

MASTER EN
**BUSINESS INTELLIGENCE
Y BIG DATA**

2019 / 2020



**PROYECTO
LADA
MODELOS
DE
COMBUSTIBLE
DINAMICOS**



Vicente **Gasull Orti**
Ignacio **Gómez Agueda**
Fernando **Arce Jiménez**
Fernando **Tejedor de Andrés**

TUTOR DEL PROYECTO:
D. Iván **Pastor**

1 Prefacio

Los incendios forestales son un fenómeno recurrente de los paisajes forestales a nivel global¹. En las regiones mediterráneas, los incendios forestales son un elemento natural de los ecosistemas que se ve favorecido por la estacionalidad del clima mediterráneo.^{2 3}

A pesar de ello, los recientes cambios en el paisaje tras el abandono rural producido en Europa en la última mitad de siglo, la falta de gestión forestal en las repoblaciones ejecutadas para evitar las pérdidas de suelo, y una política de extinción de incendios imperante frente a la prevención y gestión de los montes, han derivado en un incremento y una mayor continuidad de la superficie forestal que ha provocado la alteración de los regímenes de incendios en muchos lugares del mundo.^{4 5}

A este aumento de la superficie forestal y de su continuidad se le añade el incremento en la evapotranspiración y la intensidad de las sequías estivales debido al cambio climático, que se traduce en un aumento de la inflamabilidad de nuestros montes.⁶

Todo ello supone un impacto creciente con consecuencias económicas y ecológicas a veces irreversibles al cual, si añadimos el urbanismo caótico y salvaje de los últimos años donde la frontera entre lo urbano y lo forestal se difumina, hace que los incendios del siglo XXI dejen de ser un problema estrictamente ambiental para pasar a ser una emergencia civil.

Por tanto, los incendios forestales constituyen una de las mayores perturbaciones ecológicas a nivel mundial y, en especial, en las regiones mediterráneas. La extensión del periodo de alto riesgo, los ciclos de recurrencia cada vez más cortos de grandes incendios (mayores de 500 ha) asociados al calentamiento global, el aumento de la carga de combustible en los montes tras el éxodo rural y el éxito de las tareas de extinción en las últimas décadas suponen una amenaza creciente tanto para la población que habita el territorio de interfase urbano-forestal como para la preservación de los ecosistemas naturales.

Cuando se origina un incendio en zonas con cargas de combustible muy altas y en días con condiciones meteorológicas adversas, el comportamiento del fuego puede ser tan extremo que no exista ninguna posibilidad para que los medios de extinción puedan controlarlo. Estamos hablando de Grandes Incendios Forestales (GIF); incendios que se propagan a una velocidad e intensidad que

¹ "Large-scale, dynamic transformations in fuel moisture drive", <https://agupubs.onlinelibrary.wiley.com/>.

² "El peligro de incendios forestales derivado de la sequía ..." <https://dialnet.unirioja.es/>.

³ "Bases ecológicas para convivir con los incendios forestales." <https://www.uv.es/>.

⁴ "Análisis de la disponibilidad de combustible: índices ..." <https://www.researchgate.net/>.

⁵ "Fire regime changes in the Western Mediterranean" <https://www.researchgate.net/>.

⁶ "Changes in Fire and Climate in the Eastern Iberian Peninsula ..", <https://www.researchgate.net/>.

superan la capacidad de extinción de los medios ⁷. Para reducir los efectos negativos causados por este fenómeno es necesario contar con un sistema de prevención de incendios capaz de predecir con suficiente exactitud y precisión dónde y cuándo se puede desarrollar un incendio de estas condiciones. ⁸

Una de las políticas comunes del conjunto de países en los que el fuego juega un papel fundamental (Europa mediterránea, Estados Unidos de América, Australia, etc.) es la distribución de los recursos de extinción según la predicción del riesgo de incendio. El análisis de este factor mediante índices representativos permite optimizar la gestión de dichos recursos y anticiparse estratégicamente a la aparición del problema, concentrándose en una zona en el momento en que sea posible la ocurrencia de un GIF. ⁹

La definición de riesgo de incendio más comúnmente adoptada por las organizaciones internacionales establece lo siguiente: “*el riesgo de incendio es la posibilidad de que se inicie un incendio, según se vea afectado por la naturaleza y la incidencia de los agentes causantes*”, sin entrar en valores o en la estimación de posibles daños. ¹⁰ La probabilidad de ignición y de propagación de un fuego está inversamente relacionada con el contenido de humedad de los combustibles ¹¹. Esto es porque parte de la energía necesaria para iniciar un incendio se utiliza en el proceso de evaporación del agua retenida por la vegetación, justo antes de que el fuego comience a arder. ¹²

El comportamiento del fuego en un incendio es el resultado de la acción conjunta de un número de factores muy elevado que se encuadran en tres categorías: atmosféricos, topográficos y relativos a los combustibles. La humedad de los combustibles es una de las propiedades de estos que más influye en dicho comportamiento ¹³ y además es la que va a permitir que el combustible inicialmente se inflame o no. Puesto que el inicio, mantenimiento y posterior comportamiento del fuego se ve regulado por el contenido en humedad del combustible está plenamente justificado el interés por su determinación.

Por otro lado, desde hace más de cincuenta años, han ido surgiendo metodologías de predicción del riesgo meteorológico de incendio forestal, así como de predicción del comportamiento y propagación del fuego.

⁷ "Tipificación de los incendios forestales de Cataluña ..." <https://www.recercat.cat/>.

⁸ "Nuevas tecnologías para la estimación ... - Editorial CSIC." <https://editorial.csic.es/>.

⁹ "Prevention of Large Wildfires using the Fire Types" <https://www.researchgate.net/>.

¹⁰ "Wildland fire hazard and risk: Problems, definitions ... - naldc." <https://naldc.nal.usda.gov/>.

¹¹ "Fundamentos y Utilización de Índices Meteorológicos de ..." <https://ebuah.uah.es/>.

¹² "Fire on Earth: An Introduction | Wiley." <https://www.wiley.com/>.

¹³ "Forest fire: control and use (Book, 1973) [WorldCat.org]." <http://www.worldcat.org/>.

Lógicamente, estos sistemas de predicción del peligro de incendio y del comportamiento del fuego utilizan como variable de entrada fundamental para sus cálculos la humedad del combustible. En consecuencia, su fiabilidad depende, entre otros factores, de que el valor asignado a dicha humedad sea lo más real posible. Esto justifica doblemente la necesidad de contar con un buen método de determinación.

Existen técnicas muy variadas para medir la humedad del combustible que incluyen métodos directos e indirectos. Los directos son métodos más exactos, pero suelen precisar muestreos destructivos de la vegetación, aparatos de laboratorio, personal especializado y cierto tiempo hasta que se obtengan los resultados. Es por tanto necesaria una técnica que permita determinar con mayor rapidez y sencillez la humedad del combustible en un momento y lugar concreto con el fin de poder anticipar el comportamiento del fuego durante el desarrollo de una quema controlada o de un incendio en ese lugar, o bien para determinar si las condiciones meteorológicas favorecen el inicio y desarrollo del fuego. Los métodos indirectos, entre los que se incluye la predicción, lo permiten. Esto justifica haber seleccionado la predicción como método de determinación de la humedad en el presente trabajo.

Si diferenciamos entre combustibles vivos y combustibles muertos, los cambios en la humedad de los primeros ocurren más lentamente que en los combustibles muertos debido a que dependen no sólo de las condiciones atmosféricas sino también del contenido de humedad del suelo y la fisiología de las especies.¹⁴

Los combustibles finos arden con mayor rapidez que los más gruesos. Conocer la variación de la humedad de las partes más finas de las plantas vivas, es decir, de las hojas y las ramillas, es de máximo interés ya que dichos elementos finos son los que juegan el papel más activo en el inicio y la propagación del fuego.¹⁵ ¹⁶ Por ello, para Agresta es crucial predecir con suficiente precisión el contenido de humedad del combustible fino vivo con el objetivo de poder determinar índices de peligrosidad por incendio forestal asociados a la inflamabilidad potencial de la vegetación.

En cada lugar, la humedad de las hojas de las diferentes especies sigue un patrón condicionado por la variación estacional a lo largo del año. Este patrón "típico" se ve distorsionado cada año según las peculiaridades que el tiempo atmosférico presenta respecto al clima normal en la región (variaciones en la cantidad y distribución de las lluvias u ocurrencia de temperaturas inusualmente cálidas o frías) (Pyne et al., 1996). Por ejemplo, si en los meses previos al inicio de la brotación la precipitación ha sido escasa, el pico de humedad primaveral será menor pues el suelo tendrá pocas

¹⁴ "Large-scale, dynamic transformations in fuel moisture drive" 6 may... 2016, <https://ro.uow.edu.au/>.

¹⁵ "Fire on Earth: An Introduction | Wiley." <https://www.wiley.com/>.

¹⁶ "el peligro de incendios forestales derivado de la sequía - Dialnet." <https://dialnet.unirioja.es>.

reservas de agua. Si el suelo se mantiene seco durante el verano, la humedad del combustible decaerá más rápidamente y alcanzará valores mínimos inferiores con respecto al nivel normal de humedad del suelo.

La disminución de la humedad de las hojas y acículas no suele ser continua, sino que presenta una serie de altibajos a lo largo del tiempo. Incluso, se ha apreciado que junto con las importantes variaciones de humedad que los combustibles vivos experimentan a lo largo del año se dan variaciones de menor magnitud a lo largo del día. En los días soleados la humedad de las hojas disminuye durante la tarde y luego se recupera al anochecer. De acuerdo con Chandler et al. (1983), los cambios de temperatura a lo largo del día influyen más en los cambios de humedad de las hojas que la propia variación de la humedad atmosférica o del suelo.

El balance entre el agua cedida a la atmósfera por transpiración y la absorción radicular determina la humedad de los elementos finos en la vegetación. La existencia o no de déficit interno de humedad en un vegetal depende de las velocidades relativas de toma de agua de la planta a través de sus raíces y de pérdida de ésta a través de la transpiración (Chandler et al., 1983).

El proceso de transpiración está controlado por factores atmosféricos como la radiación solar, la temperatura, la humedad del aire y el viento y por factores propios del vegetal como la estructura foliar y el grado de apertura estomática. Durante la estación de crecimiento las hojas pueden mantener un nivel adecuado de humedad gracias a su capacidad para abrir o cerrar las estomas y así regular la velocidad de transpiración.

La absorción radicular está controlada por factores del suelo como la aireación, temperatura, tensión del agua y la concentración de solutos, así como por factores propios del vegetal como el tamaño y distribución del sistema radical.

Los niveles de temperatura y humedad que van adquiriendo tanto el suelo como el aire determinan el momento en que las plantas abandonan el período de reposo y por tanto el momento en que se incrementa su humedad. También repercuten en los niveles medios de humedad alcanzados por las hojas ya que influyen en la cantidad de biomasa foliar que se genera cada temporada.

La humedad alcanzada por los vegetales vivos también se ve influenciada por la calidad de estación en la que se desarrollan. Las localizaciones buenas, asociadas con suelos profundos con elevada capacidad de retención de agua, producen mayor crecimiento y por tanto humedades mayores que además van descendiendo más lentamente.

Por último, los vientos fuertes y secos pueden reducir, en cuestión de horas, la humedad de los combustibles vivos sobre todo si hay déficit de agua en el suelo y no pueden reponer las pérdidas por evapotranspiración.

2 Índice

1	Prefacio	1
2	Índice	5
3	Referencias de Ilustraciones	8
4	Referencias de tablas	9
5	Anexos adicionales	10
6	Introducción	11
7	Definición del problema	12
8	Análisis de las causas potenciales del problema	12
8.1	Tecnología	12
8.2	Recursos	13
8.3	Conocimiento	13
8.4	Comportamiento medioambiental variable	14
9	Diagrama Fishbone	14
10	Conclusión del problema y causas	15
11	Validación de hipótesis	15
12	Análisis preliminar de los datos	16
12.1	Fuente de datos del INIA	17
12.2	Fuente de datos abiertos “Open Source”	17
12.2.1	AEMET	18
12.2.2	SIAR	18
12.2.3	Copernicus	19
13	Análisis del entorno	20
14	Análisis competitivo	24
15	Análisis y diagnóstico	25
16	Modelo de negocio	27
16.1	Definición del modelo de negocio	28
16.1.1	Propuesta de valor	28

16.1.2	Socios clave	29
16.1.3	Actividades clave	30
16.1.4	Recursos clave	31
16.1.5	Relación y captación de clientes	31
16.1.6	Canal de distribución	32
16.1.7	Segmento de clientes	32
16.1.8	Estructura de costes	33
16.1.9	Flujo de ingresos	34
17	Plan de acción	35
17.1	Definición del alcance del proyecto	35
17.1.1	Objetivos	35
17.1.2	Métricas	36
17.2	Análisis de actividades. Modelo lógico y arquitectura técnica	38
17.2.1	Modelo lógico	38
17.2.1.1	Identificación de los datos	39
17.2.1.2	Adquisición de datos	39
17.2.1.3	Procesamiento e integración	41
17.2.1.4	Análisis descriptivo	42
17.2.1.5	Almacenamiento de los datos	44
17.2.1.6	Descubrimiento y modelado	44
17.2.1.7	Visualización de los resultados	48
17.2.2	Solución tecnológica: arquitectura técnica	48
17.3	Análisis de recursos: talento humano y recursos físicos	49
17.3.1	Talento humano	49
17.3.2	Recursos físicos y servicios externos	50
17.4	Gestión del tiempo	52
18	Rentabilidad del proyecto	54
18.1	Beneficios tangibles	55

18.1.1	Generación de ingresos	55
18.1.2	Reducción de costes	55
18.2	Beneficios intangibles	56
18.3	Beneficios estratégicos	56
18.4	Liquidez, rentabilidad y riesgos	57
19	Conclusiones	58
20	Referencias	60
20.1	Bibliografías	60
20.2	Webgrafías	61
21	Apéndices	63
21.1	Entrevistas: validación de hipótesis y obtención de información adicional	63
21.2	SIAR: Evapotranspiración, proceso y parámetros influyentes	66

3 Referencias de Ilustraciones

Ilustración 1 Diagrama Causa-Efecto del (Ishikawa) problema	14
Ilustración 2 Ubicación de las estaciones SIAR en la Comunidad de Madrid	19
Ilustración 3 Evolución de las distancias causas de deforestación en el mundo.....	21
Ilustración 4 Hectáreas quemadas en incendios forestales en España.....	22
Ilustración 5 Crecimiento de la población mundial.....	23
Ilustración 6 Diagrama DAFO	27
Ilustración 7 Diagrama CANVAS.....	28
Ilustración 8 Resumen de la propuesta de valor	29
Ilustración 9 Modelo lógico.....	39
Ilustración 10 Almacenamiento de la información en Copernicus.....	40
Ilustración 11 Visualización en la herramienta Panoply de los datos descargados por SIAR.....	41
Ilustración 12 Estacionalidad de las temperaturas	43
Ilustración 13 Análisis de correlación de las variables calculadas AEMET	43
Ilustración 14 Análisis de correlación de las variables obtenidas de SIAR.....	44
Ilustración 15 Arquitectura del proyecto	48
Ilustración 16 Diagrama de Gantt del proyecto	54
Ilustración 17 Rentabilidad del proyecto	54

4 Referencias de tablas

Tabla 1 Referencia de documentación adicional	10
Tabla 2 Listado de hipótesis planteadas	16
Tabla 3 Esquema de indicadores KPIs claves en el rendimiento del modelo	38
Tabla 4 Resultados ejecución del modelo Zona Robledo de Chavela	46
Tabla 5 Resultados ejecución del modelo en la especie Cistus Ladanifer	46
Tabla 6 Resultados ejecución del modelo en todas las zonas de estudio	47
Tabla 7 Resumen económico	58

5 Anexos adicionales

A continuación, se detallan todos los documentos utilizados para la construcción de esta memoria.

Nombre del documento	Tipo de documento	Descripción	
1. Descarga_Copernicus_API.pdf	Código fuente	Descarga de los archivos de información procedentes de la fuente de datos de COPERNICUS. Tratamiento de estos para su utilización en la ejecución de los modelos desarrollados.	
2. Combinar_Ficheros_Copernicus.pdf			
3.1 Copernicus_API.pdf			
3.2 Copernicus_AGRESTA.pdf			
4. Visualización _Copernicus.pdf		Descarga de los archivos de información de la fuente de datos SIAR a través de Web Scraping.	
5. SIAR.pdf			
6. AEMET.pdf			Descarga de los archivos de información de la fuente de datos AEMET a través de API en R.
7. Robledo_Chavela.pdf			Ejecución de los modelos tanto para la zona geográfica de Robledo de Chavela como para la especie <i>Cistus ladanifer</i> .
8. Robledo_Chavela_Clustering.pdf			
9. Cistus_Ladanifer.pdf			

Tabla 1 Referencia de documentación adicional

6 Introducción

Agresta Sociedad Corporativa es una consultora ambiental especializada en el sector forestal que principalmente se dedica a la gestión de los recursos forestales, la monitorización de procesos climáticos y ambientales y al desarrollo de proyectos de prevención y protección contra incendios.

Dicha empresa desarrolla herramientas innovadoras enfocadas al sector forestal disponiendo de una línea de proyectos orientada a la protección frente a incendios y su gestión en el territorio de forma sostenible para las generaciones futuras. Los incendios son una de las principales amenazas para la conservación de las masas forestales. Su incidencia aumenta debido a cambios en el uso del territorio y a la mayor frecuencia de condiciones meteorológicas adversas. Se exploran nuevos métodos y tecnologías para mejorar la caracterización de los combustibles vivos y secos, el comportamiento del fuego, el riesgo de incendio y la recuperación de las áreas quemadas. La empresa invierte en el desarrollo de herramientas innovadoras que ayuden en la planificación de actuaciones tanto de prevención y extinción de incendios como de restauración de las masas forestales afectadas por los mismos.

Dentro de las líneas de proyecto e investigación desarrolladas por Agresta destacan las siguientes relacionadas con incendios forestales¹⁷:

- Cartografía de modelos de combustible forestal a partir de datos LiDAR aéreos. Nueva metodología de bajo coste basada en datos LiDAR para generar mapas de modelos de combustible forestal en grandes superficies con alta resolución y capacidad de actualización.
- Selvicultura preventiva de incendios basada en simulación de incendios forestales. Optimización, basada en ciencia y conocimiento experto, de la localización de las actuaciones de selvicultura preventiva basada en la simulación del comportamiento del fuego a partir de cartografías de modelos de combustibles desarrolladas siguiendo la metodología anterior.¹⁸
- Proyecto de investigación básica en incendios forestales (GEPRIF/VIS4FIRE): Nuevas tecnologías para reducir la severidad de los incendios forestales.¹⁹ Proyectos de investigación fundamental (RTA-INIA) orientado al desarrollo de metodologías eficaces para reducir el impacto de los incendios forestales mediante una gestión integrada y la reducción de la vulnerabilidad a incendios forestales.

¹⁷ Proyectos sobre Incendios Forestales. <https://agresta.org/proyectos/proyectos-incendios-forestales/>

¹⁸ "Nueva metodología para la cartografía de combustibles ..." 24 oct... 2016, <http://chil.me/>.

¹⁹ "Proyecto GEPRIF." 17 jul. 2019, <http://proyectogeprif.es/>.

Es en esta última línea de investigación donde nuestro grupo de trabajo ha visto una oportunidad de realizar y aplicar un proyecto BI-BD.

Dentro del área de trabajo de prevención y protección contra incendios, uno de los obstáculos con que nos encontramos, es la precisión del modelo de predicción para el cálculo de la humedad relativa de combustible vivo para una determinada área objeto de estudio. La respuesta del modelo se utiliza para determinar el riesgo de incendio en dicha zona, siendo además esta variable crítica para muchos de los cálculos que llevan a cabo en otro tipo de proyectos.

7 Definición del problema

Rapidez y precisión en el cálculo estimado de la humedad del combustible vivo, variable objetivo, son los principales retos a los que se enfrenta la empresa. Actualmente el proceso de obtención de esta variable a partir de un modelo teórico en desarrollo es costoso en tiempo y recursos.

Esta problemática se debe a la suma de diversas causas potenciales que se describen en la siguiente sección de este documento. Además de contar con un número reducido de variables ambientales, la empresa no está especializada en la implementación de modelos de “Machine Learning”, ni en la utilización de herramientas o lenguajes de programación para la extracción, almacenamiento, transformación o explotación de los datos.

Para realizar esta tarea y poder validar los resultados, se necesitan recoger muestras de combustible fino en campo, que debe transportarse y analizarse en un laboratorio especializado, lo cual supone largos periodos de tiempo para su procesamiento e inversión en recursos.

8 Análisis de las causas potenciales del problema

Las causas potenciales del problema se describen a continuación en cuatro grupos principales. El equipo de trabajo ha validado estas causas e hipótesis mediante entrevistas con la empresa. Este análisis nos ayudará a centrar los esfuerzos del proyecto para la solución(es) propuestas más adelante.

8.1 Tecnología

Una de las principales causas es el número reducido de variables utilizadas para implementar el modelo y así tratar de explicar su variable objetivo.

La empresa realiza un estudio detallado utilizando únicamente información de los niveles de reflectancia de las distintas bandas de imágenes satelitales (MODIS y SENTINEL2) a través de un

sistema de información geográfica con la finalidad de poder identificar un número reducido de variables para estimar la humedad del combustible vivo objeto de estudio.

En ocasiones, cuando el área de estudio es muy reducida, encuentran cierta problemática en el tratamiento de las imágenes, ya que un único píxel puede cubrir el área de estudio por completo y por tanto la resolución espacial de las imágenes condiciona mucho su aplicabilidad. También es habitual que, al usar sensores ópticos, la presencia de nubes invalide los datos, impidiendo se puedan utilizar ciertas capturas tanto en el desarrollo del modelo como en su aplicación operativa por falta de datos.

Hoy en día, no se hace uso de las variables climatológicas disponibles en formato abierto que previsiblemente afectan a la humedad relativa de la vegetación. Los técnicos de desarrollo conocen la existencia de distintas plataformas open source (datos abiertos) que proporcionan estos datos, sin embargo, no han llegado a automatizar su extracción, tratamiento o carga y encuentran ciertas limitaciones en el uso de las herramientas y lenguajes para completar estas tareas.

8.2 Recursos

La empresa dispone de equipos propios con prestaciones suficientes para la extracción y almacenamiento de los datos para soportar los modelos, así como para la ejecución de tareas de transformación de estos.

La cualificación del personal es alta en relación con la ciencia forestal y con el uso de sistemas de información geográfica, complementándose con perfiles familiarizados con técnicas en “Machine Learning” implementadas en determinados lenguajes de programación. Sin embargo, para implementar una propuesta organizacional de uso y tratamiento eficiente de los datos, como es el caso que se propone con este proyecto, sería necesario ampliar su formación en ciertas capacidades técnicas.

Además, la empresa ha observado que, tras el muestreo en campo, algunas de las muestras vegetales presentan resultados poco fiables del contenido de humedad que pueden deberse a una falta de preparación técnica por parte de la empresa proveedora del servicio. Esta fase de campo implica llevar a cabo de una manera rigurosa el protocolo de toma de muestras.

8.3 Conocimiento

La empresa utiliza su modelo de predicción actual desde una perspectiva teórica respecto al cálculo de la humedad relativa existente en el combustible vivo que se recoge en campo y en diferentes puntos estratégicos. A esto se suma que no se usan otro tipo de técnicas de predicción

basados en valores históricos de sistemas registrados que podrían contribuir a alcanzar una mejor predicción.

Hay que señalar que, aunque conocen ciertas técnicas de “Machine Learning” y “Deep Learning” no han llegado a implementarlas de manera operativa o desconocen el alcance de estas.

8.4 Comportamiento medioambiental variable

Los factores medioambientales que influyen en la humedad del combustible vivo son complejos y no se puede asegurar a priori que unos factores influyan más que otros, ni en qué medida lo hacen. Se conoce que parámetros como las precipitaciones, temperatura, radiación solar, viento o capacidad de retención de agua en el suelo, son útiles para comprender mejor la humedad del combustible vivo, aunque es incierto tanto su impacto directo como el nivel de correlación con la variable objetivo y entre ellos. Por otro lado, la disponibilidad de uso de estos datos y sus fuentes son complejas.

9 Diagrama Fishbone

Como resumen de las causas potenciales del problema descritas en puntos anteriores, se presenta el diagrama fishbone, el cual nos permite de una manera rápida, visual y simplificada conocer las causas potenciales del problema planteado.

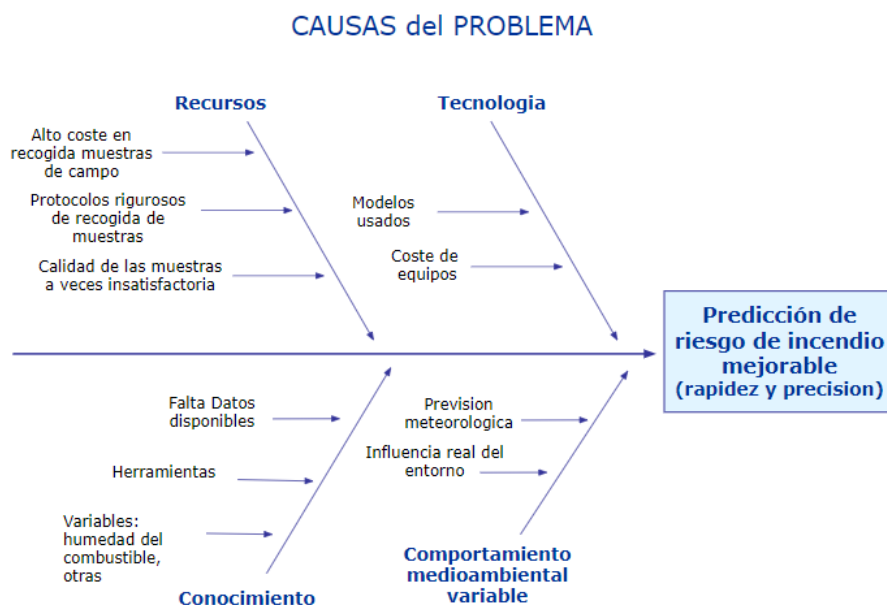


Ilustración 1 Diagrama Causa-Efecto del (Ishikawa) problema

10 Conclusión del problema y causas

El problema que definimos y en el que nos centramos en este proyecto será en una mejora de la predicción de riesgo de incendios originada por una mejor precisión del modelo desarrollado actualmente del cálculo de humedad de combustible vivo y la rapidez con que se obtiene esta predicción.

De las causas más influyentes en esta problemática destacamos la complejidad de procesado de las variables satelitales disponibles que usan en el modelo actual siendo la humedad del “combustible vivo” la variable principal.

Tras la evaluación y la validación de las causas potenciales mediante entrevistas con el grupo de trabajo de la empresa, se determina que existe una oportunidad de mejora del modelo si se incluyeran nuevas variables o información auxiliar a la información espectral satelital ya que esta presenta una limitación típica asociada a sensores ópticos. De esta forma se podrían mejorar tanto la automatización de procesos como la capacidad de predicción del modelo.

Como consecuencia de la mejora del modelo actual se pasaría de un sistema experimental a un desarrollo operativo, por tanto, la empresa podría añadir este producto a su línea de producción una vez reducidos los costes actuales directos e indirectos.

11 Validación de hipótesis

En esta sección se describen las diferentes hipótesis que este grupo de trabajo se ha planteado durante la fase de planificación del proyecto. Para ello, se han usado metodologías de “lean start-up”, “design thinking” y “brainstorming”.

Tras generar un listado de hipótesis posibles y darles una valoración de importancia según incertidumbre y criticidad, el resultado final se resume en [Apéndice 21.2](#) donde para su validación, se han realizado varias sesiones confirmatorias donde se mantienen entrevistas con la empresa y donde los puntos aclaratorios y validados se resumen por cada hipótesis planteada. Las hipótesis se listan a continuación en el orden indicado:

Hipótesis

Hipótesis 1: Aumentar la precisión y operatividad de la predicción del riesgo para la prevención de incendios mejorando el indicador de la humedad de la planta utilizado actualmente calculado a través de mediciones sobre el terreno. A través de fuentes de datos externas, permitirá predecir este indicador de humedad de la planta, mejorando en un 8% el modelo de predicción actual.

Hipótesis 2: Adaptar el modelo para incrementar la parte de negocio captando nuevas administraciones públicas durante el primer año después del lanzamiento.

Hipótesis 3: Reducir costes en un 25% al eliminar parcialmente la recogida de muestras en campo y mejorando la eficiencia en los recursos contra incidencias.

Hipótesis 4: Las CCAA y Administraciones públicas expresan una preocupación (si/no) sobre el riesgo de incendios e impacto medioambiental.

Hipótesis 5: El nuevo modelo permitirá conocer la evolución de la desertificación en su territorio con una mejora de un 5% de precisión.

Hipótesis 6: Aumentar el flujo de notificaciones / publicaciones de avisos de incendio en un 20%.

Tabla 2 Listado de hipótesis planteadas

12 Análisis preliminar de los datos

Contaremos con datos de fuentes abiertas y privadas que nos permitirán ejecutar nuestros modelos predictivos. De manera complementaria, a partir de los datos recopilados de la AEMET, Copernicus o SIAR, se podrán calcular nuevas variables, como puede ser las precipitaciones acumuladas o las temperaturas medias agregadas para un determinado intervalo de tiempo, así como otros datos ya calculados por dichos proveedores como es la humedad del suelo en su capa superficial.

Debido a la altísima variabilidad de datos de estaciones meteorológicas, coordenadas geoespaciales o variedades de combustible de bosque y matorral bajo, para el enfoque y propósito de este proyecto, se acuerdan las siguientes limitaciones.

- Área de estudio: Comunidad de Madrid
- Tipo de combustible de estudio: matorral bajo, en concreto la especie *Cistus ladanifer*, o jara pringosa
- Zonas de estudio: Inicialmente se centrará en el punto muestral 5 B de Robledo de Chavela (datos disponibles desde 2016), y se aplicarán las respectivas estaciones (datos abiertos) en función de cercanía o cálculos extrapolados con respecto a esta localización. El análisis se extenderá a otras zonas de estudio cuando sea apropiado con el propósito de análisis y observación del comportamiento de los modelos y si fuera necesario se describirán las conclusiones obtenidas de las mismas.

A continuación, se describe con mayor detalle las bases de datos mencionadas.

12.1 Fuente de datos del INIA

Valores de la humedad (%) del combustible vivo proporcionados por el INIA (periodo 2014 - actualidad). Son valores obtenidos de muestras tomadas en campo y calculados mediante un proceso de secado y pesado en laboratorio. Nos proporcionarán los valores reales de humedad de la planta para una serie de fechas y localizaciones dentro de la comunidad de Madrid. La variable objetivo recibe el nombre de FMC del inglés Fuel Moisture Content y cuyo cálculo es el siguiente:

$$FMC(\%) = \frac{W_f - W_s}{W_s} \times 100$$

donde W_f es el peso fresco y W_s es el peso seco de la planta

Dentro de los registros con los que contamos, se encuentran los datos de distintas especies de árboles y arbustos. Nuestro estudio se centra en la jara pringosa (*Cistus ladanifer*) ya que se considera un indicador de referencia para la predicción del riesgo de incendio asociado a la variabilidad de la humedad del combustible vivo.

Contaremos con datos desde el año 2014 para la especie de estudio, pero no fue hasta el año 2016 cuando estos análisis se empezaron a hacer fuera de las épocas de mayor riesgo y se recogen en la zona de Robledo de Chavela. Por tanto, para cubrir las inconsistencias y grandes periodos no estivales sin datos, será necesario realizar varios pasos en los procesos de extracción y transformación. Estos se describen en detalle en la sección de [Análisis de actividades. Modelo lógico y arquitectura técnica](#))

Estos datos se obtienen por la Comunidad de Madrid en colaboración con el laboratorio del INIA y son transferidos a Agresta de manera periódica. La periodicidad de las muestras depende de la época del año variando entre 3 días durante el verano y 15 días en las temporadas de invierno.

12.2 Fuente de datos abiertos “Open Source”

Variables meteorológicas con registros históricos en abierto a partir de los cuales se realizará el modelado para predecir los valores de humedad del combustible vivo. Se tendrá en cuenta toda la información meteorológica disponible. Estos datos abiertos se obtienen directamente de los servidores web de cada una de las plataformas mencionadas. Serán evaluados y tratados para determinar qué variables son más relevantes e impactan de manera significativa en el cálculo de la humedad del combustible vivo.

12.2.1 AEMET

Se trata de la plataforma de la Agencia Española de Meteorología con los registros de mayor relevancia para proporcionar robustez al modelo predictivo propuesto, ya que los datos obtenidos de la estación meteorológica situada en Robledo de Chavela son los más próximos a la zona de estudio donde se realizan las extracciones de muestras de campo.²⁰

En cuanto a las variables disponibles en la plataforma, son las siguientes:

- Altitud (m)
- Temperatura media (°C)
- Temperatura mínima (°C)
- Temperatura máxima (°C)
- Precipitación (mm)
- Presión atmosférica máxima (atm)
- Presión atmosférica mínima (atm)
- Velocidad media del viento (m/s)
- Dirección del viento (°)
- Racha

12.2.2 SIAR

La plataforma SIAR permite la consulta y descarga de datos meteorológicos diarios, semanales o mensuales de una o varias estaciones.²¹ En nuestro caso la estación más próxima a la zona de estudio es la de Villa del Prado que se sitúa aproximadamente a 20-30 km de Robledo de Chavela.

Además, esta fuente de datos proporciona una serie de variables que no están contenidas en el resto de las plataformas como los registros de humedad del aire, radiación solar o evapotranspiración, variables que a priori han sido consideradas relevantes y podrían resultar claves o necesarias para conseguir mayor precisión en el cálculo de la humedad de la vegetación.

²⁰ “Sistema para la difusión y reutilización de la información...” <https://opendata.aemet.es/centrodedescargas/inicio>

²¹ Consulta de Datos SIAR. <http://eportal.miteco.gob.es/websiar/SeleccionParametrosMap.aspx?dst=1>

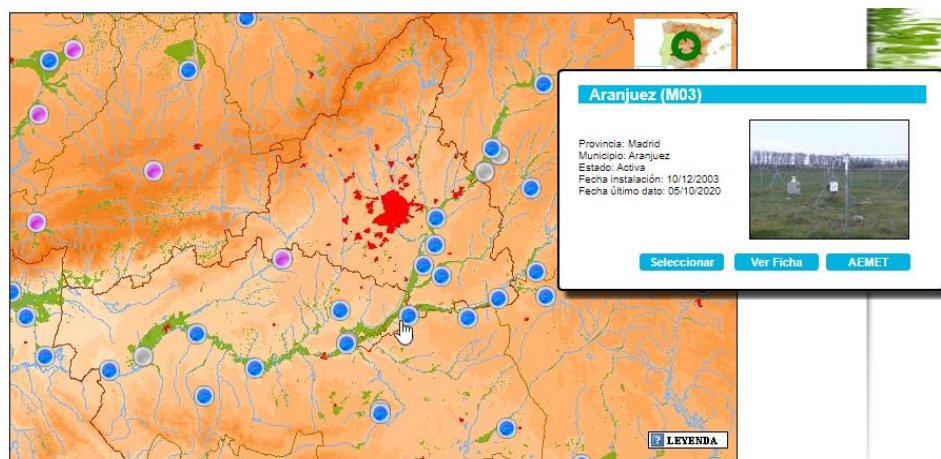


Ilustración 2 Ubicación de las estaciones SIAR en la Comunidad de Madrid

Aunque la ET se puede medir utilizando aparatos de medida como los lisímetros, el proceso de medición es largo y costoso, por lo que es mucho más frecuente su estimación en base a ecuaciones empíricas y analíticas. El detalle y características del proceso de evapotranspiración se puede consultar en el [Apéndice 22.2](#)

La ET se puede calcular con bastante precisión usando modelos matemáticos consistentes en ecuaciones complejas. Las ecuaciones empleadas para el cálculo de la evapotranspiración se pueden ver en la sección de datos calculados.

12.2.3 Copernicus

La plataforma Copernicus proporciona el conjunto de datos con las estimaciones de la humedad del suelo en todo el mundo a partir de un gran conjunto de sensores de satélite. ²²

- Humedad volumétrica del suelo (m3): Contenido de agua líquida en una capa superficial del suelo de 2 a 5 cm de profundidad, expresado en m3 de agua por m3 de suelo.
- Humedad del suelo superficial (%): Contenido de agua líquida en una capa superficial del suelo de 2 a 5 cm de profundidad expresado como porcentaje de la saturación total.

La humedad del suelo superficial (SSM) es el contenido relativo de agua de los primeros centímetros del suelo, que describe cuán húmedo o seco está el suelo en su capa superior, expresado en porcentaje de saturación. Se mide con sensores de radar satelital y permite conocer los impactos de las precipitaciones locales y las condiciones del suelo.

²² "Soil moisture gridded..." <https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-soil-moisture?tab=overview>

El SSM es un impulsor clave de los flujos de agua y calor entre el suelo y la atmósfera, regulando la temperatura y la humedad del aire. Además, en su papel de suministro de agua, es vital para la salud de la vegetación. Viceversa, el SSM es muy sensible al forzamiento externo en forma de precipitación, temperatura, radiación solar, humedad y viento. Así pues, el SSM es tanto un integrador de las condiciones climáticas como un impulsor del tiempo y el clima locales, y desempeña un papel importante en los ciclos mundiales del agua, la energía y el carbono.

El conocimiento de la dinámica de la humedad del suelo es importante para comprender los procesos en muchos ámbitos ambientales y socioeconómicos, por ejemplo, sus repercusiones en la vitalidad de la vegetación, el rendimiento de los cultivos, las sequías o la exposición a las amenazas de inundación. El Sistema Mundial de Observación del Clima (SMOC) reconoce la humedad del suelo como una variable climática esencial (VCE).

Agresta, actualmente no recoge ni guarda los registros históricos de las plataformas de datos abiertos mencionadas y por tanto el primer paso consistirá en volcar y almacenar la información descrita para posteriormente tratar y transformar de manera que resulte útil para la implementación de un modelo predictivo de la humedad del combustible vivo.

Una vez desarrollado el modelo, se automatizará los procesos de descarga, tratamiento y almacenamiento de los datos, de manera que, en el futuro si es necesario reentrenar los modelos, esta información estará disponible junto con los nuevos ficheros del INIA.

El valor del proyecto reside tanto en los procesos ETL, extracción, transformación y carga de los datos como en el desarrollo de un modelo que permita hacer predicciones con alta precisión. Aunque la mayor parte de los datos que se van a utilizar, serán datos abiertos, la transformación de estos impactará de manera significativa en los resultados del modelo. Hay que señalar que, una vez hecho el desarrollo del modelo, con los datos proporcionados por INIA, Agresta podrá verificar si el modelo sigue ejecutando las predicciones con precisión, permitiendo reducir la demanda de muestra de campo si el modelo mantiene estable dicha precisión.

13 Análisis del entorno

Con el objetivo de salvaguardar nuestros recursos, la gestión forestal ha cobrado una gran importancia durante los últimos años debido a la concienciación de la población sobre la necesidad de proteger el medio ambiente, ecosistemas y recursos socioeconómicos.

Esta mayor concienciación se ha visto enfrentada a diversas situaciones estratégicas como son la necesidad de terreno para la edificación y expansión de las zonas urbanas, la necesidad de tierras cuyo uso sea destinado a la agricultura y ganadería o la extracción de biomásas como materia

prima para la obtención de madera y papel. Esto, unido a las causas medioambientales, como las tormentas, plagas o incendios, está causando que el proceso de deforestación a nivel mundial cada vez sea más importante y significativo. Sin embargo, en España cada vez hay más superficie forestal, pero no existe una gestión adecuada de masas forestales que evite la excesiva acumulación de combustible y decaimiento por estrés hídrico, plagas, enfermedades, incendios, etc. derivados del cambio climático.²³

A continuación, se presenta un gráfico con las principales causas de deforestación a nivel mundial, donde los incendios forestales se muestran como la causa natural principal en aumento de los últimos años. Las tres causas restantes son el resultado de la actividad humana y por ello las instituciones ponen cada vez más recursos en sistemas de prevención, ya que el ser humano puede intervenir positivamente en su reducción.²⁴

Así han evolucionado las distintas causas de deforestación en el mundo

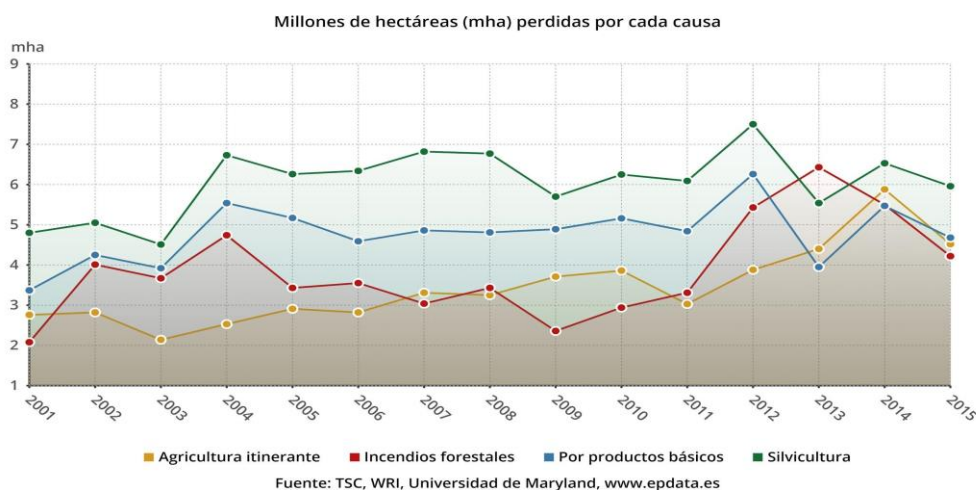


Ilustración 3 Evolución de las distintas causas de deforestación en el mundo

En el caso de España, el auge de las inversiones en materia de prevención e investigación, unido a las cada vez más restrictivas políticas agrarias, industriales y forestales, han conseguido una reducción de las hectáreas quemadas por incendios forestales²⁵, lo que paradójicamente supone una mayor acumulación de combustible vivo, aumentando así el peligro y la virulencia de futuros incendios. A continuación, se puede apreciar la reducción del número de hectáreas quemadas en

²³ “Proceso de deser...” <https://www.fundacionaquae.org/conoces-que-es-la-desertificacion-sus-causas-y-sus-soluciones/>
²⁴ “Situación de los bosques en Esp...” <https://www.epdata.es/datos/situacion-bosques-mundo-espana-datos-graficos/330>
²⁵ “Orientaciones estratégicas para la gestión de incendios forestales en...” https://www.mapa.gob.es/es/desarrollo-rural/temas/politica-forestal/orient_estrategicas_gestion_iiff-2019_tcm30-523469.pdf

España, siendo el 2018 el año con menos hectáreas quemadas desde el inicio del estudio, sin embargo, no se aprecia una tendencia clara en este aspecto.

Héctareas quemadas en incendios forestales en España

2018 es el año con menos hectáreas quemadas de este siglo

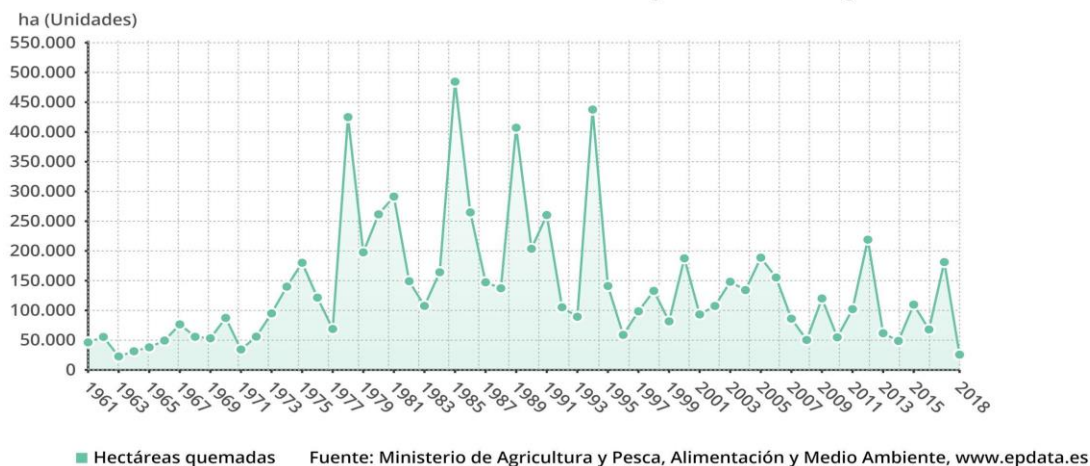


Ilustración 4 Hectáreas quemadas en incendios forestales en España

Otro de los factores clave que influye en reducir el número de incendios forestales, es la mayor concienciación de la población sobre la necesidad de cuidar el medio ambiente, lo que ha conseguido reducir en gran medida el número de incendios accidentales causados por la acción del hombre. Esto, se une a que cada vez la población presiona más a las administraciones tanto públicas como privadas para una mejor gestión forestal impactada por el abandono rural.

Sin embargo, no todos los factores son positivos en relación con el mantenimiento y cuidado de las regiones forestales. Factores sociodemográficos, incluyendo una mayor longevidad de la población, también afectan al aumento del abandono rural comentado. A continuación, se incluye un gráfico sobre la estimación de crecimiento de la población en los próximos años.²⁶

²⁶ “Perspectivas de la población mundial.” <https://www.un.org/es/sections/issues-depth/population/index.html>

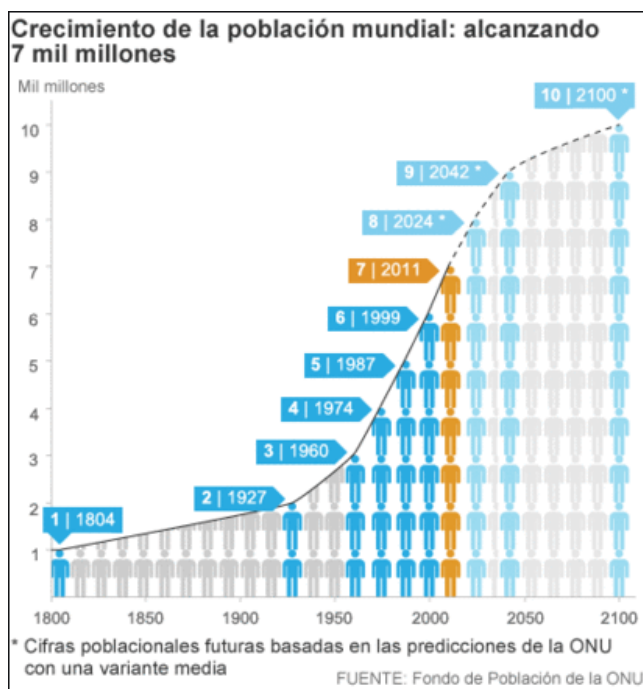


Ilustración 5 Crecimiento de la población mundial

Este crecimiento exponencial de la población, en materia económica y demográfica, supone una amenaza directa para las regiones forestales debido a la cada vez mayores necesidades alimentarias o de zonas de construcción, lo que supone un previsible aumento de las tres causas principales de deforestación causadas por el hombre, como son la agricultura, silvicultura o la producción de materias básicas.

Y es debido a estas previsiones, por las que cobra una mayor importancia para las administraciones invertir e investigar en la reducción de incendios forestales, ya que es una de las cuatro razones como hemos visto anteriormente, que más peso tiene en la deforestación, y sobre la cual el hombre puede tener una implicación positiva en materia de conservación medio ambiental.

A parte de la protección medioambiental, la extinción de incendios tiene como objetivo de reducir el impacto ecológico y socioeconómico, que como hemos visto son dos de las causas que no arrojan datos positivos en cuanto a la evolución previsible de la presión sobre las masas forestales. De ahí que la extinción siga siendo el capítulo del presupuesto ambiental en el que España hace una mayor inversión anualmente. Sin embargo, a nivel nacional hace falta repensar el modelo y empezar a aumentar la inversión en gestión forestal y ordenación de usos del suelo a escala paisajística hacia modelos más resilientes a los grandes incendios forestales.

14 Análisis competitivo

El panorama competitivo en el ámbito de predicción de incendios es bastante extenso ya que existe una gran demanda social y administrativa para combatir dicha amenaza. Esto genera cada vez más estudios que permitan incorporar las nuevas tecnologías a los conocimientos existentes para optimizar los recursos de lucha contra el fuego. Sin embargo, en el marco de la estimación de la humedad de combustible vivo, las soluciones actuales no son eficientes ya que implican un importante trabajo sobre el terreno que conlleva a que este tipo de controles sean altamente costosos.

Por otro lado, están surgiendo durante los últimos años nuevas alternativas a consecuencia de la aplicación de nuevas tecnologías y del uso de datos como imágenes satelitales, sensorización o dispositivos inteligentes. Estas nuevas vertientes permiten una reducción del coste debido a la disminución de la necesidad de tener equipos trabajando en campo y a su vez, permite el aumento de la precisión de las predicciones debido a que se tienen en cuenta muchas más variables de una manera continua y automatizada, lo que está convirtiendo estas nuevas alternativas en cada vez más comunes ante este tipo de problemas.

A continuación, indicamos algunos ejemplos de empresas y tecnologías que podrían entrar en competición con este proyecto, siendo los sistemas IA o los sistemas de teledetección por satélite los que más peso están tomando los últimos años.

- Wildfire Risk Atlas. La solución perteneciente a la empresa española Tecnosylva, permite realizar un análisis multicriterio de todos los parámetros que forman parte del concepto de riesgo de incendio. De manera complementaria, permiten realizar modelización avanzada mediante la Wildfire Analyst API, dando así soporte a todos los agentes implicados.²⁷
- La empresa española Meteogrid, está especializada en modelos climáticos, “light detection and ranging” y segmentación del territorio. Esto supone la generación de mapas de combustibles basados en estructuras forestales que permiten conocer el riesgo y el comportamiento singular del fuego para cada estructura.²⁸

²⁷ Tecnosylva Risk Atlas. Incendios Forestales. 2020. <https://tecnosylva.es/>

²⁸ Meteogrid Servicios. Incendios Forestales. 2020. <https://www.meteogrid.com/servicios/incendios-forestales/>

- La propia agencia estatal de meteorología española (AEMET) proporciona mapas de riesgo de incendios basados en los datos meteorológicos recogidos en las diferentes estaciones meteorológicas que poseen.²⁹
- La empresa española Meteosim, proporciona pronósticos meteorológicos y de riesgo de incendios utilizando dos tipos de modelaje avanzado, como son los modelos WRF (weather research and forecasting) y el modelo canadiense de riesgo de incendios, uno de los más reputados en cuanto a precisión y predicción en incendios forestales.³⁰
- BeeFireDetection. Solución que pertenece a la empresa portuguesa Compta, la cual ha integrado el uso de Inteligencia Artificial (Watson de IBM), con diferentes dispositivos IOT como pueden ser espectrómetros, cámaras digitales de video vigilancia y termografía para detectar incendios forestales a grandes distancias. De manera adicional, gracias al uso de datos provenientes de estaciones meteorológicas locales y datos de The Weather Company de IBM, son capaces de proporcionar un pronóstico de riesgo de incendio muy preciso, lo que permite la prevención de estos.³¹
- La Universidad Politécnica de Valencia en conjunto con el Centro de Coordinación de Emergencias de la Generalitat, están desarrollando un proyecto de Tecnologías de la teledetección por satélite para la gestión de alertas frente a los incendios forestales en el ámbito local de la Comunitat Valenciana. En este proyecto, se establece un nuevo modelo predictivo que actualiza y define los niveles de riesgo de incendio forestal a través de tecnologías de Teledetección por satélite. Además, se obtiene información sobre datos cuantitativos que ayuden a tomar decisiones para hacer frente al incendio forestal una vez se ha producido.³²
- La empresa finlandesa Arbonaut, también proporciona modelos de predicción de incendios de alta precisión aunando la tecnología LiDAR e información meteorológica.³³

15 Análisis y diagnóstico

A continuación, se muestra un análisis DAFO sobre la gestión del combustible para la prevención de incendios forestales en España y reducir el impacto ecológico y socioeconómico. El análisis ha

²⁹ “AEMET Predicción. Interpretación del servicios...” <http://www.aemet.es/es/eltiempo/prediccion/incendios/ayuda>

³⁰ “Meteosim Riesgo de Incendios. Sistema operativo...” <https://www.meteosim.com/es/gestion-de-riesgos-incendios/>

³¹ “Bee2FireDetection Solutions. Early fire detection and decision supp...” <http://www.bee2firedetection.com/#solution>

³² “Predecir incendi...” https://www.elperiodic.com/predecir-incendios-forestales-mediante-sistema-satelites_663197

³³ “Forest Fire Management. Forest Fire Risk...” https://www.arbonaut.com/files/pdf/Forest_Fire_Management_en.pdf

sido realizado por el Instituto Universitario de Investigación de la Universidad de Valladolid³⁴ ³⁵ y ha sido interpretado por nuestro grupo de trabajo.

Las *debilidades (D)* identificadas que limitan una adecuada gestión de los combustibles tienen una alta relevancia social y económica. Destacamos como más relevantes los conflictos que existen en planes de ordenación territorial y dentro de los presupuestos de estos planes no se tienen en cuenta partidas destinadas a la prevención. De forma genérica, las acciones de extinción prevalecen sobre las de prevención. Entre otras debilidades identificadas se encuentran la falta de incentivos para propietarios privados, el uso tradicional del fuego en áreas rurales o la falta de información pública sobre los beneficios económicos y ambientales de invertir en prevención de incendios forestales

Se destacan como *amenazas externas (A)* la baja importancia a nivel político a lo que se suma el envejecimiento de la población rural con el abandono de terrenos y el bajo valor de mercado del producto forestal. Por otro lado, hay otros factores externos con una relevancia algo menor como el escepticismo en el uso de quemas prescritas por parte de gestores y población, la falta de continuidad en la asignación de presupuestos al sector forestal o el cambio climático.

Las *fortalezas (F)* más importantes que se observan son el marco legal adecuado a nivel nacional y regional (Comunidades Autónomas), la existencia de buenas prácticas de acciones cooperativas para la prevención, tanto a nivel institucional como entre propietarios y administración y la existencia de equipos de bomberos forestales especialistas en prevención de incendios. Algunas que podrían tener menor impacto o menos relevantes son la existencia de planes de pastoreo controlado, acciones de prevención que fomentan el desarrollo rural, existencia de programas exitosos de conciliación de intereses.

Ante las técnicas disponibles, políticas desarrolladas y a nivel social, las *oportunidades (O)* con que nos encontramos incluyen los aprovechamientos energéticos de la biomasa forestal, de redes de trabajo y colaboración existentes entre los propietarios, gestores, empresas, administración e investigadores, el interés e impulso de la comunidad científica y técnica en el desarrollo de tecnologías integradas en la prevención de incendios forestales. Otras oportunidades pueden observarse fomentando el pastoreo controlado con subvenciones de la PAC y pagos por servicios ambientales, promover denominaciones de origen de productos forestales, ganaderos o agrícolas

³⁴ “La gestión de los combustibles forestales: oportunidades...” <http://sostenible.palencia.uva.es/content/la-gestion-de-los-combustibles-forestales-oportunidades-para-la-prevencion-de-incendios-en>

³⁵ “Forest fuel management for wildfire prevention in Spain: a quantitative SWOT analysis. International Journal...”

compatibles con la prevención de incendios y una producción forestal sostenible de productos maderables y no maderables.



Ilustración 6 Diagrama DAFO

Como se puede comprobar, destacamos aquellas con más relevancia e impacto con mayor tamaño y posicionadas más a la derecha del gráfico respectivamente. Dentro de las fortalezas y oportunidades existen numerosos grupos de trabajo que afrontan éstas para su mejora y optimización. Para este proyecto, destacamos la oportunidad de mayor impacto que se refiere al impulso de la comunidad científica y tecnológica, y por ello corroborar el beneficio y aprovechamiento de esta oportunidad con este proyecto.

16 Modelo de negocio

El modelo de negocio constituye la base sobre la que se definen las estrategias relevantes del proyecto. A continuación, se muestra un esquema del modelo de negocio siguiendo una estructura Canvas:

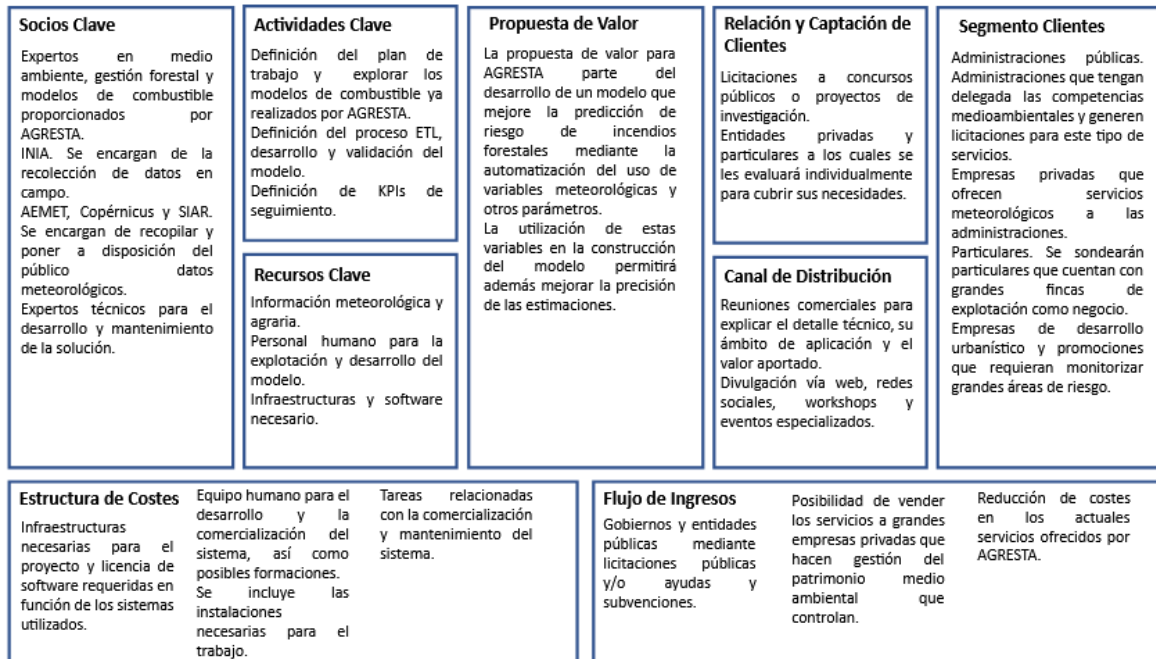


Ilustración 7 Diagrama CANVAS

16.1 Definición del modelo de negocio

16.1.1 Propuesta de valor

La propuesta de valor para Agresta parte del desarrollo de un modelo semiautomático que mejore la predicción de riesgo de incendios forestales mediante la integración de variables meteorológicas entre otras. A partir de esta información, se calculará la humedad del combustible vivo de una determinada zona, cuyo valor se relaciona directamente con el riesgo de incendio.

El modelado de esta información, junto a los servicios que ya ofrece, va a permitir a Agresta ofrecer unos servicios con un alto valor añadido en el área de incendios forestales, mejorando así su posicionamiento en el sector.

La utilización de variables meteorológicas en la construcción del modelo permitirá además mejorar la precisión de las estimaciones, automatizar los procesos de cálculo para dar resultados en tiempo casi real, garantizando una información más concreta y acertada a los equipos de prevención, que la obtenida actualmente mediante el procesado de imágenes satelitales.



Ilustración 8 Resumen de la propuesta de valor

El modelado de la información no solo va a permitir a Agresta la mejora de sus servicios, si no que su utilización permitirá reducir la necesidad de descarga y procesado de imágenes satelitales, así como la reducción de las salidas a campo para la toma de muestras, ya sea por ellos mismos o por subcontratación a otras empresas, lo que implica una reducción de costes en los servicios ofrecidos por Agresta.

Además, otro de los objetivos que se persigue en el proyecto, es poder conseguir una predicción de las condiciones del entorno lo más precisa posible, para poder adoptar medidas preventivas con antelación y reducir así los riesgos en caso de que se produzca un incendio, ofreciendo un valor diferencial con respecto a competidores.

16.1.2 Socios clave

El socio clave principal de este proyecto es la propia Agresta, ya que, mediante sus expertos en la materia, aportan la experiencia y el conocimiento necesario para guiar al equipo de proyecto en la búsqueda del valor de los datos.

De manera complementaria, se identifican los socios clave necesarios para llevar a cabo el proyecto en el ámbito de recopilación de la información a partir de la cual se realizará el modelado:

- INIA: el Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria es el encargado de recepcionar las muestras de campo recogidas por los servicios de extinción de incendios de la Comunidad de Madrid y realizar las muestras en campo y su posterior análisis de la humedad de las diferentes plantas en las zonas de estudio.
- AEMET: la Agencia Española de Meteorología aportará al proyecto toda la información meteorológica de las zonas de estudio a partir de sus datos abiertos. La AEMET cuenta con una serie de estaciones meteorológicas distribuidas por la Comunidad de Madrid que nos permiten obtener una información de alta calidad y precisión.

- SIAR: El Sistema de Información Agroclimática para el Regadío es una infraestructura que captura, registra y divulga los datos agroclimáticos necesarios para el cálculo de la demanda hídrica de las zonas de riego. Entre los datos abiertos de los que dispone, los expertos de Agresta identifican la evapotranspiración potencial (ETP) como una variable de gran importancia a la hora de construir el modelo.
- COPERNICUS: Es un proyecto dirigido por la Agencia Espacial Europea y la Agencia Europea del Medio Ambiente cuyo objetivo es lograr una completa, continua y autónoma capacidad de observación terrestre de alta calidad. Entre los datos abiertos de los que dispone, los expertos de Agresta identifican la humedad del suelo como una variable de gran importancia a la hora de construir el modelo.

De igual manera, nuestros profesores, así como la EOI serán una fuente de conocimiento para tener en cuenta en la elaboración del proyecto.

16.1.3 Actividades clave

La mayor parte de las actividades claves de este proyecto giran alrededor de los datos a partir de los cuales se realizará el modelado. A continuación, se detalla cada una de ellas:

- Exploración de los modelos de estimación de humedad de combustible vivo utilizados actualmente por Agresta. Será necesario conocer las técnicas que Agresta utiliza actualmente para ofrecer sus servicios en el área de Incendios Forestales, así como profundizar en el área de trabajo de la compañía para familiarizarse con la terminología.
- Proceso ETL para la obtención de la información a partir de la cual se realizará el modelado. Agresta no dispone de procesos que guardan en bases de datos la información requerida, por tanto, será necesario extraer, tratar y cargar la información desde cero. Esta extracción se realizará en primer lugar de manera histórica, para el desarrollo del modelo y posteriormente se configurará para la descarga de las predicciones futuras.
- Esto también nos permitirá conocer y familiarizarnos con el modelo de datos que utilizaremos en el desarrollo del modelo.
- Desarrollo del modelo. A partir de los datos obtenidos, se desarrollará la solución automatizada de manera que se obtengan resultados con mayor rapidez y mejores valores que los obtenidos por las técnicas actuales. Durante esta actividad, es posible que se puedan crear nuevas variables relevantes a partir de las que ya existen y que permitan mejorar la precisión del modelo.

- Validación del modelo. Una vez realizado el desarrollo, será necesario confirmar el correcto funcionamiento del modelo. Para ello se seguirán comparando los valores de campo proporcionados por el INIA con los valores proporcionados por el modelo.
- Seguimiento del modelo mediante la identificación de los principales KPIs. Consistirá en definir qué vamos a medir y cómo lo vamos a hacer, para obtener métricas que permitan conocer la evolución del modelo.

16.1.4 Recursos clave

El principal recurso de nuestro proyecto serán los datos. En nuestro caso, el servicio planteado consistirá en la utilización de datos meteorológicos y agrarios en gran medida para la realización del modelo de humedad de combustible vivo. Estos datos serán en gran medida abiertos, siendo proporcionado por diferentes instituciones. Otra fuente de datos clave será la “verdad terreno”, obtenida mediante datos reales de humedad de combustible tomados en campo, que una vez procesados en laboratorio permitirán validar y monitorizar el modelo.

Otro recurso clave será el factor humano. Se necesitará personal experto en la materia que permita sacar partido de la información proporcionada por el modelo y personal capacitado tecnológicamente para el desarrollo de la solución y su mantenimiento.

Por último, hacer referencia a las necesidades de infraestructura y hardware para soportar la solución desarrollada, así como el software necesario para su desarrollo y mantenimiento.

16.1.5 Relación y captación de clientes

Debido a las características y el objetivo de los servicios ofrecidos, la venta de nuestros servicios se realizará en mayor parte mediante licitaciones a concursos públicos o proyectos de inversión con carácter de investigación, ya que son los que disponen de un mayor presupuesto e invierten en este tipo de servicios.

También, aunque en menor medida, se podrá contactar con entidades privadas, empresas y particulares a los cuales se les ofrecerá el servicio adaptado a sus necesidades. Dentro de estas empresas se incluyen también las que se encuentran en relación directa con administraciones y proporcionan reportes de riesgo de incendios basados en otras técnicas y cálculos.

16.1.6 Canal de distribución

El principal canal de distribución será mediante reuniones comerciales que permitan explicar los detalles técnicos, su ámbito de aplicación y el valor añadido que aportan los servicios prestados por Agresta.

Además, se realizará divulgación del servicio vía web, redes sociales, workshops y mediante la asistencia a eventos especializados.

Durante la realización de este proyecto, se realiza una solicitud para presentar una comunicación en el 8º Congreso Forestal Español 2021 donde se presentará esta metodología innovadora aplicada en este proyecto en el ámbito forestal y los resultados obtenidos. Se hace entrega de un resumen (abstract) y quedará pendiente de la selección final para presentar la comunicación y trabajo completo al comienzo de 2021.

16.1.7 Segmento de clientes

El proyecto está planteado para satisfacer las necesidades de tres posibles tipos de clientes:

- Administraciones públicas y/o empresas representantes. Los gobiernos y administraciones públicas son los mayores clientes potenciales del servicio desarrollado. El objetivo se focalizará en todas las administraciones que tengan delegada las competencias medioambientales y generen licitaciones para este tipo de servicios. Existen empresas que proveen a las administraciones de reportes de riesgo de incendios periódicos y por tanto también serán objetivo de comercialización de este servicio ya que las metodologías que utilizan no están muy desarrolladas.
- Particulares. Se sondearán particulares que cuentan con grandes fincas de explotación como negocio. En este tipo de casos el servicio deberá ser adaptado en función de las necesidades del particular.
- Empresas de desarrollo urbanístico y promociones que requieran monitorear grandes áreas de riesgo debido a su proximidad a las zonas de negocio.

No se descarta que, transcurrido un tiempo de maduración y aprendizaje, estos servicios se puedan abrir a otros segmentos de clientes, debido a que en la actualidad la concienciación social por el cuidado medio ambiental está en continuo crecimiento.

16.1.8 Estructura de costes

En este proyecto, la mayor parte de los costes vienen asociados al desarrollo y mantenimiento del servicio, más que su puesta en producción. A continuación, se enumeran los principales costes asociados al proyecto:

- Infraestructura. Se tratan de los gastos en hardware necesarios para que la herramienta pueda funcionar, entre los que destacan el equipamiento informático y los sistemas que albergarán las bases de datos para su almacenamiento. En el proyecto, Agresta aportará la infraestructura necesaria la cual se unirá a los dispositivos propios del equipo de desarrollo.
- Equipo humano. Será necesario contratar personal especializado o bien formar al personal de Agresta para el mantenimiento y evolución de la herramienta. Para su correcto desarrollo se estiman necesarios usuarios expertos en incendios forestales y en concreto en modelos de combustibles y humedad del combustible vivo, usuarios expertos en el desarrollo del sistema y un equipo comercial que presente la herramienta a potenciales clientes.
- Formaciones. Se incluyen todos los gastos asociados que pueda ocasionar los cursos de formación al equipo humano para el manejo de la nueva herramienta.
- Mantenimiento. El sistema requerirá adaptaciones, mejoras, reentrenamiento de los modelos conforme pase el tiempo. De igual manera, será necesario adaptar el servicio ofrecido en función de las necesidades particulares de cada cliente. También se incluyen en este punto las tareas de mantenimiento que pueden surgir de la actualización de la página web o puntos de contacto digitales con el cliente.
- Licencias de software de reporte. Será necesario adquirir licencias de sistemas BI para el reporte de los resultados ofrecidos por el modelo. Estos costes serán asumidos por Agresta.
- Tareas relacionadas con la comercialización del producto. Se incluirán todas las tareas de publicidad, marketing, asistencia a eventos especializados, cuyo objetivo sea la comercialización del servicio. Este coste, al igual a lo ocurrido con los puntos anteriores, será cubierto por Agresta.
- Instalaciones. Será necesario que el equipo de trabajo disponga de un espacio de trabajo colaborativo, ya sea en las oficinas de Agresta o espacios especializados.

Entre los costes asociados, no se incluyen particularidades que puedan surgir asociadas a necesidades específicas de cada cliente, como pueden ser sistemas IoT en campo, reporte especializado etc.

16.1.9 Flujo de ingresos

Como ya se ha visto en el segmento clientes, los ingresos que puede obtener Agresta a partir de los servicios desarrollados en este proyecto, se podrían agrupar en dos posibilidades:

- Gobiernos y administraciones públicas. En mayor medida, son los gobiernos y las administraciones públicas que gestionan la materia medio ambiental, las que publican licitaciones enfocadas al estudio de los modelos de combustible y el riesgo de incendio en sus zonas de control. En este tipo de licitaciones es donde Agresta va a aportar un valor añadido al tener un servicio de bajo coste y alta precisión que permita identificar el riesgo de incendio de zonas concretas.

Al igual que las licitaciones públicas, existen numerosas ayudas y subvenciones públicas a las que Agresta va a poder acceder debido al carácter innovador y de vanguardia aportado por el nuevo servicio, lo que supondría igualmente una fuente de ingresos muy valiosa no solo para el mantenimiento del producto, si no para la mejora y evolución de sus servicios.

- Particulares y empresas. En menor medida a las instituciones públicas, en el sector privado también puede aparecer un nicho de mercado que esté interesado en este tipo de soluciones de control y seguimiento. En este caso, los beneficios esperados por Agresta son menores, ya que actualmente muy pocas corporaciones y empresas invierten en este tipo de sistemas, al igual que los servicios ofrecidos seguramente tengan que verse adaptado, con el coste que esto supone para Agresta, con el objetivo de satisfacer las necesidades específicas del sector. Sin embargo, aunque no se considere una fuente de ingresos segura, se espera que con la reducción de costes de los servicios de Agresta asociados a este proyecto, se puedan generar oportunidades de negocio que permitan introducir los servicios también en el sector privado.

Sin embargo, el mayor beneficio de este proyecto no está relacionado con los nuevos clientes que se podrían captar, si no por la reducción de costes asociados a los sistemas actuales utilizados para la generación de modelos humedad de combustible vivo. Entre los costes reducibles en servicios similares ofrecidos por Agresta se encontraría principalmente:

Procesamiento de imágenes satelitales. Actualmente Agresta dedica numerosos recursos propios al procesamiento de las imágenes satelitales (MODIS y SENTINEL 2) a partir de las

cuales genera sus modelos de combustibles. Estos procesamientos de imágenes le suponen a Agresta un esfuerzo elevado debido al alto coste que tiene este tipo de tareas.

Contratos de muestreo en campo. Actualmente la Comunidad de Madrid financia a través de sus servicios de extinción de incendios tanto el muestreo en campo como el tratamiento de las muestras por parte del Laboratorio de Incendios Forestales del INIA, para poder comparar los valores de la humedad de las plantas reales con los teóricos obtenidos a partir del procesamiento de imágenes satelitales. El coste de este sistema de toma de muestras no podría ser sufragado por Agresta en caso de poner en marcha el servicio, por lo que abaratar al máximo posible esta fase del proceso y pasar a un sistema en el que la toma de datos en campo se utilizara simplemente para validación de los resultados de los modelos es crucial para pasar a de un sistema experimental a uno operativo.

17 Plan de acción

17.1 Definición del alcance del proyecto

17.1.1 Objetivos

El objetivo que se persigue con este proyecto es mejorar el modelo actual que se utiliza para la predicción eficaz del riesgo de incendio en cuanto a su automatización y precisión.

Nuestro equipo de proyecto ha realizado varias evaluaciones e investigado aquellos factores inminentes que ayudarán a alcanzar este objetivo.

Como punto de partida, vamos a generar un modelo a partir de valores meteorológicos abiertos, que permita conocer con la mejor precisión posible la humedad del combustible vivo de la cubierta forestal, automático y que a su vez sea de bajo coste reduciendo costes externos.

Mediante la incorporación y el análisis de datos abiertos como pueden ser imágenes satelitales o datos meteorológicos se hará menos imprescindible las tomas y tratamiento de muestras en campo, siendo este último bastante costoso y dilatado en el tiempo.

Además, indirectamente, con un modelo más eficaz como el planteado, se apoyará a una gestión de los recursos que se destinan a la prevención de incendios forestales por parte de gobiernos, administraciones públicas, particulares y empresas del sector especializadas. Una vez finalizado el modelo y tras su puesta en marcha, se podrá llevar a cabo una monitorización durante los 3 meses posteriores. Durante este tiempo se podrán realizar comprobaciones reales y validaciones con los datos de campo recogidos hasta el momento.

Debido a la época del año cuando se planea el lanzamiento del nuevo modelo y debido a las condiciones meteorológicas de otoño e invierno, estimamos un periodo razonable de uso del modelo de 12 meses para testar su eficacia en la época de máximo riesgo: la campaña de incendios, de cara a garantizar que cumple con el objetivo.

Agresta podrá ofrecer servicios de modelos de humedad de combustible vivo de manera operativa, a un menor coste que el que actualmente tiene a nivel de investigación, en el que se utiliza el procesamiento de imágenes satelitales y muestras en campo.

En esta primera fase se pretende modelizar una zona concreta cercana a Robledo de Chavela, donde se presenta una cobertura forestal formada por matorral de jara pringosa (*Cistus ladanifer*), ya que es el lugar donde se están realizando diversos estudios sobre modelos de combustible vivo, y esto nos permitirá comparar la precisión del modelo actual a partir de imágenes satelitales Modis y Sentinel con los valores obtenidos con el nuevo modelo.

En fases posteriores se podrá evaluar cómo afecta sobre la precisión del modelo la existencia de estructuras forestales más complejas donde estén presentes estratos arbóreos y arbustivos con presencia de varias especies, así como realizar la extrapolación del modelo a otras zonas en las que pueda aplicar según la calidad de los resultados que generaría.

En esta primera fase será necesario obtener los datos históricos meteorológicos desde el 2014, año a partir del cual se empezaron a tomar muestras en campo, y lo que nos permitirá realizar el entrenamiento del modelo. Estos datos históricos serán tratados y enriquecidos en función de la información obtenida. Toda esta información será almacenada en la nube para que pueda ser utilizada desde cualquier dispositivo y en cualquier lugar. Una vez obtenida toda la información posible, se consultará con expertos para identificar las variables más significativas, y se acabará desarrollando el modelo. Este modelo permitirá conocer la humedad de combustible vivo a partir de los futuros datos meteorológicos, sin necesidad de tomar muestras en campo o realizar costosos procesados de imágenes.

17.1.2 Métricas

Los indicadores clave de rendimiento (KPIs) que se proponen nos permitirán controlar el progreso del proyecto. Los KPIs se van a agrupar en las siguientes categorías:

1. Análisis financiero. Su objetivo será cuantificar de una manera fiable la reducción de gastos en la elaboración de modelos de humedad de combustible vivo respecto a las técnicas utilizadas actualmente.

2. Análisis de rendimiento del modelo. Se realizará una comparativa entre los valores obtenidos mediante el procesamiento de imágenes y los obtenidos por modelaje de manera que se pueda evaluar la efectividad del desarrollo.
3. Evolución de la clientela. Con el objetivo de confirmar el aporte de valor del nuevo sistema, se evaluará la evolución del número de clientes. De igual manera servirá para evaluar la necesidad de desarrollar fases posteriores del proyecto que permitan extrapolar el modelo o incluir nueva funcionalidad.

A continuación, se muestra una tabla esquemática en la que se muestran todos los indicadores definidos en las temáticas mencionadas, indicando su cálculo y periodicidad:

Objetivo	Indicador	Cálculo	Periodicidad
Análisis Financiero	% Reducción de gasto	Se calcula la diferencia en % entre los gastos del servicio del procesamiento de imágenes y los gastos de utilización del modelo desarrollado	Mensual
	Aumento de beneficios	Se calcula el promedio de beneficio por cliente en comparación con años anteriores	Trimestral
	% Gastos fijos Vs Variables	Se calcula la diferencia de gasto fijo vs variable por proyecto	Mensual
	Gastos de explotación	Se contabiliza el gasto de implementación y explotación del servicio	Mensual
Análisis de Rendimiento	% Precisión del modelo	Se calcula el % de diferencia entre los valores resultantes del modelo y los valores del muestreo en campo	Diario

	Solicitud de adaptación	Se contabilizan el número de adaptaciones necesarias por proyecto para satisfacer las necesidades	Trimestral
	Incidencias de uso	Se contabiliza el número de incidencias en el servicio	Trimestral
Evolución de la Clientela	Aumento del número de clientes	Se evalúa la cantidad de nuevos clientes que reciben los servicios de AGRESTA	Trimestral
	Marketing y comercialización	Se calcula el beneficio según inversión (ROI) de los procesos de marketing y comercialización	Trimestral
	Retención de la clientela	Se calcula el % de clientes que renuevan los servicios de AGRESTA	Trimestral
	Satisfacción de la clientela	Se calcula el grado de satisfacción de la clientela a partir de encuestas de satisfacción sobre los servicios recibidos	Anual

Tabla 3 Esquema de indicadores KPIs claves en el rendimiento del modelo

17.2 Análisis de actividades. Modelo lógico y arquitectura técnica

17.2.1 Modelo lógico

No nos cabe duda del valor añadido que este proyecto aportará tanto a la empresa como a niveles de administraciones públicas y sociales. Este proyecto posiciona la información como el activo principal que permitirá generar un valor añadido a los servicios actuales de la empresa, y de ahí

conseguir una monetización de estos. Las fases del desarrollo del proyecto se detallan a continuación:

17.2.1.1 Identificación de los datos

Durante esta fase se han considerado aquellas actividades que permitirán obtener dicho valor de la manera más eficiente posible. Hemos de destacar que, durante esta fase de identificación, se observó que Agresta previamente no recopilaba ni almacenaba ningún tipo de información de fuentes externas más que los datos obtenidos a partir de las muestras tomadas en campo. En todos los casos se trata de datos estructurados a partir de los cuales se elaborará el modelo.

Se ha identificado el siguiente modelo lógico que permitirá extraer valor de todos los datos que el proyecto necesita para su modelo:

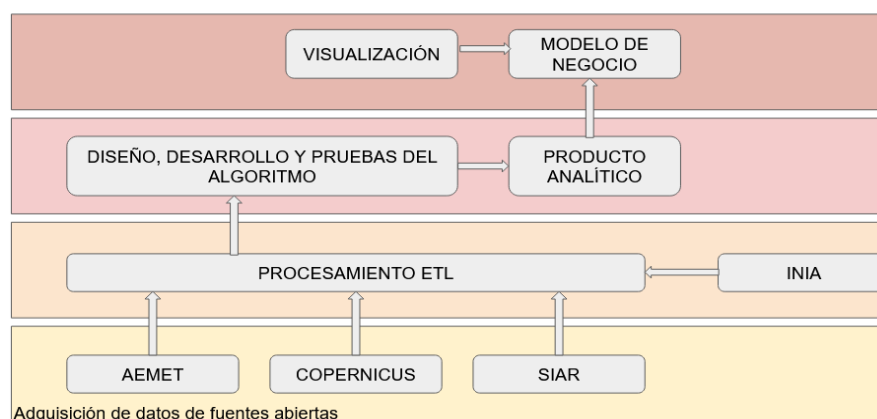


Ilustración 9 Modelo lógico

17.2.1.2 Adquisición de datos

La información de datos abiertos de las tres fuentes externas, AEMET, Copernicus y SIAR necesitan adquirirse mediante cierta estrategia, ya que cada fuente presenta un mecanismo diferente para disponer de los datos. El detalle del código que se ha llevado a cabo para cada una de las extracciones se encuentra en los apéndices correspondientes de este documento.

Para la obtención de los registros meteorológicos disponibles en la plataforma de AEMET, se ha ejecutado un script en lenguaje R que conecta de manera directa con los datos para su descarga a través de una API REST (Anexo 6 AEMET.pdf). Se han volcado registros históricos de todas las estaciones meteorológicas disponibles en la Comunidad de Madrid, que contienen información sobre variables climatológicas que a priori pueden ser relevantes para la obtención de la humedad de los combustibles vivos y que se corresponden con el periodo temporal de la recogida de muestras de estudio (datos tomados en campo por el INIA), de 2014 hasta la actualidad.

(Nota: en nuestro caso, denominamos fecha de estudio al día de la recogida de muestra en campo, y por tanto fecha referente para el resto de los registros de todas las fuentes durante los procesos ETL)

En cuanto a los datos de SIAR, estos han sido obtenidos mediante Web-scraping³⁶ (Anexo 5 SIAR.pdf), seleccionando el periodo temporal de la recogida de muestras de estudio y para las estaciones meteorológicas disponibles en la Comunidad de Madrid, ya que posteriormente se relacionarán las variables disponibles para cada estación con el posicionamiento más próximo a cada zona de estudio, de igual manera se procederá con las estaciones de AEMET.

Con respecto a los datos disponibles en la plataforma de Copernicus, estos han sido volcados a través de una API REST (Anexos 1 y 2) con conexión directa a los datos. Con respecto a los ficheros obtenidos hay que destacar que se trataba de ficheros diarios en formato NetCDF^{37 38 39}. Este tipo de archivos tienen un formato destinado a almacenar datos científicos multidimensionales (variables) como la temperatura, la humedad, la presión, la velocidad del viento o la dirección. Cada una de estas variables se puede mostrar mediante una dimensión (por ejemplo, el tiempo). Por tanto, los datos de un archivo NetCDF se almacenan en conjuntos. Adicionalmente se hace referencia a otros anexos relacionados con otros procesos llevados a cabo en Copernicus (Anexos 3.1, 3.2, 4).

En nuestro caso los datos tridimensionales (3D), como p.ej., la humedad de un área que varía con el tiempo se almacena como series de conjuntos bidimensionales tal y como se muestra en la figura siguiente:

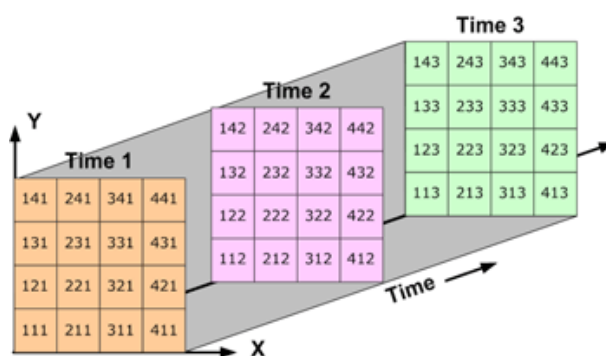


Ilustración 10 Almacenamiento de la información en Copernicus

³⁶ “Metodologías y herramientas. Como...” <https://openwebinars.net/blog/como-hacer-web-scraping-con-selenium/>

³⁷ Extracting time series data from a netCDF file into a CSV.

https://www.youtube.com/watch?v=hrm5RmsVXo0&list=PLLxyob7YmEE8S3QDs1PZQkiBxA4zn_Gx&index=3

³⁸ Unidata Python Workshop - Monday June 20 part 1 - git, conda. <https://www.youtube.com/watch?v=8x5oe-OVpoM&list=PLnXN3uoWXJ29QIHdAeCXQflx0n2yMUjRx>

³⁹ “Detailed examples for atmospheric science and...” <https://unidata.github.io/python-training/gallery/gallery-home/>

Para la comprensión de este tipo de ficheros se visualizaron en QGIS y Panoply, con la finalidad de entender cómo estaban contruidos los arrays que componen y estructuran este tipo de ficheros.

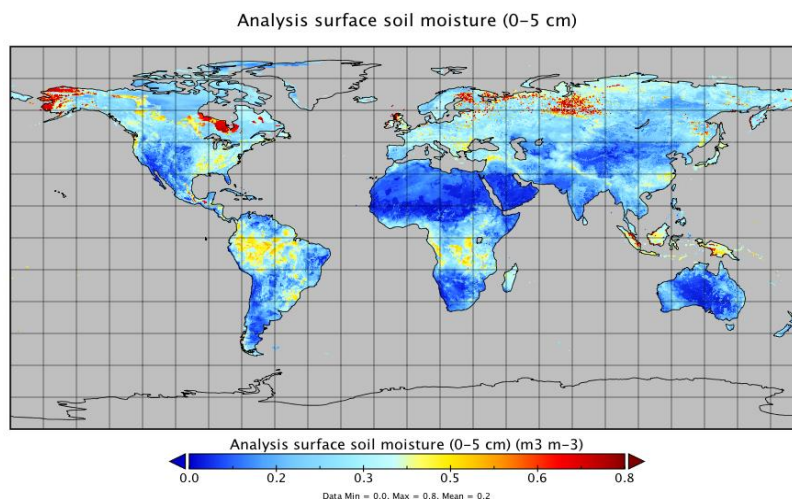


Ilustración 11 Visualización en la herramienta Panoply de los datos descargados por SIAR

Una vez volcados los ficheros diarios se procedió a compilar los mismo de manera anual, con la finalidad de conseguir la humedad del suelo para las coordenadas correspondientes a cada una de las zonas de estudio de las que se han obtenido las muestras de campo

Una vez obtenidos estos datos, se procede a examinar la calidad y ver si hay algún error a priori que supusiera tener que realizar algún paso adicional en la fase de transformación. En nuestro caso si es así, ya que se observan fallos en estaciones de la AEMET sin recogida de registros durante periodos de tiempo prolongados que además coinciden con fechas donde se han recogido muestras de campo.

17.2.1.3 Procesamiento e integración

Una vez volcados todos los ficheros de las distintas plataformas se procede a la limpieza y consolidación de los datos en un dataset de trabajo. Para la carga y la transformación de los datos se ha utilizado R como se ha mencionado anteriormente.

En primer lugar, se han cargado los ficheros para visualizar las características de los datos disponibles, seleccionando previamente entre las variables disponibles aquellas que a priori pueden resultar significativas y descartado aquellas que tienen poco valor predictivo. También se han detectado los registros vacíos y se han adoptado las primeras hipótesis para completar dichos registros. Para ello y tras consultar con la empresa nuestra propuesta, se decide incorporar o reemplazar esta falta de registros con datos de una estación próxima y que pueda ser representativa de la misma zona. Para ello se realiza una relación entre las estaciones sobre la

que se obtienen resultados vía la API Rest, teniendo en cuenta dos factores diferenciales, la cercanía y la altitud de la estación. En esta relación, se descartan las estaciones situadas en grandes urbes o en puertos de alta montaña, debido a que los datos no serían representativos para las zonas de estudio.

Una vez completados los registros vacíos siguiendo las recomendaciones de Agresta se procede a la creación de nuevas variables como: acumulados de precipitación (3, 7, 15 y 30 días anteriores a la fecha de estudio) y promedio de temperaturas mínimas, máximas y medias (3, 7, 15 y 30 días anteriores a la fecha de estudio). El cálculo de estos acumulados y valores promedios permitirán aportar al modelo información histórica de las variables más influyentes. Se decide tomar valores históricos hasta 30 días debido a que es suficientemente representativo para las formaciones de jaral estudiadas. En el caso de incluir masa forestal arbórea se ampliará hasta los 60 días.

Tras la carga y transformación de los datos, y tras consolidar un dataset de trabajo la siguiente tarea ha consistido en llevar a cabo un análisis descriptivo de las variables que conforman el dataset de trabajo intentando extraer relaciones de correlación y colinealidad entre dichas variables, así como detectar tendencias y estacionalidad para cada una de los predictores que conformarán el modelo a implementar.

La unificación final de los datos tratados se realiza de dos maneras diferentes:

1. Un primer archivo en el que únicamente se tendrá en cuenta la zona de estudio 5B determinada en el diseño de campo de la Comunidad de Madrid-desarrollado por el INIA, debido a que es de gran interés para Agresta ya que cuenta con modelos de predicción ya elaborados para dicha zona.
2. Un segundo archivo que tendrá en cuenta todas las zonas de estudio de la Comunidad de Madrid que contengan la especie de estudio (*Cistus Ladanifer*).

17.2.1.4 Análisis descriptivo

Durante la parte descriptiva de las variables se destacan las siguientes partes:

- Se ha realizado un análisis general de los datos que ha permitido conocer con mayor profundidad las diferentes variables
- Se confirma que la temporalidad de variables como las temperaturas mínimas, máximas y medias está presente. Sin embargo, no es así en las precipitaciones, aunque estas son más abundantes en ciertas épocas del año. Otras variables que presentan temporalidad son la radiación o las humedades relativas del aire, estando estas últimas influenciadas por las precipitaciones.

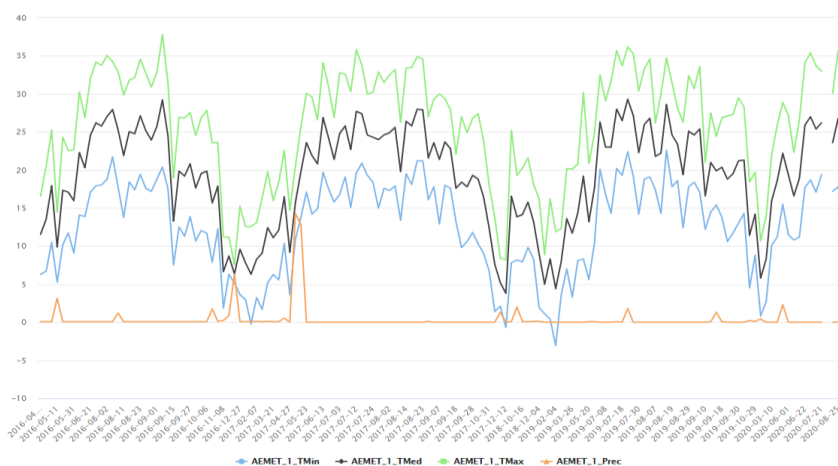


Ilustración 12 Estacionalidad de las temperaturas

- Se confirma la similitud de las variables entre estaciones próximas y de similar altitud, y con ello nos aseguramos tener una mayor representación de datos dejando el menor número de registros posibles vacíos (NA).
- Al realizar la correlación de variables, se han aplicado dos técnicas: pairs y pairwise (correlacionando las variables más relevantes con el valor de la humedad proporcionada por el INIA; INIA_FMC)

En ellas se puede observar que las variables mejor correlacionadas con la humedad de la planta (INIA_FMC) son las temperaturas y precipitaciones, resultando con un mayor índice de relación las nuevas variables calculadas a 30 días de la fecha de estudio (medias y acumuladas).

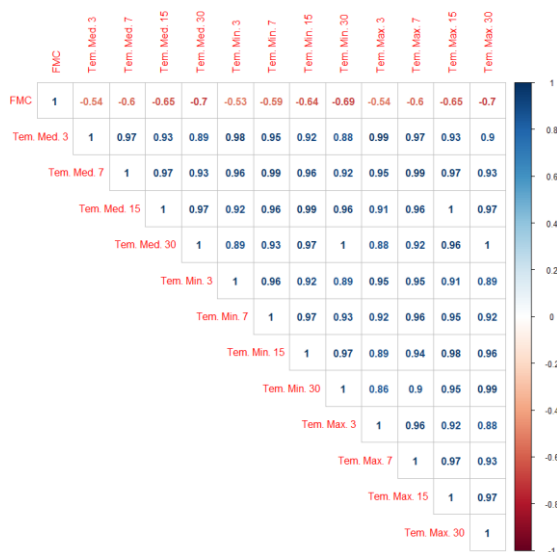


Ilustración 13 Análisis de correlación de las variables calculadas AEMET

Por otro lado, variables descritas en bibliografía y comentada por Agresta como posibles influyentes directos en la humedad como la radiación y la evapotranspiración, no parece que tenga mucha correlación en el análisis, alcanzando unos valores de -0,16 y -0,3 respectivamente.

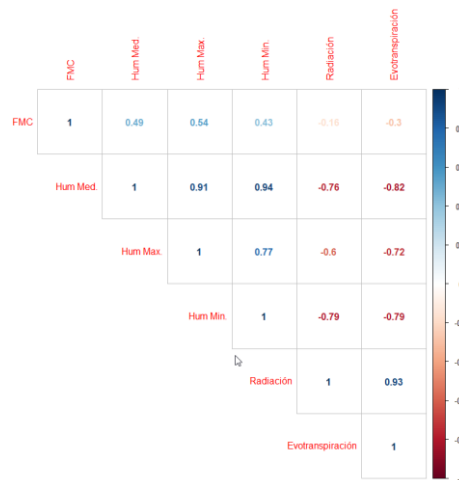


Ilustración 14 Análisis de correlación de las variables obtenidas de SIAR

Todo esto nos permite desarrollar un entendimiento más claro de nuestras variables, las cuales seleccionaremos en nuestro archivo de modelo para comenzar con esta fase

17.2.1.5 Almacenamiento de los datos

Una vez procesados e integrados todos los datos, se ha elegido la nube como el repositorio más eficaz para su análisis y distribución.

17.2.1.6 Descubrimiento y modelado

En esta fase se realiza el análisis de los datos capturados y tratados en fases anteriores.

Para ello vamos a aplicar diferentes técnicas de análisis de manera que podamos escoger un modelo que describa con la mayor precisión posible el estado de la humedad de las plantas según las variables almacenadas.

Se pretende detectar el patrón que permita conocer mediante información de datos abiertos, y con la máxima precisión posible, la humedad de la planta a partir de la cual se calculará el riesgo de incendio de una determinada zona.

Dividiremos los datos para nuestro modelo en dos grupos: de entrenamiento y test. Esta división se ha realizado de dos maneras posibles con el objetivo de maximizar la precisión de los modelos. En primer lugar, se ha hecho una división clásica 70/30 entre las muestras de entrenamiento y validación, y, en segundo lugar, debido a la baja cantidad de datos de la muestra, se ha probado

a dejar los datos del último año para la muestra de validación, siendo todos los anteriores utilizados para el entrenamiento, obteniendo así una distribución similar 75/25 aproximadamente.

Se aplican los modelos descritos a continuación para probar diferentes alternativas de modelo; hay tener en cuenta que en el código final que se entregará a la empresa solamente se dejará el modelo que mejor se ajuste, funcione y haga la predicción objeto de estudio. Las pruebas se eliminan, ya que solo nos han servido para ver la mejor aproximación y la mayor significancia de los datos de predicción del modelo.

- Stepwise: incluye modelos de regresión en los que la elección de variables predictoras se lleva a cabo mediante un procedimiento automático. El procedimiento toma la forma de una secuencia de pruebas- f en la selección o eliminación de variables explicativas.
- Multiple linear regression: permite generar un modelo lineal en el que el valor de la variable dependiente o respuesta (Y) se determina a partir de un conjunto de variables independientes llamadas predictores (X1, X2, X3...). Es una extensión de la regresión lineal simple. Se puede emplear para predecir el valor de la variable dependiente o para evaluar la influencia que tienen los predictores sobre ella
- Random forest: nos basamos en algunas características de este método como aportar estimaciones sobre qué variables son importantes en la clasificación, que genera una estimación objetiva interna de la generalización de error a medida que avanza la construcción de los árboles de decisión y tiene un método eficaz para la estimación de los datos faltantes.
- Stochastic Gradient Boosting: esta técnica persigue el objetivo de ir minimizando los residuos por iteraciones además de las ventajas de conseguir mejorar la capacidad predictiva y permitir estimar el “out of bag error”, y así facilitar la optimización de hiper parámetros. Incorpora el muestreo aleatorio de observaciones de entrenamiento. En concreto, el weak learner se ajusta empleando únicamente una fracción del set de entrenamiento, extraída de forma aleatoria y sin reemplazo (no con bootstrapping) por cada una de las iteraciones.
- Red neuronal: las redes neuronales nos permiten buscar la combinación de parámetros de entrada idóneos que permita reproducir lo más acertadamente posible la variable objetivo. En el proceso se ha realizado una normalización de los datos previa a la ejecución de la red.

A continuación, se muestra un resumen de los modelos empleados. Primero se presentan los resultados obtenidos para el estudio hecho sobre la zona 5B de Robledo de Chavela, donde se han

realizado otro tipo de simulaciones basadas en imágenes satelitales. Para cada modelo, se han calculado 4 mediciones que nos permiten conocer la calidad de los resultados.

Robledo de Chavala (5b)	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,80	0,78	0,83	0,80	0,82
MSE	344,77	411,78	236,51	351,02	319,96
RMSE	18,57	20,29	15,38	18,74	17,89
MAE	14,30	16,54	12,18	14,77	14,58

Tabla 4 Resultados ejecución del modelo Zona Robledo de Chavala

En general, podemos ver unos valores de R cuadrado similares. Esto se debe principalmente al escaso número de muestras ya que hasta 2016 no se empezaron a recoger muestras en esta zona. En este caso destaca el modelo Gradient Boosting como el modelo con mayor precisión debido a que es capaz de disminuir el impacto que tiene el reducido número de muestras de estudio.

También podemos observar, como los modelos obtenidos, tienen un valor de error aceptable para esta variable de estudio. En este caso, el modelo elaborado por Gradient Boosting será el que menos error cometa, ya que como hemos visto anteriormente, será el modelo óptimo para nuestra casuística.

Con el objetivo de obtener más conclusiones, se ha realizado el mismo estudio sobre todas las zonas en las cuales se encuentra la especie de estudio (*Cistus ladanifer*). A continuación, se muestran los resultados.

Cistus Ladanifer	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,47	0,68	0,66	0,66	0,65
MSE	657,30	406,62	417,81	447,60	418,48
RMSE	25,63	20,16	20,44	21,16	20,46
MAE	20,13	15,48	15,28	16,52	15,84

Tabla 5 Resultados ejecución del modelo en la especie Cistus Ladanifer

Podemos apreciar que la precisión ha disminuido al hacer el modelado para las diferentes zonas de estudio en la Comunidad de Madrid. Podemos obtener, así como conclusión que el error arrastrado en la estimación de los valores nulos provoca una caída de la precisión considerable que unido a los pocos valores de muestras reales en campo provocan los malos resultados. Este bajo valor puede mejorar si el número de muestras de campo aumentase y el tratamiento de los valores nulos se estimen a partir de un procesado más ajustado.

Como comentario final, una vez visto el resultado de los modelos, se observa que los resultados tienen un impacto directo con la cantidad de muestras del estudio.

El Gradient Boosting se adapta mejor a las condiciones de análisis con muestras de estudio disponibles, aunque en un futuro, con más datos disponibles podrían verse modificados estos resultados.

Por último, se ha implementado el modelo en todas las zonas de estudio del INIA en la que hace presencia la especie *Cistus Ladanifer*, con el objetivo de comprobar que el grado de precisión es similar al obtenido en la zona 5b de Robledo de Chavela.

Como se puede comprobar en la siguiente tabla, con la implementación de los mismos modelos, se obtienen resultados del mismo orden de magnitud, aunque en cuanto a la selección del modelo puede verse diferenciado respecto al modelo seleccionado para la zona 5 b de Robledo de Chavela.

Navalcarnero	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,81	0,78	0,77	0,81	0,81
MSE	369,42	205,65	296,20	218,26	202,90
RMSE	19,22	14,34	17,21	14,77	14,24
MAE	13,77	12,39	14,67	12,51	12,96

Robledo de Chavela	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,81	0,80	0,80	0,80	0,81
MSE	253,24	261,51	277,29	294,73	245,14
RMSE	15,91	16,17	16,65	17,17	15,66
MAE	12,55	12,74	12,54	13,88	12,92

Buitrago de Lozoya	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,69	0,76	0,72	0,80	0,72
MSE	569,58	466,54	387,03	284,38	539,99
RMSE	23,87	21,60	19,67	16,86	232,38
MAE	19,48	17,80	15,53	13,68	191,39

Guadarrama	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,78	0,78	0,80	0,82	0,75
MSE	410,02	330,05	273,11	520,74	372,48
RMSE	20,25	18,17	16,53	22,82	19,30
MAE	15,60	13,78	13,64	18,16	16,15

Hoyo de Manzanares	Modelo Lineal	Random Forest	Gradient Boosting	Red Neuronal	XGBoost
R2	0,81	0,80	0,84	0,81	0,82
MSE	63,74	99,92	70,22	56,65	86,49
RMSE	7,98	10,00	8,38	7,53	9,30
MAE	5,86	8,58	7,45	6,32	7,32

Tabla 6 Resultados ejecución del modelo en todas las zonas de estudio

En pasos posteriores, el servicio concluiría con elaborar cartografías diarias originadas en estos modelos predictivos.

17.2.1.7 Visualización de los resultados

No se ha realizado ninguna visualización específica para esta fase del proyecto. Sin embargo, se hace una propuesta para visualizar el tipo de riesgo en función de la humedad obtenida utilizando un modelo de clustering. Se ha utilizado la técnica de clustering no supervisada llamada self organizing maps, que es un tipo de red neuronal. Se generan grupos de observaciones para luego detectar cuál de esos grupos presenta un mayor riesgo de incendio.

Como resultado en la zona 5 b de Robledo de Chavela se han obtenido los siguientes grupos:

- FMC: 0% a 83% - Riesgo Alto
- FMC: 84% a 109% - Riesgo Medio
- FMC: 110% a 140% - Riesgo Bajo

De los 123 días empleados para el modelo, 52 de ellos han caído en la zona de riesgo coincidentes con las épocas más secas del año.

Mas detalle se explica en el anexo 8 “Robledo de Chavela Clustering”.

17.2.2 Solución tecnológica: arquitectura técnica

Por último, presentamos la arquitectura técnica propuesta para el desarrollo del proyecto en función de lo transmitido en el modelo lógico:

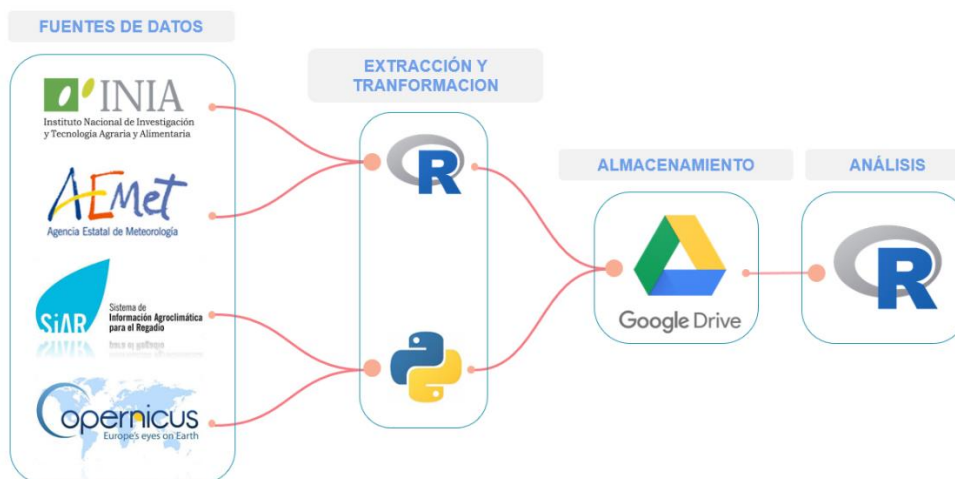


Ilustración 15 Arquitectura del proyecto

17.3 Análisis de recursos: talento humano y recursos físicos

17.3.1 Talento humano

El talento humano es uno de los factores influyentes a la hora de desarrollar correctamente este proyecto. Entre los recursos necesarios de Agresta, incluimos todos los perfiles necesarios para la viabilidad del proyecto, desde una fase inicial de diseño y planificación hasta la fase final de comercialización. No obstante, hay que puntualizar que no se anticipan o se proponen cambios estructurales como consecuencia de la implantación de este proyecto. A continuación, se describen los perfiles principales, así como las funciones que llevarán a cabo las actividades:

- Expertos en modelos de combustibles y comportamiento del fuego: Se trata de uno de los perfiles de mayor importancia en el proyecto. Se encargará de dar las pautas y el conocimiento que permita localizar la información más relevante para el modelo. De igual manera, guiarán al equipo técnico desde el inicio del proyecto con objeto de que se consiga el mayor aporte de valor posible con el producto desarrollado.

En este caso, los expertos en la materia son personal de Agresta, y estos tendrán un contacto directo con el equipo técnico que desarrolla y mantiene la herramienta. Su carga de trabajo será muy relevante en las primeras fases del proyecto ya que es necesario transmitir el conocimiento en el que se basa este proyecto.

- Expertos en procesos ETL: Se requerirá entrenar a técnicos con conocimiento en estos procesos, debido a que Agresta no dispone de ningún proceso para recoger y guardar la información necesaria para implantar la nueva herramienta. Actualmente, la empresa sí dispone de personal que realiza tareas relacionadas con programación y con conocimiento de diferentes herramientas con lo cual será necesario hacer un trabajo de transferencia para garantizar su formación en los nuevos procesos ETL. En primer lugar, se programarán los procesos para descargar y procesar un histórico de información que permita desarrollar el modelo, y posteriormente estos serán adaptados para la descarga de la información requerida con la periodicidad definida en función de las necesidades del cliente.

Las diferentes fuentes de datos a consultar vendrán definidas por los expertos en modelos de combustibles y comportamiento del fuego, quienes han determinado cuales son las variables más relevantes y su origen más fiable.

Se anticipa que estos técnicos tendrán una carga de trabajo inicial considerable ya que la obtención de la información, procesamiento y familiarización con la herramienta serán las tareas más relevantes y complejas del proyecto.

- Expertos en modelización. La principal tarea de este perfil será el análisis de todos los datos que se han podido obtener mediante los procesos ETL, con el objetivo de identificar las variables más relevantes que permitan elaborar un modelo de la máxima calidad posible. Este perfil será el encargado actualizar, entrenar y validar el modelo utilizado si se requiere modificar en un futuro y según los servicios que Agresta decida ofrecer.

Este perfil será cubierto por los miembros integrantes del grupo, ya que poseen conocimientos en diferentes lenguajes de programación y experiencia suficiente para realizar dicho análisis, el cual se apoya en los expertos en modelos de combustibles y comportamiento del fuego.

- Comerciales. Una vez el sistema esté desarrollado, se pondrá en un entorno de producción que empiece a generar valor. Este perfil se encargará de buscar clientes interesados en esta nueva herramienta y la información que proporciona ya sea por venta directa, participación en licitaciones o mediante la asociación con otras corporaciones. También se encargará de todo el proceso de promoción y comercialización, así como de la asistencia a eventos especializados que permitan obtener nuevos contactos.

En este caso, este perfil será cubierto por miembros de Agresta especialmente por el equipo del Área transversal de incendios forestales quienes capacitarán a los responsables de comercialización a nivel territorial.

- Administración y recursos humanos. Estos perfiles se encargarán de tareas secundarias pero esenciales para la optimización del valor añadido de esta herramienta como pueden ser la gestión del proyecto a nivel económico o la contratación y gestión del personal necesario. Este perfil será cubierto por personal de Agresta, en concreto por el equipo de Gestión Económica Financiera (GEFA).

En este caso, la carga de trabajo será menos significativa y equivalente a la de cualquier proyecto comercializado en Agresta, debido a que la mayor parte de las acciones serán cubiertas por el personal involucrado en el proyecto, y no se anticipa una gestión económica compleja.

17.3.2 Recursos físicos y servicios externos

Agresta cuenta con varias oficinas repartidas por el territorio nacional, sin embargo, para la ejecución del proyecto, no se ha realizado uso de ninguna de sus instalaciones, ya que cada persona integrante del equipo ha participado de manera remota mediante reuniones continuas con los responsables y técnicos partícipes. No obstante, no se descarta que, si en un futuro la herramienta

tiene aceptación y es implantada en diversos clientes, sea necesario crear un equipo de trabajo en alguna de sus sedes para dar el soporte y evolución requeridos.

Respecto a la infraestructura necesaria para la ejecución del proyecto, no ha sido requerido adquirir ningún tipo de Hardware o Software adicional. Todos los procesos se han realizado utilizando ordenadores personales de cada uno de los integrantes y software de uso libre para el desarrollo del sistema respetando la política de la empresa. Está previsto que el proyecto se pueda implementar perfectamente en la infraestructura de producción de Agresta ya que se ha confirmado el rendimiento de la infraestructura con ordenadores normales soportado por algún ordenador de alta capacidad para el procesado de datos

En el caso de una expansión rápida, es posible que se requiera de la compra de maquinaria capaz de ejecutar los procesos en un periodo de tiempo mejorable para mantener la calidad de nuestros servicios. El proyecto tiene posibilidades de ser escalado en caso de éxito a sistemas de Cloud Computing que aumentarían la capacidad de procesado. De igual manera, se entiende que los servicios de telecomunicaciones como puede ser el internet necesario para que los procesos ETL funcionen, ya ha sido contratado por Agresta y no se necesita un cambio de prestaciones. Sería necesario contabilizar servicios adicionales en la partida de gastos fijos si se considera en un futuro que se necesita ampliar la capacidad de la instalación si fuera insuficiente para ejecutar la herramienta de manera óptima.

El proyecto no ha requerido de la contratación de ningún tipo de servicio externo adicional a los ya contratados por Agresta. Actualmente en la Comunidad de Madrid, existe un convenio para la toma de muestras en campo y obtener la humedad del combustible vivo, por tanto, no se incurre en ningún gasto adicional. Sin embargo, y dado que la validación en campo es imprescindible para este servicio, el desarrollo de este en otras comunidades que no dispongan de datos de campo implicaría una inversión en toma de muestras, por lo que en dichos casos sería necesario contar con más recursos.

Por último, hay que indicar que los servicios open source a partir de los cuales el proyecto obtiene la información son todos de acceso libre y no suponen un gasto fijo.

En el caso de monetizar en algún momento dicha información, el proyecto necesitará pivotar para encontrar otras fuentes de información válidas y factibles o por el contrario contratar dichos servicios.

17.4 Gestión del tiempo

A continuación, se definen las fases y las tareas claves de cada parte del proyecto y la asignación temporal que hemos considerado más apropiada para su desarrollo:

Análisis y planificación: en esta fase inicial nos centramos en entender e interiorizar la necesidad del cliente para posteriormente ofrecer una solución diferencial que aporte valor. Las tareas que hemos definido incluyen:

- Estudio de los requerimientos del cliente. Se estudian las necesidades y características de nuestro cliente.
- Definir del problema. El equipo de proyecto junto con varios miembros clave del negocio es capaz de definir con exactitud el problema.
- Identificar hipótesis y validación. Dentro de las posibles hipótesis se elaboran criterios de prioridad y se validan con el cliente.
- Elaborar el plan de acción. Una vez bien definido el problema y validado las hipótesis, se elabora el plan de acción que nos permita mover nuestro proyecto en fases.
- Planificar el proyecto. Se correlacionan las tareas con medidas temporales de manera lógica.

Diseño. Se definen todos los procesos a llevar a cabo según el plan de acción definido. Cuenta con las siguientes tareas:

- Análisis de las fuentes de datos. Se explora y se define cómo se realizará el análisis y evaluación de las fuentes de datos disponibles, seleccionando aquellas que mejor nos ayuden a alcanzar los objetivos del proyecto.
- Definir los procesos ETL. Se definen como serán los procesos ETL de cada una de las fuentes de datos incluyendo la selección de variables relevantes y su tratamiento.
- Definir la arquitectura necesaria para la implantación del sistema.

Implantación y validación. Se desarrollarán todos los procesos necesarios para el desarrollo de la herramienta y con la conformidad de la empresa. Se podrán diferenciar las siguientes tareas:

- Ejecución de los procesos ETL. Se llevan a cabo los pasos ETL según la definición y pasos necesarios para la extracción, transformación y carga de los datos obtenidos para el modelo.

- Desarrollar y validar el modelo. El equipo del proyecto desarrolla el código necesario para plantear los modelos apropiados y valida el modelo con la empresa. En esta fase se itera varias veces volviendo al paso anterior las veces necesarias para modificar los procesos ETL tanto como sea necesario. En esta parte se incluyen los análisis descriptivo y predictivo.
- Pruebas integradas del sistema. Se realizan una serie de pruebas que validan el sistema y la arquitectura planteada funcionan correctamente. La empresa está de acuerdo con el funcionamiento de dicha herramienta tras una demostración que permita últimos cambios si fuesen oportunos.

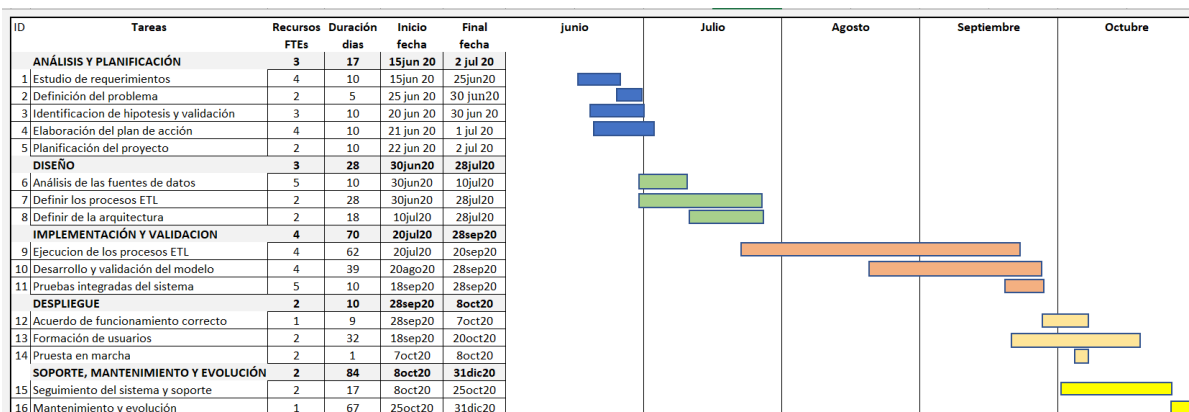
Despliegue. Se realiza el “go live” del sistema en los entornos de producción del cliente.

- Acordar con la empresa el correcto funcionamiento de la herramienta mediante la implantación del sistema. Se despliega el sistema desarrollado.
- Formación de los usuarios. La empresa forma a los técnicos y encargados de gestionar cada una de las actividades en diferentes fases, incluyendo varios técnicos durante las pruebas integradas del sistema y en fases posteriores.
- Puesta en marcha. Se realiza el “go live” de la herramienta.

Soporte, mantenimiento y evolución. Se realiza el seguimiento y soporte del sistema una vez implantado. Consta de las siguientes tareas:

- Seguimiento del sistema y soporte post implantación. El equipo de proyecto realizará una confirmación posterior a la puesta en marcha que verifique el correcto funcionamiento del sistema o aclare alguna duda con respecto al diseño o funcionamiento de la herramienta. La empresa a su vez quedará encargada de la recogida de las in
- Incidencias tras la implantación del sistema que determine si una posible actualización de la herramienta es necesaria en un futuro.
- Mantenimiento y evolución. Se realizan tareas de mantenimiento de la arquitectura y del software según las necesidades del cliente (esta fase queda fuera de este proyecto).

A continuación, se muestran las fases y tareas en una escala temporal (aproximadamente con un total de 5 meses de duración) con el tiempo planificado para cada actividad mencionada y los recursos empleados en cada una de ellas:


Ilustración 16 Diagrama de Gantt del proyecto

18 Rentabilidad del proyecto

Realizaremos un análisis de rentabilidad del proyecto, que a su vez se ha compartido con la empresa y responsables de negocio. Con ello se ha determinado el mejor enfoque posible en lo que supone y supondrá para la empresa a nivel de ingresos y costes según los objetivos marcados. En la actualidad el producto se considera que se encuentra en fase I+D+i y su viabilidad se ha realizado para una comercialización que podría comenzar en unos 4 meses.

El lanzamiento del producto se hará en forma de servicio ofrecido por Agresta al cliente final. Los resultados del proyecto se pueden ofrecer mediante diferentes vías que harán variar el valor del servicio, como pueden ser desde los reportes periódicos (cartografías de riesgo diarias) a herramientas BI para tratamiento y visualización de los modelos de combustibles.


Ilustración 17 Rentabilidad del proyecto

18.1 Beneficios tangibles

18.1.1 Generación de ingresos

Dentro de la generación de los beneficios esperados y calculados para Agresta, se encuentran aquellos que provienen de nuevos contratos. La empresa va a contar con un nuevo producto que por un lado tiene venta directa y por otro complementa los servicios que viene ofertando en el campo de simulación y prevención de incendios forestales. Por tanto, se espera que la implementación del nuevo producto conlleve la aplicación de un coste adicional en sus servicios que permita aumentar su margen de beneficio. Esto puede conducir a una reducida tasa de abandono de clientes debido a la mejora de la oferta de los servicios.

Poner en conocimiento de la comunidad científica y técnica en este campo los avances con aplicaciones tecnológicas tiene buena acogida y sin duda dará visibilidad a la empresa. El beneficio directo se hará con la captación de nuevos clientes a la vez de situarse como empresa de referencia en sistemas innovadores de teledetección.

Tras los primeros proyectos de uso de la herramienta, se podrá recoger la satisfacción del cliente y con ello optimizar cualquier parte necesaria en función de los requerimientos cambiantes del mercado.

18.1.2 Reducción de costes

Una vez realizada la comercialización del producto se puede estimar un impacto directo en ahorro con respecto a los otros productos que se encuentran en la misma línea de desarrollo.

En primer lugar, el uso de fuentes de datos abiertos no conlleva un coste para la empresa. Por otro lado, se contempla usar máquinas virtuales en la nube para optimizar el rendimiento de los procesos y ahorro en tiempo. Esto supondría igualmente un ahorro en vez de usar equipos propios que requieren de una inversión y mantenimiento o la compra de imágenes satelitales procesadas a terceros.

Otra parte importante respecto a la reducción de costes son los que se han observado al comparar los procesos utilizados en modelos anteriores. Estos últimos proponen el uso de imágenes satelitales de alto coste a lo que se añade un tiempo prolongado para realizar sus lecturas. El nuevo producto propone una alternativa mucho más eficiente y con un coste inferior.

El cálculo estimado de estos costes lo realizamos por horas en el caso de personal y en los costes actuales y conocidos en el caso de los datos obtenidos a partir del procesado de imágenes y muestras tomadas en campo y procesadas posteriormente mediante técnicas especializadas.

En relación con el personal especializado en modelos de combustibles, la automatización de todo el proceso permitirá ganar tiempo para distribuirlo en otros proyectos. De igual manera, el aumento de clientes no supondrá un aumento del personal, ya que con los expertos con los que cuenta Agresta se podrán cubrir todos los nuevos requisitos gracias a la automatización y simplicidad tecnológica que les ofrece la herramienta, lo que, a la larga, supondrá una reducción de coste y de estructura, que permitirá a Agresta ser más competitiva con el resto de los proveedores en todos sus servicios disponibles.

18.2 Beneficios intangibles

Como beneficios intangibles, podemos destacar que el procesamiento de la información se realiza casi en tiempo real y de una manera automatizada, lo cual dará una ventaja sin precedentes a las administraciones o particulares que podrían tomar medidas preventivas acorde a estos datos. Agresta contará con una base de datos meteorológicos que podrá ser utilizada de manera cruzada por todos los demás proyectos pudiendo así optimizar los resultados obtenidos hasta ahora en el resto de los servicios.

Para facilitar el acceso de esta información a sus clientes, Agresta podrá ofrecer informes periódicos detallando la información más relevante sacada de los nuevos modelos predictivos y sugerir recomendaciones basándose en su experiencia profesional en este campo, lo que podrá generar nuevas oportunidades de negocio a nivel estratégico en la rama de la consultoría forestal.

18.3 Beneficios estratégicos

Sin duda, un avance en las tecnologías aplicadas al campo de prevención de incendios conlleva un mayor crecimiento de la empresa tanto por mayor productividad como por la mejora de calidad y precisión de los servicios que ofrece. Esto hace que las empresas confíen en el servicio, aumente su buena reputación en el sector y atraigan a más clientes en un futuro.

De igual manera, Agresta se consolida como una de las consultoras ambientales española que más esfuerzos realiza en iniciativas tecnológicas y de I+D+i, logrando así que sus proyectos se diferencian por tener un fuerte carácter creativo e innovador.

18.4 Liquidez, rentabilidad y riesgos

A continuación, resumimos los cálculos de ingresos, costes y flujo de caja según lo descrito en los apartados anteriores.

Para este propósito, se contempla que el producto está en una fase última de desarrollo, y que la fase de comercialización puede llevarse a cabo en los próximos meses. Todos los cálculos son estimaciones tanto de la venta del servicio a las administraciones, posibles concursos y subvenciones, así como posibles gastos para mejoras del producto.

En este proyecto se ha considerado una tasa de interés de un 10%. El criterio que se ha seguido para escoger esta cifra es debido a la peculiaridad del producto y por ser las administraciones públicas los principales clientes potenciales. Esto conlleva un menor riesgo de inversión.

Obtenemos un VAN de 26.460,56 € y un TIR de 41,95%.

La inversión inicial para este proyecto es mínima ya que la empresa está dotada con un equipo técnico, instalaciones y personal que pueden operar inmediatamente. Sin embargo, se ha tenido en cuenta los costes en horas de capacitación del personal de Agresta con objeto de aprender a manejar la herramienta y diseñar la mejor manera de lanzar el proyecto.

En cuanto al balance de flujo de caja se ve una recuperación a partir del 2021, año en el cual se espera obtener los primeros beneficios.

Dentro de los gastos se ha incorporado personal adicional al existente en 2020 como una previsión de apoyo para los nuevos proyectos y una renovación de cierto equipo HW ya que en 3 años se estima la necesidad de adquirir un mínimo de equipo para operar. A continuación, se muestra el detalle del análisis de rentabilidad.

ANÁLISIS RENTABILIDAD DEL PROYECTO

	AÑO 2020	AÑO 2021	AÑO 2022	AÑO 2023
INVERSIÓN				
Equipo HW/SW escalado		1.500,00 €	1.500,00 €	
Desarrollos Iniciales	26.000,00 €			
TOTAL, INVERSIONES	26.000,00 €	1.500,00 €	1.500,00 €	- €
INGRESOS/BENEFICIOS				
Proyecto CA Madrid		15.000,00 €		
Subvención i+D		3.000,00 €		
Proyecto Empresa Meteorológica 1		7.500,00 €	10.000,00 €	15.000,00 €
Proyecto CA Castilla La Mancha			15.000,00 €	

Proyecto Empresa Meteorológica 2		10.000,00 €	15.000,00 €	
Proyecto Empresa Meteorológica 3			10.000,00 €	
Proyecto CA Castilla León			15.000,00 €	
Proyecto CA Comunidad Valenciana			15.000,00 €	
TOTAL, INGRESOS/BENEFICIOS	0,00	25.500,00 €	35.000,00 €	70.000,00 €

GASTOS				
Formación Agresta	2.500,00 €	500,00 €	500,00 €	500,00 €
Actualizar web	500,00 €			500,00 €
Material promocional (video, revista)		1.000,00 €	500,00 €	
Tareas comerciales		3.000,00 €	2.000,00 €	1.500,00 €
Contratación temporal/tiempo parcial		3.000,00 €	3.000,00 €	3.000,00 €
Desarrollos Adicionales y mantenimiento		6.000,00 €	6.000,00 €	13.000,00 €
Administración		500,00 €	500,00 €	500,00 €
Servicios Cloud	2.000,00 €	2.000,00 €	3.000,00 €	4.000,00 €
TOTAL, GASTOS	5.000,00 €	16.000,00 €	15.500,00 €	23.000,00 €
PERIODO	AÑO 2020	AÑO 2021	AÑO 2022	AÑO 2023
FLUJO DE CAJA OPERATIVO	- 31.000,00 €	8.000,00 €	18.000,00 €	47.000,00 €
VALOR ACTUAL	- 31.000,00 €	7.272,73 €	14.876,03 €	35.311,80 €
ACUMULADO	- 31.000,00 €	- 23.727,27 €	- 8.851,24 €	26.460,56 €

VAN:	26.460,56
TIR:	41,95%

Tabla 7 Resumen económico

19 Conclusiones

La ejecución del proyecto ha permitido confirmar la viabilidad técnica de la solución gracias a la precisión obtenida con los modelos utilizados. Por lo tanto, se confirma como alternativa entre los sistemas de cálculo teórico de humedad de combustible vivo, posicionando este servicio de manera eficiente y precisa entre los productos en desarrollo o comercializados actualmente.

Es una solución competitiva económicamente debido a los bajos costes recurrentes, sobre todo los que se asocian a desarrollos adicionales que puedan requerir los diferentes clientes. Gracias a esto, se espera que el servicio tenga una gran difusión entre las administraciones públicas debido a sus características competitivas e innovadoras.

El proyecto también cuenta con una variedad de desarrollos potenciales que permitan mejorar sus servicios y características. Entre éstas, se pueden destacar, por ejemplo, el desarrollo de un sistema BI para la muestra interactiva de los resultados, la extrapolación de la modelización a

otros tipos de combustibles vivos presentes en las áreas geográficas estudiadas y la posibilidad de modelización en entornos forestales mixtos.

20 Referencias

20.1 Bibliografías

- Aguado, I., y Camia, A. (1998). Fundamentos Y Utilización De Índices Meteorológicos De Peligro De Incendio. Serie Geográfica, 7, 49-58.
- Brown, Arthur Allen. & Davis, Kenneth P. (1973). Forest fire: control and use. New York: McGraw-Hill
- Castellnou, M.; Miralles, M.; Molina, D.; 2005. Análisis de la disponibilidad de combustible: índices meteorológicos críticos para la ocurrencia de cada patrón de grandes incendios forestales en la zona de Tivissa. IV Congreso Forestal Español. Zaragoza, (septiembre 2017)
- Castellnou, Marc & Pagés, J. & Miralles, Marta & Pique, Miriam. (2009). Tipificación de los incendios forestales de Cataluña. Elaboración del mapa de incendios de diseño como herramienta para la gestión forestal. '5° Congreso Forestal Español', 21-25 September 2009, Avila, Spain. 1-16.
- C. Chandler, P. Cheney, P. Thomas, L. Trabaud, and D. Williams, Fire in Forestry, vol. 1, New York: John Wiley & Sons, 1983.
- Chuvieso, E.; Martín, M. del P.; 2004. Nuevas tecnologías para la estimación del riesgo de incendios forestales. Consejo Superior de Investigaciones Científicas, Madrid.
- Hardy, C. C. (2005). Wildland fire hazard and risk: Problems, definitions, and context. Forest Ecology and Management, 211(1-2), 73-82.
- Marino, E.; Yebra, M.; Guillén-Climent, M.; Algeet, N.; Tomé, J.L.; Madrigal, J.; Guijarro, M.; Hernando, C. Investigating Live Fuel Moisture Content Estimation in Fire-Prone Shrubland from Remote Sensing Using Empirical Modelling and RTM Simulations. Remote Sens. 2020, 12, 2251.
- Marino E, Hernando C, Planelles R, Madrigal J, Guijarro M, Sebastián A (2014) Forest fuel management for wildfire prevention in Spain: a quantitative SWOT analysis. International Journal of Wildland Fire 23(3):373-384
- Nolan, R. H., Boer, M. M., Resco De Dios, V., Caccamo, G., y Bradstock, R. A. (2016). Large-scale, dynamic transformations in fuel moisture drive wildfire activity across southeastern Australia. Geophysical Research Letters, 43(9), 4229-4238.
- Pyne, Stephen J.; Andrews, Patricia L.; Laven, Richard D. 1996. Introduction to wildland fire. 2nd edition. New York, NY: John Wiley and Sons, Inc. 769 p.
- Scott, A. C., Bowman, D. M. J. S., Bond, W. J., Pyne, S. J., y Alexander, M. E. (2014). Fire on Earth: an introduction. John Wiley-Blackwell.

- Vélez, R. (1995). El peligro de incendios forestales derivado de la sequía. Cuadernos de La Sociedad Española de Ciencias Forestales, (2), 11.

20.2 Webgrafías

- AEMET. 2018. AEMET Open Data. Sistema para la difusión y reutilización de la información de la Agencia Estatal de Meteorología. <https://opendata.aemet.es/centrodedescargas/inicio>
- AEMET. 2020. AEMET Predicción. Interpretación del servicio de predicción de incendios. <http://www.aemet.es/es/eltiempo/prediccion/incendios/ayuda>
- Agresta. 2020. Agresta Innovación. Proyectos sobre Incendios Forestales. <https://agresta.org/proyectos/proyectos-incendios-forestales/>
- Arbonaut. 2020. Forest Fire Management. Forest Fire Risk Assessment for informative decisions. https://www.arbonaut.com/files/pdf/Forest_Fire_Management_en.pdf
- Bee2FireDetection. 2020. Bee2FireDetection Solutions. Early fire detection and decision support system. <http://www.bee2firedetection.com/#solution>
- Copernicus. 2018. Soil Moisture. Soil moisture gridded data from 1978 to present. <https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-soil-moisture?tab=overview>
- Elperiodic. 2020. Predecir incendios forestales mediante un sistema de satélites. https://www.elperiodic.com/predecir-incendios-forestales-mediante-sistemasatelites_663197
- Epdata. 2020. Situación de los bosques en España y en el mundo. <https://www.epdata.es/datos/situacion-bosques-mundo-espana-datos-graficos/330>
- Fundación Aquae. 2020. Proceso de desertificación. <https://www.fundacionaquae.org/conoces-que-es-la-desertificacion-sus-causas-y-sus-soluciones/>
- Gobierno de España. 2019. Orientaciones estratégicas para la gestión de incendios forestales en España. https://www.mapa.gob.es/es/desarrollo-rural/temas/politica-forestal/orient_estrategicas_gestion_iiff-2019_tcm30-523469.pdf
- Instituto Universitario de Investigación, Gestión Forestal Sostenible. 2014. La gestión de los combustibles forestales: oportunidades para la prevención de incendios en España. <http://sostenible.palencia.uva.es/content/la-gestion-de-los-combustibles-forestales-oportunidades-para-la-prevencion-de-incendios-en>
- Meteogrid. 2020. Meteogrid Servicios. Incendios Forestales. <https://www.meteogrid.com/servicios/incendios-forestales/>

- Meteosim. 2020. Meteosim Riesgo de Incendios. Sistema operativo de pronóstico de riesgo de incendios. <https://www.meteosim.com/es/gestion-de-riesgos-incendios/>
- Naciones Unidas. 2019. Población. Perspectivas de la población mundial. <https://www.un.org/es/sections/issues-depth/population/index.html>
- Open webinars. 2019. Metodologías y herramientas. Como hacer web scraping con Selenium. <https://openwebinars.net/blog/como-hacer-web-scraping-con-selenium/>
- SIAR. 2020. Consulta de Datos SIAR. <http://portal.miteco.gob.es/websiar/SeleccionParametrosMap.aspx?dst=1>
- Tecnosylva. 2020. Tecnosylva Risk Atlas. Incendios Forestales. <https://tecnosylva.es/>
- Unidata. 2020. Detailed examples for atmospheric science and meteorology. <https://unidata.github.io/python-training/gallery/gallery-home/>
- Youtube. 2020. Extracting time series data from a netCDF file into a CSV.
- https://www.youtube.com/watch?v=hrm5RmsVXo0&list=PLLxyyob7YmEE8S3QDs1PZQkiBxA4zn_Gx&index=3
- Youtube. 2016. Unidata Python Workshop - Monday June 20 part 1 - git, conda. <https://www.youtube.com/watch?v=8x5oe-OVpoM&list=PLnXN3uoWXJ29QlHdAeCXQfIx0n2yMUiRx>

21 Apéndices

21.1 Entrevistas: validación de hipótesis y obtención de información adicional

- Hipótesis y validación de alta criticidad e incertidumbre:
 - **Hipótesis:** Aumentar la precisión y operatividad de la predicción del riesgo para la prevención de incendios mejorando el indicador de la humedad de la planta utilizado actualmente, calculado a través de mediciones sobre el terreno. A través de fuentes de datos externas, permitirá predecir este indicador de humedad de la planta de una manera óptima estima que se mejorará en un 8% el modelo de predicción mediante muestras usado actualmente.

Discusión y validación: esta hipótesis se confirma como objetivo de mejora. Además, se verifica que la incorporación de nuevas variables al modelo existente puede conducir a esta mejora deseada. Durante la entrevista se aclaran las descripciones de las variables disponibles tanto en el modelo actual como las disponibles en otras fuentes de datos externas. Se incorporarán nuevas variables con el objetivo de evaluar la correlación que existe con la humedad del combustible vivo (variable objetivo de este proyecto) como por ejemplo las precipitaciones en las zonas monitorizadas.

- Hipótesis y validación de mediana criticidad e incertidumbre
 - **Hipótesis:** adaptar el modelo para incrementar la parte de negocio captando nuevas administraciones públicas durante el primer año después del lanzamiento.

Discusión y validación: Se confirma que los clientes interesados y que usan esta herramienta son administraciones públicas en la mayor parte de los casos. Con una mejora en la precisión del modelo se espera que aumente el interés y negocio con las administraciones.

- **Hipótesis:** reducir costes en un 25% al eliminar parcialmente la recogida de muestras en campo.

Discusión y validación: esta hipótesis no se podrá confirmar hasta que no se aplique la nueva herramienta. Durante la discusión se confirma que el laboratorio de la INIA seguiría recogiendo muestras en campo, aunque se desconoce si Agresta requerirá estos datos en un futuro (o con qué frecuencia). Se confirma que la

disminución del muestreo en campo supondría una reducción significativa de costes.

- Hipótesis y validación de baja criticidad e incertidumbre
 - Hipótesis: Las CCAA y Administraciones públicas expresan una preocupación (si/no) sobre el riesgo de incendios e impacto medioambiental.

Discusión y validación: se confirma esta hipótesis ya que existen diferentes programas de actuación sobre prevención y educación social sobre medioambiente e incendios. Empresas como Agresta siguen recibiendo solicitudes de gestión y mejoras para la prevención de incendios, así como entender mejor el comportamiento de los incendios para gestionarlos de manera más eficiente.

- Hipótesis: el nuevo modelo permitirá conocer la evolución de la desertificación en su territorio con una mejora de un 5% de precisión.

Discusión y validación: se confirma que la automatización en la recogida de datos meteorológicos y la mejora en la predicción del riesgo de incendios forestales permitirá estudios posteriores que ayudarán a conocer la evolución climática de zonas de interés con una mayor precisión, siendo los cambios meteorológicos e incendios las causas que más impacto tienen en la pérdida de la cubierta forestal.

- Hipótesis: Aumentar el flujo de notificaciones / publicaciones de avisos de incendio en un 20%.

Discusión y validación: Esta hipótesis no se podrá confirmar ya que variará en función del uso del producto que realicen los diferentes clientes, pese a todo, gracias a la mejora en la precisión y en la predicción, se espera ser capaces de aumentar tanto la cantidad como la calidad de notificaciones que emita el sistema a partir del modelo utilizado.

Por último, durante esta fase de validación y entrevistas, se han aclarado varios puntos importantes que nos ayudarán en el desarrollo del proyecto sobre todo en las fases de planificación y comprensión de los datos.

- Dentro de los datos y herramientas que la empresa utiliza para sus modelos, se encuentran imágenes satelitales (MODIS y SENTINEL2). De manera genérica, su modelo se basa en el procesado de estas imágenes conjuntamente con muestras de humedad de combustible vivo, recogidas en campo por empresas que prestan servicios de extinción de incendios a

la Comunidad de Madrid y procesadas por el laboratorio del INIA. Siendo la Comunidad de Madrid propietaria de los datos.

- A partir del procesamiento de dichas imágenes se obtienen mapas que se presentan por QGIS. Las imágenes son procesadas esporádicamente debido al alto coste que conlleva. La información de campo recogida por el INIA se comparte periódicamente vía archivo Excel cuyo volumen de muestras varían en función de la época del año (mayor frecuencia de muestreo en verano que en invierno).

No contamos con datos históricos, únicamente se utilizarán los datos de humedad de la planta (porcentaje) medido en laboratorio y en determinadas fechas de estudio (fechas de recogida de muestra en campo).

- Referente a las variables, el objetivo será estimar la humedad del combustible vivo. Dicha humedad, se presenta en % y será utilizada por las comunidades para notificar los riesgos de incendios. Estos datos son los resultados de la humedad de especies recogidas en campo desde el 2014 en diferentes zonas y épocas del año.

Además, se utilizarán variables vía API de agencias como la AEMET, SIAR o Copernicus. Estos datos servirán para modificar el modelo actual. Entre las variables disponibles, la humedad relativa del aire será una variable más, pero no se considera de gran importancia.

También se podrán crear nuevas variables mediante cálculos sencillos a partir de las variables disponibles.

Se ofrece una descripción más detallada en la sección 9. Análisis preliminar de los datos.

Se ha llevado a cabo un acuerdo de confidencialidad entre las partes para el uso de datos del INIA exclusivamente para este proyecto.

- Respecto a las herramientas que usaremos, no existe ninguna restricción o limitación que pueda suponer un problema para el desarrollo del proyecto. Los programadores de la empresa realizan códigos tanto en R o Python. Tampoco se anticipa que nuestro grupo de trabajo tenga que acceder a los servidores de la empresa y se confirma que será suficiente usar G-Drive para intercambiar información en carpetas compartidas.

21.2 SIAR: Evapotranspiración, proceso y parámetros influyentes

Las métricas disponibles en SIAR son las siguientes:

- Temperatura
 - Temperatura media(°C)
 - Temperatura máxima(°C)
 - Temperatura mínima(°C)
- Humedad
 - Humedad media (%)
 - Humedad máxima (%)
 - Humedad mínima (%)
- Velocidad del viento
 - Velocidad del viento(m/s)
 - Velocidad del viento máxima(m/s)
- Dirección del viento
 - Dirección del viento(°)
 - Dirección del viento máxima(°)
- Precipitación
 - Precipitación(mm)
 - Precipitación efectiva(mm)
- Radiación (MJ/m²)
- Temperatura del suelo(°C)
 - Temperatura del suelo 10 cm(°C)
 - Temperatura del suelo 30 cm(°C)
- Evapotranspiración(mm)

Con respecto a la evapotranspiración desde la plataforma SIAR enuncian las siguientes consideraciones que se deben resaltar:

Se conoce como evapotranspiración (ET) la combinación de dos procesos:

- Evaporación desde el suelo y desde la superficie cubierta por las plantas.
- Transpiración desde las hojas de las plantas.

Por lo tanto, el cálculo de la ET se usa para saber el agua que necesitan las plantas para su correcto desarrollo, ya sea en un jardín o en un campo de cultivo. Como consecuencia de esto, para poder diseñar todo lo relacionado con el riego, programación de riegos, cantidades de agua a aportar a un determinado cultivo, diseño y elección del mejor sistema de riego, conducciones de agua, etc., es necesario calcularlo de la manera más precisa posible.

Para que se produzca ET tienen que darse las siguientes condiciones:

- 1.- El agua tiene que estar presente.
- 2.- Tiene que haber alguna fuente de energía que convierte el agua líquida en vapor de agua.
- 3.- Se tiene que producir un fenómeno físico que separa el vapor de agua de la superficie de evaporación.

Las plantas para su crecimiento necesitan agua y otros nutrientes que toman del suelo a través de las raíces, dióxido de carbono (CO₂) que toman del aire a través de sus hojas y una fuente de energía que es la luz del Sol.

Las dos fuentes principales de agua que usan las plantas son la lluvia y el riego. En ambos casos, parte del agua que llega al suelo o las plantas queda retenido en forma de gotas sobre las hojas o sobre la superficie del suelo. Ésta es la parte de agua más fácilmente evaporable. El resto del agua se infiltra en el suelo de manera que las plantas pueden obtenerla por medio de sus raíces.

Para obtener el CO₂ de la atmósfera, las plantas abren sus estomas, poros microscópicos en la superficie de las hojas. Durante este proceso las plantas pierden agua hacia la atmósfera (transpiración).

Una vez que están disponibles para las plantas el agua y el CO₂, en las hojas se produce la fotosíntesis mediante la cual las plantas sintetizan glucosa (su alimento) a partir de agua, CO₂ y luz solar.

El agua líquida que se evapora desde la superficie foliar y desde la superficie del suelo requiere gran cantidad de energía, alrededor de 540 calorías por gramo de agua a 100 °C. Esta energía proviene de la luz del Sol en forma de energía solar.

El vapor de agua se produce en la superficie de contacto entre agua y aire produciendo un gradiente de presión de vapor enorme alrededor de dicha superficie. La dispersión del vapor de agua en la atmósfera se debe a dicho gradiente, y también debido a la influencia del viento u otros mecanismos. La evapotranspiración es continua en tanto en cuanto se mantienen las tres condiciones antes comentadas.

La evapotranspiración, o ET, se ve afectada por múltiples factores

Climatológicos	Características del suelo	Factores vegetales
Radiación solar	Textura	Tipo de plantación
Humedad relativa	Estructura	Profundidad de la raíz
Temperatura	Densidad	Densidad foliar
Velocidad del viento...	Composición Química...	Altura de las plantas Estado de crecimiento...

Aunque la ET se puede medir utilizando aparatos de medida como los lisímetros, el proceso de medición es largo y costoso, por lo que es mucho más frecuente su estimación en base a ecuaciones empíricas y analíticas.

La mayoría de las ecuaciones de ET se desarrollaron relacionando medidas de ET con medidas de parámetros climáticos que directa o indirectamente afectan a la ET.

Como hay tantos factores implicados en la ET, es extremadamente difícil formular una ecuación que permita el cálculo de ET en diferentes condiciones, es por ello por lo que los investigadores desarrollaron la idea de evapotranspiración para un cultivo de referencia expresada en milímetros (mm).

El cultivo de referencia es una superficie de gramíneas, cuyas características físicas y biológicas se han estudiado extensamente. La ET de esta superficie es conocida comúnmente como ET.

Para calcular la evapotranspiración real (ET) para un determinado cultivo en una zona con características climáticas semejantes a las de la estación meteorológica, se utiliza un factor de cultivo, conocido comúnmente como coeficiente de cultivo (K_c). Esta cantidad de agua requerida para compensar la pérdida por evapotranspiración del cultivo se define como necesidades de agua del cultivo.

Hay distintas ecuaciones que permiten el cálculo de la ET. Su elección depende del objetivo, condiciones y disponibilidad de datos.

Para superficies de referencia cuyas condiciones físicas y biológicas sean conocidas, los principales factores climáticos que afectan a la ET, son:

- Radiación solar
- Humedad relativa
- Presión de vapor
- Temperatura del aire
- Velocidad del viento

Por lo tanto, ET puede ser calculada con bastante precisión usando modelos matemáticos consistentes en ecuaciones complejas. Las ecuaciones empleadas para el cálculo de la evapotranspiración se pueden ver en la sección de datos calculados.