



2022
Lleida

27 · 1
junio · juny
julio · juliol

Cataluña
Catalunya

8º CONGRESO FORESTAL ESPAÑOL

La **Ciencia forestal** y su contribución a
los **Objetivos de Desarrollo Sostenible**

8CFE

Edita: Sociedad Española de Ciencias Forestales

Cataluña | Catalunya · 27 junio | juny - 1 julio | juliol 2022

ISBN 978-84-941695-6-4

© Sociedad Española de Ciencias Forestales



Organiza

SISREP: Sistema de Gestión de la Información del Programa de Forestación de Tierras Agrarias de Castilla y León

JOVELLAR LACAMBRA, L.C.¹, GALLEGO CIPRÉS, F.², DÍAZ-FERNÁNDEZ, S.³, RIVERA PASTOR, R.⁵, SÁNCHEZ MARTÍN, A.M.⁴, GÓMEZ CONEJO, R.² y VILLADA ROJO, D.³.

¹ Junta de Castilla y León. Servicio Territorial de Medio Ambiente. Salamanca. jovlaclu@jcyL.es

² Fundación CeseFor

³ Junta de Castilla y León. Dirección General de Patrimonio Natural y Política Forestal. Servicio de Promoción Forestal

⁴ Junta de Castilla y León. Dirección General de Patrimonio Natural y Política Forestal. Servicio de Defensa del Medio Natural

⁵ Quanticae Yewson

Resumen

Los controles de campo llevados a cabo sobre las forestaciones subvencionadas en el marco del Programa de Forestación de Tierras Agrarias en Castilla y León desde el año 1993 han generado una ingente cantidad de información. Para optimizar el complejo proceso asociado a estos controles y explotar la información obtenida, el Servicio de Promoción Forestal de la Junta de Castilla y León ha desarrollado un sistema integral de gestión.

SISREP es un sistema de información que digitaliza la recogida de datos en campo con dispositivos móviles, centraliza la información y proporciona herramientas avanzadas de consulta, actualización y explotación. Para ello, utiliza los datos de los controles sobre el terreno del Programa (1993-presente) y orienta sobre la probabilidad de éxito de futuras plantaciones utilizando técnicas de minería de datos y aprendizaje automático.

El sistema está compuesto por un gestor web de controles de campo, una aplicación SO Android para la recogida y transmisión de la información al servidor central, una base de datos espacial y herramientas de explotación e información: visor de la base de datos para consulta y panel para el uso de modelos predictivos.

Palabras clave

Forestación. Plantación forestal. Bases de datos. Inteligencia Artificial. Modelos predictivos. Aprendizaje automático. Minería de datos. Repoblación.

1. Introducción

La implantación del Programa de Forestación de Tierras Agrarias (en adelante Programa) en el año 1993 al amparo del Reglamento (CEE) N° 2080/92 supuso el comienzo de una etapa de singular relevancia en la historia de la repoblación forestal en España, especialmente en la forestación de terrenos privados. Castilla y León es la región de España con mayor superficie forestada con cargo a este Programa, con un 33 % del total de la superficie plantada en España hasta el año 2019 (VADELL *et al.*, 2019).

Desde su implantación y hasta finales de 2020 se habían forestado en la región acogidas a este Programa más de 200.000 ha, tanto en terrenos privados como de titularidad pública (JUNTA DE CASTILLA Y LEÓN, varios años).

Los retos que supuso en su día el lanzamiento del Programa y su prolongación en el tiempo han ido en paralelo con el desarrollo de múltiples aspectos técnicos, administrativos y de gestión. El empleo de nuevos métodos de preparación del terreno; la aplicación de nuevas tecnologías a los procesos de concesión, certificación y control o el fomento del asociacionismo forestal son sólo algunas de las innovaciones que han ido afectando al Programa o que él mismo ha favorecido. Estos cambios han sido tratados recientemente y con más detalle por IGLESIAS *et al.* (2021). Uno de los

aspectos más novedosos para el ámbito forestal consecuencia directa del Programa ha sido la ingente cantidad de información de campo generada a lo largo de estos últimos 29 años.

Además de las diferentes herramientas para la gestión administrativa del propio Programa, desde su inicio se ha generado información de tipología diversa: cartográfica, técnica (métodos, especies, procedencias) o administrativa (revocaciones, finalización de expedientes, tipo de propiedad). Estos datos se encontraban en ubicaciones y formatos diversos y cambiantes con los años (cartografía en soporte papel o digital, informes y actas independientes, hojas de cálculo, etc.).

La necesidad de unificar la información cartográfica y técnico-administrativa del Programa llevó a la Junta de Castilla y León en el año 2015 a promover un ambicioso proyecto para disponer de una base de cartografía digital dinámica asociada a los trabajos efectuados desde 1993. Esta cartografía, denominada “Capa de Forestación”, proporcionaría a través de una geodatabase única todas las especificaciones técnicas y parte de las administrativas de las obras certificadas en el Programa desde su comienzo así como los principales avatares históricos de las mismas. En la actualidad la geodatabase cuenta con más de 126.000 observaciones y es objeto de continuas actualizaciones.

Por otra parte, el control de los trabajos y del estado de las plantaciones para la certificación y pago de las diferentes primas (implantación, mantenimiento de la forestación y prima compensatoria) ha venido condicionado por estrictos protocolos tanto de la propia Junta de Castilla y León como de distintos organismos supervisores. La toma de datos de campo derivada de estos controles ha originado, y lo sigue haciendo, una profusa cantidad de información acerca del estado selvícola de las plantaciones y su evolución. Hasta el momento se dispone de más de 40.400 puntos de observación en toda la Comunidad Autónoma y esta cifra crece cada año con nuevas visitas de inspección.

El estado actual de las ciencias computacionales permite el desarrollo de técnicas y herramientas que transforman grandes cantidades de datos en información útil. Esto puede redundar en la mejora de la toma de decisiones técnicas, administrativas o de gestión así como en la evaluación de los programas llevados a cabo por la Administración. Lo anterior resulta fundamental para cumplir los objetivos de la Administración Pública en cuanto a la eficiencia, eficacia y transparencia que le son propias (UBALDI *et al.*, 2019) (LEY 39/2015, DEL PROCEDIMIENTO ADMINISTRATIVO COMÚN). Así mismo este análisis de la información constituye el pilar básico de la denominada gestión adaptativa y la subsiguiente mejora de la técnica forestal (RIST *et al.*, 2013).

En el caso concreto del Programa y a la luz de los datos obtenidos cabría preguntarse, por ejemplo: ¿cómo se han llevado a cabo las forestaciones durante estos casi 30 años?; ¿cuál ha sido su evolución en una zona determinada?; ¿cuál, de entre las especies utilizadas, es la más idónea para una estación concreta?; ¿cuál es la probabilidad de éxito que puede esperarse de una determinada plantación?; ¿dónde es más necesario enfatizar los controles de campo? En definitiva, cuestiones todas encaminadas a optimizar los recursos de la Administración Pública y a la mejora de la gestión forestal.

Por todo lo anterior, en 2018 se plantea desde el Servicio de Promoción Forestal de la entonces Dirección General de Medio Natural de la Junta de Castilla y León la creación de una herramienta que: (1) facilitara la toma de datos de campo, (2) permitiera el acceso flexible a toda la información disponible, (3) informara sobre la adecuación de especies y métodos utilizados y (4) orientara acerca del grado de viabilidad de plantaciones futuras. El Sistema de Gestión de la Información del Programa de Forestación de Tierras Agrarias, SISREP 1.0 pretende ser la primera versión de esta herramienta.

2. Objetivos

Los objetivos para el desarrollo de SISREP 1.0, el Sistema de Gestión de la Información del Programa de Forestación de Tierras Agrarias desarrollado por la Junta de Castilla y León que se presenta en esta comunicación, han sido:

- Diseñar y poner en funcionamiento una base de datos centralizada, georreferenciada, accesible y permanentemente actualizada con la información obtenida en las inspecciones de campo así como la correspondiente a todas las forestaciones del Programa.

- Diseñar e implementar un sistema automatizado de captura digitalizada de los datos recogidos en las inspecciones de campo, para agilizar la tramitación de los informes asociados y para alimentar la base de datos.
- Desarrollar un sistema de modelización predictiva para obtener el apoyo en la toma de decisiones técnicas (elección de especies y métodos) y para predecir el éxito de futuras plantaciones. Todo ello sobre la base de la información almacenada en el sistema.
- Desarrollar un visor que permita: (1) la consulta visual y alfanumérica de la información recogida así como su descarga para análisis específicos y (2) el acceso intuitivo y rápido a los resultados de los modelos predictivos desarrollados.

Estos objetivos se articulan sobre la base de una arquitectura completa del sistema escalable y flexible para que resulte fácilmente ampliable y adaptable a nuevos usos.

3. Metodología

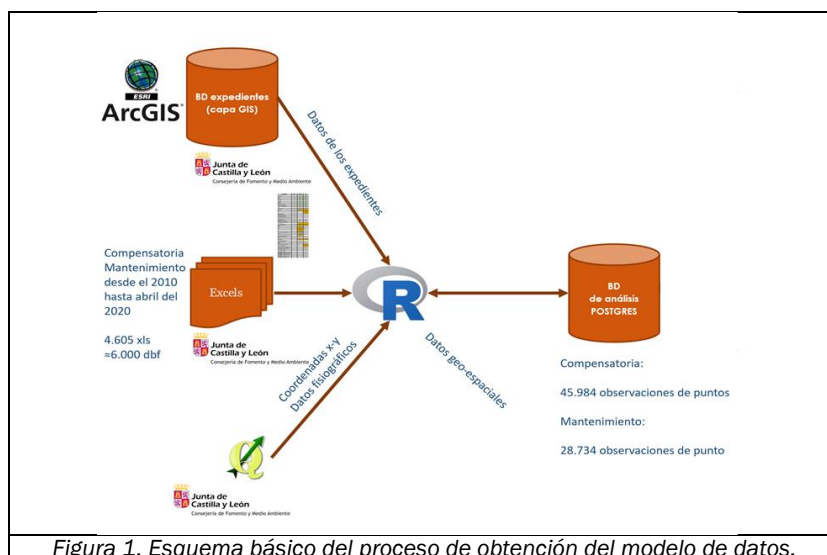
El ámbito territorial de este trabajo lo constituye la Comunidad Autónoma de Castilla y León, en concreto, las plantaciones efectuadas en el marco del Programa. En el estado actual de la aplicación, se contemplan únicamente las plantaciones en terrenos de propiedad privada.

En la primera fase del trabajo se llevó a cabo un análisis de la información existente en el Servicio de Promoción Forestal de la Junta de Castilla y León, estudiándose las diferentes fuentes de datos disponibles y la calidad de éstos. Para ello, se realizó un análisis de los estadillos resultantes de las visitas de campo de los controles, extrayendo la información relevante de supervivencia recogida en los mismos. A cada punto de observación se le añadió posteriormente la información técnico-administrativa recogida en el momento de la certificación de la forestación, especialmente centrada en las condiciones iniciales de la misma: Zona del Programa (JUNTA DE CASTILLA Y LEÓN, 2020), densidad inicial, especies, fecha de certificación, método de preparación del terreno, datos de localización, etc.). Seguidamente se obtuvieron a partir de diversas fuentes cartográficas aquellas variables del medio que, a priori, podían estar relacionadas con la supervivencia en una zona determinada: altitud, orientación, pendiente, posición fisiográfica, pH del suelo, así como varias relacionadas con la climatología. Un listado de las principales variables empleadas se esquematiza a continuación:

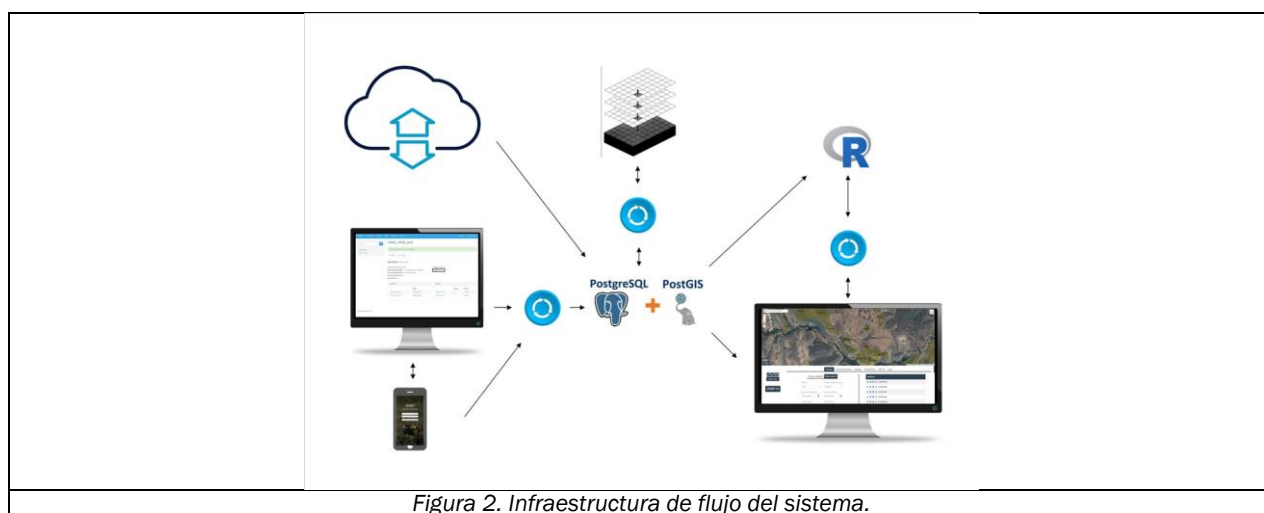
- Información de las parcelas de muestreo de las visitas de campo para los controles de los expedientes del Programa en plantaciones de propiedad privada. Las parcelas de muestreo son generalmente circulares (200 m²). Su disposición es sistemática (malla cuadrangular) de origen aleatorio. De los datos que se toman en la parcela se utilizaron los siguientes:
 - Expediente.
 - Provincia.
 - Rodal en el que se ubica el punto de control.
 - Ubicación del centro de la parcela (coordenadas UTM).
 - Densidad de planta en adecuado estado vegetativo obtenida mediante conteo en la parcela y extrapolada a pies/ha.
 - Fecha de la observación.
- Características de las plantaciones en el momento de su certificación. Estos datos se obtuvieron de la Geodatabase “Capa de Forestación” implementada sobre ArcGis (REDLANDS, CA: ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE, 2011) y fueron:
 - Zona del Programa de Forestación.
 - Fecha de certificación.
 - Especies empleadas y porcentajes (agrupadas en categorías según masas puras o mezclas).
 - Método de preparación del terreno.
 - Densidad inicial de la plantación (pies/ha).
- Datos del medio obtenidos de cartografía temática.

- Altitud (msm). A partir del modelo digital del terreno 1ª cobertura con paso de malla de 5 m (IDECYL, 2020).
- Orientación. Categorizada en 5 clases (N, S, E, O, ATV) a partir del modelo digital del terreno (IDECYL, 2020).
- Pendiente (%). A partir del modelo digital del terreno (IDECYL, 2020).
- Litología. Obtenida a partir del Mapa Litoestratigráfico de España a escala 1:200.000 (IGME, 2015).
- Posición fisiográfica. A partir del mapa de formas del terreno (mapa fisiográfico) elaborado a su vez a partir del índice de posición fisiográfica (TPI) para la Comunidad de Castilla y León según la metodología propuesta por WEISS (2001) y adaptada a las condiciones particulares de escala requerida y relieve de la zona. La resolución espacial es de 5 m.
- pH del suelo. Categorizado a partir de los mapas de propiedades químicas de la capa superior del suelo elaborados a partir de la base de datos LUCAS 2009/2012 con una resolución espacial de 500 m (BALLABIO et al., 2019).
- CRAD. Capacidad de retención de agua del horizonte superior del suelo expresada como fracción volumétrica respecto al volumen de suelo. Sus valores se han recodificado a partir de la cartografía a escala europea que se elaboró con una resolución espacial de 500 m según la base de datos LUCAS (BALLABIO et al., 2016).
- Temperatura media anual (°C). A partir del Atlas Agroclimático de Castilla y León (NAFRÍA et al., 2013).
- Índice de aridez. A partir del Atlas Agroclimático de Castilla y León (NAFRÍA et al., 2013).
- Precipitación de verano. A partir del Atlas Agroclimático de Castilla y León (NAFRÍA et al., 2013).
- Variables derivadas.
 - Edad de la plantación en el momento de la observación: Calculada como diferencia entre la fecha de la observación y la de certificación), expresada en días.
 - Porcentaje de supervivencia. Porcentaje de la densidad de planta viva en adecuado estado vegetativo en el momento de la observación respecto a la densidad inicial en el momento de la certificación.
- Otras variables.
Además, en la base de datos se incluyen otras variables climáticas y edáficas obtenidas de cartografía temática oficial y que pueden resultar de utilidad para análisis específicos.

La inclusión de toda la información en una base de datos centralizada vino precedida de un trabajo detallado de lectura y limpieza de los datos, especialmente de los resultados de visitas recogidos en 4.605 hojas de cálculo, pero también de la información contenida en los expedientes administrativos. El procesado e integración de los datos de la geodatabase, estadillos de campo y datos del medio se programaron utilizando el lenguaje R (R CORE TEAM, 2020). Gracias a este trabajo de análisis e integración se pudo establecer el modelo final de datos e introducirlo en la base definitiva (figura 1).



Este trabajo puso de relieve la necesidad de crear una estructura que hiciera posible un flujo de información continuo en el tiempo, digitalizado y automático desde la captura de datos hasta su utilización mediante las herramientas de explotación de la información. Este flujo actualizaría a su vez los modelos predictivos. Para ello se diseñó la siguiente infraestructura (figura 2).

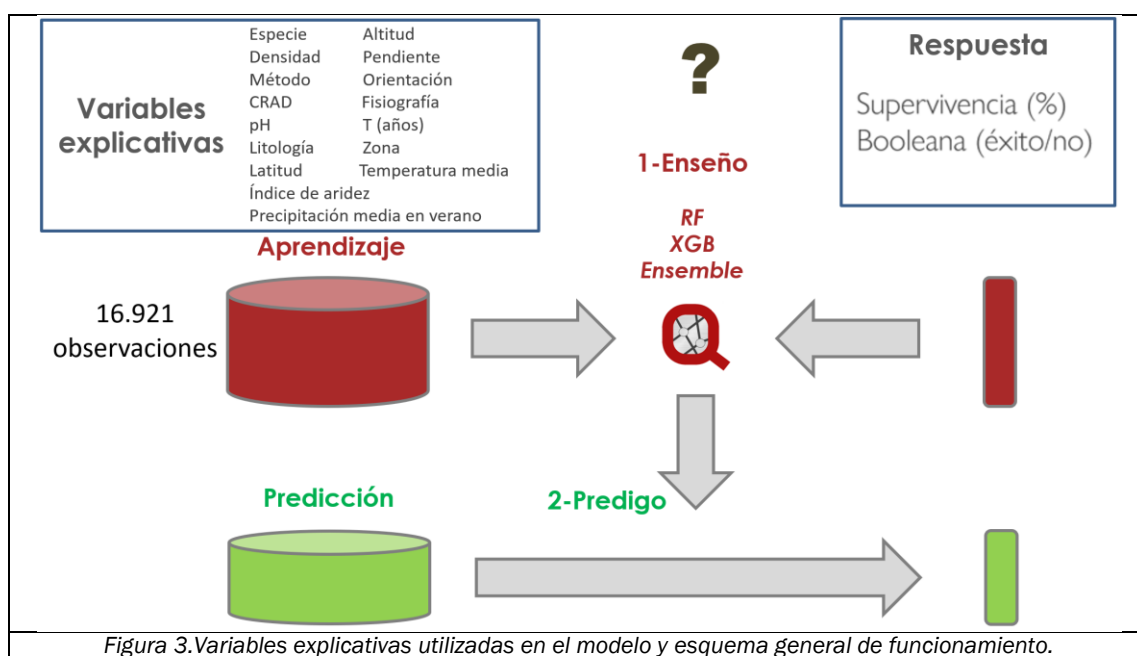


- **Base de datos.** Para el diseño de la base de datos se optó por PostgreSQL con la extensión espacial PostGIS (THE POSTGRESQL GLOBAL DEVELOPMENT GROUP, 2021) por ser un gestor de base de datos relacional de código abierto que soporta objetos geográficos.

Captura de datos. Se desarrolló una solución de software consistente en un gestor web de controles de campo y una aplicación para dispositivos móviles (Android) para la recogida de la información sobre el terreno. Para el desarrollo de la APP se utilizaron librerías basadas en Node.js, un entorno de tiempo de ejecución de JavaScript (MOZILLA DEVELOPER NETWORK, 2021) como lenguaje de programación básico y también Android nativo como lenguaje de programación para la comunicación de los sensores del dispositivo móvil con la APP móvil. También se ha implementado un servicio web desde la aplicación de gestión, que centraliza todo el proceso, para que la APP pueda comunicarse con ella.

Modelos predictivos. El objetivo de los modelos predictivos es determinar: (1) el porcentaje estimado de supervivencia de una especie en un punto del terreno y (2) la probabilidad de que dicho porcentaje supere un umbral prefijado (entre el 60% y el 90%). La muestra de partida para la elaboración de los modelos la constituyen las observaciones de campo de los controles oficiales de Prima Compensatoria (y por tanto de propiedad privada) efectuados desde el año 2007 y sobre forestaciones con más de 6 años de antigüedad. Esta restricción se justifica con el fin de asegurar la consolidación de la plantación una vez finalizados los trabajos de reposición de marras de los primeros años. De igual forma se han excluido observaciones con métodos de plantación y categorías de especies con menos de 50 observaciones. Tampoco se incluyeron aquellos puntos muestreados en rodales con mezclas de varias especies (3 o más) sin ninguna dominante o rodales no suficientemente definidos en la capa de forestación en cuanto a especies o métodos. Se han excluido también observaciones con densidades iniciales inferiores a 200 pies/ha u observaciones con datos incompletos, erróneos o incoherentes. Tras la aplicación de estos filtros la muestra útil para los modelos asciende actualmente a 16.921 observaciones, cifra que se va incrementando a medida que se incorporan nuevos datos desde la aplicación Android. Los modelos predictivos desarrollados han sido dos: uno de regresión y otro probabilístico empleando en ambos casos el ensemble de un Random Forest (LIAW & WIENER, 2002) y un eXtreme Gradient Boosting (CHEN *et al.*, 2021) utilizando el programa R (R CORE TEAM, 2020). Se han elegido ambos modelos por su contrastada calidad. El ensemble de ambos modelos se hace mediante una regresión en el caso del modelo de regresión y mediante la media en el caso del modelo probabilístico. Por lo que respecta a las variables explicativas, de la serie inicial y tras un análisis exploratorio previo, se consideraron aquellas que minimizaban los errores del modelo. Para ello se tuvo en consideración un procedimiento de eliminación recursiva de variables mediante Random Forest y Bootstrapping así como información previa acerca de posibles variables influyentes. Se empleó el paquete Caret (KUHNS, 2008) del programa R (R CORE TEAM, 2020).

En la figura 3 se relacionan las variables explicativas finalmente utilizadas y se muestra el esquema general del proceso.



Visor. Para el desarrollo de la aplicación de visualización, acceso y exportación de los datos de SISREP se utilizó el lenguaje de programación PHP (PHP DOCUMENTATION GROUP, 2021) en el desarrollo web del lado del servidor, y HTML (WHATWG, 2013), CSS (CSS WORKING GROUP, 2021) y JavaScript (MOZILLA DEVELOPER NETWORK, 2021) del lado del cliente (interfaz de usuario). Para la representación cartográfica sobre mapas, se utilizó Leaflet, una librería incluida en JavaScript (MOZILLA DEVELOPER NETWORK, 2021) que permite la representación de datos espaciales.

4. Resultados

Base de datos.

La base almacena los datos recogidos en el terreno más los datos técnico-administrativos a través del gestor web de controles de campo. También permite la centralización de la información así como la automatización de geo-procesos mediante la creación de consultas espaciales automáticas de forma que se añade información espacial derivada de su localización (fisiografía, litología, topografía, suelo y climatología) a las observaciones que llegan del campo. Actualmente, el número de observaciones que constituyen la base de datos es de 24.660. La frecuencia de actualización de la información puede llegar a ser diaria y, en todo caso, puede programarse a criterio de los administradores del sistema. Esta base de datos constituye, con los filtros que se dirán, el pilar básico de la información que se utiliza en el modelo predictivo, que se aborda más adelante.

Captura de información.

La aplicación desarrollada permite generar la documentación necesaria y crear los estadillos que posteriormente se completarán en campo a la vez que se permite la selección del equipo de trabajo que debe realizar el control. Por otro lado, la aplicación recibe la información de los puntos a visitar (codificación SIGPAC y coordenadas UTM) y la contenida en el expediente (especies plantadas, densidad inicial, área del rodal, etc.) y permite completar el estadillo de acuerdo con la realidad de la parcela. Por último, la información se envía al servidor. Estos desarrollos permiten la sincronización y almacenamiento centralizado de los datos de campo. Otro objetivo específico que se cumple es la mejora en los procesos de los controles de campo en oficina, más concretamente:

- Generación de documentación administrativa de apoyo a los equipos de campo.
- Generación de estadillos digitales a utilizar en las visitas.
- Generación semiautomática del acta de parcelas de control.

Un ejemplo de las pantallas de entrada y el estadillo generado puede observarse en la figura 4.



Modelos predictivos.

Utilizando los modelos predictivos desarrollados, esta herramienta permite estimar el porcentaje de supervivencia esperado de una forestación (modelo de regresión) o la probabilidad de

éxito a partir de una supervivencia mínima fijada por el usuario (modelo booleano) para unas coordenadas especificadas. La finalidad es ayudar a seleccionar la especie o mezcla de especies y método de plantación idóneos. En la versión actual de la aplicación (1.0) se contemplan tres tipos de análisis (i) “caso *individual*”, que calcula la probabilidad de éxito o porcentaje esperado para una especie y método seleccionados según las variables introducidas, (ii) “caso *óptimo*”, que calcula el binomio especie-método con mayor probabilidad de éxito o mayor porcentaje de supervivencia para las variables introducidas derivadas de su localización y (iii) “*todos los casos*”, que calcula todos los pares especie-método con mayores probabilidades o porcentajes (figura 5). También se ofrece la evaluación del modelo, utilizándose para el modelo booleano la curva ROC, el AUC y la matriz de confusión con la precisión, sensibilidad y especificidad. Para el modelo de regresión se calcula el error cuadrático medio y el error absoluto.

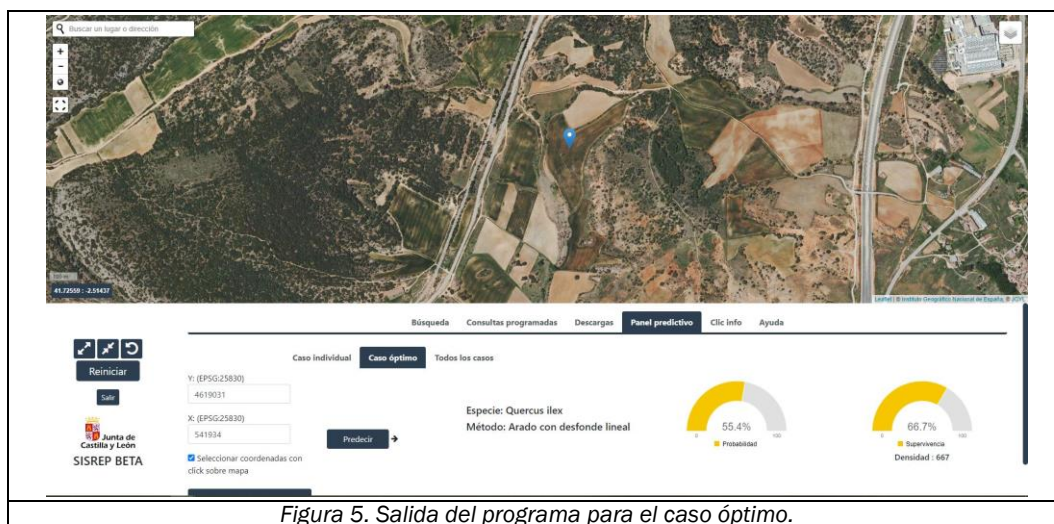
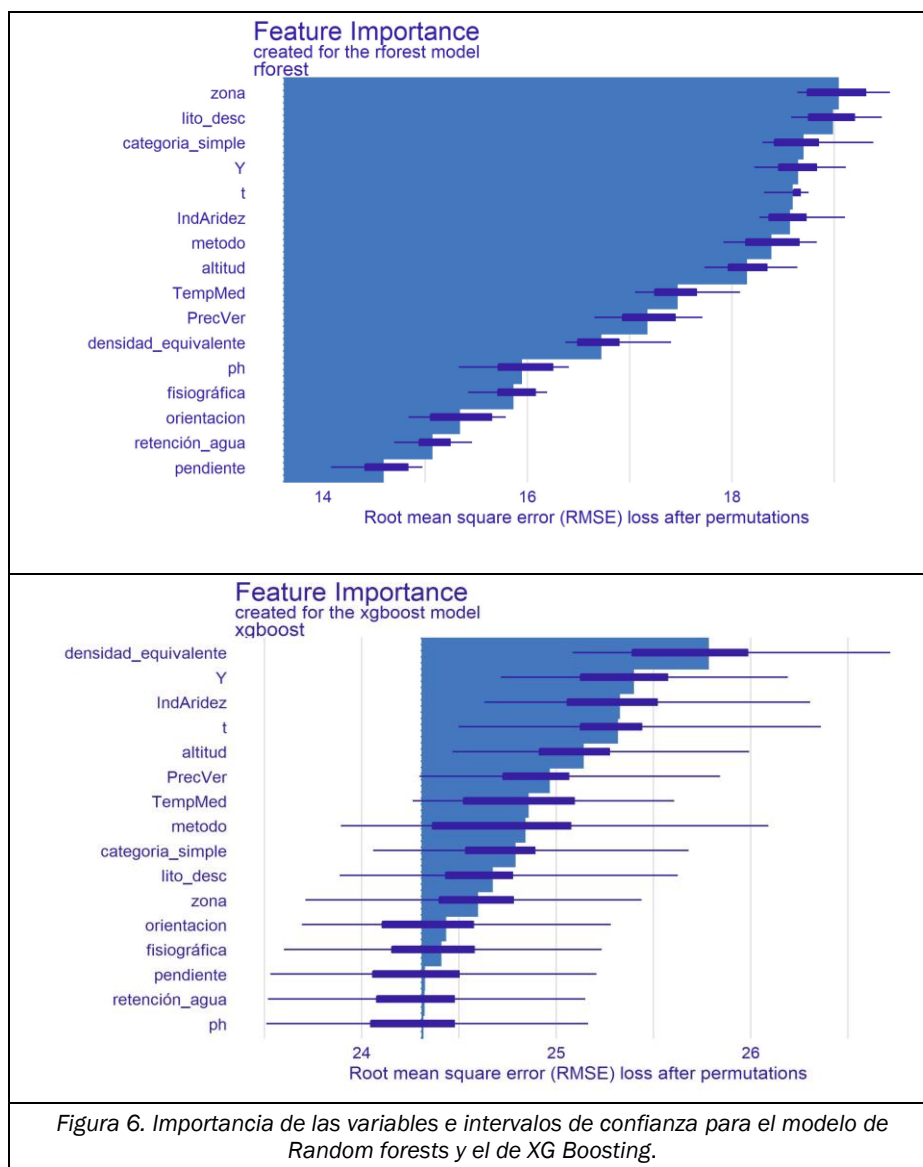
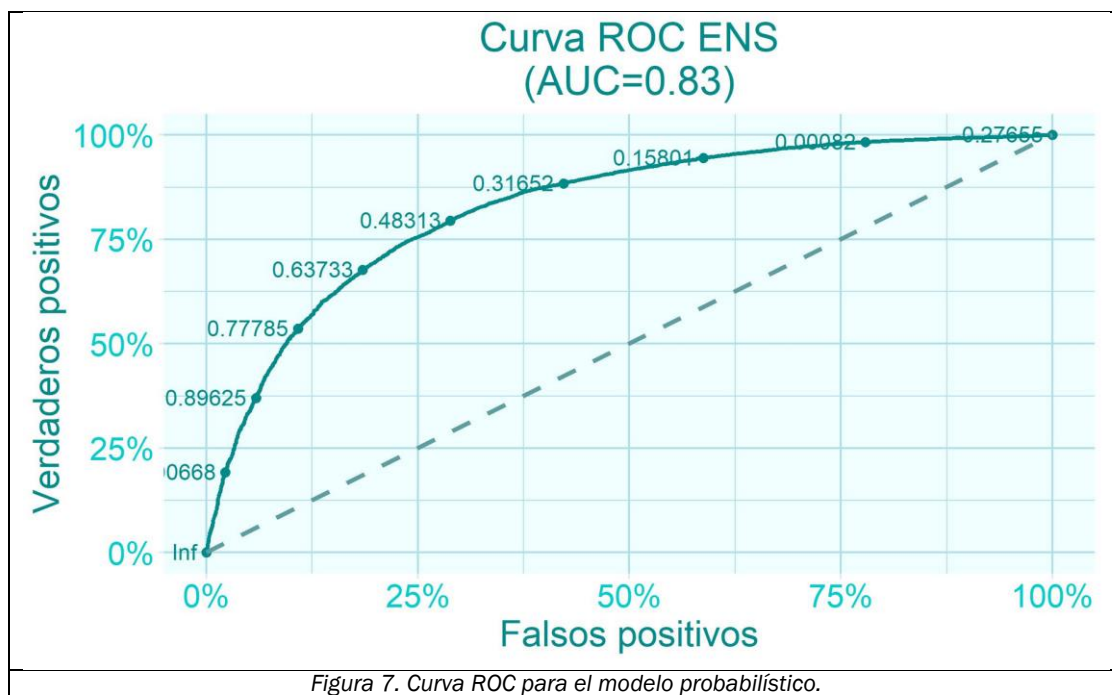


Figura 5. Salida del programa para el caso óptimo.

El programa presenta también para cada modelo, la importancia de las variables con sus intervalos de confianza calculados mediante Out of Bag Error. Se incluye un ejemplo de este análisis para ambos modelos: RF y XGB (figura 6).



Para la validación de los modelos se utilizó el método 10-fold CV, tanto para el modelo de regresión como para el modelo de probabilidad. En el primer caso se calcula la raíz del error cuadrático medio y en el segundo el AUC. El resultado obtenido para el modelo de regresión es de una RECM de 22.4 y un error absoluto medio de 16.9 comparados con 28.9 y 24.6 respectivamente del caso de asignar el valor medio y que consideramos el caso base. En el modelo probabilístico para un umbral del 70% obtenemos un AUC de 0.83 y la curva de ROC que se muestra en la figura 7.



Visor.

Constituye la interfaz con el usuario y la vía principal y más directa para la explotación de la toda la información del sistema de forma intuitiva y sencilla. Se trata de una aplicación de visualización, acceso y descarga que permite:

- Visualizar y localizar todos los puntos de control con la información de la inspección (figura 8), así como la geometría de las parcelas forestadas con la información técnico-administrativa del expediente. El visor cuenta con cartografía de apoyo de planimetría u ortofotos. También se añade información accesoria de Montes de Utilidad Pública, Oficina Virtual del Catastro y de SIGPAC.
- Acceder a cualquier expediente por codificación o mediante filtros de acuerdo con ciertos parámetros. El resultado también devuelve la información de las inspecciones realizadas (en caso de que las haya) y permite la descarga en formato “shp” tanto de la geometría de la forestación como de los puntos de control.
- Efectuar análisis descriptivos de los datos recogidos por especie. De acuerdo con diferentes filtros y consultas programadas se puede extraer información de porcentajes de supervivencia de la especie en cuestión en las condiciones especificadas.
- Acceder a los modelos predictivos.
- Descargar toda la información cartográfica y administrativa de los expedientes así como la información asociada de las inspecciones.

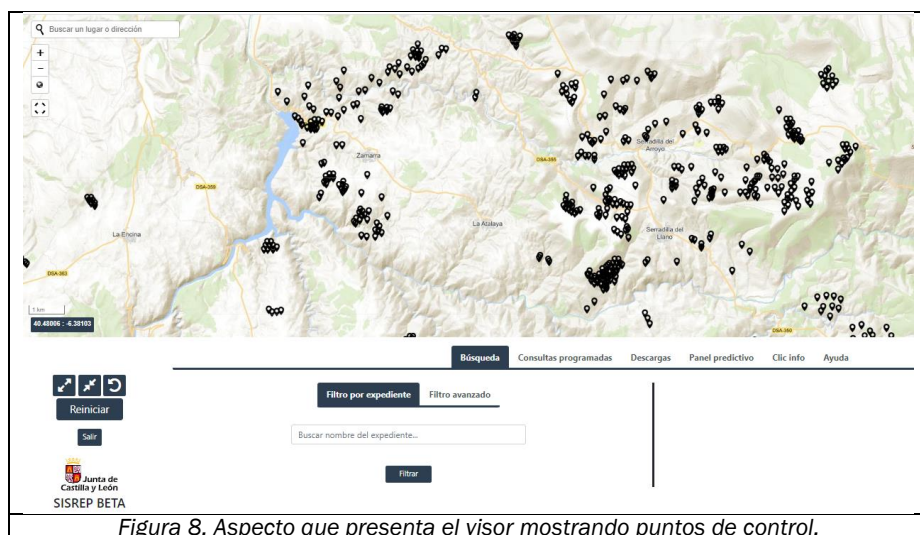


Figura 8. Aspecto que presenta el visor mostrando puntos de control.

5. Discusión

El sistema SISREP constituye, hasta donde sabemos, la primera herramienta integrada de detalle para el seguimiento y evaluación técnica de un programa de forestación a gran escala en España. Su novedad reside no sólo en el número de datos que maneja y su actualización continua sino en la integración de variables ambientales, el análisis de los resultados y su modelización. Otro aspecto destacable de la herramienta es su facilidad de explotación, mediante un visor intuitivo que permite consultas, análisis y descargas de forma sencilla y rápida.

Disponer de una base geoespacial estructurada, dinámica y fácilmente actualizable sobre un número tan amplio de forestaciones supone que el Programa de Forestación de Tierras Agrarias puede trascender ampliamente sus objetivos iniciales. Analizar las características de las plantaciones, sus resultados, su evolución o su relación con otras variables ecológicas y socioeconómicas es ahora más fácil, en línea con lo que plantean UBALDI *et al.* (2019) cuando estudian el gran valor que puede generar la gestión óptima de los datos en la Administración Pública.

El seguimiento y la evaluación de las plantaciones forestales constituyen aspectos trascendentales dentro de los planes y proyectos de forestación o reforestación, independientemente de su naturaleza y propósito (JACOBS *et al.*, 2015). De hecho, la mera evaluación del éxito de un programa de inversión pública tan importante como el Programa de Forestación de Tierras Agrarias de Castilla y León es ya relevante por sí misma. SISREP permite esta evaluación considerando no únicamente la superficie plantada, sino la persistencia en el tiempo de las masas establecidas. Esto debería ser un objetivo ineludible en la evaluación de los programas de forestación (TORRES 2001; JACOBS *et al.*, 2015; LI *et al.* 2017; ZHU *et al.* 2017).

Cada vez resulta más frecuente el uso de herramientas digitales en contextos de gestión forestal (NITOSLAWSKI *et al.*, 2021). Las herramientas integradas en SISREP agilizan la toma de datos de campo y el proceso posterior de elaboración de informes, lo que facilita el trabajo administrativo y mejora la eficiencia global en la tramitación del Programa. Además, facilitan la toma de decisiones técnicas gracias a los análisis descriptivos y predictivos que se muestran instantáneamente.

El desarrollo de herramientas de monitorización forestal basadas en imágenes de satélite o tomadas por vuelos no pilotados es hoy una tendencia creciente para reducir costes o facilitar la toma de decisiones en el ámbito forestal (véase por ejemplo: ZOU *et al.*, 2019; ALMEIDA *et al.*, 2019, 2020 o NITOSLAWSKI *et al.*, 2021). No obstante, muchas situaciones siguen necesitando el trabajo de campo. En algunos casos por la necesidad de validar los datos de los propios sensores remotos y en otros por las propias características de los sistemas a monitorizar. Es el caso de aquellas formaciones

en las que el tamaño de sus componentes, la capacidad de éstos para mimetizarse con su entorno o la precisión requerida, hacen ineludible la presencia de técnicos sobre el terreno. Esto ocurre, por ejemplo, en plantaciones jóvenes o en regenerados en competencia con la vegetación circundante, tan comunes en el ámbito mediterráneo (VILLAR-SALVADOR *et al.*, 2012). En estos casos, SISREP aporta otro valor añadido, puesto que la totalidad de las observaciones que constituyen su base de datos han sido obtenidas por técnicos especialistas sobre el terreno.

Recientemente, se han comenzado a tomar datos selvícolas básicos (clases de altura, diámetros y vegetación competidora) mediante captura directa en los controles de campo. Es de esperar que en breve se incorporen a la base de datos. Ello permitirá tener una idea más precisa del estado de la plantación y servirá para determinar la necesidad de tratamientos selvícolas en las masas jóvenes (gradeos, podas, clareos) según localizaciones, estaciones y especies así como para validar modelos selvícolas.

El uso de la inteligencia artificial es cada vez más frecuente en el ámbito forestal (véase por ejemplo: STOJANOVA *et al.*, 2010 ; VEGA & GARCÍA, 2017; SUN *et al.*, 2017; SUN & SCANLON, 2019; ZOU *et al.*, 2019). SISREP ha utilizado herramientas de minería de datos e inteligencia artificial para desarrollar los modelos predictivos que ayudan a tomar decisiones sobre la viabilidad de las diferentes especies y métodos a implantar, basándose en los datos históricos del Programa. Si bien es verdad que el sistema de caja negra de la inteligencia artificial impide ver las relaciones de causalidad, los indicadores de importancia de las variables incluidos en el panel predictivo aportan información que debe ser interpretada por técnicos o científicos. Esto resta opacidad al sistema. Es de esperar que, dentro de un rango, cuanto más representativo sea el volumen de datos recogidos en la base de datos, más realistas serán los resultados y la utilidad de los modelos predictivos. Por ello, la incorporación de datos de plantaciones en predios públicos acogidas al Programa, o aquellas llevadas a cabo por la propia Administración constituye una prioridad para las siguientes etapas del proyecto.

El panel predictivo de SISREP permite predecir los resultados para cualquier valor de las variables ambientales. Esto resulta de especial interés en el caso de las variables climáticas. En el actual contexto de incertidumbre climática, la posibilidad de crear diferentes escenarios o la integración de modelos climáticos en el propio SISREP puede ser crucial para la gestión de las plantaciones a largo plazo (REYER *et al.*, 2015).

El modelo que se presenta podría no contemplar actualmente variables de importancia para el logro de la plantación, lo que sin duda redundaría en errores de predicción. Es el caso, por ejemplo, de los múltiples factores relacionados con la calidad de los trabajos o con el mantenimiento de la plantación. A este respecto hay que tener en cuenta que los trabajos de implantación (preparación del suelo, plantación, etc.) fueron, en todos los casos, objeto de control y certificación por la Junta de Castilla y León, lo que acredita unos estándares mínimos de calidad. En el caso de los trabajos de mantenimiento (reposición de marras, eliminación de la vegetación competidora, etc.) éstos fueron también objeto de controles puntuales por parte de la Administración. Otras variables exógenas no consideradas pueden resultar definitivas para la supervivencia de la planta: presencia ilegal de ganado, daños por caza, incendios, etc. Respecto a este tipo de factores la ubicación de la plantación (determinada en la versión actual del programa por las variables “Latitud”, “Zona” y “Altitud”) puede, de forma implícita, paliar estas limitaciones u otras no consideradas y relacionadas con casuísticas de tipo local. Todo ello deberá ser objeto de análisis por los técnicos conocedores del terreno. En todo caso, como ya se indicó, los modelos tienen carácter predictivo y no analizan causalidad. En definitiva, es su capacidad predictiva final la que pretende ayudar en la toma de decisiones. Es de prever que en las siguientes fases del proyecto la incorporación de nuevos datos o la posibilidad de contrastar nuevas variables, mejoren el comportamiento de los modelos y reduzcan los errores de predicción.

Al hilo de todo lo anterior resulta fundamental interpretar las salidas de los modelos como una indicación que deberá ser objeto de valoración técnica en función de las circunstancias específicas del proyecto que se quiera llevar a cabo así como de otros condicionantes cualesquiera.

Implicaciones directas para la gestión y la investigación.

A modo de síntesis SISREP permite, entre otras posibilidades:

1. La captura de datos de nuevas plantaciones de forma normalizada y semiautomática y la agilización en la toma de decisiones acerca de la certificación de un expediente del Programa o de cualquier plantación cuya idoneidad tenga que ser determinada.
2. La evaluación del Programa de Forestación a escala regional, provincial o local.
3. El apoyo a la toma de decisiones acerca de especies, métodos y densidades a emplear en forestaciones basándose en resultados. Esto puede llevarse a cabo mediante dos vías: (1) a partir de los datos aportados directamente por la base de datos (estadística descriptiva) o (2) mediante el empleo de los modelos basados en inteligencia artificial.
4. Orientación acerca de las probabilidades de éxito de una plantación en una determinada estación, tanto en escenarios estables en el tiempo como variables.
5. El establecimiento de prioridades para la ejecución de intervenciones selvícolas según localizaciones, especies y estaciones.

Finalmente, la información contenida en SISREP puede también ser una herramienta valiosa para la investigación forestal. La cantidad y la calidad de los datos obtenidos y su facilidad de manejo abren un campo amplio para estudios de índole diversa.

Mejoras del sistema a corto y medio plazo

Entre los objetivos inmediatos para la siguiente versión de la herramienta se encuentran:

1. Incorporación de datos de control de campo en montes públicos a través de controles específicos.
2. Incorporación al sistema de datos de alturas y diámetros en los controles de campo.
3. Creación de consultas programadas por el usuario.
4. Incorporación de nuevas variables ambientales y socioeconómicas.
5. Mejora de los modelos predictivos mediante la consideración de nuevas variables y depuración de las técnicas estadísticas.

La posible ampliación del sistema a otras áreas de gestión como la silvicultura, los aprovechamientos o la gestión de flora y fauna multiplicarían las utilidades de la aplicación y su aportación a la gestión medioambiental.

6. Conclusiones

SISREP 1.0 constituye la primera versión de una herramienta valiosa para agilizar el trabajo de campo en el seguimiento y monitorización de plantaciones forestales. Aprovecha el conocimiento de lo que ha sido una parte muy importante de las plantaciones efectuadas en Castilla y León en los últimos 29 años, informa de cuáles están siendo sus resultados y contribuye a mejorar su gestión apoyando la toma de decisiones. Su ampliación mediante la incorporación de más observaciones en terrenos públicos, no incluidos hasta el momento, la propia mejora de los modelos predictivos y la posibilidad de extender su ámbito a otras facetas de la gestión forestal seguirá contribuyendo, pensamos que de manera sustancial, a la mejora del conocimiento y la gestión del medio natural en Castilla y León.

7. Bibliografía

ALMEIDA, D. R. A. de; STARK, S. C.; VALBUENA, R.; BROADBENT, E. N.; SILVA, T. S. F.; RESENDE, A. F. de; FERREIRA, M. P.; CARDIL, A.; SILVA, C. A.; AMAZONAS, N.; ZAMBRANO, A. M. A.; BRANCALION, P. H. S.; 2020. A new era in forest restoration monitoring. Restoration Ecology, 28, 1, 8–11. DOI: 10.1111/rec.13067.

ALMEIDA, D.R.A.; BROADBENT, E. N.; ZAMBRANO, A.M.A.; WILKINSON, B. E.; FERREIRA, M. E.; 2020. Monitoring the structure of forest restoration plantations with a drone. *International Journal of Applied Earth Observation*, 79, 5, 192-198. DOI: 10.1016/j.jag.2019.03.014.

ARNOLD, K.; GOSLING, J.; HOLMES, D. (2005). *The Java programming language*. Addison Wesley Professional.

BALLABIO, C.; LUGATO, E.; FERNÁNDEZ-UGALDE, O.; ORGIAZZI, A.; JONES, A.; BORRELLI, P.; MONTANARELLA, L.; PANAGOS, P.; 2019. Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. *Geoderma*, 355, 113912. DOI: 10.1016/j.geoderma.2019.113912.

BALLABIO, C.; PANAGOS, P.; MONATANARELLA, L.; 2016. Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, 261, 110–123. DOI: 10.1016/j.geoderma.2015.07.006.

CHAZDON, R.; MELI, P.; GORGENS, E. B.; SILVA, C. A.; STARK, S. C.; VALBUENA, R.; PAPA, D. A.; BRANCALION, P.H.S.; 2019. Monitoring the structure of forest restoration plantations with a drone-lidar system. *International Journal of Applied Earth Observation and Geoinformation*, 79, 05, 192–198. DOI: 10.1016/j.jag.2019.03.014.

CHEN, T.; HE, T.; BENESTY, M.; KHOTILOVICH, V.; CHO, H.; CHEN, K.; MITCHELL, R.; CANO, I.; 2021. Xgboost: Extreme Gradient Boosting. Rpackage. Versión 1.4.1.1.

CSS WORKING GROUP, 2021; CSS specifications. <https://www.w3.org/Style/CSS/>

IDECYL; 2020. Modelo Digital del Terreno de Castilla y León 2008-2011 a 5 m./pix. Infraestructura de Datos Espaciales de Castilla y León. <https://idecyl.jcyl.es/geoserver/el/wfs?request=GetCapabilities&service=WFS>.

IGLESIAS, A.; CEBALLOS, J.; JOVELLAR, L. C.; SÁNCHEZ, A.; 2021. El Programa de Forestación de Tierras Agrarias en Castilla y León: Desarrollo y avance de resultados. *Montes*, 143, 14–22.

IGME; 2015. Mapa litoestratigráfico de España 1:200.000. Ministerio de Medio Ambiente (Secretaría General para el Territorio y la Biodiversidad) y el Instituto Geológico y Minero de España.

En: http://mapas.igme.es/gis/services/Cartografia_Tematica/IGME_Litoestratigrafico_200/MapServer/WMSServer.

JACOBS, D. F.; OLIET, J. A.; ARONSON, J.; BOLTE, A.; BULLOCK, J. M.; DONOSO, P. J.; LANDHÄUSSER, S. M.; MADSEN, P.; PENG, S.; REY-BENAYAS, J. M.; WEBER, J. C.; 2015. Restoring forests: What constitutes success in the twenty-first century? *New Forests*, 46 (5-6), 601–614. <https://doi.org/10.1007/s11056-015-9513-5>

JUNTA DE CASTILLA Y LEÓN; Varios años. Anuario estadístico de Castilla y León. <https://estadistica.jcyl.es/web/es/estadisticas-temas/anuario-estadistico.html>.

JUNTA DE CASTILLA Y LEÓN; 2020. Cuadernos de zona para la realización de trabajos de forestación de tierras agrícolas. Actualización 2020. Dirección General de Patrimonio Natural y Política Forestal. En: <https://medioambiente.jcyl.es/web/es/medio-natural/actualizacion-2020-cuadernos-zona.html>.

KUHN, M.; 2008. Building predictive models in R using the caret package. Journal of statistical software, 28, 1, 1–26.

LEY 39/2015, DE 1 DE OCTUBRE, DEL PROCEDIMIENTO ADMINISTRATIVO COMÚN DE LAS ADMINISTRACIONES (02/10/2015). Jefatura del Estado. Ley 39/2015. BOE nº 236.

LIAW, A.; WIENER, M.; 2002. Classification and Regression by randomForest. R News, 2, 3, 18–22.

LI, L.; HAO, T.; CHI, T.; 2017. Evaluation on China's forestry resources efficiency based on big data. Journal of Cleaner Production, 142, 1, 513–523. DOI: 10.1016/j.jclepro.2016.02.078.

MOZILLA DEVELOPER NETWORK, 2021; JavaScript.
<https://developer.mozilla.org/es/docs/Web/JavaScript>

NAFRÍA, D.; GARRIDO, N.; ÁLVAREZ, M. V.; CUBERO, D.; FERNÁNDEZ, M.; VILLARINO, I.; GUTIÉRREZ, A.; ABIA, I.; 2013. Atlas agroclimático - Castilla y León. Editan: Ministerio de Agricultura, Alimentación y Medio Ambiente, Agencia Estatal de Meteorología; Junta de Castilla y León, Consejería de Agricultura y Ganadería, Instituto Tecnológico Agrario.

NITOSLAWSKI, S. A.; WONG-STEVENSON, K.; STEENBERG, J. W. N.; WITHERSPOON, K.; NESBITT, L.; VAN DEN KONIJNENDIJK BOSCH, C. C.; 2021. The Digital Forest: Mapping a Decade of Knowledge on Technological Applications for Forest Ecosystems. Earth's Future, 9, 8. DOI: 10.1029/2021EF002123.

PHP DOCUMENTATION GROUP; 2021. <https://www.php.net/>

R CORE TEAM; 2020. R: A language and environment for statistical computing. Viena, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

REGLAMENTO (CEE) Nº 2080/92 DEL CONSEJO, DE 30 DE JUNIO DE 1992, POR EL QUE SE ESTABLECE UN RÉGIMEN COMUNITARIO DE AYUDAS FORESTALES A LA AGRICULTURA (1992) Consejo de la Unión Europea.

RIST, L.; CAMPBELL, BRUST, M.; FROST, P.; 2013. Adaptive management: where are we now? Environmental Conservation, 40, 1, 5-18.

REYER, C. P.O.; RAMMIG, A.; BROUWERS, N.; LANGERWISCH, F.; 2015. Forest resilience, tipping points and global change processes. Journal of Ecology, 103, 1, 1–4. DOI: 10.1111/1365-2745.12342.

STOJANOVA, D.; PANOVA, P.; GJORGJIOSKI, V.; KOBLER, A.; DŽEROSKI, S.; 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. Ecological Informatics, 5, 4, 256–266. DOI: 10.1016/j.ecoinf.2010.03.004.

SUN, A. Y.; SCANLON, B. R.; 2019. How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. *Environmental Research Letters*, 14, 7, 73001. DOI: 10.1088/1748-9326/ab1b7d.

SUN, Y.; LIU, Y.; WANG, G.; ZHANG, H.; 2017. Deep Learning for Plant Identification in Natural Environment. *Computational intelligence and neuroscience*, 2017, 7361042. DOI: 10.1155/2017/7361042.

THE POSTGRESQL GLOBAL DEVELOPMENT GROUP; 2021. PostgreSQL 14.1 Documentation (14.1). <https://www.postgresql.org/files/documentation/>.

TORRES, J. M.; 2001. Evaluación de plantaciones forestales. México: Limusa-Noriega.

UBALDI, B.; VAN OOIEN, C.; WELBY, B.; 2019. A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance. *OECD Working Papers on Public Governance*, 33. <https://ideas.repec.org/p/oec/govaaa/33-en.html>.

VADELL, E.; MIGUEL, S. de; FERNÁNDEZ CENTENO, G.; ROBLA, E.; LERNER, M.; PEMÁN GARCÍA, J.; 2019. La forestación de tierras agrícolas: balance de un instrumento de política forestal para el cambio del uso de la tierra. *Cuadernos de la Sociedad Española de Ciencias Forestales*, 45, 2, 1–20. DOI: 10.31167/csecfv0i45.19497.

VEGA, C.; GARCÍA, R.; 2017 Big Data y Ordenación de Montes: ¿qué ha cambiado en realidad? *Cuadernos de la Sociedad Española de Ciencias Forestales*, 43, 31–40.

VILLAR-SALVADOR, P.; PUÉRTOLAS, J.; CUESTA, B.; PEÑUELAS, J. L.; USCOLA, M.; HEREDIA-GUERRERO, N.; REY BENAYAS, J. M.; 2012. Increase in size and nitrogen concentration enhances seedling survival in Mediterranean plantations. Insights from an ecophysiological conceptual model of plant survival. *New Forests*, 43, 5-6, 755–770. DOI: 10.1007/s11056-012-9328-6.

WHATWG (WEB HYPERTEXT APPLICATION TECHNOLOGY WORKING GROUP), 2013. HTML living standard. <https://html.spec.whatwg.org/multipage/>

WEISS, A.; 2001. Topographic position and landforms analysis. Poster presentation. ESRI Users Conference. San Diego, CA.

ZHU, J. J.; ZHENG, X.; WANG, G. G.; WU, B. F.; LIU, S. R.; YAN, C. Z.; LI, Y.; SUN, Y. R.; YAN, Q. L.; ZENG, Y.; LU, S. L.; LI, X. F.; SONG, L. N.; HU, Z. B.; YANG, K.; YAN, N. N.; LI, X. S.; GAO, T.; ZHANG, J. X.; ELLISON, A. M.; 2017. Assessment of the World Largest Afforestation Program: Success, Failure, and Future Directions (104). *BioRxiv*. DOI: <https://doi.org/10.1101/105619>. <https://www.biorxiv.org/content/10.1101/105619v1.full>.

ZOU, W.; JING, W.; CHEN, G.; LU, Y.; SONG, H.; 2019. A survey of Big Data analytics for smart forestry. *IEEE Access*, 7, 46621–46636. DOI: 10.1109/ACCESS.2019.2907999.