# Data-Driven-Based Analysis and Modeling for The Impact of Wildfire Smoke on PV Systems

Amjad Jebril Ali, *Student Member, IEEE*, Long Zhao, *Member, IEEE,* and Mohammad Heidari Kapourchali, *Member, IEEE*

***Abstract*—With the vast deployment of photovoltaic (PV) systems and the increasing number of wildfires, understanding the impact of wildfire smoke on PV generation is essential for power systems operation and planning. This paper analyzes the impact of wildfire smoke on solar spectra, radiation, spectral irradiance, and output of a grid-tied PV system based on the data collected during wildfire events. A model for the PV power output reduction caused by wildfire smoke and the significance of meteorological factors are developed and assessed using a data-driven technique. The developed model accurately estimates the power output reduction of PV systems caused by wildfire smoke. The developed model is also applied to study power reduction on different PV cell technologies. The result implies that solar cell technologies respond to wildfire smoke differently. High wildfire-risk areas should consider different solar cell materials to mitigate the power output reduction due to wildfire smoke. The outcome of this study is critical for future power systems operation decision making and management with high penetration of solar generation to ensure the stability and reliability of the power grids.**

***Index Terms*—Data-driven, PV systems, power systems operation renewable energy, solar generation, SVD, spectral response, wildfire smoke effects.**

## I. INTRODUCTION

PROPERLY utilizing renewable energy and distributed energy resources (DERs) is essential for future power grids due to the concern of climate change and the goal of decarbonization. In 2021, around 20% of the total electricity was generated by renewable energy sources in the U.S. [1]. Globally, despite the impacts of the COVID-19 pandemic, renewable energy set a record capacity in 2020 and was the only source of electricity generation to register a net increase in total capacity [2]. Due to the declining price and matured technology of solar cells, PV adoption has been significantly boosted [3-4]. In 2021, solar PV accounted for more than half of all renewable power expansion [5]. At the end of 2021, global PV installations reached 939 GWdc [4]. In addition, the Federal Energy Regulatory Commission (FERC) Order 2222 was recently approved, allowing DERs to participate in wholesale electricity markets in aggregation [6]. The US government set a goal to reach 100 percent carbon pollution-free electricity by 2035 and have a net-zero economy no later than 2050 [7-8]. Both policies will facilitate the development and deployment of renewable distributed energy generation significantly at all levels of power grids in the U.S. However, wildfires have become more frequent with the record-high temperature and drought [9]. An average of 62,805 wildfires occurred from 2011 to 2020, and 7.5 million acres were affected yearly. In 2020, 58,950 wildfires burned 10.1 million acres, the second-most acreage affected in a year since 1960. Nearly 40% of these acres were in California [10]. According to the U.S. National Interagency Coordination Center, in 2021, 58,985 wildfires burned 10,122,336 acres [11]. According to the U.S department of agriculture, catastrophic blazes in 2050 are projected to burn twice as many acres as in 2015 [12]. In addition, European Forest Fire Information System (EFFIS) estimates that within European Union countries, between 1.2 million acres and 1.6 million acres have been burned by wildfires [13]. Also, more than 60 hectares were burned in Australia in 2020 [14]. At a global level, wildfires emitted 1.76 billion tons of carbon in 2021 [15]. All of this evidence shows that climate change may increase the severity of wildfires in the future [9]. As shown in Fig. 1, wildfires produce a significant amount of smoke that can travel hundreds or even thousands of miles, covering a substantial amount of landmass. Due to the increasing penetration of solar generation in power grids, it is vital to study the impact of wildfire smoke on PV systems' power output.

Wildfire smoke can reduce the PV systems' generation output by changing the solar spectral irradiance. The authors of [22] studied the influence of the smoke from the National Park wildfire on solar spectra and radiation in Golden, Colorado. The result demonstrated that smoke significantly changes both Direct Nominal Irradiance (DNI) and Diffused Horizontal Irradiance (DHI). Similarly, [16] displays that Global Solar Radiation (GSR) follows a sine curve in clear skies with a 20% to 40% reduction in amplitude due to smoke scattering. Because of the considerable decrease in solar radiation caused by wildfire smoke, the stability and reliability of power grids with high penetration of PV systems could be jeopardized by increasing wildfires. In September 2020, California Independent System Operator (CAISO) revealed that average solar-powered electricity generation declined by nearly 30% due to the wildfires in the first two weeks, and monthly solar generation was 13.4% lower than the previous year despite a 5.3% increase in installed solar generation capacity in the past

Manuscript was submitted on February 25, 2023, for review. This work was supported in part by the U.S. National Science Foundation (NSF) under Grants RISE-2220624 and RISE-2022705. *(Corresponding author: Long Zhao).*

Amjad Jebril Ali and Long Zhao are with the Department of Electrical Engineering and Computer Science, South Dakota School of Mines and Technology, Rapid City, SD 57701 USA (e-mail: amjed.jebril@gmail.com; long.zhao@sdsmt.edu).

Mohammad Heidari Kapourchali is with the Department of Electrician Engineering, University of Alaska Anchorage, Anchorage, AK 99508 USA (e-mail: mhkapourchali@alaska.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org

year [17]. During Canadian wildfire in June 2023, according to Maxar [18], experts have determined that solar power generation in the US Northeast experienced a significant reduction to less than 10% of the expected production due to the presence of Canadian wildfire smoke. Meanwhile, PJM shows that between the hours of 1 to 2 p.m. on June 6th, a sudden drop of 1,000 MW in solar output occurred, followed by a continuous decline until sunset [19]. The authors in the paper [20] studied the impact of wildfire smoke on PV power output. Artificial wildfire was set up to measure the smoke pollutants, compared with the data collected on clear days. The results showed that PV productivity reduction caused by wildfire smoke is from 7% to 27%. A study [29] explains that the leading cause of solar radiation reduction during wildfires is the suspended particles (pollution) in the sky which absorb or scatter sunlight [21]. In the paper [22], the authors examined how air pollution affects solar power generation at different periods. The result reveals that PV output decreases as the pollution increase. The authors in the paper [31] investigated the average reduction of PV power output due to air pollution for an extended period. The study shows that the annual average energy reduction reached 20-25%, similar to clouds' impact in winter. In addition, multiple studies have revealed that PV panels with tracking systems, especially dual trackers, tend to lose more generation capacity due to aerosol pollution than fixed arrays at optimal angles [16, 23].

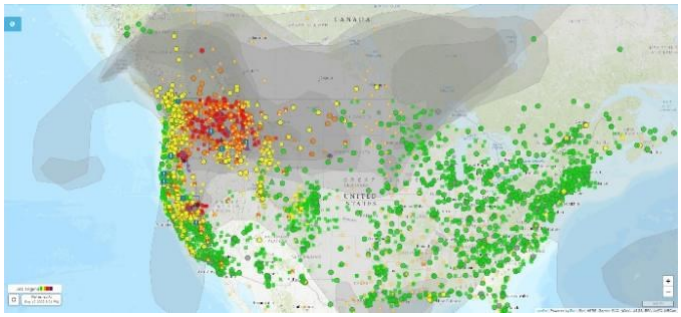


**Fig. 1.** The U.S. wildfire and smoke map on September 12, 2022, at 3:00 PM [24].

However, existing studies have yet to provide a numerical solution or model to quantify the impact of wildfire smoke on the power output of PV systems. Due to the elevation and fluid dynamics of smoke plumes, smoke quantification is highly arduous, which makes modeling wildfire smoke impact on PV systems challenging. Current methods to quantify wildfire smoke impact on PV systems are mainly developed using aerosol optical depth (AOD) or air quality index PM 2.5. The AOD is defined as the measure of extinction of the solar beam by dust and haze [21]. Considering that pollution particles block sunlight by absorbing or scattering the sunlight, the AOD has been mainly used to assess the degree of air pollution. However, the AOD only considers either short range or a specific wavelength. The power output of a solar cell depends on the spectral response (*SR*) characteristics shown in Fig. 2. The *SRs* of different solar cell materials depends on a wide range of spectral wavelengths. Therefore, the AOD is insufficient to quantify the impact of wildfire smoke on PV systems. In the paper [25], the authors use AOD to quantify the impact of wildfire smoke on a PV system. The AOD used in this paper only considers a wavelength of 550 nm, ignoring the full range of *SR* of PV devices. PM 2.5 is defined as fine inhalable particles with diameters that are generally 2.5 micrometers and smaller [27], and it is widely applied to index air quality. However, the measurement of PM 2.5 is heavily location-dependent.

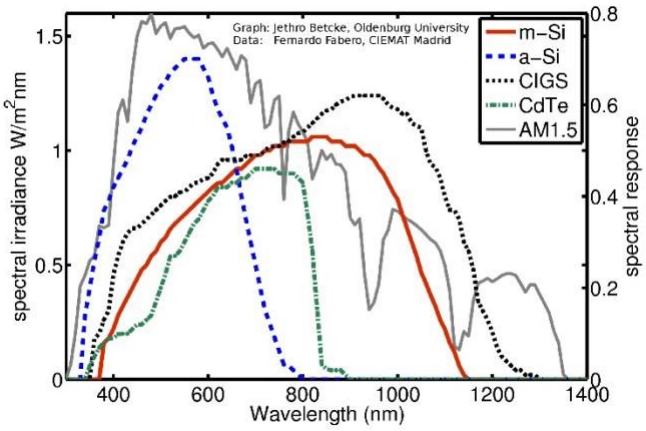



**Fig. 2.** Spectral response characteristics of different solar module technologies [26].

Due to the smoke elevation, the value of PM 2.5 can still be considered good while during a smoky day [28-30]. Multiple studies have shown that the wildfire smoke impact on the solar spectra is nonlinear [20, 23]. To properly quantify the wildfire smoke impact on solar radiation, a new quantification index parameter, Absorption Radiation Index (*ARI*), was introduced and developed in our previous work [31]. The *ARI* is determined as follows:

$$ARI = \frac{CSS - PSS}{CSS} \tag{1}$$

where *CSS* (W/m$^2$) is the solar radiation of a clear sky (Clear Sky Spectrum on a cloudless non-smoky day), and *PSS* (W/m$^2$) is the solar radiation of a polluted sky (Polluted Sky Spectrum on a cloudless day). This study analyzes the impact of wildfire smoke on solar power generation and develops a data-driven model to quantify PV power output reduction. The findings offer guidelines for selecting suitable solar cell technologies in high-risk wildfire areas, minimizing power generation losses caused by smoke.

The rest of the paper is organized as follows. Section II presents the data collection process and analyzes the wildfire smoke impact on solar spectra, solar radiation, and PV power output. Section III shows the wildfire smoke impact on solar spectral irradiance and introduces the data-driven modeling methodology. PV generation difference among four common solar cell materials under the impact of wildfire smoke is analyzed in Section IV. Section V shows the model validation, and Section VI concludes the paper.

## II. Wildfire Smoke Data Collection and Analysis

One of the biggest challenges to studying wildfire smoke impact on PV systems is the control variables. In our previous study [31], artificial woodburning smoke was generated to

simulate the wildfire smoke scenarios. However, due to the complexity of the photochemical smog formation of real wildfire smoke, artificial smoke cannot provide as valuable data as actual wildfires. Therefore, to study the wildfire smoke on PV systems, a cloudless day with constant "steady-state" wildfire smoke is required. This study seized a short window during a Montana wildfire in early September 2022, shown in Fig. 3. The multidimensional data were collected on consecutive smoky days. The section discussed wildfire impacts on the solar spectrum, solar radiation, and PV generation.

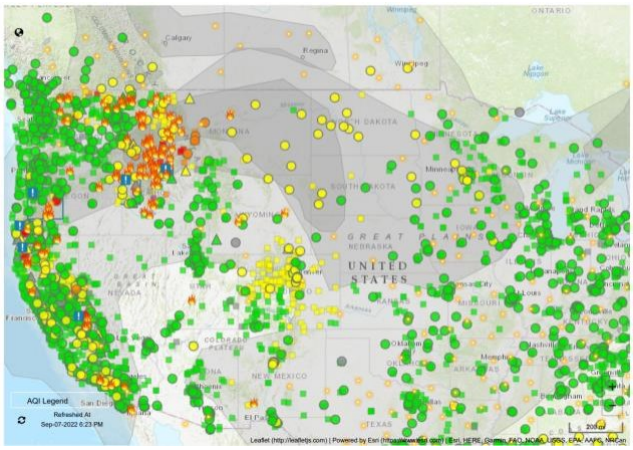



**Fig. 3.** Wildfire and smoke map on September 7, 2022, [24].

*A. Wildfire Smoke Impact on Solar Spectra*

To understand the impact of wildfire smoke on PV power generation, the impact on solar spectra needs to be investigated [31]. In this research, solar spectral data was collected using a scientific spectrometer which uses an optical fiber and a cosine corrector shown in Fig. 4. The spectrometer has a measurable range from 180nm to 1028nm [32], and it was attached to a 3-kW grid-tied rooftop solar system installed on South Dakota Mines campus. The PV system has an orientation of 215° as an Azimuth with a tilt angle of 16°. The solar spectrum is measured all day long on the Plane of Array (POA) of the PV panels. The solar spectrum was measured and recorded on September 6$^{th}$ and 7$^{th}$ from sunrise to sunset with 15 minutes time intervals. The daytime of September 6$^{th}$ was completely cloudless, with light smoke covered. September 7$^{th}$ is a severe smoky day. Most of the daytime of September 7$^{th}$ was cloudless except for a period from 2 PM to 3:15 PM. To show the impact of wildfire smoke on the solar spectrum, the spectral difference for every 15 minutes was computed.

**Fig. 4.** Solar spectra data acquisition system setup.

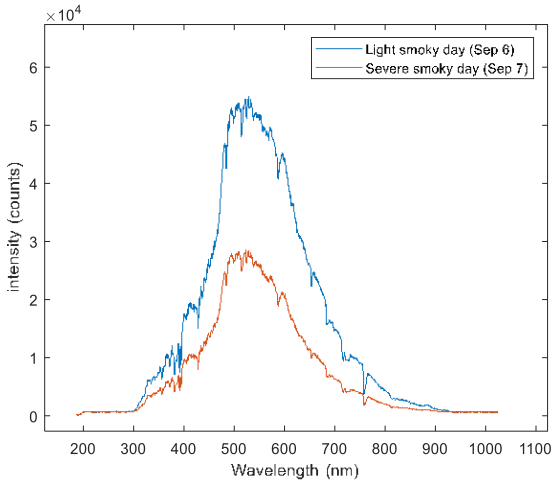



**Fig. 5.** Solar Spectral at 2:15 PM on September 6$^{th}$ and 7$^{th}$, 2022.

Due to the much heavier smoke on September 7$^{th}$, the solar spectral intensity on September 7$^{th}$ is lower than the spectral intensity on September 6$^{th}$. As shown in Fig. 5, the wildfire smoke impact on different wavelengths varies. The smoke blocks/reduces the sunlight the most from 500nm to 600nm wavelength but has much less effect on ultraviolet (UV) and infrared regions. This result explains the reddish sunlight color during wildfire days. During the wildfire smoky days, the light with longer wavelengths can still penetrate the smoke. Considering the different *SRs* for various solar cell technologies, the impact of wildfire smoke on other PV systems differs, which is discussed in Section IV.

*B. Wildfire Smoke Impact on Solar Radiation*

To investigate the wildfire impact on solar systems, 15-minute solar radiation data was also collected with a scientific weather station on South Dakota Mines campus [33]. Fig. 6 shows daily solar radiation curves for a non-smoky day (September 4$^{th}$), a light smoky day (September 6$^{th}$), and a severe smoky day (September 7$^{th}$), respectively.

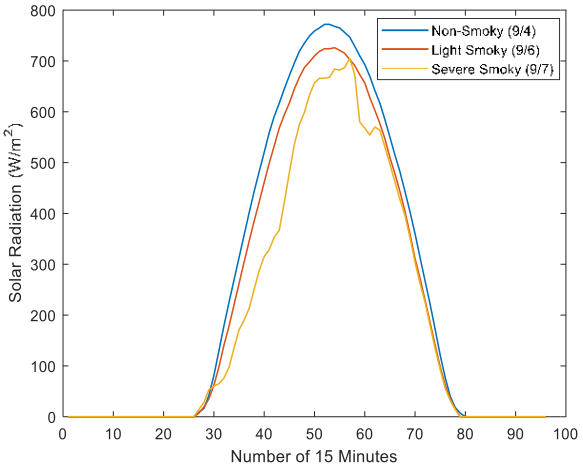



**Fig. 6.** Solar radiation for different smoke levels.

Fig. 7 shows the radiation difference between each one of them. Notably, the solar radiation on smoky days during sunrise is slightly higher than the non-smoky day. The severe smoky day has the highest solar radiation during the sunrise period, and this is mainly because wildfire smoke could scatter the dawning sunlight compared with a clear atmosphere. The PV system power output in the following subsection further confirms this observation. The most considerable solar radiation difference happened around 10:30 AM, highlighted with a red dashed line in Fig. 7.

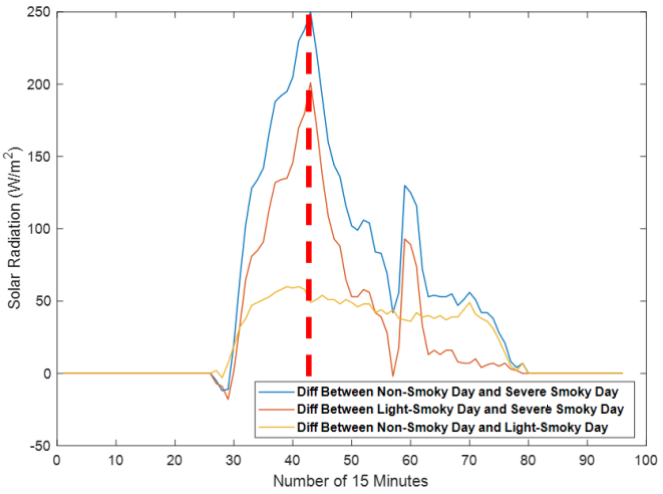



**Fig. 7.** Solar radiation difference for different smoke levels.

*C. Wildfire Smoke Impact on PV Power Output*

This study utilizes a 3-kW rooftop grid-tied solar system, which provides 5-minute power output data. To eliminate the cloud impact from 2 PM to 3:15 PM on September 7th, the data during this period is considered as outliers and replaced with the mean of adjacent data shown in Fig. 8. In this section, a non-smoky cloudless day, September 4th, 2022, is utilized as the baseline to distinguish the wildfire smoke impact on the power output of PV system.

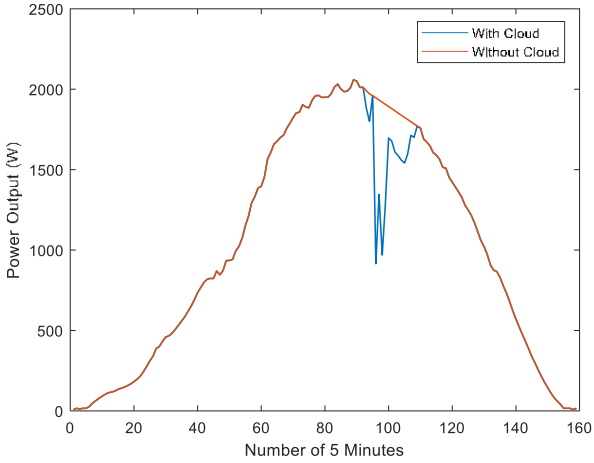



**Fig. 8.** Solar generation curve on September 7th, 2022.

To show the difference in wildfire smoke impact on the solar system among a non-smoky day (September 4th), a light smoky day (September 6th), and a severe smoky day (September 7th), Fig. 9 shows the power output curves of the 3-kW PV system of each day. Fig. 10 shows the power output difference between a non-smoky day and a severe smoky day and a light smoky day and a severe smoky day. The maximum PV power output reduction for all cases happened around the same time at 10:25 AM, highlighted with a red dash line in Fig. 10. The maximum power reduction on a severe smoky day is close to 50% of generation capacity on a clear day. This result shows that the wildfire smoke impact on solar power output differs with time. However, this peak power reduction time might vary at different months or seasons. This result indicates that a more flexible operating reserve should be considered for future operations and planning of power systems with high penetration of solar generation.

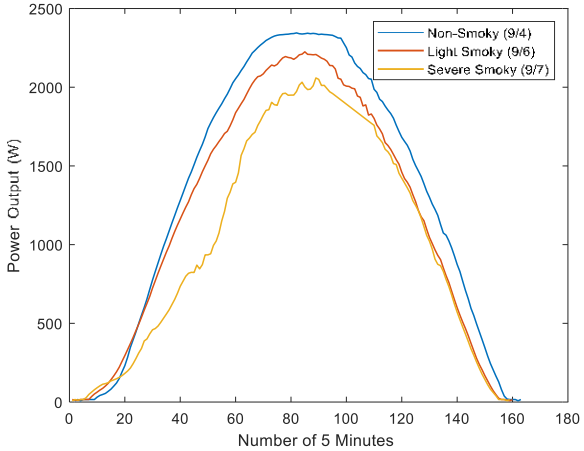



**Fig. 9.** Solar power output curves for different smoke levels.

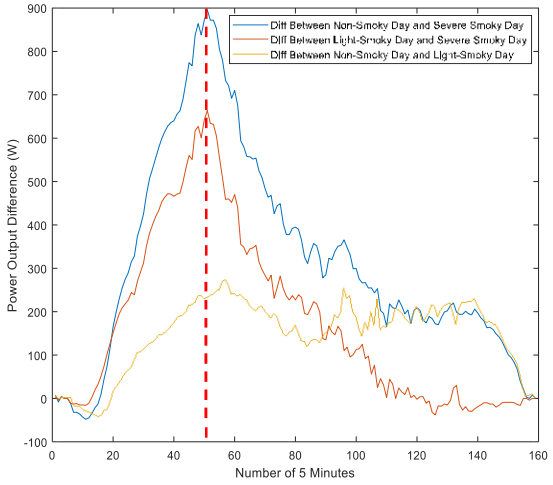



**Fig. 10.** Solar generation reductions for different smoke levels.

As discussed previously, the smoke will increase solar radiation during sunrise. As a result, the power output of PV systems during smoky days will be higher than the non-smoky days at dawning time. Therefore, for power systems with high penetration of solar generation, the system operator would anticipate slightly higher PV generation on smoky days than on non-smoky days during sunrise.

## III. Modeling Methodology

To develop a model for wildfire smoke impact on PV power output, solar spectral irradiance data from NREL was utilized in this study. Different from solar spectrum intensity and solar radiation ($W/m^2$), solar spectral irradiance ($W/m^2/nm$) measures the light energy of solar radiation per square meter at specific wavelength [34]. Due to the unique *SRs* of different solar cell technologies, solar spectral irradiance variation caused by wildfire smoke will change the power output of PV systems made of various materials differently.

### *A. Solar Spectral Irradiance Data*

The developed model in this study is built based on spectral data of DNI. The DNI spectra are obtained from the National Renewable Energy Lab (NREL) database in Golden, Colorado [35]. The DNI spectra have been captured using spectroradiometers installed in the NREL facility. The measured DNI spectrum complies with the National Institute of Standards and Technology (NIST) Standard of Spectral Irradiance Lamp Standards. The specified operating range of the PGS-100 is from -10 ºC to +45 ºC (without wind chill). The instrument has a 3.6nm spectral bandwidth, and the detector is maintained at 35ºC +/- 2ºC. The usable spectral range of the instrument is 350nm to 1050nm. Data is taken at approx. 0.7nm intervals every 5 minutes. Only spectral irradiance data from cloud-free periods are utilized in this study to extract the smoke's impact on the solar spectral irradiance. To do so, the NREL sky camera captures the sky image every10-minutes, which is applied to eliminate cloudy periods in this study. Additionally, historical meteorological data from Golden, Colorado, is acquired from [36] and applied for the modeling in this study.

### *B. Wildfire Smoke Impact on Solar Spectral Irradiance*

To understand the smoke impact on solar spectral irradiance, the hourly solar spectral irradiance data during the same period in different years are studied. A smoky wildfire period from July 30, 2020, to September 30, 2020, is compared with a non-smoky cloud-free period from July 30 to September 30, 2019. For example, Fig. 11 shows the solar spectral irradiance curves at the same time as August 13th in different years. The blue line is the solar spectral irradiance curve of a non-smoky day, and the red line is the solar spectral irradiance curve of a smoky day. Similar to the impact on solar spectra intensity, the profound reduction of solar spectral irradiance happens in the range of visible wavelengths and less reduction on UV and infrared wavelengths. This further confirms that the optical properties of wildfire smoke are more severe on visible light wavelengths but have less impact on UV and infrared wavelengths.

### *C. Data-Driven-Based Modeling*

The generation capacity of the PV system changes seasonally throughout the year due to different Azimuth and Zenith angles. Therefore, the PV power output varies with different days of the year. To compare the difference between smoky and non-smoky days and eliminate the seasonal variation impact and time dependency, historical data of 6:00 AM to 6:00 PM from 2019 and 2022 is used in this paper. Since this study aims to develop a model that estimates the reduction in PV systems due to wildfire smoke, the *ARI* and visibility range (*VR*) are considered predictor variables of the model. Because solar radiation is highly weather and location-dependent [37], meteorological parameters, including atmospheric temperature (*AT*), relative humidity (*RH*), and wind speed (*WS*) (measured 10m above the ground), are considered in the modeling.

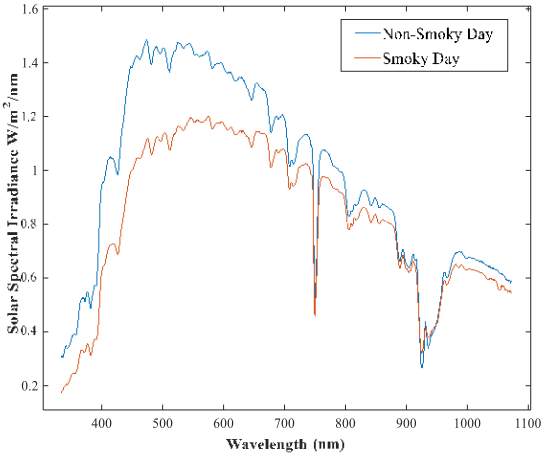



**Fig. 11.** Solar spectral irradiance at 9:30 AM between smokey and non-smoky days.

Additionally, authors in [38] discussed the dew point (*DPT*) also impacts the solar radiation on the earth's surface. Therefore, hourly data of the proceeded parameters are

considered the independent variables of the developed model. PV power output reduction percentage was developed in our previous work [31] and shown in (2).

$$P_{pv.r}\% \approx \frac{J_{s.cc} - J_{s.cs}}{J_{s.cc}} \times 100\% \quad (2)$$

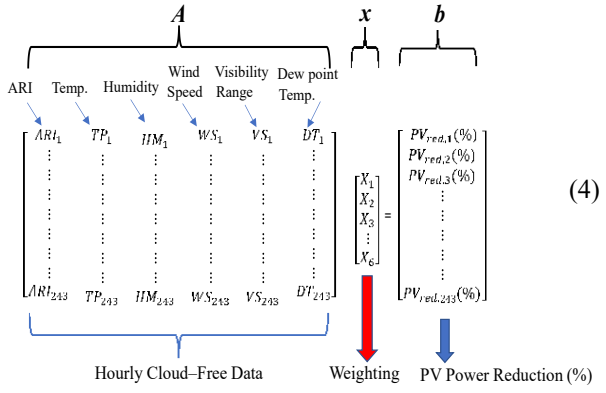
where $P_{pv.r}$ % represents the percentage of the PV power output reduction. $J_{s.c}$ is the short circuit current density which represents the short circuit current generated in ampere per unit area in $m^2$ of the PV cell. To develop a PV power output reduction model, a multilinear regression model in (3) is developed and shown below.

$$P_{pv,r}\% = f(ARI, AT, RH, WS, VR, DPT) = \alpha_1 \times ARI + \alpha_2 \times AT + \alpha_3 \times RH + \alpha_4 \times WS + \alpha_5 \times VR + \alpha_6 \times DPT \quad (3)$$

where $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, $\alpha_5$, and $\alpha_6$ are coefficients for each variable.

Multilinear regression, also known as multiple linear regression, is a statistical technique used to model the relationship between one dependent variable and two or more independent variables. It is an extension of simple linear regression, which models the relationship between one dependent variable and one independent variable. Multilinear regression can help to understand and predict the effect of multiple factors on a given outcome. To estimate the coefficients ($\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, $\alpha_5$, and $\alpha_6$), multilinear regression uses a technique called ordinary least squares (OLS). OLS minimizes the sum of squared differences between the actual and predicted values of the dependent variable. Once the coefficients are estimated, they can be used to make predictions or understand the relationship between the independent variables and the dependent variable [39]. In this study, multilinear regression is utilized to model the power reduction of PV systems and represented as $\boldsymbol{Ax}=\boldsymbol{b}$ shown in (4), where $\boldsymbol{A} \in \boldsymbol{C}^{243\times6}$ and $\boldsymbol{b} \in \boldsymbol{C}^{243\times1}$.

The meteorological measurements of cloudy periods from 2019 to 2022 were removed to eliminate the cloud impacts, and 243 measurements remained in matrix $\boldsymbol{A}$. The vector $\boldsymbol{b}$ is the corresponding power output reduction percentage calculated based on the *SR* of solar cells and solar spectral irradiance measurements.




(4)

To solve $\boldsymbol{Ax}=\boldsymbol{b}$ and determine the weighting vector $\boldsymbol{x}$, singular value decomposition (SVD) is utilized in this study. The SVD enables a systematic method to establish a low-dimensional approximation to high-dimensional data in terms of dominant patterns [40]. In this paper, the SVD of matrix $\boldsymbol{A}$ is defined as follows:

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \quad (5)$$

where $\boldsymbol{U} \in \boldsymbol{C}^{243\times243}$ can be considered as an eigen-measurements matrix of $\boldsymbol{A}$, which is hierarchically arranged to describe the variants of $\boldsymbol{A}$. $\boldsymbol{\Sigma} \in \boldsymbol{R}^{96\times96}$ are singular values ordered from largest to smallest, illustrating the importance of columns in the eigen-measurements matrix $\boldsymbol{U}$. $\boldsymbol{V} \in \boldsymbol{C}^{6\times6}$ is a matrix that corresponds to $\boldsymbol{\Sigma}$. Therefore, $\boldsymbol{Ax}=\boldsymbol{b}$ can be written in $\boldsymbol{U\Sigma V^T x}=\boldsymbol{b}$, and $\boldsymbol{x}=\boldsymbol{V\Sigma^{-1}U^T b}$. The least-square fitting of the pseudo-inverse is used to determine the matrix $\boldsymbol{x}$, which can be represented in (6)

$$min\ ||\boldsymbol{Ax}-\boldsymbol{b}||_2 \quad (6)$$

## IV. PV Generation Difference For Various Solar Cell Technologies

Based on the developed model in previous section, the PV generation difference among various solar module technologies under the impact of wildfire should is analyzed in this section. In this study, the vector $\boldsymbol{x}$ is determined using the computing tool MATLAB R2020b. The coefficients of four major solar cell technologies, m-Si, a-Si, CdTe, and CIGS are shown in Table I.

TABLE I
THE WEIGHTINGS FOR DIFFERENT PV CELL MATERIALS

| Cell Type | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\alpha_4$ | $\alpha_5$ | $\alpha_6$ |
|---|---|---|---|---|---|---|
| **m-Si** | 97.55 | -0.011 | -0.013 | -0.197 | 0.110 | -0.067 |
| **a-Si** | 106.04 | 0.107 | 0.002 | 0.044 | 0.018 | 0.067 |
| **CdTe** | 101.04 | -0.082 | -0.061 | 0.076 | 0.192 | 0.095 |
| **CIGS** | 94.70 | -0.137 | -0.058 | 0.193 | 0.165 | 0.052 |

In Table I, the coefficient of *ARI*, $\alpha_1$, is much significant than the rest of the coefficients of variables. To assess the significance of variables in matrix $\boldsymbol{A}$, each variable is investigated based on the significance probabilities (*P-value*) [41]. The *P-value* is defined as the probability that the estimated coefficient of the independent variable in the regression occurred randomly [42]. The less *P-value* is, the more important variables are. Typically, if *P-value* is less than 0.05, the variable can be considered significant. Further elaboration of the *P-value* can be found in [41-46]. As an example, the *P-value* for m-Si solar cell is shown in Table II.

TABLE II
P-VALUES OF INDEPENDENT VARIABLES FOR M-SI PV CELLS

| Predictors | *P-value* |
|---|---|
| ***ARI*** | 0.0000 |
| **Atmospheric Temp.** | 0.8738 |
| **Humidity** | 0.6056 |
| **Wind Speed** | 0.0565 |
| **Visibility** | 0.4632 |
| **Dew Point Temp.** | 0.39945 |

In Table II, the *P-value* of *ARI* is 0, which means the *ARI* is the most significant variable in the model to determine the power reduction of PV systems. The rest of the variables have relatively insignificant *P-values*, which means they are not crucial in the model development, and their variations will not change the PV power reduction caused by the wildfire smoke. Therefore, the power reduction of PV systems caused by wildfire smoke is mainly affected by the optical properties of solar spectra, and meteorological factors are insignificant. Similarly, the *P-values* of *ARI* for different solar cell materials are also equal to 0 and shown in Table III. It indicates that the *ARI* can be utilized as the only variable to model PV power output reduction caused by wildfire smoke.

TABLE III
P-VALUES OF ARI FOR FOUR MAJOR PV CELLS TECHNOLOGIES

| PV Technology | *P-value* |
|---|---|
| **m-Si** | 0.000 |
| **a-Si** | 0.000 |
| **CdTe** | 0.000 |
| **CIGS** | 0.000 |

Thus, the PV power reduction model can be solely determined by the *ARI*, which is shown in (7)

$$P_{Pv,r}\% = f(ARI) = \alpha_1 \times ARI \quad (7)$$

As the coefficient $\alpha_1$ for CIGS is the lowest among other solar cell materials, it indicates that the solar panels made by CIGS has the least impact from wildfire smoke on solar generation.

## V. MODEL VALIDATION

In this paper, the developed model is tested to avoid the risk of overfitting. Overfitting leads to a model that is incapable of generalizing [40]. To ensure accuracy, cross-validation techniques are often applied [40, 47]. Cross-validation assists with choosing the best-performing model by calculating the error using the testing dataset, which has not been used for model development. The testing dataset is applied to calculate the accuracy of the model and how it will be generalized with future input data [47]. In this paper, holdout cross-validation is used for model validation purposes. The concept of the holdout cross-validation is partitioning shuffled data randomly into two subsets of specified ratios for model development and validation. This technique performs development and testing only once, which cuts execution time on a large data set [46]. 80% of the dataset is used to build the model, and 20% is utilized to test the model. The result for each solar cell technology is plotted in Fig. 12.

To validate the model, the regression error is evaluated based on $R^2$, adjustment $R^2$, *significance F*, and root-mean-square-error (RMSE). Both $R^2$ and adjustment $R^2$ are model accuracy measurements for linear models. The *significance F* gives the probability that the model is incorrect. *Significance F* usually establishes a significance level and uses it as the cutoff point in evaluating the model. Commonly used significance levels are 1%, 5%, and 10% [42]. The dependability of the independent variable depends on a cut-off level that is decided separately [42]. The cutoff is selected based on the nature of the data studied and the different error types. *P-value* is similar in interpretation to the *significance F*. The key difference is that the *significance F* applies to the entire model, whereas the *P-value* will be used only for each corresponding coefficient [42]. The goodness of the developed model is illustrated in Table IV. The error and *significance F* show that the developed model is accurate to predict PV power output reduction purely based on the *ARI*.

In Fig. 12, few outliers can be observed due to the original abnormal/incorrect spectral irradiance data collected, and some PV power reduction points are close to 100%. This is because the solar spectral irradiance data used in this study was measured before sunrise, which is close to 0 W/m$^2$/nm. Both Table IV and Fig. 12 show that the developed model accurately performs predictive analysis using the *ARI*.

TABLE IV
HOLDOUT PARAMETERS OF THE GOODNESS OF THE TRAINED DATA FOR DIFFERENT TECHNOLOGIES OF PV PANELS

| PV Technology | Parameter | Estimated Value |
|---|---|---|
| **m-Si** | **$R^2$** | 0.9764 |
| | **adjusted $R^2$** | 0.9712 |
| | **RMSE** | 0.0689 |
| | ***Significance F*** | 0.000 |
| **a-Si** | **$R^2$** | 0.9237 |
| | **adjusted $R^2$** | 0.9237 |
| | **RMSE** | 0.0769 |
| | ***Significance F*** | 0.000 |
| **CdTe** | **$R^2$** | 0.9419 |
| | **adjusted $R^2$** | 0.9419 |
| | **RMSE** | 0.0701 |
| | ***Significance F*** | 0.000 |
| **CIGS** | **$R^2$** | 0.9379 |
| | **adjusted $R^2$** | 0.9379 |
| | **RMSE** | 0.0712 |
| | ***Significance F*** | 0.000 |

The efficiency of four commonly implemented solar cell technologies, m-Si, a-Si, CdTe, and CIGS, is also investigated based on the developed model during smoky wildfire periods. Table V shows that different *ARI* values are applied to compare the PV power reduction percentage for various solar cell technologies.

TABLE V
WILDFIRE SMOKE IMPACT ON DIFFERENT PV CELL TECHNOLOGIES

| PV Technology | *ARI* = 0.25 | *ARI* = 0.5 | *ARI* = 0.75 |
|---|---|---|---|
| **m-Si** | 23.74 % | 47.48 % | 71.21 % |
| **a-Si** | 27.93% | 55.86% | 83.79% |
| **CdTe** | 25.27 % | 50.53 % | 75.80 % |
| **CIGS** | 23.48 % | 46.95 % | 67.61 % |

Table V shows that power reduction on different PV technologies varies with various values of *ARI*. This is mainly because wildfire smoke has a profound reduction in both the short and visible ranges of the solar spectral irradiance [23]. This is because different solar panel technologies have different *SRs*, which depend on the bandgap of the solar cell material [48]. As a result, PV panels with low bandgap energy, such as CIGS with *SR* concentrated at longer wavelengths, can be considered for large implementation in high-risk wildfire areas.

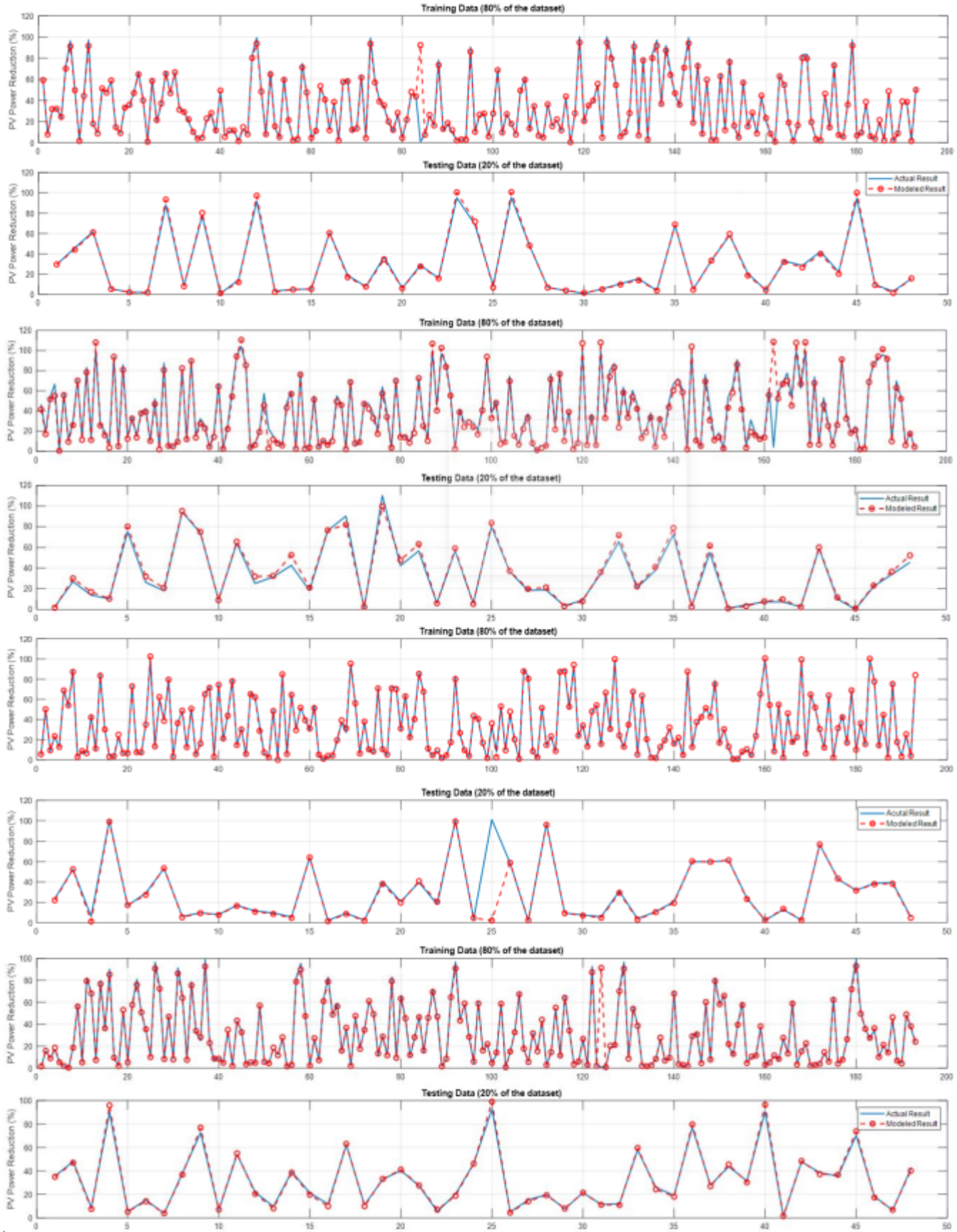



**Fig. 12.** Comparison between PV power reduction modeled results and actual results for m-Si, a-Si, CdTe, and CIGS, respectively, from July 30, 2020, to September 30, 2020 (6:00 AM to 6:00 PM)

## VI. Conclusion

The aim of this research is to investigate the impact of wildfire smoke on PV power generation through analysis of smoke optical properties. Solar spectra, radiation, and spectral irradiance data were collected and analyzed during periods of wildfire smoke, and a data-driven model was developed to quantify the effect of wildfire smoke on PV systems, enabling prediction of the generation reduction caused by smoke. Additionally, meteorological factors were evaluated for their significance during smoky days, based on the developed model. The model was also applied to investigate the impact of wildfire smoke on four commonly used solar cell technologies. The results of this study indicate that wildfire smoke reduces power generation differently for the four solar cell materials investigated, with the CIGS solar panel exhibiting the least reduction in power during smoky periods. The outcome of this paper can benefit informed decision-making for PV system operators by understanding the impact of wildfire smoke on different solar cells. Power system operators in high-risk wildfire areas can make informed decisions on which solar cell technologies to deploy to minimize the effects of smoke on solar generation. Besides, the work can be utilized by system operators to better predict and manage the effects of wildfire smoke on power grids, which can be used to maintain grid stability and reliability during smoky periods by adjusting power generation from other sources or implementing energy storage systems. This research provides valuable guidelines and identifies future research directions for assessing the impact of wildfire smoke on power systems, allowing power system operators and planners to better understand the impact of wildfire smoke on power grids with high levels of solar generation.

## References

[1] U.S. Energy Information Administration (eia), 13 September 2022. [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=427&t=3. [Accessed 4 March 2022].
[2] "Renewables 2021 Global Status Report," Paris, 2021.
[3] D. Feldman, V. Ramasamy, R. Fu, A. Ramdas, J. Desai, and R. Margolis, "U.S. Solar Photovoltaic System and Energy Storage Cost Benchmark: Q1 2020," National Renewable Energy Laboratory, January 2021.
[4] D. Feldman, K. Dummit, J. Zuboy, J. Heeter, K. Xu, and R. Margolis, "Spring 2022 Solar Industry Update," National Renewable Energy Lab (NREL), 2022.
[5] "Renewables 2021: Analysis and forecast to 2026," International Energy Agency (IEA), 2021.
[6] "Fact Sheet," FEDERAL ENERGY REGULATORY COMMISSION, 17 September 2022. [Online]. Available: https://www.ferc.gov/media/ferc-order-no-2222-fact-sheet. [Accessed 24 June 2022].
[7] "THE WHITE HOUSE," 22 APRIL 2021. [Online]. Available: https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/22/fact-sheet-president-biden-sets-2030-greenhouse-gas-pollution-reduction-target-aimed-at-creating-good-paying-union-jobs-and-securing-u-s-leadership-on-clean-energy-technologies/. [Accessed 24 September 2021].
[8] "THE LONG-TERM STRATEGY OF THE UNITED STATES: Pathways to Net-Zero Greenhouse Gas Emissions by 2050," United States Department of State and the United States Executive Office of the Presiden, Washington DC, 2021.
[9] "Climate Change Indicators in the United States: Wildfires," EPA, 2016.
[10] "Wildfire Statistics," Congressional Research Services, 2021.
[11] "Wildland Fire Summary and Statistics," National Interagency Coordination Center, 2021.
[12] "U.S. DEPARTMENT OF AGRICULTURE," 05 August 2015. [Online]. Available: https://www.usda.gov/media/press-releases/2015/08/05/forest-service-report-rising-firefighting-costs-raises-alarms#. [Accessed 23 September 2021].
[13] "2021 International Wildfires," Center for Disaster Philanthropy, October 2021. [Online]. Available: https://disasterphilanthropy.org/disaster/2021-international-wildfires/. [Accessed 23 December 2021]
[14] "Forest Fires in Europe, Middle East, and North Africa 2020," European Commission, 2021.
[15] K. Abnett, WORLD ECONOMIC FOURM, 10 December 2021. [Online]. Available: https://www.weforum.org/agenda/2021/12/siberia-america-wildfires-emissions-records-2021/. [Accessed 23 December 2021].
[16] J. Weaver and J. F. Purdom, "Some Effects of the Yellowstone Fire Smoke Plume on Northeast Colorado at the End of Summer 1988," *Journal of Climate,* vol. 117, pp. 2278-2284, 1989.
[17] S. York, "U.S. Information Administration," 30 SEPTEMBER 2020. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=45336. [Accessed 23 September 2021].
[18] J. Anderson, "US company observes solar power output declines in the Northeast due to Canadian wildfire smoke," June 12th, 2023, S&P Global Commodity Insights, [Online]. Available: https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/electric-power/061223-us-company-observes-solar-power-output-declines-in-the-northeast-due-to-canadian-wildfire-smoke# [Accessed 25 June 2023].
[19] "Quebec Wildfire Smoke Reduced Solar Output, Electricity Demand in PJM Region," June 12th, 2023, PJM Inside Lines, [Online]. Available: https://insidelines.pjm.com/quebec-wildfire-smoke-reduced-solar-output-electricity-demand-in-pjm-region/ [Accessed 25 June 2023].
[20] M. Perry and A. Troccoli, "Impact of a fire burn on solar irradiance and PV power," *Solar Energy,* vol. 114, p. 167–173, 2015.
[21] "Global Monitoring Laboratory Earth System Research Laboratories," NOAA, [Online]. Available: https://gml.noaa.gov/grad/surfrad/aod/. [Accessed 1 Jan 2022].
[22] M. Basu, N. P. Sah, C. D. Choudhuri, S. Karmakar, and T. K. Rana, "Study of the Effect of Air Pollution on Solar Power Generation Using Sun Simulator," in *2019 3rd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, India, 2019.
[23] R. L. Hulstrom and T. L. Stoffel, "Some Effects of the Yellowstone Fire Smoke Cloud on Incident Solar Irradiance," *Journal of Climate,* vol. 3, pp. 1485-1490, 1990.
[24] "Fire and Smoke Map," AirNow, [Online]. Available: https://fire.airnow.gov/. [Accessed 7 September 2022]
[25] D. L. DONALDSON, D. M. PIPER, AND D. JAYAWEERA, "Temporal Solar Photovoltaic Generation Capacity Reduction From Wildfire Smoke," *IEEE Access,* vol. 9, pp. 79841- 79852, 2021]
[26] J. J. Pe´rez-Lo´pez, F. Fabero, and F. Chenlo, "Experimental Solar Spectral Irradiance Until 2500 nm: Results and Influence on the PV Conversion of Different Materials," *PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS,* no. 15, p. 303–315, 2006
[27] "Particulate Matter (PM) Pollution," United States Enviromental Protection Agency, 18 July 2022. [Online]. Available: https://www.epa.gov/pm-pollution/particulate-matter-pm-basics. [Accessed 24 October 2022].
[28] A. Z. Bertoletti, T. Phan, and J. C. do Prado, "Wildfire Smoke, Air Quality, and Renewable Energy—Examining the Impacts of the 2020 Wildfire Season in Washington State," *sustainability,* vol. 14, no. 15, 2022.
[29] S. D. Gilletly, N. D. Jackson, and A. Staid, "Quantifying Wildfire-Induced Impacts to Photovoltaic Energy Production in the western United States," in *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, New Mexico, 2021.
[30] "Earth Observatory," [Online]. Available: https://earthobservatory.nasa.gov/images/144658/how-the-smoke-rises. [Accessed 17 12 2021].
[31] A. J. Ali, L. Zhao, M. H. Kapourchali and W. -J. Lee, "Development of A Quantification Method for The Impact of Wildfire Smoke on Photovoltaic Systems," 2023 IEEE/IAS 59th Industrial and Commercial Power Systems Technical Conference (I&CPS), Las Vegas, NV, USA, 2023, pp. 1-10, doi: 10.1109/ICPS57144.2023.10142101.
[32] "Flame Extended Range Miniature Spectrometer," Ocean Insight, [Online]. Available: https://www.oceaninsight.com/products/spectrometers/general-purpose-spectrometer/flame-series/flame-extended-range/. [Accessed 3 December 2022].

[33] Davis, [Online]. Available: https://www.davisinstruments.com/pages/vantage-pro2. [Accessed 3 December 2022].

[34] R. Garner, NASA, 27 November 2017. [Online]. Available: https://www.nasa.gov/mission_pages/sdo/science/solar-irradiance.html. [Accessed 3 December 2022].

[35] A. Andreas and T. Stoffel, "NREL Report No. DA-5500-56488," NREL Solar Radiation Research Laboratory (SRRL):Baseline Measurement System (BMS), 1981. [Online]. Available: http://dx.doi.org/10.5439/1052221.

[36] "Weather Query Builder," visualcrossing, [Online]. Available: https://www.visualcrossing.com/weather/weather-data-services. [Accessed 8 September 2022].

[37] M. G. Yazdani, M. A. Salam, and Q. M. Rahman, "Investigation of the effect of weather conditions on solar radiation in Brunei Darussalam," *International Journal of Sustainable Energy,* 2014.

[38] A. K. Shrestha, A. Thapa, and H. Gautam, "Solar Radiation, Air Temperature, Relative Humidity, and Dew Point Study: Damak, Jhapa, Nepal," *International Journal of Photoenergy,* vol. 2019, p. 7, 2019.

[39] "Multiple Linear Regression," Online. Available: http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm.

[40] S. L. Brunton and J. N. Kutz, Data-Driven Science & Engineering, Cambridge, U.K: Cambridge University Press, 2019.

[41] M. J. Schervish, "P Values: What They Are and What They Are Not," *The American Statistician,* vol. 3, pp. 203-206, 1996.

[42] "Interpreting Regression Output (Without all the Statistics Theory)," GraduateTutor.com, [Online]. Available: https://www.graduatetutor.com/statistics-tutor/interpreting-regression-output/#:~:text=Statistically%20speaking%2C%20the%20significance%20F,regression%20output%20are%20actually%20zero!. [Accessed 7 September 2022]

[43] G. K. Kanji, "Kuiper's P-value as a measuring tool and decision procedure for the goodness-of-fit test," *Journal of Applied Statistic,* vol. 15, no. 2, 1988.

[44] A. Gelman, "P Values and Statical Practice," *Epidemiology,* vol. 24, 2013

[45] G. D. Leo and F. Sardanelli, "Statistical significance: P-Value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach," *European Radiology Experimental,* 2020.

[46] S. Greenland, S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman, and D. G. Altman, "Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations," *Springer,* vol. 31, pp. 337 - 350, 2016.

[47] "Cross-Validation," MathWorks, [Online]. Available: https://www.mathworks.com/discovery/cross-validation.html#:~:text=Common%20Cross%2DValidation%20Techniques&text=k%2Dfold%3A%20Partitions%20data%20into,used%20exactly%20once%20for%20validation.. [Accessed 5 September 2022].

[48] M. A. Islam, N. M. Kassim, A. A. Alkahtani, and N. Amin, "Assessing the Impact of Spectral Irradiance on the Performance of Different Photovoltaic Technologies," *IntechOpen,* 2021.