

Desarrollo Preliminar de un Modelo de Probabilidades, para Establecer la Gravedad de los Incendios en los Bosques de California¹

John W. Benoit,² Haiganoush K. Preisler³

Resumen

En este estudio en progreso, examinamos datos históricos sobre incendios, climatología y condiciones del combustible en los bosques nacionales de California. Construimos modelos de regresión logística para determinar las probabilidades de episodios ígneos y las de que éstos se conviertan en grandes incendios. Utilizamos la misma metodología que la que desarrollamos previamente en un estudio en el que se utilizaron datos del Estado de Oregón. Nuestro objetivo es seguir desarrollando este proceso de evaluación de riesgos para que sirva de herramienta a responsables de gestión y a planificadores.

Introducción

Las autoridades responsables de gestionar los incendios en California confían en una serie de factores para determinar la gravedad de un incendio y poder adoptar decisiones referentes a la asignación de recursos y a las tácticas de extinción. Debido a la gran diversidad medioambiental de California hay que utilizar una serie de variables para describir las condiciones meteorológicas, el grado de humedad del combustible y el comportamiento del fuego. Un indicador del potencial de un incendio en una zona puede no resultar eficaz en otra. Cada experto local puede por lo tanto elegir una serie de indicadores que le resulten útiles, pero ello puede causar dificultades a la hora de comparar la gravedad de los incendios en diferentes regiones.

En los EE.UU. el Sistema Nacional de Cálculo de Peligro de Incendios (NFDRS por sus siglas en inglés) es un proceso que permite obtener una serie de indicios para cualquier región del país, dadas ciertas condiciones meteorológicas y de combustible. Estos indicios proporcionan una idea sobre el potencial de un incendio y sobre el comportamiento de éste en una zona dada. Los expertos los utilizan para establecer en qué momento hay que efectuar una quema controlada, para evaluar la necesidad de utilizar recursos de extinción y para adoptar decisiones tácticas en cuanto a ésta. Aunque estos indicios son considerados confiables, su relación con los incendios no ha sido examinada con suficiente detenimiento.

En un estudio previo recogimos el historial de incendios, los datos observados en una estación meteorológica, los valores del índice NFDRS así como otras capas geográficas correspondientes al Estado de Oregón (Preisler et al. 2004). Después construimos ecuaciones de regresión logística para estimar, en un día y en un lugar de Oregón, 1) las probabilidades de ignición y 2) las probabilidades de que una ignición

¹ Una versión abreviada de esta ponencia se presentó en el segundo simposio internacional sobre economía, política y planificación de incendios: una visión global, 19–22 de Abril de 2004, Córdoba, España.

² Especialista en Informática, USDA Forest Service, Riverside, California.

³ Estadístico, USDA Forest Service, Albany, California.

se transforme en un gran incendio (de más de 100 acres, o aproximadamente 45 hectáreas). Las probabilidades de que se produzca un gran incendio pueden luego calcularse sobre estas dos probabilidades.

En el presente estudio hemos querido aplicar y adaptar esta metodología a California. Existen algunas diferencias entre ambos estudios, sobre todo relacionadas con los datos observados recabamos. Por ejemplo, en California, podemos obtener una mayor cantidad de capas GIS, más actualizadas y de mayor resolución que las que obtuvimos en el estudio de Oregón. California tiene también una mayor diversidad geográfica y climatológica. Sin embargo, estos factores no deberían influir en nuestra metodología, ya que los modelos que construimos tienen en cuenta estas diferencias mediante variables aclaratorias, y son robustos en cualquier punto geográfico.

Datos

El Sistema de Evaluación de Incendios Forestales (WFAS) es una recopilación de datos observados y productos de previsión para establecer la gravedad de los incendios en el interior de los EE.UU. Desde enero del 2000 el WFAS publica online los datos brutos (tabulares) observados en el catálogo de las estaciones pertenecientes al Sistema de Gestión de Información Meteorológica (WIMS). Nosotros trabajamos con los datos correspondientes a cada día y a cada estación entre 2000 y 2002 incluido. En el informe diario de cada estación consta la temperatura, la humedad relativa, la velocidad del viento, la precipitación, el componente de liberación de energía (índice NFDRS), el componente de propagación (índice NFDRS), el índice de sequía Keetch-Byram y los valores de humedad del combustible a las diez, cien y mil horas. Resultó ser una gran ventaja que estos datos incluyeran información del NFDRS además de la información sobre las condiciones meteorológicas y sobre el combustible, ya que generar índices NFDRS a partir de datos históricos es un proceso complicado. Antes del 2000, los índices históricos del NFDRS no se archivaban ni tampoco eran fácilmente accesibles algunas de las variables del sistema. Por ejemplo, la temperatura observada del combustible –un dato clave- era rara vez registrada en las estaciones del WIMS.⁴

Nosotros obtuvimos los datos observados correspondientes a todos los incendios que se declararon en los bosques nacionales de California entre 2000 y 2002. Los datos recabados provenían de la base de datos NIFMID (National Interagency Fire Management Integrated Database).⁵ Incluían la fecha del incendio, el lugar, la causa sospechada y su tamaño. Como este conjunto de datos no se refiere solamente a los incendios forestales sino también a aquellos que supusieron una amenaza para los bosques (p.e. incendios que se declararon en terrenos adyacentes), utilizamos los límites de los bosques nacionales de California como una máscara para asegurarnos de disponer de una lista de incendios cuyo punto de ignición realmente se produjo dentro del bosque. Además, sólo incluimos incendios que consumieron más de 0.1 acres y que habían sido clasificados como incendios forestales. 746 incendios respondieron a estos criterios durante el período de estudio (Figura 1).

⁴ En nuestro estudio sobre Oregon utilizamos los índices NFDRS de 1989 hasta 1996 generados por un proceso diferente que el que usamos para crear los índices de las estaciones WIMS de la WFAS. En el proceso de desarrollar nuestro procedimiento para California, no quisimos combinar estos dos tipos de conjuntos de datos diferentes hasta que pudiéramos verificar que sus características generales son similares.

⁵ Consultar <http://famweb.nwcg.gov/kcfast/html/ocmenu.htm>.

Los datos de altitud de los bosques nacionales se obtuvieron por kilómetro. Como pretendemos que el procedimiento que estamos desarrollando pueda aplicarse en regiones de gran extensión, un espaciado de valores de un kilómetro nos pareció lo más adecuado.

Y por último, tomamos una capa del modelo de combustible de California del Programa del Departamento de Bosques, Protección contra Incendios y Evaluación de Recursos de California (FRAP),⁶ un mapa en el que se clasifica la vegetación por su forma de ignición. Se hicieron nuevas muestras con los datos del modelo de combustible para obtener un tamaño de celda de un kilómetro.

Todas las capas de datos fueron georegistradas (alineadas en el espacio) lado a lado. Convertimos todas las coordenadas a la proyección geográfica acimutal de Lambert, para colocar los puntos geográficos en un plano cartesiano. Esta metodología resulta conveniente para estimar el número esperado de igniciones o de grandes incendios en un día concreto.

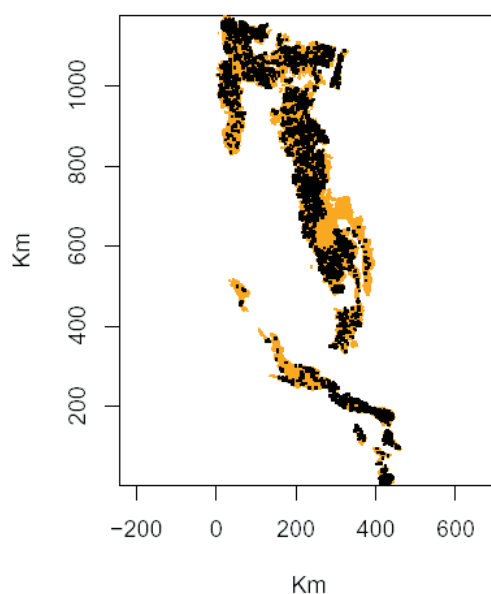


Figura 1—Incendios federales en California 2000-2003. La zona sombreada representa los bosques nacionales del estado de California. Los puntos negros representan el punto de origen de los incendios.

Metodología

En el estudio, los datos observados se representaron en una rejilla cartesiana por medio de una unidad espacial de un kilómetro y temporal de un día. La unidad básica que utilizamos en la muestra fue de un solo kilómetro cuadrado por día. Esta unidad fue denominada voxel. Los datos sobre bosques nacionales de California en los años 2000- 2002, contenían 117.438.750 voxels, en 5.317 de los cuales se produjo al menos una ignición. Como la proporción de voxels con incendio y voxels sin incendio es extremadamente baja, se implantó una técnica de muestreo en dos fases (Brillinger 2003) la cual requería definir un conjunto con todos los voxels con incendio y una

⁶ Consultar <http://frap.cdf.ca.gov/>.

pequeña muestra de voxels sin incendio⁷. A continuación, se puede realizar una estimación de los parámetros utilizando el conjunto más pequeño de voxels en vez de toda la población de datos del estudio.

Cada día, la meteorología observada, así como la humedad del combustible y los valores NFDRS registrados en las estaciones específicas son interpolados a partir de los valores de las estaciones WFAS a los voxels individuales. Puede utilizarse una interpolación ponderada inversa a la distancia para llegar a estimar los valores de las celdas de un kilómetro en un día dado. Alternativamente, otros métodos de interpolación, como el de loess o el de acanalamientos de placas delgadas, proporcionarán probablemente valores más exactos. Las rejillas con los valores interpolados se guardan para más tarde.

Consideremos ahora una variable indicativa como:

$$N_{x,y,t} = 1 \quad \text{si ha habido por lo menos una ignición en el voxel del punto } (x, y) \text{ en el día } t, y$$

$$N_{x,y,t} = 0 \quad \text{si no ha habido igniciones en el voxel}$$

A continuación podemos definir la probabilidad p_1 en la que la ignición se da en el voxel (x,y,t) mediante el modelo logístico:

$$p_{1,x,y,t} = \text{Prob}(N_{x,y,t} = 1 \mid U_{x,y,t}) = \frac{e^{\theta_{x,y,t}}}{1 + e^{\theta_{x,y,t}}}$$

en el que $U_{x,y,t}$ representa los valores históricos de las variables aclaratorias del voxel (x,y,t) y $\theta_{x,y,t}$ representa el conjunto de parámetros que queremos calcular. El *logit* de $p_{1,x,y,t}$ se define como:

$$\text{logit}(p_{1,x,y,t}) = \log\left(\frac{p_{1,x,y,t}}{1 - p_{1,x,y,t}}\right) \quad (2)$$

A partir de (1) y (2), tenemos

$$\text{logit}(p_{1,x,y,t}) = \theta_{x,y,t} = \beta_{0,x,y,t} + \beta_{1,x,y,t}x_{1,x,y,t} + \dots + \beta_{n,x,y,t}x_{n,x,y,t} \quad (3)$$

en el que n es el número de variables independientes del modelo. El valor de $\text{logit}(p_{1,x,y,t})$ puede variar entre $(-\infty, \infty)$. Asumiendo la independencia de los voxels, podemos utilizar técnicas de regresión para obtener las estimaciones de los coeficientes de (3).⁸ Su utilizamos el modelo aditivo generalizado (GAM) podríamos deducir las relaciones no-lineares que existen entre las variables aclaratorias:

$$\begin{aligned} \text{logit}(p_{1,x,y,t}) = & g_1(\mathbf{x}, \mathbf{y}) + g_2(t) + g_3(\text{Elev}_{x,y,t}) + g_4(\text{FM}_{x,y,t}) + g_4(\text{BI}_{x,y,t}) \\ & + g_5(\text{ERC}_{x,y,t}) + g_6(\text{KB}_{x,y,t}) + g_7(\text{TH}_{x,y,t}) + g_8(\text{WS}_{x,y,t}) \\ & + g_9(\text{RH}_{x,y,t}) + g_{10}(\text{Temp}_{x,y,t}) \end{aligned} \quad (4)$$

⁷ El tamaño de la muestra de voxels sin fuego se determina eligiendo una pequeña proporción (p.e. ~0.15%). Después se realiza una muestra aleatoria a este nivel con todos los voxels sin fuego.

⁸ Naturalmente, no podemos asumir fácilmente que los datos son temporal y espacialmente independientes. Por ejemplo, cuando después de un incendio se ha quemado una zona (un conjunto de voxels), no existen muchas probabilidades de que se vuelva a producir otro incendio en el mismo lugar durante cierto tiempo. Esperamos que algunas de estas relaciones sean tenidas en cuenta por las variables aclaratorias dentro de cada voxel—una zona recientemente incendiada puede presentar un cambio del grado de humedad del combustible, lo cual afectaría a la posibilidad de que se produzca otro incendio en el mismo lugar poco tiempo después.

Aquí, $Elev_{x,y,t}$ representa la elevación; $FM_{x,y,t}$ el modelo de combustible; $BI_{x,y,t}$ el índice de combustión; $ERC_{x,y,t}$ el componente de liberación de energía; $KB_{x,y,t}$ el índice de sequía de Keetch Byram; $TH_{x,y,t}$ el nivel de humedad del combustible a las mil horas; $WS_{x,y,t}$ la velocidad del viento; $RH_{x,y,t}$ la humedad relativa y $Temp_{x,y,t}$ la temperatura del voxel (x,y,t) . Las funciones $g(\bullet)$ son igualadores no-paramétricos calculados simultáneamente con el GAM. $g_1(x,y)$ representa el efecto de la localización y $g_2(t)$ es el efecto temporal del modelo.

La probabilidad condicional $p_{2,x,y,t}$ de que ocurra un gran incendio (de más de 100 acres) debido a una ignición puede ser modelada utilizando la misma técnica que nos permitió calcular $p_{1,x,y,t}$. Pueden usarse lo mismos coeficientes u otros similares. En el caso de $p_{2,x,y,t}$, sin embargo, los coeficientes se calculan utilizando solamente el conjunto de voxels en el cual *sí* hubo incendios.

Finalmente, dados $p_{1,x,y,t}$ y $p_{2,x,y,t}$, podemos ahora calcular fácilmente la probabilidad incondicional de que se produzca un gran incendio, ya que

$$p_{3,x,y,t} = P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \cdot P(\mathbf{B} | \mathbf{A}) = p_{1,x,y,t} \cdot p_{2,x,y,t} \quad (5)$$

donde

\mathbf{A} = episodio de ignición, y

\mathbf{B} = episodio de ignición convertida en gran incendio.

Una de las ventajas características de los voxels como unidad básica para calcular las probabilidades es que se puede obtener rápidamente el número esperado de incidencias de fuego de un día dado. Por ejemplo, el número esperado de incendios dentro del área de estudio o en un subconjunto de ésta en un día específico es la suma de los valores de $p_{1,x,y,t}$ en cada kilómetro en el área deseada para ese día. El número esperado de igniciones que se pueden transformar en grandes incendios por día puede calcularse de la misma forma, utilizando $p_{2,x,y,t}$ y $p_{3,x,y,t}$ respectivamente. Adicionalmente, los intervalos de seguridad pueden derivarse de estas tres posibilidades (cft. Preisler et al. 2004).

Trabajos Adicionales

Este estudio sobre California es un estudio en progreso. Aquí presentamos los resultados obtenidos hasta ahora, así como el procedimiento que pretendemos utilizar (gran parte del cual ha sido realizado para nuestro trabajo sobre Oregón). Existen algunas opciones que deseamos investigar. Nos gustaría trabajar con un período más extenso que el de los años 2000-2002. Nuestro principal obstáculo ha sido lograr obtener los valores históricos de los índices NFDRS para establecer relaciones con los casos de incendio. Actualmente, estamos investigando otros métodos posibles para generar valores históricos NFDRS.

Nuestro procedimiento nos permite testar diferentes métodos de interpolación espacial sobre los datos diarios de las estaciones. La interpolación de distancia inversa ponderada es uno de los enfoques sencillos que se pueden utilizar (se utiliza comúnmente en los productos WFAS), pero se pueden obtener resultados más exactos con métodos localmente ponderados, como las técnicas de loess o la de acanalamiento de placas delgadas.

El trabajo aquí esbozado se basa en datos observados. Sin embargo, nuestra intención es utilizar las ecuaciones de regresión logística que hemos desarrollado para

producir datos de predicción y mapas de probabilidades para calcular la gravedad de una incidencia (Figura 2). Testar los resultados de los diferentes modelos meteorológicos será de gran interés. Por lo tanto, deseamos eventualmente integrar diferentes modelos como el RSM (Roads 2000 & 2002) o el MM5, en nuestro procedimiento.

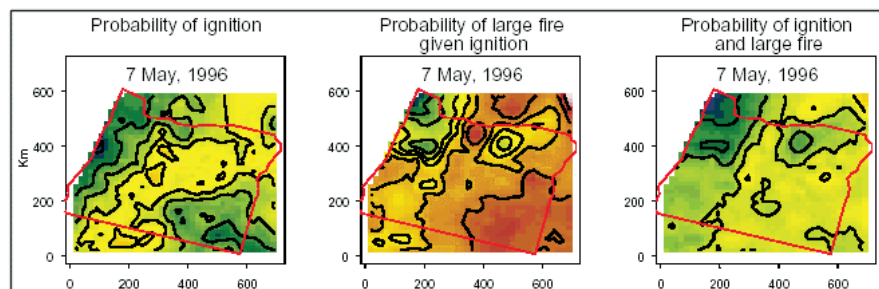


Figura 2—Mapas de probabilidad de que se produzca una ignición, un gran incendio después de una ignición y de una ignición seguida de un gran incendio en Oregon. Los colores van desde el azul/verde oscuro (baja probabilidad) hasta el marrón rojizo (alta probabilidad). Hemos creado mapas similares para California.

Con este trabajo pretendemos eventualmente crear un producto para predecir y calcular la gravedad de un incendio, basándonos en la tecnología y en las herramientas utilizadas actualmente por los profesionales. Esperamos que este sea escalable, y que pueda utilizarse tanto a nivel nacional como a nivel de distrito. Ya por último, tiene que ser sencillo de usar y de comprender. Este enfoque basado en el cálculo de probabilidades está demostrando grandes posibilidades en este sentido.

Bibliografía

- Bradshaw, Larry S.; Deeming, John E.; Burgan, Robert E.; Cohen, Jack D., compiladores. 1984. **The 1978 National Fire-Danger Rating System: technical documentation**. Gen. Tech. Rep. INT-169. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 44 pp.
- Brillinger, D.R.; Preisler, H.K.; Benoit, J. W. 2003. **Risk assessment: a forest fire example**. In: Goldstein, Darlene R., editor. Institute of Mathematical Statistics Lecture Notes – Monograph Series 40: 177-196.
- Preisler, H.K.; Brillinger, David R.; Burgan, Robert E.; Benoit, J.W. 2004. [In press]. **Probability based models for estimation of wildfire**. International Journal of Wildland Fire.
- Roads, J.O.; Chen, S.-C.; Kanamitsu, M. 2002. **US regional climate simulations and seasonal forecasts**. Journal of Geophysical Research-Atmospheres.
- Roads, J.O.; Fujioka, F.; Burgan, R. 2000. **Development of a seasonal fire weather forecast for the contiguous United States**. Preprints. Third symposium on fire and forest meteorology, Long Beach, CA. Amer. Meteor. Soc.; 99-102.
- S-PLUS 2000 guide to statistics, vol 1**. Data Analysis Products Division, MathSoft, Seattle, WA; 638 p.