Spatio-Temporal Modeling and Extrapolation of Censored Wildland Fire Intensity and Fuel Measurements

Jedidiah O. Lindborg[1*], David B. Hitchcock[2], John M. Grego[2], E. Louise Loudermilk[3], Joseph J. O'Brien[3]

[1*]Department of Mathematics and Statistics, University of Central Oklahoma, Edmond, OK, 73034, USA.
[2*]Department of Statistics, University of South Carolina, Columbia, SC, 29208, USA.
[3*]USDA Forest Service, Southern Research Station, Disturbance and Prescribed Fire Laboratory, Athens, GA, 30606, USA.

*Corresponding author(s). E-mail(s): jlindborg@uco.edu;
Contributing authors: hitchcock@stat.sc.edu; grego@stat.sc.edu; louise.loudermilk@usda.gov; Joseph.j.obrien@usda.gov;

Abstract

A spatio-temporal analysis was performed on wildland fire intensity measurements (thermal temperature data) from a prescribed burn at an ecological research center in Georgia, to better understand the variables and environmental conditions that promote fire spread in a longleaf pine ecosystem (*Pinus palustris* Mill.). This ecosystem is prevalent in the Southeastern U.S. and prescribed burning is a vital tool for maintaining the ecological function and biodiversity unique to this forest. Measuring fire behavior (spread and intensity) is critical for advancing fire behavior modeling, understanding fire effects and supporting initiatives to support public safety. The infrared cameras used to measure fire intensity had limitations allowing them to only record temperatures strictly above 300°C, where plant material combustion mainly occurs. This censoring of temperature values below 300°C posed a challenge that initiated this study, which was overcome by augmenting the data set by extrapolating temperatures down to the day's ambient temperature. This extrapolation is a major focus of this analysis. Spatio-temporal models were applied to these data, and several aspects of the models were evaluated, such as the chosen hyperparameters, the spatial covariance structure, and the prior distributions for the parameters. A sensitivity analysis tested the robustness of the model to several heuristic choices of tuning parameters allowing for multiple reconstructions of the censored data. The model fit to the reconstructed data sets containing extrapolated temperatures appeared to give relevant insight into fire behavior phenomenon. The effects of the covariates in the model, such as wind speed, wind direction, their interaction, and several vegetative fuel variables, such as coarse and fine fuels were significant. These statistical models provide more complete datasets of fire intensity that can be used to evaluate physics-based fire-atmosphere models of fire behavior.

Keywords: Controlled Burns, Functional Data, Extrapolation, Southeastern Ecosystem, Forestry, Ecological Data

1 Introduction

Many subfields in statistics have, in recent decades, gained newfound popularity due to innovative methodologies and increased technological advancement and computing power. One such area is that of spatio-temporal modeling and applications. Observations of numerous physical phenomena present themselves as spatio-temporal data, such as data from controlled burns. These types of burns, also referred to as prescribed burns, use fire as a land management tool to reduce vegetative fuels and suppress undesirable plant species from a forest, which can also maintain ecosystem structure and function. From a physics standpoint, the phenomena that govern fire behavior are deterministic, and if one had complete understanding of every physical input affecting the fire, controlled burn data could be modeled and analyzed deterministically. While researchers, such as Rothermel [1], have created mathematical models to describe fire spread, purely physics-based models incorporate the more complex fire-atmospheric dynamics that occur during a fire [2]; [3]. Spatio-temporal fire intensity data collected within wildland fires are needed to evaluate and verify such advanced models, but typical observed burn data include only basic outputs such as temperature and environmental covariates, not the intricate factors involved in the deterministic physical models. Therefore, in practice, modeling these phenomena as random events with a statistical approach can compensate for the lack of information needed for deterministic models and still provide a realistic fire behavior model [4]; [5].

Another reason for employing spatio-temporal statistical models is to gain deeper insight into the relationship between the covariates placed into the model and the response variable. These insights can be of great value and can lead to a more complete understanding of the physical phenomenon of wildland fire behavior. Knowing which covariates have a significant effect on the response variable allows researchers to focus their efforts on measuring (and potentially controlling) these specific variables. Models such as ours may contain covariates that are spatial in nature, temporal in nature, and both spatial and temporal in nature.

Understanding the effects of the covariates that fuel wildfires on the temperature and intensity of such fires is critical for managing the fire-influenced ecosystems, especially in the Southeastern United States - our region of study (see Section 1.1). In the past century, the natural ecosystem has been altered by agriculture, forestry, development, and intentional fire suppression tactics [6]; the changes to the longleaf pine (*Pinus palustris* Mill.) ecosystem are detailed by Oswalt et al. [7]. Rather than the previously prevalent open canopy, the Southeastern U.S. has become dotted with midlevel hardwoods, shrubs, and other small fuels that could lead to higher-intensity wildfires [8]. Prescribed burns serve dual purposes: They effectively thin vegetation and reduce excess fuels, and they also can be monitored to learn about the nature of fire intensity over the course of a burn and how various fuels and environmental conditions affect fire behavior [9]. In our data setting described in Section 1.1, prescribed burns have been used both to counteract the deleterious ecosystem trends mentioned above and to obtain fire temperature data for scientific studies such as the one in this article.

Spatio-temporal models have been applied to various environmental data to better understand relationships among variables in natural phenomena. Pimentel et al. [10] applied a spatio-temporal model to detect fire spots, areas where the temperature is significantly higher than in the surrounding areas, to predict where wildfires would occur in Brazil. Lin et al. [11] used spatio-temporal models based on multiple images from infrared camera sensors to detect forest fires. Spatio-temporal analysis was conducted to detect the effects of soil conditions after a controlled burn in Spain by Vega-Martínez et al. [12]. To better understand forest health, Augustin et al. [13] applied models to spatio-temporal data of tree crown defoliation, which may be an indicator of poor tree health. Fasso and Cameletti [14] proposed a general approach to spatio-temporal problems for environmental data using hierarchical models due to the increasing use of the models in this domain.

1.1 Motivating Data Set

Our motivating data set was produced from controlled wildland surface burns performed at Ichauway, home to the Joseph W. Jones Ecological Research Center in Newton, Georgia. Hiers et al. [15] detail the process for the setup and collection of this data set, which was generated during a controlled surface burn in Spring 2007. Controlled burns are performed regularly, typically every one to three years at the 29,000-acre property. The main ecosystem found here is the longleaf pine (*Pinus palustris* Mill.) ecosystem [16], which has been maintained using prescribed fire for decades at Ichauway. Without frequent fire, these ecosystems would transition to a different ecosystem dominated by hardwood plant species as well as unmanageable wildland fuel conditions seen as undesirable by managers. These surface fires, which mainly burn the grasses, shrubs, and plant debris near the ground, maintain ecosystem structure and function and support high biological diversity, including various endemic flora and fauna species. Several game species (quail, deer, and turkey) also flourish here.

Our goal was to analyze the temperature of the fire as it spread through designated longleaf pine sites. In addition, the data set contains the covariates measured within the same grid that was burned, which are described in the next section. Hiers et al. [15] and Loudermilk et al. [17] performed analyses of this data set focusing on summaries of fire temperature measurements, without accounting for temperatures below 300°C that were censored by the measurement system, which we will describe in detail in Section 1.4. Briefly, the Forward-Looking Infrared (FLIR) camera used to measure fire temperature in this study necessitates a restricted range of temperatures that the camera can measure, essentially creating a censoring mechanism in the data collection. While many fire science studies nowadays are able to use more precise measurements such as radiometers that can attain exact absolute temperature measurements, it is valuable to have a method to deal with censored temperature measurements coming from FLIR cameras to analyze older data sets or data from studies where modern radiometers are unavailable. We will model these censored trajectories through time by applying spatio-temporal models to the data. Since the temperature data are censored at 300°C, we could work only with the temperature readings above 300°C, or we could use our knowledge of the patterns by which fire temperatures rise and decay to reconstruct values below the 300°C threshold. Since we believe the usage of the whole temperature trajectory, not just the peak of it, is critical to modeling the relationship between temperature and the covariates, we undertake the reconstruction approach, while quantifying the uncertainty in our extrapolation process. In Section 4, we show that the model using the fully reconstructed curves yields more scientifically sensible coefficient estimates than the model using only the temperatures above 300°C.

1.2 Fuel Covariates

The size of the full grid where the fire temperature was measured is 4.0×4.1 meters, and several of the covariates recorded in this grid were measured spatially on $\frac{1}{3} \times \frac{1}{3}$

meter sections. The fire temperatures were measured on 0.1×0.1 meter sections of the grid. Since the fire temperature was recorded on a finer, more precise grid, we converted the temperature data to a grid of the same size as the spatial covariates by averaging the temperature data points that were closest spatially to the covariates' measurement locations. This averaging smooths the data slightly, while preserving the overall patterns within the data. As a result, both our temperature and covariate data were matched to $\frac{1}{3} \times \frac{1}{3}$ meter sections on the overall grid. On the temporal domain, the fire temperatures were measured approximately every 5 seconds over 59 discrete time steps.

Several types of vegetation were observed and recorded as binary covariates. If the fuel (vegetation) type was present in a particular spatial location, then a 1 was recorded at that location and if it was absent, a 0 was recorded. Continuous covariates, such as the depth of fuel in a particular spatial location, were recorded in centimeters. In addition to spatial covariates, there are spatio-temporal covariates of wind direction and wind speed. A full and detailed list of the specific covariates collected in this data set can be found in Appendix A along with a description of the covariates and how they were measured. However, it is worthwhile to discuss some of the main classifications of fuels in detail here; see also Dell et al. [18] for more description.

Fuels classified as 10-hour or 100-hour fuels are downed or dead branches, twigs, and wood that are one to three (10-hour) or three to eight inches (100-hour) inches in diameter. These are considered long-burning fuels that sustain the fire as it smolders. Shrubs and especially volatile shrubs tend to burn with ease and brief duration. The types of litter (deciduous, longleaf, and evergreen) are downed woody material typically less than 0.25 inches in diameter and are often called "flashy fuels" or fine fuels, promoting quick and brief flare-ups of fire. Wiregrass and graminoids (other grass-like plants) also fall into the fine-fuels category, as do forbs (ground level non-woody plants, often with deeper roots than grasses have) and pine needles. Perched pine and perched oak leaves are adapted to fire and may burn more intensely than downed fuels because of increased airflow and their lower moisture content. Mitchell et al. [19] provide a detailed study of the ecology of these fuel sources in the longleaf pine ecosystem.

1.3 Wind Covariates

The data collection for the wind involved two measuring devices (anemometers) at the north corners of the grid on the east and west sides. For each anemometer reading separately, we transformed the cyclical direction data using sine and cosine functions with the following formulas:

$$x = \sin\left(\frac{\pi\theta}{180}\right), \qquad y = \cos\left(\frac{\pi\theta}{180}\right). \tag{1}$$

Here θ represents the original angle measured by the anemometer for the wind's direction, where $0 < \theta \leq 360°$. This rescaled the range of the direction data, forcing it to be between -1 and 1, as well as preserving the cyclical nature of the data, e.g., so that angles such as 359° and 1° would be treated as similar values. The sine component indicates the extent of the easterly direction, whereas the cosine component indicates this in the northerly direction. In addition, since the effect of the wind direction on the fire naturally depends on the speed of the wind, we included terms for the interaction between the wind speed and direction variables.

1.4 Fire Data Set

The fire was manually started using a drip torch, a tool that drips flaming fuel onto the ground, along one side of the grid. The fire was then allowed to spread into the grid generally as a head fire (burning from upwind to downwind) and through it. The temperatures were recorded, approximately every four seconds, with an infrared digital thermal imagery system [20] that was positioned above the portion of the field that contained the controlled burn. In this case, that is a 4×4.1 meter grid, over which all the temperature and covariate data were measured. There are further details on data collection, instrument specifications and processing in Hiers et al. [15] and Loudermilk et al. [17]. There are certain measurement limitations associated with the FLIR camera: The detector in the camera is a focal plane array uncooled microbolometer with 320×240 pixels and is sensitive to wavelengths from 7.5-13 μm. The FLIR camera has different gain settings that restrict the set of internally calibrated temperature ranges it can record; the FLIR camera can either record temperatures of -40 to 120°C, 0 to 500°C, or between 300°C and 1500°C. In this experiment, the highest range was selected to capture the range of temperatures typical of fires in longleaf environments such as the location in this experiment, because combustion typically occurs when temperatures are approximately 300°C and above. Although there is not a single temperature at which all materials will combust, based on the fuel sources for this controlled burn, 300°C can be considered an approximate minimum combustion temperature. This restriction produced censoring of data within this data set and precluded temperature measurement during preheating and pyrolysis and cooling of fuels and soil outside of the time when the temperature exceeded 300°C. While having multiple FLIR cameras set at different ranges could mitigate this limitation, this option was impractical and cost-prohibitive at the time due to the high price of the camera.

Since the FLIR did not record any temperature below 300°C, at many spatial locations, several of the temperatures at early time steps and time steps at the end of the burn are recorded simply as ≤ 300°C. To accurately understand the factors that contribute to the fire, extrapolation of these censored data was needed. Our approach assumes that all temperatures in a fire event can be accounted for, including the rate of the increase of temperatures before combustion (pyrolysis, soil and vegetation pre-heating), the temperatures during combustion, and the rate of decreasing temperatures after combustion (cooling or heat dissipation). Therefore, it is desirable to reconstruct the censored temperatures from the data set, so that these the characteristics of the heat budget prior to and post-combustion of the fire event can be taken into account. The process of this reconstruction using spline methods will be a major focus of this paper.

2 Methods of Accounting for Censoring

To solve the issue of censoring, extrapolations were performed for the temperatures that were recorded as 300°C before and after active combustion. Natural cubic splines were used to extrapolate the temperatures at times before the true (not censored) temperatures of the fire event were recorded. These splines were selected to reflect the fact that fire tends to build in intensity after the initial ignition in a rapid exponential increase, as opposed to something linear or otherwise [21]. The extrapolation scheme described in the remainder of Section 2 was carried out separately for the temperature curves at each spatial location on the grid.

2.1 Guiding Points Before the Fire Event

Several considerations helped produce realistic extrapolated temperatures. First, all extrapolated temperatures were below 300°C. If the true temperature were actually above 300°C, then the FLIR camera would not have censored it. Next, we wanted to guide the splines and extrapolations with intentionally constructed response values that we called guiding points. Figure 1 displays a temperature curve and its censoring over time for a particular spatial location. The graph also depicts the four guiding points, two placed before the region of temperatures above 300° and two placed after.

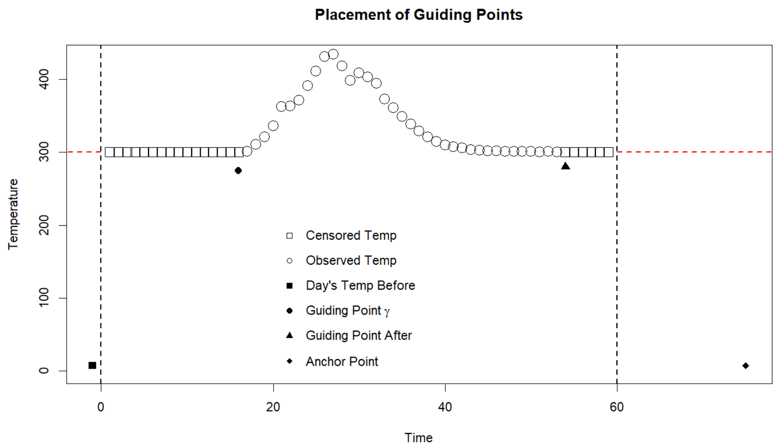

Fig. 1: Placement of all four guiding points for a single location. Open circles represent truly observed temperatures and open boxes represent temperatures censored at 300°C

These guiding points helped us to ensure that, for each location, the reconstructed temperature curves followed reasonable patterns. For example, the guiding points in the early part of the curve helped us to ensure there was monotonicity in the initial section of the curves, and that they formed a convex shape during fire buildup. Since fire has a rapid increase, a concave shape would not match our intuition of how a fire begins to burn. The guiding points also ensured that our extrapolated temperatures would not be below a certain bound (the day's ambient temperature, meaning the air temperature of the day at the time of the burn, found from historical weather records) or above 300 (the censoring temperature). We will now discuss the rationale for each of these four guiding points; specific details of their calculation may be found in Appendix B. The ambient temperature of the day (approximately 7.22°C for this burn) was placed at the time step just before the controlled burn starts. This anchors us to the reality that before the fire starts, the temperature of the location should just be the ambient temperature. This guiding point is placed at the same time step for all locations and is the leftmost point in Figure 1. Next, separately for each location one extrapolated data value was placed one time step before the first temperature recording that exceeds 300°C. The placement of these two points was constructed to help guide the cubic spline extrapolation between them. The second guiding point is not necessarily at the same time step for any two locations, since the time when combustion was initiated above 300°C was not constant across all locations. This point corresponds to the second solid point from the left in Figure 1. The method of placing this second point, called guiding point γ, depends on the behavior of the fire data at the specific spatial location in question. The first method relies on the slope of a line connecting the first two truly observed (not censored) temperatures.

Case I: If this slope is positive and steep, the guiding value was placed at the previous time step linearly extrapolating along the slope between the two first true temperature recordings, as long as the slope was steep enough to ensure that the value placed there was less than 300°C. This forced the temperature at this time step to follow the pattern of the two first true temperature recordings. At times before this point, natural cubic splines, implemented with the `ns` function in the `splines` package [22] in R, were used to extrapolate the rest of the temperatures for the remaining (earlier) time steps. It was important to follow the pattern at only the first few truly observed temperatures, as the fire would change behavior once it reached high temperatures and started burning coarser fuels, such as branches and sticks. Thus, we wanted the extrapolated values to mirror the behavior that the fire likely had as it was burning fine fuels, such as pine needles and grasses similar to kindling, at the start of the combustion.

Case II: If the line connecting the first two uncensored temperatures had a positive slope, but a value one time step previous following a linear interpolation would not be less than 300°C, then an extrapolated data value was generated using a Beta$(1, \beta)$ distribution that was scaled and shifted to have support on [7.22, 300]. See Appendix B for the exact formula for β. The parameter β was based on the steepness of the slope of the first two recorded true temperatures. If the slope was very steep, then β would be smaller. If the slope was not steep, β would be larger. The plots of two such Beta distributions in Figure 2 show how the shape of the distribution affected the chance of generating a value near 300°C.

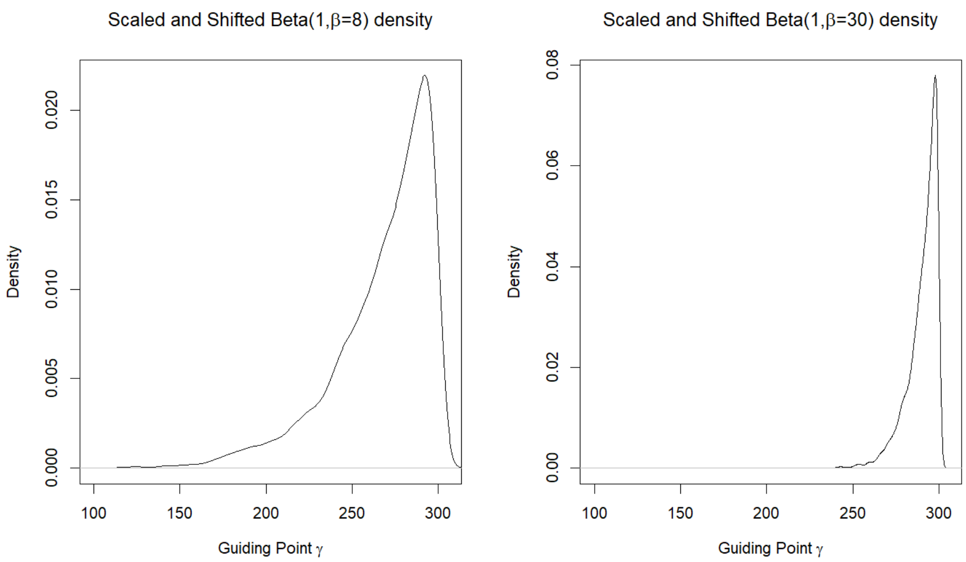

Fig. 2: Plots of two Beta density functions scaled and shifted to have support on [7.22,300], showing the difference in ranges of probable temperatures generated from Beta distribution dependent on β parameter

The maximum value in the support was 300, and the spread of the distribution was affected by the size of β. A large value of β would more likely produce an extrapolated value closer to 300, as seen in the support of the right-hand graph in Figure 2. If the slope was steep, we believed it was more likely that the temperature rose more rapidly than if the slope was gentle. These two possible behaviors correspond to either a volatile flare-up or a less rapid one.

Case III: If the slope between the first two recorded temperatures above 300 was negative, then γ was again generated from a Beta distribution. This time β was affected by the height of the first recorded temperature. See Appendix B.1 for more details. If there was a gap in time between the first uncensored temperature and the second, this indicated the temperature quickly rose above 300°C and then rapidly fell below 300°C for some amount of time before the fire gained momentum and stayed above 300°C for a sustained time. In this case, two values, called γ_1 and γ_2, would be extrapolated for this location: one before the first uncensored temperature and one midway between the first and second uncensored temperatures. The reason for this second extrapolated point was to guide the natural cubic splines to the presumed shape of the data. Again, these two extrapolated values were generated using a Beta$(1, \beta)$ distribution and β was selected based on how far above 300° the first and second recorded true temperatures were. If the first temperature was far above 300, then β was a smaller value; otherwise it was larger. The distance of the second recorded temperature from 300° was used to determine β for the second extrapolated value. Locations that displayed this flare-up behavior were rare.

2.2 Neighbors' Information

As described in Section 2.1, the guiding points were placed using the relative positions of the initial uncensored temperatures at a specific location. However, we also have information about the neighboring locations. Since it was evident that in most locations there was some degree of spatial association between fire temperatures in neighboring locations, it was desirable to use these neighbors to help place the γs. Therefore, in addition to using the fire behavior of a specific location to inform the placement of the γ, we also considered the behavior at neighboring locations. See Appendix B.2 for more details.

The neighboring information influences our reconstruction only if a neighboring location is burning above 300°C one time step before and the neighbor is upwind. Using the neighboring information, the generated Beta value will be more likely to produce an extrapolated temperature closer to 300°C than it would without the neighboring information. This permits the neighbor's known high temperature to inform the prediction of a subsequent rapid increase in temperature at the location in question. The presence of wind is an established significant factor in fire ecology and wind direction is very likely to explain heat movement between neighboring locations. However, there is a legitimate concern about using wind information as part of the reconstruction process, since a goal of the model is subsequently to interpret the wind covariates' effect on the fire temperature profile. To test whether the usage of wind and the neighboring information had any effect on the eventual inferences, we simulated 100 extrapolations that used the neighboring information and 100 that did not. We then ran the model that will be described in Section 3 on data sets resulting from both of these sets of simulations. Through that testing, we determined that using wind to inform the neighboring information had little effect upon the eventual estimated model. Thus we believe allowing wind to define the neighboring structure to be a justifiable and robust method to incorporate neighboring information into the extrapolation process. The results of these can be seen in Table B1 of Appendix B.3.

2.3 Spline Specifications

After the guiding points before the recorded true temperatures were determined, we required the locations of the knots for the natural cubic splines, which controlled the smoothness of the splines. Equivalently, we could specify the degrees of freedom for the spline fit. We initially chose a large number of degrees of freedom for every location and then iteratively lowered it until three conditions were met. First, the value just before the guiding point that was extrapolated from the splines must be lower than the guiding point. This ensures a monotonicity within the reconstructed temperatures. Second, the extrapolated values from the splines must not form a concave shape, since we believe that the fire will increase rapidly as it nears combustion. This convexity condition was verified using the `check_curve` function in the `inflection` package [23] in R. Lastly, the degrees of freedom must not drop below a prespecified lower bound of 4, to enforce a smoothness condition on the splines. At each spatial location, this approach produced satisfactory extrapolations within the region before the recorded true temperatures occurred.

2.4 Guiding Points After the Fire Event

In order to extrapolate temperature values in place of those that were recorded as 300 in the later section of the temperature curve as the fire cooled, we used a different method. At some spatial locations, the fire maintains a temperature above 300°C through the final time step of the burn and thus does not need extrapolation on the later section of the time domain. For the locations that needed extrapolation, we placed one extrapolated value one time step beyond the final recorded temperature above 300°C. This point corresponds to the guiding point second from the right in Figure 1. A process similar to the one described in Section 2.1 detected the slope of the final two uncensored temperatures and whether there was a temporal gap between the last uncensored temperature and the previous one. For these extrapolations, whenever a Beta$(1, \beta)$ distribution was used, β was larger than for the extrapolations at the start of the fire event, because the temperature should decline less steeply than it increased at the start. One key difference is that B-splines were used to extrapolate the temperatures that occur at the end of the fire event, using the `bs` function in the `splines` package [22] in R. We found that these polynomial splines allowed us to reconstruct the way a fire cools more accurately than the natural cubic splines.

We placed the knots for the splines at all of the locations of the recorded true temperatures, which worked well. When there was a large gap in the time steps between uncensored measurements, we removed the knots around the gap to force the extrapolation to be more linear, which was more consistent with a true fire's behavior.

2.5 Anchor Point

The final necessary guiding point was placed when the fire had cooled enough to reach the ambient temperature. The placement of this last guiding point was not as simple as the placement of the first, since it was not obvious when the fire returned to the ambient temperature as it cooled. The temperatures at the end of the fire event do not necessarily decrease as rapidly as they increased at the start. This is because, depending on the conditions, such as the type of fuel source, the fire's high temperatures can be sustained for an extended duration. Therefore, it is difficult to determine how many time steps in the future the fire will die off enough for the temperature to return to the day's ambient temperature. The final extrapolated value, called the anchor point, represents the return to the ambient temperature. This point corresponds to the right-most point in Figure 1. We used a time-varying anchor point at each location, which reflected the fact that some locations had fuels that promote long-term smoldering and thus a very slow cooling, while at other locations lacking such fuels, the fire should cool quickly.

We did not use covariate (fuel) information to determine the placement of the anchor point. Since a goal of our analysis was to assess the significance and effect of the covariates, it would be inappropriate to use covariate information in the extrapolation process. By keeping the anchor point placement independent of covariate information, we avoided influencing the model toward finding certain covariates significant or not significant. In addition, we wanted the method for anchor point placement to be as simple as possible and guided by expert opinion.

We believe that the following extrapolation method accurately captures the behavior at the end of the fire event. Our method used the trend of several selected uncensored temperatures that would give sufficient information about the longevity of the decay of the fire temperature as it smolders. Figure 3 shows examples of the four major behaviors in our data set for fire temperatures as they smolder just above 300°C.

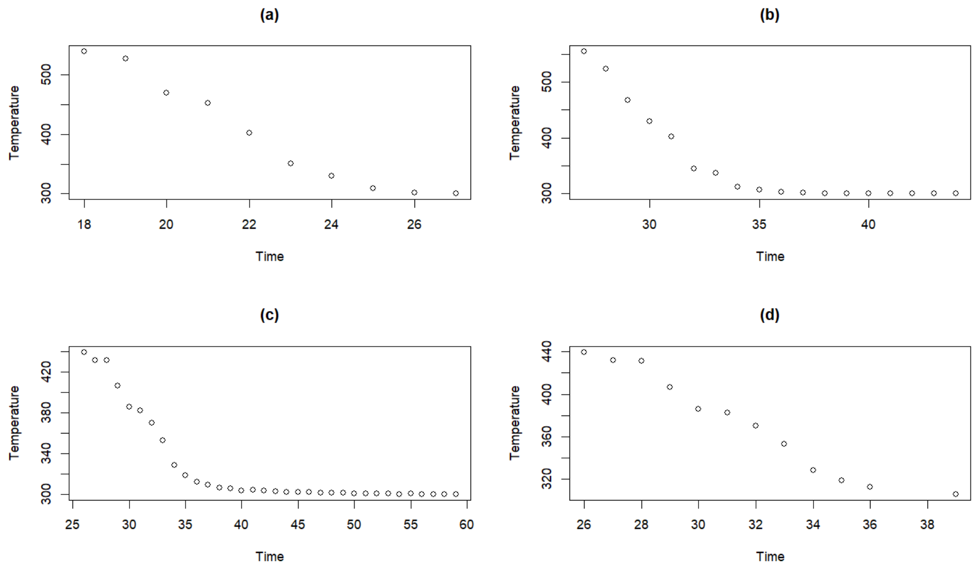

Fig. 3: Different behaviors for fire smoldering around 300 degrees. Rapid cooldown, seen in (a), or sustained smoldering, seen in (c), help determine the placement of the anchor point.

Figure 3(a) shows little to no smoldering. Only the final two temperatures should inform the location of the anchor point. This would indicate a fire that is rapidly cooling and will return to the day's ambient temperature quickly. Thus, for that location, the anchor point would not be placed far in the future. Figure 3(b) shows a longer smoldering, such that the last nine or ten temperatures should influence the placement of the anchor point. Figure 3(c) shows a long smoldering behavior, such that several temperatures should be used. In this instance, the anchor point would be placed further in the future because the fire behavior is showing a pattern of holding its high temperatures for a longer duration. Figure 3(d) shows that a gap in the time steps between the truly observed temperatures can occur and since we do not know the true behavior of the fire at this gap, the censored temperatures in the gap should not be included in the calculations.

At first it may seem unusual that several locations show a pattern in which the slope of the temperature function abruptly changes and levels off around 300°C. However, recall that around 300°C combustion should occur. That indicates that when the temperatures have decreased and are close to 300°C combustion should be almost complete. The transitions from flaming phase to solid phase or glowing combustion can be rapid; however, glowing combustion can be sustained for long periods in coarser fuels. If a fire is losing its flame, it flickers and, so long as conditions are correct for it to sustain combustion, maintains an active flaming front. Thus, it follows that we should see patterns of the temperature curves leveling off around 300°C only at locations where combustion can be sustained for an extended period of time. Therefore, it should take longer to reach the day's ambient temperature in those locations.

We fit a piecewise linear regression model to the selected data in order to determine the optimal break point where the trend changed from decaying rapidly to smoldering (i.e., the location where the two fitted lines were joined) using the `Segmented` package [24] in R. Let m be the number of temperatures after this joining location that are observed to be above 300. If m is fairly large, this indicates the fire is slowly dying, i.e., smoldering for a long time. In this case, we want to place the anchor point far into the future. Based on expert knowledge, we placed the anchor point $5m^{1.1}$ time steps after the 59th and final observed time. For example, $m = 10$ would indicate a long-smoldering fire, and the corresponding anchor point would be 63 time steps (or roughly 5 minutes) after the end of the temperature recording, which is reasonable based on fire science. In our sensitivity analysis in Section 3.4, we showed that the results were consistent for a range of choices for the multiplier and exponent in the anchor point formula.

For all locations, our method of extrapolation relied on the placement of four guiding points, two before the recorded temperatures and two after. To reconstruct the missing sections of the temperature curves, cubic splines connected the two guiding points before the true recorded temperatures and B-splines connected the latter two guiding points. We now detail the specifications of the spatio-temporal models applied to this data set.

3 Spatio-Temporal Model for the Fire Data

To specify and fit our spatio-temporal model, we used the methodology of Bakar and Sahu [25] underlying and implemented in the `spTimer` package [25] in R, which assumes a Bayesian framework and relies on MCMC methods to provide inferences [25]. The model can predict in the space and the time dimensions and allows for prediction with missing data points, including missing time steps. The covariance function can be specified to follow an exponential, Gaussian, spherical, or Matérn, which refers to the general Matérn correlation function [26], for the Gaussian process. We can also specify a different method for measuring distance between two locations if desired [25].

3.1 Distributions and Hyperparameters

The Bayesian model is represented in a hierarchical structure following Gelfand [27]. There are three levels that encapsulate the data, the process, and the parameters [25]. The hierarchical structure can be represented in the following way:

$$\begin{cases} \text{First Stage:} & [\text{data} \mid \text{process, parameters}] \\ \text{Second Stage:} & [\text{process} \mid \text{parameters}] \\ \text{Third Stage:} & [\text{parameters}] \end{cases}$$

The options for modeling the processes in the second stage are a Gaussian, an autoregressive and a Gaussian predictive process [25]. The third stage provides the hyperparameters for the model.

The Gaussian process has a hierarchical structure of the form:

$$Y(s_i, t) = O(s_i, t) + \epsilon(s_i, t),$$

$$O(s_i, t) = X(s_i, t)\boldsymbol{\beta} + \eta(s_i, t),$$

where $Y(s_i, t)$ is the observed, point-referenced data for spatial location s_i, $i = 1, \ldots, 169$, and time t, $t = 1, \ldots, 59$, $O(s_i, t)$ represent the random effects at the same location and time, X is the $169 \times p$ matrix of p covariates, $\epsilon_{s_i t} \sim N(0, \sigma_\varepsilon^2 I_n)$, and $\eta_{s_i t} \sim N(0, \Sigma_\eta)$, with the parameters represented by $\theta = (\boldsymbol{\beta}, \sigma_\eta^2, \sigma_\epsilon^2, \phi, \boldsymbol{v})$ [25]. Of particular interest is to estimate the $p \times 1$ vector of regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1 \ldots, \beta_{p-1})$ [25].

The third stage involves a prior distribution, $\pi(\theta)$, that is associated with the parameters. The parameter vector $\boldsymbol{\beta}$ is assumed to have a multivariate normal prior with a specifiable mean and variance for each element. The default is to have the distribution centered at zero with a large variance, 10^{10}, to make a diffuse prior [25]. The prior distribution for the precision parameter, the inverse of the variance, is taken to be a gamma distribution with mean $\frac{a}{b}$ [25]. To find optimal values for the correlation parameters, $\boldsymbol{v}$ and ϕ, Gibbs sampling is utilized; in addition, full Bayesian estimation of ϕ and $\boldsymbol{v}$ is also allowed in `spTimer` with discrete uniform prior distributions for both these parameters [25]. The permissible prior models for ϕ are a fixed flat prior, a discrete uniform, or a gamma.

The nugget effect, ϵ_{lt}, which is the pure error term, is assumed to be normally distributed with mean zero and variance $\sigma_\epsilon^2 I_n$, where σ_ϵ^2 is the unknown pure error variance [25]. Following the notation from Bakar and Sahu [25], η represents the spatio-temporal random effect, which follows a normal distribution with mean zero and variance $\Sigma_\eta = \sigma_\eta^2 S_\eta$, where σ_η^2 is the site-invariant spatial variance and S_η is the spatial correlation matrix obtained from the Matérn correlation function [25]. Within this function, the parameter ϕ controls the rate of decay of the correlation as the distance increases between spatial locations and $\boldsymbol{v}$ controls the smoothness of the random field [28].

The models are fitted using Gibbs sampling and, excluding the correlation parameters, all the priors are conjugate and allow standard conjugate sampling from the full conditional distributions [25]. The full conditional distributions of the correlation parameters ϕ and $\boldsymbol{v}$ are non-standard. The package provides two options for sampling these parameters corresponding to the two different prior distributions. The first is to assign a discrete uniform prior distribution that will allow the posterior distribution to be discrete [25]. The second option assigns a gamma prior distribution for the decay parameter ϕ and uses the random-walk Metropolis-Hastings algorithm to draw samples from it [25]. If this approach is used, we prefer an acceptance rate between 20% and 50% to align with theoretical justified acceptance rates [29].

3.2 Goodness of Fit

The Predictive Model Choice Criterion (PMCC) was used to determine the fit of the model [25]. This approach is derived from Gelfand and Ghosh [30] who focused on minimizing posterior predictive loss. Following this approach, a squared error loss function is justified and is particularly suitable for comparing hierarchical models with a Gaussian process involved in the first stage [25]. The general form of the PMCC is the following:

$$PMCC = \sum_i \sum_t E[Y(s_i,t)_{rep} - y(s_i,t)]^2 + Var(Y(s_i,t)_{rep})$$

The first term in the expectation denotes a future replicate of the data $y(s_i, t)$, and the expectation of this squared difference determines the goodness of fit of the model [25]. The last term of the expression is a penalty term that punishes a model that is overly complex. As the complexity increases, the last term will typically increase, but this is not theoretically guaranteed [30]. We used this criterion, along with the behavior of the residuals, to assist in tuning the model's parameters. Table C7 shown in Appendix C displays the goodness-of-fit and PMCC for a few candidate models with differing prior specifications, including the model we eventually chose based on its PMCC.

3.3 Specifications for Our Model

In the second stage of the model, a Gaussian process was used and a square root scale transformation was applied to stabilize the variance. For our data analysis, we chose a model with the following hyperparameters and specifications. The covariance function was modeled with an exponential distribution based on a series of variograms. We plotted the variograms separately for each time step. Some indicated an exponential covariance pattern, while others showed something more similar to a linear pattern. It is reasonable that fire temperature has an exponential spatial covariance structure, implying that responses spatially near each other have stronger association and that association decreases exponentially with distance. The prior for the spatial decay parameter, ϕ, was specified to be a discrete uniform distribution with parameters 0.15 and 0.20. ϕ had a strong effect on our model and we choose a uniform prior to guide ϕ to be substantially larger than zero. The Bayesian fitting procedure produced a boundary solution of 0.15 as the estimate for ϕ, and changing the prior to have a smaller lower bound consistently yielded estimates of ϕ that equaled that lower bound. Such low values for ϕ produced an overfitted model that made substantially worse predictions. In the end, we kept the lower bound of the prior at 0.15 to ensure the spatial decay parameter would be substantially greater than 0, since our prior knowledge tells us that the covariance structure should reflect an association between temperatures at spatially nearby locations.We chose the prior on ϕ and the spatial structure to comply with Tobler [31] and his first law of geography, which states that correlation is stronger for locations that are nearer to each other spatially. The effect of Tobler's first law of geography has been studied in regard to fire data sets in the past, and the evidence suggests that it holds [32]. We also analyzed the residuals of the model and Figure 4 displays these residuals plotted against time, and the residual plots indicate a sensibly performing model.

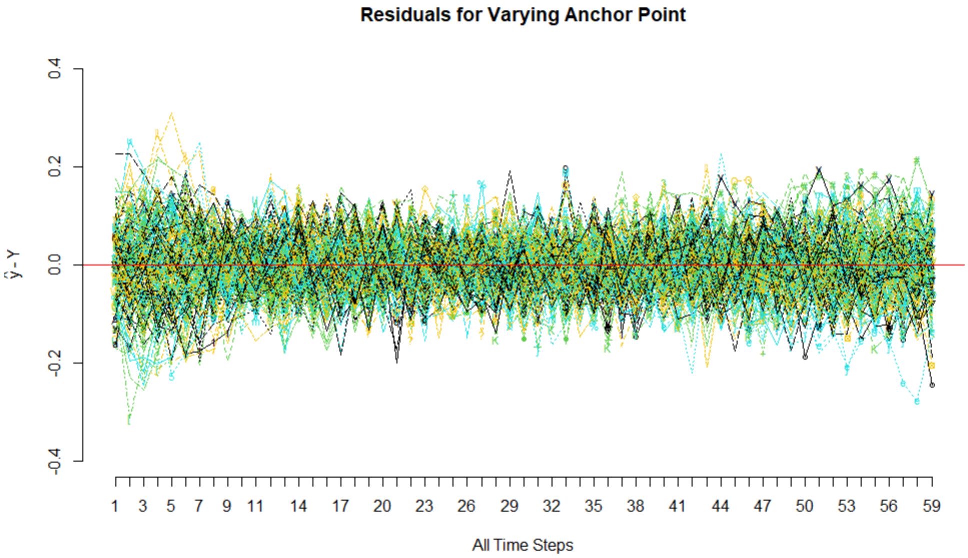

Fig. 4: Stable and unbiased residuals centered at zero generated from a model with the specifications listed above.

Incorporating a varying anchor point into the model both reduced the spread of the residuals and removed bias. We selected our final model based on the residuals and the PMCC value.

3.4 Robustness Testing

In order to test the robustness of our extrapolation method to various heuristic choices, we simulated 100 data sets while varying some key tuning parameters, the first being the placement of the anchor point. Writing a general anchor point formula as $a \times m^b$, we allowed a to vary from 3 to 8 and b to vary from 1.01 to 1.40. Also, we allowed the lower bound for the degrees of freedom for the pre-ignition spline extrapolations to vary from 5 to 10. We allowed the knot placement to vary as follows: with 0.60 probability, the knot remained at the position selected by our procedure; with 0.20 probability, it was placed one time step earlier; and with 0.20 probability, it was placed one time step later. We then fit the model described above on these simulated data sets, keeping track of how many times a covariate was significant (in the sense that a 95% credible interval for its coefficient excluded 0) out of the 100 runs. The results, for the full model including all candidate covariates in the model, can be seen in Table C3 of Appendix C. This revealed consistent assessments about significance for most of the covariates, with a surprising result that the wind speed from the West anemometer was only considered significant on a third of the runs. Inspection of the wind direction revealed that the East anemometer was the upwind anemometer for most of the burn, suggesting that the West anemometer's readings may have been influenced by the burn itself. Thus we ran the model again leaving out the West anemometer as a covariate. These results can be seen in Table C4 of Appendix C. Based on the percentage of times the covariates were significant, we fit the model a final time removing the consistently nonsignificant variables, and these results can be seen in Table C5 of Appendix C with the values for the coefficients seen in Table C6 of Appendix C.

4 Discovery and Interpretation of Significant Covariates

To validate our decision to incorporate extrapolated values for the censored temperatures that were below 300°C, we conducted a separate analysis applying a spatio-temporal model fit to only the uncensored data. The magnitudes of the estimated coefficients change between the two models, and some covariates that were significant in one model were not significant for the other. A validation of the use of extrapolated values can be seen by looking at the effects of specific covariates.

The coefficient of the 100-hour fuel source is significant and positive in the model using the extrapolated temperatures, but in the model that only uses temperatures above 300°C it is significant and negative. A 100-hour fuel source is similar to a dead tree branch (1 to 3 inches in diameter), which is harder to ignite but can burn for longer time periods than smaller live branches which hold more moisture. Figure C1 in Appendix C displays two locations that contain a 100-hour fuel source, and both locations sustain temperatures above 300°C for all time steps after combustion begins. Realistically, the presence of this covariate promotes a sustained duration of high fire temperatures. The model using the reconstructed temperatures has a positive coefficient that aligns more closely with the physical reality of the effect of a long-burning fuel source on the burn.

In addition, the interpretation of the model's significant covariates is insightful. The wind speed and wind direction are significant, along with their interaction. It is important that the model includes the interaction between the wind's direction and speed, since any effect of wind direction should be more pronounced when wind speed is stronger. Since it is well known that wind is a major contributor to fire intensity, it is reassuring that the model detects these relationships.

Also, the covariates related to coarse vegetation, such as 10-hour fuel source, 100-hour fuel source, and shrubs, all have significant positive coefficients (unlike in the model using only uncensored temperatures). Coarse fuels should have a positive influence on fire spread since they can sustain temperatures significantly higher than the ambient temperature for a long duration of time. In contrast, fine fuels, such as wiregrass, graminoids, and forbs, (with only graminoids being a significant covariate in our model) should have a negative effect on the temperature trajectory and this is reflected as they all have negative coefficients. Though these fuels can, and typically do, burn with high temperatures, they are quickly consumed and do not promote long-lasting high temperatures. Thus fine fuels require relatively low energy to ignite, but die off quickly. On the other hand, more energy would be required to ignite a coarse fuel and the maximum temperatures may not become as high as with fine fuels, but they will be sustained for substantially longer. Therefore, we believe the signs of the estimated coefficients for the fine fuels and coarse fuels in our model are sensible and reflect the model accurately capturing the physical reality of the fire behavior and the variables affecting it. Ranges based on 100 reconstructed data sets of the point estimates and 95% credible intervals for all the coefficients can be found in Table C6 of Appendix C. In addition, Table C3 of Appendix C displays the range of the mean for all covariates including non-significant covariates. Figures C2, C3, C4 and C5 of Appendix C display model diagnostics with trace plots and posterior density plots for the parameters included in the final model. The diagnostic plots show good mixing and convergence for all parameters except ϕ, which as previously discussed in Section 3.3, is de facto fixed to be at the lower limit of its prior support.

5 Discussion

Our model contributes to fire ecology in several ways. First, our spatio-temporal model relates the entire fire temperature trajectory, not merely summary statistics, to a set of covariates to assess those covariates' joint effects on the fire behavior. The findings about significant covariates both confirm existing beliefs about relevant predictors of fire behavior and illuminate less well-known factors that influence temperature curves. For example, the clear significance of wind (and in particular the interaction of wind speed and direction) is a "sanity check" for the model's performance, since fire scientists are fairly certain that wind direction catalyzes fire temperature, and that the effect is stronger with higher wind speeds. Furthermore, our model distinguishes between larger coarse fuels and more granular fine fuels, and discerns that while coarse fuels contribute positively to temperature over the whole fire trajectory, fine fuels actually have a negative association with the temperature curves, since they promote brief flare-ups rather than sustained high temperatures.

Perhaps our most novel contribution is our method for reconstructing the entire temperature trajectories despite the limitations of the FLIR technology not allowing the recording of temperatures below 300°C. While practical monetary concerns in the study at the Ichauway property precluded a more accurate method of measuring temperature, our method alleviates the major concerns with making inferential judgments despite the acknowledged data limitations. Our method uses scientific knowledge of the patterns of fire buildup and decline sensibly through spline-based extrapolations on both ends of the temperature curves and cleverly uses the pattern of observed temperature values that just exceed 300°C to inform the probable nature of the fire's cooling as it dies off. Our sensitivity analysis in which we vary our algorithm's tuning parameters allows the uncertainty in the extrapolation to propagate to the parameter estimates and yields an honest assessment of the covariate effects and their significance. From a practical standpoint, while well-funded researchers now have access to more accurate devices for measuring fire temperatures, our method can serve to allow usage of less expensive FLIR cameras for researchers unable to afford better equipment. It also allows for a more complete analysis of the US Forest Service's historical fire data which had, in past decades, mainly been recorded using a single FLIR camera. In addition, our spatio-temporal models could guide researchers who are developing computer simulations of fire behavior, an increasingly prominent area of study [2]. In short, our approach and findings contribute to the literature on statistical models for wildfire behavior alongside the works cited in Section 1.

6 Conclusion

Our study indicates that the effects of the covariates in the model align with the expected physical reality. This is an indication that our model is accurately representing the temperature phenomenon and its associations with the variables involved. Our model utilizes the whole temperature functions over time and space, rather than just a scalar summary of the temperature. Our method of extrapolation and modeling is effective for analyzing fire temperature curves that contain censoring. The extrapolation method produces a reasonable reconstruction of what the temperature was during the entire controlled burn, and can help give a more comprehensive view of the whole fire event, instead of just during combustion. This results in more reliable estimates of the covariates' effects on the fire's temperature. As a secondary application, our spatio-temporal model that relates fire intensity data to wildland fuels can be used to evaluate and calibrate "digital-twin" fire-atmosphere models [2] that simulate fire intensity and behavior using fuel variables as inputs.

Appendix A Covariate Information

The covariates measured by the researchers and used for the spatio-temporal models are the following.

- Litter Depth - measured in centimeters
- Fuel bed depth - measured in centimeters
- Ten hour fuel sources (a fuel that could sustain combustion for 10 hours)
- One hundred hour fuel sources (a fuel that could sustain combustion for 100 hours)
- Presence or absence of: perched pine, perched pine needles, perched oak leaves, wiregrass, graminoids (other grass-like plants), shrubs, volatile shrubs (volatile in terms of promoting burning), forbs (ground level non-woody plants), deciduous litter, longleaf litter, and evergreen litter
- Average wind speed from the west anemometer - measured in meters per second
- Direction of the wind measured from the west anemometer - measured in degrees
- Average wind speed from the east anemometer - measured in meters per second
- Direction of the wind measured from the east anemometer - measured in degrees

Appendix B Guiding Points

B.1 Placement for Flare-ups

The algorithm that deals with flare-ups relied on the first recorded temperature, as this behavior likely indicated that fine fuels were being consumed at these times. Then, once those fuels were consumed entirely, the temperature dropped for the second recorded temperature as the fire started to consume coarser fuels. These coarse fuels would require more energy to burn and would lower the fire's temperature as it started to consume them. We did not want to match the pattern of the fire as it burned coarse fuels but rather the behavior it has as it was building temperatures as the fire started.

Algorithm 1 Steps to handle Flare-ups

Step 1: Let s be the slope of the line connecting the first two observed temperatures, t_1 and t_2
if $s < 0$ **then**
  $D \leftarrow t_1 - 300$
end if
Place the guiding point by generating from a Beta(1, $\beta = \frac{C}{D}$) distribution. Where C is picked based on the value of the maximum flare-up for the data set (if the first recorded temperature was far above 300, then β was selected as a smaller value. This allows for the extrapolated point to potentially be below 300 by a greater amount, indicating a more rapid flare-up. If the first value was not far above 300, then β was selected to be a larger value, forcing the extrapolated point to be closer to 300, indicating a less dramatic flare-up)

B.2 Neighboring Information

If the locations where the temperatures were recorded are viewed in a discretized manner, then any interior location has eight neighboring locations. The locations on the edges of the grid have five neighboring locations. Lastly, the locations at the corners of the grid have three neighboring locations. With that we can use the following steps.

Algorithm 2 Steps to handle neighboring information

Step 1: Determine if the location in question is on an edge, corner, or is an interior point
Step 2: Determine if the wind blows through a neighbor, z, that has a temperature above 300
if Step 2 = True **then**
  $B \leftarrow w(z - 300)$, where, w is the wind's speed
  Place the guiding point by generating from a Beta(1, $\beta = B + C$) distribution
end if
if Step 2 = False **then**
  $A = \sum_{i=1}^{n} z_i$, where n is the number of neighbors, z, with temperatures above 300
  Then generate a guiding point using a Beta(1, $\beta = A + C$) distribution
end if

B.3 Robustness for Neighboring Wind Information

Table B1: Comparison of incorporating neighboring information into the creation of the extrapolated data sets and withholding that information. Percentage of simulated (varying tuning parameter values) data sets (out of 100) in which a variable was significant (i.e., a 95% credible interval for its coefficient excluded 0).The minimum and maximum value of the posterior mean for all 100 simulations is also reported for the comparison.

Using Neighbor's Information				Removing Neighbor's Information			
Covariate	**Pct Signif**	**Min Mean**	**Max Mean**	**Covariate**	**Pct Signif**	**Min Mean**	**Max Mean**
(Intercept)	100	17.828	20.441	(Intercept)	100	17.924	20.534
LITTERDP	98	-0.015	-0.009	LITTERDP	100	-0.016	-0.009
FUELBED_DP	0	0.0004	0.0013	FUELBED_DP	0	0.0002	0.0015
hr10	100	0.254	0.392	hr10	100	0.254	0.387
hr100	100	1.778	3.168	hr100	100	1.660	3.154
PERCHED_P	0	-0.005	0.100	PERCHED_P	0	0.002	0.103
PERCHED_O	71	0.309	0.541	PERCHED_O	66	0.319	0.595
WIREGRASS	0	-0.245	0.003	WIREGRASS	0	-0.233	0.018
GRAMINOIDS	95	-0.341	-0.139	GRAMINOIDS	98	-0.333	-0.146
SHRUBS	48	0.154	0.289	SHRUBS	48	0.159	0.292
FORBS	2	-0.142	-0.061	FORBS	0	-0.133	-0.066
BAREGROUND	0	-0.080	0.049	BAREGROUND	0	-0.082	0.072
DECLITTER	100	0.329	0.536	DECLITTER	100	0.328	0.541
LLPLITTER	100	-0.455	-0.313	LLPLITTER	100	-0.447	-0.306
Ave.W.m.sec	33	-0.126	0.876	Ave.W.m.sec	33	-0.136	0.818
Ave.E.m.sec	100	-3.230	-3.013	Ave.E.m.sec	100	-3.213	-2.997
W.X.sin.wind	100	-2.568	-1.847	W.X.sin.wind	100	-2.538	-1.863
W.Y.cos.wind	100	-5.381	-2.369	W.Y.cos.wind	100	-5.130	-2.385
E.X.sin.wind	100	-4.982	-4.597	E.X.sin.wind	100	-4.933	-4.609
E.Y.cos.wind	100	4.010	4.472	E.Y.cos.wind	100	3.954	4.473
Ave.W.m.sec:W.X.sin.wind	100	0.954	1.420	Ave.W.m.sec:W.X.sin.wind	100	0.983	1.396
Ave.W.m.sec:W.Y.cos.wind	100	1.021	2.080	Ave.W.m.sec:W.Y.cos.wind	100	1.0173	1.969
Ave.E.m.sec:E.X.sin.wind	100	4.390	4.631	Ave.E.m.sec:E.X.sin.wind	100	4.387	4.620
Ave.E.m.sec:E.Y.cos.wind	100	-3.722	-3.402	Ave.E.m.sec:E.Y.cos.wind	100	-3.643	-3.396

B.4 Reconstruction Workflow

For the reconstruction we do not use a fixed seed. At locations where linear interpolation cannot be done, the guiding point, γ, is picked from a single Beta draw. During the pre-ignition reconstruction there are 42 out of 169 locations that utilize this Beta draw. For the knot placement in the post-ignition splines, the knots are placed at the discrete time steps of all truly observed (above 300°C) temperature recordings. The number of these time steps varies based on the location. If the location has a gap between the last true recorded temperatures, then the knot is placed before the gap. For example, if the fire burns above 300°C at time step 40, then falls below 300°C and is censored, and then rises above 300°C at time step 43, then time step 40 is the last knot placement. In order to vary the reconstruction algorithm for the purpose of our sensitivity analysis, the upper bound for the knot is placed at this location 60% of the time; one time step before 20% of the time; and one time step after 20% of the time.

Table B2: Table of key parameters, excluding the post-ignition knot placement, for reconstruction of fire temperature trajectories. If the selection method is based on shape, then the shape of the fire temperature curve at the location of interest influences that parameter's value.

Parameter	Range	Selection Method
Beta Parameterization Pre-ignition	0 - 100	Based on Shape
Degree of Freedom LB	5 - 10	Random (uniform)
Anchor Point Coefficient	3 - 8	Random (uniform)
Anchor Point Exponent	1.01 - 1.40	Random (uniform)
Beta Parameterization Post-ignition	0 - 160	Based on Shape

Appendix C Model and Covariate Information

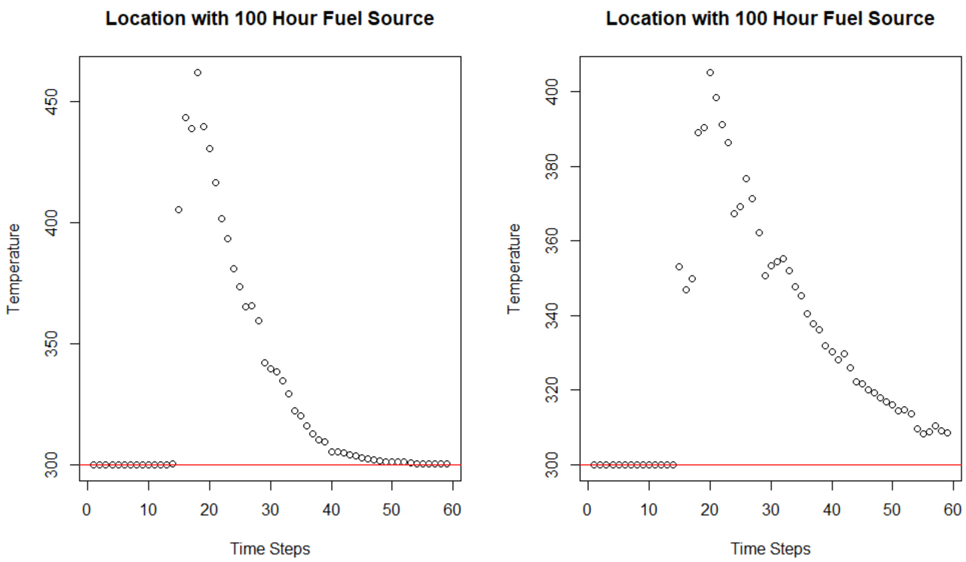

Fig. C1: Temperature functions for two locations that contain 100-hour fuel sources

Table C3: Percentage of simulated (varying tuning parameter values) data sets (out of 100) in which a variable was significant (i.e., a 95% credible interval for its coefficient excluded 0). The minimum and maximum value of the posterior mean for all 100 simulations is also reported.

Covariate	**Pct Signif**	**Min Mean**	**Mean**	**Max Mean**
(Intercept)	100	17.828	19.382	20.441
LITTERDP	98	-0.015	-0.012	-0.009
FUELBED_DP	0	0.0004	0.001	0.0013
hr10	100	0.254	0.323	0.392
hr100	100	1.778	2.362	3.168
PERCHED_P	0	-0.005	0.044	0.100
PERCHED_O	71	0.309	0.443	0.541
WIREGRASS	0	-0.245	-0.091	0.003
GRAMINOIDS	95	-0.341	-0.223	-0.139
SHRUBS	48	0.155	0.218	0.289
FORBS	2	-0.142	-0.102	-0.061
BAREGROUND	0	-0.080	-0.023	0.049
DECLITTER	100	0.329	0.428	0.536
LLPLITTER	100	-0.455	-0.380	-0.313
Ave.W.m.sec	33	-0.126	0.265	0.876
Ave.E.m.sec	100	-3.230	-3.123	-3.014
W.X.sin.wind	100	-2.568	-2.216	-1.847
W.Y.cos.wind	100	-5.381	-3.522	-2.370
E.X.sin.wind	100	-4.982	-4.795	-4.597
E.Y.cos.wind	100	4.010	4.226	4.472
Ave.W.m.sec:W.X.sin.wind	100	0.954	1.227	1.420
Ave.W.m.sec:W.Y.cos.wind	100	1.021	1.420	2.080
Ave.E.m.sec:E.X.sin.wind	100	4.390	4.516	4.631
Ave.E.m.sec:E.Y.cos.wind	100	-3.722	-3.531	-3.402

Table C4: Table showing the results for a model run with the West anemometer data removed when creating the extrapolated data sets. Percentage of simulated (varying tuning parameter values) data sets (out of 100) in which a variable was significant (i.e., a 95% credible interval for its coefficient excluded 0).The minimum and maximum value of the posterior mean for all 100 simulations is also reported.

Covariate	**Pct Signif**	**Min Mean**	**Mean**	**Max Mean**
(Intercept)	100	15.768	17.563	19.045
LITTERDP	96	-0.015	-0.012	-0.008
FUELBED_DP	0	0.0005	0.0009	0.0014
hr10	100	0.249	0.318	0.371
hr100	100	1.750	2.444	3.029
PERCHED_P	0	0.009	0.056	0.104
PERCHED_O	11	0.165	0.337	0.463
WIREGRASS	5	-0.318	-0.142	-0.017
GRAMINOIDS	95	-0.368	-0.241	-0.147
SHRUBS	31	0.149	0.214	0.287
FORBS	0	-0.145	-0.110	-0.078
BAREGROUND	0	-0.082	0.016	0.101
DECLITTER	100	0.339	0.462	0.590
LLPLITTER	100	-0.443	-0.381	-0.321
Ave.E.m.sec	100	-2.716	-1.985	-1.101
E.X.sin.wind	100	-5.252	-4.332	-3.194
E.Y.cos.wind	100	6.254	7.308	8.676
Ave.E.m.sec:E.X.sin.wind	100	3.087	4.118	4.951
Ave.E.m.sec:E.Y.cos.wind	100	-6.816	-5.718	-4.859

Table C5: Table showing the results for a model run with the West anemometer data removed when creating the extrapolated data sets. In addition, the model was run on only the variables that showed relatively high percentage of significance from the previous table. Percentage of simulated (varying tuning parameter values) data sets (out of 100) in which a variable was significant (i.e., a 95% credible interval for its coefficient excluded 0).The minimum and maximum value of the posterior mean for all 100 simulations is also reported.

Covariate	**Pct Signif**	**Min Mean**	**Mean**	**Max Mean**
(Intercept)	100	16.038	17.499	18.808
LITTERDP	100	-0.0157	-0.012	-0.009
hr10	100	0.293	0.371	0.446
hr100	100	1.855	2.511	3.185
GRAMINOIDS	97	-0.329	-0.229	-0.140
SHRUBS	100	0.250	0.309	0.375
DECLITTER	100	0.330	0.451	0.563
LLPLITTER	100	-0.464	-0.402	-0.328
Ave.E.m.sec	100	-2.615	-1.951	-1.214
E.X.sin.wind	100	-5.125	-4.289	-3.362
E.Y.cos.wind	100	6.345	7.381	8.611
Ave.E.m.sec:E.X.sin.wind	100	3.217	4.076	4.858
Ave.E.m.sec:E.Y.cos.wind	100	-6.746	-5.773	-4.937

Table C6: Table showing the results for a model run with the West anemometer data removed when creating the extrapolated data sets. The range of the mean, standard deviation, lower bound for the credible intervals and upper bound for the credible intervals for all 100 simulations are also reported.

Covariate	**Mean Range**	**SD Range**	**LB Range**	**UB Range**
(Intercept)	16.038 : 18.808	0.345 : 0.427	15.182 : 18.117	16.877 : 19.485
LITTERDP	-0.007 : -0.009	0.004 : 0.005	-0.026 : -0.017	-0.0054 : -0.0002
hr10	0.293 : 0.446	0.102 : 0.127	0.084 : 0.187	0.487 : 0.686
hr100	1.855 : 3.185	0.281 : 0.347	1.292 : 2.512	2.422 : 3.863
GRAMINOIDS	-0.329 : -0.140	0.078 : 0.097	-0.524 : -0.304	-0.136 : 0.020
SHRUBS	0.250 : 0.375	0.090 : 0.112	0.071 : 0.172	0.430 : 0.580
DECLITTER	0.330 : 0.563	0.066 : 0.081	0.198 : 0.403	0.458 : 0.719
LLPLITTER	-0.464 : -0.328	0.062 : 0.076	-0.619 : -0.459	-0.320 :-0.206
Ave.E.m.sec	-2.615 : -1.214	0.224 : 0.076	-3.058 : -1.764	-2.173 : -0.666
E.X.sin.wind	-5.125 : -3.362	0.389 : 0.481	-5.876 : -4.293	-4.375 : -2.433
E.Y.cos.wind	6.345 : 8.611	0.338 : 0.418	5.648 : 7.753	6.982 : 9.397
Ave.E.m.sec:E.X.sin.wind	3.217 : 4.858	0.258 : 0.319	2.594 : 4.354	3.834 : 5.359
Ave.E.m.sec:E.Y.cos.wind	-6.746 : -4.937	0.209 : 0.259	-7.238 : -5.335	-6.222 : -4.512
sig2eps	0.005 : 0.005	0.0001 : 0.0001	0.005 : 0.005	0.005 : 0.005
sig2eta	8.588 : 13.149	0.122 : 0.186	8.357 : 12.796	8.827 : 13.515
phi	0.15 : 0.15	0.0 : 0.0	0.15 : 0.15	0.15 : 0.15

Table C7: Goodness of Fit, Penalty, and PMCC term for a model with a Unif(0.15,0.20) prior on ϕ, a model with a Gamma(2,1) prior on ϕ, and a model with a Unif(0.01,0.99) prior on ϕ. We chose Unif(0.15,0.20) prior since it yielded the lowest PMCC. The Uniform prior was evaluated on a grid of five equally spaced points.

Prior on ϕ	Unif(0.15,0.20)	Gamma(2,1)	Unif(0.01,0.99)
GOF	39.32	37,221.71	42,247.72
Penalty	101,016.77	126,967.23	80,560.94
PMCC	101,056.09	164,188.94	122,808.66

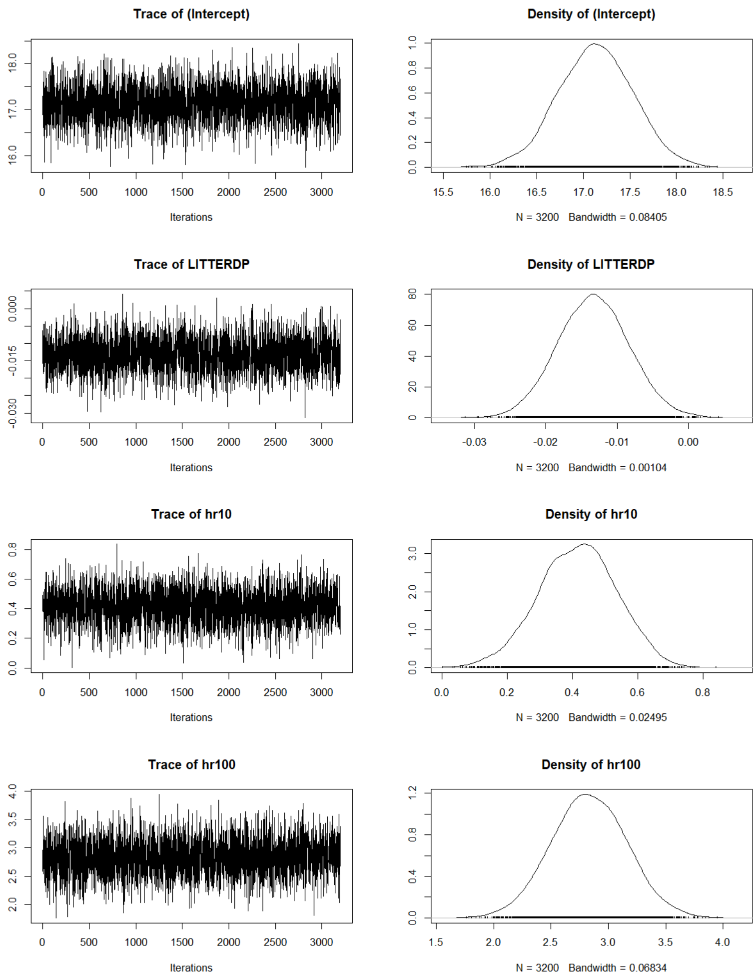

Fig. C2: Diagnostics from the final model including the trace of the intercept, litter bed depth, 10 hour fuel source and 100 hour fuel source variables.

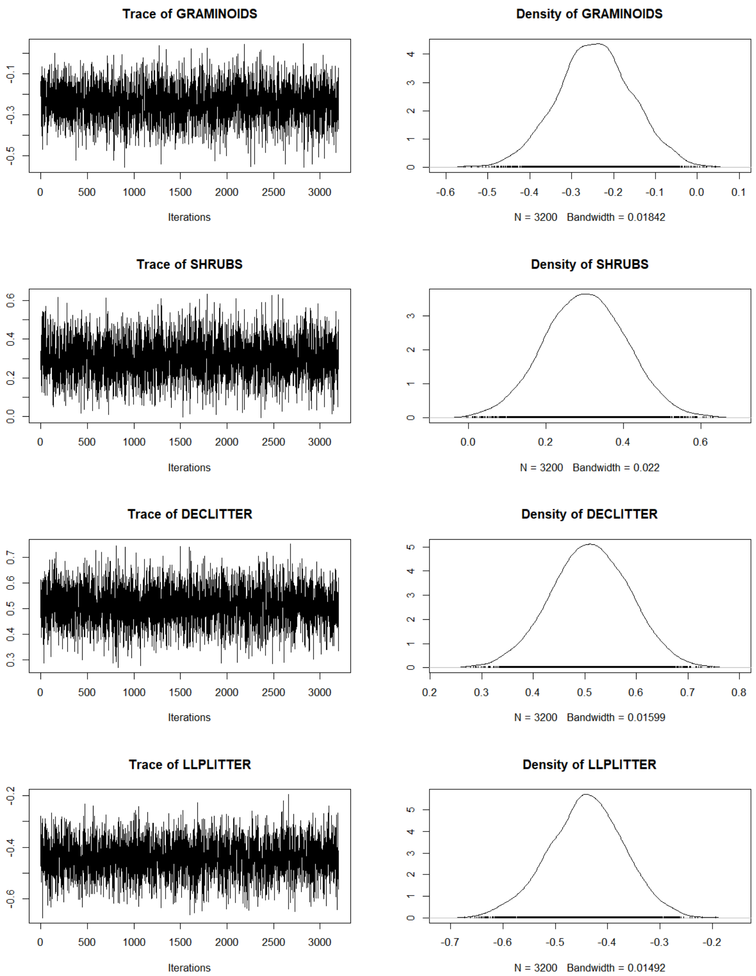

Fig. C3: Diagnostics from the final model including the trace of the graminoids, shrubs, deciduous litter and long leaf pine litter variables.

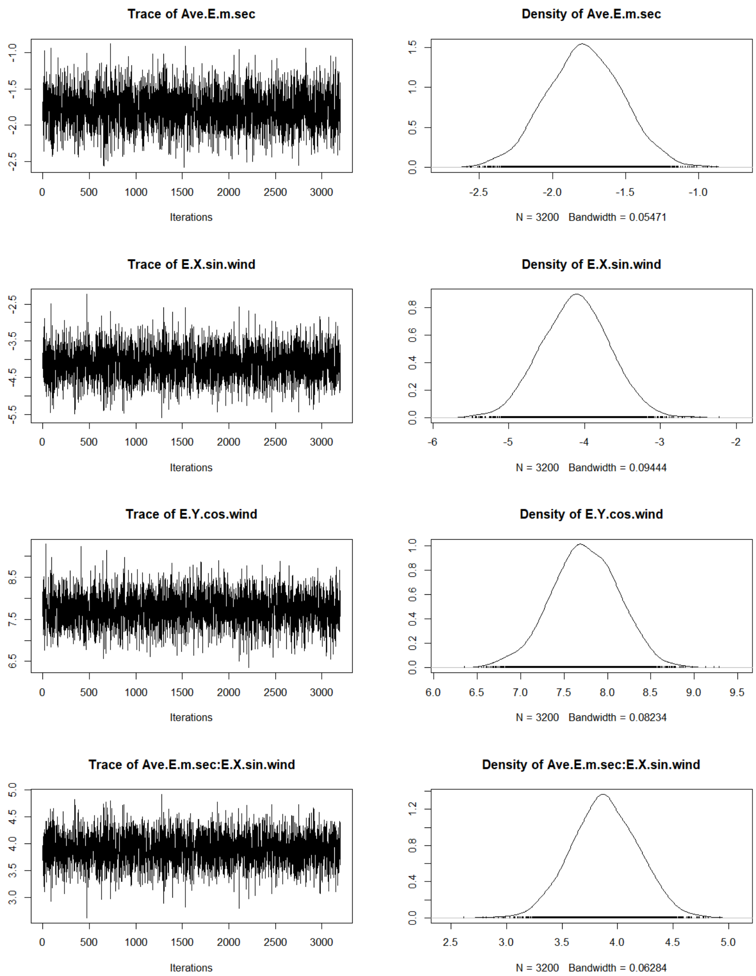

Fig. C4: Diagnostics from the final model including the trace of the wind speed and direction variables from the East anemometer and their interactions.

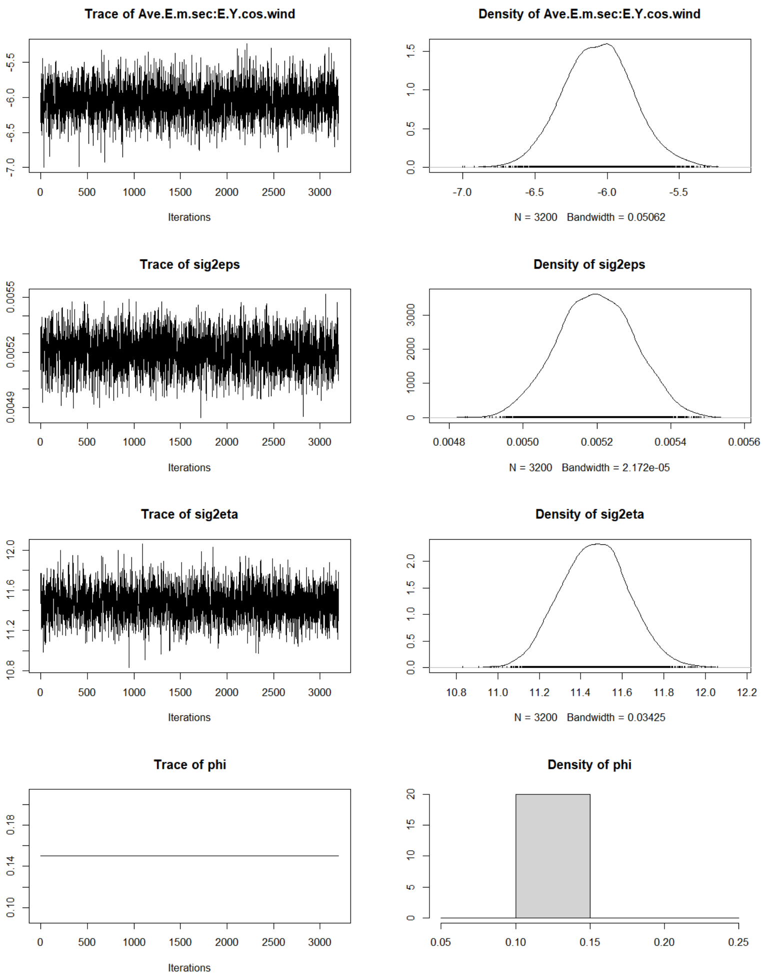

Fig. C5: Diagnostics from the final model including the trace of the last wind speed and direction interaction variable, epsilon, eta, and phi.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] Rothermel, R.C.: A Mathematical Model for Predicting Fire Spread in Wildland Fuels. Intermountain Forest & Range Experiment Station, Forest Service, U.S. Department of Agriculture, Ogden, Utah

[2] Linn, R.R., Goodrick, S.L., Brambilla, S., Brown, M.J., Middleton, R.S., O'Brien, J.J., Hiers, J.K.: Quic-fire: A fast-running simulation tool for prescribed fire planning. Environmental Modelling & Software **125**, 104616 (2020)

[3] Hoffman, C.M., Sieg, C.H., Linn, R.R., Mell, W., Parsons, R.A., Ziegler, J.P., Hiers, J.K.: Advancing the science of wildland fire dynamics using process-based models. Fire **1**(2), 32 (2018)

[4] Wikle, C.K., Zammit-Mangion, A., Cressie, N.: Spatio-temporal Statistics with R. Chapman and Hall/CRC, Boca Raton, FL (2019)

[5] Cressie, N.: The origins of kriging. Mathematical Geology **22**(3), 239–252 (1990) https://doi.org/10.1007/bf00889887

[6] Kupfer, J.A., Lackstrom, K., Grego, J.M., Dow, K., Terando, A.J., Hiers, J.K.: Prescribed fire in longleaf pine ecosystems: Fire managers' perspectives on priorities, constraints, and future prospects. Fire Ecology **18**(1), 27 (2022)

[7] Oswalt, C.M., Cooper, J.A., Brockway, D.G., Brooks, H.W., Walker, J.L., Connor, K.F., Oswalt, S.N., Conner, R.C.: History and current condition of longleaf pine in the southern united states. (2012)

[8] Varner III, J.M., Gordon, D.R., Putz, F.E., Hiers, J.K.: Restoring fire to long-unburned pinus palustris ecosystems: novel fire effects and consequences for long-unburned ecosystems. Restoration Ecology **13**(3), 536–544 (2005)

[9] Mitchell, R.J., Hiers, J., O'Brien, J.J., Jack, S., Engstrom, R.: Silviculture that sustains: the nexus between silviculture, frequent prescribed fire, and conservation of biodiversity in longleaf pine forests of the southeastern united states. Canadian Journal of Forest Research **36**(11), 2724–2736 (2006)

[10] Pimentel, J.S., Bulhões, R.S., Rodrigues, P.C.: Bayesian spatio-temporal modeling of the Brazilian fire spots between 2011 and 2022. Scientific Reports **14**(1), 21616 (2024)

[11] Lin, L., Meng, Y., Yue, A., Yuan, Y., Liu, X., Chen, J., Zhang, M., Chen, J.: A spatio-temporal model for forest fire detection using HJ-IRS satellite data. Remote Sensing **8**(5), 403 (2016)

[12] Vega-Martínez, E., Molina, J.R., Barrón, V., Rodríguez y Silva, F., Carmen del Campillo, M., Sánchez-Rodríguez, A.R.: Spatio-temporal assessment of soil properties immediately and eight months after a high intensity-controlled burn in the south of Spain. Science of the Total Environment **898**, 165368 (2023) https://doi.org/10.1016/j.scitotenv.2023.165368

[13] Augustin, N.H., Musio, M., Wilpert, K., Kublin, E., Wood, S.N., Schumacher, M.: Modeling spatiotemporal forest health monitoring data. Journal of the American Statistical Association **104**(487), 899–911 (2009)

[14] Fasso, A., Cameletti, M.: A unified statistical approach for simulation, modeling, analysis and mapping of environmental data. Simulation **86**(3), 139–153 (2010)

[15] Hiers, J.K., O'Brien, J.J., Mitchell, R., Grego, J.M., Loudermilk, E.L.: The wildland fuel cell concept: an approach to characterize fine-scale variation in fuels and fire in frequently burned longleaf pine forests. International Journal of Wildland Fire **18**(3), 315–325 (2009)

[16] Outcalt, K.W.: The longleaf pine ecosystem of the south. Native Plants Journal **1**(1), 42–53 (2000)

[17] Loudermilk, E.L., O'Brien, J.J., Mitchell, R.J., Cropper, W.P., Hiers, J.K., Grunwald, S., Grego, J., Fernandez-Diaz, J.C.: Linking complex forest fuel structure and fire behaviour at fine scales. International Journal of Wildland Fire **21**(7), 882–893 (2012)

[18] Dell, J.E., Richards, L.A., O'Brien, J.J., Loudermilk, E.L., Hudak, A.T., Pokswinski, S.M., Bright, B.C., Hiers, J.K., Williams, B.W., Dyer, L.A.: Overstory-derived surface fuels mediate plant species diversity in frequently burned longleaf pine forests. Ecosphere **8**(10), 01964 (2017)

[19] Mitchell, R.J., Hiers, J.K., O'Brien, J., Starr, G.: Ecological forestry in the southeast: understanding the ecology of fuels. Journal of Forestry **107**(8), 391–397 (2009)

[20] FLIR Inc. S60: Forward-looking infrared digital thermal imagery system. Boston, MA

[21] O'Brien, J.J., Loudermilk, E.L., Hornsby, B., Hudak, A.T., Bright, B.C., Dickinson, M.B., Hiers, J.K., Teske, C., Ottmar, R.D.: High-resolution infrared thermography for capturing wildland fire behaviour: Rxcadre 2012. International Journal of Wildland Fire **25**(1), 62 (2016) https://doi.org/10.1071/wf14165

[22] R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2024). R Foundation for Statistical Computing. https://www.R-project.org/

[23] Christopoulos, D.T.: inflection: Finds the Inflection Point of a Curve. (2022). R package version 1.3.6. https://CRAN.R-project.org/package=inflection

[24] Muggeo, V.M.R.: segmented: an R package to fit regression models with broken-line relationships. R News **8**(1), 20–25 (2008)

[25] Bakar, K.S., Sahu, S.K.: spTimer: Spatio-temporal Bayesian modeling using R. Journal of Statistical Software **63**(15), 1–32 (2015)

[26] Matérn, B.: Spatial Variation vol. 36. Springer, New York (2013)

[27] Gelfand, A.E.: Hierarchical modeling for spatial data problems. Spatial Statistics **1**, 30–39 (2012) https://doi.org/10.1016/j.spasta.2012.02.005

[28] Banerjee, S., Carlin, B.P., Gelfand, A.E.: Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC, New York (2003)

[29] Robert, C.P., Casella, G.: Monte Carlo Statistical Methods, 2nd edn. Springer, New York, NY (2004)

[30] Gelfand, A.E., Ghosh, S.K.: Model choice: a minimum posterior predictive loss approach. Biometrika **85**(1), 1–11 (1998)

[31] Tobler, W.R.: A computer movie simulating urban growth in the Detroit region. Economic Geography **46**(sup1), 234–240 (1970) https://doi.org/10.2307/143141

[32] Xu, Q., Li, W., Liu, J., Wang, X.: A geographical similarity-based sampling method of non-fire point data for spatial prediction of forest fires. Forest Ecosystems **10**, 100104 (2023) https://doi.org/10.1016/j.fecs.2023.100104