



UNIVERSITAT OBERTA DE CATALUNYA (UOC)
MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS (*Data Science*)

TRABAJO FINAL DE MÁSTER

ÁREA: BIG DATA EN SISTEMA DE INFORMACIÓN GEOGRÁFICA

El análisis del fuego perfecto **Los incendios de 6^a generación: Causas, efectos y prevención.**

Autor: Francisco Jesús Cárdenas Ruiz

Tutor: Anna Muñoz Bolas

Profesor: Albert Solé Ribalta

Almería, 10 de enero de 2024

Créditos/Copyright



Esta obra está sujeta a una licencia de Reconocimiento - NoComercial - SinObraDerivada 3.0 España de Creative Commons.

© 2023 - Francisco Jesús Cárdenas Ruiz

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

FICHA DEL TRABAJO FINAL

Título del trabajo:	El análisis del fuego perfecto
Nombre del autor:	Francisco Jesús Cárdenas Ruiz
Nombre del colaborador/a docente:	Anna Muñoz Bolas
Nombre del PRA:	Albert Solé Ribalta
Fecha de entrega (mm/aaaa):	01/2024
Titulación o programa:	Máster Universitario en Ciencia de Datos
Área del Trabajo Final:	Big Data en Sistema de Información Geográfica
Idioma del trabajo:	Español
Palabras clave	Incendio, Geoespacial, Visualización

Resumen

Un incendio forestal es la propagación de un fuego sin control sobre terreno forestal o silvestre. Pues bien, hoy en día conocemos la evolución de éste a lo que llamamos incendio de sexta generación. Este tipo de incendios también se conocen como los incendios de la era del cambio climático, ya que se atribuyen en gran medida a los efectos de este, como las condiciones más secas y calurosas, las cuales son ideales para su propagación. En este estudio analizaremos las causas y factores que han provocado la aparición de estos nuevos megaincendios que se caracterizan por ser inestables, voraces, muy destructivos y totalmente impredecibles. Se elaborará una visualización interactiva con los incendios más devastadores de los últimos años y explicaremos las fases que siguen hasta denominarse de sexta generación. Por último, haremos una investigación sobre las acciones de prevención actuales y cómo poder mejorarlas para mitigar estos incendios en unas simulaciones hipotéticas. El objetivo sería contar con una herramienta de predicción que nos ayude a estar mejor preparados ante estos nuevos fenómenos devastadores.

Palabras clave: Incendios, Fuego, Ciencia de Datos, Datos geoespaciales, Visualización, Mapa, GIS, Prevención, Trabajo Final de Máster.

Abstract

A wildfire is the uncontrolled spread of fire over forest or wild land. Well, today we know the evolution of this to what we call the sixth-generation fire. This type of fire is also known as the fires of the climate change era, since they are largely attributed to the effects of climate change, such as drier and warmer conditions, which are ideal for their spread. In this study we will analyze the causes and factors that have led to the emergence of these new mega-fires, which are characterized by being unstable, voracious, highly destructive and totally unpredictable. We will elaborate an interactive visualization with the most devastating fires of the last years, and we will explain the phases that follow until they are called the sixth generation. Finally, we will investigate current prevention actions and how they can be improved to mitigate these fires in hypothetical simulations. The objective would be to have a predictive tool that will help us to be better prepared for these new devastating phenomena.

Palabras clave: Wildfires, Fire, Data Science, Geospatial Data, Visualization, Map, GIS, Prevention, Master's Thesis.

Índice general

Resumen	III
Abstract	IV
Índice	V
Listado de Figuras	VII
Listado de Tablas	1
1. Introducción	2
1.1. Contexto y Justificación del Trabajo	2
1.2. Motivación	4
1.3. Objetivos del Trabajo	5
1.4. Impacto en sostenibilidad, ético-social y de diversidad	6
1.5. Enfoque y Metodología	9
1.6. Planificación del Trabajo	10
2. Estado del arte	13
2.1. Estudio sobre los análisis de incendios	13
2.1.1. Clasificación de las 6 generaciones	17
2.1.2. Causas generales de los incendios forestales.	18
2.1.3. Evolución de incendios en España	18
2.2. Estudios de los mapas interactivos sobre incendios	20
2.3. Tecnologías de soporte para la implementación	24
3. Diseño e Implementación del trabajo	25
3.1. Mapa interactivo del incendio de Tenerife de 2023	25
3.1.1. Primer ciclo del incendio	25
3.1.2. Segundo ciclo del incendio	26

3.1.3.	Tercer ciclo del incendio	27
3.1.4.	Espacios naturales que fueron afectados	28
3.2.	Modelo de predicción de incendios con base en los incendios de Canadá	29
3.2.1.	Análisis del conjunto de datos	30
3.2.2.	Preprocesamiento del conjunto de datos	32
3.2.3.	Visualización del conjunto de datos	34
3.2.4.	Ejecución de modelos predictivos sobre el conjunto de datos	35
4.	Conclusiones y trabajos futuros	38
4.1.	Conclusiones enfocadas al análisis de incendios	38
4.2.	Conclusiones enfocadas a los Objetivos de Desarrollo Sostenible	39
4.3.	Conclusiones basadas en la predicción del modelo con datos de Canadá	40
5.	Glosario	41
5.1.	Glosario de términos	41
5.2.	Glosario de descripciones del conjunto de datos	43
	Bibliografía	44
A.	Anexo: Código de R completo.	47

Índice de figuras

1.1. Actuación nocturna en el incendio de Nonaspe (Zaragoza), 18 de junio de 2022.	3
1.2. Los ‘baños de bosque’ o Shinrin Yoku tienen su origen en Japón.	4
1.3. Ilustración de metodología Agile a base de Sprints.	9
2.1. Ilustración del Triángulo de Fuego.	15
2.2. Ilustración de los tipos de propagación de fuego.	16
2.3. Captura de la plataforma sig.mapama.gob.es.	20
2.4. Captura de la plataforma COPERNICUS.	21
2.5. Captura de la plataforma de AEMET.	22
2.6. Captura de la plataforma de CWFIS.	23
3.1. Captura de la primera historia del mapa interactivo.	26
3.2. Captura de la segunda historia del mapa interactivo.	27
3.3. Captura de la tercera historia del mapa interactivo.	28
3.4. Captura de la cuarta historia del mapa interactivo.	29
3.5. Mapa de incendios en Canadá entre 2006-2021.	34

Índice de cuadros

1.1. Gantt Octubre.	10
1.2. Gantt Noviembre.	11
1.3. Gantt Diciembre.	11
1.4. Gantt Enero/Febrero.	12
2.1. Evolución del número de incendios y superficie afectada en España. <i>Dirección General para la Biodiversidad, Ministerio de Medio Ambiente.</i>	19
3.1. Resumen de la ejecución del modelo C50.	36
3.2. Resumen de la ejecución del modelo C50 ajustado.	37
3.3. Uso de las variables en el modelo C50 ajustado.	37

Capítulo 1

Introducción

1.1. Contexto y Justificación del Trabajo

En la actualidad, y cada vez con más frecuencia, nos informan de que se han producido nuevos incendios en el mundo. Estas noticias están empezando a ser la tónica común en las épocas de verano de diferentes zonas del planeta. Es por ello por lo que, debemos profundizar en las causas y factores que están provocando el auge de megaincendios cada año. Tenemos un gran enemigo que ha venido para quedarse y ese es el cambio climático. En paralelo a este, los ciudadanos de los países más desarrollados hemos adoptado nuevas formas de asentarnos, aumentando la capacidad de grandes urbes repletas de servicios y despoblando los pueblos y aldeas remotas. Esto lleva consigo el descuido de gran parte de terreno silvestre que queda abandonado. [7]

El fuego a veces es inevitable, porque lo puede desencadenar la simple caída de un rayo de una tormenta, por lo que el mantenimiento del monte y de las zonas forestales se vuelve esencial. Ante esta situación, el bosque se mantiene indefenso, los caminos quedan llenos de maleza y forraje seco, los cuales son un excelente combustible para el fuego.

El anteriormente llamado cambio climático nos está mostrando que las estaciones del año empiezan a ser dos en vez de cuatro y que llegar a los extremos cada vez es más cotidiano. Adoptamos 5 meses fríos y 5 meses calientes, el otoño y la primavera se disfrutan sólo 2 meses al año. Sufrimos temperaturas más bajas en el periodo frío, ocasionando grandes heladas y temporales de nieve. Y por el otro lado temperaturas más altas en el periodo cálido, batiendo récords que rozan los 50°C en muchas zonas, cuando antes esto era algo muy inusual.

Pues bien, con estos precedentes dan lugar los llamados incendios de sexta generación, éstos tienen la capacidad de modificar las condiciones meteorológicas de la zona afectada por las llamas y su entorno más inmediato. Además, tienen una circulación de aire continua, de ahí su voracidad, ya que el fuego se retroalimenta continuamente de oxígeno. [1]

Esto genera unas llamas que pueden superar los 30 metros de altura y una radiación de calor en 200 metros a la redonda que dificulta permanentemente las labores de extinción. Son, por lo tanto, fuegos inextinguibles. Sus nubes se pueden disipar o, por el contrario, caer desplomadas: un fenómeno que acontece cuando una capa de aire frío se sitúa justo encima. Además, la situación se agrava si aparecen las denominadas tormentas de fuego, capaces de generar focos secundarios. [15]



Figura 1.1: Actuación nocturna en el incendio de Nonaspe (Zaragoza), 18 de junio de 2022.

Uno de los grandes objetivos de este Trabajo Final de Máster es analizar los incendios de sexta generación para emprender las mejores acciones de prevención contra ellos. Es de vital importancia conocer que factores desencadenantes son susceptibles de ser mitigados o eliminados para que no se produzca un incendio de ciertas características.

Como habitantes de este planeta debemos de estar muy bien concienciados de todo lo que se pierde cuando hay un incendio. Y es que no es sólo vegetación y tierra quemada, hablamos de especies, fauna y ecosistemas que dejan de existir en nuestro mundo. Esto además es contraproducente porque a medida que perdemos bosque, las probabilidades de que los incendios sean mayores se acentúa ya que la zona se vuelve cada vez más árida.

Con el uso de mapas de visualización interactivos podremos ver algunas de las fases de propagación de estos fuegos y analizar si existen similitudes para definir hipótesis de extinción. Por otro lado, vamos a investigar las medidas de prevención actuales y proponer mejoras sobre ellas. Con esta información se lanzarán simulaciones y predicciones tomando como variables la temperatura, el viento, la frondosidad de la zona, la causa del incendio, la humedad relativa, ... Con la idea de poder dar una probabilidad de incendio o riesgo del mismo. Si además, en un futuro esta herramienta se alimenta de nuevos datos, las predicciones mejorarían y la precisión sería mucho mayor.

1.2. Motivación

Para mi es muy sencillo, no hay un planeta B, éste es nuestro hogar y tenemos que cuidarlo por nosotros y por los que vienen.

En este mundo cada vez más tecnológico se hace si cabe más prioritario la desconexión digital, para ello sólo conozco una forma de conseguirlo y es con un baño de bosque. Se trata de una práctica popular en Japón y el Extremo Oriente, donde se la conoce como *shinrin-yoku*, y consiste en realizar una visita a un bosque sumergiéndose en él con los cinco sentidos, a fin de obtener un bienestar para la persona o un beneficio para su salud. Recargar nuestra batería, llenar nuestros pulmones y limpiar nuestra mente. Es físicamente imposible vivir sin el oxígeno y los bosques son los generadores de este elemento.



Figura 1.2: Los ‘baños de bosque’ o Shinrin Yoku tienen su origen en Japón.

Nuestra propia evolución ha traído consigo muchos avances positivos para la sociedad y la ciencia, pero también ha traído grandes desastres y ante ello tenemos que actuar.

La naturaleza y el medio deben ser preservados, y cuando vemos la noticia de un nuevo incendio durante días, la sensación que se genera es bastante trágica. La rapidez con la que se pierde un bosque lleno de vida, y la de años que tienen que pasar para que se pueda medio recuperar... En este estudio no vamos a entrar en el hecho de que muchos de los incendios son provocados, porque esto ya es desgarrador, pero sí que vamos a analizar el impacto que el humano ha tenido en la propagación de incendios.

Por mi gremio, que es la ciencia de datos, y mi conocimiento en el análisis del dato esta es la mejor aportación que puedo hacer para combatir dentro de mis posibilidades los incendios forestales actualmente.

Por último, en este máster hemos aprendido herramientas del ámbito geoespacial como son Carto o ArcGIS, las cuales me han resultado muy interesantes para la visualización en mapas, y han motivado la elección de un trabajo enfocado en el análisis geoespacial de la información.

1.3. Objetivos del Trabajo

El objetivo principal es realizar un Trabajo Final de Máster en el que se desarrollen las capacidades adquiridas a lo largo de toda la titulación de Ciencia de datos.

Para ello voy a desarrollar una visualización interactiva en una herramienta GIS sobre las fases de los mayores incendios en los que se tienen datos. Con esto podremos estudiar algunas de sus peculiaridades y poder definir medidas que nos lleven a una mitigación o prevención de estos tipos de incendio de sexta generación.

Además, lanzaremos algunas simulaciones sobre predicción de factores de riesgo a incendios y desarrollaremos hipótesis fundamentadas en estos resultados.

Para conseguir este objetivo fijaremos unos objetivos más concretos sobre los que se desarrollará todo el trabajo final.

1. Describir y planificar el proyecto.
2. Búsqueda e investigación de las fuentes de datos sobre la temática de incendios.
3. Clasificar la información en las diferentes 3 líneas de trabajo: Análisis de incendios, Mapa interactivo y Predicción de factores desencadenantes.
4. Desarrollo de la visualización en entorno GIS.
5. Preprocesamiento de datos y análisis estadístico.

6. Ejecución de la simulación de predicción de incendios.
7. Redacción y documentación de los resultados.
8. Presentación de puntos clave y demostración.
9. Defensa del proyecto ante un tribunal.

1.4. Impacto en sostenibilidad, ético-social y de diversidad

Se podría afirmar que los incendios de sexta generación están relacionados con cada uno de los Objetivos de Desarrollo Sostenible (ODS) tal y como se explica a continuación.

[Enlace a la página de los ODS](#)

■ ODS 1 - Fin de la pobreza.

Los incendios de sexta generación pueden producir la pérdida de medios de vida. Si afectan a cultivos, ganado y otras fuentes de ingresos, llevará a la pobreza a las personas cuyo sustento se vea destruido ya que carecerán de sus medios para ganarse la vida. Se pueden mencionar como personas en especial peligro aquéllas que viven en zonas rurales.

Además, este tipo de incendios pueden quemar viviendas, dejando sin hogar a muchas personas.

Es importante también indicar que los efectos devastadores de los incendios, como es lógico, impactan más en aquellos lugares que no tienen capacidad económica suficiente para reconstruir rápidamente lo que haya sido destruido por las llamas.

■ ODS 2 - Hambre cero.

Por un lado, si los incendios destruyen cultivos, puede ocurrir una escasez de alimentos que lleven a situación de hambre, sobre todo a las personas que necesiten de las tierras quemadas para su abastecimiento.

Por otro lado, estos incendios pueden contaminar el agua potable, dificultando el acceso a alimentos seguros y nutritivos provocando incluso enfermedades y problemas de salud.

■ ODS 5 - Igualdad de género.

Los incendios de sexta generación pueden aumentar la discriminación contra las mujeres y las niñas debido a que los desastres naturales pueden crear situaciones de inseguridad, escasez de recursos e interrupción de servicios que facilitan la violencia. Además, puede verse dificultado el acceso a la protección y los servicios.

- **ODS 6 - Agua limpia y saneamiento.**

Ya se ha comentado con anterioridad que los incendios de sexta generación pueden contaminar el agua potable, con todos los problemas que acarrea ello.

Habría que añadir que, si los incendios llegan a destruir infraestructuras de agua y saneamiento, se complicará más aún el acceso a agua limpia y segura.

- **ODS 7 - Energía asequible y no contaminante.**

Principalmente, el tipo de incendios que se estudia en este trabajo, pueden provocar la liberación de contaminantes tóxicos en el aire, lo cual afectará negativamente la salud humana, animal, y al medio ambiente en general.

Y además, los incendios de sexta generación pueden considerarse una amenaza para la producción de energía renovable, pueden destruir las infraestructuras necesarias para la generación de energía solar o eólica si ocurren en lugares cercanos a dónde se encuentran situadas estas infraestructuras.

- **ODS 8 - Trabajo decente y crecimiento económico.**

Los incendios de sexta generación pueden producir pérdida de empleos, ya que al destruir todo lo que encuentran a su paso, afectarán a empresas de las que su actividad económica dependa de aquello que se ha quemado. Esto ralentizará el crecimiento económico y dificultará la generación de nuevos puestos de trabajo.

- **ODS 9 - Industria, innovación e infraestructura.**

Los incendios de sexta generación pueden impactar en el desarrollo económico y el progreso social, ralentizándolos en el caso que sean dañadas infraestructuras importantes de industria y transporte, por ejemplo. Habría que incurrir en la pérdida de puestos de trabajo así como en la paralización de servicios esenciales para la población.

- **ODS 10 - Reducción de las desigualdades.**

Los incendios de sexta generación pueden incrementar la exclusión social dejando a las personas sin hogar, sin medios de vida y sin acceso a servicios vitales tales como servicios públicos, la educación y la salud.

Los daños que causan estos desastres naturales propician una difícil recuperación económica, aumentando la desigualdades.

Una de las peores consecuencias de los incendios también será el aislamiento de las zonas donde se producen, bien durante la propagación de las llamas, o bien una vez estas han sido extinguidas por haberse quedado incomunicadas. Aumentará así la desigualdad de las personas afectadas.

- **ODS 11 - Ciudades y comunidades sostenibles.**

Del mismo modo que ocurría con el ODS 7, los incendios pueden aumentar la contaminación del aire y el agua, lo que puede afectar negativamente la salud de los residentes de un lugar, dejando esta zona de ser sostenible.

Además, los incendios también pueden provocar la destrucción de viviendas si se produce en alguna zona habitada.

- **ODS 12 - Consumo y producción responsables.**

Los incendios que aquí se analizan, son una consecuencia del consumo excesivo y la producción insostenible de la sociedad en la que vivimos. El ritmo creciente de consumo de recursos naturales y la falta de eficiencia de los procesos de producción de la economía actual, han incrementado las probabilidades de que se produzcan este tipo de incendios.

- **ODS 13 - Acción por el clima.**

El cambio climático es uno de los factores que más afecta a que los incendios sean del tipo “sexta generación”.

El aumento de las temperaturas, que cada vez son más altas, hacen que la vegetación sea más seca y propensa a la ignición.

La sequía, que también se acentúa con el paso de los años, reduce la humedad del suelo y la vegetación, lo cual facilita la propagación de los incendios.

El comportamiento de las precipitaciones, que cada vez son menos homogéneas y duraderas, dificulta la extinción de los incendios, ya que no hay agua suficiente para apagar las llamas.

- **ODS 14 - Vida submarina.**

Los incendios de sexta generación pueden liberar contaminantes en el agua, que terminan dañando los ecosistemas marinos.

- **ODS 15 - Vida de ecosistemas terrestres.**

Los incendios de sexta generación destruyen los bosques y otros ecosistemas terrestres. Al quemar la vegetación, provocan un impacto devastador en la biodiversidad del lugar donde ocurren.

Igualmente, al ser de dimensiones tan grandes, este tipo de incendios liberan cantidades de humo y cenizas a la atmósfera, contaminando el aire y el agua como ya se ha adelantado en la descripción de otros ODS.

Finalmente, mencionar que los incendios causan erosión del suelo, disminuyendo la capacidad fértil del mismo, haciendo que sea mucho más complicada la recuperación posterior del ecosistema.

- **ODS 16 - Paz, justicia e instituciones sólidas.**

De los incendios se generan conflictos sociales debido a que pueden causar daños a la propiedad de las mismas, y estas quedar desplazadas. La tensión social se verá incrementada y será más difícil lograr un estado de paz y justicia.

Además, pueden afectar a la prestación de servicios públicos y no garantizar la protección de los derechos humanos si no hay recursos suficientes como para paliar los efectos producidos por la destrucción ocasionada por las llamas.

1.5. Enfoque y Metodología

La metodología a seguir en gran medida será Agile ya que en este proyecto existen 3 líneas de trabajo. Es cierto que esta metodología no está pensada para una sola persona, pero adaptaré mis propios sprints para que así sea.

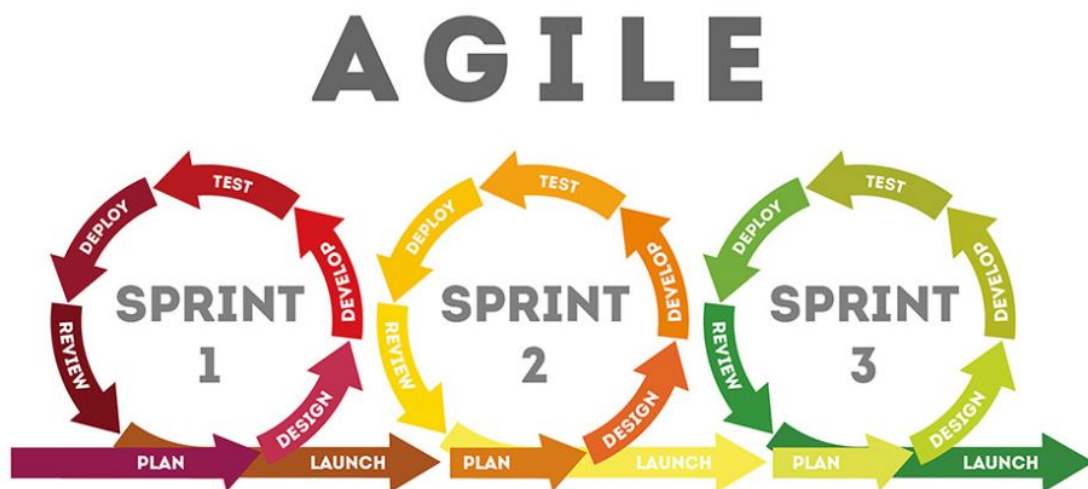


Figura 1.3: Ilustración de metodología Agile a base de Sprints.

Como haremos una investigación, lo primero será buscar información sobre el estado del arte de proyectos pasados que incluso pueden servir de inspiración.

En paralelo, deberemos buscar información sobre todo este nuevo fenómeno de los incendios de sexta generación, sus causas, factores, efectos, testimonios, ...

Una vez vayamos recopilando información ésta la iremos clasificando, dependiendo de la línea de trabajo a la que de respuesta que en algunos casos puede dar respuesta a más de una.

Aplicaremos la información de carácter geoespacial en la herramienta escogida de GIS y empezaremos a visualizar el desarrollo de algún incendio real. Al mismo tiempo iremos preprocesando datos tanto para la visualización en mapa interactivo como para la futura predicción de factores

de riesgo a incendios.

Guardaremos algunos resultados relevantes de esta primera parte, para dar explicación al fenómeno y contrastarlo con los resultados finales. Esto lo haremos con estadísticos de referencia.

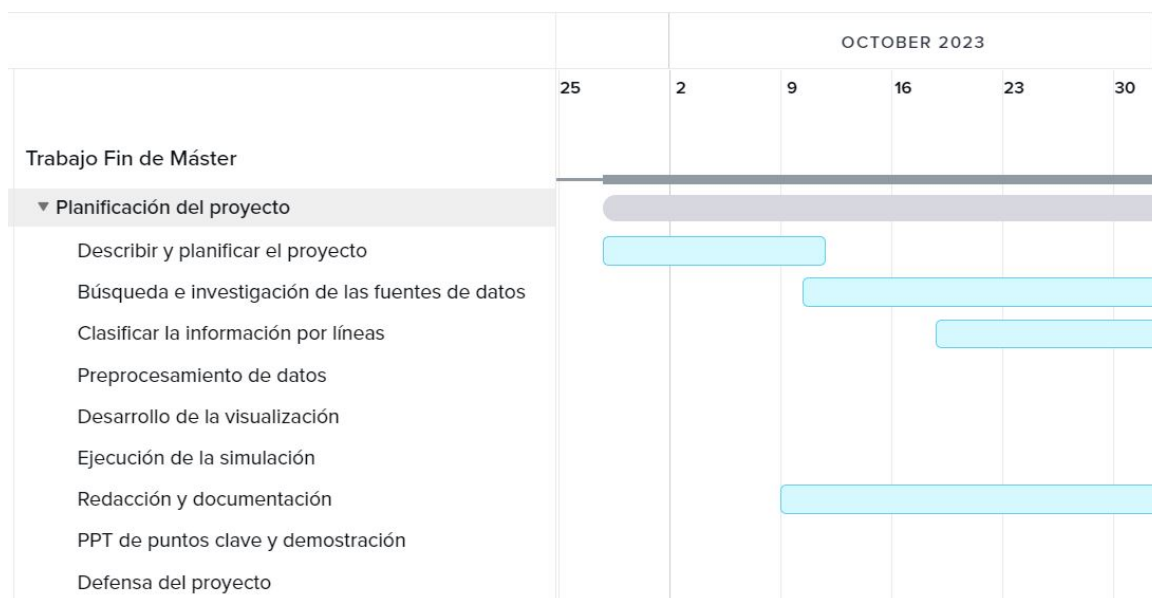
Mediante algoritmos supervisados en RStudio intentaremos ejecutar algunas simulaciones que puedan arrojaros datos sobre probabilidades de riesgo a incendios.

Iremos ajustando la cantidad y la calidad de las variables hasta conseguir resultados con precisiones por encima del 70 %.

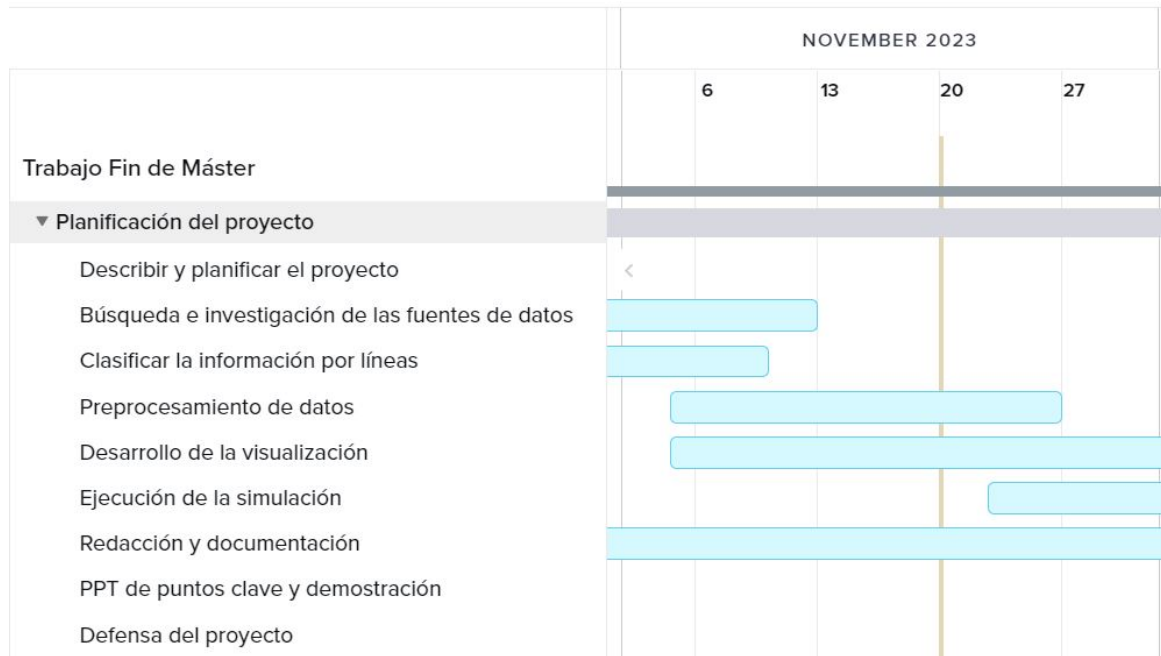
Todo el trabajo quedará redactado y documentado en esta memoria final del Trabajo Fin de Máster.

1.6. Planificación del Trabajo

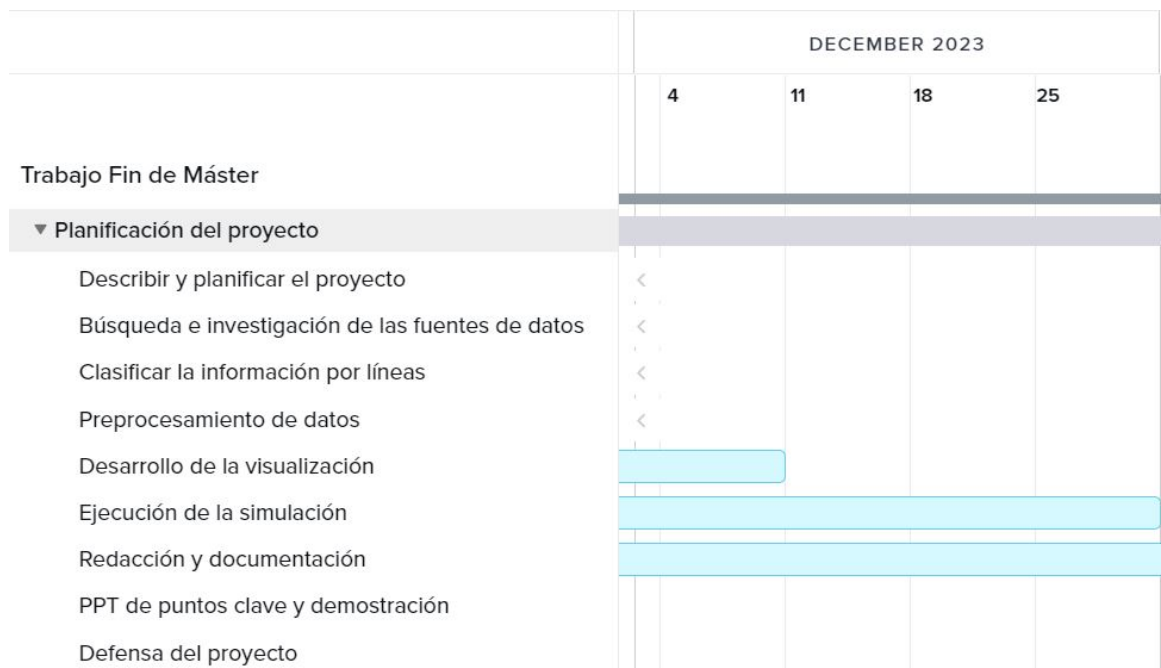
Para el desarrollo de este proyecto vamos a necesitar de datos abiertos sobre incendios, herramientas de visualización en entorno GIS como Carto o ArcGIS, un software de preprocesamiento y limpieza de datos como R-Rstudio y un procesador de textos que será Microsoft Word con la asistencia de LaTeX. En cuanto al timing o diagrama de Gant, se va a seguir la siguiente planificación según los objetivos a cumplir:



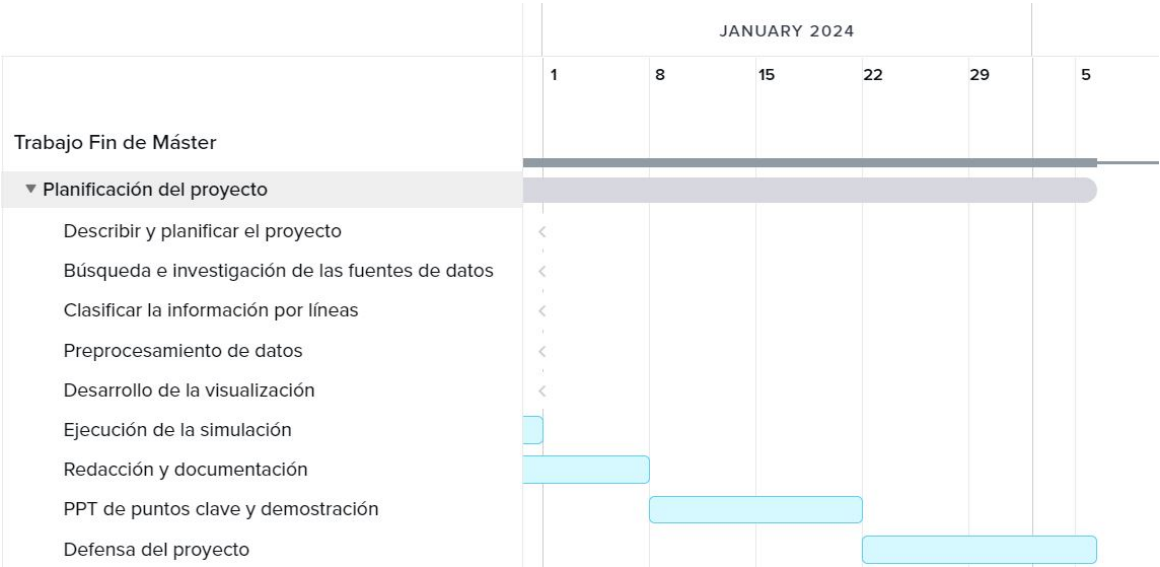
Cuadro 1.1: Gantt Octubre.



Cuadro 1.2: Gantt Noviembre.



Cuadro 1.3: Gantt Diciembre.



Cuadro 1.4: Gantt Enero/Febrero.

Capítulo 2

Estado del arte

En esta sección vamos a dar contexto al estado del arte que existe sobre las diferentes líneas de este trabajo final de máster.

Primeramente, abordaremos los estudios relacionados con el análisis de incendios. Después mostraremos algunos de los mapas y capas creados sobre esta temática, incluso veremos ejemplos de predicción de incendios y monitorización en tiempo real. Por último, veremos algunos ejemplos de causas y factores que desencadenan un incendio y como se relacionan de cara una predicción de conato.

2.1. Estudio sobre los análisis de incendios

Según la definición con la que se detalla en el artículo 6 (apartado K) de la Ley 43/2003, de 21 de Noviembre, de Montes, “Incendio forestal: el fuego que se extiende sin control sobre combustibles forestales situados en el monte”.

Una de las características fundamentales en la definición de incendio forestal es que el fuego escape al control del ser humano. Por ese motivo, no se consideran incendios la quema de pastos y matorrales o la eliminación de residuos forestales, excepto cuando las mismas se extienden a un monte de forma incontrolada. [5]

Según el artículo 5 de la ley anterior: “A los efectos de esta ley, se entiende por monte todo terreno en el que vegetan especies forestales arbóreas, arbustivas, de matorral o herbáceas, sea espontáneamente o procedan de siembra o plantación, que cumplan o puedan cumplir funciones ambientales, protectoras, productoras, culturales, paisajísticas o recreativas“. Así mismo, podemos indicar que no tienen consideración de monte ni los terrenos dedicados al cultivo agrícola ni los terrenos urbanos u otros que excluya cada Comunidad Autónoma en su normativa forestal y urbanística. [4]

Los incendios se pueden categorizar en función del tipo de terreno que se ve afectado: [11]

- Arbolada: se refiere a áreas afectadas por el fuego que están cubiertas de árboles que producen madera, frutas, leña, corcho o resina. Para que un área se considere arbolada, debe tener más del 20 % de su superficie cubierta por árboles.
- No arbolada: cuando el fuego solo afecta al matorral o al pasto situado entre los árboles sin dañar a estos. El terreno no arbolado se divide en las siguientes categorías:
 - Dehesa o monte abierto: los árboles cubren menos del 20 % de la superficie.
 - Matorral y monte bajo: incluye vegetación leñosa de porte arbustivo.
 - Pastizales: compuesto por vegetación herbácea.
 - Superficie no forestal: corresponde a áreas agrícolas o urbanas.

Las características de los incendios forestales se pueden entender a través de la interacción de varios elementos en lo que se llama el "triángulo del fuego". [8] En la Figura 2.1 tenemos una ilustración del triángulo.

En un incendio forestal, este triángulo consta de:

1. **Combustible:** Se refiere a una sustancia que, cuando se calienta en presencia de oxígeno a una temperatura específica, genera vapores que pueden encenderse simultáneamente. Es crucial notar que los combustibles no son uniformes, y la facilidad con la que se produce el fuego depende de sus propiedades individuales. Estas propiedades abarcan la forma, el contenido de humedad, el volumen y la superficie del combustible.
2. **Oxígeno:** En la atmósfera, alrededor del 21 % del aire está compuesto por oxígeno, y la mayoría de las llamas necesitan al menos un 16 % de oxígeno para mantenerse encendidas. El oxígeno desempeña un papel clave en los procesos químicos que tienen lugar durante la combustión. Cuando un material combustible arde, se combina con el oxígeno del entorno, generando calor y produciendo subproductos como gases, humo y brasas. Este fenómeno se conoce como oxidación.
3. **Calor:** Se refiere a la necesidad de una fuente de calor o energía de activación para que se produzca un incendio. La energía de activación, también conocida como calor, es esencial para iniciar la combustión de un combustible en presencia de un comburente. Los incendios forestales a menudo comienzan con un pequeño foco de calor que se propaga a otros objetos y lugares, resultando en un gran siniestro. Por lo tanto, es crucial comprender cómo se transmite el calor, que ocurre principalmente a través de tres formas: *conducción* (transmisión entre sólidos debido a la diferencia de temperatura), *convección*



Figura 2.1: Ilustración del Triángulo de Fuego.

(principalmente en líquidos y gases) y *radiación* (transferencia de energía a través del espacio mediante ondas electromagnéticas).

Además de estos tres elementos, hay un cuarto componente importante conocido como la **reacción en cadena**. Cuando un incendio forestal es muy intenso, se generan llamas y se libera una gran cantidad de calor. Esto hace que la combinación de oxígeno y combustible sea más propensa, lo que, a su vez, genera más llamas y calor. Este proceso se retroalimenta, liberando energía, gases como nitrógeno y dióxido de carbono, y nutrientes en forma de cenizas en el entorno de manera repentina.

Según la manera en que se propaga el incendio, se pueden clasificar en tres categorías: [12]

- Fuego de copas: Estos incendios ocurren en las copas de los árboles y suelen ser altamente destructivos debido al tamaño de las llamas y la rapidez con que se propagan, lo que es facilitado por los vientos más fuertes en altura.
- Fuego de suelo o superficial: Estos incendios se extienden por la superficie del suelo del bosque y afectan a la vegetación herbácea, al matorral y a los restos de vegetación. Son los incendios más comunes y pueden dar lugar a incendios de copas si no se controlan.
- Fuego de subsuelo: En este tipo de incendios, la materia orgánica seca y las raíces del subsuelo se queman. No se producen llamas visibles, emiten poco humo y se propagan muy lentamente. Este tipo de fuego suele aparecer en suelos con turba.

Es importante destacar que estos tipos de incendios pueden combinarse y, en general, los incendios de copas y de suelo son los más frecuentes en los incendios forestales. En la página web de la Asociación para la Prevención de Incendios Forestales (PREVIFOR) encontramos la Figura 2.2 para ilustrar estas categorías.

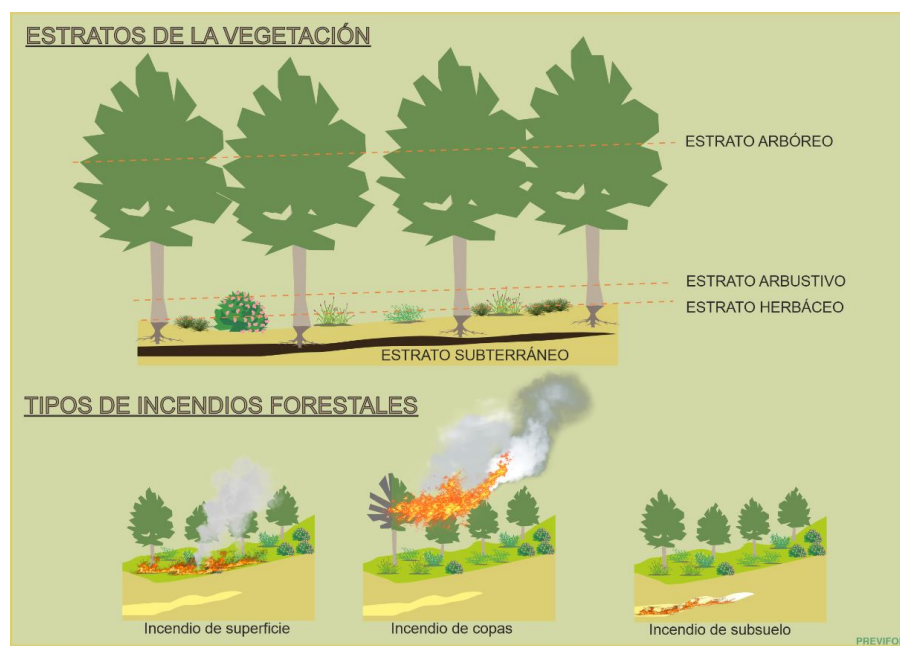


Figura 2.2: Ilustración de los tipos de propagación de fuego.
PREVIFOR

Si hablamos de los efectos que el fuego provoca en los ecosistemas forestales, podríamos dividir en tres casuísticas:

1. Consecuencias en la vegetación.

En primer lugar, se produce la destrucción completa de la parte visible de la vegetación, ya que los tejidos vegetales se someten a temperaturas superiores a 50° C durante más de 1 hora o a 60° C durante al menos 1 minuto.

Los árboles que sobreviven se debilitan, lo que varía dependiendo de las circunstancias y los individuos, lo que resulta en un crecimiento más lento y una mayor susceptibilidad a los hongos e insectos.

A medio plazo, después del incendio, las plantas herbáceas tienden a prosperar, ya que tienen acceso a una buena cantidad de luz para su regeneración. Posteriormente, las especies adaptadas al fuego, conocidas como pirófitas, ocupan el área. Finalmente, las especies de matorral que requieren mucha luz ocupan el terreno antes que los árboles.

2. Impacto en el suelo.

El incendio aumenta la erosión del suelo debido a la desaparición de la vegetación, y este

problema se agrava en terrenos con pendientes pronunciadas. La estructura del horizonte superficial del suelo se destruye, lo que resulta en una pérdida de permeabilidad.

Se produce un aumento del pH del suelo debido a las cenizas resultantes del incendio.

Se liberan repentinamente grandes cantidades de nutrientes. Los microorganismos del suelo se ven afectados negativamente, y su recuperación depende de la intensidad del fuego.

En suelos ricos en caliza, el carbonato cálcico (CaCO_3) se transforma en óxido de cal (CaO), lo que alcaliniza el suelo y afecta la asimilación de potasio y hierro, lo que lleva a la disminución de fósforo.

3. Cambios en el microclima

La ausencia de vegetación provoca un aumento en la luminosidad, lo que a su vez incrementa la variación de temperaturas, la absorción de calor por los restos carbonizados del suelo, la velocidad del viento, la evaporación, la desecación superficial y el impacto de las precipitaciones y la nieve en el suelo.

Pueden surgir plantas adaptadas a la sequedad, conocidas como xerófitas. En regiones con clima atlántico húmedo, la desaparición de árboles puede afectar los niveles freáticos y dar lugar a la formación de landas y turberas. En zonas áridas, se intensifica la erosión y la desertificación.

Se reduce la cantidad de precipitaciones ocultas, que son cruciales en áreas semiáridas y en la laurisilva canaria.

2.1.1. Clasificación de las 6 generaciones

En este punto, se detallan las características que presentan los incendios, así como, la respuesta de los dispositivos de extinción, sirviendo como base para dicha clasificación. [3]

La clasificación inicia con incendios de **primera generación** en los años cuarenta y cincuenta, marcados por un contexto de posguerra, fuerte depresión económica, y un paisaje rural manejado y poblado.

La **segunda generación**, en las décadas de los sesenta y primeros setenta, se caracteriza por el éxodo rural, el abandono de zonas cultivadas, y repoblaciones iniciales que demandan tratamientos selvícolas.

De los años setenta hasta los noventa, surge la **tercera generación**, con mayor profesionalización en los medios de extinción, incremento en el uso de medios aéreos, y continuidades forestales que propician incendios de mayores dimensiones.

La **cuarta generación**, alrededor del cambio de siglo, se asocia con la disminución de inversiones forestales, el declive rural, y la aparición de zonas de interfaz urbano-forestales, dando

lugar a incendios más extensos.

La **quinta generación**, previa a la crisis económica de 2008, se caracteriza por la inversión en dispositivos de extinción y la presencia masiva de ciudadanos en espacios naturales, junto con los efectos del cambio climático.

La **sexta generación**, aún sin una fecha precisa de inicio, se vincula a eventos como el incendio de 2017 en Pedrógão Grande, Portugal. Esta generación destaca por su capacidad para cambiar la meteorología local, con corrientes de aire convectivas que crean pirocumulonimbos.

2.1.2. Causas generales de los incendios forestales.

De acuerdo a las estadísticas recopiladas por el Ministerio de Medio Ambiente, Medio Rural y Marino, a través de la información proveniente de los Informes sobre Incendios Forestales, que son recopilados por las entidades responsables de la gestión forestal en cada región autónoma, se puede observar que el 96 % de los incendios forestales se originan debido a acciones humanas, mientras que el 4 % restante se debe a causas naturales, específicamente los rayos.

Dentro de las acciones humanas responsables de los incendios forestales, la Dirección General de Biodiversidad (DGB) identifica las siguientes categorías: [6]

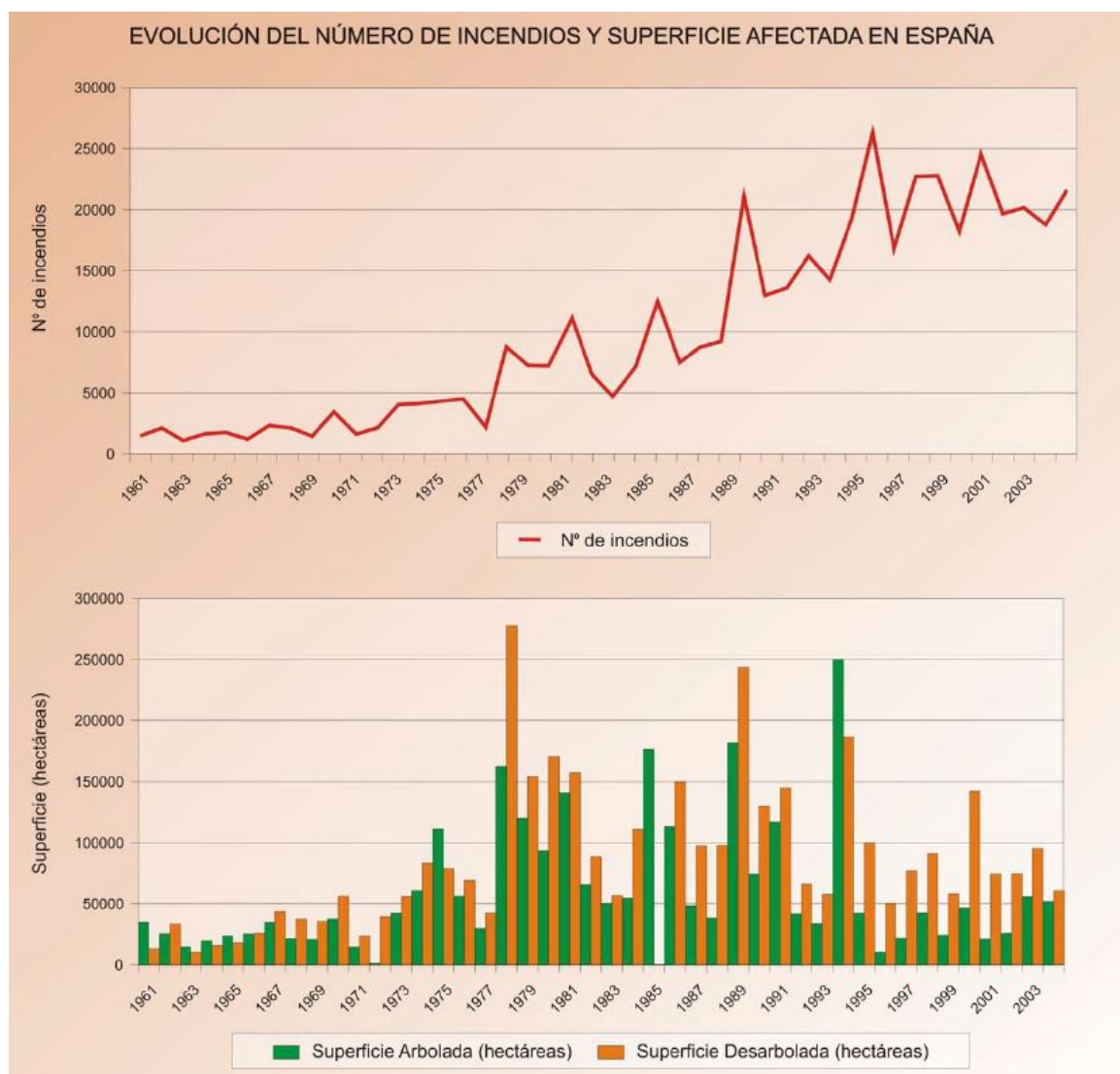
- Negligencias y causas accidentales.
- Causas intencionadas.
- Causas de origen desconocido, es decir, sin investigar la causa del incendio.
- Incendio reproducido, es decir, la reactivación de un incendio que se originó previamente.

2.1.3. Evolución de incendios en España

En España, los grandes incendios forestales se han convertido en un problema cada vez más grave. Y en el Cuadro 2.1 se puede evidenciar.

Entre los años 1963 y 2003, se produjeron 22 grandes incendios forestales en España, que afectaron a una superficie total de 1,7 millones de hectáreas.

El número de incendios y la superficie afectada por éstos en España han experimentado una evolución creciente en los últimos 50 años. Esta tendencia creciente se debe a una serie de factores, entre los que destacan el aumento de la población y la urbanización, el cambio climático y la mejora de los sistemas de detección y extinción de incendios.



Cuadro 2.1: Evolución del número de incendios y superficie afectada en España. *Dirección General para la Biodiversidad, Ministerio de Medio Ambiente.*

2.2. Estudios de los mapas interactivos sobre incendios

En la actualidad existen diseños sobre mapas interactivos que vienen a mostrar la temática de los incendios.

En algunos casos hasta muestran una predicción o posible zona de riesgo de incendio como son los mapas que se mostrarán a continuación.

■ Frecuencia de Incendios Forestales por Término Municipal.

Teniendo como fuente: Ministerio para la transición ecológica y el reto demográfico. Dirección General de Biodiversidad, Bosques y Desertificación. Subdirección General de Política Forestal y Lucha contra la Desertificación. Área de Defensa contra Incendios Forestales.

Se muestra la Figura 2.3 como ejemplo del mapa, en el que hemos incluido una capa base y la capa de Red Natura 2000. A medida que hacemos zoom vemos las zonas con más detalle.



Figura 2.3: Captura de la plataforma sig.mapama.gob.es.

Se puede acceder fácilmente desde este enlace: sig.mapama.gob.es.

- **Servicio de gestión de emergencias, sección: Incendios forestales.**

Este sería otro ejemplo, en este caso, se muestra la herramienta de COPERNICUS sobre una visualización de incendios reales, con un amplio abanico de histórico.

En la Figura 2.4 se muestra un captura con algunas de sus opciones.



Figura 2.4: Captura de la plataforma COPERNICUS.

Se puede acceder hacer uso desde este enlace: [COPERNICUS](#).

■ Servicios de Predicción desde AEMET.

Por su parte la Agencia Estatal de Meteorología también dispone de un servicio de predicción de incendios propio. En su caso, nos informa de los niveles de riesgo utilizando una barra temporal a la que podremos dar play, como se muestra en la Figura 2.5.

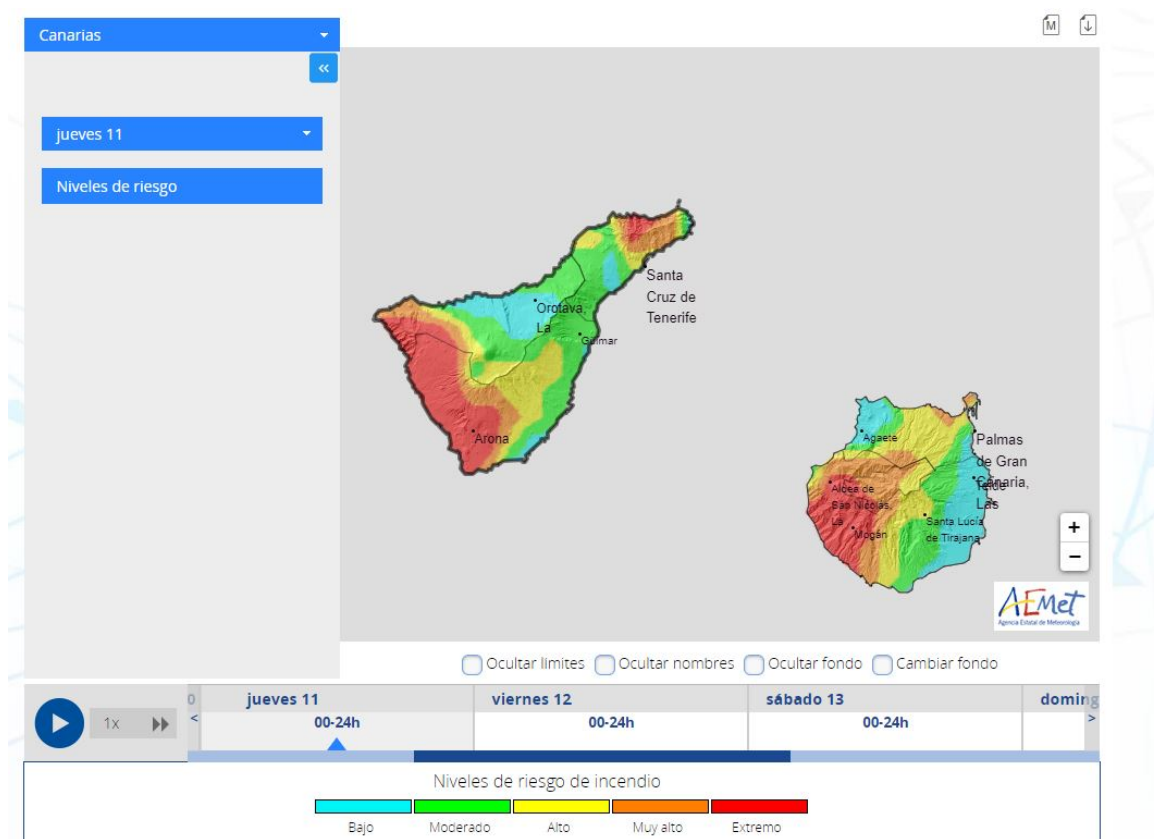


Figura 2.5: Captura de la plataforma de AEMET.

Se puede consultar desde este enlace: [AEMET](#).

- **CWFIS, Canadian Wildland Fire Information System.**

El Gobierno de Canadá nos muestra una herramienta muy completa sobre la predicción de incendios. Se muestra un ejemplo de visualización en la Figura 2.6.

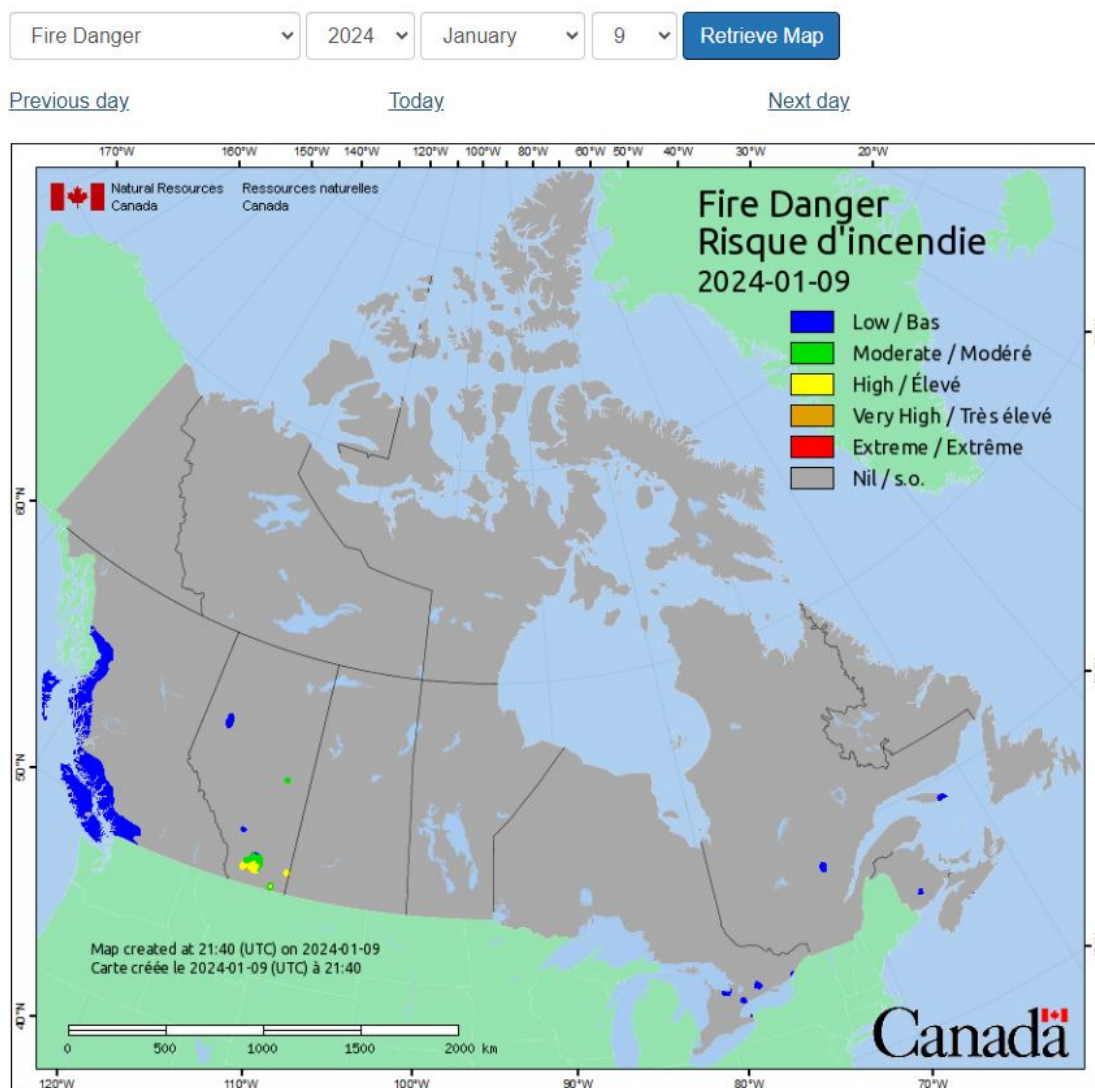


Figura 2.6: Captura de la plataforma de CWFIS.

Se puede usar desde este enlace: [CWFIS](#).

2.3. Tecnologías de soporte para la implementación

Se ha hecho uso de la plataforma de mapas del ISTAC (Instituto Canario de Estadística) que se basa en el software de **TerriaJS** [16], para implementar la visualización de un incendio. Se ha llevado a cabo haciendo agregaciones de capas con datos geográficos que el ISTAC proporciona en diferentes URLs y todas en formato XML.

Para el procesamiento de datos y generación del modelo predictivo se ha usado el lenguaje R dentro de su asistente de programación **RStudio**. El origen de datos ha sido extraído de las Bases de Datos públicas del Gobierno de Canadá, concretamente de su plataforma **Alberta**, [10] donde se aprovisiona toda la información relevante al medio ambiente. Y para un primer análisis hemos hecho uso de Microsoft Excel, al tratarse de ficheros con extensión .csv.

Y por último, para la redacción y presentación de la propia memoria, se ha utilizado el lenguaje LaTeX desde el asistente online de **Overleaf**

Capítulo 3

Diseño e Implementación del trabajo

3.1. Mapa interactivo del incendio de Tenerife de 2023

En este punto se mostrará a través de la herramienta de mapas en Open-Source de Terria.io [16], el desarrollo del incendio de la isla de Tenerife que se inició el 15 de agosto del 2023. Fue el peor de los últimos 40 años. Las llamas, que se iniciaron en las localidades de Arafo y Candelaria, arrasaron un total de 14.700 hectáreas, cerca del 7,1 % de la superficie de la mayor de las islas canarias.

El incendio afectó a 12 municipios de la isla, entre ellos Arafo, Candelaria, El Rosario, Güimar, El Sauzal, La Matanza, La Orotava, La Victoria, Los Realejos, Santa Úrsula y Tacoronte. [9] Declarado como el incendio mas extenso de España en el año 2023 [13], éste también tomó la categoría de incendio de sexta generación. Esto se debe a que en su caso fue capaz de generar un clima propio que lo hacía inextinguible durante su fase inicial y central. La particularidad de este incendio también fue la orografía del terreno, como sabemos la isla de Tenerife tiene el pico y volcán más alto de España, **el Teide** (3.715 metros sobre el nivel del mar), y el incendio tuvo lugar en las zonas escarpadas y cercanas a las coladas del Teide. Para los equipos de extinción de incendios fue muy difícil acceder a estas zonas, por lo que sólo se pudo contar con las unidades aéreas de helicópteros e hidroaviones.

Las causas del incendio aún están siendo investigadas, pero la Guardia Civil ha confirmado que fue intencionado.

3.1.1. Primer ciclo del incendio

En el primer mapa desarrollado se muestra el área, de color naranja, que comprende el fuego en los primeros días del incendio, desde el día 15 hasta el 18 de agosto.

En este ciclo el fuego avanzó rápidamente debido a las condiciones meteorológicas adversas, con

altas temperaturas, vientos fuertes y baja humedad. Las llamas obligaron a evacuar a más de 12.000 personas de sus hogares, afectando a 12 municipios de la isla.

Para obtener esta primera capa de visualización se ha hecho uso de los datos referentes al **Área quemada. 18/08/2023**, que se encuentran alojados en el [site](#) oficial del ISTAC en formato XML.

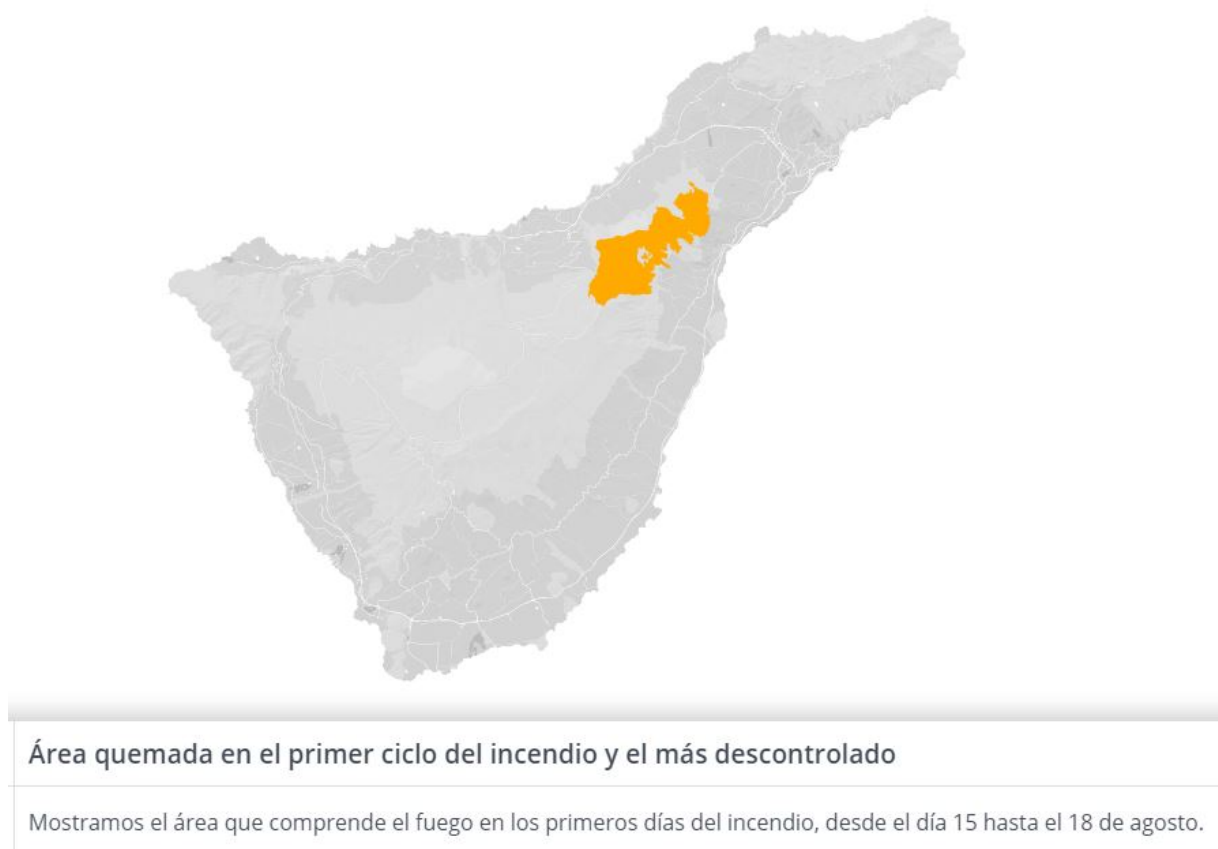


Figura 3.1: Captura de la primera historia del mapa interactivo.

3.1.2. Segundo ciclo del incendio

En esta segunda fase, el mapa desarrollado muestra, en un tono menos intenso al anterior, las nuevas zonas por las que se ha propagado el fuego entre los días 18 y 21 de agosto.

A lo largo de estos días, trabajaron en las áreas activas 381 efectivos de extinción, de los cuales 167 fueron bomberos urbanos, además de un centenar de seguridad. A los que se fueron incorporando, progresivamente, los 22 medios aéreos disponibles, los cuales se centraron en realizar descargas en las zonas más complicadas del incendio.

Para esta segunda capa de visualización se ha hecho uso de los datos referentes al **Área quemada. 21/08/2023**, que se encuentran alojados en el [site](#) oficial del ISTAC en formato XML.

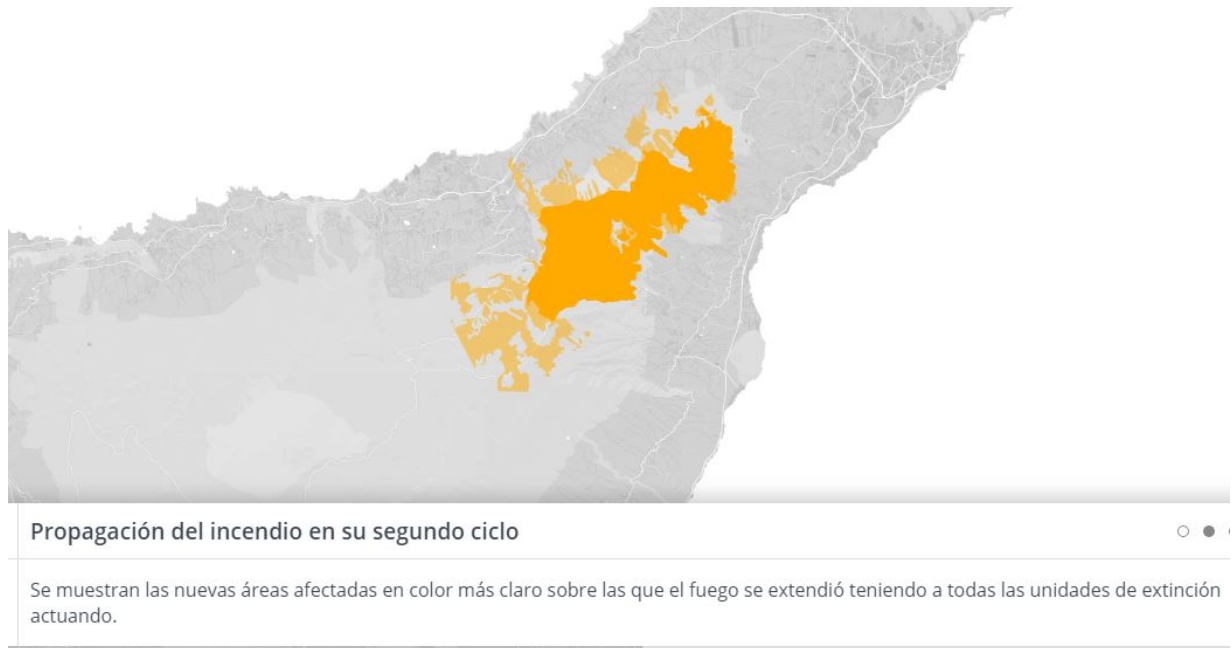


Figura 3.2: Captura de la segunda historia del mapa interactivo.

3.1.3. Tercer ciclo del incendio

Para la última fase, el mapa desarrollado muestra, las últimas zonas afectadas por las que se propagó el incendio entre el 21 y el 24 de agosto, marcando en este momento el estado de incendio controlado.

El incendio causó daños de gran magnitud en el ecosistema de la isla. Se perdieron miles de árboles, arbustos y plantas, así como animales de todo tipo. El fuego también afectó a infraestructuras, como carreteras, viviendas y cultivos.

Para esta tercera capa de visualización se ha hecho uso de los datos referentes al **Área quemada. 24/08/2023**, que se encuentran alojados en el [site](#) oficial del ISTAC en formato XML.



Figura 3.3: Captura de la tercera historia del mapa interactivo.

3.1.4. Espacios naturales que fueron afectados

En la cuarta historia que contiene la visualización, se ha querido mostrar a que espacios naturales afectó este incendio. Y para ello, el mapa ofrece la posibilidad de desplazar un cursor de izquierda a derecha que pintará sobre la base del mapa las áreas quemadas. Aquí además, se ha añadido la capa que nos muestra el perímetro de cada uno de los espacios naturales declarados en la isla.

Los parajes de *Las Lagunetas*, *Las Palomas*, *La Resbala* y *Pioneris* fueron completamente arrasados por las llamas. Por otro lado, quedaron muy afectadas las zonas de la *Corona Forestal* y *Las Coladas del Teide*.

Para esta visualización se han hecho uso de dos capas de dato referentes al **Área quemada**. **Últimos datos**, que se encuentran alojados en el [site](#) oficial del ISTAC en formato XML, junto con los datos que nos muestran el perímetro de los espacios naturales en formato de capa [wms](#).

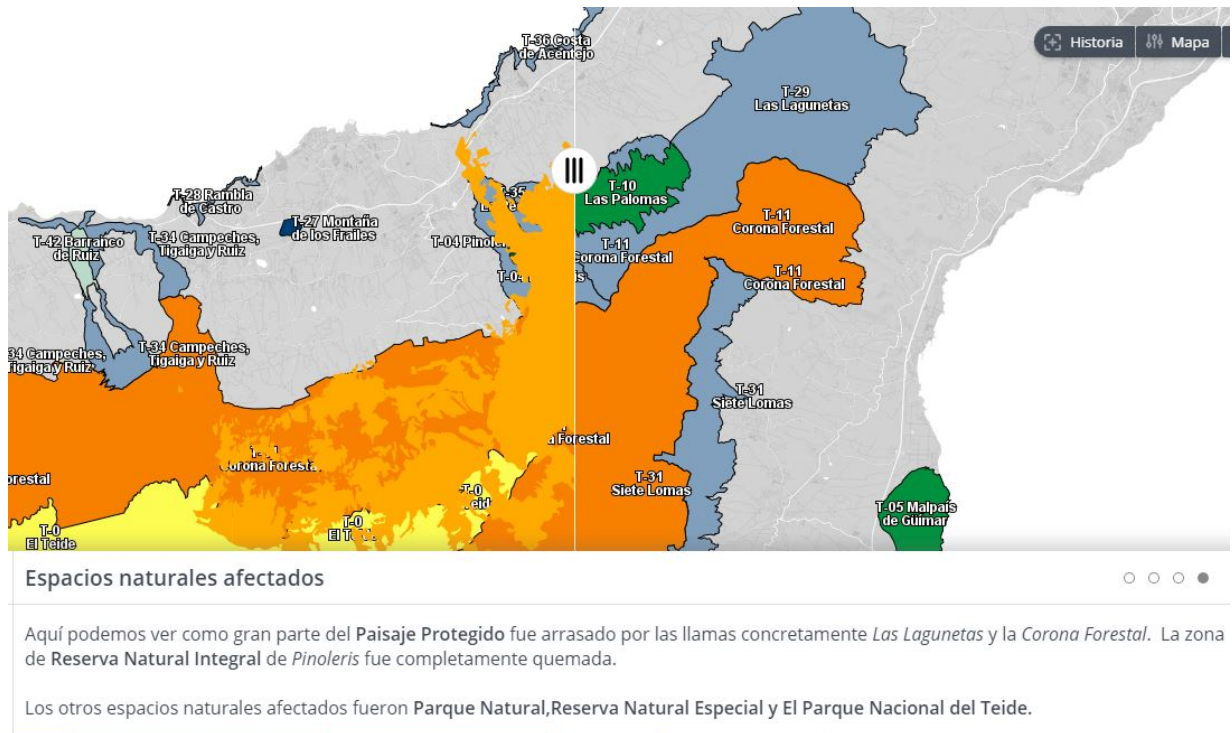


Figura 3.4: Captura de la cuarta historia del mapa interactivo.

El incendio de Tenerife es un recordatorio de la importancia de la prevención de incendios forestales. Las autoridades canarias han anunciado que reforzarán las medidas de prevención en la isla, con el objetivo de evitar que se produzcan nuevos incendios de gran magnitud. Para hacer uso de la herramienta y visualizar la historia que acabamos de redactar tenemos el siguiente enlace:

[Enlace al mapa interactivo creado sobre la plataforma de ISTAC](#)

3.2. Modelo de predicción de incendios con base en los incendios de Canadá

En esta línea vamos a desarrollar un modelo capaz de predecir cuando un incendio llegará a la categoría de sexta generación en función de las variables físicas del entorno y de la incidencia humana.

Para ello, se hará uso de los datos abiertos que proporciona **Alberta**, una base de datos del Gobierno de Canadá que contiene toda la información sobre los incendios en su territorio. [10] Se adjunta el estudio completo que ha sido desarrollado en la herramienta RStudio, como Anexo al final de este documento.

3.2.1. Análisis del conjunto de datos

La información de este conjunto de datos se relaciona con la base de datos histórica de incendios forestales de **Alberta** generada entre 2006 y 2021.

Contiene los incendios forestales que tienen un punto de origen dentro del Área de Protección Forestal (FPA), ya sean en tierras provinciales o en tierras federales. En cambio, los incendios que tienen un punto de origen fuera del FPA, se denominan incendios de ayuda mutua (MA), y no están cubiertos en este conjunto de datos.

Tiene un tamaño de 22.914 registros y 50 columnas, pero en nuestro primer cribado sobre las variables realmente valiosas del estudio, seleccionamos 21 columnas.

La relación de campos que contiene el conjunto de datos resultante es la siguiente:

1. **fire year**

Es el año del incendio forestal, comprende desde las 00:00 horas del 1 de abril hasta las 23:59 horas del 31 de marzo. Esto es porque la temporada de incendios forestales va desde el 1 de abril al 31 de marzo del año siguiente. Por ejemplo, el año de incendios 2013 comienza el 1 de abril de 2013 y finaliza el 31 de marzo de 2014.

2. **fire number**

Es el identificador único del incendio.

La primera letra indica el área forestal en la que se encuentra el incendio. C–Calgary, E–Edson, H–High Level, G–Grande Prairie, L–Lac La Biche, M–Fort McMurray, P–Peace River, R–Rocky, S–Slave Lake y W–Whitecourt. Luego, WF indica que es un incendio forestal, en inglés Wildfire. Por último los tres dígitos (con ceros a la izquierda) enumeran los incendios forestales consecutivamente dentro de cada área forestal conforme van apareciendo, es decir, es un secuencial. Importante, que si se inicia en un área forestal y se extiende a otra/as seguirá perteneciendo a el área forestal de inicio.

3. **current size**

Es el tamaño de la zona quemada por el incendio forestal en su última actualización. Expresado en hectáreas (ha).

4. **size class**

Se trata de una clasificación de los incendios forestales en función del área total de hectáreas quemadas. El desglose es el siguiente: Clase A = 0 a 0,1 ha; Clase B > 0,1 ha a 4,0 ha; Clase C > 4,0 ha a 40,0 ha; Clase D > 40,0 ha a 200 ha; Clase E > 200 ha.

5. **fire location latitude**

Ubicación de la zona de inicio del incendio forestal registrada mediante GPS. (Latitud)

6. fire location longitude

Ubicación de la zona de inicio del incendio forestal registrada mediante GPS. (Longitud)

7. fire origin

Este campo indica quién es el propietario o administrador del terreno donde arde el incendio forestal. Si el incendio abarca más de un origen, éste será el terreno donde se encuentre la zona de inicio del incendio. Las categorías son: Provincial, Private, Indian Reserve, Provincial Park, Department of National Defence, Metis Settlement.

8. general cause desc

Aquí se indica el causante que propició el incendio. Se categoriza en los siguientes tipos: AGRICULTURE INDUSTRY, FOREST INDUSTRY, GOVERNMENT, INCENDIARY, LIGHTNING, OIL AND GAS INDUSTRY, OTHER INDUSTRY, POWER LINE INDUSTRY, RAILROAD, RECREATION, RESIDENT y UNDETERMINED. En el apartado del glosario se describe el significado de cada uno de ellos.

9. fire start date

La hora y fecha en que comenzó el incendio forestal.

10. det agent

Es el identificador del agente de detección que descubrió el incendio forestal.

11. reported date

La hora y fecha en que se informó por primera vez del incendio forestal a Wildfire Management.

12. fire spread rate

La tasa de propagación del incendio forestal en el momento de la evaluación inicial, captura en metros por minuto.

13. fire type

Este es el predominante comportamiento del fuego que se observó cuando se evaluó originalmente el incendio. Se divide en las tres categorías de propagación: Ground Fire, Surface Fire y Crown Fire.

14. fire position on slope

La posición del incendio forestal en relación con la pendiente por la que avanza en el momento de la evaluación inicial. Sus categorías son: Bottom Flat, Lower 1/3, Middle 1/3, Upper 1/3.

15. **weather conditions over fire**

Descripción de las condiciones climáticas sobre el incendio forestal en el momento de la evaluación inicial. Sus categorías son: CB dry, CB wet, Clear, Cloudy y Rain showers.

16. **temperature**

La temperatura en el lugar del incendio forestal en el momento de la evaluación inicial, expresada en grados Celsius.

17. **relative humidity**

La humedad relativa en el lugar del incendio forestal en el momento de la evaluación inicial.

18. **wind speed**

La velocidad del viento en el lugar del incendio forestal en el momento de la evaluación inicial, expresada en kilómetros por hora.

19. **bucketing on fire**

Indicador de Si/NO sobre si se utilizó un helicóptero con cubeta de agua durante la acción inicial del incendio forestal.

20. **ex fs date**

Esta es la fecha y hora en que el Comandante del Incidente determinó que el incendio forestal estaba completamente extinguido (EX).

21. **ex hectares**

El tamaño del incendio forestal en el momento en que el Comandante del Incidente cambió el estado a Extinguido (EX), registrado en hectáreas totales.

3.2.2. Preprocesamiento del conjunto de datos

Haciendo uso del lenguaje de programación de R a través de RStudio y las librerías de tratamiento de datos como **dplyr**, se realizarán las transformaciones necesarias para que el conjunto de datos pueda ser objeto de modelos predictivos.

Lo primero que se ha abordado ha sido la transformación de las columnas que contienen datos de fecha, para que éstas aparezcan debidamente formateadas. Aquí hemos actuado sobre: *fire start date*, *reported date* y *ex fs date*.

El siguiente paso ha sido el tratamiento de nulos, después de analizar profundamente el conjunto de datos se ha decidido que, los registros que contienen algún valor nulo en variables marcadas como críticas” se quedarán fuera del conjunto. Estas variables son: *fire start date*, *reported date*, *wind speed* y *relative humidity*.

Para los nulos que se han encontrado en la variable *bucketing on fire*, éstos se han transformado al valor "N", equivalente a que no se ha hecho uso de un helicóptero con cubeta de agua.

Por último, se ha hecho uso de la imputación de valores para solventar las observaciones que aún seguían conteniendo nulos. La imputación se refiere al proceso de estimar o prever los valores que faltan en un conjunto de datos, y en este caso, se hará uso del algoritmo KNN (k-Nearest Neighbors).

La idea detrás de la imputación con KNN es que los valores ausentes en un conjunto de datos pueden ser estimados basándose en los valores observados más cercanos en el espacio de características. El algoritmo KNN calcula la distancia entre los registros que contienen datos ausentes y los registros vecinos que tienen datos completos. Luego, utiliza los valores de los vecinos más cercanos para predecir o imputar los valores ausentes. [2]

El proceso de imputación con KNN se puede describir en los siguientes pasos:

1. **Calcular distancias:** Calcular la distancia entre los registros con datos ausentes y todos los demás registros en función de las características compartidas.
2. **Identificar vecinos más cercanos:** Seleccionar los k registros más cercanos (vecinos) en función de la distancia calculada.
3. **Imputar valores ausentes:** Calcular un valor promedio o ponderado de los valores de los vecinos seleccionados y utilizarlo para imputar los valores ausentes en el registro de interés.

Es importante elegir un valor apropiado para el parámetro k, que representa el número de vecinos a tener en cuenta. Un valor más alto de k suavizará la imputación, mientras que un valor más bajo podría hacerla más sensible a valores atípicos.

En el conjunto de datos del estudio, hemos aplicado un valor de k=11 y las variables que han sido objeto de imputación por presentar algún valor nulo, han sido: *fire origin*, *fire spread rate*, *fire type*, *fire position on slope*, *weather conditions over fire*.

Al final de todos estos procesos, se obtiene un conjunto de datos listo para aplicar sobre él modelos predictivos. Su tamaño ahora es de 20.174 registros y 21 columnas.

3.2.3. Visualización del conjunto de datos

Haciendo uso de las variables *fire location latitude* y *fire location longitude* y después de un pequeño tratamiento de datos, se podrán obtener las coordenadas del inicio de los incendios. Con esta información, se muestran los puntos sobre un mapa base, a modo informativo en la Figura 3.5.

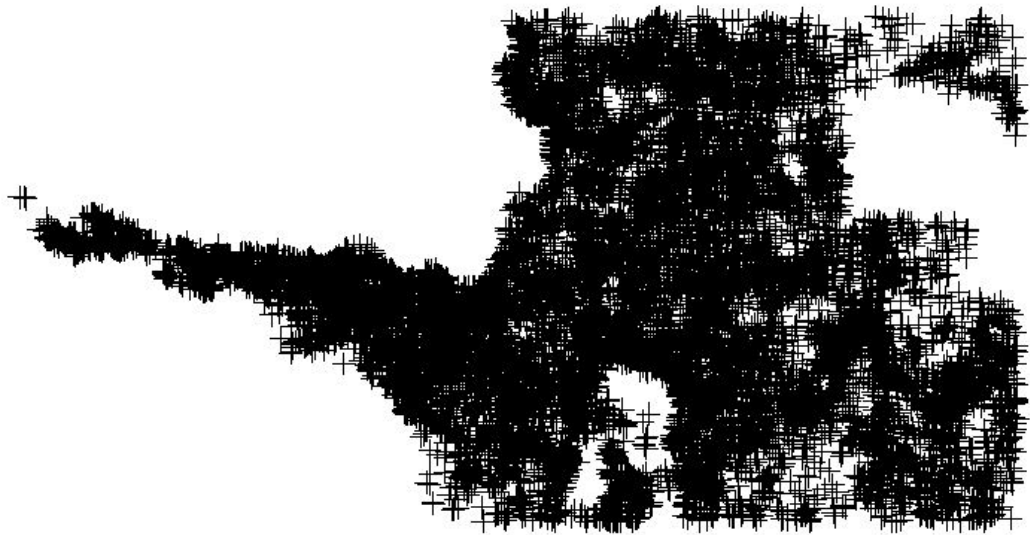


Figura 3.5: Mapa de incendios en Canadá entre 2006-2021.

3.2.4. Ejecución de modelos predictivos sobre el conjunto de datos

El objetivo de la simulación de éstos modelos va a ser la de predecir bajo que condiciones un incendio forestal puede llegar a convertirse en incendio de sexta generación. Para ello se lanzará el **algoritmo C50** que se basa en los árboles de decisión. [14]

Una vez se dispone del conjunto de datos debidamente preprocesado por los puntos anteriores, pasamos a seleccionar las variables de estudio que entraran al modelo. Como el algoritmo de C50 es un algoritmo supervisado, lo primero que se necesita es identificar cual será la variable clase del modelo. En este caso, la variable *size class* será la indicada, ya que, contiene la información del tamaño de los incendios en categorías que van desde la **A**, para los más pequeños, hasta la **D**, para los declarados megaincendios.

En cuanto a las variables que van a explicar los datos y sobre las que se lanzará la predicción son: *general cause desc*, *weather conditions over fire*, *temperature*, *relative humidity*, *wind speed*. Para que el modelo pueda ser válido necesitamos dividir en dos el conjunto de datos. Por un lado seleccionaremos 2/3 de las filas y lo llamaremos conjunto de Entrenamiento. Y por otro lado, los 1/3 restantes pasarán a ser el conjunto de Test. Esto es esencial para poder calcular la precisión del modelo después de ser entrenado.

Una vez que se ejecuta el script, obtenemos el resultado del Cuadro 3.1.

- Se observa que se han generado 128 reglas de clasificación con errores del 32,5 %.
- Las clasificaciones en las categorías **A** y **B** que corresponden con los incendios menores parecen poco precisas.
- En cuanto al peso y uso que el algoritmo ha hecho de las variables de entrada, obtenemos esta clasificación:
 - 100 % wind speed
 - 50,61 % relative humidity
 - 44,75 % general cause desc
 - 32,17 % temperature
 - 18,44 % weather conditions over fire

La precisión final que ha arrojado el modelo es del 61,83 %, por lo que está lejos del objetivo que se fijó al principio de este estudio, en el que no será válida una precisión por debajo del 70 %.

```

## Evaluation on training data (13449 cases):
##
##      Rules
##      -----
##      No      Errors
##
##      128 4369(32.5%)  <<
##
##
##      (a)  (b)  (c)  (d)  (e)  <-classified as
##      ----  ----  ----  ----  ----
##      7953  507    1      1      1  (a): class A
##      2821 1056    6      1      1  (b): class B
##      545   118   44      2      2  (c): class C
##      147   32      2      25     2  (d): class D
##      159   29      25     25     25  (e): class E
##
##
## Attribute usage:
##
## 100.00% wind_speed
## 50.61% relative_humidity
## 44.75% general_cause_desc
## 32.17% temperature
## 18.44% weather_conditions_over_fire
##
##
## Time: 0.9 secs

```

Cuadro 3.1: Resumen de la ejecución del modelo C50.

Después de estos resultados se deberá proceder a ajustar el modelo inicial.

Para ello, la medida adoptada ha sido la de eliminar los registros que corresponden a los incendios menores, clasificados como **A**. Como se ve en los resultados anteriores, estos registros son muy numerosos y distorsionan gravemente la precisión de predicción del modelo. Como al final, la idea principal va enfocada en los incendios etiquetados como **E** y **D**, eliminar estos registros de la categoría **A** puede suponer un gran beneficio. Al eliminar estos registros, el conjunto de datos pasa a tener 7.503 filas.

Ahora las variables que van a explicar los datos y sobre las que se lanzará la predicción son: *general cause desc*, *fire type*, *fire position on slope*, *weather conditions over fire*, *temperature*, *relative humidity*, *wind speed*.

Básicamente, se añaden *fire type* y *fire position on slope* a las ya existentes con anterioridad.

Una vez que se ejecuta el modelo ajustado, obtenemos el resultado del Cuadro 3.2 y Cuadro 3.3.

- Se observa que se han generado 17 reglas de clasificación con errores del 21,1 %.
- Las clasificaciones en la categoría **B** siguen reflejando mucha dispersión.
- En cuanto al peso y uso que el algoritmo ha hecho de las variables de entrada, obtenemos esta clasificación:

- 100,00 % fire type
- 59,34 % relative humidity
- 57,32 % wind speed
- 21,69 % general cause desc
- 16,31 % weather conditions over fire
- 0,66 % temperature
- 0,50 % fire position on slope

La precisión final que ha arrojado el modelo ajustado es del 78,17 %, por lo que ahora si que se cumple con el objetivo de alcanzar al menos un 70 % de precisión.

```
## Rules
## -----
## No Errors
##
## 17 1055(21.1%) <<
##
## (a) (b) (c) (d) <-classified as
## ---
## 3848 13 3 14 (a): class B
## 643 40 4 17 (b): class C
## 175 12 16 5 (c): class D
## 147 18 4 43 (d): class E
##
```

Cuadro 3.2: Resumen de la ejecución del modelo C50 ajustado.

```
## 100.00% fire_type
## 59.34% relative_humidity
## 57.32% wind_speed
## 21.69% general_cause_desc
## 16.31% weather_conditions_over_fire
## 0.66% temperature
## 0.50% fire_position_on_slope
##
##
## Time: 0.0 secs
```

Cuadro 3.3: Uso de las variables en el modelo C50 ajustado.

Por último, se intentó entrenar el modelo eliminando también los incendios categorizados como **B** pero en este caso no solo no mejoraba la precisión actual, sino que empeoraba los resultados del modelo inicial que fueron del 61,83 %.

Capítulo 4

Conclusiones y trabajos futuros

4.1. Conclusiones enfocadas al análisis de incendios

Los incendios de sexta generación son una nueva tipología de incendios forestales que se distinguen por su enorme intensidad, la capacidad de generar sus propias condiciones meteorológicas y la dificultad inherente a su extinción. Estos incendios, cada vez más frecuentes y devastadores, están estrechamente vinculados al cambio climático y al abandono de las zonas boscosas.

En los últimos años, se han registrado diversos incendios de sexta generación en diversas partes del mundo, causando considerables daños tanto en términos ambientales como económicos. En España, los incendios en Sierra Bermeja (2022) y Sierra de la Culebra (2022), junto con el estudiado incendio de la isla de Tenerife (2023) son ejemplos patentes de esta nueva realidad. En cuanto a las medidas de prevención en la actualidad sobre este tipo de incendios, se requiere que estén soportadas por una estrategia integral y compleja.

Podremos enumerar las siguientes:

- Reducción del riesgo de propagación: Esto implica la implementación de prácticas de gestión forestal sostenible, que abarquen desde la disminución de la carga combustible hasta la mejora de la conectividad del paisaje y la creación de cortafuegos.
- Mejora de la capacidad de extinción: Esto se logra mediante el desarrollo de nuevos equipos y tecnologías, así como la capacitación continua de los equipos de extinción.
- Preparación de la población: Es fundamental que la población esté debidamente preparada para enfrentar los incendios forestales, comprendiendo las medidas de auto-protección y los planes de emergencia.

Ante la amenaza persistente de los incendios de sexta generación, resulta imperativo continuar trabajando en la prevención y en el fortalecimiento de las capacidades de extinción. Asimismo, se destaca la importancia de la investigación continua para una comprensión más profunda de estos incendios y el desarrollo de estrategias innovadoras para combatirlos.

De aquí, sacamos algunas recomendaciones para el futuro, como por ejemplo:

- Integración de la prevención de incendios de sexta generación en las políticas públicas: Es crucial que los gobiernos e instituciones se comprometan activamente a reducir el riesgo de incendios forestales.
- Inversión en investigación: Es esencial seguir invirtiendo en investigaciones que profundicen en la comprensión de estos incendios, así como en el desarrollo de nuevas medidas para enfrentarlos.
- Fomento de la colaboración internacional: La lucha contra los incendios forestales constituye un desafío global, por lo que es imperativo que los países colaboren estrechamente para hacerle frente.

4.2. Conclusiones enfocadas a los Objetivos de Desarrollo Sostenible

Es importante desarrollar planes de respuesta a los incendios forestales que tengan en cuenta los impactos en la igualdad de género y la reducción de las desigualdades. Estos planes deben incluir medidas para proteger a las personas y los medios de vida, así como para restaurar los ecosistemas afectados.

- Al reducir nuestro consumo de recursos naturales y mejorar la eficiencia de nuestros procesos de producción, podemos ayudar a prevenir estos incendios.
- Reducir las emisiones de gases de efecto invernadero: Esto ayudará a mitigar el cambio climático, que es uno de los principales factores que contribuyen a los incendios de sexta generación.
- Implementar prácticas de gestión forestal sostenible: Esto ayudará a reducir la acumulación de combustible vegetal, que es un factor de riesgo de incendios.
- Educar al público sobre los incendios forestales: Esto ayudará a las personas a estar más preparadas para los incendios y a tomar medidas para prevenirlos.

4.3. Conclusiones basadas en la predicción del modelo con datos de Canadá

En este apartado se exponen los resultados que se han obtenido del estudio con el modelo predictivo.

Se obtuvieron 17 reglas de clasificación con el mejor de los modelos y sobre dichas reglas hemos obtenido la siguiente información:

- Según la regla número 13, con unas probabilidades del 71,4 %. Un incendio llegará a la categoría **E** si la causa del incendio es de **forma intencionada**, si el tipo de propagación del incendio es el de **fuego de copas** y si la velocidad del **viento supera los 14 m/s**.
- Según la regla número 8, con unas probabilidades del 80 %. Un incendio llegará a la categoría **D** si la causa del incendio es de forma accidental mientras se practicaba una **actividad recreativa en el bosque**, si el tipo de propagación del incendio es el de **fuego de copas** y si la **humedad relativa de la zona es igual o menor al 24 %**.

En paralelo, si quisiéramos sacar unas conclusiones sobre los incendios forestales que se quedan en pequeños incendios y no prosperan en su avance, podemos mostrar estos datos:

- Según la regla número 2, con unas probabilidades del 82,2 %. Un incendio no superará la categoría **B** si la **humedad relativa de la zona es mayor al 29 %** y si la velocidad del **viento no supera los 12 m/s**.
- Según la regla número 4, con unas probabilidades del 81,1 %. Un incendio tampoco superará la categoría **B** si las condiciones meteorológicas son de **chubascos con nubes** y la **humedad relativa de la zona supera el 34 %**.

Capítulo 5

Glosario

5.1. Glosario de términos

- **Herramienta GIS:** Un Sistema de Información Geográfica, o GIS por sus siglas en inglés (Geographic Information System), es una herramienta informática que permite capturar, almacenar, analizar y visualizar datos geoespaciales. Los datos geoespaciales son aquellos que están asociados a ubicaciones geográficas específicas en la Tierra. Las herramientas GIS integran hardware, software y datos para gestionar, analizar y representar información geográfica de manera eficiente.
- **Metodología Agile:** Es un enfoque de desarrollo de proyectos que se centra en la flexibilidad, colaboración y adaptabilidad a los cambios a lo largo del tiempo. Este enfoque es especialmente aplicable en entornos donde los requisitos y objetivos del proyecto pueden cambiar y evolucionar de manera significativa durante el proceso de desarrollo. La metodología ágil se basa en los principios establecidos en el "Manifiesto Ágil", un documento que destaca los valores y principios clave para guiar el desarrollo de software y otros proyectos.
- **Pirófitas:** Son plantas adaptadas al fuego, es decir, especies vegetales que han evolucionado para sobrevivir y prosperar en entornos afectados por incendios recurrentes. Estas plantas han desarrollado diversas estrategias que les permiten resistir el fuego, regenerarse después de un incendio y, en algunos casos, incluso beneficiarse de los incendios.
- **Xerófitas:** Son plantas adaptadas para sobrevivir y prosperar en ambientes secos y con escasez de agua. Estas plantas han desarrollado diversas características morfológicas y fisiológicas que les permiten resistir la sequía y conservar el agua de manera eficiente. Las xerófitas se encuentran comúnmente en regiones desérticas y otros hábitats áridos, pero

también pueden adaptarse a condiciones de sequía en otros entornos.

- **Precipitaciones ocultas:** Se refieren al proceso mediante el cual el vapor de agua en la atmósfera se condensa y forma gotas de agua o cristales de hielo sin que estos alcancen directamente la superficie terrestre en forma de lluvia o nieve. En lugar de caer directamente como precipitación visible, estas gotas o cristales pueden permanecer suspendidos en la atmósfera o ser transportados por el viento antes de llegar al suelo. Este fenómeno es también conocido como "precipitación virga".
- **Selvícola:** Se refiere a las actividades y prácticas relacionadas con el manejo y cultivo de bosques. La silvicultura desempeña un papel clave en la gestión sostenible de los recursos forestales, contribuyendo a la conservación y el uso racional de los ecosistemas forestales.
- **Corrientes de aire convectivas:** Son movimientos de aire verticales que se producen debido a diferencias de temperatura en la atmósfera. Este fenómeno está asociado con el proceso de convección, que implica el transporte de calor a través del movimiento de un fluido (en este caso, el aire). Las corrientes de aire convectivas son comunes en la atmósfera y desempeñan un papel importante en diversos fenómenos meteorológicos, incluidos los incendios forestales.
- **Pirocumulonimbos:** Son nubes de tormenta que se forman a partir de la actividad de incendios forestales o incendios en paisajes más extensos. Estas nubes tienen características similares a los cumulonimbos convencionales, pero su origen está vinculado a la intensa actividad térmica y el calor generado por el fuego.
- **Metis Settlement:** Son áreas de tierra en Alberta, Canadá, que fueron establecidas específicamente para la población Métis. Los Métis son una población originaria de Canadá con raíces tanto indígenas como europeas. Tienen una identidad cultural y étnica única que se desarrolló históricamente a través del mestizaje entre personas indígenas y europeas, principalmente durante el período de la exploración y la colonización europea en América del Norte.
- **dplyr:** Es una librería del lenguaje de R, que proporciona funciones para la manipulación de datos, como filtrado, selección, agrupación y resumen.

5.2. Glosario de descripciones del conjunto de datos

- **AGRICULTURE INDUSTRY:** Un incendio forestal resultante de una actividad agrícola o ganadera realizada por personas o máquinas (ganado o cultivos). Esta causa general se introdujo en 2012. Para los incendios forestales de 2011 y anteriores, la industria agrícola quedó incluida en la causa residente.
- **FOREST INDUSTRY:** Un incendio forestal provocado por personas o máquinas utilizadas en cualquier actividad asociada a la producción de fibra forestal. Esta categoría incluye fábricas de celulosa, aserraderos y operaciones de madera contrachapada/paneles/postes relacionadas con la inspección, topografía, protección, escalado, silvicultura, etc. de la industria de recursos.
- **GOVERNMENT:** Un incendio forestal iniciado por empleados, agentes o contratistas del gobierno. Esta causa general se introdujo en 2012. Para los incendios forestales de 2011 y anteriores, el Gobierno quedó incluido en la causa de Otras Industrias.
- **INCENDIARY:** Un incendio forestal que se ha iniciado intencionalmente en un área o bajo circunstancias donde y cuando no debería haber un incendio.
- **LIGHTNING:** Un incendio forestal provocado por un rayo.
- **OIL AND GAS INDUSTRY:** Un incendio forestal causado por cualquier máquina, empleado, agente o contratista que realiza trabajos asociados con la exploración, construcción, producción o transmisión por parte de una empresa de petróleo y gas, incluidas las empresas de arenas bituminosas.
- **OTHER INDUSTRY:** Un incendio forestal causado por una industria distinta a la de líneas eléctricas, petróleo y gas, bosques, agricultura o ferrocarriles. Esta causa general también requiere una identificación de la industria.
- **POWER LINE INDUSTRY:** Un incendio forestal causado por cualquier máquina, empleado, agente o contratista que realiza trabajos asociados con una compañía de líneas eléctricas, cualquier animal o árbol que interactúe con una línea eléctrica, o cualquier hardware o mecánico de la línea eléctrica.
- **RAILROAD:** Un incendio forestal causado por cualquier máquina, empleado, agente o contratista que realiza trabajos asociados con una operación ferroviaria.
- **RECREATION:** Un incendio forestal causado por cualquier persona o equipo involucrado en una actividad recreativa, por ejemplo: acampar, cazar, pescar, recolectar bayas, ciclistas, excursionistas, etc. Esta causa general también requiere de un grupo responsable.

- **RESIDENT:** Un incendio forestal resultante de una actividad realizada por personas o máquinas en terrenos residenciales o cualquier incendio forestal causado por una actividad asociada con la vida normal, por ejemplo: fuego en un área boscosa de residencias/edificios, quema de escombros, fumar, accidentes de vehículos que provocan incendios forestales, etc.
- **UNDETERMINED:** Un incendio forestal que ha sido investigado pero no se puede determinar la causa. Ante mayor evidencia o nueva información, se podrá determinar la causa final y esta causa podrá ser actualizada.

Bibliografía

- [1] Violeta Cubero. Incendios forestales en 2023: los inapagables fuegos de sexta generación. <https://www.plataformatierra.es/actualidad/incendios-forestales-en-2023-sexta-generacion/>, Julio 2023. Fecha de acceso: 08-10-2023.
- [2] Top Big Data. knn imputación por valores perdidos en el aprendizaje automático. https://topbigdata.es/knn-imputacion-por-valores-perdidos-en-el-aprendizaje-automatico/?utm_content=cmp-true, Julio 2020. Fecha de acceso: 25-11-2023.
- [3] Ricardo A. Salas de la Vega. Las generaciones de los incendios forestales. <https://academiamalaguenaciencias.wordpress.com/2022/11/05/las-generaciones-de-los-incendios-forestales/>, Noviembre 2022. Fecha de acceso: 18-10-2023.
- [4] Agencia Estatal Boletín Oficial del Estado. Artículo 5. ley 43/2003, de 21 de noviembre, de montes. <https://www.boe.es/eli/es/l/2003/11/21/43/cona5>, Noviembre 2003. Fecha de acceso: 09-10-2023.
- [5] Agencia Estatal Boletín Oficial del Estado. Artículo 6. ley 43/2003, de 21 de noviembre, de montes. <https://www.boe.es/eli/es/l/2003/11/21/43/cona6>, Noviembre 2003. Fecha de acceso: 09-10-2023.
- [6] Bosques y Desertificación Dirección General de Biodiversidad. Prevención y divulgación sobre incendios forestales. <https://www.miteco.gob.es/es/biodiversidad/temas/incendios-forestales/prevencion/divulgativa50.html>. Fecha de acceso: 21-10-2023.
- [7] Science Media Centre España. Incendios de sexta generación: qué son, cómo les afecta el cambio climático y formas de prevenirlos. <https://sciencemediacentre.es/incendios-de-sexta-generacion-que-son-como-les-afecta-el-cambio-climatico-y-formas-de-prevenirlos>, Agosto 2023. Fecha de acceso: 10-10-2023.
- [8] ASHES TO LIFE. Conocer para prevenir los incendios forestales. <https://www.ashestolife.es/conocer-para-prevenir-los-incendios-forestales/>, Diciembre 2019. Fecha de acceso: 20-10-2023.

-
- [9] Gobierno Local. Tenerife: el incendio más grave de 2023. <https://www.femp.es/sites/default/files/multimedia/>. Fecha de acceso: 12-11-2023.
- [10] Government of Alberta. Alberta, wildfire maps and data. <https://www.alberta.ca/wildfire-maps-and-datajumplinks-2>. Fecha de acceso: 15-11-2023.
- [11] A. Larrañaga M. Miralles P. Daniel P. Costa, M. Castellnou. La prevención de los grandes incendios forestales adaptada al incendio tipo. <https://previfor.org/wp-content/uploads/2020/05/la-prevencion-de-los-grandes-incendios-forestales-adaptada-al-incendio-tipo.pdf>, Abril 2023. Fecha de acceso: 15-10-2023.
- [12] Asociación para la Prevención de Incendios Forestales. Conocer para prevenir los incendios forestales. <https://previfor.org/tipos-de-clasificacion-de-los-incendios-forestales/>, Abril 2023. Fecha de acceso: 10-10-2023.
- [13] José Ramón Pérez. Datos y mapas sobre el incendio de tenerife, el más extenso registrado en España en 2023. <https://www.newtral.es/incendio-tenerife-superficie-quemada/20230824/>, Agosto 2023. Fecha de acceso: 10-11-2023.
- [14] Joaquín Amat Rodrigo. Árboles de decisión, random forest, gradient boosting y c5.0. https://cienciadedatos.net/documentos/33_arboles_decision_random_forest_gradient_boosting_c50, Octubre 2020. Fecha de acceso: 29-11-2023.
- [15] Bosques Sostenibles. Incendios de sexta generación, ¿qué son? <https://www.bosquessostenibles.com/blog/incendios-de-sexta-generacion-que-son/>, Julio 2022. Fecha de acceso: 08-10-2023.
- [16] TerriaJS. Herramienta de visualización geográfica terria.io. <https://terria.io/>. Fecha de acceso: 13-11-2023.

Apéndice A

Anexo: Código de R completo.

Contents

1. Carga del archivo.	1
2. Primeras transformaciones.	2
3. Tratamiento de nulos.	3
4. Imputación de valores con KNN.	5
5. Visualización geográfica.	6
6. Modelo Predictivo.	7

1. Carga del archivo.

```
# Leemos el fichero de los incendios en Canadá en formato .csv y tratamos los
# registros vacíos como NAs.
data <- read.csv('fp-historical-wildfire-data-2006-2021.csv', header=TRUE,
                 sep=';', quote = "\"", na.strings = "")
```

```
# Mostramos el número de filas y columnas.
dim(data)
```

```
## [1] 22914    50
```

```
# Mostramos la estructura del conjunto de datos.
str(data)
```

```
## 'data.frame':    22914 obs. of  50 variables:
## $ fire_year      : int  2021 2021 2021 2021 2021 2021 2021 2021 2021 2021 2021 2021 ...
## $ fire_number    : chr  "CWF001" "CWF002" "CWF003" "CWF004" ...
## $ fire_name      : chr  NA NA NA NA ...
## $ current_size   : chr  "0,01" "0,01" "0,05" "0,15" ...
## $ size_class     : chr  "A" "A" "A" "B" ...
## $ fire_location_latitude : chr  "49,410191" "49,587753" "51,1195" "51,271888" ...
## $ fire_location_longitude : chr  "-114,318843" "-114,378655" "-114,77895" "-115,025818" ...
## $ fire_origin    : chr  "Provincial Park" "Provincial Land" "Provincial Land" "Provinc:
## $ general_cause_desc : chr  "Power Line Industry" "Recreation" "Resident" "Incendiary" ...
## $ industry_identifier_desc : chr  NA NA NA NA ...
## $ responsible_group_desc : chr  NA "Hunters" NA NA ...
## $ activity_class  : chr  "Operations" "Cooking and Warming" "Structure Fire" "Unclassif:
## $ true_cause      : chr  "Line Impact" "Unsafe Fire" NA NA ...
## $ fire_start_date : chr  "13/01/2021 10:30" "28/01/2021 8:00" "22/02/2021 16:30" "14/03,
## $ det_agent       : chr  "310" "310" "GOVT" "310" ...
## $ det_agent_type  : chr  "UNP" "UNP" "UNP" "UNP" ...
## $ discovered_date : chr  "13/01/2021 11:35" NA NA NA ...
## $ discovered_size : logi  NA NA NA NA NA NA ...
## $ reported_date   : chr  "13/01/2021 11:46" "28/01/2021 17:08" "22/02/2021 16:49" "14/0:
## $ dispatched_resource : chr  "FPD Staff" "FPD Staff" "FPD Staff" "FPD Staff" ...
## $ dispatch_date   : chr  "13/01/2021 11:51" "29/01/2021 8:04" "23/02/2021 9:00" "14/03/:
## $ start_for_fire_date : chr  "13/01/2021 12:15" "29/01/2021 8:34" "23/02/2021 9:30" "14/03/:
## $ assessment_resource : chr  "IA Forces" "IA Forces" "IA Forces" "IA Forces" ...
```

```
## $ assessment_datetime      : chr "13/01/2021 13:25" "29/01/2021 10:15" "23/02/2021 10:42"
## $ assessment_hectares     : chr "0,01" "0,01" "0,05" "0,15" ...
## $ fire_spread_rate        : chr "0" "0" "0" "0" ...
## $ fire_type               : chr "Surface" "Ground" "Surface" "Surface" ...
## $ fire_position_on_slope   : chr "Flat" "Middle 1/3" "Flat" "Flat" ...
## $ weather_conditions_over_fire: chr "Cloudy" "Clear" "Cloudy" "Cloudy" ...
## $ temperature             : chr "4" "-5" "-5" "12" ...
## $ relative_humidity       : int 65 90 42 30 26 50 45 40 38 50 ...
## $ wind_direction          : chr "W" "CLM" "W" "SW" ...
## $ wind_speed              : int 30 0 5 2 14 5 5 30 25 5 ...
## $ fuel_type               : chr "C7" "C2" "C3" "01b" ...
## $ initial_action_by       : chr "FPD Staff" "FPD Staff" "Fire Department" "Fire Department" ...
## $ ia_arrival_at_fire_date  : chr "13/01/2021 13:20" "29/01/2021 10:00" NA NA ...
## $ ia_access               : chr NA NA NA NA ...
## $ fire_fighting_start_date : chr "13/01/2021 13:25" "29/01/2021 10:45" NA NA ...
## $ fire_fighting_start_size : chr "0,01" "0,01" NA NA ...
## $ bucketing_on_fire       : chr "N" "N" NA NA ...
## $ distance_from_water_source : chr NA NA NA NA ...
## $ first_bucket_drop_date   : chr NA NA NA NA ...
## $ bh_fs_date              : chr "13/01/2021 13:25" "29/01/2021 10:45" "23/02/2021 10:42"
## $ bh_hectares             : chr "0,01" "0,01" "0,05" "0,15" ...
## $ uc_fs_date              : chr "13/01/2021 13:25" "29/01/2021 10:45" "23/02/2021 10:42"
## $ uc_hectares             : chr "0,01" "0,01" "0,05" "0,15" ...
## $ to_fs_date              : chr NA NA NA NA ...
## $ to_hectares             : chr NA NA NA NA ...
## $ ex_fs_date              : chr "13/01/2021 13:35" "29/01/2021 13:15" "23/02/2021 10:52"
## $ ex_hectares             : chr "0,01" "0,01" "0,05" "0,15" ...
```

2. Primeras transformaciones.

```
# Seleccionamos las variables del estudio
columnas_estudio <- c(
  "fire_year", "fire_number", "current_size", "size_class",
  "fire_location_latitude", "fire_location_longitude", "fire_origin",
  "general_cause_desc", "fire_start_date", "det_agent", "reported_date",
  "fire_spread_rate", "fire_type", "fire_position_on_slope",
  "weather_conditions_over_fire", "temperature", "relative_humidity",
  "wind_speed", "bucketing_on_fire", "ex_fs_date", "ex_hectares")

# Creamos un nuevo dataframe con las columnas del estudio
data_TFM <- data[, columnas_estudio]

# Pasamos los campos de fechas a su formato adecuado.
data_TFM <- data_TFM %>%
  mutate(
    fire_start_date = as.POSIXct(fire_start_date, format = "%d/%m/%Y %H:%M"),
    reported_date = as.POSIXct(reported_date, format = "%d/%m/%Y %H:%M"),
    ex_fs_date = as.POSIXct(ex_fs_date, format = "%d/%m/%Y %H:%M"))
```

3. Tratamiento de nulos.

```
# Mostramos la estructura del nuevo conjunto de datos.
```

```
str(data_TFM)
```

```
## 'data.frame':    22914 obs. of  21 variables:
## $ fire_year      : int  2021 2021 2021 2021 2021 2021 2021 2021 2021 2021 ...
## $ fire_number    : chr  "CWF001" "CWF002" "CWF003" "CWF004" ...
## $ current_size   : chr  "0,01" "0,01" "0,05" "0,15" ...
## $ size_class     : chr  "A" "A" "A" "B" ...
## $ fire_location_latitude : chr  "49,410191" "49,587753" "51,1195" "51,271888" ...
## $ fire_location_longitude : chr  "-114,318843" "-114,378655" "-114,77895" "-115,025818" ...
## $ fire_origin    : chr  "Provincial Park" "Provincial Land" "Provincial Land" "Provinci
## $ general_cause_desc : chr  "Power Line Industry" "Recreation" "Resident" "Incendiary" ...
## $ det_start_date   : POSIXct, format: "2021-01-13 10:30:00" "2021-01-28 08:00:00" ...
## $ det_agent       : chr  "310" "310" "GOVT" "310" ...
## $ reported_date    : POSIXct, format: "2021-01-13 11:46:00" "2021-01-28 17:08:00" ...
## $ fire_spread_rate : chr  "0" "0" "0" "0" ...
## $ fire_type       : chr  "Surface" "Ground" "Surface" "Surface" ...
## $ fire_position_on_slope : chr  "Flat" "Middle 1/3" "Flat" "Flat" ...
## $ weather_conditions_over_fire: chr  "Cloudy" "Clear" "Cloudy" "Cloudy" ...
## $ temperature     : chr  "4" "-5" "-5" "12" ...
## $ relative_humidity : int  65 90 42 30 26 50 45 40 38 50 ...
## $ wind_speed      : int  30 0 5 2 14 5 5 30 25 5 ...
## $ bucketing_on_fire : chr  "N" "N" NA NA ...
## $ ex_fs_date      : POSIXct, format: "2021-01-13 13:35:00" "2021-01-29 13:15:00" ...
## $ ex_hectares     : chr  "0,01" "0,01" "0,05" "0,15" ...
```

```
# Y un resumen del dataset.
```

```
summary(data_TFM)
```

```
##   fire_year   fire_number   current_size   size_class
## Min.      :2006   Length:22914   Length:22914   Length:22914
## 1st Qu.:2009   Class :character   Class :character   Class :character
## Median :2013   Mode  :character   Mode  :character   Mode  :character
## Mean      :2013
## 3rd Qu.:2016
## Max.      :2021
##
## fire_location_latitude fire_location_longitude fire_origin
## Length:22914          Length:22914          Length:22914
## Class :character      Class :character      Class :character
## Mode  :character      Mode  :character      Mode  :character
##
##
##
##   general_cause_desc fire_start_date   det_agent
## Length:22914        Min.      :0201-06-10 23:18:00   Length:22914
## Class :character    1st Qu.:2009-05-25 18:40:00   Class :character
## Mode  :character    Median :2013-05-05 17:40:00   Mode  :character
##                    Mean      :2013-03-14 02:22:59
##                    3rd Qu.:2016-08-14 14:30:00
##                    Max.      :2021-12-09 10:00:00
```

```
##          NA's      :621
## reported_date      fire_spread_rate      fire_type
## Min.      :2006-04-02 14:30:00      Length:22914      Length:22914
## 1st Qu.:2009-06-12 17:51:00      Class :character      Class :character
## Median :2013-05-06 13:36:00      Mode  :character      Mode  :character
## Mean      :2013-04-25 18:42:23
## 3rd Qu.:2016-08-25 13:08:00
## Max.      :2021-12-10 15:00:00
## NA's      :1
## fire_position_on_slope weather_conditions_over_fire temperature
## Length:22914      Length:22914      Length:22914
## Class :character      Class :character      Class :character
## Mode  :character      Mode  :character      Mode  :character
##
##
##
## relative_humidity      wind_speed      bucketing_on_fire
## Min.      : 0.00      Min.      : 0.000      Length:22914
## 1st Qu.: 31.00      1st Qu.: 3.000      Class :character
## Median : 40.00      Median : 6.000      Mode  :character
## Mean      : 45.42      Mean      : 8.763
## 3rd Qu.: 56.00      3rd Qu.:12.000
## Max.      :100.00      Max.      :90.000
## NA's      :2659      NA's      :2660
## ex_fs_date      ex_hectares
## Min.      :2006-04-03 10:20:00      Length:22914
## 1st Qu.:2009-06-13 14:43:45      Class :character
## Median :2013-05-07 17:37:30      Mode  :character
## Mean      :2013-04-29 03:19:35
## 3rd Qu.:2016-08-30 13:53:00
## Max.      :2022-06-05 17:43:00
##
```

```
# Observamos si el dataset presenta nulos en cada uno de los campos.
colSums(is.na(data_TFM))
```

```
##          fire_year      fire_number
##          0          0
##          current_size      size_class
##          0          0
## fire_location_latitude      fire_location_longitude
##          0          0
##          fire_origin      general_cause_desc
##          9          0
##          fire_start_date      det_agent
##          621          1
##          reported_date      fire_spread_rate
##          1          2609
##          fire_type      fire_position_on_slope
##          2477          2649
## weather_conditions_over_fire      temperature
##          2656          2657
##          relative_humidity      wind_speed
##          2659          2660
```

```
##          bucketing_on_fire          ex_fs_date
##                6576                0
##          ex_hectares
##                0

# Eliminamos las filas en las que no viene la fecha de inicio del fuego
# debidamente informada.
data_TFM <- data_TFM[complete.cases(data_TFM$fire_start_date), ]

# Aplicamos el mismo paso para las que no informan la velocidad del viento.
data_TFM <- data_TFM[complete.cases(data_TFM$wind_speed), ]

# Aplicamos el mismo paso para los que no tienen la humedad relativa informada.
data_TFM <- data_TFM[complete.cases(data_TFM$relative_humidity), ]

# Aplicamos el mismo paso para los que no tienen la fecha de reporte informada.
data_TFM <- data_TFM[complete.cases(data_TFM$reported_date), ]

# Transformamos los NAs del campo bucketing_on_fire a N, referentes a No ha sido
# necesaria la intervención de un helicóptero.
data_TFM$bucketing_on_fire <- replace(data_TFM$bucketing_on_fire,
                                     is.na(data_TFM$bucketing_on_fire), "N")

# Observamos si hemos mejorado con la gestión de nulos.
colSums(is.na(data_TFM))
```

```
##          fire_year          fire_number
##                0                0
##          current_size          size_class
##                0                0
##          fire_location_latitude  fire_location_longitude
##                0                0
##          fire_origin          general_cause_desc
##                9                0
##          fire_start_date          det_agent
##                0                0
##          reported_date          fire_spread_rate
##                0                6
##          fire_type          fire_position_on_slope
##                1                3
## weather_conditions_over_fire          temperature
##                3                0
##          relative_humidity          wind_speed
##                0                0
##          bucketing_on_fire          ex_fs_date
##                0                0
##          ex_hectares
##                0
```

4. Imputación de valores con KNN.

```

# Imputamos los valores faltantes tomando como muestra las variables informadas.
# Generamos un nuevo dataframe.
data_TFM_imp <- kNN(data_TFM,
  variable = c('fire_origin', 'fire_spread_rate',
    'fire_type', 'fire_position_on_slope',
    'weather_conditions_over_fire'),
  dist_var = c('current_size', 'size_class',
    'general_cause_desc', 'det_agent',
    'temperature', 'relative_humidity',
    'wind_speed', 'bucketing_on_fire',
    'ex_fs_date'),
  imp_var = FALSE, impNA = TRUE, k=11 )

```

5. Visualización geográfica.

```

# Reemplazamos las comas por puntos en los campos de coordenadas.
data_TFM_imp$fire_location_latitude <- gsub(",", ".",
  data_TFM_imp$fire_location_latitude)
data_TFM_imp$fire_location_longitude <- gsub(",", ".",
  data_TFM_imp$fire_location_longitude)

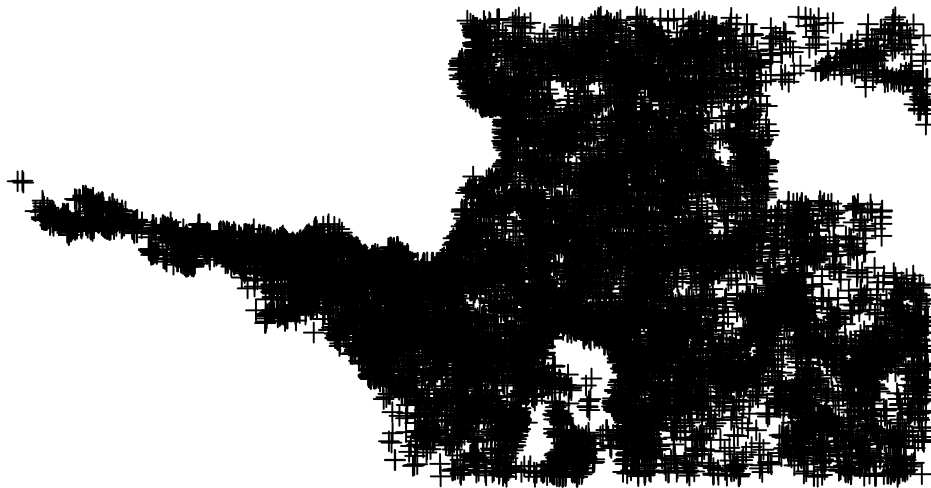
# Convertimos la columna a tipo numérico.
data_TFM_imp$fire_location_latitude <- as.numeric(data_TFM_imp$fire_location_latitude)
data_TFM_imp$fire_location_longitude <- as.numeric(data_TFM_imp$fire_location_longitude)

# Cargamos los datos geográficos.
coordenadas <- data_TFM_imp[,c(5,6)]

data_TFM_imp_geo <- SpatialPointsDataFrame(coordenadas,
  data_TFM_imp,
  proj4string = CRS("+proj=longlat +datum=WGS84 +ellps=WGS84"))

# Visualizamos en el mapa.
plot(data_TFM_imp_geo)

```



6. Modelo Predictivo.

```
# Reemplazamos las comas por puntos en los campos necesarios
data_TFM_imp$current_size <- gsub(",", ".", data_TFM_imp$current_size)
data_TFM_imp$fire_spread_rate <- gsub(",", ".", data_TFM_imp$fire_spread_rate)
data_TFM_imp$temperature <- gsub(",", ".", data_TFM_imp$temperature)
data_TFM_imp$ex_hectares <- gsub(",", ".", data_TFM_imp$ex_hectares)

# Convertimos la columna a tipo numérico.
data_TFM_imp$current_size <- as.numeric(data_TFM_imp$current_size)
data_TFM_imp$fire_spread_rate <- as.numeric(data_TFM_imp$fire_spread_rate)
data_TFM_imp$temperature <- as.numeric(data_TFM_imp$temperature)
data_TFM_imp$ex_hectares <- as.numeric(data_TFM_imp$ex_hectares)

# Nos quedamos con las variables necesarias para el estudio.
df_superv <- data_TFM_imp[,c(4,8,15,16,17,18)]

# Dividimos el dataset entre la variable clase y el resto de atributos
y <- df_superv[,1]
X <- df_superv[,2:6]

# Dividimos las filas de entrenamiento y test.
indexes = sample(1:nrow(df_superv), size=floor((2/3)*nrow(df_superv)))
```

```

trainX <- X[indexes,]
trainy <- y[indexes]
testX <- X[-indexes,]
testy <- y[-indexes]

# Aplicamos el modelo C5.0.
trainy = as.factor(trainy)
model <- C50::C5.0(trainX, trainy,rules=TRUE )
summary(model)

##
## Call:
## C5.0.default(x = trainX, y = trainy, rules = TRUE)
##
##
## C5.0 [Release 2.07 GPL Edition]      Tue Jan 09 19:07:37 2024
## -----
##
## Class specified by attribute `outcome'
##
## Read 13449 cases (6 attributes) from undefined.data
##
## Rules:
##
## Rule 1: (25/1, lift 1.5)
##   general_cause_desc = Lightning
##   weather_conditions_over_fire = CB Wet
##   temperature <= 19.5
##   relative_humidity <= 42
##   wind_speed <= 12
##   ->  class A  [0.926]
##
## Rule 2: (10, lift 1.5)
##   general_cause_desc = Lightning
##   weather_conditions_over_fire in {Cloudy, CB Wet, Rainshowers}
##   temperature <= 16.2
##   relative_humidity <= 71
##   wind_speed > 11
##   wind_speed <= 14
##   ->  class A  [0.917]
##
## Rule 3: (9, lift 1.4)
##   general_cause_desc = Resident
##   weather_conditions_over_fire = Cloudy
##   relative_humidity > 21
##   relative_humidity <= 23
##   wind_speed > 5
##   wind_speed <= 22
##   ->  class A  [0.909]
##
## Rule 4: (18/1, lift 1.4)
##   weather_conditions_over_fire = Rainshowers
##   temperature > 8.7
##   wind_speed > 22

```



```

## -> class A [0.900]
##
## Rule 5: (1369/174, lift 1.4)
## general_cause_desc = Recreation
## relative_humidity > 29
## -> class A [0.872]
##
## Rule 6: (585/91, lift 1.3)
## general_cause_desc = Recreation
## temperature > 19.1
## -> class A [0.843]
##
## Rule 7: (69/11, lift 1.3)
## general_cause_desc in {Oil & Gas Industry, Forest Industry, Resident,
##                         Incendiary, Power Line Industry, Undetermined,
##                         Agriculture Industry, Other Industry, Restart}
## weather_conditions_over_fire = CB Wet
## -> class A [0.831]
##
## Rule 8: (2135/417, lift 1.3)
## temperature <= 27.6
## wind_speed <= 1
## -> class A [0.804]
##
## Rule 9: (1374/269, lift 1.3)
## relative_humidity > 71
## -> class A [0.804]
##
## Rule 10: (74/16, lift 1.2)
## weather_conditions_over_fire = Cloudy
## temperature <= 27.5
## relative_humidity > 42
## wind_speed > 11
## wind_speed <= 13
## -> class A [0.776]
##
## Rule 11: (519/121, lift 1.2)
## weather_conditions_over_fire = Rainshowers
## relative_humidity > 29
## wind_speed <= 18
## -> class A [0.766]
##
## Rule 12: (1538/375, lift 1.2)
## general_cause_desc = Resident
## relative_humidity > 29
## wind_speed <= 12
## -> class A [0.756]
##
## Rule 13: (92/24, lift 1.2)
## weather_conditions_over_fire in {Cloudy, CB Wet, CB Dry, Rainshowers}
## relative_humidity > 42
## wind_speed > 26
## -> class A [0.734]
##

```

```

## Rule 14: (429/114, lift 1.2)
## general_cause_desc = Power Line Industry
## relative_humidity > 31
## wind_speed <= 22
## -> class A [0.733]
##
## Rule 15: (281/76, lift 1.2)
## general_cause_desc = Forest Industry
## -> class A [0.728]
##
## Rule 16: (494/135, lift 1.2)
## weather_conditions_over_fire = CB Wet
## wind_speed <= 5
## -> class A [0.726]
##
## Rule 17: (210/59, lift 1.1)
## general_cause_desc = Other Industry
## -> class A [0.717]
##
## Rule 18: (142/40, lift 1.1)
## general_cause_desc = Oil & Gas Industry
## relative_humidity > 34
## wind_speed <= 22
## -> class A [0.715]
##
## Rule 19: (464/140, lift 1.1)
## general_cause_desc = Power Line Industry
## temperature <= 22.5
## relative_humidity > 29
## wind_speed <= 39
## -> class A [0.697]
##
## Rule 20: (13220/4861, lift 1.0)
## wind_speed <= 32
## -> class A [0.632]
##
## Rule 21: (9, lift 3.1)
## general_cause_desc = Oil & Gas Industry
## temperature > 8.7
## relative_humidity > 22
## wind_speed > 27
## -> class B [0.909]
##
## Rule 22: (5, lift 3.0)
## weather_conditions_over_fire = Rainshowers
## temperature > 24.7
## temperature <= 26.5
## relative_humidity <= 71
## wind_speed > 1
## wind_speed <= 5
## -> class B [0.857]
##
## Rule 23: (5, lift 3.0)
## weather_conditions_over_fire = Cloudy

```

```

## temperature > 27.6
## wind_speed > 0
## wind_speed <= 1
## -> class B [0.857]
##
## Rule 24: (5, lift 3.0)
## general_cause_desc = Recreation
## weather_conditions_over_fire = Clear
## temperature > 9
## temperature <= 19.1
## relative_humidity > 16
## relative_humidity <= 24
## wind_speed > 10
## -> class B [0.857]
##
## Rule 25: (10/1, lift 2.9)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Cloudy
## temperature > 8.7
## relative_humidity > 34
## relative_humidity <= 36
## wind_speed > 1
## wind_speed <= 5
## -> class B [0.833]
##
## Rule 26: (4, lift 2.9)
## general_cause_desc = Resident
## weather_conditions_over_fire = Cloudy
## temperature <= 20
## relative_humidity > 27
## relative_humidity <= 29
## wind_speed > 12
## wind_speed <= 22
## -> class B [0.833]
##
## Rule 27: (4, lift 2.9)
## general_cause_desc = Oil & Gas Industry
## temperature > 19.8
## wind_speed > 3
## wind_speed <= 4
## -> class B [0.833]
##
## Rule 28: (4, lift 2.9)
## general_cause_desc = Forest Industry
## temperature <= 11.5
## relative_humidity > 39
## relative_humidity <= 40
## wind_speed > 5
## -> class B [0.833]
##
## Rule 29: (4, lift 2.9)
## general_cause_desc = Forest Industry
## temperature > 22.5
## relative_humidity > 31

```

```

## relative_humidity <= 40
## wind_speed > 5
## -> class B [0.833]
##
## Rule 30: (4, lift 2.9)
## general_cause_desc = Government
## temperature <= 14
## relative_humidity <= 40
## wind_speed > 5
## -> class B [0.833]
##
## Rule 31: (4, lift 2.9)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## temperature > 26.8
## relative_humidity > 27
## relative_humidity <= 29
## -> class B [0.833]
##
## Rule 32: (4, lift 2.9)
## weather_conditions_over_fire = CB Dry
## relative_humidity > 42
## relative_humidity <= 44
## wind_speed > 9
## wind_speed <= 14
## -> class B [0.833]
##
## Rule 33: (26/4, lift 2.8)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## temperature <= 27.5
## relative_humidity > 46
## relative_humidity <= 58
## wind_speed > 9
## wind_speed <= 14
## -> class B [0.821]
##
## Rule 34: (9/1, lift 2.8)
## general_cause_desc = Incendiary
## temperature <= 19.5
## relative_humidity <= 17
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.818]
##
## Rule 35: (9/1, lift 2.8)
## general_cause_desc = Resident
## weather_conditions_over_fire = Cloudy
## relative_humidity > 23
## relative_humidity <= 27
## wind_speed > 6
## wind_speed <= 9
## -> class B [0.818]
##

```

```

## Rule 36: (8/1, lift 2.8)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Cloudy
## temperature > 13.5
## temperature <= 23.5
## relative_humidity <= 36
## wind_speed > 1
## wind_speed <= 4
## -> class B [0.800]
##
## Rule 37: (13/2, lift 2.8)
## general_cause_desc = Undetermined
## weather_conditions_over_fire = Cloudy
## temperature <= 18.5
## relative_humidity <= 71
## wind_speed > 2
## wind_speed <= 5
## -> class B [0.800]
##
## Rule 38: (3, lift 2.8)
## general_cause_desc = Oil & Gas Industry
## relative_humidity > 30
## relative_humidity <= 31
## wind_speed > 5
## -> class B [0.800]
##
## Rule 39: (8/1, lift 2.8)
## general_cause_desc = Oil & Gas Industry
## weather_conditions_over_fire = Clear
## temperature > 24.7
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.800]
##
## Rule 40: (3, lift 2.8)
## general_cause_desc = Oil & Gas Industry
## weather_conditions_over_fire in {Cloudy, CB Dry}
## relative_humidity > 31
## relative_humidity <= 34
## wind_speed > 5
## -> class B [0.800]
##
## Rule 41: (3, lift 2.8)
## general_cause_desc = Undetermined
## weather_conditions_over_fire = Clear
## relative_humidity > 26
## wind_speed > 22
## wind_speed <= 27
## -> class B [0.800]
##
## Rule 42: (3, lift 2.8)
## general_cause_desc = Other Industry

```

```

## relative_humidity > 32
## relative_humidity <= 40
## wind_speed > 22
## -> class B [0.800]
##
## Rule 43: (23/4, lift 2.8)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## temperature > 26.9
## temperature <= 28.2
## relative_humidity > 28
## relative_humidity <= 50
## wind_speed > 1
## wind_speed <= 5
## -> class B [0.800]
##
## Rule 44: (8/1, lift 2.8)
## general_cause_desc in {Lightning, Railroad}
## weather_conditions_over_fire = Clear
## relative_humidity > 23
## relative_humidity <= 25
## wind_speed > 12
## wind_speed <= 18
## -> class B [0.800]
##
## Rule 45: (13/2, lift 2.8)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Wet
## temperature > 20.7
## temperature <= 22.2
## relative_humidity > 42
## relative_humidity <= 71
## wind_speed > 5
## wind_speed <= 14
## -> class B [0.800]
##
## Rule 46: (7/1, lift 2.7)
## general_cause_desc = Resident
## weather_conditions_over_fire = Cloudy
## relative_humidity > 28
## relative_humidity <= 29
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.778]
##
## Rule 47: (7/1, lift 2.7)
## weather_conditions_over_fire = Cloudy
## temperature <= 9
## relative_humidity > 31
## relative_humidity <= 40
## wind_speed > 19
## wind_speed <= 22
## -> class B [0.778]
##

```

```

## Rule 48: (11/2, lift 2.7)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## temperature <= 21
## relative_humidity > 33
## relative_humidity <= 37
## wind_speed > 5
## -> class B [0.769]
##
## Rule 49: (6/1, lift 2.6)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = CB Dry
## relative_humidity > 27
## relative_humidity <= 39
## wind_speed <= 5
## -> class B [0.750]
##
## Rule 50: (2, lift 2.6)
## general_cause_desc = Railroad
## wind_speed > 3
## wind_speed <= 4
## -> class B [0.750]
##
## Rule 51: (18/4, lift 2.6)
## general_cause_desc = Power Line Industry
## temperature <= 26.5
## relative_humidity > 8
## relative_humidity <= 19
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.750]
##
## Rule 52: (14/3, lift 2.6)
## general_cause_desc = Undetermined
## weather_conditions_over_fire = Cloudy
## temperature > 12.3
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.750]
##
## Rule 53: (14/3, lift 2.6)
## general_cause_desc = Agriculture Industry
## weather_conditions_over_fire = Clear
## temperature > 20.5
## relative_humidity > 33
## relative_humidity <= 40
## wind_speed > 5
## wind_speed <= 17
## -> class B [0.750]
##
## Rule 54: (2, lift 2.6)
## general_cause_desc = Other Industry

```

```

## relative_humidity <= 27
## wind_speed > 22
## -> class B [0.750]
##
## Rule 55: (22/5, lift 2.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## temperature > 22.6
## temperature <= 29.6
## relative_humidity > 30
## relative_humidity <= 50
## wind_speed > 1
## wind_speed <= 2
## -> class B [0.750]
##
## Rule 56: (24/6, lift 2.5)
## general_cause_desc = Lightning
## temperature > 24.7
## temperature <= 25.7
## relative_humidity > 48
## relative_humidity <= 57
## wind_speed > 5
## -> class B [0.731]
##
## Rule 57: (16/4, lift 2.5)
## general_cause_desc = Power Line Industry
## weather_conditions_over_fire = Cloudy
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 10
## -> class B [0.722]
##
## Rule 58: (12/3, lift 2.5)
## general_cause_desc = Resident
## weather_conditions_over_fire = Clear
## relative_humidity <= 15
## wind_speed > 8
## wind_speed <= 22
## -> class B [0.714]
##
## Rule 59: (22/6, lift 2.5)
## weather_conditions_over_fire = CB Wet
## temperature > 19.5
## temperature <= 21.5
## relative_humidity > 33
## relative_humidity <= 41
## wind_speed > 5
## -> class B [0.708]
##
## Rule 60: (14/4, lift 2.4)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Clear
## temperature <= 14.8
## relative_humidity > 21

```



```

## relative_humidity <= 31
## wind_speed > 9
## wind_speed <= 12
## -> class B [0.688]
##
## Rule 61: (17/5, lift 2.4)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## relative_humidity > 41
## relative_humidity <= 42
## -> class B [0.684]
##
## Rule 62: (25/8, lift 2.3)
## general_cause_desc = Railroad
## weather_conditions_over_fire in {Clear, CB Dry}
## wind_speed > 3
## wind_speed <= 5
## -> class B [0.667]
##
## Rule 63: (22/7, lift 2.3)
## general_cause_desc in {Forest Industry, Resident, Incendiary,
##                        Power Line Industry, Undetermined, Government}
## weather_conditions_over_fire = Cloudy
## temperature <= 20
## relative_humidity > 19
## relative_humidity <= 21
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.667]
##
## Rule 64: (13/4, lift 2.3)
## general_cause_desc = Power Line Industry
## relative_humidity > 19
## relative_humidity <= 22
## wind_speed > 12
## wind_speed <= 22
## -> class B [0.667]
##
## Rule 65: (16/5, lift 2.3)
## weather_conditions_over_fire = Clear
## relative_humidity <= 35
## wind_speed > 22
## wind_speed <= 24
## -> class B [0.667]
##
## Rule 66: (27/9, lift 2.3)
## general_cause_desc = Lightning
## relative_humidity > 42
## relative_humidity <= 43
## wind_speed > 5
## -> class B [0.655]
##
## Rule 67: (21/7, lift 2.3)
## general_cause_desc = Resident

```

```

## weather_conditions_over_fire = Cloudy
## temperature > 20
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 15
## -> class B [0.652]
##
## Rule 68: (21/7, lift 2.3)
## general_cause_desc = Power Line Industry
## weather_conditions_over_fire = Clear
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 9
## -> class B [0.652]
##
## Rule 69: (21/7, lift 2.3)
## general_cause_desc = Railroad
## weather_conditions_over_fire in {Cloudy, Clear}
## relative_humidity <= 33
## wind_speed > 5
## wind_speed <= 16
## -> class B [0.652]
##
## Rule 70: (18/6, lift 2.3)
## general_cause_desc = Agriculture Industry
## temperature > 24.5
## relative_humidity > 19
## relative_humidity <= 31
## -> class B [0.650]
##
## Rule 71: (9/3, lift 2.2)
## general_cause_desc = Incendiary
## temperature > 27.8
## relative_humidity <= 23
## -> class B [0.636]
##
## Rule 72: (14/5, lift 2.2)
## general_cause_desc = Agriculture Industry
## weather_conditions_over_fire = Clear
## relative_humidity <= 40
## wind_speed > 22
## -> class B [0.625]
##
## Rule 73: (14/5, lift 2.2)
## weather_conditions_over_fire = CB Wet
## relative_humidity <= 42
## wind_speed > 5
## wind_speed <= 9
## -> class B [0.625]
##
## Rule 74: (32/12, lift 2.1)
## general_cause_desc in {Lightning, Undetermined, Other Industry}
## temperature > 29.6

```

```

## temperature <= 30.5
## relative_humidity > 35
## wind_speed > 2
## -> class B [0.618]
##
## Rule 75: (55/21, lift 2.1)
## general_cause_desc in {Forest Industry, Resident, Incendiary,
##                        Power Line Industry, Undetermined,
##                        Agriculture Industry, Government, Other Industry}
## weather_conditions_over_fire = Cloudy
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 15
## wind_speed <= 22
## -> class B [0.614]
##
## Rule 76: (47/18, lift 2.1)
## general_cause_desc = Incendiary
## temperature <= 8.7
## relative_humidity > 39
## relative_humidity <= 71
## wind_speed > 1
## wind_speed <= 5
## -> class B [0.612]
##
## Rule 77: (16/6, lift 2.1)
## general_cause_desc = Recreation
## weather_conditions_over_fire = Clear
## temperature > 9
## temperature <= 14.8
## relative_humidity <= 29
## -> class B [0.611]
##
## Rule 78: (74/29, lift 2.1)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Clear
## temperature > 14.8
## temperature <= 21.5
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 15
## -> class B [0.605]
##
## Rule 79: (36/14, lift 2.1)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## temperature > 22.7
## relative_humidity > 33
## relative_humidity <= 37
## wind_speed > 5
## wind_speed <= 28
## -> class B [0.605]
##

```

```

## Rule 80: (20/8, lift 2.0)
## weather_conditions_over_fire = CB Dry
## temperature > 22.6
## relative_humidity > 46
## wind_speed > 3
## wind_speed <= 5
## -> class B [0.591]
##
## Rule 81: (38/16, lift 2.0)
## general_cause_desc = Incendiary
## relative_humidity > 31
## relative_humidity <= 33
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.575]
##
## Rule 82: (33/14, lift 2.0)
## general_cause_desc = Incendiary
## temperature > 17.5
## relative_humidity > 26
## wind_speed > 15
## wind_speed <= 22
## -> class B [0.571]
##
## Rule 83: (40/17, lift 2.0)
## general_cause_desc = Undetermined
## temperature <= 21.5
## relative_humidity > 19
## relative_humidity <= 31
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.571]
##
## Rule 84: (32/14, lift 1.9)
## general_cause_desc = Lightning
## weather_conditions_over_fire in {Cloudy, Clear}
## temperature <= 19.5
## relative_humidity <= 33
## wind_speed > 5
## wind_speed <= 18
## -> class B [0.559]
##
## Rule 85: (41/18, lift 1.9)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## temperature > 22.6
## temperature <= 28.2
## relative_humidity > 50
## relative_humidity <= 67
## wind_speed > 2
## wind_speed <= 5
## -> class B [0.558]
##
## Rule 86: (27/12, lift 1.9)

```

```

## general_cause_desc = Agriculture Industry
## relative_humidity > 33
## wind_speed > 13
## wind_speed <= 17
## -> class B [0.552]
##
## Rule 87: (35/16, lift 1.9)
## general_cause_desc = Incendiary
## weather_conditions_over_fire in {Cloudy, Clear}
## relative_humidity > 39
## relative_humidity <= 40
## wind_speed > 5
## wind_speed <= 22
## -> class B [0.541]
##
## Rule 88: (46/22, lift 1.8)
## relative_humidity > 35
## relative_humidity <= 40
## wind_speed > 22
## wind_speed <= 27
## -> class B [0.521]
##
## Rule 89: (54/26, lift 1.8)
## general_cause_desc = Lightning
## weather_conditions_over_fire in {Cloudy, CB Wet, CB Dry, Rainshowers}
## temperature > 21.9
## relative_humidity > 44
## relative_humidity <= 55
## wind_speed > 14
## wind_speed <= 26
## -> class B [0.518]
##
## Rule 90: (152/77, lift 1.7)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## temperature > 25.9
## temperature <= 29.6
## wind_speed > 2
## wind_speed <= 5
## -> class B [0.494]
##
## Rule 91: (45/23, lift 1.7)
## general_cause_desc = Resident
## weather_conditions_over_fire = Clear
## relative_humidity <= 39
## wind_speed > 19
## wind_speed <= 22
## -> class B [0.489]
##
## Rule 92: (154/79, lift 1.7)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## relative_humidity > 42
## relative_humidity <= 63

```

```

## wind_speed > 7
## wind_speed <= 14
## -> class B [0.487]
##
## Rule 93: (97/51, lift 1.6)
## general_cause_desc = Resident
## temperature > 8.7
## relative_humidity <= 40
## wind_speed > 22
## -> class B [0.475]
##
## Rule 94: (107/57, lift 1.6)
## relative_humidity <= 33
## wind_speed > 6
## wind_speed <= 7
## -> class B [0.468]
##
## Rule 95: (67/37, lift 1.6)
## weather_conditions_over_fire = Clear
## temperature > 24.7
## wind_speed > 5
## wind_speed <= 7
## -> class B [0.449]
##
## Rule 96: (71/41, lift 1.5)
## temperature > 26.5
## relative_humidity > 28
## wind_speed > 18
## -> class B [0.425]
##
## Rule 97: (9932/6792, lift 1.1)
## wind_speed > 3
## -> class B [0.316]
##
## Rule 98: (11208/7738, lift 1.1)
## wind_speed > 1
## -> class B [0.310]
##
## Rule 99: (4, lift 15.8)
## general_cause_desc = Lightning
## temperature > 29.6
## relative_humidity > 27
## relative_humidity <= 32
## wind_speed > 1
## wind_speed <= 4
## -> class C [0.833]
##
## Rule 100: (3, lift 15.2)
## general_cause_desc = Lightning
## relative_humidity <= 29
## wind_speed > 5
## wind_speed <= 6
## -> class C [0.800]
##

```

```

## Rule 101: (2, lift 14.2)
## general_cause_desc = Power Line Industry
## temperature > 26.5
## relative_humidity <= 19
## -> class C [0.750]
##
## Rule 102: (2, lift 14.2)
## general_cause_desc = Undetermined
## temperature > 28
## relative_humidity <= 19
## wind_speed > 5
## -> class C [0.750]
##
## Rule 103: (2, lift 14.2)
## general_cause_desc = Resident
## weather_conditions_over_fire = Cloudy
## temperature <= 20
## relative_humidity > 27
## relative_humidity <= 28
## wind_speed > 5
## wind_speed <= 12
## -> class C [0.750]
##
## Rule 104: (2, lift 14.2)
## general_cause_desc = Incendiary
## temperature > 27.8
## relative_humidity > 23
## relative_humidity <= 25
## wind_speed <= 5
## -> class C [0.750]
##
## Rule 105: (2, lift 14.2)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Cloudy
## temperature > 18.5
## relative_humidity <= 40
## wind_speed > 22
## wind_speed <= 29
## -> class C [0.750]
##
## Rule 106: (2, lift 14.2)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## temperature > 30.5
## relative_humidity > 31
## relative_humidity <= 33
## wind_speed > 5
## -> class C [0.750]
##
## Rule 107: (2, lift 14.2)
## general_cause_desc = Lightning
## temperature > 27.6
## relative_humidity <= 17
## wind_speed <= 1

```

```

## -> class C [0.750]
##
## Rule 108: (2, lift 14.2)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## temperature <= 22.7
## relative_humidity > 58
## relative_humidity <= 61
## wind_speed > 9
## -> class C [0.750]
##
## Rule 109: (4/1, lift 12.6)
## general_cause_desc = Agriculture Industry
## weather_conditions_over_fire = Clear
## temperature <= 24.5
## relative_humidity <= 31
## wind_speed > 14
## wind_speed <= 16
## -> class C [0.667]
##
## Rule 110: (4/1, lift 12.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## temperature > 22.6
## relative_humidity <= 30
## wind_speed > 1
## wind_speed <= 2
## -> class C [0.667]
##
## Rule 111: (4/1, lift 12.6)
## general_cause_desc = Agriculture Industry
## weather_conditions_over_fire = Cloudy
## temperature > 8.7
## relative_humidity > 28
## relative_humidity <= 37
## wind_speed > 22
## -> class C [0.667]
##
## Rule 112: (4/1, lift 12.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Cloudy
## relative_humidity <= 22
## wind_speed > 8
## wind_speed <= 18
## -> class C [0.667]
##
## Rule 113: (7/2, lift 12.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## temperature > 27
## relative_humidity > 25
## relative_humidity <= 27
## wind_speed > 5
## wind_speed <= 18

```



```

## -> class C [0.667]
##
## Rule 114: (4/1, lift 12.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## temperature > 30.5
## relative_humidity > 35
## wind_speed > 1
## wind_speed <= 5
## -> class C [0.667]
##
## Rule 115: (4/1, lift 12.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = CB Dry
## relative_humidity > 39
## relative_humidity <= 42
## wind_speed > 14
## wind_speed <= 17
## -> class C [0.667]
##
## Rule 116: (88/80, lift 1.9)
## general_cause_desc = Railroad
## -> class C [0.100]
##
## Rule 117: (434/401, lift 1.5)
## relative_humidity <= 19
## -> class C [0.078]
##
## Rule 118: (2, lift 55.7)
## weather_conditions_over_fire = CB Wet
## relative_humidity <= 33
## wind_speed > 17
## wind_speed <= 18
## -> class D [0.750]
##
## Rule 119: (2/1, lift 37.2)
## general_cause_desc = Under Investigation
## -> class D [0.500]
##
## Rule 120: (7, lift 56.1)
## general_cause_desc = Lightning
## temperature > 22.5
## relative_humidity <= 33
## wind_speed > 32
## -> class E [0.889]
##
## Rule 121: (3, lift 50.5)
## general_cause_desc = Incendiary
## weather_conditions_over_fire = Clear
## temperature <= 22.5
## relative_humidity <= 40
## wind_speed > 30
## -> class E [0.800]
##

```

```

## Rule 122: (3, lift 50.5)
## general_cause_desc = Lightning
## temperature > 30.5
## relative_humidity <= 28
## wind_speed > 18
## -> class E [0.800]
##
## Rule 123: (2, lift 47.4)
## general_cause_desc = Oil & Gas Industry
## weather_conditions_over_fire = Clear
## relative_humidity <= 16
## wind_speed > 5
## wind_speed <= 22
## -> class E [0.750]
##
## Rule 124: (2, lift 47.4)
## general_cause_desc = Power Line Industry
## weather_conditions_over_fire = Clear
## relative_humidity > 25
## relative_humidity <= 29
## wind_speed > 39
## -> class E [0.750]
##
## Rule 125: (10/2, lift 47.4)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## temperature > 20.5
## relative_humidity <= 28
## wind_speed > 23
## -> class E [0.750]
##
## Rule 126: (4/2, lift 31.6)
## general_cause_desc = Lightning
## temperature <= 21
## relative_humidity <= 28
## wind_speed > 18
## wind_speed <= 23
## -> class E [0.500]
##
## Rule 127: (4/2, lift 31.6)
## general_cause_desc = Lightning
## weather_conditions_over_fire = Clear
## relative_humidity > 31
## relative_humidity <= 33
## wind_speed > 18
## -> class E [0.500]
##
## Rule 128: (111/83, lift 16.2)
## general_cause_desc = Lightning
## relative_humidity <= 33
## wind_speed > 18
## -> class E [0.257]
##
## Default class: A

```

```
##
##
## Evaluation on training data (13449 cases):
##
##      Rules
##      -----
##      No      Errors
##
##      128 4369(32.5%)    <<
##
##
##      (a)  (b)  (c)  (d)  (e)  <-classified as
##      ----  ----  ----  ----  ----
##      7953  507   1    1    1  (a): class A
##      2821 1056   6    1    1  (b): class B
##      545   118  44    2    2  (c): class C
##      147   32   2    2    2  (d): class D
##      159   29   25   25   25  (e): class E
##
##
## Attribute usage:
##
## 100.00% wind_speed
##  50.61% relative_humidity
##  44.75% general_cause_desc
##  32.17% temperature
##  18.44% weather_conditions_over_fire
##
##
## Time: 0.9 secs
# A continuación comprobaremos la calidad del modelo usando el conjunto
# de datos de test.
predicted_model <- predict( model, testX, type="class" )
print(sprintf("La precisión del modelo es: %.4f %%",
              100*sum(predicted_model == testy) / length(predicted_model)))

## [1] "La precisión del modelo es: 61.8290 %"
# Hacemos la matriz de confusión para analizar la calidad de la predicción.
mat_conf <- table(testy,Predicted=predicted_model)
mat_conf

##      Predicted
## testy      A      B      C      D      E
##      A 3817  370  16    1    5
##      B 1571  335  17    1   11
##      C  268   75   4    0   2
##      D   90   28   2    0   1
##      E   81   24   4    0   2

# Vamos a eliminar los registros de los incendios menores y comprobar
# si de esta forma mejora la precisión del modelo.
data_TFM_imp_NAB <- subset(data_TFM_imp, size_class != 'A')
```

```

# Nos quedamos con las variables necesarias para el estudio.
data_TFM_imp_NAB <- data_TFM_imp_NAB[,c(4,8,13,14,15,16,17,18)]

# Dividimos el dataset entre la variable clase y el resto de atributos
y <- data_TFM_imp_NAB[,1]
X <- data_TFM_imp_NAB[,2:8]

# Dividimos las filas de entrenamiento y test.
indexes = sample(1:nrow(data_TFM_imp_NAB),
                 size=floor((2/3)*nrow(data_TFM_imp_NAB)))
trainX <- X[indexes,]
trainy <- y[indexes]
testX <- X[-indexes,]
testy <- y[-indexes]

# Aplicamos el modelo C50.
trainy = as.factor(trainy)
model_2 <- C50::C5.0(trainX, trainy, rules=TRUE )
summary(model_2)

##
## Call:
## C5.0.default(x = trainX, y = trainy, rules = TRUE)
##
##
## C5.0 [Release 2.07 GPL Edition]      Tue Jan 09 19:07:39 2024
## -----
##
## Class specified by attribute `outcome'
##
## Read 5002 cases (8 attributes) from undefined.data
##
## Rules:
##
## Rule 1: (963/153, lift 1.1)
##  general_cause_desc in {Resident, Restart, Other Industry, Railroad,
##                        Under Investigation, Prescribed Fire}
##  ->  class B  [0.840]
##
## Rule 2: (2762/491, lift 1.1)
##  relative_humidity > 29
##  wind_speed <= 12
##  ->  class B  [0.822]
##
## Rule 3: (4516/834, lift 1.1)
##  fire_type in {Surface, Ground}
##  ->  class B  [0.815]
##
## Rule 4: (721/136, lift 1.0)
##  weather_conditions_over_fire in {Rainshowers, CB Dry, CB Wet}
##  relative_humidity > 34
##  ->  class B  [0.811]
##
## Rule 5: (4/1, lift 4.7)

```

```

## fire_type = Crown
## weather_conditions_over_fire = Cloudy
## temperature <= 21
## wind_speed > 12
## wind_speed <= 15
## -> class C [0.667]
##
## Rule 6: (13/4, lift 4.7)
## general_cause_desc = Lightning
## fire_type = Crown
## fire_position_on_slope = Flat
## weather_conditions_over_fire = Clear
## temperature > 27.5
## wind_speed > 12
## wind_speed <= 24
## -> class C [0.667]
##
## Rule 7: (486/347, lift 2.0)
## fire_type = Crown
## -> class C [0.287]
##
## Rule 8: (3, lift 19.2)
## general_cause_desc = Recreation
## fire_type = Crown
## relative_humidity <= 24
## -> class D [0.800]
##
## Rule 9: (1, lift 16.0)
## general_cause_desc = Agriculture Industry
## fire_type = Crown
## -> class D [0.667]
##
## Rule 10: (6/2, lift 15.0)
## general_cause_desc = Incendiary
## fire_type = Crown
## wind_speed <= 9
## -> class D [0.625]
##
## Rule 11: (6/2, lift 15.0)
## fire_type = Crown
## weather_conditions_over_fire in {Rainshowers, CB Dry, CB Wet}
## relative_humidity <= 28
## wind_speed > 12
## -> class D [0.625]
##
## Rule 12: (12/7, lift 10.3)
## general_cause_desc = Lightning
## fire_type = Crown
## fire_position_on_slope = Flat
## weather_conditions_over_fire = Clear
## temperature <= 27.5
## relative_humidity > 25
## wind_speed > 12
## wind_speed <= 15

```

```

## -> class D [0.429]
##
## Rule 13: (5/1, lift 16.9)
## general_cause_desc = Incendiary
## fire_type = Crown
## wind_speed > 14
## -> class E [0.714]
##
## Rule 14: (4/1, lift 15.7)
## general_cause_desc = Undetermined
## fire_type = Crown
## temperature > 21
## -> class E [0.667]
##
## Rule 15: (6/2, lift 14.7)
## general_cause_desc = Power Line Industry
## fire_type = Crown
## relative_humidity <= 30
## -> class E [0.625]
##
## Rule 16: (8/3, lift 14.2)
## general_cause_desc in {Oil & Gas Industry, Government, Forest Industry}
## fire_type = Crown
## -> class E [0.600]
##
## Rule 17: (89/54, lift 9.3)
## general_cause_desc = Lightning
## fire_type = Crown
## weather_conditions_over_fire in {Clear, Cloudy}
## wind_speed > 12
## -> class E [0.396]
##
## Default class: B
##
##
## Evaluation on training data (5002 cases):
##
##           Rules
##   -----
##   No      Errors
##
##   17 1055(21.1%)  <<
##
##
##   (a)  (b)  (c)  (d)  <-classified as
##   ----  ---  ---  ---  ----
##   3848   13    3   14   (a): class B
##   643    40    4   17   (b): class C
##   175    12   16    5   (c): class D
##   147    18    4   43   (d): class E
##
##
## Attribute usage:
##

```

```
## 100.00% fire_type
## 59.34% relative_humidity
## 57.32% wind_speed
## 21.69% general_cause_desc
## 16.31% weather_conditions_over_fire
## 0.66% temperature
## 0.50% fire_position_on_slope
##
##
## Time: 0.0 secs

# A continuación comprobaremos la calidad del modelo usando el conjunto
# de datos de test.
predicted_model_2 <- predict( model_2, testX, type="class" )
print(sprintf("La precisión del segundo modelo es: %.4f %%",
              100*sum(predicted_model_2 == testy) / length(predicted_model_2)))

## [1] "La precisión del segundo modelo es: 78.1687 %"

# Hacemos la matriz de confusión para analizar la calidad de la predicción.
mat_conf_2 <- table(testy, Predicted=predicted_model_2)
mat_conf_2

##      Predicted
## testy      B      C      D      E
## B 1919      9      3     10
## C 319      14      7     14
## D 77       7      2      8
## E 82       9      1     20
```