Advances in Forest Fire Research

DOMINGOS XAVIER VIEGAS

EDITOR

2014

Generation of simulated ignitions for the continental United States

Isaac C. Grenfell[a], Mark A. Finney[b], Dianne Trethewey[c]

[a] *RTL Networks, 5775 US Highway 10 W, Missoula, Montana, USA, 59808, igrenfell@gmail.com*
[b] *Missoula Fire Sciences Laboratory, Rocky Mountain Research Station, U.S. Forest Service,5775 US Highway 10 W, Missoula, Montana, USA, 59808, mfinney@fs.fed.us*
[c] *Missoula Fire Sciences Laboratory, Rocky Mountain Research Station, U.S. Forest Service,5775 US Highway 10 W, Missoula, Montana, USA, 59808, dtrethewey@fs.fed.us*

Abstract
Fire suppression continues to be a costly endeavour in the United States and elsewhere. In order to develop a systematic approach to allocating resources to fight wildland fire, it is necessary to understand the underlying process by which ignitions occur on a continental scale with respect to fire weather. We therefore developed a technique that simulated a gridded fire weather index from historical observations and then simulated ignitions at a local scale. The procedure required two steps. First, a model for fire-days (at least one fire occurring on each 30km pixel) used the localized percentile of the ERC (energy release component) as a predictor. A logistic regression model using ERC percentile at that pixel predicted a binary response (fire-day/no-fire-day). Then, for those pixels where at least one ignition occurs, a VGAM (vector generalized additive model) with a Pareto response also used ERC percentile as the predictor. This yielded a dataset of simulated ERC percentiles and the associated number of ignitions for each pixel and each day. The related metadata for each ignition (cause, time of detection, etc.) are sampled from the historical distributions conditional on location. What results is a simulated equivalent of the Fire Occurrence Database which is ready for use in applications for strategic planning.

Keywords*: Spatial temporal simulation, ignition generation, energy release component, Pareto distribution*

1. Introduction

In order to develop a system to test various fire suppression resource allocation strategies, it is necessary to have a mechanism to simulate the distribution of ignitions across very large land areas. The main problems facing resource allocation are 1) local scarcity of resources created by ignition episodes (large number in short time frame), and 2) broad-scale scarcity of resources created by spatial synchronicity of fire ignition episodes. Firefighting resources dispatched and coordinated at a national scale, meaning that fire occurrence patterns must reflect spatial and temporal patterns at that scale.
There are several attributes that are necessary to constitute a good practical simulation method. First, the distribution of the simulated ignitions must follow approximately that of the historical distribution at the local scale. This means that the overall shape of the distribution, the rough total number of ignitions, as well as extreme behavior must be similar to a great extent. Second, the overall number and seasonal distribution over the entire country must be similar. Third, the method must make use of the historical relationship between the number of ignitions and fire weather. Regarding the first two points, there is no statistical method that systematically provides all of these desired attributes simultaneously. Thus, we must make use of expertise and intuition in order to evaluate what constitutes an acceptable or unacceptable method.

2. Methods

The method consists of statistically simulating the spatial and temporal patterns of fire occurrence based on gridded daily fire danger rating index. In (Grenfell 2010), a method was described to simulate fire weather at a continental scale that has the same seasonal variability, spatial, and temporal autocorrelation. The fire danger rating index Energy Release Component (ERC) from this modeling was used to generate ignitions. The key result from the simulation method developed there was to have simulated fire weather with similar temporal and spatial autocorrelation as the historical data. We rely on the relationship between ignition occurrence and this spatially/temporally correlated weather to provide generate the pattern of coincidence (both spatially and temporally) of ignitions. The technique relies upon a fire occurrence database (FOD) (Short 2013) which contains fire records from 1992-2011 coincident with the years generated by the weather simulation model. To choose the appropriate model, data exploration were conducted.

From the FOD, we see that the distribution of the number of ignitions at a particular pixel generally appears to follow a Pareto distribution. In Figure 1, data for a Fire Planning Unit (FPU Colorado-05) are plotted as a histogram of the observed number of ignitions, a histogram where zero ignition-days were excluded, a histogram of 1000 simulated Poisson values from the estimated mean of the historical data, and 1000 simulated Pareto values with the same shape parameter as that of the historical data.

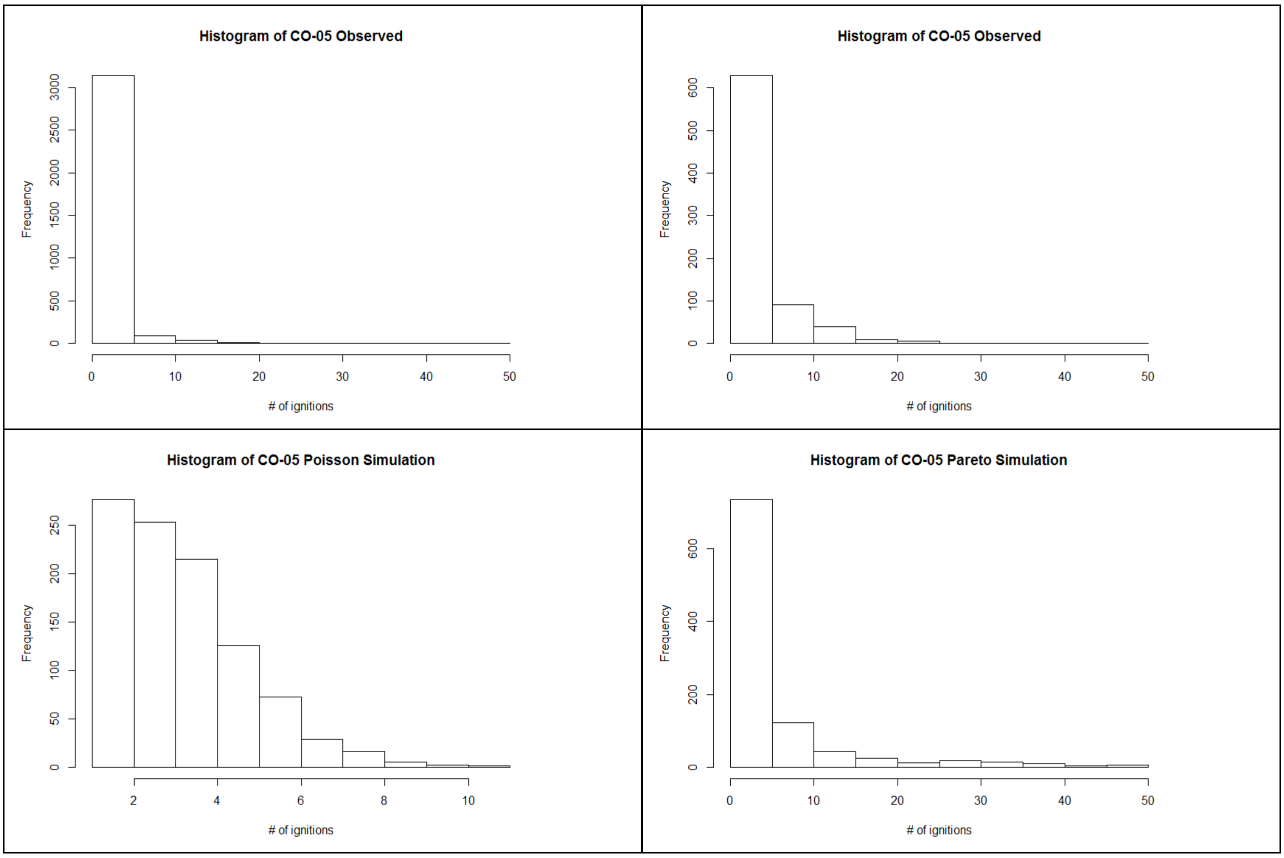

Figure 1- Above left (a), histogram of number of ignitions in FPU CO-05. Above right (b) histogram of number of ignitions on days with at least one ignition. Below left(c) histogram of 1000 simulated days using the average number of ignitions from (b). Below right (d) Histogram of 1000 simulated days using the shape parameter estimated from (b).

Notice that there are more zero-values in the historical histogram than in the simulated dataset. This is a commonly observed phenomenon known as overdispersion. In order to resolve this, we built a two stage model. For the first model, we perform a logistic regression with ERC percentile from the NARR grids being the predictor and the response being whether there are zero ignitions (0) or greater than zero (1). We use percentiles rather than just the ERC values because there are different local relationships with ERC and ignitions (a location where an ERC value of 60 would be in the 90th percentile would be indicative of high fire danger in some places, whereas a location where an ERC value of 60 would be in the 50th percentile in another location would be indicative of far more moderate fire danger).

Once we have developed the logistic models to determine whether ignitions (at least one ignition) have occurred or not, we then use the VGAM package in R in order to perform a Vector Generalized Additive Model regression. This regression method allows for a much broader set of response variables, in this case a response with a Pareto distribution. It is here that the virtues of the Pareto distribution become a hindrance, as left unsupervised the erratic nature of distribution in the right tail can come to dominate the overall phenomenon. To avoid this, we truncate the maximum number of ignitions that can occur at one pixel in one day at fifty percent above the value of the 4th from last order statistic. This allows for the simulation method to produce results greater than what was observed in the historical record (which is the reason for not simply using the historical distribution in a bootstrap-like approach), while not having values so extreme that the overall behavior of the simulated results becomes unrecognizable.

With the two sets of models in hand, we can now begin the process of simulation. It is essentially just the estimation process but in reverse. First, we derive the ERC percentiles from the simulated ERC grids. We then use the percentiles at each location on each day to determine from the logistic model whether or not there will be ignitions. If there are, we use that ERC percentile in the VGAM model to determine how many ignitions will be simulated on that day. In figure 2, we se a selected day from a simulated season (Julian day 180), with a grid of the ERC values (not percentiles). In addition, we can pass on simulated metadata values taken from the FOD, such as cause and time of day. For our purposes, we simply built tables for each location and drew a sample from that.

3. **Results**

In order to assess the overall quality of the simulation method, we evaluate the simulated tables of ignitions and consider several qualities- the overall number of ignitions, the behavior of the extreme values of the distribution, and patterns of seasonality. To get a sense of the scale and variability in the historical and simulated data, we examine a table of the mean, median, standard deviation, median absolute deviation, and maximum for both the 19 historical years and the 41 simulated years.

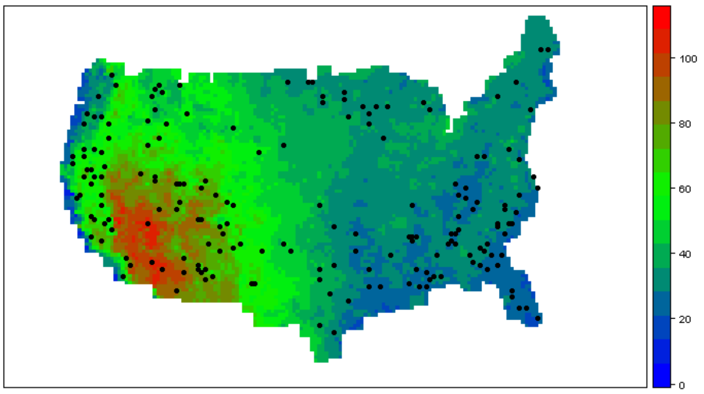

Figure 2. An example of a day's worth of simulated ERC values across the continental United States and the associated simulated ignitions for that day.

	Historical (19 yrs)	*Simulated (41 yrs)*
Mean	65,309	71,780
Median	61,225	71,556
SD	14,369	17,554
MAD	15,280	21,822
Maximum	103,120	109,930

Both the mean and median are within one standard deviation or MAD (respectively) for both the historical and simulated sets of ignitions, providing good evidence that the simulation method is doing a good job of matching the overall behaviour of the historical data.

4. References

Grenfell, I.C. , Finney, M. A, Jolly, M. Simulating spatial and temporally related fire weather. Proceedings of the VI International Conference on Forest Fire Research; 15-18 November 2010; Coimbra, Portugal. Coimbra, Portugal: University of Coimbra.

Short, K. C. A spatial database of wildfires in the United States, 1992-2011 Fort Collins, CO: USDA Forest Service, Rocky Mountain Research Station. http://dx.doi.org/10.2737/RDS-2013-0009