Power System Resilience with Wildfires and Low Carbon Technologies

by

Daniel Lyle Donaldson

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Electronic, Electrical and Systems Engineering
College of Engineering and Physical Sciences
University of Birmingham
February 2022

ACKNOWLEDGEMENTS

I would like to first thank God for providing me with the wisdom, perseverance and ability to run this PhD marathon. I am grateful to my primary supervisor, Dr. Dilan Jayaweera for his support, outstanding guidance, constructive feedback, and encouragement to push beyond my limits. Also, thanks to my second supervisor, Dr. Stuart Hillmansen for his support along the way.

I would also like to thank my colleagues Dr. Hasan Gunduz, Dr. Zafar Khan, Dr. Manuel Alvarez-Alvarado, Dr. Abdullah Altamimi, Dr. Peju Oyewole, Waleed Alabri, Bader Alharbi and Wilson Vasquez for their advice, support, and many enjoyable group meetings, dinners, and discussions along the way.

I would like to thank my former colleagues at Southern California Edison for cultivating my passion for power systems analysis, for provision of data and for the helpful conversations along the way that proved invaluable in completing this research. I would like to thank my family, friends, and those at City church and Hope church for their prayer, encouragement, and constant guidance along the way. Finally, I would like to thank my wife, Aura Donaldson, for her invaluable patience, love, and encouragement to press on when things were tough.

ABSTRACT

The world is becoming increasingly reliant on electricity as a result of decarbonisation efforts, necessitating an electric grid that is resilient to natural disasters. Wildfires can adversely impact the electric grid, requiring mitigation to limit disruption to customers. Evaluation of wildfire mitigation options in the planning horizon is necessary to provide sufficient time for implementation. The complex interaction between wildfires and low carbon technologies (LCTs) like electric vehicles (EV) and solar photovoltaic (PV) generation further complicates management of the grid during the operating horizon. As residents in the wildland-urban interface evacuate, EV charging introduces additional load volatility increasing the uncertainty. Smoke also limits solar radiation, reducing the available capacity of solar PV generation integrated power systems in affected areas. This thesis presents significant contributions to enhance wildfire resilience assessment in the planning horizon, integrating the effects of LCT adoption with geospatial visualisation to strengthen proactive mitigation efforts.

In order to identify robust mitigation plans, resilience metrics should be risk-informed, as expected value metrics are likely to under-estimate the risk of interruption during a period with extreme wildfires. This thesis primarily proposes an advanced framework to quantify wildfire resilience in the planning horizon. The uniqueness of the framework is the integration of satellite-derived empirical wildfire ignitions with grid topology to produce data for numerous synthetic wildfire seasons and presenting algorithms to quantify the risk over an entire year of events, rather than focusing on a single event. Leveraging the growth of solar PV generation and EVs in wildfire mitigation plans requires current knowledge of the location and anticipated performance of PV and EV installations. This thesis also provides an advanced agglomerative approach to robustly identify the location of solar prosumers, with net smart meter data. The approach employs piecewise-aggregate approximation to reduce the computational burden of the approach without compromising the accuracy. Wildfire smoke has far reaching effects, resulting in regional attenuation of PV production. This thesis further presents an innovative methodology to predict the derate to PV production to inform generation planning. The uniqueness of the methodology is the ability to use satellite-derived aerosol optical depth data to project the derate over a wide geographical region. Awareness of EV penetration level is important as loss of diversification in EV charging behaviour during a wildfire evacuation has the potential to increase the amount of power required from the grid. This problem is addressed in the thesis by presenting a novel algorithm that integrates behavioural factors and technical parameters of EVs to reflect the EV spatio-temporal charging demand throughout the evacuation process.

Geospatial visualisations resulting from the research findings provide practical and broad insight for power system wildfire resilience planning. The research findings presented in the thesis offer new tools to enhance wildfire resilience planning assessments in modern power systems.

THESIS BASED PUBLICATIONS

[P1] **D. L. Donaldson**, D. Jayaweera. "Effective solar prosumer identification using net smart meter data," International Journal of Electrical Power & Energy Systems, vol. 118, p. 105823, 2020.

[P2] **D. L. Donaldson**, M. S. Alvarez-Alvarado, and D. Jayaweera. Power System Resiliency During Wildfires Under Increasing Penetration of Electric Vehicles. In: 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), pp.1–6

[P3] **D. L. Donaldson**, D. M. Piper, and D. Jayaweera, "Temporal solar photovoltaic generation capacity reduction from wildfire smoke," IEEE Access, vol. 9, pp. 79 841–79 852, 2021.

[P4] **D. L. Donaldson**, M. S. Alvarez-Alvarado, and D. Jayaweera. "Integration of Electric Vehicle Evacuation in Power System Resilience Assessment," IEEE Transactions of Power Systems, (*Under 1st Revision*).

Contents

List of Figures

List of Tables

Acronyms

AOD aerosol optical depth

CAISO California Independent System Operator

CGHI Clearsky Global Horizontal Irradiance

CNN Convolutional Neural Network

CVaR Conditional Value at Risk

DER Distributed Energy Resources

DNO Distribution Network Operator

EV Electric Vehicle

EVI-Pro Electric Vehicle Infrastructure Projection Tool

HILF High Impact Low Frequency

IOU Investor Owned Utility

ISO Independent System Operator

LCT Low-Carbon Technologies

LR linear regression

MAE mean absolute error

MTTR Mean Time to Repair

NERC North American Electric Reliability Corporation

NOAA National Oceanic and Atmospheric Administration

NSRDB National Solar Radiation Database

OPF Optimal Power Flow

PAA Piecewise-Aggregate Approximation

PLR piece-wise linear regression

PM2.5 particulate matter less than 2.5 μm in diameter

PV photovoltaic

QR quantile regression

RTS-GMLC Grid Modernization Laboratory Consortium update to the IEEE reliability test network

SCADA supervisory control and data acquisition

SOM self-organizing map

SPIDC Solar Prosumer Identification Duration Curve

TSO Transmission System Operator

UK United Kingdom

VaR Value at Risk

Chapter 1

Introduction

1.1 Background

There is growing dependence on the electric grid as countries seek to reduce greenhouse gas emissions. For example, in the United Kingdom (UK), the share of energy supplied by electricity is anticipated to grow from 15-20% in 2021 to 55-65% by 2050 [1]. As this dependence grows, outages of the electric grid can propagate to other sectors including digital communications, gas, transportation, waste and water, resulting in increased risk to society from failures of the grid [1, 2]. A primary cause of large scale outages affecting power system infrastructure is weather. In the United States, a study of the North American Electric Reliability Corporation (NERC) Transmission Availability Data System from 2015-2019, found that the 29 largest transmission events (ranging from 32 to 380 automatic outages) were all weather related [3]. This is further evidenced in events such as the August 2019 Great Britain power system disruption event, where a lightning strike resulted in the loss of a transmission line and around 2,100 MW of generation affecting 1 million customers [4], and the more recent 2021 winter blackout in Texas resulting in more than 10 million people without electricity and $130 billion in damages [5].

Traditionally the sufficiency of electric power systems has been assessed in terms of reliability, defined by IEEE as "the probability that a system will perform its intended

functions without failure, within design parameters, under specific operating conditions, and for a specific period of time" [6]. However, given the disproportionate impact of extreme events compared to typical component failures, the focus is shifting towards creating a grid that is also resilient to extreme events. This shift in focus is made explicit in the United Nation's Sustainable Development Goal 9: "Build resilient infrastructure, promote sustainable industrialization and foster innovation" [7]. For this thesis the CIGRE definition [8] is used which defines resilience as: "the ability to limit the extent, severity, and duration of system degradation following an extreme event."

Decarbonisation efforts are also leading to accelerated adoption of solar photovoltaic (PV) generation and Electric Vehicles (EVs). The end of 2020 saw 10 million electric cars on the road with registrations increasing by 41% [9]. In 2030, under the International Energy Agencies' Announced Pledges Scenario, projections include "annual additions of solar PV and wind approaching 500 gigawatts", and over 190 million electric cars in the global fleet, as "EVs account for over 3% of global electricity demand" [10]. Many of these resources are connected behind-the-meter, thus creating challenges for grid operators due to the lack of visibility, potentially affecting the resilience of the system during extreme events.

Increasing adoption of Low-Carbon Technologies (LCT) such as EV and solar PV generation, results in a grid that is more reliant on large shares of inverter-based resources, variable renewable energy (VRE) and distributed energy resources (DER) [11]. LCT introduce additional uncertainty to the operating patterns of the grid, adding an extra layer of complexity to the management of the grid during extreme events [11]. To account for the increased uncertainty, Transmission System Operators (TSOs) have had to develop new reserve requirements and improve generation forecast accuracy [12]. LCT can also increase susceptibility to cascading events as demonstrated by a 2016 wildfire event where 1,200 MW of solar PV generation tripped erroneously following a 500 kV fault, requiring updates to inverter settings [13]. Therefore, development of methods to assess and enhance the resilience of low-carbon power systems must consider the behaviour and performance

of LCT in order to be robust.

1.2 Research Motivation

Across extreme events affecting power systems, wildfires are unique as they can adversely affect the operation of the grid and contribute significant amounts of carbon emissions [14]. Furthermore, grid failures can also ignite wildfires, thereby causing further damage [14]. Significant wildfires in the Western United States, Brazil, Russia, and Australia evidence the global scale of wildfire related hazards [15]. In 2016, annualised economic burden from wildfires in the United States was estimated between $71.1 and $347.8 billion by the National Institute of Standards and Technology [16]. In California alone, damage from wildfires in 2018 resulted in estimated damages of $148 billion (equivalent to 1.5% of annual California GDP) [17].

Furthermore, once a wildfire has started (due to any cause) substantial disruptions can occur to power system infrastructure as seen by the 2019 Saddle Ridge Fire that resulted in 32 outages including loss of 13 generators and 30,144 MVA of transmission capacity [3]. The likelihood of overlapping significant fires is also growing in some regions of the world with the annual number of days with synchronous fire danger (defined as "fire weather indices exceeding the local 90th percentile across $\geq$40% of forested land") in the western United States expected to increase from 20 days during 1951–1980 to a multimodel median of 47 days/year by 2051–2080 with RCP4.5 forcing [18]. Developing a methodology to assess the risks that such wildfires pose to the power transmission system in the planning horizon (1-10 years) would enable the proactive implementation of climate adaptation and mitigation plans to enhance system resilience.

It is of growing importance to understand the interrelationship between LCT and wildfires, due to growing numbers of solar PV and EVs. In 2020, significant reductions in solar PV generation were observed in California as wildfires ravaged the state [19]. As reliance on solar PV generation increases, methods to predict the derate to the performance of

solar PV production during a wildfire would be beneficial to power system operators. EV also interact with wildfires and the electric grid. From 2017 to 2019, five wildfires in California each caused the evacuation of 100,000 or more people [20]. In an electrified future, more customers will be reliant on EVs to evacuate during a wildfire. While EV charging may be diversified under normal conditions, wildfire evacuations may significantly alter charging behaviour as customers simultaneously charge their EVs to evacuate. The transmission network may already be under stress from the wildfire itself, and these abnormal charging patterns could dynamically alter the loading on transmission system components and contribute to increased risk of load shedding. Therefore, methodologies to consider interrelationships between LCT and power system wildfire resilience must be considered in planning and operational studies.

1.3 Aim and Objectives

The aim of this thesis is to investigate innovative spatiotemporal approaches for power system wildfire resilience assessment, reflecting the interrelationship with growing adoption of LCT. The critical review of literature substantiating the purpose and significance of this aim and the corresponding objectives is presented in Chapter 2. To achieve this aim, the main objectives pursued in this thesis are as follow:

1. To propose a framework for assessment of power system wildfire resilience in the planning horizon.

2. To propose an innovative algorithm to identify the location of customers with behind-the-meter solar PV installations.

3. To identify salient feature sets that could be incorporated to enhance solar PV generation forecast accuracy during a wildfire.

4. To propose a novel mobility algorithm to model changes in demand during a wildfire due to electric vehicle emergency behaviour.

5. To develop spatiotemporal visualisation techniques to enhance practical understanding of resilience for power system operators and planners.

1.4 Research Contributions

The research presented in this thesis advances existing knowledge in many ways and is summarised in the following contributions.

- **The first contribution is the innovative methodology to assess power system wildfire resilience in the planning horizon.** The significance of the methodology is that it provides a risk-informed mechanism to evaluate the efficacy of wildfire mitigation and adaptation plans with sufficient time for implementation. The methodology advances the state of the art by generating data for numerous synthetic wildfire seasons, instead of focusing on a single wildfire event, enabling use of risk-based resilience metrics. Furthermore, each season is generated from empirical satellite-derived wildfire perimeters, eliminating the need for detailed future meteorological information. The methodology advances existing knowledge by quantifying key transmission assets for reinforcement and provides geospatial visualisation of these assets contributing to better understanding of where wildfire mitigation measures are needed in the planning horizon.

- **The second contribution is a spatiotemporal approach to identify possible N-k contingencies due to wildfires for incorporation into transmission planning assessments.** The significance of the approach is the identification of overlapping outages resulting from synchronous wildfires that extend beyond the planning criteria. The uniqueness of the approach is the use of empirical wildfire data to produce a list of N-k contingencies capturing outages that may occur over numerous wildfire seasons. Thus, the approach advances the state of the art of the knowledge in wildfire resilience assessment by providing an innovative way to capture high impact low frequency wildfire outages to include in annual transmission

reliability assessments.

- **The third contribution is an innovative solar prosumer identification approach.** The significance of the approach is the ability to accurately identify solar prosumers requiring only hourly net smart meter data. In addition to requiring less data than existing methods, Piecewise-Aggregate Approximation (PAA) reduces the computational burden (by $\sim 100\times$) without sacrificing accuracy, further reflecting the effectiveness. Solar Prosumer Identification Duration Curves (SPIDCs) give insight into the duration of historical data required to reach a target confidence level, enhancing practical decision-making. This contribution resulted in the publication given in [P1] and provides a means to obtain current estimates of solar PV installations to support restoration and mitigation planning efforts.

- **The fourth contribution is a robust methodology to predict the derate to Solar PV production during a wildfire.** The methodology yields a trained quantile regression model that captures the relationship between aerosol optical depth (AOD) and PV generation capacity during a wildfire. The uniqueness comes from the use of satellite derived AOD forecasts to produce a spatial indication of the derate to PV production over entire planning areas due to wildfire smoke. Therefore, the methodology advances the knowledge of PV behaviour modelling, and provides an improved basis for generation scheduling, dispatch, and setting of reserve requirements during a wildfire. This contribution resulted in the publication given in [P3].

- **The fifth contribution is a novel EV evacuation model reflecting the anticipated behaviour of residential EV during a wildfire.** The novelty comes from the integration of EV fleet characteristics with safety, suitability, and proximity factors to reflect spatiotemporal impacts to transmission system loading. Thus the state of the art is advanced through more refined and effective EV charging profiles at each substation reflecting charging before, during and after a wildfire

evacuation with increased modelling accuracy. The proposed model enhances the ability to anticipate abnormal charging patterns (as the peak charging demand at the evacuating substation could be nearly triple its normally diversified peak charging demand), supporting proactive generation dispatch and reserve requirements. This contribution resulted in the publication given in [P2].

1.5 Thesis Outline

This thesis contains six additional chapters with Chapter 2 providing a review of literature, Chapters 3-6 providing the unique research contributions, and Chapter 7 providing conclusions and possible areas for future research. The relationship between the chapters is presented in Figure 1.1 with the content of each chapter outlined below.

Chapter 2 presents a critical review and summary of existing literature relevant to the research. First this chapter provides foundational understanding regarding the shift in focus from maintaining a reliable power system, to one that is also resilient to extreme events. Next the chapter examines the threat posed to power system infrastructure from wildfires along with existing methods to assess wildfire resilience. Finally the chapter evaluates the interrelationships between wildfires and low carbon technologies. For each area, challenges that underscore the need for the research conducted in this thesis are highlighted.

Chapter 3 presents a methodology to quantify power system wildfire resilience in the planning horizon. The methodology first generates many synthetic wildfire seasons from historical wildfire perimeters collected via satellite. Next the impacts of wildfires for each season on the transmission infrastructure are quantified using risk-based metrics. Finally, transmission lines most at risk are identified along with anticipated outages. The risk is used to inform further analysis in Chapter 6. Case studies are conducted to demonstrate the impact of infrastructure hardening and enhanced restoration speed.

Chapter 4 presents an advanced agglomerative clustering-based method to identify

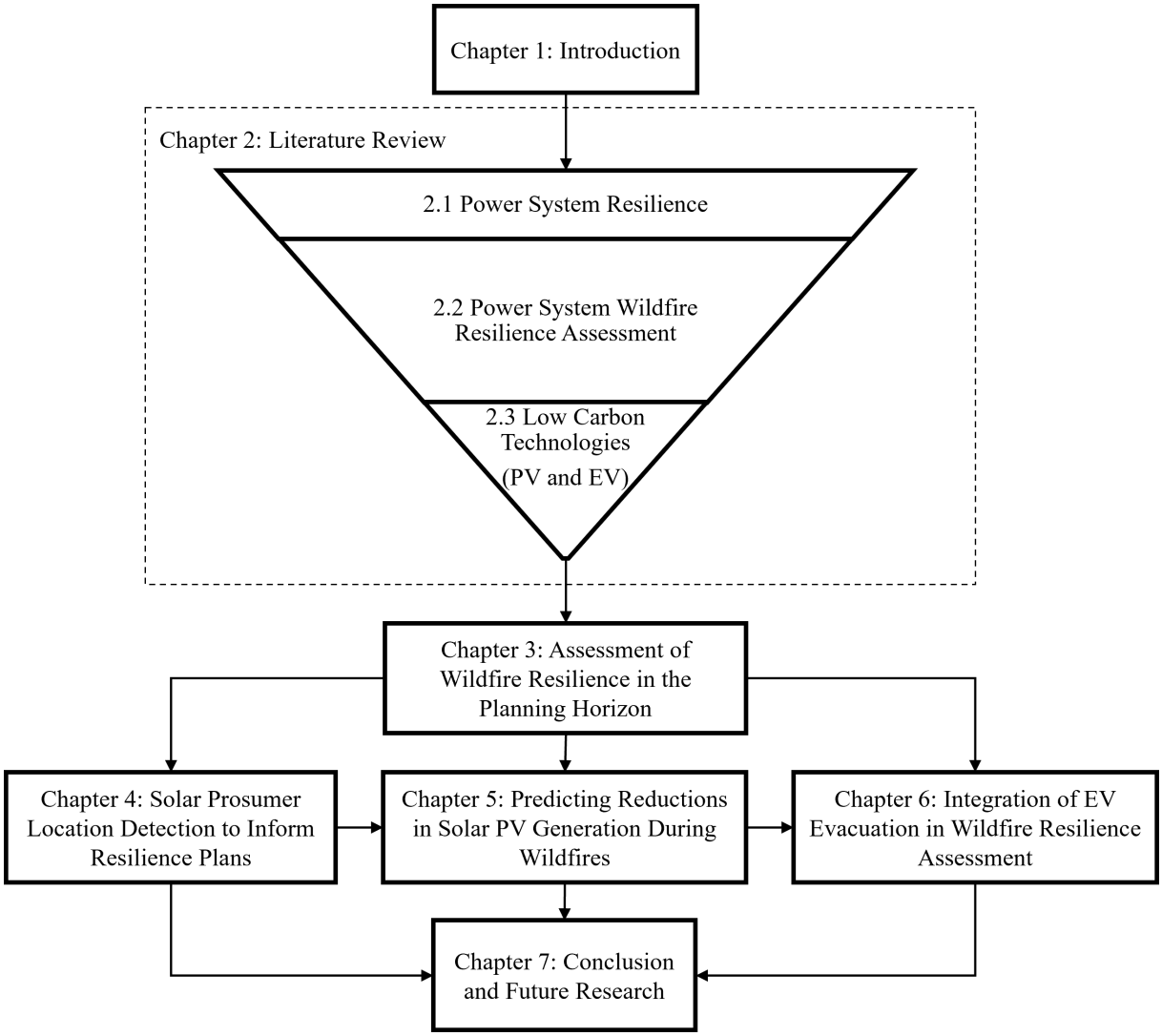


Figure 1.1: Overall Thesis Structure

solar prosumers using only net smart meter data at hourly resolution. Piecewise-aggregate approximation (PAA) is implemented to reduce the amount of data used for identification to as little as a single data point per customer to reduce the computational burden and storage requirements. A case study is used to demonstrate the accuracy of the proposed approach and improvement over existing identification approaches. The chapter closes with discussion of how the proposed identification approach can be used to inform wildfire mitigation plans.

Chapter 5 presents a quantile regression based model to predict the derate to solar PV capacity during a wildfire using measurements of aerosol optical depth (AOD). The

chapter also details how generation developers and planners can use the model to enhance the site selection process for new generation facilities. Real operating data from PV generation across four sites in California during the 2020 wildfire season are used as a case study to train the quantile regression model and validate the accuracy of the methodology.

Chapter 6 presents a new methodology to integrate electric vehicle evacuation charging demand in transmission system wildfire resilience assessments. Central to this methodology is a novel spatiotemporal EV evacuation charging demand model. The model reflects changes to EV charging demand seen by the electric grid as customers charge prior to their journey, once they arrive at their destination, as well as the charging during the remaining extent of the evacuation period. A case study informed by the wildfire risk from Chapter 3 and the potential derate to solar PV capacity from Chapter 5 is conducted to demonstrate the suitability of the proposed methodology.

Chapter 7 presents the conclusions from the research contained in Chapters 3-6. Recommendations are also provided demonstrating the practicality of the proposed methodologies for use in power system wildfire resilience assessment. Finally, several areas for future research are provided.

Chapter 2

Literature Review

This chapter critically reviews recent and relevant literature related to the topic of power system wildfire resilience with low carbon technologies (LCT). The chapter is divided into four sections: Section 2.1 examines the emergence of the concept of resilience and its application to power systems. Section 2.2 critically evaluates past and present state-of-the-art methods and techniques to assess wildfire resilience for power systems. Section 2.3 explores how LCT including solar photovoltaics (PV) and electric vehicles (EVs) impact the performance of the grid during wildfires. Finally, Section 2.4 reinforces the gaps in these areas that are the focus of this thesis.

2.1 Power System Resilience

2.1.1 Defining Resilience

Resilience can be distinguished from reliability by an expansion of the events considered, consideration of the period before and after an event in addition to the event itself, and a more holistic look at the effects across customers, emergency management personnel, and grid infrastructure [21]. Resilience and security are similar but differ in that resilience expands beyond the set of credible contingencies typically considered in a security assessment and that resilience measures are often continuous, whereas in response to a given

contingency the security is binary (secure/insecure) [22]. However, a singular definition of resilience has not been adopted. Some notable definitions include:

> "The ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions through risk management." - United Nations Office for Disaster Risk Reduction [23]

> "The ability to withstand and reduce the magnitude and/or duration of disruptive events, which includes the capability to anticipate, absorb, adapt to, and/or rapidly recover from such an event." - IEEE [21]

> "Power system resilience is the ability to limit the extent, severity, and duration of system degradation following an extreme event." - CIGRE [8]

> "Transmission system resilience is defined as the ability of the system and its components (i.e., both the equipment and human components) to minimize damage and improve recovery from non-routine disruptions, including High Impact Low Frequency (HILF) events, in a reasonable amount of time." - North American Transmission Forum [24]

Although these definitions are similar in principle, some definitions explicitly highlight HILF events [24]. While events that utilities are seeking to address may have historically been considered HILF, climate change is altering the frequency of extreme events such as drought [25] and wildfire [26]. Therefore, definitions that use more general terms such as "disruptive events" [21] or "extreme events" [8] are more robust.

2.1.2 Processes of a Resilience Assessment

Assessment of the resilience of the system to a particular hazard requires a detailed characterisation of the system processes during the timeframe leading up to the event, the

behaviour of the system during the event and the restoration process following the event [21]. Initially the processes involved were simplified to a sharp reduction in resilience due to a disruptive event, followed by a steady increase in resilience as the system is restored, termed the "resilience triangle" [27]. A refined triangle was presented in [28] to represent the response of the system resulting in a rounded point to the triangle. However, this characterisation still failed to reflect that the initial shock to the system may consist of several component failures, resulting in graduated decline rather than a single shock. To address this deficiency, the processes were expanded to form a "resilience trapezoid" consisting of 1) the disturbance process, 2) the post-disturbance degraded state, and 3) the restorative state [29]. Subsequently, the processes were further expanded by CIGRE Working Group C4.47 to seven actionable processes that reflect actions taking place before, during and after the event [8]. Authors in [30] refined the trapezoidal shape to account for these updates indicating that the preparation before an event can raise the resilience and adaptation after the event can also increase the resilience for future events. Depiction of the processes proposed by CIGRE [8] and the prior shapes (triangle [27], updated triangle [28], initial trapezoid [29], and refined trapezoid [8] are shown in Figure 2.1. Metrics to quantify the resilience are described in Section 2.1.3 with the established metrics given in [29] shown in Figure 2.1 as they are linked to the shape shown.

While considering the resilience timeline as a "trapezoid" may be useful in visualising or conceptualising resilience, it may not reflect the "shape" of actual complex disturbances of the grid. Under a real extreme event, outages and restoration can occur at different times as an event propagates across the power system, leading to irregular patterns not reflected by a trapezoid [31]. Therefore, rather than relying on a specific shape to guide assessment of the stages of resilience, it is critical to identify the underlying behaviour of the seven key processes unique to the event being considered.

Assessing the resilience of the system and determining the actions that should be taken in each process require understanding of the spatiotemporal system behaviour in response to the event. For power systems, this requires simulation of the power-flow at each time

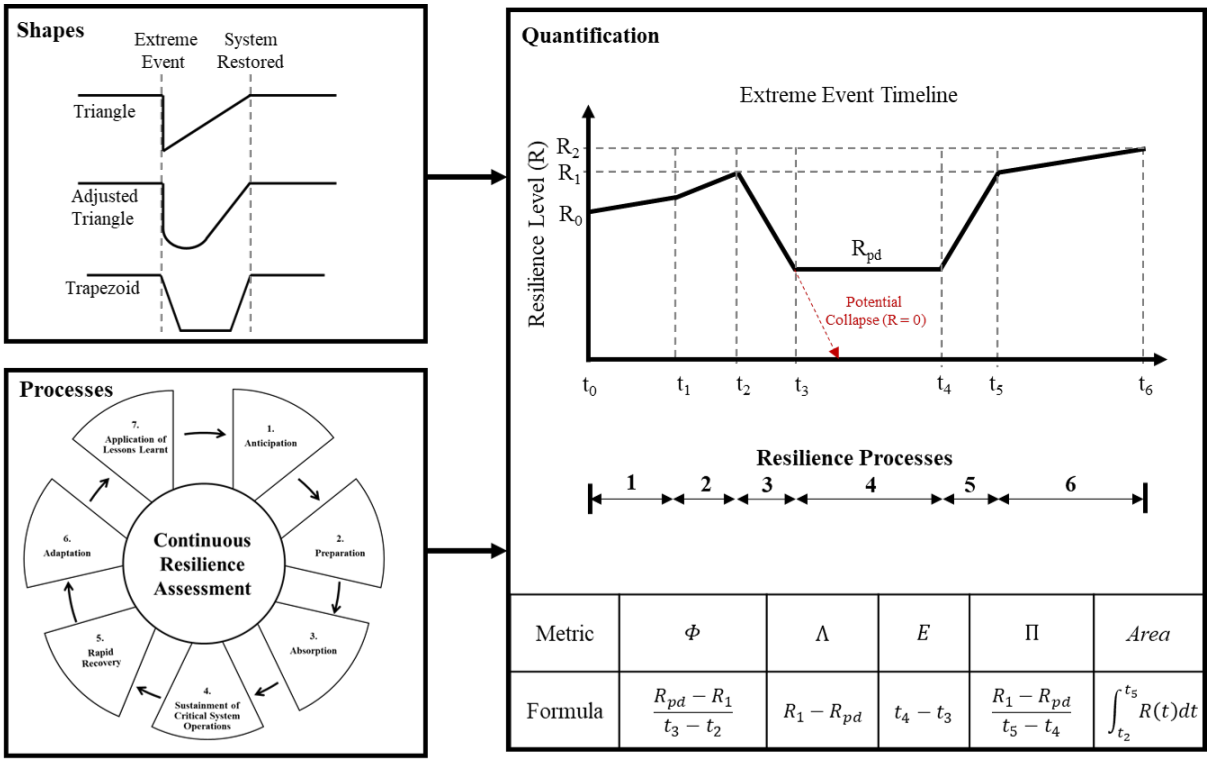


Figure 2.1: Depiction of the shapes, processes and metrics used to reflect resilience [8, 27, 28, 29]

step during the period being considered. When a hazard causes a violation of power system operating criteria, an Optimal Power Flow (OPF) model can be used to eliminate these violations via the redispatch of generation, or curtailment of load in extreme cases [32]. Two variations of OPF are traditionally simulated for resilience analyses: 1) Alternating current (AC) OPF and 2) Direct Current (DC) OPF. The difference between the two is that DC-OPF neglects the effect of reactive power, resulting in a linear approximation of the problem [33]. The lower computational burden of DC-OPF in comparison to AC-OPF can allow the simulation of a larger number of events [34]. However, DC-OPF often underestimates the impact as voltage constraints are not considered [35]. Therefore, where feasible computationally, AC-OPF should be used to identify more realistic assessments of the impact of extreme events.

For an n-bus electric power system, the AC-OPF problem can be formulated as [32]:

$$\min \sum_{i}^{NG} f_{Ci}(P_{Gi}) \tag{2.1}$$

subject to

$$P_{Gi} - P_{Di} = V_i \sum_{j=1}^{N} V_j(G_{ij}\cos(\delta_{ij}) + B_{ij}\sin(\delta_{ij})) \quad (i = 1, \dots N) \tag{2.2}$$

$$Q_{Gi} - Q_{Di} = V_i \sum_{j=1}^{N} V_j(G_{ij}\sin(\delta_{ij}) - B_{ij}\cos(\delta_{ij})) \quad (i = 1, \dots N) \tag{2.3}$$

$$P_{Gi}^{min} \leq P_{Gi} \leq P_{Gi}^{max} \quad (i = 1, \dots NG) \tag{2.4}$$

$$Q_{Gi}^{min} \leq Q_{Gi} \leq Q_{Gi}^{max} \quad (i = 1, \dots NG) \tag{2.5}$$

$$V_i^{min} \leq V_i \leq V_i^{max} \quad (i = 1, \dots N) \tag{2.6}$$

$$-S_l^{max} \leq S_l \leq S_l^{max} \quad (l = 1, \dots NB) \tag{2.7}$$

where f_C is the cost function; N, NG, NB are the number of buses, generators and branches respectively; V_i and δ_i are the magnitude and angle at bus i; δ_{ij} is the difference in angle between bus i and bus j; G_{ij} and B_{ij} are the real and imaginary parts of the corresponding element in the bus admittance matrix; P_G and Q_G are the real and reactive generation power; P_D and Q_D are the real and reactive power loads; S_l is the apparent power flow on branch l; S_l^{max} is the maximum rating of branch l; and equations (2.4)-(2.7) are the upper and lower limits of the corresponding variables. To identify the minimum amount of load shed to alleviate system constraints during an extreme event load shedding can be considered by modelling dispatchable loads as a generator with a negative output

where increasing the output has a sufficiently high cost that reflects the value of lost load (VOLL) [36]. Once the behaviour of the system in each state is computed, the resilience at each stage can be quantified using resilience metrics.

2.1.3 Metrics to Quantify Power System Resilience

Traditional indices used in power system reliability and risk assessments may be insufficient for power system resilience assessment due to their emphasis on the event probability and focus on expected or average values [37, 22]. Overall, the insufficiency of existing metrics for reliability have led to the proposal of new metrics to consider the unique characteristics of an extreme event. One such metric is '$\Phi\Lambda E\Pi$' pronounced 'FLEP', where Φ and Λ capture how fast and how low the system resilience drops, E reflects the extent of the impact in the post-event degraded state, and Π reflects the promptness of system recovery [29]. The formulation of each of these metrics alongside the resilience timeline for a theoretical extreme event can be seen in Figure 2.1. While these metrics provide a holistic view of the resilience of a system, they are focused on measuring the response to a single event rather than the resilience over a longer duration period where the system may be exposed to several extreme events.

Other studies use variations on traditional metrics such as loss of load probability (LOLP), expected demand not served (EDNS), and expected energy not served (EENS) as a means to quantify resilience [38, 39]. Authors in [37] argue against use of any metric that is based on disaster probability, under the premise that the probability of an extreme event is not able to be defined in most cases. However, the probability of a given hazard is an essential part of resilience assessments and is necessary to inform decision-making and mitigation selection. While the true event probability may be unknown, the historical frequency of occurrence of an event may be readily observable [37]. Instead of simply avoiding probabilistic metrics as suggested in [37], a better approach is to move away from simple mean based metrics which may not capture resilience [40]. One means to do this is via use of Value at Risk (VaR), and Conditional Value at Risk (CVaR) which can

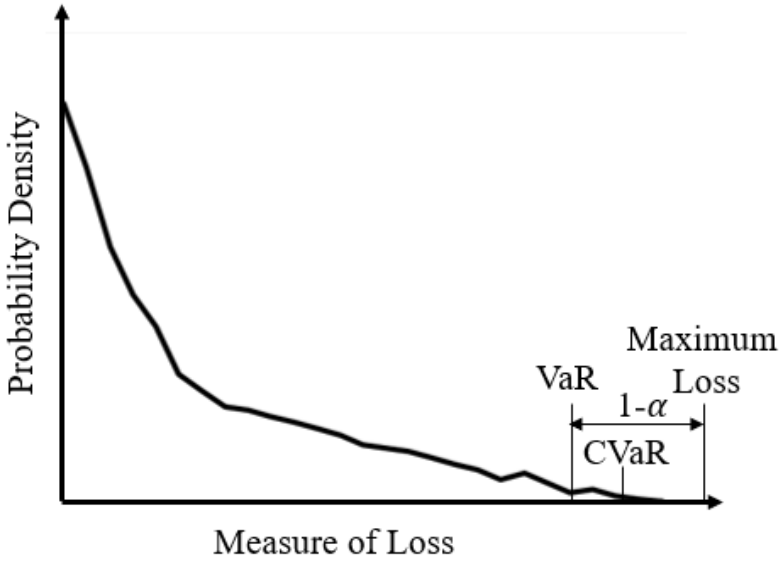


Figure 2.2: Illustrative example of VaR and CVaR for resilience analysis where HILF is defined as the top 1-α events [41].

be used to provide an indication of the expected risk under a HILF event [40, 41]. VaR_α estimates the magnitude of a loss function for a specific confidence level defined as 1-α [42]. $CVaR_\alpha(Z)$ is then a measure of expected value of the resilience indicator for the top 1-α events [41]. Both of these indicators are depicted in Figure 2.2 for a representative loss and density function [41].

$CVaR$ can be used to develop planning portfolios to constrain exposure to HILF events, enhancing resilience [43]. This is even more significant in light of climactic changes that are resulting in the increased likelihood of simultaneous natural hazards such as wildfires [18], resulting in more extreme losses under low frequency events. Therefore, use of metrics for resilience should consider both the impacts of individual events such as captured in '$\Phi\Lambda E\Pi$' while also quantifying the resilience under extreme conditions through use of $CVaR$ with multiple events occurring over the course of a given hazard season. These metrics should also be customised to reflect the specific system properties of interest for a particular type of extreme event such as a wildfire.

2.2 Power System Wildfire Resilience

Wildfires are different from other extreme events such as windstorms in that they are fuel-dependent, and increased fuel loads from wildfire suppression or other factors can result in more severe wildfires [14]. The dependence on fuel results in repeated impacts over a given period of time which can vary by area. For example, in the United States Northern Rocky Mountains, researchers found a median interval of 13 years between successive fires affecting the same area, and observed intervals as short as 0 years (indicating multiple fires in a single wildfire season) [44]. Therefore, determination of portions of the power system that may be at risk of failure due to wildfires is important to develop a resilient grid. Furthermore, modelling the resilience of a power system to wildfire related threats requires a deep understanding of the spatiotemporal extent of wildfires and the vulnerability of infrastructure. Using this understanding, power system planners, operators, and other stakeholders can identify the most suitable measures to adapt the system, or mitigate the risk posed to the system from wildfires. Each of these topics is critically reviewed below, highlighting the need for enhanced methods to assess power system wildfire resilience.

2.2.1 Simulation of Wildfire Spatiotemporal Extent

In existing literature, methods for modelling the spatiotemporal extent of wildfires for power system analysis can be grouped into two categories: 1) Thomas Formula based methods [45, 46, 47]; and 2) Cellular-Automata (C-A) based methods [48, 49]. The Thomas formula is given as:

$$R = \frac{k(1+w)}{\rho_b} \tag{2.8}$$

where R is the rate of spread (m/s); w is the wind speed; ρ_b is the bulk density of fuel; and k is a proportionality constant noted to be 0.07 kg/m^3 [50]. Authors in [45, 46, 47] all rely on the Thomas formula to assess impact of a single fire to power system infrastructure. Such studies do not incorporate elevation or varied fuel composition, which can alter the rate of spread [51]. Authors in [52] and [53] simplify this approach even further by assuming a known wildfire position relative to the line [52] rather than modelling

the spread or assume a fixed propagation speed [53]. Furthermore, across these studies [45, 46, 47, 52, 53, 54] the location of the ignition point has been fixed. The downside of approaches using the Thomas formula or fixed spread assumptions, is that the stochastic behaviour of a wildfire and the ignition location is not directly reflected.

The other approach is to use C-A models informed by the Rothermel [51] equation. In addition to the density of fuel reflected by Thomas [50], Rothermel proposed a model shown in (2.9) that incorporated the moisture content of the fuel, the temperature of ignition, and the amount of fuel:

$$R = \frac{I_R \xi (1 + \phi_w + \phi_s)}{\rho_b \varepsilon Q_{ig}} \tag{2.9}$$

where R is the rate of spread (ft./min.); I_R is the reaction intensity; ξ is the propagating flux ratio; ϕ_w and ϕ_s are factors to account for wind and slope respectively; ρ_b is the bulk density of fuel; ε is the effective heating number; and Q_{ig} is the heat of pre-ignition [51].

C-A models model the wildfire and surrounding area as a grid of cells that transition from unburned, to burning, to burned via the principles from the Rothermel equation to [55]. This can provide an accurate spatiotemporal reflection of wildfire behaviour as validated with real fires in [55, 56]. As C-A models are a stochastic process, their use enables consideration of multiple potential wildfire scenarios as the environmental parameters are varied to identify the corresponding power system impact [48, 49]. In [49], authors assigned a single propagation value to all cells, greatly simplifying the C-A model. The model employed in [48] incorporates an unvalidated "random factor" in addition to other traditional variables such as elevation, wind speed, and vegetation. While C-A models that have been validated with real fires are present in literature, the assumptions in [48] and [49] are not validated with real fires and may result in simplistic and/or inaccurate wildfire spread patterns.

Studies using the Thomas formula or un-validated C-A models may be beneficial for efficiently modelling a single wildfire for a given area but may lack the irregularities of real

fires. Furthermore, both (2.8) and (2.9) require specific meteorological and environmental variables to properly model the spread of a wildfire that may not be known in advance due to annual changes in fuel density and weather stochasticity. While such information may be available in the operating horizon, methods that can provide realistic wildfire behaviour while minimising the need for specific parameters could be beneficial for studies in the planning horizon. Therefore, given the uncertainty of these parameters in the planning horizon, this thesis models the wildfire rate of spread using historical wildfire perimeters rather than using (2.8), (2.9) or C-A models. Chapter 3 describes this process, fulfilling Objective 1 of this thesis.

2.2.2 Infrastructure Vulnerability

Wildfires can damage or disrupt the operation of electric power systems as a result of the flames, smoke, or associated wildfire mitigation activities [14]. The amount of heat produced by a wildfire varies based on fuel and other environmental properties. Surface fire temperatures are often less that 400°C; however chaparral (a type of fuel found in places such as California) can result in surface fire temperatures reaching 900°C [57]. When fires spread from the surface to the forest crown, temperatures can range from 800 to 1,200°C [57]. For overhead conductors, this heat can cause annealing, resulting in an irreversible reduction in tensile strength [58]. As a result, lines may sag below their safety clearance resulting in increased likelihood of flash-over or failure [58]. For structures, the high temperature can cause wooden poles to ignite, and in extreme cases, compromise the structural integrity of steel poles [59]. Though the specific temperature limits differ based on the exact material properties and the presence of additional coatings, authors in [59] indicate that damage can result at temperatures of 538 °C for steel, 162 °C for aluminium, 300 °C for wood and 350 °C for fiberglass. While steel poles are the most resistant to heat related failure, when failure does occur it is often catastrophic as opposed to the slower relative failure of wooden structures [59]. The smoke from wildfires can also contaminate insulators, leading to flashover and line outages [60]. The dumping

of flame retardant from aircraft as part of firefighting efforts can also result in flashover for insulators operating in the range of 115-500 kV due to the increased conductivity of the retardant [61]. As a result, lines may remain out of service until utilities are able to clean the insulators. Finally, the smoke and heat from can result in increased conductivity of the air and surrounding gases leading to line to line and line to ground faults [62].

To reflect the impact of heat on power system operations, researchers have used a solid flame model incorporating the radiative and convective heat from a progressing wildfire to identify the corresponding impact to line ratings [45, 46]. Authors in [45] focus on identifying dynamic thermal limits for distribution and transmission lines in order to aid operators in avoiding irreversible conductor damage. While this approach is beneficial for active operational damage mitigation during a wildfire, it does not predict the likelihood of a line failing if the stated limits are not observed. Authors in [46] enhance the model presented in [45] by reflecting the non-steady state heat balance equations and use stochastic programming to optimize the operation of the grid to minimise load shedding. However, the authors assume that as soon as a wildfire reaches the line failure always occurs [46]. While the focus of the work in [46] was at the distribution level, this assumption may not be valid at the transmission level depending on the line clearance, material properties, tower characteristics, and nature of the fire. While beneficial operationally, models in [45, 46] are reliant on characteristics of the wildfire flame (flame length, angle, tilt etc.) to develop an accurate measure of the corresponding radiative heat which may not be available to power system planners and operators.

The second method seeks to identify the failure of assets using a fragility curve. A fragility curve reflects the failure probability of an asset in relationship to a natural hazard and can be developed via a variety of methods including historical failure data, analytical simulation, or expert judgement [38]. While much work has been done to develop fragility curves for hazards such as hurricanes, floods, earthquakes and windstorms [38, 63], less attention has been given to wildfires. At the transmission level, initial fragility functions that reflect the risk of transmission line failure due to wildfire smoke have been proposed

recently by authors in [48] and [64]. These works model the wildfire fragility as a function of environmental variables such as temperature, humidity, smoke, and windspeed as well as the physical parameters of a given line (tower height, line spacing, and insulator shape). In application of these works to corresponding case studies, authors in [48] and [64] make multiple simplifying assumptions regarding the specific infrastructure parameters along with the exact weather conditions during the event to enable calculation of impact. Existing approaches using flame models and fragility curves are focused on the operating horizon [45, 46, 48, 64] as the detailed assumptions are unlikely to be available to power system planners multiple years in advance. Given the uncertainty regarding future hourly weather data and given that this thesis is largely focused on the planning horizon, wildfire fragility curves are not used. Instead, methods that can be applied by power system planners to simulate a range of outage behaviours in the planning horizon in absence of such high resolution data would be beneficial to inform wildfire resilience assessments. Such a method is proposed in Chapter 3, addressing this gap and fulfilling Objective 1 of this thesis.

2.2.3 Mitigation and Adaptation Measures

Both mitigation and adaptation play a role in enhancing power system resilience. Mitigation of disaster risk is defined as "The lessening of the potential adverse impacts of physical hazards (including those that are human-induced) through actions that reduce hazard, exposure, and vulnerability" and adaptation (in human systems) as "the process of adjustment to actual or expected climate and its effects, in order to moderate harm or exploit beneficial opportunities" [65]. In the context of this thesis, mitigation of the impact of wildfires refers to operational action taken to limit the extent of the damage caused by a wildfire. This can be achieved by providing operators with advanced warning of an event or the real-time risk of a fire under current weather conditions [48, 64] or by implementing methods to reconfigure the system to limit the impact [46]. Authors in [53] demonstrate how local energy resources can be used to further supply power during the

wildfire mitigating load shed requirements. Finally methods to speed up the restoration such as enhanced emergency planning, and the use of temporary microgrids can also limit the duration and extent of the damage caused to the system [66, 67]. Practical methods to aid power system planners and operators with the better understanding of the performance of the generation facilities involved in such mitigation measures would be beneficial to determine their suitability for a particular location.

The simulation done in existing wildfire resilience assessments often focuses on the effect of a single fire [46, 49] or a single day with multiple fires [48]. Single event modelling can be useful to present a new methodology, or demonstrate how new types of variables can be integrated into the event assessment. However, with growing prevalence of overlapping wildfires [18], work to assess wildfire resilience over the course of an entire wildfire season could better inform power system planners where to focus attention on resilience enhancements. This gap is addressed in Chapter 3, fulfilling Objective 5 of this thesis.

Once the resilience of the system has been quantified, suitable means of adaptation to enhance the resilience can be identified. These enhancements consist of methods of wildfire prevention, impact mitigation, and recovery [66, 67]. Approaches to wildfire prevention include vegetation management, element undergrounding, structural hardening and alteration of protection settings such as line reclosing [66, 67, 68]. Vegetation management requires regular ongoing action to ensure sufficient clearances are maintained to mitigate the risk of heat related failures [59]. Other adaptation measures such as undergrounding may provide a more permanent reduction in risk, but are much more expensive, and take multiple years to complete [67]. Therefore, methods that can identify needs for adaptation in the planning horizon would be beneficial to enable time for proactive implementation of wildfire resilience plans. This gap is addressed in Chapter 3, also fulfilling Objective 1 of this thesis.

2.3 Modelling Low Carbon Technologies During Wildfire Events

2.3.1 Detecting the Location of Rooftop PV Installations

Wildfire resilience enhancements and mitigation plans often rely on the presence of distributed generation to provide power supply to an isolated area [53, 66]. Customers around the world are rapidly installing rooftop solar PV systems to offset their energy consumption with 240 million installations anticipated globally by 2050 under the IEA net zero energy scenario [10]. Customers with solar PV installed are often referred to as 'prosumers' due to their ability to produce and consume electricity [69]. However, utilities may not be aware of the locations of each installation as many PV installations are connected via Net Energy Metering (NEM) schemes where a single meter measures the balance between import and exports [70]. Therefore, it can be difficult for operators to track the location and performance of rooftop PV installations. Furthermore, as reported in 2020, "in some countries, including the UK, small systems may legally be grid connected with no registration, meaning there may be no official documentation of their existence" [71]. In the UK at present (2022), installers of small PV facilities (a capacity of 16A per phase or less) have up to 28 days to notify Distribution Network Operators (DNOs) following installation [72]. Location uncertainty is further compounded due to unauthorised or unreported PV installations, multiple systems at one residence, and incorrect data entry [73]. Therefore, methods to detect the location of solar prosumers are essential to fully leverage existing solar PV installations as a component of resilience enhancement strategies.

There are several existing approaches to identify solar PV installations. The primary difference between approaches is the amount of input data (satellite, weather, grid measurements) and requirement of labelled training data. Classification approaches can be evaluated based on the number of True positives (TP), True negatives (TN), False positives (FP), and False negatives (FN) [74]. These indicators are combined to provide

performance measures including Precision, Recall, Accuracy and F1 Score as given in (2.10)-(2.13) [74].

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{2.10}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{2.11}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{2.12}$$

$$F1\ Score = \frac{TP}{TP + \frac{1}{2}(FN + FP)} \times 100\% \tag{2.13}$$

The first set of methods are those reliant on satellite data. An early method in this area used high resolution (0.3 meter per pixel) colour satellite orthoimagery to train a random forest (RF) classifier for identification, and was reliant on 2,700 manually labelled installations for training and testing [75]. The approach first divides the imagery into an m-dimensional feature space capturing information such as nearby colours and textures to use for identification, and then uses those features in a random forest classifier to identify whether a specific pixel is a PV array [75]. The random forest classifier works by fitting an ensemble of decision trees which each assign a probability of a pixel belonging to a specific class and then taking the average probability across all of the trees [75]. The efficacy of the approach varied with the value of Jaccard index (J) selected by the authors, only achieving an object based performance at J=0.1 of 70% Precision and 60% Recall (and lower performance at other values of J) [75].

Convolutional Neural Networks (CNNs) can provide improvements in accuracy over RF, with performance improving to 95% Precision and 80% Recall [76]. CNNs are "neural networks that use convolution in place of general matrix multiplication in at least one of their layers" [77] enabling sophisticated features to be identified [78]. A CNN based

approach, DeepSolar, was developed in 2018 [79] to detect panels for the entire contiguous United States from Google Static Maps and reported to have "a precision of 93.1% with a recall of 88.5% in residential areas and a precision of 93.7% with a recall of 90.5% in non-residential areas" [79]. Deep Solar was further improved and extended to Germany in 2020 achieving a precision of 92% and a recall of 98% [80]. While accuracy and coverage is improving, these methods remain dependent on the availability and latency of high resolution satellite imagery along with labelled training data (with [79] requiring labelling of 100,000+ images). In addition, these methods require training at a representative number of locations as CNN trained in one city may not generalise to another [81]. Finally, the end result of these methods is the identification of pixels representing a geographic location containing solar PV installations [79]. Due to the output of these models, further inaccuracies may be introduced when mapping the pixels to the exact point of the grid where the panel is connected.

Power system data (along with weather information in some cases) can also be used to identify solar prosumers. A changepoint detection approach is presented in [73], but requires information from before and after the PV installation is put into service. A support-vector classifier is proposed in [82] that uses load and meteorological data to create features including the ratio of total electricity consumption across different weather classes, a concave shape index, a concavity degree, and load ramping rate for use in solar prosumer identification. The approach is tested using 4-fold cross validation (requiring labelled training data), providing an average accuracy of 95.51% [82]. One unsupervised method to identify solar prosumers reliant solely on smart meter data utilised k-means clustering [83]. K-means is a partitioning based cluster algorithm which organises a dataset X containing n objects into k clusters $C_1, \ldots C_k$, where $C_i \subset X$ and $C_i \cap C_j = \emptyset$ for $(1 \leq i, j \leq k)$ [84]. K-means then attempts to minimize the within cluster variation given as:

$$E = \sum_{i=1}^{k} \sum_{\boldsymbol{x} \in \boldsymbol{C_i}} dist(\boldsymbol{x}, \boldsymbol{c_i})^2 \tag{2.14}$$

where E is the sum of squared error across X; $\boldsymbol{x}$ is the point in space that represents a given

object; $\boldsymbol{c_i}$ is the cluster centroid for cluster C_i; and *dist* reflects the measure of distance between $\boldsymbol{x}$ and $\boldsymbol{c_i}$ (note both $\boldsymbol{x}$ and $\boldsymbol{c_i}$ are multi-dimensional) [84]. The k-means based approach performed well with authors reporting 95% Precision, 94% recall, and 94% F1 Score [83]. However, the approach was reliant on 15-min smart meter measurement data, and required authors to manually label the resulting clusters through "visual inspection" [83]. The performance of the k-means algorithm can also be improved through careful seeding (referred to as k-means++) [85]. Alternative methods of clustering could also provide potential benefits to the accuracy. One such method, agglomerative clustering is a hierarchical method, performed in a bottoms up fashion, whereby each customer starts out as its own cluster and clusters are iteratively merged according to a specified merging criteria until either a single cluster is reached or the process is stopped by a termination criteria [84]. Criteria for merging include 1) Single-linkage: similarity is measured by the closest pair of elements; 2) Complete-linkage: similarity is determined by the maximum distance between elements; [84] and 3) Ward-linkage where the similarity is determined by the sum of squares of the clusters formed from the current clusters [86]. Artificial neural networks provide another means of clustering referred to as Self-Organising Maps (SOM) [87]. The user pre-defines the dimensions N_1 and N_2 and then a SOM projects a set of H-dimensional data into a reduced dimension space containing $N_1 \times N_2$ H-dimensional competitive units forming a competitive layer [88]. The SOM takes each of the points in the H dimensional space and identifies the best matching competitive unit in the reduced feature space based on the distance [88]. Each subsequent iteration of the SOM updates the location of the competitive units to best match the input data [88].

Furthermore, as the resolution of the data and the amount of exogenous data such as weather increases, so does the computational and storage requirements. Dimensionality reduction techniques can alleviate these concerns, reducing computational complexity and storage requirements. Dimensionality reduction techniques come with trade-offs and the best methods are able to reduce the data requirement while retaining critical data on the original series. One approach, Piecewise-Aggregate Approximation (PAA), is superior to

other established methods such as Singular Value Decomposition (SVD), Discrete Fourier transforms (DFT), and Discrete Wavelet Transforms (DWT) as PAA is much faster to compute, is not limited to arbitrary lengths of time series, and is flexible to various distance measures, while being simple to understand and implement [89]. PAA divides a times series (such as hourly smart meter data) into a series of 'frames' and calculates the mean of data within each frame with the resulting vector of means becoming the reduced time series [89]. This enables translation of a time series X of length n into a w-dimensional vector $\bar{X} = \bar{x}_1, \ldots, \bar{x}_w$ with the dimensionality reduced by a factor of $\frac{n}{w}$ where $\bar{x}_i$ is determined as [89]:

$$\bar{x}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} x_j \tag{2.15}$$

Alternative methods of aggregation within each frame such as the minimum or maximum value can be implemented as in:

$$\bar{x}_i = min\{x_j, \ldots, x_k\} \tag{2.16}$$

$$\bar{x}_i = max\{x_j, \ldots, x_k\} \tag{2.17}$$

where $j = \frac{n}{w}(i-1)+1$ and $k = \frac{n}{w}i$.

A related method is to aggregate values based on the hour of the day rather than over a frame of sequential values. This statistical approach to aggregation results in the representation of a time series as a typical daily profile. Similar to the approach in (2.15)-(2.17) the values can be aggregated based on the mean, minimum or maximum hourly values as follow:

$$EMean_h = \frac{1}{p} \sum_{d=1}^{p} x_{d_h} \tag{2.18}$$

where $h = 1, 2..., 24$ represents the hour; d represents a specific day; p represents the total number of days in the period; and $EMean_h$ is the mean electricity demand for a given

hour over the period.

$$EMax_h = max\{x_{d_h}, 1 \leq d \leq p\} \tag{2.19}$$

where $EMax_h$ is the maximum electricity demand for a given hour over the period.

$$EMin_h = min\{x_{d_h}, 1 \leq d \leq p\} \tag{2.20}$$

where $EMin_h$ is the minimum electricity demand for a given hour over the period. Application of (2.18)-(2.20) result in typical 24-hour load profiles for each customer.

Overall, when evaluating the suitability of any of these methods, the amount, sample rate and duration of data necessary, the accessibility of that data, and the frequency at which the data is updated must be considered to enable a robust means to identify candidate resources for resilience enhancement strategies. Knowledge of the presence of PV generation is also critical when performing automated switching reconfiguration or restoration as PV can mask the true load on a distribution feeder [90]. For example, when restoring a distribution feeder following an outage, for feeders with solar prosumers, the amount of load requiring restoration can be higher than pre-outage levels and can result in overloads, undervoltage, and/or tripping of protective devices [90]. Awareness of PV installations is also essential to set and maintain appropriate protection settings [90, 91, 92] as well as regulation of the voltage on distribution feeders [90, 91]. Ultimately, methods to accurately identify solar prosumers can enhance the ability of power system planners and operators to maintain a safe and resilient system, and develop robust wildfire mitigation plans. This gap is addressed in Chapter 4, fulfilling Objectives 2 and 5 of this thesis.

2.3.2 Solar PV Generation Performance

The amount of Solar PV generation present in the grid continues to increase due to favourable policies, lowering cost, and a desire to reduce greenhouse gas emissions [93]. Some regions are seeing especially high penetration such as the California Independent System Operator (CAISO), where Solar PV generation represented 14% of total supply in

2020, with an even higher share during the middle of the day [94]. As penetration grows, the main challenge for grid operators is to predict the power output due to the influence of clouds, aerosols, and other meteorological factors [93]. To account for uncertainty in the forecast output, operators integrate this uncertainty into the determination of effective load carrying capacity or operating reserves [95, 94]. While models to predict solar PV generation are growing in accuracy, difficulty remains in forecasting the output during abnormal phenomena such as fog, snow and dust [93] and wildfires.

Researchers have quantified the magnitude of impacts to solar PV production for some phenomena, leading to more accurate generation forecasts. For instance, the build-up of snow on solar PV facilities can adversely impact PV production by inhibiting solar irradiance. The magnitude of the derate to production is primarily dependent upon the climate (amount of snowfall), tilt angle, and the obstruction of snow sliding [96, 97]. A study of PV configuration in the northern United States indicated that annual snow induced losses reduced from 34% to 5% for unobstructed installations as the tilt angle increased from 0° to 45°, whereas tilt angle did not have a substantive impact on obstructed installations with losses from 29-34% [97]. As a result, models that account for the reduction in output due to snow have been integrated into industry standard forecast models to enhance accuracy [98].

Dust and pollution have also been investigated due to their impact to solar PV production. Dust from large deserts such as the Sahara can inhibit solar PV production, with events lasting several days occurring 5 to 15 times per year in Germany [99]. The magnitude of disruption can include mean daily reductions during a dust outbreak ranging from 20 to 40% and an extreme daily reduction of 79% [100]. The April 2014 dust storm in Germany highlights the significance of these reductions as failure to consider dust contributed to German TSOs over-forecasting PV output by up to 5.3 GW [99]. While the impacts from dust can be substantial, most effects are limited to areas in the proximity of large deserts such as those found in northern Africa, along with middle and eastern Asia [101]. Aerosols from air pollution have also been shown to reduce annual

PV generation in Eastern China by more than 20% [102]. The impact of pollution can be even more significant (29-34%) for PV installation with tracking capabilities as aerosols result in more attenuation of the direct irradiance that tracking seeks to maximize [102]. Authors in [103] quantify the impact of haze on solar insolation using extensive empirical data from Delhi and Singapore. A fitted exponential decay is used to reflect the correlation between particulate matter less than 2.5 μm in diameter (PM2.5) and normalised insolation, and identified insolation reductions due to haze of 2-9% for 16 cities around the world [103]. While dust and air pollution can both result in reductions to solar PV performance, the impact can be worsened by wildfire smoke.

In comparison to other phenomena such as snow, dust, and air pollution, the impact from wildfire smoke remains largely unexplored. Smoke from wildfires can span hundreds of miles affecting the performance of solar production for an entire region. In August and September 2020, California experienced five fires exceeding 300,000 acres each, with the largest (August Complex) spanning over 1 Million acres (4000 km^2) [104]. The smoke from these fires can be seen in Figure 2.3 covering large parts of the Western United States. Evidence of the widespread reduction in solar PV production was reported by the US Energy Information Administration [19]. During the first two weeks of September 2020, a nearly 30% reduction in solar PV generation was observed compared to July 2020 averages [19]. However, a simple comparison between September and July generation levels is insufficient to identify the magnitude of the derate as performance could be influenced by other factors in addition to wildfire smoke such as curtailment, scheduled outages, and seasonality. Hence further work is necessary to determine the empirical derate to solar PV generation resulting from wildfire smoke.

In initial efforts to identify the impact, a controlled burn in Australia was reported to have reduced the power output from a 1.56 kW installation by up to 27%, with an average reduction of 7% over the 140 minute study period [105]. Although analysis was only conducted for a single, relatively small (40 hectares) wildfire, it provides an initial indication of the potential for wildfire smoke to adversely affect performance. Wildfire

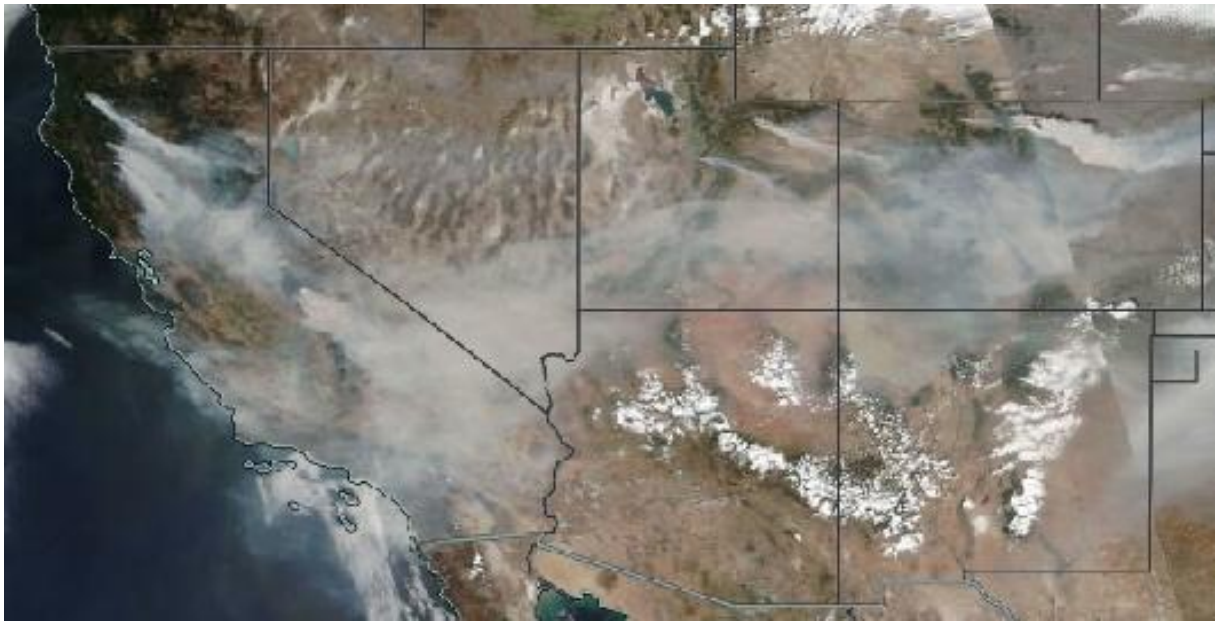


Figure 2.3: Smoke from California Wildfires on September 7th, 2020. Satellite Imagery used with permission from the NASA Worldview application. Borders from © OpenStreetMap contributors. Data available under the Open Database License.

smoke contains airborne PM2.5 and evaluation of the increase in PM2.5 emissions can provide one means to quantify the corresponding impact of wildfire smoke on PV generation [19, 106]. Regression models provide an effective and transparent means to establish the relationship between a target variable y (i.e. load or generation) and a set of predictor variable(s) $x_1, x_2, \ldots x_p$ (i.e. meteorological, natural hazard data) [107]. Once developed, these models can be used to predict the future value of a target variable given a forecast of the predictors. The most fundamental form of regression relates a target value y and a single predictor variable x_p as shown below:

$$y = \beta_0 + \beta_1 x_p + \epsilon \tag{2.21}$$

where β_0 and β_1 are unknown constants and ϵ is the error. Researchers in [106] used linear regression to identify the relationship between PM2.5, other meteorological variables, and solar PV generation and found the relationship to be statistically significant. However, the PM2.5 measurements were taken from ground based monitoring stations, limiting model coverage as plants may be located far from the nearest monitor. To extend coverage be-

yond ground based monitoring stations, an alternative means of quantifying the amount of smoke present is through aerosol optical depth (AOD), defined as "a measure of the extinction of the solar beam by dust and haze. In other words, particles in the atmosphere (dust, smoke, pollution) can block sunlight by absorbing or by scattering light. AOD tells us how much direct sunlight is prevented from reaching the ground by these aerosol particles" [108]. Average US AOD is 0.1-0.15 with 0.4 reflecting a hazy day [108] on a scale from -1 to 5 [109]. Satellites such as NOAA's geostationary Earth-observing systems (GOES-17) satellite can provide a reliable source of AOD measurements [110]. GOES-17 provides global coverage at a 10 minute refresh rate and coverage of the continental United States (CONUS) with a faster 5 minute refresh rate [109]. The development of an empirical relationship between AOD during wildfires and solar PV generation would benefit operational forecasts of solar PV production leading to better situational preparedness during an emergency. This gap is addressed in Chapter 5, fulfilling Objectives 3 and 5 of this thesis.

2.3.3 Electric Vehicle Charging Demand During Wildfire Evacuations

The loading of the grid can also be dynamically affected by the charging of EV during a wildfire as penetration grows. Adoption of EV grew by 43% in 2020 with over 10 million electric cars on the road worldwide [9]. As adoption continues to grow, power system operators face the task of ensuring the electric grid is able to safely accommodate the increased penetration. One of the most significant challenges, is to preemptively anticipate the location and magnitude of EV charging to make necessary grid upgrades and site charging infrastructure [111]. Travel patterns from GPS data and national surveys provide a key source of information for estimating the location where vehicles will charge [112]. The magnitude of charging demand for a single vehicle can vary based on the charging infrastructure, from 1.4 kW on Level 1 charger to 150 kW or more on a DC fast charger [112]. However, the average demand per vehicle is typically much less due to

diversity of charging time. For example with 100 EVs charging at a rate of 3 kW each, the average demand per vehicle seen by the grid might only be 1.2 kW [113]. The effect of diversity is further validated by field trials such as My Electric Nation in the United Kingdom which found the maximum concurrent vehicles charging on a weekday was 25% [114].

Natural disaster evacuations can alter typical charging locations and eliminate this diversity as residents concurrently seek to refuel prior to an evacuation [115]. This can cause charging infrastructure to be insufficient to meet the heightened demand, adversely affecting the ability to safely evacuate [116, 117]. Wildfires can impact residential communities, forcing evacuation of large groups of people. Between 2017 and 2019, approximately 1.1 million people were forced to evacuate as a result of 11 major wildfires in California [20]. Rapid growth in the wildland-urban interface is also placing more homeowners at risk of evacuation in the future [118].

The use of EVs for evacuation adds complexity to wildfire emergency planning, as plans must ensure the sufficiency of charging stations and ability of the electric grid to supply and transmit the necessary power. While authors in [116] and [117] explore the number of charging stations to facilitate evacuation, notably absent is analysis of the grid impacts of synchronised charging of significant numbers of electric vehicles. Unanticipated wide-scale EV charging during a natural disaster could lead to power shortages and cascading blackouts affecting regional power networks [119]. In addition to pre-departure charging, as residents evacuate, the associated EV charging demand is dispersed around the network. Therefore, further methods to model the location and magnitude of electric vehicle charging demand during a wildfire are needed. This gap is addressed in Chapter 6, fulfilling Objectives 4 and 5 of this thesis.

2.4 Summary

This chapter critically reviews the most recent and relevant frameworks, methods, models and metrics for assessing power system wildfire resilience along with the impact of growing penetration of LCT. Existing methods to identify the risk to power system infrastructure from wildfires can be improved by developing a new methodology to anticipate wildfire impacts in the planning horizon (Objective 1). This can greatly enhance flexibility and efficiency of mitigation selection as some mitigation options can take multiple years to implement. Furthermore, in the operating horizon, the growing penetration of LCT alters the behaviour of the grid during a wildfire, and the added complexity from LCT is often excluded from traditional wildfire resilience assessments. New methodologies to detect the location of solar PV installations for use in resilience and restoration planning (Objective 2), quantify the effect of wildfire smoke on PV performance (Objective 3), and predict the additional demand due to evacuating EV (Objective 4) can greatly enhance awareness and scheduling decisions in the operating horizon. Furthermore, spatiotemporal visualisation techniques are important to aid power system operators and planners in the application of these methodologies (Objective 5). These areas will be addressed in detail in the subsequent chapters of this thesis.

Chapter 3

Assessment of Wildfire Resilience in the Planning Horizon

3.1 Introduction

Climate change has resulted in increased frequency and intensity of wildfires in many parts of the world over the past several decades [26, 120]. The geographic synchronicity of wildfire danger is also increasing in areas such as the Western United States, where the number of days with synchronous fire danger is projected to double by the middle of the 21st century [18]. Therefore, it is necessary to assess the risks posed to the transmission system by multiple wildfires over an entire wildfire season. As detailed in Chapter 2.2, existing approaches to wildfire resilience typically focus on actions taking place during the operating horizon over the course of a single event period [46, 48, 49, 64]. While operational resilience measures are likely possible during the period of warning leading up to an imminent event and during the event itself, measures to enhance power system infrastructure resilience can take much longer and thus should be considered well in advance. In North America, the North American Electric Reliability Corporation (NERC) requires Transmission Planners and Planning Coordinators to perform annual planning assessments, which analyse the performance of the system under extreme events including

wildfires [121]. However, selection of the specific outages to assess can be difficult given the vast number of combinations for a large system.

The chapter presents a probabilistic methodology to assess power transmission system wildfire resilience in the planning horizon. The methodology incorporates globally available satellite information to generate many synthetic fire seasons and identify the risk to resilient operation of the transmission network. The significance of the methodology is that it delivers quantifiable measures of risk over an entire wildfire season, demonstrated using a realistic case study, that enable effective risk assessment and support mitigation plan development. This chapter also presents a selection methodology to identify N-k contingencies from wildfires which should be included into planning assessments. This facilitates more robust transmission planning assessments that underpin a utility's long-term investment strategy.

3.2 Probabilistic Methodology for the Planning Horizon

A probabilistic methodology to assess wildfire resilience in the planning horizon is proposed in this chapter with the primary aim of providing a quantifiable measure of the system resilience and identifying the transmission lines most at risk. The methodology consists of four main stages: 1) Hazard generation; 2) Network vulnerability; 3) Restoration; and 4) Resilience quantification. Each of these stages is applied for each season within a set of synthetic wildfire seasons $\mathcal{S}$, where $|\mathcal{S}|$ is calculated based on the convergence of the resilience metrics. An overview of the methodology is shown in Figure 3.1 and details for each stage are provided in the following subsections.

3.2.1 Wildfire Hazard Generation

The model leverages historical, satellite derived wildfire information to generate synthetic wildfire seasons. The spatiotemporal characteristics of wildfires can vary across fire size

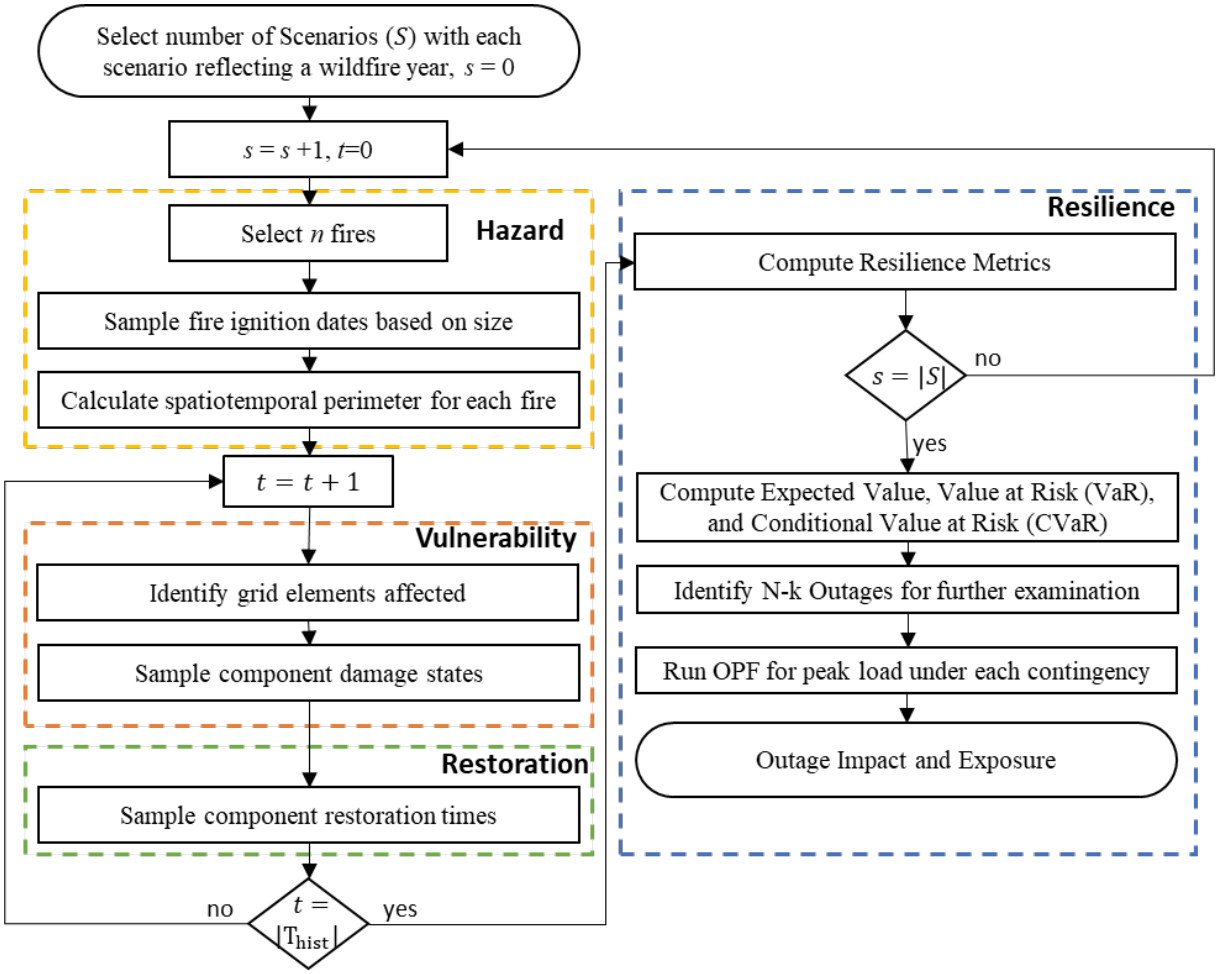


Figure 3.1: Probabilistic Wildfire Resilience Methodology

[122]. To account for this variation, the first step is to split the set of historical fires into two levels (as a two level split was used in prior literature [122]) ($\mathcal{F}$) such that:

$$\mathcal{F}_{Large} = \{f_{size} \mid f \in \mathcal{F},\ f_{size} \geq \gamma\} \tag{3.1}$$

$$\mathcal{F}_{Small} = \{f_{size} \mid f \in \mathcal{F},\ f_{size} < \gamma\} \tag{3.2}$$

where f_{size} is the size of fire f; γ is the fire size threshold; F_{Large} are large fires; and $\mathcal{F}_{Small}$ are small fires. Following separation of the historical data by size, the steps to generate each season $s \in \mathcal{S}$ consist of generating: 1) the overall number of wildfires; 2) the time at which each wildfire ignition occurs; and 3) the spatiotemporal extent of each fire.

The typical number of annual wildfires in a single wildfire season depends on two key parameters: T_{hist} and κ. T_{hist} reflects the number of historical years used to calculate the typical number of fires in a given season. This choice of length is important as the number of fires for some regions of the world has seen substantial increases in recent history [122] and therefore a longer horizon may dampen these effects. Furthermore, depending on the planning horizon selected, the projected number of fires in the future may increase. Therefore a scaling factor κ is included in (3.3) and (3.4) to enable consideration of the anticipated growth in number of fires.

$$n_{Large} = \kappa \frac{|\mathcal{F}_{Large}|}{T_{hist}} \tag{3.3}$$

$$n_{Small} = \kappa \frac{|\mathcal{F}_{Small}|}{T_{hist}} \tag{3.4}$$

$$n = n_{Large} + n_{Small} \tag{3.5}$$

where n is the total number of fires anticipated in each synthetic year; and n_{Large} and n_{Small} are the number of large and small fires anticipated in each synthetic year respectively.

The seasonal variation in wildfire ignition dates can be reflected by a truncated normal probability distribution. The truncated normal distribution is necessary to reflect the limitation that the start and end dates of wildfires bounded by the Julian calendar. The resulting truncated normal probability distribution function (PDF) is shown in (3.6) as follows:

$$\psi(\bar{\mu}, \bar{\sigma}, a, b; x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{\phi(\bar{\mu}, \bar{\sigma}^2; x)}{\Phi(\bar{\mu}, \bar{\sigma}^2; b) - \Phi(\bar{\mu}, \bar{\sigma}^2; a)} & \text{if } a < x < b \\ 0 & \text{if } b \leq x \end{cases} \tag{3.6}$$

where ϕ is the normal probability density function (PDF); Φ is the normal cumulative

density function (CDF); $\bar{\mu}$ is the mean of the parent general normal PDF; $\bar{\sigma}^2$ is the variance of the parent general normal PDF; and a and b reflect the upper and lower bounds of the truncation interval, in this case the Julian Calendar [123].

Once the distribution is calculated, the Ignition dates ($\mathcal{I}$) are determined via sampling where $\mathcal{I} \sim f(x)$. The result of the first two steps is a matrix containing the number of fires anticipated for each day of the year. For each fire ignition, the origin and spatiotemporal extent are obtained by uniformly sampling from the historical empirical distribution. This approximates the underlying size distribution of the fires while also reducing the possibility that the same fire is selected twice. Sampling is performed with replacement due to the possibility that re-burn could occur twice within the same fire season as indicated in [44] and discussed in Section 2.2. The potential for re-burn could also be excluded by sampling without replacement. Finally, the fires, ignition dates, and extent are combined for $\mathcal{F}_{Large}$ and $\mathcal{F}_{mall}$ to form a single wildfire season. This process is then repeated to form the set of synthetic fire seasons $\mathcal{S}$.

3.2.2 Network Vulnerability

The primary vulnerability considered in this work is that of transmission lines. The heat and smoke from a wildfire can result in a reduction in the breakdown voltage resulting in equipment failure or in severe cases melt the aluminium used in the lines themselves [14, 64]. The key parameter to determine the probability of breakdown of a specific line is the breakdown voltage [64]. Environmental factors such as temperature, humidity, air pressure and smoke can all effect the breakdown voltage. To correct for the impact of weather on the breakdown voltage, the following correction coefficients can be used [64]:

$$\begin{cases} \delta = \frac{(273+t_0)\rho}{(273+t)\rho_0} \\ k_\delta = \delta^m \\ k_\alpha = \frac{1}{R\alpha+1} \end{cases} \tag{3.7}$$

where δ is the relative density of air; t_0 and t are the temperature in $^{\circ}C$ under standard conditions and wildfire conditions respectively; ρ_0 and ρ are the pressure in kPa under standard conditions and wildfire conditions respectively; m is a value ranging from 0.4 to 1.0 reflecting properties of the electrodes; R is the coefficient of smoke reduction; and α is the amount of smoke present ranging from 0-1.

The adjusted breakdown voltage is then given as:

$$\hat{V}_{50} = V_{50} k_{\delta} k_{\alpha} \tag{3.8}$$

The electrical breakdown probability is most commonly represented via a Gaussian or Weibull distribution with the following probability density functions as given in [124]:

$$p(V) = \frac{1}{\sigma\sqrt{2\pi}} e^{-((V-\hat{V}_{50})^2/2\sigma^2)} \tag{3.9}$$

$$p(V) = 1 - 0.5^{[1+((V-\hat{V}_{50})^m/n\sigma)ln2]} \tag{3.10}$$

where $p(V)$ is the breakdown probability for an applied voltage V; $\hat{V}_{50}$ is the voltage which leads to a 50% discharge probability; and n for the Weibull distribution determines the value below which no flashover occurs.

Experimental testing under various voltages and environmental conditions give a range of the potential reductions in breakdown strength due to a wildfire. Authors in [125] indicate that the breakdown strength during a fire can be reduced to about 40% of that of air, whereas [126] find a reduction to 33% of air. The value is even lower for direct current (DC) lines which exhibited a breakdown strength of 20% of air under the test conditions in [127]. Authors in [48] assume a 90% reduction in the breakdown due to smoke and flame and use this to come up with a relationship between smoke, temperature and the breakdown probability.

If the smoke characteristics, historical fire temperatures, pressure and electrode properties are known, equations (3.7)-(3.9) can be used to identify specific breakdown prob-

abilities for an individual line component. However, as weather information for multiple years in the future contains a significant amount of uncertainty, exact future weather parameters are not available. Therefore, (3.7)-(3.9) are not used in this work. Instead, as the breakdown probabilities are highly dependent on the stochastic environmental parameters at the time of ignition, in the planning horizon, historical equipment failure information can be used to identify a general failure probability rate for the system. For this thesis a range of equipment failure probabilities from 0.25 to 1.0 are used as described in Section 3.4. These can be used to identify the overall system resilience with more detailed analyses conducted in the operating horizon.

3.2.3 System Restoration

The duration of wildfire induced outages can span from seconds in the case of momentary outages to sustained outages lasting several days depending on the characteristics of the fire and equipment. Whereas in normal conditions, the restoration time can be reflected by the Mean Time to Repair (MTTR), under emergencies, crews may not be able to access the affected equipment, leading to longer outage times. Two scenarios for restoration are considered in the proposed methodology to reflect the access limitations during a fire. The first is the standard restoration, which assumes that the equipment is restored after the fire is fully contained, often substantially longer than the MTTR. The second restoration scenario - Enhanced Access, which assumes that equipment is restored after the fire passes the equipment in question plus the MTTR, or following full containment of the fire, whichever is shorter. This reflects the ability of crews to reach some of the affected equipment prior to full containment of the fire. The restoration scenarios used are not based on empirical wildfire specific MTTR values, but if available, such values could be used to provide further insight into the restoration impacts.

3.2.4 Resilience Quantification

For the transmission planning horizon (1-10 years) [128], power system planners are tasked with ensuring the sufficiency of the transmission network under a range of contingencies and extreme events [121] prompting the use of infrastructure metrics to evaluate resilience. For this work, transmission lines are selected due to their susceptibility to wildfire induced failure [129, 130]. The wildfire infrastructure resilience of the system in the planning horizon is quantified using three measures: O_M: the maximum number of lines out; O_k: the frequency of an N-k outage; and the area metric LOD (Line Outage Days). The first metric O_M reflects the peak number of transmission lines out of service due to wildfires during a single fire season and is calculated as:

$$O_M = \max_{t \in T} O(t) \tag{3.11}$$

where t represents a single day of all the days T within a wildfire season; and $O(t)$ reflects the number of lines out of service on day t. This is similar to the Λ_{line} metric proposed in [29] with the key distinction that O_M considers the worst value across an entire season as opposed to the number of lines out at a given point in time.

In addition to the O_M value, a probabilistic metric (O_k) provides planners with an expectation of how many days a given level of vulnerability (k) may be reached and is calculated as follows:

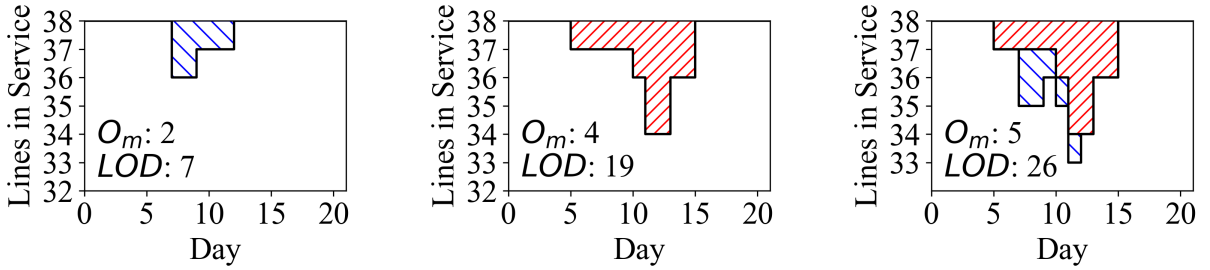
$$O_k = \frac{1}{|T|} \sum_{t \in T} \begin{cases} 1, & O(t) = k \\ 0, & \text{otherwise.} \end{cases} \tag{3.12}$$

where k is the selected level of vulnerability in terms of number of lines out of service.

Finally, LOD is proposed as shown in (3.13) to provide a quantifiable measure of the wildfire resilience over the entire year.

$$LOD = \int_t^T O\left(t\right) \tag{3.13}$$

LOD is a specific form of the generalised "Area" metric presented in [29]. The *LOD* expands the Area to consider the entire wildfire season as opposed to a single event period. As the threat of synchronous wildfire events may become more of a threat [18], *LOD* values are computed over the course of each wildfire season to captures the influence of overlapping wildfires, rather than just a single event. Examples of this metric are shown in Figure 3.2.



(a) Example Fire A

(b) Example Fire B

(c) Example of Overlapping Fires

Figure 3.2: Calculation of O_M and *LOD* for three examples

The convergence criterion for each of the resilience metrics is used to determine $|\mathcal{S}|$. The resilience metric with the lowest rate of convergence should be used as the stopping criteria. The variance of the estimated reliability index $\bar{Z}$ can be calculated as [131]:

$$V(\bar{Z}) = \frac{1}{\mathcal{S}(\mathcal{S}-1)} \sum_{s \in \mathcal{S}} (Z_s - \bar{Z})^2 \tag{3.14}$$

The coefficient of variation β which is an expression of the accuracy level of the Monte Carlo simulation can be determined as:

$$\beta = \frac{\sqrt{V(\bar{Z})}}{\bar{Z}} \tag{3.15}$$

The formulation for each of these metrics reflects the expected value of each Resilience indicator $\bar{Z}$. However, as described in Chapter 2, using the expected value may not adequately reflect the risk for High Impact Low Probability events such as an extreme wildfire season. Therefore, two additional risk informed indices of each metric, Value at

Risk (VaR) and Conditional Value at Risk (CVaR) are used [41] as described in Chapter 2.

The final step in the resilience quantification is the selection and risk-determination of outages for incorporation in long-term planning assessments. The selection of contingencies is based on the simulation of $s \in \mathcal{S}$. Contingencies which occur in at least one synthetic wildfire season are identified to produce a set of unique contingencies $\mathcal{C}$. To identify contingencies with a lower anticipated frequency, $|\mathcal{S}|$ can be increased to align with the risk tolerance of a given entity. An initial screening identification of the risk posed to the system from each contingency also requires a determination of the magnitude of impact. This magnitude is calculated by taking the peak loading of the system and assessing the projected load shed requirements via AC-OPF. For the AC-OPF, the loads are reflected in the economic dispatch as negative generators with a sufficiently large cost of operation with the formulation as discussed in Chapter 2.1.2 and [36].

3.3 Reference Transmission Network and Global Wildfire Data

A case study is used to demonstrate the robustness of the proposed methodology. The main data necessary to implement the proposed methodology are power system network model and historical wildfire data. Region 3 of the Grid Modernization Laboratory Consortium update to the IEEE reliability test network (RTS-GMLC) is used to demonstrate the proposed wildfire resilience approach [132]. The RTS-GMLC is a test network geolocated to Southern California and using the coordinates provided in the RTS-GMLC enable consideration of historical fire behaviour. The assigned geo-location of Region 3 of the RTS-GMLC spans several wildfire prone areas, making it an ideal candidate for this analysis. The geolocation of the RTS-GMLC allows realistic natural hazards, load, along with hydroelectric, solar and wind profiles to be used for the test network. Details of the transmission network for the RTS-GMLC used for this chapter are contained in Appendix

A.

After 2000, the historical wildfire data shows marked difference from prior wildfire behaviour [122]. Furthermore, the 15 largest wildfires in California history have occurred since 2003 [104]. Therefore, historical wildfire information for the case study was collected from the Global Fire Atlas which contains data for 13.3 Million individual fires greater than 0.21 km^2 for the period of 2003 to 2016 [133]. The 14 year duration provides representative coverage of wildfire patterns for this region. For each fire the Global Fire Atlas includes a shapefile with: ignition location and time; perimeter (km); size (km^2); duration (days); average daily fire line (km); average daily fire expansion ($km^2 day^{-1}$); speed (km day^{-1}); and direction of spread [134]. Figure 3.3 shows the location of historical fires [134] overlayed alongside the RTS-GMLC test network [132].

This data was benchmarked against more extensive historical wildfire perimeter data from the California Department of Forestry and Fire Protection (FRAP) [135]. Comparison of the two data sources indicates that the Fire Atlas has good coverage of fires greater than 0.21 km^2, but excludes a significant number of small fires given the limitations in satellite resolution. However, this limitation is offset by the global coverage of the Global Fire Atlas, allowing for worldwide use where areas may not have the same level of reporting or historical record keeping as that available for California.

3.4 Southern California Case Study

For the analysis presented in this case study, the methodology described in Section 3.2 is implemented via a combination of the following: numpy [136] and pandas [137] packages for the data preparation, geopandas [138], rasterio [139], rioxarray, xarray [140], and cartopy [141] for geospatial analysis; matplotlib [142] and seaborn [143] for visualisation; and scipy [144] for analysis. Analysis of the outage impact is scripted in *MATLAB ver. R2018b* with MATPOWER [145] used to conduct the AC-OPF, and the optimization solved using MATPOWER Interior Point Solver [146]. Cartographic boundaries are obtained from the

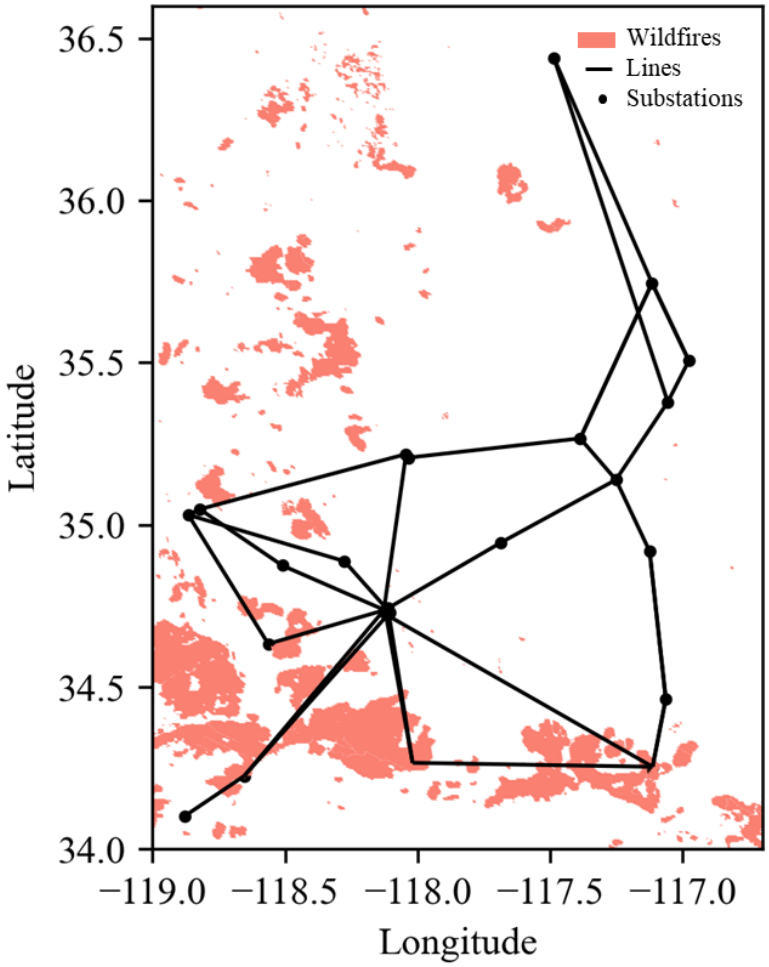


Figure 3.3: Study area mapped with IEEE RTS-GMLC [132] and historical wildfires from [134]

US Census Bureau [147].

Historical wildfire extent, points of ignition and duration are determined from satellite imagery as described in [133]. For this analysis $\mathcal{S}$ is set to 1000. A geospatial bounding box spanning longitudes of 119°W to 116.7°W and latitudes from 34°N to 36.6°N was used to reflect the spatial extent of the test network and collect historical wildfires from the Global Fire Atlas [134]. This resulted in $\mathcal{F} = 467$ wildfires over the period from 2003 to 2016 or approximately 34 fires per year. For the 135 large fires observed over the study period, 26 of those fires intersected at least one transmission line. The chronology of one example fire can be seen in Figure 3.4, burning over the course of eight days.

The next step is to identify an appropriate value of γ. This is selected by identifying

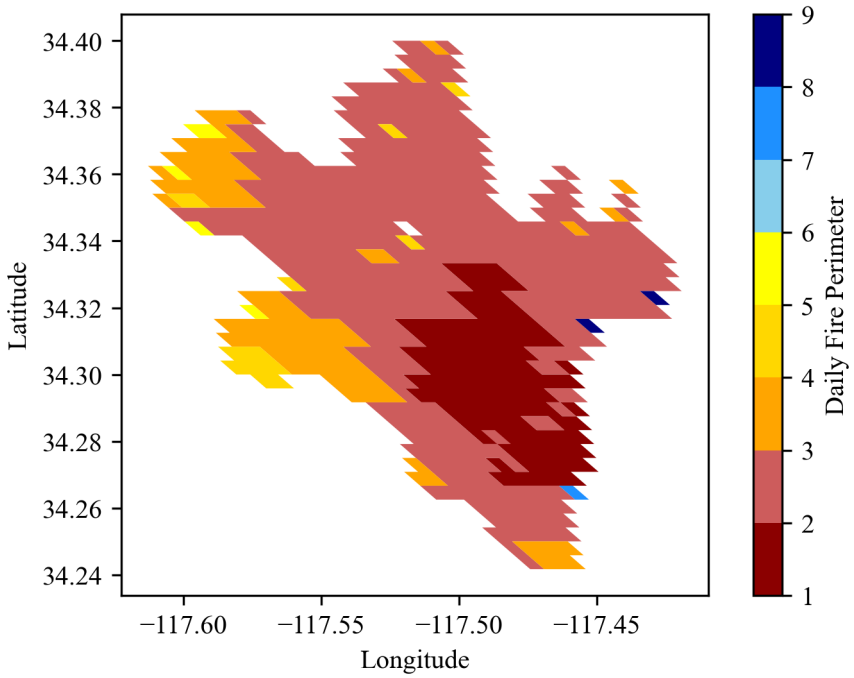


Figure 3.4: Example Fire Perimeter Burn

the inflection point on a logarithmic scale of the historical fires sorted by size resulting in $\gamma = 1000$, providing a trade-off between the number of fires in each group and the relative change in size. This can be seen in Figure 3.5.

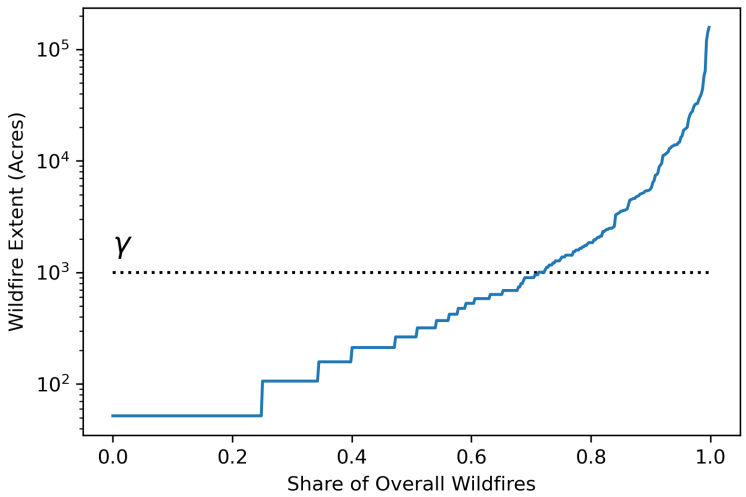


Figure 3.5: Proportion of wildfires by extent, plotted on a logarithmic scale

Next the date of occurrence for $\mathcal{F}_{Large}$ and $\mathcal{F}_{Small}$ are checked for normality using the

Kolmogorov–Smirnov test indicating that both follow a normal distribution. Therefore a truncated normal distribution is fit for each with the resulting distribution and parameters given in Figure 3.6. Comparison of the temporal distribution of $\mathcal{F}_{Large}$ and $\mathcal{F}_{Small}$ demonstrates clear distinction between these groups of fires. The mean date of occurrence for small fires is nearly a month early that that of large fires and the fire season extends substantially longer. This is significant for determining the timeframe of candidate fires within a given synthetic wildfire season s.

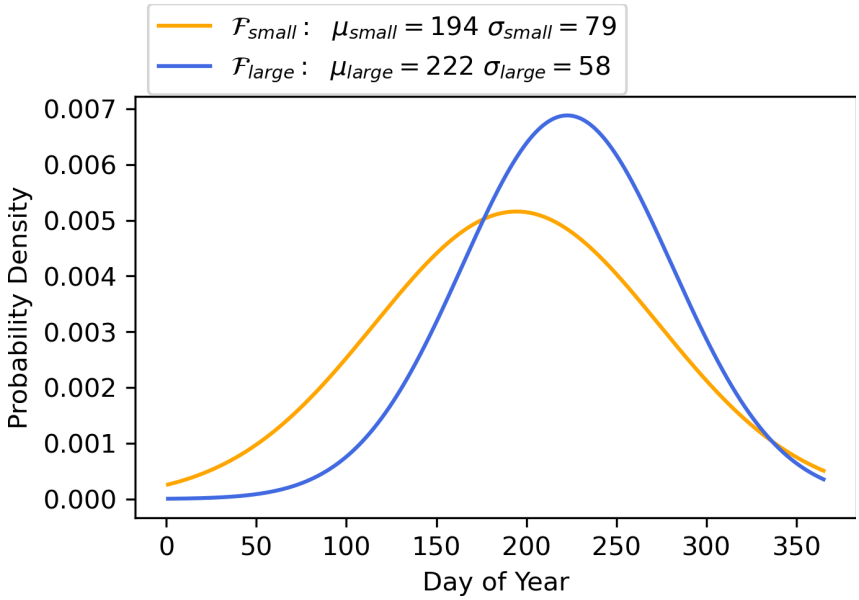


Figure 3.6: Temporal variation in wildfire occurrence by size

Given the distributions and mean annual values of $\mathcal{F}_{Large}$ and $\mathcal{F}_{Small}$, synthetic wildfire seasons can be developed for each $s \in \mathcal{S}$. A visual depiction of 100 synthetic fire years can be seen in Figure 3.7, clearly showing the clustering of fires in the spring and summer months.

As discussed in Chapter 2, the failure probability of a transmission line to a wildfire is dependent on the construction of the pylon, the clearance between the line and the environment, the presence of fuel nearby, the weather conditions, the heat of the fire, the amount of smoke. In the planning horizon (1-10 years) many of these stochastic variables are unavailable such as the exact wind speed and direction during the fire. However, the likelihood of failure of a specific line is bounded by the range of [0, 1]. Four scenarios

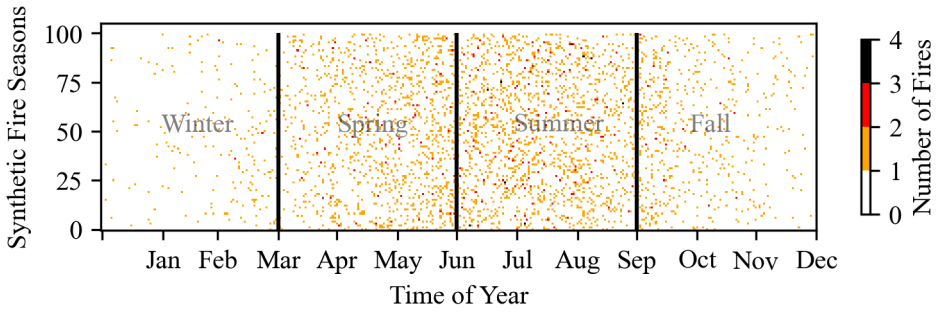


Figure 3.7: Distribution of fires across each synthetic fire season

are employed to simulate this range of risk with failure probabilities of [100%, 75%, 50%, 25%]. A base system failure rate of 50% is selected based on a study of the impact of wildfires in California from 2003 to 2016, which indicated that 44.4% of wildfires in the Southern California Mountains adversely impacted the operation of the grid [148]. The other significant variable is the time taken to restore the affected elements. The two restoration scenarios laid out in Section 3.2, are reflected for the case study. Each scenario is referred to as Scenario F[failure probability] for the base restoration speed and Scenario F[failure probability]-R for the enhanced restoration speed.

3.4.1 Base Scenario Results

A total of 1000 synthetic fire seasons were generated to quantify the resilience of the test system for the case study. An accuracy threshold (β as defined in (3.15)) of 5%, to evaluate the convergence of the resilience metrics was selected for this analysis. Figure 3.8 demonstrates the convergence of the simulation. LOD and O_M converge to 3.0% and 2.2% accuracy respectively indicating the sufficiency of 1000 fire seasons.

Analysis is conducted for $\mathcal{S}$ using the proposed methodology in Section 3.2 and the three resilience metrics are calculated. To identify the risk informed reliability metrics (VaR and CVaR) an α value of 90% is selected, resulting in the top 10% events classified as HILF, to align with the 1 in 10 scenario often used in transmission planning studies [121]. The resulting O_M and LOD are shown in Figure 3.9. Using the expected value of resilience indices can substantially underestimate the magnitude of the impact of a

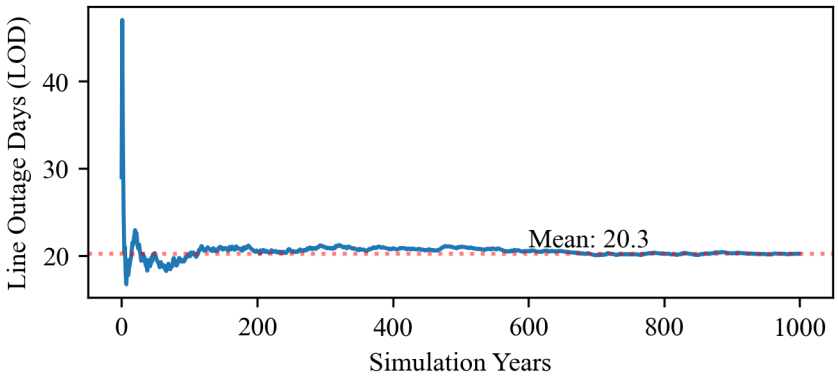
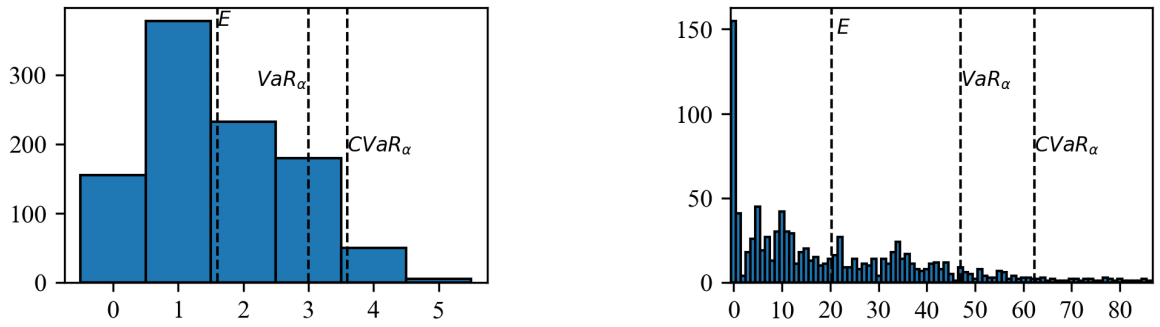


Figure 3.8: Convergence of LOD in $\mathcal{S}$

wildfire season on the transmission infrastructure. For the case study, O_M increases from an expected value of 1.6 to a CVaR of 3.6; increasing from approximately an N-2 to an N-4. Results indicate that when conducting a wildfire resilience plan for the test network, consideration should be given to N-k contingencies ranging from a single outage to outage of at least four elements. Not a single N-6 outage was identified in $\mathcal{S}$. While this does not mean an N-6 is impossible, findings suggest that N-6 outages are highly unlikely to occur and therefore can be excluded from the resilience assessment. However, $|\mathcal{S}|$ can be increased to identify more infrequent outages.



(a) Resilience Metric O_M

(b) Resilience Metric LOD

Figure 3.9: Calculation of Resilience Metrics for Base Scenario

Analysis of the frequency of each level of O_k occurring in the test network reveals that contingencies beyond an N-5 are unlikely to occur, whereas wildfire caused N-1 contingencies are expected to occur on multiple days each year. This further supports the identified range of contingencies to simulate as part of the wildfire resilience assessment.

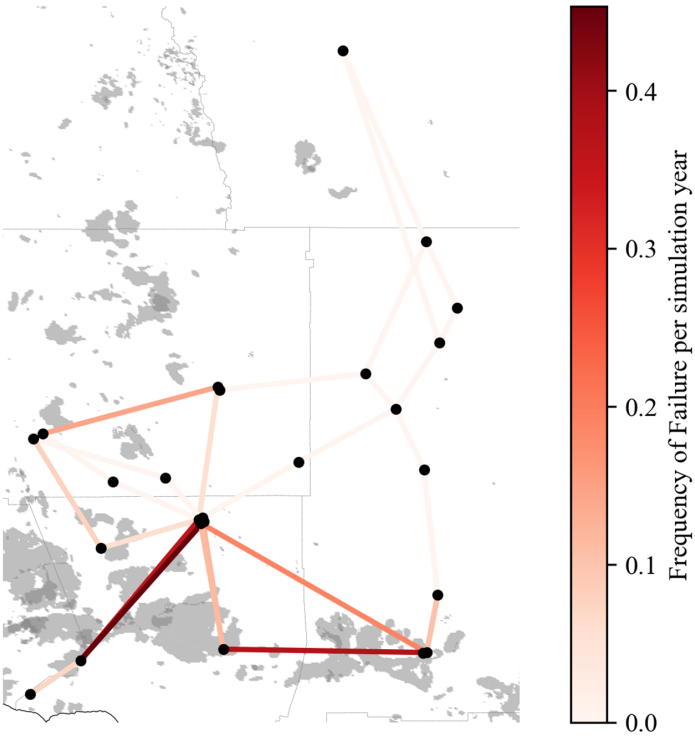


Figure 3.10: Frequency of Failure of Each Component Due to Wildfire

Specific values for each level of O_k can be seen in Table 3.1.

Table 3.1: Frequency of outages at each level of O_k (number of days per year)

	Expected (O_k)	VaR (O_k)	CVaR (O_k)
N-1	6.96	15	19.89
N-2	3.31	11	18.14
N-3	1.79	9	11.21
N-4	0.30	0	3.04
N-5	0.02	0	0.21
N-6	-	-	-

The elements which exhibited the highest frequency of outage are shown geospatially in Figure 3.10 to provide insight into the relative areas most affected by the wildfire seasons. This figure allows stakeholders to quickly identify which lines are most at risk due to wildfires, enabling the creation of effective mitigation strategies and is used to inform the wildfire assessed in Chapter 6.

3.4.2 Impact of Failure Probability and Restoration Framework

A total of eight scenarios were considered with the wildfire fragility including [0.25, 0.5, 0.75, 1] and both restoration strategies for each fragility level. The resilience indicators for each scenario are given in Table 3.2. The fragility of the system can be reduced by system hardening efforts such as replacement or reinforcement of structures, installation of automation, installation of covered conductor and enhanced vegetation management [149]. Utilities should collect historical wildfire repair statistics to further inform the estimates of restoration time following wildfire related interruptions to enhance the accuracy of restoration estimates. While lowering the failure probability and faster restoration both improve the resilience indicators, lowering the failure probability is more effective at reducing O_M, whereas enhanced restoration is more effective at reducing the LOD. Furthermore, the resulting CVaR indicates that even in the most resilient scenario (F25-R) at least N-3 contingencies should be reflected in further transmission planning studies as part of extreme event analyses, which is beyond the range of typical N-1/N-2 contingencies. Results for O_k across the range of outage magnitudes are given in Table 3.3. The maximum observed value of k ranged from 4 (F25-R and F50-R) to 7 (F75 and F100). A key element is that reduction in failure probability and enhanced access contribute to contingencies with a lower number of elements involved. Close proximity in values of CVaR for Scenarios F75 and F100, indicate that there may be a saturation in risk at higher failure probabilities. This highlights the importance of conducting adequate determination of wildfire failure probability for power system elements to determine how significant any proposed adaptation measures may be in enhancing resilience.

3.4.3 Identification of N-k Contingencies for Planning Assessments

While it is informative to have a relative measure of the system performance, the impact of the loss of transmission lines varies based on the associated generation, load profiles,

Table 3.2: Resilience Metrics for Southern California Case Study

Metric	Scenario	Standard Restoration			Enhanced Access		
		E	VaR	CVaR	E	VaR	CVaR
O_M	F25	1.2	3	3.1	1.1	2	2.5
	F50	1.6	3	3.6	1.4	3	3.2
	F75	1.8	3	4.2	1.7	3	3.9
	F100	1.9	4	4.2	1.8	4	4.1
LOD	F25	14	35	49	5	12	16
	F50	20	47	62	8	19	25
	F75	23	53	70	11	25	32
	F100	26	57	75	14	29	39

Table 3.3: O_k Metric for Southern California Case Study

		Standard Restoration			Enhanced Access		
Metric	Scenario	E	VaR	CVaR	E	VaR	CVaR
O_1	F25	6	14	19	3	8	10
	F50	7	15	20	5	10	13
	F75	8	17	22	5	11	14
	F100	8	18	23	6	12	15
O_2	F25	3	10	17	1	3	4
	F50	3	11	18	1	4	6
	F75	3	11	18	2	6	8
	F100	4	11	18	2	7	10
O_3	F25	1	1	7	0.1	0	1
	F50	2	9	11	0.3	2	2
	F75	2	10	12	1	2	4
	F100	2	11	12	1	3	6
O_4	F25	0.1	0	0.5	0.01	0	0.1
	F50	0.3	0	3	0.04	0	0.4
	F75	1	0	7	0.1	0	1
	F100	1	7	9	0.3	2	2
O_5	F25	0.01	0	0.1	-	-	-
	F50	0.02	0	0.2	-	-	-
	F75	0.04	0	0.4	<0.01	0	0.03
	F100	0.05	0	1	0.01	0	0.07
O_6	F25	-	-	-	-	-	-
	F50	-	-	-	-	-	-
	F75	0.02	0	0.2	<0.01	0	0.03
	F100	0.02	0	0.2	<0.01	0	0.04
O_7	F25	-	-	-	-	-	-
	F50	-	-	-	-	-	-
	F75	<0.01	0	0.05	-	-	-
	F100	0.01	0	0.06	-	-	-

and operating conditions of the system. Therefore, in addition to resilience metrics, the methodology identifies a set of unique contingencies for consideration in planning studies as part of extreme event analysis. For this case study, 13 transmission lines are affected by wildfires over the study period and could result in 8,191 unique contingencies. While it might be feasible to evaluate several thousand contingencies, the number of contingencies grows rapidly as the system increases in size. Therefore, methods to identify the most frequent and significant contingencies affected can be beneficial to identify a plausible set of outages to consider as part of transmission planning assessments.

Application of the proposed methodology for $\mathcal{S}$ revealed 84 unique contingencies across 13 transmission lines, with a magnitude ranging from N-1 to N-5. Evaluation revealed that outage of at least three lines due to wildfire occurred in 24% of wildfire seasons with 6% of wildfire seasons seeing four or more lines out. This methodology provides a way to simulate the worst conceivable outage (with the likelihood dependent on a set probabilistic threshold). In this case it was an N-5. While the case study found it unlikely (one day in 50 years) that five separate transmission elements would experience wildfire outages concurrently, planners should be aware of the possibility in considering the potential for system cascading or other risks as part of their security assessments for the planning horizon. With authors in [18] noting that days with synchronous fire danger are anticipated to be much more likely in the future, preparation for such events is important.

However, the risk to the transmission system cannot be directly correlated to the number of lines lost due to variation in dispatch, load patterns and redundancy. Hence, the final step in the overall resilience assessment is to simulate the identified contingencies at peak system load to identify the risk of load shedding as described in Section 3.2. Figure 3.11 presents the number of times each outage is observed over the course of $\mathcal{S}$ along with the amount of load shed during peak load on the system using the emergency line ratings given in Table A.4. Hourly analysis of a two week wildfire in the operating horizon will be presented in Chapter 6.

For this test network, much of the load shed is due to wildfires that disrupt lines connecting to substation 308 as also shown in Figure 3.10. Therefore, if the wildfire risk is sufficiently high, these lines would be good candidates for hardening via means such as installation of covered conductor, increasing insulator clearances enhanced vegetation management, undergrounding, or if risk is sufficiently high, installation of line in a new right of way [14, 150, 151]. Alternatively a local microgrid could be set up to serve these islanded substation loads during an outage as described in [152].

When considering each of these proposed mitigation solutions trade-offs between risk reduction and risk cost must be assessed against the risk. Utilities examining these trade-offs in wildfire mitigation refer to this as risk-spend efficiency [149]. For example a California utility compared covered conductor with undergrounding as part of a wildfire resilience assessment and found that although undergrounding was more effective at reducing their identified wildfire drivers (98% reduction compared to 70%), it was roughly eight times more expensive [151]. The proposed methodology in this chapter provides the foundational risk information necessary to inform such risk-spend efficiency analyses.

3.5 Summary

The number of concurrent wildfires is increasing in some areas due to climate change. As multiple fires occur simultaneously, it can be difficult for transmission planners to determine the likelihood of overlapping contingencies during a wildfire season to include into transmission planning assessments. This chapter presented a probabilistic wildfire resilience methodology enabling the quantification of the exposure and subsequent resilience of the transmission network over the course of a wildfire season.

The proposed methodology supports the mitigation selection process by providing the resilience metrics necessary to inform risk spend efficiency calculations and justify capital expenditures to mitigate risk. The case study showed that use of traditional expected value reliability metrics can lead to the underestimation of risk to the transmission net-

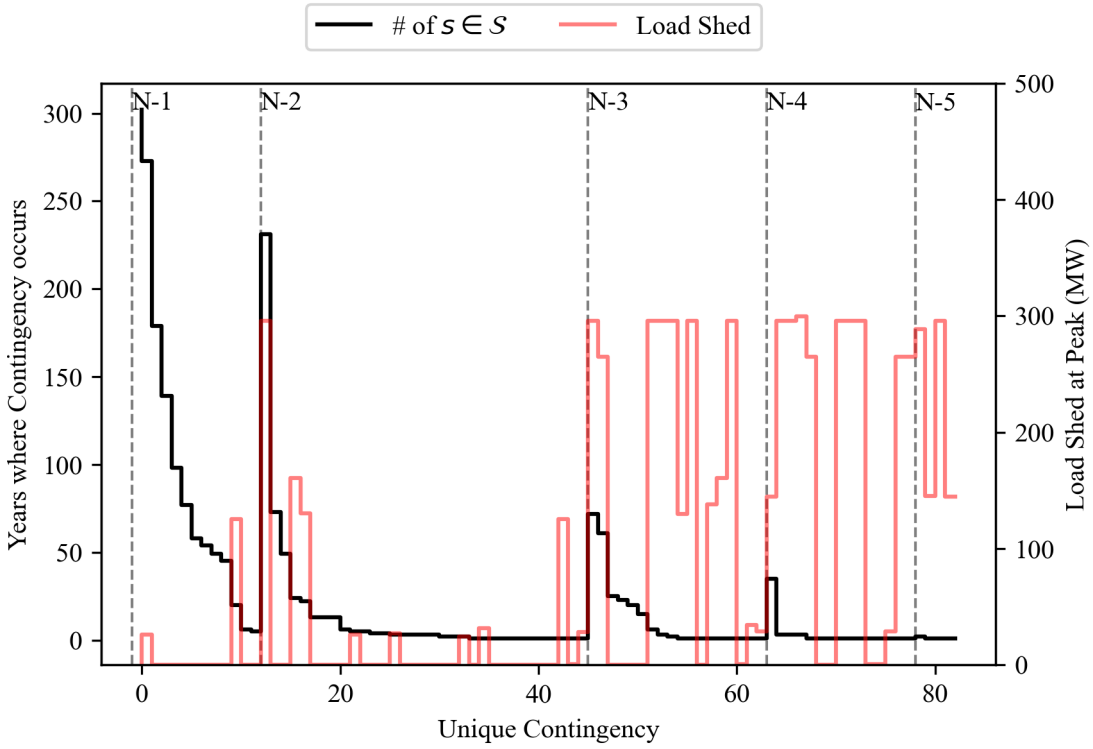


Figure 3.11: Frequency of Exposure to Wildfire Contingencies alongside the amount of load shed under peak conditions.

work during extreme wildfire seasons, whereas CVaR can reflect a means to quantify the risk over the worst expected cases. Furthermore, the case study argues that while improving the restoration time may significantly reduce the overall LOD experienced during a wildfire season, reducing restoration time had minimal impact on O_M. In order to reduce O_M, strategies such as infrastructure hardening or enhanced vegetation management are suggested to reduce the exposure to wildfires.

The proposed methodology also provides the N-k contingencies which should be included in planning assessments to proactively identify future grid needs to maintain wildfire resilience and prepare the grid for such extreme events. Finally, assessing the anticipated contingency frequency and load shed under peak loading conditions enables efficient identification of the contingencies which pose the most threat to the reliable operation of the transmission network.

Chapter 4

Solar Prosumer Location Detection to Inform Resilience Plans

Distributed Energy Resources (DER) and microgrids are becoming attractive mitigation options to enhance power system wildfire resilience especially in areas such as Australia and Southern Europe that have significant solar resources [152]. The operating status and location of existing rooftop solar PV generation installations is a critical input to assess the feasibility, cost, and time required to implement these solutions. However, as reported in 2020, "in some countries, including the UK, small systems may legally be grid connected with no registration, meaning there may be no official documentation of their existence" [71]. In the UK at present (2022), installers of small PV facilities (a capacity of 16A per phase or less) have up to 28 days to notify Distribution Network Operators (DNOs) following installation [72]. Despite existing requirements for customers to notify utilities, DNOs may be unaware of the location of solar PV facilities due to unreported or unauthorised installations or incorrect data entry [73].

This chapter presents an advanced agglomerative clustering-based approach to automatically identify solar prosumers through use of hourly net smart meter data. The hierarchical structure of agglomerative clustering provides an effective way to progressively group customers until two clusters are reached; one for prosumers and another for

consumers. The computational burden and storage requirements are reduced using PAA while retaining high accuracy. The approach also proposes the creation of SPIDCs to give insight into the duration of historical data necessary to achieve accurate identification.

The overall methodology for solar prosumer identification and SPIDC development is given in Section 4.1. The significance of the approach is that it enables effective identification of solar prosumers, demonstrated using a case study with hundreds of real prosumers providing improvements over existing methods in terms of accuracy and data required. The case study details are presented in Section 4.2 and the results are presented in Section 4.3. Applications of this approach to enhance wildfire resilience planning are provided in Section 4.4. The chapter concludes with a summary in Section 4.5. As a contribution to the state-of-the-art, the publication [P1] resulted from the research described in this chapter.

4.1 Solar Prosumer Identification Methodology

The main stages of the proposed methodology for solar prosumer identification include: 1) Data Preparation, 2) Model Selection, and 3) Solar Prosumer Identification Duration Curve (SPIDC) creation. Figure 4.1 depicts the stages in the methodology. Each stage is discussed in detail below.

4.1.1 Data Preparation

When working with real data, pre-processing is necessary as the quality of data driven methods is directly linked to the quality of the input data [84]. For smart meter measurements this consists of eliminating missing data, repeated values, and sensor errors. First, customers with missing measurements over the time period in question are removed. Next, erroneous metering measurements (Err) are calculated based on (4.1) for each customer and any customer with erroneous measurements is removed (with an erroneous

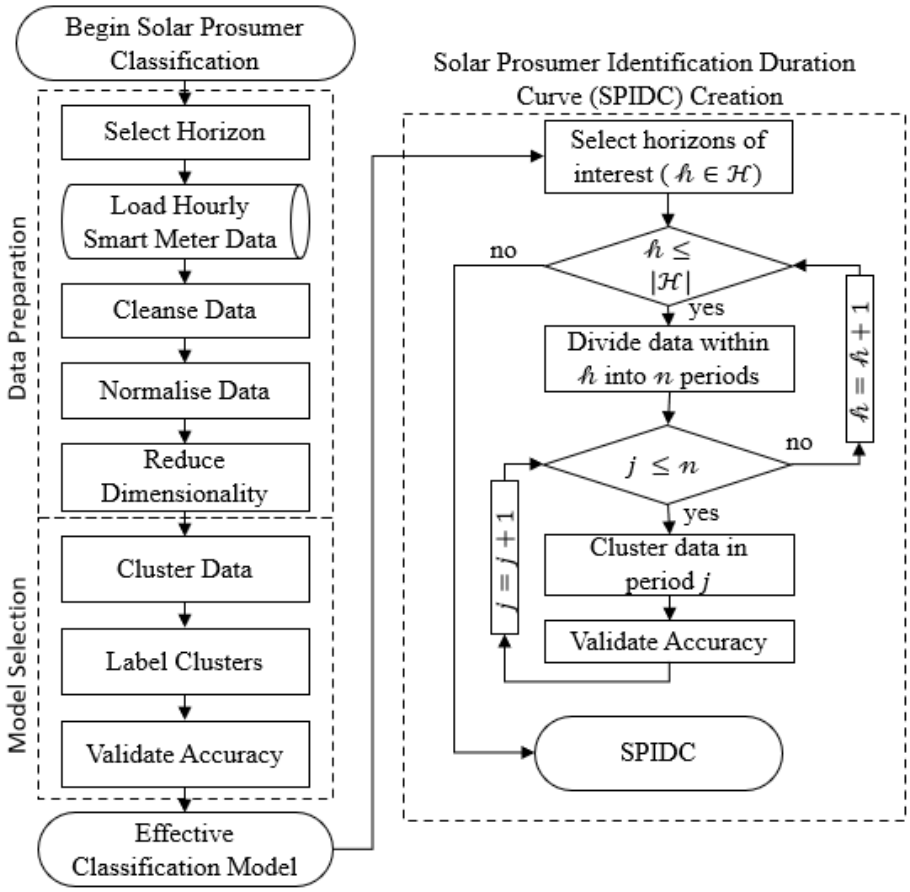


Figure 4.1: Methodology for Solar Prosumer Identification Duration Curve creation. Adapted from [P1]

measurement identified where (4.1) is true).

$$|\mathcal{L}_i - \mathcal{P}_i - \mathcal{N}_i| > 0.05 \tag{4.1}$$

where i is an hour over the period of interest; $\mathcal{L}$ is the gross metered demand; $\mathcal{N}$ is the net metered demand; and $\mathcal{P}$ is the metered generation. Finally, the standard deviation is used to identify customers with stuck readings. Any customer with a standard deviation of <0.05 kWh over the period is removed.

After data is cleansed, the smart meter data from each customer may be normalised to minimise the impact of differences in load magnitude influenced by factors such as the number of appliances installed or the presence of other energy intensive LCTs like electric vehicles or heat pumps. Application of the normalisation step depends on the use case,

and the specific model selected. For example, normalising PAA_A-min or PAA_A-max is ineffective as all customers would have the same value. Min-Max normalisation is used which converts $x = x_1, x_2, \ldots, x_n$ to a range of [0, 1] via [84]:

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{4.2}$$

where x'_i and x_i reflect the normalised and actual recorded electricity consumption at time i; x_{min} and x_{max} are the minimum and maximum demand within period n.

The final data preparation step is to apply dimensionality reduction to improve computational efficiency and reduce storage requirements. The use of PAA is proposed as it is superior to other established methods [89]. Of the six aggregation methods, given in Chapter 2.3.1, use of PAA with minimum aggregation over an annual period (PAA_A-min) is proposed in this chapter with the formulation given in Equation (2.16). PAA_min is selected based on domain knowledge that solar prosumers can feed power back to the grid (i.e. a minimum demand less than 0) whereas consumers are limited to a minimum demand of 0. The annual period (A) is selected to reduce each customer energy consumption profile from 8760 individual smart meter reads, to a single data point (the customer's minimum demand over the considered time-frame) thereby reducing the information needed to identify prosumers.

4.1.2 Model Selection

After data cleansing, and dimensionality reduction via PAA_A-min, the approach uses agglomerative clustering to group customers into two groups (solar prosumers and consumers). Ward linkage [86] is used to minimise the variance as clusters grow in size. Agglomerative clustering is an unsupervised learning technique, meaning labelled input data is not required to train the model. The output of the clustering is two unlabelled groups of customers. Domain knowledge is used to automatically label the resulting clusters as prosumers and consumers to improve process efficiency. As solar prosumers offset

a portion of their energy consumption via their on-site generation, much of the annual energy consumption is able to be offset, resulting in low net energy consumption over a given period in comparison to a traditional consumer. This informs the assumption that the group that contains the individual customer with the lowest energy consumption should be labelled 'solar prosumers'. Normalisation is not applied for the proposed approach (PAA_A-min), however for models using normalisation, after application of normalisation the normalised load for a solar prosumer is often higher than that of a typical consumer and therefore the labels are switched. Prior to use for a specific region, these assumptions should be validated through comparison of customer annual energy consumption as shown in Section 4.3.1.

Two metrics are used following cluster labelling to evaluate the method performance: accuracy and F1 Score. Use of the F1 Score in addition to accuracy allows for better evaluation of a bias in accuracy resulting from unbalanced clusters (e.g. only a few consumers or solar prosumers). The calculation of accuracy and F1 score is as given in (2.12)-(2.13).

4.1.3 Solar Prosumer Identification Duration Curve (SPIDC) Creation

After validation of the solar prosumer model performance, and running the model for a given region, it is important to understand how quickly new installations or changes in existing installations can be identified. SPIDC provide a classification accuracy nomogram to address this question. The steps to generate the SPIDC are:

1. **Select the horizon(s) of interest $\mathcal{H}$.** The important factor when conducting this step is to select a range of horizons to provide insight into variations in the model performance along with any seasonal variance in accuracy. Useful horizons to consider include: one week, one month, and one year providing examples of near-term, medium-term and long-term performance.

2. **Split the historical data by horizon.** As an example, with a monthly horizon and one year of historical data, the data would be split into twelve periods within the year.

3. **Evaluate the model over each horizon.** The classification model is run for each resulting period and evaluated using model accuracy (2.12).

4. **Identify risk tolerance for proposed use case.** A percentile should be selected that reflects the confidence in performance that meets the use case needs (for example the 50^{th} percentile would reflect the typical model performance over the horizons in the historical period). Several percentiles can be selected to reflect the confidence needs under different scenarios.

5. **Determine update regularity.** The update regularity will differ based on the use case. For example, in planning studies annual updates coinciding with the start of the plan may be sufficient, maximising historical data to achieve the highest accuracy. However, for operational studies, the ability to provide more regular updates with incremental reduction in accuracy may be desirable to enable changes to customer classification to be more quickly identified.

4.2 Data from Real Consumers and Solar Prosumers

Real hourly energy consumption data from residential customers located in the states of California, Colorado, and Texas was collected from the Pecan Street Dataport [153] for the period from 2013 to 2016. Exploratory data analysis indicated the necessity of cleansing of the smart meter data, including customers with missing, erroneous, or flatline (measurements without any change) data. Table 4.1 provides the number of customers after each stage of the data cleansing process with the final number of customers used for the annual classification ranging from 130 to 364. However, many of the customers are only missing data for specific portions of the year. Hence, as time periods shorter

than a single year are considered, the number of customers with clean data increases. This highlights the importance of providing operators with insight into how much data is necessary for solar prosumer identification.

Table 4.1: Unique customers remaining after annual data cleansing [P1]

Component	**2013**	**2014**	**2015**	**2016**
Customers (#)	439	684	626	463
Remaining customers: after removal of those with missing data	144	297	374	317
Remaining customers: after removal of those with anomalous data	138	278	365	309
Remaining customers: after removal of those with flatline data	130	278	364	305

As the Pecan St. is an active large scale research database, customer enrolment happens throughout the year and the specific customers can change from one year to the next. Therefore, the total number of unique customers across the four-year period considered (733) is larger than the number of customers in any one year. Across these customers, 269 solar PV installations and 104 electric vehicles were included. Evaluation of the power generation data for each customer is performed to determine whether a customer had a PV installation in a given year and provides a true label to compare to the label produced by the classification model. Although, true labels and PV generation are available, the proposed identification approach does not use these as inputs, but uses only customers' net energy consumption.

4.3 Case Study

The data outlined in Section 4.2 are used to validate the accuracy of the proposed agglomerative method using the aforementioned smart meter data from consumers and prosumers in the United States. Two other clustering methods are also evaluated to provide a comparison of the performance. To validate the suitability of the proposed approach, scenarios are conducted, varying the method across each decision variable as follows: 1) Normalisation: (no normalisation, normalisation); 2) Clustering method: (agglomerative, k-means++, SOM); 3) Aggregation method: (minimum, mean, maximum);

and 4) Dimensionality reduction technique (PAA, statistical method (STAT)). For PAA, the following window sizes were also evaluated 1H, 2H, 4H, 6H, 8H, 12H, 24H, Monthly (M), and Annually (A); where H is hours. This results in 180 model combinations.

The approach was implemented in `Python` using the following packages: `numpy` [136] and `pandas` [137] for data analysis; `sci-kit learn` [154] and `minisom` [155] for creation of the models; and `matplotlib` [142] and `seaborn` [143] for visualisation.

4.3.1 Effective Identification of Solar PV Prosumers

The proposed advanced agglomerative clustering approach listed in Section 4.1 is used to identify solar prosumers. Clustering using both the cleansed and normalised input data for each of the three methods (agglomerative, kmeans++, and SOM) is performed for each year. This is done using a $C \times 8760$ matrix; where C is the number of customers with complete data for a given year as shown in Table 4.1. Each row of the matrix consists of the hourly energy consumption data for a single customer. For the kmeans++ method, the number of iterations was set at 300 and the centroids are re-initialized 10 times and for the SOM the number of iterations was set at 1000 with batch training, sigma was 0.5, and the neighborhood function was Gaussian. Table 4.2 shows the classification results for each method. The results demonstrate that the proposed agglomerative clustering method outperforms the alternative methods based on the evaluation metrics for both cleansed and normalised data. Furthermore, the accuracy with cleansed-normalised data ranges from 98 to 100% demonstrating the efficacy of the proposed approach in correctly identifying solar prosumers using solely net energy consumption data.

The accuracy of the proposed approach outperforms other existing methods to identify solar prosumers while requiring only smart meter data. K-means is used in [83] and is shown in Table 4.2 to be less accurate. Authors in [82] proposed a machine learning based approach that identifies solar prosumers using weather data to identify changes in performance and the load shape, using 2015 data from Pecan St (the same database used in this chapter). The average accuracy reported in [82] is 95.51%, which is less than the

98% for 2015 using the approach proposed in this chapter. Satellite data is used in [76] reporting 95% Precision and 80% Recall yielding an F1 Score of roughly 87%. This is substantially lower than the F1 Score for the proposed approach (98 to 100%) given in Table 4.2. The comparison in accuracy with these approaches demonstrates the ability of the proposed approach to obtain higher accuracy with less data.

Table 4.2: Classification Method Evaluation [P1]

Method	**Aggregation**	**Accuracy (%)**				**F1 Score (%)**			
		2013	**2014**	**2015**	**2016**	**2013**	**2014**	**2015**	**2016**
Agglomerative	cleansed	97	98	98	97	98	98	97	98
	cleansed & normalised	100	98	98	98	100	98	98	98
k-means++	cleansed	95	96	68	94	97	95	75	94
	cleansed & normalised	98	96	96	96	98	96	96	97
SOM	cleansed	95	96	78	93	97	96	80	94
	cleansed & normalised	96	96	96	94	97	95	95	95

To validate the automated cluster labelling assumptions made in Section 4.1.2 a comparison of the annual cleansed and normalised energy of the true customers groups is calculated (with the process to obtain true labels described in Section 4.2). Across the case study period, the minimum cleansed annual energy (kWh) are -2912 kWh and 245 kWh for solar prosumers and consumers respectively, and normalised values of 99 for consumers and 1573 for prosumers as shown in Figure 4.2. This validates that prosumers have a lower minimum annual energy consumption for cleansed data, but that the relationship is inverted after normalisation, demonstrating that in absence of labelled data, minimum energy usage can provide an effective means of overall cluster labelling. While this approach is demonstrated to be effective for the case study, a few known consumers/prosumers could be used to further validate the overall cluster labelling.

4.3.2 Dimensionality Reduction for Enhanced Efficiency

In the United States, the average number of customers served by Investor Owned Utilities (IOUs) in 2017 was 654,600 [156]. However, the number can be significantly higher with largest US IOUs serving 5-5.5 million customers [156] and the largest DNO in the UK

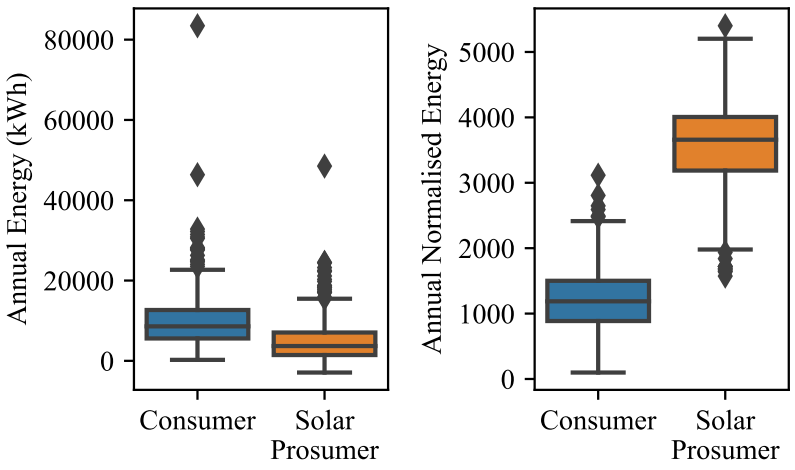


Figure 4.2: Comparison of annual energy consumption of true solar prosumers and consumers, adapted from [P1]

serving over 8.3 million customers [157]. Although the proposed prosumer classification method using agglomerative clustering of normalised smart meter data without dimensionality reduction was highly accurate (98-100%), classification with a single year of data requires storage of a $M \times 8760$ matrix where M is the number of customers being identified. The number of customers to classify can be significantly larger than the number used in the case study, leading to increases in data storage and computational burden. Attempting to cluster even the average number of customers for a US IOU using hourly resolution for a year, requires storage and computation of a matrix with over 5 billion data points. Application of dimensionality reduction reduces the size of the clustering matrix to $M \times w$ in the case of PAA where w is the selected number of dimensions and $M \times 24$ when using the statistical methods. Therefore, dimensionality reduction is significant to reduce the storage and computational requirements.

The impact of each dimensionality reduction technique on the accuracy of the proposed clustering model with and without normalisation is evaluated. The results are shown in 4.3, and demonstrate that agglomerative clustering again produces the highest accuracy across the three methods (agglomerative, kmeans++, and self-organizing map (SOM)). Overall, the performance shown in Table 4.3 demonstrates that the proposed classification

Table 4.3: Average Classification Method Accuracy (%) over the period of 2013-2016 with best in each dimensionality reduction technique underlined [P1]. (H:Hours, M:Month, A:Annual, and STAT:statistical)

Cluster	Method	1H	2H	4H	6H	8H	12H	24H	M	A	STAT
Agglom.	Max	98	76	66	52	58	47	49	50	55	57
	Mean	98	97	88	72	88	63	64	62	57	95
	Min	98	**98**	**99**	**99**	**99**	**99**	**99**	**99**	**99.7**	**99.7**
	Norm-max	**99**	97	91	94	94	94	93	78	43	83
	Norm-Mean	**99**	98	98	97	97	97	96	97	97	97
	Norm-Min	**99**	99	98	99	98	99	99	83	43	99.5
k-means++	Max	88	70	62	55	58	54	52	52	55	54
	Mean	88	88	78	72	77	64	64	62	64	78
	Min	88	95	95	95	95	94	95	97	98	98
	Norm-max	97	96	95	95	95	93	91	73	44	84
	Norm-Mean	97	97	97	97	97	96	96	96	96	97
	Norm-Min	97	97	97	97	98	96	96	84	44	98
SOM	Max	91	78	60	56	57	55	53	49	55	55
	Mean	91	91	89	79	82	65	64	64	64	89
	Min	91	92	95	95	95	95	95	97	98	98
	Norm-max	96	95	95	94	94	93	92	74	44	84
	Norm-Mean	96	96	96	96	96	95	95	95	95	95
	Norm-Min	96	96	97	97	98	96	96	73	44	97

method for solar prosumer identification (Agglomerative clustering with PAA_A-Min) can simultaneously reduce the amount of data required to identify solar prosumers to as little as a single data point per customer for each year and maintain an accuracy of nearly 100%. Dimensionality reduction can improve accuracy while using less data because some hours, such as those when the sun is not shining, may add noise to the classification. Hence, by reducing the data used and selecting data where differences between a consumer and prosumer may be more pronounced (in this case a single point representing the minimum annual energy consumption) this noise can be reduced or eliminated yielding a more accurate classification.

Although PAA_A-min shows the highest performance, for regions where consumers may have a high base load or install solar to cover only a portion of their overall energy consumption, the minimum load may be indistinguishable across customers. For such areas, Statistical dimensionality reduction methods (STAT) as described in Chapter 2 can

provide a balance between data reduction and accuracy by reflecting the daily variation in load profile rather than a single point. Figure 4.3 shows a comparison of the STAT methods of dimensionality reduction for the year 2016, demonstrating clear distinction between the solar prosumer/consumer clusters.

The results presented demonstrate that PAA does not compromise the accuracy of the model. Therefore, it serves as a means to reduce the data storage and computational burden required for identification. To compare the impact of dimensionality reduction on data storage requirements, the assumption is made that each smart meter measurement requires eight bytes of memory. Figure 4.4 shows the corresponding increase in data storage, plotted on a logarithmic scale. For an average sized IOU in the US, application of dimensionality reduction could reduce the data required from tens of gigabytes (GB) to several megabytes (MB). In addition to storage, the computational performance of the algorithm can be adversely effected if the size of the matrix for calculation grows beyond the amount of computer random access memory (RAM) available. The time it took to cluster customers for each year was determined using the `Python` 'timeit' function with 1000 trials, with the minimum time used to limit variance. The computer used for this purpose had a 1.8 GHz processor and 16 GB of RAM. The resulting times ranged from approximately 0.1 to 0.7 seconds for the full hourly data, and 1 to 4 ms after applying the PAA_A dimensionality reduction technique. Using 2013 as an example, the cluster time was reduced from approximately 0.1 seconds using the full hourly data to 0.001 seconds after application of PAA_A, a hundredfold improvement in computational burden, with larger improvements for years with more customers. Therefore, dimensionality reduction can provide substantial benefits in terms of storage and computational efficiency without compromising the classification accuracy.

4.3.3 Duration of Data Required for Identification

The last component of the case study is SPIDC creation. To assess performance over a range of horizons, the four years of data are divided into the following horizons: weekly

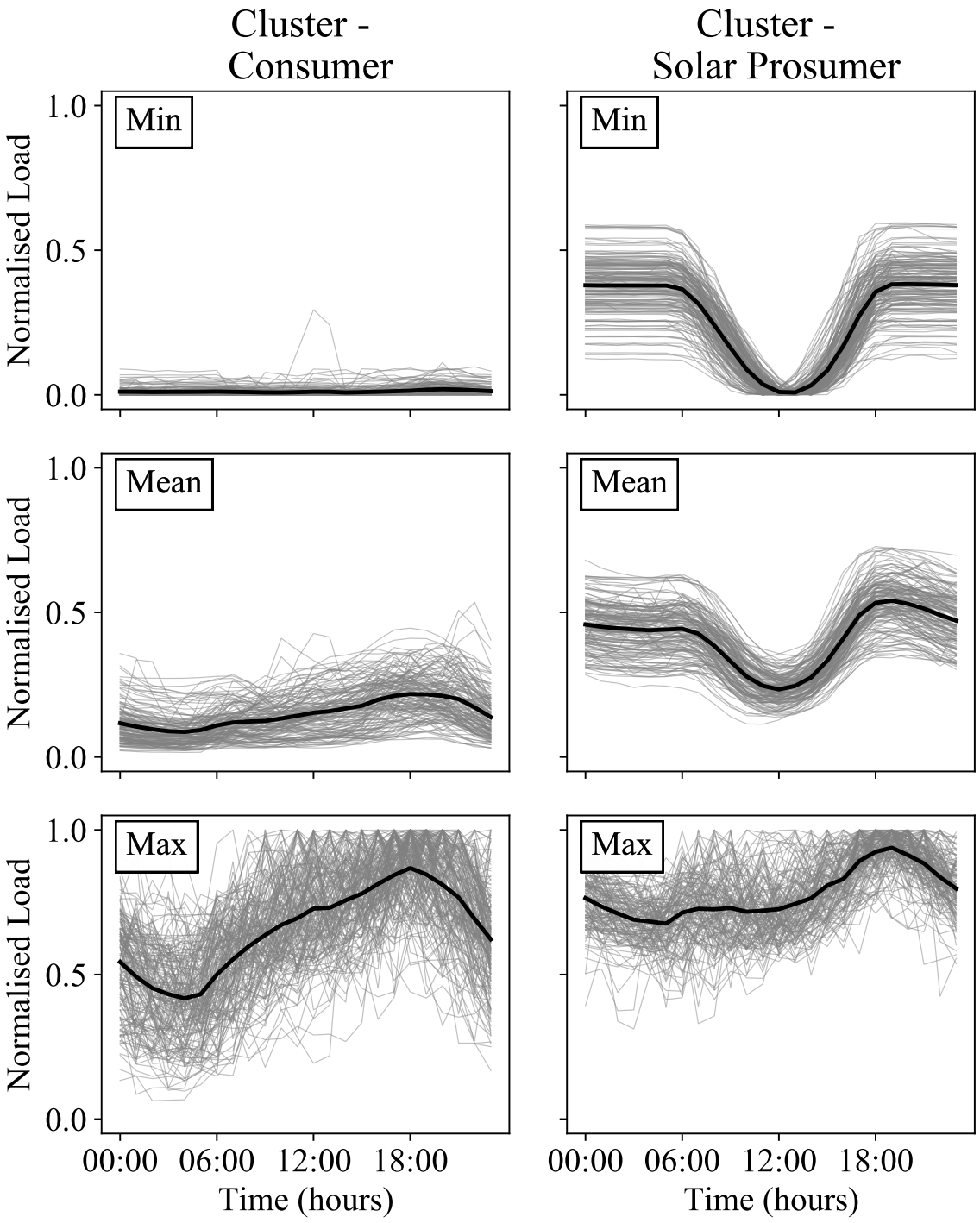


Figure 4.3: Comparison of three statistical aggregation methods for 2016 [P1]. Individual customer profiles shown in grey with the mean shown in black.

(208 periods as the partial week 53 was excluded); monthly (48 periods); and annually (four periods). If a period covers less time than the dimensionality reduction technique

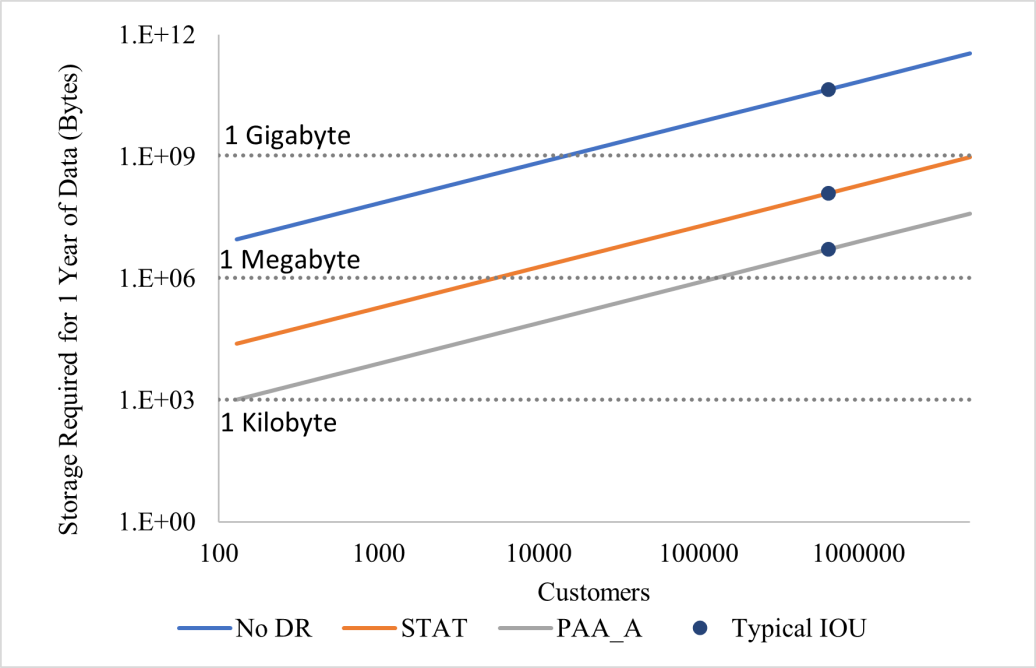


Figure 4.4: Indication of growth in storage requirements for classification at 8 bytes per individual smart meter measurement (x and y axes on logarithmic scale)

window (e.g. one week is less than the year used in PAA_A) the technique for dimensionality reduction is updated to provide a single data-point per period available ($w = 1$) (except for the special case of a period spanning two calendar months where two data-points are used for the period).

The methodology shown in Figure 4.1 was followed using the proposed agglomerative clustering approach with PAA_A-Min to evaluate the performance across these horizons. To develop the SPIDC, weekly, semi-monthly, monthly, 2-month, 3-month, 6-month, and annual horizons were evaluated. As the performance varies across the horizon selected, percentiles are used to define the acceptable historical performance. The resulting SPIDC showing a conservative risk tolerance (10th percentile) and typical tolerance (50th percentile) is provided in Figure 4.5. The selection of percentiles enable evaluation of the trade-offs between accuracy and the historical data available, capturing the risk tolerance of a particular use case.

The seasonal evaluation further confirms that the proposed solar prosumer identification method (Agglomerative PAA_A-min) is more accurate compared to kmeans++ and

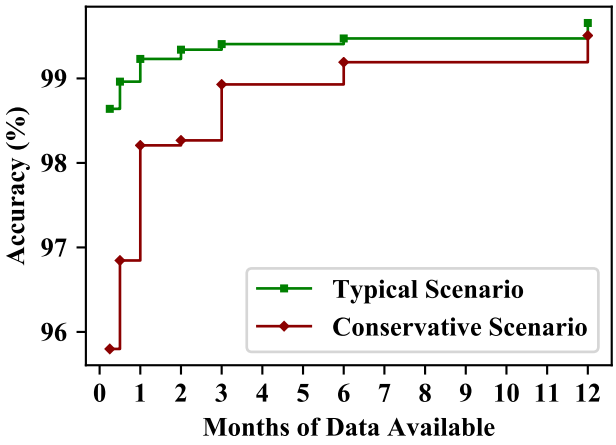


Figure 4.5: Example Solar Prosumer Identification-Duration Curve [P1]

SOM even with limited historical horizons. Figure 4.6 depicts the changing seasonal accuracy with a median weekly accuracy of 98.6% and median monthly accuracy 99.2% for the agglomerative clustering method. While SPIDC provide insight into the accuracy at different horizons, there is also a seasonal component to the accuracy as demonstrated in Figure 4.6. Comparison of the accuracy by season for weekly and monthly horizons, shows that across the four years in the case study, the summer months (July-September) where the loading is typically highest produce lower classification accuracy as a result of the lower share of daily energy consumption produced by the PV facility. This can be seen in Figure 4.6. Areas with different seasonal patterns (such as an area in another hemisphere or a winter/dual peaking area) may observe different monthly accuracy. Thus, when implementing the approach, it is important to evaluate the seasonal loading patterns specific to the particular region being assessed.

4.4 Application of Solar PV Detection to Enhance Resilience Mitigation Planning

Knowledge of the location of solar prosumer installations enables DNOs to account for their presence during switching, reconfiguration, or restoration in response to a wildfire.

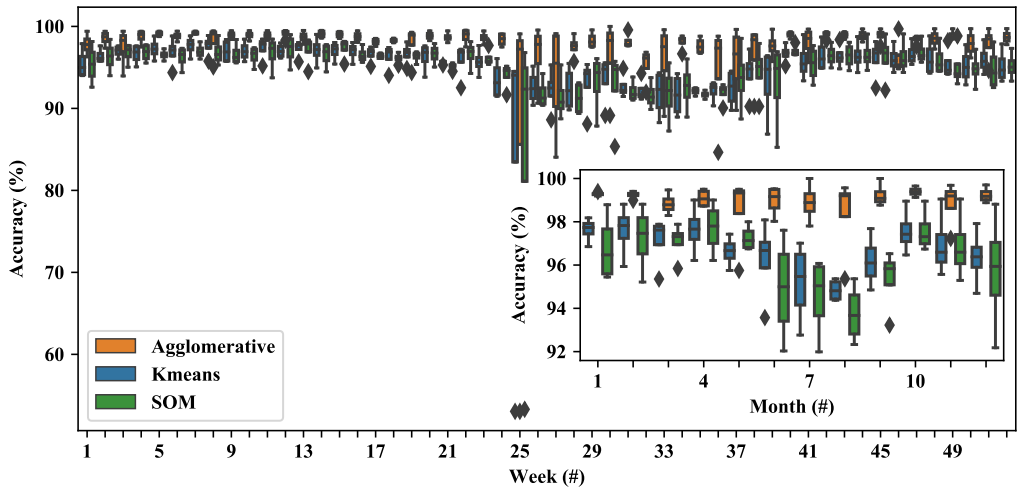


Figure 4.6: Solar prosumer classification accuracy in relation to season of the year [P1]

This knowledge is significant as PV may mask the true feeder load [90] as solar prosumer load profiles are likely to be quite different from consumers as shown in Figure 4.3. Accurately labelled customers can help to identify unauthorised installations, or clusters of solar prosumers for use in grid restoration or microgrid formation. Operator awareness can also be enhanced by using the proposed method to identify the customers that switch from one group to another over a period of time. Changes in group can reveal periods where a solar prosumer's PV installation may be offline or operating abnormally. An example of the application of this process using a monthly horizon is shown in Figure 4.7 for 60 customers who had data for the entire four year case study period. Where the model identified a change in monthly classification for more than one consecutive month, data were investigated for anomalies. This resulted in the identification of customers whose solar PV were not generating or were operating at a lower output level. Use of this approach could enable DNOs to be aware of the current state of PV installations for a given planning area when evaluating wildfire mitigation options. During periods of high wildfire risk, utilities in California are performing Public Safety Power Shutoffs (PSPSs) to pre-emptively de-energize portions of the grid in high fire risk areas [158]. Unauthorised PV installations may also lead to safety concerns for utility workers or others if not

properly installed during such conditions. The identification of such installations prior to a PSPS can help remedy such risk.

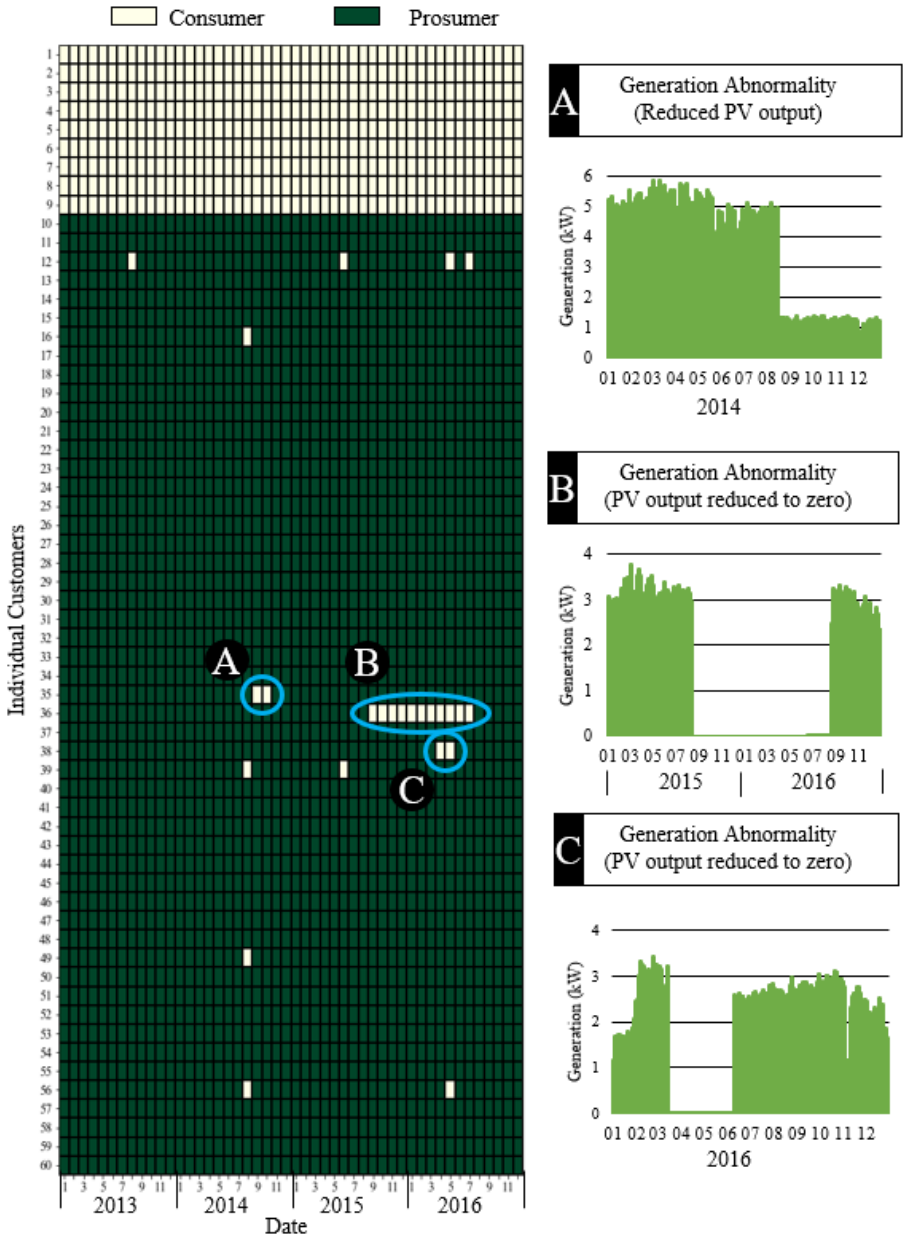


Figure 4.7: Example use of solar prosumer identification to detect anomalous behaviour [P1]

Up-to-date information regarding the number of solar PV customers on a given feeder will grow in significance as more utilities begin to pursue use of solar PV as part of efforts to enhance wildfire resilience in the form of microgrids [152]. As discussed in Chapters 2 and 3, remote microgrids are becoming a valuable tool to enhance the resilience of

the distribution system to extreme natural hazards such as wildfires. The presence of microgrids in an area enables grid operators to dynamically reconfigure the distribution network into self-sufficient islands to hasten restoration of power following a contingency, or prevent power disruption [152]. However, such approaches typically rely on the presence and location of distributed generation including solar PV [159]. In these conditions, balancing of load and generation on the microgrid is critical, and it becomes even more essential for grid operators and planners to know the location of all DGs present in a candidate distribution network. The solar prosumer identification approach presented in this chapter can enable accurate identification of solar prosumers in a natural hazard risk area, providing the operator with a projection of the number of prosumers and consumers on each feeder. Figure 4.8 shows how clusters of customers might be identified for resilience planning efforts. The location of clusters of customers can then be used to inform cost estimates of new microgrids, or the performance and scheduling existing microgrid resilience measures.

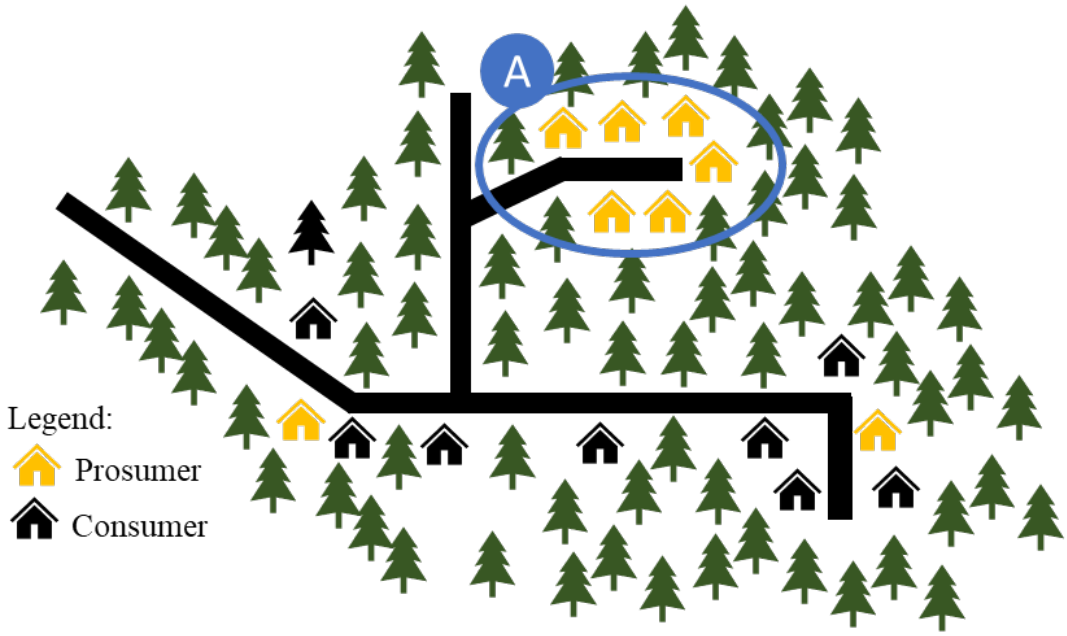


Figure 4.8: Use of solar prosumer identification to inform resilience planning. Location A is an example of a clustered group of prosumers in a remote area.

4.5 Summary

Solar prosumer identification can provide utilities with a means to enhance awareness of the location of PV systems on the distribution network, whether authorised or unauthorised. This chapter proposes an effective approach to identify solar prosumers; Agglomerative clustering and PAA_A for dimensionality reduction aggregating data for each customer based on their minimum demand over the year. This approach provides accuracy of nearly 100% when tested using real smart meter data from hundreds of customers in the United States. This equips utilities with an effective approach to use net smart meter data to directly identify solar prosumers, using less data than existing methods in the literature while also providing high identification accuracy. The approach also reflects the classification accuracy in the form of SPIDCs to provide distribution grid operators with understanding of how much historical data is needed to place confidence in the classifications. Finally practical examples demonstrate how this approach can provide the ability to flag performance issues with existing panels and update awareness of customers with PV installations for use in microgrid feasibility analysis to support resiliency during a wildfire or other natural hazard.

Chapter 5

Predicting Reductions in Solar PV Generation During Wildfires

The maximum amount of power able to be produced by a solar PV plant is a function of the plant's physical characteristics (number of PV cells, module type, and means of mounting) along with weather variables (solar irradiance, temperature, and wind speed) [160]. Wildfires can disrupt the transmission system, crippling import capability and causing risk of insufficient generation supply [3, 130]. Failure to consider reductions in solar PV capacity due to natural hazards can result in overestimating the available generation capacity and increase the risk of insufficient supply of electricity [99]. During a wildfire, smoke can obstruct the solar radiation resulting in a reduction of the output of a solar facility.

This chapter presents a methodology to predict the spatiotemporal derate to solar PV capacity from wildfire smoke, enhancing generation forecasts. This enables operators to translate forecasts of aerosol optical depth (AOD) [161] into anticipated reductions in solar capacity. Prior work to quantify the derate from wildfire smoke to solar PV performance has been limited to the effect of single events [105, 162] and use of point based PM2.5 measurements to inform performance [106]. Knowledge of how smoke might affect solar PV installations over a large geographic area can equip balancing authorities with

information to make better scheduling and dispatch decisions. Providing grid operators with a means to anticipate these effects can potentially reduce forecasting errors and the associated risk of imbalances in generation supply.

The chapter is divided into four sections. Section 5.1 proposes the method to forecast the derate to solar PV facilities. Section 5.2 demonstrates the accuracy using real PV operating data from the 2020 Southern California wildfire season. Section 5.3 provides means to practically visualise the resulting insights to inform decision making. Finally, Section 5.4 summarises the findings. As a contribution to the state-of-the-art, the publication [P3] resulted from the research described in this chapter.

5.1 Wildfire Smoke PV Derate Methodology

The methodology contains two main stages: 1) Data Preparation; and 2) Model fitting and evaluation. Figure 5.1 depicts these stages. The result of the methodology is a model that can be used to predict the derate to solar PV generation given the amount of AOD present.

5.1.1 Data Preparation

The initial step is to select sites to analyse and collect historical PV and AOD data for use in model training and evaluation. Following data collection, the methodology first controls for the influence of seasonal variation in solar irradiance, along with variation in temperature, wind speed, and PV generation nameplate capacity through a normalisation process . This produces a more robust model enhancing the model's generalisation to other facilities within a service territory. Individual PV plants often differ in the installed nameplate capacity. Min-Max normalisation converts the generation of each plant $g = g_1, g_2, \ldots, g_n$ to a range of [0, 1] via (5.1) [84]. This allows equivalent comparison of the

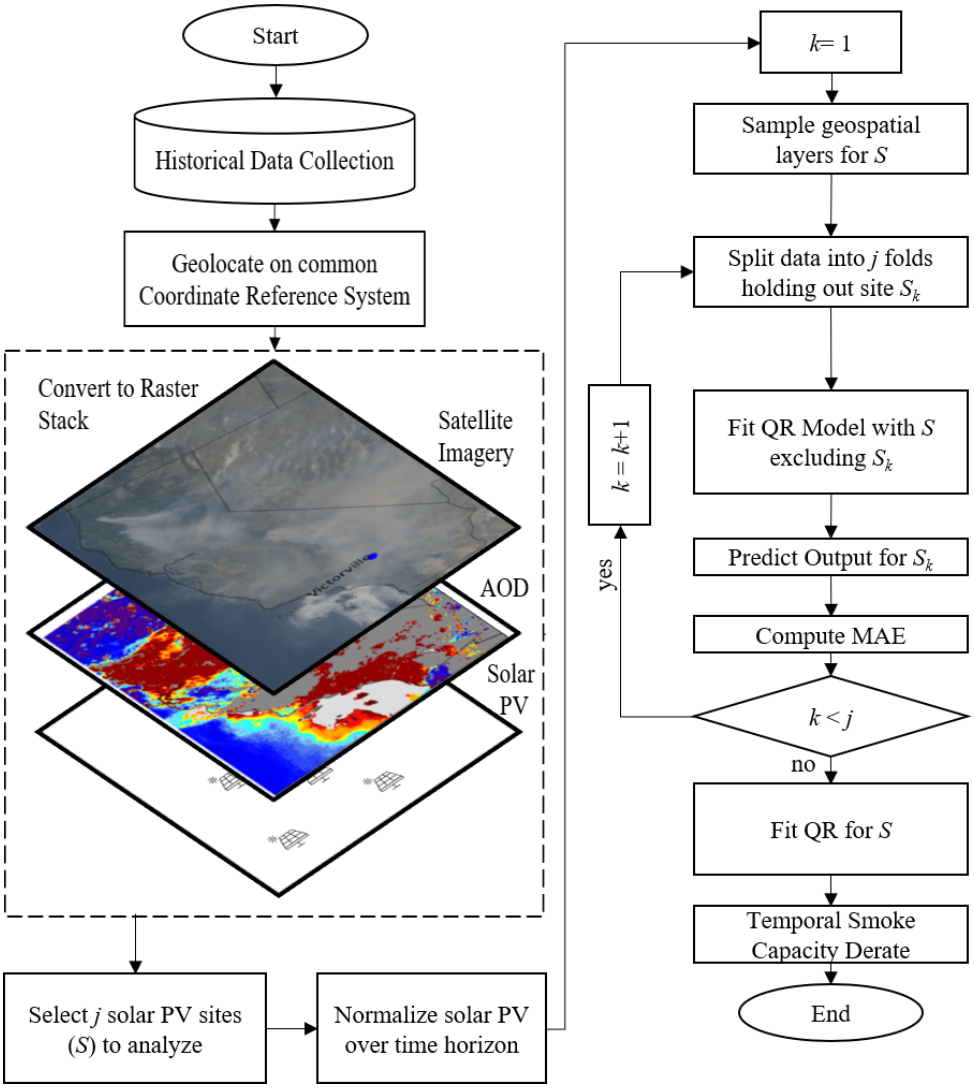


Figure 5.1: Brief overview of methodology for geospatial wildfire PV capacity model. Satellite imagery from NOAA GOES-17 satellite [163]. Figure reproduced from [P3].

effect of smoke across multiple plants while maintaining the original profile shape:

$$g'_t = \frac{g_t - g_{min}}{g_{max} - g_{min}} \tag{5.1}$$

where g'_t and g_t are the normalised and actual recorded solar power output at time t; g_{min} and g_{max} are the minimum and maximum power output within the fire study period n. When grid operators are performing assessment of the effect of smoke on a single plant's operating history, this step can be omitted to retain the nameplate value. Ambient temperature and wind speed affect performance of the solar PV capacity by affecting the

PV cell temperature. As temperature of the PV cell (τ_c) increases, the PV efficiency is reduced by a factor of 0.5% per °C for every degree above 25°C [164, 165]. Increases in wind speed (v) have the opposite effect, lowering τ_c, thereby increasing efficiency. Therefore, it is important to normalise the PV capacity for windspeed and temperature. Equations (5.2-5.3) [160] are used to determine τ_{c_i} and then (5.4) is used to calculate the weather adjustment factor (wn). Once wn is calculated, the temperature adjusted series can be calculated as in (5.5) and then normalised to retain the range of [0, 1] in (5.6) producing the weather normalised PV output (g''_t).

$$\tau_{m_t} = E_t[e^{\gamma_1+\gamma_2 v_t}] + \tau_{a_t} \tag{5.2}$$

$$\tau_{c_t} = \tau_{m_t} + \frac{E_t}{E_o}\Delta\tau_t \tag{5.3}$$

$$wn_t = \begin{cases} \frac{0.5}{100}(\tau_{c_t} - 25), & \tau_{c_t} > 25 \\ 0, & \tau_{c_t} \leq 25 \end{cases} \tag{5.4}$$

$$\delta_t = \frac{g'_t}{1 - wn_t} \tag{5.5}$$

$$g''_t = \frac{\delta_t}{\delta_{max}} \tag{5.6}$$

where τ_m is the back-surface module temperature (°C); γ_1 and γ_2 are empirically derived constants; τ_a is the ambient temperature (°C); E_o is the reference solar irradiance (1000 W/m^2); and $\Delta\tau$ is the temperature difference between the cell and the module back surface at an irradiance level of 1000 W/m^2; δ is the weather adjusted generation series; and δ_{max} reflects the maximum value of δ across observations at the same hour. If the solar plants are all located in regions with similar climates and there is limited variation in weather across the study period, then utilities may choose to eliminate this step as the

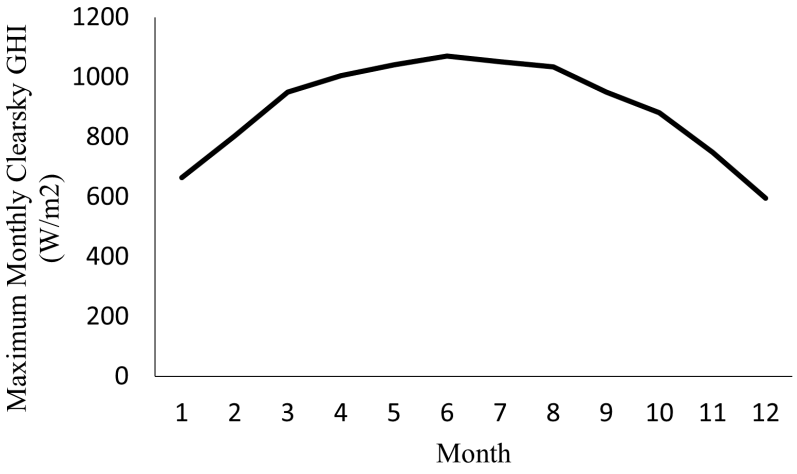


Figure 5.2: Example of monthly variance in CGHI. Data from [166].

magnitude of adjustment is likely to be small and may not warrant the collection of the additional weather variables.

Change in the relative position of the sun across the year causes seasonal variation in the Clearsky Global Horizontal Irradiance (CGHI). Figure 5.2 shows this variation for Ontario, California across the year 2019. As solar PV generation is dependent upon irradiance, the historical variation in CGHI can provide an effective means to control for differences in seasonal peak PV production, enabling comparison of the effects of wildfire smoke across seasons. A timeseries of typical CGHI (C) for a specific location is produced by taking the average hourly CGHI across the historical years available. The seasonal adjusted series is then calculated as in (5.8) and subsequently used to create the final normalised output (g_t''').

$$s_t = \frac{C_{max}}{C_t}\delta_t \tag{5.7}$$

$$g_t''' = \frac{s_t}{s_{max}} \tag{5.8}$$

where s_t is the seasonal adjustment factor at time t; C_t is the clearsky GHI for a specific hour and day of the year; and C_{max} is the maximum clearsky GHI at that hour of the

day across all days in the study period. If the time period of analysis is short, there may be limited relative variation in CGHI across the study period and the effect of seasonal normalisation may not be significant. In such cases, utilities may select to use either g'_t or g''_t instead to minimize data collection requirements.

5.1.2 Model Fitting

Once normalised the next stage of the methodology is model fitting and evaluation. A quantile regression (QR) model is proposed to characterise the relationship between AOD and solar PV generation output. This model is selected due to the robustness to outliers [167], the ability to predict the derate at different quantiles and the amount of data available for the case study. The model for linear quantile regression is given as [168]:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}(q) + \boldsymbol{\epsilon} \tag{5.9}$$

where $\boldsymbol{y}$ is the vector of responses; $\boldsymbol{X}$ is the matrix of regressors; and $\boldsymbol{\beta}(q)$ is the vector of unknown parameters for quantile q [168].

The parameters are estimated by minimizing the loss function for a particular q as in [168]:

$$\min_{\boldsymbol{\beta(q)}} \sum_{i=1}^{n} \rho_q(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}) \tag{5.10}$$

where ρ_q is the pinball loss function for a quantile q, where $\rho_q(z) = qz$ if $z \geq 0$ and $\rho_q(z) = (q-1)z$ if $z < 0$ [169]. To compare the result with linear regression or other deterministic models, the median quantile can be selected.

Other models can also be tested and easily integrated into the methodology shown in Figure 5.1 as more data becomes available. To evaluate the model performance, k-fold cross validation is used. As the number of concurrent fires differs over the course of the wildfire period, the dataset is unbalanced. Therefore, to avoid some folds being dominated by days without smoke, a stratified approach is used to ensure each fold contains a similar number of smoky days. To achieve this, the data is split into j folds where each fold

contains data for a single PV facility. Then the model is iteratively trained on $j-1$ folds and tested on the remaining fold. mean absolute error (MAE) is first computed for each fold and then MAE_{avg} computed across all folds to evaluate the error. Figure 5.3 provides an overview of this process.

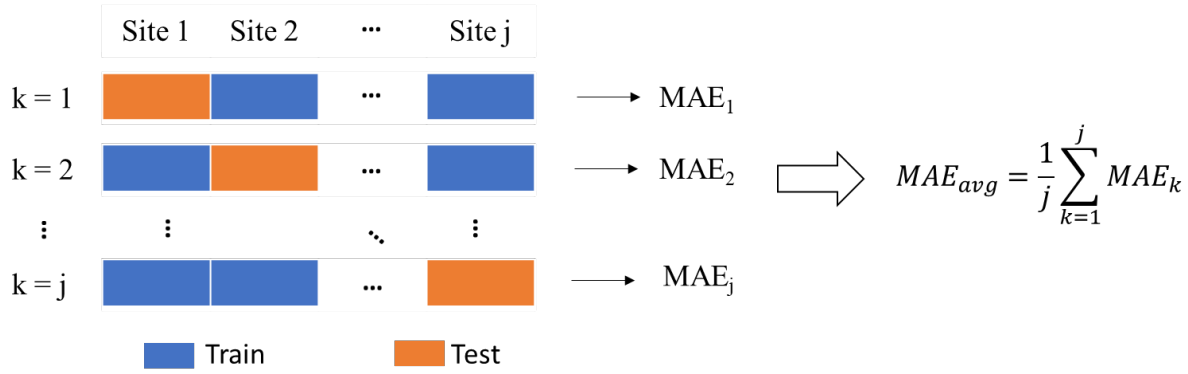


Figure 5.3: Evaluation of model performance with k-fold cross validation

5.2 Case Study for Southern California Wildfires

In 2020, California experienced an extreme wildfire season with wildfires burning over 4.2 million acres ($\approx$ 17,000 km^2) [170]. These fires affected PV production throughout the state [19]. Two scenarios with real solar PV generation and AOD data from this fire season demonstrate the efficacy of the methodology presented in Section 5.1:

- **Scenario A** covers the two-week period from September 1st to September 14th including the start of three significant fires: the Creek Fire on September 4th [171], the El Dorado Fire on September 5th [172] and the Bobcat Fire on September 6th [173].

- **Scenario B** covers an extended period from August 1st to September 28th to evaluate the impact of a longer horizon on model performance.

The accuracy of the proposed QR model is evaluated against linear regression (LR) and piece-wise linear regression (PLR) models to validate the robustness. The basic

formulation for a LR model is given in Equation (2.21). PLR describes a one-dimensional dependent variable via several discrete linear segments [174]. The PLR is formulated such that the function is continuous over the domain [174] and is as follows:

$$y(x_p) = \begin{cases} \beta_1 + \beta_2(x_p - b_1) & b_1 \leq x_p \leq b_2 \\ \beta_1 + \beta_2(x_p - b_1) + \beta_3(x_p - b_2) & b_2 < x_p \leq b_3 \\ \vdots & \vdots \\ \begin{bmatrix} \beta_1 + \beta_2(x_p - b_1) + \beta_3(x_p - b_2) \\ + \cdots + \beta_{c_b}(x_p - b_{c_b - 1}) \end{bmatrix} & b_{c-1} < x_p \leq b_{c_b} \end{cases} \tag{5.11}$$

where c_b is the number of knots in the spline; $b_0, b_1, \ldots, b_{c_b}$ are the location of the knots; and $\boldsymbol{\beta}$ are the slope coefficients. The resulting equations are solved for the unknown $\boldsymbol{\beta}$ that reduce the sum-of-square of the residuals [174]. In the future, with additional explanatory variables or extensive amounts of historical operating data, machine learning based models may also be beneficial.

5.2.1 Collection of Historical Data

Historical generation data for ten individual PV generation facilities across four distinct areas was collected via supervisory control and data acquisition (SCADA) and provided by a regional utility alongside the generalised location of each of the four geographic areas. Solar production between 12 pm and 1 pm each day were extracted from the SCADA data to assess the impact of wildfire smoke on the daily PV facility capacity. Figure 5.4 shows the location of the geographic areas alongside perimeters for large fires (>30,000 acres) obtained from [175]. Two GOES-17 satellite data products are used: 1) The ABI-L2-AODC product [109] for spatiotemporal measurements of AOD at a wavelength of 550 nm across the study region on a 2km grid [109]; 2) The Cloud and Moisture Imagery Product [109] for production of Red-Green-Blue (RGB) images. GOES-17 satellite data products are publicly available from National Oceanic and Atmospheric Administration (NOAA)

through Amazon Web Services at the following link: `https://registry.opendata.aws/noaa-goes/` [163]. The resulting SCADA PV generation measurements are combined with the AOD measurements resulting in 114 datapoints for Scenario A and 508 for Scenario B after accounting for periods with missing AOD measurements and abnormal or erroneous SCADA measurements.

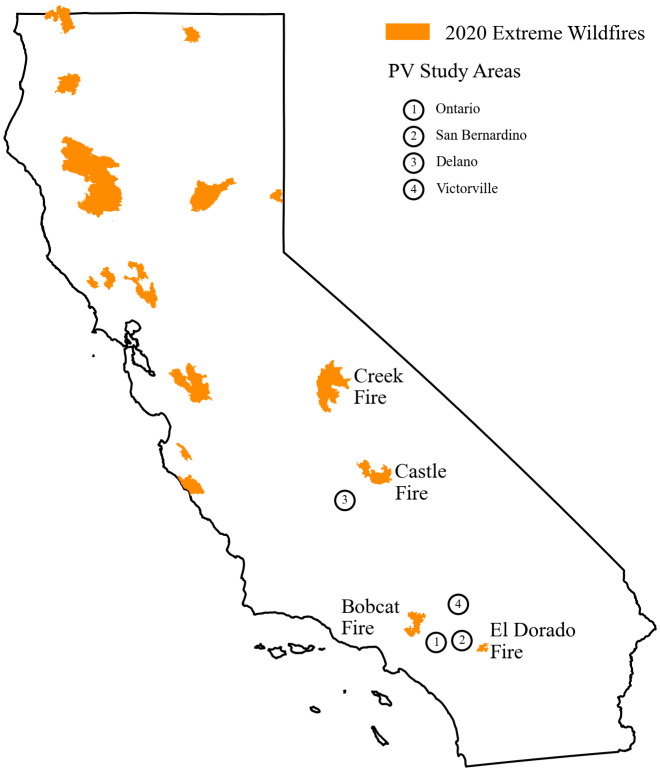


Figure 5.4: Wildfires >30,000 acres during California's 2020 wildfire season. Data for fire perimeters from [175]. Figure reproduced from [P3].

The necessary data for the three stage normalisation process was collected from the following three sources: 1) MERRA-2 renalysis for historical weather data [176]. 2) Representative physical parameters of -3.47, -.0594 and 3 from [160] for γ_1,γ_2 and ΔT respectively (used due to unavailability of actual site specific information). 3) National Solar Radiation Database (NSRDB) [166] for a fifteen year period of clear sky global horizontal irradiance (CGHI) values from 2005-2019 to calculate the mean for each site.

`Python` is used for the analysis, supported by the following packages: `numpy` [136] and `pandas` [137] for data analysis; `geopandas` [138] and `cartopy` [141] for geospatial

analysis; cartographic boundaries from the US Census Bureau [147]; code from [177] to support satellite data processing; `matplotlib` [142] for visualisation; `statsmodels` [178] for specification and fitting of the LR and QR models; and `pwlf` [174] for specification and fitting of the PLR model. The PLR model requires specification of the number of line segments and this was selected to be two for this case, in attempt to limit overfitting while still allowing variation. The `pwlf` package relies on `scipy` [144] to calculate the optimal knot location via differential evolution [174].

5.2.2 Results and Analysis

Comparison of GOES-17 satellite images and solar PV generation during the wildfire period qualitatively validates the relationship between wildfire smoke and solar PV performance and shown in Figure 5.5 . Prior to modelling, a Spearman rank-order correlation test is conducted to empirically validate the strength and significance of relationship between AOD and PV output. The results indicate that AOD and PV are highly correlated for both Scenario A and B with statistically significant ($p<0.001$) correlations of -0.76, and -0.45 respectively. The negative sign for both coefficients indicates that peak solar PV capacity declines as AOD increases. The change in the magnitude between scenarios may result from the higher proportion of smoke free days in Scenario B. As the relationship being modelled is that between smoke and PV production, as more days without smoke are included, this increases the exposure to other confounding factors such as maintenance or curtailment which reduce the output but may not be directly related to the wildfire.

The methodology given in Section 5.1 was then applied using the proposed QR model along with PLR and LR models for comparison. Comparison of the MAE demonstrates that QR is the most effective model across both scenarios. Figure 5.6 shows each of the models fit for the entire duration of each scenario. The in-sample MAE for each model as follows: Scenario A) LR: 5.7%; QR: 5.4%; and PLR: 5.4%; Scenario B) LR: 6.1%; QR:5.8%; and PLR: 6.1%. The relative positions along the y-axis of some points change across scenarios due to the differing seasonal normalisation period. The mean reduction

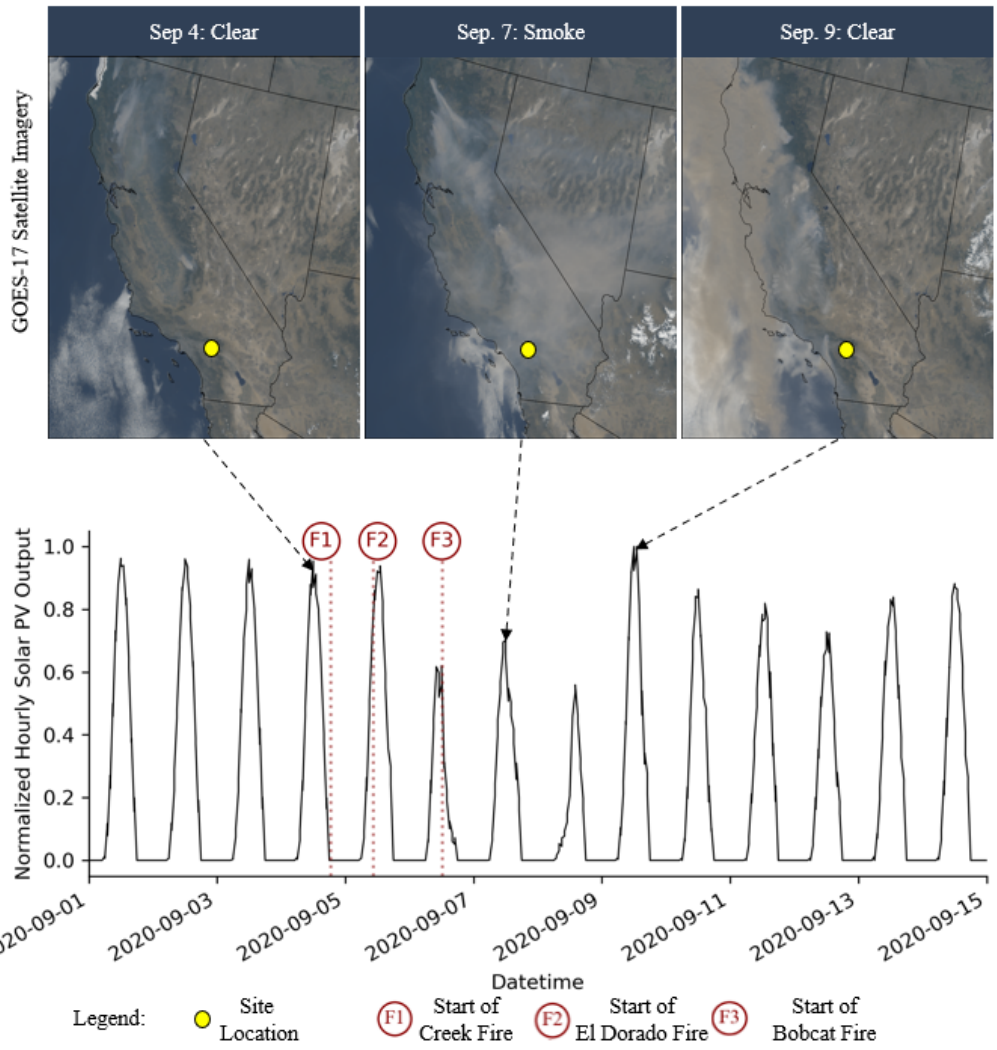


Figure 5.5: Timeline with PV generation from one region over Scenario A. Satellite imagery from [109]. Fire start dates from [171, 172, 173]. Figure reproduced from [P3].

in CGHI over the study period increased nearly three-fold from Scenario A to Scenario B increasing from 5% to 17%. This demonstrates that seasonal variation in CGHI can play a significant role in determining the overall PV capacity, and normalisation to account for this variance is needed when assessing PV performance over longer time periods.

Comparison across scenarios indicated volatility in the PLR model, as well as the importance of a sufficient training period. The PLR model breakpoints are not stable across scenarios, or even within scenarios, as multiple runs of PLR model fitting over Scenario B resulted in different breakpoints. Furthermore, the PLR model predicts that PV performance will start to increase again at AOD levels above 4; a clear example of

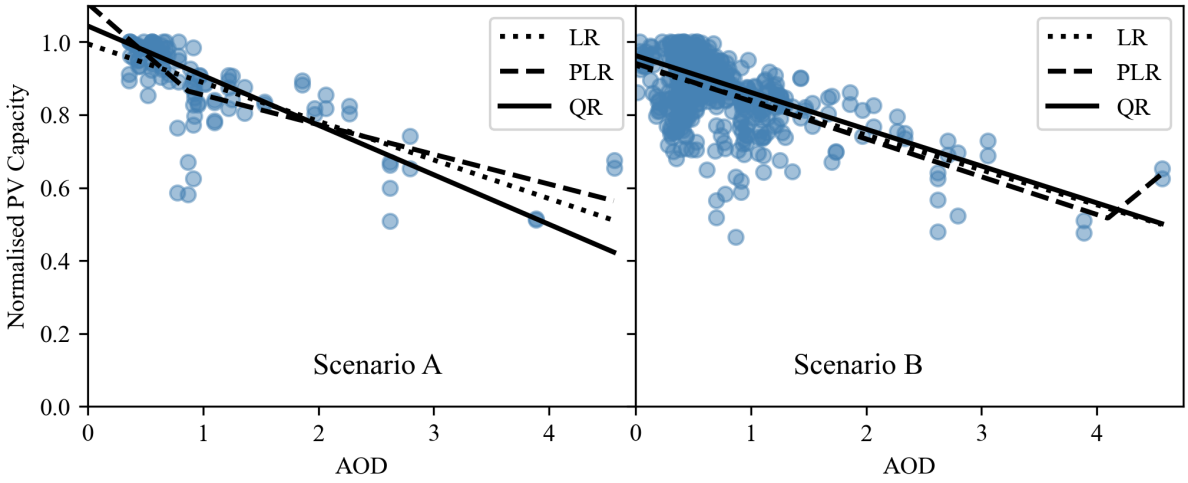
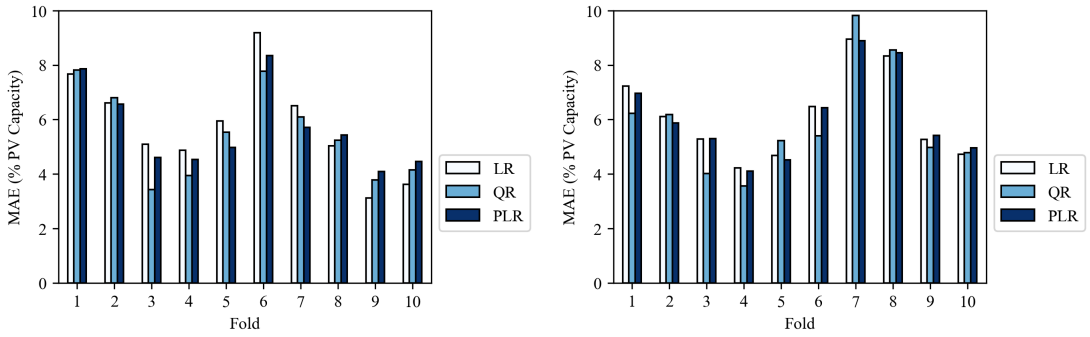


Figure 5.6: Model results on training data for each Scenario

overfitting. In practice, when the AOD is at a value of 0 indicating a perfectly clear sky, PV capacity should be at 100% with no derate from wildfire smoke. With only two weeks of training data, both the PLR and QR model provide unrealistic boundary results with the QR model predicting 104% and the PLR model predicting 111%. The inclusion of additional training data at low AOD values in Scenario B improves the ability of the models to handle the boundary conditions, showing the significance of testing models over a longer time period.

Ten iterations of model training are performed for each model based on the k-fold cross validation process described in 5.3, with each iteration using data from nine sites for training and using the remaining site for testing. The performance for each of the folds and models can be seen in Figure 5.7. For the cross-validated results, the QR model exhibited a MAE of 5.45% across the folds in Scenario A, and the highest accuracy across both Scenarios. One potential explanation for some of the variation in performance across sites is due to sparsity of the overall AOD data and the generalisation of site locations. As AOD coverage and resolution improves, the accuracy should improve further. Despite some variation across folds, the consistency in overall model accuracy across scenarios highlights the model robustness. The LR model also performs well, and outperforms the QR model for some folds. Therefore, the LR model may provide value in cases where



(a) Scenario A [P3] (b) Scenario B

Figure 5.7: Cross validation results across both scenarios indicating model accuracy.

outliers are less of a concern and there is no desire to anticipate the range of potential performance via quantiles.

The derate factor predicted by each model is calculated by subtracting the anticipated output at varying aerosol optical depths from the maximum output (100%). The derate factor at quantiles from 50-99% are shown in Figure 5.8. The quantiles produced by the proposed method enable individual operators to select a quantile that aligns with their risk modelling.

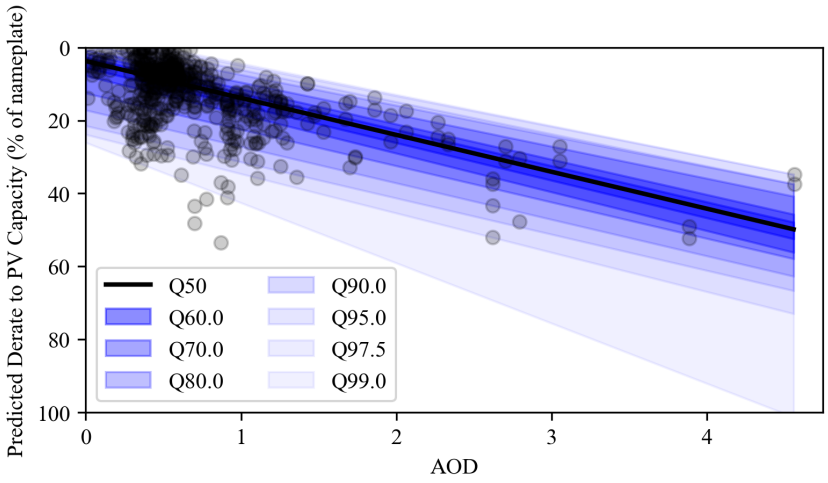


Figure 5.8: QR based solar PV capacity derate model trained on 2020 California Wildfires

Findings in [179] looking at 2020 California wildfires indicated that "hour-ahead forecasts utilized by CAISO did not include the effects of smoke and therefore overestimated the expected power production by ∼10%–50%" demonstrating the significance of antici-

pating PV production impacts from wildfire smoke. Global satellite coverage enables the methodology to be flexibly applied to ISOs around the world. The methodology provides grid operators with a means to translate historical AOD and solar PV performance data into a model to predict the spatiotemporal derate to PV capacity. As the spatial resolution and accuracy of AOD forecast products improve, grid operators will be able to place higher confidence in the anticipated reductions in solar PV capacity. Overall, the ability to quantify the projected reduction in PV output can improve generation forecasts and enhance security of energy supply to wildfire smoke related shortfalls in reserve.

5.3 Practical Applications via Geospatial Mapping

The use of spatiotemporal data in the proposed methodology also enables production of geospatial maps for practical utility decision making. One example for the operating horizon is the creation of geospatial solar PV derate maps for use by Independent System Operators (ISOs). Maps can be generated by training a model for an ISO's area using the approach in Section 5.1 alongside spatial AOD data [163]. Figure 5.9 provides an example map for three days during the 2020 California Wildfire Season using data from GOES-17 [163]. The sections in grey indicate no AOD values are available for that day. While Figure 5.9 is generated from historical data, ISOs can use AOD forecasts such as the five day ahead forecasts provided by [161] to predict the widespread effects of smoke on PV throughout their entire service area. With the high spatial resolution of AOD forecasts, an operator can also focus on a critical geographic regions where a significant number of solar PV facilities are sited.

In the planning horizon, geospatial mapping can be used to inform the siting of new weather stations or solar PV generation facilities. When siting a new PV generation facility, one consideration is the anticipated output under "typical" weather conditions such as the typical meteorological year data provided by the NSRDB [166]. Given the increasing effect of synchronous fire risk in some regions of the world such as the West-

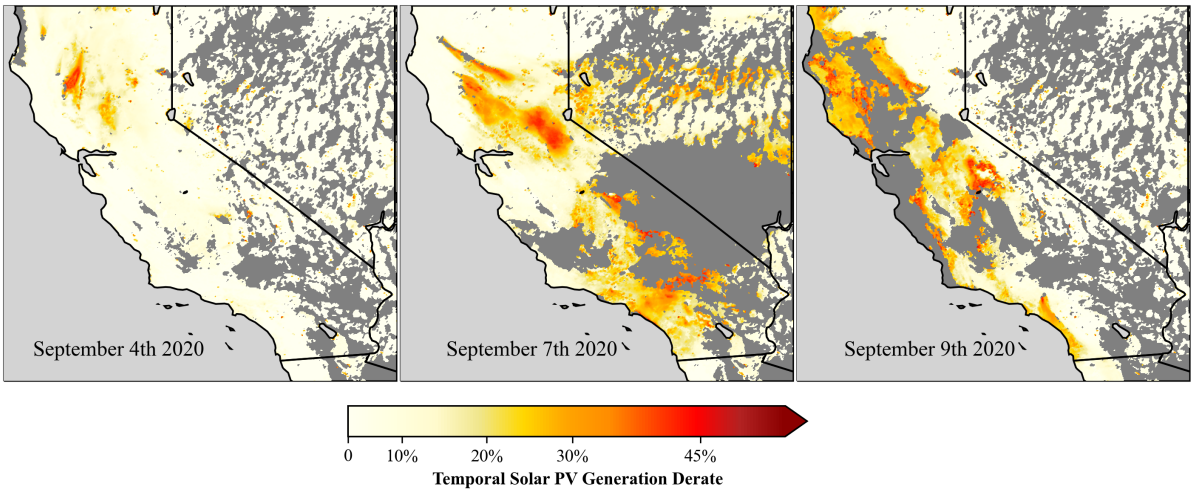
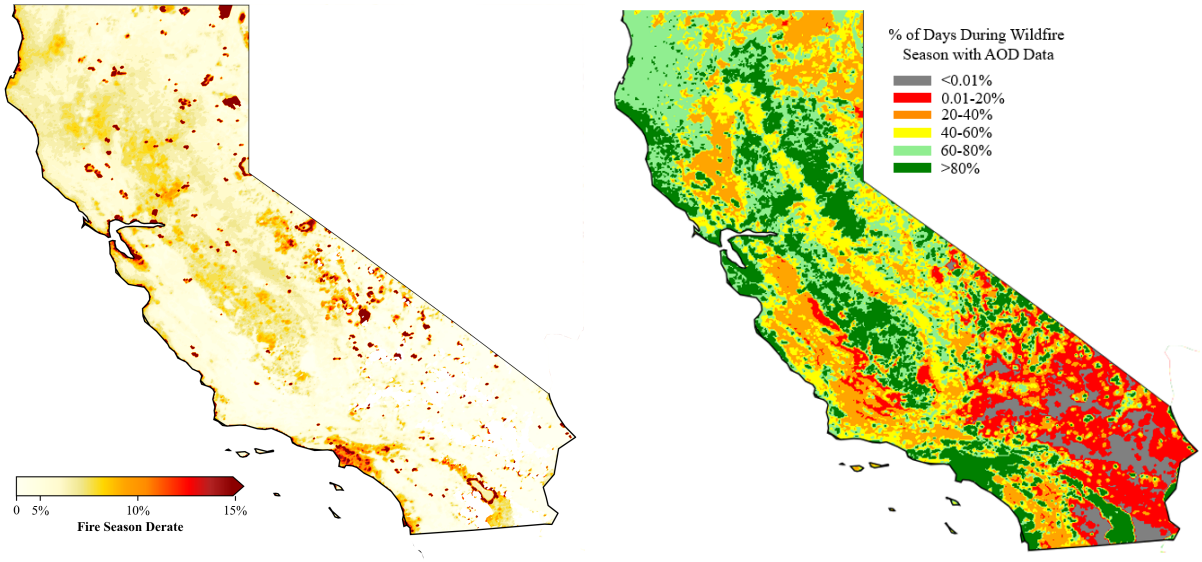


Figure 5.9: Example geospatial smoke derate map for operational use over the period from 12 pm to 1 pm on three separate days with data from GOES-17 [163].[P3]

ern United States [18] economic feasibility studies performed by generation owners may need to incorporate anticipated reductions due to wildfire smoke. A wildfire derate model trained for a specific area of interest, could be used to provide additional information for feasibility studies in high fire risk areas. A California example of such information is created by taking mean daily AOD from 12pm to 1pm over the period from April 30th to September 28th, 2020 (obtained from GOES-17 [163]) and is shown in Figure 5.10a. Generation planners could customise the timeframes and variables included in such a map and overlay additional spatial variables to meet their specific siting use case. Furthermore, a generation planner may want to exclude regions missing data for a specific percentage of time over the season (as shown in Figure 5.10b). While a single period is used for Figure 5.10a, historical duration selected could be further customised based on specific planning needs. For example, an extreme fire season could be used to provide a conservative estimate or multiple years of fire history could be used to reflect what a "typical" annual derate for a region might be.

One limitation to the use of satellite data is that GOES AOD data may not be produced under some conditions (e.g. snow, clouds, or specific reflectance values) [109] and means that data in rural or desert areas may be more sparse than data for areas with high contrast (like cities). Increases in the coverage of historical AOD data as satellite



(a) Projected seasonal capacity derate [P3] (b) AOD data coverage from GOES-17 [163]

Figure 5.10: Information over the period from April 30th to September 28, 2020 with data from GOES-17 [163]

measurements improve will serve to address the limitation of incomplete coverage of historical AOD data. Figure 5.10b shows the sparsity of AOD measurements. One method of addressing this data limitation in the future would be to install ground based stations or use a technique to fill the AOD gaps via interpolation. In this case study, AOD interpolation was not applied as any additional approximation error from the interpolation technique could mask the evaluation of model performance.

5.4 Summary

As wildfire smoke can saturate wide areas during a wildfire, anticipation of the impact to the production of solar PV generation is important to maintain an adequate supply of energy. This chapter presents a new methodology that enables grid operators to translate wildfire aerosol forecasts into PV capacity reductions, that could potentially enhance grid operator scheduling decisions and short-term forecast accuracy. The QR model performance demonstrates the robustness of the methodology validated using real solar PV operating history and data from 2020 California wildfires. Practical geospatial visuali-

sation via the proposed model, enables operators to quickly assess capacity impacts of wildfire smoke to solar facilities throughout their territory. In the planning horizon, maps of anticipated smoke exposure can inform the siting of new solar PV facilities. As forecasts of AOD are produced at higher resolution and coverage, the proposed methodology will further enhance the ability to anticipate the attenuation of PV generation during extreme wildfires.

Chapter 6

Integration of EV Evacuation in Wildfire Resilience Assessment

Wildfires have resulted in the evacuation of large groups of people [20] and more homeowners may be exposed in the future due to growth in the wildland-urban interface [118]. Unanticipated wide-scale EV charging during an extreme event could lead to power shortages and cascading blackouts [119]. To provide insight into this problem, this chapter proposes a methodology to integrate electric vehicle evacuation charging demand in transmission system wildfire resilience assessments. The proposed methodology is provided in Section 6.1. Data for the case study are given in Section 6.2 with the results and analysis presented in Section 6.3. Potential means of enhancing resilience are presented in Section 6.4. Finally, conclusions are given in Section 6.5. As a contribution to the state of the art, the research described in this chapter resulted in the publication of [P2].

6.1 Wildfire Resilience with EV Evacuation

During an emergency, the typical behaviour of a vehicle owner can be substantially altered due to a forced or voluntary evacuation. This section presents the methodology to anticipate changes in charging demand of electric vehicles during a wildfire evacuation and identify critical transmission infrastructure affected by altered loading patterns. The

proposed methodology consists of three major components: 1) novel EV wildfire evacuation model, 2) time step power flow simulation, and 3) critical element identification. Figure 6.1 shows an overview of the entire methodology.

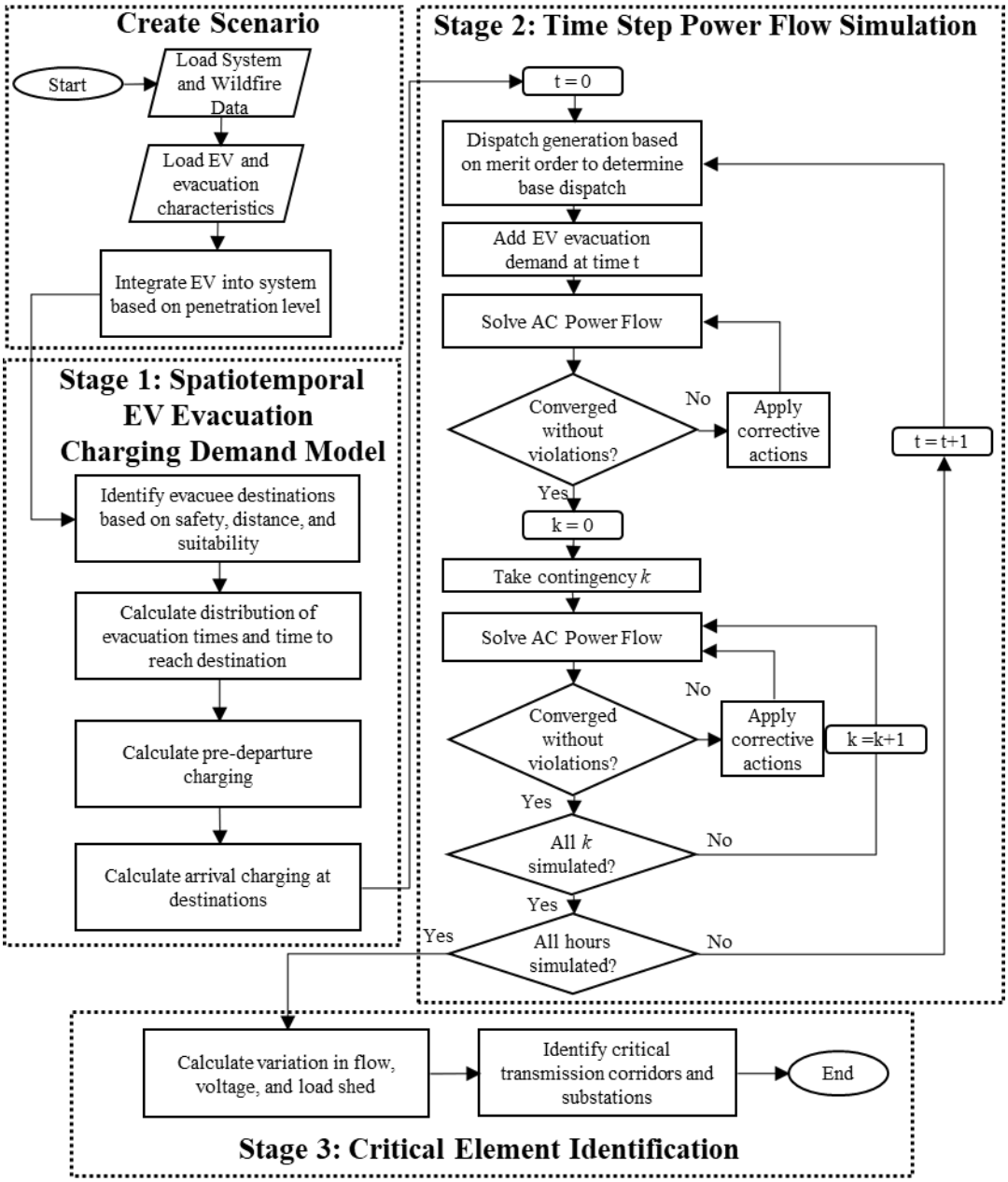


Figure 6.1: Methodology to integrate EV evacuation charging demand in transmission system wildfire resilience assessments.

6.1.1 Spatiotemporal EV Evacuation Charging Demand Model

The model reflects changes to charging demand during an evacuation as customers charge prior to their journey, once they arrive at their destination, as well as the charging during the remaining extent of the evacuation period (given in Equations (6.1)-(6.10) below). In order to identify these changes, the destination and departure time of each vehicle in the region must be determined. Then the spatiotemporal charging demand can be calculated to provide grid operators with a more holistic perspective of the potential grid impacts. An example of the timeline for an event can be seen in Figure 6.2.

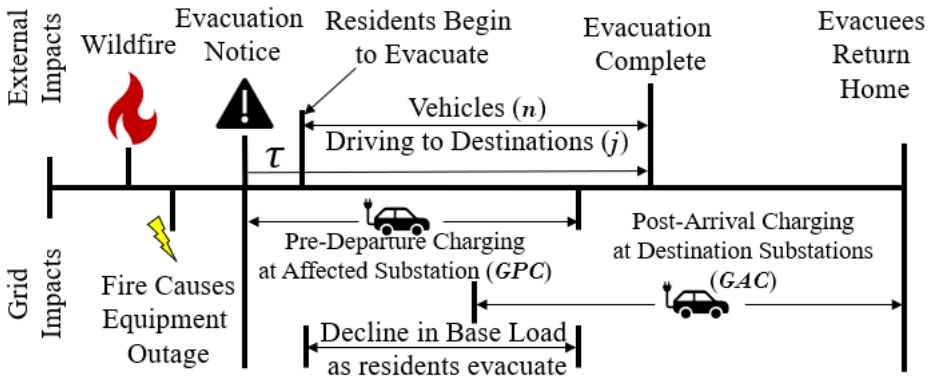


Figure 6.2: Example wildfire event timeline. Note that overlap between pre-departure charging and post-arrival charging will depend on the parameters for a given event.

Choice of Destination and Travel Time

The destination of each vehicle is identified based on three selection criteria: safety, proximity, and suitability. Some potential destinations may be inaccessible due to blocked roads or threat from the wildfire, or may be deemed unsafe due to other public health and emergency guidance. The safety criterion is used as a binary flag by the model to eliminate destinations considered to be unsafe by evacuees or local authorities.

To account for proximity, the model calculates a weight for each destination based on the relative distance. The distance based weight (EVD) of each destination j can be calculated as:

$$EVD_j = \begin{cases} \dfrac{\dfrac{1}{d_j}}{\sum_{i=1}^{B} \dfrac{1}{d_i}} & \text{if } d_j \leq R \\ 0 & \text{if } d_j > R \end{cases} \tag{6.1}$$

where R is the evacuation radius, d is the distance; and B is the total safe destinations within the radius. The radius is used to limit the evacuation to within a fixed region as wildfire evacuees may prefer proximal locations [20].

Three zones are used to group destinations of similar proximity: Z_1: $(d \leq 0.5R)$, Z_2: $(0.5R < d \leq 0.75R)$, and Z_3: $(0.75R < d \leq R)$. Suitability is then used to further refine the weighting factor among destinations within each zone to reflect the availability of food, lodging, or charging stations at each destination which may impact an evacuees choice of destination [20]. The model uses load density as a proxy for suitability, to reflect that more amenities may be found in larger load centres. The suitability adjusted allocation (EVS) at bus j within each zone is calculated as:

$$EVS_j = \sum_{m=1}^{\alpha} EVD_m \frac{L_j}{\sum_{i=1}^{\alpha} L_i} \tag{6.2}$$

where α is the number of buses within a particular zone; and L is the load in MW of a given bus.

The total number of vehicles evacuating (N) to each destination bus (j) is then given as:

$$N_j = \rho \Theta EVS_j \tag{6.3}$$

where ρ is the population of potential evacuating EV; and Θ represents the fraction of evacuating households.

After calculating the number of vehicles evacuating to each destination, the overall departure time is modelled via a Rayleigh distribution, as prior studies have identified that Rayleigh distributions are reflective of the variation in times at which residents evacuate

[180]. The total number of vehicles which have evacuated can be calculated as:

$$N_j(\tau) = (1 - e^{-0.5\left(\frac{\tau}{\beta}\right)^2})N_j \tag{6.4}$$

where τ is the time after the evacuation notice; and β is the mode of the distribution function. After this, the time the EVs arrive to their destination and resume charging is determined by the associated travel time to each destination. Once $N_j(\tau)$ has been calculated as in (6.4) the additional time to reach each destination is calculated as d_j/s where s is the speed of travel.

Charging During the Evacuation Period

Once the destination and departure time of each vehicle is known, the corresponding impact to grid demand during and after the evacuation is calculated. The duration needed to fully charge each vehicle (n) at a specific hour in time t prior to departure, can be reflected as:

$$DP_n(t) = \frac{(1 - SOC_n(t))B_n}{\eta C_n} \tag{6.5}$$

where SOC is the state of charge in per unit; B is the battery size in kWh; η reflects the overall charging efficiency of the charger and battery in per unit; and C is the rate of charge in kW. Although there is some variation in the charging rate as the battery nears full SOC, the charging rate remains fairly linear until 95% SOC, and therefore C_n is modelled as time-invariant for this analysis [181, 182].

However, not all residents may wait until a vehicle is fully charged. Therefore the power required for each vehicle at each hour $f_n(t)$ can be calculated as in (6.6) for the population of electric vehicles which have yet to evacuate (*pop*) at a specific point in time. Then, the total power seen by the grid (GPC) can be calculated as given in (6.7).

$$f_n(t) = \begin{cases} C_n & \text{if } DP_n(t) \geq 1 \\ C_n DP_n(t) & \text{if } DP_n(t) < 1 \end{cases} \tag{6.6}$$

$$GPC(t) = \sum_{n=1}^{pop} f_n(t) \tag{6.7}$$

Once vehicles arrive at their final destination, the duration of charging needed to recover the energy from their evacuation journey J can be calculated as in (6.8) with the corresponding power needed to charge each vehicle at each hour given in (6.9).

$$J_n = \frac{q_n W_n}{\eta C_n} \tag{6.8}$$

$$h_n(t) = \begin{cases} C_n & \text{if } 1 \leq J_n - \delta_n \\ (J_n - \delta_n)C_n & \text{if } 0 < J_n - \delta_n < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.9}$$

where q is the distance travelled using battery energy (may be less than the total distance for hybrids); W is the vehicle's energy consumption per unit of distance; and δ is time since arrival.

Finally, the total power required from the grid (GAC) due to the initial post-arrival charging can be calculated as the sum of the individual vehicles as presented in (6.10).

$$GAC(t) = \sum_{n=1}^{evac} h_n(t) \tag{6.10}$$

where $evac$ is the number of electric vehicles which have arrived at time t. Other EV charging assumptions specific to each case study such as the base EV load shape, load shape following post-arrival charging and vehicle characteristics are discussed in Section 6.2.

6.1.2 Time Step Power Flow Simulation

The methodology incorporates a four stage power flow simulation to identify the impact of EV charging on the operating conditions during the wildfire as well as the potential

risk to the transmission network under subsequent outages.

Step 1: *Model the generation dispatch.* A rank order approach is used to identify the generation units online at a given point in time, with units added until a 10% reserve requirement is satisfied, or all available units are online, with the ranking given in Appendix A. The approach is flexible such that a variety of dispatch scenarios can be included to reflect anticipated or conservative dispatch patterns. Other factors specific to a wildfire such as a reduction in Solar PV generation due to wildfire smoke as highlighted in Chapter 5 may also be included.

Step 2: *Add evacuation demand.* The additional load due to evacuating electric vehicles calculated using the steps in Section 6.1.1 is included to reflect the charging behaviour of electric vehicles at each stage of the evacuation.

Step 3: *Conduct AC power flow at each time step.* When assessing the resilience, temporal power flow analyses are necessary to account for the hourly fluctuations in load and generation in relationship to the hazard.

Step 3a: *Check for convergence.* If the power flow solution diverges, loads are shed beginning with the bus with the largest power mismatch. This approach is selected as divergence can often result from voltage collapse and the worst mismatch bus would then be indicative of the contribution to the solution's divergence [183].

Step 3b: *Identify Violations* Upon reaching a converged solution, any additional violations of bus voltage limits and branch thermal ratings are identified.

Step 3c: *Mitigate Violations* Mitigation options can include generation re-dispatch, switching of shunt devices, on-load tap changing of transformers, network re-configuration, demand response, and load shedding. Specific options considered should be tailored to the network and regulatory criteria of the area under analysis.

Step 4: *Evaluate Subsequent Outages.* The likelihood of concurrent wildfires during periods of high risk is growing for some regions [18]. Furthermore utilities may ex-

perience normal outages during a wildfire event. Therefore, this step evaluates any potential cascading or degradation in system performance resulting from the additional electric vehicle evacuation load. The specific outages in this list can be further informed by each local planning area's internal and external criteria. To identify the impact, Repeat Steps 3a. - 3c. for the new network with the outage condition. The impact of evacuation on the system's ability to withstand further outages which might occur during the wildfire period is determined to identify the effect of EV evacuation on the system's overall resilience.

6.1.3 Critical Element Identification

Risk matrices provide an intuitive means of prioritizing risk, accounting for probability and impact [184]. After assessing the performance of the transmission system under the wildfire and subsequent outages, an outage criticality matrix is used to quantify risk for each contingency and identify the effect of EV evacuation on the overall risk. Outage criticality is a risk informed metric relating the probability of load shed resulting from an outage (Γ) and the magnitude of load shed (E).

$$E_k = \sum_{t \in T} LS_k(t) \tag{6.11}$$

$$\Gamma = \frac{100\%}{T} \sum_{t \in T} [LS_k(t) > 0] \tag{6.12}$$

where t is an hour over the event period T; and LS_k is the amount of load shed for a contingency k. These two factors are combined as shown in Figure 6.3 to form a qualitative outage criticality matrix to readily identify the outages which result in the most risk.

Each of the risk zones in the matrix is given a criticality label of 'Minimal', 'Moderate', 'Elevated', or 'Extreme'. The rationale for labelling each of the risk zones is based on the concept that elements which have the highest frequency and magnitude of load shed following a contingency are the most critical. The criticality is skewed to the right as

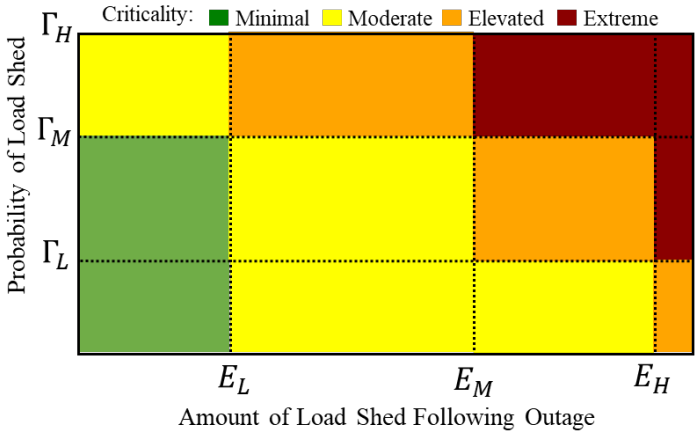


Figure 6.3: Example Criticality Matrix.

minimal amounts of load shed even for outages which are more likely to occur may not represent a substantial concern. To identify the impact of changes in demand from electric vehicle evacuation, symbols are used to indicate whether the presence of additional electric vehicle evacuation demand makes that outage worse, better, or is not anticipated to have an impact. This helps prioritise assessment of charging behaviour to specific locations. The criticality matrix is transformed to a geospatial map of critical elements, described in Section 6.3.2 to further aid operational awareness.

6.2 Electric Vehicle and Network Parameters

In order to demonstrate the affects of EV evacuation on the risk and resilience of the grid during a wildfire, the 2019 update to the IEEE Reliability Test System, RTS-GMLC, is used [132]. RTS-GMLC includes a modern generation mix and is geolocated to the south-western United States [132]. Further details surrounding data for the RTS-GMLC are given in Appendix A. While wildfires can affect many components of the power system, this case study focuses on concerns at a transmission level, where widespread load transfers are assessed. At the transmission level, wildfires can result in the failure or de-rate of transmission lines and for this case study, the assumption is that the 230 kV transmission line between Bus 313 and Bus 323 which passes directly through the wildfire area, is

disconnected for the duration of the fire. This is informed by the analysis in Chapter 3, that demonstrates this line may be subject to frequent wildfire related contingencies. The fire and associated outage are assumed to occur at 1:00am with the evacuation notice given at 9:00am. Assumptions for each of the model parameters are found in Table 6.1. The test system [132] and wildfire [175] are depicted in Figure 6.4 along with the evacuation proximity zones and destinations.

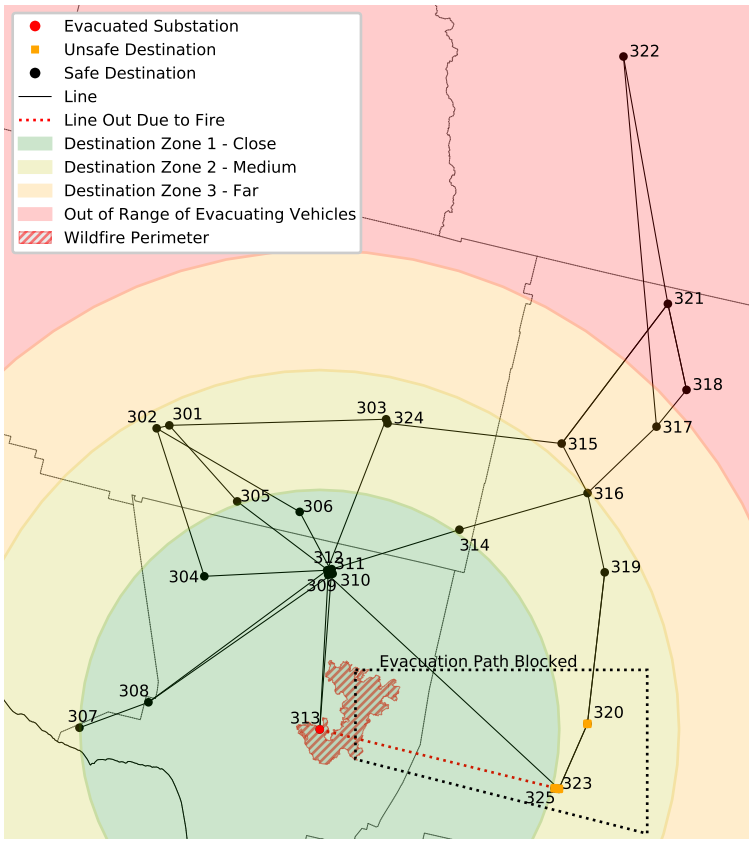


Figure 6.4: IEEE RTS-GMLC test system [132] fitted into Southern California geography with wildfire perimeter [175] and evacuation zones for an evacuation of Bus 313.

EV Behavioural Assumptions

Four types of vehicles are modelled: Plug in Hybrid Electric Vehicles with ranges of 20 and 50 miles (PHEV20 and PHEV50) and Battery Electric Vehicles with ranges of 100 and 250 miles (BEV100 and BEV250). Values of W, C, B and η for each vehicle are assumed consistent with those given in Appendix B, Table B.1. PHEV can travel beyond these ranges using gasoline, with the range being used to identify the charging requirements.

EV load shapes can be seen in Figure 6.5 are created using Electric Vehicle Infrastructure Projection Tool (EVI-Pro) Lite [112]. Further details around the base EV assumptions and EVI-Pro Lite are given in Appendix B.

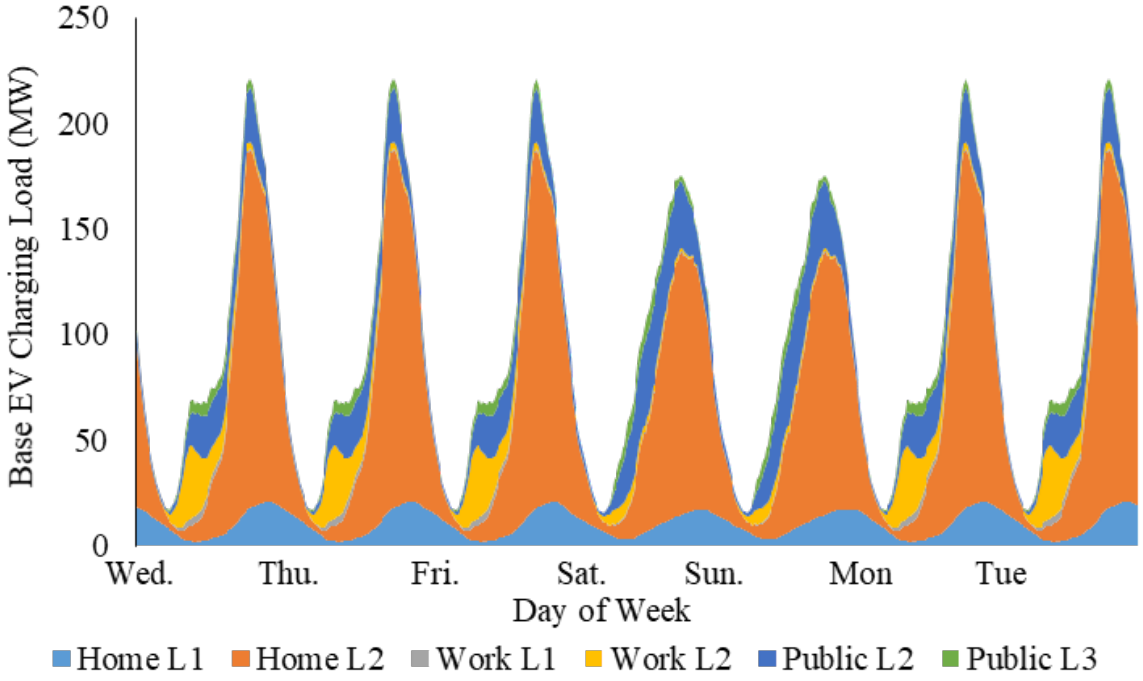


Figure 6.5: Base EV charging load shape for the test system (repeated over two weeks for the case study) created using EVI-Pro Lite [112].

To capture the response of households after receiving the evacuation notice, evacuating customers are modelled to plug in their vehicles from the point at which the notice is received until they evacuate. The remaining customers follow the base EV charging shape. The model reflects customers leaving based on the evacuation distribution, with some customers leaving before fully charged. It is assumed that sufficient charge is received to reach their destination or customers are aware of charging on the way. Upon reaching their destination, customers are modelled to recharge their battery at a public charger based on the distance they travelled to evacuate. On subsequent days a public Level 2 charging shape derived from the base shapes in Figure 6.5 is modelled. This shape is selected as these customers may no longer have access to chargers in the evacuation area, and instead be dependent on charging stations near their destination. Further Wildfire

and EV characteristics used for the Case Study are given in Table 6.1.

Table 6.1: Case Study Assumptions for Wildfire and Electric Vehicles

	Variable	Value	Rationale
Fire	Duration	2 weeks	Typical fire duration (weighted by fire size) for North America of 13.4 days based on 13.3 million individual fires from 2003-2016 [133]
	Time period	Summer peak	Worst wildfires in California recorded history typically occurred in July-September [185] which coincides with the periods of peak load in CAISO [186]
	Fire Extent	Bobcat Fire	2020 wildfire in the study area [175]
	Outages from fire	Line 313 to 323	Line passes directly through wildfire
	Evacuation Bus	313	Proximity to wildfire
	Unsafe Destinations	320, 323, 325	Evacuation path blocked by fire
	Residual Load at Evacuation Bus	20%	This captures the 11% of customers remaining as well as residual load, and load to serve fire-fighters and emergency personnel
EV	EV Fleet Size	128,000	20% EV penetration with the number of vehicles calculated using census data as in B.3
	Power Factor	0.98 inductive	Typical value determined in [187]
	Evac. Range (R)	100 mi	Surveys indicated wildfire evacuees preferred to travel shorter distances [20]. It is assumed that PHEV20 and PHEV40 rely on some gasoline if evacuating beyond their electric range.
	Evac. Speed (s)	30 mi/hr	Typical travel speed in absence of detailed routing information.
	Evacuation distribution (β)	117 minutes	Study for hurricanes in [188]
	Share Evacuating (Θ)	0.89	sample mean of the wildfire evacuation survey conducted in [189].

Outage and Criticality Matrix Assumptions

Following the base case evaluation, under the proposed framework, single and common corridor branch contingencies on top of the wildfire are considered. Overall the combination of the transmission line loss due to the wildfire, redistribution of load and subsequent single and common corridor contingencies, results in evaluation of several distinct N-k

events. For this case the post contingency operator adjustments consisted of adjustment of the shunt component, redispatch of online generation, and shedding of load. If post-convergence load shed is necessary, the amount is identified via AC optimal power flow to simultaneously evaluate generation re-dispatch and minimise load shed by modelling loads as generation with negative cost as given in [36] and described in Chapter 2. In assessing the feasibility of generation redispatch, changes in output were not limited by ramp rate or any storage use limits. For this analysis the following values of Γ are used: Γ_L: 1%, Γ_M: 10%, and Γ_H: 100%. The values 10 MW, 100 MW and 1000 MW are used for E_L, E_M, and E_H respectively. These values follow a logarithmic scale to enable consideration of an exponential increase in risk rather than a linear relationship, and should reflect the overall size of the study area.

Tools used for Analysis of the Case Study

To assess the case study, the process described in Section 6.1 is implemented via code using a combination of the following: MATPOWER [145] to conduct the power flow and determine the generation dispatch with the optimization solved using MATPOWER Interior Point Solver [146]; time step analysis scripted in *MATLAB ver. R2018b*; and the remainder of the analysis is conducted in python through the following packages: `scipy` [144], `numpy` [136] and `pandas` [137] packages for the data preparation, `geopandas` [138] and `cartopy` [141] for geospatial analysis and visualization using `matplotlib` [142]. Cartographic boundaries are obtained from the US Census Bureau [147].

6.3 Case Study

The data presented in Section 6.2 are assessed to provide insight into the significance of integrating EV evacuation charging demand in wildfire resilience assessments. Prior to analysis of evacuation, the changes from the wildfire itself are considered. Loss of transmission infrastructure due to a wildfire can cause significant redistribution of flow

on the transmission network. In this case study, loss of Line 313-323 was simulated due to the wildfire. As a result, without consideration of any additional EV evacuation charging demand, flow on the remaining transmission lines changed by up to 20% over the course of the case study period. Figure 6.6 shows the range of flow changes on each line.

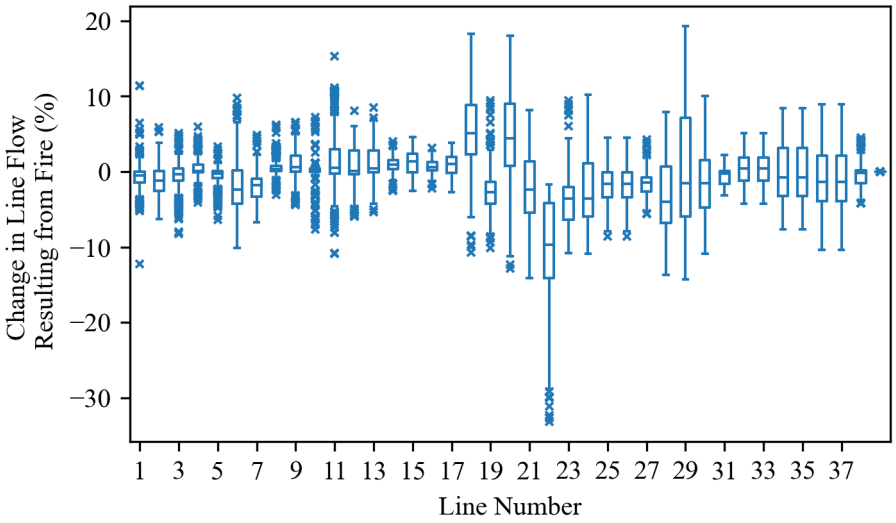


Figure 6.6: Changes in flow resulting from the wildfire over the course of the event period. Length of whiskers reflect 1.5 IQR (Inter-Quartile Range).

Such changes can require grid operators to quickly adapt to abnormal flow patterns, re-dispatch generation, and prepare for any subsequent outages which may occur. EV evacuation can add further complexity to these decisions by altering the anticipated loading patterns on the transmission network. The impact of these loading patterns is assessed for the base network as well as subsequent outages as part of the wildfire resilience assessment.

6.3.1 EV Evacuation Charging Demand Impacts

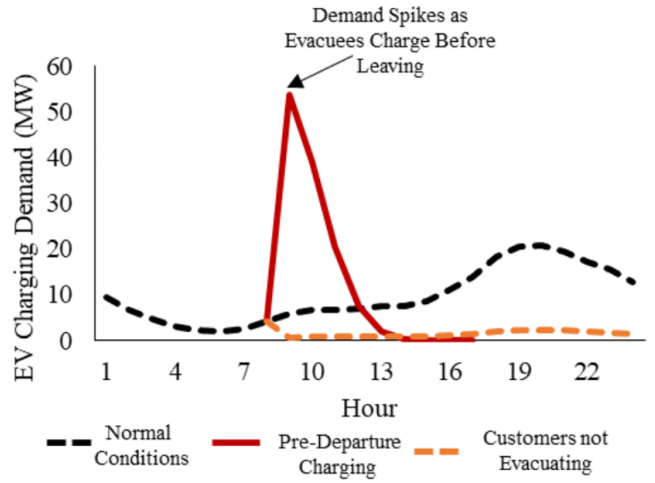
In the period following the evacuation notice, many residents may simultaneously charge their EV prior to evacuating. This synchronised charging can cause spikes in demand that may adversely affect transmission infrastructure in the area. For the case study, pre-departure EV charging demand at Bus 313 spiked to 53 MW. This reflects a near tripling of the normal diversified peak EV charging demand and is roughly 20% of the substation's

peak load. Figure 6.7a depicts the changes in electricity demand at Bus 313 from EV charging. For this case study, the evacuation occurred prior to the daily peak demand and prior to the system peak later in the study period. However, if the evacuation was coincident with the daily or annual system peak, a 20% increase in peak load could further exacerbate risk of load shed. Therefore, in an electrified future, when issuing evacuation orders and developing evacuation plans, it is critical to consider temporal charging effects.

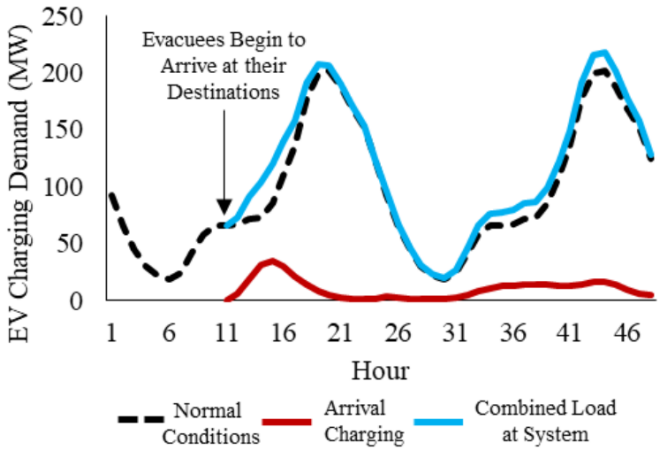
EV charging and evacuation further alters loading on the transmission system as vehicles are dispersed throughout the network. The change in EV charging demand at the destination substations can be seen in Figure 6.7b. This additional charging demand also reduced the overall voltage profile of the system. The impacts could be even higher with higher levels of EV penetration, DC fast charging or less dispersed evacuation. Awareness of these spatiotemporal changes in transmission loading are important to properly position the transmission network to maintain safe and reliable operation during a wildfire. This awareness also enables performance of real time contingency analysis for the conditions identified to be most adversely affected by evacuation demand. The redistribution of load due to EV charging during an evacuation can also reduce the flow seen on some corridors, potentially improving system resilience to subsequent contingencies affecting those lines. Therefore, it is important to further examine the effect of EV charging on subsequent outages during a wildfire event.

6.3.2 Resilience to Subsequent Outages

During the wildfire evacuation period, additional outages may occur due to the wildfire, other wildfires occurring in the region, or other causes. For the case study, evaluation of the 38 single element and seven common-corridor contingencies beyond the initial loss of Line 313-323 was conducted. Eight contingencies resulted in load shed for at least one hour during the event period. The impact to the transmission network from EV evacuation varies dynamically across contingencies. The case study results reveal that EV evacuation heightens the risk of cascading outages by increasing line loading and load



(a) Pre-Departure Charging at Bus 313.



(b) Charging Demand at Destination Substations.

Figure 6.7: EV Evacuation Impact to Charging Demand.

shed exposure. For some contingencies the change in load shedding exposure is minor, whereas for others EV evacuation increases the frequency and magnitude.

The proposed criticality matrix can identify the most significant contingencies to assess further and is depicted in Figure 6.8 for the case study. Note that consequential load loss is not included in the criticality matrix as these contingencies would already be known. The arrows are used as symbols to indicate whether integration of EV evacuation charging demand reduces, increases or has inconsequential impact on the load shed for each outage. Inconsequential for purposes of the case study was determined by a change in load at risk less than 5 MW over the course of the event period. The resulting matrix indicates that for

the most significant contingencies, EV evacuation charging demand increases the risk and should be accounted for when developing mitigation plans. An example of the magnitude of risk change resulting from the additional EV evacuation charging demand is given for the loss of line 306-310 where E increases from 711 to 781 MW and Γ increases from 9.2% to 10.1% (3 additional hours). This also causes the contingency to move from elevated to extreme risk zone, demonstrating that failure to consider evacuation load can result in underestimation of risk. While the case study evaluates EV evacuation charging demand due to wildfires, other events that cause mass movement of electric vehicles outside of their normal charging locations such as hurricanes or sporting events could result in similar disruptions to the grid.

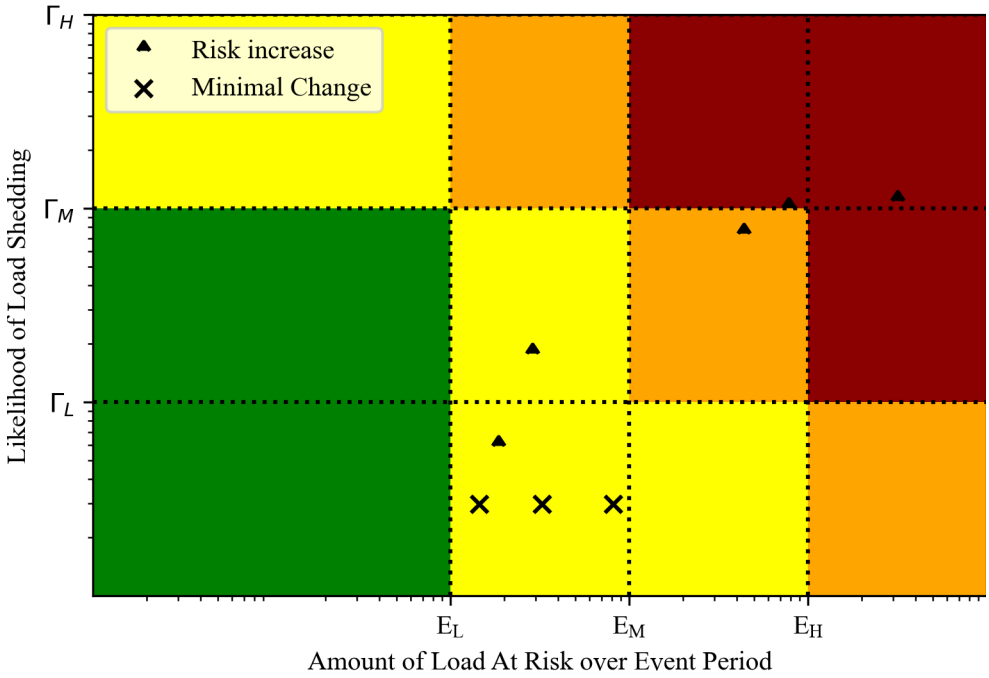


Figure 6.8: Outage criticality matrix indicating the impact of EV evacuation on risk.

The criticality matrix can also be transformed into a map to provide geospatial indication of the criticality of each element and the impact of EV evacuation. This is depicted in Figure 6.9. The lines which are affected most by electric vehicle evacuation (indicated by the upwards arrow) are those surrounding Bus 310, as most vehicles evacuate to this area (as seen in Figure 6.10. Figure 6.9 enables the most critical lines to be quickly identified along with their location and likelihood of resulting in load shed. Awareness of EV

evacuation charging location and magnitude can enable utilities to better plan to serve load during an emergency and proactively prepare the transmission system, improving its resilience.

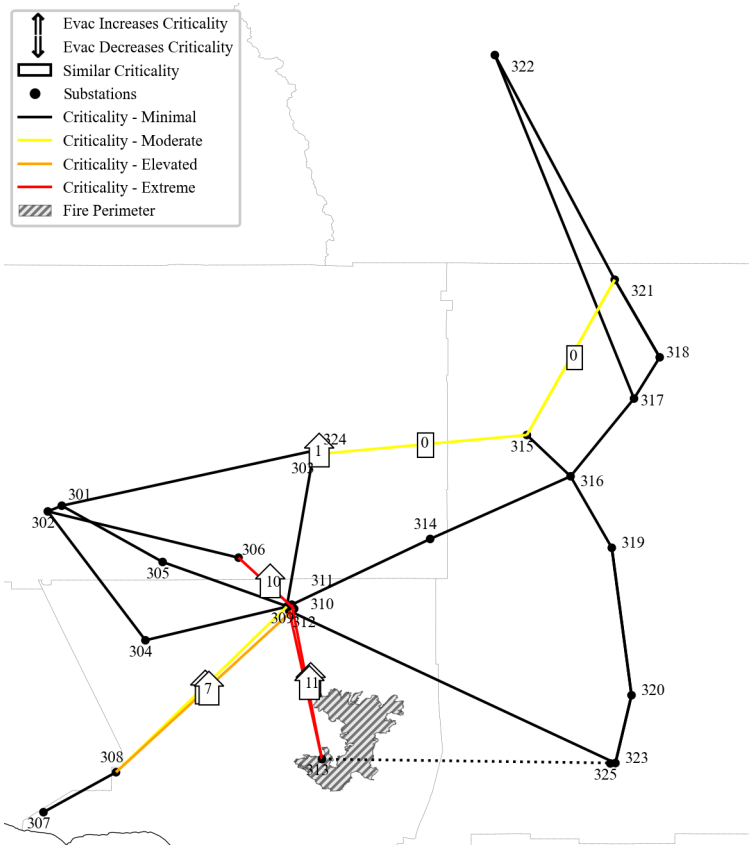


Figure 6.9: Identification of Critical Transmission Elements. Worst outage each element involved in is shown with the number inside the arrow or box representing the % of hours that load shed will occur after that outage during the wildfire period.

Results using the proposed methodology for this case study demonstrate that evacuation of mobile loads during wildfires can place additional stress on specific transmission components, increasing risk of load shedding, and adversely effecting overall system resilience. This demonstrates the need to provide electric vehicle penetration levels to power system operators and planners. The demand profiles of the evacuating vehicles based on the evacuation model can be generated ahead of time for different evacuation scenarios and could therefore be used to inform online decision making. This information should be shared with emergency planning agencies to assess whether sufficient charging is available to support the evacuation and to identify ideal or anticipated locations for primary

evacuation zones.

6.4 Recommended Means to Enhance Resilience

The results of the proposed approach can enable proactively decision-making to enhance wildfire resilience by improving operator preparedness and inform the need proactively site temporary EV charging stations. Once the anticipated EV evacuation charging demand is obtained, there are two routes to mitigate any adverse effects: The first route is to eliminate or minimise the additional charging load which will be seen by the grid. This can be done by deploying temporary charging stations, which are not tied to the main grid [190]. These charging stations can be sited in high fire risk areas to mitigate the initial spike in demand from evacuating EV, if an evacuation were to take place or to aid evacuees in charging along the way. Use of Equation 6.3 also could enable siting of off-grid charging stations at evacuation destination locations. This could preemptively mitigate any performance issues identified as a result of the increased charging. An example map such as the one depicted in Figure 6.10 could be used to inform such efforts.

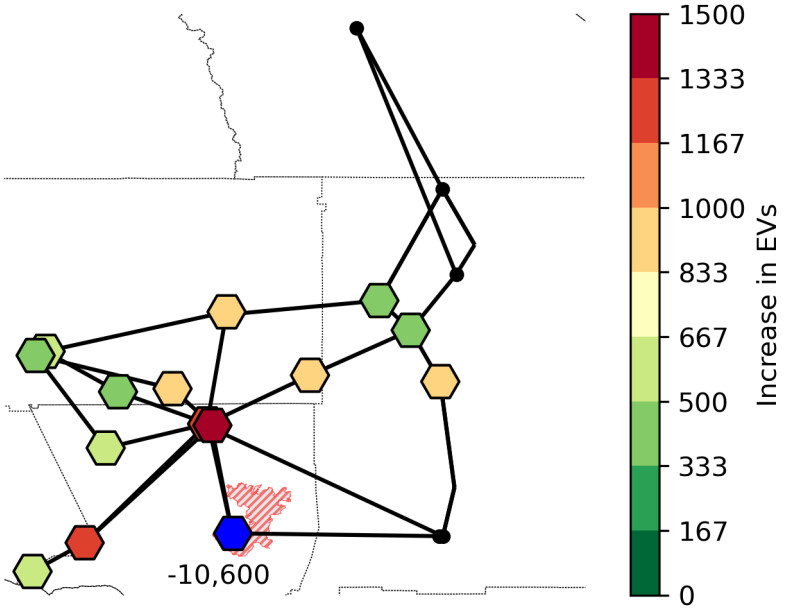


Figure 6.10: Change in Electric Vehicles at each location as a result of Evacuation.

The second route for mitigation is to enhance the readiness of the grid to withstand the additional load. This could be achieved by increasing the real and reactive reserve margin of the grid to account for the anticipated EV evacuation load. Furthermore, operators could incorporate the changes in load into their contingency analysis to ensure that any proposed resiliency mitigation are sufficient to handle evacuation. As electrification grows, this information will become increasingly important for other types of events beyond natural disasters that could involve large temporary shifts in population including tourism and major sporting events. This could be achieved via a screen or indicator as part of the energy management system which details the anticipated penetration of electric vehicles at that specific bus.

6.5 Summary

EV evacuation during wildfires can cause dynamic operational stress even at 20% EV penetration, adding complexity to operator decision-making. The proposed methodology provides a reproducible means of simulating the effects of EV evacuation on the transmission network and is extensible to reflect a range of operating conditions, grid topologies and EV penetration levels. As evacuees synchronously charge, the EV charging demand can increase to three times the normal diversified charging demand. This additional demand during a wildfire evacuation can contribute to increased line loading and risk of load shed in some portions of the system, while reducing line loading in others. Therefore, grid operators should be involved in emergency planning efforts to prepare for the impact of EV evacuation.

While many potential solutions exist to remediate grid stress resulting from evacuation charging load, awareness of the magnitude of EV charging impacts throughout the system is a pre-requisite to identification of appropriate mitigation measures. As the changes in demand introduced by evacuation vary geographically, the critical element identification method presented in this chapter indicates the impact of EV evacuation

charging on the most critically affected transmission areas, enabling the establishment of targeted adaptation and mitigation plans. This research provides a solid foundation for future researchers to investigate the optimal remediation strategy for extreme wildfires considering the effects of evacuating vehicles.

Chapter 7

Conclusion and Future Research

Extreme wildfires can result in damage, failure, or destruction of transmission infrastructure, posing a significant threat to the safety and resiliency of the electric grid. Failure to consider the interactions between wildfires and LCT can result in over-forecasting PV generation output due to wildfire smoke, or under-forecasting the electricity demand due to EV charging during an evacuation, further compounding the threat. This chapter presents key findings of the research in this thesis to inform assessments of power system wildfire resilience in the planning and operating horizon, reflecting the behaviour of LCT, along with potential future areas of research.

7.1 Conclusions

Predicting adverse effects from wildfires enables proactive development of strategies to enhance the resilience of the system. However, some wildfire mitigation measures require significant lead time to implement due to construction time and regulatory requirements. Chapter 3 presented an advanced methodology to assess power system wildfire resilience in the planning horizon, enabling risk quantification in advance of a disaster. The methodology also quantifies critical transmission assets that require reinforcement and visualises assets for rapid mitigation development geospatially. The uniqueness of the methodology is the integration of satellite-derived empirical wildfire ignitions with grid topology

to produce data of synthetic wildfire seasons and quantify the risk over an entire year, rather than focusing on a single event. The viability of the methodology was assessed using scenarios considering numerous synthetic wildfire seasons in Southern California. A significant finding from the case study was that for outage magnitude, the $CVaR$ was double the expected value. Line Outage Days showed even more significant difference with the $CVaR$ triple the expected value. The use of expected value metrics (traditionally used in reliability studies) mask the risk during extreme wildfires seasons, potentially leading to insufficient mitigation measures. The proposed methodology addresses this issue through use of robust risk-based metrics. Results also confirm that the outage impact is not directly correlated with the number of elements offline, requiring analysis of each outage to identify the severity.

Residential solar PV masks the true load of a distribution feeder, making awareness of PV installations important for operators during restoration planning efforts. Solar PV generation also provides a resource for use in resilience mitigation measures such as local area microgrids. However, utilities may not be aware of the location of all installations, compromising the integrity of planning efforts. To overcome this limitation, Chapter 4 presented a new agglomerative clustering based approach to automatically identify solar prosumers. The uniqueness of the approach is that it only requires hourly smart meter data and is able to reduce the data required for classification to a single data-point through PAA, significantly reducing the computational burden without compromising accuracy. The approach is validated using four years of real smart meter data from the United States, yielding 99.7% classification accuracy. Therefore, hourly smart meter data provides an accessible source of data to effectively locate solar PV installations. The proposed SPIDCs quantify the amount of data needed to meet a specific accuracy threshold to provide confidence in decision-making. A significant finding from the case study is that only one week of historical data was sufficient to classify solar prosumers with greater than 95% accuracy. Customer classification should be tracked over time to identify new prosumer installations, anomalous solar PV performance, or panel removal. The proposed approach

results in a current snapshot of the location and status of PV installations, supporting effective wildfire mitigation decision-making.

During a wildfire, smoke inhibits the solar irradiance, adversely impacting solar PV generation in the affected areas. Knowledge of how this smoke will impact performance is critical prior to relying on solar PV generation at the system level or as part of a microgrid. Chapter 5 presented a methodology incorporating quantile regression to predict the derate to solar PV capacity during a wildfire. The uniqueness of the methodology comes from using satellite-derived aerosol optical depth (AOD) measurements to capture the quantity of particulate in the air during a wildfire over large regions and translate those measurements into reductions in generation. The efficacy of the proposed approach is validated using real PV operating data from California's 2020 wildfire season. An important finding is that days with smoke yielded substantial reductions in PV performance across vast geographical areas. As a result, even if a wildfire may be far away, assessments of wildfire resilience in the operating horizon should account for the potential disruption to capacity across entire planning areas. Another finding is that real-time AOD forecasts are a key source data in the operation horizon. The geospatial derate maps resulting from the proposed methodology provide a practical mechanism to translate AOD forecasts into anticipated PV performance, enhancing generation scheduling and dispatch during a wildfire.

During a wildfire, EV charging demand can adversely affect grid performance by altering loading patterns. As entire communities are forced to evacuate their homes and flee extreme wildfires, evacuees may synchronously charge, resulting in a spike in electricity demand. Chapter 6 presented a novel model to predict the changes to EV charging demand during an evacuation. The novelty comes from the model's ability to reflect the geospatial distribution of charging load over the entire event horizon, including charging before, during and after evacuation. The viability of the methodology was assessed through a case study reflecting 20% EV penetration and a 2020 wildfire perimeter. Results indicated that localised EV load seen at the transmission level can triple as customers evacuate, and

increase exposure to load shed. The proposed model can predict changes in load ahead of time, enabling operators to re-dispatch the grid to minimise risk of overloads, thereby enhancing the resilience of the network. As more EVs are equipped with vehicle to grid capabilities, the proposed model can also support the development of resilience plans to leverage EVs in restoration efforts.

The investigation in this thesis has made significant contributions to inform assessments of power system wildfire resilience in the planning and operating horizon, reflecting the behaviour of LCT. First, a methodology to proactively quantify wildfire risk in the planning horizon is presented. Use of this methodology provides the ability to evaluate mitigation measures well in advance to increase the time available for implementation. Second, the ability to effectively identify solar prosumers further enhances mitigation planning. The proposed approach was robust, efficient, and provides a mechanism to obtain a current snapshot of PV installations in an area of interest. Third, quantification of the relationship between wildfire smoke and PV generation capacity is achieved. Real-time AOD forecasts enable use of this relationship to predict the impact to PV generation over entire planning areas rather than single locations. Finally, a methodology to integrate EV evacuation charging demand provides operators with a means to reflect the spatiotemporal variance in power system wildfire resilience assessments. Foresight of significant increases in charging demand provides the opportunity to proactively dispatch generation and mitigate increased stress to power system infrastructure. For each of these contributions, visualisation varying over space and time is given to provide practical insight to support decision-making across the planning and operating horizon.

The research presented in this thesis demonstrates that wildfires have wide-reaching impact to power system infrastructure, causing outages of transmission lines, reduction in solar PV generation, and altering demand. Although the focus of this thesis has been on wildfire resilience, the spatiotemporal methods presented in this thesis can be extended to other natural disasters, leading to more integrated overall resilience plans. Ultimately, robust assessments that consider grid performance over an entire wildfire season along

with the effect of growing penetration of LCT are essential to develop a modern power grid that is resilient to wildfires.

7.2 Future Work

The work in this thesis presents foundational methodology for the analysis of power system wildfire resilience in the presence of LCT. The following section provides possible avenues for future research work.

- **Post-Wildfire Landslide Resilience.** This thesis provided a methodology to identify the transmission corridors most at risk from wildfires in the planning horizon. After a wildfire burns the foliage in a specific area, even if the transmission system remains intact, transmission lines may be at risk due to increased risk of landslide in the period following [191]. Initial work to assess the resilience of power system equipment during a landslide has been carried out in [192], using regression to identify the influence of factors such as slope, soil, and geology. Future researchers could assess the propensity of damage to power system infrastructure to create composite fragility curves to represent the landslide risk before and after a wildfire. Use and collection of empirical hazard data in areas prone to both wildfires and landslides could inform this effort. The resulting fragility curves could be used to coordinate emergency planning efforts across natural disaster response agencies.

- **Integrated Prosumer Evaluation Tool.** As more customers adopt electric vehicles, distributed energy storage and heat pumps alongside residential solar PV, new methods for comprehensive detection of customer LCT adoption of will be needed. Future research could integrate the method to identify solar PV prosumers presented in this thesis with other methods for EV, energy storage, and heat pump detection to develop a tool that comprehensively identifies the technologies customers have adopted in real-time and predicts their behaviour. Many methods are being developed to disaggregate solar PV consumption from net load data [193] and provide

opportunities for application to prosumer detection. This is made more complex by rising energy storage adoption [194] and understanding the control behaviour, pricing and operating modes will be important to ensure robust identification. Identifying data sources that can enable differentiation between prosumer sources of generation will also be important to support this work. Overall, development of such a tool could inform real-time restoration and enhance operational resilience during extreme events. Further research could also explore how this tool could be integrated into existing grid management software to enhance operator LCT visibility.

- **Empirical EV Charging Behavioural Assessment during Abnormal Events.** While the behaviour of EV charging under typical conditions is becoming more well known, EV performance during natural disasters is less understood. The work in Chapter 6 of this thesis demonstrated via simulated behaviour that EV charging profiles during a wildfire may differ markedly from those during normal operating conditions. The emergence of new geo-data sources have enabled researchers to track human mobility with [195] detailing many sources of data applied to tracking mobility dynamics during the COVID-19 pandemic. Such sources of mobility data including empirical EV connectivity data could be used to identify how EV charging diversity factors differ across extreme events including hurricanes, floods, and earthquakes as well as sporting events such as the Commonwealth games. Such analysis could enable better capacity utilisation of the distribution network and charging infrastructure development.

Appendix A

IEEE RTS-GMLC

This section presents the information of Area 3 of the 2019 Update to the IEEE Reliability Test System from the Grid Modernization Laboratory Consortium (RTS-GMLC) [132] which was used for demonstrating the case study in Chapters 3 and 6. Data for the RTS-GMLC is available at `https://github.com/GridMod/RTS-GMLC` provided by National Renewable Energy Laboratory ("NREL"), which is operated by Alliance for Sustainable Energy, LLC ("ALLIANCE") for the U.S. Department Of Energy ("DOE") [132]. The buses, peak load and voltage limits used are presented in Table A.1 with underlying data from [132]. Authors in [132] note that the system is overbuilt to allow users to consider a range of generation configurations. The overbuilt nature is exemplified as the original Region 3 contains only 2850 MW of load but $\approx$ 6800 MW of Generation capacity of which $\approx$ 3900 MW is wind and solar. Hence for this analysis, the wind and solar generation capacity is reduced by 50%. This reduction is also informed by insight from Chapter 5 which indicates that solar PV generation capacity can adversely affect solar PV generation capacity. The corresponding generation information from [132] with these reductions and the rank order used for the case study in Chapter 6 added is shown in Tables A.2 and A.3. Finally, the transmission line information is presented in Table A.4 with underlying data from [132].

Table A.1: Bus information used for Case Study

Bus	Type	Load (MW)	Load (MVaR)	Voltage (kV)	Vmin (pu)	Vmax (Pu)
301	2	108	22	138	0.95	1.1
302	2	97	20	138	0.95	1.1
303	1	180	37	138	0.95	1.1
304	1	74	15	138	0.95	1.1
305	1	71	14	138	0.95	1.1
306	1	136	28	138	0.95	1.1
307	2	125	25	138	0.95	1.1
308	1	171	35	138	0.95	1.1
309	1	175	36	138	0.95	1.1
310	1	195	40	138	0.95	1.1
311	1	0	0	230	0.95	1.1
312	1	0	0	230	0.95	1.1
313	2	265	54	230	0.95	1.1
314	2	194	39	230	0.95	1.1
315	2	317	64	230	0.95	1.1
316	2	100	20	230	0.95	1.1
317	1	0	0	230	0.95	1.1
318	3	333	68	230	0.95	1.1
319	1	181	37	230	0.95	1.1
320	1	128	26	230	0.95	1.1
321	2	0	0	230	0.95	1.1
322	2	0	0	230	0.95	1.1
323	2	0	0	230	0.95	1.1
324	1	0	0	230	0.95	1.1
325	1	0	0	230	0.95	1.1

Type 1 = PQ, Type 2 = PV, and Type 3 = Swing

Table A.2: Generation information used for Case Study (Oil, Natural Gas, Coal, Synchronous Condenser, and Hydro)

ID	Bus	Type	P Max (MW)	Pmin (MW)	Q Max (MVaR)	Qmin (MVaR)	Rank Order
301_CT_1	301	Oil	20	8	10	0	65
301_CT_2	301	Oil	20	8	10	0	66
301_CT_3	301	NG	55	22	19	-15	51
301_CT_4	301	NG	55	22	19	-15	52
302_CT_1	302	Oil	20	8	10	0	67
302_CT_2	302	Oil	20	8	10	0	68
302_CT_3	302	NG	55	22	19	-15	53
302_CT_4	302	NG	55	22	19	-15	54
307_CT_1	307	NG	55	22	19	-15	49
307_CT_2	307	NG	55	22	19	-15	50
313_CC_1	313	NG	355	170	150	-25	56
315_STEAM_1	315	Oil	12	5	6	0	60
315_STEAM_2	315	Oil	12	5	6	0	61
315_STEAM_3	315	Oil	12	5	6	0	62
315_STEAM_4	315	Oil	12	5	6	0	63
315_STEAM_5	315	Oil	12	5	6	0	64
315_CT_6	315	NG	55	22	19	-15	44
315_CT_7	315	NG	55	22	19	-15	45
315_CT_8	315	NG	55	22	60	0	46
316_STEAM_1	316	Coal	155	62	80	-50	69
318_CC_1	318	NG	355	170	150	-25	59
321_CC_1	321	NG	355	170	150	-25	55
322_CT_5	322	NG	55	22	19	-15	47
322_CT_6	322	NG	55	22	19	-15	48
323_CC_1	323	NG	355	170	150	-25	57
323_CC_2	323	NG	355	170	150	-25	58
314_SYNC_COND_1	314	Sync_Cond	0	0	200	-50	1
322_HYDRO_1	322	Hydro	50	0	16	-10	2
322_HYDRO_2	322	Hydro	50	0	16	-10	3
322_HYDRO_3	322	Hydro	50	0	16	-10	4
322_HYDRO_4	322	Hydro	50	0	16	-10	5

Table A.3: Generation information used for Case Study (Solar, Wind, and Storage)

ID	Bus	Type	P Max (MW)	Pmin (MW)	Q Max (MVaR)	Qmin (MVaR)	Rank Order
320_PV_1	320	Solar	25.8	0	0	0	6
314_PV_1	314	Solar	25.8	0	0	0	7
314_PV_2	314	Solar	25.8	0	0	0	8
313_PV_1	313	Solar	47.55	0	0	0	9
314_PV_3	314	Solar	46.35	0	0	0	10
314_PV_4	314	Solar	25.8	0	0	0	11
313_PV_2	313	Solar	46.65	0	0	0	12
310_PV_1	310	Solar	25.85	0	0	0	13
324_PV_1	324	Solar	24.85	0	0	0	14
312_PV_1	312	Solar	47.05	0	0	0	15
310_PV_2	310	Solar	25.8	0	0	0	16
324_PV_2	324	Solar	25.8	0	0	0	17
324_PV_3	324	Solar	25.5	0	0	0	18
319_PV_1	319	Solar	94.1	0	0	0	19
308_RTPV_1	308	Solar	50.45	0	0	0	20
313_RTPV_1	313	Solar	50.85	0	0	0	21
313_RTPV_2	313	Solar	31.55	0	0	0	22
313_RTPV_3	313	Solar	32.7	0	0	0	23
313_RTPV_4	313	Solar	33.5	0	0	0	24
313_RTPV_5	313	Solar	32.4	0	0	0	25
313_RTPV_6	313	Solar	31.9	0	0	0	26
313_RTPV_7	313	Solar	32.05	0	0	0	27
313_RTPV_8	313	Solar	33.3	0	0	0	28
313_RTPV_9	313	Solar	31.2	0	0	0	29
313_RTPV_10	313	Solar	33.45	0	0	0	30
313_RTPV_11	313	Solar	32.6	0	0	0	31
313_RTPV_12	313	Solar	13.9	0	0	0	32
320_RTPV_1	320	Solar	13.65	0	0	0	33
320_RTPV_2	320	Solar	13.5	0	0	0	34
320_RTPV_3	320	Solar	14.15	0	0	0	35
313_RTPV_13	313	Solar	13.6	0	0	0	36
320_RTPV_4	320	Solar	13.5	0	0	0	37
320_RTPV_5	320	Solar	14.1	0	0	0	38
320_RTPV_6	320	Solar	4.7	0	0	0	39
309_WIND_1	309	Wind	74.15	0	0	0	40
317_WIND_1	317	Wind	399.55	0	0	0	41
303_WIND_1	303	Wind	423.5	0	0	0	42
313_STORAGE_1	313	Storage	50	0	0	0	43

Table A.4: Branch information (Lines and Transformers) used for Case Study

ID	From Bus	To Bus	R (pu)	X (pu)	B (pu)	Continuous Rating (MVA)	Emergency Rating (MVA)	MTTR (hr)
C1	301	302	0	0.01	0.46	175	193	16
C2	301	303	0.06	0.21	0.06	175	208	10
C3	301	305	0.02	0.09	0.02	175	208	10
C4	302	304	0.03	0.13	0.03	175	208	10
C5	302	306	0.05	0.19	0.05	175	208	10
C6	303	309	0.03	0.12	0.03	175	208	10
C7	303	324	0	0.08	0	400	510	768
C8	304	309	0.03	0.1	0.03	175	208	10
C9	305	310	0.02	0.09	0.02	175	208	10
C10	306	310	0.01	0.06	2.46	175	193	35
C11	307	308	0.02	0.06	0.02	175	208	10
C12-1	308	309	0.04	0.17	0.05	175	208	10
C13-2	308	310	0.04	0.17	0.05	175	208	10
C14	309	311	0	0.08	0	400	510	768
C15	309	312	0	0.08	0	400	510	768
C16	310	311	0	0.08	0	400	510	768
C17	310	312	0	0.08	0	400	510	768
C18	311	313	0.01	0.05	0.1	500	600	11
C19	311	314	0.01	0.04	0.09	500	600	11
C20	312	313	0.01	0.05	0.1	500	600	11
C21	312	323	0.01	0.1	0.2	500	600	11
C22	313	323	0.01	0.09	0.18	500	600	11
C23	314	316	0.01	0.06	0.08	500	600	11
C24	315	316	0	0.02	0.04	500	600	11
C25-1	315	321	0.01	0.05	0.1	500	600	11
C25-2	315	321	0.01	0.05	0.1	500	600	11
C26	315	324	0.01	0.05	0.11	500	600	11
C27	316	317	0	0.03	0.06	500	600	11
C28	316	319	0	0.02	0.05	500	600	11
C29	317	318	0	0.01	0.03	500	600	11
C30	317	322	0.01	0.11	0.22	500	600	11
C31-1	318	321	0	0.03	0.06	500	600	11
C31-2	318	321	0	0.03	0.06	500	600	11
C32-1	319	320	0.01	0.04	0.08	500	600	11
C32-2	319	320	0.01	0.04	0.08	500	600	11
C33-1	320	323	0	0.02	0.05	500	600	11
C33-2	320	323	0	0.02	0.05	500	600	11
C34	321	322	0.01	0.07	0.14	500	600	11
C35	323	325	0	0.01	0	722	893	768

Appendix B

Electric Vehicle Information

Base EV Characteristics

To reflect a realistic vehicle composition, four types of vehicles are modelled: Plug in Hybrid Electric Vehicles with ranges of 20 and 50 miles (PHEV20 and PHEV50) and Battery Electric Vehicles with ranges of 100 and 250 miles (BEV100 and BEV250). Values of Nominal Efficiency W, Maximum AC Charging Power C, and Onboard Charger Efficiency (η) for each vehicle are assumed consistent with those given in Table B.1. The nominal range and nominal efficiency are used to calculate the battery size.

Table B.1: EV Characteristics from EVI-Pro-Lite [112]

EV Characteristics	Units	PHEV20	PHEV50	BEV100	BEV250
Nominal Electric Driving Range	mi	20	50	100	250
Sedan Nominal Efficiency (excludes charger effic.)	Wh/mi	325	325	325	325
SUV Nominal Efficiency (excludes charger effic.)	Wh/mi	450	450	450	450
Onboard Charger Efficiency	%	90	90	90	90
Maximum AC Charging Power	kW	3.6	3.6	7.2	11.5
Maximum DC Charging Power	kW	0	0	50	150

Vehicle State of Charge Prior to Evacuation

The state of charge of electric vehicles can vary substantially and has a direct impact on the charging demand required prior to evacuation. Authors in [196] indicate that vehicles should maintain a SOC in the range of 20-80% to minimize losses. However, empirical SOC data from [187] presented in blocks of 8.3% would indicate that in practice vehicle

owners may operate outside of these ranges. For this work the empirical data from [187] for a weekday between the hours of 6am and 10am is used as a foundation, but the share of vehicles which have a SOC of either 0/12 or 1/12 are removed and the remaining shares scaled so that the values sum to 100%. The resulting values can be seen in Table B.2.

Table B.2: State of Charge of Electric Vehicles in the Evacuation Area Upon Receiving Evacuation Notice. Adapted from [187]

SOC	% of Vehicles
0.00%	0%
8.33%	0%
16.67%	7%
25.00%	10%
33.33%	10%
41.67%	10%
50.00%	11%
58.33%	12%
66.67%	10%
75.00%	7%
83.33%	8%
91.67%	7%
100%	7%

Base EV Load Shapes

To generate typical load shapes for a particular fleet of residential electric vehicles, physical characteristics should be combined with be behavioural factors including the typical vehicle miles travelled and preferences for charging location (i.e. work or home). The National Renewable Energy Laboratory (NREL) and the California Energy Commission (CEC) in partnership with the U.S. DOE's Vehicle Technologies Office have developed the Electric Vehicle Infrastructure Projection Tool (EVI-Pro) as a bottoms-up tool to generate electric vehicle load shapes [112]. The U.S. DOE has made an open source simplified version of this tool EVI-Pro-Lite available to generate load profiles for customised vehicle fleets and locations [112]. Tools such as EVI-Pro Lite can be used by grid operators and city planners to identify typical EV charging demand for regions of their service territory.

For creation of the base profiles the NREL EVI-Pro Lite API script was used (avail-

able at `https://github.com/NREL/EVI-Pro-Lite` [197]) with the following parameters: fleet_size: 128000; mean_dvmt: 35; temp_c: 35 with short temp list; pev_dist: EQUAL; class_dist: Sedan; home_access_dist: HA100; home_power_dist and work_power_dist: MostL2; pref_dist: Home100; and res_charging and work_charging: min_delay. BEV250 are modelled to have 11.5 kW chargers to enable a full charge. Further description of each of these variables is given in [112].

EV Population Calculation

California household vehicle ownership statistics from the U.S. Census Bureau for California [198] shown in Table B.3 were used, conservatively assuming that no household has more than 3 vehicles to obtain an average of 1.8 vehicles per household. Assumptions surrounding the share of residential load and number of households in the test network are shown in Table B.4. This yields 640,000 vehicles in the entire test system and 128,000 electric vehicles at 20% EV penetration. The number of vehicles at each individual bus is assumed to be proportional to the overall load of the bus. For example, bus 313 contains approximately 9% of the system load and is allocated a total of almost 60,000 vehicles (almost 12,000 of which are electric).

Table B.3: California Household Vehicle Ownership [198]. Table reproduced from [P2]

Category	Estimate	Margin of Error
No Vehicle	939,034	±6,018
1 Vehicle	3,993,143	±12,465
2 Vehicles	4,838,980	±15,184
3 Vehicles	3,194,278	±14,420
3+ Vehicles	12,965,435	±19,785

Table B.4: Calculation of Total Vehicles in Test Network. Adapted from [P2]

Assumption	Value
Total Area Peak Load	2,850 MW
Residential Load Share	50%
Individual Household Peak	4 kW
Number of Households	356,250
Number of Vehicles	640,000
Number of Electric Vehicles	128,000

References

[1] Climate Change Committee, "Independent assessment of uk climate risk," *Advice to Government For the UK's third Climate Change Risk Assessment (CCRA3)*, 2021.

[2] S. Thacker, R. Pant, and J. W. Hall, "System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures," *Reliability Engineering & System Safety*, vol. 167, pp. 30–41, 2017.

[3] S. Ekisheva, R. Rieder, J. Norris, M. Lauby, and I. Dobson, "Impact of extreme weather on north american transmission system outages," *2021 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, 2021.

[4] Energy Emergencies Executive Committee, "GB POWER SYSTEM DISRUPTION – 9 AUGUST 2019 Energy Emergencies Executive Committee: Interim Report ," Department for Business, Industry, and Industrial Strategy, Report, 2019.

[5] J. W. Busby *et al.*, "Cascading risks: Understanding the 2021 winter blackout in texas," *Energy Research & Social Science*, vol. 77, p. 102106, 2021.

[6] IEEE, "The authoritative dictionary of IEEE standards terms, seventh edition," *IEEE Std 100-2000*, pp. 1–1362, 2000.

[7] United Nations General Assembly, "Transforming our world: the 2030 agenda for sustainable development, 21 october 2015," A/RES/70/1, Report, 2015.

[8] E. Ciapessoni, D. Cirio, A. Pitto, M. Panteli, M. Van Harte, and C. Mak, "Defining power system resilience," CIGRE C4.47 reference paper, Report, 2019.

[9] IEA, "Global EV Outlook 2021," IEA, Paris, Report, 2021. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2021 [Accessed: Jan. 19, 2022].

[10] IEA, "World energy outlook 2021," IEA, Paris, Report, 2021. [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2021 [Accessed: Jan. 19, 2022].

[11] J. Eggleston, C. Zuur, and P. Mancarella, "From security to resilience: Technical and regulatory options to manage extreme events in low-carbon grids," *IEEE Power and Energy Magazine*, vol. 19, no. 5, pp. 67–75, 2021.

[12] M. Matos, R. Bessa, A. Botterud, and Z. Zhou, *11 - Forecasting and setting power system operating reserves*. Woodhead Publishing, 2017, pp. 279–308.

[13] North American Electric Reliability Coorporation, "1,200 MW Fault Induced Solar Photovoltaic Resource Interruption Disturbance Report, Southern California 8/16/2016 Event," NERC, Report, 2017.

[14] W. Jahn, J. L. Urban, and G. Rein, "Powerlines and wildfires: Overview, perspectives, and climate change: Could there be more electricity blackouts in the future?" *IEEE Power and Energy Magazine*, vol. 20, no. 1, pp. 16–27, 2022.

[15] WWF and BCG, "Fires forests and the future: A crisis raging out of control," 2020. [Online]. Available: https://wwfeu.awsassets.panda.org/downloads/wwf_fires_forests_and_the_future_report.pdf [Accessed: Feb. 2020].

[16] D. Thomas, D. Butry, S. Gilbert, D. Webb, and J. Fung, "The costs and losses of wildfires," *NIST Special Publication*, vol. 1215, no. 11, 2017.

[17] D. Wang *et al.*, "Economic footprint of california wildfires in 2018," *Nature Sustainability*, vol. 4, no. 3, pp. 252–260, 2021.

[18] J. T. Abatzoglou, C. S. Juang, A. P. Williams, C. A. Kolden, and A. L. Westerling, "Increasing synchronous fire danger in forests of the western united states," *Geophysical Research Letters*, vol. 48, no. 2, p. e2020GL091377, 2021.

[19] S. York, "Smoke from california wildfires decreases solar generation in caiso," Sep. 30, 2020 2020. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=45336 [Accessed: May 5, 2021].

[20] S. D. Wong, J. C. Broader, and S. A. Shaheen, "Review of california wildfire evacuations from 2017 to 2019," UC Office of the President: University of California Institute of Transportation Studies., Report, 2020.

[21] IEEE PES Industry Technical Support Task, "The definition and quantification of resilience," IEEE Power and Energy Society, Report PES-TR65, April 2018.

[22] A. Gholami, T. Shekari, M. H. Amirioun, F. Aminifar, M. H. Amini, and A. Sargolzaei, "Toward a consensus on the definition and taxonomy of power system resilience," *IEEE Access*, vol. 6, pp. 32 035–32 053, 2018.

[23] United Nations Office for Disaster Risk Reduction, "Terminology - resilience," 2022. [Online]. Available: https://www.undrr.org/terminology/resilience [Accessed: January 24, 2022].

[24] North American Transmission Forum, "Transmission resilience overview," 2021. [Online]. Available: https://www.natf.net/docs/natf/documents/resources/resiliency/transmission-resilience-overview.pdf [Accessed: Jan. 2022].

[25] N. S. Diffenbaugh, D. L. Swain, and D. Touma, "Anthropogenic warming has increased drought risk in california," *Proceedings of the National Academy of Sciences*, vol. 112, no. 13, pp. 3931–3936, 2015.

[26] J. T. Abatzoglou, A. P. Williams, and R. Barbero, "Global emergence of anthropogenic climate change in fire weather indices," *Geophysical Research Letters*, vol. 46, no. 1, pp. 326–336, 2019.

[27] K. Tierney and M. Bruneau, "Conceptualizing and measuring resilience: A key to disaster loss reduction," *TR news*, no. 250, 2007.

[28] L. Carlson, G. Bassett, W. Buehring, M. Collins, S. Folga, B. Haffenden, F. Petit, J. Phillips, D. Verner, and R. Whitfield, "Resilience: Theory and applications," ANL/DIS-12-1, Report, 2012.

[29] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017.

[30] S. Skarvelis-Kazakos *et al.*, "Resilience of electric utilities during the covid-19 pandemic in the framework of the cigre definition of power system resilience," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107703, 2022.

[31] N. K. Carrington, I. Dobson, and Z. Wang, "Extracting resilience metrics from distribution utility data using outage and restore process statistics," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5814–5823, 2021.

[32] W. Li, *Risk assessment of power systems: models, methods, and applications.* Hoboken, New Jersey: John Wiley & Sons, 2014.

[33] K. Baker, "Solutions of dc opf are never ac feasible," *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, p. 264–268, 2021.

[34] S. Espinoza, A. Poulos, H. Rudnick, J. C. de la Llera, M. Panteli, and P. Mancarella, "Risk and resilience assessment with component criticality ranking of electric power systems subject to earthquakes," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2837–2848, 2020.

[35] S. LaRocca, J. Johansson, H. Hassel, and S. Guikema, "Topological performance measures as surrogates for physical flow models for risk and vulnerability analysis for electric power systems," *Risk Analysis*, vol. 35, no. 4, pp. 608–623, 2015.

[36] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower's extensible optimal power flow architecture," *2009 IEEE Power & Energy Society General Meeting*, pp. 1–7, 2009.

[37] H. Raoufi, V. Vahidinasab, and K. Mehran, "Power systems resilience metrics: A comprehensive review of challenges and outlook," *Sustainability*, vol. 12, no. 22, 2020.

[38] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2017.

[39] B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic load forecasting via quantile regression averaging on sister forecasts," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 730–737, 2017.

[40] Y. Zhou, M. Panteli, B. Wang, and P. Mancarella, "Quantifying the system-level resilience of thermal power generation to extreme temperatures and water scarcity," *IEEE Systems Journal*, vol. 14, no. 1, pp. 749–759, 2020.

[41] S. Poudel, A. Dubey, and A. Bose, "Risk-based probabilistic quantification of power distribution system operational resilience," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3506–3517, 2020.

[42] M. Shahidehpour and Z. Yamin, Hatim Li, *Market Operations in Electric Power Systems: Forecasting, Scheduling and Risk Management.* New York: John Wiley and Sons, Inc., 2002.

[43] R. Moreno and G. Strbac, "Integrating high impact low probability events in smart distribution network security standards through cvar optimisation," in *IET International Conference on Resilience of Transmission and Distribution Networks (RTDN) 2015*, Conference Proceedings, pp. 1–6.

[44] B. J. Harvey, D. C. Donato, and M. G. Turner, "Burn me twice, shame on who? interactions between successive forest fires across a temperate mountain region," *Ecology*, vol. 97, no. 9, pp. 2272–2282, 2016.

[45] M. Choobineh, B. Ansari, and S. Mohagheghi, "Vulnerability assessment of the power grid against progressing wildfires," *Fire Safety Journal*, vol. 73, pp. 20–28, 2015.

[46] D. N. Trakas and N. D. Hatziargyriou, "Optimal distribution system operation for enhancing resilience against wildfires," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2260–2271, 2018.

[47] M. Choobineh and S. Mohagheghi, "Power grid vulnerability assessment against wildfires using probabilistic progression estimation model," in *2016 IEEE Power and Energy Society General Meeting (PESGM).* IEEE, Conference Proceedings, pp. 1–5.

[48] S. Dian *et al.*, "Integrating wildfires propagation prediction into early warning of electrical transmission line outages," *IEEE Access*, vol. 7, pp. 27 586–27 603, 2019.

[49] T. Tapia, A. Lorca, D. Olivares, M. Negrete-Pincetic, and A. J. Lamadrid, "A robust decision-support method based on optimization and simulation for wildfire resilience in highly renewable power systems," *European Journal of Operational Research*, vol. 294, no. 2, pp. 723–733, 2021.

[50] P. H. Thomas, "Rates of spread of some wind-driven fires," *Forestry: An International Journal of Forest Research*, vol. 44, no. 2, pp. 155–175, 1971.

[51] R. C. Rothermel, *A mathematical model for predicting fire spread in wildland fuels*. Intermountain Forest & Range Experiment Station, Forest Service, US, 1972, vol. 115.

[52] Y. Guo, R. Chen, J. Shi, J. Wan, H. Yi, and J. Zhong, "Determination of the power transmission line ageing failure probability due to the impact of forest fire," *IET Generation, Transmission & Distribution*, vol. 12, no. 16, pp. 3812–3819, 2018.

[53] M. Abdelmalak and M. Benidris, "Enhancing power system operational resilience against wildfires," *IEEE Transactions on Industry Applications*, pp. 1–1, 2022.

[54] M. Nazemi, P. Dehghanian, M. Alhazmi, and Y. Darestani, "Resilient operation of electric power distribution grids under progressive wildfires," *IEEE Transactions on Industry Applications*, pp. 1–1, 2022.

[55] A. Alexandridis, D. Vakalis, C. I. Siettos, and G. V. Bafas, "A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990," *Applied Mathematics and Computation*, vol. 204, no. 1, pp. 191–201, 2008.

[56] J. G. Freire and C. C. DaCamara, "Using cellular automata to simulate wildfire propagation and to assist in fire management," *Nat. Hazards Earth Syst. Sci.*, vol. 19, no. 1, pp. 169–179, 2019.

[57] A. C. Scott, D. M. J. S. Bowman, W. J. Bond, S. J. Pyne, and M. E. Alexander, *Fire on Earth : An Introduction*. Hoboken, UNITED KINGDOM: John Wiley & Sons, Incorporated, 2014.

[58] H. Wan, J. D. McCalley, and V. Vittal, "Increasing thermal rating by risk analysis," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 815–828, 1999.

[59] B. W. Butler, J. Webb, J. Hogge, and T. Wallace, "Vegetation clearance distances to prevent wildland fire caused damage to telecommunication and power transmission infrastructure." in *Proceedings of the large wildland fires conference;*, R. E. Keane, M. Jolly, R. Parsons, and K. Riley, Eds. Missoula, MT.: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Conference Proceedings, pp. 35–40.

[60] J. A. Sathaye *et al.*, "Rising temps, tides, and wildfires: Assessing the risk to california's energy infrastructure from projected climate change," *IEEE Power and Energy Magazine*, vol. 11, no. 3, pp. 32–45, 2013.

[61] L. Cui, R. S. Gorur, and D. Chipman, "Evaluating flashover performance of insulators under fire fighting conditions," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 2, pp. 1051–1056, 2017.

[62] S. Jazebi, F. de León, and A. Nelson, "Review of wildfire management techniques—part i: Causes, prevention, detection, suppression, and data analytics," *IEEE Transactions on Power Delivery*, vol. 35, no. 1, pp. 430–439, 2020.

[63] Federal Emergency Management Agency (FEMA), "Hazus user and technical manuals." 2021. [Online]. Available: https://www.fema.gov/flood-maps/tools-resources/flood-map-products/hazus/user-technical-manuals [Accessed: Jan. 28, 2022].

[64] Y. Liu, B. Li, C. Wu, B. Chen, and T. Zhou, "Risk warning technology for the whole process of overhead transmission line trip caused by wildfire," *Natural Hazards*, vol. 107, no. 1, pp. 195–212, 2021.

[65] IPCC, "Glossary of terms," in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC).*, C. Field *et al.*, Eds. Cambridge, UK and New York, USA: Cambridge University Press, 2012, pp. 555–564.

[66] A. Arab, A. Khodaei, R. Eskandarpour, M. P. Thompson, and M. Wei, "Three lines of defense for wildfire risk management in electric power grids: A review," *IEEE Access*, vol. 9, pp. 61 577–61 593, 2021.

[67] J. W. Muhs, M. Parvania, and M. Shahidehpour, "Wildfire risk mitigation: A paradigm shift in power systems planning and operation," *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 366–375, 2020.

[68] M. Davoudi, B. Efaw, M. Avendaño-Mora, J. L. Lauletta, and G. B. Huffman, "Reclosing of distribution systems for wildfire prevention," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2298–2307, 2021.

[69] A. Toffler, *The Third Wave*. Pan in association with Collins, 1981.

[70] A. Gautier, J. Jacqmin, and J.-C. Poudou, "The prosumers and the grid," *Journal of Regulatory Economics*, vol. 53, no. 1, pp. 100–126, 2018.

[71] D. Stowell *et al.*, "A harmonised, high-coverage, open dataset of solar photovoltaic installations in the uk," *Scientific Data*, vol. 7, no. 1, p. 394, 2020.

[72] Energy Networks Association, "Distributed generation connection guides: G98 for single premises full version," Version 3.0, Report, November 2021.

[73] X. Zhang and S. Grijalva, "A data-driven approach for detection and estimation of residential pv installations," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2477–2485, 2016.

[74] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[75] J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied Energy*, vol. 183, pp. 229–240, 2016.

[76] J. M. Malof, L. M. Collins, and K. Bradbury, "A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Conference Proceedings, pp. 874–877.

[77] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[78] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[79] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, "Deepsolar: A machine learning framework to efficiently construct a solar deployment database in the united states," *Joule*, vol. 2, no. 12, pp. 2605–2617, 2018.

[80] K. Mayer, Z. Wang, M. L. Arlt, D. Neumann, and R. Rajagopal, "Deepsolar for germany: A deep learning framework for pv system mapping from aerial imagery," in *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, Conference Proceedings, pp. 1–6.

[81] R. Wang, J. Camilo, L. M. Collins, K. Bradbury, and J. M. Malof, "The poor generalization of deep convolutional networks to aerial imagery from new geographic locations: an empirical study with solar array detection," in *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Conference Proceedings, pp. 1–8.

[82] F. Wang, K. Li, X. Wang, L. Jiang, J. Ren, Z. Mi, M. Shafie-khah, and J. P. S. Catalão, "A distributed pv system capacity estimation approach based on support vector machine with customer net load curve features," *Energies*, vol. 11, no. 7, 2018.

[83] C. M. Cheung, W. Zhong, C. Xiong, A. Srivastava, R. Kannan, and V. K. Prasanna, "Behind-the-meter solar generation disaggregation using consumer mixture models," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Conference Proceedings, pp. 1–6.

[84] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2012.

[85] D. Arthur and S. Vassilvitskii, "k-means++: the advantages of careful seeding," *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, p. 1027–1035, 2007.

[86] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963.

[87] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[88] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, "Load pattern-based classification of electricity customers," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1232–1239, 2004.

[89] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra, "Dimensionality reduction for fast similarity search in large time series databases," *Knowledge and Information Systems*, vol. 3, no. 3, pp. 263–286, 2001.

[90] R. Seguin, J. Woyak, D. Costyk, J. Hambrick, and B. Mather, "High-penetration pv integration handbook for distribution engineers," Report, 2016.

[91] S.-E. Razavi *et al.*, "Impact of distributed generation on protection and voltage regulation of distribution systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 157–167, 2019.

[92] A. M. Tsimtsios and V. C. Nikolaidis, "Towards plug-and-play protection for meshed distribution systems with dg," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 1980–1995, 2020.

[93] A. Tuohy *et al.*, "Solar forecasting: Methods, challenges, and performance," *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 50–59, 2015.

[94] E. Hildebrandt *et al.*, "2020 annual report on market issues and performance," Department of Market Monitoring – California ISO, Report, 2021. [Online]. Available: http://www.caiso.com/Documents/2020-Annual-Report-on-Market-Issues-and-Performance.pdf [Accessed: Jan. 2022].

[95] Australian Energy Market Operator (AEMO), "Integrated System Plan (ISP) Methodology," AEMO, Report, 2021. [Online]. Available: https://aemo.com.au/-/media/files/major-publications/isp/2021/2021-isp-methodology.pdf?la=en [Accessed: Jan. 27, 2022].

[96] R. E. Pawluk, Y. Chen, and Y. She, "Photovoltaic electricity generation loss due to snow – a literature review on influence factors, estimation, and mitigation," *Renewable and Sustainable Energy Reviews*, vol. 107, pp. 171–182, 2019.

[97] N. Heidari, J. Gwamuri, T. Townsend, and J. M. Pearce, "Impact of snow and ground interference on photovoltaic electric system performance," *IEEE Journal of Photovoltaics*, vol. 5, no. 6, pp. 1680–1685, 2015.

[98] P. Gilman, A. Dobos, N. DiOrio, J. Freeman, S. Janzou, and D. Ryberg, "Sam photovoltaic model technical reference update." NREL/TP-6A20-67399., Report.

[99] D. Rieger, A. Steiner, V. Bachmann, P. Gasch, J. Förstner, K. Deetz, B. Vogel, and H. Vogel, "Impact of the 4 april 2014 saharan dust outbreak on the photovoltaic power generation in germany," *Atmos. Chem. Phys.*, vol. 17, no. 21, pp. 13 391–13 415, 2017, aCP.

[100] I. Neher, T. Buchmann, S. Crewell, B. Pospichal, and S. Meilinger, "Impact of atmospheric aerosols on solar power," *Meteorologische Zeitschrift*, vol. 28, no. 4, pp. 305–321, 2019.

[101] J. X. L. Wang, "Mapping the global dust storm records: Review of dust data sources in supporting modeling/climate study," *Current Pollution Reports*, vol. 1, no. 2, pp. 82–94, 2015.

[102] X. Li, F. Wagner, W. Peng, J. Yang, and D. L. Mauzerall, "Reduction of solar photovoltaic resources due to air pollution in china," *Proceedings of the National Academy of Sciences*, vol. 114, no. 45, pp. 11 867–11 872, 2017.

[103] I. M. Peters, S. Karthik, H. Liu, T. Buonassisi, and A. Nobre, "Urban haze and photovoltaics," *Energy & Environmental Science*, vol. 11, no. 10, pp. 3043–3054, 2018.

[104] CAL FIRE, "Top 20 largest california wildfires," 2021. [Online]. Available: https://www.fire.ca.gov/media/4jandlhh/top20_acres.pdf [Accessed: October 25, 2021].

[105] M. Perry and A. Troccoli, "Impact of a fire burn on solar irradiance and pv power," *Solar Energy*, vol. 114, pp. 167–173, 2015.

[106] S. D. Gilletly, N. D. Jackson, and A. Staid, "Quantifying wildfire-induced impacts to photovoltaic energy production in the western united states," in *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*, Conference Proceedings, pp. 1619–1625.

[107] R. J. Hyndman and G. Athanasopoulos, *Forecasting Principles and Practice*. OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/ [Accessed: Nov. 15, 2021].

[108] NOAA Global Monitoring Laboratory, "SURFRAD Aerosol Optical Depth," 2021. [Online]. Available: https://gml.noaa.gov/grad/surfrad/aod/ [Accessed: Feb. 27, 2021].

[109] M. Carlomusto, *GOES-R Series Product Definition and Users' Guide (PUG)*. Harris Corp., US DOC, NOAA, NESDIS, NASA, Dec. 2019, vol. 5, no. 416-R-PUG-L2 Plus-0349, Revision 2.2,.

[110] NOAA/NESDIS, "Office of satellite and product operations - hazard mapping system." [Online]. Available: https://www.ospo.noaa.gov/Products/land/hms.html#about [Accessed: Dec. 15, 2021].

[111] M. Pagani, W. Korosec, N. Chokani, and R. S. Abhari, "User behaviour and electric vehicle charging infrastructure: An agent-based model assessment," *Applied Energy*, vol. 254, p. 113680, 2019.

[112] U.S. Department of Energy, "Alternative Fuels Data Center: Electric Vehicle Infrastructure Projection Tool (EVI-Pro) Lite," 2021. [Online]. Available: https://afdc.energy.gov/evi-pro-lite/load-profile [Accessed: Sep. 1, 2021].

[113] A. Navarro-Espinosa and L. F. Ochoa, "Probabilistic impact assessment of low carbon technologies in lv distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2192–2203, 2016.

[114] Electric Nation, "Powered up - charging evs without stressing the electricity network," 2019. [Online]. Available: http://www.electricnation.org.uk/wp-content/

uploads/2019/10/Electric-Nation-Powered-Up-Report-WEB.pdf [Accessed: Aug. 25, 2021].

[115] The Florida Department of Transportation (FDOT), "Hurricane irma's effect on florida's fuel distribution system and recommended improvements," Report, 2018. [Online]. Available: https://www.fdot.gov/docs/default-source/info/CO/news/newsreleases/020118-FDOT-Fuel-Report.pdf [Accessed: Dec. 4, 2021].

[116] S. A. Adderly, D. Manukian, T. D. Sullivan, and M. Son, "Electric vehicles and natural disaster policy implications," *Energy Policy*, vol. 112, pp. 437–448, 2018.

[117] C. D. MacDonald, L. Kattan, and D. Layzell, "Modelling electric vehicle charging network capacity and performance during short-notice evacuations," *International Journal of Disaster Risk Reduction*, vol. 56, p. 102093, 2021.

[118] V. C. Radeloff *et al.*, "Rapid growth of the us wildland-urban interface raises wildfire risk," *Proceedings of the National Academy of Sciences*, vol. 115, no. 13, pp. 3314–3319, 2018.

[119] K. Feng, N. Lin, S. Xian, and M. V. Chester, "Can we evacuate from hurricanes with electric vehicles?" *Transportation Research Part D: Transport and Environment*, vol. 86, p. 102458, 2020.

[120] A. P. Williams *et al.*, "Observed impacts of anthropogenic climate change on wildfire in california," *Earth's Future*, vol. 7, no. 8, pp. 892–910, 2019.

[121] North American Electric Reliability Coorporation, "Transmission system planning performance requirements - TPL-001-4," 2014.

[122] S. Li and T. Banerjee, "Spatial and temporal pattern of wildfires in california from 2000 to 2019," *Scientific Reports*, vol. 11, no. 1, p. 8779, 2021.

[123] J. Burkardt, "The truncated normal distribution," 2014. [Online]. Available: https://people.sc.fsu.edu/~jburkardt/presentations/truncated_normal.pdf [Accessed: Jan. 12, 2022].

[124] E. Kuffel, W. S. Zaengl, and J. Kuffel, *High Voltage Engineering Fundamentals*, 2nd ed. Newnes, 2000.

[125] P. Li *et al.*, "Influence of forest fire particles on the breakdown characteristics of air gap," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 4, pp. 1974–1984, 2016.

[126] T. Wu, J. Ruan, C. Chen, and D. Huang, "Field observation and experimental investigation on breakdown of air gap of ac transmission line under forest fires," in *2011 IEEE Power Engineering and Automation Conference*, vol. 2, Conference Proceedings, pp. 339–343.

[127] Z. Pu, J. Ruan, D. Huang, T. Wu, and P. Li, "Study on the breakdown characteristics of the transmission line gap under forest fire conditions," *International Transactions on Electrical Energy Systems*, vol. 25, no. 11, pp. 2731–2744, 2015.

[128] North American Electric Reliability Coorporation, "Glossary of terms used in nerc reliability standards," 2021. [Online]. Available: https://www.nerc.com/files/glossary_of_terms.pdf [Accessed: Feb. 3, 2022].

[129] J. R. Fonseca *et al.*, "Effects of agricultural fires on the performance of overhead transmission lines," *IEEE Transactions on Power Delivery*, vol. 5, no. 2, pp. 687–694, 1990.

[130] California Independent System Operator, "System conditions bulletin," 2021. [Online]. Available: http://www.caiso.com/Documents/SystemConditionsBulletin.pdf [Accessed: Nov. 15, 2021].

[131] R. Billinton and W. Li, *Reliability Assessment of Electric Power Systems Using Monte Carlo Methods*. New York: Springer Science+Business Media, LLC, 1994.

[132] C. Barrows *et al.*, "The ieee reliability test system: A proposed 2019 update," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 119–127, 2020.

[133] N. Andela *et al.*, "The global fire atlas of individual fire size, duration, speed and direction," *Earth Syst. Sci. Data*, vol. 11, no. 2, pp. 529–552, 2019, eSSD.

[134] N. Andela, D. Morton, L. Giglio, and J. Randerson, *Global Fire Atlas with Characteristics of Individual Fires, 2003-2016*. ORNL Distributed Active Archive Center, 2019. [Online]. Available: 10.3334/ORNLDAAC/1642

[135] California Department of Forestry and Fire Protection, "Fire perimeters through 2020," 2021. [Online]. Available: https://frap.fire.ca.gov/mapping/gis-data/ [Accessed: Sep. 23, 2021].

[136] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The numpy array: A structure for efficient numerical computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

[137] W. McKinney, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, vol. 445. Austin, TX, Conference Proceedings, pp. 51–56.

[138] K. Jordahl *et al.*, "geopandas/geopandas v0.8.1," 2020.

[139] S. Gillies *et al.*, "Rasterio: geospatial raster i/o for Python programmers," Mapbox, 2013–. [Online]. Available: https://github.com/rasterio/rasterio

[140] S. Hoyer and J. Hamman, "xarray: N-d labeled arrays and datasets in python," *Journal of Open Research Software*, vol. 5, no. 1, p. 10, 2017.

[141] M. Office, "Cartopy 18.0: a cartographic python library with a matplotlib interface," 2020. [Online]. Available: https://scitools.org.uk/cartopy/docs/latest/

[142] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[143] M. Waskom *et al.*, "mwaskom/seaborn: v0.9.0 (july 2018)," 2018.

[144] P. Virtanen *et al.*, "Scipy 1.0: fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020.

[145] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[146] H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185–1193, 2007.

[147] US Census Bureau, "Cartographic boundary files," 2018. [Online]. Available: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html

[148] L. Dale, M. Carnall, M. Wei, G. Fitts, and S. Lewis McDonald, "Assessing the impact of wildfires on the california electricity grid," Lawrence Berkeley National Laboratory, California's Fourth Climate Change Assessment. California Energy Commission, Report CCCA4-CEC-2018-002, 2018.

[149] D. A. Zuniga Vazquez, F. Qiu, N. Fan, and K. Sharp, "Wildfire mitigation plans in power systems: A literature review," *IEEE Transactions on Power Systems*, pp. 1–1, 2022.

[150] D. Sharafi *et al.*, "Wildfires down under: Impacts and mitigation strategies for australian electricity grids," *IEEE Power and Energy Magazine*, vol. 20, no. 1, pp. 52–63, 2022.

[151] B. Chiu, R. Roy, and T. Tran, "Wildfire resiliency: California case for change," *IEEE Power and Energy Magazine*, vol. 20, no. 1, pp. 28–37, 2022.

[152] R. Moreno *et al.*, "Microgrids against wildfires: Distributed energy resources enhance system resilience," *IEEE Power and Energy Magazine*, vol. 20, no. 1, pp. 78–89, 2022.

[153] Pecan Street Inc, "Pecan street inc. dataport," 2019. [Online]. Available: https://www.pecanstreet.org/about/ [Accessed: 2019].

[154] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[155] G. Vettigli, "Minisom: minimalistic and numpy-based implementation of the self organizing map," 2018. [Online]. Available: https://github.com/JustGlowing/minisom

[156] A. Lindstrom and S. Hoff, "Investor-owned utilities served 72% of U.S. electricity customers in 2017," 2019. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=40913 [Accessed: Nov. 29, 2021].

[157] UK Power Networks, "UK Power Networks Annual Review 2020/21," 2022. [Online]. Available: https://annualreview2021.ukpowernetworks.co.uk/ [Accessed: Feb. 2022].

[158] Pacific Gas and Electric, "Pacific gas and electric company 2020 wildfire mitigation plan report updated," Report, 2020. [Online]. Available: https://www.pge.com/pge_global/common/pdfs/safety/emergency-preparedness/natural-disaster/wildfires/wildfire-mitigation-plan/2020-Wildfire-Safety-Plan.pdf [Accessed: Sep. 15, 2021].

[159] E. C. Venizelos, D. N. Trakas, and N. D. Hatziargyriou, "Distribution system resilience enhancement under disastrous conditions by splitting into self-sufficient microgrids," in *2017 IEEE Manchester PowerTech*, Conference Proceedings, pp. 1–6.

[160] J. A. Kratochvil, W. E. Boyson, and D. L. King, "Photovoltaic array performance model," Report, 2004-08-01 2004.

[161] Copernicus Atmosphere Monitoring Service (CAMS), "Global forecast plots," 2021. [Online]. Available: https://atmosphere.copernicus.eu/index.php/global-forecast-plots [Accessed: Nov. 10, 2021].

[162] J. L. Gómez-Amo *et al.*, "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*, vol. 235, pp. 1226–1234, 2019.

[163] NOAA, "Registry of Open Data on AWS - NOAA Geostationary Operational Environmental Satellites (GOES) 16 & 17." [Online]. Available: https://registry.opendata.aws/noaa-goes/ [Accessed: 2020].

[164] S. Dubey, J. N. Sarvaiya, and B. Seshadri, "Temperature dependent photovoltaic (pv) efficiency and its effect on pv production in the world – a review," *Energy Procedia*, vol. 33, pp. 311–321, 2013.

[165] H.-D. Mohring *et al.*, "Outdoor performance of polycrystalline thin film pv modules in different european climates," in *19th European PV Solar Energy Conference and Exhibition*, Conference Proceedings.

[166] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, 2018.

[167] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[168] C. Davino, *Quantile regression : theory and applications / Cristina Davino, Marilena Furno, Domenico Vistocco.* Chichester, England : Wiley, 2014., 2014.

[169] R. Koenker and K. F. Hallock, "Quantile regression," *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 143–156, 2001.

[170] CAL FIRE, "2020 fire season," 2021. [Online]. Available: https://www.fire.ca.gov/incidents/2020/ [Accessed: Nov. 10, 2021].

[171] National Wildfire Coordinating Group, "Creek fire information - inciweb the incident information system," 2021. [Online]. Available: https://inciweb.nwcg.gov/incident/7147/ [Accessed: January 2021].

[172] National Wildfire Coordinating Group, "El dorado fire information - inciweb the incident information system," 2021. [Online]. Available: https://inciweb.nwcg.gov/incident/7148/ [Accessed: January 2021].

[173] National Wildfire Coordinating Group, "Bobcat information - inciweb the incident information system," 2021. [Online]. Available: https://inciweb.nwcg.gov/incident/7152/ [Accessed: Jan. 2021].

[174] C. F. Jekel and G. Venter, "pwlf: A python library for fitting 1d continuous piecewise linear functions," 2019. [Online]. Available: https://github.com/cjekel/piecewise_linear_fit_py

[175] National Interagency Fire Center, "2020 perimeters to date - public nifs archive perimeters," 2021. [Online]. Available: https://data-nifc.opendata.arcgis.com/ [Accessed: Jan. 2021].

[176] Global Modeling and Assimilation Office, *MERRA-2 tavg1_2D_slv_Nx: 2D, 1-hourly, time-averaged, single-level, assimilation, single-level diagnostics*. Goddard Earth Sci. Data Inf. Services Center, 2015.

[177] A. Huff and R. Theurer, "Iss3 workshop on new generation satellite products for operational fire and smoke applications," 2020. [Online]. Available: https://www.star.nesdis.noaa.gov/smcd/spb/aq/STAR_Python_Hub/index.php [Accessed: January 2021].

[178] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with python," in *Proceedings of the 9th Python in Science Conference*, vol. 57. Austin, TX, Conference Proceedings, p. 61.

[179] T. W. Juliano, P. A. Jiménez, B. Kosović, T. Eidhammer, G. Thompson, L. K. Berg, J. Fast, A. Motley, and A. Polidori, "Smoke from 2020 united states wildfires responsible for substantial solar energy forecast errors," *Environmental Research Letters*, vol. 17, no. 3, p. 034010, feb 2022.

[180] S. W. Tweedie, J. R. Rowland, S. J. Walsh, R. P. Rhoten, and P. I. Hagle, "A methodology for estimating emergency evacuation times," *The Social Science Journal*, vol. 23, no. 2, pp. 189–204, 1986.

[181] K. Qian, C. Zhou, M. Allan, and Y. Yuan, "Modeling of load demand due to ev battery charging in distribution systems," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 802–810, 2011.

[182] A. M. A. Haidar and K. M. Muttaqi, "Behavioral characterization of electric vehicle charging loads in a distribution power grid through modeling of battery chargers," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, pp. 483–492, 2016.

[183] D. Jayaweera, "Security enhancement with nodal criticality-based integration of strategic micro grids," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 337–345, 2015.

[184] P. Garvey and Z. Lansdowne, "Risk matrix: An approach for identifying, assessing, and ranking program risks." *Air Force Journal of Logistics*, vol. 22, pp. 18–21, 1998.

[185] California Department of Forestry and Protection, "Welcome to stats and events," 2021. [Online]. Available: https://www.fire.ca.gov/stats-events/ [Accessed: Dec. 15, 2021].

[186] California Independent System Operator, "California iso peak load history 1998 through 2020," 2021. [Online]. Available: https://www.caiso.com/Documents/CaliforniaISOPeakLoadHistory.pdf [Accessed: Dec. 15, 2021].

[187] J. Quirós-Tortós, A. N. Espinosa, L. F. Ochoa, and T. Butler, "Statistical representation of ev charging: Real data analysis and applications," in *2018 Power Systems Computation Conference (PSCC)*, Conference Proceedings, pp. 1–7.

[188] M. K. Lindell and C. S. Prater, "Critical behavioral assumptions in evacuation time estimate analysis for private vehicles: Examples from hurricane research and planning," *Journal of Urban Planning and Development*, vol. 133, no. 1, pp. 18–29, 2007.

[189] P. Mozumder, N. Raheem, J. Talberth, and R. P. Berrens, "Investigating intended evacuation from wildfires in the wildland–urban interface: Application of a bivariate probit model," *Forest Policy and Economics*, vol. 10, no. 6, pp. 415–423, 2008.

[190] M.-S. Răboacă, I. Băncescu, V. Preda, and N. Bizon, "An optimization model for the temporary locations of mobile charging stations," *Mathematics*, vol. 8, no. 3, 2020.

[191] F. K. Rengers, L. A. McGuire, N. S. Oakley, J. W. Kean, D. M. Staley, and H. Tang, "Landslides after wildfire: Initiation, magnitude, and mobility," *Landslides*, vol. 17, no. 11, pp. 2631–2641, 2020.

[192] R. Ghorani, S. Fattaheian-Dehkordi, M. Farrokhi, M. Fotuhi-Firuzabad, and M. Lehtonen, "Modeling and quantification of power system resilience to natural hazards: A case of landslide," *IEEE Access*, vol. 9, pp. 80 300–80 309, 2021.

[193] K. Pan, Z. Chen, C. S. Lai, C. Xie, D. Wang, Z. Zhao, X. Wu, N. Tong, L. L. Lai, and N. D. Hatziargyriou, "A novel data-driven method for behind-the-meter solar generation disaggregation with cross-iteration refinement," *IEEE Transactions on Smart Grid*, pp. 1–1, 2022.

[194] F. Wang, X. Ge, Z. Dong, J. Yan, K. Li, F. Xu, X. Lu, H. Shen, and P. Tao, "Joint energy disaggregation of behind-the-meter pv and battery storage: A contextually supervised source separation approach," *IEEE Transactions on Industry Applications*, vol. 58, no. 2, pp. 1490–1501, 2022.

[195] X. Li, H. Xu, X. Huang, C. Guo, Y. Kang, and X. Ye, "Emerging geo-data sources to reveal human mobility dynamics during covid-19 pandemic: opportunities and challenges," *Computational Urban Science*, vol. 1, no. 1, p. 22, 2021.

[196] E. D. Kostopoulos, G. C. Spyropoulos, and J. K. Kaldellis, "Real-world study for the optimal charging of electric vehicles," *Energy Reports*, vol. 6, pp. 418–426, 2020.

[197] NREL, "Github - evi-pro-lite," 2021. [Online]. Available: https://github.com/NREL/EVI-Pro-Lite [Accessed: Sep. 1, 2021].

[198] United States Census Bureau, "American Community Survey 5-year Estimates Table DP04," 2018.