



Lisbon School
of Economics
& Management
Universidade de Lisboa

MASTER
DATA ANALYTICS FOR BUSINESS

MASTER'S FINAL WORK
DISSERTATION

THE IMPACT OF WILDFIRE SMOKE ON SOLAR
PHOTOVOLTAIC POWER GENERATION IN ALBERTA,
CANADA

SAMANTHA TREACY

MARCH – 2024



Lisbon School
of Economics
& Management
Universidade de Lisboa

MASTER **DATA ANALYTICS FOR BUSINESS**

MASTER'S FINAL WORK **DISSERTATION**

THE IMPACT OF WILDFIRE SMOKE ON SOLAR
PHOTOVOLTAIC POWER GENERATION IN ALBERTA,
CANADA

SAMANTHA TREACY

SUPERVISION:
ALEXANDRA BUGALHO DE MOURA

MARCH – 2024

GLOSSARY

AESO - Alberta Energy System Operator

AUC – Alberta Utilities Commission

ANN - Artificial Neural Networks: A computational model inspired by the structure and functioning of the human brain, used in machine learning for pattern recognition and decision-making.

BC – British Columbia

CAAQS - Canadian Ambient Air Quality Standards

DL - Deep Learning: A subset of machine learning that involves training artificial neural networks with multiple layers (deep neural networks) on extensive datasets to perform complex tasks, enabling hierarchical feature representation and abstraction.

Extreme Fire Weather - Conditions conducive to the rapid spread and intensification of wildfires. Conditions include high temperatures, low humidity levels, strong winds, and dry vegetation, creating an elevated risk of wildfires igniting and spreading rapidly.

Fire Regime – The patterns, frequency, and intensity of wildfires.

Fire Season – The period during which environmental conditions, such as dryness and temperature, are conducive to the occurrence and spread of wildfires. In Alberta, the official fire season takes place from April 1st to Oct 31st.

FWI – Fire Weather Index: A widely used system for assessing fire weather conditions, giving a numerical rating of fire danger.

MSE - Mean Squared Error: A metric used to measure the average squared difference between predicted and actual values in regression analysis, indicating the model's accuracy.

PM2.5 – Fine particulate matter, usually consisting of smoke or pollution, with a diameter of 2.5 micrometers or smaller, measurements are expressed in micrograms per cubic meter ($\mu\text{g}/\text{m}^3$) of air.

POWER - National Aeronautics and Space Administration (NASA) Langley Research Center (LaRC) Prediction of Worldwide Energy Resource (POWER). Project funded through the NASA Earth Science/Applied Science Program.

PV – Photovoltaic: Relating to the conversion of light into electricity, commonly associated with solar power generation.

R² - The coefficient of determination, quantifies the proportion of the variance in the dependent variable that is predictable from the independent variables, indicating the model's goodness of fit.

RFR – Random Forest Regression: An ensemble learning method that builds multiple decision trees and merges their predictions to improve accuracy and reduce overfitting.

SAM – System Advisory Model: A software used for solar PV power generation simulation.

Solar Irradiation (GHI) – Global Horizontal Irradiance: The amount of solar radiation (sunlight) received per horizontal unit area over a specified period, typically expressed in units such as watts per square meter (kW/m²)

Solar Resource Potential - The amount of solar energy that can be harnessed from sunlight.

SVR – Support Vector Regression: A machine learning algorithm that utilizes support vector machines for regression tasks, seeking to find an optimal hyperplane that best fits the data points with a specified margin.

XGB - XGBoost: A machine learning algorithm that uses an ensemble of decision trees to achieve high predictive accuracy and is widely used for regression and classification tasks.

ABSTRACT, KEYWORDS AND JEL CODES

This study investigates the impact of wildfire smoke on solar photovoltaic (PV) power generation in Alberta, Canada, over a five-year period. As Alberta increasingly leverages its high solar potential to meet renewable energy demand, it is confronting the effects of climate change, particularly the escalating frequency and severity of wildfires and associated smoke emissions. Smoke has the potential to impede solar irradiation, thereby posing a significant challenge to solar energy production. Understanding the complex interaction between wildfire smoke and solar energy production is crucial for effective renewable energy planning and electricity grid management in the province. This study develops a Random Forest Regression model to forecast solar energy generation in Alberta. It incorporates PM2.5 measurements as indicators of wildfire smoke, alongside environmental parameters with strong predictive potential such as Global Horizontal Irradiance, and simulated solar energy generation data from current and future grid-connected solar sites in the province. The model is later used to isolate for the impact of smoke on solar power generation within the province. It is important to note that this study encompasses completed, proposed, and under-construction projects, as well as estimated financial data. Its aim is to assess the potential impact of smoke on Alberta's solar photovoltaic production potential, rather than provide a retrospective evaluation of historical events. Findings reveal an average 6.37% decrease in solar power production during periods of moderate to severe smoke, as compared to periods without smoke present, and that severe smoke levels are responsible for 76% of all losses. These results offer critical insights into the challenges and opportunities for renewable energy development in wildfire-prone regions and emphasize the need for proactive strategies to mitigate the impacts of extreme smoke concentrations on solar energy generation.

JEL CODES: C63, N72, P18, Q42, Q47, Q54

TABLE OF CONTENTS

Glossary	ii
Abstract, Keywords and JEL Codes	iv
Table of Tables	vii
Table of Figures	viii
Acknowledgements	ix
1. Introduction	1
1.1 Research Objectives	4
2. A Review of Wildfires & Solar PV Power Production	5
2.1 Wildfire Regime Changes & Atmospheric Indicators	5
2.3 Simulating Power Production for Research	8
2.4 Predicting Solar PV Power Production	9
2.5 Quantification Methods for Calculating PV Power Losses	11
3. Datasets & Data Preparation	13
3.1 Datasets	13
3.1.1 Solar Site Selection & PM2.5 Datasets	13
3.1.2 Simulated Solar PV Production Data	15
3.1.3 Final Dataset Aggregation	16
3.1.4 Other Data	16
3.2 Data Preparation	17
3.2.1 PM2.5	17
3.2.2 Global Horizontal Irradiance, Aerosol Optical Depth & Fire Season	19
3.2.3 Simulated Solar PV Power Production Data	19
4. The Predictive Model	20
4.1 Initial Model Selection	20
4.2 Model Creation and Tuning	21
4.3 Isolating Wildfire Smoke Impacts	22
4.4 Comparison with Innisfail Onsite Data	23
4.5 Financial Implications	23
5. Results & Discussion	23
5.1 Exploratory Analysis	23
5.2 Comparison of Innisfail Simulated & Actual Production	25
5.3 Isolating Wildfire Impacts	26

6. Conclusions & recommendations30

 6.1 Limitations31

 6.2 Recommendations & Potential for Future Research.....32

Appendix34

References39

TABLE OF TABLES

Table I - Studies quantifying the reduction in solar production from particulate matter...	34
Table II- Locations of PM2.5 stations and solar sites.....	34
Table III – System design and environmental inputs for SAM	35
Table IV – Final dataset description	35
Table V – Model training.....	36
Table VI – Leave one out cross validation results	36
Table VII – Solar production losses in periods of moderate to severe smoke.....	38

TABLE OF FIGURES

Figure 1 - Transport of smoke from BC wildfires to Alberta	2
Figure 2 - Installed generation capacity by technology in Alberta	3
Figure 3 – Solar resource potential of Canada & Alberta	3
Figure 4 – Map of Solar and PM2.5 Stations.....	15
Figure 5 – Simplified flowchart of data sources & integration, excluding financial data	17
Figure 6 - Missing Values at PM2.5 sampling stations	18
Figure 7 - Monthly average PM2.5 for each station in study	18
Figure 8 – Edmonton, May 30, 2019	19
Figure 9 - Model feature importance	22
Figure 10 - Canadian Ambient Air Quality Standards for PM2.5	23
Figure 11 – Simulated and actual PM2.5 values at Innisfail, 2021	23
Figure 12 - Correlation matrix (Spearman’s correlation coefficient)	24
Figure 13 - Weekly average of PM2.5 for all solar sites, by CAAQS.....	25
Figure 14 – Daily sum of solar power generation for all sites, scaled by daily mean PM2.5 levels.	25
Figure 15 - Comparison of monthly Innisfail satellite & onsite GHI measurements	26
Figure 16 – Comparison of monthly Innisfail simulated & onsite production data	26
Figure 17 - Yearly estimated monetary losses (\$CAD) from smoke.....	27
Figure 18 – Yearly PM2.5 fluctuations during fire season.....	27
Figure 19 - Yearly financial losses due to smoke, by PM2.5 severity.....	28
Figure 20 – Differences in power production, sized by normalized losses, coloured by financial losses	29
Figure 21 - Comparison of normalized and real losses (\$) in solar power production from smoke by location	29
Figure 22 - Reduction in power production (kW) and PM2.5 levels for Georgetown, July – August 2021	29
Figure 23 - Total financial loss in production attributable to smoke, by project stage.....	30
Figure 24 - Highest production losses due to wildfire smoke	37
Figure 25 - Highest \$ monthly losses from smoke	37
Figure 26 – Cumulative sitewide losses from smoke over the 2018-2022 fire seasons	37

ACKNOWLEDGEMENTS

First and foremost, I extend my gratitude to Professor Alexandra Bugalho De Moura for her guidance and invaluable feedback, throughout the research process. Additionally, I express heartfelt thanks to my mother for providing me background on the solar industry in Alberta, assistance in compiling datasets and for her endless patience for proofreading. My sincere thanks also go to Elemental Energy for providing essential data for this project. Finally, I am grateful for the support of my friends and family during this project.

1. INTRODUCTION

In Alberta, and across Western Canada, the escalating frequency and severity of wildfires, attributed to anthropogenic climate change, is evident (Alberta Environment and Parks, 2019; Hanes et al., 2019; X. Wang et al., 2017). The fire season, defined as “the annual period during which forest fires are likely to start, spread, and cause damage”, is extending in length (Natural Resources Canada, 2015). For instance, over a 43-year period from 1961 –2003, fire season in Alberta increased by a total of 51 days (Albert-Green et al., 2013). Fire weather, encompassing weather conditions favourable for wildfires and impacting fire behaviour, is also worsening in severity (Jain et al., 2017; Natural Resources Canada, 2015). By the end of this century, projections indicate that the area burned by wildfires globally is expected to roughly double (Flannigan et al., 2005).

Air pollution from wildfires, in the form of smoke aerosols and fine particulate matter, is transported over vast distances to areas downwind of the wildfire. In some cases smoke plumes have been observed to circumnavigate the earth (Damoah et al., 2004; Sokolik et al., 2019). Recent research suggests that the mass of aerosols and fine particulate matter released into the stratosphere from extreme fires in Western North America on a singular day in 2017 was equal to that of a moderate volcanic eruption (Peterson et al., 2018). Direct yearly carbon emissions from wildfires in Canada were equivalent to 18% of the country’s annual carbon emissions from the energy sector from 1959 – 1999, with this figure projected to rise due to the observed escalation in fire severity in recent years (Amiro et al., 2001).

Alberta, a province in Western Canada, is substantially affected by changes in the Western Canadian fire regime, referring to the patterns, frequency, and intensity of wildfires. For example, the 2016 Fort McMurray wildfire in Alberta was the most expensive Canadian disaster in modern history, costing \$5.96 billion CAD in insured losses (Hanes et al., 2019; Insurance Bureau of Canada, 2023). Figure 1 illustrates that even during periods of reduced fire activity within the province, Alberta is susceptible to the transport of smoke from fires in the neighbouring province of British Columbia and the Western USA. Wildfire smoke, from both internal and external wildfires, has emerged as a crucial air quality issue in Alberta (Alberta Environment and Parks, 2019).

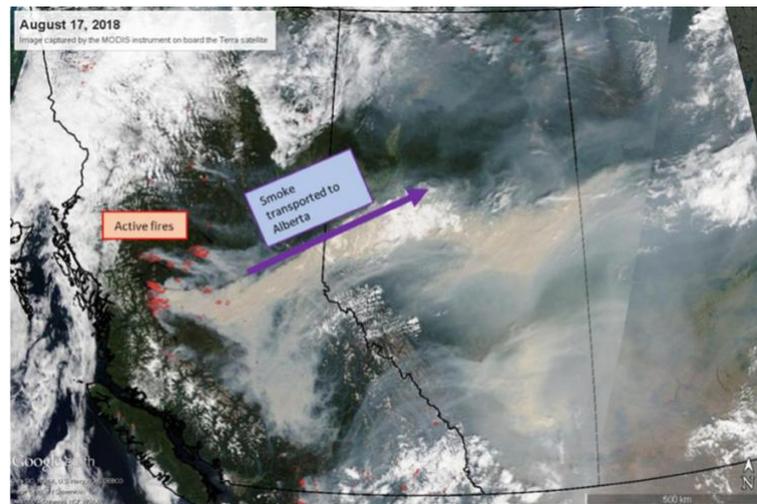


FIGURE 1 - Transport of smoke from BC wildfires to Alberta

Source: Alberta Environment and Parks, 2019

Amidst intensified wildfire events linked to climate change, a considerable global effort is underway to rapidly decarbonize and mitigate future climate disasters, including catastrophic wildfires. Decarbonization is primarily taking place through the adoption of renewable energy systems, harnessing energy from natural sources such as sunlight, wind, flowing water, biomass, waves, tides, and geothermal energy (Lai et al., 2020). Encouraging progress in the global transition is already apparent; in 2022 renewable energy contributed to 90% of the world's growth in electricity generation (Chen *et al.*, 2023). However, the transition in Alberta faces a challenging starting point. As of 2020, Alberta generated only 10% of its electricity from renewables, with the remaining 90% sourced from fossil fuels (Davis et al., 2020). Despite Alberta representing less than 10% of Canada's electricity generation in 2021, it contributed nearly half of Canada's electricity emissions (Noel & Jeyakumar, 2023).

Nevertheless, Alberta is committed to the renewable energy transition, aiming for a 30% share of electricity generation from renewables by 2030, as stated in the Renewable Energy Act (Government of Alberta, 2020). The province's shift to renewables is supported by increasing cost-effectiveness and rapid growth in renewable energy generation, Alberta accounted for 75% of Canada's utility-scale wind and solar growth in 2022 (Canadian Renewable Energy Association, 2022). Figure 2 highlights solar power generation growth in the province, while Figure 3 demonstrates Alberta's high solar resource potential, representing the amount of energy that can be captured from sunlight in a given area. The future growth of solar energy generation in Alberta is dependent on its high solar resource potential.

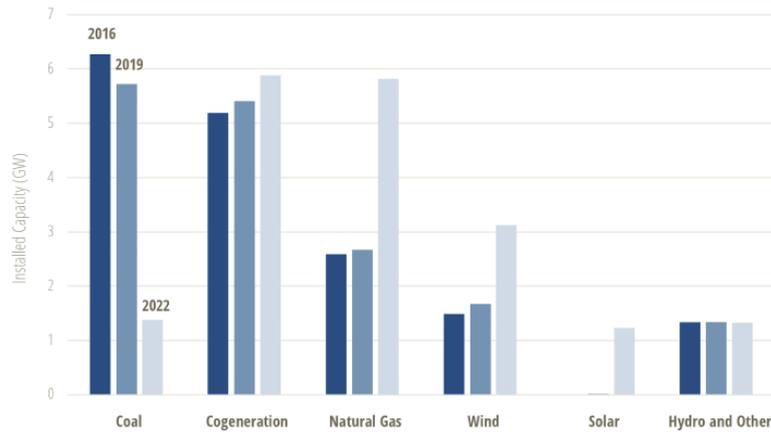


FIGURE 2 - Installed generation capacity by technology in Alberta

Note: Sourced from Singh et al., 2023, Actual 2023 data not available at time of writing.

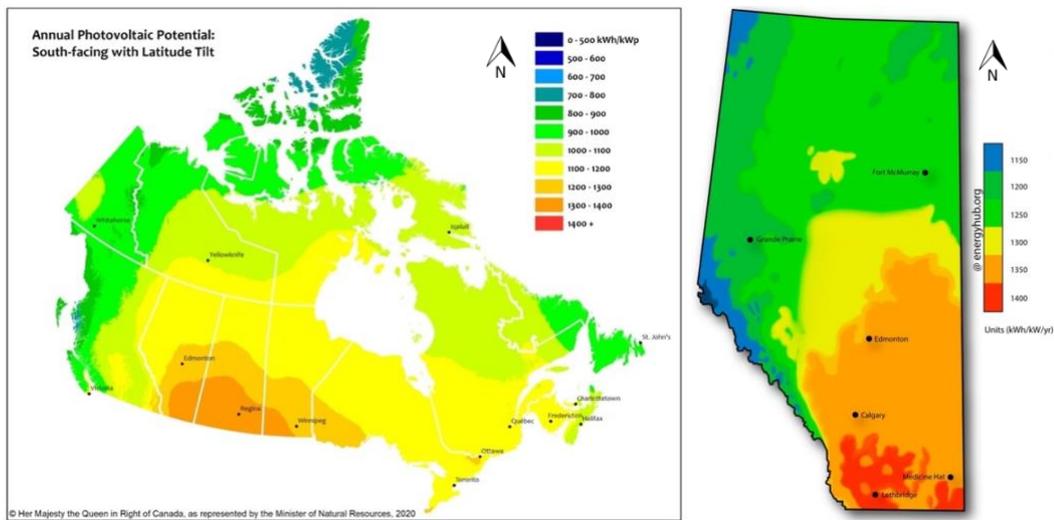


FIGURE 3 – Solar resource potential of Canada & Alberta

Note: Sourced from Natural Resources Canada, 2016, Rylen Urban, 2018

Wildfire smoke has the potential to adversely affect Alberta’s solar resource potential, and therefore, its ability to generate solar PV power - impacting one of the province’s primary paths to the decarbonization of its electricity grid. Fine particulate matter (PM2.5), composed of particles 2.5 micrometers or smaller, is the primary smoke by-product monitored in Alberta and is the largest contributor to poor air quality events in the province (Alberta Environment and Parks, 2019; H. Canada, 2021).

PM2.5 from wildfire smoke can remain in the atmosphere for months, and has been known to diminish solar radiation, the key input in solar PV power production (Keelin et al., 2021; Sokolik et al., 2019; United Nations Environment Programme, 2022). Heavy levels of smoke reduce the amount of solar irradiation that can reach solar PV panels by scattering and

absorbing sunlight (Li *et al.*, 2017; Gilletly *et al.*, 2023). In other geographic areas, wildfire smoke has been shown to reduce the amount of solar PV power produced. For example, in September 2020, during a period of severe wildfires, California solar power production was reduced 13.4% from the prior year due to smoke, despite an increase in total system capacity (California Independent System Operator, 2020). However, such an evaluation has never been carried out in Alberta, nor in Western Canada.

1.1 Research Objectives

The objective of this study is to investigate the potential impact of wildfire smoke on solar PV power generation in Alberta, Canada. Specifically, I seek to answer this question by examining smoke's effects on power production and their financial implications. Given Alberta's provincial commitment to increase renewable energy production to 30% of electricity generation, an assessment of potential barriers in achieving this goal is imperative (Government of Alberta, 2020). Considering Alberta's high potential for other renewable sources, such as wind, if wildfire smoke is found to have a significant impact on solar power production, perhaps other renewable energy methods should be considered and prioritized (Davis *et al.*, 2020).

On August 3rd, 2023, the Government of Alberta announced a seven-month pause on approvals for renewable energy projects over 1 megawatt (MW), affecting 118 different projects, including 64 GW worth of solar projects (J. Wang & Noel, Will, 2023). This decision was controversial and was deemed by many to have been made for purely political reasons. Within this context it is interesting to understand if the government had any valid reasoning for pausing solar projects within the larger context of climate change and catastrophic wildfires.

Additionally, models employed to forecast the supply and demand of power play a critical role in the management and optimization of electricity grids. While current models for forecasting solar PV power production typically incorporate weather and atmospheric conditions as inputs, they often overlook polluting aerosols such as PM_{2.5}. The omission of these aerosols can create forecasting inaccuracies, potentially leading to significant consequences for electricity grid management (Lai *et al.*, 2020). Understanding the effect that wildfire smoke has on the generation of PV solar power will lead to improvements in forecasting accuracy and a more efficient allocation of resources across the electricity grid. It

will also allow for a more accurate understanding of the long term viability of solar power production in certain geographic areas (Gómez-Amo et al., 2019).

While the effects of wildfire smoke and atmospheric aerosols (dust, pollutants) on solar PV power production have been examined in areas such as Asia and the United States, it is important to examine their effects in varied climates where solar PV technology is experiencing large scale adoption, such as Alberta. In this study I will examine the effects of wildfire smoke on solar PV power production in Alberta. Following an analysis of the current research outlook, I will construct a model to predict solar power production in Alberta, specifically focusing on isolating the effects of smoke within the model, and its hypothetical financial consequences. To the best of my knowledge, this study is the first time the impacts of wildfire smoke on solar PV production have been investigated in Canada, and specifically in Alberta.

2. A REVIEW OF WILDFIRES & SOLAR PV POWER PRODUCTION

Rapidly changing fire regimes in Western North America are contributing to an increase in the amount of smoke present in Alberta during the fire season. Research indicates that fine particulate matter from wildfire smoke adversely affects solar PV power production across geographic boundaries, and that increases in smoke may potentially diminish the amount of solar PV power generated in the future. These reductions can potentially be estimated using machine learning. In the following sections, I review the current state of research on wildfire regime changes in Alberta including their atmospheric effects, the simulation of solar power generation, commonly used machine learning techniques for predicting solar PV production and quantification methods for calculating losses from particulate matter.

2.1 Wildfire Regime Changes & Atmospheric Indicators

The report "Spreading like Wildfire" from the United Nations Environment Programme emphasizes the global nature of increasing wildfire intensity and frequency (United Nations Environment Programme, 2022). The report recognizes that by the end of the century, the likelihood of catastrophic wildfire events will increase from a factor of 1.31 to 1.57 (United Nations Environment Programme, 2022). Complimenting this, Jain et al. (2009) conducted a comprehensive review of global fire indicators, with a specific focus on North America. Their analysis reveals an uptick in extreme fire weather conditions and the lengthening of fire seasons. Notably, they report a rise in the mean value of the 99th percentile of the Canadian Fire Weather Index (FWI) from 26.1 to 27 between 1979 and 2002 (Jain et al., 2017).

Additionally, their time series analysis indicates an average increase in the length of fire seasons across North America at a rate of 3.8 days per decade, increasing from a total of 191.2 days in 1979 to 204.8 days in 2002 (Jain et al., 2017).

Examining fire season further, Albert-Green et al., (2013) investigated the seasonality of lightning caused fires in Alberta and Ontario. In agreement with previous studies, they find a significant increase in fire season length. Over a 43-year period from 1961 –2003, fire season in Alberta increased by 51 days in total, adding 19 days at the beginning of the season, and 32 at the end (Albert-Green et al., 2013). Hanes *et al.* (2019) also discuss fire season length in their analysis of fire regime changes in Canada from 1959-2015, finding that the fire season in 2015 began approximately one week earlier and ended one week later than in 1959 (Hanes et al., 2019). Together, these studies come to the undeniable conclusion that fire season in Alberta is both starting earlier and ending later.

Hanes et al. (2019) discuss changes in several other indicators including characteristics of large fires and fire cause from 1959 – 2015 in Canada. Results demonstrate that the overall size of large fires increased: fire size at the 95th percentile was approximately 57% larger in 2015 than in 1959. Human-caused fires over the time period declined, while an increase in fire starts was attributed to lightning strikes resulting from heightened severe weather events linked to climate change (Hanes et al., 2019). By the end of this century, climate scenario projections conducted by Flannigan et al. (2005) show a doubling of the area burned by wildfires.

Expanding on the theme of climate change, Parisien et al. (2023) delve into the significant surge in climate climate-induced wildfires in British Columbia since the mid-2000s. Their study reveals the compound effects of climate-induced moisture changes and altered fuels, leading to more frequent and intense wildfire seasons in British Columbia. They establish that four of the most severe wildfire seasons of the last century occurred in the past 7 years: 2017, 2018, 2021, and 2023 (Parisien et al., 2023). Two of these years, specifically 2018 and 2022, will be incorporated into my study. In addition, certain wildfires displayed extreme behaviours, including fire whirls resembling tornadoes, and the generation of lightning that sparked additional fires (Parisien et al., 2023). These occurrences are indicative of intensifying wildfire behavior under the influence of climate change.

Parisien *et al.* (2023) further illustrate that uncontained wildfires in British Columbia can burn for months at a time, resulting in widespread smoke coverage over British Columbia,

and its eastward dispersion towards Alberta, Central Canada, and parts of the USA. Wildfire-induced air pollution exacerbates climate change; in 2021, 76 billion tonnes of carbon emissions were directly attributable to wildfires (World Economic Forum, 2021). The significance of wildfire smoke's impact on air quality is underscored by the Government of Canada (2023), which states: "smoke from wildfires is a major air pollutant for Canadians, especially those in Alberta and British Columbia, and is capable of traversing hundreds, or even thousands of kilometres from the wildfire source".

The increasing severity of wildfires, and their far-travelling smoke emissions pose a significant threat to solar production by reducing the amount of sunlight able to reach the ground, impeding solar PV power generation. An assessment of smoke levels in the atmosphere is usually conducted by using fine particulate matter (PM_{2.5}) as an indicator. The United Nations Environment Programme Report, "Spreading like Wildfire"(2022), found PM_{2.5} to be the major measurable smoke by-product of concern. In Alberta, wildfires are the primary emission source of PM_{2.5}, frequently escalating to hazardous levels during the fire season, and posing a severe threat to human health (Government of Alberta, 2023c).

Aerosol Optical Depth (AOD) serves as a secondary indicator widely employed to track wildfire smoke. Sioris et al. (2017) analyzed AOD and PM_{2.5} concentrations across Canadian sites utilizing satellite measurements. They observed a strong correlation between AOD and PM_{2.5}, particularly fine-grained AOD (Sioris et al., 2017). Both Sioris et al. (2017) and Baibakov et al. (2021) agree on the classification of fine-grained aerosols as particulate matter emitted by wildfires. Baibakov et al. (2021) further investigated AOD values in Fort McMurray, located in Northern Alberta, and found mid/fine-visible AOD in the 440 nm (mid) to 550 nm (fine) wavelength range particularly sensitive to wildfire smoke. Significant increases were observed during wildfire events where "observed AOD values range from 0.1 for rural sites with background aerosols to over 3 for aerosol plumes from intense forest fires, dust storms, or volcanic eruptions" (Baibakov et al., 2021). A notable monthly record of AOD in Alberta took place during the disastrous Fort McMurray fire of 2016, with the highest hourly and daily averages being 2.10 and 1.80 respectively (Baibakov et al., 2021). Also in the vicinity of Fort McMurray, Shinozuka and Redemann (2011) measured AOD₄₉₉ values that frequently exceeded 1 and at times reached 4, due to wildfire smoke. Similarly, Calinoiu *et al.* (2013) found that aerosols resulting from biomass burning usually result in high values of AOD greater than 0.6.

A departure from conventional indicators is presented in the work of Ali et al. (2023), who questioned the reliability of PM_{2.5} and AOD as indicators of wildfire smoke. In their analysis, they introduced a novel indexing parameter, ARI (Amount of Radiation Impacted), designed to measure the solar radiation affected by suspended particles in the sky. I will not follow this approach, and instead use the conventional indicators presented above. Although there is an established correlation between PM_{2.5} from forest smoke and fine-grained AOD (0.55µm), the field lacks consensus on the comprehensive atmospheric effects of wildfire smoke, particularly its interaction with solar radiation. The extent to which smoke absorbs or scatters solar radiation remains a subject of ongoing investigation, with Sokolik et al. (2019) finding smoke to both absorb or scatter solar radiation depending on the situation.

In conclusion, the reviewed studies demonstrate a clear consensus regarding the lengthening of fire seasons and the correlation between a changing Canadian fire regime and climate change. Wildfire smoke resulting from increased fire occurrence has the potential to diminish solar power production by reducing sunlight, although the extent to which wildfire smoke's interaction with solar radiation is not yet fully understood. PM_{2.5} and AOD are established as reliable indicators of wildfire smoke and will be used in this study to quantify the levels of smoke in the atmosphere.

2.3 Simulating Power Production for Research

Simulating solar power involves modelling solar PV systems based on engineering principles to understand current, past, and future behaviour under different conditions, aiding in system design, research, financial analysis, and feasibility assessments. One publicly available simulation software, System Advisory Model (SAM), developed by the National Renewable Energy Laboratory was recommended to me by an industry expert (National Renewable Energy Laboratory, 2022). SAM is noted for its user-friendliness and accuracy, with a demonstrated accuracy rate within 4% of actual solar production levels, as shown by Gurupira & Rix (2017), comparable to other simulation methods. In this study, SAM will be employed to simulate historical solar power by incorporating historical environmental conditions and actual site specifications for solar PV generation sites in Alberta. While simulation can also be used for future predictions, its effectiveness may be limited by the extensive input data required, seen in Appendix Table III, which is frequently unavailable or challenging to obtain. Other methods, as elaborated below, are better tailored for predictive tasks.

2.4 Predicting Solar PV Power Production

The growing integration of solar power into the energy mix has prompted the need for prediction models, often using machine learning, to accurately forecast future power generation. Predictive forecasts are essential for effective grid management, optimization, and scenarios where machine learning can effectively capture complex data patterns, such as understanding the influence of individual parameters on predictions. Predicting solar PV power using machine learning typically uses historical weather and solar irradiation data. This section will outline environmental parameters with predictive potential, followed by potential machine learning models, including Support Vector Regression (SVR), Random Forest (RFR), and Extreme Gradient Boosting (XGB).

Meral & Dinçer (2011) examine factors influencing the operation and efficiency of solar PV power generation, including system technology, environmental parameters, and equipment selection. They identify sunlight, or solar irradiation (GHI), as the primary energy source for solar systems, noting temperature's significant influence on system efficiency and output (Meral & Dinçer, 2011). Mekhilef et al. (2012) reinforce the impact of temperature on solar PV power generation, emphasizing the significance of maintaining lower cell temperatures to enhance efficiency. Other atmospheric parameters, such as dust/particulate matter, humidity and wind speed were identified as key influencers of PV system performance by both Meral & Dinçer (2011) and Mekhilef et al. (2012). For instance, fine particulate matter, can accumulate on the surface of a PV module, obstructing sunlight access, and negatively impacting PV efficiency, with excessive dust deposition leading to solar cell quality deterioration; thus optimal performance is achieved by keeping the module dust-free (Mekhilef et al., 2012). Mekhilef *et al.* (2012) further observed that elevated humidity levels typically decrease solar PV panel performance. Higher wind speeds can be both beneficial and detrimental; while they reduce humidity, they also lead to more dust deposition (Mekhilef et al., 2012). In a further review of atmospheric conditions, Mellit et al. (2020) also note the dependence of solar PV power output on solar irradiance, air temperature, cloud variation, wind speed, and relative humidity.

Transitioning to the examination of machine learning strategies for predicting solar PV power output, Chahboun and Maaroufi (2021) conducted a comparison of machine learning (ML) techniques for hourly predictions of solar PV power, including MLR (Multiple Linear Regression), SVR, and RFR. They found that RFR outperformed SVR and MLR, achieving

an R-squared (R^2) value of 96% in the testing phase (Chahboun & Maaroufi, 2021). The study emphasized the effectiveness of ensemble methods (such as RFR and XGB) in predicting solar PV output. Torres-Barrán et al. (2019) extended the analysis by incorporating Gradient Boosted Regression (GBR) and XGB alongside SVR, RF, and Multilayer Perceptrons (MLPs) in their comparison, with GBR and XGB tying for first place. In a further recommendation of ensemble methods, Das et al. (2018) conducted a review of PV power forecasting models and demonstrated that ensemble methods consistently outperformed classical regression techniques. The authors recommended SVM, regression trees, and random forests as go-to methods for solar power prediction due to their promising results and suggested caution in using Artificial Neural Networks (ANN) due to high time and data costs with potential for overfitting (Das et al., 2018). Additionally, in some cases, RFR has been shown to outperform ANN in predicting solar PV power output (Kim et al., 2019).

Kim et al. (2019) proposed a two-step approach connecting unannounced weather variables with announced weather forecasts for solar power generation prediction. Their study encompassed various ML algorithms with RFR emerging as the best-performing method with an R^2 value of 70.5% in the test data (Kim et al., 2019). This study showcased the potential of ML algorithms for predicting solar PV power and emphasized the importance of considering weather information in solar power prediction. For example, most prediction methods in the reviewed research demonstrated greater effectiveness on sunny days, with forecasting accuracy diminishing on cloudy days (Mellit et al., 2020). Furthermore, Isaza et al. (2023) found that aerosol optical depth is a reliable predictor of hourly PV energy production, but only during clear sky conditions. Additionally, Rieger et al. (2017) examined the impact of Saharan mineral dust particles on the accuracy of solar production forecasts in Germany. Their study identified that including aerosol conditions as inputs into the solar PV forecast improved forecasting accuracy for 65% of stations. Although not commonly used in current models, incorporating aerosol conditions could enhance model performance (Isaza et al., 2023).

Overall, the literature highlights the complex nature of factors influencing PV cell and system efficiency. Studies identified solar irradiance, temperature, windspeed, dust deposition, humidity, clouds, and aerosol conditions as key influencers on solar power production. While aerosol conditions (indicated with PM_{2.5} or AOD) are identified as

influencing solar PV power generation, they are often not included in prediction models. The literature demonstrates that their inclusion improves the accuracy of forecasts. In addition, recent research highlights the effectiveness of ML techniques, particularly RFR and XGB, in solar PV power prediction. These ensemble methods outperform classical regression techniques, and in some cases ANNs, and showcase versatility in addressing different predictive challenges (Kim et al., 2019). While SVR has demonstrated effectiveness in predicting solar PV output, several studies indicate that RFR outperforms SVR (Chahboun & Maaroufi, 2021; Kim et al., 2019). A majority of literature considers the results from RFR, SVM and XGB to be equivalent (Das et al., 2018; Torres-Barrán et al., 2019). In my study, I aim to integrate predictive variables noted in the literature and will assess RFR, XGB, and SVR as potential model candidates.

2.5 Quantification Methods for Calculating PV Power Losses

Existing research uses various methods to reach the consensus that wildfire smoke, and other polluting atmospheric particulate matter have a negative impact on solar production, although the extent to which, and the quantification methods used differ. Studies primarily take place in the Western USA, Asia, Spain, the Sahel zone, and Australia.

In the Western USA, Gilletly et al. (2023) utilize satellite weather data, solar PV production data, and PM_{2.5} particulate matter data in training a RFR algorithm to predict solar PV production. After model training, they replaced observed PM_{2.5} with the non-wildfire average in a new dataset and assessed the change in predicted production. Their study, conducted during the severe wildfire season of 2018, found that high levels of wildfire smoke decreased solar PV production by 8.3% on average (Gilletly et al., 2023). I will follow this methodology closely in my study on Alberta. Another study in California during September 2020 reported a 13.4% decrease in solar power generation compared to the previous year, despite an increase in grid capacity, attributable to smoke (California Independent System Operator, 2020).

In China, Li et al. (2017) investigated the impact of pollutant particles on solar PV power generation using satellite AOD data and simulated solar production data. Their findings demonstrate that aerosols in the heavily polluted regions of northern and eastern China could reduce solar power production by up to 30% (Li et al., 2017). Bergin et al. (2017) also examined the effect of pollution on solar PV production, focusing on India and China, with a specific emphasis on the deposition of fine particulate matter onto solar panels. Reductions in

available solar energy (GHI) were noted to be 5-15%, with solar energy generation decreased by 17-25% (Bergin et al., 2017). In South Korea, Son et al. (2020) found that solar PV power generation diminished with deteriorating air quality, indicated by fine particulate matter (PM_{2.5}). They examined cloudless days to isolate the impact of pollution, noting a reduction in power generation of up to 29% on heavily polluted days (Son et al., 2020). All three studies observed power reductions due to particulate pollution in Asia, and used either PM_{2.5} or AOD as indicators of pollution, with Li et al. (2017) establishing a basis for using simulated solar power production data in research. While the techniques used by Bergin et al. (2017) are valid, I am unable to utilize particulate deposition data in my study due to data unavailability. Similarly, Son et al. (2020) pursued a different methodology to what I plan to implement in my study.

In Spain, Gómez-Amo et al. (2019) utilized a natural experiment involving an extreme dust episode and a nearby wildfire to assess a solar PV power plant's performance in different conditions. Compared to two previous benchmark clear days, the plant experienced an average 20% loss in production, with AOD being a strong predictor of energy loss (Gómez-Amo et al., 2019). The authors noted that current predictive models inadequately consider aerosols, which contributes to forecasting error (Gómez-Amo et al., 2019). In my study, I aim to build upon their research by integrating aerosol conditions into my model. In the African Sahel zone, spanning from the Atlantic to the Red Sea, Neher et al. (2017) simulated solar production on both clear sky and "aerosol loaded" days using a chained model approach. They observed an average 8% reduction in solar production on clear days and further established the utility of simulated production data in research (Neher et al., 2017).

In Australia, Perry & Troccoli (2015) used a planned fire burn event on a clear sky day to study the impact of smoke on both PV production and solar radiation, compared to previous clear sky days and a clear sky model. Their findings revealed a 7% average reduction in solar power generation, with a peak reduction of 27% (Perry & Troccoli, 2015). Consistent with earlier research on atmospheric effects of smoke, they observed the most significant aerosol (AOD) impact on solar irradiation at the 400 to 500 nm level (Perry & Troccoli, 2015). Also in Australia, Isaza et al. (2023) examined the effects of PM_{2.5} and AOD from wildfire on cloudless days in 2019-2020, by segmenting the data into bins of polluted and non-polluted hours. They noted that PV energy production decreased by an average of 20% in polluted conditions compared to clear conditions, with hourly reductions in output up to 65% (Isaza et

al., 2023). Again, Isaza et al. (2023) provided a basis for using PM_{2.5} and AOD as an indicator of smoke for determining power losses.

To conclude, all of the reviewed studies highlight the adverse effects of wildfire smoke on solar PV power production, with the results summarized in Appendix, Table I. Work by Bergin et al. (2017) presents further negative consequences from the deposition of fine particulate matter onto solar PV panels. However, due to limited data availability, this aspect cannot be incorporated into my research. I will adopt methodology from Gilletly et al. (2023), with Li et al. (2017) and Neher et al. (2017) providing a basis for using simulated solar production data in research. Collectively, the reviewed studies provide compelling evidence supporting the use of PM_{2.5} and/or AOD as indicators of wildfire smoke, which I aim to incorporate.

3. DATASETS & DATA PREPARATION

3.1 Datasets

In this chapter, I will elaborate on the datasets used in this study, their sources and the methodologies used for data preparation. It is worth noting that all datasets, except for the Innisfail production dataset, were acquired from publicly available resources. My study will span from 2018 to 2022, specifically chosen to encompass recent wildfire trends in Western Canada, including years marked by record breaking fire and smoke activity, such as 2018 or 2021, and comparatively milder years like 2020 (Parisien et al., 2023). While incorporating data from 2023 would have provided valuable insights, given it was the worst fire season on record with highly elevated levels of smoke in the province, the data is not available at the time of writing (Z. Wang et al., 2023). Another rationale for selecting this timeframe is the implementation of province-wide standardized PM_{2.5} sensor measurements after 2017, as detailed in subsequent sections (Government of Alberta, 2023c). All data integration and preparation was completed using Python, and associated packages, including: Geopandas, Pandas, Plotly, Matplotlib, Numpy, Seaborn, Scikit-learn and XGBoost (Chen & Guestrin, 2016; Harris, C.R. et al., 2020; John D. Hunter, 2007; Joris Van den Bossche, 2023; McKinney, 2010; Pedregosa et al., 2011; Plotly Technologies Inc., 2015; Waskom, 2021).

3.1.1 Solar Site Selection & PM_{2.5} Datasets

In my study I use PM_{2.5}, as a proxy for wildfire smoke, consistent with established research practices. PM_{2.5} is monitored by the Government of Alberta as part of the Canadian Ambient Air Quality Standards (CAAQS) (Government of Alberta, 2023c). Initial PM_{2.5}

sampling site candidates were sourced from the Alberta Air Quality Data Warehouse (2023), with selection criteria requiring data availability for the entire study period. Solar site candidates were identified using the Alberta Major Projects database, which lists infrastructure projects with budgets exceeding \$5 million, encompassing completed, proposed, or under-construction projects with their geographic coordinates (Government of Alberta, 2023b). At the time of download there were 42 PM_{2.5} station candidates and 77 solar site candidates. All solar sites, regardless of status, were included in the analysis, except for one cancelled due to the moratorium and another lacking sufficient system design information (see following section). Including both current and future solar sites will help to provide a comprehensive understanding of smoke's broader implications on solar PV power production in the province, recognizing that a larger dataset provides more representative samples and improves the reliability of findings.

To determine the inclusion of PM_{2.5} stations and solar PV sites in my study, I conducted a geographical cross-referencing of both datasets. Solar PV sites were included if they fell within a 50km radius of a PM_{2.5} sensor, while PM_{2.5} stations were selected if they were the closest to a solar station in the study. A 50km radius ensures reasonably accurate values at solar sites, and is consistent with the methodology employed in a similar study conducted by Gilletly et al., 2023. One PM_{2.5} station, Red Deer Riverside, was removed due to data quality issues and its proximity (9.4km) to another station, which could provide data for the same region. After completing the cross-referencing process, 22 solar sites and 10 PM_{2.5} sampling stations were identified, with an average distance of 22.57km between paired sites. Unsurprisingly, selected solar sites are concentrated in Southern Alberta due to its high solar PV potential, with a gap in Southeastern Alberta where solar sites are not within close range of PM_{2.5} stations. Please refer to Figure 4 for a site map, and to Appendix Table II for a detailed description of solar sites and respective PM_{2.5} stations.

Hourly PM_{2.5} data for selected stations was sourced from the Alberta Air Data Warehouse, for 2018-2022 (Government of Alberta, 2023a). Provincial measurement standards mandate that PM_{2.5} sampling methods should have an operational range of up to 500 µg m⁻³, and that sensors must not be located near other pollution sources, such as busy roads (Government of Alberta, 2016). Various sampling methods are used based on site-specific criteria such as maintenance requirements; samples obtained with each approved

method are deemed comparable by the province, and were standardized after 2017 (Government of Alberta, 2017, 2023c).



FIGURE 4 – Map of Solar and PM2.5 Stations

3.1.2 Simulated Solar PV Production Data

Due to historical production data being unavailable, I chose to simulate solar PV power production at all 22 sites. Simulation was conducted using the publicly available SAM software, specifically the PVWatts Model. Using simulated data provides several benefits. Primarily, it allows for a much more robust dataset, including sites that would not have data yet (those under construction or in the proposal phase). Secondly, this method requires much less data preparation, as real-world production data often includes site down-time and recording error. To simulate power generation using SAM, I required inputs of environmental (weather) parameters and system design data for each site, as detailed in Appendix Table III. Hourly environmental data was obtained from the NASA Langley Research Center (LaRC) POWER project funded through the NASA Earth Science/Applied Science Program, using the POWER Project's Hourly 2.5.4 version on 2022/10/30. POWER offers satellite data tailored for the renewable energy sector, and was preformatted for SAM (NASA POWER, 2021). The weather data corresponds to the location of each solar site based on the nearest satellite grid cell. Grid resolution is 1.0° latitude by 1.0° longitude for the radiation data sets and $\frac{1}{2}^{\circ}$ latitude by $\frac{5}{8}^{\circ}$ longitude for the meteorological data sets (NASA POWER, 2020).

SAM's PVWatts model required additional system design input data such as module type to simulate actual production accurately; refer to Appendix Table III for a full description. Data for each site was retrieved and compiled from application information using the Alberta Utilities Commission (AUC) application portal. As mentioned previously, one site was excluded due to the absence of system design parameters as of Fall 2023. After collecting the environmental and system design inputs, hourly solar production in kW was generated. It is important to note that SAM doesn't produce output data for Feb 29 on leap years, which will be addressed in the following section.

3.1.3 Final Dataset Aggregation

To analyze the effect of smoke on solar power production using a predictive model, I needed historical weather data at each solar site, in addition to the PM2.5 and simulated production data. For the weather data, I used parameters based upon the reviewed literature, including GHI, humidity, windspeed, temperature and precipitation. Appendix Table IV can be referenced for a full description of each parameter. Again, the data was obtained from the POWER Project's Hourly 2.5.4 version dated 2022/11/28. After all datasets were available, I first combined the PM2.5 data and SAM simulated production data, ensuring each solar station production entry had a corresponding PM2.5 value. This involved retrieving the PM2.5 measurement from the closest station at the same timestamp. Following that, I integrated this dataset with the POWER satellite weather data, merging based on location and time stamp. Before integration, the POWER weather data had to be converted to the appropriate time zone, and data on February 29 on leap years was excluded, as SAM does not provide corresponding values for this date. After aggregation the dataset had a total of 963,600 rows, with data recorded at an hourly granularity. Figure 5 displays a flowchart showing data sources, parameters, usages, and the overall aggregation process.

3.1.4 Other Data

Actual Innisfail production data, obtained from Elemental Energy, the site's owner, and operator, served to assess the accuracy of SAM's simulated solar power generation (Elemental Energy, 2024). This dataset consists of actual production values and GHI data for the Innisfail location, between July 2020 and Dec 2022.

Lastly, I needed financial data to estimate the effects in financial terms. To assess the financial impact of solar PV power generation losses due to smoke, I obtained monthly average pool price data from AESO (Alberta Energy Systems Operator), representing the

selling price of electricity per megawatt hour (AESO, 2018). The data was converted from \$/MWh to cents/kWh to match the hourly format of the other datasets and was adjusted for inflation to Canadian dollars as of December 2022, using the National Power Selling Price Index from Statistics Canada (Statistics Canada, 2024). This data will be used to estimate monetary losses in this study.

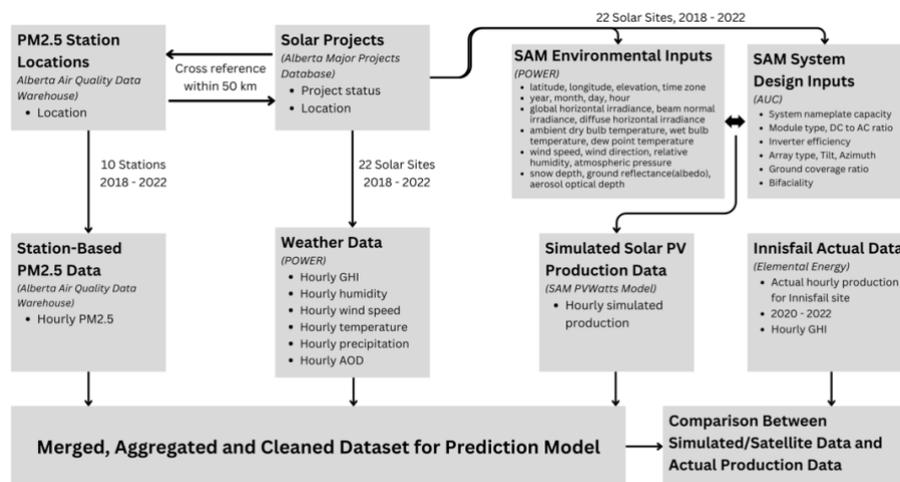


FIGURE 5 – Simplified flowchart of data sources & integration, excluding financial data

3.2 Data Preparation

In this section I will outline the steps for data cleaning and preparation of the aggregated dataset created in the previous section. This preparation is crucial for utilizing the data in a prediction model, where I will evaluate the influence of wildfire smoke on solar PV power production.

3.2.1 PM2.5

Beginning with the PM2.5 variable, I initially addressed duplicate and missing values. Some stations had multiple different PM2.5 values for a given hour, especially during periods of sampling method changes (assuming equivalence across all methods, as explained in the Data section). In these cases, the average was taken. Additionally, each PM2.5 station contained a small percentage of missing values, as seen in Figure 6, with Lethbridge having the highest number of missing values at 2.12%, which is equivalent to approximately two weeks' worth of data. Valid PM2.5 values at each station exhibited synchronized movements over time: when one station displayed high values, nearby stations typically showed similar readings (Figure 7). Given this correlation, I chose to estimate missing values using a weighted average rather than removing them, filling missing points contextually. To calculate the weighted average, I used values from other stations that had valid data at the same time

stamp. The closest station was assigned the highest weight (0.5), with weights decreasing linearly to 0 as distance increased.

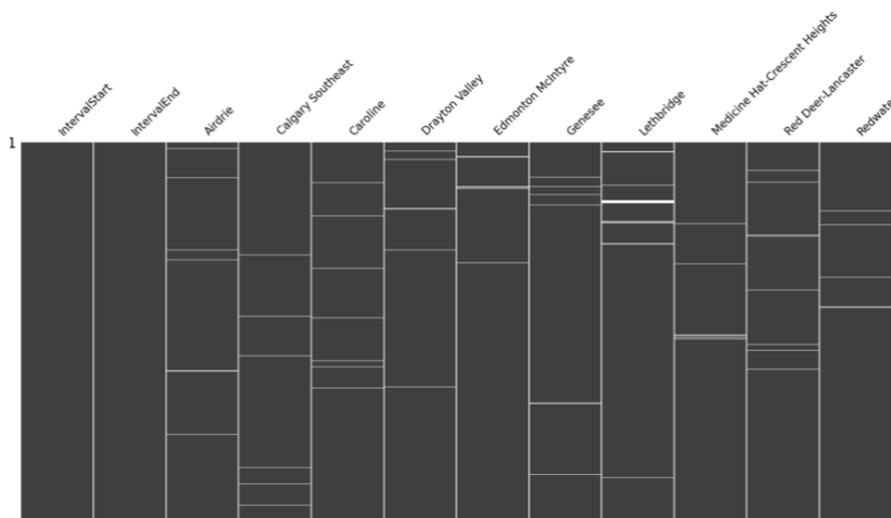


FIGURE 6 - Missing Values at PM2.5 sampling stations

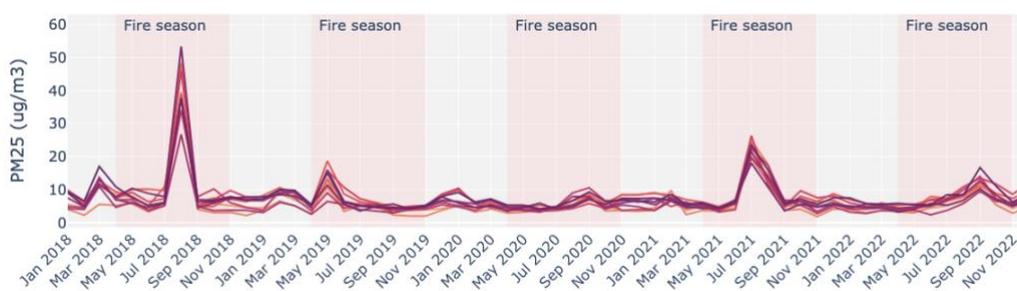


FIGURE 7 - Monthly average PM2.5 for each station in study

After addressing duplicate and missing values in the PM2.5 data, I examined its distribution and identified some outliers. While the required operating range maximum for PM2.5 measurements is 500 $\mu\text{g m}^{-3}$, there were several readings above this threshold in the dataset (Government of Alberta, 2016). These outliers all occurred on May 30, 2019, in Northern Alberta, when a severe wildfire event was taking place. Figure 8 shows imagery of the smoke levels at the time. Although these readings likely represent extreme PM2.5 levels, they fall outside the acceptable range, raising the possibility of measurement error. To address this, I replaced the outliers outside of the maximum range with the 99.99th percentile value of PM2.5 values for that location, preserving the high measurement while mitigating the risk of measurement error.



FIGURE 8 – Edmonton, May 30, 2019

Source: Ian Kucerak, 2019

3.2.2 Global Horizontal Irradiance, Aerosol Optical Depth & Fire Season

Next, I focused on cleaning and preparing the GHI and AOD columns. Since both variables are measured only in daytime hours, NASA POWER filled the unavailable nighttime values with -999, which is unsuitable for analysis. I replaced these values with 0. Additionally, after analyzing AOD, I observed data quality issues similar to Gilletly et al., (2023), who also used POWER data. As noted in academic literature, mid/fine grain AOD displays variability, ranging from under 0.1 for rural locations with minimal aerosols to over 3 during intense wildfire aerosol plumes (Baibakov et al., 2021). Most studies do not find values above 4 (Shinozuka & Redemann, 2011). Considering the research, I examined the AOD distribution in my dataset and observed 126 values exceeding 4, with 614 instances above 3. While these high values occurred during fire season, their frequency raised concerns about data quality. Furthermore, some values during peak fire season appeared unusually low, such as those in August 2018. Due to these inconsistencies, I removed AOD data from the dataset entirely. Lastly, to easily differentiate whether a point fell within fire season in visual analysis, I added a column with a categorical variable representing fire season, taking place April 1 – October 31 (Government of Alberta, 2024b).

3.2.3 Simulated Solar PV Power Production Data

Transitioning to the simulated solar PV power production data, an initial hurdle was to ensure comparability across the varied solar sites, with each featuring different equipment and production capacities. To achieve this, I applied the Min-Max normalization technique, following the methodology outlined by Gilletly et al. (2023) This process scales the values between 0 and 1 representing each site's production as a proportion of its maximum observed production during the analyzed period and creates a way to compare the production of different sites. The normalized power production will serve as the predicted variable by the

model. Next, I removed rows with non-zero GHI (sunlight) but no corresponding power generation, which could indicate potential anomalies in the simulation model. This led to the removal of 24,654 rows from the dataset, approximately 2.5% of the original. The final dataset has 938, 946 rows.

4. THE PREDICTIVE MODEL

In this chapter I will highlight my process for selecting and developing a predictive machine learning model to forecast solar PV power production in Alberta and isolating the impact of smoke on the model's predictions. I will use Python and previously listed packages, and the model will leverage the aggregated and prepared dataset detailed in the preceding sections. Six predictive parameters will be used in the model to predict normalized power - GHI, temperature, humidity, wind, PM25 and precipitation. These parameters are further detailed and noted in bold in Appendix, Table IV.

4.1 Initial Model Selection

A review of academic literature revealed various methodologies employed to predict solar PV power generation, including Support Vector Regression (SVR), Random Forest Regression (RFR), and XGBoost (XGB), Artificial Neural Networks (ANN) and Deep Learning (DL). In the literature RFR tends to outperform SVR, and in some cases, the performance of RFR, SVM, and XGBoost is deemed equivalent (Chahboun & Maaroufi, 2021; Das et al., 2018; Kim et al., 2019; Torres-Barrán et al., 2019). ANN and DL models were excluded from my study due to interpretability concerns: despite their ability to handle labeled data and perform supervised learning tasks, these models are often considered "black box" systems, making it difficult to understand the reasoning behind their predictions. Since understanding the rationale behind predictions is crucial for my study, I removed these models. Furthermore, RFR has been shown to outperform ANN in some cases of predicting solar PV output (Kim et al., 2019).

In initial model training, the first ~80% of the data, representing the first four years was designated as the training dataset, while the remaining 20%, corresponding to the last year (2022), served as the test dataset. This approach in splitting the data by years ensured a balanced representation in the training data, and prevented potential bias that could arise from only including a portion of the year in model training. The performance of the baseline model, and further iterations were evaluated using key metrics such as Mean Squared Error

(MSE) and R^2 . These metrics provide a comprehensive understanding of the model's predictive accuracy and ability to capture variance in the data.

To compare the remaining methods, I created baseline models for RFR, SVR and XGB. In this process, SVR proved impractical due to its slow training speed; in Scikit Learn SVR is only designed for datasets up to 10,000 rows, and my dataset is much larger (Pedregosa et al., 2011). For this reason, SVR was eliminated from the algorithm selection process. When evaluating the baseline model scores for RFR and XGB, RFR performed slightly better, although the differences were negligible. Despite minor variations in model performance, I selected RFR for its established effectiveness and basis in academic research.

4.2 Model Creation and Tuning

After selecting RFR as the preferred approach, the next steps were to create the model, analyze its initial effectiveness, optimize hyperparameters, and perform group-based leave one out cross validation. Throughout this process I used the previously described train-test split and evaluation metrics. The baseline model scores are presented Appendix, Table V, offering a quantitative assessment of initial model performance.

After creating the basic model, I utilized Randomized Search Cross-Validation (RandomizedSearchCV), an optimization method for hyperparameter selection, to explore model configuration settings. RandomizedSearchCV randomly samples from a defined search space, unlike exhaustive methods, reducing time and computational costs (Pedregosa et al., 2011). I used search spaces established by Torres-Barrán et al. (2019) for similar hyperparameter optimization problems. The second-ranked model, which displayed comparable accuracy to the top-ranking model but trained significantly faster, was chosen for its optimal performance within a reasonable time frame. Refer to Appendix, Table V for further details.

After optimizing hyperparameters, I assessed the model's ability to generalize on unseen location data using a site-level leave-one-group-out strategy, following methodology outlined by Gilletly et al. (2023). Instead of using a temporal train-test split, each iteration involved excluding one site as a test set, while training on the remaining sites. This technique evaluates the model's predictive accuracy for sites not encountered during the training process and gives a measure of how well the model will adapt to new sites. This process was repeated for all 22 sites, resulting in an average MSE of 0.005761 and an average R^2 of 0.94173, indicating that the model generalized well to new location data. Notably, the model

performed worst at Chappice Lake and Barlow and best at Dunmore, as seen in Appendix, Table VI. Following this satisfactory performance analysis, I trained the model on all available data.

In the final model, global horizontal irradiance is the most important feature, followed by temperature. The remaining features, as illustrated in Figure 9, demonstrate relatively lesser predictive value within the model.

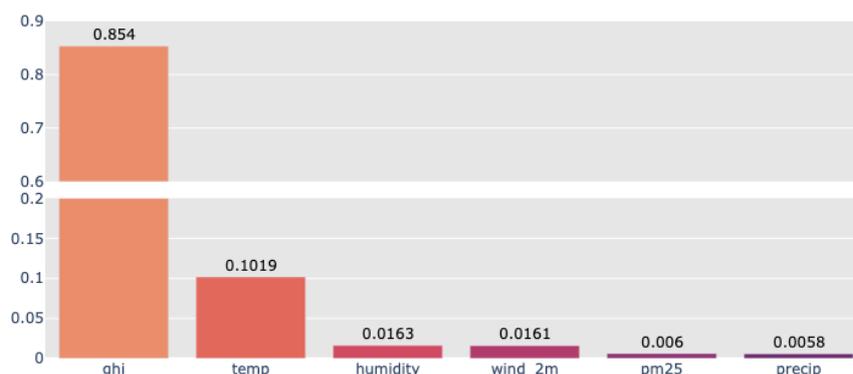


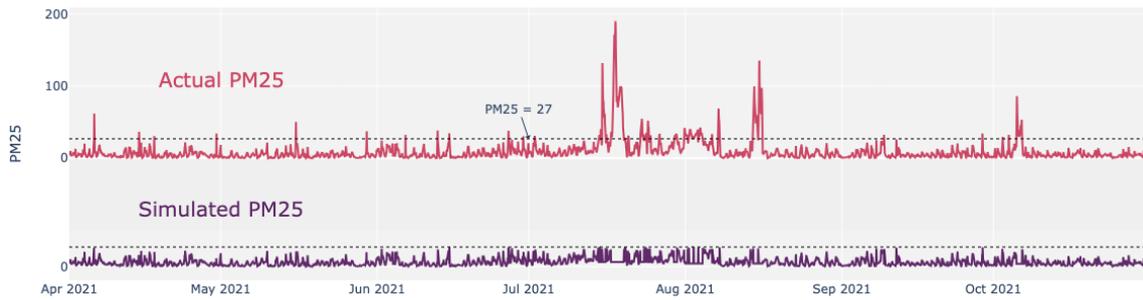
FIGURE 9 - Model feature importance

4.3 Isolating Wildfire Smoke Impacts

To analyze the impact of wildfire smoke using my model, I filtered the original dataset based on high observed PM2.5 values corresponding to the red and orange levels in the Canadian Ambient Air Quality Standards (Figure 10), and positive GHI values (theoretical production hours). Next, I predicted the amount of power produced for these smoky periods using my model. After initial prediction on the smoke dataset, I calculated the monthly mean for "clear" conditions ($PM_{2.5} \leq 10$) and substituted high observed PM2.5 values in the smoke dataset with the monthly clear value. This process yielded a dataset with synthetic PM2.5 values, effectively eliminating the presence of smoke. Lastly, predictions were made on the synthetic dataset, keeping all else constant, allowing for isolation of PM2.5’s effects on the model’s predicted power output. Figure 11 compares the appearance of actual and simulated PM2.5 values for Innisfail, an individual solar site.

Management Level	PM _{2.5} 24-hour ($\mu\text{g m}^{-3}$)		PM _{2.5} annual ($\mu\text{g m}^{-3}$)	
	2015	2020	2015	2020
Red (CAAQS)	> 28	> 27	> 10.0	> 8.8
Orange	20 to 28	20 to 27	6.5 to 10.0	6.5 to 8.8
Yellow	11 to 19	11 to 19	4.1 to 6.4	4.1 to 6.4
Green	≤ 10	≤ 10	≤ 4.0	≤ 4.0

FIGURE 10 - Canadian Ambient Air Quality Standards for PM2.5



Source: (Government of Alberta, 2024a)

FIGURE 11 – Simulated and actual PM2.5 values at Innisfail, 2021

4.4 Comparison with Innisfail Onsite Data

I visually compared the simulated production and satellite GHI data, with onsite data sourced from Elemental Energy to assess the reliability of the data used in my study and its suitability for drawing generalized conclusions on the impact of wildfire smoke on solar PV production in Alberta.

4.5 Financial Implications

Additionally, I aim to explore not only the impact of wildfire smoke on solar PV power production but also its financial repercussions. In Alberta's deregulated electricity market, participants engage in buying and selling electricity, with the pool price representing the cost per megawatt hour paid to generators (AESO, 2018). Not all power producers sell at the pool price, as some have confidential power purchase with private clients, but in this study, I will assume that all generators (solar PV producers) sell at the pool price to generate an estimate of solar PV power generation losses from wildfire smoke in financial terms. To estimate losses, hourly pool prices, adjusted for inflation, were multiplied by the predicted amount of power lost per hour, and summarized per period with all figures in Canadian dollars as of December 2022.

5. RESULTS & DISCUSSION

In this section I will discuss the outcomes of my study, which includes several components. To begin I will conduct an exploratory analysis of the data. Secondly, I will investigate the feasibility of using simulated solar production data, contrasted with actual production data. Lastly, a detailed analysis of the impact of wildfire smoke on solar power production will be completed.

5.1 Exploratory Analysis

I first aimed to understand relationships between variables by creating a correlation matrix, Figure 12, using Spearman’s correlation coefficient. My focus was primarily on examining relationships between the target variable and predictor variables. As expected, GHI and temperature both show a strong positive relationship with power_norm. Meanwhile, PM2.5 exhibits a weak negative relationship, indicating that solar power production decreases slightly responding to increases in PM2.5 levels.

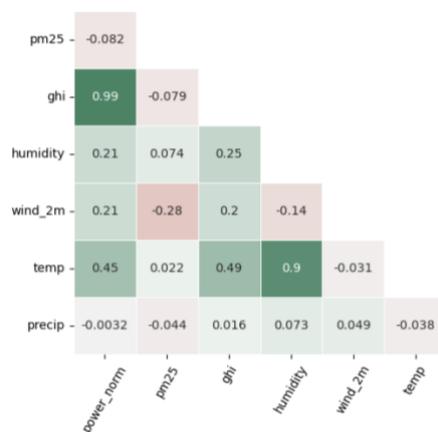


FIGURE 12 - Correlation matrix (Spearman’s correlation coefficient)

Next, I observed the fluctuations of PM2.5 throughout the study duration. Figure 13 depicts weekly average PM2.5 levels across stations throughout the study period. In this visual, spikes in PM2.5 levels during the fire season are evident, with the highest levels occurring in 2018, 2021, and some severe periods in 2019. Weekly average PM2.5 values in Alberta exceeded orange and red levels exclusively during the fire season, indicating, as expected, that the province’s most severe air pollution from PM2.5 originates from wildfire smoke. Although 2018 witnessed more pronounced spikes in PM2.5, 2021 exhibited longer durations of moderate smoke.

An analysis of simulated solar generation data from SAM reveals that days characterized by elevated PM2.5 levels rarely reach the high end of the production spectrum, as illustrated in Figure 14. This indicates a negative correlation between high PM2.5 levels and reduced solar power production, as previously illustrated by the correlation matrix (Figure 12). Solar PV power generation exhibits a clear seasonal pattern, with higher output in the summer months, coinciding with the fire season. The alignment of smoke periods and peak production times presents potential challenges for solar power production.

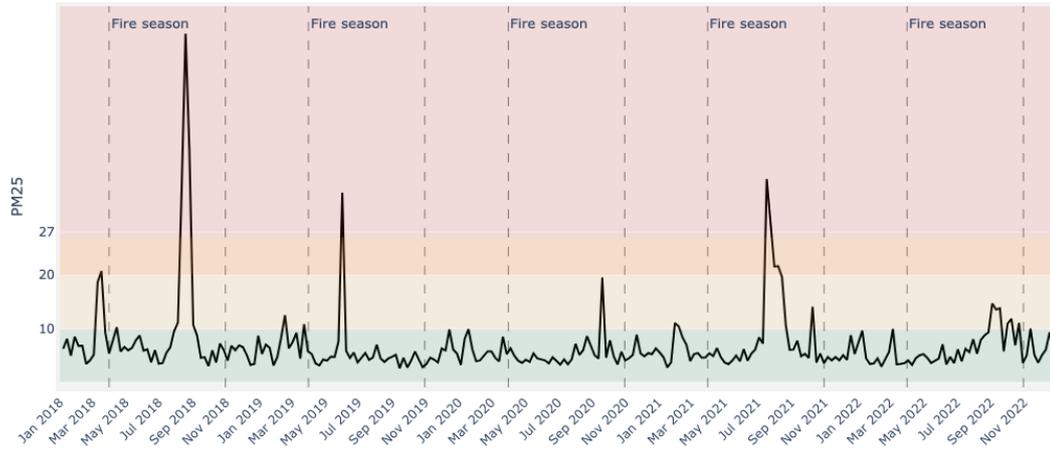


FIGURE 13 - Weekly average of PM2.5 for all solar sites, by CAAQS

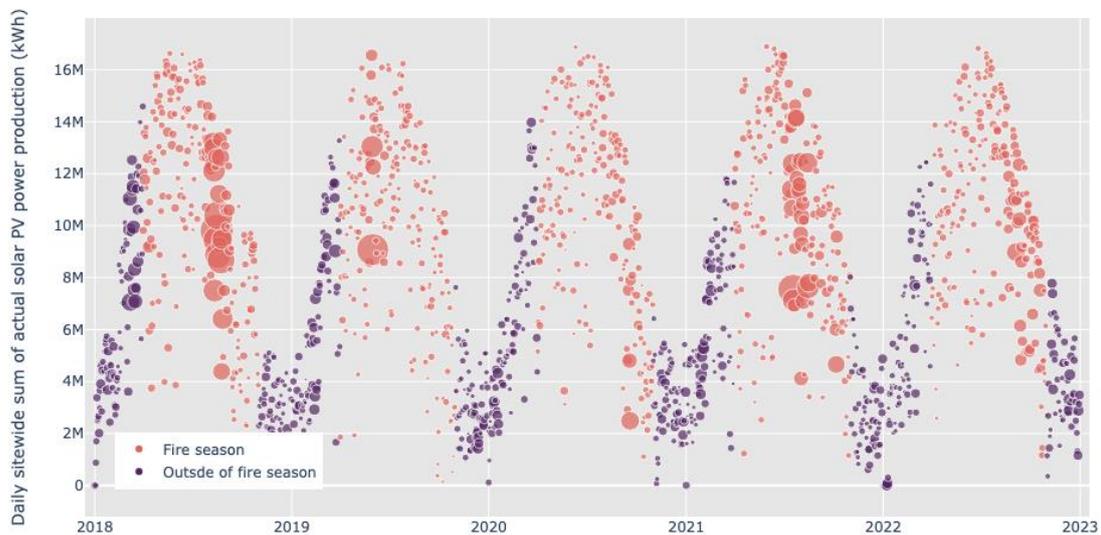


FIGURE 14 – Daily sum of solar power generation for all sites, scaled by daily mean PM2.5 levels

5.2 Comparison of Innisfail Simulated & Actual Production

Utilizing actual Innisfail production data obtained from Elemental Energy enabled me to conduct comparisons between on-site data, and data used in my study for GHI and production. Notably, the trends observed in Figure 15, displaying monthly mean GHI from satellite (POWER) data and onsite data, closely mirrored each other, with the on-site values being slightly lower, displaying minimal differences. Discrepancies may result from the grid cell not precisely aligning with the location, resulting in slight measurement variations.

Additionally, comparisons were made between monthly on-site production data and simulated production data from SAM, where there were greater variations, as illustrated in Figure 16. During the winter months, the simulated data displayed significant inaccuracies,

overestimating the power produced. This can perhaps be attributed to adverse weather conditions such as snow, affecting operational days onsite, compared to within the simulation. In contrast, the margin of error during the fire season was considerably smaller, although actual output was often underestimated. Overall, despite some inaccuracies in the production data, the narrow margin of error between actual and simulated power production during the critical period of interest for this study (fire season) suggests that the methods used in place of onsite data are sufficiently accurate for drawing conclusions.

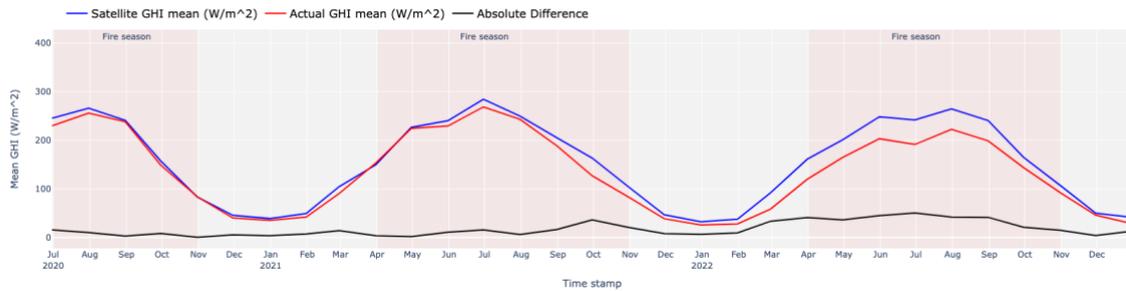


FIGURE 15 - Comparison of monthly mean Innisfail satellite & onsite GHI measurements

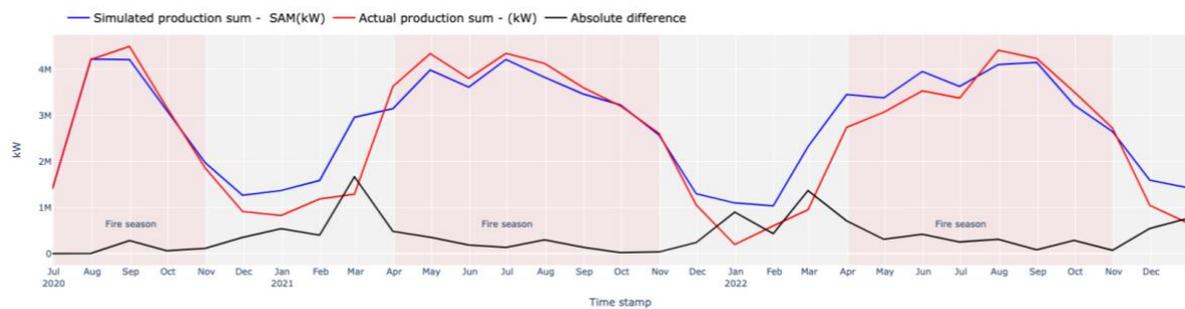


FIGURE 16 – Comparison of monthly sum of Innisfail simulated & onsite production

5.3 Isolating Wildfire Impacts

To isolate the impact of smoke on solar PV power production I compared the predictions from the smoke dataset, containing actual and simulated PM2.5 values during the fire season, keeping all else constant, as detailed in section 4.3. During hours characterized by moderate to extreme levels of smoke ($PM_{2.5} \geq 20$) in Alberta, there was a mean reduction in power generation of 6.37% when wildfire smoke was present. The cumulative effect over the fire season of all 5 years, including non-smoky periods, resulted in an overall average reduction in power generation of 0.35%, equating to an estimated hypothetical monetary loss of approximately \$3,380,965 (Appendix Table VII). These differences were found to be significant using a paired T-test at a 0.05 significance level.

2021 emerged as the year with the most significant losses in power production, both in terms of output and financial impact, despite having lower average PM2.5 levels than 2018; reference Appendix Table VII. This suggests that other factors, such as the timing of smoke events could play a role in affecting solar PV production. For example, Figure 17 shows that in 2021, periods of high smoke occurred during the longest days of the year when the most sunlight was available. Analyzing seasonal trends and variations in solar PV power production could provide further insights into the interaction between wildfire smoke and power reductions. In terms of percentage losses, as expected, 2021 and 2018 had the highest reductions in power generation over the fire season. The highest monthly production loss due to smoke took place in August 2018 at 3.39%, seen in Appendix Figure 18. In financial terms the highest hypothetical monthly reduction was \$1.1 million in July 2021 (Appendix, Figures 24 & 25), differing from the highest production loss due to the fluctuating cost of power.

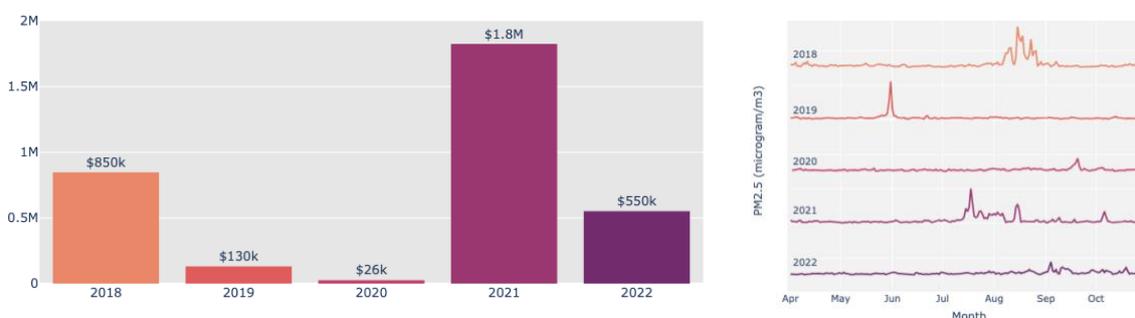


Figure 17 - Yearly estimated monetary losses (\$CAD) from smoke
Figure 18 – Yearly PM2.5 fluctuations during fire season

Next, I further examined production declines during periods of heavy smoke ($PM_{2.5} > 27$), the results of which can be examined in Appendix Table VII. In compared periods, solar PV power production exhibited an average decrease of 6.9% in the presence of severe smoke, surpassing reductions observed during moderate smoke conditions (6.36%). This suggests a correlation between smoke severity and larger power production declines. Severe smoke periods accounted for most of the overall hypothetical financial losses and power reductions, approximately 76% for both (\$2,575,691 in losses). Figure 19 displays a yearly figure of the proportional losses from severe and moderate periods of smoke. With consensus in the scientific community indicating a worsening Canadian fire regime, these power reductions are expected to escalate as severe smoke levels occur within the province more frequently.

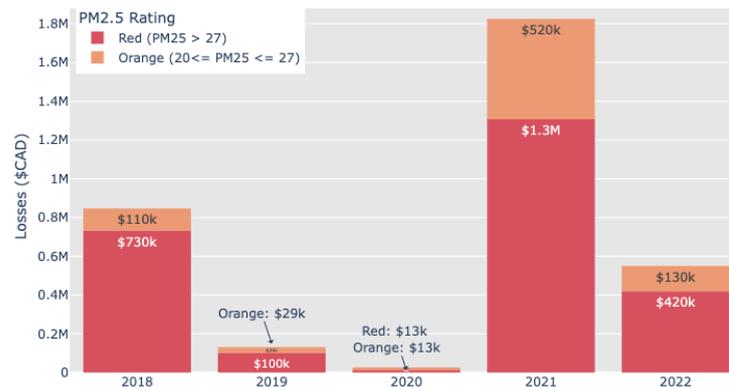


FIGURE 19 - Yearly financial losses due to smoke, by PM2.5 severity

Next, I investigated power loss by location, beginning with normalized power production, to isolate effects from the production capacity of the solar PV production site. Most sites saw comparable outcomes with those in Southern Alberta being more affected. This is evident in Figure 20, where southern solar sites generally have larger normalized production reductions and financial losses. Possible explanations include elevated levels of smoke present in Southern Alberta for the duration of the study, or the elevated solar PV potential of Southern Alberta providing stations with a higher possibility for production loss.

When examining the context of financial losses, the largest production sites were obviously more affected, and had the highest monetary losses, as depicted in Figure 21. Georgetown emerged as the site with the most substantial power loss, both in financial terms and normalized power output. In July and August 2021, Figure 22 demonstrates that as PM2.5 levels spiked, power production dropped dramatically at the site, revealing the loss in power generation attributable to wildfire smoke. In contrast, when PM2.5 levels were low there were few losses. Georgetown is currently in the proposal stage. Given the observed effects, it may be wise to consider the potential impact of smoke on this plant.

The cumulative effect of wildfire smoke on solar power generation in the fire season across all 5 years of the study, including non-smoky periods, resulted in sitewide losses ranging from 0.09% (Sollair) to 0.54% (Georgetown), as seen in Appendix, Figure 26. This is fairly comparable to the sitewide comparison of normalized losses in Figure 20 and demonstrates that all sites fall into a similar range in terms of overall losses.

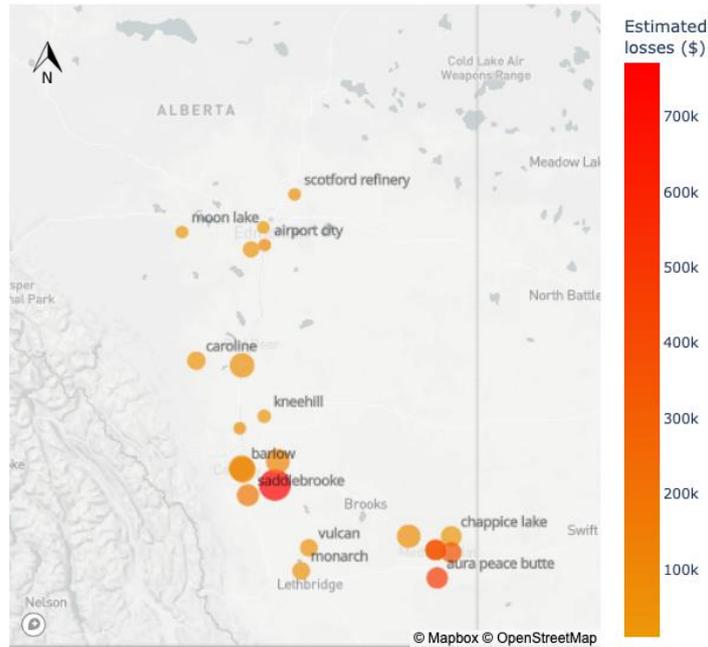


FIGURE 20 – Differences in power production, sized by normalized losses, coloured by financial losses

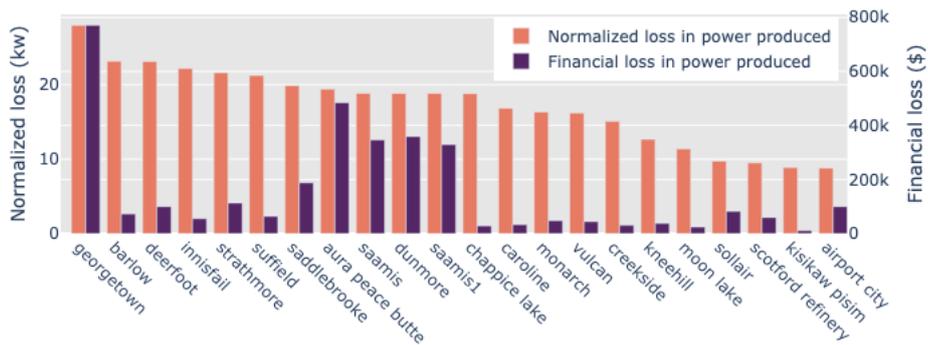


FIGURE 21 - Comparison of normalized and real losses (\$) in solar power production from smoke by location

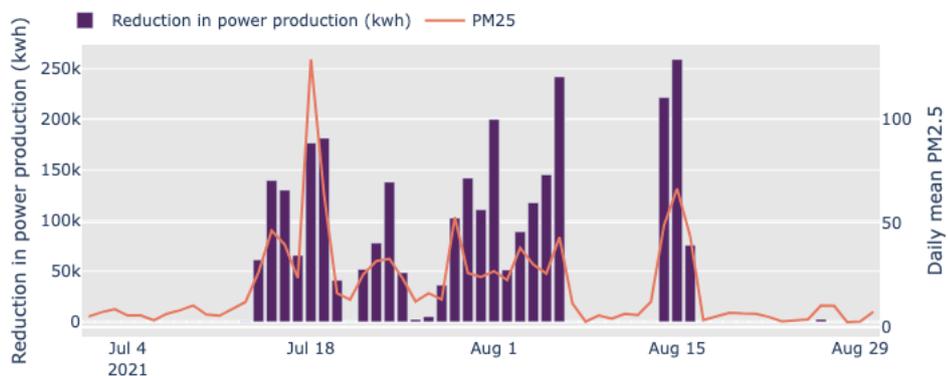


FIGURE 22 - Reduction in power production (kW) and PM2.5 levels for Georgetown, July – August 2021

In fact, several stations that are highly susceptible to potential financial losses are presently in the proposal or under construction stages (Figure 23). This can be attributed to the ongoing installation of numerous large-scale solar projects in Alberta, as solar capacity continues to grow. Previously constructed projects tended to be smaller, often serving experimental purposes, so it stands to reason that proposed or under construction projects are more financially affected due to their larger size. Nonetheless, it is advisable for project evaluators to carefully consider the potential impact of smoke on proposed projects during their assessments.

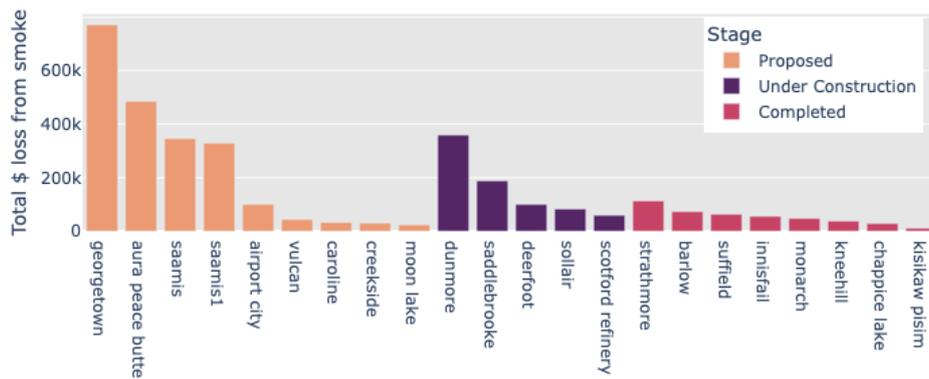


FIGURE 23 - Total financial loss in production attributable to smoke, by project stage

The findings and discussions presented in this chapter shed light on how wildfire smoke affects solar power production in Alberta. The comparative analysis at the Innisfail location reveals the viability of utilizing simulated solar and satellite data as effective substitutes for actual data. Notably, periods of wildfire smoke resulted in average power reductions of 6.37% and total financial losses amount to approximately \$3 million, with periods of heavy smoke accounting for 76% of all losses. In addition, periods of heavy smoke resulted in higher comparative losses at 6.9%. The most severe decreases occurred during the intense fire seasons of 2018 and 2021, where PM2.5 spikes from wildfire smoke were highly correlated with reductions in solar power generation. Furthermore, sites located in Southern Alberta experienced more substantial losses. Lastly, larger sites, including some that are still under construction, experienced more significant financial losses in this analysis scenario. These findings underscore the importance of accounting for the potential impacts of smoke on solar projects across Alberta.

6. CONCLUSIONS & RECOMMENDATIONS

This research examines and quantifies the impact of wildfire smoke on solar PV power generation in Alberta, finding a reduction in solar power production during periods of

wildfire smoke, which is consistent with other findings seen in Appendix, Table I. However, the study is also subject to several limitations, described below. Given the consensus on the worsening fire regime in Alberta due to climate change, coupled with significant focus on integrating renewable energy to mitigate carbon emissions, it is important to consider these results, as they could potentially impact energy policy in the province moving forward. In the following sections, policy recommendations and areas for future research are also discussed.

6.1 Limitations

Given that this study encompasses completed, proposed and under-construction projects, its purpose lies in assessing the potential impact of smoke on Alberta's solar PV power potential, rather than providing an evaluation of actual historical events. If the study had exclusively considered completed projects over the period studied, the effects in terms of power production losses in MW and financial implications would likely be smaller, as many of the largest contributing projects are not complete. Nonetheless, considering the overall similarities among sites in terms of normalized and percentage-wise production loss, the losses would likely be similar when compared in unitless terms. Additionally, the use of satellite environmental data, offsite PM_{2.5} measurements, and simulated solar PV production data will also have an impact on the precision of the results. However, given the comparative analysis conducted between simulated and actual data for Innisfail, this is not a major concern. Additionally, this study does not consider the effects of particulate matter deposition onto solar panels, which have also been shown to contribute to significant reductions in production by Bergin et al. (2017). This may lead to an underestimation of the reductions in solar PV power generation attributed to wildfire smoke demonstrated in the study.

Another limitation arises from data quality issues experienced with AOD; a commonly utilized indicator of wildfire smoke often paired with PM_{2.5}. Despite AOD's intended use in this study, it was omitted due to these data quality issues. This omission may have led to less accurate measurement of smoke in the atmosphere, although PM_{2.5} is still considered reliable on its own. Furthermore, relying on monthly pool price averages to estimate financial losses lacks precision, as certain solar PV producers negotiate special agreements for power sales at rates different from the pool price. Moreover, using a daytime average pool price for financial estimates would enhance accuracy, considering that solar PV power generation occurs exclusively during daylight hours. Unfortunately, such data was not available at the time of writing.

Despite these limitations, this study provides a functional estimate on the potential impacts of wildfire smoke on solar power generation in Alberta. While further refinement and exploration are needed to address identified constraints, the presented findings serve as a critical starting point to understand the interaction between wildfire smoke events and solar energy production in Alberta, as well as contribute to the overall body of knowledge on how wildfire smoke affects solar PV power generation.

6.2 Recommendations & Potential for Future Research

The finding that 2021 experienced the most significant losses despite having lower average PM_{2.5} levels compared to 2018 suggests that the timing of smoke events, along with seasonal and daily variations, influences reductions in solar PV power. This temporal variability should be considered in future assessments. The coinciding time of fire season and peak production periods also presents potential challenges for solar power production. Additionally, the complex relationships between smoke and other environmental factors, as observed in both the literature and this study, suggests that a deeper understanding of these dynamics could improve the ability to isolate the impact of smoke. Further research targeting these relationships, and their incorporation into similar analyses has the potential to enhance the ability to isolate the presence of smoke and predict its effects more accurately.

Additionally, while most solar PV power stations experienced similar losses in terms of normalized solar PV power production, the heightened impact in Southern Alberta raises questions about regional sensitivity to smoke. Further investigation into atmospheric patterns and transport mechanisms in Alberta could help explain why certain areas are more susceptible to smoke-induced reductions in power production. The revelation that larger production sites, notably Georgetown, suffered the most substantial losses demonstrates the vulnerability of major solar PV producers to wildfire smoke. This highlights the importance of implementing measures to enhance the resilience of critical power production facilities, a role that many solar facilities are poised to fulfill in the future amid the growing adoption of renewable power generation.

A substantial decline in solar PV power generation during severe smoke episodes, with severe smoke being the primary cause of most losses, suggests a progressively larger impact on power generation as smoke levels intensify and escalate. Policymakers and stakeholders should prioritize developing strategies to mitigate the effects of extremely high smoke

concentrations, and plan for them to happen more frequently in the future, given the worsening nature of the changing fire regime in Alberta.

The study's findings on the reductions in solar PV power generation caused by wildfire smoke in Alberta over a five-year period emphasizes the long-term implications of smoke on renewable energy planning, and electricity generation. Integrated energy planning strategies are necessary to address potential disruptions from environmental factors, ensuring the resilience and sustainability of solar PV power generation and the Albertan electricity grid.

While the study did not extensively explore the impact of the moratorium on solar power and its implications for the province, it is evident that despite the clear detrimental effects of wildfire smoke on solar power production, transitioning away from fossil fuel energy generation in Alberta remains imperative. Failure to do so would have more severe implications for the province and its future. Leveraging Alberta's substantial solar resource potential is crucial to mitigate the adverse effects of climate change to the greatest extent possible.

APPENDIX

TABLE I - Studies quantifying the reduction in solar production from particulate matter

Study	Authors	Location	Reduction in power	Data
Evaluating the impact of wildfire smoke on solar photovoltaic production	(Gilletly et al., 2023)	Western USA	~8.3%	53 sites, April 2018 – April 2019
Smoke from California wildfires decreases solar generation in CAISO	(California Independent System Operator, 2020)	California	13.4%	First two weeks of September 2020
Reduction of solar photovoltaic resources due to air pollution in China	(Li et al., 2017)	China	Up to 30%	2003 - 2014
Large Reductions in Solar Energy Production Due to Dust and Particulate Air Pollution	(Bergin et al., 2017)	China & India	~17–25%	Samples in winter 2016
Empirical estimates of the radiative impact of an unusually extreme dust and wildfire episode on the performance of a photovoltaic plant in Western Mediterranean	(Gómez-Amo et al., 2019)	Spain	~20%	26–28 and 29–30, June 2012
The effect of particulate matter on solar photovoltaic power generation over the Republic of Korea	(Son et al., 2020)	South Korea	Up to 29%	2015 - 2017
Impact of atmospheric aerosols on photovoltaic energy production Scenario for the Sahel zone	(Neher et al., 2017)	Sahel Zone	~8%	69 days in 2006
Impact of a fire burn on solar irradiance and PV power	(Perry & Troccoli, 2015)	Australia	~7%, up to 27%	4th March, 2014
Air quality impacts on rooftop photovoltaic energy production during the 2019–2020 Australian bushfires season	(Isaza et al., 2023)	Australia	Up to 20%	November 2019 – January 2020

TABLE II- Locations of PM2.5 stations and solar sites

Solar Project	Stage	Closest PM2.5 Station	Distance (km)
Airport City Solar	Proposed	Edmonton McIntyre	23.30
Aura Peace Butte Solar	Proposed	Medicine Hat-Crescent Heights	31.87
Barlow Solar Project	Completed	Calgary Southeast	0.94
Caroline Solar Project	Proposed	Caroline	16.93
Chappice Lake Solar Storage Project	Completed	Medicine Hat-Crescent Heights	25.36
Creekside Solar Project	Proposed	Genesee	25.75
Deerfoot Solar Project	Under Construction	Calgary Southeast	1.77
Dunmore Solar Project	Under Construction	Medicine Hat-Crescent Heights	17.14
Georgetown Solar Project	Proposed	Calgary Southeast	43.37
Innisfail Solar Project	Completed	Red Deer-Lancaster	25.93
Kisikaw-pisim	Completed	Edmonton McIntyre	10.07
Kneehill Solar Project	Completed	Airdrie	40.43
Monarch	Completed	Lethbridge	18.86
Moon Lake Solar Project	Proposed	Drayton Valley	22.52
Saamis Solar Farm	Proposed	Medicine Hat-Crescent Heights	3.04
Saamis Solar Farm 1	Proposed	Medicine Hat-Crescent Heights	3.04

Saddlebrook Solar and Storage Project	Under Construction	Calgary Southeast	32.88
Scotford Refinery Solar Farm	Under Construction	Redwater	16.45
Sollair Solar Energy Project	Under Construction	Airdrie	12.90
Strathmore Solar Farm	Completed	Calgary Southeast	42.77
Suffield Solar Project	Completed	Medicine Hat-Crescent Heights	39.11
Vulcan Solar Project	Proposed	Lethbridge	42.07

TABLE III – System design and environmental inputs for SAM

Variable name	Description
System nameplate capacity, kWdc	The maximum rated electrical output capacity of the solar power system under standard test conditions, measured in kilowatts direct current (kWdc).
Module type	The categorization of solar panels into standard, premium, or thin-film modules, representing different technologies and efficiency levels.
DC to AC ratio	The ratio of the direct current (DC) power generated by the solar panels to the alternating current (AC) power delivered to the electrical grid.
Inverter efficiency	The efficiency of the inverter in converting DC power generated by the solar panels into usable AC power for the electrical grid.
Array type	The configuration of the solar panel array, including options such as fixed open rack, fixed roof mount, 1-axis tracking, 1-axis backtracking, and 2-axis
Tilt, degrees	The angle at which the solar panels are tilted from the horizontal plane, measured in degrees, influencing exposure to sunlight and energy production.
Azimuth, degrees	The compass direction the solar panels face, measured in degrees, indicating the orientation of the panels towards the sun for optimal energy capture.
Ground coverage ratio (GCR)	The ratio of the total surface area covered by solar panels to the total ground area, influencing power density and land utilization. In some cases this information was not available, and a value of 0.3 is used, which is the default value in the software, and also that recommended by an industry expert.
Bifaciality	A binary variable indicating whether the solar panels are bifacial, capable of capturing sunlight from both the front and rear sides. Options include "Yes" or "No."
List of environmental inputs for SAM: latitude, longitude, time zone, elevation, year, month, day, hour, global horizontal irradiance, beam normal irradiance, diffuse horizontal irradiance, ambient dry bulb temperature, wet bulb temperature, dew point temperature, wind speed, wind direction, relative humidity, atmospheric pressure, snow depth, ground reflectance(albedo), and aerosol optical depth.	

Source: (National Renewable Energy Laboratory, 2022)

TABLE IV – Final dataset description

Parameter	Description	Abbreviation	Source	Use
Time stamp	Time	Time_stamp		Time of observation. Not used in model.
Power	Amount of power produced in kW	power	SAM	To create power norm column. Not used in model.

Location	Name of solar production site	location	Alberta Major Projects Database	To differentiate between locations. Not used in model.
PM2.5	The amount of PM2.5 particulate matter in the air	PM2.5	Alberta Air Quality Warehouse	Model
All Sky Surface Shortwave Downward Irradiance	“The total solar irradiance incident (direct plus diffuse) on a horizontal plane at the surface of the earth under all sky conditions. An alternative term for the total solar irradiance is the "Global Horizontal Irradiance" or GHI.”	GHI	NASA POWER	Model
Specific Humidity at 2M	“The ratio of the mass of water vapor to the total mass of air at 2 meters (kg water/kg total air).”	humidity	NASA POWER	Model
Wind Speed at 2 Metres	“The average of wind speed at 2 meters above the surface of the earth.”	WS2M	NASA POWER	Model
Temperature at 2M	“The average air (dry bulb) temperature at 2 meters above the surface of the earth.”	T2M	NASA POWER	Model
Precipitation Corrected	“The bias corrected average of total precipitation at the surface of the earth in water mass (includes water content in snow).”	precip	NASA POWER	Model
Fire Season	Conditional column on if it was fire season	Fire_season		To differentiate between values taking place in fire season and those not. Not used in model
Normalized Power Production	The amount of power generated at each station normalized between stations using Min-Max normalization	power_norm	SAM	Model

Note: All variable descriptions from NASA power sourced from NASA POWER (2020).

TABLE V – Model training

Model	Base (default Scikit learn parameters)	1 st place random search	2 nd place random search	3 rd place random search
n_estimators	100	1000	155	577
min_samples_split	2	40	10	15
Min_samples_leaf	2	12	8	4
max_features	1.0	0.55	0.55	0.7
max_depth	None	50	60	100
bootstrap	True	True	True	True
random_state	28	28	28	28
Mean test score (MSE), random search	NA	0.038371	0.038441	0.038495
On test data				
MSE	0.00660	0.00632	.00638	0.00641
R2	0.93453	0.93735	0.93679	0.93642
Approx. fit time	3.5 Min	15 Min	5 Min	11 Min

TABLE VI – Leave one out cross validation results

	Worst performance	Best performance
Mean MSE: 0.005761	Max MSE: Chappice Lake 0.011351	Min MSE: Dunmore 0.002441
Mean R2: 0.94173	Min R2: Barlow 0.89283	Max R2: Dunmore 0.97484

FIGURE 24 - Highest production losses due to wildfire smoke

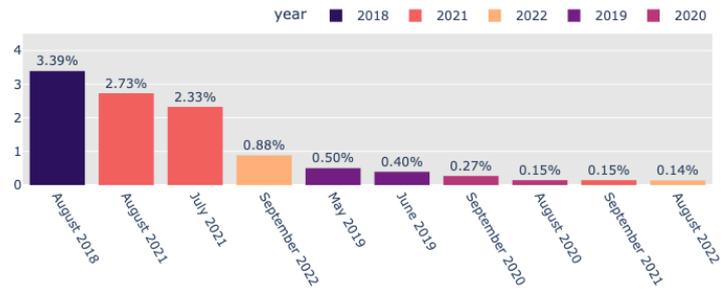


FIGURE 25 - Highest \$ monthly losses from smoke

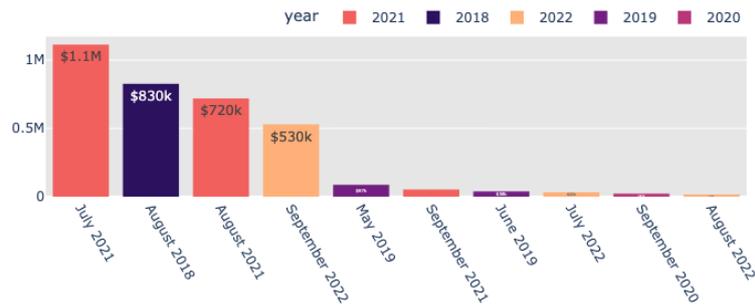


FIGURE 26 – Cumulative sitewide losses from smoke over the 2018-2022 fire seasons

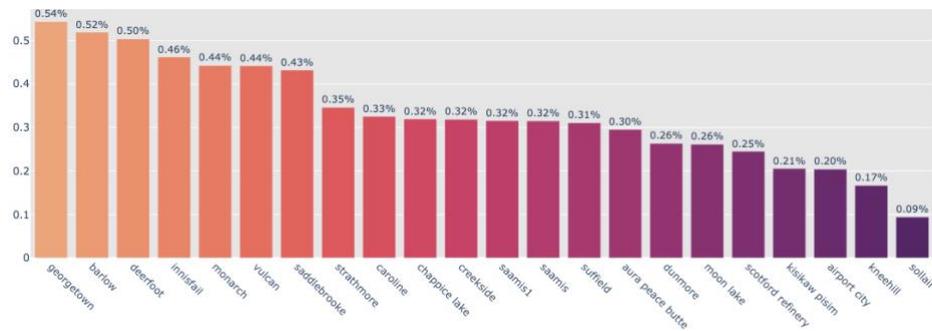


Table VII – Solar production losses in periods of moderate to severe smoke

Year	2018	2019	2020	2021	2022	Entire study
Average PM2.5 in fire season	10.46	5.63	4.87	9.1	6.27	7.26
Mean % production loss compared periods when (PM2.5 ≥ 20)	5.70%	6.86%	3.43%	8.11%	4.42%	6.37%
Mean % production loss when compared periods when PM2.5 > 27	6.29%	5.79%	3.14%	8.42%	6.57%	6.9%
Observations meeting criteria PM2.5 ≥ 20	6457	1487	1223	6707	2305	18,179
Observations meeting criteria PM2.5 > 27	4718	4706	1125	915	590	12,053
Mean reduction over entire fire season when PM2.5 ≥ 20	0.56%	0.16%	0.06%	0.84%	0.16%	0.35%
Total power loss in MW when PM2.5 ≥ 20	10,129	1727	637	15,315	1947	29,755
Proportion of losses (MW) attributable to severe smoke when PM2.5 > 27	83.37%	75.5%	47.88%	72%	75.4%	76.79%
Hypothetical \$ losses when PM2.5 ≥ 20	\$847,139	\$131,313	\$26,469	\$1,825,264	\$550,781	\$3,380,965
Hypothetical \$ losses attributable to severe smoke PM2.5 > 27	\$732,928	\$101,916	\$13,095	\$1,307,751	\$420,001	\$2,575,691
Proportion of hypothetical \$ losses attributable to severe smoke	86.52%	77.61%	49.48%	71.75%	76.26%	76.12%

REFERENCES

- AESO. (2018). *How the pool price is determined*. <https://www.aeso.ca/assets/listedfiles/How-the-Pool-Price-is-Determined-2018.pdf>
- Alberta Environment and Parks. (2019, July 1). *Wildfire smoke impacts on air quality in Alberta—Open Government*. <https://open.alberta.ca/publications/wildfire-smoke-impacts-on-air-quality-in-alberta>
- Albert-Green, A., Dean, C. B., Martell, D. L., & Woolford, D. G. (2013). A methodology for investigating trends in changes in the timing of the fire season with applications to lightning-caused forest fires in Alberta and Ontario, Canada. *Canadian Journal of Forest Research*, 43(1), 39–45. <https://doi.org/10.1139/cjfr-2011-0432>
- Ali, A. J., Zhao, L., Kapourchali, M. H., & Lee, W.-J. (2023). 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)*, 1–10. <https://doi.org/10.1109/ICPS57144.2023.10142101>
- Amiro, B. D., Todd, J. B., Wotton, B. M., Logan, K. A., Flannigan, M. D., Stocks, B. J., Mason, J. A., Martell, D. L., & Hirsch, K. G. (2001). Direct carbon emissions from Canadian forest fires, 1959-1999. *Canadian Journal of Forest Research*, 31(3), 512–525. <https://doi.org/10.1139/x00-197>
- Baibakov, K., LeBlanc, S., Ranjbar, K., O’Neill, N. T., Wolde, M., Redemann, J., Pistone, K., Li, S.-M., Liggio, J., Hayden, K., Chan, T. W., Wheeler, M. J., Nichman, L., Flynn, C., & Johnson, R. (2021). Airborne and ground-based measurements of aerosol optical depth of freshly emitted anthropogenic plumes in the Athabasca Oil Sands Region. *Atmospheric Chemistry and Physics*, 21(13), 10671–10687. <https://doi.org/10.5194/acp-21-10671-2021>

- Bergin, M. H., Ghoroi, C., Dixit, D., Schauer, J. J., & Shindell, D. T. (2017). Large Reductions in Solar Energy Production Due to Dust and Particulate Air Pollution. *Environmental Science & Technology Letters*, 4(8), 339–344.
<https://doi.org/10.1021/acs.estlett.7b00197>
- California Independent System Operator. (2020, September 30). *Smoke from California wildfires decreases solar generation in CAISO*.
<https://www.eia.gov/todayinenergy/detail.php?id=45336>
- Calinoiu, D., Paulescu, M., Ionel, I., Stefu, N., Pop, N., Boata, R., Pacurar, A., Gravila, P., Paulescu, E., & Trif-Tordai, G. (2013). Influence of aerosols pollution on the amount of collectable solar energy. *Energy Conversion and Management*, 70, 76–82.
<https://doi.org/10.1016/j.enconman.2013.02.012>
- Canada, E. and C. C. (2023, June 29). *Wildfire smoke, air quality and your health* [Education and awareness]. <https://www.canada.ca/en/environment-climate-change/services/air-quality-health-index/wildfire-smoke.html>
- Canada, H. (2021, April 14). *Infographic: What is fine particulate matter (PM2.5)?* [Promotional material]. <https://www.canada.ca/en/health-canada/services/publications/healthy-living/infographic-fine-particulate-matter.html>
- Canadian Renewable Energy Association. (2022, January 31). *NEWS RELEASE: Canada added 1.8 GW of wind and solar in 2022 – Canadian Renewable Energy Association*.
<https://renewablesassociation.ca/news-release-canada-added-1-8-gw-of-wind-and-solar-in-2022/>
- Chahboun, S., & Maaroufi, M. (2021). Performance Comparison of Support Vector Regression, Random Forest and Multiple Linear Regression to Forecast the Power of Photovoltaic Panels. *2021 9th International Renewable and Sustainable Energy Conference (IRSEC)*, 1–4. <https://doi.org/10.1109/IRSEC53969.2021.9741154>

- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Damoah, R., Spichtinger, N., Forster, C., James, P., Mattis, I., Wandinger, U., Beirle, S., Wagner, T., & Stohl, A. (2004). Around the world in 17 days—Hemispheric-scale transport of forest fire smoke from Russia in May 2003. *Atmospheric Chemistry and Physics*, 4(5), 1311–1321.
- Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I., Van Deventer, W., Horan, B., & Stojcevski, A. (2018). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, 81, 912–928. <https://doi.org/10.1016/j.rser.2017.08.017>
- Davis, M., Moronkeji, A., Ahiduzzaman, M., & Kumar, A. (2020). Assessment of renewable energy transition pathways for a fossil fuel-dependent electricity-producing jurisdiction. *Energy for Sustainable Development*, 59, 243–261. Scopus. <https://doi.org/10.1016/j.esd.2020.10.011>
- Elemental Energy. (2024). *Innisfail Solar*. Elemental Energy. <https://elementalenergy.ca/project/innisfail-solar/>
- Flannigan, M. D., Krawchuk, M. A., Groot, W. J. de, Wotton, B. M., & Gowman, L. M. (2009). Implications of changing climate for global wildland fire. *International Journal of Wildland Fire*, 18(5), 483–507. <https://doi.org/10.1071/WF08187>
- Flannigan, M. D., Logan, K. A., Amiro, B. D., Skinner, W. R., & Stocks, B. J. (2005). Future Area Burned in Canada. *Climatic Change*, 72(1), 1–16. <https://doi.org/10.1007/s10584-005-5935-y>

Gilletly, S. D., Jackson, N. D., & Staid, A. (2023). Evaluating the impact of wildfire smoke on solar photovoltaic production. *Applied Energy*, 348, 121303.

<https://doi.org/10.1016/j.apenergy.2023.121303>

Gómez-Amo, J. L., Freile-Aranda, M. D., Camarasa, J., Estellés, V., Utrillas, M. P., & Martínez-Lozano, J. A. (2019). Empirical estimates of the radiative impact of an unusually extreme dust and wildfire episode on the performance of a photovoltaic plant in Western Mediterranean. *Applied Energy*, 235, 1226–1234. Scopus.

<https://doi.org/10.1016/j.apenergy.2018.11.052>

Government of Alberta. (2016). *Air Monitoring Directive Chapter 3: Ambient Monitoring Site Selection, Siting Criteria and Sampling System Requirements. 1.*

Government of Alberta. (2017). *Air Monitoring Directive Chapter 4: Monitoring Requirements and Equipment Technical Specifications. 1.*

Government of Alberta. (2020, January 1). *Renewable Electricity Act.*

<https://open.alberta.ca/publications/r16p5#summary>

Government of Alberta. (2023a). *Alberta Air Data Warehouse.*

<https://www.alberta.ca/alberta-air-data-warehouse.aspx>

Government of Alberta. (2023b). *Alberta Major Projects.* Alberta Major Projects.

<https://majorprojects.alberta.ca/>

Government of Alberta. (2023c, April). *Air indicators – Fine particulate matter | Alberta.ca.*

<https://www.alberta.ca/air-indicators-fine-particulate-matter>

Government of Alberta. (2024a, January 17). *Canadian Ambient Air Quality Standards |*

Alberta.ca. <https://www.alberta.ca/canadian-ambient-air-quality-standards>

Government of Alberta. (2024b, January 17). *Wildfire preparedness | Alberta.ca.*

<https://www.alberta.ca/wildfire-preparedness>

- Gurupira, T., & Rix, A. (2017, January 30). *PV SIMULATION SOFTWARE COMPARISONS: PVSYST, NREL SAM AND PVLIB*.
- Hanes, C. C., Wang, X., Jain, P., Parisien, M.-A., Little, J. M., & Flannigan, M. D. (2019). Fire-regime changes in Canada over the last half century. *Canadian Journal of Forest Research*, 49(3), 256–269. <https://doi.org/10.1139/cjfr-2018-0293>
- Harris, C.R., Millman, K.J., & van der Walt, S.J. (2020). Array programming with NumPy. *Nature*, 585,(357–362). <https://doi.org/10.1038/s41586-020-2649-2>.
- Ian Kucerak. (2019). *A view from Forest Heights Park of Edmonton as forest fire smoke leads to an air quality warning on May 30, 2019*.
<https://edmontonjournal.com/news/local-news/edmonton-weather-air-quality-advisory-in-effect-as-wildfire-smoke-blankets-city>
- Insurance Bureau of Canada. (2023, January 18). *Severe Weather in 2022 Caused \$3.1 Billion in Insured Damage – making it the 3rd Worst Year for Insured Damage in Canadian History*. Insurance Bureau of Canada. <https://www.ibc.ca/news-insights/news/severe-weather-in-2022-caused-3-1-billion-in-insured-damage-making-it-the-3rd-worst-year-for-insured-damage-in-canadian-history>
- Isaza, A., Kay, M., Evans, J. P., Prasad, A., & Bremner, S. (2023). Air quality impacts on rooftop photovoltaic energy production during the 2019–2020 Australian bushfires season. *Solar Energy*, 257, 240–248.
- Jain, P., Wang, X., & Flannigan, M. D. (2017). Trend analysis of fire season length and extreme fire weather in North America between 1979 and 2015. *International Journal of Wildland Fire*, 26(12), 1009–1020. <https://doi.org/10.1071/WF17008>
- John D. Hunter. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>

Joris Van den Bossche. (2023). *geopandas/geopandas: V0.14.3*”.

<https://doi.org/10.5281/zenodo.10601680>

Keelin, P., Kubiniec, A., Bhat, A., Perez, M., Dise, J., Perez, R., & Schlemmer, J. (2021).

Quantifying the solar impacts of wildfire smoke in western North America. *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 1401–1404.

<https://doi.org/10.1109/PVSC43889.2021.9518440>

Kim, S.-G., Jung, J.-Y., & Sim, M. K. (2019). A Two-Step Approach to Solar Power Generation Prediction Based on Weather Data Using Machine Learning.

Sustainability, *11*(5), Article 5. <https://doi.org/10.3390/su11051501>

Lai, J.-P., Chang, Y.-M., Chen, C.-H., & Pai, P.-F. (2020). A Survey of Machine Learning Models in Renewable Energy Predictions. *Applied Sciences*, *10*(17), Article 17.

<https://doi.org/10.3390/app10175975>

Li, X., Wagner, F., Peng, W., Yang, J., & Mauzerall, D. L. (2017). Reduction of solar photovoltaic resources due to air pollution in China. *Proceedings of the National Academy of Sciences*, *114*(45), 11867–11872.

<https://doi.org/10.1073/pnas.1711462114>

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61.

<https://doi.org/10.25080/Majora-92bf1922-00a>

Mekhilef, S., Saidur, R., & Kamalisarvestani, M. (2012). Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. *Renewable and Sustainable Energy Reviews*, *16*(5), 2920–2925.

<https://doi.org/10.1016/j.rser.2012.02.012>

Mellit, A., Massi Pavan, A., Oglari, E., Leva, S., & Lughì, V. (2020). Advanced Methods for Photovoltaic Output Power Forecasting: A Review. *Applied Sciences*, *10*(2), Article 2.

<https://doi.org/10.3390/app10020487>

- Meral, M. E., & Dinçer, F. (2011). A review of the factors affecting operation and efficiency of photovoltaic based electricity generation systems. *Renewable and Sustainable Energy Reviews*, 15(5), 2176–2184. <https://doi.org/10.1016/j.rser.2011.01.010>
- NASA POWER. (2020, October 14). *NASA POWER / Docs*.
<https://power.larc.nasa.gov/docs/methodology/>
- NASA POWER. (2021, December 15). *NASA POWER / Docs*.
<https://power.larc.nasa.gov/docs/>
- National Renewable Energy Laboratory. (2022). *System Advisor Model Version 2022.11.29 (SAM 2022.11.21)* [Computer software]. <https://sam.nrel.gov/>
- Natural Resources Canada. (2015, September 14). *Fire weather*. Natural Resources Canada.
<https://natural-resources.canada.ca/climate-change/impacts-adaptations/climate-change-impacts-forests/forest-change-indicators/fire-weather/17776>
- Natural Resources Canada. (2016, April 7). *Photovoltaic potential and solar resource maps of Canada*. Natural Resources Canada. <https://natural-resources.canada.ca/our-natural-resources/energy-sources-distribution/renewable-energy/solar-photovoltaic-energy/tools-solar-photovoltaic-energy/photovoltaic-potential-and-solar-resource-maps-canada/18366>
- Neher, I., Buchmann, T., Crewell, S., Evers-Dietze, B., Pfeilsticker, K., Pospichal, B., Schirrmeister, C., & Meilinger, S. (2017). *Impact of atmospheric aerosols on photovoltaic energy production Scenario for the Sahel zone*. 125, 170–179. Scopus.
<https://doi.org/10.1016/j.egypro.2017.08.168>
- Noel, W., & Jeyakumar, B. (2023, June). *Zeroing In*. Pembina Institute.
[//www.pembina.org/pub/zeroing-in](https://www.pembina.org/pub/zeroing-in)

- Olivia Chen, Víctor García Tapia, & Arthur Rogé. (2023, March). *CO2 Emissions in 2022 – Analysis—IEA*. International Energy Agency. <https://www.iea.org/reports/co2-emissions-in-2022>
- Parisien, M.-A., Barber, Q. E., Bourbonnais, M. L., Daniels, L. D., Flannigan, M. D., Gray, R. W., Hoffman, K. M., Jain, P., Stephens, S. L., Taylor, S. W., & Whitman, E. (2023). Abrupt, climate-induced increase in wildfires in British Columbia since the mid-2000s. *Communications Earth & Environment*, 4(1), Article 1. <https://doi.org/10.1038/s43247-023-00977-1>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830.
- Perry, M., & Troccoli, A. (2015). Impact of a fire burn on solar irradiance and PV power. *Solar Energy*, 114, 167–173. <https://doi.org/10.1016/j.solener.2015.01.005>
- Peterson, D., Campbell, J., Hyer, E., Fromm, M., Kablick, G., Cossuth, J., & Deland, M. (2018). Wildfire-driven thunderstorms cause a volcano-like stratospheric injection of smoke. *Npj Climate and Atmospheric Science*, 1. <https://doi.org/10.1038/s41612-018-0039-3>
- Plotly Technologies Inc. (2015). Collaborative data science. *Plotly Technologies Inc.* <https://plot.ly>
- Rieger, D., Steiner, A., Bachmann, V., Gasch, P., Förstner, J., Deetz, K., Vogel, B., & Vogel, H. (2017). Impact of the 4 April 2014 Saharan dust outbreak on the photovoltaic power generation in Germany. *Atmospheric Chemistry and Physics*, 17(21), 13391–13415. <https://doi.org/10.5194/acp-17-13391-2017>

- Rylen Urban. (2018, May 12). *Solar Energy Maps Canada 2023 (Every Province)*.
Energyhub.Org; Rylan Urban. <https://www.energyhub.org/solar-energy-maps-canada/>
- Shinozuka, Y., & Redemann, J. (2011). Horizontal variability of aerosol optical depth observed during the ARCTAS airborne experiment. *Atmospheric Chemistry and Physics*, *11*(16), 8489–8495. <https://doi.org/10.5194/acp-11-8489-2011>
- Singh, K., Noel, W., & MacDougall, S. (2023). *Grid locked: Risks of unabated gas-fired electricity for a clean grid in Alberta*.
- Sioris, C. E., Abboud, I., Fioletov, V. E., & McLinden, C. A. (2017). AEROCAN, the Canadian sub-network of AERONET: Aerosol monitoring and air quality applications. *Atmospheric Environment*, *167*, 444–457.
<https://doi.org/10.1016/j.atmosenv.2017.08.044>
- Sokolik, I. N., Soja, A. J., DeMott, P. J., & Winker, D. (2019). Progress and Challenges in Quantifying Wildfire Smoke Emissions, Their Properties, Transport, and Atmospheric Impacts. *Journal of Geophysical Research: Atmospheres*, *124*(23), 13005–13025.
<https://doi.org/10.1029/2018JD029878>
- Son, J., Jeong, S., Park, H., & Park, C.-E. (2020). The effect of particulate matter on solar photovoltaic power generation over the Republic of Korea. *Environmental Research Letters*, *15*(8), 084004. <https://doi.org/10.1088/1748-9326/ab905b>
- Statistics Canada. (2024, January 9). *Electric power selling price index, monthly*.
<https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810020401>
- Torres-Barrán, A., Alonso, Á., & Dorronsoro, J. R. (2019). Regression tree ensembles for wind energy and solar radiation prediction. *Neurocomputing*, *326–327*, 151–160.
<https://doi.org/10.1016/j.neucom.2017.05.104>
- United Nations Environment Programme. (2022, February 22). *Spreading like Wildfire: The Rising Threat of Extraordinary Landscape Fires*. UNEP - UN Environment

- Programme. <http://www.unep.org/resources/report/spreading-wildfire-rising-threat-extraordinary-landscape-fires>
- Wang, J. & Noel, Will. (2023, August 24). *Investment Impact of Alberta's Renewable Energy Moratorium*. Pembina Institute. [//www.pembina.org/pub/investment-impact-albertas-renewable-energy-moratorium](http://www.pembina.org/pub/investment-impact-albertas-renewable-energy-moratorium)
- Wang, X., Parisien, M.-A., Taylor, S. W., Candau, J.-N., Stralberg, D., Marshall, G. A., Little, J. M., & Flannigan, M. D. (2017). Projected changes in daily fire spread across Canada over the next century. *Environmental Research Letters*, 12(2). Scopus. <https://doi.org/10.1088/1748-9326/aa5835>
- Wang, Z., Wang, Z., Zou, Z., Chen, X., Wu, H., Wang, W., Su, H., Li, F., Xu, W., Liu, Z., & Zhu, J. (2023). Severe Global Environmental Issues Caused by Canada's Record-Breaking Wildfires in 2023. *Advances in Atmospheric Sciences*. <https://doi.org/10.1007/s00376-023-3241-0>
- Waskom, M. L. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- World Economic Forum. (2021, December 10). *This is how much carbon wildfires have emitted this year*. World Economic Forum. <https://www.weforum.org/agenda/2021/12/siberia-america-wildfires-emissions-records-2021/>