

## ABSTRACT

O'DONNELL, ERICA THI. A Comparison of the Performance and Degradation of the c-Si and a-Si Photovoltaic Systems on the NCSU Solar House. (Under the direction of Dr. Stephen D. Terry).

The NCSU Solar House serves as a research test bed for building-integrated passive and active solar features. This study focuses on the two photovoltaic (PV) arrays installed on the house and the associated garage. The first PV system, located on the house, is a 5.4 kW crystalline silicon (c-Si) array installed in 2008. The garage system is a 3.1 kW amorphous silicon (a-Si) array installed in 2003. The former array's performance has been continuously monitored since 2018, and the latter's since 2020. This, combined with data from previous studies, was used to evaluate the current performance of the arrays under different conditions, as well as their long-term degradation.

This study first compared the performance of the two arrays under different weather conditions. The c-Si array was found to produce more power during sunny weather, while the a-Si array performed better under low light conditions. This includes cloudy, rainy, and snowy weather, as well as times near sunrise and sunset. An analysis of power production as a function of solar position was also conducted. This revealed that both of the arrays were negatively impacted by shading from nearby trees.

As the data available for calculating the degradation rate of the arrays is limited, a new methodology was developed to find the optimal degradation model parameters. This was used in conjunction with several different analytical methods in order to evaluate which was best suited for each array. This methodology successfully lowered the model uncertainty and increased its accuracy when given only a few years of data. Additionally, it resulted in closer agreement of the  $R_d$  values produced by the different analytical methods.

The year-over-year (YoY) method resulted in the lowest model uncertainty for both arrays. The a-Si array was found to have a higher degradation rate than the c-Si array. The c-Si array degraded at a rate of  $-0.79 \pm 0.52\%$ , while the a-Si degraded at  $-1.26 \pm 0.93\%$ . While these results have large uncertainties, they are supported by the results of past studies. Overall, these findings provide preliminary information on the degradation of the two arrays, and open up many avenues for future study of the Solar House.

A Comparison of the Performance and Degradation of the c-Si and a-Si Photovoltaic Systems on  
the NCSU Solar House

by  
Erica Thi O'Donnel

A thesis submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the degree of  
Master of Science

Mechanical Engineering

Raleigh, North Carolina

2021

APPROVED BY:

---

Dr. Stephen D. Terry  
Chair of Advisory Committee

---

Dr. Herbert M. Eckerlin

---

Dr. Brendan T. O'Connor

## **BIOGRAPHY**

Erica O'Donnel was born on September 2<sup>nd</sup>, 1997 in Indianapolis, Indiana. She grew up in Virginia, where she discovered her love of nature and the outdoors. In 2015, she started school at Virginia Tech. During that time, she was an active member of the Marching Virginians, conducted undergraduate research, and served as an executive officer for the Society of Women Engineers at Virginia Tech. She graduated in 2019 with a Bachelor of Science in Mechanical Engineering and minors in Green Engineering and Music.

That fall, she started her master's program at NC State University. While completing her thesis, she had a research assistantship in the Industrial Assessment Center and worked part-time as a Development Engineer at Hexagon Energy. Post-graduation, she looks forward to a long career helping bring renewable energy to the world.

## **ACKNOWLEDGEMENTS**

I would like to thank Dr. Terry for his continued support and advice. The opportunity he gave me to work both on the NC State Solar House and in the Industrial Assessment Center has given me the tools I need to succeed in my career. I am grateful for all that I have learned these past two years. I would also like to thank Dr. Eckerlin and Dr. O'Connor for serving on my advisory committee.

Thank you to all of the members of the IAC for your friendship and support. Anvay, Elizabeth, Emily, Kristine, Nicole, and Sarath, your good humor and perpetual willingness to commiserate over grad school has made the past couple of years bearable.

Patrick, thank you for being the most patient and encouraging partner I could ask for. Thank you for always listening as I ramble about my work, and for calming my anxieties about grad school, the pandemic, and life in general.

A million thanks to my family for being with me every step of the way. Dad, thank you for your invaluable advice from one engineer to another, and for always taking an interest in my work and encouraging me to do more. Mom, thank you for always reassuring me and going above and beyond to brighten my day. I couldn't have done this without all of you.

# TABLE OF CONTENTS

List of Tables .....	iv
List of Figures.....	v
<b>1 Introduction and Background.....</b>	<b>1</b>
1.1 Solar House PV System .....	2
1.2 Garage PV System .....	4
1.3 Previous Work and Available Data.....	6
<b>2 Literature Review .....</b>	<b>8</b>
2.1 Types of Solar Panels.....	8
2.2 Solar Panel Degradation.....	8
2.3 Calculating $R_d$ .....	10
2.3.1 Data Processing.....	11
2.3.2 Performance Metrics .....	12
2.3.3 Analytical Methods.....	14
2.3.4 Model Limitations.....	16
<b>3 Methods .....</b>	<b>18</b>
3.1 Measurement System .....	18
3.2 Plane of Array Irradiance .....	18
3.3 Cell Temperature.....	19
3.4 Calculation of Power Factor and Real Power .....	20
3.5 Data Validation .....	21
3.6 Degradation Models and Uncertainty .....	26
<b>4 Development of Degradation Model .....</b>	<b>29</b>
4.1 Performance Metric.....	31
4.2 Filters.....	33
4.3 Aggregation.....	38
4.4 Selection of Modeling Criteria.....	39
<b>5 Comparison of a-Si and c-Si Arrays .....</b>	<b>42</b>
5.1 Dependence of Performance on Weather Conditions .....	42
5.2 Dependence of Performance on Solar Position.....	48
5.3 Degradation Rates .....	49

5.3.1	Comparison to Previous Studies .....	55
<b>6</b>	<b>Conclusions.....</b>	<b>60</b>
6.1	Recommendations and Future Work.....	61
<b>References .....</b>	<b>.....</b>	<b>62</b>

## LIST OF TABLES

<b>Table 3.1:</b>	The specifications of the measurement devices.....	18
<b>Table 3.2:</b>	Filtering criteria used for base degradation model. ....	27
<b>Table 4.1:</b>	Summary of optimized filtering criteria. ....	37
<b>Table 4.2:</b>	Summary of degradation model results for c-Si test data. ....	40
<b>Table 5.1:</b>	Performance of a-Si panels under different weather conditions.....	47
<b>Table 5.2:</b>	Performance of c-Si panels under different weather conditions.....	47
<b>Table 5.3:</b>	The optimized modeling parameters for the c-Si array. ....	50
<b>Table 5.4:</b>	The optimized modeling parameters for the a-Si array. ....	50
<b>Table 5.5:</b>	Summary of degradation model results for c-Si array. ....	54
<b>Table 5.6:</b>	Summary of degradation model results for a-Si array. ....	54
<b>Table 5.7:</b>	a-Si array performance data from 2003-2021.....	55

## LIST OF FIGURES

<b>Figure 1.1:</b> The house PV system.....	2
<b>Figure 1.2:</b> The house inverter and AC disconnect. ....	3
<b>Figure 1.3:</b> The components of the house PV system. ....	4
<b>Figure 1.4:</b> The garage PV system. ....	5
<b>Figure 1.5:</b> The components of the garage PV system. ....	5
<b>Figure 1.6:</b> The garage inverter, AC disconnect, utility meter, and circuit breaker box. ....	6
<b>Figure 2.1:</b> Degradation modes for (a) all years and (b) systems built after 2007. ....	10
<b>Figure 3.1:</b> The power factor of the house as a function of the apparent power. ....	21
<b>Figure 3.2:</b> A comparison of the weather station and NSRDB values for GHI. ....	22
<b>Figure 3.3:</b> A comparison of the weather station and NSRDB values for POA irradiance. ....	23
<b>Figure 3.4:</b> A comparison of the weather station and NSRDB values for wind speed. ....	24
<b>Figure 3.5:</b> A comparison of the weather station and NSRDB values for ambient temperature. ....	25
<b>Figure 3.6:</b> Daily $PR$ s for (a) the c-Si array and (b) the a-Si array. ....	26
<b>Figure 3.7:</b> Model uncertainty as a function of number of sampling repetitions used in bootstrap model.....	28
<b>Figure 4.1:</b> $R_d$ and uncertainty as a function of the number of years of data using the base model parameters. ....	30
<b>Figure 4.2:</b> The effect of performance metric on $R_d$ and uncertainty. ....	32
<b>Figure 4.3:</b> The effect of changing the (a) lower and (b) upper limit of the performance metric on $R_d$ and uncertainty. ....	34
<b>Figure 4.4:</b> The effect of changing the (a) lower and (b) upper limit of the irradiance filter on $R_d$ and uncertainty. ....	35
<b>Figure 4.5:</b> The effect of changing the (a) lower and (b) upper limit of the temperature filter on $R_d$ and uncertainty. ....	36
<b>Figure 4.6:</b> The effect of aggregation period on $R_d$ and uncertainty.....	38
<b>Figure 4.7:</b> $R_d$ and uncertainty as a function of the number of years of data using the optimized model parameters. ....	39
<b>Figure 5.1:</b> The (a) $P_N$ and (b) $PR$ of the two solar panel systems on a sunny day. ....	43

<b>Figure 5.2:</b> The (a) $P_N$ and (b) $PR$ of the two solar panel systems on a cloudy day. (c) The correlation of $P_N$ with solar irradiance.....	44
<b>Figure 5.3:</b> The (a) $P_N$ and (b) $PR$ of the two solar panel systems on a rainy day.....	45
<b>Figure 5.4:</b> The (a) $P_N$ and (b) $PR$ of the two solar panel systems on a snowy day.....	46
<b>Figure 5.5:</b> $P_N$ as a function of solar position for the (a) c-Si and (b) a-Si arrays.....	48
<b>Figure 5.6:</b> Shading from a tree on the western side of the Solar House. ....	49
<b>Figure 5.7:</b> $R_d$ and uncertainty for the c-Si array using the base and optimized model parameters. ....	52
<b>Figure 5.8:</b> $R_d$ and uncertainty for the a-Si array using the base and optimized model parameters. ....	53
<b>Figure 5.9:</b> $R_d$ for the a-Si array based on the power at $G_{ref}$ .....	56
<b>Figure 5.10:</b> $R_d$ for the a-Si array based on the daily yield. ....	57
<b>Figure 5.11:</b> $R_d$ for the a-Si array based on $PR_{corr}$ .....	57
<b>Figure 5.12:</b> $R_d$ for the c-Si array based on panel efficiency. ....	59
<b>Figure 5.13:</b> $R_d$ for the a-Si array based on panel efficiency. ....	59

# 1 Introduction and Background

As concerns over global warming grow and fossil fuel supplies dwindle, scientists and engineers are looking towards alternative sources of energy. One of the fastest growing energy sources is solar power. Once prohibitively expensive, the installed cost of photovoltaic (PV) solar systems has dropped over 70% in the last decade [1]. With the help of these competitive prices and the stability provided by the Solar Investment Tax Credit (ITC), solar has experienced an average annual growth rate of 42% in the last 10 years [1]. Solar power offers a number of key benefits over fossil fuels, including lower greenhouse gas emissions, increased energy security, and lower overall environmental impact [2]. It also is unique in that it allows for the formation of distributed generation systems and mini-grids, allowing communities and individual buildings to produce their own power.

The US currently has 97.7 GWdc of total installed capacity, a figure that is expected to more than double in the next five years [3]. Solar is growing especially fast in North Carolina, which placed among the top states for most new capacity additions over the past decade [1]. North Carolina currently has 7,038 MWdc of installed capacity, more than any other state except California and Texas.

Solar power is advancing rapidly not only in the marketplace, but also technologically. Module efficiencies have improved considerably, and are still continuing to rise. Between 2006 and 2019, the efficiency of the average monocrystalline silicon PV panel increased from 14.7% to 18%, a trend that is expected to continue through 2030 [4]. In addition, emerging technologies continue to change the solar landscape [4]. Newer types of PVs, such as perovskite and organic cells, show promise and may revolutionize the solar market. Improvements to manufacturing and installation, including bifacial modules, solar trackers, and building-integrated PVs, have already transformed the solar industry. With such drastic changes taking place, it remains important not only to study PV performance, but to develop better methods for doing so. In addition to the standard laboratory tests, research testbeds that demonstrate how PV systems perform in the real world play a crucial role in our understanding of PV performance.

The NCSU Solar House, located at North Carolina State University, is one such place. It was built in 1981 under the direction of Professor Herbert M. Eckerlin as a demonstration of how passive solar features can be incorporated into a traditional home. It incorporates a number of

passive solar and energy conservation features, including Trombe walls, a sun space with thermal storage, and high R-value insulation. Throughout the years, new features have been added to the house, most notably a PV system.

### 1.1 Solar House PV System

In 2008, an array of 24 SunPower SPR-225 PV panels was installed on the roof of the Solar House. The panels are arranged in three rows of eight and occupy a 21'x15' rectangular area. The panels lay flush with the south-facing roof at 36° from horizontal [5]. The PV system, shown in Figure 1.1, is located in the center of the house, above the sunspace. The weather station can also be seen to the right of the array. Each monocrystalline silicon (c-Si) panel has a peak power rating of 225 W, is rated at 41.0 V and 5.49 A, and has a rated efficiency of 18.1% [6]. Each row of eight panels is wired in series and then combined in parallel in the combiner box [5]. Overall, the system has a peak power of 5.4 kW and is rated at 328 V and 16.47 A.



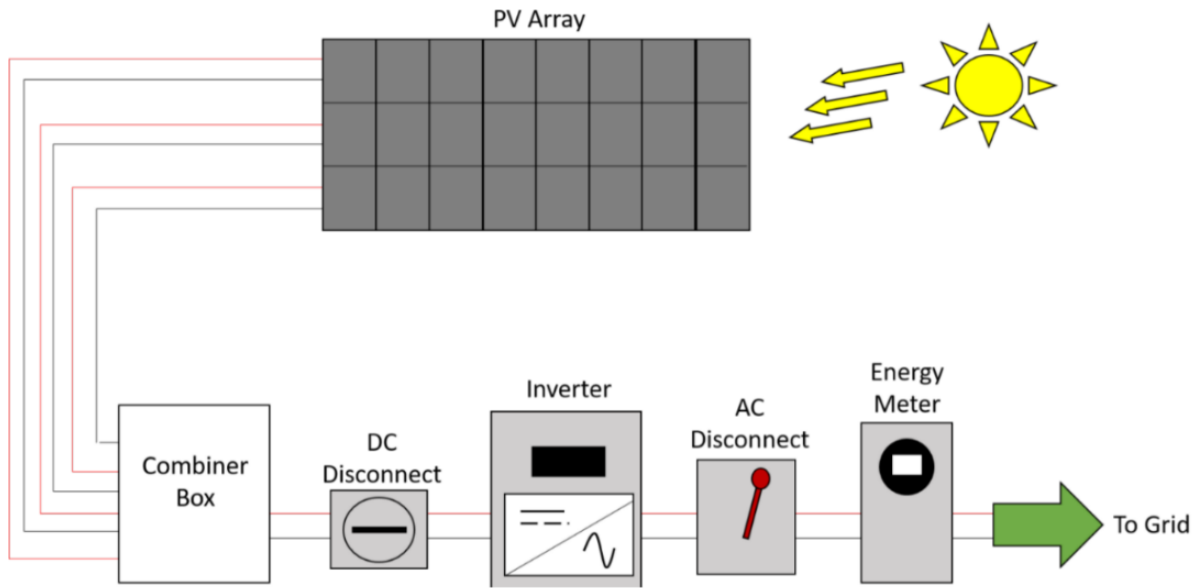
*Figure 1.1: The house PV system.*

Power from the array is fed into a SunPower SPR-6000m inverter, where it is converted from DC to AC. The inverter has a peak rated conversion efficiency of 97%, an output power factor range of 0.95-1.0, a maximum DC input current of 25 A, and an input voltage range of 250-600 V [7]. It has a maximum AC power output of 6 kW. With an inverter of this size, the DC/AC ratio is 0.9, meaning the inverter is oversized for the system. This eliminates the risk of the power output being limited by the inverter, known as inverter clipping. The inverter and AC disconnect are shown in Figure 1.2.



**Figure 1.2:** *The house inverter and AC disconnect.*

After the power is inverted, it is fed directly into the grid, as shown in Figure 1.3 [5]. There is currently no battery connected to the system.



**Figure 1.3:** The components of the house PV system.

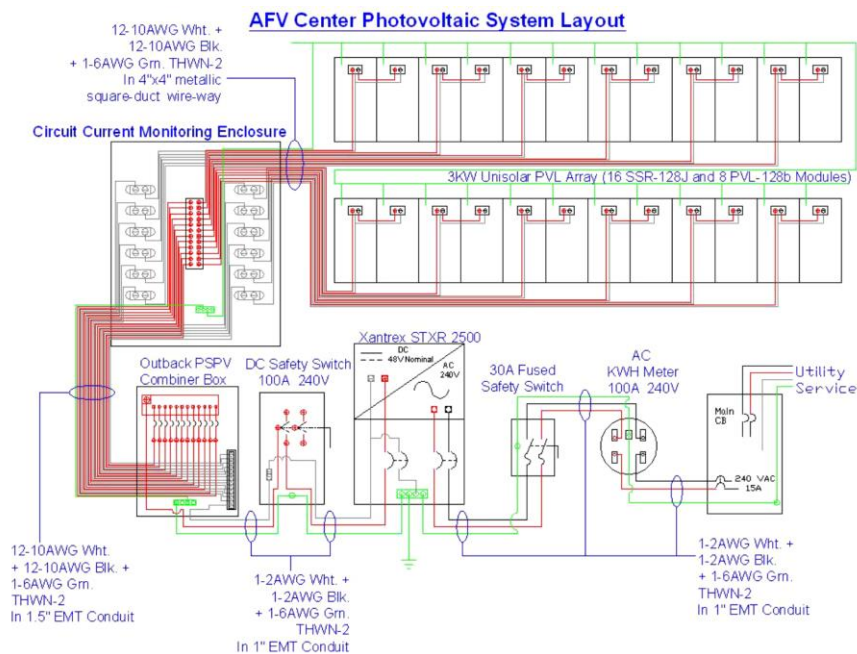
## 1.2 Garage PV System

Next to the Solar House is the Solar House garage, which once housed equipment for research on hydrogen production for vehicles. Today the garage is used as classroom space for K-12 STEM outreach activities. The garage is equipped with 24 amorphous-silicon (a-Si) solar panels, occupying an area of 31'x18'. These panels are meant to emulate conventional roofing materials and are more lightweight and flexible than the traditional c-Si panels installed on the house [8]. Of the 24 panels, sixteen are Uni-Solar SSR-128J panels and eight are Uni-Solar PVL-128B panels. As shown in Figure 1.4, the eight PVL-128B panels are on the left and the 16 SSR-128J panels are on the right. While the two types have identical electrical specifications, they were installed using different methods. The SSR-128J panels were attached to the roofing material before shipping and were in storage for over 3 years prior to installation [9]. The PVL-128B panels were field applied soon after purchasing. The array was made operational in 2003.



**Figure 1.4:** The garage PV system.

Each panel has a rated power of 128 W and is rated at 33.0 V and 3.88 A [8]. Overall, the system has a rated power of approximately 3.1 kW. Each pair of panels is wired in series in order to raise the output voltage to an acceptable level. The rated efficiency of the panels is approximately 6.8%. The system as it was in 2007 is shown in Figure 1.5. While the solar panels and combiner box are part of the original system, the other system components, including the inverter and safety switch, were replaced in 2019.



**Figure 1.5:** The components of the garage PV system.

The garage inverter is a SolarEdge SE7600A-US. The inverter has a maximum conversion efficiency of 98%, an output power factor range of 0.85-1.0, a maximum DC input current of 23.5 A, and an input voltage range of 270-500 V [10]. It has a maximum AC power output of 8.35 kW. Like the house system, this inverter is oversized, with a DC/AC ratio of only 0.36. The garage inverter, along with the AC disconnect, utility meter, and circuit breaker box are shown in Figure 1.6.



*Figure 1.6: The garage inverter, AC disconnect, utility meter, and circuit breaker box.*

### 1.3 Previous Work and Available Data

A number of works have assessed the performance of the NCSU Solar House, although they mainly focused on the performance of the house's passive solar features in the years immediately following its construction. While some data is available on the PV systems' performance, further analysis is needed to understand how they are currently performing and how they have changed over time. Nicholas Llewellyn evaluated the performance of the house's PV and solar water heating systems for data collected in 2018 and 2019 [5]. Analysis revealed that the system's performance is comparable to that of similar systems, although it did not perform as well

as predicted by the PVWatts simulation [5]. In addition to this study, raw data for the system is available from October 2018 to present. Initial system performance can be estimated from the PV panel specification sheets, although it should be noted that the initial behavior of the system is expected to be significantly different from its long term behavior [11].

A performance study was conducted on the garage PV system beginning with its construction in November 13, 2003 and continuing through April 9, 2004 [12]. While this study focused on determining the relationship between the panel short circuit current and the global irradiance, other useful data was also collected. Using this data and additional data collected in 2006-2007, Daniel Christy performed a follow-up investigation [9]. In his thesis, he evaluated the inverter efficiency, the individual performance of each PV panel, and the effect of solar flux on the system's performance. He also performed preliminary analysis of the system degradation and found a 9% power decrease from 2003 to 2007. However, these two data points are insufficient to fully assess the degradation rate of the panels. Further study is needed to gain a clearer picture. Continuous performance data collected beginning in January 2020 allows for this work to be extended. As a-Si solar panels are a newer technology whose degradation rates are not yet well established, this information is valuable in understanding how these systems behave in the long-term.

## 2 Literature Review

### 2.1 Types of Solar Panels

There are many different types of PV cells, each with their own advantages and disadvantages. While some are made of cadmium telluride (CdTe), copper indium gallium diselenide (CIGS), or organic polymers, the most common material is silicon (Si) [13]. Si solar cells can be either crystalline, multi-crystalline, or amorphous. c-Si panels are most commonly used due to their higher rated efficiencies and lower cost per rated Watt; the efficiency of a typical a-Si panel is only about half that of a typical c-Si panel [14]. In the solar house's case, the efficiency of the c-Si modules is nearly three times that of the a-Si modules. As a-Si panels have lower efficiencies, they require a larger area to capture the same amount of energy.

The performance characteristics are calculated when the module is subjected to standard test conditions (STC): a cell temperature of 25°C, a solar irradiance of 1000 W/m<sup>2</sup>, and an air mass of 1.5. Field conditions rarely meet these conditions, so PV panels often perform significantly differently in the field than they do in a laboratory setting. In these real-world conditions, a-Si panels often outperform c-Si panels as they are effective in a wider range of conditions [14]. This can be attributed to the fact that a-Si panels typically have larger bandgaps [15]. The bandgap is the energy needed to change a valence electron in an atom to a conduction electron. A photon with energy below the bandgap cannot be absorbed, while a photon with energy above the bandgap will have its excess energy dissipated as heat. Jansen et al. found that the larger bandgap in a-Si panels allows them to absorb more energy throughout the day, especially in overcast conditions where diffuse, blue light dominates [14]. They also perform better under high temperatures as they have temperature coefficients roughly half those of c-Si panels. This is significant as the hottest part of the day is often the time when the most solar insolation is available and electrical demand is at its peak. The study found that these effects cause a-Si panels to have a higher energy yield overall, especially during the hottest part of the day.

### 2.2 Solar Panel Degradation

Solar panels are expected to lose efficiency over time. The rate at which the system's performance decreases is represented by the degradation rate,  $R_d$ . The standard  $R_d$  value used by many manufacturers is 0.5% performance loss per year, represented as -0.5% [16]. However, a number of factors can cause the actual rate to differ from its stated value. As the panel performance

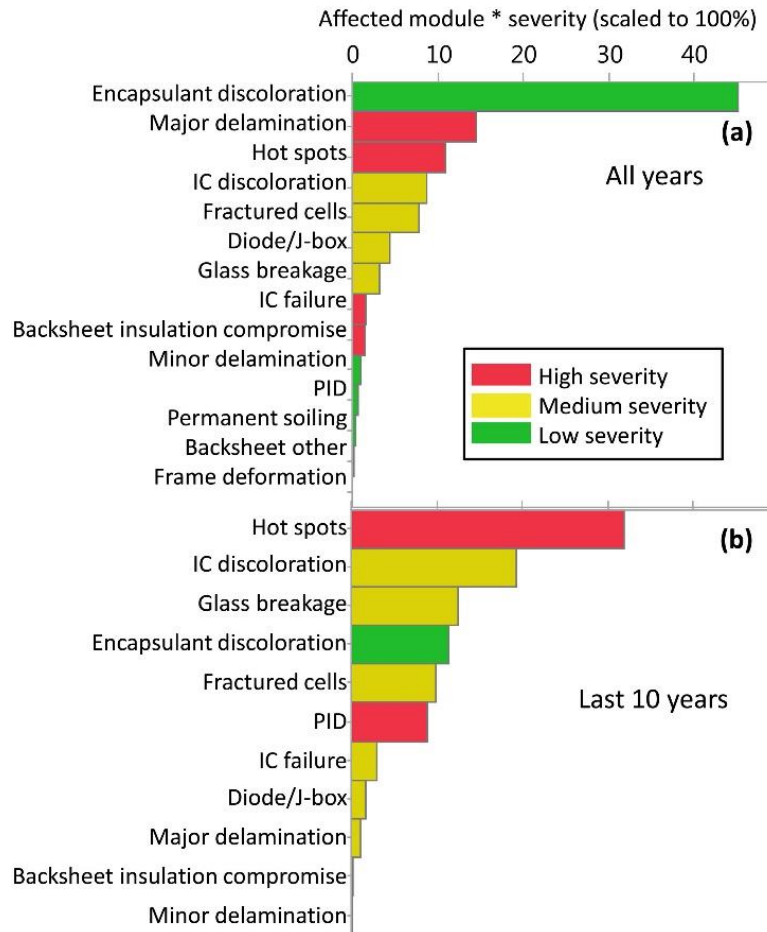
is key in determining the environmental and financial impacts of the system, establishing the true  $R_d$  is important for the successful integration of PV systems into the electrical grid.

A recent review paper analyzed over 11,000 PV degradation rates aggregated from nearly 200 studies [16]. When considering every data point, the median degradation rate was -0.90%. For only c-Si systems, they found a median degradation rate of -0.91%. The median degradation rates for thin-film systems, like the a-Si system on the garage, are considerably higher at -1.38%. These values are significantly higher than the value reported by manufacturers, indicating that PV panels may perform differently under field conditions than they do in a laboratory setting.

Degradation can happen in a number of different ways. One of the most commonly reported degradation modes is encapsulant discoloration. Another is the development of hot spots, which occurs when one cell in a series is negatively biased, causing it to dissipate the power produced by the other cells as heat [17]. Hot spots can be caused by damage to the cell or interconnections, partial shading, and current mismatch between cells. Delamination can occur between the encapsulant and the cell or between the cell and the glass, increasing light reflection and allowing water to penetrate the cell [18]. Moisture penetration can in turn lead to corrosion, which increases leakage current and can cause the cells to separate from their frame [11]. Other degradation modes include light-induced degradation (LID), bubbling, potential-induced degradation (PID), mechanical failure, and soiling [11], [19]–[21].

A literature review of degradation modes found that 56% of papers report discoloration, 42% report delamination, and 37% report hot spots [22]. However, while discoloration is prevalent, it has a much smaller effect on power loss than hot spots. This is important to consider as the most prevalent degradation mode may not necessarily be the most significant.

The age of the installation is an important factor in which modes are of most concern. Jordan et al. performed a review of the number of times a degradation mode was mentioned in literature, then weighted each mode by its severity [23]. A summary of their results is shown in Figure 2.1. When looking at systems of all ages, discoloration dominates overall, even when accounting for its low impact. But when analyzing only systems built after 2007, hot spots are the dominant degradation mode.



**Figure 2.1:** Degradation modes for (a) all years and (b) systems built after 2007.

### 2.3 Calculating $R_d$

There are many methods for calculating  $R_d$ , and as of yet there is no standardized methodology. The most controlled and well-documented method is indoor IV curve testing at STC. However, this method is not frequently used as it is time-consuming, and is uneconomical for larger systems [24]. More often, PV systems are tested in the field. In this case, the methods for collecting, filtering, and aggregating data are left to the researcher's discretion. There are also a variety of performance metrics and analytical methods that can be used to calculate  $R_d$ . Unsurprisingly, this inconsistent methodology results in a high degree of interanalyst variability. In one study, six analysts were given raw data for five systems and asked to use their preferred method to calculate  $R_d$ . The resulting variance among the analysts was 0.745% [25]. Such a large variance indicates that many methods may not accurately calculate  $R_d$ . Additionally, comparing

the performance of panels analyzed by different groups may prove difficult or impossible. From this, it is clear that the different methods must be evaluated in order to determine which can produce the most reliable results.

### 2.3.1 Data Processing

In order to determine  $R_d$ , one must collect data on energy production and environmental conditions. However, the exact data collected can vary widely. Energy production data can be measured at the panel (DC) or after the inverter (AC). Cell temperature can be measured directly, modeled based on atmospheric conditions, or not taken into account at all. Irradiance data can be collected locally, from a nearby weather station, or from satellites. If measured locally, the sensor can be a pyranometer or reference cell, either tilted to the plane of the array or mounted horizontally. All of these differences in technique can have an impact on the data, and ultimately the final result. For example, pyranometers are less spectrally selective than most reference cells, and thus report an annual irradiance value 2-4% higher than that of the reference cells [26].

Much of the data collected will not be representative of the system's performance. A temporary drop in performance caused by soiling, inclement weather, outages, or shading can create noise and obscure long-term trends. Nighttime data is likewise not an accurate representation of the system's performance. Applying filters can remove this erroneous data and reduce noise. One of the most common filters excludes data from the night and cloudy days by filtering on the measured irradiance. This has been shown to minimize model uncertainty [27]. However, care must be taken not to over-filter the data, as this could introduce bias, reduce the statistical significance of the results, and artificially lower the degradation rate [24], [28]. The data can also be filtered to remove the effects of inverter clipping, extreme cell temperatures, unstable irradiance and temperature measurements, and other outliers [28].

Paudyal et al. investigated the effects of data filtering on  $R_d$  by varying each criterion individually. They found that increasing the irradiance filter from 500 to 800 W/m<sup>2</sup> decreased  $R_d$  by up to 0.8%, while decreasing the clear-sky index filter from  $\pm 20\%$  to  $\pm 5\%$  decreased  $R_d$  by up to 1.3% [29]. Jordan and Kurtz applied irradiance, outlier, and stability filters to their data, then examined the correlation of  $R_d$  and uncertainty with the percentage of data points left in the dataset. They found that while the minimum uncertainty occurred at different points for different systems, minimizing uncertainty resulted in an  $R_d$  value in agreement with indoor IV measurements for all

systems [28]. Thus, selecting the appropriate filters is important not only in reducing uncertainty, but also in obtaining a more accurate result.

Lastly, the data must be aggregated. The selection of the aggregation period is important as too small of an interval leaves the model more sensitive to outliers and increases variability, while too large of an interval reduces the number of available data points. For this reason, data is often aggregated on a daily, weekly, or monthly basis [24]. One study found that while the standard deviation of the data continues decreasing as the aggregation interval increases, a weekly interval was able to reduce standard deviation significantly while maintaining a sufficient number of data points [30]. Another found that reducing the aggregation interval from monthly to weekly allows  $R_d$  to be determined in a shorter period of time when using an analytical method that is robust against outliers [31]. Other studies have found little to no correlation between  $R_d$  and aggregation period [29], [32].

### 2.3.2 Performance Metrics

$R_d$  is defined as the annual percent decrease of some performance metric, but the exact metric used can vary. One of the most informative metrics is the IV curve of the system. This gives information on the degradation of each of the individual electrical parameters (short circuit current, open circuit voltage, maximum power, etc.), which can help identify the main degradation modes [33]. However, tracing the IV curve involves performing a voltage sweep, which means measurements cannot be performed during regular operation. Thus, this method is only appropriate for use in research installations [24].

A more commonly used performance metric is the performance ratio ( $PR$ ), defined as

$$PR = \frac{Y_f}{Y_r} \quad (1)$$

where  $Y_f$  is the final yield and  $Y_r$  is the reference yield. These parameters are described by the equations

$$Y_f = \frac{E}{P_{rated}} \quad (2)$$

$$Y_r = \frac{H}{G_{ref}} \quad (3)$$

where  $E$  is the energy generated within a specified time period,  $P_{rated}$  is the rated power of the system,  $H$  is the total in-plane irradiance over the same time period, and  $G_{ref}$  is the reference irradiance, typically  $1000 \text{ W/m}^2$ . This metric is simple and can be calculated for any time interval. However, while it does take into account variations in solar radiation, it does not account for temperature effects. This leads to seasonal variations in the data which can obscure long-term trends. One way to account for this is to use a temperature-corrected performance ratio,

$$PR_{corr} = \frac{Y_f}{Y_r (1 - \gamma(T_c - T_{c,ref}))} \quad (4)$$

where  $\gamma$  is the temperature coefficient,  $T_c$  is the cell temperature, and  $T_{c,ref}$  is the reference cell temperature [34]. This has been shown to improve the accuracy and precision of the  $R_d$  calculations and allow its determination in a shorter period of time [28]. Normalized metrics, such as  $P_{DC}/P_{rated}$ ,  $P_{AC}/P_{rated}$ , and  $P_{DC}/G$  are also frequently used [24].

More advanced options for performance metrics include regression models such as the Photovoltaics for Utility Scale Applications (PVUSA) model and the Sandia Array Performance Model (SAPM). Both models estimate the system's expected output based on the irradiance, ambient temperature, and wind speed [35], [36]. The two models have been found to be in good agreement with each other [35]. While they are able to account for environmental effects with an uncertainty of around 5%, they are not as accurate at low irradiance values, and building the models requires time and a larger amount of data [36].

Analysis was conducted using the SAPM to normalize the system's power. As the SAPM estimates the expected power, system performance can be evaluated as a fraction of the expected power, producing a normalized power value,

$$P_{N,SAPM} = \frac{P_{AC}}{P_{SAPM}} \quad (5)$$

where  $P_{AC}$  is the measured array power and  $P_{SAPM}$  is the SAPM modeled power. A  $P_{N,SAPM}$  value of 1 indicates that the system is performing as modeled, while a smaller value indicates that the system is producing less power than expected, either due to degradation or to other losses.

Studies that compare the  $R_d$  values for different performance metrics have found mixed results. Phinikarides et al. found that  $PR$  produces lower degradation rates than  $PR_{corr}$  [32], while Jordan and Kurtz found the opposite [28]. Marion et al. found good agreement between  $R_d$  values calculated using  $PR$  and the PVUSA model [37]. However, Makrides et al. found significant deviations [38]. Some of these contradictions may be due to differences in data filtering and aggregation, as discussed above, as well as the use of different analytical methods.

### 2.3.3 Analytical Methods

There are a variety of ways to extract  $R_d$  from the data. Commonly used methods include ordinary least squares regression (OLS), classical decomposition (CD), seasonal-trend decomposition (STL), Holt-Winters (HW) exponential smoothing, and the year-over-year method (YoY) [24], [30]. In recent years, more advanced techniques have emerged, such as Autoregressive Integrated Moving Average (ARIMA) models, machine learning algorithms, and new statistical methods [31], [39]–[41].

The simplest and most common method for trend extraction is OLS, or linear regression. This method attempts to minimize the sum of squared residuals to fit a linear model to the data [24].  $R_d$  is then simply the slope of the resulting equation. Since this model does not account for seasonal variations, it can have a very large model uncertainty [24]. It is also very sensitive to outliers, especially those towards the beginning and end of the dataset [25]. Thus, it is important that the proper filters be applied to remove outliers when using this method.

The CD method breaks the model into its trend, seasonal, and residual components. The trend component is extracted using a 12-month centered moving average [32]. While this method can account for seasonality, it assumes that the seasonal component does not vary with time, and therefore may not capture all the characteristics of the data. Additionally, the application of a 12-month moving average means that the model excludes the first and last six months of the data [42]. This is especially of concern when only a few years of data are available. Lastly, it is not robust to data shifts [42].

Similarly to CD, STL breaks the model down into trend, seasonal, and residual components. It differs in that it uses LOcally wEighted Scatterplot Smoothing (LOESS) to extract the trend [43]. This method has several advantages over CD. It allows the seasonal component to

change over time and is robust to outliers [42]. The user can control the smoothness of the trend, as well as how quickly the trend and seasonal components change [42]. This may result in a more detailed model, but it also requires the user to have a greater understanding of the data.

Another method which breaks the model into seasonal and trend components is the HW method. This method combines exponential smoothing equations for the level, trend, and seasonal components to construct the model [44], [45]. It is robust against outliers and accounts for seasonality in the data [46]. The seasonal component can be either additive or multiplicative. The additive method is appropriate for datasets where the seasonal component remains relatively constant, as is often the case for PV performance data. The multiplicative method may be used if the seasonal component varies over time. This allows the model to have greater flexibility. However, it also requires input from the analyst, which can affect the results and requires a greater understanding of the data.

The YoY method takes a different approach by comparing each data point to its corresponding points in the previous years. This generates a distribution of  $R_d$  values, the median of which is taken as a representative value for the system [47]. This method is independent of seasonality and is somewhat robust against outliers [47]. One drawback is that it cannot be used to examine short-term trends.

Several studies have investigated the effects of the different methods on model accuracy and uncertainty. One study performed an analysis of the OLS, CD, HW, STL, and ARIMA methods for a variety of PV technologies [48]. It was found that the OLS method produced the lowest value of  $R_d$  for most systems. However, it was the least robust as there were significant autocorrelations in the residuals due to seasonal fluctuations. All of the other analytical methods were able to adequately account for these seasonal variations. The CD and HW methods were in good agreement with each other, as were the STL and ARIMA methods.

A study by Lindig et. al. found similar results [46]. The degradation rates for one c-Si and one a-Si system were calculated using the OLS, CD, HW, STL, and ARIMA methods. The OLS method was found to be insufficient for calculating  $R_d$  due to its high uncertainty and sensitivity to outliers. The CD method also had some drawbacks as it is more sensitive to outliers and excluded the first and last datapoints. As a result, it tended to overestimate  $R_d$ . The results obtained using the HW, STL, and ARIMA methods were found to be in good agreement. The STL and

ARIMA methods gave the most accurate results and had the lowest model uncertainty. However, between the two, the STL method was preferred due to its ease of implementation.

A comparison between the OLS and YoY methods revealed that YoY is more robust to seasonality, outliers, and soiling [47]. Despite this, the degradation rates obtained from each method were in good agreement when irradiance, outlier, and stability filters were applied. While these studies provide a baseline for how different analytical methods can be expected to perform, they do not take into account the effects of the performance metric and data processing decisions made by the analyst, the variation of which can have a profound impact on the results.

#### 2.3.4 Model Limitations

In order to build an accurate degradation model, there must be a sufficient amount of high-quality data available, typically around 5 years [49]. This can be a major limitation in some situations. In the case of the Solar House, the observation period is less than three years, which is insufficient for producing an accurate result using some modeling techniques. The requirement of long observation periods is also a concern with new PV technologies and products, as it is not ideal to have to wait many years before gaining information on long-term performance and degradation. There are also major economic implications, as a larger than expected degradation rate increases the levelized cost of energy (LCOE) for the PV system. Furthermore, as PV power's share of the energy landscape continues to grow, reliable methods for determining  $R_d$  will become even more important. Successful grid integration will require the accurate prediction of energy output over a long period of time. These future power predictions are highly dependent on the degradation rate.

The length of observation period necessary to get an accurate  $R_d$  value varies based on the degradation modeling method used. However, there has not been much research into how the required observation period can be minimized. One study investigated this, comparing the results from the OLS, CD, and ARIMA + CD methods using a simulated dataset with a known degradation rate [31]. Given a long enough observation period, five or six years, all methods converged on the true degradation rate. However, with fewer years of data available, both the OLS and CD methods demonstrated a large amount of bias. In contrast, the ARIMA + CD method was able to accurately model the degradation rate with as few as two years of data. This study demonstrates that it is possible to accurately determine  $R_d$  in a short period of time if the appropriate modeling method is used. However, it does not factor in the effect of performance

metric or filter selection, nor does it consider other common analytical methods like HW, STL, and YoY. The latter part of this study addresses these additional factors with the goal of further minimizing the amount of data required to determine  $R_d$ .

### 3 Methods

#### 3.1 Measurement System

For both PV systems, the array AC power is measured with current transformers and recorded by energy meters. Weather data is taken from a weather station located immediately behind the house. The station records irradiance, ambient temperature, relative humidity, wind direction, and wind speed. Energy data is logged once every minute and weather data once every five minutes. All measurement equipment was manufactured by Onset Computer Corp. The data is stored on a remote server, then downloaded locally for further analysis. A summary of the specifications for each device is given in Table 3.1.

*Table 3.1: The specifications of the measurement devices.*

Device	Model Number	Range	Resolution	Uncertainty
Energy Meter Data Logger	EG4115	9-60 Vdc, 5-6900 A		
Current Transformers	ACT-0750-050	5-250 A		$\pm 0.75\%$
Weather Station Data Logger	RX3002			$\pm 8$ s/month
Pyranometer	S-LIB-M003	0-1280 W/m <sup>2</sup> , 300-1100 nm	1.25 W/m <sup>2</sup>	$\pm 10$ W/m <sup>2</sup> or $\pm 5\%$
Temperature Sensor	S-THB-M008	-40-75°C	0.02°C at 25°C	$\pm 0.21$ °C
RH Sensor	S-THB-M008	0-100% RH	$\pm 2.5\%$ from 10% to 90% RH, $\pm 5\%$ otherwise	0.1% RH
Solar Radiation Shield	RS3-B			
Wind Direction Sensor	S-WDA-M003	0-355°	1.4°	$\pm 5$ °
Wind Speed Sensor	S-WSB-M003	0-76 m/s	0.5 m/s	$\pm 1.1$ m/s or $\pm 4\%$

#### 3.2 Plane of Array Irradiance

The Solar House is equipped with a horizontally mounted pyranometer, which measures the global horizontal irradiance (*GHI*). This value was transposed to the plane-of-array (POA)

irradiance ( $G$ ) in order to account for the difference in angle between the array and the pyranometer. The POA irradiance is defined as

$$G = G_b + G_g + G_d \quad (6)$$

where  $G_b$  is the beam irradiance,  $G_g$  is the ground-reflected diffuse irradiance, and  $G_d$  is the sky diffuse irradiance [50]. The beam irradiance is represented by the equation

$$G_b = DNI * \cos(AOI) \quad (7)$$

where  $DNI$  is the direct normal irradiance and  $AOI$  is the angle of incidence between the Sun's rays and the array [50].  $DNI$  was calculated from  $GHI$  using the DIRINT model function of the pvlib python library [51], [52].

The ground-reflected diffuse irradiance is represented by the equation

$$G_g = GHI * albedo * \frac{1 - \cos\theta_T}{2} \quad (8)$$

where  $albedo$  is the reflectivity of the ground and  $\theta_T$  is the angle of the array from horizontal. The albedo of a surface can range from 0 for a black body, to 1 for a completely reflective surface. The albedo for the Solar House was estimated to be 0.25, which is the value for grass [50].

Lastly, the sky diffuse irradiance can be calculated using the isotropic sky diffuse model:

$$G_d = DHI * \frac{1 + \cos\theta_T}{2} \quad (9)$$

In this model,  $DHI$  is the diffuse horizontal irradiance,

$$DHI = GHI - DNI * \cos\theta_z \quad (10)$$

where  $\theta_z$  is the solar zenith angle [50].

### 3.3 Cell Temperature

Cell temperature can be measured by attaching a thermocouple or thermometer to the back of the solar panel. However, it is not currently possible to access the back side of the panels, so the cell temperature must be modeled based on available data. It can be estimated to within  $\pm 5^\circ\text{C}$  using

a simple empirical model [36]. The temperature of the module back surface  $T_m$  is a function of ambient temperature, irradiance, and wind speed,

$$T_m = G \cdot e^{a+b \cdot WS} + T_a \quad (11)$$

where  $a$  and  $b$  are empirical coefficients,  $WS$  is the wind speed in m/s, and  $T_a$  is the ambient temperature in °C. The empirical coefficients were retrieved from the Sandia module database [53].  $T_c$  is related to  $T_m$  by

$$T_c = T_m + \frac{G}{G_{STC}} \Delta T \quad (12)$$

where  $\Delta T$  is the temperature difference between the cell and the module when  $G = G_{STC}$ , also retrieved from the Sandia module database.

### 3.4 Calculation of Power Factor and Real Power

For several months in 2019, the data logger failed to collect any data for phase B of the house current, causing the calculation of energy usage to be incorrect. To account for this, the apparent power  $V$  of the house was calculated using the equation

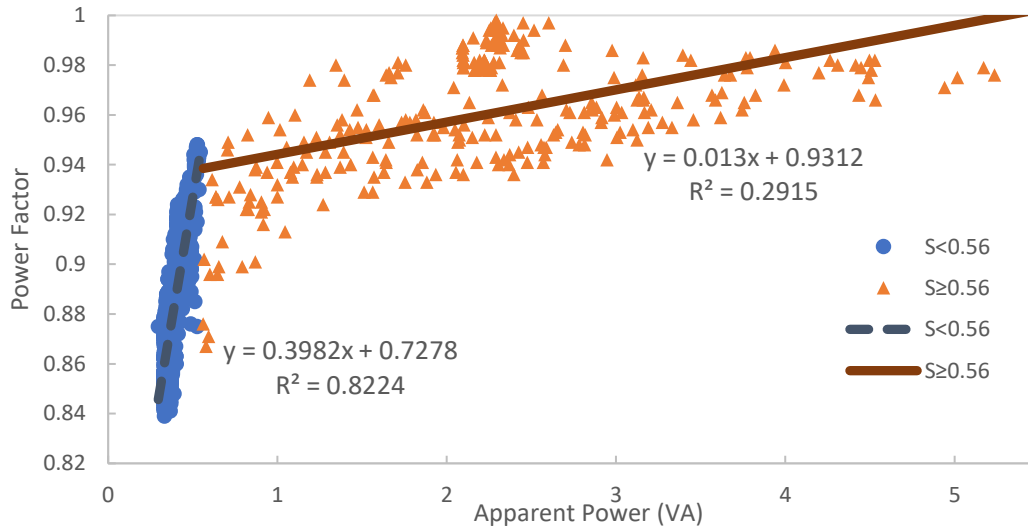
$$S = (I_A + I_B)V \quad (13)$$

where  $I_A$  and  $I_B$  are the currents for phases A and B of the house, respectively, and  $V$  is the supplied voltage, assumed to be constant at 123V. This equation, however, is not an accurate representation of the real power absorbed by the load. Reactive and/or non-linear loads cause the voltage and current drawn to be out of phase, resulting in a portion of the power flowing back to the source. This is known as the reactive power ( $Q$ ). The apparent power calculated in Equation 13 is a vector sum of the reactive power and the real power  $P$  used by the load. To isolate the real power from the apparent power, we must determine the relationship between the real and apparent power, given by the power factor

$$PF = \frac{P}{S} \quad (14)$$

The power factor was calculated using power data from September 3<sup>rd</sup> to September 10<sup>th</sup>, 2019. Data from this time period was collected in a way that correctly accounted for the power factor. The power factor changes based on which loads are currently drawing power, and thus is a function of apparent power (Figure 3.1). The data is split into two regimes at  $S = 0.56\text{VA}$ . As the

equation for the later regime has a small slope and a low  $R^2$  value,  $PF$  was taken to be constant for  $S \geq 0.56VA$ . These equations were used to calculate the real power for the months with missing data.

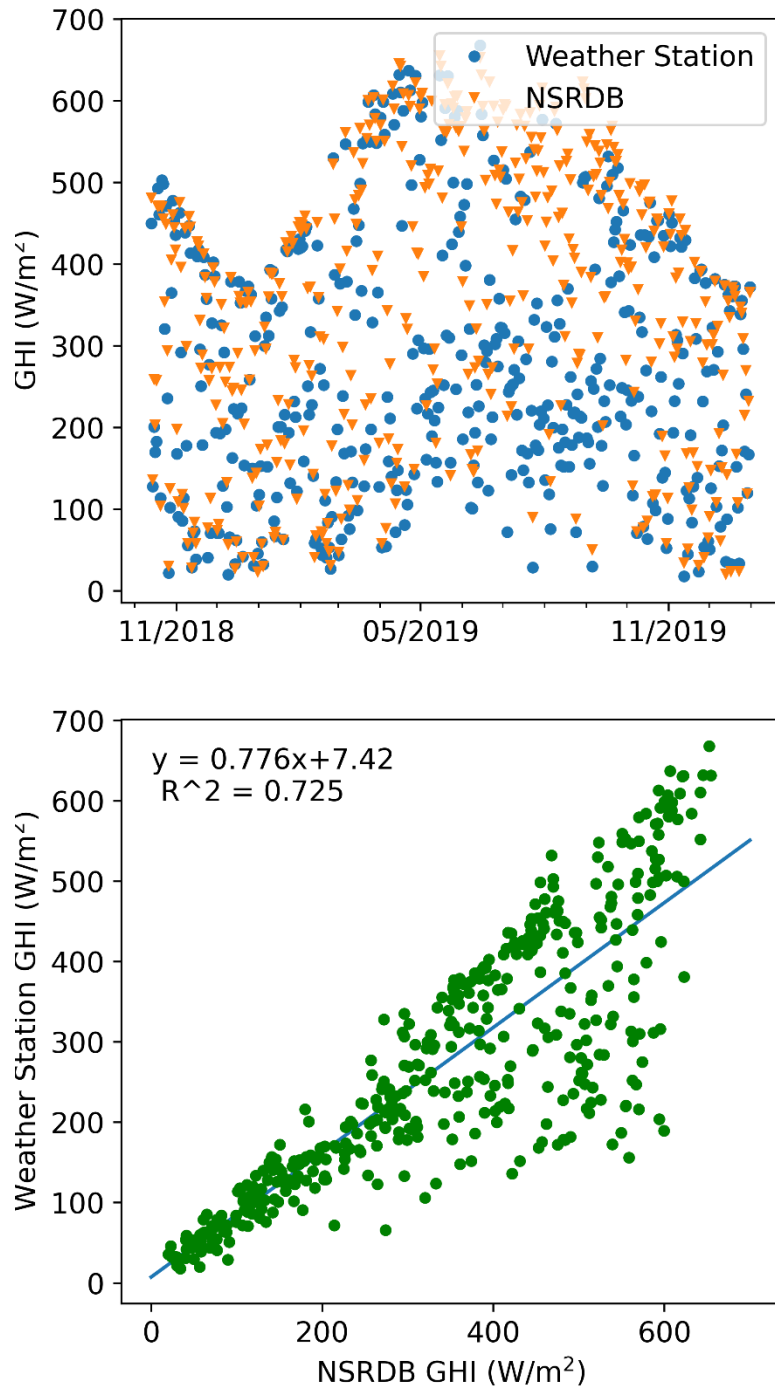


**Figure 3.1:** The power factor of the house as a function of the apparent power.

### 3.5 Data Validation

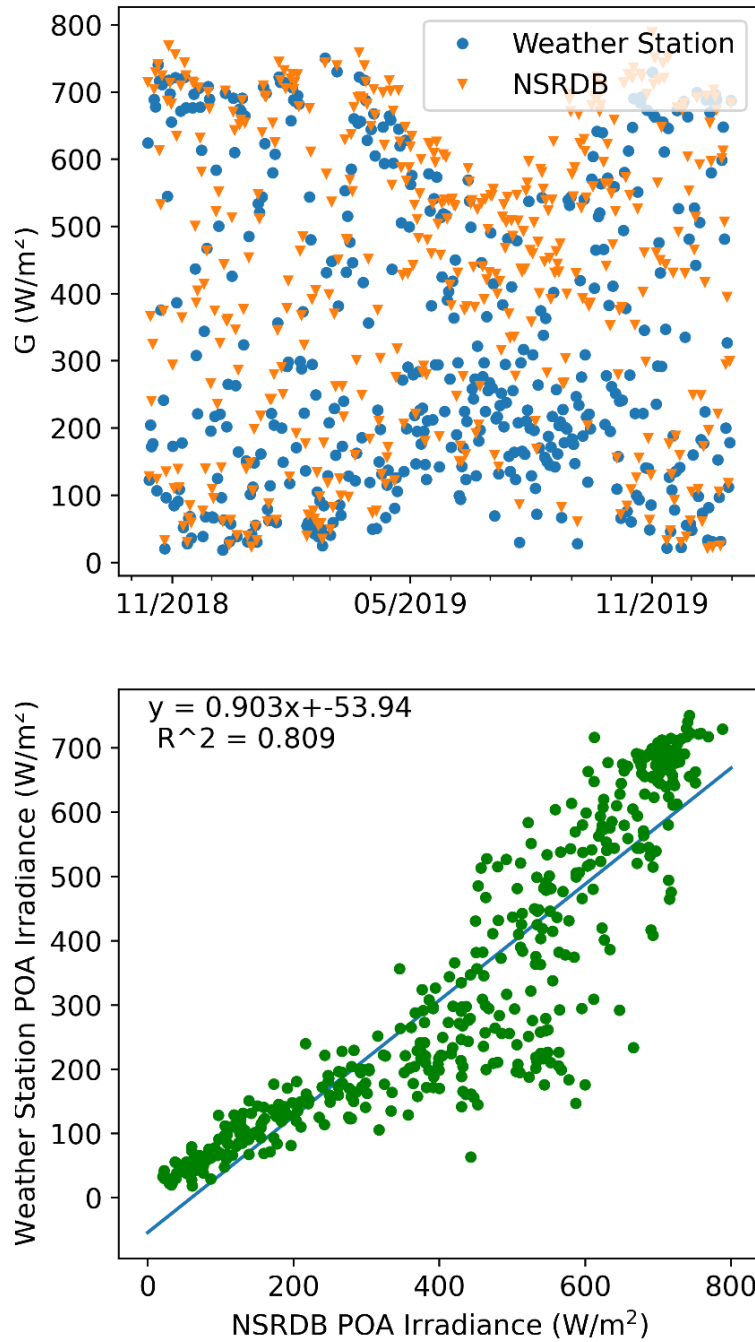
To ensure that the weather station was operating as expected, the station's weather data was compared to data from the National Solar Resource Database (NSRDB) [54]. The NSRDB provides data in 30 minute intervals and has a horizontal resolution of 4 km. The  $GHI$  and  $DNI$  data have been shown to have mean percentage biases of less than 5% and 10%, respectively. In comparing these two datasets, the daily median value for each measurement was used.

Figure 3.2 shows a comparison of the  $GHI$  data for both sources over time, as well as the linear relationship between them. While the data from each source follows similar trends, they are not a perfect match. For many days, the measured  $GHI$  is much lower than that given by the NSRDB. These discrepancies may be explained by the different measurement techniques used. While the weather station uses a local pyranometer to measure irradiance, the NSRDB uses satellite data to model irradiance. There may also be smaller spatial variances that the NSRDB model does not take into account due to its lower spatial resolution.



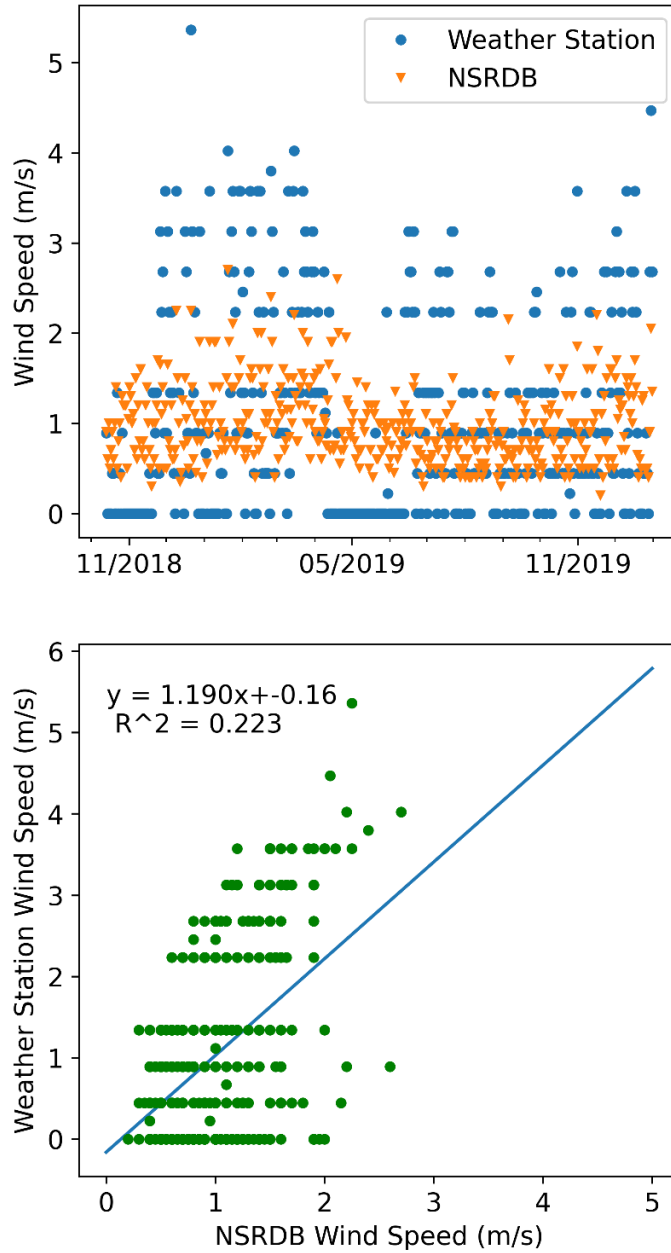
**Figure 3.2:** A comparison of the weather station and NSRDB values for GHI.

A similar trend is shown in the comparison of POA irradiance values (Figure 3.3). While the NSRDB values are often higher than the weather station values, the two are in reasonable agreement, and in fact match more closely than the values for *GHI* alone. This verifies that the process used in Section 3.2 to model *DHI*, *DNI*, and *G* is sufficiently accurate.



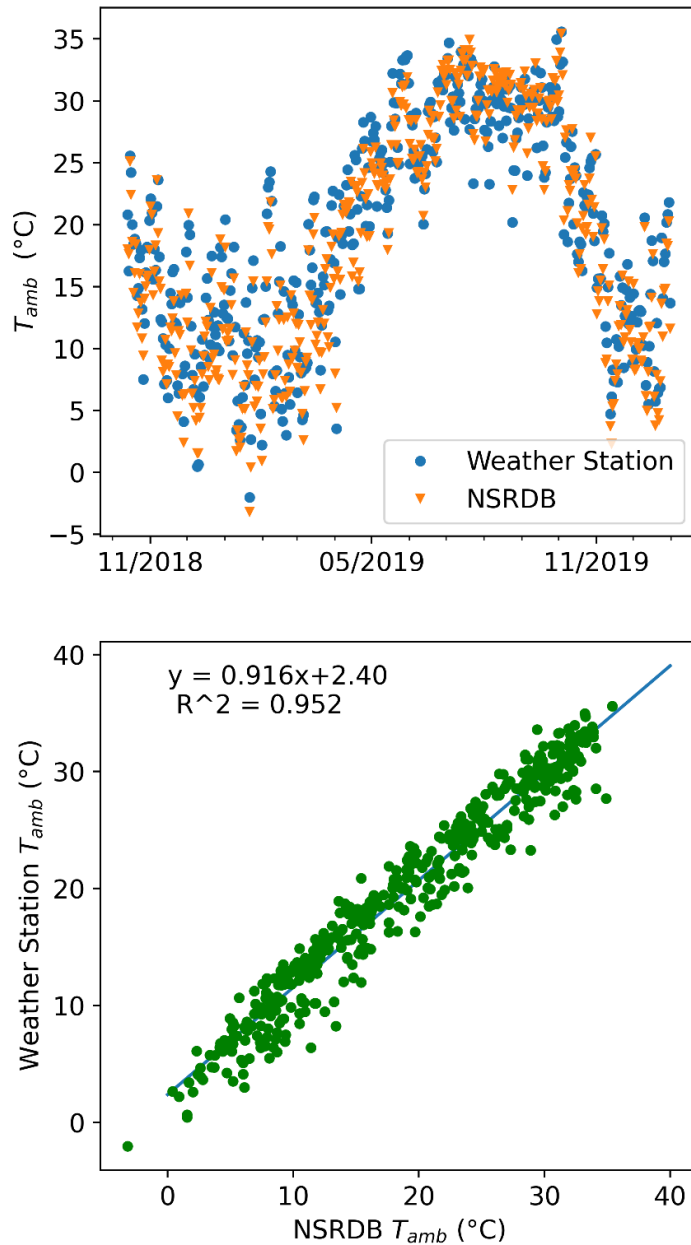
**Figure 3.3:** A comparison of the weather station and NSRDB values for POA irradiance.

Figure 3.4 shows the same analysis for wind speed. There was significant deviation between these two datasets. This may be partly attributed to the fact that the weather station's resolution is much lower. Additionally, the anemometer on the weather station is partially shielded from the wind by the house and surrounding trees. As wind speed is not necessary for these calculations, this data was omitted.



**Figure 3.4:** A comparison of the weather station and NSRDB values for wind speed.

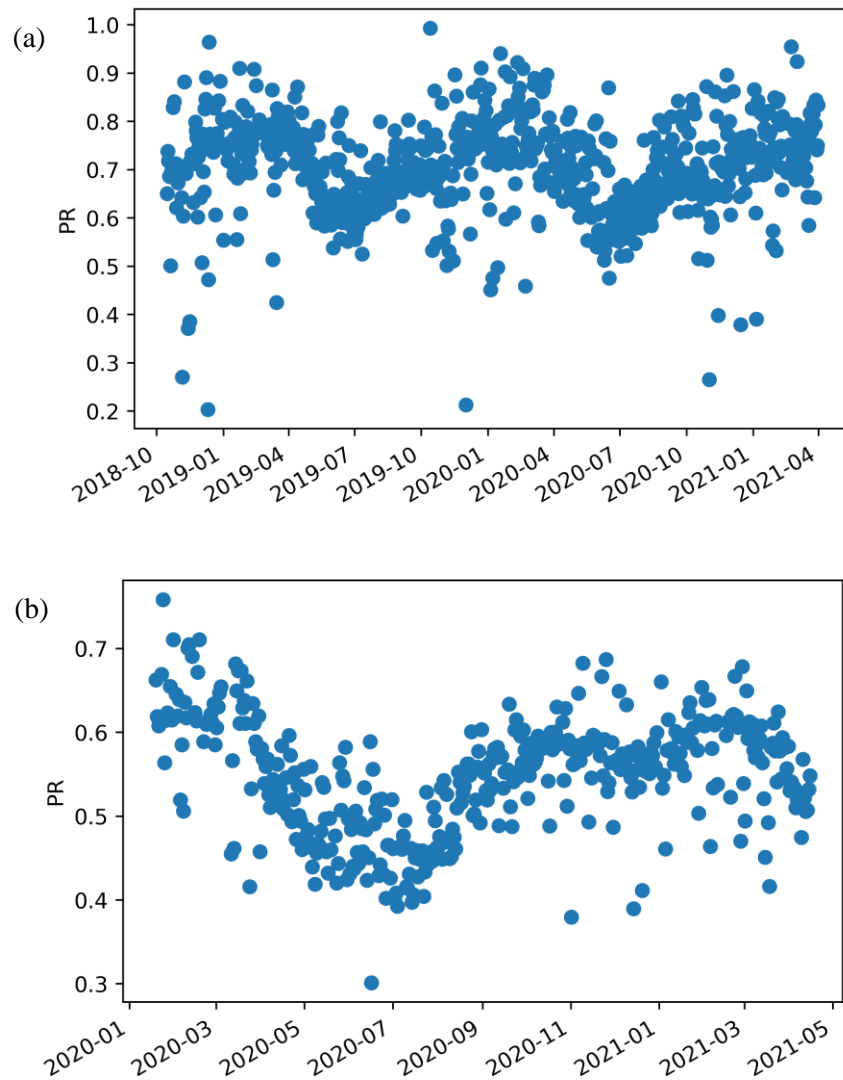
The last parameter analyzed was ambient temperature (Figure 3.5). Very good agreement was found between the two datasets, with the weather station values being on average only slightly higher than the NSRDB values. This indicates that the weather station thermometer is accurate and properly calibrated.



**Figure 3.5:** A comparison of the weather station and NSRDB values for ambient temperature.

### 3.6 Degradation Models and Uncertainty

The daily  $PRs$  for both arrays are shown in Figure 3.6. As can be seen from this figure, the observation periods for both arrays are very short. The fact that so little data is available presents a substantial challenge to accurately determining the degradation rate. As discussed in Section 2.3, the method used to calculate  $R_d$  can have a large impact on the accuracy and uncertainty of the results. Thus, in order to maximize the accuracy of the calculations, various degradation models were built and compared.



**Figure 3.6:** Daily  $PRs$  for (a) the  $c$ -Si array and (b) the  $a$ -Si array.

For the base model parameters, *PR* was selected as the performance metric. The data was filtered based on *PR*, cell temperature, and POA irradiance, as described in Table 3.2. Data was then aggregated into weekly values using the insolation weighted average.

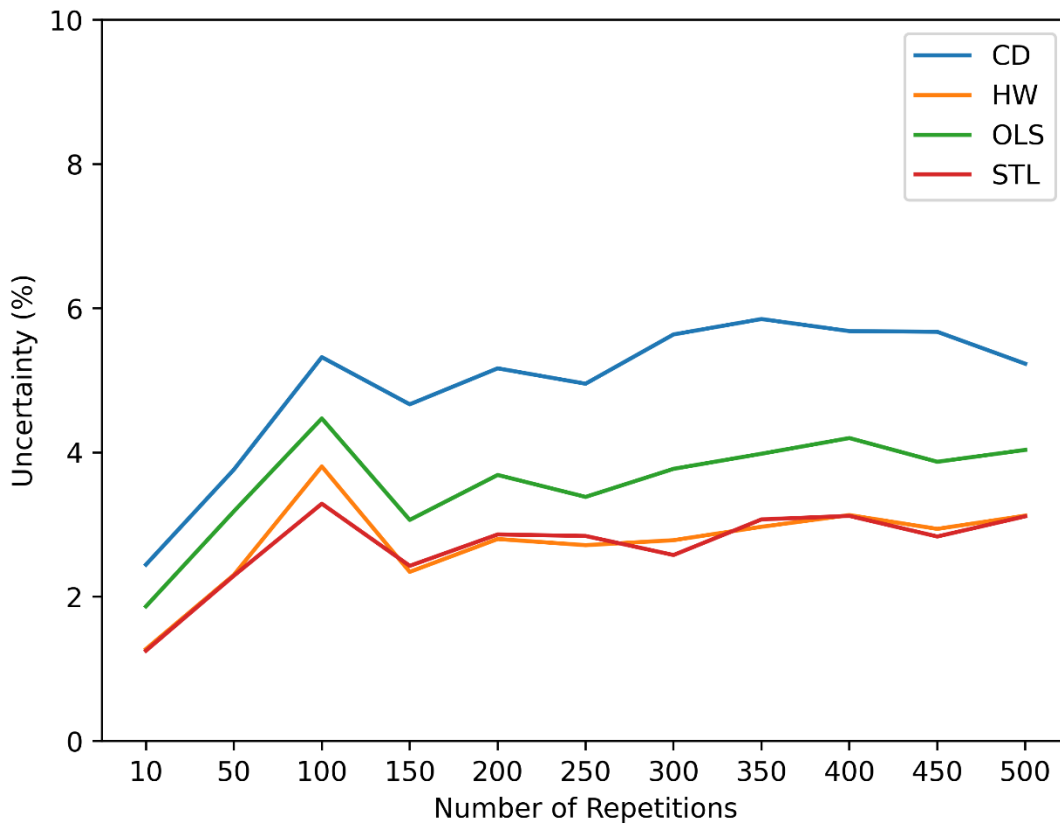
*Table 3.2: Filtering criteria used for base degradation model.*

<b>Filtering Criteria</b>	<b>Value</b>
<b><i>PR</i> Lower Limit</b>	0.01
<b><i>PR</i> Upper Limit</b>	1.5
<b>Cell Temperature Lower Limit</b>	-50°C
<b>Cell Temperature Upper Limit</b>	110°C
<b>POA Irradiance Lower Limit</b>	200 W/m <sup>2</sup>
<b>POA Irradiance Upper Limit</b>	1,200 W/m <sup>2</sup>

The analytical methods explored were ordinary least squares regression (OLS), classical decomposition (CD), seasonal-trend decomposition (STL), Holt-Winters (HW) exponential smoothing, and the year-over-year method (YoY). OLS and YoY calculations were performed using the RdTools library in Python [55]. For the CD, STL, and HW methods, OLS was applied to the models' trends to extract the linear degradation rate.

Bootstrapping was used to calculate the uncertainty of the results. Using this method, the dataset is resampled, and the degradation rate calculations are performed on that subset of the data [56]. This process is repeated many times until the distribution of results approximates the sampling distribution, allowing the 95% confidence interval to be calculated. As the OLS, CD, STL, and HW methods use time series data, the traditional bootstrap method cannot be used since it does not capture the autocorrelation of the data. A block bootstrap method, which samples blocks of data rather than individual data points, was used to correct this issue [57]. The optimal block length and bootstrap uncertainties were determined using the arch Python library [58].

The data must be sampled a large number of times in order to produce an approximately normal distribution. As this is computationally expensive, it was desired to determine the minimum number of sampling repetitions required to obtain an accurate result. To do this, uncertainty calculations were performed for an increasing number of sampling repetitions (Figure 3.7). It was found that with less than 250 repetitions, the uncertainty varies considerably. Once the number of repetitions increases to 300 or above, there is little change in the uncertainty value. Thus, the number of repetitions was set to 300.



**Figure 3.7:** Model uncertainty as a function of number of sampling repetitions used in bootstrap model.

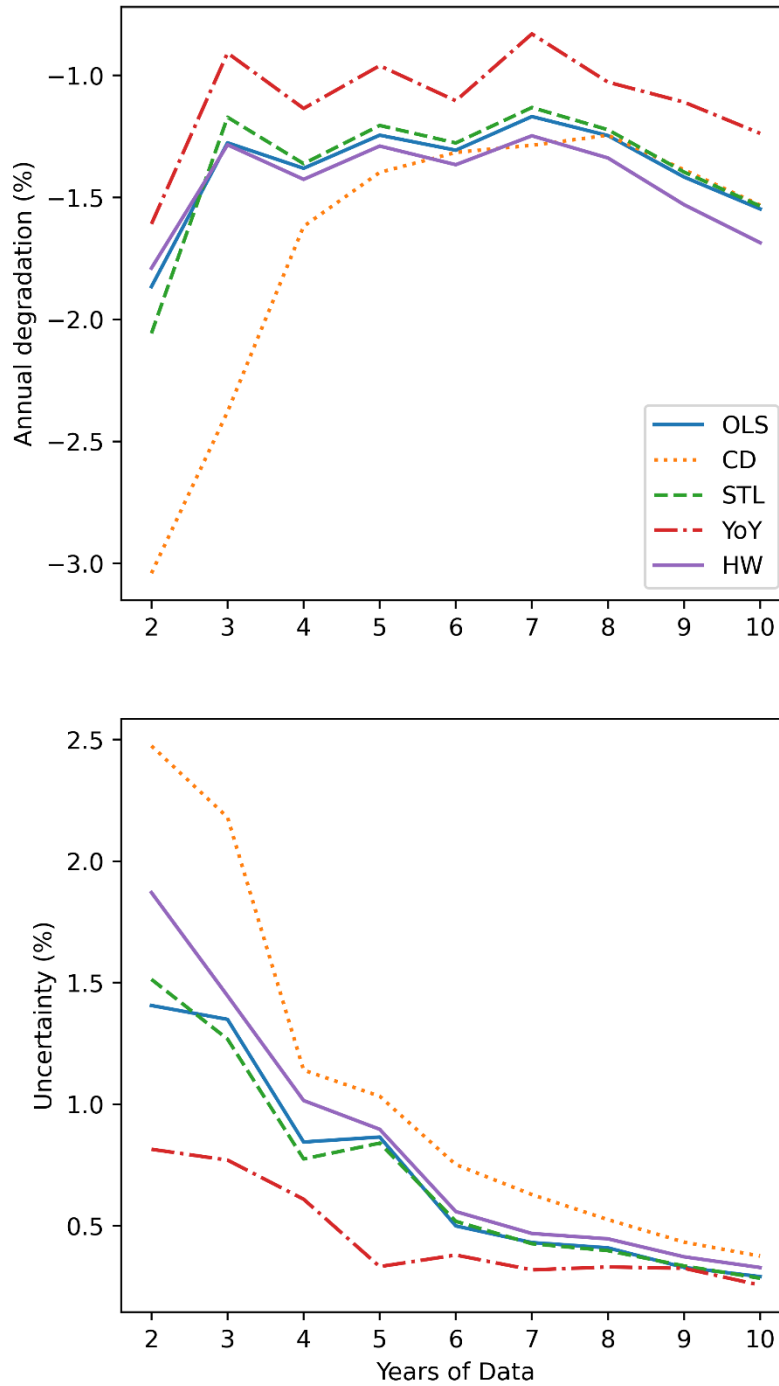
## 4 Development of Degradation Model

For the purposes of this analysis, a successful model not only has a low uncertainty but can also determine the true degradation rate with only a few years' worth of data. Most well-designed models will converge upon a single value of  $R_d$  given a large enough timespan. The challenge lies in developing a model that will converge upon that value in as short a time as possible. By observing how the modeled  $R_d$  changes as the number of years of data increases, such a model can be identified. This analysis has applications not only for the Solar House, but for any researcher attempting to determine  $R_d$  with limited data or in a shorter period of time. Even when working with a large dataset, this methodology is useful as its main goal is to reduce the uncertainty of the calculations.

Determining whether a model is accurate requires a performance dataset that spans many years. Thus, data was sourced from another solar array so that the model optimization method could be validated before being applied to the Solar House arrays. The data for this study was sourced from the Desert Knowledge Australia Solar Centre (DKA), which has an array with similar properties to the c-Si system [59]. The installation was completed in November 2009 and has been continuously monitored since that time. The array consists of 27 SunPower SPR-215 c-Si panels arranged in two rows. The ground-mounted fixed array faces due North and has a tilt angle of  $20^\circ$  from horizontal. Each PV panel has a peak power rating of 215 W, is rated at 39.8 V and 5.40 A, and has a rated efficiency of 17.3% [60]. The total system power is 5.805 kW. The array is connected to a SMA Sunny Mini Central 6000A inverter with a maximum DC power rating of 6.3 kW and a peak rated conversion efficiency of 96.1%.

Model efficacy was assessed by examining how the  $R_d$  value produced with only 2 years of data compares to that produced using the full dataset, the deviation between values produced by the different analytical methods, and the model uncertainty. Figure 4.1 shows a visualization for the c-Si array which demonstrates how  $R_d$  and uncertainty change based on the number of years of data available. With only a few years of data, the different analytical methods produce disparate values of  $R_d$ , and all result in a high level of uncertainty. As the models are given more data, the models tend to vary less with the addition of more data and converge towards a common value. In particular, the values generated by the OLS, CD, and STL methods are all in close agreement. Conversely, the YoY method yielded a  $R_d$  value that was significantly smaller than those modeled

by the other methods. Unlike the other methods, YoY is not a time series decomposition method. Because of this different approach, it is more robust against outliers. This may account for the disparity as there are some outliers not filtered out in the base model parameters.



**Figure 4.1:**  $R_d$  and uncertainty as a function of the number of years of data using the base model parameters.

Uncertainty for each model decreased dramatically with the addition of more data. The YoY method consistently yielded the lowest uncertainty, while the values for the OLS, STL, and HW methods are similar to one another. On the other hand, the CD method resulted in the highest uncertainty. This method also took the longest for the value of  $R_d$  to level out and approach that of the other models. This supports the idea that uncertainty can be taken as an indicator of model accuracy.

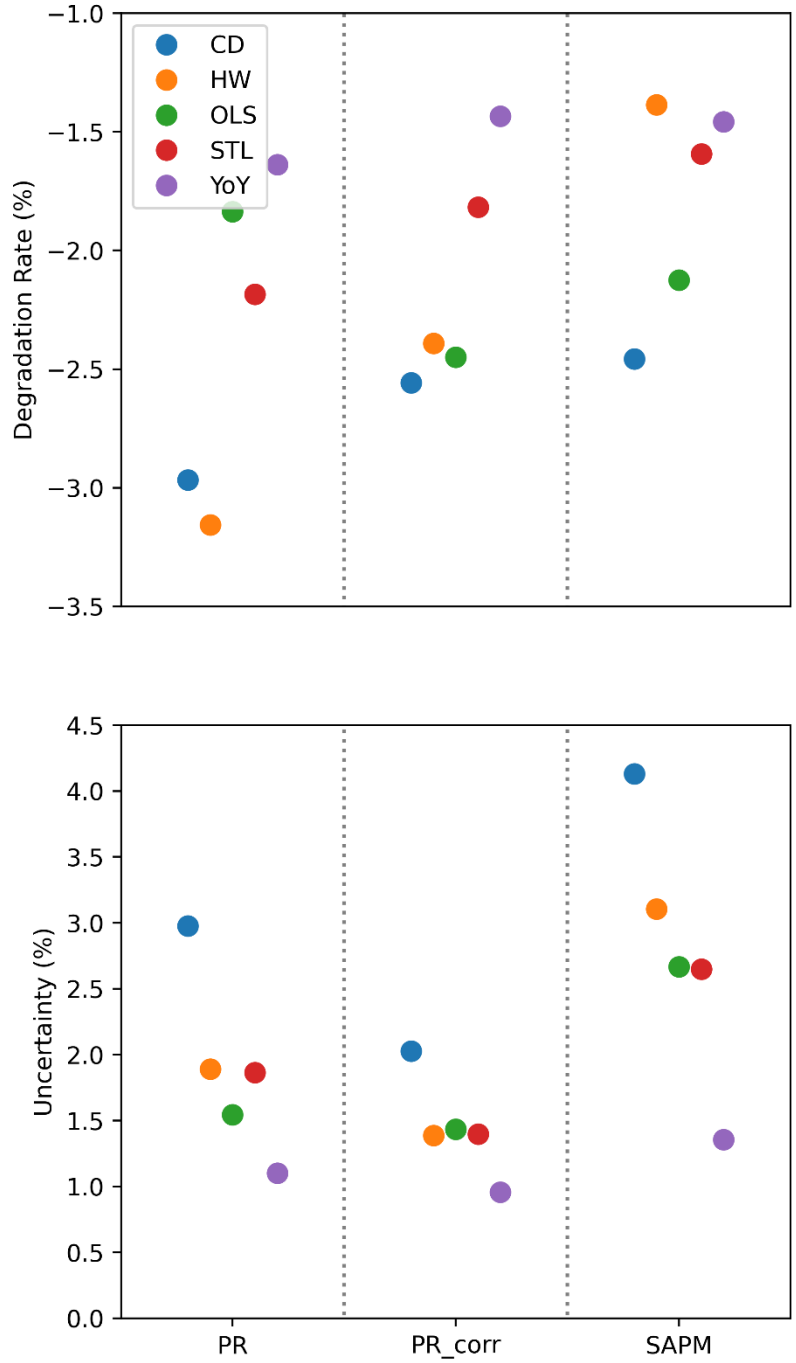
This analysis clearly shows that the base model parameters are insufficient for accurately determining  $R_d$  with less than three years of data. Given two years of data, all of the analytical models show a high degree of uncertainty. Such values are of limited use when assessing PV panel degradation. Additionally, the values produced by the different analytical methods are in disagreement with each other. With values of  $R_d$  ranging from -3.04% to -1.61%, corresponding to a standard deviation of 0.51%, the usefulness of these calculations is further thrown into question. Changes to the base parameters must be made in order to produce useful results. The following sensitivity analysis examines what affect these changes have on  $R_d$  and model uncertainty and seeks to find the optimal parameters for calculating  $R_d$ .

The first parameter to be varied was the performance metric. For each analytical method, the performance metric with the lowest uncertainty was selected. This metric was then used in the subsequent calculations to find the optimal filtering criteria. Finally, the optimal aggregation period was identified. In Sections 4.1–4.4, this process is shown for the test data. Calculations were performed using the first two years of data to better represent the available Solar House datasets. This allows for the methodology to be assessed not only for uncertainty minimization, but also for how well it approximates the “true” degradation rate with limited data.

#### 4.1 Performance Metric

The performance metrics examined in this analysis are  $PR$ ,  $PR_{corr}$ , and  $P_{N,SAPM}$  (Figure 4.2). The choice of performance metric had a significant impact on  $R_d$  for every analytical model. Using  $PR$  tended to result in a larger  $R_d$  value, while using  $P_{N,SAPM}$  resulted in a smaller one. The OLS method is the exception, as it produced the smallest value when using  $PR$  as the performance metric. There were also significant changes in the model uncertainty. Using  $P_{N,SAPM}$  resulted in the highest uncertainty, while the lowest uncertainty for every analytical methods was obtained

using  $PR_{corr}$ . Based on these results,  $PR_{corr}$  was selected for use in the subsequent analysis of the filtering criteria and aggregation period.



**Figure 4.2:** The effect of performance metric on  $R_d$  and uncertainty.

## 4.2 Filters

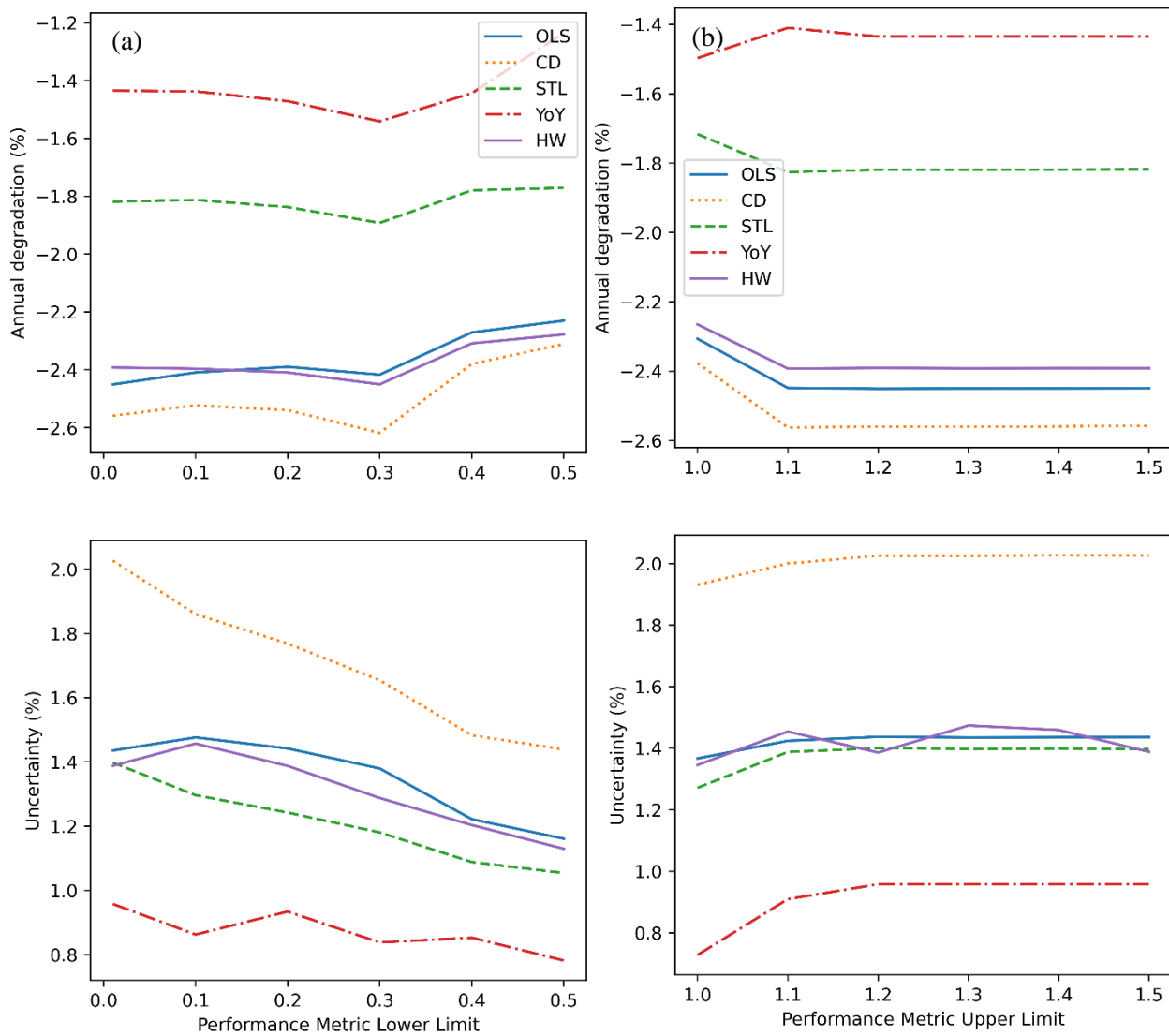
While filtering criteria can vary widely between analysts, there exist some commonalities. Filters on solar irradiance are frequently used to remove unrepresentative data [49]. Upper limits, used to remove outliers and measurement errors, are commonly set at  $1200 \text{ W/m}^2$ . There is less of a consensus on the lower limit, however. Some studies set it at  $800 \text{ W/m}^2$  in order to simulate STC, while others set the limit as low as  $100 \text{ W/m}^2$ , just enough to remove nighttime and cloudy weather data. Power filtering is also common. The filter is often set as a percentage of maximum power. Similarly, the filter can also be based on the value of the performance metric. The last filter explored in this study is a cell temperature filter. Similarly to the irradiance filter, there are a wide range of temperature limits used by researchers. While some set the filter to capture only values near STC (e.g. a range of  $23\text{--}27^\circ\text{C}$ ), others set a looser limit to filter out only the extreme outliers (e.g.  $-50\text{--}110^\circ\text{C}$ ).

There are many other filters that may be applied to PV performance data. Also commonly used is an inverter clipping filter, which removes data near the saturation limit of the inverter. However, all of the systems examined in this study have DC/AC ratios far below 1, and thus are not subject to significant inverter clipping. Other filtering criteria, such as clear-sky and stability filters, are less common and were not included in this analysis.

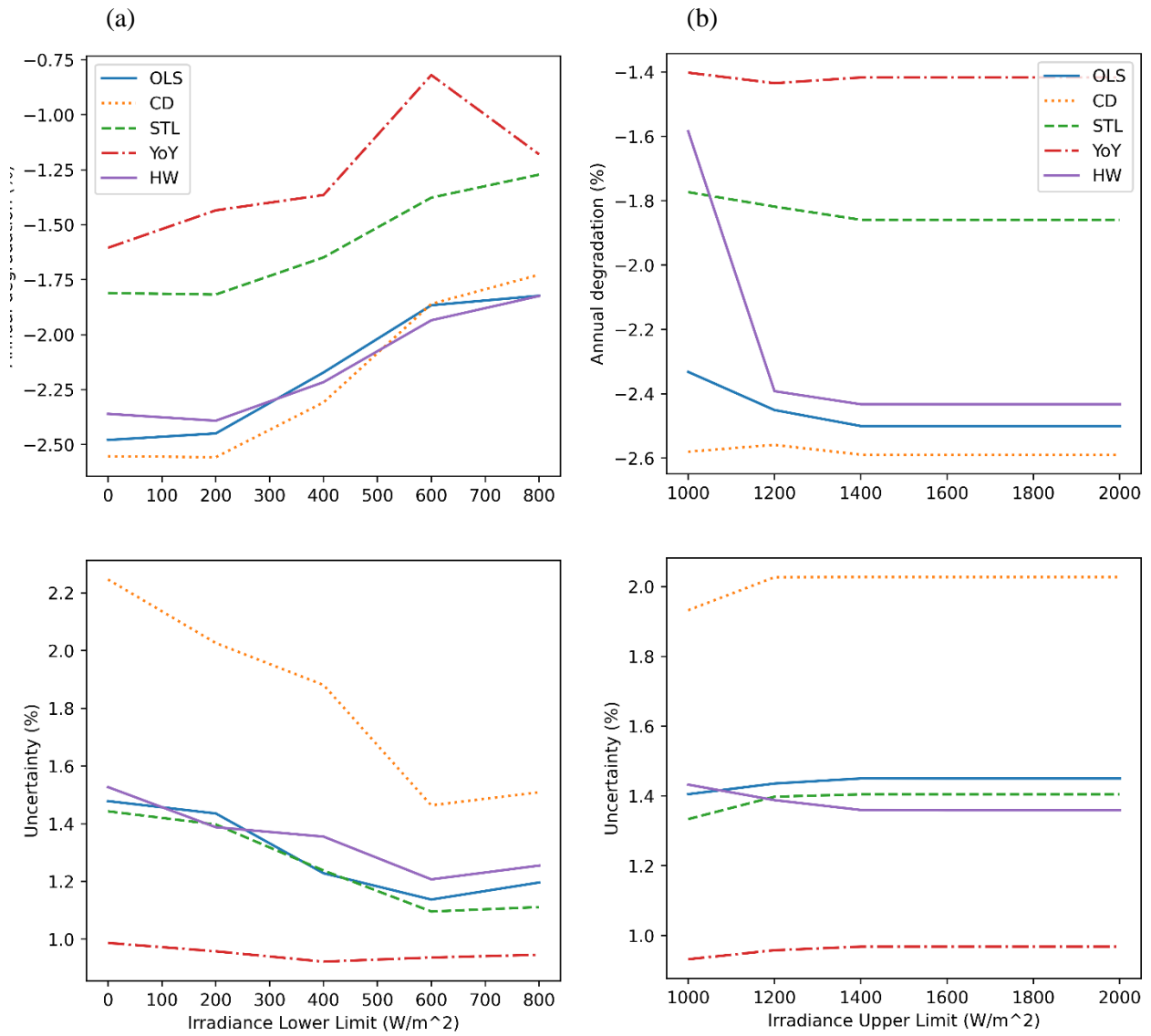
The result of varying these filters is shown in Figures 4.3–4.5. Changing the filters for the performance metric upper limit, irradiance upper limit, and temperature had very little effect on either  $R_d$  or uncertainty. This is due to the fact that there were few outliers for these datasets, making the exact bounds of the filter unimportant. The exception to this is when the strictest filtering criteria were applied. Implementing too tight of a filter generally resulted in poor results, with much higher values of uncertainty and drastically different  $R_d$  values. This indicates that a filtering regime aimed at replicating STC may not result in the most reliable degradation model.

Unlike the previously discussed filters, the lower performance metric did significantly impact results. While increasing the lower bound of this filter had little impact on  $R_d$ , it considerably reduced uncertainty. As lower values of the performance metric are often not representative of the array's typical operation, filtering them out produced more accurate results. Similarly, modifying the lower irradiance filter had a large effect on both  $R_d$  and the uncertainty. Irradiance values frequently drop below  $1000 \text{ W/m}^2$  due to weather or time of day, meaning this

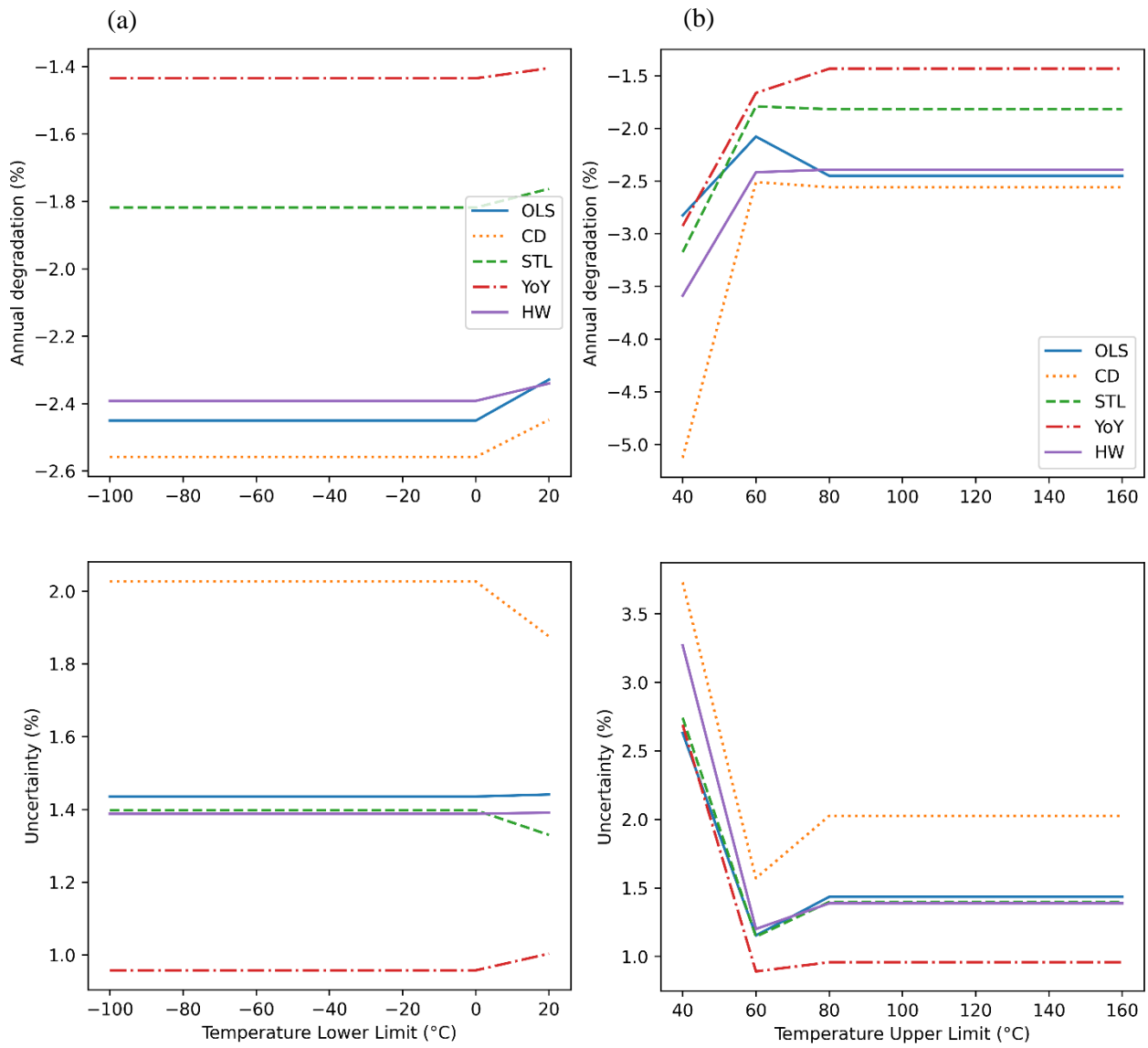
filter has the potential to remove large amounts of data and have a greater impact on the final result. Increasing the lower bound resulted in a smaller degradation rate, as the periods of poor performance caused by low irradiance were filtered out. However, raising the lower bound past a certain point also caused uncertainty to increase, as a large portion of the data was filtered out. To minimize uncertainty for these models, the proper balance must be found between removing unrepresentative data and filtering out too many datapoints.



**Figure 4.3:** The effect of changing the (a) lower and (b) upper limit of the performance metric on  $R_d$  and uncertainty.



**Figure 4.4:** The effect of changing the (a) lower and (b) upper limit of the irradiance filter on  $R_d$  and uncertainty.



**Figure 4.5:** The effect of changing the (a) lower and (b) upper limit of the temperature filter on  $R_d$  and uncertainty.

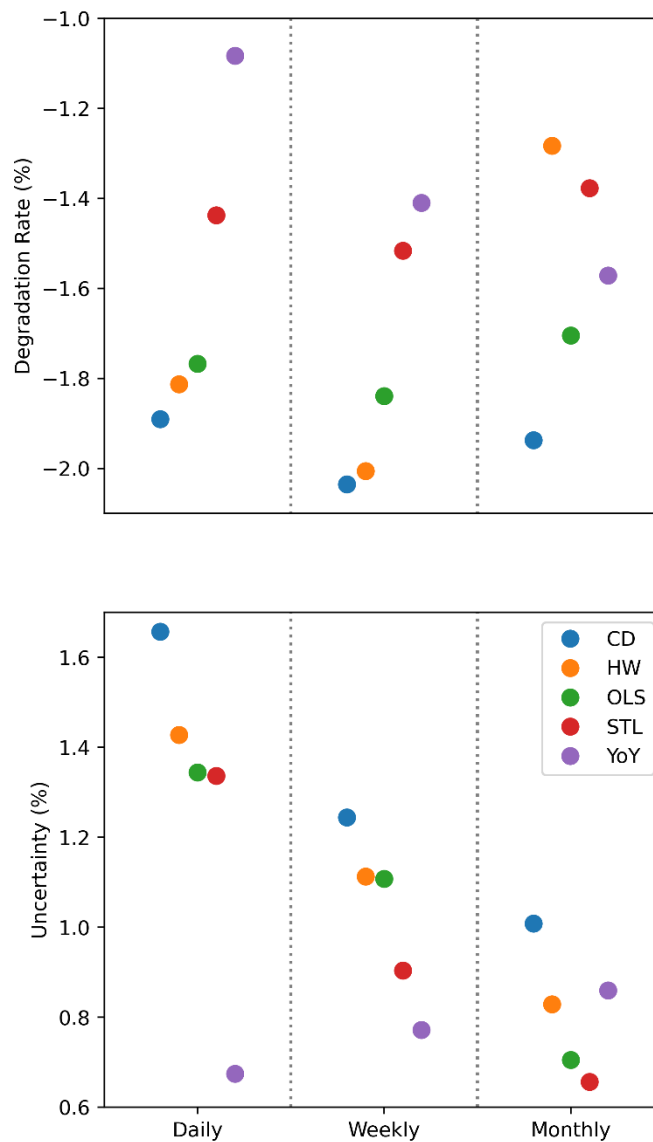
A summary of the optimal filtering criteria for each analytical method is presented in Table 4.1.

*Table 4.1: Summary of optimized filtering criteria.*

<b>Method</b>	<b>PR Lower Limit</b>	<b>PR Upper Limit</b>	<b>Irradiance Lower Limit (W/m<sup>2</sup>)</b>	<b>Irradiance Upper Limit (W/m<sup>2</sup>)</b>	<b>Cell Temperature Lower Limit (°C)</b>	<b>Cell Temperature Upper Limit (°C)</b>
<b>OLS</b>	0.5	1	600	1000	20	60
<b>CD</b>	0.5	1	600	1600	-100	60
<b>STL</b>	0.5	1	600	1000	-100	60
<b>YoY</b>	0.5	1	600	1000	-100	60
<b>HW</b>	0.5	1	400	1000	-100	60

### 4.3 Aggregation

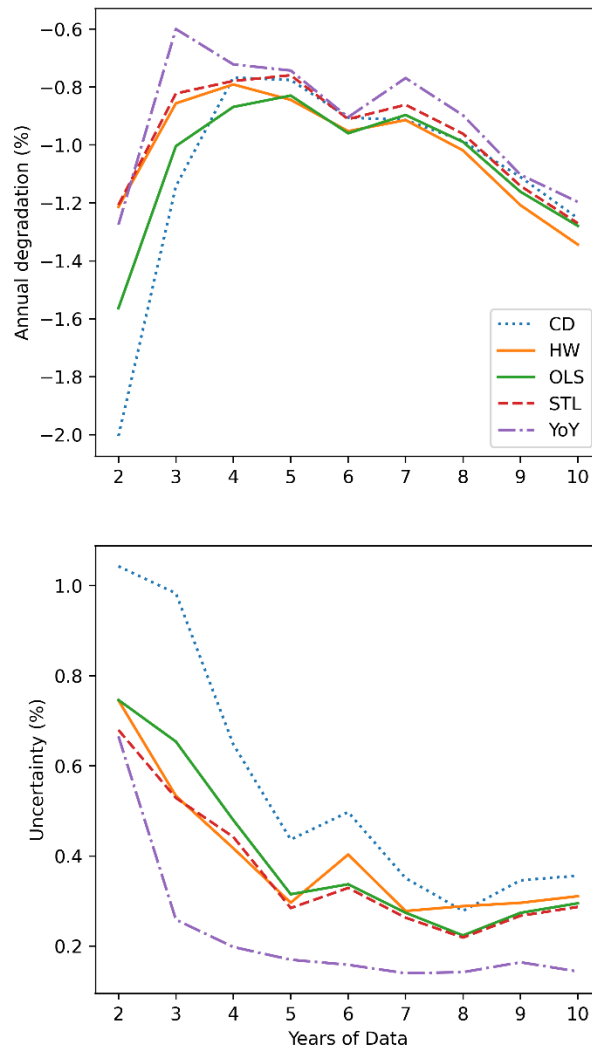
Using the parameters identified above, the data was aggregated on a daily, weekly, and monthly basis (Figure 4.6). The choice of aggregation period had a significant effect on uncertainty. For the YoY method, the daily aggregation period produced the lowest uncertainty. This method is capable of accounting for seasonality and outliers, and is not affected by variability as long as the data is normally distributed. Thus, it benefited from an increased number of datapoints. The other methods performed better with a monthly aggregation period as it reduced variability.



**Figure 4.6:** The effect of aggregation period on  $R_d$  and uncertainty.

#### 4.4 Selection of Modeling Criteria

Combining the modeling criteria developed in Sections 4.1-4.3, new optimized models were built and assessed using the same criteria as for the base model (Figure 4.7). Using the optimized modeling criteria, the different analytical models converged with fewer years of data and were in better agreement at the two year mark compared to the base model. Additionally, the model uncertainty was much lower across the board. This data shows that the above method for selecting modeling criteria is effective at determining  $R_d$  with a higher degree of accuracy and in a shorter period of time than the base model. This allows for the degradation of the arrays on the Solar House to be assessed with a much higher degree of confidence.



**Figure 4.7:**  $R_d$  and uncertainty as a function of the number of years of data using the optimized model parameters.

Table 4.2 shows the results for the base and optimized models. In addition to the uncertainty being reduced for each analytical method when using the optimized models, there was also closer agreement between the different methods. The range of  $R_d$  values dropped from 0.633% to only 0.105%, and the standard deviation was reduced from 0.219% to 0.038%. This effect can also be seen for the shorter observation period. Given two years of data, the base models produced  $R_d$  values with a range of 1.432% and a standard deviation of 0.505%. Comparatively, the optimized models produced  $R_d$  values with a range of 0.798% and a standard deviation of 0.305%. The optimization process also resulted in  $R_d$  values that more closely match those produced using the full dataset. This can be seen by taking the difference between the  $R_d$  for the full dataset and for the two year dataset ( $R_{d,full} - R_{d,2}$ ).

**Table 4.2:** Summary of degradation model results for c-Si test data.

Method	Model	Full Dataset		First Two Years		$R_{d,full} - R_{d,2}$
		$R_d$ (%)	Uncertainty (%)	$R_d$ (%)	Uncertainty (%)	
CD	Base	-1.35	0.347	-1.87	1.41	0.517
	Optimized	-1.11	0.346	-2.00	1.04	-0.894
HW	Base	-1.69	0.344	-3.04	2.48	-1.355
	Optimized	-1.21	0.296	-1.21	0.745	-0.006
OLS	Base	-1.55	0.299	-2.06	1.51	-0.510
	Optimized	-1.16	0.274	-1.56	0.746	-0.402
STL	Base	-1.54	0.293	-1.61	0.815	-0.073
	Optimized	-1.14	0.267	-1.21	0.680	-0.066
YoY	Base	-1.05	0.199	-1.79	1.87	-0.738
	Optimized	-1.10	0.164	-1.28	0.666	-0.173

Of the analytical methods evaluated in this study, the YoY method produced the lowest uncertainty given two years of data, with an  $R_d$  value of  $-1.28 \pm 0.67\%$ . The HW and STL methods produced slightly higher uncertainties but were still able to accurately predict  $R_d$  given two years of data. The OLS method resulted in a comparable uncertainty value, but with a  $R_{d,full} - R_{d,2}$  value of -0.4, it was not accurate in calculating  $R_d$  given limited data. The CD method yielded the

highest uncertainty value. Additionally, there is a large discrepancy between the  $R_d$  calculated using two years of data and the values produced by the other methods, as well as the value calculated using the full dataset. This demonstrates that the level of uncertainty can be used to identify whether a model based on limited data is accurate.

## 5 Comparison of a-Si and c-Si Arrays

### 5.1 Dependence of Performance on Weather Conditions

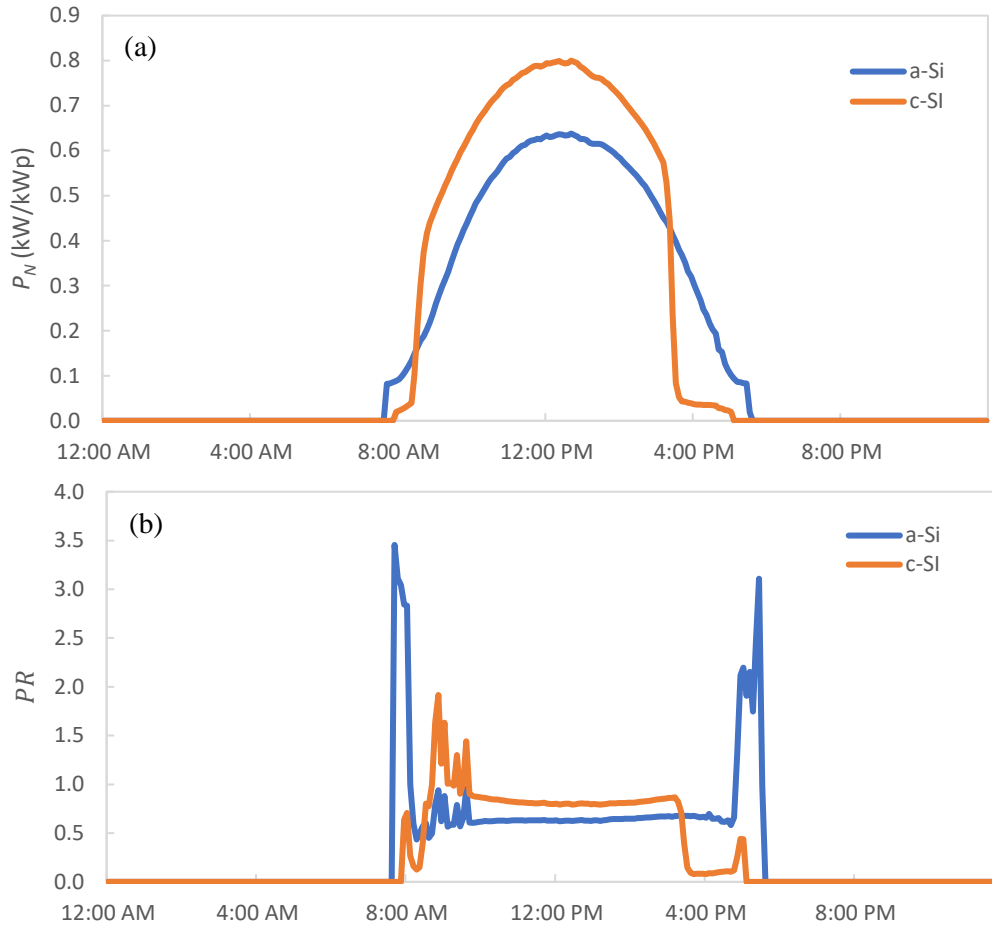
Solar panel performance is strongly dependent on the weather. Sky conditions change not only the amount of solar irradiance available, but also the diffusion and spectral distribution of the available light. Understanding this weather dependence is essential for accurately predicting how a PV system will behave throughout the year.

Two metrics were used to compare the weather-dependent performance of the a-Si and c-Si panels: normalized power and  $PR$ . The normalized power is defined as

$$P_N = \frac{P_{AC}}{P_{rated}} \quad (13)$$

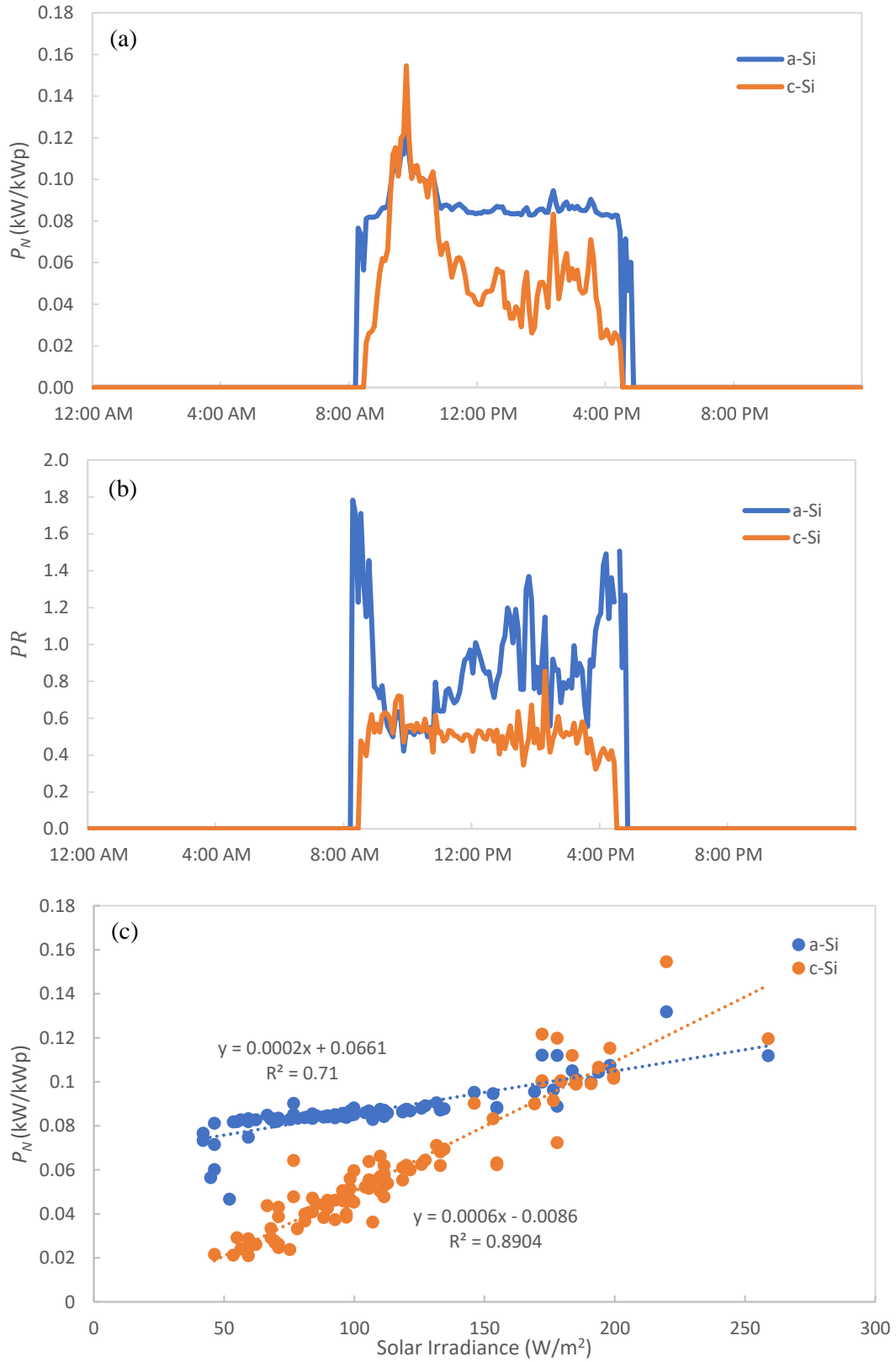
where  $P_{rated}$  is the rated power of the array. Since the c-Si panels have a larger rated power, they nearly always generate more energy than the a-Si panels. Normalizing the power allows for a more direct comparison of the panels relative to their ideal performance.

The results for a sunny day (February 9<sup>th</sup>, 2020) are shown in Figure 5.1. The normalized power plot shows that the c-Si panels have a higher peak power and outperform the a-Si panels for most of the day. This is supported by the  $PR$  plot, which reveals a higher  $PR$  for the c-Si panels than the a-Si panels. This is not surprising as the c-Si array is newer and has not had as much time to degrade. Despite this, the a-Si panels had the better performance during the morning and evening. During these times, the a-Si panels have a higher normalized power than the c-Si panels. They also generate energy for a longer period of time. The  $PR$  in the morning and evening spikes sharply, reaching values above 3. The pyranometer is mounted horizontally, rather than tilted to the plane of the array. As a result, it may not be able to capture as much irradiance as the arrays do when the sun is low in the sky, causing  $PR$  to be overestimated. Regardless, this demonstrates that the a-Si panels are capable of utilizing a wider spectrum of light than the c-Si panels.



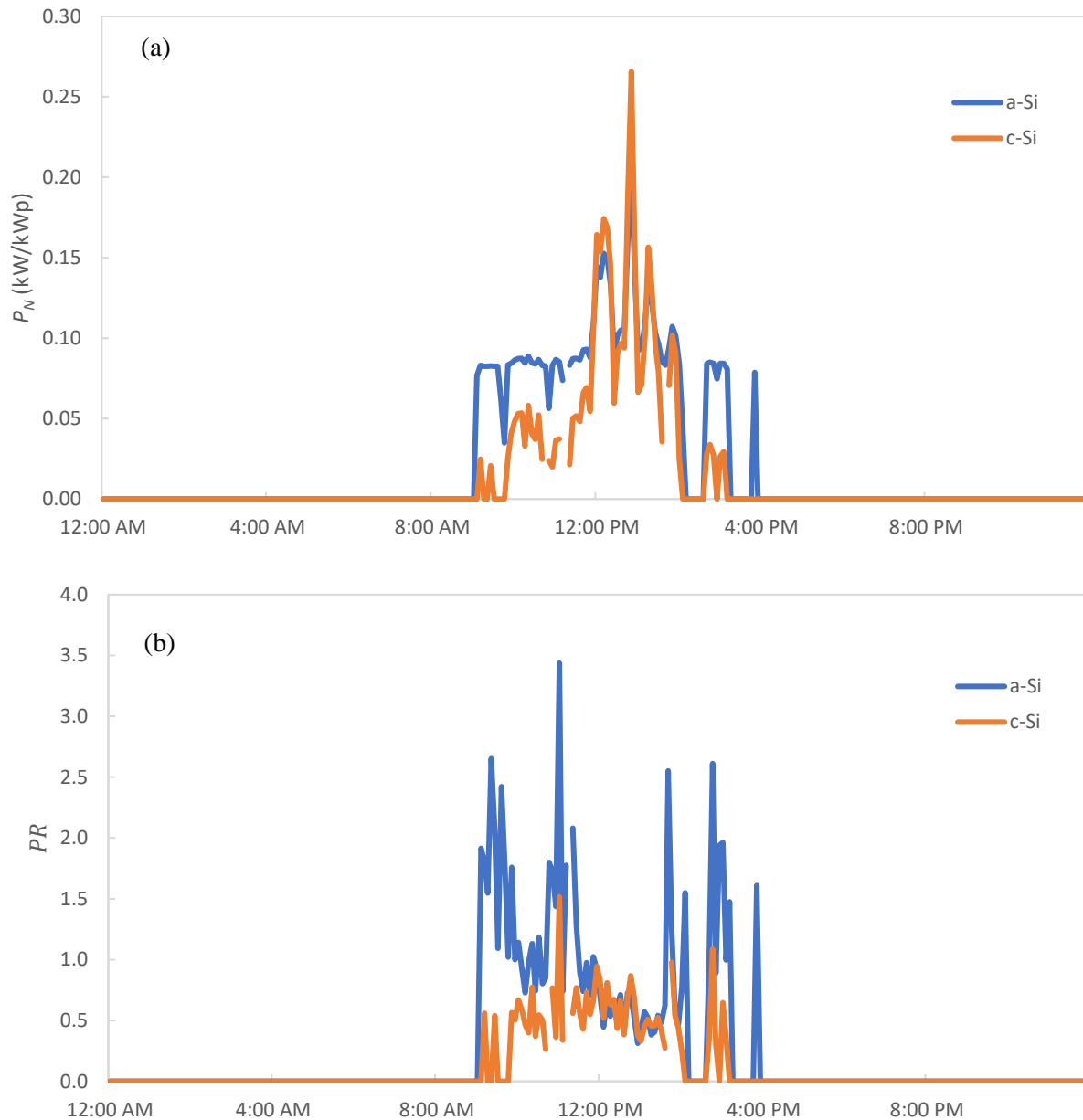
**Figure 5.1:** The (a)  $P_N$  and (b)  $PR$  of the two solar panel systems on a sunny day.

The results for a cloudy day (February 1st, 2020) are shown in Figure 5.2. As the clouds block solar radiation, power output for both arrays was significantly reduced. Unlike the sunny day case, the normalized power plots do not follow a smooth curve. The c-Si power fluctuated throughout the day, while the a-Si  $P_N$  remained relatively constant at around 0.08. Conversely, the  $PR$  for the c-Si panels did not vary much throughout the day, while the a-Si panels experienced large swings. These differences can be explained by examining the relationship between power generation and solar irradiance. The linear regression model of the cloudy day data reveals that the slope of the normalized power versus solar irradiance function is much steeper for the c-Si panels, indicating that the performance of the c-Si panels is much more dependent on fluctuations in the irradiance. This also demonstrates the a-Si panels' ability to excel in low-light conditions. On a  $P_N$  basis, they outperformed the c-Si panels when irradiance is less than approximately  $185 \text{ W/m}^2$ .



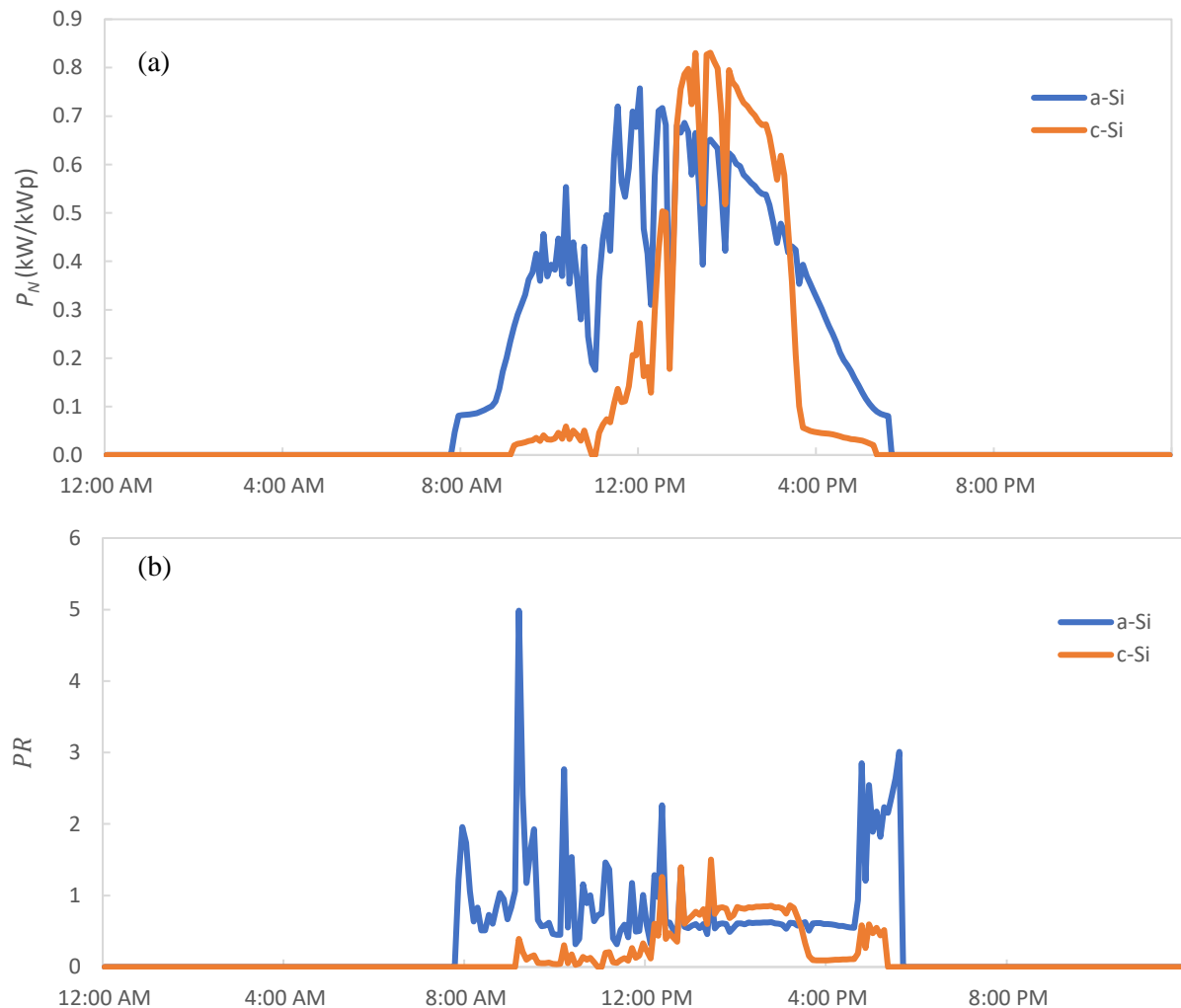
**Figure 5.2:** The (a)  $P_N$  and (b) PR of the two solar panel systems on a cloudy day. (c) The correlation of  $P_N$  with solar irradiance.

The rainy day performance of the panels (February 6th, 2020), shown in Figure 5.3, followed similar trends to the cloudy day performance data. Again, the a-Si panels generated more consistent power, while the c-Si panels had a more consistent  $PR$ . However, the performance of both systems varied more throughout the day. This is likely because there was more variation in the solar irradiance.



**Figure 5.3:** The (a)  $P_N$  and (b)  $PR$  of the two solar panel systems on a rainy day.

Last is a partly cloudy day with snow on the ground (February 21st, 2020), shown in Figure 5.4. The c-Si panels had a low  $PR$  and generated almost no energy before noon on this day. In contrast, the a-Si panel performance was similar to the sunny day performance, although with more variation. This can be partially attributed to the different mounting configurations of the two systems. The a-Si panels are laminated onto the metal roof of the garage, allowing them to absorb more heat from the roof and melt the snow faster. On the other hand, the c-Si panels are closely mounted to the roof, but not in direct contact. Thus, they absorbed less heat and the snow melted slowly. Additionally, the a-Si panels picked up more of the light reflected off of the snow. The c-Si panels are less capable of picking up this diffuse light, and so could not generate as much energy in the morning. In the afternoon, the snow began to melt and the solar irradiance increased, causing the panels to match their expected performance more closely.



**Figure 5.4:** The (a)  $P_N$  and (b)  $PR$  of the two solar panel systems on a snowy day.

Table 5.1 and Table 5.2 summarize the systems' performance for each day. Both systems performed well on the sunny days, with high values for PR and  $P_N$ . The c-Si panels produced much more energy, which is to be expected given their larger rated power. They also had a larger peak power, which reflects how their power fluctuated more throughout the day. Comparing the PRs for each system, the c-Si panels showed better performance. This is likely because the a-Si panels are five years older and thus have degraded more.

On the cloudy and rainy days,  $P_N$  for both systems was much lower as there was less solar energy available. The c-Si panels also had a much lower PR and generated very little energy. Conversely, the a-Si panels had excellent performance, producing nearly as much energy as the c-Si panels despite their lower capacity. They had a very high PR on both of these days. On the rainy day, the PR exceeded a value of 1, meaning the panels surpassed their rated efficiency. Likewise, the PR for the snowy day is close to 1. This is a testament to the a-Si panels' ability to efficiently convert diffuse light to electricity.

**Table 5.1:** Performance of a-Si panels under different weather conditions.

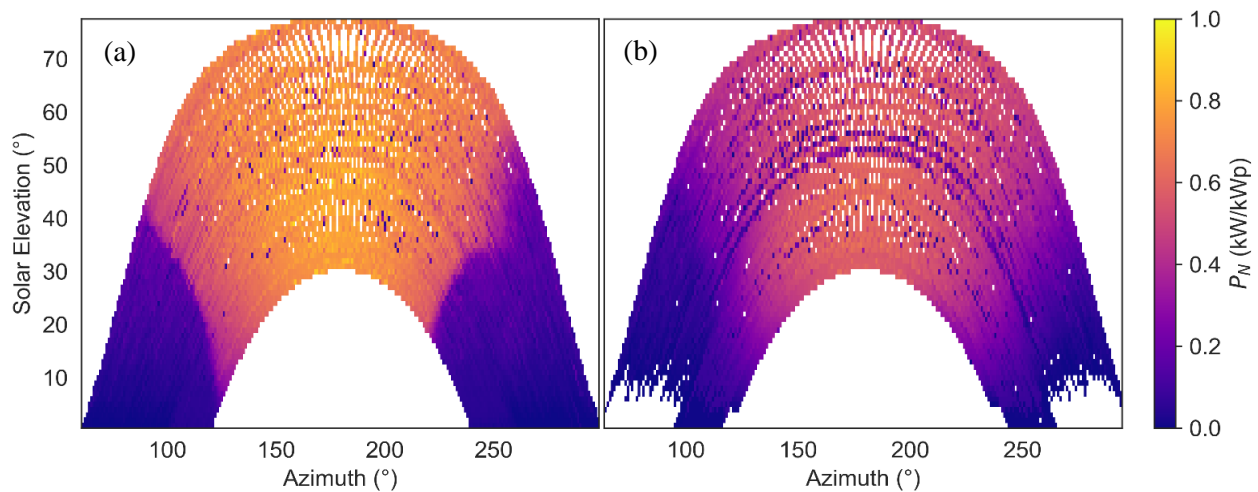
Weather Condition	Average $\eta_p$	Average $P_N$ (kW/kWp)	Maximum $P_N$ (kW/kWp)	Average PR	Energy Produced (kWh)
Sunny	5.0%	0.42	0.64	0.85	12.68
Cloudy	6.4%	0.09	0.13	0.89	2.28
Rainy	7.3%	0.09	0.21	1.14	1.65
Snowy	5.1%	0.38	0.76	0.95	11.54

**Table 5.2:** Performance of c-Si panels under different weather conditions.

Weather Condition	Average $\eta_p$	Average $P_N$ (kW/kWp)	Maximum $P_N$ (kW/kWp)	Average PR	Energy Produced (kWh)
Sunny	16.8%	0.52	0.80	0.72	25.75
Cloudy	11.9%	0.06	0.15	0.51	2.59
Rainy	16.0%	0.07	0.27	0.57	1.87
Snowy	10.9%	0.29	0.83	0.42	13.10

## 5.2 Dependence of Performance on Solar Position

Figure 5.5 visualizes the relationship between  $P_N$  and the sun's position in the sky over the course of a year. The position of the sun is represented by its angle from north (azimuth) and its angle from the horizon (solar elevation). The arrays produced the maximum amount of power when the sun is positioned such that its rays are perpendicular to the plane of the array. As the arrays are oriented to the South, power was maximized when the azimuth was equal to  $180^\circ$ , or due South. Similarly,  $P_N$  was greatest when the solar elevation was equal to the array tilt, or  $36^\circ$ . The further the sun moved from this position, the more  $P_N$  decreased.



**Figure 5.5:**  $P_N$  as a function of solar position for the (a) c-Si and (b) a-Si arrays.

The plot for the c-Si array is asymmetrical, with more power loss occurring when the sun was to the west. This indicates that the panels experienced shading from the western side which degrades their performance. As shown in Figure 5.6, this shading is due to a tree which is located near the panels [61]. The a-Si plot is also asymmetrical as the panels experienced shading from the trees to the east of the array.



*Figure 5.6: Shading from a tree on the western side of the Solar House.*

### 5.3 Degradation Rates

Repeating the steps outlined in Chapter 4, degradation models were developed for the a-Si and c-Si Solar House arrays. The optimized parameters for both are shown in Table 5.3 and Table 5.4. For both arrays, the optimal modeling parameters do not vary much between the analytical methods. The exception is the YoY method, which largely performs better with less restrictive filtering criteria. This is because it is resistant to outliers, and therefore benefits from the inclusion of more data points.

*Table 5.3: The optimized modeling parameters for the c-Si array.*

Method	CD	HW	OLS	STL	YoY
Performance Metric	$PR_{corr}$	$PR_{corr}$	$PR_{corr}$	$PR_{corr}$	$PR_{corr}$
Performance Metric Lower Limit	0.5	0.5	0.5	0.5	0.5
Performance Metric Upper Limit	1.2	1.2	1.2	1.2	1.1
Irradiance Lower Limit (W/m <sup>2</sup> )	200	800	800	800	200
Irradiance Upper Limit (W/m <sup>2</sup> )	1400	1200	1200	1400	1800
Cell Temperature Lower Limit (°C)	-100	-100	-100	-100	-100
Cell Temperature Upper Limit (°C)	100	100	100	100	100
Aggregation	Daily	Monthly	Monthly	Monthly	Daily

*Table 5.4: The optimized modeling parameters for the a-Si array.*

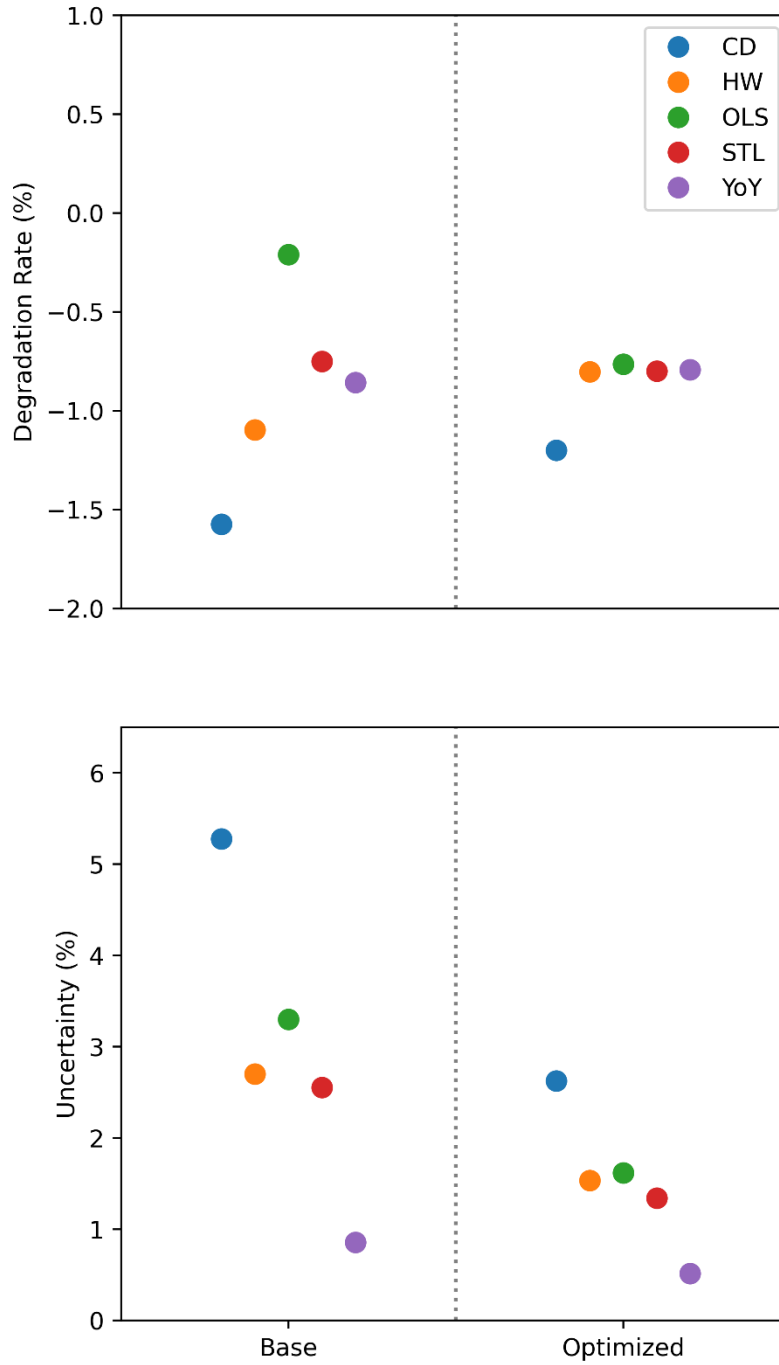
Method	CD	HW	OLS	STL	YoY
Performance Metric	$PR_{corr}$	$PR_{corr}$	$PR_{corr}$	$PR_{corr}$	$P_{N,SAPM}$
Performance Metric Lower Limit	0.1	0.5	0.4	0.1	0.3
Performance Metric Upper Limit	1	1.1	1.3	1.3	1
Irradiance Lower Limit (W/m <sup>2</sup> )	800	800	800	800	200
Irradiance Upper Limit (W/m <sup>2</sup> )	1200	1400	1200	1200	1200
Cell Temperature Lower Limit (°C)	20	20	20	20	-100
Cell Temperature Upper Limit (°C)	100	120	100	100	100
Aggregation	Weekly	Weekly	Weekly	Weekly	Monthly

The optimal parameters vary significantly between the two arrays. Generally, the lower limits for performance metric are higher for the c-Si array compared to the a-Si array. As c-Si panels perform worse than a-Si panels under low light conditions, these higher filters are necessary to remove datapoints that are unrepresentative of the array's typical performance. Conversely, the limits on the temperature filter were narrower for the a-Si array. This is counter-intuitive as the a-Si panels have a smaller rated temperature coefficient and should not be as affected by temperature as the c-Si panels. That the optimal temperature filter range is smaller indicates that the a-Si panels may have a larger temperature coefficient than given in their data sheet. Mahmood et. al. established that temperature coefficients can increase over time [62]. As the a-Si array is older by five years, its temperature coefficient has likely surpassed that of the c-Si array, necessitating a smaller temperature filter range.

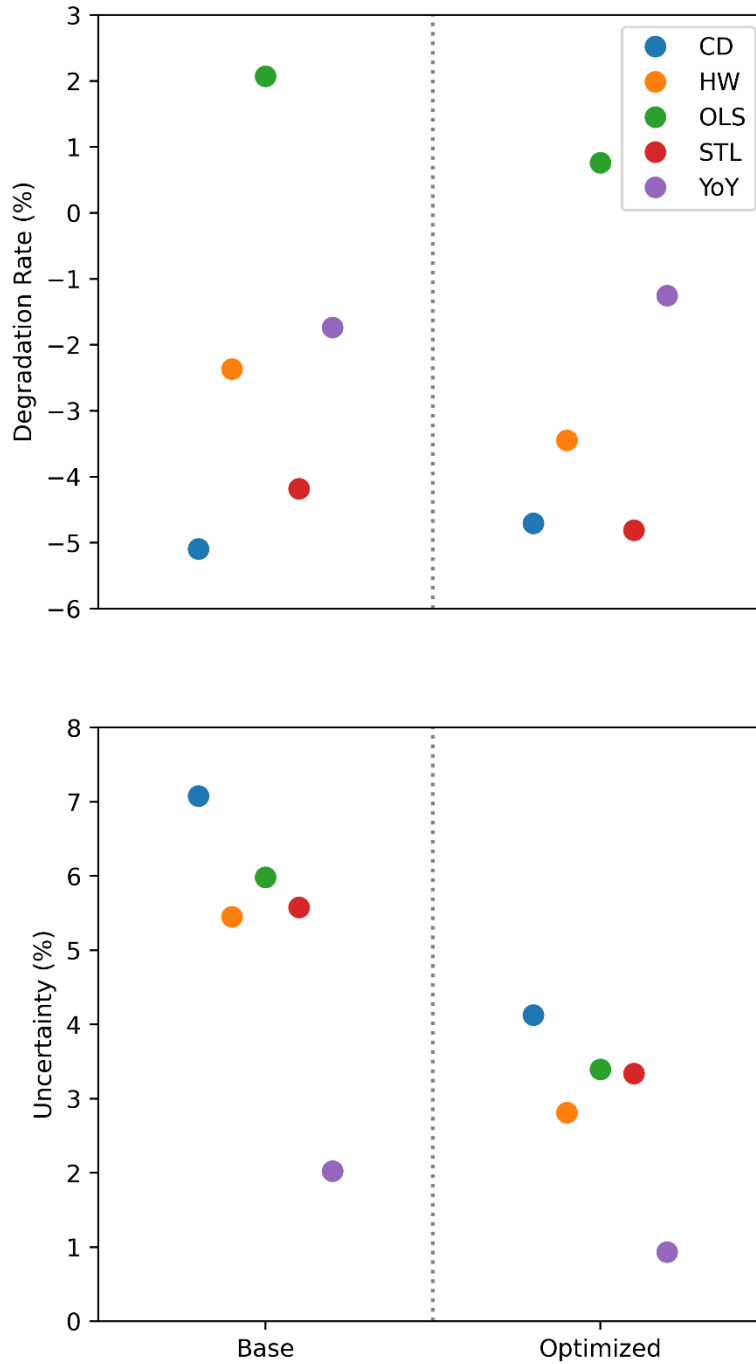
The optimal aggregation period also differs between the two arrays. With the exception of the CD and YoY method, monthly aggregation was optimal for the c-Si array, while weekly aggregation offered better results for the a-Si array. This is likely due to the differences in data collection. With 2.5 years of data collected for the c-Si array, there was enough data available to make monthly aggregation feasible. As a larger aggregation period reduces variability, aggregating the data by month lowered the model uncertainty. On the other hand, the a-Si array has less than 1.5 years of data available. Monthly aggregation would result in a very small number of data points, making weekly aggregation a better choice for building a robust model.

The optimal modeling parameters differ not only between the a-Si and c-Si arrays, but also between the c-Si array and the comparable test array. This indicates that the model optimization process is specific to the individual array. The optimized parameters developed for one array cannot be assumed to be the optimal parameters for another, even if both arrays use the same type of PV panel.

Figure 5.7 and Figure 5.8 show the result of using these optimized model parameters compared to the base model. Uncertainty decreased in each case, sometimes by a large amount. Most notably, the uncertainty for the c-Si array when using the CD and OLS methods was more than halved. The models also were generally in closer agreement after optimization. In particular, the HW, OLS, STL, YoY models were in very close agreement for the c-Si array. The standard deviation of  $R_d$  values for this array dropped from 0.45% to 0.16%. The standard deviation for the a-Si array remained relatively high at 2.15%, but was still reduced from the initial value of 2.48%.



*Figure 5.7:  $R_d$  and uncertainty for the c-Si array using the base and optimized model parameters.*



**Figure 5.8:**  $R_d$  and uncertainty for the a-Si array using the base and optimized model parameters.

A summary of the above data is presented in Table 5.5 and Table 5.6. For both arrays, the YoY method resulted in the lowest uncertainty. Using this method, the degradation rates of the c-Si and a-Si arrays are -0.79% and -1.26%, respectively. While both are higher than the rule-of-

thumb rate of -0.5%, they align more closely with the median values for those technologies found in other studies (-0.91% for c-Si and -1.38% for thin-film panels) [16]. Such similar values suggest that reasonable results can be obtained for observation periods shorter than 2 years. However, the results of the other methods must also be examined. Using these methods,  $R_d$  for the c-Si array ranges from -0.76% to -1.20%. The range of  $R_d$  values for the a-Si array is considerably higher as a result of the shorter observation period, varying between 0.76% and -4.81%. The uncertainty values also vary widely, ranging as high as 4.13%. This demonstrates the importance of properly selecting modeling parameters and analytical methods, and shows the utility of this model optimization process.

*Table 5.5: Summary of degradation model results for c-Si array.*

Method	Model	$R_d$ (%)	Uncertainty (%)
CD	Base	-1.57	5.28
	Optimized	-1.20	2.62
HW	Base	-1.10	2.70
	Optimized	-0.80	1.53
OLS	Base	-0.21	3.30
	Optimized	-0.76	1.62
STL	Base	-0.75	2.55
	Optimized	-0.80	1.34
YoY	Base	-0.86	0.85
	Optimized	-0.79	0.52

*Table 5.6: Summary of degradation model results for a-Si array.*

Method	Model	$R_d$ (%)	Uncertainty (%)
CD	Base	-5.09	7.08
	Optimized	-4.70	4.13
HW	Base	-2.36	5.45
	Optimized	-3.45	2.81
OLS	Base	2.07	5.98
	Optimized	0.76	3.39
STL	Base	-4.18	5.58
	Optimized	-4.81	3.34
YoY	Base	-1.74	2.02
	Optimized	-1.26	0.93

Although the model optimization universally decreased model uncertainty, the level of uncertainty is still very high even for the YoY method. This is also reflected in the wide range of  $R_d$  values produced by the different models. Therefore, while this process produces results with a much higher degree of confidence, it is not sufficient to definitively establish the degradation rate of the arrays given the short observation period. More data is necessary to properly assess the accuracy of these models and to make a comparison between the two arrays.

### 5.3.1 Comparison to Previous Studies

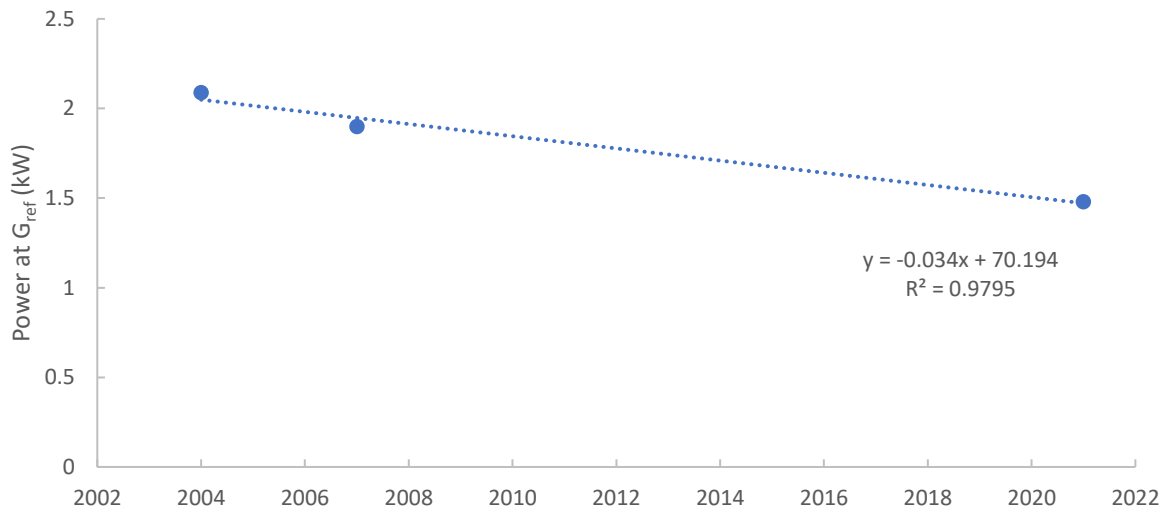
While continuous data is not available for most of the lifetime of the PV systems, previous studies of the Solar House contain additional information that can shed light on the systems' long term degradation rates. In his graduate thesis, Daniel Christy analyzed the degradation of the a-Si array from 2003 to 2007 [9]. This study compared the array's performance from December 2, 2003-March 2, 2004 with its performance during the same period in 2006-2007. More recent data, from December 2, 2020-March 2, 2021, was used to compare these figures with the current array performance. One performance metric used in Christy's thesis was the power generated at the reference irradiance ( $G_{ref}$ ) of 1000 W/m<sup>2</sup>. For this study, the power was averaged across all datapoints whose irradiance value was 1000±10 W/m<sup>2</sup>. Weather data used to calculate the cell temperature and average irradiance was sourced from two local weather stations whose data is available from the North Carolina State Climate Office [63]. These values were used to calculate the average  $PR_{corr}$  for each available year. A summary of this data is shown in Table 5.7. Note that the values for  $G$  and  $PR$  are low as the original study did not filter out nighttime data.

**Table 5.7:** a-Si array performance data from 2003-2021.

Years	Average $G$ (W/m <sup>2</sup> )	Average $T_c$ (°C)	Power at $G_{ref}$ (kW)	Daily Yield (kWh/day)	Average $PR_{corr}$
<b>2003-2004</b>	369.79	25.66	2.09	8.54	0.077
<b>2006-2007</b>	322.42	25.61	1.90	7.79	0.080
<b>2020-2021</b>	323.52	24.29	1.48	5.86	0.062

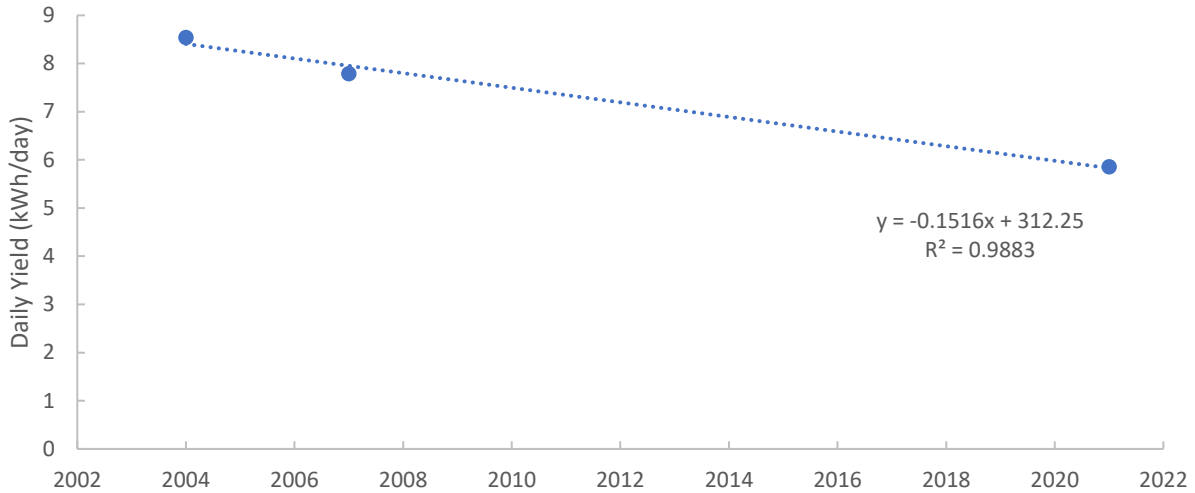
Linear regression models were fit to each of the above parameters in order to determine  $R_d$ . Using the power at  $G_{ref}$  as the performance metric,  $R_d$  is equal to -1.63% (Figure 5.9). This value most closely matches that obtained by the YoY method, which provides evidence that this method is best suited for modeling the degradation of the a-Si array. As the values obtained using the YoY method also had the lowest uncertainty, this also lends credence to the idea that minimizing uncertainty is a valid approach for obtaining the most accurate value of  $R_d$ .

This degradation rate is 0.37% higher than the YoY value obtained using the model optimization method described previously. However, this is to be expected as it has been shown that degradation rates obtained using only a few data points tend to be much larger than those obtained using higher quality data [16]. It is also important to note that these calculations do not account for temperature effects. However, as all measurements were taken during the same time of year and under the same weather conditions (a clear day during peak sunlight hours), these variations are likely minimal.



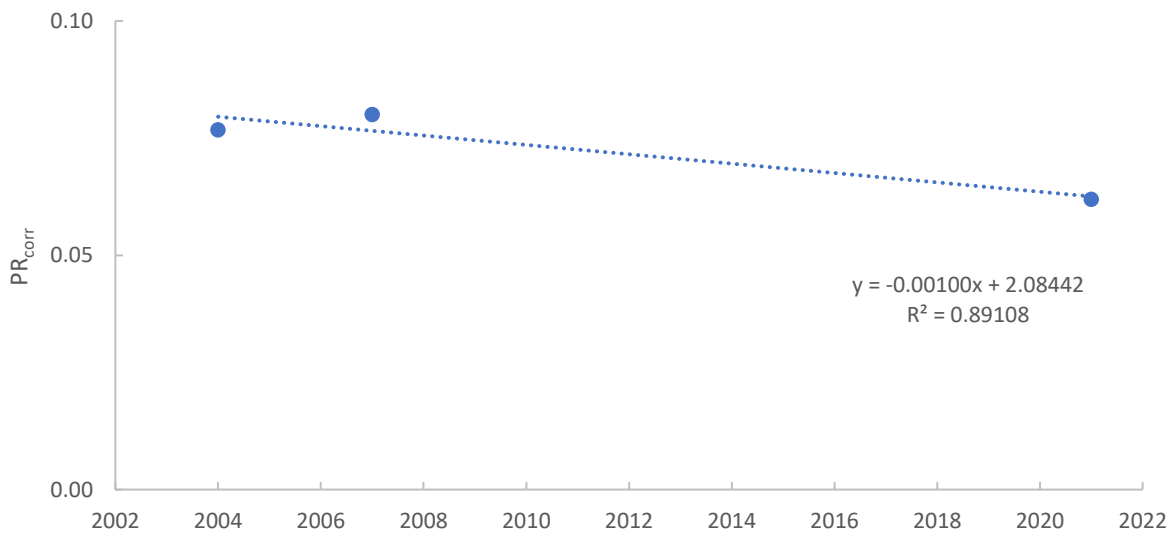
**Figure 5.9:**  $R_d$  for the a-Si array based on the power at  $G_{ref}$ .

Using the daily yield,  $R_d$  is -1.78% (Figure 5.10). This is in fairly close agreement with the values obtained using the power at  $G_{ref}$  method. One limitation of using the daily yield is that it does not account for differences in temperature or average  $G$ , which has a significant impact on power output. From Table 5.7, it can be seen that  $G$  decreases over time, causing this method to somewhat overestimate the degradation rate.



**Figure 5.10:**  $R_d$  for the a-Si array based on the daily yield.

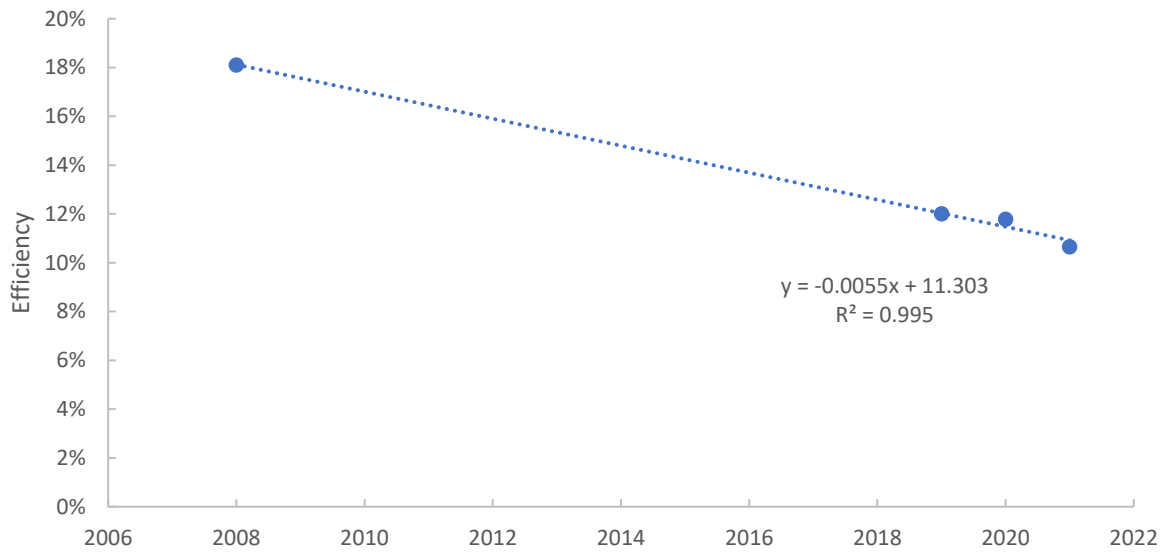
To account for this change in average  $G$ ,  $PR_{corr}$  was calculated for each datapoint (Figure 5.11). Using  $PR_{corr}$  to calculate  $R_d$  results in a value of -1.30%. As expected, this is considerably lower than the  $R_d$  calculated using the power at  $G_{ref}$  and the daily yield as it accounts for both variations in solar irradiation and temperature. This metric is the most accurate of the three as it takes into account these weather effects. There is a difference of only 0.04% between this value and that obtained in the previous section. This provides supporting evidence for the validity of the model optimization method developed in this work, and allows the  $R_d$  value for the a-Si array to be stated with a higher degree of certainty.



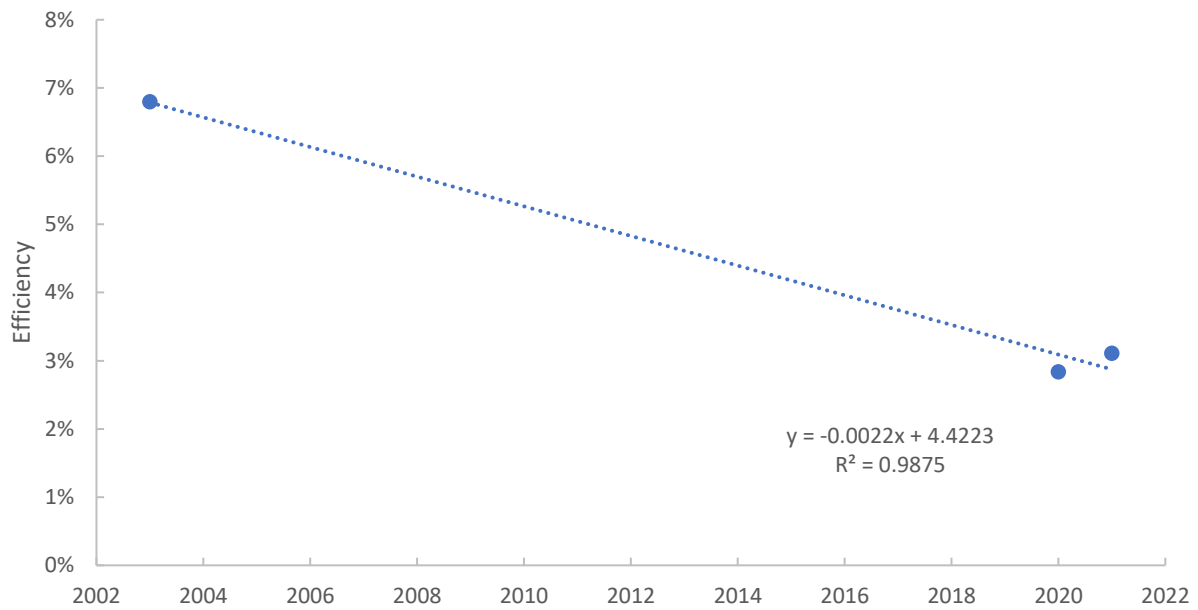
**Figure 5.11:**  $R_d$  for the a-Si array based on  $PR_{corr}$ .

Comparable data is unfortunately not available for the c-Si array. However, a similar analysis can be conducted using information from the panels' specification sheet. The panels' rated efficiency in 2008 was 18.1% at STC. By using data from times when the weather was near STC ( $1000 \pm 10 \text{ W/m}^2$ ), a comparison can be made with subsequent years, as shown in Figure 5.13. Using this method, the degradation rate for the c-Si array is -3.04%. For comparison, this analysis was repeated for the a-Si array, as shown in Figure 5.12. For the a-Si array,  $R_d$  is larger at -3.24%. In addition to the effect of the low quality data as discussed previously, this method likely overestimates the degradation rate as efficiency ratings obtained in a laboratory cannot be directly related to those obtained in the field. The field efficiency will likely be significantly lower as the panels are not in a controlled environment at exactly STC. Additionally, there are many system losses that are not represented in the rated efficiency. Inverter and wiring losses, soiling, LID, variations in manufacturing, and other factors mean that the efficiency of the panels soon after installation was likely lower than the rated efficiency. Thus, the degradation rates obtained from this method are likely greater than the true degradation rates.

While this method may not provide the most accurate results, it does indicate that the a-Si array is degrading faster than the c-Si array. This further supports the findings of the previous section, which showed a difference of 0.47% between the  $R_d$  values of the two arrays.



**Figure 5.13:**  $R_d$  for the c-Si array based on panel efficiency.



**Figure 5.12:**  $R_d$  for the a-Si array based on panel efficiency.

## 6 Conclusions

The performance of the a-Si and c-Si arrays at the NCSU Solar House was compared for different weather conditions and solar positions. It was found that while the c-Si array had a higher normalized power during sunny days, the a-Si array outperformed it during inclement weather, including cloudy, rainy, and snowy conditions. Additionally, the a-Si array produced more power at dawn and dusk, proving that it performed better than the c-Si array under low-light conditions. An analysis of power production as a function of solar position revealed shading that impacted both of the arrays' performance.

This study also examined different methods for calculating the degradation rate of the arrays. The methodology developed for optimizing the degradation rate models proved successful at significantly lowering model uncertainty and allowing more accurate calculations when fewer years of data are available. It was found that changing the performance metric and aggregation period used in the model had a significant impact on the results, while the effect of changing the data filtering criteria was minimal. The choice of analytical method also played an important role. Different analytical methods produced not only different model uncertainties, but also different values of  $R_d$ . This model optimization process helped to minimize these differences, resulting in better agreement between the analytical methods.

It was found that the YoY analytical method produced the lowest uncertainty for both arrays.  $R_d$  was found to be significantly different between the two, with the c-Si array having a value of  $-0.79 \pm 0.52\%$ , while the a-Si array had a value of  $-1.26 \pm 0.93\%$ . Although the uncertainty of these values is much lower than for the base models, it is not low enough to have a high degree of confidence in the results. From this, it is clear that more data is needed to produce an accurate  $R_d$  value.

While the above method on its own is insufficient to definitively state the degradation rate of the arrays, corroborating evidence can be found using the data from the panel specification sheets as well as from past studies. The  $R_d$  value for the a-Si array calculated using data from previous studies were found to be in good agreement with the results of the YoY method. Additionally, calculations using the specification sheets indicate that the a-Si array is degrading at a faster rate. These results strengthen the findings of the previous section, and the accuracy of the  $R_d$  calculations will only improve as the available amount of data continues to grow.

## 6.1 Recommendations and Future Work

The methodology developed in this thesis lays the groundwork for much future study on the Solar House. As more data is collected, future studies can validate the results of this work, and continue to improve upon the current degradation models. This may include addressing other modeling parameters not included in this study, such as stability data filters and the effect of wind speed, or expanding to include more advanced analytical methods like ARIMA and machine learning. Examining the data for potentially non-linear degradation is another area that may be explored once more data is available.

Measurements of the array IV curves would provide valuable information on the PV panels' performance. Additionally, it would allow for some analysis of the degradation modes of the panels. This, combined with a physical inspection of the arrays would give insight into the causes of performance degradation and how they differ between a-Si and c-Si PV panels. Monitoring of the power produced by each string or individual panel would also help determine sources of performance degradation by identifying which panels are underperforming.

Several improvements could be made to the data acquisition setup that would improve future results. Installing a temperature monitor on the back of the panels would greatly increase the accuracy of the cell temperature model, and therefore improve the accuracy of degradation rate calculations. Similarly, moving the anemometer so that it can record reliable wind speed data would allow for more accurate performance modeling. Tilting the pyranometer so it is at the same angle as the arrays would eliminate the need to computationally model the POA irradiance, removing another source of uncertainty from the models. Evaluating the effect of these changes on measurement uncertainty would provide insight into best practices for PV performance evaluation. As available data grows and the data acquisition and modeling methodologies improve, there is a wealth of knowledge that can be gained from the Solar House. With so much work left to be done, it is clear that the Solar House will remain a valuable resource for many years to come.

## REFERENCES

- [1] “Solar Industry Research Data,” *SEIA*, 2019. <https://www.seia.org/solar-industry-research-data> (accessed Oct. 08, 2019).
- [2] “U.S. Renewable Energy Factsheet,” *University of Michigan Center for Sustainable Systems*, 2018. <http://css.umich.edu/factsheets/us-renewable-energy-factsheet> (accessed Oct. 08, 2019).
- [3] “U.S. Solar Market Insight,” *SEIA*, Sep. 17, 2019. <https://www.seia.org/us-solar-market-insight> (accessed Oct. 18, 2019).
- [4] E. Asmelash and G. Prakash, “Future of Solar Photovoltaic: Deployment, investment, technology, grid integration and socio-economic aspects,” International Renewable Energy Agency, Abu Dhabi, Nov. 2019. Accessed: May 10, 2021. [Online]. Available: [https://irena.org/media/-/Files/IRENA/Agency/Publication/2019/Nov/IRENA\\_Future\\_of\\_Solar\\_PV\\_2019.pdf](https://irena.org/media/-/Files/IRENA/Agency/Publication/2019/Nov/IRENA_Future_of_Solar_PV_2019.pdf).
- [5] N. Llewellyn, “An Evaluation of the Rooftop PV and Solar Water Heating Systems at the NCSU Solar House,” North Carolina State University, Raleigh, NC, 2019.
- [6] “SunPower 225 Solar Panel Specification Sheet.” SunPower Corporation, Mar. 2008, Accessed: Sep. 20, 2019. [Online]. Available: [https://www.pocosolar.com/wp-content/themes/twentyfifteen/pdfs/Sunpower%20Solar%20Panels/sunpower\\_225blk\\_spec\\_sheet.pdf](https://www.pocosolar.com/wp-content/themes/twentyfifteen/pdfs/Sunpower%20Solar%20Panels/sunpower_225blk_spec_sheet.pdf).
- [7] “SunPower SPR-5000m / SPR-6000m / SPR-7000m / SPR-8000m Installation Guide.” SunPower Corporation, 2010, Accessed: Sep. 20, 2019. [Online]. Available: <http://spectrum.sunpower.com/sites/default/files/uploads/resources/mn-sunpower-spr-5000m-spr-6000m-spr-7000m-spr-8000m-installation-guide.pdf>.
- [8] “Uni-Solar Field Applied Roofing PV Laminate With Top Termination For Steel Roofs Data Sheet.” United Solar Ovonic, 2000, Accessed: Sep. 24, 2019. [Online].
- [9] D. Christy, “An Experimental Evaluation of the Performance of the Amorphous Silicon PV Array on the NCSU AFV Garage,” North Carolina State University, Raleigh, NC, 2007.
- [10] “SolarEdge Single Phase Inverters.” SolarEdge Technologies, Inc., Feb. 2014, [Online]. Available: <https://www.solaredge.com/sites/default/files/se-single-phase-us-inverter-datasheet.pdf>.

- [11] M. Kumar and A. Kumar, "Performance assessment and degradation analysis of solar photovoltaic technologies: A review," *Renew. Sustain. Energy Rev.*, vol. 78, pp. 554–587, Oct. 2017, doi: 10.1016/j.rser.2017.04.083.
- [12] J. Stanley, "An Analysis of Short-Circuit Current Versus Irradiance in Uni-Solar PVL-128 Amorphous Silicon Photovoltaic Cells," North Carolina State University, Raleigh, NC, May 2004.
- [13] "Solar Photovoltaic Cell Basics," *Energy.gov*, Aug. 16, 2013. <https://www.energy.gov/eere/solar/articles/solar-photovoltaic-cell-basics> (accessed Oct. 01, 2019).
- [14] K. W. Jansen, S. B. Kadam, and J. F. Groelinger, "The Advantages of Amorphous Silicon Photovoltaic Modules in Grid-Tied Systems," in *2006 IEEE 4th World Conference on Photovoltaic Energy Conference*, May 2006, vol. 2, pp. 2363–2366, doi: 10.1109/WCPEC.2006.279666.
- [15] C. N. Jardine, G. J. Conibeer, and K. J. Lane, "PV-COMPARE: Direct Comparison of Eleven PV Technologies at Two Locations in Northern and Southern Europe," presented at the Sixteenth European Photovoltaic Solar Energy Conference, 2001.
- [16] D. C. Jordan, S. R. Kurtz, K. VanSant, and J. Newmiller, "Compendium of photovoltaic degradation rates," *Prog. Photovolt. Res. Appl.*, vol. 24, no. 7, pp. 978–989, 2016, doi: 10.1002/pip.2744.
- [17] E. Molenbroek, D. W. Waddington, and K. A. Emery, "Hot spot susceptibility and testing of PV modules," in *The Conference Record of the Twenty-Second IEEE Photovoltaic Specialists Conference - 1991*, Oct. 1991, pp. 547–552 vol.1, doi: 10.1109/PVSC.1991.169273.
- [18] A. Ndiaye, A. Charki, A. Kobi, C. M. F. Kébé, P. A. Ndiaye, and V. Sambou, "Degradations of silicon photovoltaic modules: A literature review," *Sol. Energy*, vol. 96, pp. 140–151, Oct. 2013, doi: 10.1016/j.solener.2013.07.005.
- [19] S. Pingel *et al.*, "Potential Induced Degradation of solar cells and panels," in *2010 35th IEEE Photovoltaic Specialists Conference*, Jun. 2010, pp. 002817–002822, doi: 10.1109/PVSC.2010.5616823.

- [20] S. Pingel, Y. Zemen, O. Frank, T. Geipel, and J. Berghold, “Mechanical Stability of Solar Cells within Solar Panels,” *Proc 24th Eur. Photovolt. Energy Conf*, Jan. 2009, doi: 10.4229/24thEUPVSEC2009-4AV.3.49.
- [21] M. S. El-Shobokshy and F. M. Hussein, “Degradation of photovoltaic cell performance due to dust deposition on to its surface,” *Renew. Energy*, vol. 3, no. 6, pp. 585–590, Sep. 1993, doi: 10.1016/0960-1481(93)90064-N.
- [22] C. Deline, N. DiOrio, D. Jordan, and F. Toor, “Progress & Frontiers in PV Performance,” presented at the Solar Power International 2016, Las Vegas, Nevada, Sep. 12, 2016, Accessed: Sep. 20, 2019. [Online]. Available: <https://www.nrel.gov/docs/fy16osti/67174.pdf>.
- [23] D. C. Jordan, T. J. Silverman, J. H. Wohlgemuth, S. R. Kurtz, and K. T. VanSant, “Photovoltaic failure and degradation modes,” *Prog. Photovolt. Res. Appl.*, vol. 25, no. 4, pp. 318–326, 2017, doi: 10.1002/pip.2866.
- [24] A. Phinikarides, N. Kindyni, G. Makrides, and G. E. Georghiou, “Review of photovoltaic degradation rate methodologies,” *Renew. Sustain. Energy Rev.*, vol. 40, pp. 143–152, Dec. 2014, doi: 10.1016/j.rser.2014.07.155.
- [25] D. C. Jordan *et al.*, “Reducing Interanalyst Variability in Photovoltaic Degradation Rate Assessments,” *IEEE J. Photovolt.*, vol. 10, no. 1, pp. 206–212, Jan. 2020, doi: 10.1109/JPHOTOV.2019.2945191.
- [26] A. Woyte, M. Richter, D. Moser, S. Mau, N. Reich, and U. Jahn, “Monitoring of Photovoltaic Systems: Good Practices and Systematic Analysis,” in *28th European Photovoltaic Solar Energy Conference and Exhibition*, Paris, France, Oct. 2013, p. 9.
- [27] A. Kimber *et al.*, “Improved test method to verify the power rating of a photovoltaic (PV) project,” in *2009 34th IEEE Photovoltaic Specialists Conference (PVSC)*, Jun. 2009, pp. 000316–000321, doi: 10.1109/PVSC.2009.5411670.
- [28] D. C. Jordan and S. R. Kurtz, “The Dark Horse of Evaluating Long-Term Field Performance—Data Filtering,” *IEEE J. Photovolt.*, vol. 4, no. 1, pp. 317–323, Jan. 2014, doi: 10.1109/JPHOTOV.2013.2282741.
- [29] B. Paudyal, M. Bolen, and D. Fregosi, “PV Plant Performance Loss Rate Assessment: Significance of Data Filtering and Aggregation,” in *2019 46th IEEE Photovoltaic*

- Specialists Conference (PVSC)*, Jun. 2019, pp. 0866–0869, doi: 10.1109/PVSC40753.2019.8981247.
- [30] D. C. Jordan, C. Deline, S. R. Kurtz, G. M. Kimball, and M. Anderson, “Robust PV Degradation Methodology and Application,” *IEEE J. Photovolt.*, vol. 8, no. 2, pp. 525–531, Mar. 2018, doi: 10.1109/JPHOTOV.2017.2779779.
- [31] D. C. Jordan and S. R. Kurtz, “Analytical improvements in PV degradation rate determination,” in *2010 35th IEEE Photovoltaic Specialists Conference*, Jun. 2010, pp. 002688–002693, doi: 10.1109/PVSC.2010.5617074.
- [32] A. Phinikarides, G. Makrides, and G. E. Georghiou, “Comparison of Analysis Methods for the Calculation of Degradation Rates of Different Photovoltaic Technologies,” 2013, pp. 3973–3976, [Online]. Available: <http://foss.ucy.ac.cy/projects/pvdegradation/pdf/5BV.4.39.pdf>.
- [33] A. R. Gxasheka, E. E. van Dyk, and E. L. Meyer, “Evaluation of performance parameters of PV modules deployed outdoors,” *Renew. Energy*, vol. 30, no. 4, pp. 611–620, Apr. 2005, doi: 10.1016/j.renene.2004.06.005.
- [34] H. Haeberlin, C. Beutler, and I. Burgdorf, “Normalized Representation of Energy and Power for Analysis of Performance and On-line Error Detection in PV-Systems,” in *13th EU PV Conference on Photovoltaic Energy Conversion*, Nice, France, 1995, p. 4, [Online]. Available: <https://pdfs.semanticscholar.org/69e5/d9080ffd1d09af8b5f250870a0a140387fa6.pdf>.
- [35] C. M. Whitaker *et al.*, “Application and validation of a new PV performance characterization method,” in *Conference Record of the Twenty Sixth IEEE Photovoltaic Specialists Conference - 1997*, Sep. 1997, pp. 1253–1256, doi: 10.1109/PVSC.1997.654315.
- [36] J. A. Kratochvil, W. E. Boyson, and D. L. King, “Photovoltaic Array Performance Model,” Sandia National Laboratories, Albuquerque, New Mexico, SAND2004-3535, Dec. 2004. doi: 10.2172/919131.
- [37] B. Marion *et al.*, “Performance parameters for grid-connected PV systems,” in *Conference Record of the Thirty-first IEEE Photovoltaic Specialists Conference, 2005.*, Jan. 2005, pp. 1601–1606, doi: 10.1109/PVSC.2005.1488451.

- [38] G. Makrides, B. Zinsser, G. E. Georghiou, M. Schubert, and J. H. Werner, “Degradation of different photovoltaic technologies under field conditions,” in *2010 35th IEEE Photovoltaic Specialists Conference*, Jun. 2010, pp. 002332–002337, doi: 10.1109/PVSC.2010.5614439.
- [39] A. Phinikarides, G. Makrides, B. Zinsser, M. Schubert, and G. E. Georghiou, “Analysis of photovoltaic system performance time series: Seasonality and performance loss,” *Renew. Energy*, vol. 77, pp. 51–63, May 2015, doi: 10.1016/j.renene.2014.11.091.
- [40] B. Meyers, M. Deceglie, C. Deline, and D. Jordan, “Signal Processing on PV Time-Series Data: Robust Degradation Analysis Without Physical Models,” *IEEE J. Photovolt.*, vol. 10, no. 2, pp. 546–553, Mar. 2020, doi: 10.1109/JPHOTOV.2019.2957646.
- [41] A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, “Definition and Computation of the Degradation Rates of Photovoltaic Systems of Different Technologies With Robust Principal Component Analysis,” *IEEE J. Photovolt.*, vol. 5, no. 6, pp. 1698–1705, Nov. 2015, doi: 10.1109/JPHOTOV.2015.2478065.
- [42] R. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018.
- [43] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, “STL: A Seasonal-Trend Decomposition Procedure Based on Loess,” *J. Off. Stat.*, vol. 6, no. 1, pp. 3–73, 1990.
- [44] C. C. Holt, “Forecasting seasonals and trends by exponentially weighted moving averages,” *Int. J. Forecast.*, vol. 20, no. 1, pp. 5–10, 1957, doi: 10.1016/j.ijforecast.2003.09.015.
- [45] P. R. Winters, “Forecasting Sales by Exponentially Weighted Moving Averages,” *Manag. Sci.*, vol. 6, no. 3, pp. 324–342, Apr. 1960, doi: 10.1287/mnsc.6.3.324.
- [46] S. Lindig, I. Kaaya, K.-A. Weiss, D. Moser, and M. Topic, “Review of Statistical and Analytical Degradation Models for Photovoltaic Modules and Systems as Well as Related Improvements,” *IEEE J. Photovolt.*, vol. 8, no. 6, pp. 1773–1786, Nov. 2018, doi: 10.1109/JPHOTOV.2018.2870532.
- [47] D. C. Jordan, M. G. Deceglie, and S. R. Kurtz, “PV degradation methodology comparison — A basis for a standard,” in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, Jun. 2016, pp. 0273–0278, doi: 10.1109/PVSC.2016.7749593.

- [48] A. Phinikarides, G. Makrides, N. Kindyni, and G. E. Georghiou, “Comparison of trend extraction methods for calculating performance loss rates of different photovoltaic technologies,” in *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, Jun. 2014, pp. 3211–3215, doi: 10.1109/PVSC.2014.6925619.
- [49] S. Lindig *et al.*, “International collaboration framework for the calculation of performance loss rates: Data quality, benchmarks, and trends (towards a uniform methodology),” *Prog. Photovolt. Res. Appl.*, pp. 1–30, Mar. 2021, doi: <https://doi.org/10.1002/pip.3397>.
- [50] “Plane of Array (POA) Irradiance,” in *PVPMC Modeling Steps*, Albuquerque, NM: Sandia National Laboratories.
- [51] P. Ineichen, R. Perez, R. Seal, E. L. Maxwell, and A. Zalenka, “Dynamic global-to-direct irradiance conversion models,” in *ASHRAE Transactions*, 1992, vol. 1, p. 98.
- [52] W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, “pvlib python: a python package for modeling solar energy systems,” *J. Open Source Softw.*, vol. 3, no. 29, p. 884, Sep. 2018, doi: 10.21105/joss.00884.
- [53] *System Advisor Model*. Golden, CO: National Renewable Energy Laboratory, 2020.
- [54] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, “The National Solar Radiation Data Base (NSRDB),” *Renew. Sustain. Energy Rev.*, vol. 89, pp. 51–60, Jun. 2018, doi: 10.1016/j.rser.2018.03.003.
- [55] M. Deceglie, A. B. Shinn, K. Anderson, M. Mikofski, J. Yan, and W. Vining, *RdTools*. National Renewable Energy Laboratory, 2020.
- [56] B. Efron, “Bootstrap Methods: Another Look at the Jackknife,” *Ann. Stat.*, vol. 7, no. 1, pp. 1–26, Jan. 1979, doi: 10.1214/aos/1176344552.
- [57] S. N. Lahiri, “Theoretical comparisons of block bootstrap methods,” *Ann. Stat.*, vol. 27, no. 1, pp. 386–404, Mar. 1999, doi: 10.1214/aos/1018031117.
- [58] K. Sheppard *et al.*, *bashtage/arch: Release 4.11*. Zenodo, 2019.
- [59] “Download Data: Array 10 SunPower, 5.8kW, mono-Si, Fixed, 2009, 215W,” Desert Knowledge Australia Centre, Alice Springs. Accessed: Feb. 25, 2021. [Online]. Available: <http://dkasolarcentre.com.au/source/alice-springs/dka-m7-a-phase>.

- [60] “SunPower 215 Solar Panel Specification Sheet.” SunPower Corporation, Oct. 2007, Accessed: Mar. 25, 2021. [Online]. Available: <https://pdf.archiexpo.com/pdf/sunpower/215-solar-panel-datasheet/54500-12401.html>.
- [61] “North Carolina State University Solar House/N.C. Clean Energy Technology Center,” *Visit Raleigh*. <https://www.visitraleigh.com/listing/north-carolina-state-university-solar-house-n-c-clean-energy-technology-center/59323/> (accessed Apr. 16, 2021).
- [62] F. Mahmood *et al.*, “Temperature coefficient of power (Pmax) of field aged PV modules: impact on performance ratio and degradation rate determinations,” in *Reliability of Photovoltaic Cells, Modules, Components, and Systems X*, Aug. 2017, vol. 10370, p. 1037008, doi: 10.1117/12.2281840.
- [63] “Cardinal [data retrieval interface],” *North Carolina State Climate Office*. <https://products.climate.ncsu.edu/cardinal/request> (accessed Apr. 19, 2021).