

TRABAJO FIN DE MÁSTER

MODELADO DINÁMICO DEL COMPONENTE ANTRÓPICO EN EL PELIGRO DE INCENDIO: ANÁLISIS DE LA DIMENSIÓN TEMPORAL COMO FACTOR EXPLICATIVO EN LA CAUSALIDAD HUMANA

Autor: Martín González, Yago

Director: Dr. Rodrigues Mimbrero, Marcos

Máster Universitario en

**Tecnologías de la información geográfica para la ordenación del
territorio: sistemas de información geográfica y teledetección**

Septiembre de 2015



**Universidad
Zaragoza**

**Departamento de
Geografía y Ordenación del
Territorio**



Agradecimientos

El trabajo que aquí presento pone fin a una etapa muy importante de mi vida. Cinco años de formación dentro de un departamento lleno de profesionales magníficos y, por encima de todo, personas extraordinarias que me han ayudado y guiado en todo lo posible. Es imposible citar a todos en estas líneas, pero me gustaría recordar y agradecer el gran impacto que han tenido todos ellos en mi desarrollo personal.

En primer lugar, a Marcos Rodrigues, director de este trabajo, por toda la ayuda prestada y el conocimiento compartido conmigo, mostrándome el mapa sin darme el tesoro.

A Maite Echeverría y María Zúñiga, por su ayuda y consejos desde que comencé el grado, siendo al mismo tiempo mentoras y modelos. Sin duda son las dos personas que más me han enseñado en mi vida académica y parte esencial de los éxitos cosechados.

A mi familia y amigos, que siempre han estado ahí para apoyarme en las importantes decisiones tomadas y que hacen más fácil conseguir las metas propuestas. Mención especial a mi madre, mi heroína, sólo nosotros sabemos lo superado para llegar hasta aquí. Gracias, una vez más.

Por último a Darío Domingo, amigo y compañero de viaje desde hace ya bastantes años. Juntos hemos aprendido y mejorado, pero lo más importante es que tu amistad me ha hecho mejor persona.

Resumen:

Los modelos de peligro de ignición derivado de la acción humana se han centrado hasta la fecha fundamentalmente en la identificación de las variables explicativas, desde un punto de vista estático o estructural, sin tener en cuenta los ciclos temporales que rigen el comportamiento humano. Este trabajo analiza en detalle la influencia de la dimensión o variabilidad temporal del comportamiento humano en el peligro de incendio, dando un paso más en la precisión de las predicciones, desarrollando modelos predictivos más adecuados para la gestión y prevención. Los modelos se han creado mediante el algoritmo de Máxima Entropía (MaxEnt), utilizando distintos escenarios temporales (mensuales y estacionales) basados en la ocurrencia histórica de incendios en el periodo 2008-2011, que permiten analizar la variación temporal de los factores socio-económicos (interfases, vías de comunicación e infraestructuras y asentamientos) y ambientales (temperatura) asociados con la ocurrencia. Las variables explicativas han sido seleccionadas a partir de un análisis exploratorio previo mediante regresión logística, que determina su participación en cada escenario. Los modelos obtenidos han sido evaluados utilizando un método de validación cruzada con los datos de ocurrencia 2008-2011 y mediante la comparación de la predicción generada para el año 2012. Los resultados sugieren que la calidad de los modelos generados es suficiente con valores de AUC en torno a 0,8, corroborando la validez de éste nuevo enfoque para el análisis del peligro de ignición antrópico, y confirmando la existencia de variabilidad temporal intraanual en los factores explicativos.

Palabras clave: incendios forestales, peligro de ignición, causa humana, dimensión temporal, MaxEnt, España

Abstract

Up to know, human caused ignition risk models have essentially been focused on the identification of explanatory variables, from a static or structural point of view, regardless the time cycles that drive human behavior. This research comprehensively analyzes the influence of temporal dimension or variability of human behavior in wildfire risk, moving one step farther towards more accurate predictions by developing more suitable predictive models for management and prevention. Models have been elaborated through the Maximum Entropy algorithm (MaxEnt), using different temporal scenarios (monthly and seasonally) based on fire occurrence in the 2008-2011 period. This allows analysis of the temporal variation of socio-economic (interfaces, transportation networks, infrastructures and human settlements) and environmental (temperature) factors associated with occurrence. The explanatory variables were selected by a previous exploratory analysis consisting of logistic regression, which determines its participation in each scenarios. The obtained models have been assessed using a cross validation method with 2008-2011 occurrence data and by the comparison with the generated prediction for 2012. Results suggest that the quality of the models is enough, with AUC values around 0.8, supporting the validity of this new approach for the analysis of human caused ignition risk, and proving the existence of intra-annual temporal variability in explanatory factors.

Key words: wildfires, ignition risk, human caused, temporal dimension, MaxEnt, Spain

Contenidos

1. Introducción y antecedentes	9
1.1. Interés del tema de estudio.	9
1.2. Evaluación del riesgo de incendio y variable temporal	10
2. Área de estudio	12
3. Objetivos.....	13
4. Metodología y materiales	13
4.1. Definición de escenarios temporales	14
4.2. Obtención de información	15
4.2.1. Variable dependiente. Presencia de incendio.	15
4.2.2. Variable dependiente. Ausencia de incendio	16
4.2.3. Variables independientes.....	16
4.2.3.1. Interfases	17
4.2.3.2. Infraestructuras y vías de comunicación.....	18
4.2.3.3. Factores socioeconómicos	18
4.2.3.4. Factores ambientales	19
4.3. Espacialización de las variables dependientes e independientes	20
4.3.1. Variable dependientes	20
4.3.2. Variables independientes.....	21
4.3.2.1. Interfases	23
4.3.2.2. Infraestructuras y vías de comunicación y factores socioeconómicos.....	23
4.3.2.3. Factores ambientales	24
4.3.3. Análisis de colinealidad de las variables independientes	24
4.4. Análisis exploratorio de variables independientes. Técnicas de regresión logística.....	26
4.5. Modelado predictivo del peligro de incendio. Modelo de Máxima Entropía (MaxEnt)	26
4.6. Selección del modelo y obtención del peligro para un determinado día.	27
5. Resultados y valoración.....	27
5.1. Patrones temporales de la ocurrencia de incendios causados por el ser humano	27
5.1.1. Evolución histórica de los incendios en el área de estudio, 1988-2011.	28
5.1.2. Patrones temporales mensuales y diarios.....	29
5.2. Evaluación cuantitativa de las variables independientes en diferentes rangos temporales. Proceso de selección de variables explicativas.	31
5.3. Modelos predictivos generados con MaxEnt	39
5.3.1. Predicción e incertidumbre de los escenarios temporales	39
5.3.2. Validación de los modelos	41
5.3.3. Curvas de respuesta y contribución de variables en los modelos predictores	44
5.3.3.1. Invierno	45
5.3.3.2. Primavera	49
5.3.3.3. Verano	51
5.3.3.4. Otoño.....	55
5.4. Validación operativa. Predicción del peligro de incendio para 2012.	58

5.4.1. Caracterización del año 2012.	58
5.4.2. Predicción del peligro para 2012 y evaluación estadística de los resultados.	62
6. Conclusiones y líneas de trabajo futuras.....	66
6.1. Conclusiones	66
6.2. Potenciales líneas de desarrollo del trabajo	68
7. Bibliografía	69

Índice de figuras

<i>Figura 1. Número de incendios en la región Mediterránea. Extraído de San-Miguel y Camia, 2009.</i>	<i>9</i>
<i>Figura 2. Marco de trabajo para la evaluación del riesgo de incendio integrado (Chuvienco et al., 2012)</i>	<i>11</i>
<i>Figura 3. Localización del área de estudio.....</i>	<i>12</i>
<i>Figura 4. Esquema metodológico.</i>	<i>14</i>
<i>Figura 5. Proceso de depuración de la base de datos.</i>	<i>16</i>
<i>Figura 6. Máscaras de zona susceptible de incendio. Arriba-izq: mapa forestal, arriba-dcha.: CLC, abajo: VCF</i>	<i>20</i>
<i>Figura 7. Flujo de trabajo de obtención de interfases.....</i>	<i>23</i>
<i>Figura 8. Número total de incendios (causa humana) por año.....</i>	<i>28</i>
<i>Figura 9. Desviación porcentual del total de incendios (causa humana) respecto a la media por estaciones para el área de estudio (1988-2011).</i>	<i>29</i>
<i>Figura 10. Distribución mensual de incendios (causa antrópica) para el área de estudio en el periodo 1988-2011 y 2008-2011.</i>	<i>30</i>
<i>Figura 11. Izquierda: distribución semanal de incendios (causa antrópica) para los periodos 1988-2011 y 2008-2011. Derecha: porcentaje respecto a una distribución homogénea ideal.</i>	<i>31</i>
<i>Figura 12. Variables independientes finales.....</i>	<i>34</i>
<i>Figura 13. Representación gráfica de los resultados de regresión logística.....</i>	<i>36</i>
<i>Figura 14. Gráficos jackknife de MaxEnt.</i>	<i>38</i>
<i>Figura 15. Promedio de probabilidad de ignición de los modelos individuales de los escenarios temporales.</i>	<i>40</i>
<i>Figura 16. Incertidumbre en la predicción de los modelos de escenarios temporales.....</i>	<i>41</i>
<i>Figura 17. Medias de omisión y predicción para los distintos escenarios.</i>	<i>43</i>
<i>Figura 18. Gráficos de área bajo la curva ROC (Receiver Operating Curve o curva de funcionamiento del receptor).....</i>	<i>44</i>
<i>Figura 19. Gráficos jackknife de Inv.</i>	<i>45</i>
<i>Figura 20. Curvas de respuesta de variables predictoras de Inv.</i>	<i>46</i>
<i>Figura 21. Gráficos jackknife de Inv_FF.</i>	<i>47</i>
<i>Figura 22. Histogramas de frecuencias de variables predictoras del modelo de verano.</i>	<i>48</i>
<i>Figura 23. Curvas de respuesta de variables predictoras de Inv_FF.</i>	<i>48</i>
<i>Figura 24. Gráficos jackknife de Pri.....</i>	<i>49</i>
<i>Figura 25. Curvas de respuesta de variables predictoras de Pri.</i>	<i>50</i>
<i>Figura 26. Gráficos jackknife de Pri_FF.</i>	<i>51</i>
<i>Figura 27. Curvas de respuesta de variables predictoras de Pri_FF</i>	<i>51</i>
<i>Figura 28. Gráficos jackknife de Ver.....</i>	<i>52</i>
<i>Figura 29. Curvas de respuesta de variables predictoras de Ver.....</i>	<i>53</i>
<i>Figura 30. Gráficos jackknife de Ver_FF.....</i>	<i>54</i>

<i>Figura 31. Curvas de respuesta de variables predictoras de Ver_FF.</i>	<i>55</i>
<i>Figura 32. Gráficos jackknife de Oto.....</i>	<i>56</i>
<i>Figura 33. Curvas de respuesta de variables predictoras de Oto.....</i>	<i>57</i>
<i>Figura 34. Gráficos jackknife de Oto_FF..</i>	<i>57</i>
<i>Figura 35. Curvas de respuesta de variables predictoras de Oto_FF.</i>	<i>58</i>
<i>Figura 36. Carácter de la temperatura en 2012. Fuente: AEMET (2013).....</i>	<i>59</i>
<i>Figura 37. Número total de incendios por causa humana en 2012 por mes.</i>	<i>60</i>
<i>Figura 38. Izquierda: distribución semanal de incendios (causa antrópica) en 2012.</i>	
<i>Derecha: porcentaje respecto a una distribución homogénea ideal de 2012.</i>	<i>61</i>
<i>Figura 39. Distribución espacial de los incendios de 2012 y 2008-2011.</i>	<i>61</i>
<i>Figura 40. Comportamiento térmico de agosto y septiembre de 2012. Fuente: AEMET</i>	
<i>(2012a, 2012b)</i>	<i>62</i>
<i>Figura 41. Mapas de peligro e incendios para días concretos de 2012. Los círculos</i>	
<i>amarillos representan la ocurrencia de un incendio.</i>	<i>63</i>
<i>Figura 42. Boxplot de peligro de incendio en cada localización de incendio forestal</i>	
<i>en 2012 por meses.....</i>	<i>64</i>
<i>Figura 43. Boxplot de peligro de incendio en cada localización de incendio forestal</i>	
<i>en 2012 por estación.....</i>	<i>65</i>
<i>Figura 44. Boxplot de peligro de incendio en cada localización de incendio forestal</i>	
<i>en 2012 por causas.</i>	<i>66</i>

Índice de tablas

<i>Tabla 1. Escenarios temporales. Aclaración: _FF (Fin de semana y/o Festivo)</i>	<i>15</i>
<i>Tabla 2. Variables independientes seleccionadas.</i>	<i>22</i>
<i>Tabla 3. Resultados del análisis de colinealidad. Índice de correlación de la Rho de Spearman.</i>	<i>25</i>
<i>Tabla 4. Análisis exploratorio de regresión logística 1. En rojo relación positiva, en azul relación negativa.</i>	<i>32</i>
<i>Tabla 5. Análisis exploratorio de regresión lineal 2. En azul: variables con sentido explicativo negativo, en rojo: sentido explicativo positivo.</i>	<i>35</i>
<i>Tabla 6. Porcentaje de contribución y de permutación de variables independientes producido por MaxEnt</i>	<i>37</i>
<i>Tabla 7. Selección final de variables representativas.</i>	<i>38</i>
<i>Tabla 8. Valores de AUC de los modelos predictivos.</i>	<i>42</i>
<i>Tabla 9. Porcentajes de contribución y permutación de variables predictoras de Inv.</i>	<i>45</i>
<i>Tabla 10. Porcentajes de contribución y permutación de variables predictoras de Inv_FF.</i>	<i>47</i>
<i>Tabla 11. Porcentajes de contribución y permutación de variables predictoras de Pri.</i>	<i>49</i>
<i>Tabla 12. Porcentajes de contribución y permutación de variables predictoras de Pri_FF.</i>	<i>51</i>
<i>Tabla 13. Porcentajes de contribución y permutación de variables predictoras de Ver.</i>	<i>52</i>
<i>Tabla 14. Porcentajes de contribución y permutación de variables predictoras de Ver_FF.</i>	<i>54</i>
<i>Tabla 15. Porcentajes de contribución y permutación de variables predictoras de Oto.</i>	<i>56</i>
<i>Tabla 16. Porcentajes de contribución y permutación de variables predictoras de Oto_FF.</i>	<i>58</i>

1. Introducción y antecedentes

Los incendios forestales se constituyen como uno de los riesgos naturales con mayor potencial devastador, provocando grandes pérdidas anualmente en todo el mundo.

1.1. Interés del tema de estudio.

La intensa relación entre el medio mediterráneo, del cual la Península Ibérica forma parte, y el fuego es intrínseca (Le Houerou, 1973, 1977; Naveh, 1975). Las características climáticas de este espacio condicionan y favorecen la aparición ocasional de incendios forestales de origen natural (Hötl, 2008), que en condiciones ideales mantendrían un equilibrio con el medio, contribuyendo a la propia "salud" de los ecosistemas. Sin embargo, la actividad del hombre ha modificado este equilibrio favoreciendo la aparición de incendios, siendo responsable de más del 90% de los fuegos (FAO, 2007), más de un 95% en España de acuerdo a las estadísticas nacionales (Martínez *et al.* 2009).

Las consecuencias de los incendios forestales han sido muy estudiadas, destacando la evidente afección sobre las especies florísticas y sobre la fauna que habita en el medio natural, con la posible reducción de biodiversidad en el ecosistema, la modificación de las características físico-químicas y biológicas del suelo, y con ello su respuesta erosiva (Wagenbrenner *et al.*, 2006; Ben-Hur *et al.*, 2008) o, a una mayor escala, la emisión de gases de efecto invernadero (van der Werf *et al.*, 2010). La pérdida de propiedades e infraestructuras e incluso de vidas humanas es otro de los efectos adversos de este tipo de fenómenos.

Diversos autores han demostrado que la frecuencia de incendios así como la superficie quemada, se ha incrementado en la región mediterránea desde 1960 (Kliot, 1996; Pausas and Vallejo, 1999; Wittenberg and Malkinson, 2009, San-Miguel y Camia, 2009). Esto se puede achacar a la acumulación de combustible en áreas rurales abandonadas (Shakesby, 2011; Pausas and Fernández-Muñoz, 2012), a la silvicultura monoespecífica de especies inflamables (Shakesby, 2011) y al cambio climático (Pausas, 2004; Pausas and Fernández-Muñoz, 2012). En este contexto, España se erige como uno de los grandes afectados dentro del continente europeo, habiendo registrado en el periodo 1988-2011 cerca de 400.000 incendios. Especialmente dañina es la concentración de los incendios en tiempo y espacio, algo que resulta frecuente tanto en nuestro país como en nuestros vecinos europeos, dando lugar a temporadas de incendios especialmente cruentas.

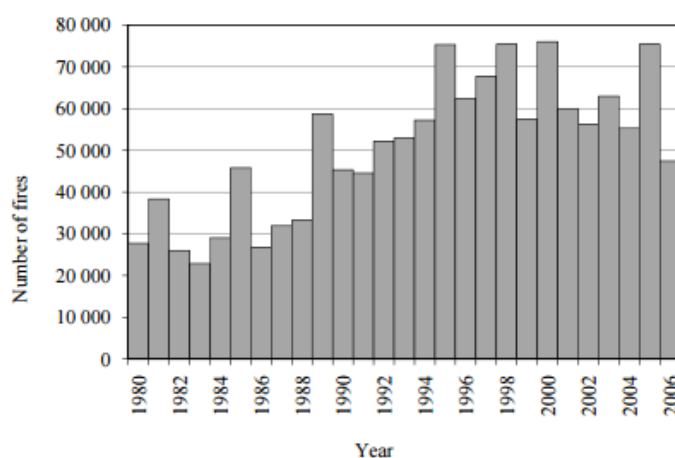


Figura 1. Número de incendios en la región Mediterránea. Extraído de San-Miguel y Camia, 2009.

Esta yuxtaposición de pérdidas ecológicas, económicas y humanas junto con el aumento de la frecuencia e intensidad de estos fenómenos y ante las expectativas climáticas poco halagüeñas para las próximas décadas hacen que investigaciones sobre las causas y consecuencias de los incendios forestales sean todavía muy necesarias para tratar de minimizar los efectos negativos de estos eventos. Las autoridades competentes deben basar sus decisiones en estudios sólidos y detallados, lo que justifica la inversión en el mayor conocimiento del origen y la dinámica de los incendios forestales.

El trabajo de numerosos equipos de investigación en todo el mundo, fundamentalmente a partir de la década de los 60, ha conseguido arrojar luz al conocimiento de sus causas, consecuencias, dinámicas y patrones y es en este contexto en el que se enmarca el presente Trabajo Fin de Máster. Se trata de continuar una línea de trabajo ya iniciada, incluyendo un enfoque poco desarrollado anteriormente, la variación espacio-temporal de la causalidad humana de incendios.

1.2. Evaluación del riesgo de incendio y variable temporal

La terminología en los riesgos naturales no es siempre clara y globalmente aceptada. En el caso del riesgo de incendios forestales lo es todavía menos, habiéndose empleado diferentes términos (riesgo, peligro de incendio, probabilidad de incendio etc.) para expresar el mismo concepto o variaciones muy ligeras del mismo.

Con objeto de encuadrar este estudio dentro del amplio espectro que engloba el riesgo de incendio se va a emplear el esquema de riesgo integrado de incendio disponible en Chuvieco *et al.* (2012) y recogido en la figura 2. En Chuvieco *et al.* (2012) se entiende el riesgo de incendio como un producto de la probabilidad de ocurrencia de un fuego, de su propagación y de los daños potenciales que éste pueda ocasionar. Dentro de este marco conceptual, el presente trabajo se encuadra dentro de la evaluación del peligro de incendio, concretamente en la probabilidad de ignición derivada de la acción humana, aunque utiliza una variable relacionada con la humedad del combustible por su elevada relación con las posibilidades de ignición y su potencial relación con la ocurrencia derivada de negligencias o accidentes.

En este sentido, han sido numerosas las investigaciones que han tratado de identificar los factores humanos responsables del origen de los incendios además de sus distribución espacial (Kalabokidis *et al.*, 2002; Martínez *et al.*, 2004 y 2009; Syphard *et al.* 2007; Rodrigues y de la Riva, 2014b; Rodrigues *et al.*, 2014). Estos estudios han obtenido resultados positivos, logrando identificar las variables más influyentes y así consiguiendo modelizar el peligro de ignición por causa humana. Sin embargo, hasta ahora, en la mayor parte de los casos, se ha analizado la causalidad humana de forma estática, invariable o estructural, sin discriminar que la actividad del ser humano está sujeta a ciclos temporales bien conocidos que pueden tener influencia en el peligro de ignición. Parametrizar el comportamiento humano en tiempo y espacio resulta complejo, fundamentalmente por la no disponibilidad o ausencia de datos (Martell *et al.*, 1987) desagregados temporalmente. Por este motivo, el presente estudio se apoya en la dimensión temporal de los incendios para derivar el comportamiento humano que los origina, es decir, puede no disponerse de información espacializada de, por ejemplo, el número de personas que utilizan los caminos forestales cada día, pero si se conoce la localización espacio-temporal precisa del inicio de un incendio y la distancia del mismo a un camino forestal.

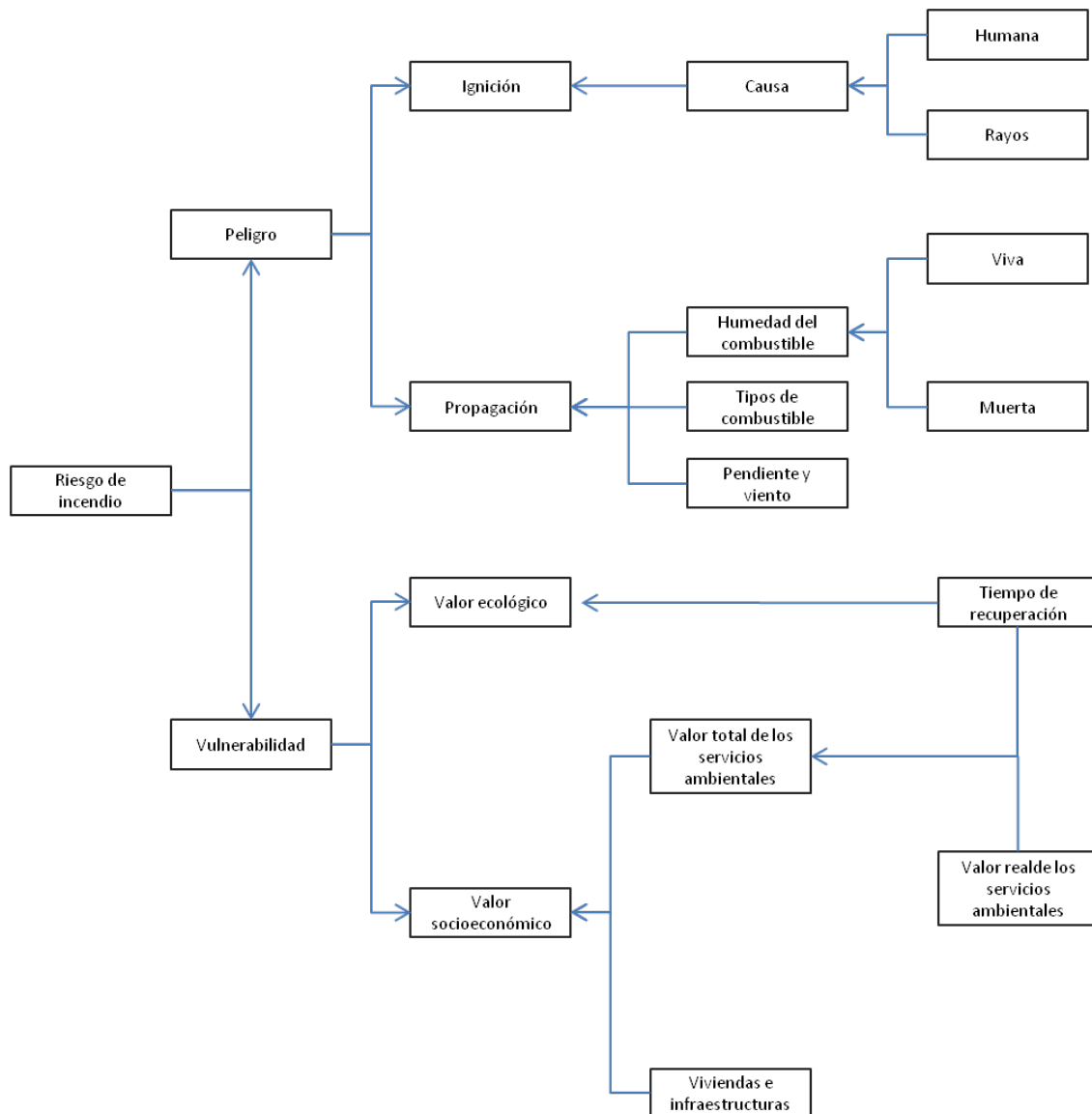


Figura 2. Marco de trabajo para la evaluación del riesgo de incendio integrado (Chuvieco *et al.*, 2012)

El comportamiento humano en relación a factores relacionados con el peligro de ignición puede ser, en ocasiones, fácilmente predecible, como son las actividades de recreo o vacacionales o los trabajos agrarios. Sin embargo, en otras, la complejidad es muy elevada, como es el caso de los incendios intencionados o cierta clase de accidentes. Este es un hecho que todo intento de modelización del peligro de ignición por causa humana debe asumir y que puede dificultar la investigación. Hasta el momento, son pocos los autores que han tratado de modelar el riesgo de ignición por causa humana desde una perspectiva dinámica, muy probablemente condicionados por los aspectos comentados anteriormente. En algunos casos si se ha tratado de ajustar un modelo que diera respuesta a la ocurrencia diaria de incendios por causa humana (Vega-García *et al*, 1995), aunque sus objetivos en aquel momento fueron menos ambiciosos y se centraron en pronosticar la aparición o no de un incendio en una zona determinada mediante un modelo de regresión logística (*logit*) alimentado por cuatro variables independientes. De este modo consiguieron predecir el 79% de los días en los que se originó un incendio y un 81% de los que no se inició un incendio por causa humana. En la actualidad, la disponibilidad de productos de alta resolución temporal (como MODIS) y la utilización masiva del posicionamiento por GPS facilita la investigación de patrones temporales.

Esta investigación destaca por tanto por la innovación que supone la introducción de modelos temporales de probabilidad de ignición por causa humana, siendo ésta su principal contribución y que puede además abrir una posible línea de investigación.

2. Área de estudio

El área de estudio seleccionada ha pretendido abarcar un área suficientemente extensa para tener una muestra de incendios relevante pero sin demorar en exceso los cálculos necesarios de los algoritmos empleados. La elección de las provincias de Huesca, Zaragoza, Lérida, Barcelona y Tarragona responde al intento de aunar distintas realidades dentro de la zona a estudiar, desde el carácter más rural que puede predominar en la provincia oscense hasta la gran actividad urbano-industrial del litoral barcelonés, desde las grandes cumbres pirenaicas hasta las costas mediterráneas.



Figura 3. Localización del área de estudio

Esta variabilidad de ecosistemas que se puede encontrar dentro del área de estudio se cree que puede ser una sub-muestra representativa para conocer el comportamiento de la causalidad humana en los incendios forestales. Este espacio responde por tanto a la gran diversidad biofísica de la Península Ibérica, albergando climas con patrones semi-oceánicos, de alta montaña, esteparios/semi desérticos y propiamente mediterráneos. Además, como ya se ha comentado, es buen representante de la diversidad de condiciones socioeconómicas y de los diferentes sistemas de producción y estructuras de población y del territorio. El área de estudio tiene un total de 59081 km², lo que representa un 11,7% del total del territorio nacional. Una particularidad del área de estudio es que comparte dos husos (30 y 31) en la proyección UTM, lo que implica forzar las coordenadas a un único huso para poder representar de forma continua el territorio en una aplicación informática. En este caso se ha forzado las coordenadas del huso 31 (Cataluña y la parte más oriental de Aragón) al huso 30.

3. Objetivos

El principal objetivo de este trabajo es evaluar la influencia de la dimensión o variabilidad temporal del comportamiento humano en el riesgo de ignición. Con ello se pretende poder dar un paso más en las predicciones de peligro de incendio, superando el estatismo de los modelos de peligro humanos disponibles y adaptándose, en la medida de lo posible, a la compleja realidad de la actividad humana, facilitando así la prevención y gestión del riesgo a las administraciones competentes. Este gran objetivo se ha dividido en un conjunto de subobjetivos de a cara a la consecución del mismo:

- Describir los patrones temporales de la ocurrencia de incendios causados por el ser humano.
- Analizar qué factores humanos tienen mayor peso explicativo dentro de la zona y periodo de estudio.
- Estimar la variabilidad temporal de los factores explicativos identificados, generando modelos predictivos para los distintos escenarios temporales.
- Prever el peligro por factor humano para cualquier día del año en base a unos modelos característicos predefinidos.

La hipótesis de partida es que la dimensión temporal de las actividades antrópicas tiene gran importancia en el peligro de incendio por causas humanas, por lo que se espera obtener diferencias significativas en los modelos generados para distintos escenarios temporales. Con ello se pretende no sólo precisar el "tiempo" sino también el "lugar", ya que se espera que las distintas variables explicativas actúen de forma heterogénea a lo largo del año, produciendo diferencias espaciales en el nivel de probabilidad de ignición.

4. Metodología y materiales

Este apartado describe detalladamente la metodología empleada en este estudio, describiendo el proceso de obtención tanto de las variables dependientes como independientes, así como la determinación de los escenarios temporales. La figura 4 ilustra el proceso seguido hasta la obtención de los distintos modelos temporales.

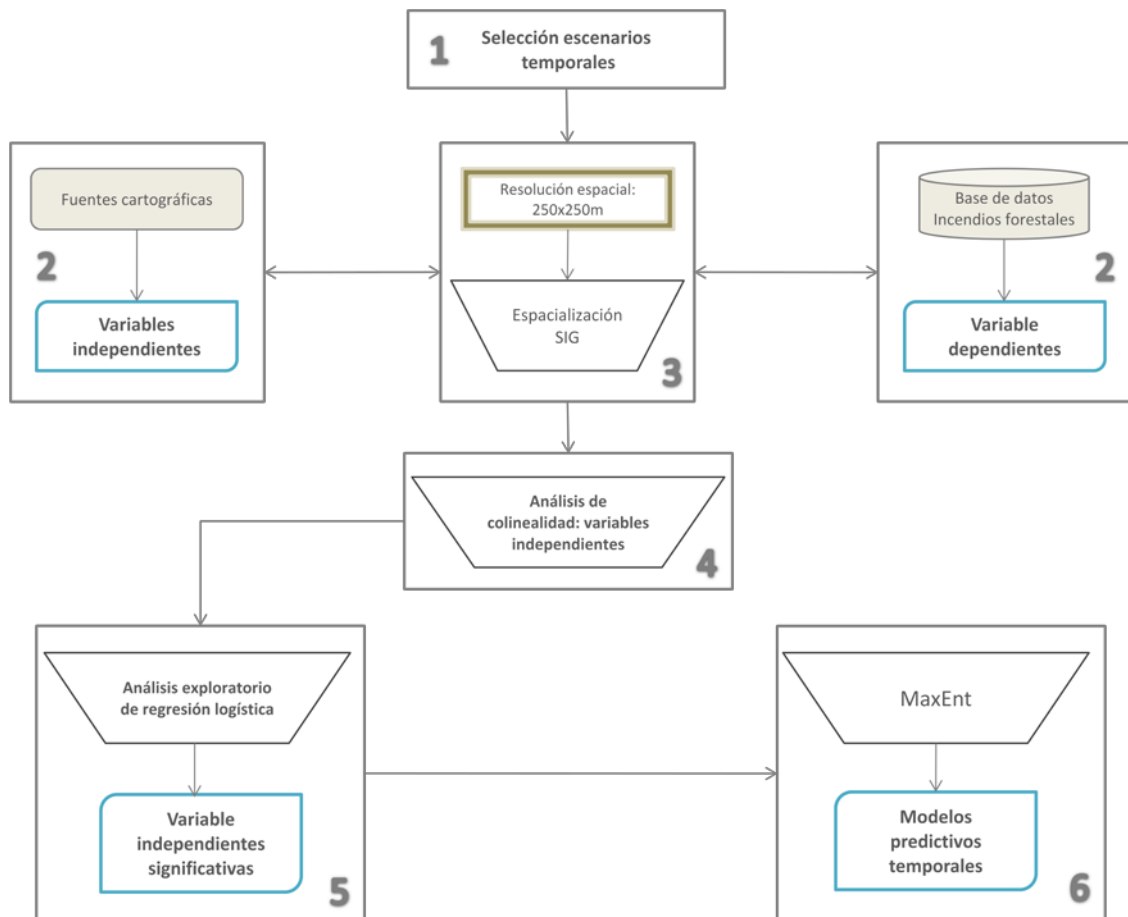


Figura 4. Esquema metodológico.

1. Definición de escenarios temporales
2. Obtención de información
3. Espacialización de variables dependientes e independientes
4. Análisis de colinealidad de las variables independientes
5. Análisis exploratorio y selección de las variables independientes mediante regresión logística
6. Ajuste y validación de modelos predictivos por medio del modelo de máxima entropía (MaxEnt).

4.1. Definición de escenarios temporales

En un estadio previo al análisis exploratorio que posteriormente se detallará se seleccionan cuatro meses representativos de cada estación (enero, abril, agosto y noviembre) además de las cuatro estaciones (invierno, primavera, verano y otoño) con el fin de poder apreciar el grado de desagregación temporal que puede abstraerse de la serie de datos de ocurrencia y por tanto ser empleado para la modelización.

En base a la hipótesis de partida -el peligro de incendio por origen humano varía en función de ciclos temporales a los que está sujeta la actividad humana- se decide separar para cada una de las submuestras anteriores aquellos incendios ocurridos en días festivos, fines de semana y días "puente" entre un festivo y un sábado o domingo. De este modo se obtienen dieciséis escenarios temporales (tabla 1). La definición de estos escenarios temporales se ha llevado a cabo ante la imposibilidad de obtener información sobre las actividades humanas desagregada temporalmente (intranualmente). Por ello, en lugar de incluir variables socioeconómicas dinámicas

se ha optado por compartimentar temporalmente la ocurrencia de incendios y analizar diferencias en la participación y capacidad explicativa de dichas variables.

Tabla 1. Escenarios temporales. Aclaración: *_FF* (Fin de semana y/o Festivo)

Escenarios temporales	
Enero (<i>Ene</i>)	Enero_FF (<i>Ene_FF</i>)
Abril (<i>Abr</i>)	Abril_FF (<i>Abr_FF</i>)
Agosto (<i>Ago</i>)	Agosto_FF (<i>Ago_FF</i>)
Noviembre (<i>Nov</i>)	Noviembre_FF (<i>Nov_FF</i>)
Invierno (<i>Inv</i>)	Invierno_FF (<i>Inv_FF</i>)
Primavera (<i>Pri</i>)	Primavera_FF (<i>Pri_FF</i>)
Verano (<i>Ver</i>)	Verano_FF (<i>Ver_FF</i>)
Otoño (<i>Oto</i>)	Otoño_FF (<i>Oto_FF</i>)

En azul se muestra la agregación temporal mensual y en verde la estacional. El análisis exploratorio mediante regresión logística (4.4) servirá también para evaluar la idoneidad de los escenarios temporales.

4.2. Obtención de información

En este apartado se detalla el proceso de obtención de la información necesaria en la generación de la variable dependiente y de aquellas que actuarán como predictores -variables independientes-.

4.2.1. Variable dependiente. Presencia de incendio.

En este trabajo se ha utilizado la base de datos de Estadística General de Incendios Forestales (EGIF) compilada por el Ministerio de Alimentación, Agricultura y Medio Ambiente del Gobierno de España como resultado de la compilación de los reportes realizados por las Comunidades Autónomas (Moreno *et al.*, 2011). Esta serie comienza en 1968, siendo una de las más antiguas recopiladas a nivel europeo (Vélez, 2001), aunque los datos no se consideran totalmente fiables hasta 1988 (Martínez *et al.* 2009). En el presente estudio se requieren las localizaciones exactas (pares de coordenadas) de cada incendio registrado, por lo que se escoge el periodo 2008-2011 -ambos inclusive- para el desarrollo de los modelos y el 2012 como validación operativa. Esta selección de años recientes dentro de la serie histórica garantiza la precisión espacial en la localización de los incendios.

Para el estudio, de los múltiples campos que contiene esta base de datos, se ha seleccionado como información relevante la que se describe a continuación:

- **Huso.** Es necesaria dado que la zona occidental del área de estudio se sitúa en el huso 30 mientras que la oriental se encuentra en el huso 31 del sistema de coordenadas *European Terrestrial Reference System 1989*.
- **Coordenadas.** Los incendios utilizados (serie 2008-2011) se encuentran geolocalizados mediante un par de coordenadas recogidas mediante GPS. Este par de coordenadas se hace indispensable para la creación de las nubes de puntos y poder extraer los datos de las variables independientes.
- **Grupo de causa.** Bajo esta denominación se recoge la causa originaria del fuego. La Unión Europea recomienda cuatro categorías (1-Desconocido, 2-Causas naturales, 3. Causa accidental o negligencia y 4-Deliberado). Para este trabajo se ha empleado todo aquel incendio que no haya tenido un

origen natural, ya que de aquellos de origen desconocido se asume que la mayor parte es consecuencia de la actividad humana.

- **Fecha de detección.** A partir de la fecha de inicio de fuego se podrá determinar la influencia que esta tiene sobre el peligro total de incendio.

Los datos extraídos de la base de datos compondrán pues la variable dependiente como presencia de incendio, siendo sometidos previamente a un proceso de depuración que descarte errores espaciales significativos.

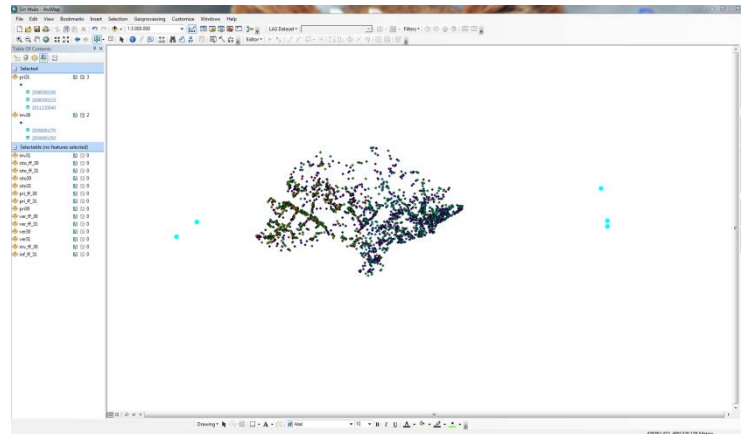


Figura 5. Proceso de depuración de la base de datos.

Éste es un proceso esencial para asegurar la calidad de los resultados obtenidos, por lo que resulta importante prestar atención a este procedimiento.

4.2.2. Variable dependiente. Ausencia de incendio

Como resulta lógico, la base de datos general de incendios únicamente recoge aquellas localizaciones en las que efectivamente se ha producido un fuego. Sin embargo, en los modelos del peligro de incendio por regresión logística se requiere información sobre la ausencia del fenómeno analizado, en este caso, una muestra de localizaciones en las que no se ha producido un incendio (Chuvieco et al., 2010; Padilla et al., 2011), por lo que la generación de estos se hace indispensable para llevar a cabo el análisis exploratorio de las variables explicativas mediante regresión logística, si bien esta muestra no será necesaria para el desarrollo de los modelos predictivos con MaxEnt.

Se asume, por tanto, que una muestra aleatoria para el área de estudio y el periodo de tiempo analizado puede actuar como información de ausencia del fenómeno. Esta ausencia es únicamente válida para el periodo estudiado, no excluye la posibilidad de que en esa localización se haya producido un incendio en el pasado o lo pueda hacer en un futuro.

Estas muestras de ausencia de fuego deben generarse para cada uno de los escenarios temporales propuestos anteriormente, y deben estar ceñidas a localizaciones en las que exista la posibilidad de originarse un incendio forestal (excluyendo zonas antrópicas, láminas de agua, suelos desnudos etc.).

4.2.3. Variables independientes

En este apartado se describen las decisiones y procesos empleados para la evaluación y selección de las variables independientes que formarán parte de los modelos finales.

Para la selección de las variables explicativas se ha acudido a estudios previos que han trabajado en la modelización del peligro y la ocurrencia con origen antrópico

en los últimos años en España, como es el caso de Chuvieco *et al.* (2010), Martínez *et al.* (2009) o Rodríguez *et al.* (2014). Dada la precisión espacial que permite la localización de los incendios mediante GPS se ha decidido excluir variables de origen estadístico que mermarían esta resolución espacial y que sí han sido empleadas en los estudios citados.

La resolución espacial se ha establecido en 250x250m por un criterio de velocidad y volumen de cálculo en el proceso de creación de los modelos, pudiendo haberse trabajado con una resolución inferior en otras condiciones tecnológicas y temporales.

4.2.3.1. Interfases

Las interfases o áreas de vecindad entre vegetación natural y actividades humanas, pueden ser buenos factores explicativos del peligro de incendio (Martínez *et al.*, 2009 y Chuvieco *et al.* 2012). Por este motivo se han seleccionado siete posibles interfases, cinco de ellas de forma desagregada, aislando diferentes actividades humanas o infraestructuras, y dos de forma conjunta, agregando actividades

En un primer momento se evaluaron las distintas fuentes cartográficas ya existentes que permitieran conocer los distintos usos de suelo. El mapa forestal de España permitía trabajar con algunas de estas variables, pero los distintos criterios en la elaboración de la cartografía entre las distintas autonomías impedían la obtención de todas las interfases (en concreto de la que incluía los prados y pastizales, categoría no existente en Cataluña). Debido a ello se decide emplear las delimitaciones de usos de suelo del *Corine Land Cover* de 2006.

Se trabaja con distancia a la interfase con objeto de aprovechar al máximo la precisión espacial de la localización de la ocurrencia, descartando por este hecho la utilización de buffers (y la decisión arbitraria de su radio) si empleados en otros estudios (Chuvieco *et al.*, 2010 y Rodríguez *et al.* 2014).

- **Interfase entre vegetación natural y cultivos (VCI).** Las actividades agrarias son responsables del origen de numerosos incendios forestales a lo largo del año, generalmente por negligencia/imprudencia en el desarrollo de esta labor. Por tanto, se espera una relación positiva entre esta variable y el peligro de ignición.

- **Interfase entre vegetación natural y pastizales (VPI).** Las zonas de pastizal también pueden tener un uso que los hace espacios sensibles a la aparición de un incendio.

- **Interfase entre vegetación natural y actividades industriales (VHI).** La manipulación de sustancias inflamables, maquinaria, vehículos etc. es un factor de riesgo en aquellas zonas en las que existe contacto con vegetación natural. Es por ello esperable una relación positiva con el peligro de incendio.

- **Interfase entre vegetación natural e infraestructuras de riesgo (VRI).** Dentro de este grupo de infraestructuras de riesgo se incluyen los vertederos y escombreras, las canteras y minas y espacios en construcción. Sus actividades también pueden suponer un riesgo en la aparición de un incendio forestal.

- **Interfase entre vegetación natural y espacios urbanos (VUI).** La concentración de un foco de población en general supone un foco de origen de incendios, ya que un 90% de los fuegos tienen origen humano.

- **Interfase entre vegetación natural-pastizales y espacios urbanos-actividades industriales-infraestructuras de riesgo (WUI).** En esta interfase se agrupa toda vegetación natural además de los pastizales con algún uso humano

- **Interfase entre vegetación natural + pastizales y cultivos (WAI).** En este caso se añaden los pastizales a las zonas potencialmente inflamables.

4.2.3.2. Infraestructuras y vías de comunicación

Es conocida la importante influencia que tienen distintas construcciones humanas en el origen de incendios. Concretamente se han considerado diferentes infraestructuras (ferrocarril y líneas eléctricas) y vías de comunicación (carreteras, pistas forestales, itinerarios, sendas y vías pecuarias). Éstas variables han sido extraídas de la Base Cartográfica Numérica 1:25.000, disponible en el Centro Nacional de Información Geográfica.

- **Carreteras (ROADS).** La cercanía a éstas puede ser origen de incendios por accidentes o negligencias tales como objetos arrojados desde los vehículos.
- **Líneas ferroviarias (TRAM).** Se sabe que ha sido origen de algunos incendios, por lo que se espera tenga una relación positiva con el peligro de ignición.
- **Itinerarios (IT).** Es tráfico de personas en estos caminos es mayor que en otras zonas naturales, con lo que el peligro de incendio se espera que sea mayor.
- **Líneas eléctricas (PWL).** Distancia a líneas eléctricas. Según la bibliografía constituyen un riesgo de incendio de primer orden por chispas que se pueden desprender, caída de tendidos eléctricos con fuertes vientos o accidentes de aves.
- **Pistas (PIS).** La densidad de población en estos caminos es mayor que en otras zonas naturales, con lo que el peligro de incendio se espera que sea mayor.
- **Sendas (SEM).** La densidad de población en estos caminos es mayor que en otras zonas naturales, con lo que el peligro de incendio se espera que sea mayor.
- **Vías pecuarias (VPE).** La densidad de población en estos caminos es mayor que en otras zonas naturales, con lo que el peligro de incendio se espera que sea mayor.
- **Caminos (TRACKS).** En esta variable se agregan todas aquellas vías que potencialmente pueden ser usadas por viandantes o excursionistas. Tienen cabida: itinerarios, pistas, sendas y vías pecuarias.

4.2.3.3. Factores socioeconómicos

Aquí se agrupan diferentes variables relacionadas con el medio socioeconómico que algunos autores han destacado como relevantes a la hora de explicar los orígenes de los fuegos de carácter antrópico.

- **Zonas protegidas (PROT_A).** En este caso se trata de una variable binaria (categórica) que alberga la existencia o no de un área protegida. Es esperable que la precaución de las personas que visitan estos espacios y la prohibición de algunas actividades de riesgo se relacionen con un menor peligro de incendio, es decir, de forma negativa.

Otra variable socioeconómica que según la bibliografía podría tener importancia son los potenciales demográficos, en tanto que a mayor presencia humana (presión antrópica) se presupone o espera un mayor peligro de incendio.

- **Potencial demográfico por vía de comunicación (*POT_POB*)**. Esta variable, basada en el método de Calvo y Pueyo (2008) y empleada en otros trabajos (Rodriguez y de la Riva, 2014), recoge los potenciales poblacionales por kilómetro cuadrado, siendo un buen reflejo de la presión humana en cada punto del espacio del área de estudio.

4.2.3.4. Factores ambientales

Estas variables pueden ayudar a indicar si existe un comportamiento diferencial en función de las características del terreno, las condiciones de la vegetación etc. En este trabajo se pretende analizar el peligro de incendio humano, pero estas variables juegan un papel como condicionantes que no se puede desdeñar.

- **Altitud (*ALT*)**. Metros sobre el nivel del mar. La altitud resulta un factor de primer orden para entender la distribución de las diferentes comunidades vegetales así como de su estado hídrico, por lo que puede tener un papel importante dentro de las variables independientes. Además, en algunas casos, la actividad humana está también condicionada por ella.
- **Orientación (*ASP*)**. Otro factor que influye de gran manera sobre la humedad edáfica y con ello sobre la vegetación es la orientación. Es por ello por lo que debe estar presente como variable independiente aunque su interpretación puede ser bastante compleja.
- **Pendiente (*SLO*)**. De nuevo un factor con clara influencia en las condiciones edáficas y con ello en la vegetación. También tiene una gran relevancia en las actividades humanas.

Por otra parte, se ha decidido incluir dos productos MODIS que sirven como aproximación al hídrico de la vegetación o combustible, además de añadir variabilidad temporal más allá de los escenarios temporales seleccionados. En este sentido, se entiende que un estado hídrico favorable a la ignición complementa la influencia de las actividades humanas en la medida en que favorece o determina si una posible negligencia o accidente desemboque en ignición o no.

- **MOD16 global evapotranspiration (*MOD_ET*)**. Tiene una resolución espacial de 1km² y una resolución temporal mensual, que permitirá incorporar los datos medios de evapotranspiración para cada mes y cada estación seleccionados como escenarios temporales.
- **MOD11A2. MODIS/Terra Land Surface Temperature and Emissivity (*MOD_LST*)**. Este producto proporciona valores de temperatura y emisividad para cada pixel de 1km² en una secuencia temporal de 8 días. Para este trabajo se ha optado por agregar la información a nivel mensual y adjudicar a cada incendio los valores de su mes y año correspondiente, con objeto de obtener una estimación de su estado hídrico y con ello de la susceptibilidad a un posible incendio.

Para la descarga masiva de los productos MODIS en el periodo 2008-2011 se ha desarrollado un script en R a partir del algoritmo ModisDownload (disponible en <http://r-gis.net/?q=ModisDownload> a 01/09/2015). Dicho algoritmo permite la descarga (de todas o parte de las bandas del producto), mosaicado y reproyección de imágenes MODIS distribuidas por el *Land Processes Distributed Active Archive Center (LP DAAC)*. Esta función permite acelerar significativamente el proceso de obtención de la información, posibilitando trabajar con grandes conjuntos de datos, como es el caso.

4.3. Espacialización de las variables dependientes e independientes

A continuación se detallan los procedimientos realizados en la espacialización de la variable dependiente y de las variables predictoras.

4.3.1. Variable dependientes

La presencia u ocurrencia de incendio se ha obtenido de la base de datos EGIF, por lo que su espacialización pasa por la representación espacial en su correspondiente huso a partir de pares de coordenadas para cada escenario temporal seleccionado. Para ello se utiliza el software propietario ArcGIS mediante su función *Add X/Y data*, teniendo en cuenta que la zona de estudio queda dividida en dos husos distintos en el sistema de coordenadas que se va a utilizar, el ETRS 1989.

La creación de las dieciséis nubes de puntos de ausencia de incendio (ver apartados 4.1 y 4.2.2) se realiza mediante la herramienta *Create Random Points* de ArcGIS. Esta herramienta permite la generación de nubes de puntos aleatorios de tamaño definido por el usuario (en este caso el mismo número de incendios que cada uno de los escenarios temporales anteriormente citados) y con la opción de utilizar una "máscara" con objeto de que los puntos creados estén inscritos dentro de la misma.

Esta máscara es la que nos permite "dirigir" los puntos aleatorios hacia espacios donde potencialmente podría originarse un incendio, evitando así incluir localizaciones inoportunas (láminas de agua) o que se encuentren fuera del perímetro del área de estudio. En la elaboración de la máscara se han estudiado tres opciones de máscara como muestra la figura 6.

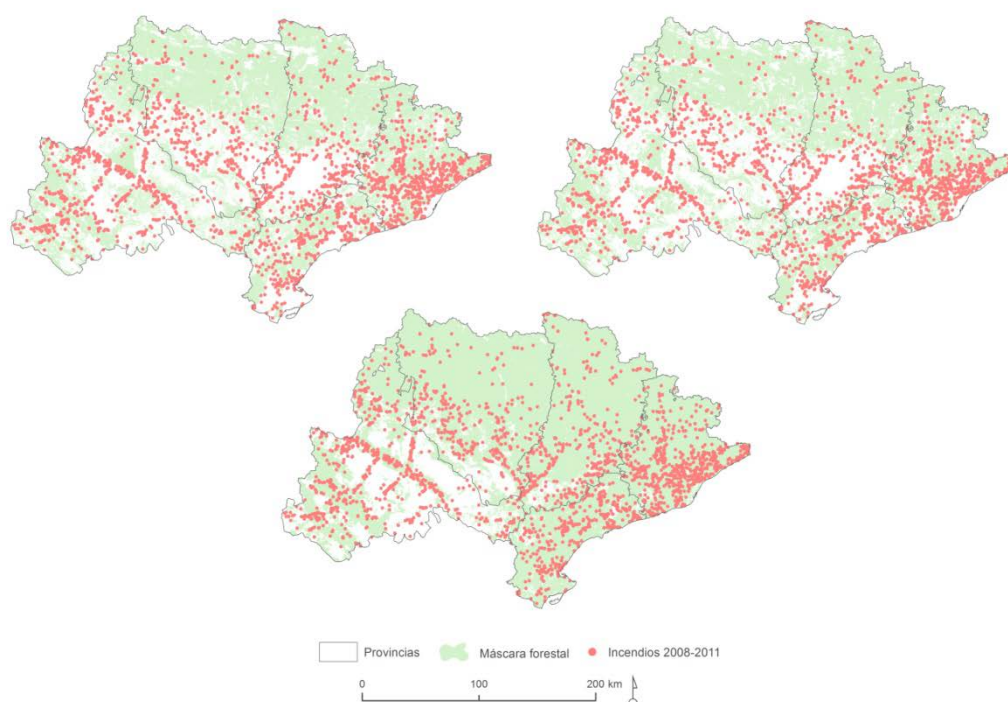


Figura 6. Máscaras de zona susceptible de incendio. Arriba-izq: mapa forestal, arriba-dcha.: CLC, abajo: VCF

La primera de estas máscaras (arriba-izquierda) se corresponde con una fusión de categorías del mapa forestal de España 1:50.000 disponible en la web del Ministerio de Alimentación, Agricultura y Medio Ambiente. Aquí se han excluido los campos de cultivos que no poseen ningún retazo de vegetación natural, además de los espacios en que no resulta posible la aparición de un incendio. Para la segunda se realizó el mismo procedimiento con los datos del *Corine Land Cover* de 2006. Ante la baja asociación entre el conjunto de incendios y las máscaras creadas (muchos de los mismos se localizaban fuera de la misma) se decidió acudir a un producto MODIS, el *Vegetation Continuous Fields*, que contiene una estimación de la cubierta vegetal en valores de 0-100 derivada de las siete bandas del sensor MODIS-Terra. A partir de este producto de 250m de resolución y estableciendo un umbral a partir del cual se considera que existe vegetación suficiente susceptible de ser incendiada (1% en base al análisis el histograma del producto y la distribución espacial de incendios) se consigue un polígono que, si bien no cuenta con la precisión de las máscaras obtenidas mediante Corine o el mapa forestal, si recoge la práctica totalidad de los incendios originados en el periodo estudiado.

De este modo se dispone ya de dieciséis nubes de puntos proyectadas a partir de sus coordenadas (presencia) y dieciséis nubes de puntos aleatorios (ausencia).

4.3.2. Variables independientes

El proceso de espacialización de las variables independientes resulta algo más complejo que la variable dependiente. La tabla 2 resume las variables seleccionadas, los métodos empleados para su espacialización así como otra información complementaria.

Tabla 2. Variables independientes seleccionadas.

	Num.	Nombre de variable	Descripción	Unidades	Fuente de datos	Relación esperada con riesgo ignición
Interfases	1	VCI	Interfase entre vegetación natural y cultivos	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	2	VPI	Interfase entre vegetación natural y pastizales	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	3	VHI	Interfase entre vegetación natural y actividades industriales	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	4	VRI	Interfase entre vegetación natural e infraestructuras de riesgo	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	5	VUI	Interfase entre vegetación natural y espacios urbanos	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	6	WUI	Interfase entre vegetación natural-pastizales y actividades industriales-infraestructuras de riesgo-espacios urbanos	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
	7	WAI	Interfase entre vegetación natural-pastizales y cultivos	Distancia en metros a contacto	Corine Land Cover 2006	Positiva
Socioeconómicos	8	ROAD	Red de carreteras	Distancia en metros	BCN 1:25.000	Positiva
	9	TRAIN	Red ferroviaria	Distancia en metros	BCN 1:25.000	Positiva
	10	ITI	Itinerarios	Distancia en metros	BCN 1:25.000	Positiva
	11	PWL	Líneas y tendidos eléctricos	Distancia en metros	BCN 1:25.000	Positiva
	12	PIS	Pistas forestales	Distancia en metros	BCN 1:25.000	Positiva
	13	SEN	Senderos	Distancia en metros	BCN 1:25.000	Positiva
	14	VPE	Vías pecuarias	Distancia en metros	BCN 1:25.000	Positiva
	15	PROT_A	Áreas protegidas	Categoría (1-SI, 0-NO)	BCN 1:25.000	Negativa
	16	TRACKS	Caminos: ITI-PIS-SEN-VPE	Distancia en metros	BCN 1:25.000	Positiva
	17	POT_POB	Potencial de población de 2011	Habitantes por km ²	Atlas Nacional de España	Positiva
Medio natural	18	ALT	Altitud	Metros sobre el nivel del mar	MDE	?
	19	ASP	Orientación	Grados (°)	MDE	?
	20	SLO	Pendiente	Porcentaje (%)	MDE	?
	21	MOD16	Producto MODIS de evapotranspiración	mm/día	MODIS	Positiva
	22	LST	Producto MODIS de Temperatura de superficie y emisividad	T ^a (K)	MODIS	Positiva

4.3.2.1. Interfases

El procedimiento de espacialización de estas variables es idéntico, por lo que se han utilizado funciones *batch* en el software ArcGIS siempre que ha sido posible, acelerando de este modo la creación de las variables. El proceso seguido se recoge en la siguiente figura:

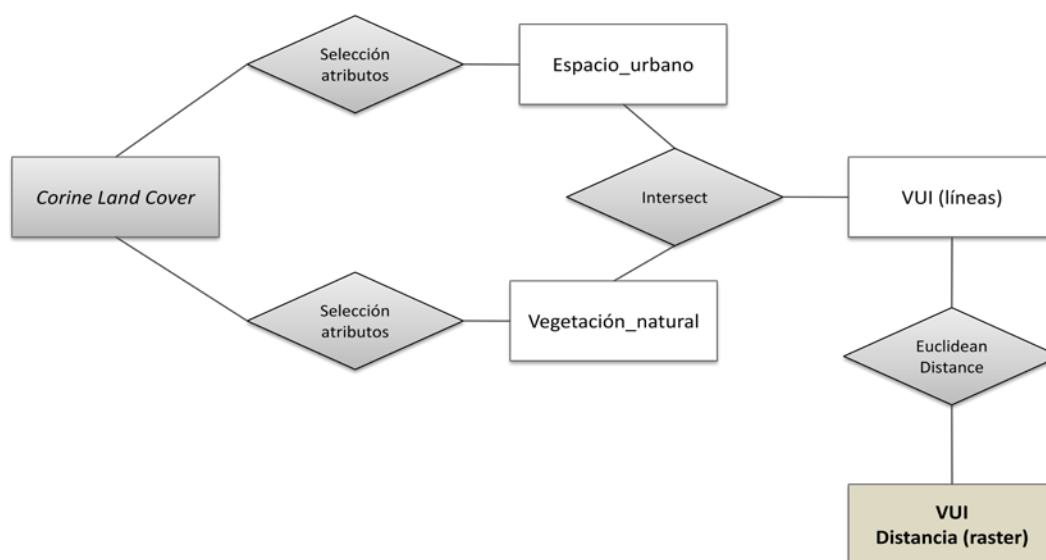


Figura 7. Flujo de trabajo de obtención de interfases.

En un primer paso se seleccionan los dos componentes de la interfase (en el ejemplo los espacios urbanos y la vegetación natural), creando capas temporales que servirán para realizar el *intersect* que producirá como resultado las líneas de contacto (de nuevo según el ejemplo entre los espacios urbanos y la vegetación natural). A partir de ésta y mediante la función de ArcGIS *Euclidean Distance* se halla la distancia de cada pixel a la línea de contacto más cercana, quedando así espacializadas las interfases seleccionadas.

4.3.2.2. Infraestructuras y vías de comunicación y factores socioeconómicos

El proceso de obtención pasa por la descarga de todas las hojas de la Base Cartográfica Numérica 1:25.000 (BCN25) que componen el área de estudio (625 hojas en total) y la unión de cada elemento que posteriormente será empleado como una variable. La depuración de los errores detectados es también parte de este proceso.

Una vez obtenidas en formato vectorial cada una de las variables (carreteras, red ferroviaria, pistas, sendas, itinerarios, vías pecuarias, caminos, líneas eléctricas y áreas protegidas) se procede al cálculo de la distancia a cada una de ellas siguiendo el mismo procedimiento de las interfases, es decir, mediante la herramienta *Euclidean Distance*. Al ser un proceso similar para todas las variables excepto para las zonas protegidas se utiliza una función *batch* que ayuda a reducir el tiempo empleado.

Las áreas protegidas se presentan como una variable categórica, por lo únicamente se debe realizar un proceso de rasterización (*Polygon to Raster*) de la capa vectorial y una reclasificación (*Reclass*) de los valores, asignando el 0 a los píxeles de zonas no protegidas y el 1 a espacios protegidos.

4.3.2.3. Factores ambientales

Entre estas variables, la altitud, la orientación y la pendiente se extraen fácilmente del Modelo Digital de Elevaciones. La primera de ellas es el producto en bruto (altitud en metros sobre el nivel del mar), mientras que para la orientación y la pendiente es necesario realizar sendos procesos. Respectivamente se aplican las herramientas de ArcGIS *Aspect* y *Slope*.

Para el análisis exploratorio de las variables independientes se ha asignado a cada incendio los valores de su imagen MODIS (descargadas y procesadas por el script descrito con anterioridad) correspondiente, mientras que las muestras aleatorias, al carecer de fecha, han tenido que recoger sus valores del promedio mensual o estacional de dichas imágenes. En la aplicación MaxEnt también se hace necesario facilitar estos promedios, puesto que no es posible asignar manualmente cada imagen a las fechas de los incendios.

La agregación de imágenes MODIS para el cálculo de su media es un proceso local (celda a celda) que se puede llevar a cabo en el software ArcGIS con la herramienta *Cell Statistics*. Una vez obtenido el promedio y transformado a formato .asc (como requiere MaxEnt) se adapta su *extent* en QGIS al del resto de variables predictoras.

4.3.3. Análisis de colinealidad de las variables independientes

Se ha realizado un análisis de correlación entre las variables independientes que asegure la consistencia de los resultados de la regresión. El sentido de este análisis es evitar que en el caso que las variables independientes tengan mucha relación entre sí, el modelo de regresión logística no pueda discernir que parte de la variable dependiente se explica con cada variable independiente. Los modelos calibrados con MaxEnt no se ven afectados por colinealidad o autocorrelación espacial de las variables explicativas.

Para ello se ha empleado el coeficiente no paramétrico de *Spearman* (la mayor parte de las variables no se ajustan a una distribución normal), que devuelve un valor ρ (rho) que es una medida de la asociación entre dos variables aleatorias continuas. Es un estadístico no paramétrico, es decir es independiente de la distribución probabilística.

Como se muestra en la tabla 3 los valores están próximos a 0, lo que implica la baja asociación entre las variables independientes. Las variables MODIS sí presentan altos valores de colinealidad, lo que no resulta un problema dado que en el modelo no actuarán nunca al mismo tiempo, ya que cada una lo hará en su mes o estación correspondiente.

Tabla 3.Resultados del análisis de colinearidad. Índice de correlación de la Rho de Spearman.

	POWLINES_DIST	FOR-AGR_DIST	FOR-URB_DIST	TRACK_DIST	ROAD_DIST	PROT_A	MOD_LST_oto	MOD_LST_ver	MOD_LST_pri	MOD_LST_inv
POWLINES_DIST	1.000	-0.003	0.315	0.005	0.122	0.028	-0.118	0.118	-0.064	-0.262
FOR-AGR_DIST	-0.003	1.000	0.224	-0.002	0.025	0.040	0.186	0.132	0.148	0.086
FOR-URB_DIST	0.315	0.224	1.000	0.140	0.227	-0.002	-0.026	0.292	0.038	-0.266
TRACK_DIST	0.005	-0.002	0.140	1.000	0.045	-0.035	0.022	0.091	0.036	-0.049
ROAD_DIST	0.122	0.025	0.227	0.045	1.000	0.062	-0.055	-0.019	-0.077	-0.041
PROT_A	0.028	0.040	-0.002	-0.035	0.062	1.000	-0.084	-0.155	-0.103	-0.027
MOD_LST_oto	-0.118	0.186	-0.026	0.022	-0.055	-0.084	1.000	0.705	0.766	0.673
MOD_LST_ver	0.118	0.132	0.292	0.091	-0.019	-0.155	0.705	1.000	0.724	0.131
MOD_LST_pri	-0.064	0.148	0.038	0.036	-0.077	-0.103	0.766	0.724	1.000	0.598
MOD_LST_inv	-0.262	0.086	-0.266	-0.049	-0.041	-0.027	0.673	0.131	0.598	1.000

4.4. Análisis exploratorio de variables independientes. Técnicas de regresión logística

La regresión logística ha sido profusamente empleada para determinar qué variables tienen mayor importancia en el peligro de ignición a distintas escalas (Martínez et al., 2004; Vasconcelos et al., 2001; Vega-García et al., 1995; Chuvieco et al., 2010). En el presente estudio se utiliza esta técnica para llevar a cabo el análisis exploratorio de la capacidad explicativa de las variables independientes consideradas, así como la selección de los escenarios temporales (mensual o estacional) más adecuados. En comparación con los modelos MaxEnt, la regresión logística resulta más eficiente desde el punto de vista del análisis de la carga explicativa de las variables independientes, además de ser una alternativa más viable en términos de computación.

El modelo de regresión logística se define:

$$P_i = \frac{1}{1 + e^{-z}}$$

$$z = B_0 + B_1X_1 + B_pX_p + \dots + B_pX_p$$

Donde P_i es la probabilidad de ocurrencia de incendio, z la combinación de variables independientes con sus coeficientes de regresión (B), X el valor de cada variable independiente y e la base del logaritmo natural (Pew y Larsen, 2001 citando a Afifi y Clark, 1990; McGrew y Monroe, 1993).

En concreto, de entre las posibilidades de modelos de regresión logística binaria, se aplica el modelo *logit*:

$$\log\left(\frac{p}{1-p}\right) = x^T \beta$$

Siendo x^T el vector de las variables explicativas y β el vector de los parámetros.

Este análisis ha sido realizado en R, obteniendo una serie de resultados que permiten evaluar la representatividad de las variables predictoras. Para entender la relación con la variable dependiente es especialmente importante atender a su sentido explicatorio (*Estimate*) y a la probabilidad Z de que la variable pueda explicar la variable dependiente, habiendo establecido como umbral el 95% ($\Pr|z| < 0,05$).

4.5. Modelado predictivo del peligro de incendio. Modelo de Máxima Entropía (MaxEnt)

MaxEnt es un algoritmo o método de aprendizaje artificial de propósito general que trabaja únicamente con datos de presencia por lo que no es necesario suministrarle las nubes de puntos aleatorias utilizadas en la regresión logística como datos de ausencia. Los modelos MaxEnt se fundamentan en la comparación iterativa de los valores de las variables predictoras -independientes- en las localizaciones de ocurrencia con una gran submuestra que extrae del área de estudio y que actúa como valores de no ocurrencia (Phillips et al. 2006; Elith et al. 2011). Ha sido empleado en estudios de modelización del peligro de ignición en Estados Unidos (Parisien and Moritz, 2009), en la India (Renard et al. 2012) o en China (Chen et al., 2015) entre otras regiones.

MaxEnt estima la probabilidad de distribución de un evento ajustando la distribución de probabilidad a la máxima entropía, es decir, a aquella que es más uniforme a las variables explicativas en cada punto de observación (Phillips, 2006). Una de las características esenciales de MaxEnt es la capacidad de ajustar funciones de respuesta muy complejas combinando varios tipos de funciones (linear, cuadrática, producto, umbral y bisagra), lo que permite modelizar respuestas discontinuas que no pueden ser ajustadas los modelos de regresión más flexibles. El algoritmo MaxEnt se caracteriza fundamentalmente por ser un método que trabaja únicamente con datos de ocurrencia, lo que es particularmente interesante en el caso que nos ocupa, dada la controversia existente en cuanto a los datos de ausencia de incendio, habiendo sido utilizado y validado como método en otros estudios como Bar Massada *et al.* (2012). Dentro de las diferentes opciones de modelado disponibles que trabajan únicamente con datos de presencia, MaxEnt ha demostrado una mayor precisión en la predicción, especialmente con tamaños muestrales pequeños (Elith *et al.* 2006; Pearson *et al.* 2007)

Para el cálculo de los distintos modelos de los escenarios temporales se ha utilizado el modelo de máxima entropía (MaxEntversion 3.3.3k; Phillips *et al.*, 2006; <http://www.cs.princeton.edu/wschapire/Maxent/>)

Para la validación de los modelos, la aplicación ofrece diferentes posibilidades (validación cruzada, *bootstrapping* y *subsampling*). En este estudio se ha optado por el método de validación cruzada mediante k-fold. Se han utilizado cuatro grupos (*folds*) para los que MaxEnt compara iterativamente los resultados (validación cruzada), dejando en cada iteración uno de los grupos como *background* o muestra de validación, de forma que todos y cada uno de los puntos intervienen tanto en observación como en validación. De este modo, MaxEnt produce cuatro modelos con sus correspondientes salidas gráficas (AUC y *jackknife*) y cartográficas en formato .asc (una por cada grupo establecido) asociadas a cada iteración. También facilita una cuantificación de la importancia de las variables empleadas en el modelo basada en el incremento de la ganancia de entrenamiento regularizada a través del desarrollo del modelo, donde la ganancia significa el incremento de la probabilidad de ignición en las localizaciones empleadas como *training*. Dicha ganancia se evalúa utilizando un proceso denominado *jackknife* en el que la importancia de cada variable se estima a partir de modelos con una única variable y modelos con el resto de variables.

Para más información sobre las características y cualidades de MaxEnt se recomienda la lectura de Phillips *et al.* (2004, 2006) y Elith *et al.* (2011).

4.6. Selección del modelo y obtención del peligro para un determinado día.

Una vez los modelos han sido ajustados con MaxEnt, es posible hallar el peligro diario de incendio en función del producto MODIS con resolución temporal de 8 días y seleccionando el modelo adecuado dependiendo la estación del año (invierno, primavera, verano u otoño) y día determinado (día laboral o fin de semana- festivo).

5. Resultados y valoración.

En este apartado se presentan los resultados del estudio y la valoración general de los mismos, respondiendo a los objetivos planteados en el segundo apartado de este trabajo.

5.1. Patrones temporales de la ocurrencia de incendios causados por el ser humano

Este apartado resulta de utilidad como contexto para entender tanto la evolución del número de incendios en España como su distribución temporal a lo largo del año.

5.1.1. Evolución histórica de los incendios en el área de estudio, 1988-2011.

Algunos autores han demostrado la existencia de una tendencia al alza en el número total de incendios y en la superficie calcinada desde 1960 (Kliot, 1996; Pausas and Vallejo, 1999; Wittenberg and Malkinson, 2009). En nuestro caso en concreto, se aprecia un ligero incremento en el número total de incendios (Moreno *et al.*, 2014 y Rodríguez *et al.*, 2013), aunque se echa en falta una serie temporal más amplia para poder extraer conclusiones más sólidas.

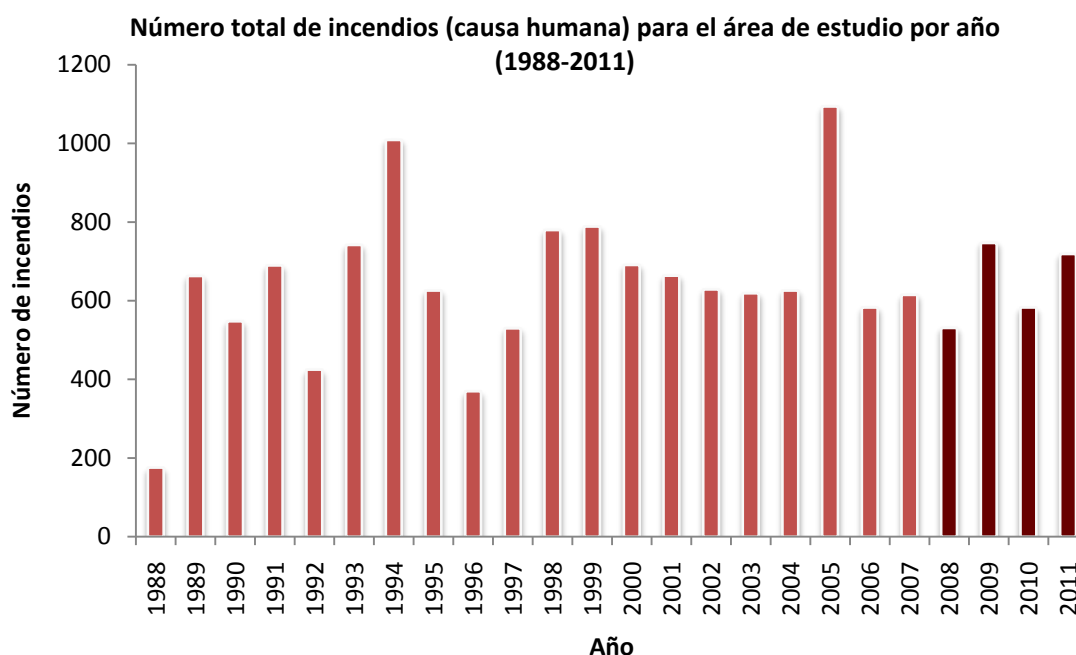


Figura 8. Número total de incendios (causa humana) por año. En granate el periodo de estudio.

Una característica destacable que muestra la gráfica anterior es la amplia variabilidad interanual en el número de incendios antropogénicos, pudiendo originarse menos de 400 en un año como 1994 y casi cuatro veces más en 2005. Este patrón es observable también a una escala estacional, tal y como demuestra la serie histórica del número de incendios por estaciones entre 1988 y 2011 (figura 8). En esta figura es fácilmente observable este comportamiento heterogéneo, llegándose a registrar en el mismo año grandes desviaciones positivas en alguna estación y desviaciones negativas en las siguientes (ver año 2005). Otro aspecto que se puede destacar observando este gráfico es la mayor irregularidad de los incendios de invierno y otoño, siendo el verano la estación que más estable se muestra en el número total de incendios (dentro de una cierta variabilidad intrínseca).

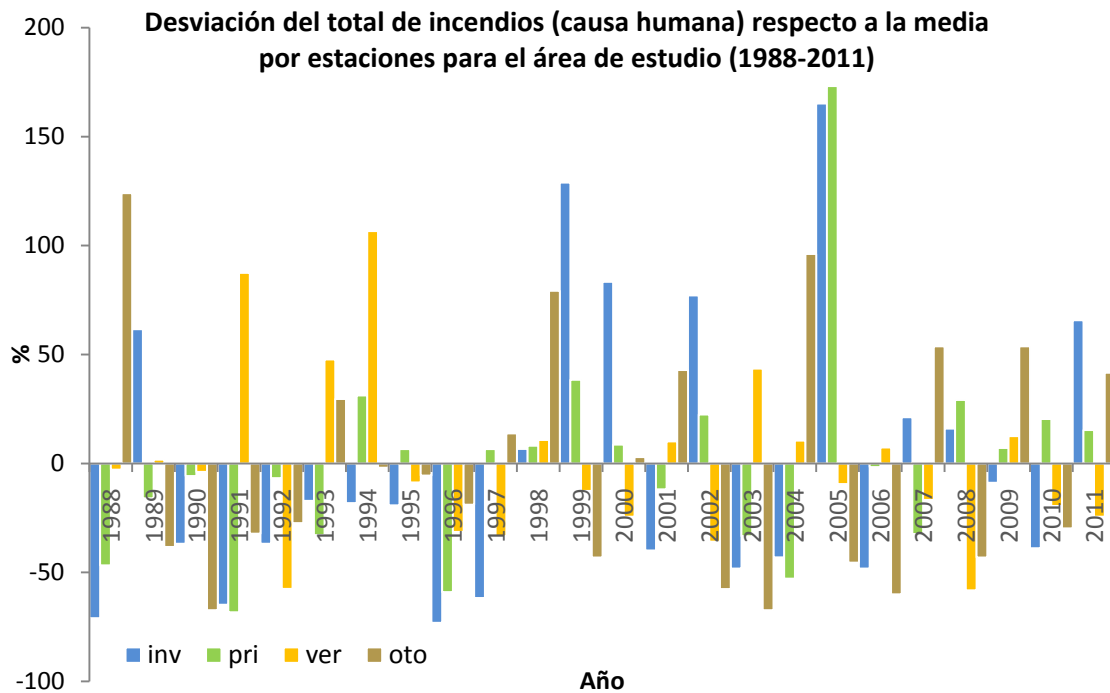


Figura 9. Desviación porcentual del total de incendios (causa humana) respecto a la media por estaciones para el área de estudio (1988-2011).

5.1.2. Patrones temporales mensuales y diarios

La hipótesis de partida de este trabajo destacaba el diferente comportamiento en el origen del fuego por causa humana a escala mensual y diaria, es decir, se entiende, *a priori*, que en función del momento del año en el que nos encontremos, varían los factores responsables (variables independientes) y por tanto, los patrones espaciales de probabilidad.

Las actividades humanas responden a criterios espacio-temporales relacionados con ciclos diarios (*commuting*), semanales (fin de semana), mensuales y/o estacionales (invierno-verano, temporada de incendios), por lo que resulta lógico relacionar esto con el origen de los incendios. Sirva como ejemplo la gráfica mostrada a continuación, que refleja un comportamiento excéntrico en el patrón mensual de los incendios.

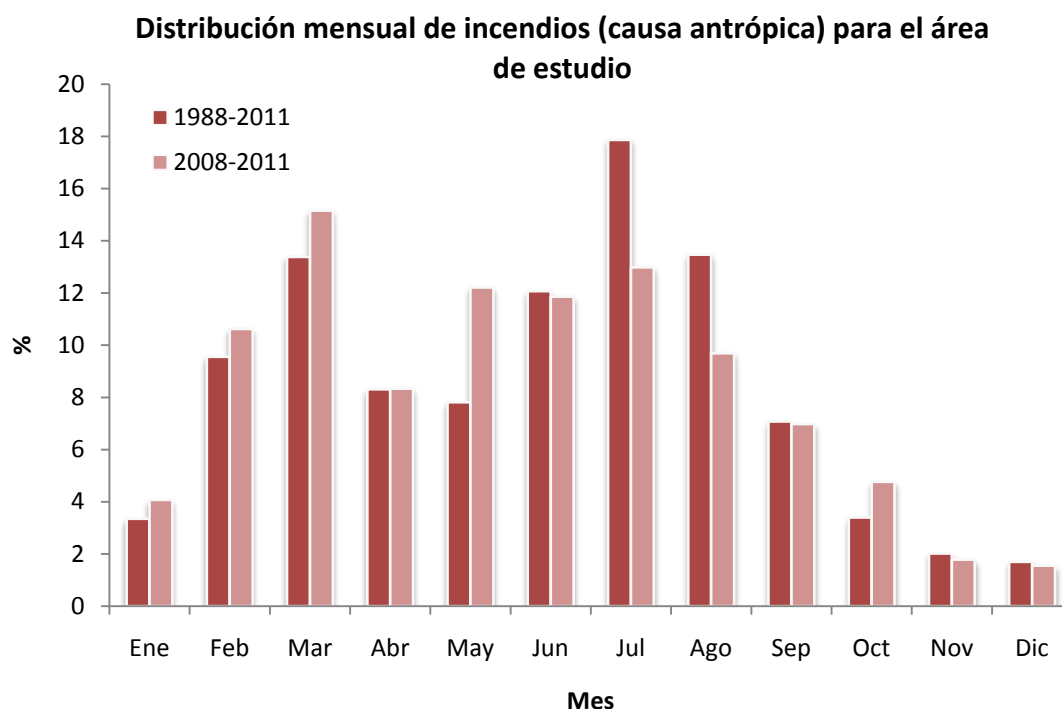


Figura 10. Distribución mensual de incendios (causa antrópica) para el área de estudio en el periodo 1988-2011 y 2008-2011.

La gráfica muestra una distribución bimodal con máximos en finales del invierno-principio de la primavera y en verano. La lógica llevaría a suponer que la distribución debería ser unimodal y concentrada exclusivamente en el verano, cuando la vegetación es más propensa a incendiarse debido a su estrés hídrico y las condiciones meteorológicas son más proclives al inicio del fuego. La temporada de incendios en España se extiende desde finales de primavera hasta principios de otoño, donde se concentran los mayores esfuerzos de vigilancia y prevención. Sin embargo, si bien es cierto que la gran mayoría de los fuegos se originan en el trimestre más seco y cálido (junio-julio-agosto), la aparición de un segundo máximo en el mes de marzo nos habla de un comportamiento humano que podría ser el origen de estos incendios (probablemente la utilización del fuego en superficies agrícolas o forestales para la eliminación de matorrales, pastizales, rastrojos o restos forestales). La hipótesis de que las prácticas agrícolas como origen de este submáximo parece plausible a tenor de los datos que arrojan los meses posteriores (abril o mayo), en principio más propensos a los incendios forestales que el mes de marzo pero en los que las limitaciones de ciertas actuaciones en el medio natural están en vigor. Por otra parte, el elevado número de incendios del trimestre estival puede responder, además de a criterios climáticos (mayor estrés hídrico de la vegetación), al carácter vacacional de estos meses, en los excursionistas se multiplican en los espacios naturales y se produce un mayor número de desplazamientos por carretera.

Del mismo modo que podemos observar un patrón mensual en los incendios forestales antrópicos, la distribución semanal de los incendios vuelve a mostrar un claro sesgo hacia los días en los que se espera que exista mayor contacto entre el ser humano y el medio natural (los días no laborables -sábados y domingos-).

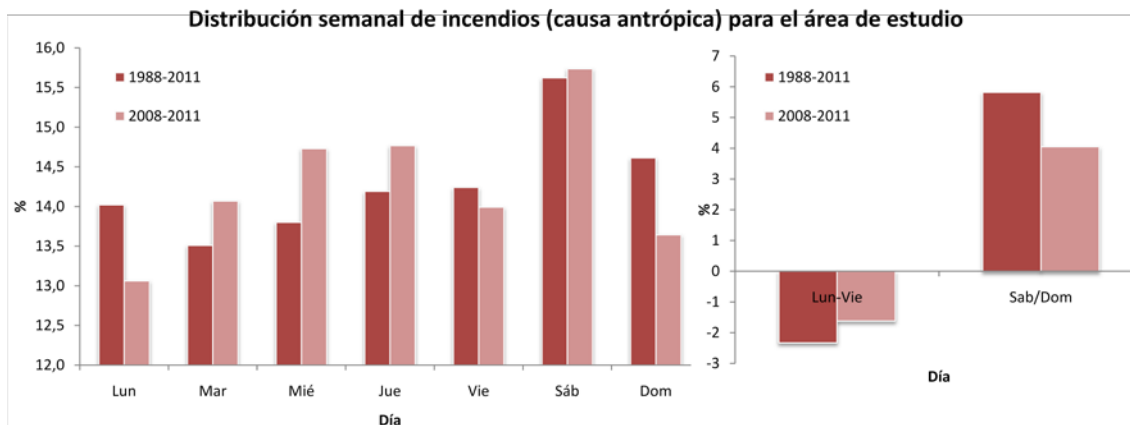


Figura 11. Izquierda: distribución semanal de incendios (causa antrópica) para los periodos 1988-2011 y 2008-2011. Derecha: porcentaje respecto a una distribución homogénea ideal.

Como muestra la figura 11, existe una mayor concentración de incendios durante los fines de semana, siendo un 6% superior (periodo 1988-2011) a lo esperado si la distribución del total de incendios fuera homogénea entre los días de la semana.

El fin de semana es el periodo de la semana en el que se producen un mayor número de desplazamientos hacia espacios no-urbanos, además de ser días en los que agricultores a tiempo parcial, es decir, aquellas personas que tienen su puesto laboral en otros sectores económicos, aprovechan para realizar algunos trabajos agrarios que pudieran ser origen de incendios (Cruz, J. *et al.*, 1985). Por tanto, se puede asumir que la actividad en áreas de riesgo se intensifica en sábados y domingos, lo que puede estar detrás del mayor número de fuegos en estos días. En principio, los días festivos tendrían un comportamiento similar a los sábados y domingos debido a los mismos factores ya destacados.

5.2. Evaluación cuantitativa de las variables independientes en diferentes rangos temporales. Proceso de selección de variables explicativas.

Aquí se muestran los resultados conducentes a la selección de las variables predictoras finales. La tabla 4 corresponde a una de las pruebas iniciales de regresión logística. En ella se puede apreciar la probabilidad Z de que la variable independiente pueda explicar la variable dependiente. Aquellas casillas sombreadas con color representan las que superan el umbral establecido del 95% ($p < 0,05$), reservando el color rojo para aquellas con una relación positiva y el azul para las que tienen una relación negativa.

Tabla 4. Análisis exploratorio de regresión logística 1. En rojo relación positiva, en azul relación negativa.

Pr(> z)	Abr	Abr_FF	Ago	Ago_FF	Nov	Nov_FF	Inv	Inv_FF	Pri	Pri_FF	Ver	Ver_FF	Oto	Oto_FF
ASP	0.379	0.567	0.106	0.362	0.591	0.975	0.877	0.899	0.725	0.519	0.363	0.531	0.354	0.097
VCI	0.076	0.043	0.762	0.454	0.883	0.975	0.560	0.847	0.125	0.070	0.008	0.123	0.928	0.020
VHI	0.403	0.566	0.007	0.507	0.362	0.976	0.231	0.411	0.909	0.002	0.051	0.326	0.404	0.627
VRI	0.281	0.512	0.692	0.704	0.104	0.976	0.486	0.238	0.066	0.168	0.721	0.132	0.912	0.915
VPI	0.024	0.000	0.162	0.439	0.247	0.975	0.000	0.000	0.000	0.000	0.447	0.249	0.000	0.596
VUI	0.258	0.564	0.210	0.689	0.771	0.975	0.008	0.292	0.014	0.580	0.038	0.318	0.239	0.640
ROADS	0.097	0.841	0.001	0.004	0.543	0.976	0.000	0.001	0.000	0.000	0.000	0.002	0.000	0.002
TRAIN	0.152	0.046	0.231	0.041	0.139	0.976	0.522	0.934	0.017	0.095	0.025	0.101	0.974	0.485
ITI	0.076	0.339	0.949	0.890	0.772	0.976	0.987	0.617	0.522	0.397	0.429	0.621	0.301	0.467
PWL	0.007	0.962	0.272	0.480	0.795	0.975	0.183	0.019	0.001	0.481	0.662	0.084	0.468	0.427
ALT	0.010	0.064	0.110	0.108	0.021	0.975	0.861	0.804	0.000	0.339	0.081	0.024	0.477	0.078
PIS	0.083	0.436	0.525	0.682	0.351	0.976	0.140	0.735	0.024	0.063	0.313	0.540	0.226	0.215
POT_POB	0.459	0.984	0.199	0.868	0.779	0.981	0.778	0.290	0.244	0.316	0.084	0.332	0.180	0.801
SEN	0.152	0.853	0.923	0.643	0.650	0.975	0.255	0.419	0.576	0.966	0.544	0.488	0.934	0.830
SLO	0.918	0.024	0.241	0.009	0.043	0.975	0.383	0.846	0.309	0.054	0.176	0.822	0.629	0.040
VPE	0.062	0.033	0.897	0.787	0.446	0.976	0.008	0.004	0.026	0.358	0.000	0.928	0.001	0.062
PROT_A	0.984	0.994	0.680	0.499	0.994	0.991	0.254	0.673	0.072	0.897	0.950	0.111	0.027	0.594
MOD_ET	0.634	0.554	0.366	0.252	0.759	0.975	0.504	0.972	0.473	0.400	0.000	0.060	0.231	0.014

Como se refleja en la tabla, los resultados obtenidos tienen una interpretación compleja y en ocasiones contradictoria, lo que habla de la baja representatividad de las variables. Por ejemplo, resulta llamativa la relación positiva de la interfase VPI (vegetación natural y prados) así como de las vías pecuarias (VPE), es decir, que a mayor distancia mayor peligro de incendio, lo que *a priori* no debiera ser así.

Por otra parte, los resultados arrojados en los escenarios temporales mensuales hacen pensar que una desagregación temporal de este detalle no es la mejor opción, por lo menos con los datos de que se disponen. Destaca, por ejemplo, el reducido número de incendios registrados en noviembre o enero en comparación con otros meses.

Por tanto, esta primera aproximación conduce a descartar los escenarios temporales mensuales dado el problema que el tamaño muestral de algunos meses y aconseja una reorientación de las variables. De este modo, se conservan sin modificaciones las variables PWL, ROADS y PROT_A y se procede agregar variables como TRACKS, WUI y WAI, además de descartar la variable MOD_ET de evapotranspiración a favor de la de temperatura de superficie y emisividad (MOD_LST). Este conjunto de variables puede observarse espacializado en la figura 12. El resto de variables quedan fuera.

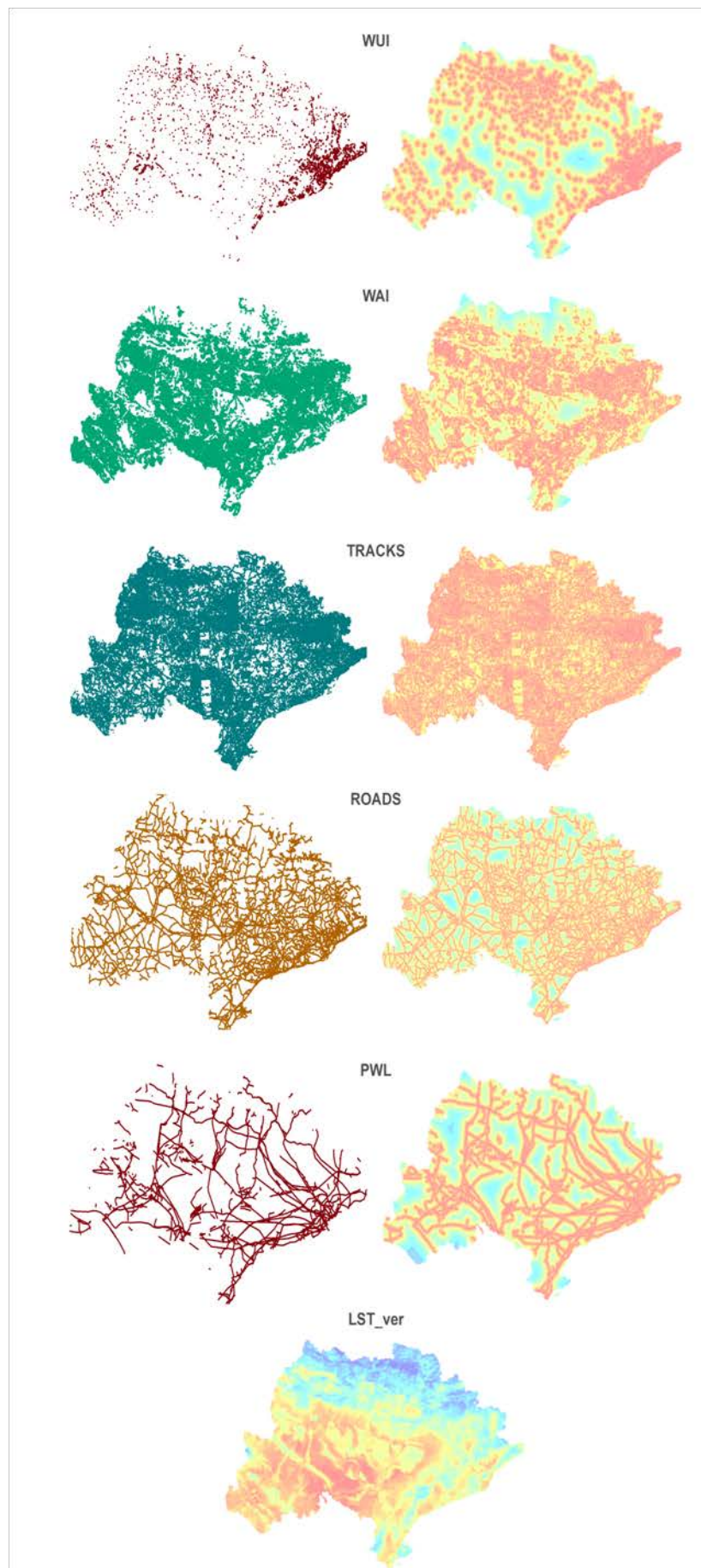


Figura 12. Variables independientes finales.

Tabla 5. Análisis exploratorio de regresión lineal 2. En azul: variables con sentido explicativo negativo, en rojo: sentido explicativo positivo.

Pr(> z)	Inv	Pri	Ver	Oto	Inv_FF	Pri_FF	Ver_FF	Oto_FF
PWL	0.297	0.000	0.000	0.789	0.004	0.723	0.012	0.306
ROADS	0.050	0.000	0.000	0.003	0.010	0.000	0.000	0.000
TRACKS	0.310	0.487	0.041	0.096	0.150	0.207	0.009	0.825
WUI	0.975	0.025	0.004	0.702	0.048	0.149	0.003	0.716
WAI	0.243	0.005	0.051	0.044	0.440	0.004	0.430	0.159
MOD_LST	0.000	0.000	0.000	0.000	0.000	0.109	0.000	0.000
PROT_A	0.595	0.575	0.829	0.238	0.891	0.065	0.226	0.872

La tabla 5 presenta los resultados del análisis de regresión utilizando escenarios estacionales en lugar de mensuales, conservando la desagregación en días festivos y fines de semana. En este caso los resultados parecen más coherentes. Los distintos tonos de azul muestran la distinta capacidad de las variables para explicar la variable dependiente (más intenso, más importancia). Como se puede apreciar, tanto las carreteras como la nueva variable MODIS se configuran en estos escenarios como las más representativas o con mayor carga explicativa en los modelos. Esto quiere decir que en líneas generales la ocurrencia de incendios por causa humana se relaciona con la proximidad a las principales vías de comunicación, acentuado además por el incremento en la temperatura de superficie. Por otra parte, variables como TRACKS, resultado de la agregación de pistas, itinerarios, sendas y vías pecuarias, muestran un comportamiento diferencial a lo largo del año, contribuyendo únicamente en los meses de verano, algo razonable dada la mayor afluencia de turistas hacia zonas forestales en esta época del año, siendo más importante durante los festivos. La interfase vegetación natural - espacio urbano/industrial/riesgo (*WUI*) muestra sus valores representativos más altos durante el estío, cuando estas zonas de contacto se erigen como puntos calientes en el inicio de fuegos. Por su parte, la variable *WAI* muestra sus picos de representatividad durante la primavera, cuando los trabajos agrarios suelen ser más intensos. *PWL* es la variable con un comportamiento más complejo y sin explicación lógica aparente, mientras que las áreas protegidas no presentan representatividad en ninguno de los escenarios temporales analizados, aunque se aproxima al umbral en los fines de semana y festivos de primavera.

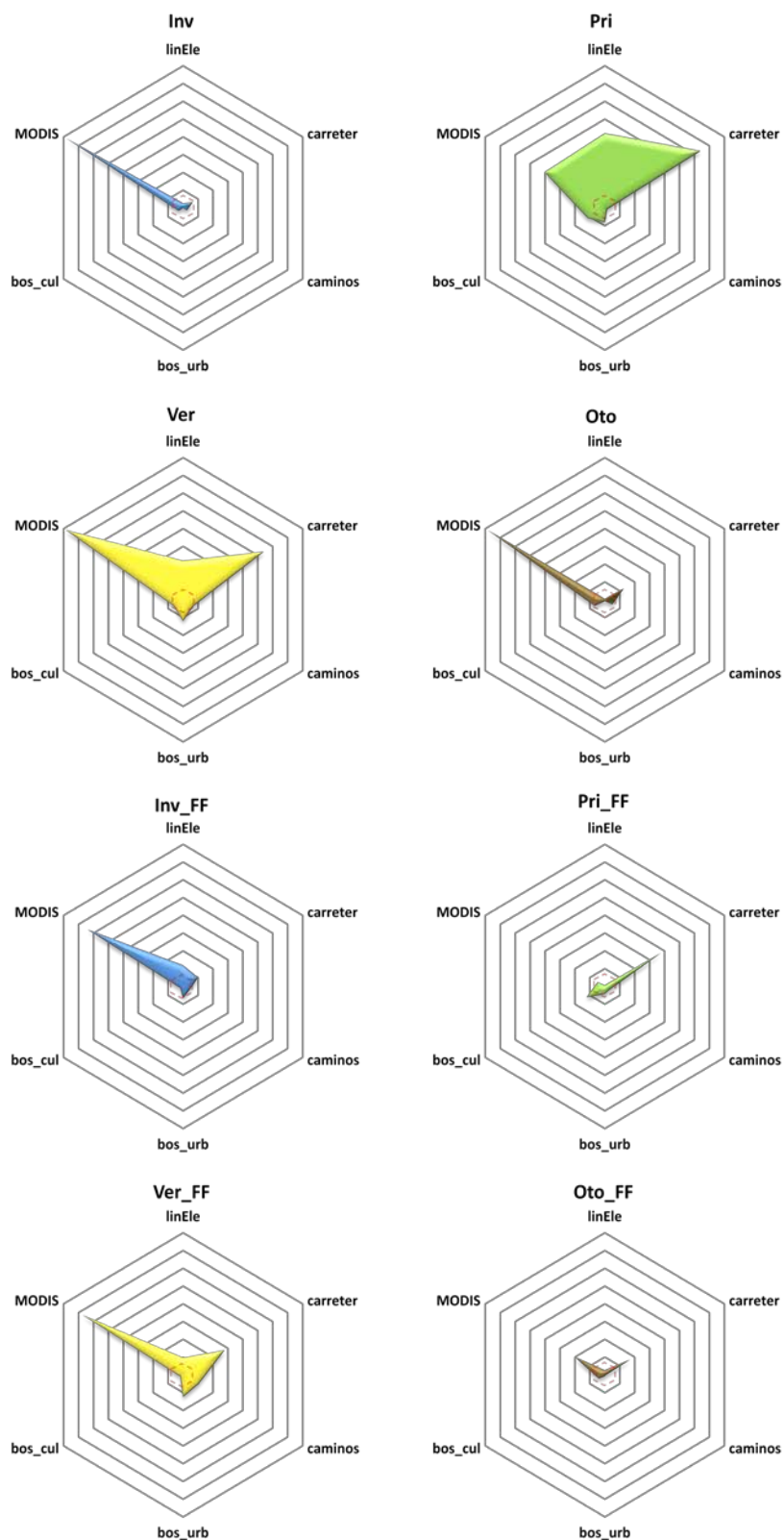


Figura 13. Representación gráfica de los resultados de regresión logística. Azul: invierno, verde: primavera, amarillo: verano, marrón: otoño. El hexágono rojo marca el $P|z| < 0,05$.

La gráfica superior facilita la comprensión de los resultados presentados en la tabla 5. De este modo es más sencillo interpretar el juego de los tamaños

muestrales, ya que como se ve en la parte derecha de la figura, la significación de las variables es menor en los fines de semana y festivos.

Con el fin de comparar y analizar la consistencia de los resultados del proceso de análisis exploratorio de regresión logística con aquellos que se obtendrían directamente en MaxEnt (utilizando todas las variables para el cálculo de cada escenario, no sólo las significativas), se ha realizado el mismo procedimiento en este último, exponiendo sus resultados a continuación.

Tabla 6. Porcentaje de contribución y de permutación de variables independientes producido por MaxEnt

Percent contribution	Inv	Pri	Ver	Oto	Inv_FF	Pri_FF	Ver_FF	Oto_FF
ROAD	35,2	36,7	40,8	29,6	32,2	34,5	37,7	44,9
WAI	31,4	19,7	30	24,7	28,9	32,7	19,7	23
LST_MOD	13,6	24,7	9,9	17,7	15,7	23,2	13,3	10,6
WUI	10,8	3,6	10,8	19,7	9,6	5,4	13,6	16,7
PWL	6,3	14,3	6	5,1	13,4	3,3	13,5	4,3
TRACK	1,5	0,9	1,5	3,2	0,1	0,7	1,1	0,4
PROT_A	1,1	0	1	0	0	0,2	1,1	0

Permutation importance	Inv	Pri	Ver	Oto	Inv_FF	Pri_FF	Ver_FF	Oto_FF
ROAD	51	25,8	47,8	26,5	32,9	30,2	40	43,1
WAI	10,1	6,4	11,2	7,8	19,6	21,9	17	19,5
LST_MOD	17,5	34,2	11,8	32	17,7	30,1	18,9	16,3
WUI	8,6	7,8	7	12,3	8,6	9,3	6,1	11,9
PWL	12,1	16,4	19,4	11,9	19,1	7,7	15,1	9,1
TRACK	0,3	9,3	1,2	9,5	2,1	0,6	0,9	0,1
PROT_A	0,5	0,1	1,6	0	0	0,3	2	0

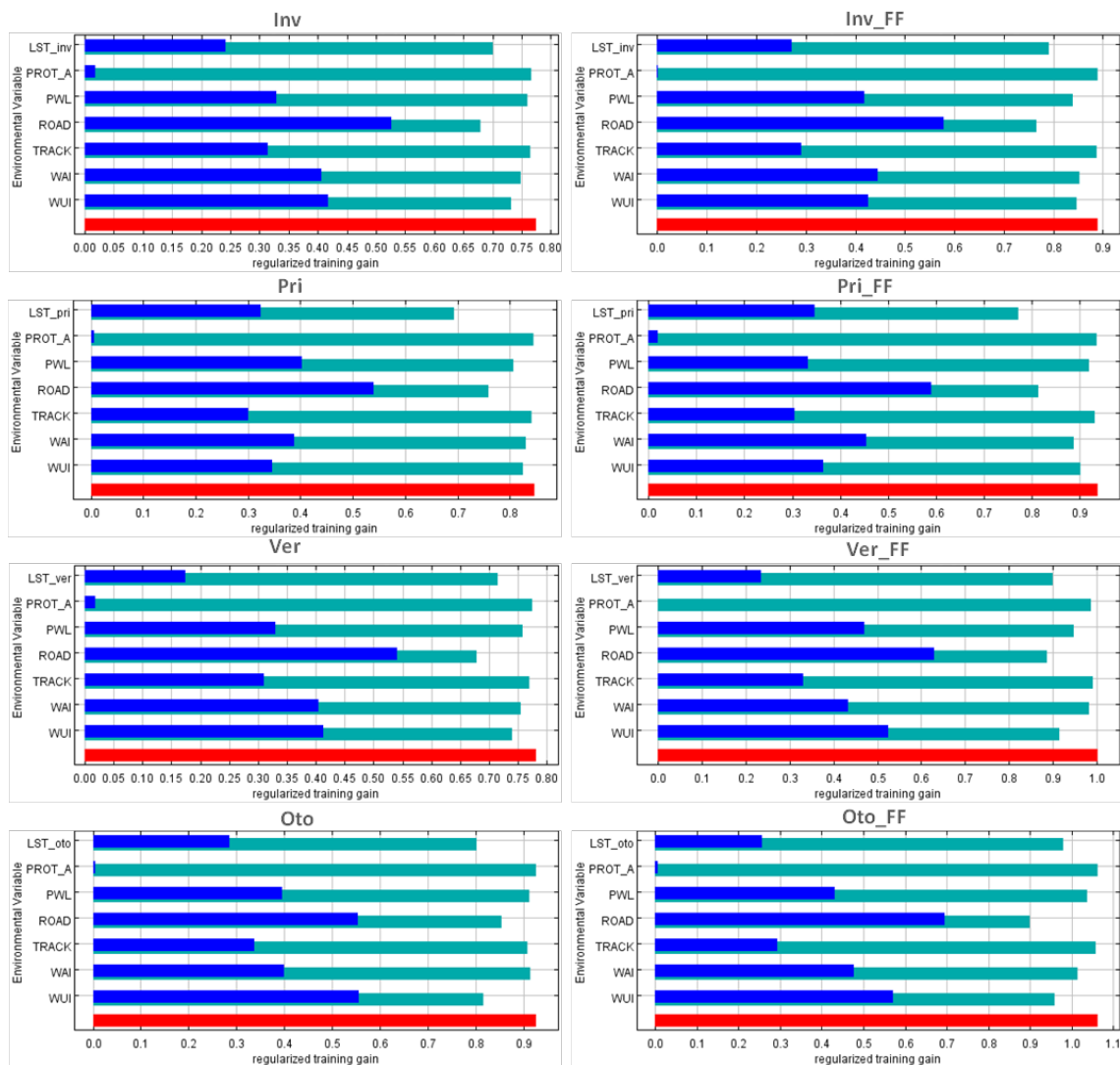


Figura 14. Gráficos *jackknife* de MaxEnt. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Estos resultados de MaxEnt coinciden a grandes rasgos con los obtenidos previamente por la regresión logística, confirmando su validez como método exploratorio de variables predictoras. En general, destacan las carreteras el factor decisivo en todos escenarios temporales, teniendo también un peso destacado el producto MODIS y la interfase WAI.

De acuerdo con estos resultados, se utilizan las variables representativas para cada escenario temporal en el cálculo del algoritmo de MaxEnt para desarrollar los modelos temporales, es decir, para cada escenario temporal las variables empleadas serán las siguientes:

Tabla 7. Selección final de variables representativas.

	Inv	Pri	Ver	Oto	Inv_FF	Pri_FF	Ver_FF	Oto_FF
ROAD	X	X	X	X	X	X	X	X
WAI		X	X	X		X		
MOD_LST	X	X	X	X	X		X	X
WUI		X	X		X		X	
PWL		X	X		X		X	
TRACK			X				X	
PROT_A								

En el caso de *WAI* en verano se ha decidido su inclusión por estar prácticamente en el umbral (0,05086), pudiendo ser un factor explicativo importante.

5.3. Modelos predictivos generados con MaxEnt

Se presentan a continuación los modelos predictivos hallados con MaxEnt junto con la validación de los mismos y la implicación de las variables independientes en cada modelo creado.

5.3.1. Predicción e incertidumbre de los escenarios temporales

La figura 15 presenta las salidas cartográficas derivadas de la aplicación MaxEnt tras un tratamiento en ArcGIS que mejora la calidad de la cartografía presentada por la aplicación. Mediante una leyenda común se puede comparar las diferencias entre los distintos escenarios, todos muy marcados por el predominio de las carreteras como factor explicativo. A grandes rasgos, se puede destacar el litoral catalán y la zona oriental de la provincia de Huesca como las áreas con un mayor nivel de peligro, mientras zonas con una baja presión demográfica menor, como el Pirineo, muestran niveles de peligro sensiblemente menores.

La figura 16, por su parte, muestra la incertidumbre de cada modelo utilizando el coeficiente de variación (desviación estándar del conjunto de los modelos individuales obtenidos en la validación cruzada dividida por la media de los mismos). El resultado es el grado de incertidumbre expresado en tantos por 1 de variación de la probabilidad por encima o por debajo del valor promedio obtenido (figura 16). Esto ayuda a comprender las zonas en las que los modelos se comportan de una forma más estable y por ende fiable o dónde, por el contrario, encuentran mayor dificultad a la hora de prever el peligro, variando las predicciones en función de la muestra de datos de ocurrencia utilizada. Visualmente se puede apreciar que la mayor incertidumbre se concentra en zonas de peligro bajo, coincidiendo en los distintos modelos temporales las mismas áreas con alta incertidumbre.

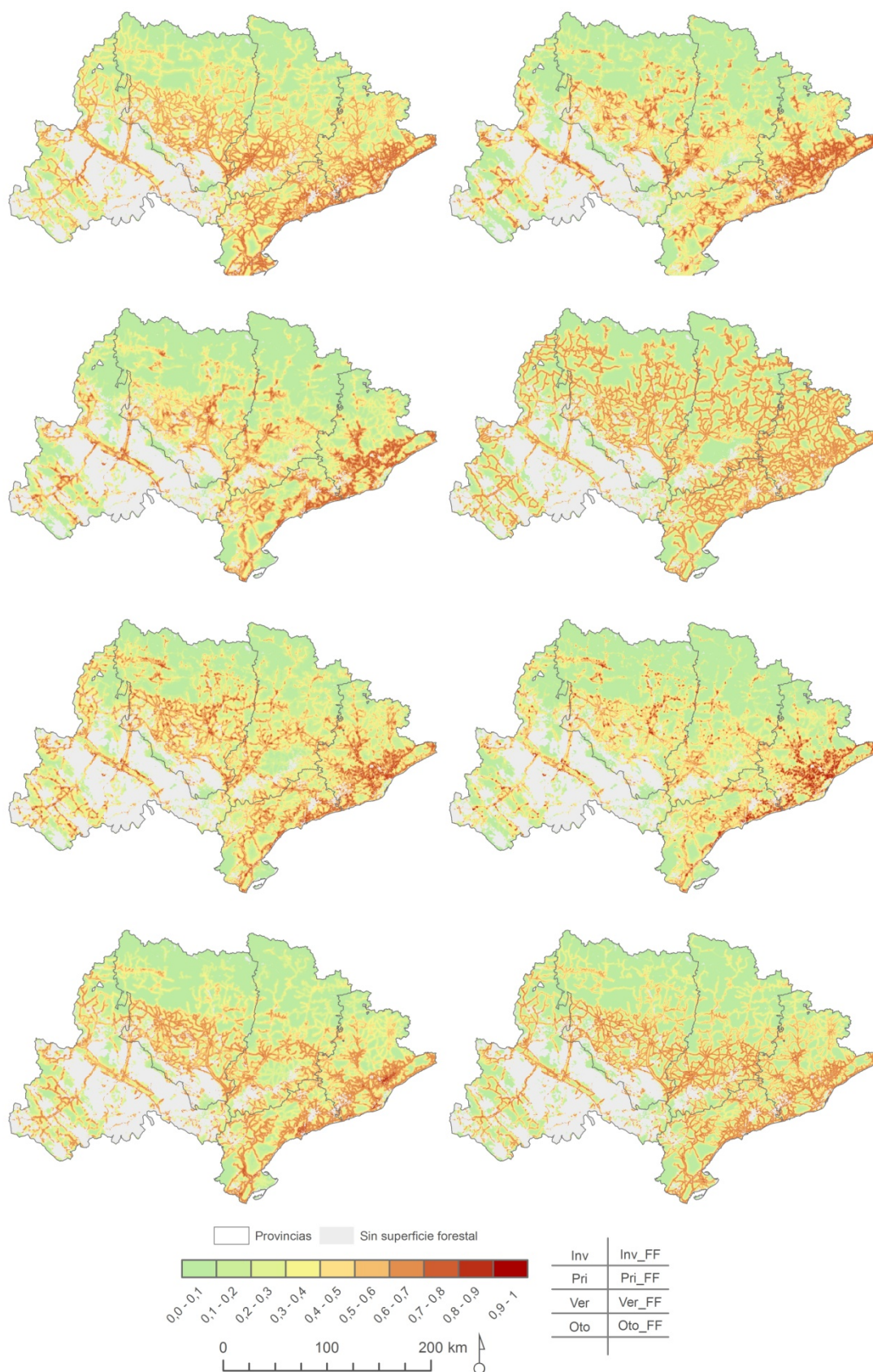


Figura 15. Promedio de probabilidad de ignición de los modelos individuales de los escenarios temporales.

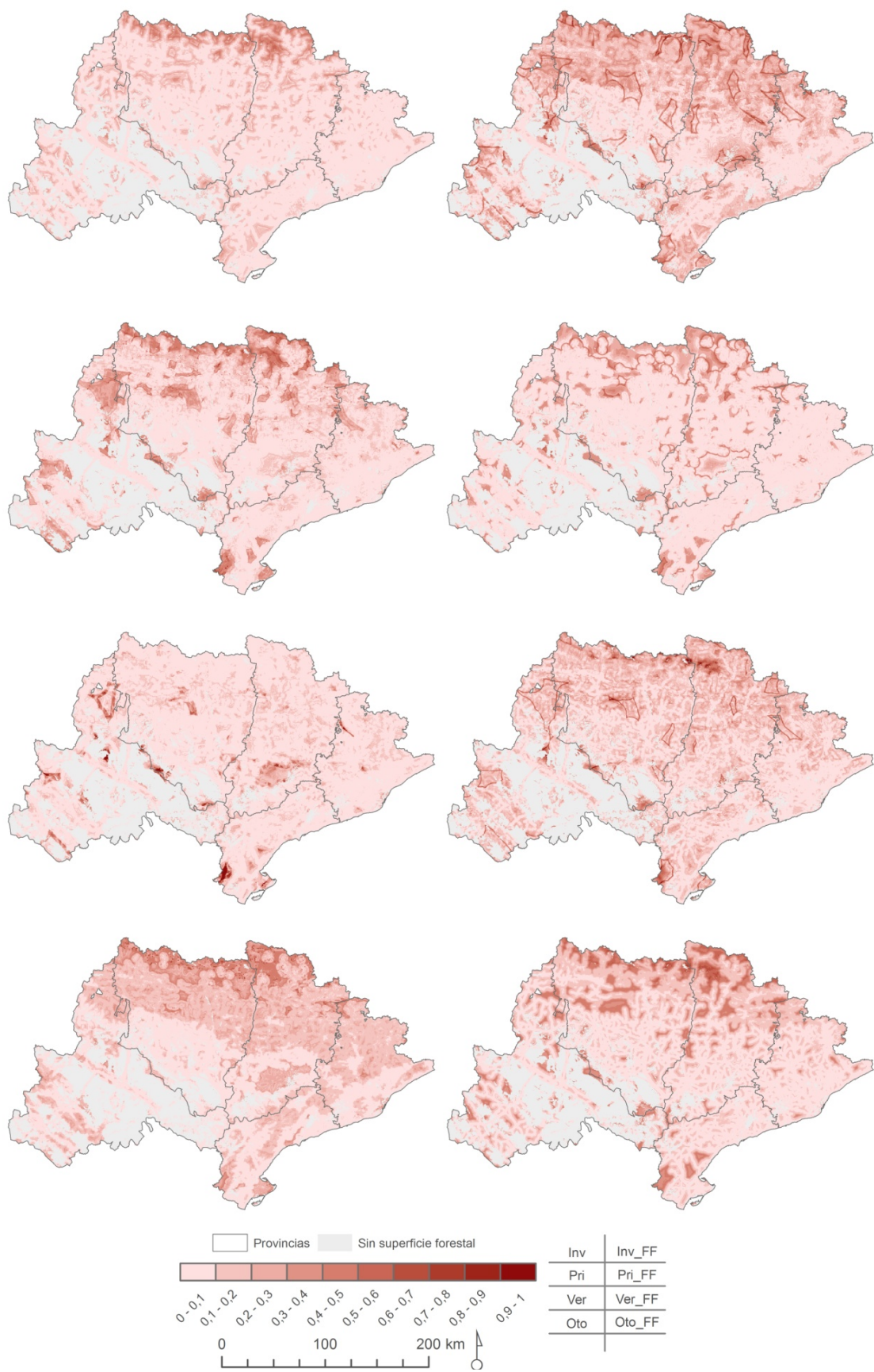


Figura 16. Incertidumbre en la predicción de los modelos de escenarios temporales

5.3.2. Validación de los modelos

La figura 17 recoge las medias de las tasas de omisión y el área predicha como función del umbral acumulado para cada escenario de los cuatro modelos (*k-fold*) individuales generados (modo de validación escogido entre las opciones de MaxEnt). La tasa de omisión, representada por la línea azul turquesa y el sombreado amarillo que recoge la varianza cubierta por una desviación estándar, debiera estar, idealmente, próxima a la tasa de omisión predicha (línea diagonal desde el origen de coordenadas). En las gráficas se identifica como la tasa de omisión se adapta armónicamente a la omisión predicha, teniendo una desviación mayor en la estación otoñal. Esto significa que no existe sesgo en ningún intervalo, es decir, que no se aprecia una gran desviación en valores bajos, medios o altos, según la distribución esperada.

Tabla 8. Valores de AUC de los modelos predictivos.

	AUC (min)	AUC (max)	AUC (mean)	StdDev (mean)
Inv	0,764	0,817	0,799	0,021
Inv_FF	0,811	0,854	0,835	0,016
Pri	0,829	0,851	0,843	0,009
Pri_FF	0,829	0,851	0,843	0,008
Ver	0,822	0,857	0,842	0,014
Ver_FF	0,827	0,870	0,855	0,017
Oto	0,810	0,846	0,823	0,014
Oto_FF	0,815	0,847	0,834	0,012

Por otra parte, la tabla 8 y los gráficos recogidos en la figura 18, reflejan la calidad de los modelos obtenidos. En las gráficas se puede observar la gran distancia que existe desde la línea que representa la media del AUC (*Area Under the ROC Curve*) de los modelos individuales y la línea recta que representa como actuaría un modelo aleatorio, lo que implica que los modelos generados están prediciendo razonablemente bien la ocurrencia de incendio.

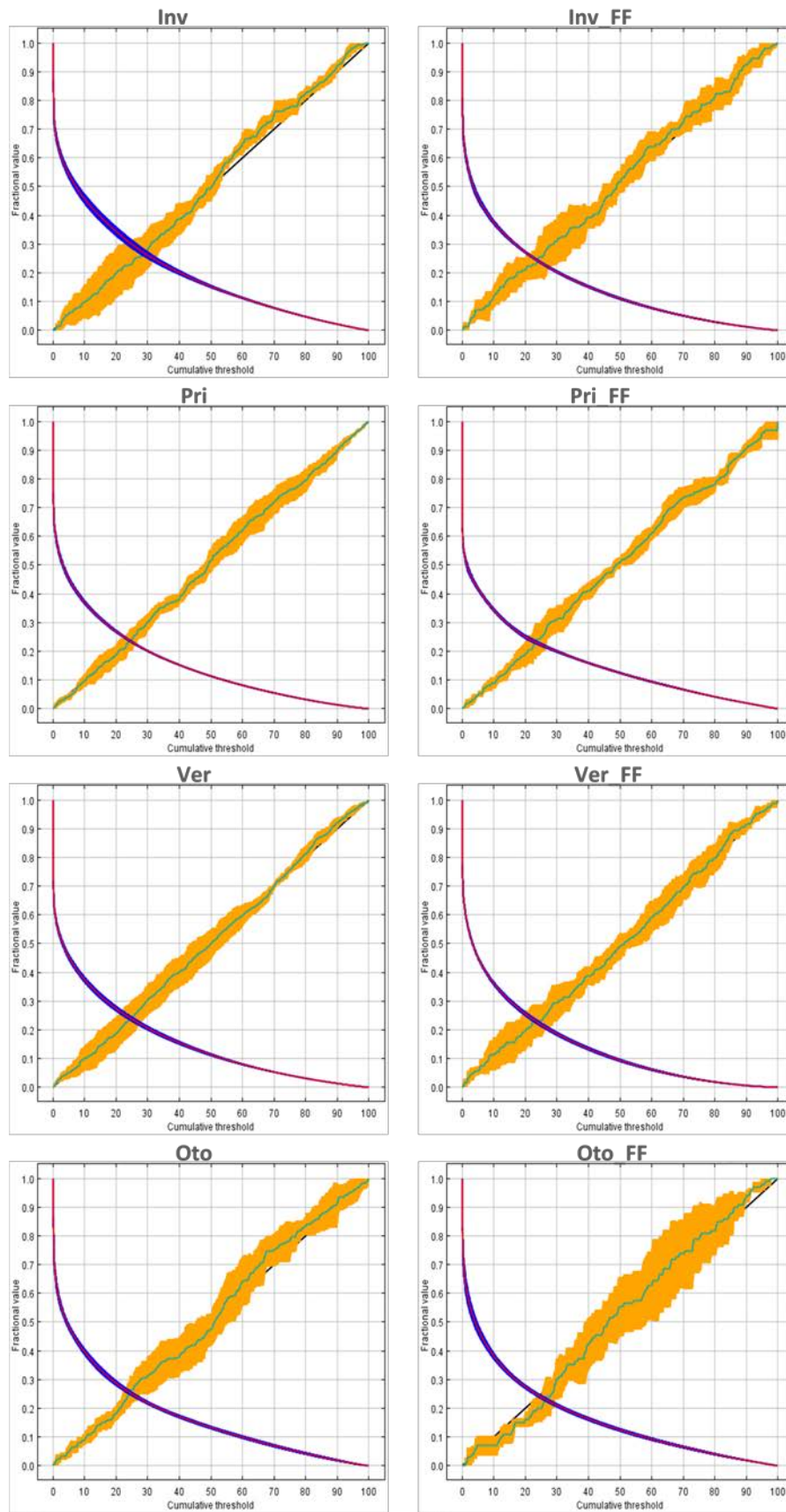


Figura 17. Medias de omisión y predicción para los distintos escenarios. Rojo: área media, azul: área media \pm una desviación estándar, turquesa: media de omisión en los datos test, naranja: media de omisión \pm una desviación estándar, negro: omisión predicha.

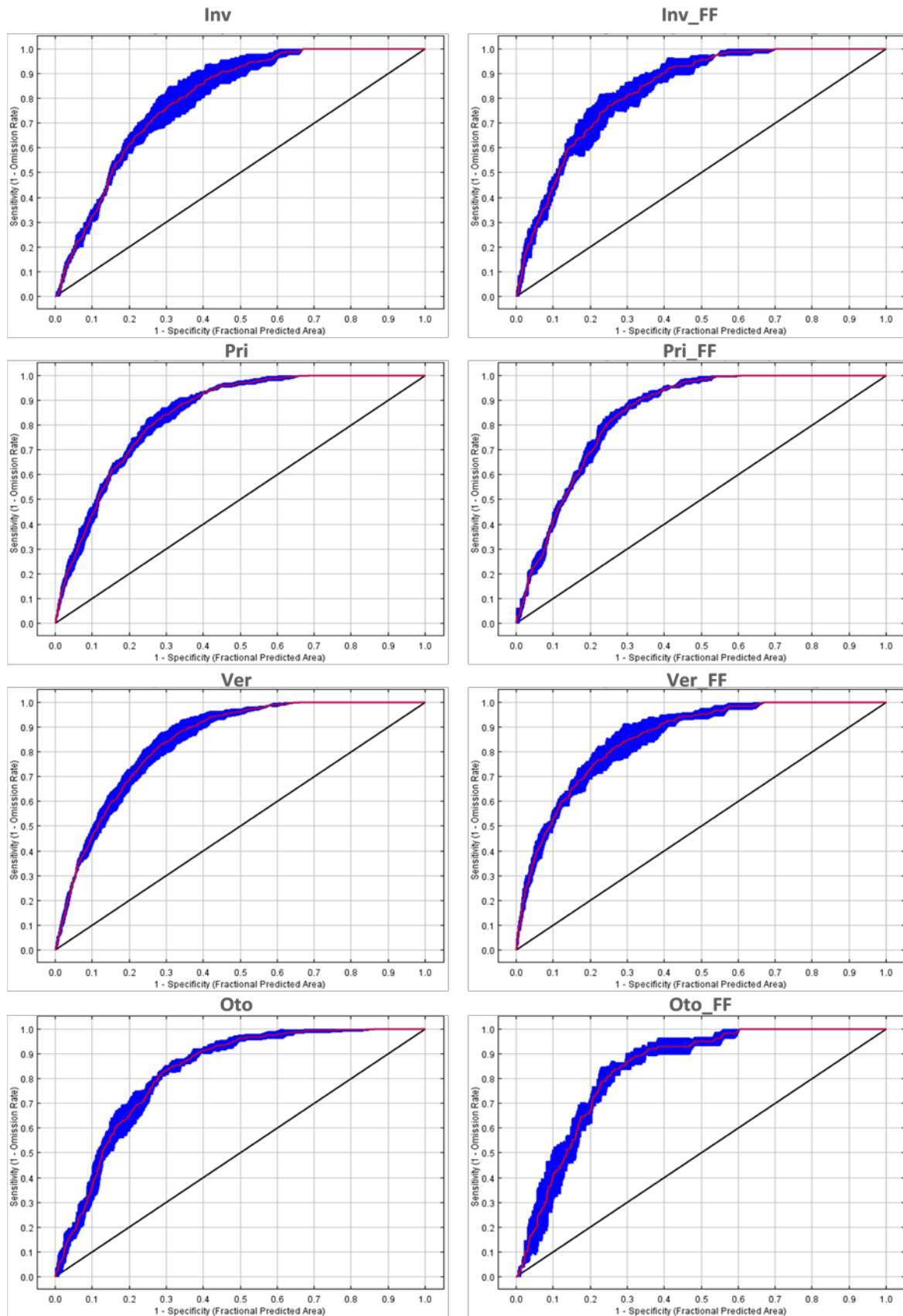


Figura 18. Gráficos de área bajo la curva ROC (Receiver Operating Curve o curva de funcionamiento del receptor). En rojo: media, en azul: media +/- una desviación estándar, negro: predicción aleatoria.

5.3.3. Curvas de respuesta y contribución de variables en los modelos predictores

En este apartado se van a mostrar los resultados con detalle de las variables explicativas para cada escenario temporal.

5.3.3.1. Invierno

En el caso de los días laborables de la estación invernal, únicamente las variables MOD_LST y carreteras forman parte del modelo, lo que facilita la interpretación de la siguiente figura. En ella se puede apreciar la contribución de cada variable en el modelo, con un peso muy superior de las carreteras, cuyo porcentaje de contribución llega hasta el 79%.

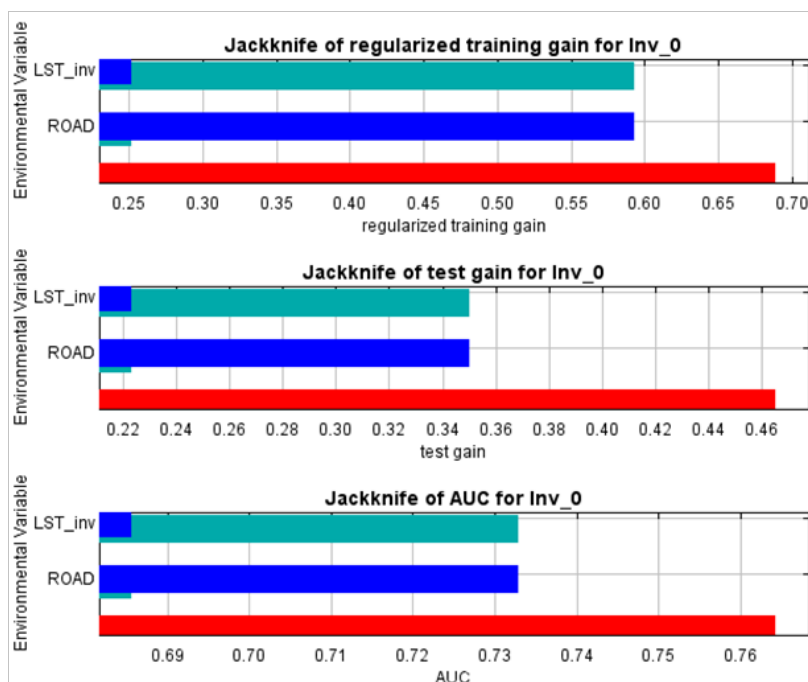


Figura 19. Graficos *jackknife* de Inv. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 9. Porcentajes de contribución y permutación de variables predictoras de Inv.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	78,8	77,4
LST_inv	21,2	22,6

Por otra parte, en la figura 18 se observa como las variables explicativas afectan la predicción de MaxEnt. Estas curvas muestran cómo cambia la predicción conforme a la variación de la variable independiente.

La variable MODIS se relaciona positivamente con el peligro, de forma que a mayor temperatura de superficie mayor peligro. Por su parte, la interacción con las carreteras es inversa y mucho más sesgada hacia valores 0, es decir, la acción de las carreteras como factor explicativo decae drásticamente conforme nos alejamos de ellas o lo que es lo mismo, a menos distancia mayor riesgo.

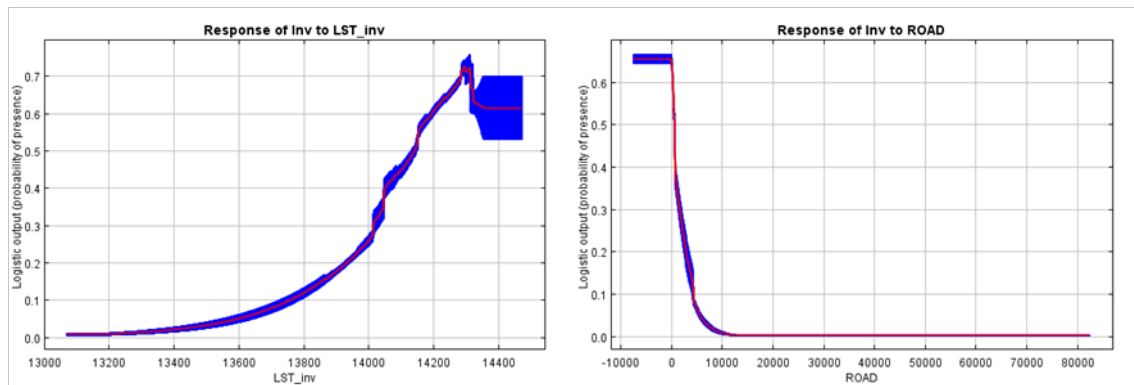


Figura 20. Curvas de respuesta de variables predictoras de Inv.

En el modelo que recoge los días festivos y fin de semana en invierno intervienen cuatro variables: MOD_LST, ROAD, PWL y WUI. Entre estas, vuelve a ser las carreteras el factor más importante, aunque no con tanta preponderancia como en el caso anterior. Las carreteras son responsables de un 45% del modelo, con un importante descenso en el porcentaje de permutación, lo que indica que al asignar valores aleatorios a esta variable el modelo se resiente. Este hecho es también observable en el *jackknife* AUC donde se muestra que con la ausencia de las carreteras el modelo pierde capacidad explicativa, aunque de forma muy inferior a lo que sucedía en los valores de invierno-días laborables. Un aspecto común al conjunto de curvas de respuesta de carreteras presentadas en los siguientes gráficos es que el radio de acción de estas es constante, en torno a los 5000 metros. Este patrón ha sido ya analizado en otros estudios que han concluido la presencia de este umbral (Vega Garcia et al.1993, Vasconcelos et al., 2001, Mollicone *et al.*, 2006).

La variable MODIS no se erige como una variable clave a la hora de explicar el modelo, ya que como se muestra en la figura su capacidad explicativa en ausencia de otras variables es muy baja, sin embargo si sirve como matiz necesario en el conjunto de las variables, pues en su ausencia merma la capacidad predictiva de la distribución de ocurrencia. Las líneas eléctricas y la interfase WUI, si bien son capaces de explicar una buena parte del modelo por sí mismas, si se omiten el porcentaje total AUC (que mide la capacidad predictiva) disminuye muy ligeramente.

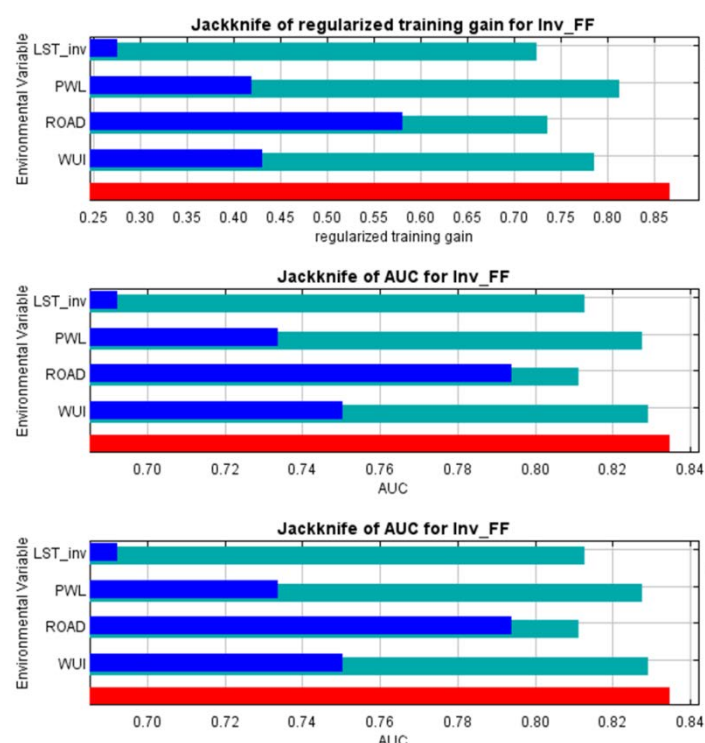


Figura 21. Gráficos *jackknife* de Inv_FF. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 10. Porcentajes de contribución y permutación de variables predictoras de Inv_FF.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	45	34,6
LST_inv	20,3	28,9
PWL	20,2	18,7
WUI	14,6	17,8

El comportamiento de las variables predictoras es el esperado tal y como presentan sus curvas de respuesta (figura 23). Sin embargo, es especialmente reseñable en la variable MODIS y en la distancia a las líneas eléctricas las discontinuidades en la curva que las recoge, resultado especialmente de la distribución propia de la variable dentro de la zona de estudio (predominan los valores medios) y a la resolución espacial empleada (250m), lo que limita una mayor precisión del modelo. Esto último es más fácilmente entendible consultando la figura siguiente, que recoge una serie de histogramas del modelo de verano que pueden ayudar a entender estas discontinuidades en las curvas. Globalmente el comportamiento es lógico, con una progresión positiva en la variable MODIS (mayor temperatura, mayor peligro) e inversa en las variables de carreteras, líneas eléctricas y la interfase WUI (menor distancia, mayor peligro).

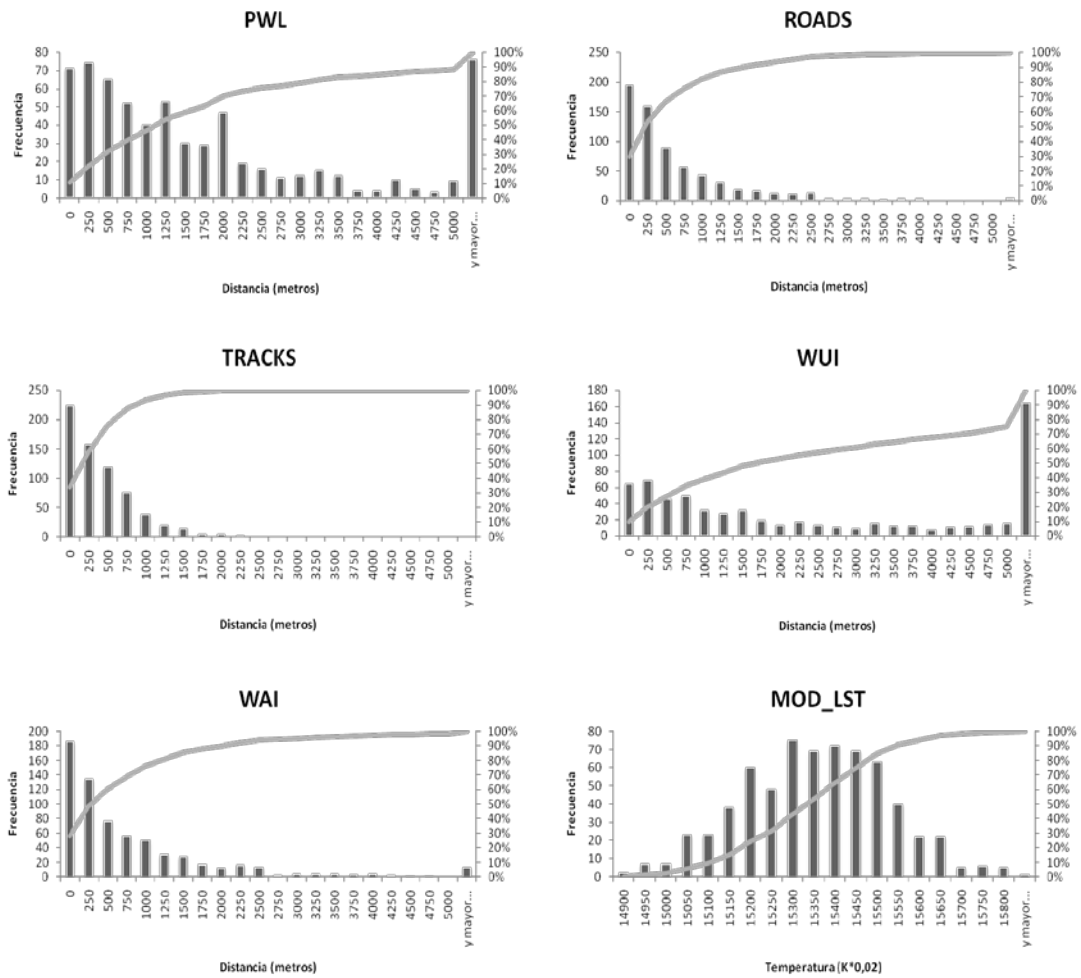


Figura 22. Histogramas de frecuencias de variables predictoras del modelo de verano.

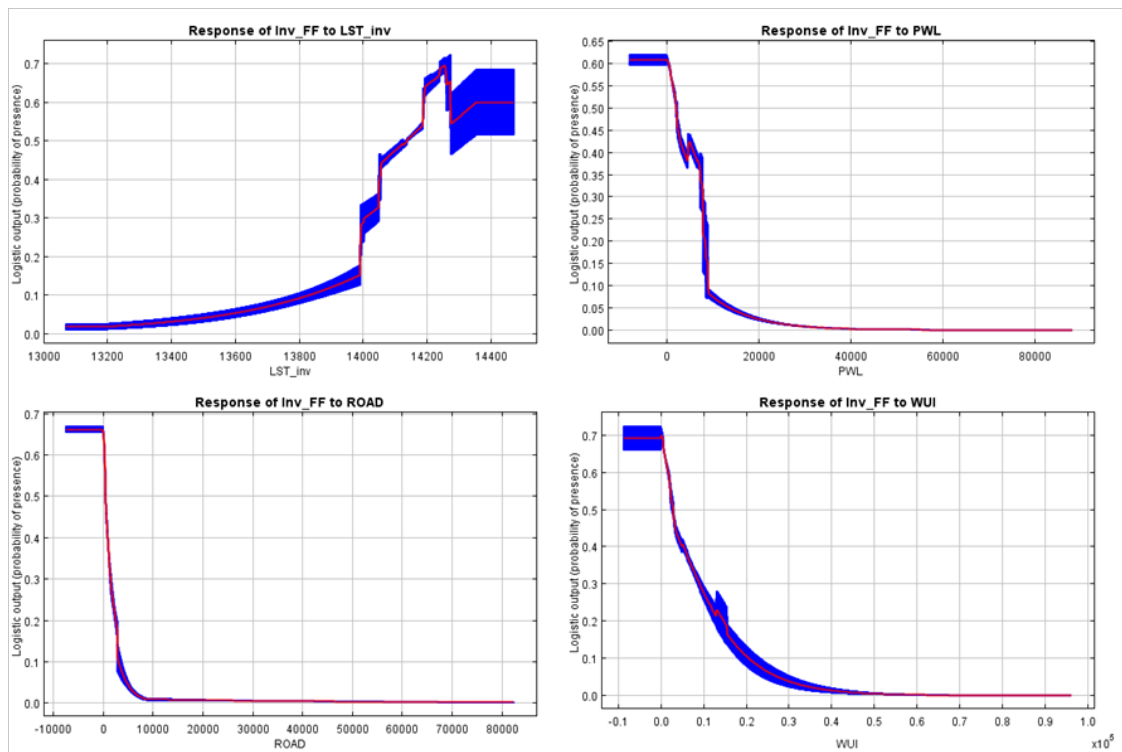


Figura 23. Curvas de respuesta de variables predictoras de Inv_FF.

5.3.3.2. Primavera

En primavera se amplía el número de variables implicadas, pero vuelven a ser la distancia a las carreteras el factor con un mayor peso explicativo en el modelo, aunque en este caso la variable LST, en caso de ser excluida, es la que produce un menor AUC total del modelo. En este caso intervienen ambas interfases, siendo la de mayor importancia la agraria, responsable de casi un 20% de la capacidad predictiva y con una importante reducción del porcentaje en la permutación, lo que implica que el modelo depende de forma importante en esta variable. Este hecho es significativo de la mayor actividad en las parcelas agrarias en los meses de primavera respecto al invierno. Las líneas eléctricas parecen tener mayor peso en este modelo respecto a otras estaciones, pero encontrar una causalidad resulta complejo.

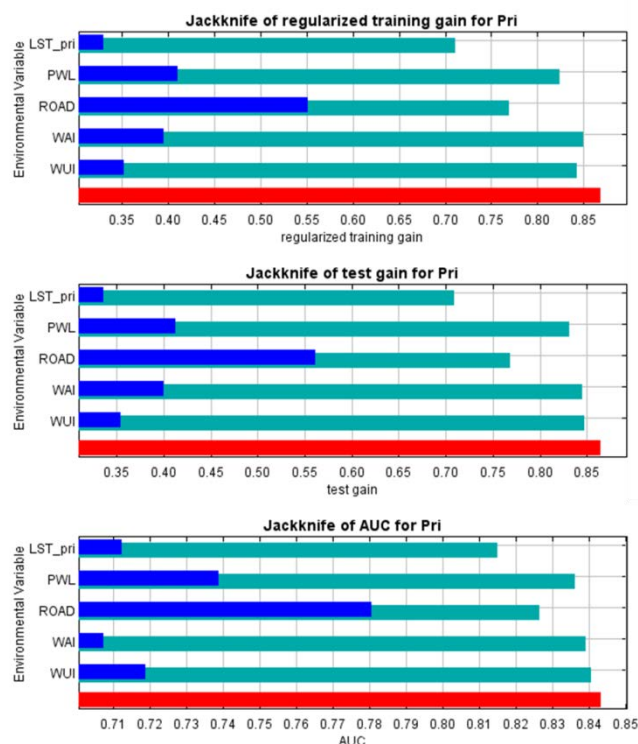


Figura 24. Gráficos jackknife de Pri. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 11. Porcentajes de contribución y permutación de variables predictoras de Pri.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	37,6	34,4
LST_pri	25,3	31,5
WAI	19,9	10,4
PWL	12,3	16,6
WUI	4,9	7,1

Las curvas de respuesta nos ayudan a comprender el comportamiento del peligro en función de las distintas variables. Las discontinuidades, en este caso, destacan en la variable WUI, mientras que PWL y especialmente ROAD y WAI muestran un perfil descendente muy definido, desde un peligro muy elevado en los primeros metros de proximidad hasta valores prácticamente nulos cuando nos alejamos de estas localizaciones. El peligro en función de la variable MODIS, por su parte, crece de forma exponencial con los valores de esta, desde un peligro nulo con los valores más bajos (situados en las áreas más frías y umbrías) hasta alcanzar un máximo entre los

14.800 y los 15.000 -cabe recordar que para obtener los valores en grados *celsius* hay que multiplicar estas cifras por 0,02 y restar 273-. A partir de este umbral, el peligro decae como consecuencia del descenso en la frecuencia, no por la disminución real de la ocurrencia de fuegos en estas zonas.

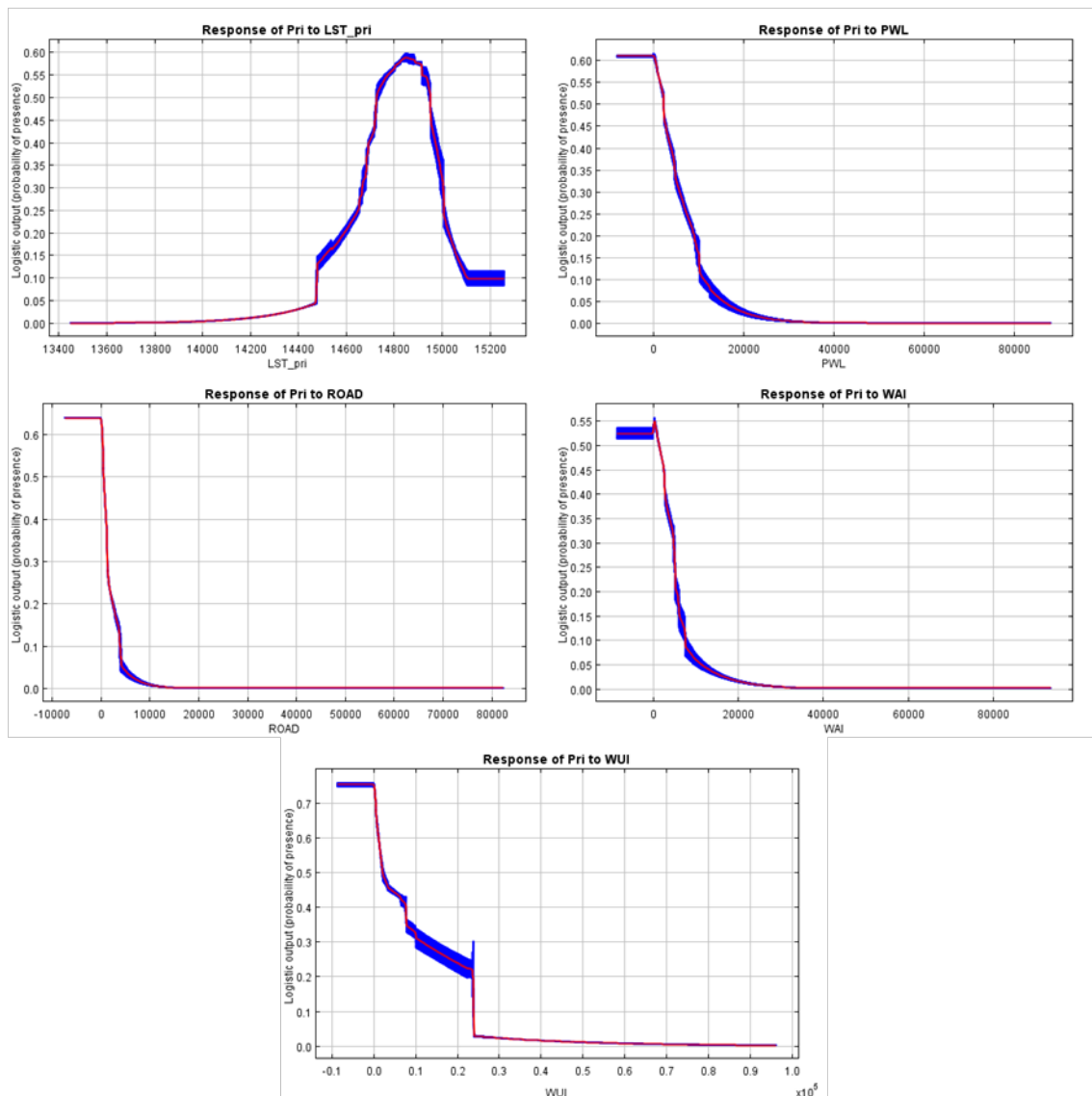


Figura 25. Curvas de respuesta de variables predictoras de Pri.

El caso de los fines de semana y festivos de primavera es distinto a cualquier otro modelo generado por la ausencia de la variable MODIS en la predicción. En este caso es la variable carreteras la que tiene un mayor peso explicativo, aunque se reparte prácticamente al 50% la influencia en la predicción final con la interfase WAI. Este hecho, recalca de nuevo la mayor actividad agraria en este periodo del año, y destaca como los fines de semana y festivo la importancia de estos trabajos aumenta, consecuencia probable del alza de la agricultura a tiempo parcial, que encuentra en estos días no laborables el momento de trabajo en el campo.

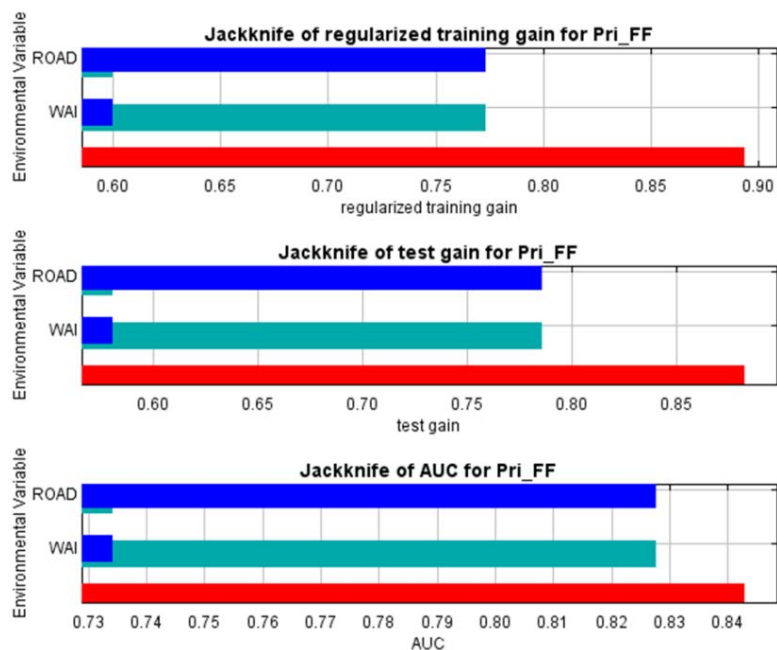


Figura 26. Gráficos *jackknife* de *Pri_FF*. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 12. Porcentajes de contribución y permutación de variables predictoras de *Pri_FF*.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	53,2	61,9
WAI	46,8	38,1

Las curvas de respuesta de ambas variables son muy similares, marcadas por el descenso brusco y sin discontinuidades desde los valores 0 que se relacionan con el mayor peligro en la predicción generada.

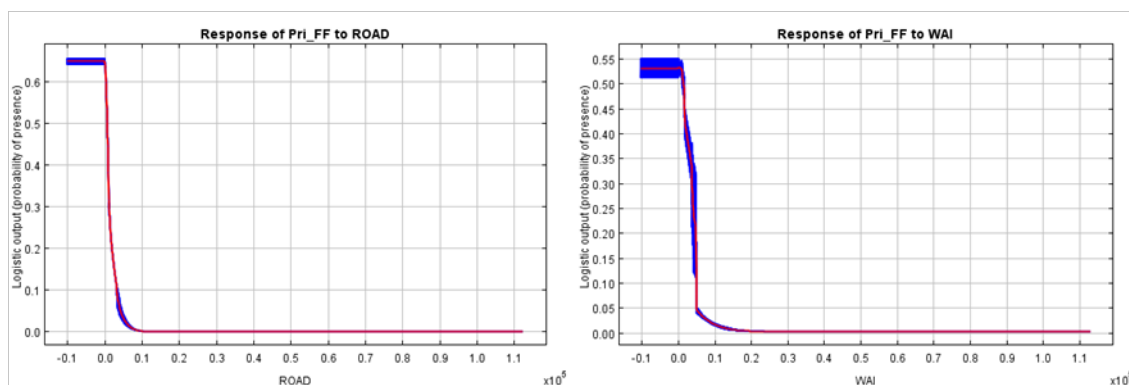


Figura 27. Curvas de respuesta de variables predictoras de *Pri_FF*

5.3.3.3. Verano

El verano tiene una particularidad importante respecto al resto de escenarios, la entrada de los caminos como variable explicativa, y es que es en este periodo en el que la presión humana es mayor sobre las zonas forestales, con miles de excursionistas cada día en estos espacios. Pese a ello, la importancia de esta variable como factor explicativo es baja, siendo de nuevo las carreteras la variable con mayor capacidad predictiva. De nuevo, como lo era en los meses de primavera, la

importancia de la interfase WAI se relaciona con una actividad elevada en los espacios agrícolas, rozando el 29% de la capacidad explicativa del modelo final y con una fuerte incidencia si sus valores son sustituidos aleatoriamente (bajo porcentaje de permutación). Por último, la facilidad de ignición en estos meses estivales puede ser la causa del aumento de los fuegos consecuencia de las líneas y tendidos eléctricos y de su correspondiente peso en el modelo de verano.

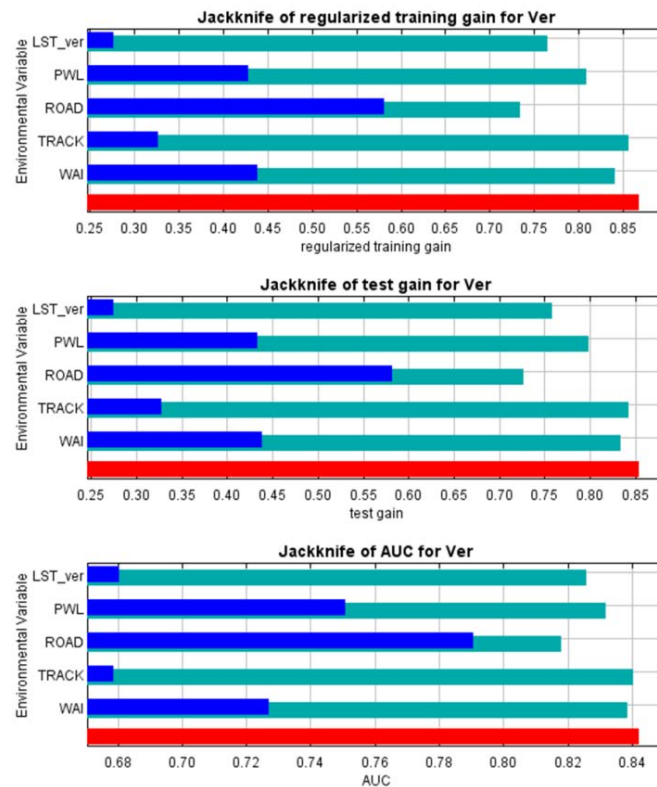


Figura 28. Gráficos *jackknife* de Ver. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 13. Porcentajes de contribución y permutación de variables predictoras de Ver.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	37,4	22,1
WAI	28,9	17,1
LST_ver	18,1	30,3
PWL	13,3	16,3
TRACK	2,3	14,3

Las curvas de respuesta no tienen comportamiento distinto al esperado y al ya mostrado en los modelos presentados anteriormente. La probabilidad de origen de un incendio por causa humana decae rápidamente conforme nos alejamos de las causas principales de incendios: los trabajos agrícolas (variable WAI), las carreteras, las líneas eléctricas y los caminos. Aumenta con la temperatura exponencialmente con la temperatura de superficie, hasta que la frecuencia de los valores de la misma es responsable del descenso de la curva.

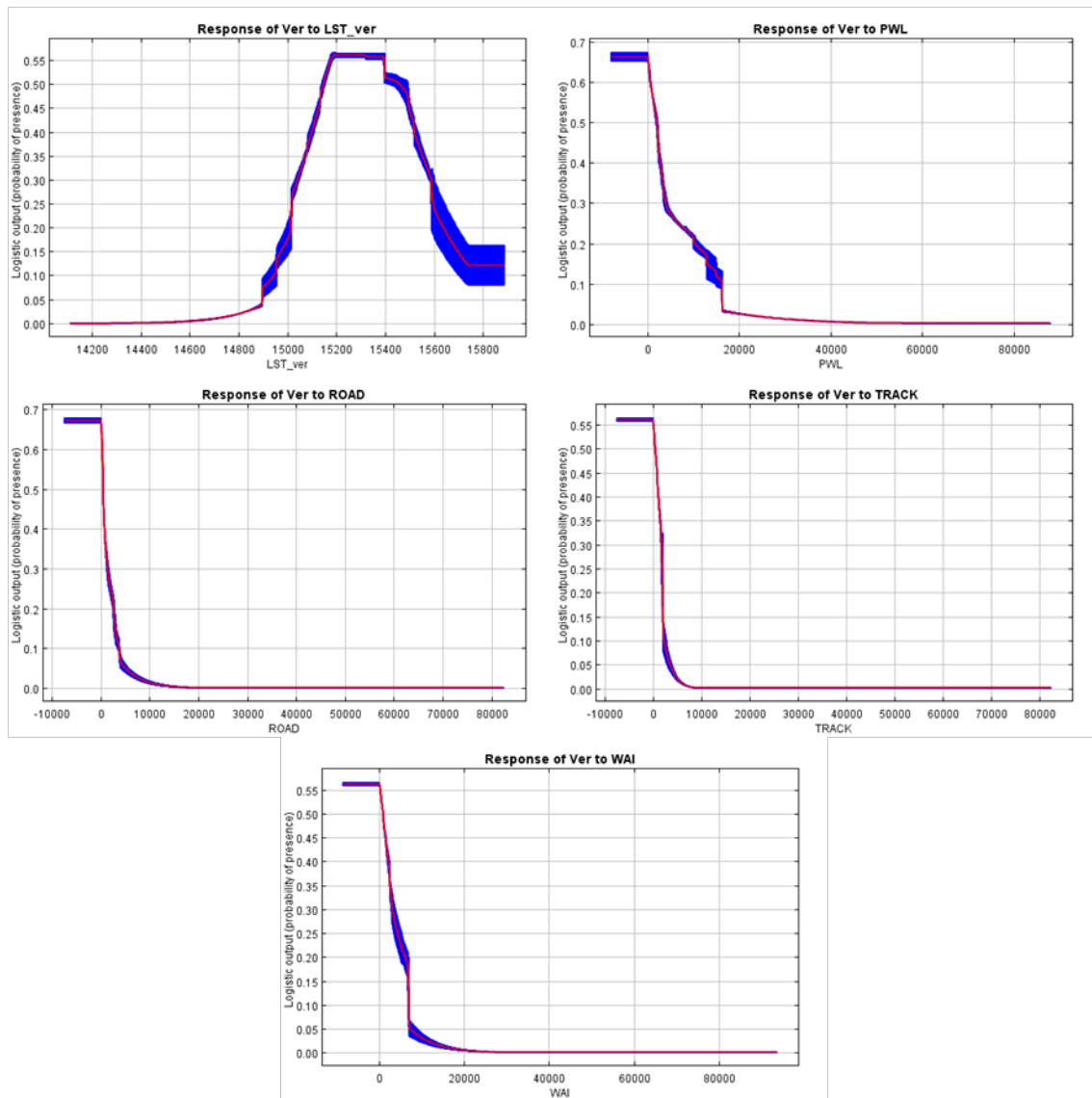


Figura 29. Curvas de respuesta de variables predictoras de Ver.

El modelo que recoge los festivos y fines de semana del verano presenta un comportamiento de difícil interpretación. La aparición de la interfase WUI como factor explicativo importante (17,1%) puede tener explicación en el mayor movimiento en el entorno de zonas pobladas y segundas residencias durante el fin de semana. Por otra parte, el peso de la variable PWL asciende notablemente respecto a los días laborables, lo que no tiene una posible causa aparente. Pese a estos cambios, la variable independiente fundamental continúa siendo las carreteras, en este caso de forma muy destacada al resto de variables, y es que los fines de semana y festivos en verano son de las fechas con mayor volumen de desplazamientos, lo que maximiza las probabilidades de accidente o negligencia. De nuevo los caminos aparecen como factor explicativo aunque de nuevo con un porcentaje marginal. Si bien confirma que en este periodo los excursionistas son una fuente de incendios importante.

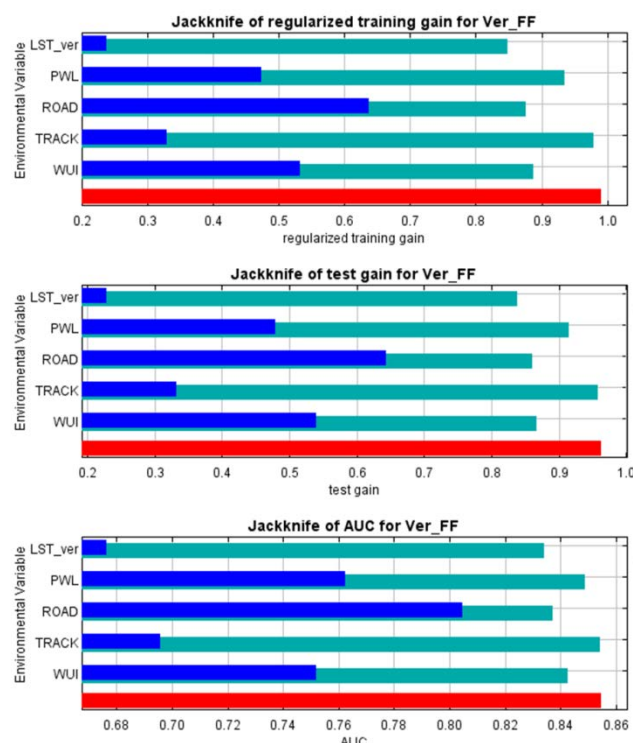


Figura 30. Gráficos *jackknife* de Ver_FF. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 14. Porcentajes de contribución y permutación de variables predictoras de Ver_FF.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	46,7	43,3
PWL	17,3	15,9
WUI	17,1	7,9
LST_ver	16,9	27,4
TRACK	2	5,6

Las curvas tienen de nuevo un comportamiento normal, aunque con alguna discontinuidad observable en las líneas eléctricas y la interfase WUI. Cabe destacar como la variable WUI es la que mayor capacidad explicativa tiene en las probabilidades altas (ver los valores del eje de ordenadas). Así, el peligro se relaciona negativamente con todas las variables a excepción de la temperatura de superficie, donde a mayor valor de esta el peligro aumenta.

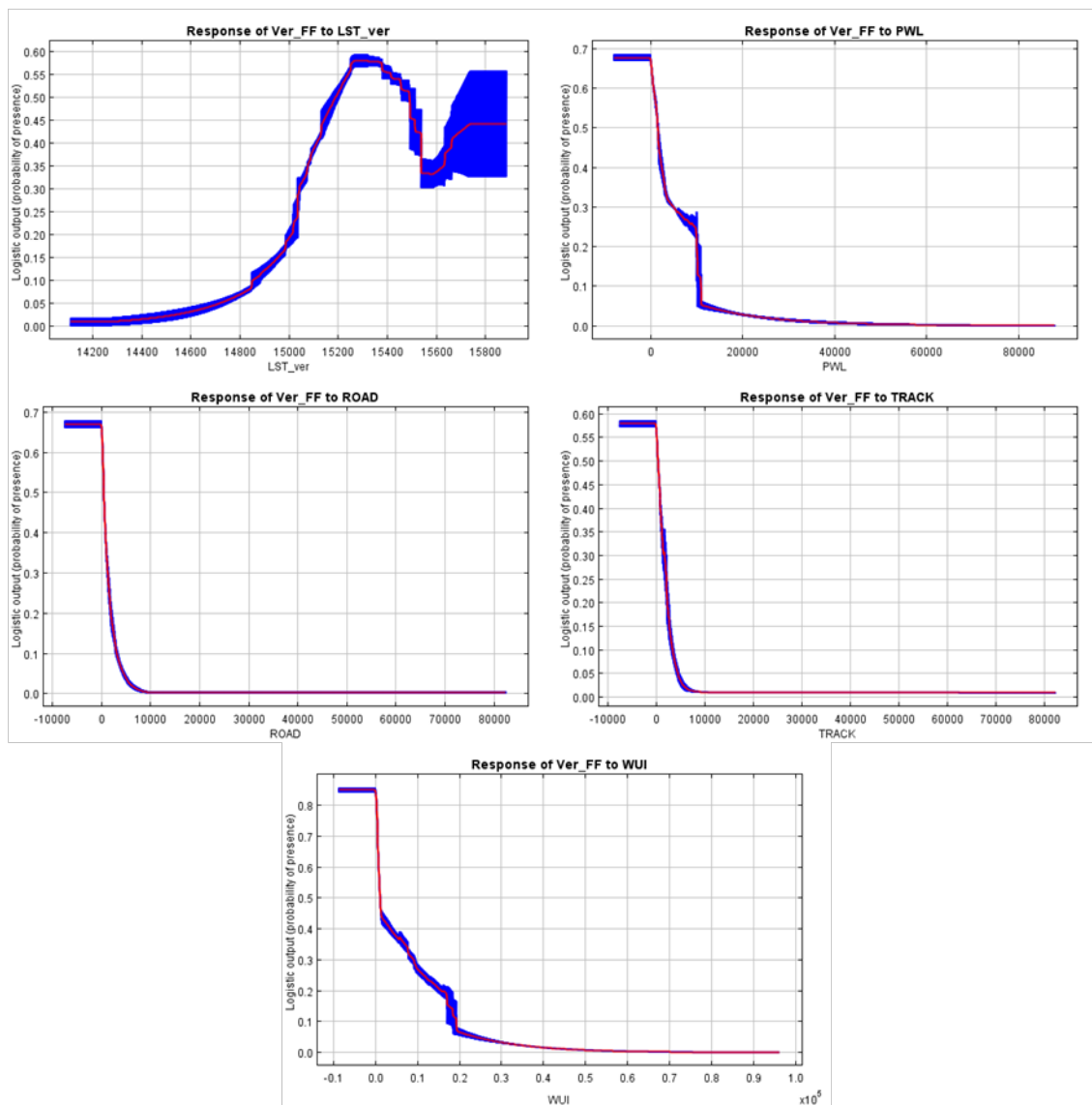


Figura 31. Curvas de respuesta de variables predictoras de Ver_FF.

5.3.3.4. Otoño

En el modelo de otoño, el número de variables explicativas vuelve a reducirse considerablemente, entrando como predictoras únicamente el componente de MODIS, las carreteras y la interfase agrícola. Es en este periodo, y especialmente durante fines de semana y festivos, cuando el tamaño muestral puede jugar un efecto más decisivo, puesto que el último trimestre del año es el que menor cantidad de incendios forestales presenta (ver figura 10). Las carreteras, tal y como se puede en el último de los *jackknife*, es la variable clave, pudiendo llegar a superar un 0,78 en el valor de AUC por si misma (sin la colaboración de otras variables). En este sentido, la interfase agrícola es la primera que se podría prescindir en caso de ser necesario, ya que de acuerdo a lo observado en la figura 32 presenta el menor decrecimiento del valor total de AUC en su ausencia.

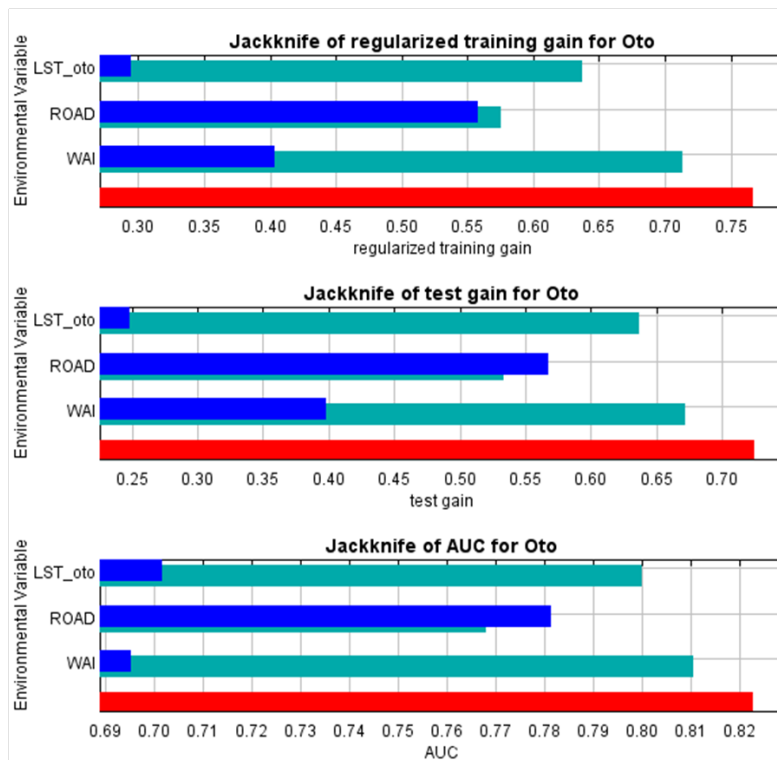


Figura 32. Gráficos *jackknife* de Oto. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 15. Porcentajes de contribución y permutación de variables predictoras de Oto.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	46	54,9
WAI	31,5	23,9
LST_oto	22,5	21,2

La curva de respuesta de la variable MODIS, en esta ocasión, no presenta el descenso tras alcanzar un valor máximo que si se presentaba anteriormente (ver figuras 29 y 31), creciendo de forma continuada hasta estabilizarse en un valor de 15000. Además, de las variables que forman parte del modelo, es la que explica los valores más altos de probabilidad, atendiendo a los valores del eje de ordenadas. Las carreteras tienen en la curva marcada por el descenso brusco, del mismo modo que en todos los escenarios ya analizados. Por último, la interfase WAI, aunque presentando un perfil algo quebrado, responde a la lógica de disminuir el peligro conforme se gana distancia con el contacto entre los espacios agrícolas y la vegetación natural.

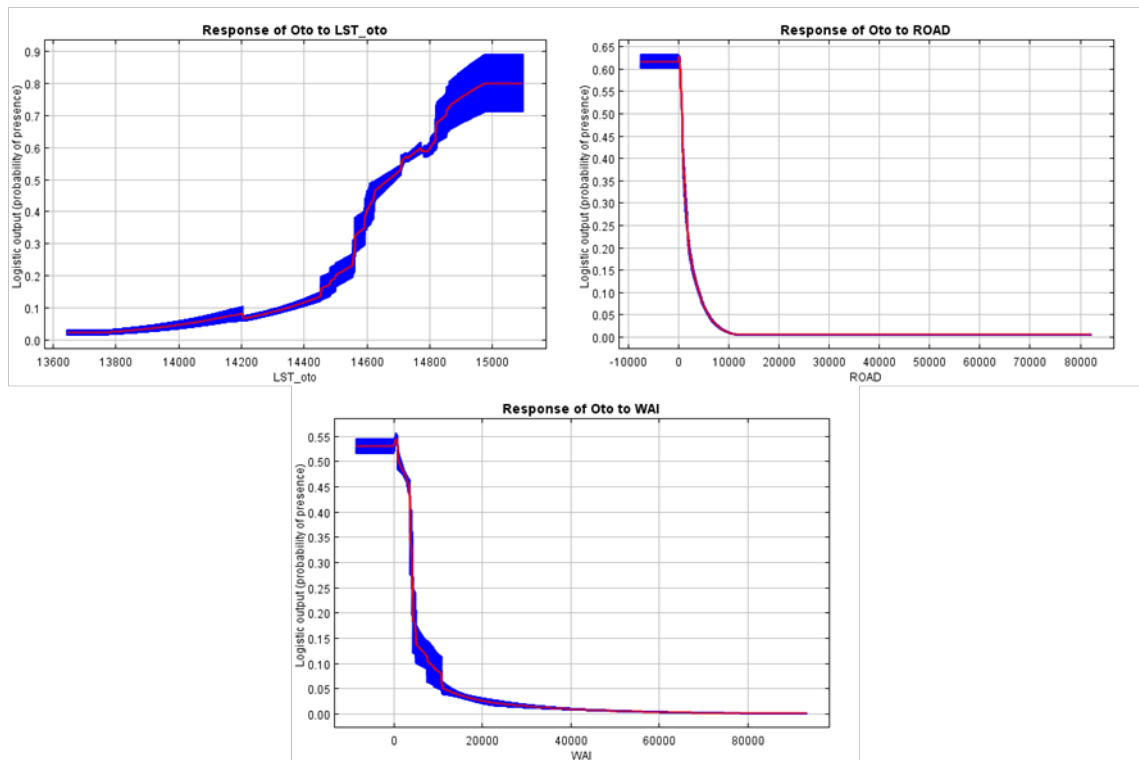


Figura 33. Curvas de respuesta de variables predictoras de Oto.

El último de los modelos, el de los fines de semana y festivos de otoño, únicamente se explica por la acción de dos variables, con una predominancia absoluta de las carreteras como factor explicativo (86%) sobre el producto MODIS. Las curvas de comportamiento responden al mismo esquema que el modelo de otoño, sin observar cambios significativos que merezcan un comentario diferente.

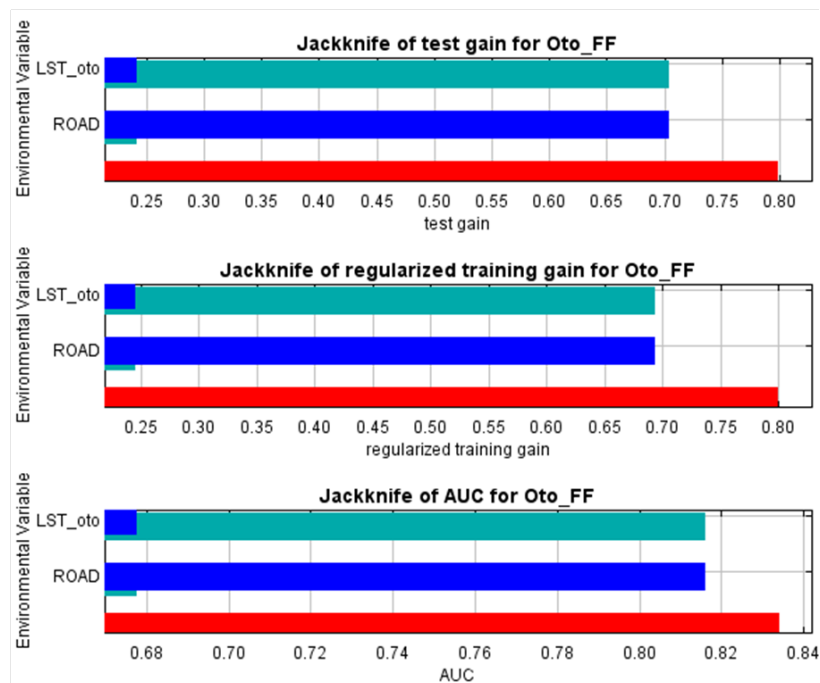


Figura 34. Gráficos *jackknife* de Oto_FF. Turquesa: Sin la variable, azul: sólo con la variable, rojo: con todas las variables.

Tabla 16. Porcentajes de contribución y permutación de variables predictoras de Oto_FF.

Variable	Porcentaje de contribución	Porcentaje de permutación
ROAD	86	81,6
LST_oto	14	18,4

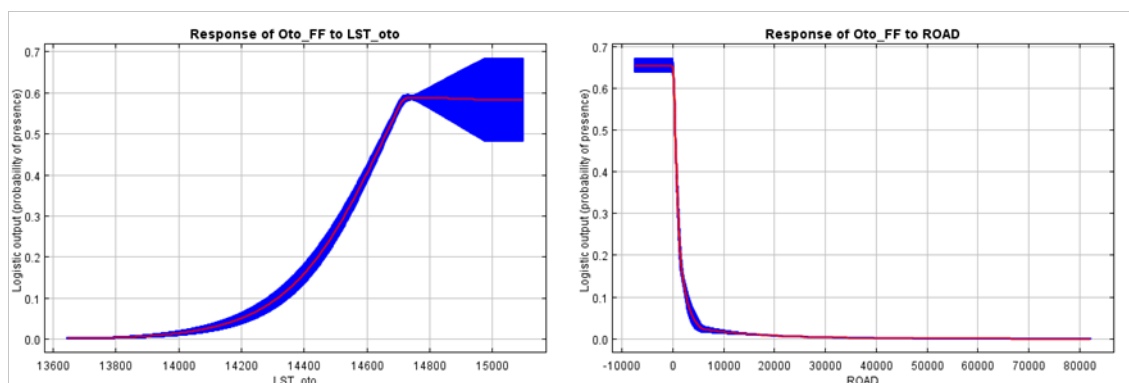


Figura 35. Curvas de respuesta de variables predictoras de Oto_FF.

5.4. Validación operativa. Predicción del peligro de incendio para 2012.

En este apartado se procede a validar operativamente los modelos creados en MaxEnt, es decir, a analizar los resultados obtenidos en base a los incendios acontecidos en 2012, tratando de dilucidar si los modelos logran predecir con precisión el peligro de incendio. Para ello, primeramente, se realiza una caracterización del comportamiento climático y de la distribución de incendios forestales de 2012.

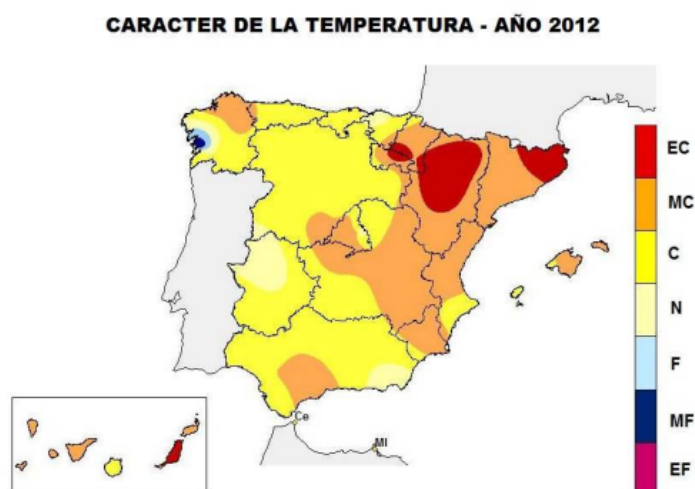
Esta validación se ha desarrollado a partir de rutinas programadas en R, a partir del mismo algoritmo de predicción (MaxEnt), utilizando datos del producto MODIS MOD11A2 descargados y post-procesados (mosaicado, reproyección y recorte) para el año 2012. Concretamente se han ajustado modelos diarios utilizando los modelos estacionales, sustituyendo la variable LST por el producto MODIS correspondiente a la fecha analizada.

5.4.1. Caracterización del año 2012.

Para poder comprender y contextualizar adecuadamente los resultados de la validación operativa realizada para el año 2012 es indispensable conocer en contexto climático de dicho año, así como caracterizar el comportamiento y distribución espacio-temporal de los incendios registrados.

Comenzando por el aspecto climático, el año 2012 presentó un comportamiento térmico anómalo muy importante en gran parte del área de estudio (figura 36). En concreto, todas las provincias objeto de estudio registraron una temperatura media, como mínimo, muy cálida, clasificándose dentro del 20% de años más cálidos en la serie histórica de referencia 1971-2000. Destaca el centro de las provincias de Huesca y Zaragoza, para los que 2012 constituyó el año más cálido tomando como referencia el periodo antes citado. Las grandes anomalías se concentraron en los meses más cálidos, especialmente en mayo, junio y agosto, donde las temperaturas llegaron a superar por más de 2°C los valores medios habituales. Julio y septiembre, sin tener un comportamiento extraordinario, estuvieron por encima de los valores

medios, del mismo modo enero, marzo, abril, octubre, noviembre y diciembre. Únicamente febrero tuvo una media por debajo de lo normal para este mes.



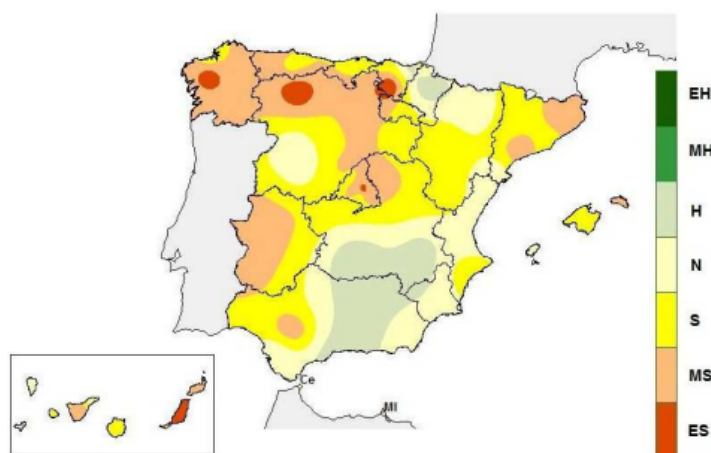
FUENTE: Agencia Estatal de Meteorología. Ministerio de Agricultura, Alimentación y Medio Ambiente.

EC =Extremadamente Cálido: Las temperaturas sobrepasan el valor máximo registrado en el periodo de referencia 1971 – 2000.
MC =Muy cálido: $f < 20\%$. Las temperaturas registradas se encuentran en el intervalo correspondiente al 20% de los años más cálidos.
C =Cálido: $20\% \leq f < 40\%$.
N =Normal: $40\% \leq f \leq 60\%$. Las temperaturas registradas se sitúan alrededor de la mediana.
F =Frio: $60\% \leq f < 80\%$.
MF =Muy Frio: $f \geq 80\%$.
EF =Extremadamente frío: Las temperaturas no alcanzan el valor mínimo registrado en el periodo de referencia 1971 – 2000

Figura 36. Carácter de la temperatura en 2012. Fuente: AEMET (2013)

Por otra parte, la pluviometría de 2012 estuvo por debajo de lo normal en gran parte del área de estudio, a excepción de la provincia de Huesca, donde la precipitación registrada se mantuvo dentro de los valores normales. En buena parte de la provincia de Tarragona, 2012 puede calificarse como muy seco, situándose entre el 20% de años más secos en el periodo de referencia 1971-2000. En la distribución mensual de la precipitación vuelve a destacar que los valores más bajos respecto a la media se concentraron en el semestre más cálido, con especial incidencia en agosto, que registro valores históricos de baja precipitación en numerosos observatorios. También enero y febrero estuvieron por debajo de la media de precipitaciones. Sólo abril y los meses otoñales registraron precipitaciones superiores a la media.

CARACTER DE LA PRECIPITACIÓN - AÑO 2012



FUENTE: Agencia Estatal de Meteorología. Ministerio de Agricultura, Alimentación y Medio Ambiente.

EH =Extremadamente húmedo: Las precipitaciones sobrepasan el valor máximo registrado en el periodo de referencia 1971–2000.
 MH = Muy húmedo: $f < 20\%$. Las precipitaciones se encuentran en el intervalo correspondiente al 20% de los años más húmedos.
 H = Húmedo: $20\% \leq f < 40\%$.
 N = Normal: $40\% \leq f < 60\%$. Las precipitaciones registradas se sitúan alrededor de la mediana.
 S = Seco: $60\% \leq f < 80$
 MS = Muy seco: $f \geq 80\%$.
 ES =Extremadamente seco: Las precipitaciones no alcanzan el valor mínimo registrado en el periodo de referencia 1971–2000.

En cuanto al comportamiento de los incendios de 2012 se puede destacar que fue un año muy por encima de la media en cuanto a número de incendios (644 incendios de media en el periodo 1988-2011 y 645 en 2008-2011). Se registraron 970 fuegos de origen humano a lo largo del año, constituyéndose como el tercer año con más incendios de este tipo en la serie 1988-2011 (ver figura 8).

Distribución mensual de incendios (causa humana) en 2012

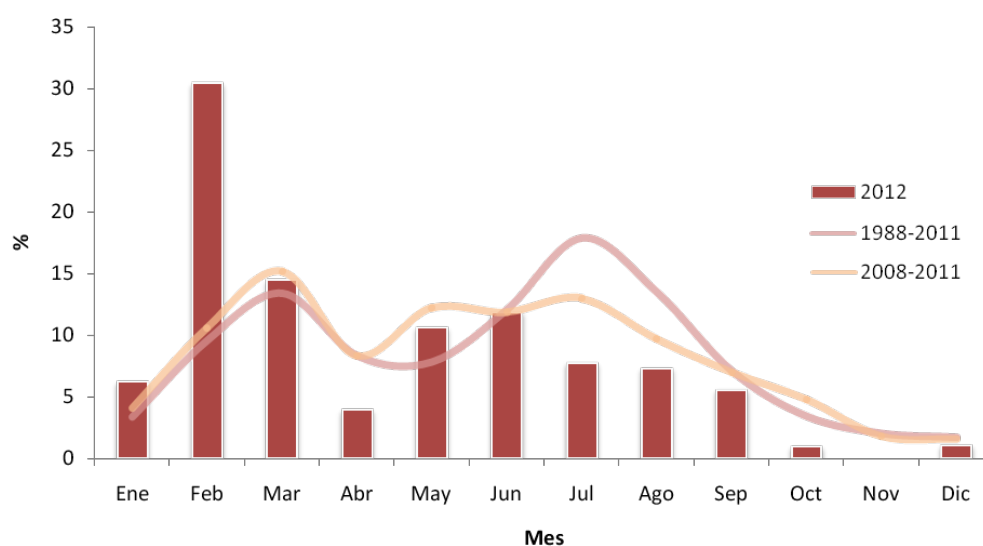


Figura 37. Número total de incendios por causa humana en 2012 por mes.

La figura 37 muestra un comportamiento atípico en la distribución de incendios en 2012. Febrero (curiosamente el único mes con un comportamiento térmico por debajo de la media, lo que evidencia la importancia del factor humano) copó el 30% de los incendios de todo el año, provocando que 2012 tuviera en conjunto una

cantidad de incendios extraordinaria para la serie de referencia. La distribución de incendios en el resto de los meses se ajusta algo más a lo esperado, aunque destaca el número de incendios de mayo y junio, que supera a los registrados en los dos meses *a priori* más conflictivos (julio y agosto). La contribución del otoño al total de los incendios fue muy baja, quedándose en 0 en el mes de noviembre. La distribución semanal si presenta unos valores en consonancia con los recogidos en el apartado 5.1.1. Se identifica una concentración en el fin de semana superior en un 8% a una distribución homogénea ideal.

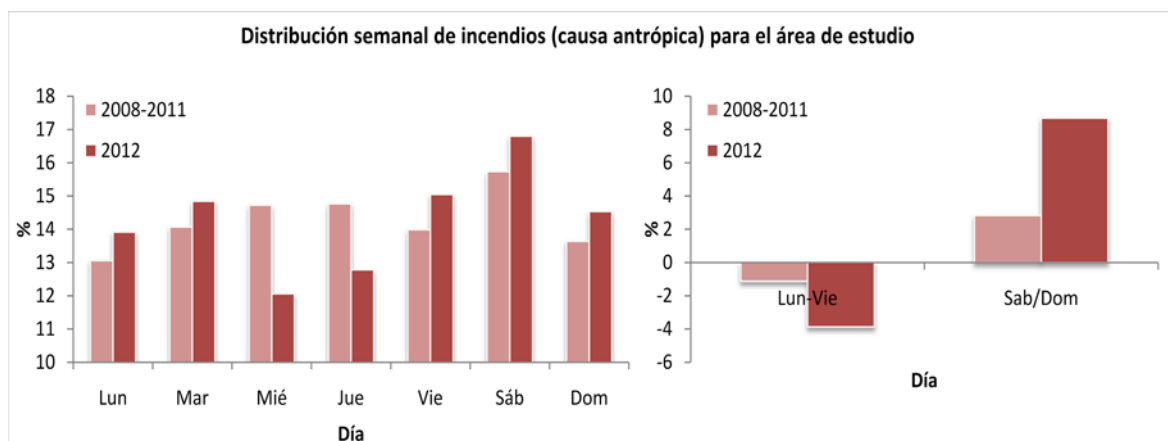


Figura 38. Izquierda: distribución semanal de incendios (causa antrópica) en 2012. Derecha: porcentaje respecto a una distribución homogénea ideal de 2012.

Por último, y con el fin de excluir cualquier posibilidad de que la validación pudiera estar condicionada por la distribución espacial de los incendios (si existiera una clara concentración en una zona determinada) se presenta el mapa de distribución de los incendios de 2012 y de los años 2008-2011. Como se puede observar, el patrón espacial de los incendios de 2012 es similar al registrado en el periodo 2008-2011.

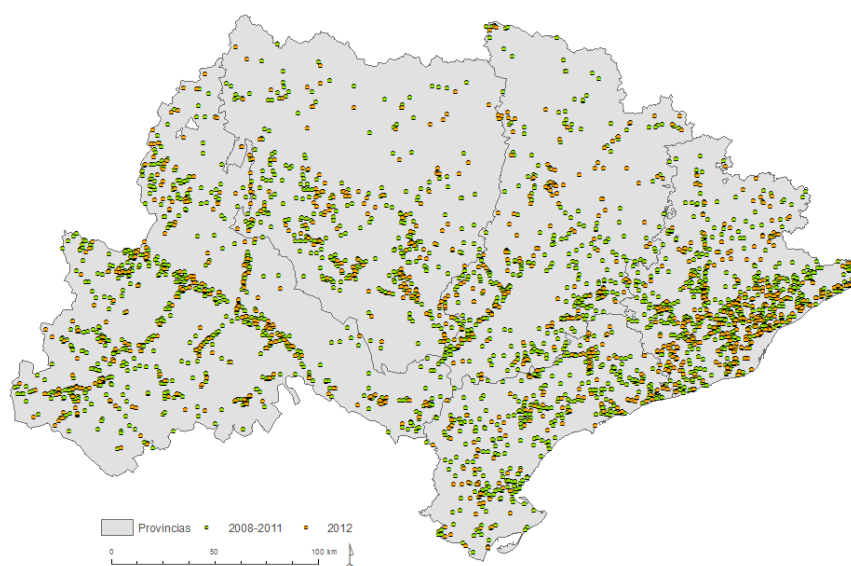


Figura 39. Distribución espacial de los incendios de 2012 y 2008-2011.

5.4.2. Predicción del peligro para 2012 y evaluación estadística de los resultados.

La figura 41 presenta la cartografía de peligro de incendio para determinados días de cada mes junto con la localización de los incendios ocurridos en esas fechas. Los días de marzo, junio y noviembre se corresponden con modelos de fin de semana y/o festivo. Se puede observar como la distribución espacial del riesgo varía en cada fecha, fruto de la temporalidad que introduce el producto MODIS y del modelo (de los ocho generados) que determina el peligro en función de la fecha. En la siguiente *url* se puede ver una animación la previsión del peligro para cada día del año 2012:

http://155.210.62.105/Riesgo_2012/

Destaca una posible incoherencia presente en los primeros días de octubre debido a una anómala interacción entre el producto MODIS y la interfase cultivo-forestal (WAI) del modelo de octubre, originada por un periodo de temperaturas inusualmente cálidas durante agosto y parte de septiembre de 2012 (figura 40).

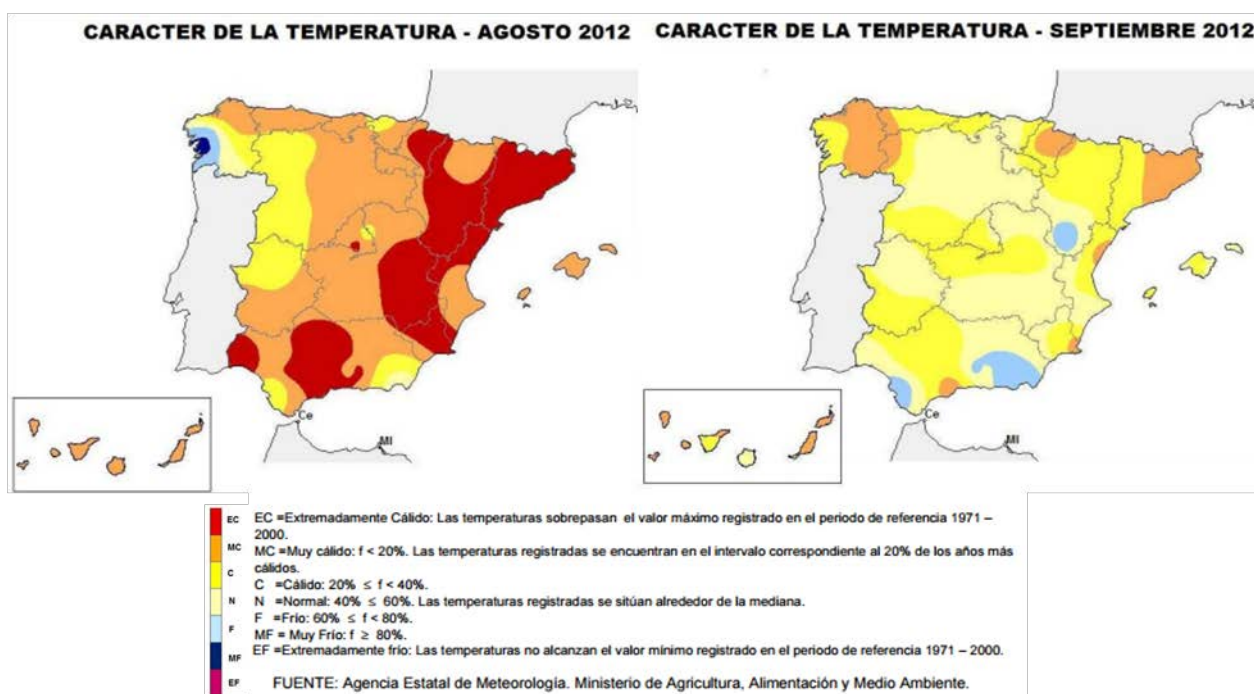


Figura 40.Comportamiento térmico de agosto y septiembre de 2012. Fuente: AEMET (2012a, 2012b)

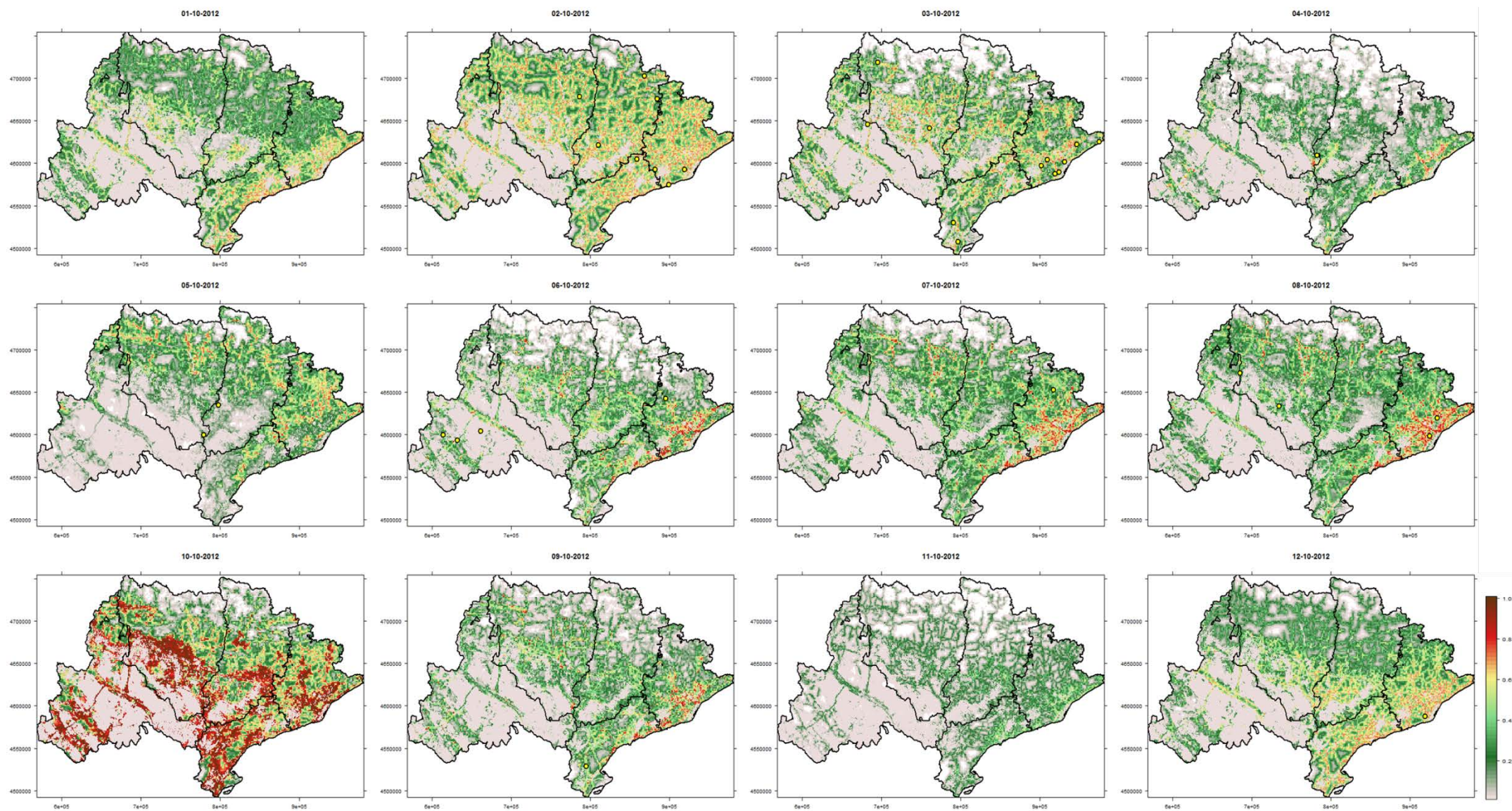


Figura 41. Mapas de peligro e incendios para días concretos de 2012. Los círculos amarillos representan la ocurrencia de un incendio.

Visualmente se puede apreciar que, salvo alguna excepción ya comentada, la predicción se comporta de forma normal y consistente. Se presenta a continuación la validación de la predicción del peligro para el año 2012, que se ha llevado a cabo mediante la extracción de la predicción diaria para cada uno de los incendios registrados en dicho año. Para facilitar la interpretación, los resultados obtenidos se muestran en forma de diagramas de cajas en los que se puede observar el rango y distribución de los valores de probabilidad de ocurrencia (peligro) agregados a distintas escalas temporales y por causalidad.

En la figura 42, se presenta un *boxplot* mensual que demuestra que la mayor parte de los incendios (cajas verdes) se sitúa en probabilidades de peligro entre el 0,4 y el 0,6, siendo algo superiores si consideramos el máximo de la vecindad de 250 m que podría ser un valor más acorde en algunos casos por la resolución de los modelos. Los valores presentan una cierta homogeneidad apreciable por los valores de mediana, con la única excepción de mayo, lo que habla de que podría ser el mes que peor quedara caracterizado por los modelos estacionales. Noviembre no ha podido ser evaluado puesto que no se registro ningún incendio en el área de estudio en ese año.

En cuanto a los valores estacionales, recogidos en la figura 43, muestran también valores de mediana próximos al nivel de peligro de 0,5, siendo la primavera, condicionada por los bajos valores de mayo, de nuevo la que presenta unos resultados más bajos.

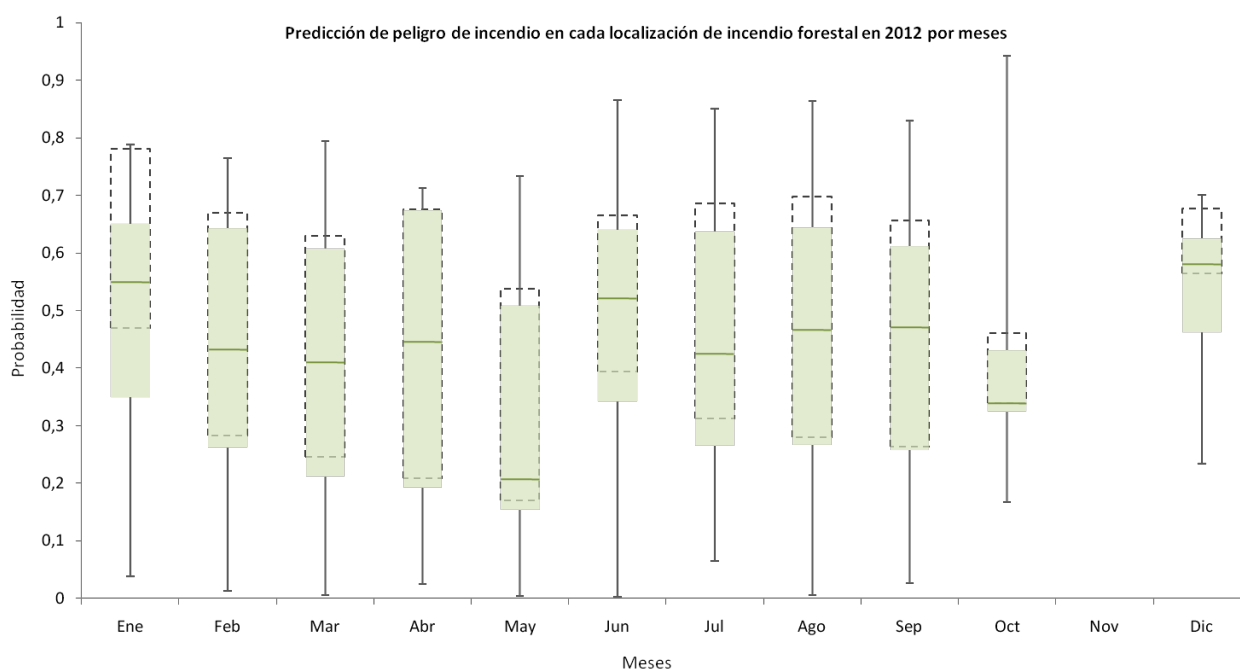


Figura 42. *Boxplot* de peligro de incendio en cada localización de incendio forestal en 2012 por meses. La caja verde representa la distribución entre el segundo y el tercer cuartil. La línea horizontal se corresponde con la mediana mientras que la caja discontinua representa los valores máximos de riesgo en una vecindad de 250m respecto al punto de incendio. Los "bigotes" de las cajas muestran el máximo y el mínimo de la serie.

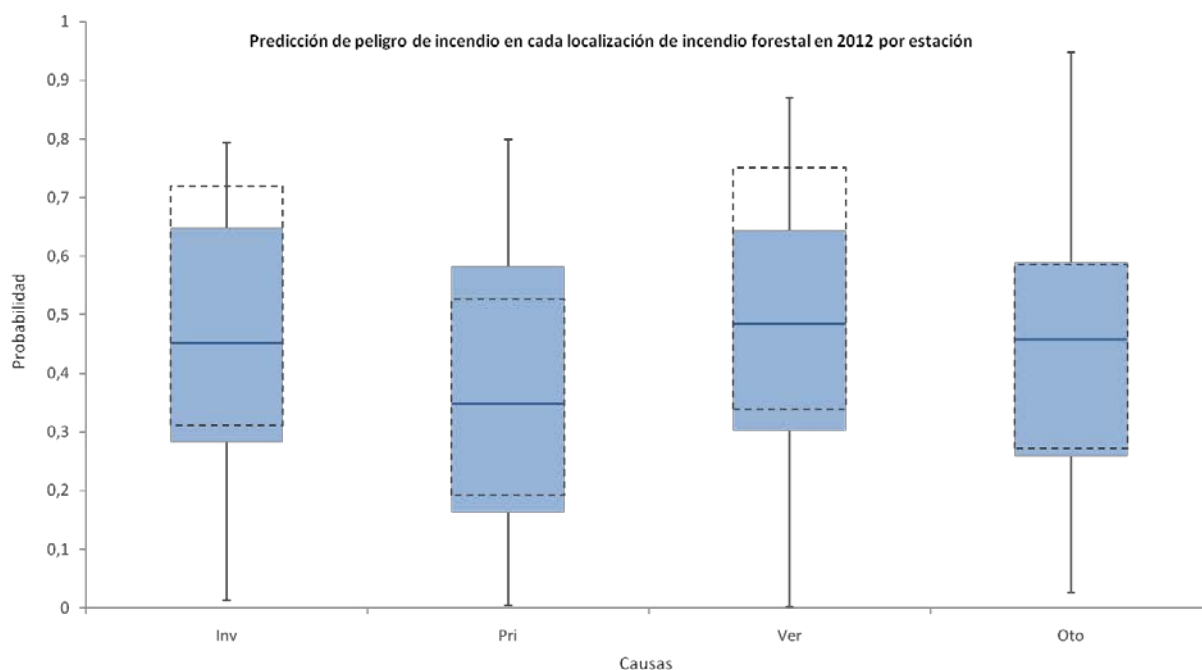


Figura 43. Boxplot de peligro de incendio en cada localización de incendio forestal en 2012 por estación. La caja azul representa la distribución entre el segundo y el tercer cuartil. La línea horizontal se corresponde con la mediana mientras que la caja discontinua representa los valores máximos de riesgo en una vecindad de 250m respecto al punto de incendio. Los "bigotes" de las cajas muestran el máximo y el mínimo de la serie.

A partir de las figuras 42 y 43, se puede concluir que, a pesar del atípico comportamiento climático del año 2012 y de su inusual distribución intraanual de la ocurrencia de incendios, los modelos consiguen predecir relativamente bien dicha ocurrencia, alcanzando un umbral de peligro entorno a un valor de probabilidad de 0,4 o superior.

Por último, se ha querido mostrar la relación entre la probabilidad de incendio predicha y los grupos de causalidad de los incendios. La figura 44 recoge esta distribución, en la que podemos observar que los incendios intencionados son aquellos que mejor predicen los modelos, acercándose su mediana al valor 0,5. Por el contrario, los incendios por causa accidental son los que muestran valores de probabilidad más bajos, aunque no resultan distantes a los valores de los tres grupos restantes, que varían mayoritariamente entre valores de 0,3 y 0,6, con medianas próximas a 0,5.

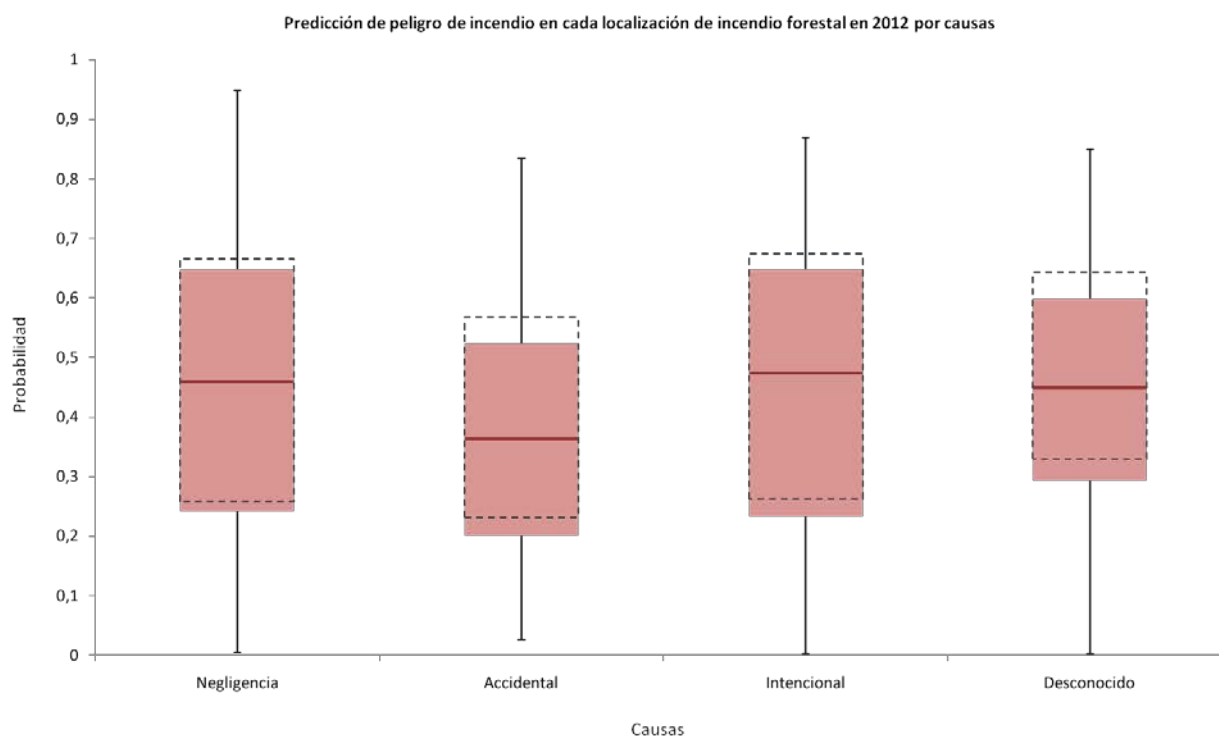


Figura 44. Boxplot de peligro de incendio en cada localización de incendio forestal en 2012 por causas. La caja color salmón representa la distribución entre el segundo y el tercer cuartil. La línea horizontal se corresponde con la mediana mientras que la caja discontinua representa los valores máximos de riesgo en una vecindad de 250m respecto al punto de incendio. Los "bigotes" de las cajas muestran el máximo y el mínimo de la serie.

6. Conclusiones y líneas de trabajo futuras

En este apartado final se recogen las principales conclusiones de este Trabajo Fin de Máster así como posibles vías de continuación de este estudio que pueden desarrollarse en un futuro próximo.

6.1. Conclusiones

La consideración de la variable temporal en los modelos de predicción del riesgo de ignición humano supone un paso adelante en el intento por optimizar la calidad y precisión de los mismos, integrando los ciclos temporales que rigen la actividad humana dentro del análisis del riesgo de incendio, superando así la modelización que únicamente trabaja con la temporalidad de los factores ambientales.

La metodología desarrollada en este trabajo ha permitido evaluar la conveniencia de diferentes escenarios temporales y la creación de modelos de peligro basados en éstos, logrando predecir el peligro de incendio humano para 2012 con notable consistencia. La utilización de scripts ha permitido el trabajo con grandes volúmenes de datos que requiere un estudio de estas características.

Se exponen a continuación de forma sintética las conclusiones extraídas de esta investigación y que dan respuesta a los objetivos planteados en su inicio:

- La distribución interanual de los incendios derivados de la acción humana, así como la intraanual, está sujeta a una acusada variabilidad temporal. Dentro de esta heterogeneidad se puede identificar una distribución mensual bimodal con máximos en marzo y julio-agosto y mínimos durante el otoño. La

distribución de los incendios a escala diaria demuestra los distintos ciclos de actividad humana, con una concentración de los fuegos muy superior en días festivos y de fin de semana con respecto a los días laborables.

- La desagregación temporal escogida -estacional- ha sido el máximo grado de detalle posible en función de lo permitido por el conjunto de datos empleado en el estudio. Este hecho se comprobó en los primeros análisis exploratorios mediante regresión logística, descartando así la agregación mensual. No obstante, se han obtenido diferencias en los factores explicativos. También se ha logrado una diferenciación entre días "laborables" y días "festivos" (sábados, domingos y festividades).
- Las variables independientes que han resultado ser significativas según el análisis exploratorio y, por tanto, se han incluido en los modelos predictivos han sido la interfase cultivo-forestal (WAI), la interfase urbano-forestal (WUI), las líneas y tendidos eléctricos (PWL), la red de caminos (TRACKS), las carreteras (ROADS) y el producto MODIS de temperatura de superficie y emisividad (MOD_LST).
- El algoritmo de Máxima Entropía (MaxEnt) ha permitido generar los ocho modelos que responden a los escenarios temporales seleccionados con una calidad notable, tal y como se extrae de los valores AUC derivados de la validación cruzada (k-fold), cercanos o superiores en todos los casos a 0,8.
- La distancia a carreteras (ROADS) es la variable con mayor peso explicativo en los modelos generados, siendo fundamental en la predicción del peligro de ignición en todos los escenarios.
- La variable MODIS condiciona la posibilidad de ignición, teniendo un peso importante en los modelos, lo que es especialmente patente al prescindir de ella en la modelización, lo que ocasiona un descenso importante en la calidad de los modelos.
- Las dos interfases presentan comportamientos lógicos y colaboran en la generación de los modelos en distintos escenarios temporales.
- La variable independiente TRACKS presenta influencia durante el verano, cuando estos caminos son más utilizados por excursionistas. Por último, la variable PWL, aunque significativa, tiene un comportamiento de difícil interpretación.
- La predicción del peligro para 2012 ha revelado la buena precisión de los modelos a pesar de las anomalías térmicas, pluviométricas y de distribución intraanual de los incendios de dicho año. La mediana del valor de probabilidad del conjunto de los incendios se acerca a los 0,45, superándose en algunos meses. Mayo es el único mes que presenta unos valores de probabilidad bajos, lo que puede indicar que los modelos no describen bien el comportamiento de este mes, o al menos no lo hacen para los incendios de 2012.

6.2. Potenciales líneas de desarrollo del trabajo

Este trabajo abre el camino hacia otra serie de investigaciones que pueden mejorar sus resultados y profundizar en el método aquí propuesto. El tamaño muestral empleado, tanto en extensión temporal como geográfica, puede ser ampliado, lo que contribuiría a la consistencia a los resultados obtenidos y al mismo tiempo abriría la posibilidad de nuevos escenarios temporales más desagregados (escala mensual).

Otro aspecto que puede ser interesante mejorar es la resolución espacial de los modelos, aprovechando precisión que permite la localización mediante GPS de los incendios y la precisión espacial de las variables predictoras derivadas de cartografía a escala 1:25.000. Los problemas a salvar en este caso particular son la baja resolución espacial de imágenes satelitales cuando su resolución temporal es alta (el caso de MODIS) y el gran volumen de cálculo de los algoritmos empleados en este trabajo, que implicaría quizás utilizar métodos de paralelización del cálculo.

Atendiendo a las variables predictivas empleadas, esta investigación ha espacializado las mismas mediante la distancia a cada elemento (vías de comunicación, líneas eléctricas etc.), mientras que en otros trabajos (Vilar *et al.*, 2011; Rodrigues *et al.* 2014) se ha optado por emplear un *buffer* que recogiera la distancia a la que se espera que una variable pueda tener influencia, quedando de este modo de forma categórica la relación entre cada incendio y las variables (dentro o fuera del *buffer*). Se cree que esta espacialización favorece los resultados de los modelos al conseguir aprovechar de un mejor modo la resolución espacial de la variable dependiente. Sin embargo, en futuras investigaciones podría resultar de interés combinar ambas formas de espacialización, de forma que sea un *buffer* el que limite el radio de acción de cada variable predictora, pero conservando la distancia específica de cada pixel al elemento en cuestión.

Otro aspecto que puede mejorar la predicción es la inclusión de un índice que mejore a la única variable que introduce factores ambientales (la variable MOD_LST). En los inicios de esta investigación se trató de utilizar el *Fire Weather Index* como variable ambiental, llegando incluso a contactar con la institución encargada de facilitar dicha información. La información disponible no concordaba en tiempo y espacio con la base de datos que se pretendía utilizar, por lo que no se pudo incorporar a la modelización.

7. Bibliografía

- Agencia Estatal de Meteorología (AEMET). 2013. Resumen anual climatológico. Año 2012. Ministerio de Agricultura, Alimentación y Medio Ambiente
- Agencia Estatal de Meteorología (AEMET). 2012a. Informe mensual climatológico. Agosto de 2012. Ministerio de Agricultura, Alimentación y Medio Ambiente
- Agencia Estatal de Meteorología (AEMET). 2012b. Informe mensual climatológico. Septiembre de 2012. Ministerio de Agricultura, Alimentación y Medio Ambiente
- Bar Massada, A., Syphard, A.D., Stewart, S.I., Radeloff, V.C., 2012. Wildfire ignition-distribution modelling: a comparative study in the Huron- Manistee National Forest, Michigan, USA. *Int. J. Wildl. Fire* -. doi:<http://dx.doi.org/10.1071/WF1117> 8
- Ben-Hur, M., 2008. Seal formation effects on soil infiltration and runoff in arid and semiarid regions under rainfall and sprinkler irrigation conditions. In: Zereini, F., Hötzi, H.(Eds.), *Climatic Changes and Water Resources in the Middle East and North Africa*. Environmental Science and Engineering. Springer, Berlin, Heidelberg, pp. 429-452.
- Calvo, J., & Pueyo, A. (2008). *Atlas Nacional de España: Demografía*. Madrid: Centro Nacional de Información Geográfica.
- Chen, F., Yongsheng, D., Shukui, N., Jinlong, Z. 2015. Modeling Forest Lightning Fire Occurrence in the Daxinganling Mountains of Northeastern China with MAXENT. *Forests* 2015, 6, 1422-1438; doi:10.3390/f6051422
- Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R., Herrera, M.A., Zamora, R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecol. Modell.* 221, 46-58. doi:10.1016/j.ecolmodel.2008.11.017
- Chuvieco, E., Aguado, I., Jurdao, S., Pettinari, M.L., Yebra, M., Salas, J., Hantson, S., de la Riva, J., Ibarra, P., Rodrigues, M., Echeverría, M., Azqueta, D., Román, M. V, Bastarrika, A., Martínez, S., Recondo, C., Zapico, E., Martínez-Vega, F.J. 2012. Integrating geospatial information into fire risk assessment. *Int. J. Wildl. Fire*. doi:<http://dx.doi.org/10.1071/WF12052>
- Cruz, J., Fajardo, A., Gil, B., Gutiérrez, F. Jañez, B., Toledo, F. y Vera, A.L. 1985. La agricultura a tiempo parcial como sistema de vida en el medio rural andaluz. En Etxezarreta, M (Dir.), *La Agricultura Insuficiente. Una aproximación a la Agricultura a Tiempo Parcial en España*. Madrid. Ministerio de Agricultura. Pp 121-151.
- Dirección General de Conservación de la Naturaleza. DGCN, 2006. Los incendios forestales en España. In: Informe anual y decenal. Ministerio de Medio Ambiente, Madrid.
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity&Distributions* 17(1), 43-57. doi:10.1111/J.1472-4642.2010.00725.X
- FAO. Fire Management - Global Assessment 2006. 2007. A Thematic Study Prepared in the Framework of the Global Forest Resources Assessment 2005. FAO, Rome.
- Hötzi, H.(Eds.), *Climatic Changes and Water Resources in the Middle East and North Africa*. 2008. Environmental Science and Engineering. Springer, Berlin, Heidelberg, pp. 429-452.
- Kalabokidis, K.D., Gatzojannis, S., Galatsidas, S., 2002. Introducing wildfire into forest management planning: towards a conceptual approach. *For. Ecol. Manage.* 158, 41-50. doi:10.1016/s0378-1127(00)00715-5

- Kliot, N. 1996. Forests and Forest Fires in Israel. Haifa University, Department of Geography, Haifa, Israel 2-6 (IFFN No. 15, September 1996)
- Le Houerou, N.H. 1973. Fire and vegetation in Mediterranean Basin. *Proc. Ann. Tall. Timb. Fire Ecol.* Pp 1120-1127.
- Le Houerou, N.H. 1977. Fire and vegetation in North Africa. *En: Proc. Symp. Envir. Cons. Fire and fuel Manag. Medit. Ecosys. Forest Ser. Dep. Agric. U.S.*
- Martell, D.L., Otukol, S., Stocks, B.J., 1987. A logistic model for predicting daily people-caused forest fire occurrence in Ontario. *Canadian Journal of Forest Research* 17, 394-401.
- Martínez, J., Chuvieco, E., Martín, M.P. 2004. Estimating human risk factors in wildland fires in Spain using logistic regression, in: *II International Symposium on Fire Economics, Planning and Policy: A Global Visio Cordoba.*
- Martínez, J., Vega-García, C., Chuvieco, E. 2009. Human-caused wildfire risk rating for prevention planning in Spain. *J. Environ. Manage.* 90, 1241-1252. doi:10.1016/j.jenvman.2008.07.005
- Martínez-Fernández, J., Koutsias, N., 2011. Modelling fire occurrence factors in Spain. National trends and local variations, in: *San-Miguel Ayanz J Camia A, Oliveira S, G.I. (Ed.), Advances in Remote Sensing and GIS Applications in Forest Fire Management From Local to Global Assessments. JRC66634 Scientific and Technical Reports, Luxemburg, pp.203-208.*
- Mollicone, D., Eva, H.H., Achard, F. 2006. Ecology-human role in Russian wild fires. *Nature* 440;436-437.
- Moreno, M. V., Malamud, B.D., Chuvieco, E.A., 2011. Wildfire Frequency-Area Statistics in Spain. *Procedia Environ. Sci.* 7, 182-187. doi:10.1016/j.proenv.2011.07.032
- Naveh Z. 1975. The evolutionary significance of fire in the mediterranean region *Vegetatio* 21: 199-208.
- Padilla, M., Vega-García, C. 2011. On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. *Int. J. Wildl. Fire* 20, 46-58. doi:http://dx.doi.org/10.1071/WF09139
- Parisien, M.A.; Moritz, M.A. 2009. Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecol. Monogr.* 79, 127-154.
- Pausas, J.G. 2004. Changes in fire and climate in the Eastern Iberian Peninsula (Mediterranean Basin). *Clim. Chang.* 63, 337-350.
- Pausas, J.G., Fernández-Muñoz, S. 2012. Fire regime changes in the Western Mediterranean Basin: from fuel-limited to drought-driven fire regime. *Clim. Chang.* 110, 215-226.
- Pausas, J.G., Vallejo, V.R. 1999. The role of fire in European Mediterranean ecosystems. In: Chuvieco, P.D.E. (Ed.), *Remote Sensing of Large Wildfires.* Springer, Berlin, Heidelberg, pp. 3 -16.
- Pearson R.G., Raxworthy C.J., Nakamura M., Townsend Peterson A. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34(1), 102-117. doi:10.1111/J.1365-2699.2006.01594.X
- Pew, K.L., Larsen, C.P.S. 2001. "GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada", *Forest Ecology and Management*, 140, pp. 1-18
- Phillips, S., Dudík, M., Schapire, R. 2004. A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-First International Conference on Machine Learning*, pages 655-662.
- Phillips, S., Anderson, R., Schapire, R. 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190 (3e4), 231e259. http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026.

- Renard, Q.; Péliissier, R.; Ramesh, B.R.; Kodandapani, N. 2012. Environmental susceptibility model for predicting forest fire occurrence in the Western Ghats of India. *Int. J. Wildland Fire* 21, 368-379.
- Rodrigues M and de la Riva J. 2014. Assessing the effect on fire risk modeling of the uncertainty in the location and cause of forest fires. In Viegas DX (ed.) *Advances in ForestFireResearch*. Coimbra, Imprensa da Universidade de Coimbra, 1061-1072. http://dx.doi.org/10.14195/978-989-26-0884-6_116.
- Rodrigues, M. and de la Riva, J. 2014b. An insight into machine-learning algorithms to model human-caused wildfire occurrence. *EnvironmentalModelling& Software* 57: 192-201. doi:10.1016/j.envsoft.2014.03.003.
- Rodrigues, M., de la Riva, J., Fotheringham, S. 2014. Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography* 48:52-63. doi:10.1016/j.apgeog.2014.01.011
- San-Miguel Ayanz, J., Camia, A., 2009. Forest fires at a glance: facts, figures and trends in the EU, in: Birot, Y. (Ed.), *Living with Wildfires: What Science Can Tell Us. A Contribution to the Science-Policy Dialogue*. European Forest Institute, pp. 11-18.
- Shakesby, R.A. 2011. Post-wildfire soil erosion in the Mediterranean: review and future research directions. *Earth Sci. Rev.*, 105, 71-100.
- van der Werf G.R., Randerson J.T., Giglio L., Collatz G., Mu M., Kasibhatla P.S., Morton DC, DeFries RS, Jin Y, van Leeuwen TT. 2010. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009). *Atmospheric Chemistry and Physics* 10, 11 707-11 735. doi:10.5194/ACP-10-11707 2010
- Syphard, A.D., Radeloff, V.C., Keely, J.E., Hawbaker, T.J., Clayton, M.K., Stewart, S.I., Hammer, R.B. 2007. Human influence on California fire regimes. *Ecological Applications* 17(5), 1388-1402. doi:10.1890/06-1128.1
- Vasconcelos, M.J.P., Silva, S., Tomé, M., Alvim, M., Pereira, J.M.C. 2001. Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. *Photogramm Eng Rem S* 5, 101-111.
- Vega Garcia, C., P.M. Woodard, and B.S. Lee. 1993. Mapping risk of wildfires from human sources of ignition with a GIs. Pages 419-426 In *Proceedings of the Thirteenth Annual ESRI User's Conference*, Palm Springs, California, USA, May 24-28th 1993, Vol. 1. 608 pp.
- Vega-Garcia, C. V, Woodard, P.M., Titus, S.J., Adamowicz, W.L., Lee, B.S. 1995. A Logit Model for Predicting the Daily Occurrence of Human Caused Forest-Fires. *Int. J. Wildl. Fire* 5, 101- 111. doi:<http://dx.doi.org/10.1071/WF9950101>
- Vélez, R., 2001. Fire Situation in Spain, in: Goldammer, J.G., Mutch, R.W., Pugliese, P. (Eds.), *Global Forest Fire Assessment 1990-2001*. FAO, Roma.
- Vilar, L., Martín, M.P., Martínez, J. 2011. Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. *European Journal of Forest Research (Impact Factor: 1.68)*. 11/2011; 130(6). DOI: 10.1007/s10342-011-0488-2
- Wagenbrenner, J.W., MacDonald, L.H., Rough, D. 2006. Effectiveness of three post-fire rehabilitation treatments in Colorado Front Range. *Hydrol. Process.* 20, 2989-3006
- Wittenberg, L., Malkinson, D., 2009. Spatio-temporal perspectives of forest fires regimes in a maturing Mediterranean mixed pine landscape. *Eur. J. For. Res.* 128, 297-304.

