

Index for the estimation of the occurrence of forest fires in large areas

Índice para la estimación de ocurrencia de incendios forestales en superficies extensas

Juan M. Torres-Rojo

Centro de Investigación y Docencia Económica (CIDE). Carretera México-Toluca 3655, Lomas de Santa Fe. C. P. 01210. Álvaro Obregón, Ciudad de México, México. e-mail: juanmanuel.torres@cide.edu; tel.: +52 55 4850 4207.

Centro Transdisciplinario Universitario para la Sustentabilidad (CENTRUS), Universidad Iberoamericana. Prolongación Paseo de la Reforma 880, Santa Fe. C. P. 01219. Álvaro Obregón, Ciudad de México, México.

Abstract

Introduction: Estimating the risk of occurrence of a fire contributes to reducing human, infrastructure and natural resource losses; promoting activities to maintain and restore fire regimes; and optimizing resources for suppression.

Objective: To develop an index of occurrence of forest fires on large areas, called Area at risk of fire (SeR).

Materials and methods: The index corresponds to the area associated with a probability level measured at the right tail of the density distribution of the area affected annually by forest fires. The density distribution was estimated from the history of the area affected (1970-2018) in Mexico by state. The fit was performed by minimizing the Kolmogorov-Smirnov statistic with four models: exponential, gamma, lognormal and Weibull. Two related indicators are proposed: proportion of forest area affected by wildfires (PSeR) and incremental area at risk (ISeR).

Results and discussion: all models showed a statistically significant fit ($P < 0.05$); the lognormal model performed the best. The SeR discriminates territorial units with the largest area affected by fires; additionally, it efficiently predicts the area to be affected by fires. The PSeR facilitates the comparison of the risk of fire occurrence between territorial units of different sizes, while the ISeR estimates the change in the maximum area affected by fires over a period.

Conclusion: SeR is an extreme event risk index that provides useful information and has a statistically acceptable predictive power.

Resumen

Introducción: La estimación del riesgo de ocurrencia de un incendio contribuye a reducir pérdidas humanas, de infraestructura y recursos naturales; promover actividades para mantener y restaurar regímenes de fuego; y optimizar los recursos destinados a la supresión.

Objetivo: Desarrollar un índice de ocurrencia de incendios forestales en superficies extensas, denominado superficie en riesgo de incendio (SeR).

Materiales y métodos: El índice corresponde a la superficie asociada a un nivel de probabilidad medido en la cola derecha de la distribución de densidad de la superficie afectada anualmente por incendios forestales. La distribución de densidad se estimó a partir del historial de superficie siniestrada (1970-2018) en México por entidad federativa. El ajuste se realizó minimizando el estadístico Kolmogorov-Smirnov con cuatro modelos: exponencial, gama, lognormal y Weibull. Se proponen dos indicadores relacionados: proporción de la superficie forestal afectada (PSeR) y superficie incremental en riesgo (ISeR).

Resultados y discusión: Todos los modelos mostraron un ajuste significativo ($P < 0.05$); el modelo lognormal fue el de mejor desempeño. La SeR discrimina unidades territoriales con mayor superficie afectada por incendios; adicionalmente, predice eficientemente la superficie a ser afectada por incendios. La PSeR facilita la comparación del riesgo de ocurrencia de incendios entre unidades territoriales de tamaño diferente, mientras que la ISeR estima el cambio en la máxima superficie afectada por incendios en un periodo.

Conclusión: La SeR es un índice de riesgo de eventos extremos que brinda información útil y tiene un poder predictivo estadísticamente aceptable.

Keywords: fire risk; fuel accumulation; fire prediction; density distribution; burned area.

Palabras clave: riesgo de incendio; acumulación de combustibles; predicción de incendios; distribución de densidad; superficie siniestrada.

Introduction

Forest fires are considered the most important cause of loss of vegetation cover (Moore, 2019). In tropical regions of Asia and America, the occurrence of fires in wooded areas, is the second most important factor of loss of vegetation only behind Land-Use Change (Food and Agricultural Organization [FAO], 2007). In Mexico, the effects of these incidents are an important cause of loss and degradation of forest cover. An average of $7\,087 \pm 2\,538$ fires are recorded each year (1970-2018) with an average annual burned area of $253\,950 \pm 189\,500.597$ ha and an average burned area by fire of 109 ± 107 ha (Comisión Nacional Forestal [CONAFOR], 2019).

In recent years, forest fires show greater dispersion, frequency, size and severity (International Panel on Climate Change [IPCC], 2007). This increase in the occurrence rate is attributed both to the generation of greater fuel volume, as a result of fire suppression actions, and to climate change (Marlon et al., 2009; Westerling, Hidalgo, Cayan, & Swetnam, 2006). Both drivers have led to significant changes in precipitation and drought patterns, and fuel accumulation processes, which together have changed the dynamics of forest fires (IPCC, 2007). As a result, globally, governments spend a large budget on fire prevention and firefighting. While most of the government expenditure is spent on suppression, the design of prevention and management mechanisms, among which the generation of information to qualify the risk or anticipate the occurrence of a fire, accounts for a significant proportion of the expenditure (FAO, 2007). In this context, the estimation of the occurrence of a fire is significant, since it allows to plan measures to minimize loss of human lives, infrastructure and natural resources (Adab, Kanniah, Solaimani, & Sallehuddin, 2015); to promote activities to maintain and restore fire regimes (Keeley, Fotheringham, & Morais, 1999); and to optimize the distribution of available resources in suppression activities (Podschwit, Larkin, Steele, Cullen, & Alvarado, 2018).

The occurrence of a fire has been estimated through two strategies. The first is to estimate indicators of risk of occurrence based on indirect measures of climatic, physical, environmental attributes or the characteristics of the same fires (*i. e.* statistics of occurrence). This strategy is derived from the absence or low quality and quantity of data on the causal variables of a fire or from the limited knowledge of fire behavior in a particular region (Miller & Ager, 2013; Thompson & Calkin, 2011). The estimates are made through indicators derived from probabilistic analysis of the available information (distributions of sizes) with refinements through regression models, neural networks, mathematical programming and Markovian models (Ager, Vaillant, & Finney, 2010; Miller & Ager, 2013; Podur, Martell, & Stanford,

Introducción

Los incendios forestales se consideran la causa más importante de pérdida de cobertura vegetal (Moore, 2019). En regiones tropicales de Asia y América, la ocurrencia de incendios, en áreas arboladas, es el segundo factor más importante de pérdida de vegetación tan solo detrás del cambio de uso del suelo (Food and Agricultural Organization [FAO], 2007). En México, las afectaciones de estos siniestros son una causa importante de pérdida y degradación de la cobertura forestal. Cada año se registran en promedio $7\,087 \pm 2\,538$ incendios (1970-2018) con una superficie anual afectada promedio de $253\,950 \pm 189\,500.597$ ha y una superficie de afectación promedio por incendio de 109 ± 107 ha (Comisión Nacional Forestal [CONAFOR], 2019).

En años recientes, los incendios forestales muestran mayor dispersión, frecuencia, tamaño y severidad (International Panel on Climate Change [IPCC], 2007). Este aumento en la siniestralidad se atribuye tanto a la generación de mayor volumen de combustible, resultado de las acciones de supresión de incendios, como al cambio climático (Marlon et al., 2009; Westerling, Hidalgo, Cayan, & Swetnam, 2006). Ambos factores han ocasionado cambios importantes en los patrones de precipitación y sequía, y en los procesos de acumulación de combustibles, que en conjunto han modificado la dinámica de los incendios forestales (IPCC, 2007). Por lo anterior, a nivel mundial, los gobiernos gastan un presupuesto importante en labores de prevención y combate de incendios. Si bien la mayor parte del gasto se destina a labores de supresión, el diseño de mecanismos de prevención y manejo, dentro de los cuales destaca la generación de información para calificar el riesgo o anticipar la ocurrencia de un incendio, representa una proporción importante del gasto (FAO, 2007). En este contexto, la estimación de la ocurrencia de un incendio es relevante, ya que permite planear medidas para minimizar pérdidas humanas, de infraestructura y de recursos naturales (Adab, Kanniah, Solaimani, & Sallehuddin, 2015); promover actividades para mantener y restaurar los regímenes de fuego (Keeley, Fotheringham, & Morais, 1999); y optimizar la distribución de los recursos disponibles en las labores de supresión (Podschwit, Larkin, Steel, Cullen, & Alvarado, 2018).

La ocurrencia de un incendio se ha estimado a través de dos estrategias. La primera consiste en estimar indicadores de riesgo de ocurrencia con base en medidas indirectas de atributos climáticos, físicos, ambientales o de las características de los mismos incendios (*i. e.* las estadísticas de ocurrencia). Esta estrategia se deriva de la ausencia o poca calidad y cantidad de datos de las variables causales de un incendio o del conocimiento

2009; Preisler et al., 2004). The second strategy is the prediction of the occurrence or characteristics of a fire (size, intensity and duration), through models that predict the probability of observing an affected area (or fire attributes) depending on the site characteristics (v. g. slope, exposure, elevation, road distance and population density) (Adab et al., 2015; Adab, Kanniah, & Solaimani, 2013; Avila et al., 2010; Munn, Zhai, & Evans, 2003). Estimates are made through a wide variety of techniques ranging from statistical prediction models, to Bayesian networks and methods, among others (Stojanova, Panov, Kobler, Džeroski, & Taškova, 2006).

The estimation of the occurrence of a fire by one or another strategy is complex, not only due to the multitude of factors that determine its presence (Rodríguez, 2015), but also because they have their own dynamics. Therefore, forest fires have been categorized as “complex events, whose occurrence is the result of the interaction between natural processes and human factors in a social, economic and cultural context” (Vasilakos, Kalabokidis, Hatzopoulos, Kallos, & Matsinos, 2007).

The estimation of indicators of occurrence of fires or the burned area (annual burned area) in a season, based on the analysis of distribution of fire sizes, has been a recurrent practice to analyze their general behavior or to generate long-term indicators of occurrence (Malamud, Morein, & Turcotte, 1998; Torres-Rojo, Magaña-Torres, & Ramírez-Fuentes, 2007). The basic assumption is that burned surfaces follow a stationary process, so their size distribution does not change (Federal Emergency Management Agency [FEMA], 2002; Finney, 2005). These analyses have been extended to the study of large-scale fires, events that have acquired great relevance, because their frequency has increased and because they cause the greatest negative impact on forest cover (Jones et al., 2016). Studies have shown that large burned areas are a non-stationary process; that is, their frequency distribution does change, making their estimation more complicated (Holmes, Huggett, & Westerling, 2008; Liu, Stanturf, & Goodrick, 2010; Sun & Tolver, 2012).

In this context, the present work is a contribution to the study of size distribution and its application to estimate indicators of occurrence of large burned areas. The study aimed to present an indicator of the maximum area at risk of being affected by forest fires (SeR) in a territorial unit and over a season.

Materials and methods

Development of the maximum area at risk of fire indicator

The random variable “annual burned area” (y) is defined as the sum of the area affected by all fires

limitado del comportamiento del fuego en una región particular (Miller & Ager, 2013; Thompson & Calkin, 2011). Las estimaciones se hacen a través de indicadores derivados de análisis probabilísticos de la información disponible (distribuciones de tamaños) con refinamientos a través de modelos de regresión, redes neuronales, programación matemática y modelos Markovianos (Ager, Vaillant, & Finney, 2010; Miller & Ager, 2013; Podur, Martell, & Stanford, 2009; Preisler et al., 2004). La segunda estrategia es la predicción de la ocurrencia o características de un incendio (tamaño, intensidad y duración), a través de modelos que predicen la probabilidad de observar una superficie afectada (o atributos del incendio) en función de las particularidades del sitio (v. g. pendiente, exposición, elevación, distancia a caminos y densidad de población) (Adab et al., 2015; Adab, Kanniah, & Solaimani, 2013; Ávila et al., 2010; Munn, Zhai, & Evans, 2003). Las estimaciones se hacen a través de una variedad amplia de técnicas que van desde modelos estadísticos de predicción, hasta redes y métodos bayesianos, entre otras (Stojanova, Panov, Kobler, Džeroski, & Taškova, 2006).

La estimación de ocurrencia de un incendio por una u otra estrategia es compleja, no solo por la multitud de factores que determinan su presencia (Rodríguez, 2015), sino también porque estos tienen dinámica propia. Por ello, los incendios forestales han sido catalogados como “eventos complejos, cuya ocurrencia es el resultado de la interacción entre procesos naturales y factores humanos en un entorno social, económico y cultural” (Vasilakos, Kalabokidis, Hatzopoulos, Kallos, & Matsinos, 2007).

La estimación de indicadores de ocurrencia de incendios o del área siniestrada (*annual burned area*) en una temporada, con base en el análisis de distribución de tamaños de incendio, ha sido una práctica recurrente para analizar su comportamiento general o generar indicadores de ocurrencia de largo plazo (Malamud, Morein, & Turcotte, 1998; Torres-Rojo, Magaña-Torres, & Ramírez-Fuentes, 2007). El supuesto básico es que las superficies siniestradas siguen un proceso estacionario, por lo que su distribución de tamaños no cambia (Federal Emergency Management Agency [FEMA], 2002; Finney, 2005). Estos análisis se han extendido al estudio de incendios de gran magnitud, eventos que han adquirido gran relevancia, debido a que su frecuencia ha aumentado y a que causan el mayor impacto negativo en la cobertura vegetal (Jones et al., 2016). Los estudios han demostrado que las grandes superficies incendiadas constituyen un proceso no estacionario; es decir, su distribución de frecuencias si cambia, por lo que su estimación es más complicada (Holmes, Huggett, & Westerling, 2008; Liu, Stanturf, & Goodrick, 2010; Sun & Tolver, 2012).

occurring over a year in a territorial unit and, like any random variable, it has a density distribution $f(y)$ and a distribution function $F(y)$. This variable is known in the English-language literature as “annual burned area” (ABA) and should not be confused with the variable “area affected by a single fire” (SAI) in a territorial unit, which is most commonly used to model fire size distributions. The frequency distributions of the variable-SAI exhibits positive skewness, similar to that illustrated in Figure 1 for the random variable y ; however, they are characterized by having heavy-tailed distribution (Jiang, Zhuang, Flannigan, & Little, 2009; Katz, Brush & Parlange, 2005), which complicates the modeling, even with the use of extreme-value distributions (Alvarado, Sandberg, & Pickford, 1998; Jiang & Zhuang, 2011).

For the remainder of the document, the random variable y will be referred to as the area burned and is equivalent to the sum of areas affected by fires over a year in a territorial unit. It should be noted that this variable is the most used indicator to qualify the severity of a fire season, the risk of the area under timber harvesting or to associate air quality and health in the presence of fires (Wiitala & Carlton, 1994). The variable also has a positive skewness (Figure 1), but the weight of the right tail is lower, because individual fires with very large effects are a plus in the sum of affected areas in a year and do not represent a single value at the end of the distribution, as happens in individual event distributions.

En este contexto, el presente trabajo es una contribución al estudio de la distribución de tamaños y su aplicación para estimar indicadores de ocurrencia de superficies extensas afectadas por incendios. El estudio tuvo por objetivo presentar un indicador de la superficie máxima en riesgo de ser afectada por incendios forestales (SeR) en una unidad territorial y a lo largo de una temporada.

Materiales y métodos

Desarrollo del indicador de máxima superficie en riesgo de incendio

La variable aleatoria “superficie anual afectada por incendios forestales” (y) se define como la suma de la superficie afectada por todos los incendios ocurridos a lo largo de un año en una unidad territorial y, como toda variable aleatoria, posee una distribución de densidad $f(y)$ y una función de distribución $F(y)$. Dicha variable se conoce en la literatura anglosajona como “área quemada anual” (*annual burned area*) y no debe confundirse con la variable “superficie afectada por incendio” (SAI) en una unidad territorial, la cual es la más usada comúnmente para modelar distribuciones de tamaño de incendios. Las distribuciones de frecuencia de la variable SAI exhiben asimetría positiva, similar a la que se ilustra en la Figura 1 para la variable aleatoria y ; sin embargo, se caracterizan por tener una cola muy pesada (Jiang, Zhuang, Flannigan, & Little, 2009; Katz, Brush, & Parlange, 2005), que dificulta el modelaje,

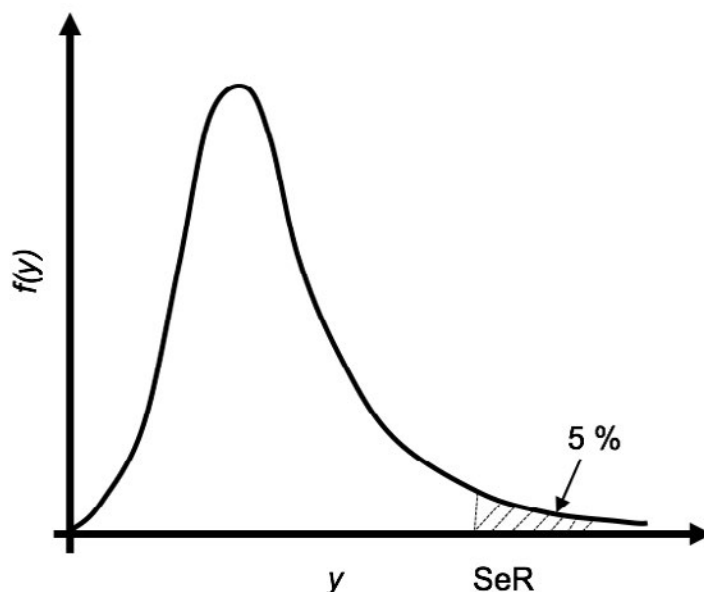


Figure 1. Random variable y (annual area affected by forest fires) with density distribution $f(y)$. The maximum area at risk of fire (SeR) with maximum probability (P_{max}) of 5 % is illustrated.

Figura 1. Variable aleatoria y (superficie anual afectada por incendios forestales) con distribución de densidad $f(y)$. Se ilustra la máxima superficie en riesgo de incendio (SeR) con probabilidad máxima (P_{max}) de 5 %.

If $f(y)$ is known, then the maximum area at risk of fire (SeR) in a territorial unit, given a level of probability P_{max} , is defined as the area corresponding to the value of y in the right tail of $f(y)$ for a level of probability P_{max} ; that is, $SeR = F^{-1}(1 - P_{max})$. SeR is an indicator analogous to Value at Risk (Jorion, 2000) popularly known as VaR and used as a measure of risk in the financial sector. SeR is illustrated in Figure 1 with $P_{max} = 5\%$. This indicator, measured in area, has the attribute of isolating very high value events (total area affected in a year) from those occurring in $1 - P_{max}$ of all occasions. This allows the estimation of the magnitude of a rare but large-scale loss without the effect of the most common events (area affected in one year).

The determination of the P_{max} value depends on the risk willing to be taken in the estimation of the maximum burned area. Financial VaR uses a probability level of 5 %, but occasionally uses 2.5 % and 1 % (Dowd, 2003). For the estimation of SeR, the value $P_{max} = 5\%$ seems appropriate, since it approximates the relative frequency of large-scale fires (Taylor et al., 2013) and can also be associated with a return period ($PR = \frac{1}{1-F(y)}$) of 20 years.

Other related indicators

Large - scale forest fires are regularly associated with climate change, increased fuel accumulation or poor incident management (Fernandes, Pacheco, Almeida, & Claro, 2016); therefore, fires follow a non-stationary pattern derived from the same characteristic of their main causes, which causes their probability distribution to change from period to period (Read & Vogel, 2016). This change in the distribution of large fires and, consequently, of the affected areas between year $t-k$ and year t , can be assessed by the difference in the value of SeR in both times. This measure will be called in the subsequent Incremental SeR (ISeR); its calculation for a period of k years, given a P_{max} ($ISeR_{k,P_{max}}$) is performed as $ISeR_{k,P_{max}} = SeR_{t,P_{max}} - SeR_{t-k,P_{max}}$, where $SeR_{t,j}$ corresponds to SeR in period i calculated with a $P_{max} = j$.

ISeR provides a metric of the magnitude and sense in which the risk of occurrence of an burned area changes over the k -year period. An increase in ISeR over a period of k years indicates that risk increases at increasing rate (nonlinear), given the nonlinearity at the tail of the frequency distribution (Dowd, 2003).

The ratio between the area affected by fires in a year and the total forest area, in each territorial unit, is a relative expression of the area affected with respect to the forest area. This proportion constitutes a new random variable x ($x = \text{total burned area/forest area}$), which also has a density function $f(x)$. Similarly, to SeR, the “maximum proportion of forest area at risk

incluso con el uso de distribuciones de valores extremos (Alvarado, Sandberg, & Pickford, 1998; Jiang & Zhuang, 2011).

En lo que resta del documento, la variable aleatoria y será referida como la superficie afectada y equivale a la suma de superficies afectadas por incendios a lo largo de un año en una unidad territorial. Cabe señalar que esta variable es el indicador más usado para calificar la severidad de una temporada de incendios, el riesgo de la superficie bajo aprovechamiento maderable o para asociar la calidad del aire y salud en presencia de incendios (Wiitala & Carlton, 1994). La variable también tiene una asimetría positiva (Figura 1), pero el peso de la cola derecha es menor, debido a que los incendios individuales con afectaciones muy grandes son un sumando más en la suma de superficies afectadas en un año y no representan un valor único en el extremo de la distribución, como sucede en distribuciones de eventos individuales.

Si $f(y)$ es conocida, entonces la máxima superficie en riesgo de incendio (SeR) en una unidad territorial, dado un nivel de probabilidad P_{max} , se define como la superficie que corresponde al valor de y en la cola derecha de $f(y)$ para un nivel de probabilidad P_{max} ; es decir, $SeR = F^{-1}(1 - P_{max})$. La SeR es un indicador análogo al valor en riesgo (Jorion, 2000) conocido popularmente como VaR y usado como medida de riesgo en el sector financiero. La SeR se ilustra en la Figura 1 con $P_{max} = 5\%$. Este indicador, medido en superficie, tiene el atributo de aislar los eventos de valores muy altos (superficie total afectada en un año) de los que ocurren en $1 - P_{max}$ de las ocasiones. Con ello se logra estimar la magnitud de una pérdida poco frecuente, pero de gran escala, sin el efecto de los eventos (superficie afectada en un año) más comunes.

La determinación del valor de P_{max} depende del riesgo que se esté dispuesto a tomar en la estimación de la superficie de máxima afectación. El VaR financiero usa un nivel de probabilidad de 5 %, aunque ocasionalmente utiliza 2.5 % y 1 % (Dowd, 2003). Para la estimación de la SeR, el valor $P_{max} = 5\%$ parece apropiado, ya que aproxima la frecuencia relativa de incendios de gran magnitud (Taylor et al., 2013) y también puede asociarse a un periodo de retorno ($PR = \frac{1}{1-F(y)}$) de 20 años.

Otros indicadores relacionados

Los incendios forestales de gran magnitud regularmente se asocian al cambio climático, mayor acumulación de combustibles o al manejo deficiente de incidentes (Fernandes, Pacheco, Almeida, & Claro, 2016); por ello, los incendios siguen un patrón no estacionario derivado de la misma característica de sus principales causales, lo cual origina que su distribución de probabilidades

of fire" (PSeR) with a confidence level P_{max} , was defined as the proportion in the right tail of $f(x)$ given a P_{max} probability level.

If $f(x)$ is known for two consecutive moments, it is possible to calculate a difference in the maximum proportion of area at risk of fire (PSeR) in a period of k years and given a P_{max} (ISeR $_{k,P_{max}}$). This difference will be referred to as Incremental PSeR (IPSeR), which is estimated as $IPSeR_{k,P_{max}} = PSeR_{i,P_{max}} - PSeR_{i-k,P_{max}}$, where $PSeR_{i,j}$ corresponds to the PSeR in the period i calculated with $P_{max} = j$.

Data

The area affected by forest fires per entity and per year was obtained from CONAFOR for the period 1970-2018. The area covered by forests varies year after year as a result of land-use change processes. The estimation of this coverage has a variable level of accuracy. Therefore, the forest area was considered constant over time and was determined based on the last available land use and vegetation cover (Series VI of the Instituto Nacional de Estadística y Geografía [INEGI], 2015). This assumption makes the form of the empirical distribution of the variable PSeR $[f(x)]$ similar to that of the variable SeR $[f(x)]$, since $X = Y/k$, where k is constant in all years.

Method

Each of the forest burned area data series, by state, was adjusted to several distribution function models to identify the best model fit. The models tested were: gama, exponential, lognormal and Weibull, selection based on the recommendations of several authors (Alvarado et al., 1998; Cui & Perera, 2008; Cumming, 2001; Jiang & Zhuang, 2011). The probability function was adjusted from the estimation of an empirical distribution $[\tilde{F}(y)]$ (Forbes, Evans, Hastings, & Peacock, 2011). This distribution was adjusted to the test model (gamma, exponential, lognormal and Weibull) using the Dmax statistic from the Kolmogorov-Smirnov test, as a criterion for minimization of deviations. The statistic is defined as $D_{max} = \sup_y |\tilde{F}(y) - \hat{F}(y)|$, where represents the distribution function predicted under the hypothetical model. This methodology ensured that estimates met the minimum Dmax criterion from the fit. The fit was made using the Risk Simulator™ application, developed for Excel.

The performance of each model was evaluated with the significance of the Dmax statistic and two qualitative criteria. The first criterion qualified the model with the best significance with a value of 1, the model with the next best significance with 2 and so on until 4, corresponding to the model with the

cambie de periodo a periodo (Read & Vogel, 2016). Este cambio en la distribución de incendios grandes y, por consecuencia, de las superficies afectadas entre el año $t-k$ y el año t , se puede evaluar por la diferencia en el valor de la SeR en ambos tiempos. A esta medida se le denominará en lo subsecuente SeR incremental (ISeR); su cálculo para un periodo de k años, dada una P_{max} (ISeR $_{k,P_{max}}$) se realiza como $ISeR_{k,P_{max}} = SeR_{t,P_{max}} - SeR_{t-k,P_{max}}$, donde $SeR_{i,j}$ corresponde a la SeR en el periodo i calculada con una $P_{max} = j$.

La ISeR brinda una métrica de la magnitud y el sentido en que el riesgo de ocurrencia de una superficie afectada cambia en el periodo de k años. Un aumento en la ISeR en un periodo de k años indica que el riesgo aumenta a tasa creciente (no lineal), dada la no linealidad de la cola de la distribución de frecuencias (Dowd, 2003).

La razón entre la superficie afectada por incendios en un año y la superficie forestal total, en cada unidad territorial, es una expresión relativa de la superficie afectada respecto a la superficie forestal. Esta proporción constituye una nueva variable aleatoria x (x = superficie total incendiada/superficie forestal), la cual también tiene una función de densidad $f(x)$. De manera análoga a la SeR, la "máxima proporción de superficie forestal con riesgo de ser afectada por incendios" (PSeR) con un nivel de confianza P_{max} , se definió como la proporción en la cola derecha de $f(x)$ dado un nivel de probabilidad P_{max} .

Si $f(x)$ es conocida para dos momentos consecutivos, es posible calcular una diferencia en la proporción máxima de superficie en riesgo de incendio (PSeR) en un periodo de k años y dada una P_{max} (ISeR $_{k,P_{max}}$). Esta diferencia se denominará PSeR incremental (IPSeR), la cual se estima como $IPSeR_{k,P_{max}} = PSeR_{i,P_{max}} - PSeR_{i-k,P_{max}}$, donde $PSeR_{i,j}$ corresponde a la PSeR en el periodo i calculada con una $P_{max} = j$.

Datos

La superficie afectada por incendios forestales por entidad y por año se obtuvo de la CONAFOR para el periodo 1970-2018. La superficie con cobertura forestal varía año con año derivado de los procesos de cambio de uso del suelo. La estimación de esta cobertura tiene un nivel variable de precisión. Por ello, la superficie forestal se consideró constante en el tiempo y se determinó con base en la última cobertura de vegetación disponible (serie VI del Instituto Nacional de Estadística y Geografía [INEGI], 2015). Este supuesto hace que la forma de la distribución empírica de la variable PSeR $[f(x)]$ sea similar a aquella de la variable SeR $[f(x)]$, dado que $X = Y/k$, donde k es constante en todos los años.

least significance. The rating of each model across all states was summed up and used as a first indicator of the most recommended model. The second criterion rated the model that provided the best fit with 1 and 0 for the others. The sum of these values, divided by the number of features, defines the proportion of times a model provided the best fit.

Once the most significant model was chosen, the SeR was calculated for P_{max} values of 1 %, 5 %, 20 % and 90 % for the series 1970-2008, 1970-2013 and 1970-2018, following the best fit model for each entity (*i. e.* a generic model was not used for all entities). These P_{max} values were chosen since they correspond to return periods of 100, 20, 5 and 1 year, respectively.

In order to identify the predictive power of the developed indicators, the prediction for year $t+1$ of a model composed by SeR and ISeR was contrasted with the prediction for the same year based on the affected area in the previous year variable (SAA), which theoretically has greater predictive power. The prediction model with SeR and ISeR is based on a linear model of two components: i) a systematic or inevitable risk estimator related to the inherent risk of the territorial unit and ii) an estimator of the trend of the area at risk, which, as noted, approximates the accumulation of fuels. The systematic risk component was defined by SeR with $P_{max}=90\%$ (SeR_{0.90}), because this area is the expected in a one-year return period, while the trend component was approached with the ISeR of the last $k=5$ años (ISeR_{5,0.90}). The prediction model has the basic form $SAA_{t+1} = \beta_0 + \beta_1 \text{SeR}_{t,0.90} + \beta_2 \text{ISeR}_{5,0.90} + e$, where the β_i correspond to model parameters and e corresponds to the error term. This model is analogous to a search algorithm in which, given a reference point, the next point is approached with a direction and a step size. The step size is defined by the value of the SeR _{P_{max}} , while the direction is approximated by the ISeR _{k,P_{max}} estimated with the same P_{max} . Variants of this model were compared with the lag prediction model, defined by $SAA_{t+1} = \beta_0 + \beta_1 SAA_t + e$. The comparison was made with the criterion of R^2 .

Results

Functional form of $f(y)$

The fits had high statistical significance for the Dmax statistic in all states, with the exception of Baja California Sur, where it was necessary to eliminate observations with zero burned area to improve the quality of the fits. The lognormal model had the best rating according to qualitative criteria, followed by the gamma, exponential and finally the Weibull models (Table 1). This shows that burned area distribution follows a lognormal function in most cases, as several

Métodos

Cada una de las series de datos de superficie afectada a causa de incendios forestales, por entidad, se ajustó a varios modelos de función de distribución para identificar el modelo de mejor ajuste. Los modelos probados fueron: gama, exponencial, lognormal y Weibull, selección realizada con base en las recomendaciones de varios autores (Alvarado et al., 1998; Cui & Perera, 2008; Cumming, 2001; Jiang & Zhuang, 2011). La función de probabilidad se ajustó a partir de la estimación de una distribución empírica $\tilde{F}(y)$ (Forbes, Evans, Hastings, & Peacock, 2011). Esta distribución se ajustó al modelo de prueba (gama, exponencial, lognormal y Weibull) con el uso del estadístico Dmax de la prueba Kolmogorov-Smirnov, como criterio de minimización de desviaciones. El estadístico se define como $D_{max} = \sup_y |\tilde{F}(y) - \hat{F}(y)|$, donde representa la función de distribución predicha bajo el modelo hipotético. Esta metodología aseguró que los estimadores cumplieran con el criterio Dmax mínimo desde el ajuste. La estimación se realizó con la aplicación Risk Simulator™, desarrollada para Excel.

El desempeño de cada modelo se evaluó con la significancia del estadístico Dmax y dos criterios cualitativos. El primer criterio calificó con un valor de 1 al modelo con mejor significancia, 2 al modelo con la siguiente mejor significancia y así sucesivamente hasta 4 que corresponde al modelo de menor significancia. La calificación de cada modelo en todas las entidades se sumó y se usó como un primer indicador del modelo más recomendable. El segundo criterio calificó con 1 al modelo que brindó el mejor ajuste y 0 para los demás. La suma de estos valores, dividido entre el número de entidades, define la proporción de ocasiones en que un modelo brindó el mejor ajuste.

Una vez elegido el modelo de mayor significancia se procedió a calcular la SeR para valores de P_{max} de 1 %, 5 %, 20 % y 90 % para las series 1970-2008, 1970-2013 y 1970-2018, siguiendo el modelo de mejor ajuste para cada entidad (*i. e.* no se usó un modelo genérico para todas las entidades). Estos valores de P_{max} se eligieron dado que corresponden a periodos de retorno de 100, 20, 5 y 1 año, respectivamente.

A fin de identificar el poder predictivo de los indicadores desarrollados, la predicción para el año $t+1$ de un modelo compuesto por la SeR y la ISeR se contrastó con la predicción para el mismo año basada en la variable superficie afectada en el año anterior (SAA), variable que en teoría tiene mayor poder predictivo. El modelo de predicción con la SeR y la ISeR se basa en un modelo lineal de dos componentes: i) un estimador de riesgo sistemático o inevitable relacionado con el riesgo inherente de la unidad territorial y ii) un estimador de la

Table 1. Rating of the distribution function models tested in the burned area data series by entity.
Cuadro 1. Calificación de los modelos de función de distribución probados en las series de datos de superficie afectada a causa de incendios forestales por entidad.

Model/Modelo	Ordered significance*/ Significancia ordenada*	Entities with better fit (%)/ Entidades con mejor ajuste (%)
Exponential/Exponencial	75	3
Gama	67	16
Lognormal	38	81
Weibull	120	0

* Lower value means greater significance.
*Valor menor significa mayor significancia.

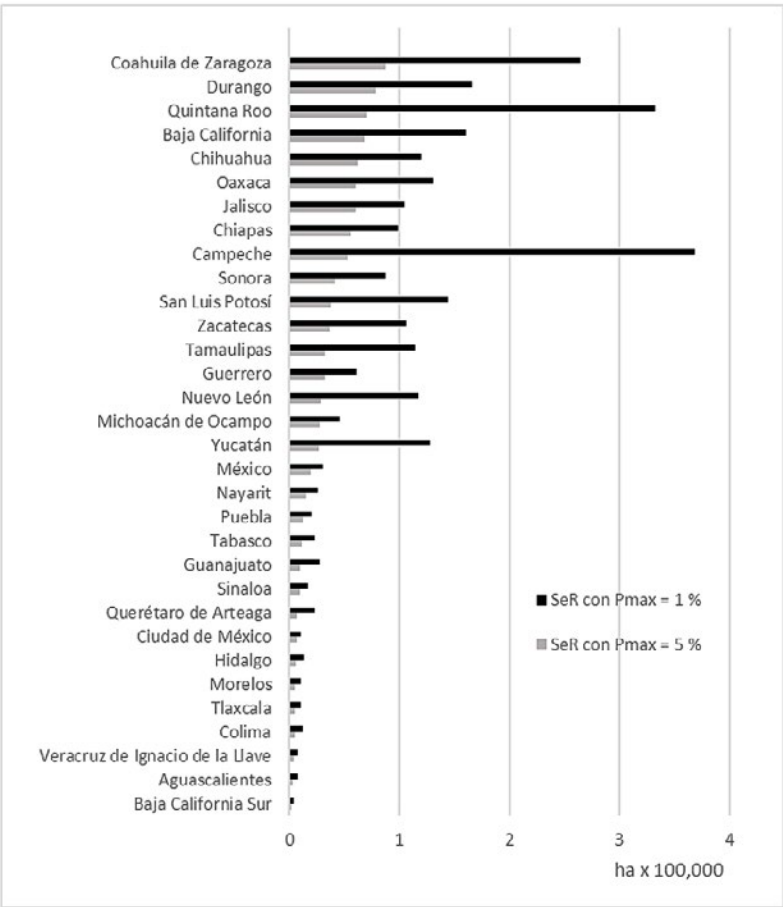


Figure 2. Maximum area at risk of fire (SeR) estimates for maximum probability values (P_{max}) at 1 and 5 %.
Figura 2. Estimaciones de superficie máxima en riesgo de incendio (SeR) para valores de probabilidad máxima (P_{max}) al 1 y 5 %.

authors have mentioned (Holmes et al., 2008; Reed, 2001; Strauss, Bednar, & Mees, 1989).

Maximum area at risk of fire (SeR)

Figure 2 shows the maximum area at risk of fire (SeR) for P_{max} values at 1 % ($SeR_{1\%}$) and 5 % ($SeR_{5\%}$). In this graph

tendencia de la superficie en riesgo, el cual como se ha señalado, aproxima la acumulación de combustibles. El componente de riesgo sistemático se definió por la SeR con $P_{max} = 90\%$ ($SeR_{0.90}$), debido a que esta superficie es la esperada en un periodo de retorno de un año, mientras que el componente de tendencia se aproximó con la ISeR de los últimos $k = 5$ años ($ISeR_{5,0.90}$). El modelo de

the states with the largest forest area (Chihuahua, Durango, Jalisco, Quintana Roo, Oaxaca, Chiapas and Guerrero) stand out, derived from their greater exposure to risk (greater forest area). Some states such as Coahuila, Sonora, Baja California, Zacatecas and Nuevo Leon showed a significantly large SeR; however, the result may be misleading since the area reported as affected is not necessarily related to the wooded area.

Figure 2 highlights the differences between $SeR_{1\%}$ and $SeR_{5\%}$ for states such as Quintana Roo, Coahuila, Sonora, Oaxaca, Campeche, Chiapas and Querétaro. The differential shows a greater occurrence of large effects in return periods of 100 years ($SeR_{1\%}$). Presumably, the differential could be associated with particular fuel accumulation conditions, climatic conditions (extreme temperatures and specific winds), fire regimes, efficiency of suppression actions or even socio-economic conditions that create an environment of increased risk for the spread of large-scale events. These differences contrast with states such as Mexico City, Mexico State or Sinaloa, where it is evident that extreme events (*i. e.* big total effects in a year) in a 100-year return period, are not proportionally larger in relation to those that could be observed in a 20-year period.

Maximum proportion of area at risk of fire (PSeR)

Unlike SeR, PSeR is comparable between territorial units, regardless of its size and amount of forest area; therefore, it is a relative index. The PSeR evaluated with $P_{max} = 5\%$ and $P_{max} = 1\%$ in each entity, for the period 1970-2018, is shown in Figure 3. In this figure, considering $P_{max} = 5\%$, the values of entities with PSeR lower than 3% were divided from those with higher proportion. The results showed some interesting patterns. The states with the largest forest area are within the group of entities with $PSeR < 3\%$; that is, the proportionally largest effects do not occur in the states with the largest forest area. In this way, states such as Coahuila, Baja California, Nuevo Leon and Tabasco, with a low proportion of forest area, have high PSeR values. Also, the case of Baja California stands out with very high values of PSeR (26%). However, this value is misleading since PSeR is relative to forest cover and much of the fires in this state are grassland or semi-arid vegetation. In a similar situation are the states of Coahuila and Nuevo Leon.

Another relevant aspect in the results is the PSeR with $P_{max} = 1\%$ ($PSeR_{1\%}$) for Quintana Roo, a state with a high incidence of large events; its $PSeR_{1\%}$ is proportionally higher than the $PSeR_{5\%}$, suggesting greater probability of observing more isolated events of large-magnitude than events of medium magnitude. The same case is observed for states such as Campeche, Tamaulipas, Zacatecas, Yucatan, Queretaro and Baja California.

predicción tiene la forma básica $SAA_{t+1} = \beta_0 + \beta_1 SeR_{t,0.90} + \beta_2 ISeR_{5,0.90} + e$, donde las β_i corresponden a parámetros del modelo y e corresponde al término de error. Este modelo es análogo a un algoritmo de búsqueda en el cual, dado un punto de referencia, se aproxima el siguiente punto con una dirección y un tamaño de paso. El tamaño de paso está definido por el valor de la $SeR_{P_{max}}$, mientras que la dirección se aproxima por la $ISeR_{k,P_{max}}$ estimada con la misma P_{max} . Variantes de este modelo se compararon con el modelo de predicción de un retraso, definido por $SAA_{t+1} = \beta_0 + \beta_1 SAA_t + e$. La comparación se realizó con el criterio de R^2 .

Resultados

Forma funcional de $f(y)$

Los ajustes tuvieron significancia alta para el estadístico Dmax en todos los estados, a excepción de Baja California Sur, entidad en la que fue necesario eliminar las observaciones con nula superficie afectada para mejorar la calidad de los ajustes. El modelo lognormal brindó la mejor calificación de acuerdo con los criterios cualitativos, seguido del modelo gama, exponencial y al final el modelo Weibull (Cuadro 1). Lo anterior muestra que la distribución de incendios sigue una función lognormal en la mayoría de los casos, tal y como lo han referido varios autores (Holmes et al., 2008; Reed, 2001; Strauss, Bednar, & Mees, 1989).

Superficie máxima en riesgo de incendio (SeR)

La Figura 2 muestra la superficie máxima en riesgo de incendio (SeR) para valores de P_{max} al 1% ($SeR_{1\%}$) y al 5% ($SeR_{5\%}$). En este gráfico sobresalen los estados de mayor superficie forestal (Chihuahua, Durango, Jalisco, Quintana Roo, Oaxaca, Chiapas y Guerrero), derivado de su mayor exposición al riesgo (mayor superficie arbolada). Algunos estados como Coahuila, Sonora, Baja California, Zacatecas y Nuevo León mostraron una SeR significativamente grande; sin embargo, el resultado puede ser engañoso dado que la superficie reportada como afectada no necesariamente está relacionada con la superficie arbolada.

La Figura 2 resalta las diferencias entre la $SeR_{1\%}$ y la $SeR_{5\%}$ para estados como Quintana Roo, Coahuila, Sonora, Oaxaca, Campeche, Chiapas y Querétaro. El diferencial muestra una mayor ocurrencia de grandes afectaciones en periodos de retorno de 100 años ($SeR_{1\%}$). Presumiblemente, el diferencial podría estar asociado a condiciones particulares de acumulación de combustibles, condiciones climáticas (temperaturas extremas y vientos específicos), regímenes de fuego, eficiencia de las acciones de supresión o incluso de condiciones socioeconómicas que crean un entorno de mayor riesgo para la propagación de eventos de gran escala. Estas diferencias contrastan con entidades

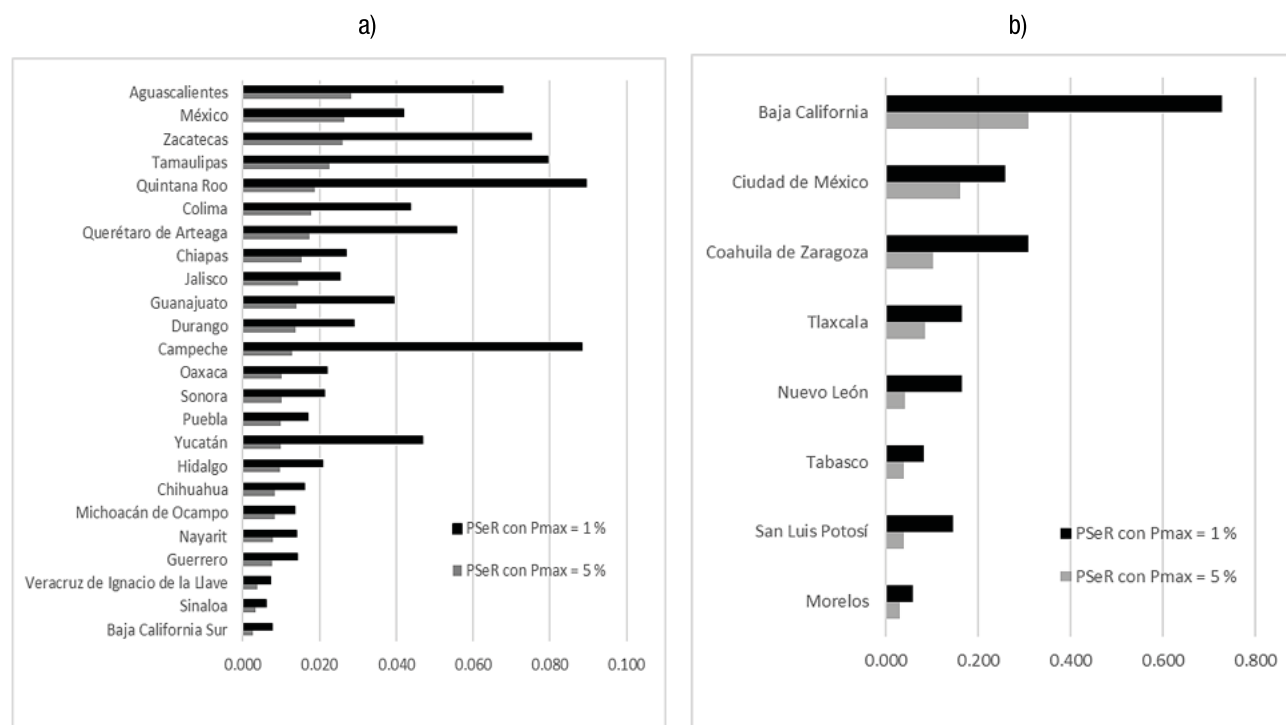


Figure 3. Estimates of the maximum proportion of area at fire risk (PSeR) for maximum probability values (P_{max}) of 1 % and 5 %. Entities with PSeR < 3 % considering $P_{max} = 5$ % (a) and PSeR ≥ 3% considering $P_{max} = 5$ % (b), in the period 1970-2018.

Figura 3. Estimaciones de la máxima proporción de la superficie en riesgo de incendio (PSeR) para valores de probabilidad máxima (P_{max}) de 1 % y 5 %. Entidades con PSeR < 3 % considerando $P_{max} = 5$ % (a) y PSeR ≥ 3 % considerando $P_{max} = 5$ % (b), en el periodo 1970-2018.

Proportion of area at incremental risk (IPSeR)

To exemplify the incremental PSeR (IPSeR) a period $k = 10$ years was used. The difference was obtained by comparing the PSeR for the period 1970-2018 with that for the period 1970-2008 (*i. e.* the evaluated period is 2009-2018). In all cases, the comparison was made with the functional form $f(x)$ that had the best fit for each state. Figure 4 shows three groups of entities: the first composed of entities where IPSeR is increasing, the second group with entities where there is practically no change, and the third group where IPSeR is decreasing for the period 2009-2018. Increases in IPSeR above 2 % are only observed in Mexico City, Baja California, Quintana Roo, and Aguascalientes. In contrast, reductions in IPSeR are only observed in Coahuila; the rest of the entities appear to have very small changes during the analyzed period.

IPSeR is an indicator of the direction in which the PSeR changes during the analyzed period and is most likely related to the amount of fuel available or the higher incidence of causative agents in a territorial unit. An example of their relationship with fuel availability are the states of Coahuila, Chiapas, Mexico and Nuevo Leon, which had significant burned areas in the previous

como la Ciudad de México, Estado de México o Sinaloa, donde es evidente que los eventos extremos (*i. e.* afectaciones totales grandes en un año) en un periodo de retorno de 100 años, no son proporcionalmente más grandes en relación con aquellos que podrían observarse en un periodo de 20 años.

Máxima proporción de la superficie en riesgo de incendio (PSeR)

A diferencia de la SeR, la PSeR es comparable entre unidades territoriales, independientemente de su tamaño y cantidad de superficie forestal; por ello es un índice relativo. La PSeR evaluada con $P_{max} = 5$ % y $P_{max} = 1$ % en cada entidad, para el periodo 1970-2018, se muestra en la Figura 3. En esta figura, considerando $P_{max} = 5$ %, los valores de las entidades con PSeR menor de 3 % se dividieron de aquellas con mayor proporción. Los resultados mostraron algunos patrones interesantes. Los estados de mayor superficie forestal se encuentran dentro del grupo de entidades con PSeR < 3 %; esto es, las afectaciones proporcionalmente mayores no ocurren en las entidades de mayor superficie forestal. De esta forma, entidades como Coahuila, Baja California, Nuevo León y Tabasco, con baja proporción de superficie forestal, presentan altos valores de PSeR.

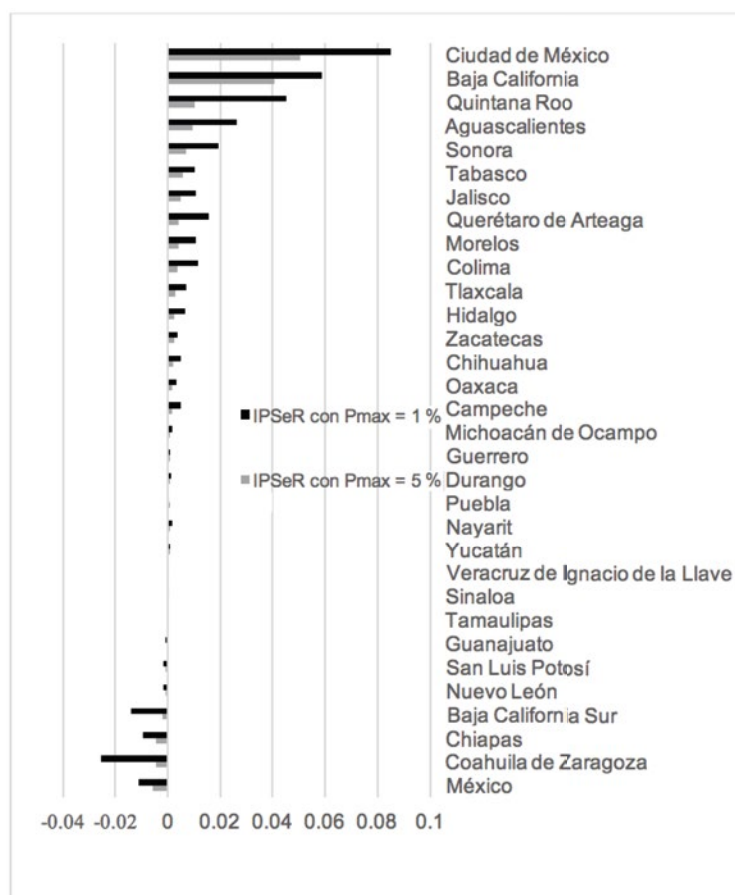


Figure 4. Values of proportion of area at incremental risk ($IPSeR_{10, 5\%}$) of fires between 2009 and 2018.

Figura 4. Valores de proporción de la superficie en riesgo incremental ($IPSeR_{10, 5\%}$) de incendios entre 2009 y 2018.

period (2009-2018) and caused a reduction in fuel volume. Such a reduction presumably decreased the post-period effects, which is reflected in an $IPSeR < 0$. On the other hand, states with positive $IPSeR$ (Mexico City, Baja California, Quintana Roo and Aguascalientes) are characterized by having presented extensive areas affected in a period greater than 10 years, before 2009, so they most likely recovered the inventory of fuels along the evaluated period (2009-2018) and show greater risk of having extensive burned areas after that year. A result that reinforces this argument is that, in these entities, the average number of fires in the analyzed period was statistically higher than the average for the whole series.

Prediction of affected areas

Table 2 shows the results of the adjustments made for the prediction of SAA in 2014, based on variants of the models indicated in the methodological section. The $ISeR$ corresponds to that between the periods 2009-2013 ($k = 5$). In all cases a logarithmic transformation was used, because the best fit model is a lognormal model. Model C provided estimators whose sum of β_i is close to the unit, which is consistent with the theoretical model

Asimismo, el caso de Baja California destaca con valores muy altos de $PSeR$ (26 %). No obstante, este valor es engañoso dado que la $PSeR$ es relativa a la cobertura forestal y gran parte de los incendios en esta entidad corresponden a pastizales o vegetación semiárida. En una situación similar se encuentran los estados de Coahuila y Nuevo León.

Otro aspecto relevante en los resultados es la $PSeR$ con $P_{max} = 1\%$ ($PSeR_{1\%}$) para Quintana Roo, entidad con alta incidencia de eventos grandes; su $PSeR_{1\%}$ es proporcionalmente más alta que la $PSeR_{5\%}$, lo que sugiere mayor probabilidad de que se observen eventos más aislados de gran magnitud que eventos de mediana magnitud. El mismo caso se observa para entidades como Campeche, Tamaulipas, Zacatecas, Yucatán, Querétaro y Baja California.

Proporción de la superficie en riesgo incremental ($IPSeR$)

Para ejemplificar la $PSeR$ incremental ($IPSeR$) se usó un periodo $k = 10$ años. La diferencia se obtuvo comparando la $PSeR$ del periodo 1970-2018 con aquella del periodo 1970-2008 (i. e. el periodo evaluado es 2009-2018). En

Table 2. Models used to predict the area affected by fires in the previous year compared to 2014.**Cuadro 2. Modelos usados para la predicción de la superficie afectada por incendios en el año anterior con respecto al 2014.**

Variable	Models / Modelos			
	A	B	C	D
Intercept / Intercepto	14.3744 ns			
ln(Affected area 2013)/ln(Superficie afectada 2013)		0.8351*		
ln(ISeR 2009-2013)	1.3237*		0.4871*	0.8189*
ln(SeR ($P_{\max} = 90\%$))	-0.3916 ns		0.4170*	1.9051*
ln[ln(SeR ($P_{\max} = 90\%$))]	1.2691 ns			-7.0846*
R ² adjusted / R ² ajustado	0.5402	0.9431	0.9349	0.9416
Observations / Observaciones	32	32	32	32

ns: no significant; *significant ($P \leq 0.05$).ns: no significativo; *significativo ($P \leq 0.05$).

$SAA_{t+1} = \beta_0 + \beta_1 \text{SeR}_{t,0.90} + \beta_2 \text{ISeR}_{5,0.90} + e$ and a β_0 restricted to zero. In contrast, model A did not restrict $\beta_0 = 0$, although the adjustment showed that this estimator is statistically equal to zero ($\beta = 0$), which confirmed the expected result for model 3. Model D included a nonlinear component for $\text{SeR}_{0.90}$ (double logarithm of $\text{SeR}_{0.90}$) and shows that the sum of β_1 moved away from the unit, but with more robust estimators. Finally, the autoregressive model (B) showed an extraordinary fit for the 2014 prediction. However, this model has an obvious restriction on predicting infrequent events, as it requires year-over-year calibration, unlike the more consistent SeR-based models. In summary, model D showed the best fit; however, model C showed the best parsimony and reliability.

Discussion

Adjustments for all distributions of annual burned area over a year, and for each of the states, showed high statistical significance ($P \leq 0.05$). Several authors have pointed out that the Pareto distribution is the one that best adjusts the size distributions of individual fires, since these are very heavy-tailed (Alvarado et al., 1998; Cui & Perera, 2008; Irland, 2013; Schoenberg, Peng, & Woods, 2003); however, the distributions of the total burned area for a year in a territorial unit did not show such heavy tail, and therefore did not require the use of extreme value models (*i. e.* Pareto, Gumbel, Fréchet or Weibull of more than two parameters) for its fit or the application of special fitting techniques (*i. e.* probabilistically weighted model). In all cases, the models tested to adjust the size distributions of the effects in each of the states showed high goodness of fit. The lognormal model stood out for being the best

todos los casos, la comparación se hizo con la forma funcional $f(x)$ que tuvo mejor ajuste para cada estado. La Figura 4 muestra tres grupos de entidades: el primero compuesto por entidades donde la IPSeR es creciente, el segundo grupo con entidades donde prácticamente no hay cambio, y el tercer grupo donde la IPSeR es decreciente para el periodo 2009-2018. Los aumentos en la IPSeR superiores a 2 % solo se observan en la Ciudad de México, Baja California, Quintana Roo y Aguascalientes. Por el contrario, reducciones en la IPSeR solo se observan en Coahuila; el resto de las entidades parece tener cambios muy pequeños durante el periodo analizado.

La IPSeR es un indicador de la dirección en que la PSeR cambia durante el periodo analizado y muy probablemente se relaciona con la cantidad de combustible disponible o la mayor incidencia de agentes causales en una unidad territorial. Un ejemplo de su relación con la disponibilidad de combustibles son los estados de Coahuila, Chiapas, México y Nuevo León, que tuvieron superficies incendiadas importantes en el periodo previo (2009-2018) y causaron una reducción en el volumen de combustibles. Tal reducción presumiblemente disminuyó las afectaciones posteriores a este periodo, lo cual se refleja en una IPSeR < 0 . Por el contrario, entidades con IPSeR positivos (Ciudad de México, Baja California, Quintana Roo y Aguascalientes) se caracterizan por haber presentado superficies extensas afectadas en un periodo mayor de 10 años, antes de 2009, por lo que muy probablemente recuperaron el inventario de combustibles en el periodo evaluado (2009-2018) y muestran mayor riesgo de tener superficies extensas afectadas después de dicho año. Un resultado que refuerza este argumento es que, en

fit in 75 % of states, in agreement with Taylor et al. (2013), who propose this model for the same variable (sum of areas burned in a season).

The SeR approximates the occurrence of very large events with a level of probability, making it an indicator of possible large-scale impact. The indicator is relevant since large events constitute less than 5 % of occurrence (Taylor et al., 2013); however, they are the ones that cause the greatest damage (Strauss et al., 1989). The SeR does not consider factors such as intensity of the event or the effects of fire on the structure or composition of the affected vegetation; variables that are more linked to short-term risk indices and of local application as defined by Adab et al. (2015), Ager et al. (2010), Keeley et al. (1999), Podur et al. (2009) y Preisler et al. (2004). Nevertheless, an estimator of the occurrence of extreme events in a territorial unit is useful for estimating a budget scenario in the presence of atypical (large) events, as well as estimating the opportunity cost of a forest area's lack of protection. As far as is known, there are no risk indicators similar to those proposed in this study that have been reported in the literature. The closest is Jiang and Zhuang (2011), who used extreme value distributions to estimate the return periods of large forest fires using individual fire data. As noted above, the return period is related to the SeR as the size distribution is estimated and the extreme values are identified. The other close estimator is that of Podur, Martell, and Knight (2002), who used the distribution of the sum of annually burned areas (annually burned area), to assess the change in fire damage in Canada; their approximation uses moments of distribution to identify differences in burned areas, but did not estimate a probability of occurrence of burned areas over large expanses.

The PSeR is a particularly useful indicator for a central planner, such as CONAFOR, as it provides information to prioritize regions or states in the face of the possibility of having an atypical event (total area burned) of great magnitude. The knowledge of the regions or entities that may have the greatest risk of presenting an event of this nature improves the planning of the needs of equipment, human and financial resources that must be invested in each entity to face the contingencies.

Several authors point out that large, atypical events are more related to the amount of fuel available than to weather conditions, combined with poor incident management (Fernandes et al., 2016; Taylor et al., 2013). Therefore, recurring recommendations to avoid large, atypical events are to reduce the amount of fuels where possible and to increase monitoring of this material, particularly in regions near the urban interface (Fernandes et al., 2016). In this context, the ISeR is a relevant indicator, since it shows the dynamics of the

estas entidades, el número promedio de incendios en el periodo analizado fue estadísticamente superior al promedio de toda la serie.

Predicción de superficies afectadas

El Cuadro 2 muestra los resultados de los ajustes realizados para la predicción de la SAA en el año 2014, a partir de variantes de los modelos señalados en la sección metodológica. La ISeR corresponde a aquella entre los periodos 2009-2013 ($k = 5$). En todos los casos se usó una transformación logarítmica, debido a que el modelo de mejor ajuste es un modelo lognormal. El modelo C brindó estimadores cuya suma de β_i es cercana a la unidad, lo cual es consistente con el modelo teórico $SAA_{t+1} = \beta_0 + \beta_1 SeR_{t,0.90} + \beta_2 ISeR_{5,0.90} + e$ y una β_0 restringida a cero. Por el contrario, en el modelo A no se restringió $\beta_0 = 0$, aunque el ajuste mostró que este estimador es estadísticamente igual a cero ($\beta = 0$), lo que confirmó el resultado esperado para el modelo 3. El modelo D incluyó un componente no lineal para la $SeR_{0.90}$ (doble logaritmo de $SeR_{0.90}$) y muestra que la suma de β_i se alejó de la unidad, pero con estimadores más robustos. Finalmente, el modelo autorregresivo (B) mostró un ajuste extraordinario para la predicción de 2014. No obstante, este modelo tiene una evidente restricción para predecir eventos poco frecuentes, ya que requiere de una calibración año con año, a diferencia de los modelos basados en SeR que son más consistentes. En resumen, el modelo D mostró el mejor ajuste; sin embargo, el modelo C presentó mejor parsimonia y fiabilidad.

Discusión

Los ajustes de todas las distribuciones de superficie anual afectada por incendios a lo largo de un año, y para cada una de las entidades federativas, mostraron alta significancia estadística ($P \leq 0.05$). Varios autores han señalado que la distribución Pareto es la que mejor ajusta las distribuciones de tamaño de incendios individuales, ya que estas tienen cola muy pesada (Alvarado et al., 1998; Cui & Perera, 2008; Irland, 2013; Schoenberg, Peng, & Woods, 2003); sin embargo, las distribuciones de la superficie total afectada por incendios durante un año en una unidad territorial no mostraron colas tan pesadas, por lo que no requirieron el uso de modelos de valores extremos (*i. e.* Pareto, Gumbel, Fréchet o Weibull de más de dos parámetros) para su ajuste o la aplicación de técnicas de ajuste especial (*i. e.* modelo ponderado probabilísticamente). En todos los casos, los modelos probados para ajustar las distribuciones de tamaño de las afectaciones en cada una de las entidades mostraron alta bondad de ajuste. El modelo lognormal destacó por ser el de mejor ajuste en 75 % de los estados, en concordancia con Taylor et al. (2013), quienes proponen dicho modelo para la

SeR over time and, thus, the probable accumulation of fuel given the greater or lesser occurrence of large, atypical events. A small (negative) ISeR is related to recent large, atypical events, which have probably consumed fuel; so a smaller occurrence of such events would be expected in the short term. The relationship seems to have theoretical bases; however, more research is needed to calibrate the fuel accumulation time with the periodicity of the calculation of ISeR, since this variable could be related to others non-observables such as extreme climatological events, increased anthropogenic activity linked to the occurrence of fires, differentiated suppression actions or poor incident management, as indicated by Fernandes et al. (2016).

Based on estimates, the states with the largest forest area have smaller PSeR but relatively larger SeR. The inverse relationship can be attributed to factors such as: 1) states with less forest area have forests more vulnerable to fire; 2) states with higher tree cover presumably have more area under timber management, which leads to the development of numerous activities for fuel control or improvement in the timeliness and effectiveness of the suppression activities being carried out. In any case, more research is needed on this subject, the result of which can be of great value in the planning of fire prevention and suppression activities.

The SeR, PSeR and IPSeR indicators are practical tools for estimating the occurrence of large fire effects. The three indicators can be recalculated year by year on the basis of new fire statistics and used in a public application to update such important variables in the planning of fire seasons as: 1) the estimate of the maximum area expected to be burned in each state or region; 2) the expected burned area at the beginning of the season and which can be used as an early warning approach and a tool for planning fire prevention and control activities; 3) the prioritization of regions according to their risk to suffer large damage, which is a tool for budgeting of central planners; and 4) the estimation of priorities for fire prevention and fire management activities, which results from the potential that IPSeR has for estimating fuel accumulation.

The analysis of frequency distributions of fire sizes is an area of study that can provide more information about the dynamics of these events and their impact on the environment and society. Topics such as the dynamics of joint distributions of surface and number of fires or the analysis of the spatial and temporal distribution of such distributions, undoubtedly, open up a wide space for research.

Conclusions

The distribution of the sum of areas affected by forest fires year on year in a state can be modeled with an

misma variable (suma de superficies incendiadas en una temporada).

La SeR aproxima la ocurrencia de eventos muy grandes con un nivel de probabilidad, por lo que es un indicador de una posible afectación de gran escala. El indicador es relevante dado que los eventos grandes constituyen menos de 5 % de ocurrencia (Taylor et al., 2013); sin embargo, son los que ocasionan daños mayores (Strauss et al., 1989). La SeR no considera factores como intensidad del evento o los efectos del incendio sobre la estructura o composición de la vegetación afectada; variables que están más ligadas a los índices de riesgo de corto plazo y de aplicación local como los definidos por Adab et al. (2015), Ager et al. (2010), Keeley et al. (1999), Podur et al. (2009) y Preisler et al. (2004). No obstante, un estimador de la ocurrencia de eventos extremos en una unidad territorial resulta útil para estimar un escenario presupuestal en presencia de eventos atípicos (grandes), así como estimar el costo de oportunidad de la falta de protección de un área forestal. Hasta donde se tiene conocimiento, no hay indicadores de riesgo similares a los propuestos en este estudio que hayan sido reportados en la literatura. El más cercano es el de Jiang y Zhuang (2011), quienes usaron distribuciones de valor extremo para estimar los periodos de retorno de grandes incendios forestales con datos de incendios individuales. Como se señaló anteriormente, el periodo de retorno se relaciona con la SeR en tanto se estima la distribución de tamaños y se identifican los valores extremos. El otro estimador cercano es el de Podur, Martell, y Knight (2002), quienes usaron la distribución de la suma de áreas incendiadas anualmente (área quemada anualmente), para evaluar el cambio en los daños derivados de incendios en Canadá; su aproximación usa momentos de la distribución para identificar diferencias en las áreas quemadas, pero no estimaron una probabilidad de ocurrencia de áreas quemadas en grandes extensiones.

La PSeR es un indicador especialmente útil para un planificador central, como podría ser la CONAFOR, ya que brinda información para priorizar regiones o estados ante la posibilidad de tener un evento (superficie total incendiada) atípico de gran magnitud. El conocimiento de las regiones o entidades que podrían tener más riesgo de presentar un evento de estas características mejora la planeación de las necesidades de equipo, recursos humanos y financieros que deberán apostarse en cada entidad para enfrentar las contingencias.

Varios autores señalan que los eventos atípicos grandes están más relacionados con la cantidad de combustible disponible que con las condiciones meteorológicas, aunado al manejo deficiente de incidentes (Fernandes et al., 2016; Taylor et al., 2013). Por ello, las recomendaciones recurrentes para evitar

asymmetric distribution. The lognormal model was the model that best adjusted such distributions. The area at risk (SeR) approximates the forest area potentially affected by forest fires with a certain level of probability. The SeR can be transformed in proportion of the area potentially affected (PSeR) by forest fires and both indicators are useful for the annual planning of fire prevention and suppression activities. In addition, the study showed that the SeR and its trend indicator (ISeR) are good predictors of the area expected to be burned in the next year. This makes SeR not only an indicator of risk, but also an approximate variable of a short-term hazard index.

Acknowledgments

To Dr. Dante Arturo Rodríguez Trejo for his valuable comments during the integration of the first drafts of this document, as well as to the Gerencia de Incendios Forestales of CONAFOR for the data provided.

End of English version

References / Referencias

- Adab, H., Kanniah, K. D., & Solaimani, K. (2013). Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. *Natural Hazards*, 65(3), 1723–1743. doi: 10.1007/s11069-012-0450-8
- Adab, H., Kanniah, K. D., Solaimani, K., & Sallehuddin, R. (2015). Modelling static fire hazard in a semi-arid region using frequency analysis. *International Journal of Wildland Fire*, 24(6), 763–777. Retrieved from <http://www.publish.csiro.au/wf/WF13113>
- Ager, A. A., Vaillant, N. M., & Finney, M. A. (2010). A comparison of landscape fuel treatment strategies to mitigate wildland fire risk in the urban interface and preserve old forest structure. *Forest Ecology and Management*, 259(8), 1556–1570. doi: 10.1016/j.foreco.2010.01.032
- Alvarado, E., Sandberg, D. V., & Pickford, S. G. (1998). Modeling large forest fires as extreme events. *Northwest Science*, 72, 66–75. Retrieved from https://www.researchgate.net/publication/312916018_Modeling_large_forest_fires_as_extreme_events
- Ávila, F. D., Pompa, G. M., Antonio, N. X., Rodríguez, T. D. A., Vargas, P. E., & Santillán, P. J. (2010). Driving factors for forest fire occurrence in Durango State of Mexico: A geospatial perspective. *Chinese Geographical Science*, 20(6), 491–497. doi: 10.1007/s11769-010-0437-x
- Comisión Nacional Forestal (CONAFOR). (2019). Incendios forestales: Serie histórica de datos. Retrieved from <https://datos.gob.mx/busca/organization/conafor>
- Cui, W., & Perera, A. H. (2008). What do we know about forest fire size distribution, and why is this knowledge
- eventos atípicos grandes son reducir la cantidad de combustibles donde sea posible y aumentar el monitoreo de este material, particularmente en las regiones cercanas a la interfaz urbana (Fernandes et al., 2016). En este contexto, la ISeR es un indicador relevante, dado que muestra la dinámica de la SeR en el tiempo y, por tanto, la probable acumulación de combustible dada la mayor o menor ocurrencia de eventos atípicos grandes. Una ISeR pequeña (negativa) está relacionada con eventos atípicos grandes recientes, que seguramente han consumido combustible; por tanto, se esperaría una menor ocurrencia de este tipo de sucesos en el corto plazo. La relación parece tener bases teóricas; sin embargo, es necesaria más investigación que ayude a calibrar el tiempo de acumulación de combustibles con la periodicidad del cálculo de la ISeR, ya que esta variable podría estar relacionada con otras no analizadas como eventos climatológicos extremos, mayor actividad antropogénica ligada a la aparición de incendios, acciones de supresión diferenciadas o un pobre manejo de incidentes, como lo indican Fernandes et al. (2016).
- Con base en las estimaciones, las entidades con mayor superficie forestal tienen PSeR más pequeñas, aunque SeR relativamente más grandes. La relación inversa puede atribuirse a factores como: 1) entidades con menos superficie forestal tienen bosques más vulnerables a la aparición de incendios; 2) entidades con mayor cobertura arbórea presumiblemente tienen más superficie bajo manejo maderable, lo que conduce al desarrollo de numerosas actividades para el control de combustibles o la mejora en la oportunidad y eficacia de las actividades de supresión que se realizan. De cualquier forma, es necesaria más investigación sobre este tema, cuyo resultado puede tener alto valor en la planeación de actividades de prevención y supresión de incendios.
- Los indicadores SeR, PSeR e IPSeR son herramientas prácticas para estimar la ocurrencia de grandes afectaciones producidas por incendios. Los tres indicadores pueden recalcularse año con año con base en las nuevas estadísticas de incendios y usarse en una aplicación de uso público para actualizar variables tan importantes en la planeación de temporadas de incendio como: 1) la estimación de la superficie máxima que se espera sea afectada en cada entidad o región; 2) la superficie que se espera sea afectada al inicio de la temporada y que puede ser usada como aproximación de alerta temprana e instrumento de planeación de actividades de prevención y control de incendios; 3) la priorización de las regiones por su riesgo a sufrir grandes afectaciones, la cual es una herramienta para la planeación presupuestal de los planeadores centrales; y 4) la estimación de prioridades para realizar actividades de prevención y manejo del

- useful for forest management? *International Journal of Wildland Fire*, 17(2), 234–244. Retrieved from <http://www.publish.csiro.au/wf/wf06145>
- Cumming, S. G. (2001). A parametric model of the fire-size distribution. *Canadian Journal of Forest Research*, 31(8), 1297–1303. doi: 10.1139/x01-032
- Dowd, K. (2003). *An introduction to market risk measurement*. West Sussex, England: John Wiley & Sons.
- Federal Emergency Management Agency (FEMA). (2002). Guidelines and specifications for flood hazard mapping partners. Appendix D: guidance for coastal flooding analyses and mapping. Retrieved from <https://www.fema.gov/es/media-library/assets/documents/13948>
- Fernandes, P. M., Pacheco, A. P., Almeida, R., & Claro, J. (2016). The role of fire-suppression force in limiting the spread of extremely large forest fires in Portugal. *European Journal of Forest Research*, 135(2), 253–262. doi: 10.1007/s10342-015-0933-8
- Finney, M. A. (2005). The challenge of quantitative risk analysis for wildland fire. *Forest Ecology and Management*, 211(1-2), 97–108. doi: 10.1016/j.foreco.2005.02.010
- Food and Agricultural Organization (FAO). (2007). Fire management global assessment 2006. A thematic study prepared in the framework of the Global Forest Resources Assessment 2005. FAO Forestry Paper 151. Retrieved from <http://www.fao.org/3/a0969e/a0969e00.htm>
- Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions*. New Jersey, Estados Unidos: John Wiley & Sons.
- Holmes, T. P., Huggett, R. J., & Westerling, A. L. (2008). Statistical analysis of large wildfires. In T. P. Holmes, J. P. Prestemon, & K. L. Abt (Eds.), *The economics of forest disturbances. Forestry sciences* (vol. 79, pp. 59–77). Dordrecht, Holanda: Springer. doi: 10.1007/978-1-4020-4370-3_4
- Instituto Nacional de Estadística y Geografía (INEGI). (2015). Carta de uso del suelo y vegetación de México. Retrieved from <http://www.beta.inegi.org.mx/app/mapas/default.html?>
- International Panel on Climate Change (IPCC). (2007). *Climate Change 2007. Impacts, adaptation, and vulnerability. Contribution of working group II to the fourth assessment report of the IPCC*. Cambridge, UK: Cambridge University Press. Retrieved from https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf
- Irland, L. C. (2013). Extreme value analysis of forest fires from New York to Nova Scotia, 1950–2010. *Forest Ecology and Management*, 294, 150–157. doi: 10.1016/j.foreco.2012.09.004
- Jiang, Y., & Zhuang, Q. (2011). Extreme value analysis of wildfires in Canadian boreal forest ecosystems. *Canadian Journal of Forest Research*, 41(9), 1836–1851. doi: 10.1139/x11-102
- Jiang, Y., Zhuang, Q., Flannigan, M. D., & Little, J. M. (2009). Characterization of wildfire regimes in Canadian fuego, la cual resulta del potencial que la IPSeR tiene para estimar acumulación de combustibles.
- El análisis de distribuciones de frecuencias de tamaños de incendios es un área de estudio que puede brindar más información sobre la dinámica de estos eventos y su impacto en el ambiente y la sociedad. Temas como la dinámica de distribuciones conjuntas de superficie y número de incendios o el análisis de la distribución espacial y temporal de tales distribuciones, sin duda, abren un espacio amplio de investigación.

Conclusiones

La distribución de la suma de superficies afectadas por incendios forestales año con año en una entidad se puede modelar con una distribución asimétrica. El modelo lognormal fue el modelo que mejor ajustó tales distribuciones. La superficie en riesgo (SeR) aproxima el área forestal potencialmente afectada por incendios forestales con un determinado nivel de probabilidad. El SeR puede transformarse en proporción de la superficie potencialmente afectada (PSeR) por incendios forestales y ambos indicadores son útiles para la planeación anual de las actividades de prevención y supresión de incendios. Adicionalmente, el estudio mostró que la SeR y su indicador de tendencia (ISeR) son buenos predictores de la superficie que se espera sea afectada al siguiente año. Esto hace que la SeR no solo sea un indicador de riesgo, sino también una variable aproximada de un índice de peligro de corto plazo.

Agradecimientos

Al Dr. Dante Arturo Rodríguez Trejo por sus valiosos comentarios durante la integración de los primeros borradores de este documento, así como a la Gerencia de Incendios Forestales de la CONAFOR por los datos proporcionados.

Fin de la versión en español

- boreal terrestrial ecosystems. *International Journal of Wildland Fire*, 18(8), 992–1002. Retrieved from <http://www.publish.csiro.au/WF/WF08096>
- Jones, G. M., Gutiérrez, R. J., Tempel, D. J., Whitmore, S. A., Berigan, W. J., & Peery, M. Z. (2016). Megafires: an emerging threat to old-forest species. *Frontiers in Ecology and the Environment*, 14(6), 300–306. doi: 10.1002/fee.1298
- Jorion, P. (2000). *Value at risk*. New York, USA: McGraw-Hill.
- Katz, R. W., Brush, G. S., & Parlange, M. B. (2005). Statistics of extremes: modeling ecological disturbances. *Ecology*, 86(5), 1124–1134. doi: 10.1890/04-0606
- Keeley, J. E., Fotheringham, C. J., & Morais, M. (1999). Reexamining fire suppression impacts on brushland fire regimes. *Science*, 284(5421), 1829–1832. doi: 10.1126/science.284.5421.1829
- Liu, Y., Stanturf, J., & Goodrick, S. (2010). Trends in global wildfire potential in a changing climate. *Forest Ecology & Management*, 259(4), 685–697. doi: 10.1016/j.foreco.2009.09.002
- Malamud, B. D., Morein, G., & Turcotte, D. L. (1998). Forest fires: An example of self-organized critical behavior. *Science*, 281(5384), 1840–1842. doi: 10.1126/science.281.5384.1840
- Marlon, J. R., Bartlein, P. J., Walsh, M. K., Harrison, S. P., Brown, K. J., Edwards, M. E., Brunelle, A., ...Whitlock, C. (2009). Wildfire responses to abrupt climate change in North America. *Proceedings of the National Academy of Sciences*, 106(8), 2519–2524. doi: 10.1073/pnas.0808212106
- Miller, C., & Ager, A. A. (2013). A review of recent advances in risk analysis for wildfire management. *International Journal of Wildland Fire*, 22(1), 1–14. doi: 10.1071/WF11114
- Moore, P. F. (2019). Global wildland fire management research needs. *Current Forestry Reports*, 5, 210–225. doi: 10.1007/s40725-019-00099-y
- Munn, I. A., Zhai, Y., & Evans, D. L. (2003). Modeling forest fire probabilities in the south central United States using FIA data. *Southern Journal of Applied Forestry*, 27(1), 11–17. doi: 10.1093/sjaf/27.1.11
- Podschwit, H., Larkin, N., Steel, E., Cullen, A., & Alvarado, E. (2018). Multi-model forecasts of very-large fire occurrences during the end of the 21st century. *Climate*, 6(4), 100. doi: 10.3390/cli6040100
- Podur, J., Martell, D. L., & Knight, K. (2002). Statistical quality control analysis of forest fire activity in Canada. *Canadian Journal of Forest Research*, 32(2), 195–205. doi: 10.1139/x01-183
- Podur, J. J., Martell, D. L., & Stanford, D. (2009). A compound Poisson model for the annual area burned by forest fires in the province of Ontario. *Environmetrics*, 21(5), 457–469. doi: 10.1002/env.996
- Preisler, H. K., Brillinger, D. R., Burgan, R. E., & Benoit, J. W. (2004). Probability based models for estimation of wildfire risk. *International Journal of Wildland Fire*, 13(2), 133–142. doi: 10.1071/WF02061
- Read, L. K., & Vogel, R. M. (2016). Hazard function theory for nonstationary natural hazards. *Natural Hazards and Earth System Sciences*, 16, 915–925. doi: 10.5194/nhess-16-915-2016
- Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters*, 74(1), 15–19. doi: 10.1016/S0165-1765(01)00524-9
- Rodríguez, T. D. A. (2015). *Incendios de vegetación: su ecología, manejo e historia* (vol. 2). Texcoco, México: Colegio de Posgraduados.
- Schoenberg, F. P., Peng, R., & Woods, J. (2003). On the distribution of wildfire sizes. *Environmetrics*, 14(3), 583–592. doi: 10.1002/env.605
- Stojanova, D., Panov, P., Kobler, A., Džeroski, S., & Taškova, K. (2006). Learning to predict forest fires with different data mining techniques. Retrieved from https://www.researchgate.net/publication/228527438_Learning_to_predict_forest_fires_with_different_data_mining_techniques
- Strauss, D., Bednar, L., & Mees, R. (1989). Do one percent of the forest fires cause ninety-nine percent of the damage? *Forest Science*, 35(2), 319–328. doi: 10.1093/forestscience/35.2.319
- Sun, C., & Tolver, B. (2012). Assessing the distribution patterns of wildfire sizes in Mississippi, USA. *International Journal of Wildland Fire*, 21(5), 510–520. doi: 10.1071/WF10107
- Taylor, S. W., Woolford, D. G., Dean, C. B., & Martell, D. L. (2013). Wildfire prediction to inform management: statistical science challenges. *Statistical Science*, 28(4), 586–615. doi: 10.1214/13-STS45
- Thompson, M. P., & Calkin, D. E. (2011). Uncertainty and risk in wildland fire management: a review. *Journal of Environmental Management*, 92(8), 1895–1909. doi: 10.1016/j.jenvman.2011.03.015
- Torres-Rojo, J. M., Magaña-Torres, O. S., & Ramírez-Fuentes, G. A. (2007). Índice de peligro de incendios forestales de largo plazo. *Agrociencia*, 41(6), 663–674. Retrieved from <http://www.scielo.org.mx/pdf/agro/v41n6/1405-3195-agro-41-06-663.pdf>
- Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., Kallos, G., & Matsinos, Y. (2007). Integrating new methods and tools in fire danger rating. *International Journal of Wildland Fire*, 16(3), 306–316. doi: 10.1071/WF05091
- Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science*, 313(5789), 940–943. doi: 10.1126/science.1128834
- Wiitala, M. R., & Carlton, D. W. (1994). Assessing long-term fire movement risk in wilderness fire management. *Proceedings of the International Conference on Fire and Forest Meteorology*, 12, 187–194.