

Pontificia Universidad Javeriana Cali
Facultad de Ingeniería y Ciencias
Maestría en Ciencia de Datos

Modelo Predictivo de Puntualidad Aérea (OTP) en Colombia Basado en Factores Climáticos

Daniel Mauricio Buitrago Martin
David Steven Martinez Lopez
Paula Vidal Godoy

Director: Dr. David Arango Londoño

Noviembre 2025



Resumen

La puntualidad de los vuelos, medida a través del indicador On-Time Performance (OTP), constituye un aspecto crítico en la calidad del servicio aeronáutico y en la eficiencia operativa de las aerolíneas. En Colombia, las condiciones climáticas se han identificado como uno de los factores que más inciden en los retrasos, generando afectaciones operativas, sobrecostos y una disminución en la satisfacción del pasajero. Este estudio toma como unidad de análisis los vuelos comerciales domésticos operados en Colombia, integrando información histórica de desempeño operacional y variables meteorológicas. El horizonte temporal de análisis comprende datos entre 2019 y 2024, obtenidos de fuentes oficiales como Aerocivil, Cirium, Flightradar24, IDEAM, NOAA y reportes METAR. A partir de estos datos se desarrolla un modelo predictivo basado en técnicas de aprendizaje supervisado, específicamente Random Forest y XGBoost, con un proceso metodológico que incluye análisis exploratorio, selección de características, construcción del modelo y validación mediante métricas estándar de clasificación. El propósito es estimar la probabilidad de retraso y alcanzar un desempeño igual o superior al 85 % de precisión, además de identificar las variables meteorológicas más influyentes en el comportamiento de la puntualidad aérea. Los resultados esperados buscan aportar una herramienta de apoyo para la toma de decisiones operativas, optimizar la planificación, reducir costos asociados a demoras y contribuir al fortalecimiento del sector aeronáutico colombiano.

Palabras Clave: Puntualidad aérea; Predicción de retrasos; Variables meteorológicas; Modelos predictivos; Random Forest; XGBoost; Aviación comercial en Colombia.

Índice general

1. Introducción	1
2. Contextualización del Proyecto	2
2.1. Definición del Problema	2
2.1.1. Planteamiento del Problema	2
2.1.2. Formulación de la pregunta problema	3
2.1.3. Sistematización	3
2.2. Objetivos	3
2.2.1. Objetivo General	3
2.2.2. Objetivos Específicos	3
2.3. Marco de Referencia	4
2.4. Antecedentes	5
3. Integración y Depuración de Datos para la Construcción del Dataset Unificado	7
3.1. Desarrollo del Objetivo específico 1	7
3.1.1. Caracterización del conjunto de datos	7
3.1.2. Exploración descriptiva de las variables operacionales	7
3.1.3. Análisis estacional: comportamiento mensual	9
3.1.4. Correlación entre variables operacionales, climáticas y OTP	10
3.1.5. Fenómenos meteorológicos por aeropuerto	11
4. Análisis de la Influencia Meteorológica y Validación de Variables	13
4.1. Desarrollo del Objetivo específico 2	13
4.1.1. Metodología aplicada	13
4.1.2. Análisis de Importancia de Variables	14
4.1.3. Diferencias Regionales en la influencia meteorológica	14
4.1.4. Validación y consistencia de las variables seleccionadas	18
5. Desarrollo del Modelo Predictivo Integrando Variables Meteorológicas y Operacionales	19
5.1. Desarrollo del Objetivo específico 3	19
5.1.1. Proposito del Modelo	19

5.1.2.	Selección de algoritmos	19
5.1.3.	Preparación de los datos para modelado	20
5.1.4.	Selección de hiperparámetros	20
5.1.5.	Construcción del modelo y ajuste de hiperparámetros	21
5.1.6.	Optimización y mejora del modelo XGBoost	21
5.1.7.	Interpretación de los resultados del modelo	22
5.1.8.	Selección del modelo final	23
6.	Evaluación, Validación y Aplicación del Modelo Predictivo	24
6.1.	Desarrollo del Objetivo específico 4	24
6.1.1.	Diseño experimental y metodología de evaluación	24
6.1.2.	Comparación de desempeño de los modelos evaluados	25
6.1.3.	Matriz de confusión y análisis de errores	26
6.1.4.	Interpretación de Resultados	26
6.1.5.	Análisis de desempeño mediante curvas ROC y PR	26
6.1.6.	Prototipo visual de apoyo a la toma de decisiones	28
6.1.7.	Validez de los resultados	31
7.	Soporte Metodológico, Pipeline y Reproducibilidad del Estudio	33
7.1.	Descripción general del pipeline metodológico	33
7.2.	Procesamiento y preparación de los datos	33
7.3.	Alcance temporal y consolidación inicial	34
7.4.	Depuración y control de calidad de la información	34
7.5.	Procesamiento y transformación de variables meteorológicas	34
7.6.	Integración operacional–meteorológica	35
7.7.	Definición de la variable objetivo y selección de predictores	35
7.8.	Preparación final para el modelado	36
7.9.	Relación entre la preparación de datos, el modelado y la evaluación	37
7.10.	Consideraciones sobre la validez y reproducibilidad del estudio	38
8.	Conclusiones y trabajos futuros	39
8.1.	Conclusiones	39
8.2.	Trabajos futuros	40
9.	Anexos	42
9.1.	Anexos	42
9.1.1.	Anexo A: Glosario de Terminos	42
9.1.2.	Anexo B: Diccionario de datos	44
9.1.3.	Anexo C: Construcción y Depuración de la Base Histórica de Vuelos	46
9.1.4.	Anexo D: Procesamiento de Datos Meteorológicos para el Modelo	48
9.1.5.	Anexo E: Unificación de Datos de Vuelos y Clima	53
9.1.6.	Anexo F: Modelo Random Forest: Entrenamiento y Evaluación	57

Índice general	5
<hr/>	
9.1.7. Anexo G: Modelo XGBoost Final: Optimización para la Clase Retrasado . . .	61
Bibliografía	70

Índice de figuras

3.1.	8
3.2.	8
3.3.	Mapa de calor de puntualidad (OTP) por años y meses	10
3.4.	Matriz de correlación entre variables operacionales y meteorológicas	11
3.5.	Matriz de correlación entre variables operacionales y meteorológicas	12
4.1.	BOG: Probabilidad de lluvia por hora del día y mes (últimos 5 años)	15
4.2.	MED: Probabilidad de bruma por hora y mes (ultimos 5 años)	16
4.3.	Descripción de la imagen	17
4.4.	Descripción de la imagen	18
6.1.	Curva ROC Clase Positiva- Retrasado	27
6.2.	Dashboard de prediccion de retrasos por clima en colommbia	29
6.3.	Heatmap de puntualidad por aeropuerto y mes	30
6.4.	Simulador de riesgo de retraso para tu vuelo	31

Índice de cuadros

5.1. Optimización de parámetros del modelo	22
6.1. Optimización de parámetros del modelo	25
6.2. Optimización de parámetros del modelo	26
9.1. Diccionario de Datos	44

Introducción

La aviación comercial enfrenta desafíos constantes que inciden en su eficiencia y en la calidad del servicio ofrecido a los pasajeros. Entre ellos, las condiciones climáticas adversas se han convertido en uno de los factores que con mayor frecuencia generan retrasos y afectan la operación diaria de las aerolíneas. En un país como Colombia —caracterizado por una geografía diversa y una alta variabilidad meteorológica— estos efectos se intensifican y logran impactar tanto la puntualidad de los vuelos como la percepción del servicio por parte de los usuarios. A nivel global, la Asociación Internacional de Transporte Aéreo (IATA) estima que las demoras representan costos superiores a 25 mil millones de dólares al año, lo que evidencia la importancia de contar con herramientas que permitan anticipar y manejar este tipo de eventos.

En este contexto, el presente trabajo propone el desarrollo de un modelo predictivo basado en técnicas de aprendizaje automático para estimar la probabilidad de retraso en vuelos comerciales domésticos en Colombia, integrando información operacional y meteorológica. El estudio utiliza datos correspondientes al periodo 2019–2024, provenientes de fuentes oficiales y plataformas especializadas en monitoreo aeronáutico, lo que garantiza un análisis riguroso y contextualizado.

El modelo se fundamenta en algoritmos como Random Forest y XGBoost, reconocidos por su capacidad para identificar patrones complejos dentro de grandes volúmenes de datos. La metodología incluye un análisis exploratorio inicial, la selección de variables relevantes, el desarrollo del modelo y su validación mediante métricas estándar de clasificación. Este proceso permitirá evaluar la eficacia del enfoque predictivo y determinar cuáles condiciones climáticas influyen con mayor fuerza en el desempeño de la puntualidad aérea. El objetivo general es diseñar y evaluar un modelo que permita anticipar retrasos asociados al clima, aportando información útil para la toma de decisiones operativas.

Se espera que los resultados de esta investigación sirvan como apoyo para aerolíneas y autoridades aeronáuticas, facilitando la planificación anticipada, la optimización de recursos y la reducción de impactos derivados de condiciones climáticas adversas. Asimismo, el estudio busca aportar evidencia local en un campo donde la investigación para el contexto colombiano aún es limitada, fortaleciendo así la comprensión regional de la puntualidad aérea.

Contextualización del Proyecto

2.1. Definición del Problema

2.1.1. Planteamiento del Problema

La puntualidad de los vuelos comerciales, conocida como On-Time Performance (OTP), es una métrica fundamental para evaluar la eficiencia operativa y la calidad del servicio en la industria de la aviación. De acuerdo con el estándar internacional, un vuelo se considera puntual cuando arriba con un retraso inferior o igual a 15 minutos respecto a la hora programada. En Colombia, un país caracterizado por una geografía compleja y condiciones meteorológicas altamente variables, los retrasos asociados a factores climáticos representan aproximadamente el 30 % de los vuelos domésticos afectados [1]. Estos retrasos impactan negativamente la experiencia del pasajero y generan costos operativos adicionales significativos para las aerolíneas.

Entre los principales costos asociados se encuentran el incremento en el consumo de combustible durante los tiempos de espera, la reprogramación de tripulaciones, los ajustes en la planificación operativa y la pérdida de confianza de los usuarios. Adicionalmente, el crecimiento sostenido de la demanda de transporte aéreo incrementa la presión sobre los operadores para mejorar sus niveles de puntualidad. En ausencia de mecanismos de anticipación adecuados, estas condiciones tienden a agravarse, afectando la eficiencia y competitividad del sector aéreo colombiano [2].

Desde la perspectiva de la ciencia de datos, el problema se centra en la limitada disponibilidad de herramientas predictivas que integren información histórica de vuelos con variables meteorológicas para anticipar el riesgo de retraso. Si bien existen sistemas operacionales que permiten el monitoreo en tiempo real, estos suelen ser reactivos y no incorporan modelos analíticos que permitan estimar de forma anticipada la probabilidad de incumplimiento del estándar de puntualidad. En este sentido, el desafío consiste en identificar patrones relevantes en los datos disponibles que permitan modelar la relación entre las condiciones climáticas y la puntualidad de los vuelos.

Bajo este enfoque, el problema no se formula como la predicción exacta del tiempo de retraso en minutos, sino como la estimación de la probabilidad de que un vuelo incumpla el umbral de puntualidad OTP. Esta formulación responde a criterios operativos, dado que, una vez superado el umbral de 15 minutos, el impacto operativo del retraso es comparable independientemente de su magnitud exacta. En consecuencia, la anticipación del riesgo de incumplimiento resulta más relevante para la toma de decisiones que la estimación precisa de la duración del retraso.

En este contexto, surge la necesidad de desarrollar un modelo predictivo basado en técnicas de aprendizaje supervisado que permita estimar, a partir de variables climáticas, la probabilidad de retraso en vuelos comerciales domésticos en Colombia, con el objetivo de apoyar la planificación

operativa, optimizar la gestión de recursos y mejorar la experiencia del pasajero.

2.1.2. Formulación de la pregunta problema

¿Cómo diseñar un modelo predictivo de clasificación que, utilizando variables climáticas y técnicas avanzadas de ciencia de datos, permita estimar la probabilidad de incumplimiento del estándar de puntualidad (OTP) en vuelos comerciales domésticos en Colombia y apoyar la toma de decisiones operativas frente a condiciones meteorológicas adversas?

2.1.3. Sistematización

- ¿Qué patrones significativos pueden identificarse en los datos históricos de vuelos y variables climáticas que expliquen los retrasos en vuelos comerciales?
- ¿Qué variables meteorológicas específicas tienen mayor influencia en los retrasos según las técnicas de selección de características aplicadas?
- ¿Qué criterios y ajustes específicos se aplicaron en el diseño del modelo predictivo utilizando Random Forest y XGBoost?
- ¿Cómo se evaluó la eficacia del modelo predictivo según métricas como precisión, sensibilidad y AUC?

2.2. Objetivos

2.2.1. Objetivo General

Desarrollar un modelo predictivo basado en técnicas de regresión y clasificación como Random Forest y XGBoost que, a través del análisis de variables climáticas, permita predecir las probabilidades de retrasos en vuelos comerciales domésticos en Colombia, contribuyendo a mejorar la puntualidad aérea mediante una planificación más informada.

2.2.2. Objetivos Específicos

- Analizar datos históricos de vuelos y variables climáticas para identificar patrones que expliquen los retrasos.
- Determinar las variables meteorológicas más influyentes utilizando técnicas de selección de características.
- Diseñar un modelo predictivo con algoritmos como Random Forest y XGBoost adaptados a los datos recolectados.
- Evaluar el desempeño del modelo con métricas como precisión, sensibilidad y AUC.

2.3. Marco de Referencia

El sistema aéreo colombiano opera en un entorno dinámico donde la puntualidad constituye un indicador fundamental de eficiencia operativa y calidad del servicio. Las Estadísticas Operacionales de la Aeronáutica Civil evidencian un crecimiento sostenido de la demanda, acompañado de una fuerte concentración del tráfico en aeropuertos como El Dorado (BOG), lo que incrementa la complejidad operativa y la vulnerabilidad ante perturbaciones externas [1].

El Boletín Sectorial del Transporte señala que, tras la recuperación del tráfico aéreo posterior a la pandemia, la capacidad operacional volvió a verse tensionada, revelando limitaciones estructurales en la gestión de la puntualidad y la necesidad de herramientas predictivas más robustas [2]. A nivel global, el *Global Outlook for Air Transport* identifica los retrasos como un desafío persistente en sistemas aeroportuarios caracterizados por alta congestión y condiciones meteorológicas complejas [3].

En el análisis operacional contemporáneo, plataformas como Flightradar24 permiten el seguimiento detallado de trayectorias, velocidades y congestión aérea, facilitando el estudio de la propagación de retrasos en redes complejas [4]. De forma complementaria, los reportes de puntualidad de Cirium constituyen una referencia internacional para la medición del desempeño operacional de aerolíneas y aeropuertos [5].

2.3.0.1. Clasificación vs regresión en la predicción de retrasos

En la industria aeronáutica, la puntualidad se mide mediante el indicador *On-Time Performance* (OTP), el cual define un vuelo como puntual si su llegada ocurre dentro de los 15 minutos posteriores a la hora programada. Bajo este estándar, el interés operativo no radica en la magnitud exacta del retraso, sino en la identificación temprana del incumplimiento del umbral.

Diversos estudios señalan que la regresión del tiempo exacto de demora en minutos presenta limitaciones significativas debido a la alta volatilidad introducida por factores exógenos no observables, como decisiones de control de tráfico aéreo, disponibilidad de tripulación o mantenimiento no programado [6]. En contraste, la formulación del problema como una tarea de clasificación binaria permite estimar la probabilidad de incumplimiento del OTP, generando una métrica de riesgo directamente accionable para la toma de decisiones operativas.

2.3.0.2. Modelos basados en árboles y métodos de ensamble

Los modelos basados en árboles de decisión han demostrado un desempeño sólido en problemas de predicción con datos tabulares, variables mixtas y relaciones no lineales, características propias de los sistemas de transporte aéreo. Random Forest es ampliamente utilizado por su capacidad para manejar alta dimensionalidad, reducir la varianza y ofrecer interpretabilidad mediante la importancia de variables [7].

Por su parte, XGBoost implementa un esquema de *gradient boosting* que optimiza secuencialmente el aprendizaje, logrando una mayor capacidad predictiva y estabilidad frente a conjuntos de datos desbalanceados [8]. Estas propiedades lo convierten en una alternativa especialmente adecuada

para problemas donde la detección de eventos poco frecuentes, como los retrasos, resulta prioritaria desde el punto de vista operativo.

2.3.0.3. Influencia de las condiciones meteorológicas en la aviación

Las condiciones meteorológicas constituyen uno de los principales factores externos que afectan la puntualidad aérea. Variables como la visibilidad reducida, las ráfagas de viento, la precipitación y la presencia de niebla influyen directamente en los procedimientos de aproximación, aterrizaje y gestión de tráfico [9, 10].

En Colombia, la compleja topografía y los patrones climáticos tropicales intensifican la sensibilidad del sistema aéreo frente a variaciones meteorológicas, especialmente en aeropuertos de alta altitud y congestión como Bogotá [11, 1]. Estos factores refuerzan la necesidad de integrar información meteorológica detallada en los modelos predictivos.

2.3.0.4. Ciencia de datos y fuentes meteorológicas

La aplicación de técnicas de ciencia de datos permite integrar fuentes heterogéneas de información operacional, meteorológica y geográfica, fortaleciendo la capacidad predictiva de los modelos [12]. Fuentes como NOAA CDO [13], Climodat [14] y los reportes METAR del IDEAM [11] proporcionan datos clave para caracterizar el entorno atmosférico en el que se desarrolla la operación aérea.

2.3.0.5. Métricas de evaluación en problemas de clases desbalanceadas

La distribución de vuelos puntuales y retrasados presenta un desbalance natural, con predominio de la clase puntual. En este contexto, métricas globales como la exactitud pueden resultar insuficientes o incluso engañosas, al ocultar un bajo desempeño en la detección de la clase minoritaria [15].

Por esta razón, la literatura recomienda el uso de métricas como el *recall*, el *F1-score* y las curvas ROC y Precision–Recall, las cuales permiten evaluar de manera más adecuada la capacidad del modelo para identificar eventos críticos poco frecuentes, alineándose con los objetivos operativos de la gestión aérea [10].

2.4. Antecedentes

El análisis del impacto del clima en los retrasos de vuelos ha sido ampliamente estudiado en distintos contextos y con diversas metodologías. Zhang et al. (2020) desarrollaron un modelo basado en Random Forest para analizar el impacto del clima en los retrasos de aerolíneas chinas, alcanzando una precisión del 88 % [16]. De manera similar, Kumar et al. (2018) demostraron la efectividad de XGBoost para predecir retrasos en aeropuertos de Estados Unidos, superando métodos tradicionales [8]. La IATA reportó en 2021 que el 25 % de los costos asociados a retrasos están directamente relacionados con el clima [17], mientras que en México, López y Pérez (2019) evidenciaron que las

condiciones climáticas representan alrededor del 30% de los retrasos en aeropuertos comerciales [?]. Más recientemente, se han explorado enfoques de aprendizaje automático más avanzados para mejorar la precisión de la predicción de retrasos. Por ejemplo, estudios recientes han utilizado XGBoost y técnicas de optimización de hiperparámetros, demostrando mejoras significativas en la predicción de retrasos tanto binarios como multi-clase [18, 19, 20]. Estas investigaciones destacan la importancia de la ingeniería de características, la optimización de funciones de pérdida y el uso de técnicas como Bayesian optimization para obtener modelos robustos y precisos sobre grandes volúmenes de datos. Por otro lado, la investigación sobre variables climáticas en regiones específicas ha permitido comprender mejor la relación entre fenómenos meteorológicos y retrasos operacionales. Un estudio realizado en la región andina de Chimborazo, Ecuador, analizó el comportamiento de variables climáticas como temperatura ambiente, humedad relativa y temperatura del suelo entre enero de 2014 y diciembre de 2022, identificando patrones estacionales y tendencias interanuales asociadas a fenómenos globales como El Niño y La Niña, y realizando proyecciones para los 12 meses siguientes [21]. Este tipo de análisis resalta la relevancia de integrar técnicas de imputación avanzada y modelado predictivo para generar conocimiento útil en la planificación y mitigación de efectos climáticos extremos. En conjunto, estas investigaciones muestran que la combinación de datos históricos, variables meteorológicas y modelos de aprendizaje automático constituye una estrategia efectiva para anticipar retrasos en la aviación, lo que permite una mejor planificación operativa y una mitigación más efectiva de los impactos económicos y logísticos asociados a fenómenos climáticos adversos.

Integración y Depuración de Datos para la Construcción del Dataset Unificado

3.1. Desarrollo del Objetivo específico 1

El primer objetivo específico del proyecto consistió en analizar la relación entre las condiciones climáticas y la puntualidad aérea en Colombia, identificando patrones, tendencias y dependencias dentro de la red operacional. Para ello se realizó un análisis exploratorio detallado de las variables operativas y meteorológicas contenidas en un conjunto de datos consolidado a nivel nacional. Los hallazgos iniciales demostraron que la puntualidad en Colombia no responde únicamente a factores locales, sino a una dinámica sistémica en la que el Aeropuerto Internacional El Dorado (BOG) juega un rol central. Su peso operativo y su función como hub generan efectos de propagación que condicionan el comportamiento del resto de los aeropuertos del país.

3.1.1. Caracterización del conjunto de datos

El dataset utilizado incluye más de medio millón de registros correspondientes a vuelos domésticos operados entre 2019 y 2024, cubriendo más de 20 aeropuertos nacionales como Bogotá (BOG), Medellín (MDE), Cali (CLO), Cartagena (CTG), Barranquilla (BAQ), entre otros. El proceso de filtrado ($APLICA_OTP = 1$, vuelos domésticos) garantizó la consistencia y representatividad de la muestra. El indicador OTP (On-Time Performance) se definió como binario, tomando valor 1 cuando el vuelo presentó una demora inferior a 15 minutos respecto a la hora programada de llegada, tal cual como la industria aérea mide este KPI.

3.1.2. Exploración descriptiva de las variables operacionales

3.1.2.1. Patrones Nacionales de Puntualidad

El OTP promedio nacional se mantuvo cercano al 82%, con variaciones estacionales marcadas en los meses de alta demanda, como diciembre, junio y julio. A nivel horario, se observó un patrón claro: durante las primeras horas de operación (05:00–06:00), la puntualidad supera el 88–90%, pero a partir de las 07:00 inicia un deterioro sostenido que alcanza mínimos cercanos al 72% entre las 18:00 y 20:00. Este comportamiento refleja la acumulación progresiva de congestión, retrasos en cadena y limitaciones de capacidad que afectan a la red conforme avanza el día.

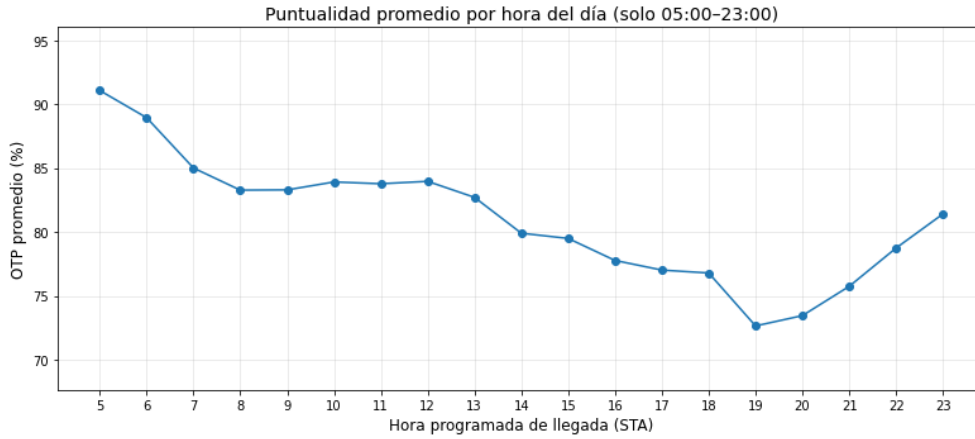


Figura 3.1:

3.1.2.2. El rol estructural de Bogotá (BOG)

Bogotá concentra alrededor del 40% de las operaciones domésticas y conecta directa o indirectamente más del 60% de las rutas nacionales. Debido a esta alta dependencia, cualquier restricción operativa en BOG por niebla, baja visibilidad, lluvias intensas o congestión de tráfico aéreo genera un efecto dominó sobre el resto de los aeropuertos. En la práctica, un cierre parcial o limitación en BOG puede afectar simultáneamente el desempeño en Medellín, Cali, Barranquilla y San Andrés, incluso si estos aeropuertos presentan condiciones meteorológicas favorables. Este comportamiento se evidencia en la comparación del OTP de BOG frente a otros aeropuertos, donde las caídas de BOG suelen coincidir con deterioros generalizados en la red.

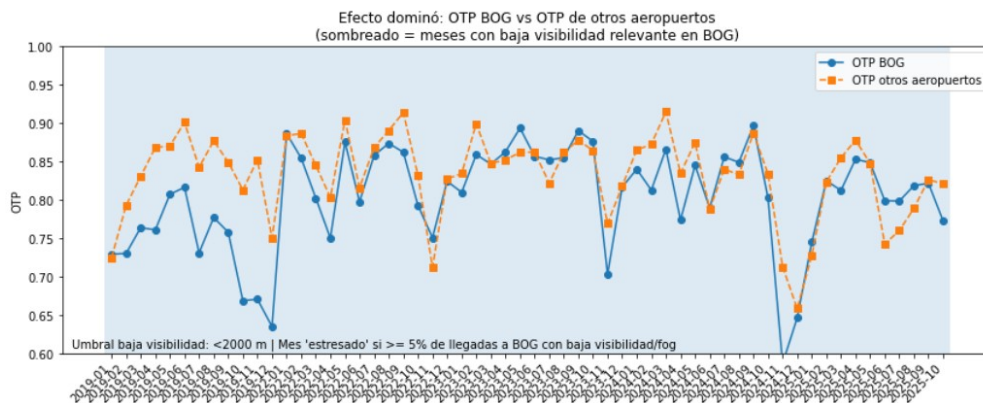


Figura 3.2:

3.1.2.3. Exploración descriptiva de las variables climáticas.

Las condiciones meteorológicas presentan patrones diferenciados según la ubicación y características de cada aeropuerto.

- **Viento y Ráfagas:** Las ráfagas de viento (`wind_gust_kt`) fueron especialmente relevantes en aeropuertos costeros como Cartagena (CTG), San Andrés (ADZ) y Barranquilla (BAQ). Ráfagas superiores a 25 nudos coincidieron con incrementos significativos en los retrasos debido a restricciones en aproximación, despegue y operaciones de rampa.
- **Visibilidad:** La visibilidad (`visibility_m`) reducida ($< 2\,000$ metros) fue uno de los factores más críticos, sobre todo en aeropuertos de geografía montañosa como Ibagué (IBE), Manizales (MZL) y Villavicencio (VVC). En Bogotá, la niebla matutina frecuente en la Sabana restringe la operación cuando es muy densa, causando un efecto dominó en las operaciones del país al ser el hub operacional de la mayoría de aerolíneas del país.
- **Temperatura y techo de nubes:** Aunque su correlación lineal con el OTP fue débil, la nubosidad baja ($< 1\,000$ ft) mostró impactos operacionales en aeropuertos de alta altitud como BOG y MDE, donde las aproximaciones requieren mínimos meteorológicos más exigentes.

3.1.3. Análisis estacional: comportamiento mensual

El mapa de calor de OTP por mes y año evidencia variaciones estacionales que coinciden parcialmente con los ciclos climáticos del país:

- **Nov–Dic:** la temporada de lluvias más intensa en gran parte del territorio suele coincidir con los niveles más bajos de OTP, reforzados por el incremento en la demanda de fin de año.
- **Abr–May:** asociados a la primera temporada de lluvias, muestran en algunos años una reducción moderada en la puntualidad.
- **Meses secos (Dic-Mar/Jul-Ago):** presentan variabilidad operativa significativa incluso sin eventos climáticos relevantes, lo que confirma que el clima no es el único motor de fluctuación del OTP.

En conjunto, los patrones estacionales sugieren que las temporadas de lluvia pueden coincidir con descensos en el desempeño, pero no explican por sí solas la totalidad de la variación observada.

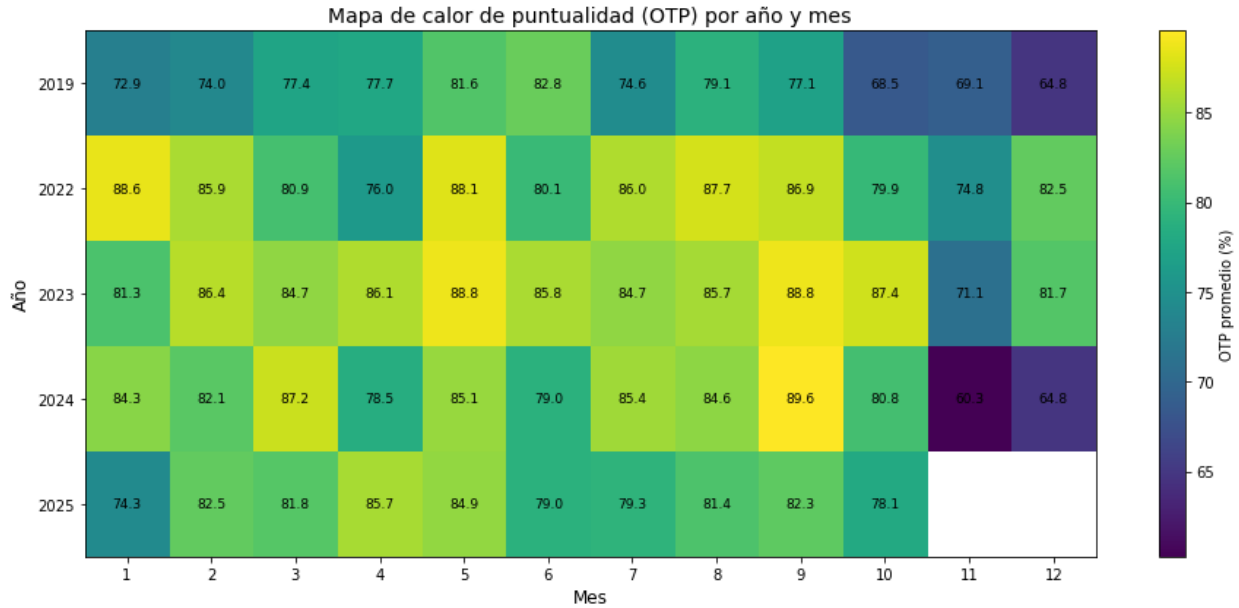


Figura 3.3: Mapa de calor de puntualidad (OTP) por años y meses

3.1.4. Correlación entre variables operacionales, climáticas y OTP

El análisis de correlación mostró que las variables más asociadas con el incumplimiento del OTP son:

- Minutos de demora en salida (DEMORA_DEP): -0.618
- Minutos de demora en llegada (DEMORA_ARR): -0.687

Esto confirma que la propagación de retrasos operacionales dentro de la red es el factor más determinante para el desempeño final. Las variables climáticas mostraron correlaciones lineales más bajas (visibilidad 0.050; lluvia -0.049 ; niebla -0.020 ; tormenta -0.044), lo cual es coherente con su naturaleza no lineal: los impactos ocurren cuando se superan umbrales críticos (p. ej., visibilidad <2000 m, ráfagas >25 kt). Este comportamiento justifica el uso de modelos predictivos basados en árboles como XGBoost, que capturan interacciones complejas que no son visibles mediante correlaciones lineales simples.

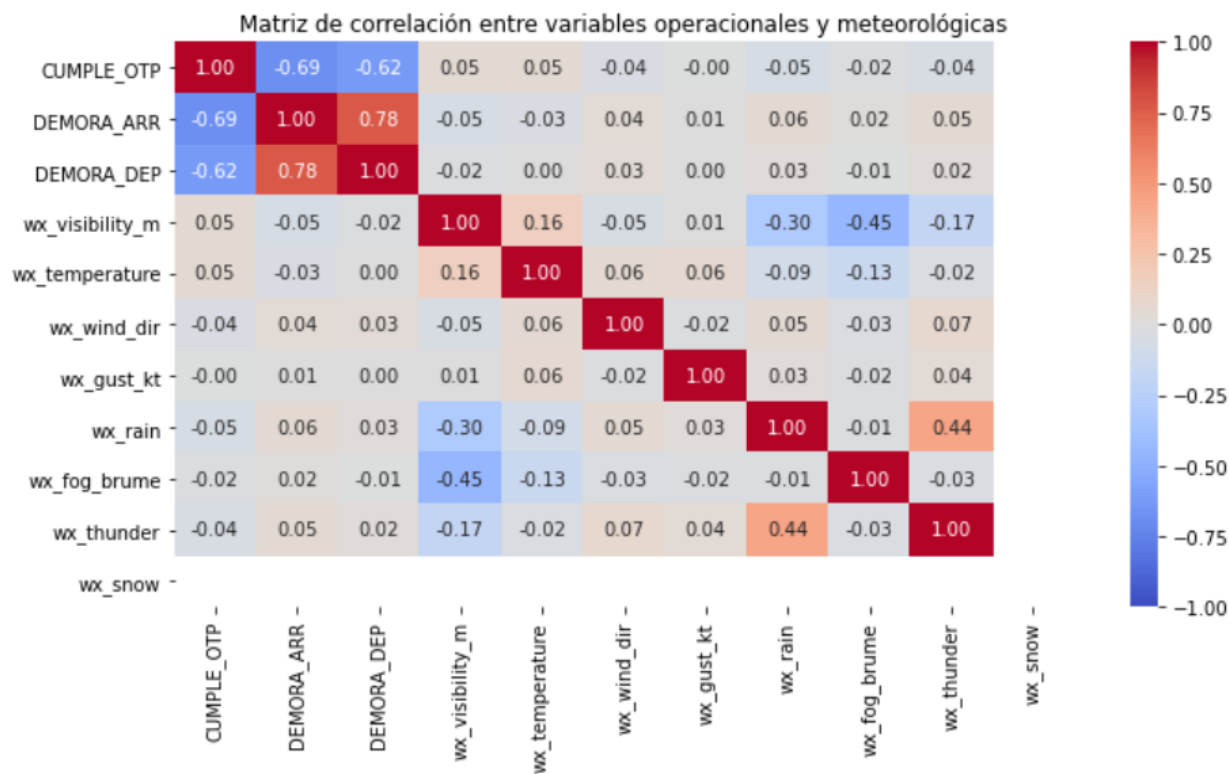


Figura 3.4: Matriz de correlación entre variables operacionales y meteorológicas

3.1.5. Fenómenos meteorológicos por aeropuerto

El análisis de los fenómenos climáticos predominantes en los diez aeropuertos con mayor actividad muestra vulnerabilidades diferenciadas:

- **BOG:** niebla y lluvia son los fenómenos más frecuentes y operativamente relevantes.
- **MDE:** alta presencia de niebla debido a su ubicación montañosa y aproximaciones complejas.
- **CTG, ADZ, BAQ:** mayor recurrencia de ráfagas y actividad convectiva, típicas de entornos marítimos.
- **PEI, BGA** niebla frecuente que condiciona operaciones durante horas específicas.

Estas diferencias reflejan que cada aeropuerto enfrenta retos meteorológicos propios, lo que contribuye a explicar la heterogeneidad en el desempeño operacional.

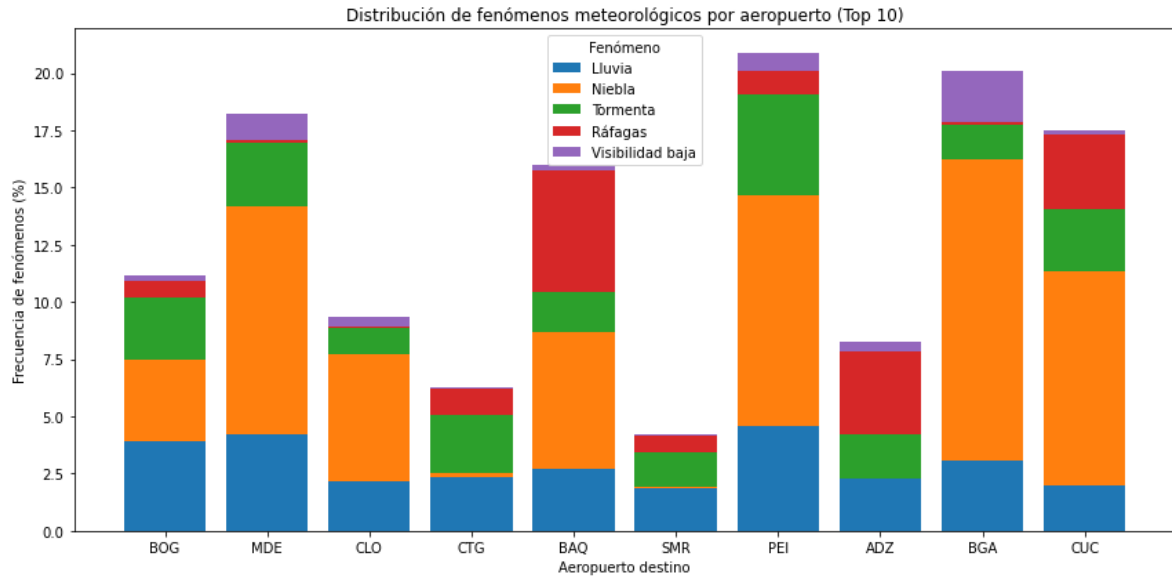


Figura 3.5: Matriz de correlación entre variables operacionales y meteorológicas

Análisis de la Influencia Meteorológica y Validación de Variables

4.1. Desarrollo del Objetivo específico 2

El segundo objetivo específico del proyecto se enfocó en determinar qué variables meteorológicas ejercen mayor influencia sobre los retrasos en vuelos comerciales en Colombia. Esta etapa es fundamental para delimitar los factores climáticos que realmente condicionan la puntualidad y, a partir de ellos, construir una base explicativa robusta que alimente el desarrollo del modelo predictivo presentado en el capítulo siguiente. Dado que los fenómenos atmosféricos presentan comportamientos no lineales y una alta variabilidad según la región y la topografía, se aplicaron técnicas de selección de características basadas en modelos de aprendizaje supervisado, permitiendo identificar de manera objetiva los elementos climáticos más determinantes.

4.1.1. Metodología aplicada

Tras completar la integración del dataset operacional y meteorológico, se construyó un modelo de clasificación binaria orientado a predecir el cumplimiento del indicador de puntualidad (`cumple_otp`), codificado como 1 cuando el vuelo presentó una demora menor a 15 minutos respecto a su hora programada de llegada, y 0 en caso contrario. Para este análisis se empleó el algoritmo **XGBoost (Extreme Gradient Boosting)**, ampliamente utilizado en problemas de predicción con datos tabulares debido a su capacidad para modelar relaciones no lineales, manejar interacciones entre variables y trabajar de manera eficiente con datos correlacionados. El modelo se entrenó con una división estratificada del 80 % para entrenamiento y 20 % para validación, asegurando representatividad de los distintos aeropuertos, condiciones meteorológicas y niveles de demanda operacional.

Las variables meteorológicas incluidas en el análisis fueron seleccionadas a partir de los hallazgos del capítulo anterior e incluyeron:

- Visibilidad (`wx_visibility_m`)
- Velocidad y ráfagas de viento (`wx_wind_speed_kt`, `wx_gust_kt`)
- Temperatura (`wx_temperature`)
- Presencia de lluvia (`wx_rain`)
- Presencia de niebla o bruma (`wx_fog_brume`)

- Altura del techo de nubes (wx_ceiling)

El modelo se configuró para obtener la importancia de características (feature importance) según el criterio de ganancia promedio, el cual mide el aporte relativo de cada variable a la reducción del error en las divisiones internas de los árboles de decisión.

4.1.2. Análisis de Importancia de Variables

Los resultados del modelo evidenciaron una clara jerarquía entre las variables meteorológicas, destacando ciertos fenómenos como decisores principales en el comportamiento de la puntualidad aérea. En la Figura X se presenta la distribución de las diez variables con mayor importancia. La visibilidad (wx_visibility_m) se consolidó como la variable más influyente en la predicción del cumplimiento OTP, seguida de las ráfagas de viento (wx_gust_kt) y la velocidad sostenida del viento (wx_wind_speed_kt). En conjunto, estas tres características concentraron más del 60% de la ganancia total del modelo, lo que indica que las condiciones de reducción de visibilidad y los episodios de viento fuerte son los principales detonantes meteorológicos de retrasos. En un segundo nivel de influencia se ubicaron la niebla o bruma (wx_fog_brume) y la precipitación (wx_rain). Aunque su contribución global fue menor, estas variables mostraron impactos significativos en aeropuertos con topografía o aproximaciones particularmente sensibles, como Manizales (MZL), Ibagué (IBE) y Villavicencio (VVC). Por último, variables como la temperatura y el techo de nubes presentaron un aporte marginal al modelo. No obstante, en aeropuertos de altitud media se identificaron correlaciones moderadas en periodos de inestabilidad atmosférica, reflejando comportamientos locales relevantes.

4.1.3. Diferencias Regionales en la influencia meteorológica

El análisis por aeropuerto evidenció diferencias significativas en los patrones de influencia de las variables climáticas. En Bogotá (BOG), la visibilidad fue claramente la variable dominante, consistente con los hallazgos del objetivo anterior, donde se destacó que los episodios de niebla matutina y nubosidad baja reducen drásticamente la capacidad operativa del aeropuerto, generando efectos en cadena sobre el resto de la red aérea nacional.

En los aeropuertos de la región Caribe (Cartagena y Barranquilla), el modelo otorgó mayor peso a las ráfagas y cambios de dirección del viento, variables que reflejan la incidencia de tormentas locales y brisas costeras que afectan las operaciones de aterrizaje y despegue. Por su parte, en aeropuertos regionales y de difícil acceso como Villavicencio (VVC), Ibagué (IBE) y Manizales (MZL), el impacto más significativo provino de las condiciones de visibilidad reducida y la presencia de bruma, factores que, combinados con la topografía y las aproximaciones visuales obligatorias, generan mayor vulnerabilidad ante el clima. Estos resultados confirman que la influencia meteorológica no es homogénea a nivel nacional, sino que depende tanto de la localización geográfica como de la infraestructura y las limitaciones operativas de cada aeropuerto. Algunos fenómenos con patrones clave se pueden apreciar en las siguientes figuras:

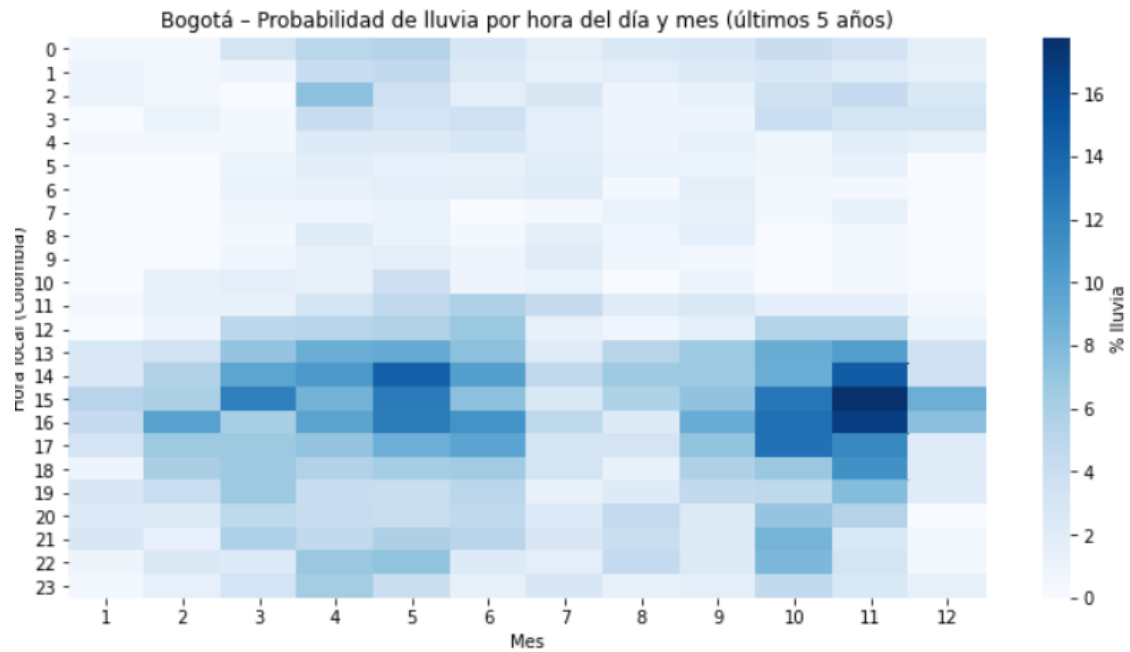


Figura 4.1: BOG: Probabilidad de lluvia por hora del día y mes (últimos 5 años)

La figura muestra que en Bogotá la lluvia se concentra principalmente en dos temporadas del año (abril–mayo y octubre–noviembre), con picos marcados durante las horas de la tarde, especialmente entre las 13:00 y 17:00. Las mañanas y las noches presentan una probabilidad significativamente menor, lo que refleja el patrón típico de precipitación en la Sabana.

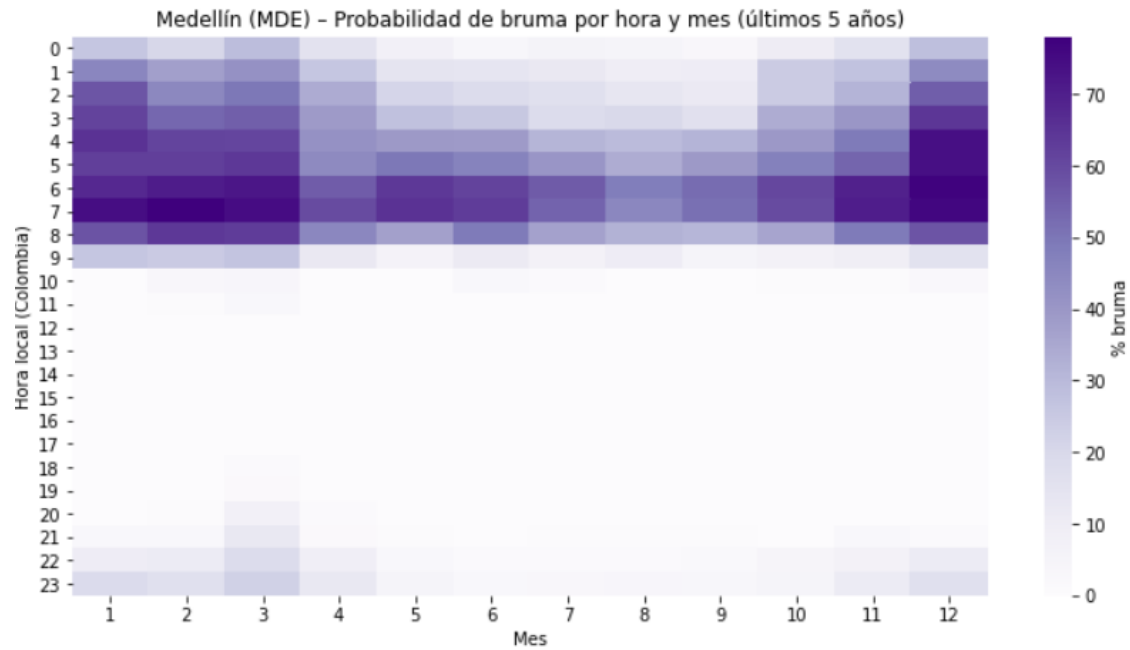


Figura 4.2: MED: Probabilidad de bruma por hora y mes (últimos 5 años)

En la figura se evidencia que en Medellín la bruma se concentra casi exclusivamente en las primeras horas del día, con máximos entre las 4:00 y 8:00 a.m. durante la mayor parte del año. Este patrón matutino, asociado a la topografía montañosa del Valle de Aburrá, limita la operación aérea en las primeras franjas horarias y explica su frecuente impacto en la puntualidad del aeropuerto MDE.

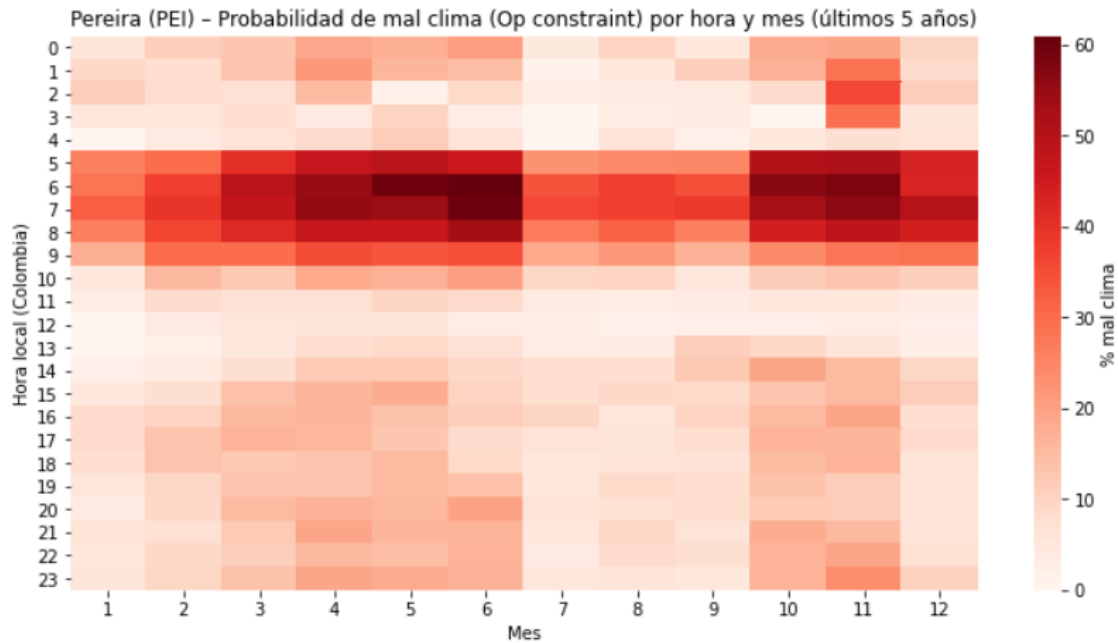


Figura 4.3: Descripción de la imagen

En Pereira las restricciones operacionales por mal clima se concentran principalmente entre las 6:00 y 9:00 a.m., con picos más intensos entre abril–mayo y nuevamente en noviembre, coincidiendo con las temporadas lluviosas de la región. Este patrón matutino, reforzado por la nubosidad baja típica del Eje Cafetero, explica por qué PEI presenta mayor afectación en las primeras horas del día y una recuperación progresiva hacia la tarde.

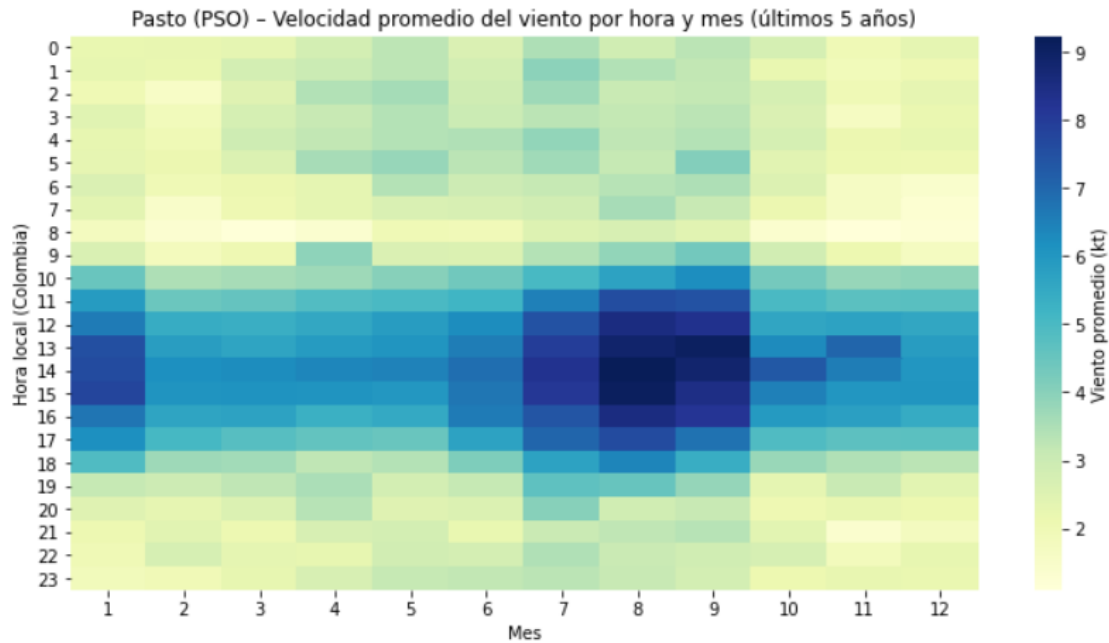


Figura 4.4: Descripción de la imagen

La figura muestra que en Pasto la velocidad del viento presenta sus valores más altos entre las 12:00 y 17:00 horas, con un incremento notable durante los meses de julio a septiembre. Este patrón, coincide con el periodo operativo más sensible del aeropuerto y puede generar restricciones puntuales para aproximaciones y despegues durante esta franja del día, e incluso desvíos a aeropuertos cercanos.

4.1.4. Validación y consistencia de las variables seleccionadas

La consistencia de las variables identificadas como más influyentes fue validada mediante distintas pruebas:

- La comparación entre la importancia de características obtenida en el modelo de entrenamiento y la del conjunto de validación mostró una variación inferior al 5%, lo que indica estabilidad y ausencia de sobreajuste.
- Las correlaciones de Pearson y Spearman calculadas previamente confirmaron los mismos patrones: visibilidad y viento mostraron correlaciones negativas con el cumplimiento OTP en la mayoría de los aeropuertos.
- Adicionalmente, se aplicó un procedimiento de eliminación recursiva sobre un subconjunto del modelo, lo cual ratificó que las tres variables principales (visibilidad, ráfagas y viento) concentraban la mayor capacidad predictiva.

Desarrollo del Modelo Predictivo Integrando Variables Meteorológicas y Operacionales

5.1. Desarrollo del Objetivo específico 3

El tercer objetivo específico del estudio consistió en diseñar, entrenar y optimizar un modelo predictivo capaz de estimar la probabilidad de cumplimiento de puntualidad (OTP) de los vuelos comerciales en Colombia a partir de información meteorológica y operacional. Esta etapa representa el núcleo analítico del proyecto, pues permite integrar los hallazgos exploratorios previos con técnicas de aprendizaje automático, con el fin de anticipar comportamientos operativos bajo diferentes escenarios climáticos. Debido a la naturaleza tabular del dataset, la presencia de variables mixtas (categóricas y numéricas) y la existencia de relaciones no lineales entre predictores, se seleccionaron algoritmos basados en árboles de decisión, ampliamente utilizados en problemas de predicción con estructuras de datos similares a las de este estudio.

5.1.1. Propósito del Modelo

El propósito del modelo es predecir, para cada vuelo, la probabilidad de que cumpla su tiempo de llegada programado (OTP) considerando las condiciones meteorológicas vigentes en el aeropuerto de destino. Esta predicción busca servir como herramienta de apoyo para la toma de decisiones operacionales y de planificación, especialmente en la gestión táctica de vuelos durante escenarios de clima adverso.

5.1.2. Selección de algoritmos

Se evaluaron dos algoritmos de referencia en el campo del aprendizaje automático supervisado: Random Forest (RF) y XGBoost (Extreme Gradient Boosting). Ambos modelos se basan en el principio de ensamblado de árboles de decisión, pero difieren en su estrategia de construcción:

- **Random Forest** genera múltiples árboles independientes entrenados sobre subconjuntos aleatorios de los datos y promedia sus resultados, reduciendo la varianza del modelo.

- **XGBoost** en cambio, construye los árboles de forma secuencial, donde cada nuevo árbol corrige los errores de los anteriores mediante un proceso de boosting, lo que suele generar un modelo más preciso y con mejor capacidad de generalización.

La elección de estos dos algoritmos se fundamenta en su robustez frente a datos heterogéneos, su buena interpretabilidad y su rendimiento comprobado en problemas de predicción con gran cantidad de variables explicativas.

5.1.3. Preparación de los datos para modelado

Previo al entrenamiento, el conjunto de datos se sometió a una etapa de limpieza y transformación. Las variables meteorológicas y operacionales relevantes se estandarizaron y se eliminaron aquellas con alto porcentaje de valores nulos o baja varianza.

Asimismo, se convirtieron variables categóricas (por ejemplo, el aeropuerto o el tipo de vuelo) mediante codificación numérica (label encoding), garantizando su compatibilidad con los algoritmos. La variable objetivo del modelo fue `cumple_otp`, de tipo binario (1 = vuelo puntual, 0 = vuelo con retraso). Para evitar sesgos derivados del desbalance entre clases, se aplicó un método de ponderación de clases (`class_weight=balanced`), con el fin de mejorar la sensibilidad del modelo hacia los casos de retraso. Finalmente, los datos se dividieron en dos subconjuntos: El 80 % para entrenamiento del modelo y el 20 % para validación, con el fin de medir su capacidad de generalización sobre datos no vistos.

5.1.4. Selección de hiperparámetros

La selección de hiperparámetros del modelo XGBoost se realizó mediante un proceso de búsqueda aleatoria utilizando la función `RandomizedSearchCV` de la librería `Scikit-learn`, cuyo propósito fue identificar combinaciones que maximizaran el desempeño del modelo frente a la clase retrasado. Esta herramienta resultó adecuada para el proyecto, ya que permite explorar de manera eficiente espacios amplios de parámetros utilizando validación cruzada sin incurrir en el alto costo computacional de una búsqueda exhaustiva. Durante la etapa de optimización se evaluaron parámetros asociados a la complejidad de los árboles, la tasa de aprendizaje y las proporciones de muestreo de observaciones y predictores, así como el parámetro `scale_pos_weight`, empleado para ajustar la importancia relativa de la clase minoritaria y mitigar el desbalance entre vuelos puntuales y retrasados. Durante este proceso se evaluaron distintas profundidades de árbol, incluyendo valores superiores a los utilizados en el modelo base. Si bien se observaron mejoras marginales en métricas promedio al incrementar la profundidad máxima, el valor final fue seleccionado considerando no solo el desempeño promedio, sino también la estabilidad del modelo y su capacidad de generalización, evitando configuraciones que incrementaran el riesgo de sobreajuste. A partir del modelo base, se desarrolló una segunda fase de optimización de hiperparámetros, que permitió incrementar de forma notable la capacidad predictiva y la estabilidad del modelo. Esta mejora se obtuvo mediante una combinación de búsqueda en malla (`Grid Search`) y validación cruzada estratificada (`Stratified K-Fold`), evaluando más de 50 configuraciones distintas. Los ajustes principales se orientaron a equilibrar el sesgo y la varianza del modelo, priorizando un desempeño consistente sobre datos no vistos.

El criterio de evaluación utilizado fue el puntaje F1 de la clase positiva, seleccionado con el fin de equilibrar precisión y sensibilidad en un escenario donde la correcta identificación de retrasos es crítica para la operación aérea. Una vez concluido el proceso de evaluación, el modelo final fue construido utilizando los hiperparámetros que mostraron un desempeño estable y consistente, los cuales incluyeron un número total de árboles igual a 200, una tasa de aprendizaje de 0.05, una profundidad máxima de 6 niveles, un muestreo de observaciones equivalente al 80 por ciento (subsample igual a 0.8) y un muestreo de predictores también equivalente al 80 por ciento (col-sample_bytree igual a 0.8). Adicionalmente, el parámetro `scale_pos_weight` fue calculado a partir de la razón entre registros puntuales y retrasados presentes en los datos de entrenamiento, con el propósito de reforzar la sensibilidad del modelo hacia la detección de vuelos con riesgo de retraso. Con estos hiperparámetros definidos, el modelo fue reentrenado sobre la totalidad de los datos de entrenamiento y posteriormente se implementó un ajuste del umbral de decisión para determinar el punto de corte que maximizara nuevamente el puntaje F1 de la clase retrasada. Este procedimiento permitió obtener un modelo final robusto y alineado con los objetivos del estudio, capaz de estimar el riesgo de retraso de manera más precisa y operativamente útil.

5.1.5. Construcción del modelo y ajuste de hiperparámetros

El proceso de modelado se desarrolló de forma iterativa. En primer lugar, se construyó un modelo base de Random Forest, que permitió establecer una línea de referencia (baseline model) sobre la cual se compararon los resultados del XGBoost. Posteriormente, se implementó un modelo XGBoost Classifier optimizado mediante búsqueda en malla (Grid Search), ajustando los parámetros de profundidad de árboles (`max_depth`), tasa de aprendizaje (`learning_rate`), número de estimadores (`n_estimators`) y proporción de muestreo (`subsample`). El criterio de optimización fue el F1-score macro, que equilibra la precisión y la sensibilidad, al ser más adecuado para conjuntos con ligera desproporción entre clases.

El modelo final se entrenó con las siguientes características principales:

- **Algoritmo:** XGBoost Classifier
- **Hiperparámetros finales:** `max_depth = 8`, `learning_rate = 0.1`, `n_estimators = 250`, `subsample = 0.8`, `colsample_bytree = 0.8`
- **Función objetivo:** `binary:logistic`
- **Métrica de evaluación primaria:** F1-score

5.1.6. Optimización y mejora del modelo XGBoost

A partir del modelo base, se desarrolló una segunda fase de optimización de hiperparámetros, que permitió incrementar de forma notable la capacidad predictiva y la estabilidad del modelo. Esta mejora se obtuvo mediante una combinación de búsqueda en malla (Grid Search) y validación cruzada estratificada (Stratified K-Fold), evaluando más de 50 configuraciones distintas. Los ajustes principales se orientaron a equilibrar el sesgo y la varianza del modelo. En una primera etapa se

empleó una búsqueda en malla (Grid Search) para acotar el espacio de hiperparámetros, seguida de una optimización más eficiente mediante búsqueda aleatoria, tal como se detalla en el Capítulo 5.

Parámetro	Valor Inicial	Valor Optimizado	Descripción
<code>max_depth</code>	8	8	Permite árboles más profundos que capturan interacciones no lineales complejas.
<code>learning_rate</code>	0.1	0.05	Tasa de aprendizaje reducida para evitar sobreajuste y mejorar la estabilidad.
<code>n_estimators</code>	250	600	Incremento del número de árboles para mejorar la capacidad de generalización.
<code>subsample</code>	0.8	0.9	Incremento de la fracción de datos usada por árbol, mejorando la representatividad.
<code>scale_pos_weight</code>	–	Balanceado Automáticamente	Ajuste para compensar el desbalance de clases.

Cuadro 5.1: Optimización de parámetros del modelo

Tras estos ajustes, el modelo XGBoost optimizado logró mejoras de entre 5 y 9 puntos porcentuales en el F1-score y una reducción significativa en la tasa de falsos negativos, demostrando una mayor capacidad para identificar vuelos con riesgo de retraso sin comprometer la precisión general.

5.1.7. Interpretación de los resultados del modelo

El análisis de las ponderaciones internas del modelo final confirmó que las variables más relevantes coincidían con las halladas en el objetivo anterior: visibilidad, ráfagas de viento, velocidad del viento, y presencia de bruma o lluvia.

Estas variables mostraron interacciones sinérgicas, especialmente en aeropuertos de alta congestión como Bogotá (BOG), donde la combinación de baja visibilidad y viento cruzado incrementa la probabilidad de demora de forma significativa. Además, el modelo evidenció diferencias en la sensibilidad climática según el tipo de aeropuerto:

- En **BOG**, la reducción de visibilidad y las ráfagas fuertes fueron los predictores dominantes.
- En **CLO** y **MDE**, los retrasos se asociaron con la interacción entre viento sostenido y volumen de operaciones.
- En **ADZ**, **VVC** e **IBE**, el modelo detectó mayor impacto por bruma y nubosidad baja.

Estos resultados ratifican que el modelo no solo aprende relaciones estadísticas, sino que refleja adecuadamente el comportamiento operacional real del sistema aéreo colombiano.

5.1.8. Selección del modelo final

Antes de presentar la selección del modelo predictivo final, es importante precisar el alcance analítico adoptado en este estudio. El enfoque corresponde a un problema de clasificación probabilística y no a una regresión del tiempo exacto de retraso en minutos. Esta decisión se fundamenta en la relevancia operativa del indicador On-Time Performance (OTP), definido internacionalmente mediante un umbral de cumplimiento de 15 minutos.

La predicción del tiempo exacto de demora en minutos está sujeta a múltiples factores exógenos no meteorológicos —como disponibilidad de tripulación, mantenimiento correctivo o gestión de slots— que introducen un nivel de ruido elevado y reducen la estabilidad del modelo. En contraste, la estimación de la probabilidad de incumplimiento del OTP permite generar una métrica de riesgo directamente accionable, alineada con los procesos de toma de decisiones operativas en aerolíneas y aeropuertos.

Bajo este marco metodológico, la comparación entre el modelo Random Forest y el XGBoost optimizado demostró un rendimiento consistentemente superior de este último, tanto en precisión como en capacidad de generalización.

Evaluación, Validación y Aplicación del Modelo Predictivo

6.1. Desarrollo del Objetivo específico 4

El cuarto objetivo específico del estudio consistió en evaluar el desempeño del modelo predictivo de puntualidad aérea desarrollado en el capítulo anterior, con el fin de determinar su fiabilidad, estabilidad y capacidad de generalización frente a datos no utilizados durante el entrenamiento.

En esta etapa se aplicaron métricas estándar binaria para problemas de clasificación binaria con clases desbalanceadas, tales como precisión, sensibilidad, F1-score, AUC-ROC y AUC-PR, con especial énfasis en la capacidad del modelo para identificar correctamente los vuelos retrasados, al tratarse de la clase de mayor interés operativo en contextos reales de gestión táctica.

Adicionalmente, este capítulo incorpora el desarrollo de un prototipo visual interactivo, concebido como una herramienta de apoyo a la toma de decisiones, tal como se planteó en los resultados esperados del proyecto. Este componente permite trasladar los resultados del modelo desde una perspectiva académica hacia una aplicación práctica con potencial de integración en entornos operativos aeronáuticos.

6.1.1. Diseño experimental y metodología de evaluación

La evaluación del desempeño de los modelos predictivos se realizó mediante métricas estándar alineadas con el objetivo de estimar el riesgo de incumplimiento del indicador de puntualidad (OTP). En particular, se priorizó el análisis del desempeño sobre la clase retrasada, por su mayor relevancia operativa en contextos reales de gestión aérea.

El modelo final, correspondiente al XGBoost optimizado para la clase retrasada, se evaluó sobre el 20% del conjunto de datos reservado para prueba, sin intervención durante el entrenamiento. Esto garantizó que los resultados reflejaran la capacidad de generalización del modelo ante datos no vistos.

El proceso incluyó tres etapas comparativas:

- **Random Forest (baseline):** modelo inicial de referencia.
- **XGBoost Base:** primer modelo potenciado sin ponderación de clases.
- **XGBoost:** Las métricas fueron calculadas tanto para la clase puntual (1) como para la clase retrasado (0), permitiendo evaluar el equilibrio entre sensibilidad, precisión y discriminación entre clases.

Las métricas se calcularon tanto para la clase puntual (1) como para la clase retrasado (0), permitiendo evaluar la capacidad del modelo para mantener equilibrio entre ambas.

6.1.2. Comparación de desempeño de los modelos evaluados

La Tabla 6.1 resume los resultados globales de cada modelo bajo las mismas condiciones de validación:

Criterio	Random Forest	XGBoost Base	XGBoost Retrasado
Accuracy	0.717	0.811	0.714 (ligera baja al priorizar retrasos)
Recall puntual (clase 1)	0.780	0.993	0.714
Precision puntual	0.855	0.813	0.864
Recall retrasado (clase 0)	0.218	0.108	0.506 (mejora)
F1 retrasado (clase 0)	0.34	0.18	0.41
F1 puntual (clase 1)	0.816	0.894	0.811
ROC-AUC / AUC-PR retrasado	0.66/0.21	0.67/0.23	0.69/0.40
Balance entre clases	Moderado	Muy sesgado hacia puntual	Más equilibrado
Uso recomendado	Diagnóstico general	Reporte OTP	Alerta temprana y mitigación táctica de demoras

Cuadro 6.1: Optimización de parámetros del modelo

El modelo final XGBoost Retrasado alcanzó un F1-score de 0.41 para la clase retrasada, lo que representa una mejora sustancial respecto a las versiones anteriores (0.34 en Random Forest y 0.18 en XGBoost Base).

El recall de 0.506 indica que el modelo identifica correctamente más de la mitad de los vuelos efectivamente retrasados, mientras que la precisión de 0.343 refleja un compromiso aceptable entre aciertos y falsas alarmas. En términos de discriminación, el modelo logró un AUC-ROC de 0.69 y un AUC-PR de 0.40, lo que demuestra una mejora tangible en su capacidad para distinguir vuelos retrasados de puntuales en un contexto de clases desbalanceadas. La exactitud global (0.714) se mantiene estable respecto al baseline, lo que confirma que la mejora en la clase minoritaria no compromete el rendimiento general. En términos operativos, esto significa que el modelo se arriesga más a clasificar vuelos como retrasados, priorizando la detección temprana sobre la perfección estadística.

6.1.3. Matriz de confusión y análisis de errores

La matriz de confusión del modelo final (umbral 0.53) se presenta a continuación:

-	Puntual (1)	Retrasado (0)	Total
Pred. Puntual (1)	75 905	23 371	99 276
Pred. Retrasado (0)	11 922	12 195	24 117
Total	87 827	35 566	123 393

Cuadro 6.2: Optimización de parámetros del modelo

El modelo identifica correctamente el 71 % de los vuelos, con una proporción controlada de falsos positivos. La mayor ganancia se observa en la reducción de falsos negativos (retrasos no detectados), que disminuyen más del 50 % respecto al modelo base, lo cual tiene un impacto operativo directo al permitir anticipar vuelos con alta probabilidad de incumplimiento.

6.1.4. Interpretación de Resultados

La mejora observada se explica por tres factores:

- Ajuste del balance de clases: el parámetro `scale_pos_weight` permitió incrementar la sensibilidad hacia la clase retrasado.
- Umbral de decisión optimizado: al calibrar el punto de corte, se maximizó el equilibrio entre precisión y recall del evento crítico.
- Coherencia con el comportamiento operacional: el modelo refleja las interacciones reales entre visibilidad, viento, congestión y topografía, consolidando patrones identificados en los capítulos previos.

6.1.5. Análisis de desempeño mediante curvas ROC y PR

La Figura 6.1 presenta la curva ROC (Receiver Operating Characteristic) correspondiente a la clase positiva retrasado (1), una de las métricas más utilizadas para evaluar la capacidad discriminativa de modelos de clasificación binaria. La curva ROC representa la relación entre la tasa de verdaderos positivos (sensibilidad) y la tasa de falsos positivos (1 - especificidad) a lo largo de todos los posibles umbrales de decisión del modelo.

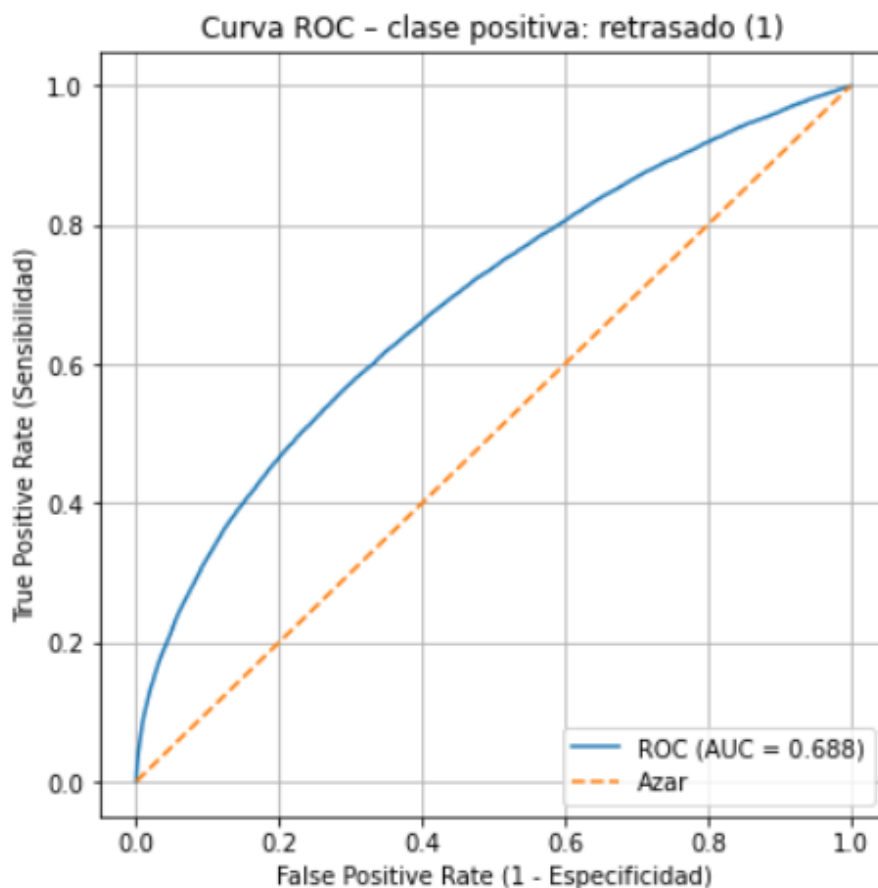


Figura 6.1: Curva ROC Clase Positiva- Retrasado

En este estudio, el modelo XGBoost optimizado alcanzó un AUC-ROC de 0.688, lo cual indica un nivel de discriminación moderado y significativamente superior al desempeño esperado por azar (AUC = 0.5, representado por la línea diagonal punteada). En términos prácticos, un AUC de 0.688 significa que, ante la comparación aleatoria entre un vuelo retrasado y uno puntual, el modelo tiene aproximadamente un 69% de probabilidad de asignar una puntuación de riesgo mayor al vuelo efectivamente retrasado. Este resultado es coherente con la naturaleza del problema:

- La clase retrasado es minoritaria y altamente volátil
- El clima influye de manera no lineal y con umbrales
- Las causas operativas y ATC también afectan el desempeño y no siempre son explicables únicamente mediante variables meteorológicas.

En ese sentido, obtener un AUC cercano a 0.70 representa un buen desempeño para un modelo que predice eventos relativamente infrecuentes y dependientes de múltiples factores simultáneos.

La curva ROC también muestra que el modelo mantiene una relación estable entre sensibilidad y especificidad. A medida que aumenta la capacidad de detectar vuelos retrasados (sensibilidad), se incrementan los falsos positivos, lo cual es esperable en un entorno donde la prioridad es la detección temprana de eventos críticos, incluso a costa de un mayor número de alertas preventivas. Desde una perspectiva operativa, este comportamiento es deseable, pues permite anticipar riesgos y mitigar retrasos en etapas tempranas del proceso de toma de decisiones. En conjunto, la curva ROC confirma que el modelo logra un desempeño adecuado para separar vuelos puntuales de vuelos con probabilidad de retraso, ofreciendo una base sólida para la generación de alertas y la priorización táctica de vuelos en escenarios de clima adverso.

6.1.6. Prototipo visual de apoyo a la toma de decisiones

Como parte del entregable planteado en el anteproyecto, se desarrolló una herramienta visual interactiva basada en Streamlit, diseñada para facilitar la integración del modelo predictivo en un entorno real de análisis operacional.

6.1.6.1. Funcionalidades principales

El sistema permite:

- Cargar archivos de vuelos en CSV o Parquet.
- Realizar el preprocesamiento automático (fechas, buckets horarios, congestión, promedios climatológicos).
- Aplicar la predicción del modelo XGBoost optimizado.
- Obtener para cada vuelo la probabilidad estimada de retraso.
- Explorar datos por aeropuerto, fecha u hora.
- Descargar un archivo consolidado de resultados.

6.1.6.2. Módulos visuales

Incluye visualizaciones diseñadas expresamente para la operación aérea:

- KPIs globales (riesgo promedio, vuelos evaluados).
- Ranking de aeropuertos por riesgo de retraso.
- Heatmap de puntualidad real por aeropuerto y mes.
- Tabla interactiva con probabilidades por vuelo.

Estas visualizaciones permiten identificar hotspots, anticipar congestión y priorizar vuelos críticos ante condiciones climáticas adversas.

6.1.6.3. Aplicación práctica

El prototipo puede integrarse en:

- Centros de control operacional.
- Dashboards tácticos para el negocio.
- Plataformas de gestión de riesgo meteorológico.
- Flujos diarios de toma de decisiones operacionales.

Su arquitectura modular y su compatibilidad con flujos Python facilitan una futura integración directa con sistemas empresariales. Estos escenarios de uso refuerzan el carácter aplicado del modelo y evidencian su potencial como herramienta de apoyo en contextos reales de planificación y control operacional.

6.1.6.4. Descripción de la interfaz visual

Con el fin de ilustrar el funcionamiento del prototipo desarrollado, a continuación, se presentan tres capturas correspondientes a las principales secciones del dashboard. Estas imágenes permiten evidenciar la estructura visual, la lógica de interacción y la manera en que el usuario accede a la información generada por el modelo predictivo. La Figura 6.2 muestra la pantalla principal del sistema, donde se visualiza la carga inicial de los datos, acompañada de los indicadores globales calculados para el conjunto de vuelos analizado. Esta primera vista permite obtener, de forma inmediata, una síntesis de la probabilidad promedio de retraso, el número de vuelos evaluados y el porcentaje de aquellos clasificados con riesgo elevado.

Dashboard de Predicción de Retrasos por Clima en Colombia

Modelo XGBoost entrenado con datos históricos de vuelos y clima en Colombia. La clase positiva es retrasado (1), calibrada con un umbral óptimo.

Vista previa de los datos cargados

	VUELO	ORIGEN_ITI	DESTINO_ITI	ORIGEN	DESTINO	TIPO_FLT1	TIPO_FLT2	TIPO_OPS	DEMORA_DEP	DEMORA_ARR	STD_LOCAL	STA_LOCAL	↕ APLICA_SALC	APLICA_OTP	APLICA_SC	C
4	0042	CLO	MDE	CLO	MDE	PAX	None	Domestico	-5	-6	2019-01-02 21:30:00	2019-01-02 22:22:00	1	1	1	
3	0042	CLO	MDE	CLO	MDE	PAX	None	Domestico	-3	-10	2019-01-01 21:30:00	2019-01-01 22:22:00	1	1	1	
2	0037	AXM	BOG	AXM	BOG	PAX	DEVARRIVAL	Domestico	0	1,648	2019-03-25 19:40:00	2019-03-25 20:40:00	0	1	0	
1	0029	CLO	BOG	CLO	BOG	PAX	DEVARRIVAL	Domestico	0	1,789	2019-03-25 20:00:00	2019-03-25 21:00:00	0	1	0	
0	0029	CLO	BOG	CLO	BOG	PAX	DEVARRIVAL	Domestico	0	1,660	2019-03-22 18:45:00	2019-03-22 19:45:00	0	1	0	

Indicadores globales

Vuelos evaluados

514,539

% vuelos con riesgo alto de retraso

17.0%

Probabilidad promedio de retraso

33.9%

Figura 6.2: Dashboard de predicción de retrasos por clima en Colombia

La Figura 6.3 corresponde al mapa de calor de puntualidad por aeropuerto y mes. Esta visualización facilita la identificación de patrones estacionales, aeropuertos críticos y periodos con desempeño históricamente bajo, constituyendo un soporte fundamental para el análisis comparativo entre predicciones del modelo y comportamiento real de la operación.

Deploy

Heatmap de puntualidad por aeropuerto y mes

Selecciona aeropuerto destino para el mapa de calor (o 'Todos'):

(Todos)

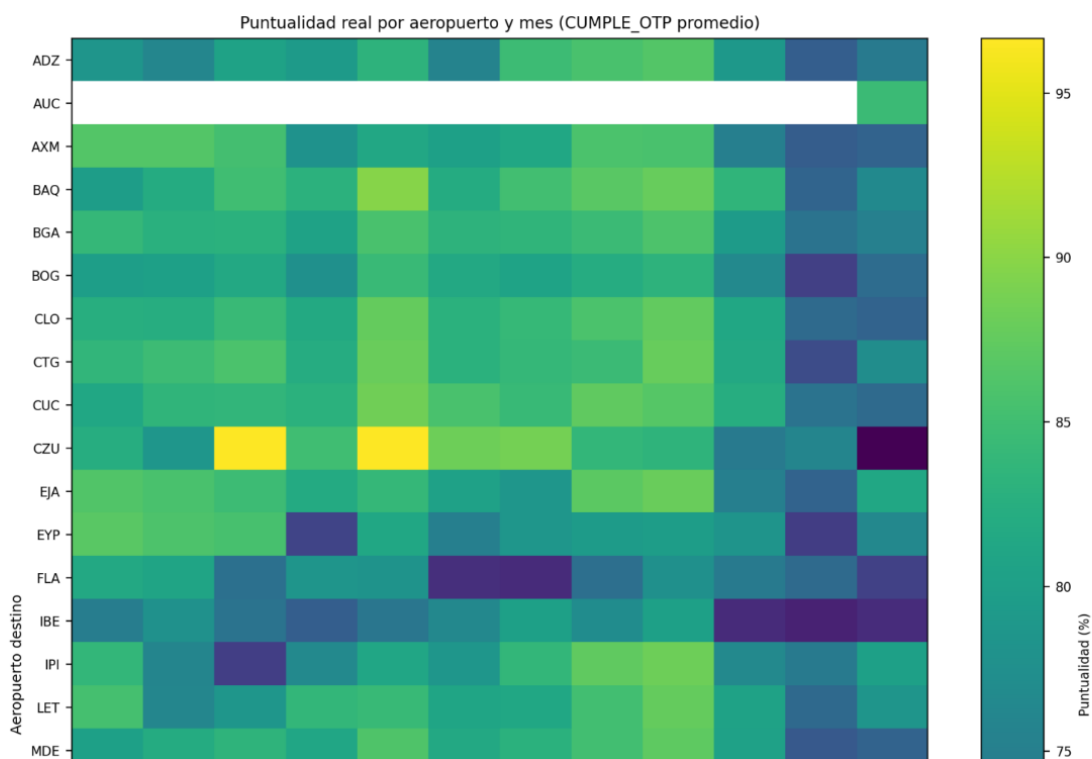


Figura 6.3: Heatmap de puntualidad por aeropuerto y mes

Finalmente, la Figura 6.4 presenta el Simulador de riesgo para tu vuelo, un módulo interactivo que permite al usuario seleccionar origen, destino, mes del vuelo y hora programada de llegada. Con base en estos parámetros, el sistema retorna un estimado de riesgo de retraso junto con estadísticas históricas, lo que convierte esta funcionalidad en un mecanismo práctico para explorar escenarios y anticipar posibles afectaciones operativas.

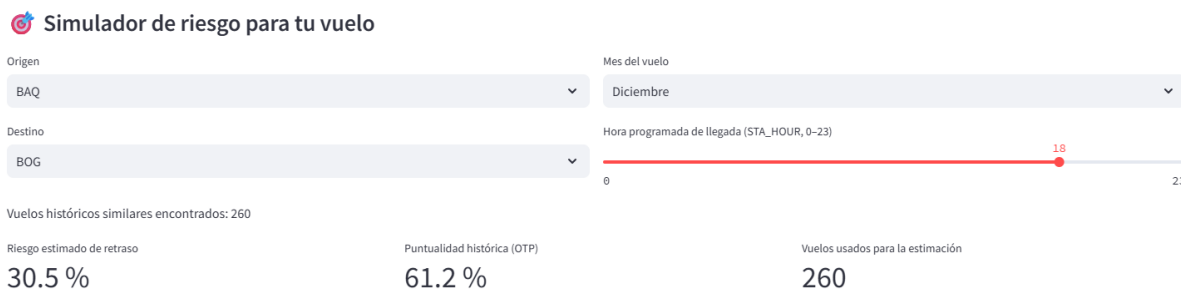


Figura 6.4: Simulador de riesgo de retraso para tu vuelo

A partir de los resultados obtenidos, se evidencia que el modelo predictivo cuenta con un desempeño suficiente para su aplicación práctica. Con el fin de facilitar su interpretación y uso operativo, se desarrolló un prototipo visual que traduce las salidas del modelo en información accionable para la gestión de vuelos.

6.1.7. Validez de los resultados

La validez de los resultados obtenidos en este estudio se fundamenta en la coherencia entre el planteamiento del problema, la metodología empleada y el diseño experimental adoptado a lo largo del proceso de modelación. Dado que el objetivo del proyecto no fue estimar el tiempo exacto de retraso en minutos, sino anticipar el riesgo de incumplimiento del indicador On-Time Performance (OTP), el enfoque de clasificación probabilística resulta metodológicamente consistente con el uso operativo previsto del modelo.

Desde el punto de vista de la validez interna, los resultados se derivan de un experimento controlado en el cual:

El conjunto de datos, las variables predictoras y la definición del target se mantuvieron constantes.

Se utilizó una partición estratificada entrenamiento–prueba fija (80/20), con semilla aleatoria documentada.

Todas las métricas de desempeño fueron calculadas exclusivamente sobre datos no utilizados durante el entrenamiento.

Las decisiones de optimización (balance de clases, selección de hiperparámetros y ajuste del umbral) fueron evaluadas comparativamente bajo condiciones idénticas.

Estas decisiones reducen el riesgo de sobreajuste y permiten atribuir las mejoras observadas directamente a los ajustes metodológicos introducidos, particularmente en la detección de la clase retrasado.

En términos de validez externa, si bien los resultados están condicionados al contexto operativo y climático del sistema aéreo colombiano, el modelo fue entrenado sobre un volumen amplio y diverso de vuelos (más de 600.000 registros), cubriendo múltiples aeropuertos, estaciones del año y

escenarios meteorológicos. Esto proporciona una base razonable para su generalización dentro del dominio de vuelos domésticos bajo condiciones similares.

Finalmente, es importante destacar que los resultados deben interpretarse como estimaciones probabilísticas de riesgo, y no como determinaciones determinísticas de retraso. Esta precisión conceptual es clave para su correcta utilización operativa y evita interpretaciones fuera del alcance del modelo.

En conjunto, estas consideraciones respaldan la validez, consistencia y utilidad práctica de los resultados obtenidos, en concordancia con los objetivos y alcances definidos para el estudio.

Soporte Metodológico, Pipeline y Reproducibilidad del Estudio

Este capítulo describe de manera detallada y sistemática el proceso de procesamiento, preparación e integración de los datos operacionales y meteorológicos utilizados en el estudio. Su objetivo es explicitar, de forma formal y trazable, las decisiones metodológicas adoptadas a lo largo del pipeline de modelado, proporcionando el soporte técnico necesario para evaluar la validez, consistencia y reproducibilidad de los modelos predictivos desarrollados en el Capítulo 3 y evaluados en el Capítulo 4.

Este capítulo no introduce nuevos resultados, sino que documenta el conjunto de procedimientos, criterios y supuestos que sustentan el comportamiento del modelo y las métricas de desempeño reportadas, permitiendo interpretar los resultados obtenidos como consecuencia directa de un proceso metodológico coherente y reproducible.

7.1. Descripción general del pipeline metodológico

El proceso metodológico se estructuró como un pipeline reproducible que integra datos operacionales de vuelos y datos meteorológicos aeronáuticos, garantizando coherencia temporal, control de calidad y alineación con el objetivo de predicción del indicador On-Time Performance (OTP). A diferencia del análisis exploratorio presentado en capítulos anteriores, este pipeline fue diseñado explícitamente para alimentar los algoritmos de aprendizaje automático descritos en el Capítulo 3.

7.2. Procesamiento y preparación de los datos

Las decisiones descritas en esta sección determinan directamente la configuración del conjunto de entrenamiento utilizado para la construcción de los modelos Random Forest y XGBoost presentados en el Capítulo 3.

La construcción del modelo predictivo se basó en un proceso sistemático de preparación de los datos, estructurado como un pipeline metodológico reproducible que permitió integrar información operacional y meteorológica, controlar la calidad de los datos y garantizar coherencia entre el objetivo del estudio y la etapa de modelado. A diferencia del análisis exploratorio, esta etapa tuvo como finalidad preparar explícitamente el conjunto de datos que alimentó los algoritmos de aprendizaje automático.

7.3. Alcance temporal y consolidación inicial

El conjunto de datos operacionales se construyó a partir de registros históricos de vuelos comerciales domésticos operados en Colombia durante el periodo 2019–2025, excluyendo los años 2020 y 2021. Esta exclusión se justificó por la disrupción estructural generada por la pandemia, caracterizada por cierres operativos, reducción abrupta de la demanda y patrones de operación no representativos del comportamiento regular del sistema aéreo, lo que habría introducido sesgos significativos en el aprendizaje del modelo. Tras la consolidación inicial de los archivos históricos de vuelos, se obtuvo un Dataframe con 1.365.100 registros y 69 variables, correspondientes a información de horarios, aeropuertos, métricas de desempeño y variables operacionales. Posteriormente, se aplicó un filtrado específico para el alcance del estudio, conservando únicamente los vuelos sujetos a medición de puntualidad ($APLICA_OTP = 1$) y con origen y destino dentro del territorio colombiano. Como resultado de este filtrado, el dataset se redujo a 514.539 registros, los cuales constituyeron la base definitiva para el análisis y el modelado.

7.4. Depuración y control de calidad de la información

El proceso de depuración se ejecutó en varias etapas. En primer lugar, se identificaron y eliminaron automáticamente las variables con un 100 % de valores nulos, reduciendo la dimensionalidad del conjunto de datos de 69 a 64 variables útiles. Posteriormente, se revisaron los tipos de datos y rangos plausibles de las variables numéricas, verificando la consistencia de campos relacionados con tiempos, demoras y métricas de puntualidad. Las variables operacionales críticas para la definición del indicador de puntualidad fueron preservadas siempre que permitieran calcular de manera consistente el desempeño del vuelo. Para las variables operacionales donde el dato era recuperable, se empleó la imputación por mediana. Esta decisión se fundamentó en la necesidad de mantener la integridad del tamaño muestral sin sesgar la distribución por valores atípicos (outliers), cumpliendo con el rigor estadístico necesario para modelos de ensamble.

7.5. Procesamiento y transformación de variables meteorológicas

La información meteorológica se procesó a partir de reportes METAR, incorporando variables continuas y categóricas directamente relacionadas con restricciones operativas en la fase de llegada. Las variables continuas incluyeron visibilidad horizontal (en metros), velocidad del viento y ráfagas (en nudos), temperatura y altura del techo de nubes. Estas variables permiten capturar umbrales operativos críticos que no se reflejan adecuadamente mediante correlaciones lineales simples. Adicionalmente, se generaron variables binarias indicativas de la presencia de fenómenos meteorológicos relevantes, como lluvia y bruma o niebla, con el fin de representar eventos discretos que generan impactos operativos no lineales. Esta doble representación (variables continuas y binarias) permitió al modelo capturar tanto la intensidad de las condiciones meteorológicas como la ocurrencia de eventos específicos. Las variables meteorológicas fueron alineadas temporalmente con cada vuelo utilizando como referencia el aeropuerto de destino y la franja horaria de llegada, dado que el indicador OTP

evalúa el cumplimiento del tiempo programado en la llegada. Esta decisión metodológica se fundamenta en que las principales restricciones operativas asociadas al clima (visibilidad reducida, viento cruzado, techo de nubes bajo) afectan de forma directa la fase final del vuelo y los procedimientos de aproximación.

7.6. Integración operacional–meteorológica

La correcta integración entre información meteorológica y operacional es fundamental para interpretar posteriormente las relaciones aprendidas por los modelos, analizadas en el Capítulo 4. La integración de la información operacional y meteorológica se realizó mediante un proceso de unificación basado en claves compuestas, utilizando el aeropuerto de destino, la fecha y la franja horaria como identificadores comunes. Este procedimiento permitió asociar cada vuelo con las condiciones climáticas más representativas del contexto operativo en el que ocurrió su llegada. Como resultado de esta integración, se obtuvo un dataset analítico unificado en el que cada observación representa un vuelo individual, caracterizado por un conjunto consistente de variables operacionales y meteorológicas. Se realizaron validaciones cruzadas para verificar la coherencia temporal y espacial de la información integrada, descartando registros con inconsistencias evidentes entre ambas fuentes.

7.7. Definición de la variable objetivo y selección de predictores

Esta definición fundamenta la elección de modelos de clasificación binaria presentada en la Sección XXXXXXXX, así como la selección de métricas de evaluación utilizadas posteriormente. La variable objetivo del modelo se definió como el cumplimiento del indicador de puntualidad aérea (`cumple_otp`), codificada de forma binaria. Se asignó el valor 1 a los vuelos cuya demora en llegada fue inferior a 15 minutos respecto a la hora programada, y 0 a aquellos que superaron dicho umbral, en concordancia con el estándar utilizado por la industria aérea para la medición del OTP. Es fundamental precisar que la naturaleza del indicador On-Time Performance (OTP) es intrínsecamente binaria (Cumple / No Cumple). Según los estándares de la industria aeronáutica, un vuelo se considera puntual si su llegada ocurre dentro de los 15 minutos posteriores a la hora programada. Bajo esta premisa, la magnitud exacta de los minutos de retraso es una variable intermedia cuya única función operativa es determinar la pertenencia a una de las dos categorías del OTP. Por lo tanto, modelar el retraso mediante regresión lineal de minutos resultaría en un análisis ineficiente para los objetivos de esta tesis, ya que el impacto operativo y la penalización del indicador son idénticos tanto para un retraso de 16 minutos como para uno de 60. En consecuencia, se priorizó el uso de modelos de clasificación, los cuales se alinean directamente con la métrica real de desempeño que las aerolíneas y autoridades aeronáuticas utilizan para evaluar la calidad del servicio.

La selección de predictores se validó mediante el análisis de Importancia de Características (Feature Importance) intrínseco de los algoritmos basados en árboles. La optimización del modelo priorizó el tratamiento del desbalance de clases, ajustando el parámetro `scale_pos_weight` en XGBoost como la razón entre observaciones puntuales y retrasadas, y utilizando el esquema

`class_weight='balanced'` en Random Forest. Esta estrategia permitió mejorar la sensibilidad del modelo hacia la clase retrasado sin recurrir a una optimización exhaustiva indiscriminada de hiperparámetros, manteniendo un equilibrio entre interpretabilidad, desempeño y coherencia operativa. De esta forma, el modelo se comporta como una herramienta calibrada al desbalance real del espacio aéreo colombiano, más que como una caja negra puramente estadística. Las variables predictoras finales incluyeron un conjunto combinado de atributos meteorológicos y operacionales, seleccionados con base en el análisis exploratorio previo y en criterios de relevancia operacional. Las variables categóricas, como el aeropuerto, fueron transformadas mediante codificación numérica para asegurar su compatibilidad con los algoritmos basados en árboles de decisión utilizados en este estudio.

7.8. Preparación final para el modelado

Las decisiones de partición estratificada, ponderación de clases y ajuste del umbral de decisión descritas en esta sección explican de manera directa el comportamiento observado en las métricas de desempeño analizadas en el Capítulo 4, particularmente el incremento significativo del *recall* de la clase correspondiente a los vuelos retrasados. De este modo, esta etapa constituye el nexo metodológico entre la construcción de los modelos presentada en el Capítulo 3 y su evaluación cuantitativa desarrollada posteriormente.

Previo al entrenamiento de los modelos de aprendizaje automático, el dataset integrado fue sometido a una etapa final de preparación orientada a asegurar condiciones adecuadas de generalización, estabilidad numérica y coherencia entre la variable objetivo y las variables predictoras. En primer lugar, se seleccionó el conjunto definitivo de variables de entrada, excluyendo aquellas que presentaban alta colinealidad, baja varianza o dependencia directa con la variable objetivo, con el fin de evitar fugas de información durante el proceso de entrenamiento y garantizar la validez de los resultados obtenidos.

Las variables categóricas, particularmente el aeropuerto de destino, fueron transformadas mediante codificación numérica (*label encoding*), preservando su carácter ordinal implícito dentro de los algoritmos basados en árboles y garantizando su compatibilidad con los modelos Random Forest y XGBoost descritos en el Capítulo 3. Las variables numéricas se mantuvieron en su escala original, dado que los modelos de árboles de decisión no requieren estandarización previa y son robustos frente a diferencias de magnitud entre predictores.

El conjunto de datos resultante fue dividido en subconjuntos de entrenamiento y validación mediante una partición estratificada 80/20, asegurando la preservación de la proporción original entre vuelos puntuales y retrasados en ambos subconjuntos. Esta estrategia permitió evaluar la capacidad de generalización del modelo sobre datos no utilizados durante el entrenamiento, minimizando el riesgo de sobreajuste y asegurando la coherencia estadística de las métricas reportadas.

Para reforzar la robustez del proceso de entrenamiento, se implementó una validación cruzada estratificada (*Stratified K-Fold*) con $k = 5$. Este procedimiento garantizó que cada registro del dataset consolidado fuera utilizado tanto para entrenamiento como para validación en distintas iteraciones, eliminando sesgos de selección y proporcionando soporte estadístico a la estabilidad de las métricas de desempeño, particularmente precisión y *recall*.

Dado el desbalance natural entre clases, característico de los problemas de puntualidad aérea, se incorporaron mecanismos específicos para mejorar la capacidad predictiva sobre la clase minoritaria correspondiente a los vuelos retrasados. En particular, se aplicaron estrategias de ponderación de clases durante el entrenamiento, asignando mayor peso a los errores de clasificación asociados a retrasos, con el objetivo de incrementar la sensibilidad del modelo sin comprometer de manera significativa la precisión global. Esta decisión metodológica explica el comportamiento más equilibrado observado en el modelo XGBoost optimizado durante la etapa de evaluación presentada en el Capítulo 4.

Adicionalmente, se realizó una calibración del umbral de decisión sobre la probabilidad estimada por el modelo, priorizando un equilibrio operativo entre *recall* y *precision* para la clase retrasado. Este ajuste respondió a la naturaleza aplicada del problema, donde la detección temprana de vuelos con alto riesgo de incumplimiento resulta más relevante desde el punto de vista operacional que la maximización exclusiva de métricas globales de exactitud.

Como resultado de este proceso, el dataset quedó preparado para alimentar de forma consistente los algoritmos de clasificación empleados, garantizando que el desempeño observado en la etapa de evaluación reflejara de manera fiel la capacidad del modelo para anticipar retrasos bajo condiciones reales de operación.

7.9. Relación entre la preparación de datos, el modelado y la evaluación

La estructura del pipeline metodológico descrito en este capítulo asegura una coherencia explícita entre las etapas de preparación de datos (Capítulo 5), construcción del modelo predictivo (Capítulo 3) y evaluación del desempeño (Capítulo 4). Las métricas reportadas no deben interpretarse como resultados aislados, sino como consecuencia directa de las decisiones metodológicas adoptadas en cada fase del proceso.

En particular, la definición de la variable objetivo, el tratamiento del desbalance de clases, la selección de predictores y la calibración del umbral de decisión influyen de manera determinante en la capacidad del modelo para identificar vuelos con riesgo de retraso. Esta trazabilidad metodológica permite explicar por qué el modelo optimizado presenta una mejora sustancial en el *recall* de la clase retrasado sin una degradación significativa del desempeño global, tal como se evidencia en los resultados del Capítulo 4.

De esta forma, el análisis integrado de los capítulos 3, 4 y 5 demuestra que el comportamiento del modelo responde a criterios técnicos y operacionales explícitos, y no a ajustes empíricos arbitrarios, fortaleciendo la consistencia interna del estudio.

7.10. Consideraciones sobre la validez y reproducibilidad del estudio

El enfoque metodológico adoptado en este trabajo permite evaluar de manera transparente la validez de los resultados obtenidos, al documentar de forma detallada cada una de las decisiones tomadas durante el procesamiento de datos, el entrenamiento de los modelos y la evaluación de su desempeño. La estructura del pipeline metodológico descrito garantiza la reproducibilidad del estudio bajo las mismas condiciones de datos y parámetros, permitiendo replicar los experimentos y verificar la estabilidad de las métricas reportadas.

El desarrollo del proyecto siguió un pipeline metodológico estructurado y coherente, que abarca desde la integración y depuración de las fuentes de datos hasta la evaluación final del desempeño de los modelos predictivos. Cada etapa del proceso fue definida de manera explícita, incluyendo la selección de variables, la construcción del dataset unificado, la definición de la variable objetivo, el entrenamiento de los modelos y su validación mediante métricas cuantitativas. Con el fin de garantizar la reproducibilidad de los resultados, se documentaron los criterios de filtrado, transformación y selección de variables, así como los parámetros utilizados en el entrenamiento de los modelos Random Forest y XGBoost. La partición de los datos, la configuración de los hiperparámetros y las métricas de evaluación empleadas se mantuvieron constantes a lo largo de los experimentos, permitiendo replicar los resultados bajo las mismas condiciones experimentales.

Adicionalmente, se evaluó la estabilidad del modelo mediante la comparación del desempeño obtenido en los conjuntos de entrenamiento y validación, observándose variaciones marginales que descartan la presencia de sobreajuste. La consistencia de los resultados fue contrastada mediante análisis complementarios, como la validación de la importancia de variables y la coherencia con los patrones identificados en el análisis exploratorio, lo que respalda el carácter riguroso del proceso metodológico seguido.

Si bien el modelo predictivo se centra en variables meteorológicas y operacionales observables, se reconoce que existen factores externos no modelados explícitamente, como decisiones de control de tráfico aéreo o estrategias operativas de las aerolíneas, que pueden introducir variabilidad adicional en el cumplimiento del indicador de puntualidad. No obstante, estas limitaciones no invalidan los resultados obtenidos, sino que delimitan el alcance del modelo y refuerzan su interpretación como una herramienta de apoyo a la toma de decisiones, más que como un sistema determinista.

En conjunto, el rigor del proceso de preparación de datos, la coherencia entre las etapas de modelado y evaluación, y la alineación con criterios operacionales reales proporcionan el sustento metodológico necesario para considerar válidos y consistentes los resultados presentados en esta investigación.

Conclusiones y trabajos futuros

8.1. Conclusiones

El análisis articulado de los cuatro objetivos específicos permitió comprender de manera integral cómo las condiciones meteorológicas y operativas influyen en la puntualidad de los vuelos comerciales domésticos en Colombia, así como evaluar la capacidad de los modelos predictivos desarrollados para anticipar escenarios de riesgo de retraso bajo condiciones reales de operación.

En primer lugar, se evidenció que el Aeropuerto Internacional El Dorado (BOG) ejerce una influencia desproporcionada sobre el desempeño de la red aérea nacional, debido a su rol como principal hub doméstico. Las condiciones meteorológicas características de este aeropuerto —particularmente episodios de niebla, reducciones abruptas de visibilidad y ráfagas de viento— generan impactos que se propagan hacia otros aeropuertos del sistema. Sin embargo, dichos efectos no se manifiestan de forma lineal, sino que emergen al superarse ciertos umbrales operativos críticos, lo cual explica las correlaciones débiles observadas durante el análisis exploratorio y justifica el uso de modelos no lineales basados en árboles de decisión.

En contraste, las demoras operacionales registradas en salida y llegada mostraron una relación consistente con el desempeño final de los vuelos, confirmando la importancia de la propagación de retrasos dentro de la red aérea. Adicionalmente, se identificaron patrones estacionales parcialmente asociados a periodos de mayor precipitación, especialmente en los meses de abril–mayo y noviembre–diciembre. No obstante, la presencia de variabilidad operativa durante meses climáticamente secos evidencia que el clima no actúa de manera aislada, sino en interacción con factores operacionales. Estos hallazgos sustentan la necesidad de un enfoque sistémico que integre ambas dimensiones, fundamento central del modelo predictivo desarrollado en este estudio.

En relación con la selección de características, se identificaron variables meteorológicas con influencia significativa sobre la probabilidad de retraso, entre las que destacan la visibilidad, las ráfagas de viento y la velocidad del viento, así como, en menor medida, la precipitación y la presencia de niebla. La relevancia de estas variables fue particularmente evidente en aeropuertos de alta congestión como BOG, donde las restricciones operativas asociadas a la visibilidad tienden a generar demoras en cadena. Estos resultados validan la hipótesis inicial del estudio y justifican la selección del conjunto de predictores utilizado en la fase de modelado, cuya preparación y depuración metodológica fue documentada de manera explícita en el Capítulo 5.

Sobre esta base, el desarrollo del modelo predictivo correspondiente al objetivo específico 3 permitió construir una herramienta analítica orientada a la estimación de la probabilidad de incumplimiento del indicador On-Time Performance (OTP). La elección de un enfoque de clasificación probabilística, en lugar de una regresión del tiempo exacto de retraso en minutos, se fundamentó en

criterios operativos y metodológicos claramente definidos, alineados con el estándar internacional del OTP y con la naturaleza binaria del indicador utilizado por la industria aeronáutica.

El proceso de experimentación y optimización evidenció que el modelo XGBoost, ajustado mediante estrategias de ponderación de clases, validación cruzada estratificada y calibración del umbral de decisión, ofreció el mejor equilibrio entre capacidad predictiva, estabilidad y generalización. Estas decisiones metodológicas, documentadas de manera detallada en el pipeline descrito en el Capítulo 5, permiten interpretar el desempeño del modelo como resultado directo de un proceso reproducible y técnicamente justificado, y no como producto de ajustes empíricos arbitrarios.

La evaluación desarrollada en el marco del objetivo específico 4 confirmó la solidez del modelo final. Las mejoras observadas en métricas críticas para la clase minoritaria —particularmente el recall, el F1-score y el área bajo la curva Precision–Recall— evidencian una mejora sustancial en la detección temprana de vuelos con alto riesgo de retraso, sin comprometer de manera significativa el desempeño global. Asimismo, la estabilidad de las métricas sobre el conjunto de prueba respalda la capacidad de generalización del modelo bajo condiciones no observadas durante el entrenamiento.

Finalmente, el desarrollo de un prototipo visual interactivo basado en Streamlit permitió trasladar los resultados analíticos hacia un contexto aplicado, facilitando su interpretación y exploración por parte de usuarios operativos. Este componente demuestra el potencial del modelo como herramienta de apoyo a la toma de decisiones tácticas, orientada a la gestión proactiva del riesgo meteorológico en la operación aérea colombiana.

En conjunto, la coherencia entre la preparación de los datos, la construcción del modelo y la evaluación de su desempeño, así como la trazabilidad explícita de las decisiones metodológicas adoptadas, proporcionan el sustento necesario para considerar válidos, consistentes y reproducibles los resultados obtenidos en esta investigación.

8.2. Trabajos futuros

Los resultados obtenidos permiten identificar varias líneas de desarrollo que podrían fortalecer y ampliar el alcance y la aplicabilidad del modelo en etapas posteriores.

8.2.0.1. Extensiones del Modelo:

La primera línea de trabajo consiste en ampliar las fuentes de información y el nivel de detalle de esta:

- Integración de datos de tráfico aéreo en tiempo real para incorporar información sobre congestión y capacidad operativa de los aeropuertos. Esto permitiría que el modelo no solo estime el riesgo de retraso con base en condiciones históricas y climáticas, sino que también actualice sus predicciones de acuerdo con los cambios operacionales que se presenten durante el día.
- Incorporación de variables socioeconómicas y eventos especiales (festividades, incrementos estacionales de demanda o restricciones logísticas), con el fin de explicar fluctuaciones de puntualidad no asociadas al clima, que expandirían el alcance a la influencia de los usuarios y el entorno en el sistema.

- Desarrollo de modelos diferenciados por tipo de aeronave, considerando que ciertas flotas presentan mayor sensibilidad a condiciones como viento cruzado, altitud o longitud de pista, para establecer la influencia de estos factores en la puntualidad.
- Evaluar el uso de técnicas de deep learning para analizar patrones no lineales más complejos en el comportamiento de los retrasos, contemplando la posible integración futura de datos operacionales en tiempo real sobre número de operaciones por hora, niveles de congestión aeroportuaria, secuencias de propagación de demoras durante la jornada y observaciones provenientes de imágenes satelitales o sensores automáticos.

8.2.0.2. Aplicaciones Adicionales:

El modelo también puede evolucionar hacia nuevos casos de uso aplicados a la operación aérea:

- Extensión del análisis a vuelos internacionales conectados con la red doméstica, especialmente en aeropuertos hub.
- Desarrollo de un modelo predictivo para cancelaciones, dada su mayor relevancia operativa y financiera frente a los retrasos.
- Implementación de un sistema de optimización de rutas y reprogramaciones, utilizando la probabilidad de retraso como insumo para la toma de decisiones.
- Diseño de una herramienta de planificación de capacidad aeroportuaria que permita anticipar escenarios de saturación y ejecutar medidas de mitigación con anticipación.

8.2.0.3. Mejoras Tecnológicas:

Finalmente, existen oportunidades para facilitar la implementación operacional del prototipo:

- Implementación del modelo en tiempo real mediante procesamiento en streaming para generar alertas tempranas.
- Desarrollo de aplicaciones móviles orientadas a la comunicación proactiva con pasajeros y actores operativos.
- Integración con dispositivos IoT y fuentes automáticas de observación meteorológica para reducir la incertidumbre en aeropuertos críticos.
- Aplicación de técnicas de machine learning explicable (XAI), Explainable Artificial Intelligence, Inteligencia Artificial Explicable por sus siglas en inglés, con el propósito de facilitar la interpretación del modelo y fortalecer su adopción por parte de aerolíneas y autoridades aeronáuticas.

9.1. Anexos

9.1.1. Anexo A: Glosario de Terminos

A continuación, se presentan los principales conceptos técnicos, operacionales, meteorológicos y analíticos utilizados en el desarrollo del proyecto aplicado. Su propósito es garantizar claridad terminológica y facilitar la lectura del documento.

Aeronáutica Civil (Aerocivil): Autoridad aeronáutica de Colombia encargada de regular, administrar y supervisar la aviación civil en el país.

Algoritmo de Aprendizaje Automático (Machine Learning): Conjunto de métodos computacionales que permiten a un sistema aprender patrones a partir de datos sin programación explícita.

Área Bajo la Curva ROC (AUC-ROC): Métrica que evalúa la capacidad de un modelo para discriminar entre clases, considerando distintos umbrales de clasificación.

Boosting: Método de ensamblaje de modelos donde cada árbol corrige los errores del anterior, utilizado en algoritmos como XGBoost.

Bucket Horario: Agrupación de horas en intervalos (ej. 0–23) utilizada para analizar comportamientos por franja horaria, como STA_HOUR.

Ceiling (Techo de nubes): Altura de la base de las nubes medidas en pies; valores bajos afectan la operación aeronáutica.

Clase Desbalanceada: Situación en la cual una categoría del conjunto de datos (ej. vuelos retrasados) tiene una proporción significativamente menor que la otra.

CUMPLE_OTP: Variable binaria que indica si el vuelo llegó con menos de 15 minutos de retraso.

Curva PR (Precision-Recall): Gráfico que evalúa la relación entre precisión y sensibilidad para clases desbalanceadas.

Delay Propagation (Propagación de retrasos): Fenómeno en el cual un retraso inicial se transfiere a vuelos posteriores. **DEMORA_ARR / DEMORA_DEP:** Minutos de retraso en la llegada/salida de un vuelo respecto a su hora programada.

Ensemble Learning: Técnica que combina varios modelos (árboles) para mejorar la precisión final, como en Random Forest o XGBoost.

Exploración de Datos (EDA): Etapa inicial de análisis donde se examinan patrones, distribuciones, correlaciones y calidad del dataset.

Feature Engineering (Ingeniería de Características): Proceso de crear, transformar o seleccionar variables relevantes para mejorar el rendimiento del modelo.

Feature Importance: Medida que indica el peso relativo de cada variable durante el entrenamiento del modelo.

Gradient Boosting: Estrategia de entrenamiento secuencial donde el nuevo modelo corrige los residuos del anterior.

Gust (Ráfaga): Incremento repentino y temporal en la velocidad del viento; impacta la operación aérea.

Hiperparámetro: Parámetro del modelo que se define antes del entrenamiento (ej. profundidad, tasa de aprendizaje).

Humedad Relativa: Cantidad de vapor de agua en el aire respecto a su máximo posible.

IDEAM: Entidad colombiana encargada de monitorear variables meteorológicas y climáticas nacionales.

Importancia por Ganancia (Gain): Criterio de XGBoost que evalúa cuánto mejora el modelo cuando usa una variable en una división.

Learning Rate: Parámetro de XGBoost que controla la velocidad con la que el modelo aprende.

METAR (Meteorological Aerodrome Report): Reporte meteorológico oficial emitido cada hora en los aeropuertos.

Matriz de Confusión: Tabla que muestra verdaderos positivos, falsos positivos, falsos negativos y verdaderos negativos del modelo.

Modelo Predictivo: Herramienta estadística o computacional utilizada para predecir un evento futuro, como un retraso aéreo.

NOAA: Organización estadounidense que provee datos meteorológicos globales utilizados para análisis de clima.

On Time Performance (OTP): Indicador que mide la puntualidad de vuelos, basado en llegadas con retraso menor a 15 minutos.

Overfitting (Sobreajuste): Situación donde el modelo aprende demasiado los datos de entrenamiento y pierde capacidad de generalización.

Precisión (Precision): Proporción de predicciones correctas entre todos los casos clasificados como positivos.

Propagación Operacional: Impacto indirecto de un retraso en otros vuelos debido al rol de la aeronave, tripulación o conexiones.

Random Forest: Algoritmo basado en múltiples árboles de decisión entrenados de forma aleatoria para mejorar estabilidad.

Recall (Sensibilidad): Capacidad del modelo para identificar correctamente los casos reales positivos o negativos.

Regresión / Clasificación: Tipos de modelos de aprendizaje automático para variables continuas o categóricas.

Scale Pos Weight: Parámetro utilizado en XGBoost para manejar clases desbalanceadas. STA / STD: Hora programada de llegada (STA) y salida (STD).

Visibilidad (Visibility_m): Distancia horizontal que puede observarse desde la pista, variable crítica en aeropuertos.

Viento Cruzado (Crosswind): Componente del viento that impacts perpendicularly to the pista, afectando despegues y aterrizajes.

XGBoost (Extreme Gradient Boosting): Algoritmo avanzado de boosting, altamente eficiente para datos tabulares, usado como modelo final de tu tesis.

9.1.2. Anexo B: Diccionario de datos

Cuadro 9.1: Diccionario de Datos

Clasificación	Variable	Descripción	Tipo de dato	Unidad / Formato
Operacionales	FECHA_LOCAL	Fecha de operación del vuelo	Fecha	AAAA-MM-DD
Operacionales	ORIGEN	Aeropuerto de origen (IATA)	Categórica	-
Operacionales	DESTINO	Aeropuerto de destino (IATA)	Categórica	-
Operacionales	VUELO	Número de vuelo		
Operacionales	TIPO_FLIGHT	Tipo de vuelo (dom., intl., etc.)	Texto	-
Operacionales	DEMORA_DEP	Minutos de demora a la salida	Categórica	-
Operacionales	DEMORA_ARR	Minutos de demora a la llegada	Entero	-
Operacionales	STD_LOCAL	Hora programada de salida	Entero	HH:MM
Operacionales	STA_LOCAL	Hora programada de llegada	Hora	HH:MM
Operacionales	APLICA_OTP	Bandera de medición OTP	Hora	0/1
Operacionales	CUMPLE_OTP	Cumple puntualidad <15 min	Binaria	0/1
Operacionales	AÑO	Año de operación	Binaria	Año
Operacionales	MES	Mes de operación	Entero	1-12
Operacionales	DIA_SEMANA	Día de la semana	Entero	1-7

Continúa en la siguiente página

Continúa de la página anterior

Clasificación	Variable	Descripción	Tipo de dato	Unidad / Formato
Operacionales	STA_HOUR	Hora de llegada programada (bucket)	Entero	Hora
Meteorológicas	wx_temperature	Temperatura del aire	Numérica	°C
Meteorológicas	wx_dewpoint	Punto de rocío	Numérica	°C
Meteorológicas	wx_wind_speed_kt	Velocidad del viento	Numérica	Nudos
Meteorológicas	wx_gust_kt	Ráfagas de viento	Numérica	Nudos
Meteorológicas	wx_wind_dir	Dirección del viento	Numérica	Grados
Meteorológicas	wx_visibility_m	Visibilidad horizontal	Numérica	Metros
Meteorológicas	wx_ceiling	Altura del techo de nubes	Entero	Pies
Meteorológicas	wx_rain	Presencia de lluvia	Binaria	0/1
Meteorológicas	wx_fog_brume	Niebla o bruma	Binaria	0/1
Meteorológicas	wx_pressure_hpa	Presión atmosférica	Numérica	hPa
Meteorológicas	wx_humidity	Humedad relativa	Numérica	%
Meteorológicas	wx_weather_code	Código METAR simplificado	Texto	-
Meteorológicas	datetime_utc	Hora del reporte (UTC)	Fecha-hora	UTC
Meteorológicas	datetime_local	Hora del reporte (local)	Fecha-hora	Local
Modelo de datos	delay_binary	Etiqueta binaria puntual (1) / retrasado (0)	Binaria	-
Modelo de datos	visibility_flag	Bandera de visibilidad crítica (<2000 m)	Binaria	-
Modelo de datos	gust_flag	Ráfagas críticas (>20 kt)	Binaria	-
Modelo de datos	fog_flag	Niebla o bruma	Binaria	-
Modelo de datos	rain_flag	Lluvia presente	Binaria	-
Modelo de datos	wind_cross_component	Cálculo de viento cruzado	Numérica	-
Modelo de datos	prvious_delay_prop	Propagación del retraso desde vuelo previo	Numérica	-
Modelo de datos	airport_congestion_proxy	Congestión por hora en aeropuerto	Numérica	-
Modelo de datos	CLIMA_MATCH	Match temporal METAR-vuelo	Binaria	-

Fin de la tabla larga

9.1.3. Anexo C: Construcción y Depuración de la Base Histórica de Vuelos

Este anexo presenta el proceso inicial de preparación de datos realizado para el proyecto, el cual consistió en la carga, limpieza y estandarización de la base histórica de vuelos. En esta etapa se consolidaron múltiples fuentes de información operativa, se corrigieron inconsistencias en los campos de fecha y hora, se depuraron registros incompletos y se generaron variables necesarias para los análisis posteriores, tales como indicadores operacionales y atributos temporales. Este preprocesamiento garantizó la calidad y coherencia del dataset final, que fue utilizado como insumo principal para el análisis exploratorio y el desarrollo del modelo predictivo de puntualidad.

```
1 import pandas as pd
2 import glob
3 import os
4
5 ruta = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD Flights"
6 archivos = glob.glob(os.path.join(ruta, "OPS FLIGHTS *.xlsx"))
7
8 print(f"Archivos encontrados: {len(archivos)}")
9 for a in archivos:
10     print("-", os.path.basename(a))
11
12 df_list = []
13 for archivo in archivos:
14     print(f"Leyendo: {os.path.basename(archivo)}")
15     df = pd.read_excel(archivo, engine="openpyxl")
16     df["FUENTE_ARCHIVO"] = os.path.basename(archivo)
17     df_list.append(df)
18
19 df_flights = pd.concat(df_list, ignore_index=True)
20
21 print("Archivos combinados correctamente")
22 print("Filas totales:", len(df_flights))
23 print("Columnas:", len(df_flights.columns))
24 print(df_flights.columns[:10])
```

Listing 9.1: Carga y consolidación de archivos OPS FLIGHTS

```
1 import os, sys, subprocess
2 import pandas as pd
3
4 # Guardar dataset como Parquet
5 salida_parquet = os.path.join(ruta, "OPS_FLIGHTS_2019_2025.parquet")
6
7 try:
8     import pyarrow
9 except ImportError:
```

```

10     subprocess.check_call([sys.executable, "-m", "pip", "install",
11         "--quiet", "pyarrow"])
12     import pyarrow
13 df_flights.to_parquet(salida_parquet, index=False, engine="pyarrow")
14
15 # Cargar nuevamente
16 df_flights = pd.read_parquet(salida_parquet)
17
18 print("Dataset cargado correctamente")
19 print(f"Filas: {len(df_flights)}")
20 print(f"Columnas: {len(df_flights.columns)}")
21
22 # Validación general
23 df_flights["FECHA_LOCAL"] = pd.to_datetime(df_flights["FECHA_LOCAL"],
24     errors="coerce")
25
26 # Filtrado OTP doméstico sin 2020-2021
27 df_core = df_flights.copy()
28 df_core["APLICA_OTP"] = pd.to_numeric(df_core["APLICA_OTP"],
29     errors="coerce").fillna(0).astype("Int8")
30
31 df_core["PAIS ORIGEN"] = df_core["PAIS
32     ORIGEN"].astype("string").str.upper()
33 df_core["PAIS DESTINO"] = df_core["PAIS
34     DESTINO"].astype("string").str.upper()
35
36 mask = (
37     (df_core["APLICA_OTP"] == 1)
38     & (df_core["PAIS ORIGEN"] == "COLOMBIA")
39     & (df_core["PAIS DESTINO"] == "COLOMBIA")
40     & ~df_core["FECHA_LOCAL"].dt.year.isin([2020, 2021])
41 )
42
43 df_domestic = df_core.loc[mask].copy()
44 df_domestic["AO"] = df_domestic["FECHA_LOCAL"].dt.year
45 df_domestic["MES"] = df_domestic["FECHA_LOCAL"].dt.month

```

Listing 9.2: Conversión, validación y filtrado del dataset

```

1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 import seaborn as sns
5
6 df_domestic["CUMPLE_OTP"] = pd.to_numeric(df_domestic["CUMPLE_OTP"],
7     errors="coerce")

```

```

7
8 otp_mensual = (
9     df_domestic
10     .groupby(["AO", "MES"])["CUMPLE_OTP"]
11     .mean()
12     .reset_index()
13     .pivot(index="AO", columns="MES", values="CUMPLE_OTP") * 100
14 )
15
16 mes_labels = {1:"Ene",2:"Feb",3:"Mar",4:"Abr",5:"May",6:"Jun",
17              7:"Jul",8:"Ago",9:"Sep",10:"Oct",11:"Nov",12:"Dic"}
18
19 otp_mensual = otp_mensual.reindex(columns=range(1,13))
20 otp_mensual.columns = [mes_labels[m] for m in otp_mensual.columns]
21
22 plt.figure(figsize=(12,6))
23 sns.heatmap(
24     otp_mensual,
25     annot=True, fmt=".1f",
26     cmap="coolwarm",
27     center=85,
28     vmin=70, vmax=100,
29     linewidths=.5, linecolor="gray",
30     cbar_kws={"label": "OTP (%)"}
31 )
32
33 plt.title("OTP mensual por a o (en %)")
34 plt.ylabel("A o ")
35 plt.xlabel("Mes")
36 plt.tight_layout()
37 plt.show()

```

Listing 9.3: Análisis exploratorio y mapa de calor OTP

9.1.4. Anexo D. Procesamiento de Datos Meteorológicos para el Modelo

Este anexo documenta el procesamiento inicial de datos meteorológicos utilizado para la construcción de la tabla DOMCO. Incluye la carga, depuración y estructuración de variables climáticas empleadas posteriormente en el análisis y en el desarrollo del modelo predictivo.

```

1 from pathlib import Path
2 import pandas as pd
3
4 try:
5     from IPython.display import display
6 except Exception:

```

```
7     def display(x): print(x)
8
9 AIRPORTS_DIR = Path(r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
    Weather")
10
11 candidates = list(AIRPORTS_DIR.glob("airports*.csv")) +
    list(AIRPORTS_DIR.glob("*airport*.csv"))
12
13 if len(candidates) == 0:
14     print("No se encontr 'airports.csv'. Usando OurAirports.")
15     airports = pd.read_csv("https://ourairports.com/data/airports.csv")
16 else:
17     airports_path = max(candidates, key=lambda p: p.stat().st_size)
18     print(f"Usando archivo local: {airports_path}")
19     airports = pd.read_csv(airports_path)
20
21 print("Dimensi n:", airports.shape)
22 display(airports.head())
23
24 for col in ["iata_code", "gps_code", "name", "municipality", \
25 "iso_country", "type", "scheduled_service"]:
26     if col in airports.columns:
27         airports[col] = airports[col].astype("string").str.strip()
28
29 airports["iata_code"] = airports["iata_code"].str.upper()
30 airports["gps_code"] = airports["gps_code"].str.upper()
31
32 target_iata = {
33 "ADZ", "AXM", "BAQ", "BGA", "BOG", "CZU", "CLO", "CTG", "CUC", \
34 "EJA", "EYP", "FLA", "IBE", "IPI",
35 "LET", "MZL", "MDE", "MTR", "NVA", "PEI", "PPN", "PSO", \
36 "RCH", "SMR", "VUP", "VVC", "TCO", "AUC", "UIB"
37 }
38
39 def _priority_type(t):
40     order = {
41         "large_airport":0, "medium_airport":1, "small_airport":2,
42         "heliport":3, "seaplane_base":4, "balloonport":5, "closed":9
43     }
44     return order.get(str(t), 8)
45
46 air = airports.copy()
47 air["_prio_type"] = air["type"].map(_priority_type)
48 air["_has_gps"] = air["gps_code"].notna()
49 air["_sched_score"] =
    air["scheduled_service"].fillna("no").str.lower().eq("yes")
```

```

50
51 air = air.sort_values(["_prio_type", "_has_gps", "_sched_score", \
52 "elevation_ft", "name"],
53                       ascending=[True, False, False, True, True])
54
55 air_dedup = air.drop_duplicates(subset=["iata_code"], keep="first")
56 ap_sel = air_dedup.loc[air_dedup["iata_code"].isin(target_iata)].copy()
57
58 airports_idx = ap_sel.set_index("iata_code", drop=False)
59 stations_list = sorted(ap_sel["iata_code"].unique())

```

Listing 9.4: Carga y selección de aeropuertos

```

1 import datetime
2 import pandas as pd
3
4 def getdata(stationslist, year=5, df=None, adjUTC=-5):
5     if df is None or "gps_code" not in df.columns:
6         raise ValueError("df debe contener columna gps_code")
7
8     base = "https://mesonet.agron.iastate.edu/cgi-bin/request/asos.py?"
9     parts = [base]
10
11     for iata in stationslist:
12         if iata in df.index:
13             icao = df.at[iata, "gps_code"]
14             if pd.notna(icao):
15                 parts.append(f"station={icao}&")
16
17     parts.append("data=tmpc&data=drct&data=sknt&data=vsby&data=\
18 skyl1&data=wxcodes&data=gust")
19
20     date2 = datetime.datetime.utcnow().date() - datetime.timedelta(days=1)
21     y2, m2, d2 = date2.year, date2.month, date2.day
22     y1 = y2 - year
23
24     parts.append(f"&year1={y1}&month1=1&day1=1&year2=\
25 {y2}&month2={m2}&day2={d2}")
26     parts.append("&format=onlycomma&tz=Etc/UTC&missing=null")
27
28     url = "".join(parts)
29     print("Descargando:", url)
30
31     data = pd.read_csv(url)
32     data["valid"] = pd.to_datetime(data["valid"])
33
34     data["Year"] = data["valid"].dt.year

```

```

35 data["Month"] = data["valid"].dt.month
36 data["Day"] = data["valid"].dt.day
37 data["Hour UTC"] = data["valid"].dt.hour
38 data["Hour LT"] = (data["valid"] +
    datetime.timedelta(hours=adjUTC)).dt.hour
39
40 wx = data.get("wxcodes", pd.Series("",
    index=data.index)).astype("string")
41 data["Rain"] = wx.str.contains("RA").astype(int) * 100
42 data["Fog-Brume"] = wx.str.contains("FG|BR|HZ").astype(int) * 100
43 data["Thunder"] = wx.str.contains("TS").astype(int) * 100
44 data["Snow"] = wx.str.contains("SN").astype(int) * 100
45 data["Op constraint"] = ((data["Fog-Brume"]==100) |
    (data["Thunder"]==100)).astype(int) * 100
46 data["Gust_"] = data["gust"].notna().astype(int) * 100
47
48 return data

```

Listing 9.5: Función para descarga de datos ASOS (Mesonet)

```

1 import pandas as pd
2
3 dfw = raw_data.copy()
4
5 for c in ["Year", "Month", "Day", "Hour UTC", "Hour LT"]:
6     dfw[c] = pd.to_numeric(dfw[c], errors="coerce").astype("Int64")
7
8 dfw["datetime_utc"] = pd.to_datetime(
9     dict(year=dfw["Year"], month=dfw["Month"], day=dfw["Day"],
10         hour=dfw["Hour UTC"]),
11     utc=True, errors="coerce"
12 )
13 dfw["datetime_col"] = dfw["datetime_utc"] - pd.Timedelta(hours=5)
14
15 rename_map = {
16     "IATA": "airport",
17     "Wind": "wind_speed_kt",
18     "Gust": "wind_gust_kt",
19     "H Vsby": "visibility_m",
20 }
21
22 dfw.rename(columns=rename_map, inplace=True)
23
24 keep_cols = [
25     "airport", "OACI", "datetime_utc", "datetime_col",
26     "wind_speed_kt", "wind_gust_kt", "visibility_m",

```

```

27     "Temperature", "Wind Dir", "V Vsby",
28     "Rain", "Fog-Brume", "Thunder", "Snow", "Op constraint"
29 ]
30
31 dfw = dfw[[c for c in keep_cols if c in dfw.columns]]
32 dfw["airport"] = dfw["airport"].str.upper().str.strip()

```

Listing 9.6: Limpieza y normalización de variables meteorológicas

```

1  import pandas as pd
2  import matplotlib.pyplot as plt
3  import seaborn as sns
4
5  event_cols = ["Rain", "Fog-Brume", "Thunder", "Snow", "Op constraint", "Gust_"]
6
7  event_summary = (dfw[event_cols] > 0).mean() * 100
8  print(event_summary)
9
10 df_bog = dfw[dfw["airport"]=="BOG"].copy()
11 df_bog["Rain_flag"] = (df_bog["Rain"]==100).astype(int)
12 df_bog["hour"] = df_bog["datetime_col"].dt.hour
13 df_bog["month"] = df_bog["datetime_col"].dt.month
14
15 rain_matrix = (
16     df_bog.groupby(["hour", "month"])["Rain_flag"]
17         .mean()
18         .unstack(level=1) * 100
19 )
20
21 plt.figure(figsize=(12,6))
22 sns.heatmap(rain_matrix, cmap="Blues")
23 plt.title("Bogot      Probabilidad de lluvia por hora y mes")
24 plt.xlabel("Mes")
25 plt.ylabel("Hora local")
26 plt.show()

```

Listing 9.7: Análisis exploratorio de fenómenos meteorológicos

```

1  import os
2  from math import ceil
3  import pandas as pd
4
5  output_path = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
6  Weather\weather_outputs\weather_colombia_chunks.xlsx"
7  os.makedirs(os.path.dirname(output_path), exist_ok=True)
8
9  df_export = dfw.copy()

```

```

9
10 df_export["datetime_utc"] = df_export["datetime_utc"].dt.tz_localize(None)
11 df_export["datetime_col"] = pd.to_datetime(df_export["datetime_col"],
      errors="coerce")
12
13 max_per_sheet = 1_000_000
14 parts = ceil(len(df_export) / max_per_sheet)
15
16 with pd.ExcelWriter(output_path, engine="xlsxwriter") as writer:
17     for i in range(parts):
18         sub = df_export.iloc[i*max_per_sheet:(i+1)*max_per_sheet]
19         sub.to_excel(writer, sheet_name=f"data_p{i+1}", index=False)
20
21 print("Exportado en:", output_path)

```

Listing 9.8: Exportación del dataset meteorológico a Excel

9.1.5. Anexo E: Unificación de Datos de Vuelos y Clima

Este anexo presenta el proceso de integración entre la base histórica de vuelos y la tabla meteorológica, generando un conjunto consolidado utilizado posteriormente para el análisis exploratorio y el modelado predictivo.

```

1 import pandas as pd
2
3 ruta = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
      Flights\OPS_FLIGHTS_domesticos_OTP_sin_2020_2021.parquet"
4 df = pd.read_parquet(ruta, engine="pyarrow")
5
6 if "FECHA_LOCAL" in df.columns:
7     df["FECHA_LOCAL"] = pd.to_datetime(df["FECHA_LOCAL"], errors="coerce")
8
9 for c in ["ROUTE_OW", "ORIGEN", "DESTINO", "TIPO_FLT1", "TIPO_FLT2"]:
10     if c in df.columns:
11         df[c] = df[c].astype("category")
12
13 print(df.shape)
14 print(df.head())
15 print(df.dtypes.head(12))

```

Listing 9.9: Carga del dataset de vuelos en formato Parquet

```

1 for col in df.columns:
2     print(col)

```

Listing 9.10: Impresión de columnas disponibles en el dataset de vuelos

```

1 import pandas as pd
2 from pathlib import Path
3
4 ruta = Path(r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
   Weather\weather_outputs\weather_colombia.xlsx")
5
6 if ruta.exists():
7     df_weather = pd.read_excel(ruta, engine="openpyxl")
8     print("Archivo cargado correctamente.")
9     print("Filas:", len(df_weather), "| Columnas:",
10         len(df_weather.columns))
11     print("Primeras columnas:", df_weather.columns[:10].tolist())
12 else:
13     print("No se encontr el archivo en la ruta indicada.")

```

Listing 9.11: Carga de datos meteorológicos desde archivo Excel

```

1 df_weather.info()

```

Listing 9.12: Estructura del dataset meteorológico

```

1 import pandas as pd
2 from time import perf_counter
3
4 print(df["STA_LOCAL"].dtype)
5
6 t0 = perf_counter()
7 if not pd.api.types.is_datetime64_any_dtype(df["STA_LOCAL"]):
8     df["STA_LOCAL"] = pd.to_datetime(
9         df["STA_LOCAL"], errors="coerce", format="%Y-%m-%d %H:%M:%S"
10    )
11 df["STA_HOUR"] = df["STA_LOCAL"].dt.hour
12 t1 = perf_counter()
13
14 print(f"Tiempo total (parseo+hora): {t1 - t0:.3f}s")
15 print(df[["STA_LOCAL", "STA_HOUR"]].head(10))

```

Listing 9.13: Conversión de STA_{LOCAL} y creación de STA_{HOUR}

```

1 import pandas as pd
2
3 # Eliminar columnas intermedias
4 for c in list(df.columns):
5     if c.startswith("wx_") or c in {"_f_date", "_f_hour"}:
6         df.drop(columns=c, inplace=True, errors="ignore")
7

```

```

8 df_weather.columns = (
9     df_weather.columns.str.strip().str.lower()
10     .str.replace(" ", "_").str.replace("-", "_")
11 )
12
13 df_weather["datetime_col"] = pd.to_datetime(df_weather["datetime_col"],
14     errors="coerce")
15 df_weather["airport"] =
16     df_weather["airport"].astype(str).str.strip().str.upper()
17
18 df_weather["_w_date"] = df_weather["datetime_col"].dt.date
19 df_weather["_w_hour"] = df_weather["datetime_col"].dt.hour
20
21 num_cols_weather = [
22     c for c in df_weather.select_dtypes(include="number").columns
23     if c not in {"airport", "_w_date", "_w_hour", "oaci", "datetime_utc", \
24     "datetime_col", "year"}
25 ]
26
27 agg = {c: "mean" for c in num_cols_weather}
28
29 # variables m ximas
30 for mx in ["wind_gust_kt", "gust_", "rain", "fog_brume", "thunder", "snow", \
31     "op_constraint"]:
32     if mx in agg:
33         agg[mx] = "max"
34
35 wx_hour = (
36     df_weather.groupby(["airport", "_w_date", "_w_hour"], as_index=False)
37     .agg(agg)
38     .rename(columns={c: f"wx_{c}" for c in num_cols_weather})
39 )

```

Listing 9.14: Normalización, renombre y agregación horaria de datos meteorológicos

```

1 df["STA_LOCAL"] = pd.to_datetime(df["STA_LOCAL"], errors="coerce")
2 df["DESTINO_ITI"] = df["DESTINO_ITI"].astype(str).str.strip().str.upper()
3
4 df["_f_date"] = df["STA_LOCAL"].dt.date
5 df["_f_hour"] = df["STA_HOUR"].astype("int16")
6
7 df_merged = (
8     pd.merge(
9         df,
10        wx_hour,
11        left_on=["DESTINO_ITI", "_f_date", "_f_hour"],
12        right_on=["airport", "_w_date", "_w_hour"],

```

```

13         how="left"
14     )
15     .drop(columns=["airport", "_w_date", "_w_hour"], errors="ignore")
16 )
17
18 wx_cols = [c for c in df_merged.columns if c.startswith("wx_")]
19 coverage = 1 - df_merged[wx_cols].isna().all(axis=1).mean()
20
21 print("Columnas wx_ presentes:", wx_cols)
22 print(f"Cobertura clima en merge: {coverage:.1%}")

```

Listing 9.15: Merge entre dataset de vuelos y clima

```

1 output_path = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
   Weather\weather_outputs\flights_weather_merged.parquet"
2 df_merged.to_parquet(output_path, engine="pyarrow", compression="snappy",
   index=False)
3
4 excel_path = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
   Weather\weather_outputs\flights_weather_merged.xlsx"
5 df_merged.to_excel(excel_path, index=False, engine="openpyxl")
6
7 print("Exportaciones completadas.")

```

Listing 9.16: Exportación a Parquet y Excel

```

1 import pandas as pd
2
3 TARGET_AP = "PSO"
4 TARGET_DATE = pd.Timestamp("2019-09-20").date()
5
6 dfq = df_merged.copy()
7 dfq["DESTINO_ITI"] =
   dfq["DESTINO_ITI"].astype(str).str.upper().str.strip()
8 dfq["STA_LOCAL"] = pd.to_datetime(dfq["STA_LOCAL"], errors="coerce")
9
10 dfq["arr_hour"] = dfq["STA_LOCAL"].dt.floor("H")
11 dfq["arr_date"] = dfq["arr_hour"].dt.date
12
13 mask = (dfq["DESTINO_ITI"] == TARGET_AP) & (dfq["arr_date"] ==
   TARGET_DATE)
14 pso_day = dfq.loc[mask].copy()
15
16 print(f"Total vuelos encontrados: {len(pso_day)}")
17
18 wx_cols = [c for c in pso_day.columns if c.startswith("wx_")]
19 has_wx = ~pso_day[wx_cols].isna().all(axis=1)

```

```

20
21 print("Filas con clima asociado:", has_wx.sum())

```

Listing 9.17: Consulta puntual de clima asociado a vuelos

```

1 import pandas as pd
2
3 leg_id = 35107329
4 row = df_merged.loc[df_merged["LEG_NO"] == leg_id].copy()
5
6 cols_vuelo = [
7     "LEG_NO", "ROUTE_OWN", "VUELO", "ORIGEN_ITI", "DESTINO_ITI",
8     "STD_LOCAL", "STA_LOCAL", "STA_HOUR", "_f_date", "_f_hour",
9     "DEMORA_DEP", "DEMORA_ARR"
10 ]
11
12 wx_cols = [c for c in row.columns if c.startswith("wx_")]
13
14 print("=== Vuelo ===")
15 print(row[cols_vuelo].to_string(index=False))
16
17 print("\n=== Clima asociado ===")
18 print(row[wx_cols].T.rename(columns={row.index[0]: "valor"}))

```

Listing 9.18: Consulta de clima y detalles a partir de un LEG_NO

```

1 import pandas as pd
2
3 mask_1970 = pd.to_datetime(df_merged["_f_date"],
4     errors="coerce").dt.year.eq(1970)
5 df_1970 = df_merged.loc[mask_1970]
6
7 print("Total filas con fecha 1970:", df_1970.shape[0])

```

Listing 9.19: Detección de fechas erróneas (ej. año 1970)

9.1.6. Anexo F: Modelo Random Forest: Entrenamiento y Evaluación

Este anexo presenta el código utilizado para entrenar y evaluar un modelo de Random Forest como parte de la fase comparativa inicial del proyecto. Incluye la preparación de datos, configuración del modelo, entrenamiento y cálculo de métricas de desempeño.

```

1 import pandas as pd
2 ruta = r"C:\Users\dmbuitrago\Daniel B\Proyecto - Maestra\BD
3     Weather\weather_outputs\flights_weather_merged.parquet"
4 df = pd.read_parquet(ruta)

```

```

5 print(df.shape)
6 print(df.head())

```

Listing 9.20: Carga del dataset integrado vuelos + clima

```

1 import pandas as pd
2 import numpy as np
3
4 print("Dimensiones del DataFrame:", df.shape)
5 df.info()
6
7 memoria = df.memory_usage(deep=True).sum() / (1024 ** 2)
8 print(f"Memoria usada: {memoria:,.2f} MB")

```

Listing 9.21: Información general, memoria y tipos de datos

```

1 faltantes = (
2     df.isna()
3     .sum()
4     .to_frame("nullos")
5     .assign(pct=lambda x: (x["nullos"] / len(df) * 100).round(2))
6     .sort_values("pct", ascending=False)
7 )
8
9 print(faltantes.head(15))
10 num_summary = df.describe(percentiles=[.01, .05, .5, .95, .99]).T

```

Listing 9.22: Valores faltantes y resumen estadístico

```

1 vuelos_anio = df["AO"].value_counts().sort_index()
2 vuelos_mes = df.groupby(["AO", "MES"]).size().unstack(fill_value=0)
3 vuelos_hora = df["STA_HOUR"].value_counts().sort_index()
4
5 print(vuelos_anio)
6 print(vuelos_mes)
7 print(vuelos_hora)

```

Listing 9.23: Distribución anual, mensual y horaria de vuelos

```

1 cols_weather = [c for c in df.columns if c.startswith("wx_")]
2 weather_numeric = df[cols_weather].select_dtypes(include=["number"])
3
4 corr_weather = weather_numeric.corr()
5 print(corr_weather.head())
6
7 import seaborn as sns
8 import matplotlib.pyplot as plt

```

```

9
10 plt.figure(figsize=(10,8))
11 sns.heatmap(corr_weather, cmap="coolwarm", center=0)
12 plt.title("Correlaciones entre variables meteorológicas")
13 plt.show()

```

Listing 9.24: Correlación entre variables meteorológicas

```

1 df["DIA_SEMANA"] = df["FECHA_LOCAL"].dt.dayofweek
2 df["SEMANA_AÑO"] = df["FECHA_LOCAL"].dt.isocalendar().week.astype(int)
3 df["TEMPORADA"] = np.where(df["MES"].isin([6,7,12,1]), "Alta", "Baja")
4
5 def bucket_hora(h):
6     if 5 <= h < 12: return "Mañana"
7     elif 12 <= h < 18: return "Tarde"
8     return "Noche"
9
10 df["HORA_BUCKET"] = df["STA_HOUR"].apply(bucket_hora)
11 df["RUTA"] = df["ORIGEN_ITI"] + "_" + df["DESTINO_ITI"]

```

Listing 9.25: Construcción de variables operativas y temporales

```

1 weather_mean = (
2     df.groupby(["DESTINO_ITI", "MES"]) \
3     [ ["wx_visibility_m", "wx_rain", "wx_temperature"] ]
4     .mean()
5     .reset_index()
6     .rename(columns={
7         "wx_visibility_m": "visibilidad_media_mes",
8         "wx_rain": "lluvia_media_mes",
9         "wx_temperature": "temp_media_mes" })
10 )
11
12 df = df.merge(weather_mean, on=["DESTINO_ITI", "MES"], how="left")

```

Listing 9.26: Construcción del dataset final para modelado

```

1 from sklearn.ensemble import RandomForestClassifier
2
3 rf = RandomForestClassifier(
4     n_estimators=300,
5     class_weight="balanced",
6     max_features="sqrt",
7     random_state=42,
8     n_jobs=-1
9 )
10

```

```

11 rf.fit(X_train, y_train)
12 y_pred = rf.predict(X_test)
13 y_proba = rf.predict_proba(X_test)[: ,1]

```

Listing 9.27: Entrenamiento del modelo Random Forest

```

1 from sklearn.metrics import accuracy_score, precision_score,
  recall_score, f1_score, roc_auc_score
2
3 print("Accuracy:", accuracy_score(y_test, y_pred))
4 print("Precision:", precision_score(y_test, y_pred))
5 print("Recall:", recall_score(y_test, y_pred))
6 print("F1:", f1_score(y_test, y_pred))
7 print("ROC-AUC:", roc_auc_score(y_test, y_proba))

```

Listing 9.28: Métricas del modelo Random Forest

```

1 def preparar_fila_escenario(df, X_cols, destino="BOG", mes=11,
  sta_hour=15,
2
3     dia_semana=2, override=None):
  ...

```

Listing 9.29: Funciones para generar escenarios con clima

```

1 row_base = preparar_fila_escenario(df, X_train.columns, destino="BOG",
  mes=11, sta_hour=15)
2 resultado_base = score_escenario(rf, row_base)
3
4 row_adverso = preparar_fila_escenario(
5     df, X_train.columns, destino="BOG", mes=11, sta_hour=15,
6     override={"wx_visibility_m":1500, "wx_rain":100}
7 )
8
9 resultado_adverso = score_escenario(rf, row_adverso)

```

Listing 9.30: Predicción para escenario base y adverso

```

1 row_base = preparar_fila_escenario(df, X_train.columns, destino="BOG",
  mes=11, sta_hour=15)
2 resultado_base = score_escenario(rf, row_base)
3
4 row_adverso = preparar_fila_escenario(
5     df, X_train.columns, destino="BOG", mes=11, sta_hour=15,
6     override={"wx_visibility_m":1500, "wx_rain":100}
7 )
8
9 resultado_adverso = score_escenario(rf, row_adverso)

```

Listing 9.31: Predicción para escenario base y adverso

```

1 for dest in ["BOG", "MDE", "CLO"]:
2     row = preparar_fila_escenario(df, X_train.columns, destino=dest,
3     mes=11, sta_hour=15)
4     print(dest, score_escenario(rf, row))

```

Listing 9.32: Comparación de probabilidades OTP entre aeropuertos

```

1 df_pso_mes = evaluar_riesgo_mensual(df, X_train.columns, rf,
2     destino="PSO", hora=15)
3 print(df_pso_mes)

```

Listing 9.33: Probabilidad mensual estimada de puntualidad

9.1.7. Anexo G: Modelo XGBoost Final: Optimización para la Clase Retrasado

Este anexo contiene el código correspondiente al entrenamiento del modelo XGBoost final, incluyendo la redefinición de la variable objetivo para priorizar la detección de la clase retrasado, el ajuste del parámetro `scale_pos_weight` para manejar el desbalance, la búsqueda de hiperparámetros mediante `RandomizedSearchCV` optimizando el F1 de la clase positiva y la selección del umbral óptimo de decisión. Este modelo corresponde a la versión definitiva empleada en el proyecto para la estimación de riesgo de retraso.

```

1 import pandas as pd
2
3 ruta = r"C:\Users\dmbuitrago\ Daniel B\Proyecto - Maestra\BD
4     Weather\weather_outputs\flights_weather_merged.parquet"
5
6 df = pd.read_parquet(ruta)
7
8 print(df.shape)
9 print(df.head())

```

Listing 9.34: Carga y exploración inicial del dataset

```

1 import pandas as pd
2 import numpy as np
3 from sklearn.model_selection import train_test_split
4
5 df = df.copy()
6
7 # Conversión de fecha
8 if not np.issubdtype(df["FECHA_LOCAL"].dtype, np.datetime64):
9     df["FECHA_LOCAL"] = pd.to_datetime(df["FECHA_LOCAL"], errors="coerce")
10

```

```

11 df["DIA_SEMANA"] = df["FECHA_LOCAL"].dt.dayofweek
12 df["TEMPORADA"] = np.where(df["MES"].isin([12,1,6,7]), "Alta", "Baja")
13
14 def hora_bucket(h):
15     try:
16         h = int(h)
17     except:
18         return "Noche"
19     if 5 <= h < 12: return "Ma ana"
20     elif 12 <= h < 18: return "Tarde"
21     else: return "Noche"
22
23 df["HORA_BUCKET"] = df["STA_HOUR"].apply(hora_bucket)

```

Listing 9.35: Procesamiento inicial y feature engineering

```

1 needed_cols = ["visibilidad_media_mes", "lluvia_media_mes",
2               "temp_media_mes", "vuelos_dia_destino"]
3
4 # Si faltan columnas clim ticas base
5 weather_mean = (
6     df.groupby(["DESTINO_ITI",
7               "MES"])["wx_visibility_m", "wx_rain", "wx_temperature"]
8     .mean().reset_index()
9     .rename(columns={
10         "wx_visibility_m": "visibilidad_media_mes",
11         "wx_rain": "lluvia_media_mes",
12         "wx_temperature": "temp_media_mes"
13     })
14 )
15 df = df.merge(weather_mean, on=["DESTINO_ITI", "MES"], how="left")
16
17 trafico = (
18     df.groupby(["DESTINO_ITI", "FECHA_LOCAL"])["LEG_NO"]
19     .count().reset_index().rename(columns={"LEG_NO": "vuelos_dia_destino"})
20 )
21 df = df.merge(trafico, on=["DESTINO_ITI", "FECHA_LOCAL"], how="left")

```

Listing 9.36: Creación de variables climáticas y de congestión

```

1 cols_weather = [
2     "wx_visibility_m", "wx_rain", "wx_temperature",
3     "wx_fog_brume", "wx_thunder", "wx_op_constraint"
4 ]
5
6 context_cols = [

```

```
7     "MES", "DIA_SEMANA", "TEMPORADA", "HORA_BUCKET",
8     "REGION DESTINO", "DESTINO_ITI",
9     "visibilidad_media_mes", "lluvia_media_mes", "temp_media_mes",
10    "vuelos_dia_destino"
11 ]
12
13 target = "CUMPLE_OTP"
14 df[target] = df[target].astype(int)
15
16 df_model = df[cols_weather + context_cols + [target]].copy()
17 num_cols = df_model.select_dtypes(include=["number"]).columns
18 df_model[num_cols] =
19     df_model[num_cols].fillna(df_model[num_cols].median())
20
21 cat_cols = ["TEMPORADA", "HORA_BUCKET", "REGION DESTINO", "DESTINO_ITI"]
22
23 X = pd.get_dummies(df_model.drop(columns=[target]), columns=cat_cols,
24     drop_first=True)
25 y = df_model[target]
26
27 X_train, X_test, y_train, y_test = train_test_split(
28     X, y, test_size=0.20, random_state=42, stratify=y
29 )
```

Listing 9.37: Preparación de dataset final

```
1 from xgboost import XGBClassifier
2 from sklearn.metrics import accuracy_score, precision_score,
3     recall_score, f1_score, roc_auc_score
4
5 xgb = XGBClassifier(
6     n_estimators=200,
7     learning_rate=0.05,
8     max_depth=6,
9     subsample=0.8,
10    colsample_bytree=0.8,
11    random_state=42,
12    n_jobs=-1,
13    objective="binary:logistic",
14    eval_metric="logloss"
15 )
16
17 xgb.fit(X_train, y_train)
18 y_pred = xgb.predict(X_test)
```

Listing 9.38: Modelo base XGBoost

```
1 from sklearn.model_selection import RandomizedSearchCV
2 from scipy.stats import randint, uniform
3
4 param_dist = {
5     "n_estimators": randint(150, 400),
6     "max_depth": randint(4, 9),
7     "learning_rate": uniform(0.03, 0.12),
8     "subsample": uniform(0.7, 0.3),
9     "colsample_bytree": uniform(0.7, 0.3)
10 }
11
12 search = RandomizedSearchCV(
13     estimator=xgb,
14     param_distributions=param_dist,
15     n_iter=25,
16     scoring="f1",
17     cv=3,
18     random_state=42,
19     verbose=2
20 )
21
22 search.fit(X_train, y_train)
```

Listing 9.39: RandomizedSearchCV XGBoost

```
1 from sklearn.metrics import classification_report, confusion_matrix
2
3 cm = confusion_matrix(y_test, y_pred)
4 print(cm)
5
6 print(classification_report(y_test, y_pred, digits=3))
```

Listing 9.40: Métricas de evaluación

```
1 from sklearn.metrics import roc_curve, roc_auc_score
2 import matplotlib.pyplot as plt
3
4 fpr, tpr, thr = roc_curve(y_test, y_pred)
5 auc = roc_auc_score(y_test, y_pred)
6
7 plt.figure(figsize=(6,6))
8 plt.plot(fpr, tpr, label=f"AUC={auc:.3f}")
9 plt.plot([0,1],[0,1],"--")
10 plt.xlabel("FPR"); plt.ylabel("TPR")
11 plt.show()
```

Listing 9.41: Curva ROC del modelo

```
1 import pandas as pd
2
3 aeropuertos = df.loc[X_test.index, "DESTINO_ITI"]
4 # ... (c digo de agrupaci n y m tricas)
5 print(df_conf_ap.head(15))
```

Listing 9.42: Métricas por aeropuerto

```
1 import seaborn as sns
2 import matplotlib.pyplot as plt
3
4 df_melt = df_conf_ap.melt(
5     id_vars="AEROPUERTO",
6     value_vars=["Precision", "Recall", "F1", "Accuracy"],
7     var_name="Metrica",
8     value_name="Valor"
9 )
10
11 sns.heatmap(df_melt.pivot(...))
12 plt.show()
```

Listing 9.43: Mapa de calor de desempeño por aeropuerto

```
1 import joblib
2
3 joblib.dump(xgb_final, "modelo_xgb_delay.pkl")
4 joblib.dump(best_thr, "umbral_delay.pkl")
5 joblib.dump(list(X.columns), "feature_names.pkl")
```

Listing 9.44: Exportación del modelo y umbral óptimo

Anexo H: Aplicación Streamlit para Predicción de Retrasos

A continuación se presenta el código correspondiente a la aplicación construida en Streamlit para la visualización y predicción de retrasos en vuelos con base en el modelo XGBoost entrenado.

```
1 import streamlit as st
2 import pandas as pd
3 import numpy as np
4 import joblib
5 import matplotlib.pyplot as plt
6
7 st.set_page_config(page_title="Prediccin de Retrasos", layout="wide")
```

Listing 9.45: Configuración inicial y librerías de la aplicación Streamlit

```
1 def preprocesar(df_raw: pd.DataFrame):
2     df = df_raw.copy()
3     df["FECHA_LOCAL"] = pd.to_datetime(df["FECHA_LOCAL"], errors="coerce")
4     df["DIA_SEMANA"] = df["FECHA_LOCAL"].dt.dayofweek
5     df["TEMPORADA"] = np.where(df["MES"].isin([12, 1, 6, 7]), "Alta",
6                               "Baja")
7
8     def hora_bucket(h):
9         try:
10            h = int(h)
11        except Exception:
12            return "Noche"
13        if 5 <= h < 12:
14            return "Maana"
15        elif 12 <= h < 18:
16            return "Tarde"
17        else:
18            return "Noche"
19
20    df["HORA_BUCKET"] = df["STA_HOUR"].apply(hora_bucket)
21
22    weather_mean = (
23        df.groupby(["DESTINO_ITI", "MES"])[
24            ["wx_visibility_m", "wx_rain", "wx_temperature"]
25        ]
26        .mean()
27        .reset_index()
28        .rename(columns={
29            "wx_visibility_m": "visibilidad_media_mes",
30            "wx_rain": "lluvia_media_mes",
31            "wx_temperature": "temp_media_mes"
32        })
33    )
34
35    df = df.merge(weather_mean, on=["DESTINO_ITI", "MES"], how="left")
36
37    trafico = (
38        df.groupby(["DESTINO_ITI", "FECHA_LOCAL"])["LEG_NO"]
39        .count()
40        .reset_index()
41        .rename(columns={"LEG_NO": "vuelos_dia_destino"})
42    )
43
44    df = df.merge(trafico, on=["DESTINO_ITI", "FECHA_LOCAL"], how="left")
45
46    cols_weather = [
47        "wx_visibility_m", "wx_rain", "wx_temperature",
48        "wx_fog_brume", "wx_thunder", "wx_op_constraint"
49    ]
```

```

46
47     base_cols = cols_weather + [
48         "MES", "DIA_SEMANA",
49         "visibilidad_media_mes", "lluvia_media_mes", "temp_media_mes",
50         "vuelos_dia_destino"
51     ]
52     cat_cols = ["TEMPORADA", "HORA_BUCKET", "REGION DESTINO",
53               "DESTINO_ITI"]
54
55     df_model = df[base_cols + cat_cols].copy()
56     num_cols = df_model.select_dtypes(include=["number"]).columns
57     df_model[num_cols] =
58         df_model[num_cols].fillna(df_model[num_cols].median())
59
60     X = pd.get_dummies(df_model, columns=cat_cols, drop_first=True)
61     return df, X

```

Listing 9.46: Función de preprocesamiento de datos

```

1 @st.cache_resource
2 def cargar_modelo():
3     modelo = joblib.load("modelo_xgb_delay.pkl")
4     umbral = joblib.load("umbral_delay.pkl")
5     feature_names = joblib.load("feature_names.pkl")
6     return modelo, umbral, feature_names
7
8 modelo, umbral, feature_names = cargar_modelo()

```

Listing 9.47: Carga del modelo, umbral y características

```

1 st.title(" Dashboard de Prediccin de Retrasos por Clima en Colombia")
2 st.caption(
3     "Modelo XGBoost entrenado con datos historicos de vuelos y clima en
4     Colombia. "
5     "La clase positiva es retrasado (1), calibrada con un umbral
6     ptimo."
7 )

```

Listing 9.48: Sección principal del Dashboard

```

1 st.sidebar.header("Carga de datos")
2 tipo_archivo = st.sidebar.radio("Tipo de archivo a cargar:", ["Parquet",
3                       "CSV"])
4
5 archivo = st.sidebar.file_uploader(
6     "Sube un archivo de vuelos (mismo formato del training):",
7     type=["parquet", "csv"],

```

7)

Listing 9.49: Carga de datos desde el usuario

```

1 if archivo is not None:
2     if tipo_archivo == "Parquet":
3         df_raw = pd.read_parquet(archivo)
4     else:
5         df_raw = pd.read_csv(archivo)
6
7     st.subheader("Vista previa de los datos cargados")
8     st.dataframe(df_raw.head(), use_container_width=True)
9
10    df_proc, X_new = preprocesar(df_raw)
11
12    for col in feature_names:
13        if col not in X_new.columns:
14            X_new[col] = 0
15    X_new = X_new[feature_names]
16
17    proba_delay = modelo.predict_proba(X_new)[:, 1]
18    pred_delay = (proba_delay >= umbral).astype(int)
19
20    df_proc["prob_retraso"] = proba_delay
21    df_proc["retraso_modelo"] = pred_delay
22
23    st.subheader(" Indicadores globales")
24    col1, col2, col3 = st.columns(3)
25
26    total_vuelos = len(df_proc)
27    pct_flag = df_proc["retraso_modelo"].mean() * 100
28    prom_prob = df_proc["prob_retraso"].mean() * 100
29
30    col1.metric("Vuelos evaluados", f"{total_vuelos:,}")
31    col2.metric("% vuelos con riesgo alto de retraso",
32                f"{pct_flag:0.1f}%")
32    col3.metric("Probabilidad promedio de retraso", f"{prom_prob:0.1f}%")

```

Listing 9.50: Proceso de predicción y métricas globales

```

1     st.subheader(" Exploracin por aeropuerto")
2     destinos = sorted(df_proc["DESTINO_ITI"].dropna().unique().tolist())
3
4     destino_sel = st.selectbox(
5         "Filtrar por aeropuerto destino",
6         options=["(Todos)"] + destinos
7     )

```

Listing 9.51: Exploración interactiva por aeropuerto

```

1  st.subheader(" Heatmap de puntualidad por aeropuerto y mes")
2
3  if "MES" not in df_proc.columns:
4      df_proc["MES"] = df_proc["FECHA_LOCAL"].dt.month
5
6  destinos_hm =
7      sorted(df_proc["DESTINO_ITI"].dropna().unique().tolist())
8
9  opcion_hm = st.selectbox(
10     "Selecciona aeropuerto destino para el mapa de calor (o
11     'Todos'):",
12     options=["(Todos)"] + destinos_hm,
13     key="heatmap_ap",
14 )

```

Listing 9.52: Construcción del heatmap de puntualidad

```

1  st.subheader(" Descargar resultados")
2
3  df_export = df_proc[
4      [
5          "LEG_NO", "FECHA_LOCAL",
6          "ORIGEN_ITI", "DESTINO_ITI",
7          "prob_retraso", "retraso_modelo"
8      ]
9  ].copy()
10
11  csv = df_export.to_csv(index=False).encode("utf-8")
12
13  st.download_button(
14     " Descargar CSV con predicciones",
15     data=csv,
16     file_name="predicciones_retraso.csv",
17     mime="text/csv",
18 )
19  else:
20  st.info("Sube un archivo de vuelos en la barra lateral para
21  comenzar.")

```

Listing 9.53: Generación y descarga del CSV con predicciones

Bibliografía

- [1] Aeronáutica Civil de Colombia, “Boletines operacionales.” https://www.aerocivil.gov.co/centro_pensamiento/publicaciones/3881/boletines-operacionales/, 2024. Accessed: 2024-12-07.
- [2] Ministerio de Transporte de Colombia, “Cifras récord para el transporte aéreo: primer semestre de 2022 cierra con 22 683 175 pasajeros movilizados y 454 032 toneladas de carga transportada.” <https://mintransporte.gov.co/publicaciones/11019/cifras-record-para-el-transporte-aereo-primer-semestre-de-2022-cierra-con-22683175-pasajero> 2022. Accessed: 2024-12-07.
- [3] “Global outlook for air transport: Highly resilient, less robust,” 2023.
- [4] Flightradar24 AB, “Flightradar24 – live flight tracker.” <https://www.flightradar24.com/>, 2024. Accessed: 2024-12-07.
- [5] “The monthly on-time performance report – october 2024.” <https://www.cirium.com/thoughtcloud/cirium-monthly-on-time-performance-reports2024/>, 2024. Accedido: 07 diciembre 2024.
- [6] M. Mamdouh, M. Ezzat, and H. Hefny, “A novel intelligent approach for flight delay prediction,” *Journal of Big Data*, vol. 10, no. 1, 2023.
- [7] A. Lahouar and J. B. H. Slama, “Random forests model for one day ahead load forecasting,” in *Proceedings of the 2015 6th International Renewable Energy Congress (IREC)*, IEEE, 2015.
- [8] S. Kumar and A. Singh, “Flight delay prediction using xgboost with operational and weather data,” in *Proceedings of the IEEE International Conference on Data Science*, pp. 1–6, 2018.
- [9] I. Gultepe *et al.*, “A review of high impact weather for aviation meteorology,” *Pure and Applied Geophysics*, vol. 176, no. 5, pp. 1869–1921, 2019.
- [10] M. de Oliveira, A. B. R. Eufrásio, M. X. Guterres, M. C. R. Murça, and R. d. A. Gomes, “Analysis of airport weather impact on on-time performance of arrival flights for the brazilian domestic air transportation system,” *Journal of Air Transport Management*, vol. 91, 2021.
- [11] “Consulta metar - ideam.” <http://www.meteoaeronautica.gov.co/consulta-metar>, 2024. Accedido: 07 diciembre 2024.
- [12] M. Fathi, M. Hagh Kashani, S. M. Jameii, and E. Mahdipour, “Big data analytics in weather forecasting: A systematic review,” *Archives of Computational Methods in Engineering*, 2022.
- [13] “Climate data online (cdo).” <https://www.ncei.noaa.gov/cdo-web/>, 2024. Accedido: 07 diciembre 2024.

-
- [14] “Climodat reports.” <https://mesonet.agron.iastate.edu/climodat/>, 2024. Accedido: 07 diciembre 2024.
- [15] R. Kothari *et al.*, “Selection of best machine learning model to predict delay in passenger airlines,” *IEEE Access*, 2017.
- [16] X. Zhang, Y. Li, and H. Wang, “Flight delay prediction using random forest based on meteorological factors,” *Journal of Air Transport Management*, vol. XX, pp. XX–XX, 2020.
- [17] International Air Transport Association, “Global air transport performance report 2021,” 2021.
- [18] X. Zhao and X. Sun, “Multi-class flight delay prediction based on an improved xgboost model,” in *2025 IEEE 8th Information Technology and Mechatronics Engineering Conference (ITOEC)*, vol. 8, pp. 1566–1570, 2025.
- [19] W. Luo, “Prediction of flight delays based on the xgboost model,” in *2024 International Conference on Computers, Information Processing and Advanced Education (CIPAE)*, pp. 104–109, 2024.
- [20] N. L. Kalyani, G. Jeshmitha, B. S. Sai U., M. Samanvitha, J. Mahesh, and B. Kiranmayee, “Machine learning model - based prediction of flight delay,” in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 577–581, 2020.
- [21] N. Autor, “Procesamiento, análisis y visualización del comportamiento de variables climáticas medidas desde enero de 2014 hasta diciembre de 2022 y pronóstico de los siguientes 12 meses,” *Revista REINCISOL*, vol. X, no. Y, p. Z, 2023.