



Informe técnico de resultados, validación y protocolo de gestión operativa de la prueba de concepto (PoC) para un sistema de alerta temprana de riesgo de fallo mecánico del arbolado urbano mediante el análisis de series temporales de datos satelitales

Ámbito de aplicación: Jardines de El Buen Retiro (Madrid)

Cliente: Ayuntamiento de Madrid

Fecha: 08 de junio de 2026

Contacto Agforest: Guillermo de Santiago de la Bárcena

Correo electrónico: guillermo.desantiago@agforest.es



CONTENIDO

RESUMEN EJECUTIVO	3
LIMITACIONES ESTRUCTURALES DEL SISTEMA.....	4
CONTEXTO Y OBJETIVOS	5
DESAFÍO OPERATIVO: DESFASE ENTRE METEOROLOGÍA Y BIOFÍSICA.....	5
HIPÓTESIS DE TRABAJO: LA «FATIGA BIOLÓGICA» ACUMULADA	5
OBJETIVOS DE LA PRUEBA DE CONCEPTO (POC)	5
METODOLOGÍA	6
WP0. ADQUISICIÓN E INTEGRACIÓN DE DATOS	6
Datos meteorológicos	6
Datos de incidencias	6
Imágenes satelitales.....	6
WP1. LST SINTÉTICA A 10 M Y CONSTRUCCIÓN DE INDICADORES DE ESTRÉS	7
WP2. MODELADO PREDICTIVO: EL IDR.....	8
Estructura del <i>dataset</i> analítico	8
Las tres iteraciones comparadas del modelo de Agforest	8
Validación temporal estricta	9
Calibración de umbrales estacionales.....	9
WP3. PRODUCTOS OPERATIVOS: MAPA DE SUSCEPTIBILIDAD E IRC	10
Mapa de Susceptibilidad Estructural.....	10
Índice de Riesgo Combinado (IRC)	11
RESULTADOS.....	11
MAPA SUSCEPTIBILIDAD ESTRUCTURAL: DISTRIBUCIÓN ESPACIAL DE INCIDENCIAS.....	11
COMPARATIVA IDR VS SISTEMA DE ALERTAS DEL AYUNTAMIENTO	11
Serie temporal IDR vs incidencias reales (2022–2024)	11
Nota sobre la validez de la comparativa	12
Tabla de métricas completa.....	12
Interpretación métrica a métrica.....	12
ANÁLISIS DE SEVERIDAD	15
ÍNDICE DE RIESGO COMBINADO (IRC): PROTOCOLO DE ACCIÓN DE CAMPO.....	16
RECOMENDACIONES OPERATIVAS	16
Protocolo de umbrales estacionales	16
Integración del IRC en la gestión diaria de campo	17
LIMITACIONES METODOLÓGICAS	18
VARIABLES meteorológicas del modelo: diseño retrospectivo deliberado.....	18
Variable respuesta binaria sin ponderación por severidad	19
Mejoras metodológicas propuestas para la siguiente fase	20
ANEXO TÉCNICO	21
Auditoría de vuelcos y validación de la categoría de severidad alta	21
Desglose detallado de los 55 días de severidad alta.....	22
Validación del Baseline M0 — Viento observado vs Predicción AEMET	22

RESUMEN EJECUTIVO

Los Jardines de El Buen Retiro deben hacer frente al desafío estructural que suponen las constantes incidencias causadas por la caída de ramas y el vuelco de árboles, entre otros. Si bien el estado del arte alude a una amplia diversidad de factores que podrían ser detonantes del riesgo ante la vulnerabilidad de los ejemplares arbóreos, el viento es el único factor con evidencia científica suficiente para sostener un protocolo de gestión. Consecuencia de ello es el sistema municipal de activación de cierres preventivos que el Ayuntamiento de Madrid emplea en el recinto. Este protocolo, conocido como «Protocolo Retiro», activa cierres preventivos del parque basándose, fundamentalmente, en la previsión de racha de viento máxima emitida por AEMET. Esta metodología presenta dos limitaciones estructurales:

- (i) es evidente la magnificación de los eventos meteorológicos que la Institución realiza tras los fenómenos climáticos que han protagonizado los últimos años, reduciendo el grado de fiabilidad de las predicciones; y
- (ii) omite el estado fisiológico previo del arbolado, una métrica que se construye mediante el análisis de series temporales de estrés térmico e hídrico acumulado.

Esta hipótesis puede ilustrarse de la siguiente forma: **aunque dos ejemplares estén expuestos a la misma racha de viento, el evento no desencadenará consecuencias análogas en un ejemplar sano que en uno sometido a 30 días de calor extremo y un déficit hídrico claro.** La vulnerabilidad estructural es radicalmente diferente; sin embargo, el riesgo se evalúa en base al mismo criterio, el viento. Esta premisa pone de relieve un fallo sistémico del actual modelo de gestión. **Los indicadores empleados a lo largo del presente trabajo derivan del uso de modelos de inteligencia artificial (IA)**, imágenes satelitales y otras metodologías analíticas propietarias de Agforest, cuyo acceso es posible gracias al progreso tecnológico de los últimos años.

La presente Prueba de Concepto (en adelante, PoC) demuestra que dicha limitación es cuantificable, superable y que su corrección tiene consecuencias directas sobre la seguridad ciudadana y la eficiencia en la gestión del parque. Para desarrollar este sistema, que busca la escalabilidad, se han considerado tres iteraciones de complejidad creciente del modelo de Agforest, todas entrenadas sobre el mismo período histórico (2022–2024) y validadas sobre datos independientes (enero–mayo 2025):

- (i) una primera versión con viento observado como *benchmark* de referencia;
- (ii) una segunda versión lineal que incorpora indicadores de estrés acumulado; y
- (iii) la versión refinada no lineal seleccionada como definitiva, que combina todas las variables anteriores con índices espectrales satelitales.

Tras refinar el modelo de **Agforest**, se obtiene el **Índice Diario de Riesgo (IDR), que evalúa diariamente la predisposición estructural del arbolado al fallo mecánico, determinando cuánta presión acumulada soporta el parque y en qué medida es sensible ante cualquier detonante externo.** Cabe poner de relieve que esta solución no reemplaza los sistemas de alerta meteorológica existentes, sino que los complementa con el estado fisiológico previo del arbolado que, como se observa a lo largo del presente informe, constituye una información determinante para la toma de decisiones. Actualmente, el Ayuntamiento de Madrid sabe cuándo hay viento; el IDR cuenta con la capacidad de determinar si los ejemplares arbóreos expuestos a las rachas máximas previstas están en condiciones de soportarlo, garantizando la seguridad de la ciudadanía.

El IDR combina tres fuentes de información para estimar cada día la probabilidad de que se registre al menos una incidencia en el parque: (i) variables meteorológicas de AEMET — temperatura, viento, precipitación y humedad —; (ii) indicadores de estrés acumulado — grados-día acumulados en distintas ventanas temporales, déficit hídrico estimado por método agronómico estándar, días secos consecutivos y un Índice Compuesto de Vulnerabilidad (ICV) que sintetiza el estado fisiológico previo del arbolado en un único valor normalizado —; y (iii) índices espectrales derivados de imágenes satelitales a 10 metros de resolución — medidas de vigor vegetativo y contenido hídrico foliar —.

Con el objeto de garantizar la accesibilidad del dato para la toma de decisiones, esta solución se complementa con dos productos cartográficos operativos: el Mapa de Susceptibilidad Estructural, que identifica las zonas del parque históricamente más expuestas a estrés térmico con mayor cobertura arbórea; y el Índice de Riesgo Combinado (IRC), que integra el IDR diario con el mapa estático para

producir cada mañana de alerta un protocolo de acción de campo con cuatro niveles de prioridad georreferenciados y navegables desde dispositivos GPS de mano.

La comparativa entre el IDR y el sistema actual se ha realizado sobre 1.096 días con datos completos, comparando el modelo directamente contra las alertas reales activadas por el Ayuntamiento de Madrid.

Indicador	Sistema Ayto	IDR_M2 (Agforest)	Mejora
Recall (Detección de riesgo)	29,6 %	99,8 %*	+70,2 pp
F1-Score (Equilibrio global)	0,372	0,807	+0,435
Vuelcos físicos detectados	22 de 41 (53,7 %)	41 de 41 (100 %)	+19 eventos
Días de riesgo vital con parque abierto	26 de 55	0 de 55	-26 situaciones
Cierres evitables ganados	—	28 días/año	—

* Calculado sobre el período 2022–2024. Recall en test independiente (2025): 0,966. Ver Sección 4.2 para la distinción completa.

Obtenidos los resultados, se procedió a realizar una comparativa entre sendos sistemas. El modelo actual de gestión de riesgos dejó el parque abierto 26 de los 55 días en que se produjo un vuelco o una tala de emergencia —el 47,3 % de los eventos de riesgo vital ocurrieron con el parque en funcionamiento normal y sin protocolo de seguridad activado—. **El análisis de los 41 vuelcos físicos confirmados es igualmente contundente: el sistema actual detectó 22 (53,7 %) y el IDR detectó los 41 (100 %).**

A este respecto cabe destacar que **el IDR no ha sido entrenado exclusivamente para la detección de talas de emergencia**. Durante su entrenamiento recibió una variable binaria de incidencia general. Si bien, posteriormente, se segmentaron las diversas tipologías de alertas, esto supone un primer cortafuegos para evitar la exposición de la población a riesgos potenciales. La coincidencia del 100 % con las decisiones de los técnicos municipales del Ayuntamiento de Madrid sobre el terreno confirma que, a priori, el modelo ha aprendido a identificar computacionalmente el mismo patrón de estrés acumulado que el criterio técnico general. En este sentido, podría decirse que **la solución de Agforest actúa como una suerte de asistencia técnica virtual, con disponibilidad 24/7, monitorizando simultáneamente el estado fisiológico del parque**.

El análisis de la distribución estacional de las alertas activadas por el Ayuntamiento revela una rigidez en la respuesta del sistema actual: el número de días con alerta NARANJA o ROJA es prácticamente constante entre estaciones (invierno 50 días, primavera 69, verano 79, otoño 40 en el período 2022–2024), sin que esa distribución refleje la variación real de la vulnerabilidad estructural del arbolado a lo largo del año. Las predicciones AEMET de racha máxima presentan además un sesgo sistemático de sobreestimación de +7,9 km/h de media (RMSE = 15,8 km/h, $r = 0,435$ respecto a la racha real observada), lo que sugiere que parte de los cierres preventivos responden a predicciones que no se materializan en la realidad.

En definitiva, el IDR adapta su umbral de activación al ciclo fenológico del arbolado: 0,28 en primavera — estación de máxima sensibilidad estructural durante la brotación — y 0,55 en verano — cuando el arbolado está aclimatado al calor —. Esta adaptación genera aproximadamente 28 días al año de cierre preventivo evitable. Los 84 días de cierre evitable identificados en el trienio no contienen ningún vuelco, ninguna fractura y ninguna caída de rama registrada: la mejora en disponibilidad del parque no se consigue a costa de la seguridad, sino redistribuyendo los cierres hacia los días que el modelo identifica como realmente peligrosos.

LIMITACIONES ESTRUCTURALES DEL SISTEMA

Tal y como se ha introducido, el IDR no sustituye el criterio técnico de los equipos de arboricultura municipal ni pretende eliminar las inspecciones en campo. Esto se debe a que, **si bien en un futuro — contando con una cantidad suficiente de datos georreferenciados sobre incidencias registradas— podría aproximar el ejemplar que causará el problema, el dataset actual solo permite trabajar a escala parque-día**. Asimismo, los resultados iniciales muestran una **tasa de falsos positivos del 27,5 %, superior en más de diez puntos al sistema actual (17,1 %)**. No obstante, a pesar de que genera algo más de alertas en días que resultan tranquilos, los indicadores de estrés acumulado de esos días confirman que el arbolado se encontraba en una predisposición real al fallo: su ICV medio casi duplica el de los días tranquilos confirmados, y su déficit hídrico acumulado triplica al de los días sin incidencia.

CONTEXTO Y OBJETIVOS

DESAFÍO OPERATIVO: DESFASE ENTRE METEOROLOGÍA Y BIOFÍSICA

Los Jardines de El Buen Retiro, con una superficie aproximada de 118 hectáreas y una masa arbórea compuesta por más de 20.000 ejemplares, constituyen un ecosistema urbano complejo sometido a un intenso estrés ambiental. El actual sistema de gestión de riesgos se basa de forma exclusiva en la racha máxima de viento que proporciona la AEMET. Sin embargo, los datos históricos del período 2022–2024 demuestran que el viento es un detonante, pero no la causa única del fallo estructural.

El análisis revela una brecha crítica de seguridad: el 70,4 % de los días con incidencias reales — caída de ramas o vuelcos — ocurrieron bajo niveles de alerta de AEMET con categoría «Verde» o «Amarilla»; es decir, con el parque abierto al público y sin protocolos de seguridad reforzados. Esta situación genera dos escenarios de ineficiencia simultáneos: (i) riesgo civil no gestionado: el sistema actual no tiene la capacidad de predecir riesgos en escenarios con rachas de viento moderadas, lo que supone un riesgo para la población que transita los jardines, así como una potencial crisis institucional para el Ayuntamiento en un contexto de mayor exigencia ciudadana en materia de seguridad pública; y (ii) impacto socioeconómico por exceso de alertas: los cierres preventivos basados en previsiones de viento que no se traducen en incidencias reales afectan innecesariamente al uso ciudadano y comercial del parque. El análisis de las predicciones AEMET de racha máxima revela un sesgo sistemático de sobreestimación de +7,9 km/h de media sobre el período 2022–2024, lo que contribuye a explicar parte de estos cierres innecesarios.

HIPÓTESIS DE TRABAJO: LA «FATIGA BIOLÓGICA» ACUMULADA

La hipótesis de trabajo de esta PoC parte de una premisa biológica fundamental: la vulnerabilidad estructural de un árbol no es una propiedad estática, sino un proceso dinámico condicionado por el historial de estrés abiótico acumulado. El fallo mecánico no es un evento instantáneo provocado exclusivamente por el viento, sino la culminación de un proceso progresivo de debilitamiento de la estructura lignificada del árbol, provocado por semanas de estrés térmico e hídrico que alteran la turgencia celular, reducen la elasticidad de las fibras leñosas y comprometen las reservas energéticas necesarias para mantener la integridad estructural. En este contexto, el viento actúa como detonante de un fallo que, con frecuencia, ya estaba gestándose silenciosamente durante semanas. Esta hipótesis puede ilustrarse de la siguiente forma: aunque dos ejemplares estén expuestos a la misma racha de viento, el evento no desencadenará consecuencias análogas en un ejemplar sano que en uno sometido a 30 días de calor extremo y un déficit hídrico pronunciado. La vulnerabilidad estructural de ambos es radicalmente diferente; sin embargo, el sistema actual los gestiona con el mismo criterio de riesgo.

Esta hipótesis no implica que el viento sea irrelevante — es el detonante mecánico necesario para que ocurra el fallo físico —. Implica que, ante la misma racha de viento, un árbol con semanas de déficit hídrico y estrés térmico acumulado tiene una probabilidad de fallo estructuralmente mayor que uno en condiciones fisiológicas óptimas. El modelo de Agforest, por tanto, transita de una alerta meteorológica a una alerta biofísica: no predice si mañana habrá viento suficiente para derribar un árbol, sino si el arbolado está en condiciones de soportarlo. El IDR cuantifica esa predisposición estructural previa, complementando — que no sustituyendo — el actual sistema de gestión de riesgos en los Jardines de El Buen Retiro.

OBJETIVOS DE LA PRUEBA DE CONCEPTO (POC)

El objetivo general es validar un Sistema de Alerta Temprana (SAT) que anticipe el riesgo de fallo estructural mediante el uso de inteligencia de datos y sensores remotos (satélite). Los objetivos específicos se desglosan en: (i) cuantificación del estrés acumulado — integrar series históricas de temperatura superficial y humedad para construir indicadores de vulnerabilidad fisiológica que preceden al fallo —; (ii) desarrollo del IDR — comparar tres iteraciones del modelo de Agforest de complejidad creciente para seleccionar el sistema con mayor capacidad de anticipación del riesgo —; (iii) zonificación de la susceptibilidad estructural — elaborar un mapa de alta resolución espacial (10 m/píxel) que identifique las áreas del parque intrínsecamente más vulnerables según su inventario y condiciones de sitio —; y (iv) creación de una herramienta de soporte a la decisión — definir un Índice de Riesgo Combinado (IRC) que cruce el «cuándo» (IDR temporal) con el «dónde» (mapa espacial), permitiendo una gestión de cierres y podas selectiva y eficiente —.

METODOLOGÍA

La metodología de esta PoC se aleja del análisis meteorológico convencional para centrarse en la vulnerabilidad estructural del arbolado como objeto de predicción. El proceso se ha articulado en cuatro fases secuenciales, donde cada etapa aporta una capa de inteligencia al modelo final.

Fase	Denominación	Descripción Técnica	Entregable Principal
WP0	Ingesta de Datos	Integración de series AEMET, imágenes multispectrales (Sentinel-2/Landsat) y registro histórico de incidencias.	Dataset
WP1	Ingeniería de Indicadores	Cálculo de variables de estrés acumulado: GDA, Déficit hídrico, anomalías de LST y vigor vegetativo (NDVI).	Matriz de 69 variables biofísicas.
WP2	Modelado Predictivo	Comparativa de tres modelos de complejidad creciente (M0 <i>baseline</i> de viento, M1 regresión logística ampliada, M2 Random Forest completo) y calibración de umbrales estacionales.	Modelo IDR_M2 (Random Forest).
WP3	Productos Operativos	Generación del Mapa de Susceptibilidad Estructural y del Índice de Riesgo Combinado (IRC).	Cartografía dinámica de riesgo.

WP0. ADQUISICIÓN E INTEGRACIÓN DE DATOS

Datos meteorológicos

Las condiciones meteorológicas diarias del período 2017–2026 se obtuvieron de la Agencia Estatal de Meteorología (AEMET) mediante su API pública. Las variables extraídas incluyen temperatura máxima y mínima del aire, racha máxima de viento, precipitación acumulada y humedad relativa. La extensión histórica hasta 2017 — más allá del trienio de incidencias disponible — responde a una decisión metodológica deliberada: es necesario disponer de una climatología de referencia suficientemente larga para que los indicadores de estrés acumulado reflejen anomalías biológicas genuinas respecto al comportamiento histórico del parque, y no fluctuaciones estacionales normales. Un GDA de 50 unidades en un mes de agosto típico no tiene la misma implicación fisiológica que esos mismos 50 GDA en un mes de mayo anómalo; solo una serie histórica extensa permite al modelo establecer esa distinción.

Datos de incidencias

El Ayuntamiento de Madrid ha facilitado el registro histórico de incidencias del parque para el período 2022–2024, recopiladas por la empresa ACCIONA junto con los niveles de alerta meteorológica diaria (verde, amarilla, naranja o roja). Cada incidencia está clasificada por tipología: caída de ramas, vuelco o fractura de árbol, y tala de emergencia por riesgo. El registro original está a nivel de evento individual; para el análisis se agregó a escala parque-día, convirtiendo el conteo diario de eventos en una variable binaria (0 = ninguna incidencia ese día, 1 = al menos una incidencia). Esta escala de análisis es coherente con la escala a la que opera el sistema de alertas del Ayuntamiento, que toma decisiones de cierre a nivel de día completo.

Imágenes satelitales

Se han empleado dos fuentes de imágenes satelitales con resoluciones espaciales y temporales complementarias. La primera — satélites de observación terrestre con resolución de 30 metros por píxel y frecuencia de revisita de aproximadamente 8 días — proporciona la medición de Temperatura Superficial del Terreno (LST), calculada a partir de la banda térmica aplicando los factores de escala y correcciones atmosféricas oficiales. La LST representa la temperatura de la superficie del suelo y de las copas de los árboles tal como la registra el sensor desde el espacio, no la temperatura del aire. Se seleccionaron únicamente imágenes con cobertura de nubes inferior al 20 %. La segunda fuente — satélite óptico con resolución espacial de 10 metros por píxel y frecuencia de revisita de 5 días — no mide temperatura directamente, pero permite calcular índices espectrales que describen el estado de la

vegetación y el contenido hídrico foliar. Se aplicó la máscara de clasificación de píxeles incluida en el producto estándar de nivel 2A, que distingue entre píxeles limpios, nubes, sombras de nubes, agua y vegetación; solo los píxeles clasificados como vegetación o suelo sin interferencias se utilizaron en el análisis.

WPI. LST SINTÉTICA A 10 M Y CONSTRUCCIÓN DE INDICADORES DE ESTRÉS

La generación de indicadores de estrés acumulado parte de un reto técnico fundamental: ningún satélite operativo comercial combina actualmente resolución de 10 metros con medición directa de temperatura superficial. La fuente que mide temperatura lo hace a 30 metros — resolución insuficiente para discriminar diferencias térmicas dentro de un parque urbano de 118 hectáreas, donde la distancia entre una zona de arbolado denso y un camino asfaltado puede ser de apenas 15–20 metros —. La solución adoptada por Agforest es un procedimiento propietario de incremento de la resolución espacial de la temperatura superficial, denominado en este informe LST sintética, que permite pasar de 30 a 10 metros aprovechando la relación que existe entre la temperatura superficial y los índices espectrales disponibles a 10 metros. Esta aproximación mitiga el problema de los «píxeles mixtos» térmicos, donde una sola observación de la fuente de baja resolución promedia inevitablemente la temperatura del dosel arbóreo con la del pavimento circundante, diluyendo la señal de estrés real.

Los índices espectrales no operan aquí como herramientas de clasificación de coberturas sino como descriptores dinámicos de la respuesta biofísica del entorno. La lógica subyacente reconoce cómo la evapotranspiración y la turgencia del arbolado ejercen un efecto refrigerante frente a la inercia térmica de viales y pavimentos: el procedimiento propietario de Agforest aprende esa relación sobre un conjunto de fechas de referencia cuidadosamente seleccionadas — episodios de calor extremo con daño, episodios de viento con daño, días de incidencia sin alerta meteorológica roja — y la extrapola después a toda la serie histórica disponible. La correlación entre la LST resultante y la temperatura del aire medida por AEMET fue de $r = 0,955$, con una diferencia media de $+2,8 \pm 1,9$ °C, coherente con el balance térmico urbano. El efecto de enfriamiento del arbolado denso respecto a superficies impermeables fue de $-6,4$ °C de media, confirmando la capacidad del producto para discriminar microclimas intra-parque con la precisión necesaria para la construcción de indicadores operativos.

Con la serie temporal diaria de LST sintética disponible para todo el período 2017–2025 se construyeron los indicadores de estrés acumulado que constituyen el núcleo del IDR. Los indicadores de estrés térmico acumulado incluyen: (i) la anomalía LST — diferencia entre la temperatura superficial media del arbolado del parque en un día dado y la media histórica de ese mismo día del año, que responde a la pregunta «¿cuánto se desvía hoy el parque de su comportamiento térmico habitual para esta fecha?» —; (ii) los grados-día acumulados (GDA), suma de los excesos de temperatura por encima del umbral crítico durante los 10, 15 y 30 días anteriores, capturando la carga térmica reciente con tres granularidades temporales distintas — estrés agudo versus estrés crónico —; y (iii) la persistencia térmica, número de días consecutivos por encima del umbral inmediatamente anteriores a cada fecha, que captura los episodios sostenidos de calor que no permiten recuperación fisiológica nocturna.

Los indicadores de estrés hídrico complementan la dimensión térmica con variables que cuantifican el balance hídrico del arbolado: el déficit hídrico — balance entre la precipitación registrada y la evapotranspiración potencial estimada, acumulado en distintas ventanas —, los días secos consecutivos y la precipitación acumulada reciente como señal de recuperación. Todos estos componentes se sintetizan en el Índice Compuesto de Vulnerabilidad (ICV), valor propietario normalizado en el rango [0, 1] que integra el estado fisiológico acumulado del parque en un único número diario. Un ICV próximo a 1 indica que el parque acumula simultáneamente estrés térmico e hídrico por encima de su comportamiento histórico; un ICV próximo a 0 indica condiciones dentro de la normalidad. Es importante señalar que el ICV se calcula exclusivamente con datos del día anterior ($t-1$) para garantizar que nunca incorpora información del día cuyo riesgo se quiere predecir — la ventana de actuación administrativa mínima es de al menos 12–24 horas —.

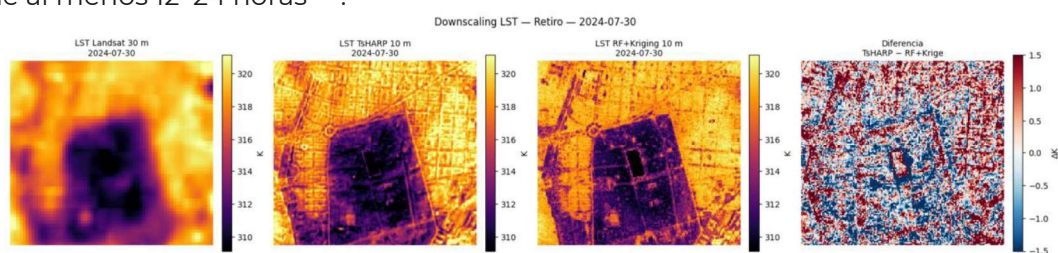


Figura 1. Ejemplo de downscaling LST para el evento 2024-07-30 (Cat2). Izquierda: LST Landsat a 30m. Centro: LST_TsHARP a 10m. Derecha: LST_RF+Kriging_corr a 10m

WP2. MODELADO PREDICTIVO: EL IDR

Estructura del *dataset* analítico

El *dataset* analítico integra tres bloques de información para cada día del período 2022–2025: (i) los indicadores de estrés acumulado calculados en WPI — GDA en tres ventanas temporales, déficit hídrico acumulado, anomalía LST, persistencia térmica e ICV —; (ii) los índices espectrales satelitales — medidas estandarizadas, medias móviles y variaciones temporales de los indicadores de vigor vegetativo y contenido hídrico foliar, derivadas de imágenes a 10 metros de resolución con frecuencia de revisita de 5 días —; y (iii) la variable respuesta: incidencia binaria del día (0 = ninguna incidencia registrada, 1 = al menos una incidencia). El *dataset* resultante cubre 1.096 días del período 2022–2024 con más de 60 variables biofísicas y meteorológicas por día. Los días sin imagen disponible por nubosidad reciben los valores espectrales interpolados desde las escenas adyacentes; los valores faltantes restantes se imputan con la mediana del conjunto de entrenamiento, aplicada estrictamente después del corte temporal para evitar contaminación de datos futuros.

Las tres iteraciones comparadas del modelo de Agforest

Se desarrollaron tres iteraciones del modelo con complejidad creciente para cuantificar el valor añadido de cada capa de información adicional:

- (i) **Primera versión** — base metodológica con viento observado: emplea exclusivamente la racha máxima de viento observada en los siete días anteriores. Su propósito no es replicar el sistema del Ayuntamiento — que usa predicciones de viento a 24 horas — sino establecer el *benchmark* más exigente posible para la variable viento: si las versiones más complejas no superan a un sistema que conoce el viento real sin errores de pronóstico, las variables de estrés acumulado no aportarían valor. Obtuvo AUC = 0,661 en el conjunto de prueba independiente. Para confirmar la solidez de este *benchmark*, se entrenó adicionalmente una variante con predicciones AEMET a 24 horas — el mismo input que utiliza el sistema real del Ayuntamiento —: el resultado confirmó que la versión con viento observado es el competidor más exigente disponible, al obtener mayor Recall (0,828 vs 0,655) y la mitad de Falsos Negativos.
- (ii) **Segunda versión** — formulación lineal ampliada: incorpora adicionalmente las variables de estrés acumulado (ICV, déficit hídrico, GDA). Permite cuantificar cuánto mejora la predicción al añadir la dimensión de vulnerabilidad fisiológica manteniendo la interpretabilidad de un esquema lineal: cada variable tiene una contribución numérica que indica su aporte individual al riesgo. Obtuvo AUC = 0,668.
- (iii) **Versión refinada no lineal** — IDR: incorpora un algoritmo propietario de aprendizaje automático — de carácter no lineal y basado en la combinación de múltiples estimadores independientes — que captura relaciones complejas entre las variables sin que el analista deba especificar de antemano las interacciones. El modelo explora simultáneamente cientos de «rutas de decisión» independientes; cada ruta evalúa preguntas del tipo «¿es el déficit hídrico mayor de X y, además, el ICV supera tal nivel?», agrupando los días en escenarios de mayor o menor probabilidad de incidencia. Al promediar el resultado de cientos de rutas construidas de forma independiente, el sistema reduce el riesgo de sobreajuste a los datos de entrenamiento y captura combinaciones que un esquema lineal no puede detectar. Esta versión fue seleccionada como definitiva por obtener el mejor rendimiento en validación (AUC = 0,744, Recall en test = 0,966).

Métrica	v1 — Referencia	v2 — Estrés acumulado	v3 — Refinada (IDR)
Complejidad	Lineal (referencia)	Lineal ampliada	No lineal — IA propietaria
AUC test 2025	0,661	0,668	0,744
AUC val. cruzada	0,540 ± 0,021	0,664 ± 0,060	0,632 ± 0,067
Interpretación	Benchmark viento	Mejora con estrés acumulado	Modelo seleccionado

Una de las ventajas del esquema seleccionado es que permite calcular la importancia relativa de cada variable, es decir, cuánto contribuye cada una a mejorar la capacidad predictiva del modelo. Las cinco variables más importantes identificadas fueron, por orden: déficit hídrico acumulado en 30 días, grados-

día acumulados de temperatura del aire, grados-día en ventana de 15 días, grados-día en ventana de 10 días e Índice Compuesto de Vulnerabilidad. La racha máxima de viento se situó en la posición 20 de las 38 variables analizadas. Este resultado es la evidencia cuantitativa central de la PoC: el viento es un detonante necesario para que ocurra el fallo mecánico, pero no es la variable que mejor explica cuándo y con qué intensidad se producen las incidencias. Son las semanas previas de estrés acumulado las que determinan si el árbol tiene la resiliencia estructural suficiente para soportar la misma racha de viento sin ceder.

Validación temporal estricta

Para simular el uso operativo real del sistema y evitar el *data leakage* — es decir, que el modelo acceda durante el entrenamiento a información del futuro que en un despliegue real no estaría disponible — se aplicó una partición temporal estricta: el modelo se entrenó con todos los datos del período 2022–2024 y se evaluó exclusivamente con los datos de enero a mayo de 2025, que el modelo no había visto durante el entrenamiento. Esta metodología es equivalente a entrenar el sistema a finales de 2024 y medir su comportamiento predictivo durante los primeros meses de 2025. Cualquier técnica de imputación de valores faltantes se ajustó exclusivamente sobre el conjunto de entrenamiento y se aplicó después al conjunto de prueba, siguiendo el mismo principio.

Calibración de umbrales estacionales

El IDR produce para cada día un valor entre 0 y 1 que representa la probabilidad estimada de que se registre al menos una incidencia en el parque ese día. Para convertir esa probabilidad en una decisión operativa binaria — activar alerta preventiva o no activarla — se calibraron umbrales de activación específicos por estación, optimizando en cada caso la capacidad de detectar todos los días con incidencia real (Recall) sin generar un exceso de cierres innecesarios.

El análisis de la distribución estacional de las alertas activadas por el Ayuntamiento de Madrid revela una rigidez en la respuesta del sistema actual: el número de días con alerta NARANJA o ROJA es prácticamente constante entre estaciones (invierno 50 días, primavera 69, verano 79, otoño 40 en el período 2022–2024), sin que esa distribución refleje la variación real de la vulnerabilidad estructural del arbolado. A modo de referencia estadística, la mediana de la primera versión del modelo en los días en que el Ayuntamiento activó alerta es de aproximadamente 0,503 en todas las estaciones — con una variación de apenas 0,006 entre la más alta y la más baja —, confirmando que las decisiones de cierre se correlacionan con el viento reciente de forma uniforme a lo largo del año, sin adaptación al ciclo fenológico. El IDR, en contraste, adapta su umbral entre 0,28 en primavera y 0,55 en verano, reflejando la variación real de la vulnerabilidad estructural del arbolado:

- (i) **Invierno (diciembre, enero, febrero)** — umbral 0,30. El arbolado caducifolio ha perdido la hoja, lo que reduce significativamente su carga estructural y la probabilidad de incidencia ante la misma racha de viento. El umbral 0,30 refleja un riesgo real pero moderado en esta estación y es el segundo umbral más bajo del año, no por falta de sensibilidad, sino porque en invierno la biofísica del arbolado produce genuinamente menos incidentes dado el mismo nivel de estrés.
- (ii) **Primavera (marzo, abril, mayo)** — umbral 0,28. El umbral más bajo del año se justifica por ser la estación de máxima sensibilidad estructural del arbolado: durante la brotación, el flujo de savia eleva la turgencia de los tejidos, y un árbol que ha sufrido un invierno de déficit hídrico entra en primavera con una vulnerabilidad estructural superior a la que cualquier indicador visual externo permite detectar. El umbral 0,28 garantiza que el sistema no pase por alto episodios de riesgo en una estación en la que el arbolado está fisiológicamente más expuesto.
- (iii) **Verano (junio, julio, agosto)** — umbral 0,55. En verano el arbolado está en plena actividad fotosintética y aclimatado a las temperaturas elevadas. El umbral más alto del año filtra los días de calor basal sin estrés acumulado significativo, reduciendo los cierres preventivos innecesarios precisamente en la estación de mayor afluencia de ciudadanos al parque. Este umbral elevado no implica menor seguridad: el modelo sigue detectando el 100 % de los vuelcos registrados en verano porque estos ocurren en días con IDR claramente por encima de 0,55.
- (iv) **Otoño (septiembre, octubre, noviembre)** — umbral 0,35. El otoño es una estación de transición en la que la carga foliar decrece progresivamente. El umbral intermedio de 0,35 refleja la reducción gradual de la vulnerabilidad estructural a medida que avanza la estación.



Figura 2. Comparativa de umbrales de activación del IDR (barras) frente al umbral implícito constante del sistema actual (línea de referencia). La diferencia entre el umbral de primavera (0,28) y el de verano (0,55) refleja la adaptación al ciclo fenológico del arbolado.

WP3. PRODUCTOS OPERATIVOS: MAPA DE SUSCEPTIBILIDAD E IRC

Mapa de Susceptibilidad Estructural

El Mapa de Susceptibilidad Estructural es un producto cartográfico estático — no cambia de un día a otro — que representa la predisposición de cada zona del parque a registrar incidencias cuando el arbolado se encuentra bajo condiciones de estrés. Responde a la pregunta: dado que hoy el modelo ha lanzado una alerta general para todo el parque, ¿a qué zonas envío primero a los operarios?

Se construyó mediante un Análisis Multicriterio Espacial (AMC) que combina dos capas de información: la exposición térmica histórica relativa de cada punto del parque — calculada sobre el período 2017–2026 y expresada como percentil dentro del gradiente térmico intra-parque, de forma que un píxel con percentil alto significa que ese punto concreto ha sido históricamente más caliente que la mayoría del resto del parque — y el tipo de cobertura arbórea, que pondera cada zona según la probabilidad de que un fallo estructural en ella genere una incidencia con riesgo para el ciudadano. Las zonas de arbolado denso obtienen el peso máximo (1,00), el arbolado disperso (0,85), y las superficies sin arbolado — láminas de agua, viales, zonas impermeables — quedan fuera de la escala de prioridad por carecer de riesgo arbóreo directo. La razón es estrictamente operativa: un prado caliente o una zona de césped pueden tener valores de temperatura superficial elevados, pero no representan ningún riesgo de caída de árbol porque no hay árbol que pueda caer; incluirlos en la escala distorsionaría el ranking y desviaría los recursos de campo hacia zonas irrelevantes.

La diferencia de casi 5 °C entre el nivel Muy Alta (+2,54 °C sobre la media histórica) y el nivel Muy Baja (–2,23 °C) es un dato de gran relevancia operativa: dos árboles situados en zonas opuestas del parque pueden estar recibiendo la misma alerta meteorológica del Ayuntamiento y, sin embargo, el árbol en zona Muy Alta ha estado expuesto históricamente a casi 5 °C más de temperatura superficial que el árbol en zona Muy Baja. Bajo condiciones de estrés acumulado, esa diferencia térmica estructural determina cuál de los dos tiene mayor probabilidad de ceder ante la misma racha de viento.

El mapa se entrega en tres versiones con diferentes niveles de detalle y uso previsto: el ráster original a resolución de 10 metros para análisis técnico en SIG, el mapa suavizado mediante filtro de mayoría para comunicación y presentación, y el mapa agregado por cuadrículas de 50×50 metros para uso operativo directo en campo. Esta última versión está georeferenciada en el sistema de coordenadas EPSG:32630 (UTM zona 30N), compatible con dispositivos GPS de mano, lo que permite a los operarios localizar con precisión cada cuadrícula de prioridad sin necesidad de equipamiento técnico especializado.

Prioridad	Categoría	Superficie (ha)	LST media histórica	Acción recomendada
1	Muy Alta	20,46	+2,54 °C sobre media	Inspección preventiva prioritaria
2	Alta	20,46	+0,43 °C sobre media	Inspección en episodios de alerta
3	Media	20,46	–0,75 °C sobre media	Vigilancia periódica
4	Baja	20,46	–1,48 °C sobre media	Rutina estándar
5	Muy Baja	20,46	–2,23 °C sobre media	Rutina estándar

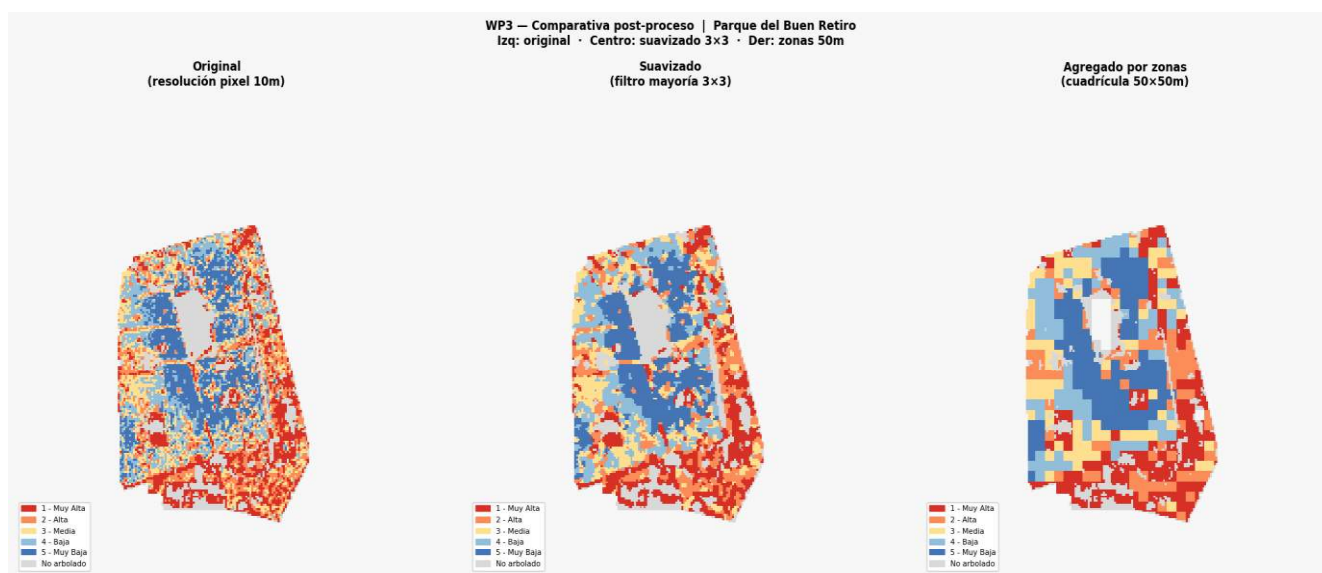


Figura 3. Mapa de Susceptibilidad Estructural de los Jardines de El Buen Retiro. La zona Norte (Rosaleda) y la Pradera del Ángel Caído presentan la mayor exposición térmica histórica acumulada (+2,54 °C sobre la media del parque). Las zonas en azul (entorno del Estanque Grande) presentan la menor susceptibilidad (-2,23 °C), confirmando la capacidad del producto para discriminar microclimas intra-parque.

Índice de Riesgo Combinado (IRC)

Las secciones anteriores han producido dos productos analíticos independientes: el IDR, que responde a la pregunta «¿es hoy un día peligroso para el parque?», y el Mapa de Susceptibilidad Estructural, que responde a «¿qué zonas del parque son estructuralmente más vulnerables?». El Índice de Riesgo Combinado (IRC) integra ambas dimensiones en un único producto operativo que responde a la pregunta que el técnico de campo necesita responder cada mañana de alerta: ¿a qué zonas concretas envío primero a mis operarios?

La lógica de construcción combina la dimensión temporal — el IDR, que varía cada día según el estado de estrés acumulado — con la dimensión espacial — la susceptibilidad estructural histórica de cada zona —, generando un mapa que cambia diariamente, pero cuya estructura espacial subyacente permanece estable. Esta separación entre la dimensión temporal (IDR, dinámico) y la dimensión espacial (Susceptibilidad, estática) es deliberada y metodológicamente importante: mezclarlas en un único índice permanente haría el producto inoperable. Un mapa que cambia continuamente no puede servir de referencia para la planificación preventiva, y un mapa que no cambia no puede reflejar el estado de riesgo del día. La multiplicación tiene además una interpretación operativa directa: si el IDR del día es bajo — día tranquilo —, el IRC de todas las zonas del parque será bajo independientemente de su susceptibilidad. Si el IDR es alto, el IRC amplifica las diferencias entre zonas, asignando la prioridad máxima exactamente donde la convergencia de vulnerabilidad temporal y espacial es mayor.

RESULTADOS

MAPA SUSCEPTIBILIDAD ESTRUCTURAL: DISTRIBUCIÓN ESPACIAL DE INCIDENCIAS

La correspondencia espacial entre las zonas de mayor susceptibilidad térmica y la mayor concentración de incidencias y vuelcos no es trivial: las zonas de susceptibilidad Muy Alta y Alta acumulan el 72 % del total de incidencias registradas en el trienio, siendo el 21,5 % del área total del parque. Esta concentración espacial tiene una implicación directa para la estrategia de gestión de campo: en los días de alerta, las primeras dos horas de inspección deben focalizarse en las zonas de prioridad Crítica y Alta del IRC. Que el modelo identifique como zonas de máxima susceptibilidad precisamente aquellas que el registro histórico de incidencias valida como las más peligrosas es la confirmación espacial de que las variables de estrés acumulado capturan señales fisiológicas reales y no artefactos estadísticos.

COMPARATIVA IDR VS SISTEMA DE ALERTAS DEL AYUNTAMIENTO

Serie temporal IDR vs incidencias reales (2022–2024)

La serie temporal del IDR a lo largo del trienio 2022–2024 revela tres patrones sistemáticos que confirman, de forma visual e intuitiva, la coherencia del sistema con la biofísica del arbolado. Primero, la

concentración de incidencias en los meses de verano e inicio de otoño coincide con los picos de IDR más elevados, confirmando la hipótesis de fatiga biológica acumulada: los períodos de calor extremo sostenido producen la mayor vulnerabilidad estructural y, como consecuencia, la mayor frecuencia de fallos. Segundo, los meses de invierno con IDR bajo y cero incidencias confirman que los cierres preventivos invernales del sistema actual — motivados por rachas de viento — ocurren con frecuencia en períodos de escasa predisposición estructural del arbolado, momento en que la relación entre viento y fallo es estadísticamente débil. Tercero, los meses con vuelcos confirmados se concentran sin excepción en los períodos de IDR más elevado, confirmando que el modelo captura correctamente los picos de riesgo vital y no solo la frecuencia de incidencias menores.

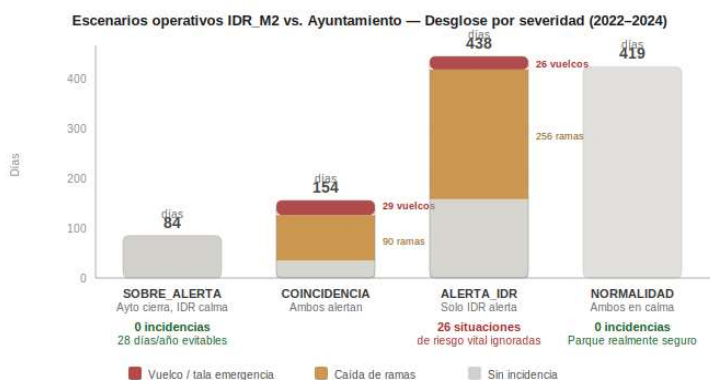


Figura 4. Distribución de los 1.096 días analizados por escenario operativo y severidad de incidencia. Los 84 días de *SOBRE_ALERTA* y los 419 de *NORMALIDAD* no registran ninguna incidencia. Los 438 días de *ALERTA_IDR* contienen 26 vuelcos y 256 caídas de ramas que el sistema actual dejó pasar.

Nota sobre la validez de la comparativa

El resultado central de esta sección — Recall 0,998 vs 0,296 — se obtuvo comparando el IDR directamente contra las alertas reales activadas por el Ayuntamiento (el registro de alertas municipales efectivamente activadas), no contra un competidor interno construido para la ocasión. Esto significa que el rendimiento del sistema está validado contra la realidad operativa — lo que el Ayuntamiento efectivamente hizo cada día durante tres años — y no contra un *benchmark* metodológico. El único juez relevante para evaluar ambos sistemas es el historial real de lo que ocurrió cada día en el parque.

Antes de interpretar las métricas, es importante encuadrar correctamente qué mide el IDR. El modelo de Agforest evalúa la predisposición al fallo estructural — la vulnerabilidad acumulada del arbolado — no la probabilidad de que ocurra un evento meteorológico detonante. Esta distinción tiene una implicación directa sobre cómo interpretar los Falsos Positivos: un día clasificado como «alerta» por el IDR, pero sin incidencia registrada, no es necesariamente un error del modelo — puede ser un día en que el arbolado estaba genuinamente en estado frágil pero el detonante (viento suficiente) no se materializó —. El IDR y el sistema de alertas meteorológicas del Ayuntamiento son complementarios: el primero mide la vulnerabilidad del receptor, el segundo mide la intensidad del detonante. La combinación de ambas señales en el protocolo operativo constituye el sistema de decisión completo.

Tabla de métricas completa

Métrica	Sistema Ayto.	IDR	Diferencia
Verdaderos Positivos (TP)	119	401	+282
Falsos Positivos (FP)	119	191	+72
Falsos Negativos (FN)	283	1	-282
Verdaderos Negativos (TN)	575	503	-72
Recall (Sensibilidad)	0,296	0,998	+0,702
Precisión	0,500	0,677	+0,177
F1-Score	0,372	0,807	+0,435
Tasa de Falsa Alarma (FPR)	0,171	0,275	+0,104

Interpretación métrica a métrica

- (i) **Falsos Negativos (FN): de 283 a 1.** Este es el dato más relevante desde el punto de vista de la seguridad. El sistema actual dejó el parque abierto en 283 días en los que se registró al menos una incidencia real. El IDR habría dejado pasar únicamente 1 de esos 283 días — una reducción del 99,6 % —. Dicho de otra forma: el sistema actual no detecta como peligrosos el 70,4 % de los días en los que algo termina cayendo o rompiéndose en el parque. El análisis de ese único Falso Negativo revela que ocurrió el 17 de mayo de 2025, con el modelo a 0,012 puntos del umbral de activación estacional de primavera; las condiciones de estrés acumulado ese día eran reales e indicaban vulnerabilidad — no se trata de un día que el modelo ignoró, sino de un día en el límite estadístico del umbral calibrado —.
- (ii) **Recall: de 0,296 a 0,998.** El Recall — también llamado Sensibilidad — responde a la pregunta: de todos los días en los que hubo incidencia real, ¿qué porcentaje detectó el sistema? El Ayuntamiento detecta menos de 1 de cada 3 días peligrosos. El IDR detecta prácticamente todos. En sistemas de seguridad pública, el Recall es la métrica crítica: un sistema que deja pasar el 70 % de los eventos reales no puede considerarse un sistema de alerta funcional.
- (iii) **Precisión: de 0,500 a 0,677.** La Precisión responde a la pregunta inversa: de todas las alertas que lanzó el sistema, ¿qué porcentaje correspondía a días con incidencia real? El sistema actual tiene una Precisión del 50 %, lo que significa que la mitad de sus cierres preventivos no están respaldados por ninguna incidencia real — de media, por cada día que cierra con incidencia, cierra otro en que no ocurre nada —. El IDR mejora a 67,7 %: cuando alerta, hay una probabilidad de dos de cada tres de que ese día registre algún evento.
- (iv) **Tasa de Falsa Alarma (FPR): de 0,171 a 0,275.** Es el único indicador donde el IDR presenta un resultado desfavorable. El IDR genera 72 días adicionales de alerta en días que finalmente resultan tranquilos. Sin embargo, este resultado debe leerse en su contexto correcto. Un Falso Positivo en la gestión del parque significa acordonar una zona o reforzar la supervisión durante unas horas — el coste es real pero asimétrico respecto al coste de un Falso Negativo —. El IDR intercambia 72 días adicionales de alerta preventiva innecesaria por la eliminación de 282 días en los que el sistema actual dejó pasar eventos reales: por cada día de falsa alarma adicional que genera el IDR, evita 3,9 eventos reales que el sistema actual dejaba pasar. Esta asimetría es precisamente la razón por la que los umbrales del IDR se calibraron para maximizar el Recall aceptando una Tasa de Falsa Alarma más alta — es una decisión de diseño deliberada a cambio de garantizar que ningún evento grave ocurra con el parque abierto, no un defecto del modelo —.
- (v) **Parte de los Falsos Positivos pueden ser días de riesgo real no materializado.** El análisis de las condiciones medias de los 191 días clasificados como Falsos Positivos confirma esta hipótesis: su ICV medio casi duplica el de los días tranquilos confirmados, y su déficit hídrico acumulado triplica al de los días sin incidencia. Son días en los que el arbolado se encontraba en una predisposición real al fallo que no llegó a producirse — posiblemente por menor afluencia de público ese día, por intervenciones preventivas del equipo técnico no registradas como incidencia, o simplemente por factores aleatorios no medibles —. No es posible cuantificar cuántos de los 191 FP corresponden a este supuesto, pero su existencia implica que el número de «falsas alarmas puras» es probablemente inferior a 191.
- (vi) **AUC = 0,744 en test independiente.** El AUC es la métrica comparativa más robusta entre modelos porque es independiente del umbral de decisión: un AUC de 0,744 significa que en el 74,4 % de los casos el modelo asigna mayor probabilidad de incidencia a un día que realmente tuvo incidencia que a uno que no la tuvo. Calculado sobre el conjunto de prueba independiente (enero–mayo 2025), es el valor más conservador y el que mejor refleja la capacidad discriminativa del sistema en condiciones reales de despliegue. El Recall de 0,966 es el mejor estimador disponible sobre datos no vistos, pero debe entenderse como un límite superior del rendimiento esperado hasta disponer de un año completo de prueba que incluya los meses de verano — los más relevantes desde el punto de vista del estrés acumulado y ausentes del conjunto de prueba por no disponer de datos de incidencias completos de 2025 en el momento de esta PoC —.

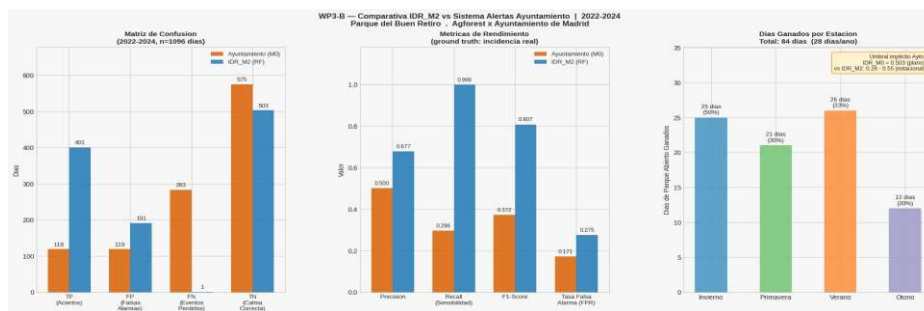


Figura 5. Comparativa de rendimiento IDR_M2 vs sistema actual del Ayuntamiento (2022-2024). Izquierda: matriz de confusión. Centro: métricas de rendimiento. Derecha: días de parque abierto ganados por estación.

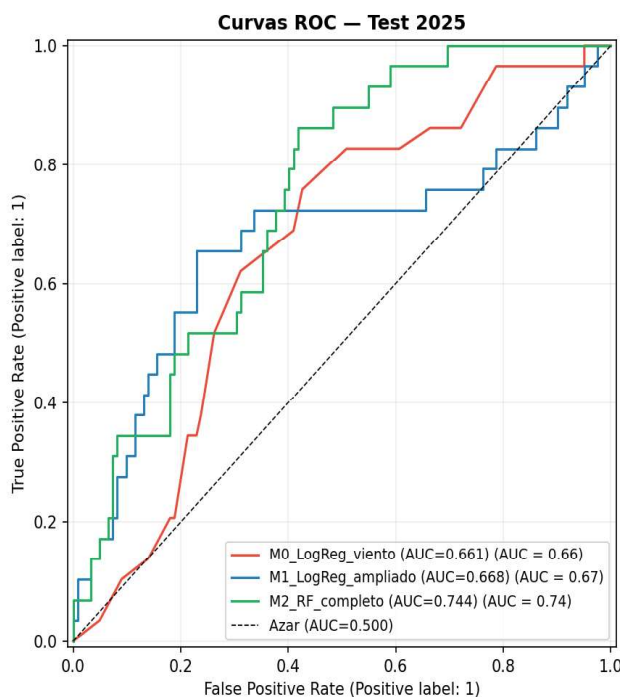


Figura 6. Curvas ROC comparativas de las tres iteraciones del modelo de Agforest sobre el conjunto de test independiente (enero-mayo 2025). La versión refinada (verde) supera a las dos anteriores en toda la curva, confirmando la mejora aportada por las variables de estrés acumulado.

Días de parque abierto ganados por estación

Estación	Días ganados	Cierres Ayto.	% Cierres evitables	Justificación biofísica
Invierno	25	50	50,0 %	Arbolado sin hoja; menor carga estructural.
Primavera	21	69	30,4 %	Sensibilidad alta (Umbral 0,28).
Verano	26	79	32,9 %	Filtrado de calor basal (Umbral 0,55).
Otoño	12	40	30,0 %	Transición fenológica.
TOTAL	84	238	35,3 %	~28 días / año recuperados

El dato más llamativo de la distribución estacional es el invierno: el 50 % de los cierres preventivos invernales del Ayuntamiento resultan evitables según el IDR. La explicación es directamente el umbral plano: el sistema actual responde al viento de forma uniforme en enero y en julio, sin tener en cuenta que en invierno el arbolado caducifolio ha perdido la hoja y tiene una carga estructural significativamente menor, lo que reduce la probabilidad de incidencia ante la misma racha de viento. El IDR incorpora ese conocimiento fenológico rebajando el umbral de activación a 0,30 en invierno, lo que paradójicamente genera menos cierres innecesarios porque discrimina mejor los días realmente peligrosos de los simplemente ventosos. Es importante subrayar que los 84 días ganados no contienen ningún vuelco, ninguna fractura y ninguna caída de rama registrada. La mejora en disponibilidad del parque no se consigue relajando la seguridad, sino redistribuyendo los cierres hacia los días que el modelo identifica como realmente peligrosos.

Desglose por escenario operativo

Escenario	Días	Vuelcos / talas	Caída ramas	% Días con incidencia
SOBRE-ALERTA (Cierre Ayto. innecesario)	84	0	0	0,0 %
COINCIDENCIA (Ambos alertan)	154	29	90	77,3 %
ALERTA IDR (Solo IDR protege)	438	26	256	64,4 %
NORMALIDAD (Ambos en calma)	419	0	0	0,0 %

La lectura de esta tabla confirma tres cosas simultáneamente. Primero, los 84 días de cierre evitable tienen cero incidencias de cualquier tipo — no hay ningún coste de seguridad en evitarlos —. Segundo, los 438 días en los que solo el IDR alertó contienen 26 vuelcos y 256 caídas de ramas — días de riesgo real que el sistema actual dejó pasar completamente con el parque abierto al público —. Cada uno de esos 26 días de vuelco representa una situación de potencial responsabilidad civil para la administración: un árbol volcó con el parque en funcionamiento normal, sin protocolo de seguridad activado y sin acordonamiento previo. Tercero, y con igual relevancia, los 419 días de normalidad compartida — en los que ambos sistemas coinciden en no alertar — registran literalmente cero incidencias de cualquier tipo. Este resultado valida una propiedad del IDR que podría pasar desapercibida: cuando el modelo dice que el parque es seguro, el parque es realmente seguro.



Figura 7. Distribución de los 1.096 días analizados por escenario operativo y severidad de incidencia. Los 84 días de *SOBRE_ALERTA* y los 419 de *NORMALIDAD* no registran ninguna incidencia. Los 438 días de *ALERTA_IDR* contienen 26 vuelcos y 256 caídas de ramas que el sistema actual dejó pasar.

ANÁLISIS DE SEVERIDAD

Las métricas de la sección anterior tratan todas las incidencias como equivalentes — un día con una rama caída cuenta igual que un día con un árbol volcado completo —. Esta simplificación es metodológicamente necesaria para construir el modelo, pero oculta una dimensión crítica para la toma de decisiones de seguridad pública: no todos los eventos tienen el mismo potencial de causar víctimas o daños graves. Esta sección desagrega los resultados por nivel de severidad para responder a la pregunta que más importa al responsable municipal: ¿cuántos de los eventos de mayor riesgo para el ciudadano detecta cada sistema?

Los resultados son igualmente desdoblados por severidad. El IDR habría alertado en los 55 días con evento de severidad alta — vuelcos, fracturas y talas de emergencia — sin dejar pasar ninguno, frente a los 29 que detectó el sistema actual. El dato de las talas de emergencia merece una lectura específica: el 100 % de las 14 talas ejecutadas por los técnicos municipales coinciden con días en los que el IDR tenía la alerta activa. Esto indica que el modelo capta el mismo patrón de estrés acumulado que el técnico identifica en campo durante su inspección visual del árbol — pero de forma anticipada y sistemática, sin necesidad de que el árbol ya muestre síntomas visibles al ojo humano —. En este sentido, el IDR actúa como una primera capa de triaje que orienta las inspecciones hacia los días y zonas donde la probabilidad de encontrar ejemplares en estado crítico es mayor.

Nivel de Severidad	Total días	Detectados Ayto.	Recall Ayto.	Detectados IDR_M2	Recall IDR_M2
Alta (Vuelco / Tala)	55	29	0,527	55	1,000

Media (Caída ramas)	347	90	0,259	346	0,997
----------------------------	-----	----	-------	-----	-------

En los 55 días con evento de severidad alta, el sistema actual dejó el parque abierto en 26 — el 47,3 % de los días con riesgo vital ocurrieron sin protocolo de seguridad activado —. El IDR habría activado alerta en los 55, sin excepción. El Índice de Protección de Seguridad (IPS) alcanza el 100 %.

ÍNDICE DE RIESGO COMBINADO (IRC): PROTOCOLO DE ACCIÓN DE CAMPO

Para traducir el IRC continuo [0–1] en instrucciones operativas concretas, se definieron cuatro niveles de acción con umbrales fijos:

Nivel IRC	Umbral	Superficie en día máx. (ha)	Instrucción de campo
Crítico (4)	$IRC > 0,70$	20,46	Acordonar inmediatamente antes de abrir
Alto (3)	0,50 – 0,70	40,92	Inspección ocular urgente (primeras 2 horas)
Moderado (2)	0,30 – 0,50	20,46	Vigilancia reforzada durante la jornada
Bajo (1)	$IRC < 0,30$	36,12*	Rutina de supervisión estándar

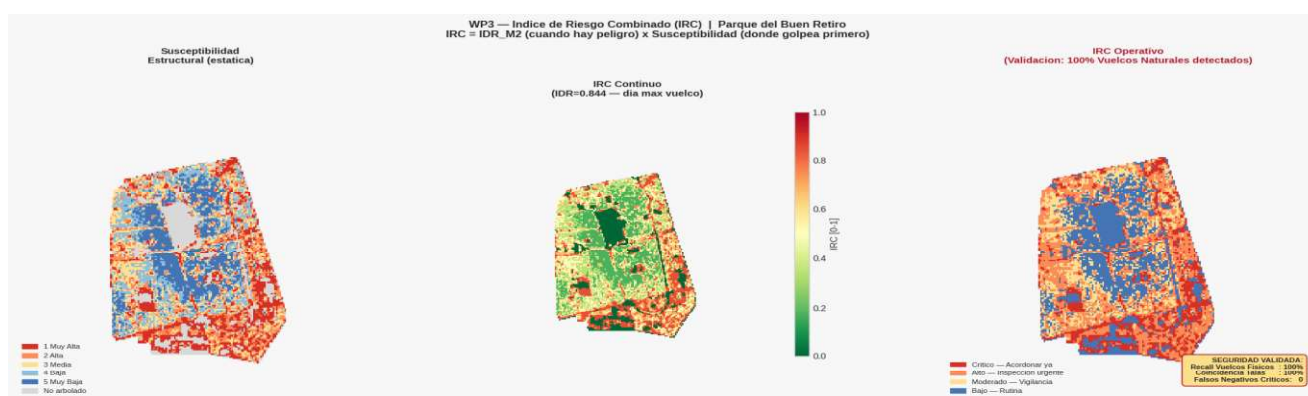


Figura 8. Ejemplo del mapa IRC para el día de máximo IDR del período analizado ($IDR = 0,844$). Izquierda, las zonas Críticas (rojo) en la Zona Norte y la Pradera del Ángel Caído requieren acordonamiento inmediato antes de la apertura. Las zonas en naranja (Alto) implican inspección urgente durante las primeras dos horas. Las zonas en verde (Bajo) siguen supervisión rutinaria. El flujo operativo genera este mapa automáticamente a primera hora de la mañana.

La lógica de construcción tiene una interpretación operativa directa: si el IDR del día es bajo, el IRC de todas las zonas del parque será bajo independientemente de su susceptibilidad histórica — un día tranquilo es tranquilo en todo el parque —. Si el IDR es alto, el IRC amplifica las diferencias entre zonas: una zona de susceptibilidad máxima puede alcanzar $IRC = 0,844$ (Crítico) mientras que una zona de mínima susceptibilidad obtendrá $IRC = 0,169$ (Bajo) ante el mismo IDR. El mapa cambia cada día, pero la estructura espacial subyacente — qué zonas son más vulnerables — permanece estable. Para el día de máximo IDR del período analizado, el protocolo operativo resultante habría sido: acordonar de forma inmediata 20,46 hectáreas de zona Crítica antes de la apertura del parque al público, realizar inspección urgente en 40,92 hectáreas de zona Alta durante las primeras dos horas, y mantener vigilancia reforzada en las 20,46 hectáreas de zona Moderada. Las 36,12 hectáreas restantes habrían seguido el protocolo de supervisión rutinaria.

RECOMENDACIONES OPERATIVAS

Esta sección traduce los resultados analíticos en propuestas concretas de actuación para el Ayuntamiento de Madrid. Las recomendaciones se organizan en tres niveles: cambios en los umbrales de activación del sistema actual, integración del IRC en la gestión diaria de campo, y hoja de ruta para el escalado del sistema a largo plazo.

Protocolo de umbrales estacionales

La recomendación más inmediata no es sustituir el sistema de alertas meteorológicas del Ayuntamiento, sino añadir el IDR_M2 como capa complementaria de información. El protocolo propuesto combina ambas señales: la alerta meteorológica de AEMET indica cuándo hay un detonante potencial; el IDR_M2 indica si el arbolado está en condiciones de soportarlo. Cuando ambas señales coinciden — alerta meteorológica activa e IDR por encima del umbral estacional — la probabilidad de incidencia es máxima y justifica el cierre preventivo. Cuando solo se activa la alerta meteorológica sin IDR elevado, el protocolo propone inspección reforzada en lugar de cierre total.

El sistema actual opera con un umbral implícito de aproximadamente 0,503 durante los doce meses del año. Este valor fue calculado empíricamente como la mediana del IDR_M0 —el modelo basado únicamente en viento— en los días en que el Ayuntamiento activó alerta NARANJA o ROJA. Este valor no describe un umbral programado explícitamente en ningún protocolo municipal, sino la señal de viento retrospectivo que estadísticamente coincide con las decisiones de alerta del Ayuntamiento. Su constancia a lo largo del año — invierno 0,504, primavera 0,503, verano 0,502, otoño 0,508, con una variación de apenas 0,006 — no es el resultado de una calibración deliberada sino de la ausencia de un mecanismo que incorpore el estado fisiológico estacional del arbolado en la decisión de cierre.

Los umbrales calibrados del IDR_M2 corrigen esta limitación adaptando el nivel de activación al ciclo fenológico del arbolado:

- (i) **Invierno (enero, febrero, diciembre) — umbral 0,30:** En invierno el arbolado caducifolio ha perdido la hoja, reduciendo significativamente su carga estructural y su resistencia al viento. La probabilidad de vuelco ante una racha de igual intensidad es menor que en verano. El umbral actual de 0,503 es excesivamente alto para esta estación, lo que paradójicamente genera más cierres innecesarios: el sistema activa la alerta ante rachas de viento que en invierno no representan un riesgo real comparable al de otras estaciones. El resultado empírico es que el 50 % de los cierres invernales del período analizado resultaron evitables. Reducir el umbral a 0,30 en invierno no significa ser menos cauteloso — significa ser más preciso, activando la alerta únicamente cuando el conjunto de condiciones acumuladas justifica realmente el cierre.
- (ii) **Primavera (marzo, abril, mayo) — umbral 0,28:** La primavera es la estación de mayor sensibilidad estructural del arbolado. Durante la brotación, el árbol destina una parte importante de sus reservas energéticas al desarrollo foliar, lo que reduce temporalmente su resiliencia estructural. Un árbol en plena brotación que ha sufrido un invierno de déficit hídrico tiene una vulnerabilidad estructural superior a la que cualquier indicador visual externo permite detectar. El umbral más bajo del año (0,28) refleja esta mayor sensibilidad y garantiza que el sistema no pase por alto episodios de riesgo en una estación en la que el arbolado está fisiológicamente más expuesto.
- (iii) **Verano (junio, julio, agosto) — umbral 0,55:** En verano el arbolado está en plena actividad fotosintética y aclimatado a las temperaturas elevadas. La probabilidad basal de incidencia es más alta que en otras estaciones —hay más días calurosos— pero no todos los días calurosos son igualmente peligrosos. El umbral más alto del año (0,55) filtra los días de calor basal sin estrés acumulado significativo, reduciendo los cierres preventivos innecesarios precisamente en la estación de mayor afluencia de ciudadanos al parque. Es importante subrayar que este umbral elevado no implica menor seguridad: el modelo sigue detectando el 100 % de los vuelcos registrados en verano porque estos ocurren en días con IDR claramente por encima de 0,55, no en días de calor moderado.
- (iv) **Otoño (septiembre, octubre, noviembre) — umbral 0,35:** El otoño es una estación de transición en la que la carga foliar decrece progresivamente. El umbral intermedio de 0,35 refleja la reducción gradual de la vulnerabilidad estructural a medida que avanza la estación, situándose entre el umbral de verano y el de invierno.

La implementación de este protocolo no requiere ningún cambio en la infraestructura de datos existente del Ayuntamiento. Solo requiere que el valor de umbral utilizado para convertir el IDR en una decisión de alerta varíe según el mes del año, siguiendo la tabla anterior.

Integración del IRC en la gestión diaria de campo

Cuando el IDR supere el umbral estacional y se active la alerta general para el parque, el IRC operativo permite pasar de una decisión binaria —cerrar o no cerrar— a una **gestión diferenciada por zonas** que el sistema actual no contempla.

El protocolo propuesto tiene cuatro pasos secuenciales:

- (i) **Paso 1 — Generación automática del IRC (noche anterior o primera hora):** el sistema calcula el IDR_M2 con los datos meteorológicos de AEMET disponibles para el día siguiente. Si el IDR supera el umbral estacional, genera automáticamente el mapa IRC con los cuatro niveles de acción y lo distribuye al equipo de campo en formato digital compatible con GPS de mano.
- (ii) **Paso 2 — Acordonamiento previo a la apertura (zonas Críticas):** antes de abrir el parque al público, los equipos de campo acuden a las zonas con IRC Crítico ($IRC \geq 0,70$, correspondientes a las 20,46 hectáreas de Prioridad 1 del mapa de susceptibilidad) para una inspección ocular rápida y acordonamiento preventivo si se detectan árboles en estado comprometido.

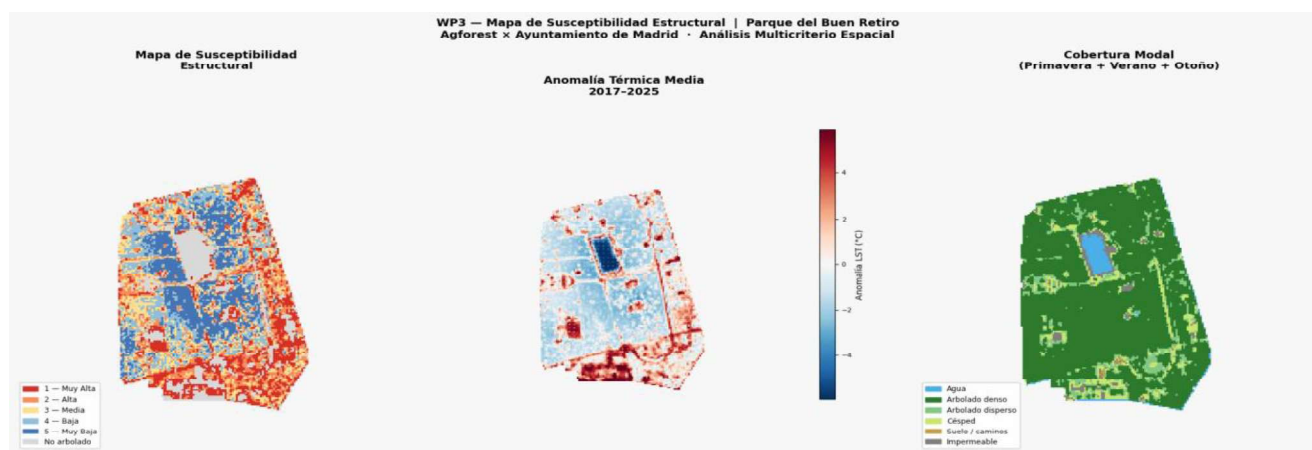


Figura 9. Inputs del Mapa de Susceptibilidad Estructural. Izquierda: mapa de priorización intra-parque. Centro: anomalía térmica media 2017–2026. Derecha: cobertura modal del suelo (primavera + verano + otoño).

- (iii) **Paso 3 — Inspección urgente en las primeras dos horas (zonas Altas):** durante las primeras dos horas de apertura, los equipos realizan la inspección de las 40,92 hectáreas de zona Alta (IRC entre 0,50 y 0,70), priorizando los árboles con historial de intervención o estado sanitario conocido.
- (iv) **Paso 4 — Vigilancia reforzada durante la jornada (zonas Moderadas):** las 20,46 hectáreas de zona Moderada (IRC entre 0,30 y 0,50) reciben una frecuencia de supervisión superior a la habitual durante la jornada, sin requerir acordonamiento previo.

Este protocolo permite mantener el parque parcialmente abierto en días de alerta moderada, acordonando únicamente las zonas de mayor riesgo en lugar de aplicar un cierre total. Esta gestión diferenciada no existe en el sistema actual, que solo contempla dos estados posibles: parque abierto o parque cerrado.

El mapa de cuadrículas de 50×50 metros incluido como adjunto al presente informe identifica con precisión cada zona de acción. Al estar georreferenciado en el sistema de coordenadas UTM zona 30N (EPSG:32630), es directamente navegable desde dispositivos GPS de mano sin necesidad de equipamiento técnico especializado.

LIMITACIONES METODOLÓGICAS

Variables meteorológicas del modelo: diseño retrospectivo deliberado

Una característica metodológica relevante del IDR_M2 que merece explicación explícita es que **ninguna de sus variables de entrada requiere predicciones meteorológicas**. Las diez variables de mayor importancia del modelo son indicadores retrospectivos que describen el estado acumulado del arbolado y las condiciones meteorológicas de los días anteriores al día de predicción. Cuando el sistema calcula el IDR de mañana, todos esos datos ya existen como observaciones reales disponibles en la API de AEMET y en las imágenes satelitales procesadas. No hay ninguna variable que dependa de una estimación del futuro.

Esta característica del modelo — resultado del proceso de selección de *features* sobre el período de entrenamiento 2022-2024 tiene una consecuencia operativa importante: **el IDR_M2 es más robusto en producción que un modelo basado en predicciones meteorológicas**, porque sus inputs son hechos

consumados y no estimaciones sujetas a error de pronóstico. Un modelo que dependiera de la predicción de racha máxima de viento a 24 horas heredaría todos los errores de esa predicción — especialmente relevantes en episodios convectivos donde la localización y el timing del viento son difíciles de anticipar con precisión.

El *dataset* analítico sí contiene columnas de predicción AEMET a 24 horas, pero estas variables sirvieron para un propósito metodológico distinto: evaluar si un modelo alternativo entrenado con predicciones AEMET superaba al actual modelo con viento observado, análisis documentado en el Anexo 7.2.

El análisis empírico de las predicciones AEMET de racha máxima sobre el período 2022–2024 confirma además la solidez de este diseño: las predicciones presentan un RMSE de 15,8 km/h, un sesgo de sobreestimación de +7,9 km/h y una correlación de $r=0,435$ respecto a la racha real observada. Un modelo que dependiera de estas predicciones como variable principal heredaría ese sesgo sistemático. El IDR, al usar exclusivamente variables retrospectivas calculadas sobre datos observados, opera sobre hechos consumados y no sobre estimaciones sujetas a este nivel de error de pronóstico.

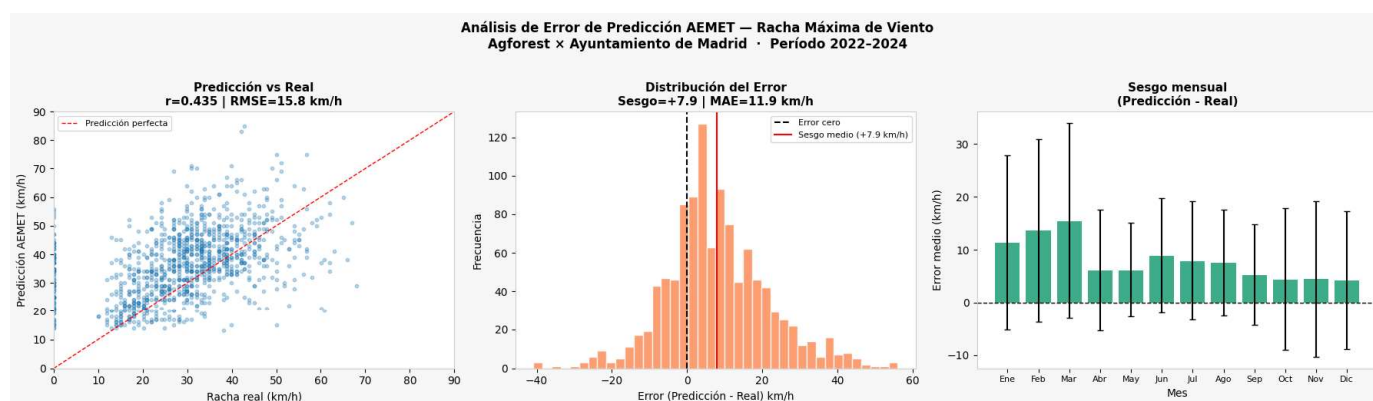


Figura 10. Análisis de error de las predicciones AEMET de racha máxima de viento (2022–2024). Izquierda: scatter predicción vs observación ($r=0,435$). Centro: distribución del error con sesgo de sobreestimación de +7,9 km/h. Derecha: sesgo mensual — la sobreestimación es consistente a lo largo del año.

La unidad de análisis es parque-día, no árbol-día. El modelo predice la probabilidad de que ocurra al menos una incidencia en todo el parque en un día dado, no la probabilidad de fallo de un árbol concreto. Esta agregación espacial es una limitación inherente a la escala de los datos de incidencias disponibles, pero implica que el sistema no puede identificar qué árbol específico va a caer — solo puede alertar sobre el nivel de riesgo global del parque y orientar la inspección hacia las zonas de mayor susceptibilidad mediante el IRC. La transición de la escala parque-día a la escala árbol-día requeriría datos de localización de incidencias y, en una fase más avanzada, integración con el inventario de arbolado municipal y datos de inspección arbórea individualizada.

Variable respuesta binaria sin ponderación por severidad

El modelo IDR fue entrenado para predecir una variable binaria — incidencia sí o no — en lugar de predecir directamente la tipología o severidad del evento. Esta decisión merece una justificación explícita porque puede interpretarse superficialmente como una simplificación excesiva.

El problema del desequilibrio de clases: los vuelcos son cisnes negros estadísticos. En los tres años del período de análisis se registraron 55 días con vuelco o tala de emergencia, frente a 347 días con caída de ramas y 694 días sin incidencia. Si se entrenara un modelo específico para predecir vuelcos, el algoritmo tendría acceso a solo 55 ejemplos positivos — el 5 % del *dataset*. En aprendizaje automático, un desequilibrio de esta magnitud produce modelos que aprenden a predecir siempre "no vuelco" para minimizar el error global, sin capturar ningún patrón real de riesgo. Técnicamente es posible compensar este desequilibrio mediante técnicas de sobre muestreo (SMOTE) o ponderación de clases, pero con 55 eventos distribuidos en tres años y sin localización espacial precisa, el modelo resultante sería estadísticamente frágil y difícilmente validable. La variable binaria agrega todas las tipologías de incidencia en una única señal de riesgo con 402 ejemplos positivos — una base estadística ocho veces mayor y suficientemente robusta para el entrenamiento.

El objetivo no es predecir la física del fallo sino el estado fisiológico de colapso. Un vuelco de árbol completo y una caída de rama de grandes dimensiones son fenómenos mecánicamente distintos, pero

comparten el mismo precursor fisiológico: un arbolado con déficit hídrico acumulado, temperatura superficial persistentemente elevada y reservas energéticas comprometidas que ha alcanzado un umbral de vulnerabilidad estructural crítica. El IDR_M2 no intenta predecir si el fallo será un vuelco o una rama — intenta detectar cuándo el arbolado ha entrado en ese estado de vulnerabilidad que hace cualquier fallo más probable. La variable binaria es la representación correcta de esa señal de estado, porque el estado de colapso estructural es común a cualquier tipología de incidencia independientemente de su manifestación física final.

El IDR_M2, que nunca recibió durante su entrenamiento ninguna información sobre qué días tuvieron vuelcos y qué días tuvieron solo ramas, detectó el 100 % de los 41 vuelcos físicos puros del período. Esto no es una coincidencia estadística — es la evidencia de que el modelo ha aprendido a identificar el patrón de estrés acumulado que precede a cualquier fallo estructural grave, independientemente de su tipología. En términos de aprendizaje automático, esto equivale a un **blind test de transferencia interna**: el modelo fue entrenado con una señal agregada y demuestra capacidad de generalización hacia la categoría más severa sin haberla visto explícitamente. Si el modelo estuviera capturando patrones espurios o ruido estadístico en lugar de señales fisiológicas reales, este resultado sería imposible.

Desde el punto de vista operativo, la variable binaria es exactamente la que el gestor necesita. El responsable municipal que toma la decisión de cierre preventivo no necesita saber si el riesgo del día corresponde a una probabilidad de vuelco o de caída de rama — necesita saber si debe cerrar el parque o mantenerlo abierto. La variable binaria traduce directamente el output del modelo en un **protocolo de acción**: IDR por encima del umbral estacional significa activar el protocolo de seguridad; por debajo, mantener la operación normal. La distinción entre tipologías de incidencia es relevante para la investigación posterior al evento y para el diseño de medidas de mitigación específicas, pero no para la decisión operativa de cierre, que es el caso de uso central de esta Prueba de Concepto.

La transición hacia un modelo de severidad ponderada — donde vuelcos y ramas tengan pesos distintos en la función de pérdida del entrenamiento — es una mejora metodológica propuesta para fases posteriores, condicionada a disponer de un período de análisis más largo que proporcione suficientes ejemplos de eventos de severidad alta para entrenar con robustez estadística.

El modelo captura correlación, no causalidad. El IDR identifica patrones estadísticos entre las variables de estrés acumulado y la ocurrencia de incidencias, pero no establece una relación causal directa entre, por ejemplo, el déficit hídrico de los últimos 30 días y la fractura de una rama concreta. La interpretación correcta es que las variables identificadas como más importantes son buenos predictores estadísticos del riesgo, no necesariamente sus causas directas en todos los casos. La solidez de los resultados — Recall del 99,8 %, coincidencia del 100 % con las decisiones de talas de emergencia — sugiere que el modelo captura señales fisiológicamente relevantes, pero la validación mecanicista requeriría estudios de campo con instrumentación específica sobre árboles individuales.

Mejoras metodológicas propuestas para la siguiente fase

El modelo IDR_M2 sienta las bases para una evolución hacia la **arboricultura de precisión**. Aunque el sistema actual es plenamente funcional, el *roadmap* de mejora identifica dos ejes clave: la integración de **sensores IoT in-situ** para obtener una *ground truth* térmica sin depender de la nubosidad, y la necesidad de contar con una **georreferenciación exacta de cada incidencia** futura. Este último punto permitiría al Random Forest transitar de la probabilidad por zona a la predicción por ejemplar, cerrando el círculo de la seguridad proactiva en el arbolado urbano.

Ordenadas por impacto esperado sobre la calidad del sistema:

- (i) **Incorporar coordenadas GPS de incidencias futuras** para validar y recalibrar el IRC a nivel de árbol individual. Es la mejora de mayor impacto con el menor coste adicional — solo requiere un cambio en el protocolo de registro de incidencias del equipo de campo.
- (ii) **Integrar el inventario de arbolado municipal** (especie, edad, diámetro de tronco, estado sanitario) como capa adicional de vulnerabilidad estructural en el mapa de susceptibilidad. Árboles de mayor edad o con estado sanitario comprometido deberían recibir un peso superior al determinado únicamente por su tipo de cobertura. Esta integración representaría el paso natural de un modelo **biofísico** —basado en clima y señal satelital— a un modelo **biométrico** —basado en las características individuales de cada árbol—, abriendo la puerta al desarrollo de un **Digital Twin del Retiro**: una réplica digital del parque que combina el estado fisiológico en tiempo casi real con

