International Journal of Energy Engineering 2020, 10(3): 80-89
DOI: 10.5923/j.ijee.20201003.02

A Comprehensive Forecasting, Risk Modelling and Optimization Framework for Electric Grid Hardening and Wildfire Prevention in the US

Soham Ghosh[1,*], Sreejata Dutta[2]

[1]School of Engineering, University of Kansas, Overland Park, USA
[2]Department of Biostatistics & Data Science, University of Kansas, Kansas City, USA

Abstract Historical wildfire patterns have experienced a recent shift in terms of its scale and intensity. Through the continuing advancements in electrical protection technology, statistical forecasting methodologies, availability of meteorological field data, and regional risk-modelling, wildfire management practices can be made more proactive in the United States and around the globe. To create a comprehensive and practical operating framework, an advanced seasonal autoregressive integrated moving average time series modelling technique for wildfire forecasting is explored. These regressive models, due to their mathematical accuracy has been used in many engineering and scientific applications. The study presented here was done using a qualitative investigation approach to wildfire data. Computer automated grid search techniques were developed to determine suitable seasonal regressive model hyper-parameters. With the usage of power transforms to fit skewed statistical models under study, it is found that a much more accurate and computationally efficient model can be generated. Statistical forecasts and regional risk mapping techniques can influence strategic operational practices for regional and local fire authorities. Concepts that can enhance power system protection and electrical grid hardening are explored and practical guidelines to help electrical utilities improve electrical grid operations are provided. Many benefits of using distributed energy resources are discussed and an optimal power flow involving these resources is formulated to help grid operators preserve system stability under these wildfire scenarios.

Keywords Grid maintenance planning, SARIMA, Auto-regressive models, Forecasting, Wildland-urban interface, Risk-model, Distributed energy resources

1. Introduction

Effective wildfire policies are of paramount importance for natural resource management and to ensure public health and safety. In the US, historical patterns of these fires were mostly caused due to lightning or burning by the native population. However, recently these historical patterns have shown significant changes in terms of the size, intensity, and duration of the wildfires [1,2]. The changes in the historical patterns are mostly attributed to a synergy of natural and human causes. Changes in rainfall patterns and snowpack, increased periods of drought, and replacement of native grasslands with invasive species are rapidly increasing the potential of future wildfires. To add to the problem, population growth is pushing residential development into forested and semi forested areas which are recurrent spots for wildfires [3–5]. One key issue of these wildland-urban developments is the growing need for electrification of these areas with transmission, distribution, and substation assets. The inherent risk of fires due to electrical faults and other maintenance issues of these assets makes the management of the wildfires in these urban fringe areas challenging and costly. For instance, from 2000-2016, the electrical asset-related damages for utilities in California is in upwards of $700 million [6] mostly reported along these wildlife-urban fringe areas. Understanding and statistical forecasting of these wildfire time-series trends can demonstrate a pathway for future energy policies and utility level decision making on electrical maintenance and planning.

A risk event frequency study [7] from 2015-2017 over the high fire threat districts (HFTD) of California, US has revealed the key contributing element for over four hundred wildfire events as summarized in Table 1. As one can observe, contact of live conductors with vegetation, and equipment and hardware failures are reasons that are

* Corresponding author:
sghosh27@ku.edu (Soham Ghosh)
Published online at http://journal.sapub.org/ijee

potentially controllable with human interference yet have contributed greatly to the cause of massive wildfires. From this observation greater focus has been given to the role of vegetation in increasing forest fire risk and preventive maintenance solutions have been likewise investigated.

A vast amount of literature has been focused on the application of autoregressive integrated moving average (ARIMA) models or somewhat of its improved form in areas ranging from wind speed forecasting [8], to commercial commodities and electricity price forecasting [9–11]. On the contrary, most of the studies that have been done in the area of wildfire prediction have been from a statistical standpoint, without any traceable link to future energy policy recommendations. For instance, in reference [12], a comparison has been made between the relative performance of an ARIMA approach versus a Holt-Winter (HW) approach, for forecasting meteorological factors conducive for forest fires. The authors of this study have established a causal relationship between forest fires and meteorological parameters as wind speed, ambient temperature, precipitation rate, ambient pressure, and relative humidity.

In another interesting study [13], the combination of meteorological and remote sensing satellite data was used to detect potential terrestrial hot-spots, a precursor of most forest fires. A series of hot-spots were observed from a combination of sources and a time series hot-spot model was created. As such, the time-series forecasting model in this study applies the Box-Jenkins approach to build regressive models. The authors in this study investigate the different models and their corresponding information criterion scores. Relative model performance was investigated by dividing the data set into training and testing data and mean absolute error performance metric was used to evaluate overall model performance.

Table 1. Key contributing elements for wildfire events under study

Factors	Percentage contribution	Nature of the event
Vegetation contact	49	Contact of line conductor with vegetation
Equipment or other failure	27	Failure of equipment like power transformers, conductors
Other inadvertent contact	13	Accidents/ sabotage
Animal contact	8	Bird flights and burrowing animal contact
Unknown	2	Causes of unknown origin
Fuse operation	1	Fire ignition from fuse blowout

Usage of remote sensing and geographic information system technologies was outlined in [14] for a case study in Hubei province in China. The study implemented historical ignition points along the power line corridor of Hubei province from 2009-2014 as its training set and used logistical regression to produce ignition probability maps. The conclusion found from this study confirms high dependence of wildfire behaviour on the type of ground vegetation, the dynamic and challenging nature of the input variables as multi-temporal satellite imaging and the importance of accurate meteorological data and field stations for data acquisition.

The authors in [15] used a combination of data science and big data techniques for a wildfire emergency management plan. The wildfire monitoring and emergency management plan tie meaningful data from people on the ground and earth-observing satellites in near-earth orbit for an early fire warning system. Some of the other decision inputs discussed by the authors are data from the weather forecast, visible-infrared earth mapping and risk impact mapping, with the final goal of providing a continuous stream of information both to the authorities and the public.

A theoretical stochastic study as in [16] has assessed the vulnerability of power grids using a probabilistic progression estimation model. The underlying idea presented is the de-rating of the at-risk transmission line to counter the heat gained from the fire plume. Effect of wind direction and speed has been factored in to develop a probabilistic optimal power flow subject to constraints as voltage, generator capacity, and spinning reserve.

From the literature review, a clear research gap is found that no research study has provided a satisfactory framework to analyse the risk and trends of these wildfires and related it to fire mitigation strategies that utilities, fire departments, and distributed energy resource owners can collectively implement in order to slash these fire risks dramatically. This paper establishes a statistical seasonal autoregressive integrated moving average (SARIMA) prediction structure and a regional fire risk modelling-framework to understand the implication of these wildfires on energy policies and is organized as follows. Section 2 focuses on statistical forecasting and develops a comparison between an automated grid search approach for selection of the SARIMA model hyperparameters versus power transformation on the data prior to fitting the model. Merits of using power transform on similarly skewed data set has been presented. Section 3 presents the numerical testing results obtained from Section 2. Finally, Section 4 is inspired by the high importance of the domain of wildfire predictions on governing future energy policies and is approached from three different angles: from the standpoint of local and regional fire departments, from the standpoint of transmission and distribution asset owners, and from the standpoint of distributed energy system asset owners. A DC optimal power flow formulation have been proposed in this section, that combines de-rated transmission line operation, and dispatch support from distributed energy systems. Finally, conclusions are given in Section 5.

2. Automated Grid Search and the Use of Power Transforms

2.1. Data Acquisition and Pre-processing

The study employs a time series data spanning from 1992 to 2015, collected out of Forest Service Research Data Archive, United States Department of Agriculture [17]. The wildfire records of this data set were acquired from federal, state and local fire organizations. The available data-set has over one million forty-eight thousand entries with the granularity at a day level. From the given data-set, the dates at which a given wildfire was observed and the size of the burnt area are the primary area of interest for this study. Most commonly available data-sets suffer from random missing observations. The common approach to deal with random missing observations is to insert an estimated value. For this study, the weighted average of p observation points before and after the missing values has been used, as shown in Equation 1. Furthermore, for a discussion on observations with systematic missing patterns [18] provides an excellent reference.

$$\hat{z}_t = -d_0^{-1} \sum_{j=1}^{p} d_j \left(z_{T-j} + z_{T+j} \right) \quad (1)$$

where $d_j = \sum_{i=j}^{p} \emptyset_i \emptyset_{i-j}, \emptyset_0 = -1$, and $d_0 = 1 + \sum_{i=1}^{p} \emptyset_i^2$.

Following the data pre-processing steps, some descriptive statistical tools were used to visualize the data. Figure 1 shows a box and whisker plot of the monthly wildfire trend, with a clear indication of high median count for the summer and early fall season, and a near-zero median count during the winter season. This observation would be of some importance for the structuring the energy policy recommendations later in Section 4 of this study.

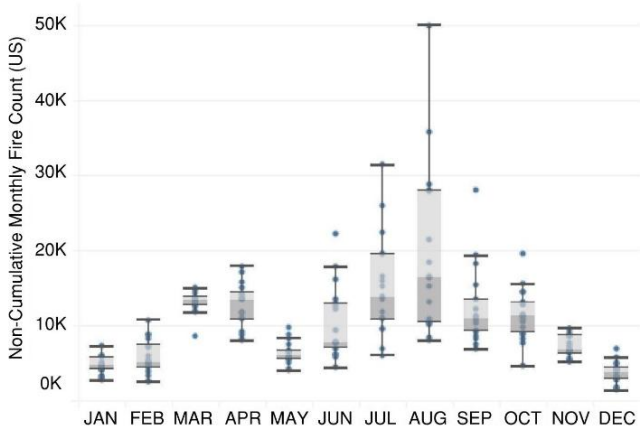


Figure 1. Box and whisker plot summary of the US monthly wildfire count

2.2. Automated Grid Search Approach to Determine SARIMA Model Hyperparameters

The building blocks of an ARIMA (p,d,q) model are the autoregressive order p: AR (p) model and moving average order q: MA (q) model. An AR (p) model is a discrete time series linear equation, as described by Equation 2, where p is the model order and α_i are the coefficients with ε_t as error term.

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \varepsilon_t \quad (2)$$

With L defined as a lag operator $L^k x_t = x_{t-k}$, we can generalize AR (p) model from Equation 2 as shown in Equation 3.

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) X_t = \varepsilon_t \quad (3)$$

A MA (q) model is a linear regression of the current value of a time series with current and previous random shocks, which are assumed to be mutually independent. This can be generalized into Equation 4 with L being a lag operator as defined earlier.

$$X_t = \left(1 + \sum_{k=1}^{r} \beta_k L^k\right) \varepsilon_t \quad (4)$$

An ARIMA (p,d,q) model, a particular case of ARMA (p,q) model, is a discrete time linear equation, as shown in Equation 5, where p, d, and q are the ARIMA model orders, α and β are the real number coefficients, ε_t as error term.

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) (1 - L)^d X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t \quad (5)$$

The classical approach [19] of identifying the ARIMA model hyper-parameters (p,d,q) involves performing stationary tests as augmented Dickey-Fuller (ADF) or Kwiatkowski–Phillips–Schmidt–Shin (KPSS) [20,21] with an application of necessary differencing to make the series stationary. The d term in ARIMA (p,d,q) indicates the order of differencing required to make the series stationary. Further, an AR (p) order is detected if the partial auto-correlation plot (PACF) cuts off after p lags, with the auto-correlation plot (ACF) tailing off. On the other hand, a MA(q) order is detected, if the ACF plot cuts off after q lags, with the PACF plot tailing off. Tailing off of both ACF and PACF plots confirms an ARMA (p,q) order.

Selecting an ARMA (p,q) model order from the ACF and PACF plot is not straightforward and these plots at best suggest that the series under investigation is not a pure AR (p) or MA (q) series. Some of the other academic approaches involves the usage of extended sample auto-correlation function (ESACF) and inverse auto-correlation function (IACF) [22].

On the other hand, the objective of the automated grid search of these ARIMA model hyper-parameters is to avoid manually looking at the ACF and PACF plots, a process which is cumbersome, time-consuming and prone to human errors. The grid search described here deploys a nested loop approach to determine p, d, and q, by running a maximum of $p \times d \times q$ cases. Improvement in computation time has been observed by using selection criterions which penalizes more complex models if they fail to show significantly better performance over simpler models. From the commonly available criteria in the literature, the Akaike information criterion (AIC) and Bayesian information criterion (BIC), Equations 6a and 6b has been selected to gauge model performance in this study.

$$AIC = 2k - 2\ln(\hat{L}) \quad (6a)$$

$$BIC = \ln(n)k - 2\ln(\hat{L}) \quad (6b)$$

In both Equations 6a and 6b, $\hat{L}$ is the maximum value of the likelihood function of the model, *n* is the number of data points and *k* is the number of free parameters to be estimated. The second term in Equations 6a and 6b penalizes the ARIMA model for too many parameters. As such, to improve computation efficiency the approach used here is to exit the nested grid search loop if the numerical value of the AIC or BIC converges to a high number after several iteration steps.

A final step before a grid search simulation is executed to establish if the given data has a seasonal trend. From several decomposition techniques that are discussed in the literature [23,24] a seasonal and trend decomposition using loess (STL) technique was used to verify seasonality. The relative merit of this decomposition technique is that its robust to outliers, a property ideal to analyse wildfire data. With seasonality observed in the data-set three more hyperparameters, *P*, *D*, and *Q* are defined for the seasonal component of the model. Thus, the complete SARIMA *(p,d,q)(P,D,Q,m)* model under investigation have trend component parameters of *p*, *d* and, *q* for the autoregression, difference and moving average portions and *P*, *D*, *Q* and, *m* due to the seasonal component for autoregression, difference, moving average and seasonal period.

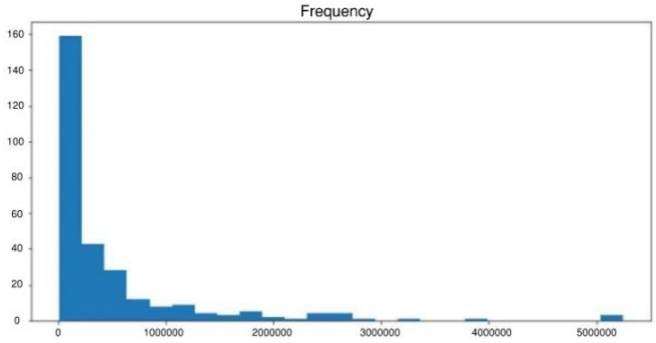
With a grid search simulation on the wildfire data-set with 288 entries, with an aggregated granularity at a monthly level, the hyper-parameters for this SARIMA model are found to be (0,1,1)(1,0,1,12) along with a total computation time of 38.76 seconds on standard Intel i5 hardware. The model performance in terms of standard statistical metrics and numerical values of AIC, BIC are summarized in Table 2. Figure 3 shows the prediction performance for 24 months with seasonal high during the summer months.

2.3. SARIMA Model Parameter Selection Using Power Transforms

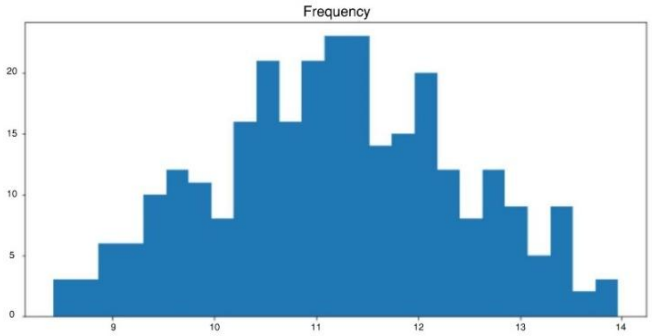
Real-life data as obtained from [17] rarely tend to be normally distributed and is typically skewed. The issue with trying to fit a regression model on a skewed data is often complicated and not encouraged since the regression assumptions are violated. Decision tree-based classification techniques can be used to fit this kind of data, without requiring the data to be transformed, as noted in reference [25].

For auto-regressive based models, power transforms like Box-Cox or Yeo-Johnson transformations [26,27] are typically used to normalize a skewed distribution. We can use these transforms to our advantage to normalize our data-set for better fitting our model. A two-parameter Box-Cox transformation as defined in Equation 7 is used to transform the wildfire study data. The transformed data is sequentially fitted into a SARIMA model using the same grid search technique described in Section 2.2. Once the transformation is applied and the grid-search algorithm is executed, it is imperative to re-transform the data back to get to the original scale.

$$y(\lambda) = \begin{cases} \frac{((y+\lambda_2)^{\lambda_1}-1}{\lambda_1} & if\ \lambda_1 \neq 0 \\ \log(y+\lambda_2) & if\ \lambda_1 = 0 \end{cases} \quad (7)$$



(a) Time series statistics of the untransformed data-set



(b)Time series statistics of the Box-Cox transformed data-set

Figure 2. Critical statistics of the time series data

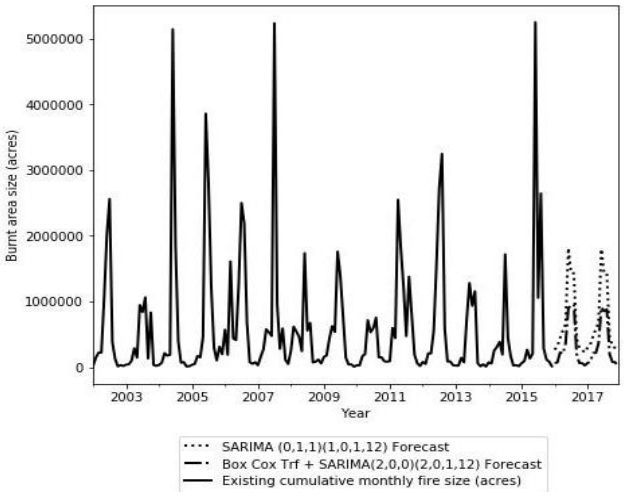


Figure 3. Forecast performance of SARIMA and transformed SARIMA models

In Equation 7, *y* is the response variable and λ_1, λ_2 are the transformation parameter. It should be noted that $\lambda_1 = 0$ requires a special definition and is the log of the data itself. Statistical analysis on the original data-set from [17] reveals skewness in Figure 2a, as expected with phenomenon like wildfires. The given data-set is normalized with a Box-Cox transformed, as described by Equation 7 and is shown in

Figure 2b. Readers should be aware that due to the nature of the Box-Cox transformation, only positive valued data-set or data-set with non-negative values can be transformed which is ideal for wildfire data-set. With a grid search simulation on the Box-Cox transformed wildfire data-set with 288 entries, aggregated to a monthly granularity level, the hyper-parameters for the Box-Cox transformed SARIMA model were found to be (2,0,0)(2,0,1,12) with a total computation time of 101.13 seconds on standard Intel i5 hardware. The model performance in terms of standard statistical metrics and numerical values of AIC, BIC are summarized in Table 2. Figure 3 shows the prediction performance for 24 months with a seasonal high during the summer months.

3. Numerical Performance Evaluation

Before evaluating model performance, some notes on forecasting very long and extremely granular times series are presented. It is a general observation that most time series model do not work well for a long and extremely granular data-set. Wildfire data obtained from [17] consists of multiple entry for each date-time index and recorded at a twenty-four-hour level of granularity. These individual entries should be aggregated to an acceptable level of granularity, to improve computation time and retain forecasting objective. For this study, the time stamp data were aggregated to the level of granularity at a monthly level, which aligns well with the objective of this study.

Table 2. Relative model performance and forecast error

Simulation type	AIC	BIC	Mean	RMSE	Simulation time (s)
Grid search	8541	8560	211230	6802	38.75
Box-Cox & Grid search	590	620	211230	1144	101.13

For model evaluation, the entire data-set is broken into training and testing data and the model is trained on the initial eighty percentage of the data. The forecasting performance of both the SARIMA (0,1,1)(1,0,1,12) model and the Box-Cox transformed SARIMA (2,0,0)(2,0,1,12) model has been measured by root mean square error (RMSE) on the test set of the wild fire data-set. For a detailed discussion of these evaluation metrics, readers are encouraged to refer [28]. The improvement of the proposed Box-Cox transformed SARIMA (2,0,0)(2,0,1,12) model over the benchmark regular SARIMA (0,1,1)(1,0,1,12) model can be seen from Table 2. It is observed that without the transformation, a straightforward nested grid search algorithm yields a relatively high information criterion and higher root mean square error on the test data. Also, the overall performance of the two models is superior, given the low RMSE compared to the data mean. From these observations, it can be suggested that suitable transformation methods should be explored before fitting naturally skewed data-set for robust model fitting. An initial descriptive statistical analysis should also be conducted, as discussed above, on skewed data-sets to establish whether the data is left or right-skewed.

4. Recommendations on Future Energy Policies

4.1. Operating Framework for Local and Regional Fire Departments

This section deals with developing a workable risk-model for the local and regional fire departments. The proposed risk-model for regional risk assessment is divided into three parts: the working voltage level of the electrical asset being analysed, the position of the active zone of the flame in regard to the electrical conductors, and the effect of temperature, relative humidity and ionization in form of air pollution.

The idea of the first component of the risk-model is to establish transmission security for important 500, 345 and 138 kV lines since an outage on these lines have major impacts on grid stability and on public welfare. Since the voltage levels and power carrying capacity of electrical networks are typically non-linear, a logarithmic scale is used for this model. In Equation 8, α_i is the sensitivity factor to be chosen by fire analyst and *n* represents a scaled integer corresponding to a voltage level.

$$Risk\ component\ due\ to\ electrical\ voltage = \alpha_i \times log(n)\ \forall n \in [1,10] \quad (8)$$

The second part of the model evaluates the breakdown risk of the electrical conductors/bus of electrical assets based on the relative position of the flame. For example, a 345 kV transmission line with a typical height ranging between 49 to 180 ft has a lower risk than a bus at a switching substation at the same voltage level due to its proximity to the general terrain. In Equation 9, β_i is the sensitivity factor to be chosen by fire analyst and H_{c-t} represent the height of the energized electrical asset from the terrain level.

$$Risk\ component\ due\ to\ spatial\ position\ of\ fire = \beta_i \times \frac{1}{H_{c-t}} \quad (9)$$

The last portion of the risk-model deals with the risk components due to environmental factors: temperature, relative humidity, and ionization in the form of air pollution. This model considers the result from the study [29] which found the effect of temperature on electric field breakdown to have an inverse relation. Also, areas with higher forest coverage (ν_{fc}) typically tend to increase the temperature factor by providing forest fuel for a wildfire burn and is included as a risk-parameter. Further, higher relative humidity (*RH*) tends to reduce wildfire risk by a square inverse factor. Last, experimental studies in a controlled environment have shown a higher risk of electrical breakdown in the presence of atmospheric contaminants (τ_{ppm}) [30] and is included in the risk-model. The combined

model with all the above parameters included is shown in Equation 10. Like the sub-models discussed an operator adjustable sensitivity factor of Υ_i is added into the equation.

$$Risk\ component\ due\ to\ environmental\ factor = \Upsilon_i \times \nu_{fc} \times T_{ambient} \times \frac{K_1}{RH^2} \times \tau_{ppm} \quad (10)$$

Combining the three components of this risk-model, a heat map was generated for a 138 kV transmission line using a pole tower of 90' height for the state lower 48 states of the US with atmospheric conditions from major regional cities. It should be noted that due to the nature of the risk components in Equations 8 and 9 multiple simulations are necessary to capture transmission assets at different voltage levels and to differentiate transmission and distribution assets from substation assets. Also, given to the dynamic nature of the environmental component of the risk-model, as in Equation 10, regular model updates with current environmental parameters is necessary.

Figure 4 shows high risk areas in US, with county wise mapping of the lower 48 states using Federal Information Processing Standards (FIPS) codes. The mapping uses resources like Google Maps and global historical climatology network (GHCN) data-set from National Centers for Environmental Information [31] for this mapping. Equal weigh was given to all three sensitivity factors α_i, β_i, and Υ_i. The model used a scoring system and the heat-zones in Figure 4 represents a unique score for each county for all the forty-eight states. High risk areas are detected whenever the combination of high forest coverage (ν_{fc}), low relative humidity (RH), high ambient temperature ($T_{ambient}$) and high atmospheric contamination (τ_{ppm}) are favourable from a multiplicative standpoint. As such, from Figure 4, US states like California, Nevada, Arizona and Oregon having the right mix of these factors are more susceptible to forest fires as compared to other US states like Virginia and Tennessee, where high relative humidity creates an unfavourable wildfire condition.

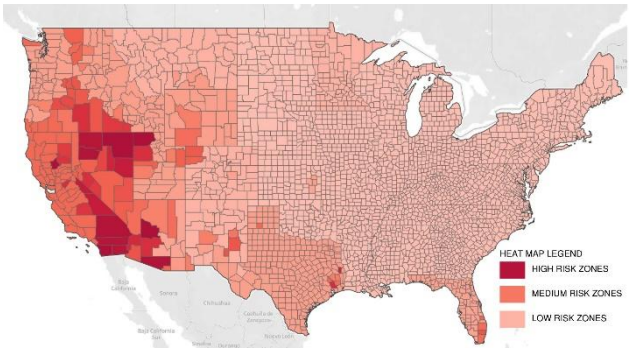


Figure 4. Heat map showing modelled risk of wildfires in US for 138 kV transmission lines

4.2. Operating Framework for Entities Managing Transmission and Distribution

4.2.1. Upgrade System Protection and Relaying Schemes

The strategical operational practices: combining forecasting wildfires from time series analysis of historical data and accurate regional risk mapping can help shape local and regional fire management policies. However, grid operators can do much more than solely relying on these strategic practices for grid operation and can invest to upgrade system protection and harden the overall system, as would be discussed next.

The issue of wildfires as seen in the introductory section of this article revealed that a major portion of these incidents is caused due to broken or downed conductors. Fault scenarios arising from these faults are typically single line to ground (SLG) in nature. These SLG faults are unbalanced and must be analysed using symmetrical component analysis technique. An excellent discussion of symmetrical component analysis can be found in [32], and the results from such an analysis is summarized in Equation 11a-c, where V_f is the pre-fault voltage, Z_f is the fault impedance and Z_0, Z_1, Z_2 are the zero, positive and negative sequence impedance.

$$I_a = I_0 + I_1 + I_2 = 3I_1 = \frac{3V_f}{Z_0+Z_1+Z_2+3Z_f} \quad (11a)$$

$$I_b = 0 \quad (11b)$$

$$I_c = 0 \quad (11c)$$

As can be seen from Equation 11a, the problem with these SLG faults is the fact that high fault impedance can make these faults seem benign, though they have a huge potential of causing a wildfire. Adding to the complication of detecting SLG faulted conductor and subsequent tripping under these high impedance fault (HIF) scenarios is the fact that poor insulation maintenance or insulator degradation can also cause a high-impedance fault scenario, with current trickling out of the conductors at these weak spots. Typically, time and frequency domain signatures are analysed to detect high impedance faults [33], which are otherwise difficult to detect. The challenge is to design a protection system that can combine the detection of high-impedance faults along with some signals from the field of an actual downed conductor. To accomplish this, a feeder protection relay with an internal logic set to detect high-impedance fault will have to be combined with wireless signals obtained from its downstream feeder lines. This combination logic can be used to send a block signal to the reclosers controlling the feeder. In general, the above combination logic must detect both high impedance fault and a downed conductor by analysing the negative sequence current signature obtained from the wireless signal of its downstream feeder lines. As such, a broken phase conductor will have a high negative to positive sequence current ratio [34].

To prevent the cost of maintenance and ensure reliable operation without an additional power source, field clampable current transformers module (FCCTM) can be used to generate these wireless signals, using the power from a distribution feeder line itself. Use of an internal Lithium cell battery would make them a standalone sensor and adding a wireless antenna can report loss of current to its parent station.

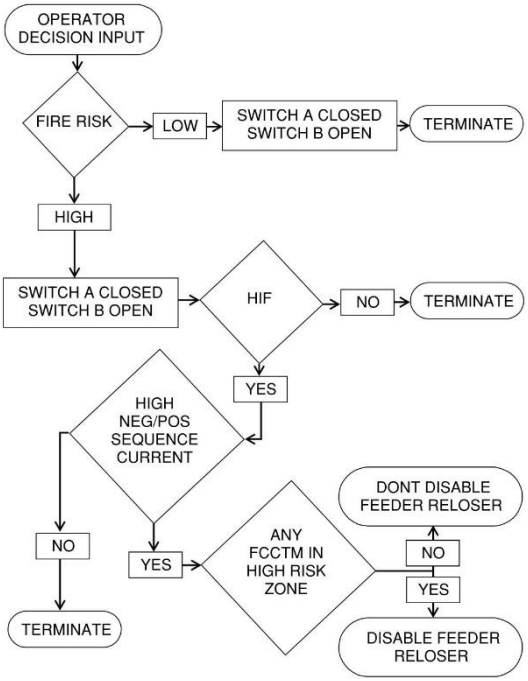


Figure 5. Flowchart to guide disabling reclosure and fast acting fuse switching

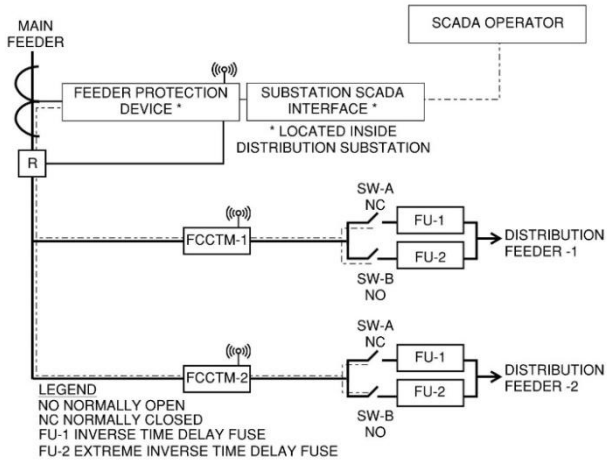


Figure 6. Guided reclosing using field FCCTM signals

The method under discussion of disabling reclosure operation can be implemented in conjunction with slight configuration of the distribution feeder relay to either use or bypass an extreme time inverse fuse: a decision that the SCADA operator can make remotely, as can be seen from the working flowchart in Figure 5 and network configuration in Figure 6. The overall protection philosophy is to detect down conductor signal (high negative to positive sequence current ratio) from these FCCTM devices and combine it with a high impedance fault combination logic for feeder relay operation on the reclosure. A benefit of having unique identification signatures for each of the FCCTM devices is that the SCADA operator can decide to engage reclosing if the FCCTM's negative to positive sequence high current signal is not from a high-risk fire zone and to disable reclosing otherwise.

4.2.2. Other Effective Techniques to Harden the Overall Electrical Grid

A few non-conventional approaches are discussed in this section, which has the potential to harden the electrical network against wildfires.

To enhance electrical grid resilience regular oil filled instrumentation transformers can be replaced with its optical counterpart, in high wildfire risk zones. These compact optical transformers have superior seismic performance due to its light weight and footprint and can be used to adapt to a wide range of electrical turns ratio. Other major advantages of these systems are its linear performance due to the absence of iron core saturation and galvanic isolation of high voltage lines. Unlike a capacitor coupled voltage transformer (CCVT) an optical transformer relies on the physical phenomenon of Pockels effect [35] and can be used to measure harmonic contents at interfacing substations and point of interconnections.

Though instrumentation transformers can be replaced by optical transformers in limited scenarios, oil cooled power transformers possess its unique challenges in terms of wildfire risk. Engineered K-class ester based mineral oil [36] with extremely low flammability can be used to replace existing transformer oil during routing or emergency maintenance, to reduce fire propagation through utility transformer assets in high risk areas. These engineered ester-based oils can have fire points as high as 680°F and can be tuned to comply with industry testing standards as IEEE C57.154 [37] for high temperature insulation systems.

4.3. Operating Framework for Entities Managing Distribution Energy Storage Assets

Distributed energy storage systems (DESS) can contribute in forming a bare backbone transmission and distribution network with decentralized energy pools. Since the reliability of the transmission and distribution network is ensured around peak wildfire period of summer and early fall, due to the maintenance and upgrades done in the previous season as outlined in Section 4.1, developers can use this timeframe to line up DESS projects.

During these wildfire events, reasons to promote the use of DESS in enhancing grid flexibility are many. During a wildfire incident a DESS resource can support to maintain security constrained power flow, help prevent anti-islanding by real power support, and help system operators to reconfigure the network remotely and still maintain overall grid stability. After the course of a wildfire, a DESS cluster of distributed generation (DG)/ demand responsive load (DR) can be used to black start tripped generators and for other load restoration. As such these DESS clusters can help in a much faster recovery of the affected electrical grid. Readers should be aware however of the potential risks involved with energy storage technologies from a fire safety standpoint and should have these resources installed at site in accordance to NFPA 855: standard for the installation of stationary energy storage system [38].

Furthermore, during daily grid operations, often time in practice a combination of reduced transmission line loading along with support from distributed generation and demand-responsive loads can help the grid ride-through an emergency condition. The rest of this section presents a discussion of a DC optimal power flow formulation under a combination of these stressed situations, to ensure the most economical allocation and dispatch of generations and loads. The formation discussed henceforth considers the presence of both demand-responsive loads and distributed generation in micro-grid nodes which are connected to the main electric grid. There are many DC optimal power flow formulation available in the literature. For instance [39] presents a DC optimal power flow formulation (DCOPF) to analyse N-1 contingency analysis with optimal transmission switching. Each DC optimal power flow formulation has its own set of objective function and boundary conditions, and hence the corresponding primal-dual interpretation of every problem is unique to the situation.

The optimization problem assumes that all generation costs are linear and a lower bound of zero active power production from these generating resources. The optimization problem is developed as follows:

$$min \sum_{t=1}^{T} \left(\left(\sum_{g} C_g(t) P_g(t) \right) + \left(\sum_{DR} C_{DR}^{i}(t) P_{DR}^{i}(t) \right) + \left(\sum_{DG} C_{DG}^{i}(t) P_{DG}^{i}(t) \right) \right) \quad (12a)$$

$$st - P_g^{min} \geq -P_g \geq -P_g^{max} \quad (12b)$$

$$0 \geq -P_{DG} \geq -P_{DG}^{max} \quad (12c)$$

$$0 \geq -P_{DR} \geq -P_{DR}^{contract} \quad (12d)$$

$$\left(\frac{\theta_a - \theta_b}{-X_k} \right) - \psi P_k = 0 \quad (12e)$$

$$\sum_{\forall inputflow} P_k - \sum_{\forall outputflow} P_k + \sum_{\forall g} P_g = d_n \quad (12f)$$

$$(1 - \omega_i) \left(\sum_{DG} P_{DG}^{i}(t) + \sum_{DR} P_{DR}^{i}(t) \right) \geq (1 - \omega_i) d_i \quad (12g)$$

Nomenclature

i	:	nodes containing distributed generation (DG)/ demand responsive (DR) assets
n	:	nodes containing no distributed generation (DG)/ demand responsive (DR) assets
k	:	transmission element
c_g^k	:	linear operating cost of generator g at node k
c_{DR}^i	:	linear unit cost associated with controllable load DR at node i
c_{DG}^i	:	linear operating cost of distributed generator DG at node i
P_g^k	:	real power from generator g at node k
P_{DR}^i	:	real power from controllable load DR at node i
P_{DG}^i	:	real power from distributed-generator DR at node i
P_k	:	real power flow in transmission element k
X_k	:	reactance of the given transmission element
$\theta_{a,b}$	:	voltage phase angle
ψ	:	operator chosen line derating factor during fire season
ω_i	:	binary islanding operator switch
d_n	:	real power load demand at node n
d_i	:	real power load demand at node i

The objective function and constraints of the of the above optimal dispatch formulation are explained below:

Equation 12a: Objective function - The objective function of this formulation is to minimize generation cost and hence find the most economical generation level, that can provide a feasible solution during grid disturbance as these. The first term of the minimization objective function represents the cost of generation from convention generating units. The second term of the objective function represents the cost of purchasing demand response reserve from a demand response load. Typically, these can be done through contractual agreements and can be renewed as required. The last term represents the cost of generation from the distributed resources which can look very different from its utility generation counterpart.

Equation 12b: Utility generation constraint - This constraint captures the practical loading of utility generators. By its very nature these generators are often bound in their minimum and maximum real power output levels set-forth by manufacturer warranty agreements.

Equation 12c: Distributed generation constraint - The given constraint captures the maximum power limit of the distributed generators. Note that since these are usually modular power electronic controlled devices, the lower operating limit is not of concern and is chosen to be zero.

Equation 12d: Demand response contractual constraint - The demand response contractual constraint captures the contractual agreement made between the system operators and the owners of the demand responsive load. This constraint can be changed based on season agreements.

Equation 12e: Transmission element thermal rating constraint - The power flow constraint is a linear approximation of the actual line flow constraint and consists of the reactance X_k of the transmission element and the voltage phase angle between its two ends $\theta_{a,b}$. The term ψ is an active derating factor that can be chosen by the grid operator to de-rate certain transmission elements in the path of high fire risk zones.

Equation 12f: Node power balance constraint - The node power balance constraint involves boundary condition between power inflow and outflow in a transmission element, generation and load levels at a node.

Equation 12g: Island operation constraint - A feasible operational islanded microgrid needs to balance its load in case the operator must island an electric region to preserve grid stability. The active load demand of this microgrid should be less than or equal to the available real power from its distributed generation and demand responsive load resources. The term ω_i is a binary state indicator and equals zero in an active island condition.

As can be seen from this formulation, the system operator can ensure better ride-through of the electrical grid under system disturbance with a combination of strategies as reduced line loading, contractual demand responsive load agreements and islanded microgrid operations in a small scale in probable high fire risk zones. On the other hand, DESS developers can actively engage to increase the preparedness of the electric network to disturbance as these through simulation studies and strategic placement of DESS resources.

5. Conclusions

A combination of statistical forecasting, risk-modelling, and power flow optimization framework has great potential to improve the impacts of wildfires on public welfare and safety. A proactive multi-faceted approach at different levels and implementation of design codes and standards can make the bulk power system resilient to these phenomena. This paper explores the practical challenges of numerical forecasting of wildfire data and provides solutions to reproducible accurate forecasting. Fire operators can use the framework discussed in this paper at a level of granularity of their geographical region for future planning and related investment decisions. Electrical utilities can implement a multi-year approach to systematically upgrade its protection schemes with advanced protection and relaying schemes presented in this paper. Along with these policy recommendations, a feasible optimization framework of combined distributed generation, demand-responsive loads and adjusted line capacities can help system operators ride-through these challenging situations.

REFERENCES

[1] Susan M Stein, J Menakis, MA Carr, SJ Comas, SI Stewart, H Cleveland, L Bramwell, and VC Radeloff. Wildfire, wildlands, and people: understanding and preparing for wildfire in the wildland-urban interfacea forests on the edge report. *Gen. Tech. Rep. RMRS-GTR-299. Fort Collins, CO. US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 36 p.*, 299, 2013.

[2] Fantina Tedim, Vittorio Leone, and Gavriil Xanthopoulos. A wildfire risk management concept based on a social-ecological approach in the european union: Fire smart territory. *International Journal of Disaster Risk Reduction,* 18:138–153, 2016.

[3] A Andreu and L Annie Hermansen-Baez. Fire in the south 2: the southern wildfire risk assessment. *A report by the Southern Group of State Foresters, 32 p.*, 2008.

[4] David E Calkin, Alan Ager, Matthew P Thompson, Mark A Finney, Danny C Lee, Thomas M Quigley, Charles W McHugh, Karin L Riley, and Julie M Gilbertson-Day. A comparative risk assessment framework for wildland fire management: the 2010 cohesive strategy science report. 2011.

[5] Timothy J Brown, Beth L Hall, and Anthony L Westerling. The impact of twenty-first century climate change on wildland fire danger in the western united states: an applications perspective. *Climatic change*, 62(1-3): 365–388, 2004.

[6] Larry Dale, Michael Carnall, G Fitts, SL McDonald, and M Wei. Assessing the impact of wildfires on the California electricity grid. *California's Fourth Climate Change Assessment, California Energy Commission. Publication Number: CCCA4-CEC-2018-002*, 2018.

[7] PG&E safety plan 2019. Pacific Gas and Electric Company Amended 2019 Wildfire Safety Plan, Feb 2019, Also available as: https://www.pge.com/pge_global/common/pdfs/safety/emergency-preparedness/natural-disaster/wildfires/Wildfire-Safety-Plan.pdf.

[8] Saurabh S Soman, Hamidreza Zareipour, Om Malik, and Paras Mandal. A review of wind power and wind speed forecasting methods with different time horizons. In *North American Power Symposium 2010*, pages 1–8. IEEE, 2010.

[9] Tina Jakaˇsa, Ivan Androˇcec, and Petar Sprˇci´c. Electricity price forecasting—arima model approach. In *2011 8th International Conference on the European Energy Market (EEM)*, pages 222–225. IEEE, 2011.

[10] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3): 1014–1020, 2003.

[11] Menzie David Chinn, Michael LeBlanc, and Olivier Coibion. The predictive characteristics of energy futures: Recent evidence for crude oil, natural gas, gasoline and heating oil. *Natural Gas, Gasoline and Heating Oil (October 2001). UCSC Economics Working Paper*, (490), 2001.

[12] Sakulkitbanjong Kanyaphorn, Pongchavalit Chunchom, and Garivait Savitri. Time series analysis and forecasting of forest fire weather. *International Journal of Management and Applied Science (IJMAS)*, 3: 35–40, 2017.

[13] Athaya Putri Slavia, Edi Sutoyo, and Deden Witarsyah. Hotspots forecasting using autoregressive integrated moving average (arima) for detecting forest fires. In *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*, pages 92–97. IEEE, 2019.

[14] Kan Xu, Xiaozhi Zhang, Zhiguo Chen, Wenhao Wu, and Tao Li. Risk assessment for wildfire occurrence in high-voltage power line corridors by using remote-sensing techniques: a case study in hubei province, china. *International journal of remote sensing*, 37(20):4818–4837, 2016.

[15] Conrad Bielski, V O'Brien, C Whitmore, K Ylinen, I Juga, P Nurmi, J Kilpinen, I Porras, JM Sole, P Gamez, et al. Coupling early warning services, crowdsourcing, and modelling for improved decision support and wildfire emergency management. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 3705–3712. IEEE, 2017.

[16] Moein Choobineh and Salman Mohagheghi. Power grid vulnerability assessment against wildfires using probabilistic progression estimation model. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2016.

[17] Karen C Short. Spatial wildfire occurrence data for the united states, 1992-2015 [fpa fod 20170508]. (4th Edition), 2017.

Also available as https://doi.org/10.2737/RDS-2013-0009.4.

[18] Lingling Li, Changyu Shen, Xiaochun Li, and James M Robins. On weighting approaches for missing data. *Statistical methods in medical research*, 22(1): 14–30, 2013.

[19] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[20] Richard ID Harris. Testing for unit roots using the augmented dickey-fuller test: Some issues relating to the size, power and the lag structure of the test. *Economics letters*, 38(4): 381–386, 1992.

[21] Yongcheol Shin and Peter Schmidt. The kpss stationarity test as a unit root test. *Economics Letters*, 38(4): 387–392, 1992.

[22] Christopher Chatfield. Inverse autocorrelations. *Journal of the Royal Statistical Society: Series A (General)*, 142(3): 363–377, 1979.

[23] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.

[24] Robert B Cleveland, William S Cleveland, Jean E McRae, and Irma Terpenning. Stl: A seasonal-trend decomposition. *Journal of official statistics*, 6(1): 3–73, 1990.

[25] David A Cieslak and Nitesh V Chawla. Learning decision trees for unbalanced data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 241–256. Springer, 2008.

[26] Remi M Sakia. The box-cox transformation technique: a review. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 41(2): 169–178, 1992.

[27] Sanford Weisberg. Yeo-johnson power transformations. *Department of Applied Statistics, University of Minnesota. Retrieved June*, 1: 2003, 2001.

[28] Samaher Al Janabi, Ibrahim Al Shourbaji, and Mahdi A Salman. Assessing the suitability of soft computing approaches for forest fires prediction. A*pplied computing and informatics*, 14(2): 214–224, 2018.

[29] Haixiang Chen, Yun Zhang, and Linhe Zhang. A method to assess the wildfire induced breakdown of high-voltage transmission lines. In *Journal of Physics: Conference Series*, volume 1074, page 012152. IOP Publishing, 2018.

[30] AA Al-Arainy, NH Malik, and MI Qureshi. Influence of sand/dust contamination on the breakdown of asymmetrical air gaps under lightning impulses. *IEEE transactions on electrical insulation*, 27(2): 193–206, 1992.

[31] N.O.A.A. Global historical climatology network. 2019. NOAA GHCN (Global Historical Climatology Network) – Daily Documentation, 2019, Also available as https://www.ncdc.noaa.gov/cdo-web/data-sets# GHCND.

[32] Sajeeb Saha, M Aldeen, and Chee Pin Tan. Unsymmetrical fault diagnosis in transmission/distribution networks. *International Journal of Electrical Power & Energy Systems*, 45(1): 252–263, 2013.

[33] B Mike Aucoin and B Don Russell. Distribution high impedance fault detection utilizing high frequency current components. *IEEE Transactions on Power Apparatus and Systems*, (6): 1596–1606, 1982.

[34] R Weerapun and P Anantachai. A novel detection system for broken distribution conductor on radial scheme. *CIRED 21st International Conference on Electricity Distribution*, 1(2): 1, 2011.

[35] Lars Hofmann Christensen. Design, construction, and test of a passive optical prototype high voltage instrument transformer. *IEEE Transactions on Power Delivery*, 10(3): 1332–1337, 1995.

[36] C Patrick McShane. Relative properties of the new combustion-resist vegetable-oil-based dielectric coolants for distribution and power transformers. *IEEE Transactions on industry applications*, 37(4):1132–1139, 2001.

[37] I.E.E.E. Standard for the design, testing, and application of liquid-immersed distribution, power, and regulating transformers using high-temperature insulation systems and operating at elevated temperatures. pages 1–49, 2012. IEEE Std C57. 154-2012.

[38] N.F.P.A. Standard for the installation of stationary energy storage systems. 2020. NFPA 855.

[39] Kory W Hedman, Richard P O'Neill, Emily Bartholomew Fisher, and Shmuel S Oren. Optimal transmission switching with contingency analysis. *IEEE Transactions on Power Systems*, 24(3): 1577–1586, 200.