



Convolutional Non-Homogeneous Poisson Process and its Application to Wildfire Ignition Risk Quantification for Power Delivery Networks

Guanzhou Wei, Feng Qiu & Xiao Liu

To cite this article: Guanzhou Wei, Feng Qiu & Xiao Liu (15 Jul 2024): Convolutional Non-Homogeneous Poisson Process and its Application to Wildfire Ignition Risk Quantification for Power Delivery Networks, Technometrics, DOI: [10.1080/00401706.2024.2365729](https://doi.org/10.1080/00401706.2024.2365729)

To link to this article: <https://doi.org/10.1080/00401706.2024.2365729>

 View supplementary material 

 Published online: 15 Jul 2024.

 Submit your article to this journal 

 Article views: 112

 View related articles 

 View Crossmark data 



Convolutional Non-Homogeneous Poisson Process and its Application to Wildfire Ignition Risk Quantification for Power Delivery Networks

Guanzhou Wei^a, Feng Qiu^b, and Xiao Liu^a

^aH. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA; ^bAdvanced Grid Modeling, Optimization and Analytics, Argonne National Lab, Lemont, IL

ABSTRACT

To quantify wildfire ignition risks on power delivery networks, the current practice predominantly relies on the empirically calculated fire danger indices, which may not well capture the effects of dynamically changing environmental factors. This article proposes a spatio-temporal point process model, known as the Convolutional Non-homogeneous Poisson Process (cNHPP), and applies the model to quantify wildfire ignition risks for power delivery networks. The proposed model captures both the current (i.e., instantaneous) and cumulative (i.e., historical) effects of key environmental processes (i.e., covariates) on wildfire risks, as well as the spatio-temporal dependency among different segments of the power delivery network. The computation and interpretation of the intensity function are thoroughly investigated. We apply the proposed approach to estimate wildfire ignition risks on major transmission lines in California, using historical fire data, meteorological and vegetation data obtained from the National Oceanic and Atmospheric Administration and National Aeronautics and Space Administration. A comprehensive comparison study is performed to show the applicability and predictive capability of the proposed approach.

ARTICLE HISTORY

Received February 2023
Accepted May 2024

KEYWORDS

Non-homogeneous Poisson process; Power delivery networks; Spatio-temporal point process; Wildfire ignition risks

1. Introduction

Wildfires, ignited from power delivery systems, have become a major threat to energy infrastructure and public safety. In 2018, the Camp Fire, that is, the deadliest wildfire in California history, was ignited by a faulty electric transmission line. The Camp Fire killed 85 people, destroyed over 15,000 structures, and caused a total insured loss of \$12.5 billion (Schulze et al. 2020). Other examples of wildfires ignited by power lines include the Grass Valley Fire, Malibu Canyon Fire, Rice Fire, Sedgwick Fire, and Witch Fire, and these power-line fires burned a total area of more than 334 square miles (CPUC 2014). Recently, the Department of Energy (DOE) announced up to \$3.46 billion across 44 states to strengthen electric grid resilience and reliability across the United States, all while improving climate resilience (U.S. DOE 2023). The current projection shows that much of the continental U.S. will have significantly hotter and drier days due to climate change (Brown, Wang, and Feng 2021), leading to more wildfire hazards.

To estimate the wildfire ignition risk on electric grid, utilities rely on fire indices calculated from a set of environmental variables using some empirical equations (NFDRS 2002; SDG&E 2021). These indices may well capture the trend of fire ignition risks on a large spatio-temporal scale, but become less adequate in capturing the dynamic real-time relationship between fire risks and key environmental processes, such as the dynamically changing meteorological conditions like temperature and wind.

However, the capability to dynamically update the predicted fire ignition risks for power delivery infrastructures, given the latest information on meteorological and vegetation conditions is critical for utilities to facilitate real-time operational decisions. For example, in response to the devastation of power-line fires, the California Public Utilities Commission (CPUC) launched the Public Safety Power Shutoff (PSPS) activity in 2018 which authorizes electric utilities to shut off electric power for public safety. Pacific Gas and Electric (PG&E), one of the major utilities in California, also took a series of PSPS activities in 2019. Although these PSPS activities decreased the ignition of power-line fires, about 1,848,000 customers were impacted by a series of power outages, and 44% of respondents reported power loss for three or more days (Mildenberger et al. 2022).

To better plan for PSPS against wildfires and prioritize power grid protection measures given limited resources, this article proposes a spatio-temporal point process model, known as the Convolutional Non-homogeneous Poisson Process (cNHPP), and applies this model to quantify daily fire risks at different segments/locations of a subset of power grids in California. Ideally, the model needs to be capable of capturing the (i) current effects (i.e., short-term) of environmental covariates on fire risks, (ii) cumulative effects (i.e., long-term) of historical covariates information on current fire risks, and (iii) spatio-temporal interactions among neighboring power transmission lines. Building such a model is made possible as data become readily available.

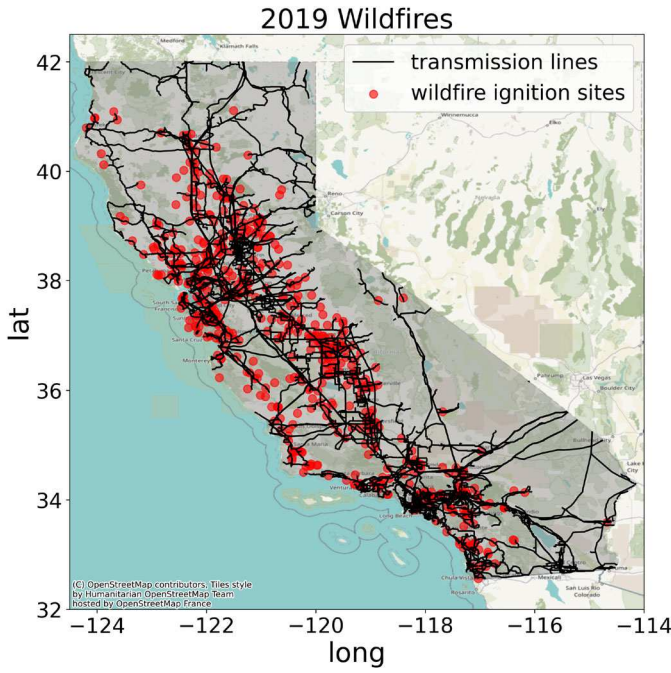


Figure 1. Fires ignited from major power delivery networks in California in 2019.

For example, the CPUC requires electric companies to report fire incident data from their power facilities. Figure 1 shows the reported power-line fires in 2019 and the network of power transmission lines in California from the U.S. Energy Information Administration (EIA). It is seen that most of the fire incidences are on or in close proximity to the major power delivery lines and are related to power delivery infrastructures. Combined with the meteorological and vegetation variables from the National Oceanic and Atmospheric Administration (NOAA) and National Aeronautics and Space Administration (NASA), we have the datasets needed for constructing and validating the statistical model to be described in this article, which estimates and predicts wildfire event intensities on segments of a network of power transmission lines.

1.1. Literature Review

Assessment of wildfire hazards has a long history since the Canadian Forest Fire Danger Rating System (CFFDRS) was developed in 1968 and the U.S. National Fire Danger Rating System (NFDRS) was created in 1972. These rating systems generate fire indices that reflect the potential wildfire hazards based on weather, fuels, and topography information (NFDRS 2002); for example, the Canadian Fire Weather Index (CFFWI) and Fire Potential Index (FPI). As discussed above, these fire indices are often used to capture the overall long-term trend of wildfire risks on a large spatio-temporal scale. Given the observed wildfire incident data, point process models have also been investigated (Taylor et al. 2013; Holbrook Ji). For example, Peng, Schoenberg, and Woods (2005) developed a spatio-temporal point process for modeling wildfire risks in Los Angeles County. The authors incorporated both the burning index (one of the indices generated from the NFDRS) and the spatio-temporal trend obtained from the historical wildfire data into the intensity function of the proposed point process. Xu and Schoenberg (2011) later

showed that the incorporation of weather and fuel information into the intensity function, proposed in Peng, Schoenberg, and Woods (2005), enhances the performance of the spatio-temporal point process. Serra et al. (2014) adopted a spatio-temporal log-Gaussian Cox process for modeling Catalonia wildfire occurrences. Opitz, Bonneau, and Gabriel (2020) leveraged a log-Gaussian Cox process model for forest fires in Mediterranean France, which incorporates covariates such as land use and weather conditions through a linear model with random effects.

Although various spatio-temporal point processes have been investigated for modeling wildfire risks, most of them consider a continuous two-dimensional space. Because power delivery infrastructures are distributed on a linear network that consists of segments of power lines, the support of the process is constrained by a linear network, and the spatio-temporal covariance can depend on the topology of the network such as the distances between transmission lines. Hence, there is a need to construct spatio-temporal point processes on a linear network of power transmission lines. Uppala and Handcock (2020) adopted a separable temporal linear point process to model wildfire ignition on a road network, in which the Papangelou conditional intensity function is assumed to be a log-linear function of covariates. Zhu et al. (2022) used the Hawkes process to model power outages on power grids under extreme weather conditions. The authors developed the background intensity by employing a deep neural network that incorporates the cumulative weather effects in time, while the triggering effects were introduced by considering power grid connectivity and power outage history. Note that, point process models on linear networks have also been found in the modeling of street crimes and traffic accidents (Baddeley et al. 2021; D'Angelo et al. 2024), visitors' stops at tourist attractions (D'Angelo et al. 2022), ambulance interventions on a road network (Gilardi, Borgoni, and Mateu 2024).

1.2. Overview and Contributions

The contributions of this work are summarized as follows:

We propose a spatio-temporal point process model, known as the Convolutional Non-homogeneous Poisson Process (cNHPP), on a linear network. Based on the proposed model, the event process on each segment of the linear network is modeled as an NHPP with its log-intensity being given by the sum of an infinite series. For each segment i of the network, the model captures (i) how current covariates, associated with segment i , affect the event intensity of segment i (i.e., the short-term instantaneous effect), and (ii) how historical covariates, associated with segment i and its neighboring segments, together affect the event intensity of segment i (i.e., the long-term cumulative effects) given network topology and the spatio-temporal dependency among segments. In particular, the current covariates affect the event intensity through a log-linear model, while the historical covariates affect the current event intensity through a convolution approach. This extends the conventional log-linear NHPP model with its intensity function only depending on the current covariates; see Section 2.1.

In Sections 2.2 and 2.3, we provide additional insights and discussions on the proposed cNHPP. In particular, we present detailed investigations on the computation of the proposed

intensity function, the graphical representation of the proposed model, and the continuous-time representation of the model. In supplement B, we provide additional discussions on an interesting connection between the proposed cNHPP and the architecture of a simple Recurrent Neural Network (RNN).

In Section 3, using the environmental data obtained from NOAA and NASA, we apply the proposed approach to model and predict the wildfire ignition risks on major transmission lines in part of California. By investigating the estimated effects of different covariates on the ignition of fires on transmission lines, we obtain useful insights and recommendations that potentially enhance power grid resilience against fires. This section also includes a comprehensive comparison study among different approaches. Section 4 concludes the article.

2. Convolutional NHPP

2.1. Basic Model Formulation

Consider a linear network L with N segments; for example, power transmission lines. Let l_i be the i th segment, and the network is represented by $L = \bigcup_{i=1}^N l_i$. In this article, we consider events that occur on segments of a network, and the event process on the i th segment is modeled as a point process with a conditional intensity function:

$$\lambda(i, t | \mathcal{H}_t) = \lim_{\Delta \rightarrow 0} \frac{\mathbb{E}[N_i([t, t + \Delta)) | \mathcal{H}_t]}{\Delta}, \quad (1)$$

where $N_i([t, t + \Delta))$ is a counting measure on the i th segment over the time interval $[t, t + \Delta)$, and $\mathcal{H}_t$ represents the event history on the entire network L (which is omitted thereafter). In particular, we assume that the event process on each segment i is an NHPP with the intensity function $\lambda(i, t)$. Hence, for any $\tilde{L} \subseteq L$, the total number of events on the sub-network $\tilde{L}$ over a time interval $[t, t + \Delta)$ follows a Poisson distribution with a parameter $\int_t^{t+\Delta} \sum_{i \in \tilde{L}} \lambda(i, u) du$. Because segments l_i , $i = 1, \dots, N$, in a network are disjoint, we have

$$\Pr(N_i[t, t + \Delta) = n_i, i = 1, 2, \dots, N) = \prod_i \frac{\Lambda^{n_i}}{n_i!} e^{-\Lambda}, \quad (2)$$

where n_i is a nonnegative integer and $\Lambda = \int_t^{t+\Delta} \lambda(i, u) du$. For power line fire risk quantification, for example, the equation above allows us to evaluate the probabilities of different fire scenarios over a network.

Hence, the construction of the intensity function $\lambda(i, t)$ becomes critical. The goal of this article is to construct a statistical model that adequately explains the intensity function $\lambda(i, t)$ by taking into account both the short-term (i.e., instantaneous) and long-term (i.e., cumulative) effects of covariates, as well as the spatio-temporal dependency among multiple segments. The following model is proposed,

$$\log \lambda(i, t) = \underbrace{c(i, t)}_{\text{current effects}} + \underbrace{h(i, t)}_{\text{cumulative historical and spatial effects}} \quad (3)$$

which decomposes the log-intensity into two additive components. The first component $c(i, t)$ captures the current effects of covariates at time t , while the second component $h(i, t)$ captures

the cumulative effects of historical covariate information (before time t) through the spatio-temporal interactions among neighboring network segments. To elaborate,

- $c(i, t)$ incorporates the current effects of covariates at time t for segment i through a linear model $c(i, t) = \mathbf{x}^T(i, t)\boldsymbol{\beta}$, where $\mathbf{x}(i, t) = (1, x_1(i, t), x_2(i, t), \dots, x_q(i, t))^T$ denote the covariates associated with segment i at time t , $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_q)^T$ is a vector of coefficients, and q is the number of covariates. In the application presented in Section 3, potential covariates for fire events include vegetation and meteorological variables.

- $h(i, t)$ captures the long-term cumulative effects as well as the spatio-temporal interactions among segments. In other words, it explains how historical covariates (before time t) associated with the neighboring segments of segment i influence the intensity of segment i at time t . We model $h(i, t)$ in the following way:

$$h(i, t) = \xi \sum_{i' \in \Omega_i} w_{ii'} \log \lambda(i', t - \Delta), \quad (4)$$

where Ω_i is the set that contains pre-defined neighboring segments of the i th segment, $w_{ii'}$ is the contribution (i.e., weight) to $h(i, t)$ from the i' th segment from time $t - \Delta$, and $\xi \in [0, 1]$ is the decay factor that controls the rate of decay of the cumulative effects. A smaller ξ indicates that the spatial cumulative effect quickly decays, while a larger ξ makes the current intensity to be dependent more on historical intensities. A larger ξ makes $\lambda(i, t)$ smoother in time. This is similar to the weighted moving average that controls the smoothness by distributing weights to current and historical observations (Perry 2010).

Note that, the model (4) implies that the intensity at time t depends not only on the intensity at time $t - \Delta$, but also on the intensities at times $t - 2\Delta, t - 3\Delta, \dots, -\infty$. One may see this by replacing $\log \lambda(i', t - \Delta)$ in (4) by $c(i', t - \Delta) + h(i', t - \Delta)$. By iterating this process, the intensity of each segment depends on that of neighboring segments over the entire history. To show how the spatio-temporal dependency structure among $\{\lambda(i, t)\}_{i=1}^N$ is established by (4), for a function $f(i, t): \{1, 2, \dots, N\} \times [0, T] \mapsto \mathbb{R}^+$, we introduce a *Network Convolution* operator $\mathcal{NC}$ as follows:

$$\mathcal{NC}\{f\}(i, t) = \sum_{i' \in \Omega_i} w_{ii'} f(i', t), \quad (5)$$

where $\mathcal{NC}\{f\}(i, t): \{1, 2, \dots, N\} \times [0, T] \mapsto \mathbb{R}^+$ is a new function generated by the operator $\mathcal{NC}$. Note that, $\mathcal{NC}$ is not strictly a mathematical convolution operation, but rather a weighted sum of the outputs from a set of functions (i.e., $h(i, t)$ in (3) is given by a weighted sum of the log intensities from a set of segments at time $t - \Delta$). Recall a Convolutional Neural Network (CNN) where the input to a node is the weighted sum of the outputs from a set of nodes in the previous layer. Although this is not strictly a mathematical convolution neither, the term convolution is used. For this reason, we still call $\mathcal{NC}$ a network convolution operator, given the network structure that determines the set Ω_i . Because the operator $\mathcal{NC}$ is linear, we may also define $\mathcal{NC}^{(n)}\{f\}(i, t)$ as the n -fold network convolution for a function $f(i, t)$. For example, applying $\mathcal{NC}$ to $f(i, t)$ twice (i.e.,

a 2-fold operation) yields:

$$\begin{aligned}\mathcal{NC}^{(2)}\{f\}(i, t) &= \mathcal{NC}\{\mathcal{NC}\{f\}\}(i, t) = \mathcal{NC}\left\{\sum_{i' \in \Omega_i} w_{ii'} f(i', t)\right\} \\ &= \sum_{i' \in \Omega_i} w_{ii'} \mathcal{NC}\{f\}(i', t) \\ &= \sum_{i' \in \Omega_i} w_{ii'} \sum_{i'' \in \Omega_{i'}} w_{i'i''} f(i'', t).\end{aligned}\tag{6}$$

Then, it follows from (5), $h(i, t)$ in (4) can be written as $\sum_{n=1}^{\infty} \xi^n \mathcal{NC}^{(n)}\{c\}(i, t - n\Delta)$. The derivation of $h(i, t)$ involves a long equation that is provided in the supplement A. Finally, by substituting the expression of $h(i, t)$ into (3), we obtain the proposed statistical model as follows:

Convolutional NHPP. Consider a linear network $L = \cup_{i=1}^N l_i$ with N disjoint segments. The event process on each segment i is modeled as a non-homogeneous Poisson process (NHPP) with its log intensity being given by the sum of an infinite series:

$$\begin{aligned}\log \lambda(i, t) &= c(i, t) + h(i, t) \\ &= \mathcal{NC}^{(0)}\{c\}(i, t) + \sum_{n=1}^{\infty} \xi^n \mathcal{NC}^{(n)}\{c\}(i, t - n\Delta) \\ &= \sum_{n=0}^{\infty} \xi^n \mathcal{NC}^{(n)}\{c\}(i, t - n\Delta),\end{aligned}\tag{7}$$

where $\mathcal{NC}^{(0)}\{c\}(i, t) \triangleq c(i, t)$.

It is seen that, (i) the intensity of segment i at time t depends on not only the covariates associated with segment i at time t (i.e., the current effect), but also the historical covariates associated with the neighboring segments that are involved through a sequence of network convolution operations $\{\mathcal{NC}^{(n)}\}_{n=1}^{\infty}$ prior to time t (i.e., the cumulative historical and spatio-temporal dependency). (ii) $\log \lambda(i, t)$ is represented by the sum of an infinite series. Each term $\xi^n \mathcal{NC}^{(n)}\{c\}(i, t - n\Delta)$ corresponds to the contribution to $\log \lambda(i, t)$ from $c(\cdot, t - n\Delta)$ at the current or a historical time $t - n\Delta$, $n = 0, 1, 2, \dots, \infty$. Because the contribution from $t - n\Delta$ decays to zero as n increases for $\xi \in [0, 1)$, it is possible to approximate the series by only retaining the first K terms. In the next section, we present a graphical

representation of the model and illustrate the computation of $\log \lambda(i, t)$ leveraging such a graphical representation.

2.2. Computation, Graphical Representation and Likelihood

In this section, we show how the computation of $\log \lambda(i, t)$ can leverage a natural graphical representation of (7), and present the likelihood function needed for parameter estimation. Some notation and network operations are first introduced.

For any segment i , let $\Omega_i^{(m)}$ be a set that contains the indices of the m th generation neighbors of segment i ($m = 0, 1, 2, \dots, K$). As illustrated in Figure 2, $\Omega_i^{(0)} = \{i\}$ contains segment i itself, $\Omega_i^{(1)}$ contains the immediate neighbors of segment i , $\Omega_i^{(2)}$ contains the neighbors of the neighbors of segment i , and so on.

In addition, for any segment j in the set $\Omega_i^{(m)}$, we define the ancestor operator $\mathcal{A}^{(1)}(j)$ that returns the parent segment of j . Similarly, we may introduce the 2-fold ancestor operation $\mathcal{A}^{(2)}(j)$ that returns the grand parent of j . By extending this idea, we let $\mathcal{A}^{(n)}(j)$ denote the n -fold ancestor operation, and let $\mathcal{A}^{(0)}(j)$ return j itself; see Figure 2. It is easy to see that, for any line segment j that belongs to the m th generation neighbor set $\Omega_i^{(m)}$ of segment i , the m -fold ancestor operation of j returns i , that is, $\mathcal{A}^{(m)}(j) = i$ for $j \in \Omega_i^{(m)}$.

Based on the notation and operation defined above, we show that the computation of $\log \lambda(i, t)$ takes a natural graphical representation.

- (contribution from time t) The contribution to $\log \lambda(i, t)$ at time t directly comes from $c(i, t)$, which is the first term of the series (7).
- (contribution from time $t - \Delta$) The contribution to $\log \lambda(i, t)$ from $c(\cdot, t - \Delta)$ at time $t - \Delta$ is associated with all neighboring segments in $\Omega_i^{(1)}$, that is, the second term of the series (7) can be computed by

$$\begin{aligned}&\xi \mathcal{NC}^{(1)}\{c\}(i, t - \Delta) \\ &= \xi \sum_{j \in \Omega_i^{(1)}} w_{\mathcal{A}^{(1)}(j), \mathcal{A}^{(0)}(j)} c(j, t - \Delta).\end{aligned}\tag{8}$$

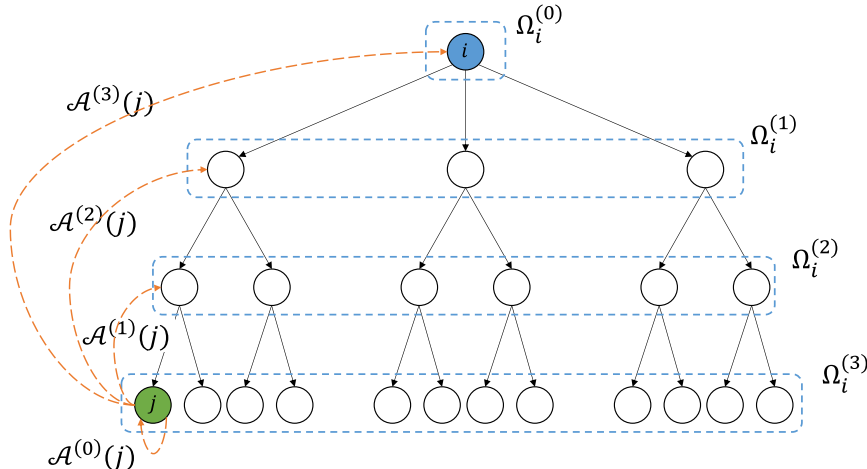


Figure 2. A graphical illustration of the neighbor set and ancestor operation with each node representing a segment: (i) $\Omega_i^{(0)}$ is segment i itself, $\Omega_i^{(1)}$ contains the neighbors of segment i , $\Omega_i^{(2)}$ contains the neighbors of the neighbors of i , and so on; (ii) $\mathcal{A}^{(0)}(j)$ returns j itself, $\mathcal{A}^{(1)}(j)$ returns the parent of j , $\mathcal{A}^{(2)}(j)$ returns the grandparent of j , and so on.

- (contribution from time $t - 2\Delta$) The contribution to $\log \lambda(i, t)$ from $c(\cdot, t - 2\Delta)$ at time $t - 2\Delta$ is associated with all neighboring segments in $\Omega_i^{(2)}$, that is, the third term of the series (7) can be computed by

$$\xi^2 \mathcal{N}C^{(2)}\{c\}(i, t - 2\Delta) = \xi^2 \sum_{j \in \Omega_i^{(2)}} w_{\mathcal{A}^{(2)}(j), \mathcal{A}^{(1)}(j)} w_{\mathcal{A}^{(1)}(j), \mathcal{A}^{(0)}(j)} c(j, t - 2\Delta). \quad (9)$$

- (contribution from time $t - n\Delta$) For any given $n \in \mathbb{N}^+$, it is seen that the generic expression of the contribution to $\log \lambda(i, t)$ from $c(\cdot, t - n\Delta)$ is associated with all neighboring segments in $\Omega_i^{(n)}$ and can be written as

$$\xi^n \mathcal{N}C^{(n)}\{c\}(i, t - n\Delta) = \xi^n \sum_{j \in \Omega_i^{(n)}} \left\{ \prod_{p=1}^n w_{\mathcal{A}^{(p)}(j), \mathcal{A}^{(p-1)}(j)} \right\} c(j, t - n\Delta). \quad (10)$$

The generic expression (10) above provides a way to compute each term in the series (7). When the series (7) is approximated by only retaining the first K terms, we have

$$\begin{aligned} \log \lambda(i, t) &= \sum_{n=0}^{\infty} \xi^n \mathcal{N}C^{(n)}\{c\}(i, t - n\Delta) \\ &\approx c(i, t) + \sum_{n=1}^K \xi^n \sum_{j \in \Omega_i^{(n)}} \\ &\quad \times \left\{ \prod_{p=1}^n w_{\mathcal{A}^{(p)}(j), \mathcal{A}^{(p-1)}(j)} \right\} c(j, t - n\Delta). \end{aligned} \quad (11)$$

Figure 3 provides a graphical illustration of (11) that depicts the discussions above.

Let $\log \lambda(t) = (\log \lambda(1, t), \log \lambda(2, t), \dots, \log \lambda(N, t))^T$ be a vector that contains the log intensity functions on all N segments of a network L . For $k = 0, 1, \dots, K$, let $c(t - k\Delta) = (c(1, t - k\Delta), c(2, t - k\Delta), \dots, c(N, t - k\Delta))^T$. Then, the approximated log intensity $\log \lambda(t)$ (by retaining the first K terms in the series (7)) admits the following matrix form:

$$\log \lambda(t) \approx c(t) + \xi W c(t - \Delta) + \dots + \xi^K W^K c(t - K\Delta), \quad (12)$$

where W is an $N \times N$ weight matrix with its (i, j) th entry, w_{ij} , being the contribution weight to $h(i, t)$ from $\log \lambda(j, t - \Delta)$, and W^K is the K th power of W . Note that,

(i) because the intensity on a segment i is only affected by its neighboring segments, W is sparse with $w_{ij} = 0$ for j that is not a neighbor of segment i . Hence, the sizes of the neighboring sets Ω control the sparsity of the problem;

(ii) the proposed model provides some flexibility about how the weights are chosen. For example, $w_{ii'} = |\Omega_i|^{-1}$, that is, all neighboring segments of segment i equally contribute to the i th segment. The parametric form such as $w_{ii'} = \frac{\exp(-d_{ii'}/l)}{\sum_{i' \in \Omega_i} \exp(-d_{ii'}/l)}$, where $d_{ii'}$ is the Euclidean distance between the centers of segments i and i' and l is the distance scale parameter;

(iii) the matrix form (12) can be understood from the perspective of graph theory. Note that, a nonzero weight $w_{ij}^{(k)}$ in W^k implies that segment j can reach segment i with a k -step walk in the network. In addition, the value of $w_{ij}^{(k)}$ in W^k is the sum of the contribution weights that are the multiplications of $w_{\cdot, \cdot}$ of the linked segments along all possible k -step walks from segment j to segment i . Thus, the contribution of $c(t - k\Delta)$ to $\log \lambda(t)$ naturally admits the expression of $\xi^k W^k c(t - k\Delta)$. For example, $w_{ij}^{(2)} \neq 0$ in W^2 indicates that segment j can walk to segment i with two steps, and the value of $w_{ij}^{(2)}$ is the total contribution from all possible two-step walks from segment j to segment i in the network. Then, the corresponding contribution of $c(t - 2\Delta)$ to $\log \lambda(t)$ is $\xi^2 W^2 c(t - 2\Delta)$.

Finally, when $c(i, t)$ is modeled by a linear function of covariates, we obtain a linear model for $\log \lambda(t)$ that incorporates covariate information:

$$\begin{aligned} \log \lambda(t) &\approx X(t)\beta + \xi W X(t - \Delta)\beta + \dots + \xi^K W^K X(t - K\Delta)\beta \\ &= \left(\sum_{k=0}^K \xi^k W^k X(t - k\Delta) \right) \beta \triangleq \tilde{X}(t)\beta, \end{aligned} \quad (13)$$

where $X(t) = (x(1, t), \dots, x(N, t))^T$ is the covariate matrix at time t , and $\tilde{X}(t) = (\tilde{x}(1, t), \dots, \tilde{x}(N, t))^T$ is the transformed covariate matrix after the convolution operation, and we call it the Convolutional Covariate Matrix (CCM) in this article.

Note that, (i) if W is an identity matrix, the intensity of each line segment only depends on its own historical information,

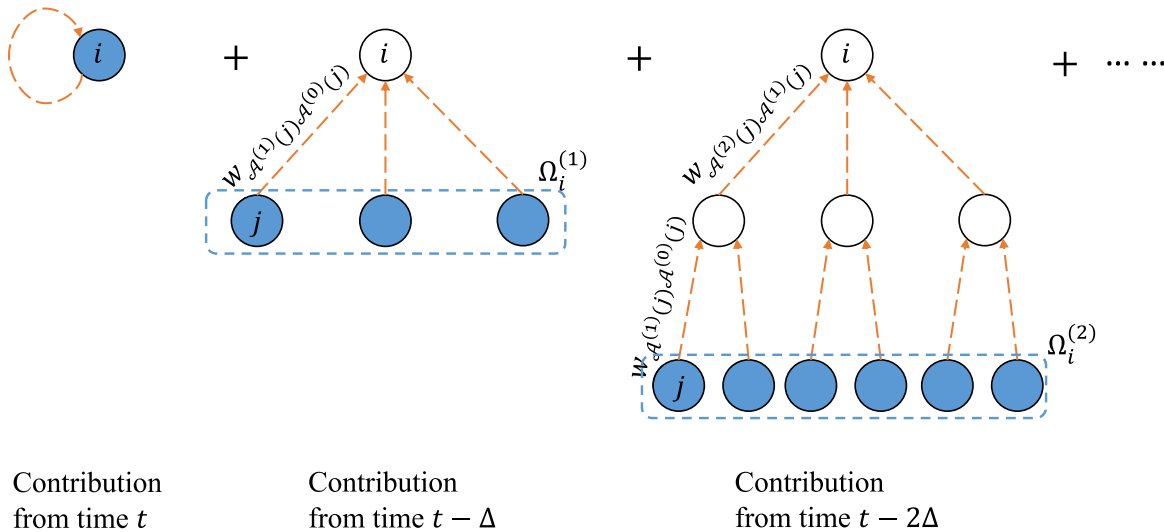


Figure 3. A graphical illustration of how the intensity $c(i, t)$ is contributed from the neighbor sets of i at times $t, t - \Delta, t - 2\Delta, \dots$

and there is no spatial dependency among line segments; (ii) if all elements in $\mathbf{W}$ are zeros, there exist no historical effects nor spatial dependency among the intensity functions. In this case, the intensity functions can be completely explained by current covariates, and the proposed model (3) degenerates to $\log \lambda(i, t) = \mathbf{x}^T(i, t)\boldsymbol{\beta}$, which is widely used in existing NHPP models; (iii) the spatio-temporal dependency is embedded in $\mathbf{W}$, and such a dependency spans over all segments given the $\mathcal{NC}$ operator defined in (5). In particular, $\{\mathbf{W}^m\}_{m=1}^K$ in (13) are induced by $\mathcal{NC}$ on the linear network considering the past K time steps.

Based on (13), we obtain the log-likelihood function for the unknown parameters, including the decay factor ξ and coefficients $\beta_0, \beta_1, \dots, \beta_q$. Let $\boldsymbol{\theta} = (\xi, \beta_0, \beta_1, \dots, \beta_q)^T$, we have (Peng, Schoenberg, and Woods 2005):

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^N \sum_{j=1}^{b_i} \log \lambda(i, t_j) - \sum_{i=1}^N \int_0^T \lambda(i, t) dt \\ &= \sum_{i=1}^N \sum_{j=1}^{b_i} \tilde{\mathbf{x}}^T(i, t_j)\boldsymbol{\beta} - \sum_{i=1}^N \int_0^T \exp[\tilde{\mathbf{x}}^T(i, t)\boldsymbol{\beta}] dt, \end{aligned} \quad (14)$$

where b_i is the number of events on segment i , and T is the length of the observation period.

Finally, the linear form in (13) suggests that the intensity at time t depends on the historical and current covariates, and this structural characteristic enables us to draw an interesting connection between the proposed cNHPP and a simple Recurrent Neural Networks (RNN). Details are provided in supplement B.

2.3. Two Additional Remarks

In the section, we make two important remarks on the proposed cNHPP model.

2.3.1. The Continuous-Time Form of the Intensity Function (7)

For the proposed cNHPP model, it is extremely important to note that ξ , $c(i, t)$, and $\mathcal{NC}$ all implicitly depend on the choice of Δ (i.e., scale-dependent). For example, if Δ is defined as a day, then ξ and $c(i, t)$ are, respectively, the daily decay and daily generation, and the operator $\mathcal{NC}$ operates over a *one-day interval*. Hence, to obtain the continuous-time form of the intensity function (7), we first need to make such a dependency explicit by letting ξ_δ and $c_\delta(i, t)$, respectively, be the total amount of decay and generation over a δ -interval, and letting $\mathcal{NC}_\delta = \sum_{i' \in \Omega_{i,\delta}} w_{i,i'}^\delta f(i', t)$ be the network convolution operator that operates over a δ -interval. Note that, the neighbor set $\Omega_{i,\delta}$ also depends on the length of the interval as we naturally expect the size of this set to be nondecreasing in δ . In fact, a stronger condition needs to be imposed on $\Omega_{i,\delta}$ as to be discussed below. When the neighbor set depends on δ , the weight $w_{i,i'}^\delta$ is also scale-dependent.

Next, let $\Delta = n\delta$, two assumptions are needed: (i) $\xi_\Delta = \xi_\delta^n = \exp(-bn\delta)$ for some $b > 0$, and (ii) $\mathcal{NC}_\Delta\{f\}(i, t) = \mathcal{NC}_{n\delta}\{f\}(i, t) = \mathcal{NC}_\delta^{(n)}\{f\}(i, t)$. The second assumption immediately imposes two consistency conditions on how neighbor sets are defined over the network, as well as how the weights

associated with the operator $\mathcal{NC}$ are determined for different scales of the time intervals.

Condition 1. the n th generation neighbor set corresponding to a time interval δ needs to be the same as the first generation neighbor set corresponding to a time interval Δ , that is, $\Omega_{i,\delta}^{(n)} = \Omega_{i,\Delta}$.

Condition 2. for any segment i , the weight w_{ij}^Δ between i and its first generation neighbor $j \in \Omega_{i,\Delta}$ (given a time interval Δ) needs to be the same as the weight between i and its n th generation neighbor $j \in \Omega_{i,\delta}^{(n)}$ (given a time interval δ), that is, $w_{ij}^\Delta = \prod_{p=1}^n w_{\mathcal{A}^{(p)}(j), \mathcal{A}^{(p-1)}(j)}^\delta$.

Having made the two assumptions above, the intensity function (7) can be written as

$$\begin{aligned} \log \lambda(i, t) &= \xi_\Delta \sum_{i' \in \Omega_{i,\Delta}} w_{i,i'}^\Delta \log \lambda(i', t - \Delta) + c_\Delta(i, t) \\ &= \xi_\delta^n \mathcal{NC}_\delta^{(n)}\{\log \lambda\}(i, t - n\delta) \\ &\quad + \sum_{k=1}^{n-1} \xi_\delta^k \mathcal{NC}_\delta^{(k)}\{c_\delta\}(i, t - k\delta) + c_\delta(i, t). \end{aligned} \quad (15)$$

Following an approach described in Brown et al. (2000), we let $n \rightarrow \infty$, and the first term on the right hand side, $\xi_\delta^n \mathcal{NC}_\delta^{(n)}\{\log \lambda\}(i, t - \Delta)$, goes to zero, that is,

$$\log \lambda(i, t) \approx \sum_{k=1}^{\infty} \xi_\delta^k \mathcal{NC}_\delta^{(k)}\{c_\delta\}(i, t - k\delta) + c_\delta(i, t). \quad (16)$$

Then, by letting $\delta \rightarrow 0$, we obtain the continuous-time representation of the intensity function (7)

$$\log \lambda(i, t) = \int_{v=0}^{\infty} \exp(-bv) \mathcal{NC}_v\{\tilde{c}\}(i, t - v) dv$$

where $\tilde{c}(i, t) = \lim_{\delta \rightarrow 0} \delta^{-1} c_\delta(i, t)$ is the rate.

2.3.2. The Difference between cNHPP and Other Models

The proposed cNHPP model not only distinguishes itself from the linear and nonlinear Hawkes processes (defined in (17)), but also extends the conventional NHPP model with a log-linear intensity function (18):

$$\lambda(i, t) = \phi_i(\mu_i + \sum_{t_j: t_j < t} \gamma_i(t - t_j)) \quad (17)$$

$$\lambda(i, t) = \exp(\mathbf{x}^T(i, t)\boldsymbol{\beta}), \quad (18)$$

where $\phi_i(\cdot)$ is a nonnegative link function that can be either linear or nonlinear, μ_i is the baseline intensity, $\gamma_i(\cdot)$ is the triggering kernel function (e.g., the exponential kernel), and $\{t_j : t_j < t\}$ contains the event times before time t .

In the linear and nonlinear Hawkes processes, the intensity function only considers the influence of past events rather than the covariate information (Hawkes 2018). In contrast, the proposed cNHPP directly establishes an explanatory model for the intensity function that incorporates the current and historical covariate information. This yields more interpretable outcomes for assessing wildfire risks, aiding utilities in enhancing wildfire mitigation strategies and power grid operation planning based on real-time environmental data. For this reason, the Hawkes process is known as a self-exciting process, but the proposed

cNHPP is not a self-exciting process. Specifically when $n = 0$ in (7) or $\xi = 0$ in (4), the proposed cNHPP degenerates to the conventional NHPP model with a log-linear intensity function, which only depends on the current covariate information (Cox 1972).

3. Application: Wildfire Ignition Risks on California Power Transmission Lines

We apply the proposed approach for modeling and predicting wildfire risks on networks of power transmission lines in part of California. Section 3.1 provides detailed descriptions of the datasets. Section 3.2 presents the model outputs and provides some useful insights on the impact of wildfires on power delivery infrastructures. Comparison studies are presented in Section 3.3.

3.1. Data

This application example involves four major datasets: (i) power delivery infrastructure data, (ii) wildfire incident data, (iii) meteorological data, and (iv) vegetation data. Figure 4(a) shows the main power transmission lines in California. This dataset is obtained, in the shapefile format, from the U.S. Energy Information Administration (EIA). In particular, we focus on a spatial area indicated by the square in Figure 4(a). This spatial area is defined by $[120^\circ\text{W}, 119^\circ\text{W}] \times [36^\circ\text{N}, 37^\circ\text{N}]$, and there is a total number of 6398 power transmission lines within this area; see Figure 4(b). Figure 4(c) shows the histogram of the lengths of these 6398 segments, and most of these line segments are less than 1000 meters.

The fire incident dataset contains the information about overhead power-line fires, including fire locations and dates. The wildfire incident data are obtained from the California Public Utilities Commission (CPUC), which is a government agency that regulates public utility companies including privately owned electric, natural gas, and telecommunications companies. In response to increasingly severe overhead power-line wildfires, CPUC requires electricity companies to report data on power-line fires. We obtain the fire incident data reported from the Southern California Edison (SCE), Pacific Gas and Electric (PG&E), and San Diego Gas and Electric (SDG&E) companies. From June 1 to June 30, 2019, a total number of 15

wildfires were reported within the study area. On average, there was a fire incident every two days due to the high temperature and dry weather at the beginning of the summer season.

Meteorological data are obtained from the High-Resolution Rapid Refresh (HRRR) model maintained by NOAA. The HRRR model provides hourly meteorological data with a spatial resolution of 3 kilometers. Although HRRR contains 170 meteorological variables in 2D surface levels, most of these variables are the same meteorological conditions but in different pressure regions. Hence, we select three representative variables, including temperature (2 meters above ground), specific humidity (2 meters above ground), and wind speed (10 meters above ground). For convenience, we denote temperature as TMP ($^\circ\text{C}$), specific humidity as SPFH ($\text{kg} \cdot \text{kg}^{-1}$), and wind speed as WIND ($\text{m} \cdot \text{s}^{-1}$). Because power lines do not always locate in a regular grid, the meteorological data at the nearest grid points are assigned to each power line segment. As an illustration, Figure 5(a)–(c) show the meteorological conditions at 12:00 p.m. UTC on June 01, 2019.

Vegetation data are obtained from the Normalized Difference Vegetation Index (NDVI) that indicates the vegetation-water status. A higher NDVI corresponds to a denser and healthier vegetation canopy and vice versa. This dataset is obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board NASA's Aqua and Terra satellites. Because MODIS only has 8-day NDVI data products, we manually calculate daily NDVI values using the daily land surface reflectance products by $\text{NDVI} = (\rho_{\text{NIR}} - \rho_{\text{red}}) / (\rho_{\text{NIR}} + \rho_{\text{red}})$ where ρ_{red} and ρ_{NIR} , respectively, denote the reflectances of near-infrared and red spectral regions. Detailed descriptions about the reflectance products can be found in MODIS (2015). Here, ρ_{red} and ρ_{NIR} have a 250-m spatial resolution in MODIS land surface reflectance products, and so do the processed NDVI data. Figure 5(d) gives an example of the processed NDVI values that range from 0 to 1.

3.2. Results and Discussions

Based on the power transmission line data above, $\lambda(t)$ is a column vector that contains the intensity functions of the 6398 power lines. The convolutional covariate matrix $\tilde{X}$ is a 6398×6 matrix, and the sparse weighted adjacency matrix W has a

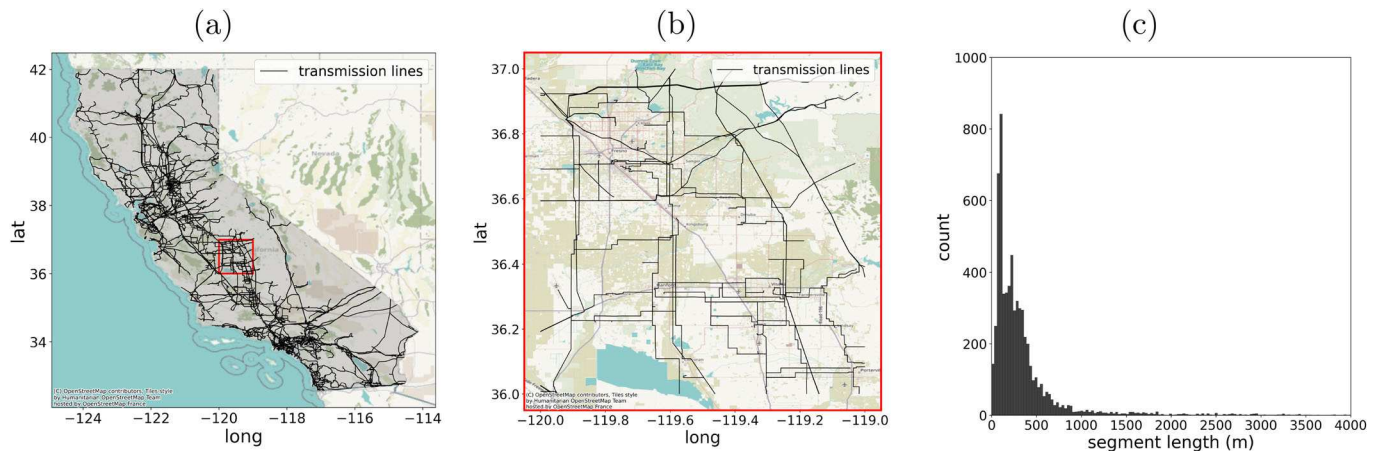


Figure 4. (a) Main power transmission lines in California; (b) Power transmission lines in the study area; (c) Histogram of the length of line segments in the study area.

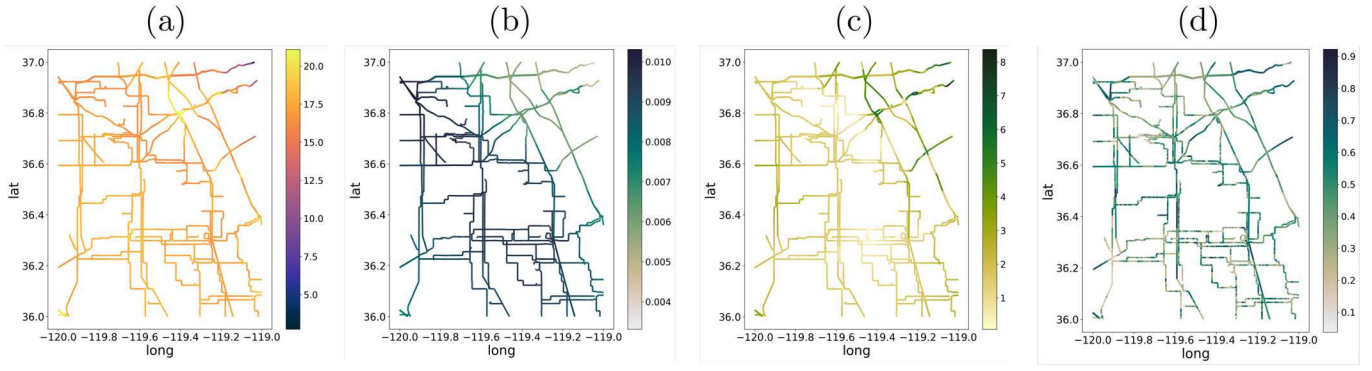


Figure 5. Illustration of three meteorological variables and the processed vegetation data on June 01, 2019: (a) TMP ($^{\circ}\text{C}$); (b) SPFH ($\text{kg} \cdot \text{kg}^{-1}$); (c) WIND ($\text{m} \cdot \text{s}^{-1}$); (d) Normalized Difference Vegetation Index (NDVI).

dimension of 6398×6398 . In this application, we let $w_{ii'} = 1/|\Omega_i|$ for $i' \in \Omega_i$, implying that all neighbors of segment i equally contribute to the fire intensity of the i th segment. Here, our focus is on daily wildfire risk, and Δ equals one day.

We let $x_1(i, t)$, $x_2(i, t)$, $x_3(i, t)$, $x_4(i, t)$, and $x_5(i, t)$, respectively, denote the NDVI, TMP, WIND, SPFH, and the LENGTH of the i th segment at day t . Specifically, we use the TMP, WIND, and SPFH from the HRRR model at 12:00 p.m. UTC to represent their daily quantities. All covariates are standardized so as to facilitate the comparison between their effects β_1 , β_2 , β_3 , β_4 , and β_5 . An intercept β_0 is also included. The parameters $\theta = (\xi, \beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$ in (13) are estimated by maximizing the log-likelihood function (14). Note that, the log-likelihood function (14) is concave with respect to β , but the computational cost of $\tilde{X}(t)$ for different values of ξ is extremely high. Hence, we adopt a more practical approach by considering a finite number of ξ (for computing $\tilde{X}(t)$), and compare the corresponding maximized log-likelihoods with respect to $\hat{\beta}$. In particular, we consider a finite set $\{0.1, 0.2, \dots, 0.9\}$ for ξ , and let $K = 7$ meaning that the historical covariate information from the past one week is considered. For each ξ , $\tilde{X}(t)$ is computed, and the MLE of β is found using the `SciPy` package using the L-BFGS-B method in Python. The maximum log-likelihood is found when $\xi = 0.7$, and the corresponding $\hat{\beta}$ is found as $\hat{\beta}_0 = -2.760$ ($-3.109, -2.412$), $\hat{\beta}_1 = -1.521$ ($-1.795, -1.248$), $\hat{\beta}_2 = 0.541$ ($0.058, 1.025$), $\hat{\beta}_3 = 0.860$ ($0.407, 1.312$), $\hat{\beta}_4 = -0.470$ ($-0.871, -0.069$), and $\hat{\beta}_5 = 2.716$ ($2.356, 3.076$), where the values in (,) are the large-sample approximate 95% confidence bounds.

It is seen that higher temperature and stronger wind speed increase the wildfire risks by factors of 1.72 (i.e., $\exp(\hat{\beta}_2)$) and 2.36 (i.e., $\exp(\hat{\beta}_3)$), respectively. Such findings can be well justified as follows: (i) a higher temperature makes the ignition of the underlying fuels easier (e.g., grasses, shrubs, dead leaves, etc.), (ii) with an increased wind speed, the electrical conductors and surrounding vegetation are more likely to result in arcing, increasing the probability of wildfire ignition (Mitchell 2013; Vazquez et al. 2022). Both NDVI and SPFH have negative impacts on wildfire risks, while SPFH has a relatively weaker effect compared with that of NDVI. Note that, a higher NDVI indicates healthier vegetation with more water conditions and fewer potential fuels that can be ignited. Similarly, a high SPFH gives potential fuels more moisture, reducing the wildfire risk.

In fact, NDVI (with $\hat{\beta}_1 = -1.521$) is found to be more influential than the other three dynamical environmental variables (recall that all covariates are standardized). Activities such as weeding are highly recommended to eliminate unhealthy vegetation around power delivery infrastructures. Finally, LENGTH is found to be the most influential factor, with $\hat{\beta}_5 = 2.783$, which is consistent with our experience that longer lines are more susceptible to initiating a fire.

The proposed model can also be used for short-term forecasting of wildfire intensities. The forecasted wildfire intensities are obtained once the convolutional matrix $\tilde{X}(t+k)$ can be computed given future covariate values, that is, $\log \hat{\lambda}(t+k) = \tilde{X}(t+k)\hat{\beta}$. To demonstrate the predictive capabilities of our model, we treat the future covariate values as if they were known exactly and use their real values from our datasets. Figure 6 shows the estimated and forecasted wildfire event intensities based on the proposed approach. It is seen that different power line segments are associated with different wildfire risks due to the spatially- and temporally-varying covariate information. It is also seen from the second row of Figure 6 that the forecasted wildfire risks change smoothly over time. This is because the proposed model incorporates the cumulative long-term effects of covariates and the estimated decay factor $\hat{\xi} = 0.7$. As a result, the wildfire intensities do not dramatically change in a short period even if the current covariates change abruptly. Note that, if a smaller decay value ξ is obtained, the cumulative effects decay faster and the forecasted wildfire risks are more sensitive to current environmental conditions.

3.3. Comparison Studies

In this section, we first present the comparison based on the real datasets. Because there is a limited number of fire events, a simulation-based comparison is also presented.

3.3.1. Comparison based on the Real Data

Based on the real wildfire data, we compare the proposed cNHPP model with three models: (i) HPP: Homogeneous Poisson Process (HPP) that models the wildfire intensity as a constant over the entire network, that is, $\log \lambda(i, t) = \log \lambda$ for $i = 1, 2, \dots, N$; (ii) NHPP: Conventional NHPP model for which the wildfire intensity is only determined by the current covariates without

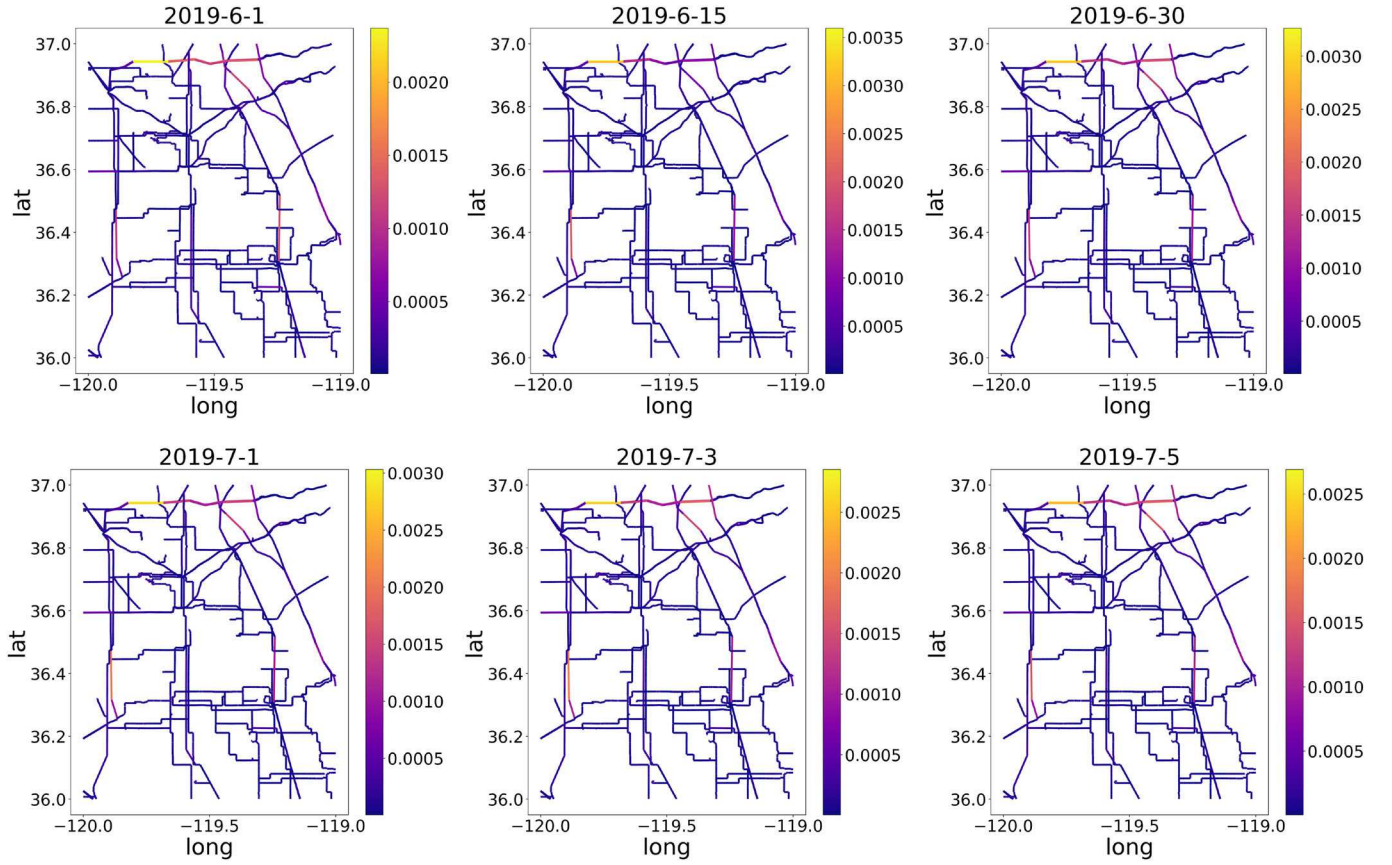


Figure 6. (Top row): estimated wildfire event intensities for power lines at June 01, 2019, June 15, 2019, and June 30, 2019; (Bottom row): one-day, three-day, and five-day ahead forecasts of wildfire event intensities.

Table 1. Estimated parameters from different models.

	HPP	NHPP	mRNN	cNHPP
$\hat{\lambda}$ (rate $\times 10^{-5}$)	7.815	—	—	—
$\hat{\xi}$ (decay)	—	—	0.670	0.700
$\hat{\beta}_0$ (intercept)	—	−9.023	−2.926	−2.760
$\hat{\beta}_1$ (NDVI)	—	−2.744	−1.538	−1.521
$\hat{\beta}_2$ (TMP)	—	0.598	0.830	0.541
$\hat{\beta}_3$ (WIND)	—	1.362	0.737	0.860
$\hat{\beta}_4$ (SPFH)	—	−0.343	−0.652	−0.470
$\hat{\beta}_5$ (LENGTH)	—	4.399	2.783	2.716
log-likelihood	−156.853	−153.695	−151.923	−151.864

accounting for the cumulative effects of covariates and spatial dependency, that is, $\log \lambda(i, t) = \mathbf{x}^T(i, t)\boldsymbol{\beta}$ for $i = 1, 2, \dots, N$; (iii) mRNN: The model-inherited RNN (mRNN) described in Supplement B and implemented by PyTorch with the Adam optimizer (learning rate = 0.001) in training the model. Convergence of the loss function and unknown parameters are provided in Supplement C, where a total number of 20,000 epochs are employed.

Table 1 presents the estimated model parameters using the same training data described in Section 3.2. We see from this table that cNHPP and mRNN yield larger log-likelihood than that of HPP and NHPP. Although NHPP, cNHPP, and mRNN give different estimated values for the effects $\boldsymbol{\beta}$, the signs of these estimates remain consistent. Based on the estimates in Table 1, we obtain the estimated intensity functions for all power-line segments from NHPP, mRNN, and cNHPP. Figure 7 shows the corresponding density plots of the estimated wildfire intensities

over the power lines on selected days. It is seen that density plots are all centered around the average $\hat{\lambda}$ obtained from the HPP model (the vertical dashed line), suggesting that all three approaches perform reasonably well in terms of estimating the underlying wildfire risks over the network of power transmission lines. It is also noted that the proposed cNHPP generates longer upper tails of the estimated intensity on two of the three days. This observation indicates that the proposed model may better capture some extreme cases and identify transmission lines with high risks.

We further validate the proposed approach in quantifying wildfire risks over the network of power lines. Note that, there are two challenges associated with validating the proposed approach: (i) because only 15 fire incidents are included in the training datasets and the majority of the power lines do not experience any fires, traditional measures (such as the mean prediction error, C-index, etc.) become less effective unless we have a much bigger dataset with a much larger number of fire incidents; (ii) power lines with high fire intensity may not always have fire incidents, while power lines with relatively low fire intensity can occasionally catch fire. Hence, we validate the capabilities of the proposed model by using a straightforward but interpretable procedure: for each fire incident, we first order the estimated fire intensities from the lowest to the highest for all power transmission lines, and then compute the percentile of the estimated intensity associated with the power line where the wildfire incident occurs. If the model works well, we expect the calculated percentiles associated with those power lines with

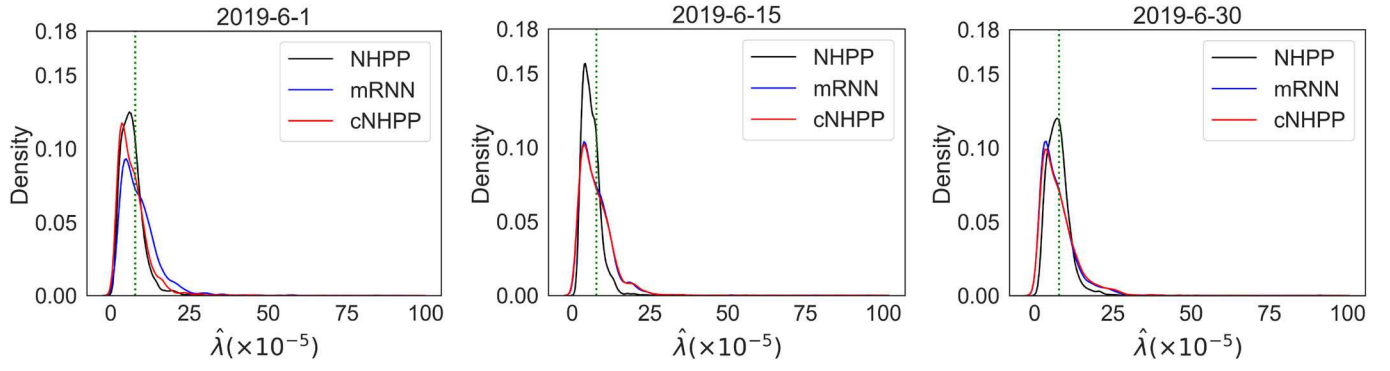


Figure 7. Distribution of estimated wildfire intensities on power transmission lines by different models at selected days (the dashed line denotes the estimated average wildfire intensity based on the NHPP model).

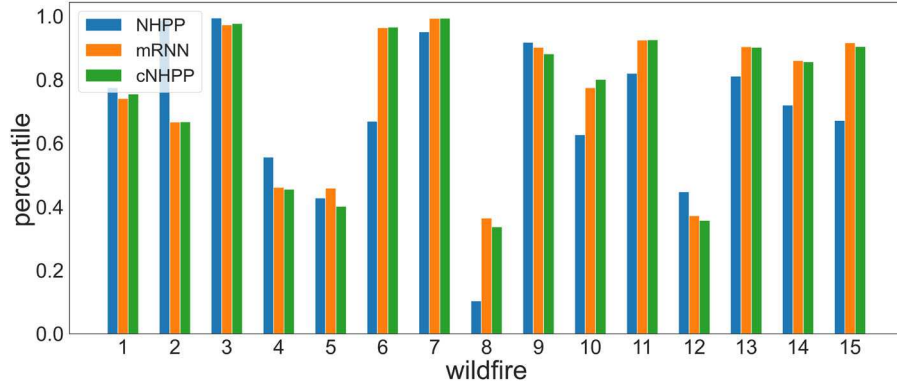


Figure 8. Percentiles of the estimated intensities associated with those power lines where fire events occur.

fire events to be high. The results are shown in Figure 8. It is seen that, among 8 out of the 15 wildfire incidents, cNHPP and mRNN yield higher percentiles than NHPP. This suggests that incorporating historical cumulative covariate effects and spatial dependency has the potential to improve the conventional NHPP model.

3.3.2. Comparison with the Simulated Data

Because the number of fire events in this dataset is not sufficient for a comprehensive comparison, a simulation-based comparison is performed in this section. To perform a comprehensive comparison study, a sufficiently large number of wildfire events are simulated, based on the parameters shown in Table 1, over the entire power transmission network and within a window of 37 days. In particular, we first compute two intensity functions, respectively, for the proposed cNHPP and the conventional NHPP. Then, we scale up these two estimated intensity functions by a factor of 5000 to generate two wildfire incident datasets. This simulation is based on the thinning approach for generating the NHPP (Lewis and Shedler 1979), and the two simulated datasets are, respectively, denoted as SimDataA and SimDataB.

In this simulation-based comparison, we compare the proposed approaches to the convectional NHPP model, as well as two models based on the Hawkes processes. The first Hawkes process, denoted as `Hawkes`, is the linear Hawkes process with an exponential triggering kernel, while the second Hawkes process, denoted as `nl-Hawkes`, is a nonlinear Hawkes process recently proposed by Zhou et al. (2022). To compare the

predictive capabilities of different models, data from the first 30 days are used for training while the data from the last 7 days are used for testing purposes.

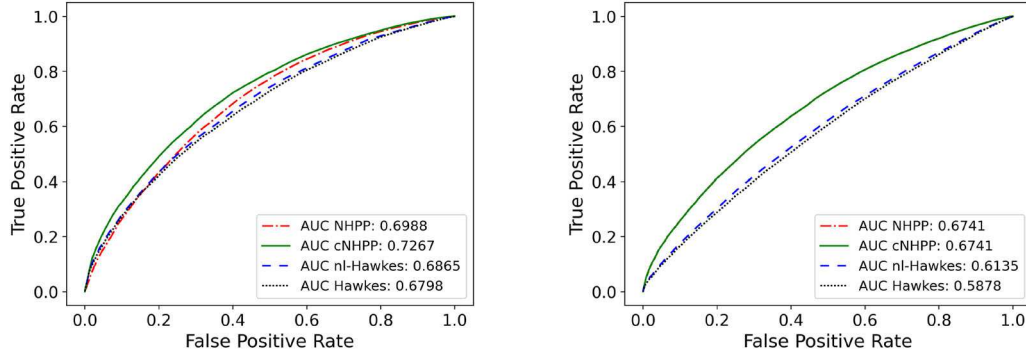
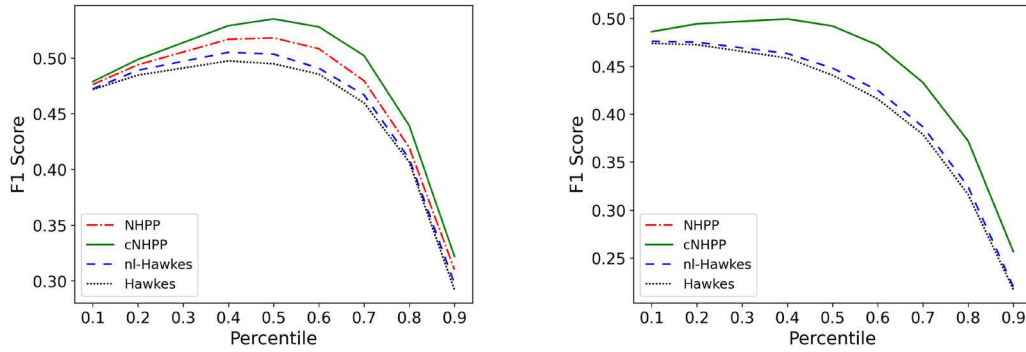
The Mean Absolute Prediction Error (MAPE) is adopted to evaluate the performance of different approaches. Here, the MAPE is defined as $MAPE = \frac{1}{N} \sum_{i=1}^N |N_i([t, t+1)) - \int_t^{t+1} \hat{\lambda}(i, u) du|$, where $N_i([t, t+1))$ represents the number of observed wildfires on day $t+1$ on the i th segment, and $\hat{\lambda}(i, u)$ is the predicted intensity function at time u for the i th segment. Table 2 shows the MAPE obtained from the four approaches based on the two simulated datasets. It is seen that the proposed cNHPP yields the lowest MAPE for 6 out of the 7 days in SimDataA, and ties for the lowest MAPE for 6 out of the 7 days with the NHPP in SimDataB. It is also extremely important to note that the proposed cNHPP and conventional NHPP approaches have exactly the same MAPE on dataset SimDataB. This is because the proposed cNHPP includes the NHPP as a special case (see Section 2.3.2), and SimDataB is generated from an NHPP. Hence, the proposed cNHPP can achieve a comparable prediction performance even when the simulated wildfire data are generated from the an NHPP model.

If a wildfire event is predicted to occur when the predicted wildfire intensity surpasses a specific threshold Xu and Schoenberg (2011), the Receiver Operating Characteristic (ROC) curves can also be constructed to assess different modeling approaches. The False Positive Rate (FPR) denotes the proportion of transmission lines without wildfires for which the predicted intensity exceeds the threshold, while the True Positive

Table 2. Comparison of MAPE for different approaches.

		Day 31	Day 32	Day 33	Day 34	Day 35	Day 36	Day 37	Total
cNHPP	(SimDataA)	0.460	0.492	0.459	0.453	0.433	0.423	0.526	1.242
NHPP		0.478	0.521	0.467	0.478	0.453	0.447	0.659	1.484
Hawkes		0.489	0.510	0.490	0.474	0.455	0.460	0.534	1.521
nl-Hawkes		0.472	0.496	0.472	0.470	0.455	0.446	0.524	1.442
cNHPP	(SimDataB)	0.468	0.511	0.449	0.464	0.465	0.444	0.658	1.305
NHPP		0.468	0.511	0.449	0.464	0.465	0.444	0.658	1.305
Hawkes		0.485	0.510	0.479	0.496	0.492	0.505	0.694	1.513
nl-Hawkes		0.492	0.515	0.479	0.486	0.490	0.474	0.685	1.479

NOTE: The “bold” highlights some good model performance.

**Figure 9.** ROC curves for different approaches based SimDataA (left) and SimDataB (right).**Figure 10.** F₁ scores for different approaches on the simulated dataset SimDataA (left) and simulated dataset SimDataB (right).

Rate (TPR) represents the proportion of transmission lines with fires for which the predicted intensity also exceeds the threshold. Both FPR and TPR are used to generate ROC curves shown in Figure 9, and the Area Under the Curve (AUC) is also reported. We observe that the proposed cNHPP achieves the highest AUC and a higher TPR for any given FPR in SimDataA, and it ties for the highest AUC and a higher TPR for any given FPR with the NHPP in SimDataA. Additionally, nl-Hawkes slightly outperforms Hawkes.

Finally, from the perspective of power grid operations, another performance metric can be used for model comparison. In practice, given the limited resources available for preventive actions, utilities often prioritize power lines with higher predicted intensities. If the model performs well, this strategy is expected to achieve a greater accuracy in directing preventive efforts to power lines with actual wildfires. In this context, we define True Positive (TP) as the percentage of power lines, for which preventive actions are scheduled, that actually catch fire; False Positive (FP) as the power lines where no fires occur but for which preventive actions are scheduled; and False Negative (FN) as the percentage of power lines that catch fires but for which preventive actions are not taken. Then, the F₁ score can

be calculated as follows: $F_1 = \frac{2TP}{2TP+FP+FN}$, and is shown in Figure 10 for different approaches based on both SimDataA and SimDataB. The horizontal axis represents the percentile threshold at which preventive actions are performed. It is seen that the proposed cNHPP model achieves the highest F₁ scores across all percentile thresholds on both simulated datasets, SimDataA and SimDataB. This finding shows the potential of the proposed model for prioritizing limited resources for power grid fire prevention.

4. Conclusions

This article proposed a Convolutional Non-homogeneous Poisson Process (cNHPP) on a linear network. On each network segment, the intensity function is given by the sum of two components. The first component is used to capture the effects of current covariates, while the second term is used to capture the effects of covariates in the previous time step due to spatial-temporal dependency among neighboring network segments. The article showed that the intensity function can be given by the sum of an infinite series, where each term of the series captures the effects of either the current or historical covariates.

The article provided detailed discussions on how the two components can be constructed, the computation of the intensity function, the graphical representation of the proposed cNHPP, the continuous-time representation of the model, and how the proposed approach is different from the existing self-exciting process. In the application example, the proposed approach has been successfully applied to model and predict the wildfire risks over a network of power transmission lines in California. The model captured how weather and vegetation variables affect the wildfire risks on power transmission lines and provided some useful insights for mitigating wildfire risks. Comprehensive comparison studies have been demonstrated the predictive capabilities of the proposed cNHPP in terms of MAPE, area under the ROC curves, and the F_1 score. It is noted that the proposed approach is computationally intensive (primarily due to the convolution operations). Hence, one important future research is to investigate computationally efficient method so that the proposed model can be applied to a large power line network.

Supplementary Materials

The online supplementary materials include: the derivation of long-term cumulative effects $h(i, t)$ (Supplement A), representation of cNHPP using the architecture of a Recurrent Neural Network (Supplement B), as well as the parameter estimation results of the RNN for Section 3.3.1 (Supplement C). We also provide the Python code for reproducing Table 1, Figures 6–8 in this article.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

- Baddeley, A., Nair, G., Rakshit, S., McSwiggan, G., and Davies, T. M. (2021), “Analysing Point Patterns on Networks—a Review,” *Spatial Statistics*, 42, 100435. [2]
- Brown, E. K., Wang, J., and Feng, Y. (2021), “US Wildfire Potential: A Historical View and Future Projection Using High-Resolution Climate Data,” *Environmental Research Letters*, 16, 034060. [1]
- Brown, P. E., Roberts, G. O., Kåresen, K. F., and Tonellato, S. (2000), “Blur-Generated Non-separable Space-Time Models,” *Journal of the Royal Statistical Society, Series B*, 62, 847–860. [6]
- Cox, D. R. (1972), “The Statistical Analysis of Dependencies in Point Processes,” in *Stochastic Point Processes*, ed. P. A. W. Lewis, pp. 55–66, Wiley: New York. [7]
- CPUC. (2014), “Decision Adopting Regulations to Reduce the Fire Hazards Associated with Overhead Electric Utility Facilities and Aerial Communications Facilities,” available at <https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M087/K892/87892306.PDF>. [1]
- D’Angelo, N., Adelfio, G., Abbruzzo, A., and Mateu, J. (2022), “Inhomogeneous Spatio-Temporal Point Processes on Linear Networks for Visitors’ Stops Data,” *The Annals of Applied Statistics*, 16, 791–815. [2]
- D’Angelo, N., Payares, D., Adelfio, G., and Mateu, J. (2024), “Self-Exciting Point Process Modelling of Crimes on Linear Networks,” *Statistical Modelling*, 24, 139–168. [2]
- Gilardi, A., Borgoni, R., and Mateu, J. (2024), “A Nonseparable First-Order Spatiotemporal Intensity for Events on Linear Networks: An Application to Ambulance Interventions,” *The Annals of Applied Statistics*, 18, 529–554. [2]
- Hawkes, A. G. (2018), “Hawkes Processes and their Applications to Finance: A Review,” *Quantitative Finance*, 18, 193–198. [6]
- Holbrook, A. J., Ji, X., and Suchard, M. A. (2022), “Bayesian Mitigation of Spatial Coarsening for a Hawkes Model Applied to Gunfire, Wildfire and Viral Contagion,” *The Annals of Applied Statistics*, 16, 573–595. [2]
- Lewis, P. W., and Shedler, G. S. (1979), “Simulation of Nonhomogeneous Poisson Processes by Thinning,” *Naval Research Logistics Quarterly*, 26, 403–413. [10]
- Mildenberger, M., Howe, P. D., Trachtman, S., Stokes, L. C., and Lubell, M. (2022), “The Effect of Public Safety Power Shut-Offs on Climate Change Attitudes and Behavioural Intentions,” *Nature Energy*, 7, 736–743. [1]
- Mitchell, J. W. (2013), “Power Line Failures and Catastrophic Wildfires under Extreme Weather Conditions,” *Engineering Failure Analysis*, 35, 726–735. [8]
- MODIS. (2015), “MODIS Surface Reflectance User’s Guide Collection 6,” available at https://modis-land.gsfc.nasa.gov/pdf/MOD09_UserGuide_v1.4.pdf. [7]
- NFDRS. (2002), “Gaining an Understanding of the National Fire Danger Rating System,” available at <https://www.nwcg.gov/sites/default/files/publications/pms932.pdf>. [1,2]
- Opitz, T., Bonneau, F., and Gabriel, E. (2020), “Point-Process based Bayesian Modeling of Space-Time Structures of Forest Fire Occurrences in Mediterranean France,” *Spatial Statistics*, 40, 100429. [2]
- Peng, R. D., Schoenberg, F. P., and Woods, J. A. (2005), “A Space-Time Conditional Intensity Model for Evaluating a Wildfire Hazard Index,” *Journal of the American Statistical Association*, 100, 26–35. [2,6]
- Perry, M. B. (2010), “The Weighted Moving Average Technique,” in *Wiley Encyclopedia of Operations Research and Management Science*, ed. J. J. Cochran, Hoboken, NJ: Wiley. [3]
- Schulze, S. S., Fischer, E. C., Hamideh, S., and Mahmoud, H. (2020), “Wild-fire Impacts on Schools and Hospitals Following the 2018 California Camp Fire,” *Natural Hazards*, 104, 901–925. [1]
- SDG&E. (2021), “San Diego Gas & Electric Company 2020–2022 Wildfire Mitigation Plan Update,” available at <https://www.sdge.com/sites/default/files/regulatory/SDG.pdf>. [1]
- Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C., and Rue, H. (2014), “Spatio-Temporal Log-Gaussian Cox Processes for Modelling Wildfire Occurrence: The Case of Catalonia, 1994–2008,” *Environmental and Ecological Statistics*, 21, 531–563. [2]
- Taylor, S. W., Woolford, D. G., Dean, C., and Martell, D. L. (2013), “Wildfire Prediction to Inform Fire Management: Statistical Science Challenges,” *Statistical Science*, 28, 586–615. [2]
- U.S. DOE. (2023), “Protecting Our Electric Grid from Wildfire.” [1]
- Uppala, M., and Handcock, M. S. (2020), “Modeling Wildfire Ignition Origins in Southern California Using Linear Network Point Processes,” *The Annals of Applied Statistics*, 14, 339–356. [2]
- Vazquez, D. A. Z., Qiu, F., Fan, N., and Sharp, K. (2022), “Wildfire Mitigation Plans in Power Systems: A Literature Review,” *IEEE Transactions on Power Systems*, 37, 3540–3551. [8]
- Xu, H., and Schoenberg, F. P. (2011), “Point Process Modeling of Wildfire Hazard in Los Angeles County, California,” *The Annals of Applied Statistics*, 5, 84–704. [2,10]
- Zhou, F., Kong, Q., Deng, Z., Kan, J., Zhang, Y., Feng, C., and Zhu, J. (2022), “Efficient Inference for Dynamic Flexible Interactions of Neural Populations,” *Journal of Machine Learning Research*, 23, 1–49. [10]
- Zhu, S., Yao, R., Xie, Y., Qiu, F., Qiu, Y. L., and Wu, X. (2022), “Quantifying Grid Resilience Against Extreme Weather Using Large-Scale Customer Power Outage Data,” arXiv preprint arXiv:2109.09711. [2]