

Evaluación de covariables exógenas en un modelo de series temporales para la predicción de eventos de Dengue en Cali, Valle del Cauca.

Evaluation of exogenous covariables in a time series model for the prediction of dengue events in Cali, Valle del Cauca.

Trabajo de grado para optar al título de Biólogo

Autor: Juan David Castillo Hurtado ^{1*}

Director: David Arango ²

1 Programa de Biología, Facultad de Ingeniería y Ciencias, Pontificia Universidad Javeriana Cali, Colombia.

2 Departamento de Ciencias Naturales y Matemáticas, Facultad de Ingeniería y Ciencias, Pontificia Universidad Javeriana Cali, Colombia.

* Correo electrónico: jdcastillo@javerianacali.edu.co

Resumen Este trabajo evalúa si la incorporación de covariables exógenas mejora la predicción semanal de casos de dengue en Cali (Colombia). Se construyeron y compararon cuatro modelos de series de tiempo: SARIMA univariado (A) y tres SARIMAX con predictores exógenos (B: meteorología semanal agregada de NASA POWER; C: índices de búsqueda en Google Trends; y D: Combina B y C). El periodo de datos fue 2007-2024; la validación fuera de muestra se realizó año a año para 2021-2024 mediante un esquema rolling-origin con ventana móvil. El ajuste consideró transformación Box-Cox y selección automática de órdenes; la evaluación incluyó métricas puntuales (MAE, RMSE, MAPE, MASE), desempeño probabilístico (coberturas 80/95 % e *interval score*) y pruebas de contraste: Diebold-Mariano contra una línea base seasonal naive estacional y Ljung-Box sobre residuos. En 2021-2022 los modelos con exógenas redujeron errores y mostraron DM favorables frente a snaive, evidenciando valor añadido en regímenes estacionales “típicos”. En 2023, año epidémico atípico, el rendimiento se degradó y la línea base igualó o superó a los modelos, apuntando a la necesidad de capturar no linealidades y cambios de régimen. En 2024 se recuperó la calibración (altas coberturas) a costa de intervalos más anchos, lo que sugiere incertidumbre ampliada tras el quiebre de 2023. Globalmente, clima y búsquedas aportan señales complementarias (físico-biológica y conductual), y su integración es la vía más prometedora para un sistema de alerta temprana robusto; metodológicamente se recomienda modelar umbrales/no linealidades, incorporar forzantes de gran escala (ENSO) y reentrenar tras años extremos.

Palabras clave: Dengue; Series temporales; ARIMA; Variables exógenas; Clima; Google Trends; Cali (Colombia); ENSO; Seasonal naive.

Abstract This study evaluates whether the incorporation of exogenous covariates improves the weekly prediction of dengue cases in Cali (Colombia). Four time series models were constructed and compared: univariate SARIMA (A) and three SARIMAX models with exogenous predictors (B: weekly aggregate meteorology from NASA POWER; C: search indices from Google Trends; and D: combines B and C). The data period was 2007-2024; out-of-sample validation was performed year by year for 2021-2024 using a rolling-origin scheme with a moving window. The adjustment considered Box-Cox transformation and automatic order selection; the evaluation included point metrics (MAE, RMSE, MAPE, MASE), probabilistic performance (80/95% coverage and interval score), and contrast tests: Diebold-Mariano against a seasonal naive baseline and Ljung-Box on residuals. In 2021-2022, models with exogenous variables reduced errors and showed favorable DM compared to naive models, demonstrating added value in “typical” seasonal regimes. In 2023, an atypical epidemic year, performance deteriorated and the baseline equaled or surpassed the models, pointing to the need to capture nonlinearities and regime shifts. In 2024, calibration (high coverage) recovered at the expense of wider intervals, suggesting increased uncertainty following the break in 2023. Globally, climate and searches provide complementary signals (physical-biological and behavioral), and their integration is the most promising path to a robust early warning system; methodologically, it is recommended to model thresholds/nonlinearities, incorporate large-scale forcings (ENSO), and retrain after extreme years.

Keywords: Dengue; Time series; ARIMA; Exogenous variables; Climate; Google Trends; Cali (Colombia); ENSO; Seasonal naive.

INTRODUCCIÓN

El dengue es una enfermedad viral causada por el virus del dengue (género *Flavivirus*) y transmitida por mosquitos de acuerdo con las directrices de la Organización Mundial de la Salud (OMS, 2009). Se presenta con síntomas como fiebre alta, dolor de cabeza, dolor muscular y articular, etc., y, en casos graves, puede evolucionar a dengue hemorrágico (Chagas et al., 2022). La transmisión se produce a través de la picadura de mosquitos del género *Aedes*—en particular *Aedes aegypti* (Linnaeus, 1762)—vector predominante en áreas urbanas y suburbanas, que se reproduce en recipientes con agua estancada y presenta mayor actividad en horas de la mañana y finales de la tarde.

Según Bhatt et al. (2013), la enfermedad es endémica en más de 128 países, siendo Asia la región más afectada; el dengue representa un problema significativo de salud pública a nivel mundial, cerca de 390 millones de personas se infectan con el virus del dengue cada año. En las últimas décadas se ha observado un aumento en la incidencia de la enfermedad, asociado a factores como el cambio climático, la urbanización no planificada y la globalización, que han facilitado la expansión del virus hacia nuevas áreas. (Areeb Hussain & Dhiman, 2022). En América, el dengue es la arbovirosis—enfermedad transmitida por artrópodos hematófagos—más frecuente, con ciclos epidémicos cada 3 a 5 años y más de 3 millones de casos reportados en 2019 y 2023, junto con incrementos recientes en incidencia y letalidad (Instituto Nacional de Salud, 2024). En Colombia, el dengue es considerado un problema recurrente, con patrones epidémicos cíclicos y picos importantes en 2010, 2013, 2016, 2019 y 2023 (Instituto Nacional de Salud, 2024). En este contexto, territorios urbanos como Cali, en el Valle del Cauca, se han visto afectados de manera recurrente, lo que resalta la necesidad de contar con herramientas que apoyen la anticipación de picos de transmisión y la planificación temprana de acciones de control.

De acuerdo con el CDC—Centro para el Control y Prevención de Enfermedades—(2006), la vigilancia en salud pública implica la recolección, análisis, interpretación y diseminación sistemática y continua de datos, proceso que muchas veces se ve obstaculizado por retrasos en la notificación, baja sensibilidad en la detección de casos y escasa cobertura poblacional. Los sistemas tradicionales de vigilancia epidemiológica son la base del monitoreo en salud pública, y en la actualidad estos sistemas de vigilancia enfrentan serias limitaciones para responder a enfermedades como el dengue, por ejemplo, operan con tiempos de respuesta lentos, lo cual afecta su capacidad para proporcionar información útil en tiempo real (Weir & Mykhalovskiy, 2009). En el caso colombiano, se ha descrito que pueden transcurrir varias semanas entre la aparición del caso y su reporte definitivo, lo que reduce la utilidad de la información para la toma de decisiones en tiempo real y limita la capacidad de los modelos para emitir alertas tempranas, estos problemas de calidad, oportunidad y completitud de los datos impactan directamente la construcción y calibración de modelos predictivos. (Reich et al., 2016; Instituto Nacional de Salud, 2024).

Frente a estas limitaciones, aumenta el interés por modelos estadísticos capaces de aprovechar series temporales de eventos de dengue e incorporar factores contextuales, usando modelos SARIMA—Modelo Autorregresivo Integrado de Media Móvil Estacional—para capturar patrones de tendencia y estacionalidad, en algunos casos integrando variables meteorológicas como temperatura, humedad y precipitación, que influyen sobre la transmisión del virus (Hasan et al., 2024; Alam et al., 2025). Otros trabajos han mostrado que modelos de aprendizaje y enfoques híbridos pueden mejorar el desempeño predictivo cuando se incluyen factores climáticos y otros determinantes contextuales (Jain et al., 2019; Cheng & Moraga, 2024; Sebastianelli et al., 2024). No obstante, estas metodologías enfrentan desafíos relacionados con la calidad y disponibilidad de los datos, lo que evidencia la necesidad de desarrollar modelos robustos,

adaptables y capaces de operar con datos incompletos o retrasados (Reich et al., 2016; Sebastianelli et al., 2024).

En este escenario, ha surgido una línea de investigación centrada en el uso de datos alternativos que complementen a los registros clínicos. Entre estos se incluyen las tendencias de búsqueda en Google, variables meteorológicas de alta resolución, indicadores socioeconómicos, movilidad poblacional y actividad en redes sociales (Jain et al., 2019; Tuan et al., 2024; Xiao et al., 2024). Estos enfoques se desarrollan en el concepto de nowcasting, que busca estimar en tiempo casi real la incidencia de enfermedades integrando fuentes de información que están disponibles de manera más oportuna que reportes oficiales (Reich et al., 2016). Las tendencias de búsqueda en motores de búsqueda—en particular Google—pueden funcionar como indicadores tempranos de actividad epidemiológica, dado que las personas tienden a buscar información sobre síntomas o sobre la enfermedad antes o en paralelo a la consulta médica (Xiao et al., 2024). Diversos estudios han mostrado el potencial de combinar datos de vigilancia con datos de búsquedas en línea para mejorar las estimaciones de dengue; Xiao et al. (2024)—por ejemplo—desarrollaron en Brasil un modelo de nowcasting integrando casos oficiales y tendencias de búsqueda en Google, mejorando la precisión de las estimaciones semanales frente a modelos tradicionales. Sebastianelli et al. (2024) propusieron un modelo de aprendizaje automático capaz de predecir brotes en Brasil y otros países como Perú. Pese a la evidencia internacional, son escasos los estudios que—en el contexto colombiano y en particular en Cali—integren de manera explícita datos biológicos (eventos notificados), variables meteorológicas y tendencias en motores de búsqueda, dentro de un mismo modelamiento de series temporales.

En este contexto, la presente investigación busca desarrollar varios modelos predictivos—basados en series temporales—que combinen eventos de dengue con variables meteorológicas e índices de búsqueda en Google; permitiendo así comparar la precisión entre modelos que combinen de diferente manera las variables, y evaluando cada una en el contexto local de Cali, Valle del Cauca.

MATERIALES Y MÉTODOS

Diseño del estudio y periodo

Se realizaron pronósticos de los eventos de dengue—con temporalidad semanal—en el municipio de Santiago de Cali (Colombia), con horizonte fuera de muestra para los años calendario 2021-2024 en el software R, versión 4.5.1 (R Core Team, 2014). Para cada año t se entrenaron los modelos con todo el historial disponible desde 2007 hasta $t - 1$ —año previo a nuestro año de interés—y posteriormente se generaron pronósticos rolling origin para todas las semanas comprendidas entre el primer y el último lunes de t . (Tashman, 2000; Hyndman & Athanasopoulos, 2021).

Fuentes de información y preprocesamiento

Para el desarrollo del estudio se toman como datos de interés—variables y covariables—la cantidad de eventos de dengue, temperatura (máxima, mínima y promedio), humedad relativa, precipitación y el índice de interés de Google Trends (<https://trends.google.com/trends/>).

Los registros de dengue se obtuvieron del Sistema Nacional de Vigilancia en Salud Pública de Colombia (SIVIGILA) mediante el paquete sivirep (Gómez-Millán et al., 2019). Las fechas de inicio de síntomas se transformaron a clase Date y se agregaron a semanas epidemiológicas ancladas al lunes mediante funciones de lubridate (Spinu et al., 2024), posteriormente se construyó un calendario de referencia con todos los lunes desde el primer lunes de 2007 hasta el último lunes de 2024 y se realizó una unión (*left join*) con los

conteos semanales; las semanas sin notificación se imputaron como 0 casos, garantizando una serie semanal completa y coherente (Gómez-Millán et al., 2019).

Las variables meteorológicas se obtuvieron del proyecto Prediction Of Worldwide Energy Resources (POWER) de la NASA—desarrollado por el NASA Langley Research Center y la División de Ciencias de la Tierra (NASA POWER Project, 2023)— a partir del servicio de datos diarios por punto de la API NASA POWER, que entrega series temporales diarias para coordenadas geográficas específicas, solicitando datos para el municipio de Cali con las coordenadas (3.44° N, 76.52° O) en el periodo 2007-2024. Las series diarias se agregaron a frecuencia semanal utilizando el mismo calendario de lunes definido para la serie de dengue, se tomaron promedios semanales para temperaturas y humedad y sumas semanales para la precipitación. Eventos vacíos se rellenaron mediante procedimientos simples de *carry-forward* y *carry-backward* (arrastre hacia adelante/atrás) con el objetivo de evitar cortes en las series utilizadas como covariables y en las visualizaciones (Xiao et al., 2024; Alam et al., 2025). Dado que la transmisión del dengue está mediada por la dinámica del vector y los tiempos de incubación, se asume que la influencia de temperatura, humedad y precipitación sobre los casos notificados se manifiesta con un rezago de varias semanas. Diversos estudios han documentado asociaciones significativas entre variables meteorológicas y la incidencia de dengue con lags cortos, por lo general entre 2 y 8 semanas (Xiao et al., 2024; Alam et al., 2025). En este trabajo, esa evidencia motivó la construcción de covariables meteorológicas con rezagos de 2 a 8 semanas.

La información sobre búsquedas relacionadas con dengue se obtuvo de la interfaz web de Google Trends (GT), restringiendo las consultas al término “dengue” (Khairul et al., 2021) y a la región geográfica correspondiente al Valle del Cauca para el periodo 2007-2024. La plataforma entrega un índice de interés de búsqueda entre 0 y 100, normalizado en función del máximo valor observado en el intervalo analizado, de modo que representa la intensidad relativa de las búsquedas en el tiempo y no conteos absolutos (Gluskin et al., 2024). Los datos se descargaron en formato CSV. Dado que, para ventanas largas de tiempo, GT entrega series agregadas a frecuencia mensual, cada valor mensual de *gt_dengue* se asignó a todas las semanas cuyo lunes pertenece a ese mes, generando una serie semanal definida sobre el mismo calendario de lunes utilizado para los eventos de dengue. Semanas sin información explícita se completaron mediante procedimientos simples de arrastre hacia adelante y hacia atrás (*carry-forward/carry-backward*), con el fin de evitar interrupciones en la serie utilizada como covariable. A partir de esta serie semanal se construyeron covariables exógenas con rezagos de 2 a 8 semanas, en concordancia con la evidencia que muestra que los patrones de búsqueda en línea se correlacionan con la incidencia de dengue y pueden anticipar la dinámica de transmisión en ventanas de pocas semanas. (Gluskin et al., 2014; Guo et al., 2017; Yang et al., 2017; Xiao et al., 2024).

Especificación de los modelos de series de tiempo

Se construyeron cuatro modelos sobre la misma base de eventos semanales, un modelo base univariante (Modelo A), un modelo con covariables meteorológicas (Modelo B), un modelo con covariables correspondiente a índices de búsqueda de GT (Modelo C), y un modelo que combina los variables del modelo B y C (Modelo D). Todos se enmarcan en la familia ARIMA—Modelo Autorregresivo Integrado de Media Móvil—estacional y se ajustaron siguiendo los principios expuestos por Hyndman & Athanasopoulos (2021).

El Modelo A se define como un modelo SARIMA—Modelo Autorregresivo Integrado de Media Móvil Estacional—aplicado a la serie semanal de casos y_t sin covariables exógenas. Se aplicó una transformación Box-Cox a $(y_t + 1)$, con el parámetro λ estimado automáticamente por máxima verosimilitud a $\lambda \geq 0$. Esta transformación se utiliza para estabilizar la varianza y aproximar la distribución de los errores a la normalidad, mientras que el término $+1$ evita problemas numéricos en semanas con conteos cero (Hyndman &

Athanasopoulos, 2021; Hyndman et al., 2025). Sobre la serie transformada se aplicaron las funciones `ndiffs()` y `nsdiffs()` del paquete `forecast` para sugerir el número de diferenciaciones ordinarias y estacionales necesarias para lograr estacionariedad. Con base en esas sugerencias, se ajustó un modelo $SARIMA(p, d, q) \times (P, D, Q)_{52}$ utilizando `auto.arima()`, imponiendo al menos una diferenciación estacional ($D \geq 1$), sin término de *drift* ni media, y utilizando búsqueda exhaustiva (`stepwise = FALSE`, `approximation = FALSE`). La selección del orden se realizó minimizando el Criterio de Información de Akaike (AICc) tal como se recomienda para modelos ARIMA en tareas de pronóstico (Hyndman et al., 2025). Para obtener la distribución predictiva se generaron, para cada año objetivo, entre 5000 trayectorias futuras mediante simulación bootstrap a partir del modelo ajustado, en la escala transformada. Las trayectorias simuladas se transformaron de regreso a la escala original mediante la inversa de Box-Cox y se restó 1, truncando valores negativos a cero. El pronóstico puntual para cada semana se definió como la mediana de las simulaciones, mientras que los intervalos de predicción del 80 % y 95 % se obtuvieron a partir de los cuantiles correspondientes de la distribución simulada. Este enfoque de intervalos basados en simulación es coherente con la implementación del paquete `forecast` y con prácticas modernas de pronóstico probabilístico (Hyndman & Athanasopoulos, 2021; Hyndman et al., 2025; Gneiting & Raftery, 2007).

El Modelo B extiende el Modelo A incorporando covariables exógenas derivadas de las variables meteorológicas, resultando en un modelo de tipo SARIMAX (Modelo Estacional Autorregresivo Integrado de Media Móvil Exógena). A partir de las series semanales de las covariables— temperatura (máxima, mínima y promedio), humedad relativa y precipitación— se generó una matriz de regresores con rezagos para cada variable, de forma que cada fila X_t representa las condiciones meteorológicas ocurridas en las semanas anteriores al tiempo t . Para evitar fuga de información, la estandarización de las covariables se realizó utilizando únicamente el conjunto de entrenamiento de cada año: se calcularon la media y la desviación estándar de cada columna de X en el periodo 2007 a $t - 1$, y estos parámetros se emplearon para centrar y escalar tanto las covariables de entrenamiento como las correspondientes a todo el periodo. Las filas con valores faltantes debidos a los rezagos se excluyeron del ajuste, pero se conservaron para la fase de pronóstico, asegurando que—para el año t —las covariables utilizadas provinieran solo de observaciones anteriores o contemporáneas a la semana pronosticada. El ajuste del modelo se llevó a cabo sobre la serie transformada con Box-Cox, utilizando la función `auto.arima` bajo la misma lógica de selección de órdenes que en el Modelo A. La predicción se realizó mediante simulación condicional sobre las covariables futuras ya estandarizadas, generando miles de trayectorias semanales y aplicando la misma estrategia de mediana e intervalos por cuantiles para regresar a la escala original. Esta estructura SARIMAX con covariables climáticas es coherente con trabajos previos que han empleado modelos ARIMA/SARIMA con variables meteorológicas y/o digitales para predecir dengue (Zhang, 2019; Xiao et al., 2024; Alam et al., 2025).

El Modelo C extiende el Modelo A incorporando como covariable exógena un índice de búsqueda de GT relacionado con dengue, dando lugar a una estructura SARIMAX orientada a capturar señales digitales tempranas de actividad epidémica. Cada valor mensual del índice se asignó a todas las semanas cuyo lunes pertenecía a dicho mes, generando una serie semanal alineada con el mismo calendario de lunes utilizado para los eventos de dengue. Semanas sin información se completaron mediante procedimientos simples de arrastre hacia adelante y hacia atrás (`carry-forward/carry-backward`), evitando cortes en la serie de covariables. Sobre esta serie semanal se construyó una matriz de regresores exógenos con rezagos de 2 a 8 semanas, de modo que cada vector X_t resume la intensidad de búsquedas de dengue observada en las semanas previas al tiempo t ; la elección de este rango de lags se fundamenta tanto en consideraciones biológicas—el intervalo entre la percepción de síntomas, la búsqueda de información en línea y la notificación—como en evidencia empírica que muestra asociaciones entre índices de Google y casos de

dengue en ventanas de rezago (Guo et al., 2017; Yang et al., 2017; Khairul et al., 2021; Xiao et al., 2024). Al igual que en el Modelo B, la estandarización de las covariables se realizó utilizando únicamente el conjunto de entrenamiento de cada año. Utilizando la función `auto.arima()` sigue las mismas directrices que en los Modelos A y B (Hyndman & Athanasopoulos, 2021; Hyndman et al., 2025). La predicción se realizó mediante simulación condicional sobre las covariables futuras ya estandarizadas, generando miles de trayectorias semanales y aplicando la misma estrategia de mediana e intervalos por cuantiles para regresar a la escala original. (Gluskin et al., 2014; Guo et al., 2017; Yang et al., 2017; Xiao et al., 2024).

El modelo D combina los dos bloques de covariables de los modelos anteriores: variables meteorológicas del proyecto NASA POWER y el índice de búsqueda de GT. Al igual que en los otros modelos, se trabajó sobre la serie semanal de casos y_t para Cali en el periodo 2007-2024. Sobre la base de datos diarios de NASA POWER y de datos de GT, se construyeron series semanales tal como en el modelo B y modelo C. A partir de estas series se generó una matriz conjunta de regresores exógenos X_t , construyendo para cada variable climática y para el índice de GT un conjunto de rezagos de 2 a 8 semanas, de modo que cada fila de X_t resume las condiciones ambientales y digitales observadas en las semanas previas al tiempo t . La estandarización de las covariables se realizó por año utilizando únicamente el conjunto de entrenamiento 2007-t-1: se calcularon la media y desviación estándar de cada columna de X en ese periodo y estos parámetros se aplicaron para centrar y escalar tanto las covariables de entrenamiento como las empleadas en el horizonte de pronóstico. Las filas con valores faltantes originados por los rezagos se excluyeron del ajuste, manteniéndose sin embargo las observaciones necesarias para la fase de predicción. Sobre la serie transformada y las covariables estandarizadas se ajustó un modelo SARIMAX(p, d, q) \times (P, D, Q)₅₂ utilizando la función `auto.arima()` siguiendo las mismas directrices que en los Modelos A, B y C (Hyndman & Athanasopoulos, 2021; Hyndman et al., 2025). La predicción se realizó mediante simulación condicional sobre las covariables futuras ya estandarizadas, generando miles de trayectorias semanales y aplicando la misma estrategia de mediana e intervalos por cuantiles para regresar a la escala original. (Gluskin et al., 2014; Guo et al., 2017; Yang et al., 2017; Xiao et al., 2024). Para reducir los tiempos de computo se paralelizó utilizando los paquetes `future` y `furrr`, configurando un plan `multisession` con cuatro `workers` y aplicando `future_map2()` sobre los pares año de entrenamiento-año de pronóstico. (Bengtsson, 2025; Vaughan & Dancho, 2022).

Estrategia de evaluación y validación

La evaluación se organizó por año calendario t , se compararon—semana a semana—los valores observados y predichos, construyendo tablas de comparación con casos observados, casos predichos, residuos (observado - predicho), y un pronóstico de referencia `seasonal naive` (`snaive`). Este esquema de evaluación, conocido como `rolling-origin with expanding window`, se considera una buena práctica para cuantificar el desempeño de modelos de pronóstico en contextos donde los datos llegan secuencialmente en el tiempo (Tashman, 2000; Hyndman & Athanasopoulos, 2021).

Como línea base se utilizó un modelo `snaive` con periodo $m = 52$, en el cual el pronóstico para una semana dada replica el valor observado 52 semanas antes (misma semana del año previo), esta línea base sirve como punto de comparación para determinar si modelos más complejos aportan una mejora real en términos de precisión (Hyndman & Athanasopoulos, 2021). Para contrastar formalmente el desempeño de los modelos frente a la línea base se aplicó la prueba de Diebold-Mariano (DM), que evalúa la hipótesis de igualdad de precisión de dos modelos a partir de la serie de diferencias de pérdida (en este caso, pérdida cuadrática MSE) (Diebold & Mariano, 1995). Se utilizó horizonte $h = 1$ semana y una alternativa unilateral en la que se plantea que el modelo tiene menor error esperado que el `snaive`. Un valor de $p < 0.05$ se interpretó

como evidencia de que el modelo supera significativamente a la línea base estacional. Además, se evaluó la adecuación de los residuos del ajuste ARIMA/SARIMAX mediante la prueba de Ljung-Box, aplicada a los residuos en la escala transformada, con un lag de 52 semanas y corrigiendo los grados de libertad según el número de parámetros autoregresivos y de media móvil del modelo (Ljung & Box, 1978; Hyndman & Athanasopoulos, 2021). Un valor de p elevado se interpretó como ausencia de autocorrelación, lo que sugiere que el modelo captura adecuadamente la dependencia temporal de la serie.

Para cada combinación modelo-año se calcularon métricas de error que resumen el desempeño puntual del pronóstico: error absoluto medio (MAE), raíz del error cuadrático medio (RMSE), error porcentual medio absoluto (MAPE), y error absoluto medio escalado (MASE) con estacionalidad $m = 52$ (Hyndman & Koehler, 2006; Hyndman & Athanasopoulos, 2021; Cheng & Moraga, 2024; Xiao et al., 2024; Alam et al., 2025). Otras métricas como el error porcentual medio absoluto simétrico (sMAPE), error porcentual medio absoluto ponderado (WMAPE), mediana del error porcentual absoluto (MdAPE), raíz del error cuadrático medio logarítmico (RMSLE) como indicadores descriptivos que permiten caracterizar el error en distintos rangos de incidencia (Hyndman & Koehler, 2006; Cheng & Moraga, 2024). Además se calcularon dos conjuntos de medidas sobre los intervalos de predicción del 80 % y 95 %, la cobertura empírica (porcentaje de observaciones que caen dentro del intervalo) y el *interval score*, que es una regla de *scoring* estrictamente propia que penaliza tanto la anchura excesiva como las observaciones fuera del intervalo, promoviendo intervalos bien calibrados y lo más estrechos posible (Gneiting & Raftery, 2007). La comparación entre modelos—junto con la referencia *snaive*—permite cuantificar el efecto de incorporar variables exógenas sobre la precisión de los pronósticos de dengue y valorar si la complejidad añadida por dichas covariables se traduce en mejoras significativas (Zhang, 2019; Xiao et al., 2024; Alam et al., 2025).

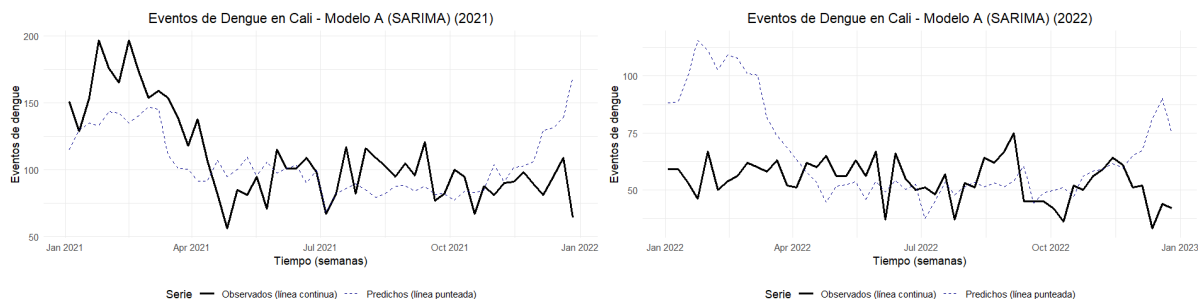
Paquetes de R

Para la importación y depuración de registros de SIVIGILA se empleó *sivirep* 1.0.1 (Gómez-Millán et al., 2024). El manejo de fechas y semanas epidemiológicas se realizó con *lubridate* 1.9.4 (Spinu et al., 2024), mientras que la transformación y organización de datos tabulares se efectuó con *tidyverse* 2.0.0 (Wickham, 2023b). El ajuste de modelos ARIMA, la selección automática de órdenes y la generación de pronósticos se hicieron con *forecast* 8.24.0 (Hyndman et al., 2025), y la visualización de series, pronósticos e intervalos de predicción con *ggplot2* 3.5.2 (Wickham et al., 2025). Las métricas de error se calcularon con *Metrics* 0.1.4 (Hamner et al., 2018), la descarga y el parseo de datos meteorológicos vía API con *httr* 1.4.7 y *jsonlite* 2.0.0 (Wickham, 2023; Ooms et al., 2025), y la exportación de tablas de resultados a hojas de cálculo con *writexl* 1.5.1 (Ooms, 2025b). Para reducir los tiempos de cómputo asociados se usaron los paquetes *furrr* 0.3.1 (Vaughan & Dancho, 2022) y *future* 1.67.0 (Bengtsson, 2025).

Resultados

Modelo A

Comportamiento gráfico observado vs. Predichos



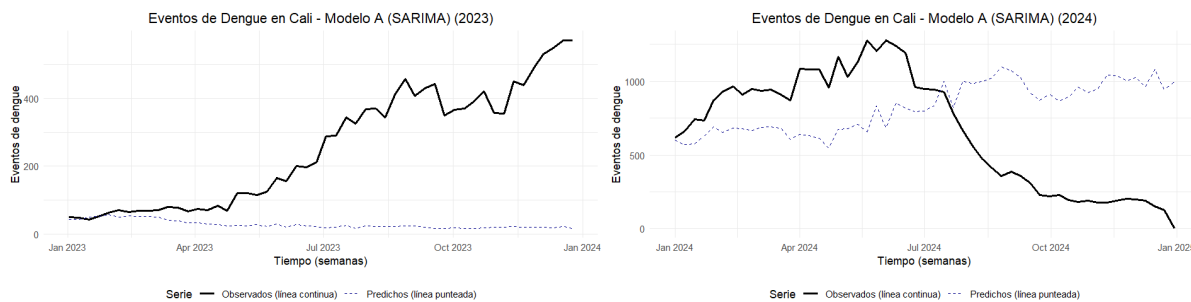


Figura 1-4. Casos semanales de dengue en Cali y pronóstico fuera de muestra del Modelo A para 2021-2024. La línea negra continua muestra los casos observados y la línea azul punteada los valores pronosticados.

En 2021 (Fig. 1) el modelo reproduce de forma razonable la dinámica general de la serie, con valores predichos que siguen el nivel promedio y la tendencia descendente observada a partir del segundo trimestre del año. En el gráfico se aprecia que la línea punteada (pronóstico) permanece próxima a la serie observada, aunque tiende a suavizar las fluctuaciones semanales y a infraestimar algunos picos tempranos. Para 2022 (Fig. 2) el comportamiento es similar: el modelo captura el nivel medio de incidencia y la relativa estabilidad de los casos semanales, con desviaciones moderadas alrededor de la serie real. La trayectoria pronosticada presenta oscilaciones más suaves que las observaciones, pero mantiene una buena correspondencia en la mayor parte del año. En 2023 (Fig. 3) se observa un cambio marcado en la dinámica de la serie, con un incremento sostenido de los casos desde mediados de año y la aparición de un brote de gran magnitud en el segundo semestre. En contraste, el modelo mantiene valores pronosticados relativamente bajos y casi planos; la serie predicha se sitúa muy por debajo de los valores observados durante la mayor parte del año, lo que visualmente se traduce en una fuerte subestimación de la epidemia. Por último, en 2024 (Fig. 3) la discrepancia entre observados y predichos también es notable. Los casos notificados alcanzan niveles superiores a los de años previos durante el primer semestre y posteriormente descienden de forma abrupta hacia final de año, mientras que el modelo pronostica una trayectoria más estable, con un aumento gradual, pero de menor magnitud y sin reproducir el colapso de casos observado al cierre del año. Esto se refleja en una brecha amplia y persistente entre la línea continua (observados) y la punteada (predichos).

Métricas de precisión por año

Año	2021	2022	2023	2024
MAE	22.28	18.03	226.54	470.91
RMSE	29.72	25.08	288.79	538.87
MAPE (%)	21.54	34.99	72.86	1.87E+11
MASE	0.153	0.124	1.627	3.291
Cov80 (%)	100	90.4	30.8	75.5
Cov95 (%)	100	98.1	42.3	98.1
IS80	322.63	218.56	1673.98	2107.47
IS95	632.64	407.37	3246.54	3615.19
DM_stat	-4.03672	-6.22548	8.07488	-2.0789
DM_pvalue	9.11E-05	4.54E-08	1	0.021288
LB_stat	47.14168	44.9831	44.53397	47.1809
LB_pvalue	0.466762	0.556461	0.57527	0.46516

Tabla 1. Métricas de evaluación y resultados de las pruebas para el modelo A.

En 2021 el error absoluto medio (MAE) fue de 22.3 casos semanales y la raíz del error cuadrático medio (RMSE) de 29.7 casos, con un MAPE cercano al 22 % y un MASE de 0.15. Estos valores indican errores moderados y muestran que el modelo es, en promedio, más preciso que un pronóstico ingenuo estacional. En 2022 el desempeño es incluso algo mejor: el MAE desciende a 18.0 casos y el RMSE a 25.1 casos, con un MAPE de alrededor del 35 % y un MASE de 0.12, lo que sugiere una buena capacidad del modelo para reproducir el nivel medio de la serie en un año de incidencia relativamente baja y estable. En 2023 las métricas se deterioran de forma marcada: el MAE se eleva a 226.5 casos y el RMSE a 288.8 casos, con un MAPE de 72.9 % y un MASE de 1.63, lo que refleja errores grandes y un desempeño globalmente peor que el de la línea base estacional. En 2024 los errores aumentan aún más, con un MAE de 470.9 casos y un RMSE de 538.9 casos, y un MASE de 3.29. El MAPE resulta numéricamente inestable (del orden de 10^{11}) debido a semanas con denominadores muy pequeños, por lo que solo se interpreta de manera cualitativa como indicio de errores porcentuales extremos en este año epidémico.

Cobertura de intervalos e Interval score

Los intervalos de predicción del 80 % y 95 % muestran una muy buena cobertura empírica en 2021 y 2022. En 2021, el modelo alcanza coberturas de 100 % tanto para el intervalo del 80 % (Cov80) como para el del 95 % (Cov95), mientras que en 2022 las coberturas son de aproximadamente 90 % y 98 %, respectivamente. Esto indica que los intervalos son conservadores, pero razonablemente calibrados en años de incidencia moderada, lo que se corresponde con puntajes de intervalo relativamente bajos (IS80 e IS95) en comparación con los años siguientes. En 2023 la cobertura cae de forma notable, con Cov80 \approx 31 % y Cov95 \approx 42 %, acompañadas de valores muy altos de IS80 e IS95, lo que evidencia intervalos demasiado estrechos o mal centrados frente al brote observado. En 2024 la cobertura mejora parcialmente (Cov80 \approx 76 % y Cov95 \approx 98 %), pero sigue sin alcanzar el nivel nominal del 80 % para el primer intervalo y mantiene interval scores elevados, lo que señala dificultades del modelo para cuantificar adecuadamente la incertidumbre en contextos de incidencias muy altas.

Comparación con la línea base estacional y diagnóstico de residuos

La prueba de Diebold-Mariano frente a la línea base estacional *snaive* muestra que en 2021 y 2022 el modelo A reduce significativamente el error cuadrático medio respecto a la referencia. Los estadísticos DM son negativos (-4.04 y -6.23, respectivamente) con valores p muy pequeños ($p \approx 9 \times 10^{-5}$ y $p \approx 4.5 \times 10^{-8}$), lo que indica una mejora clara frente al pronóstico ingenuo en esos dos años. En 2023 el estadístico DM es positivo (8.07) con $p = 1$, lo que refleja que la línea base estacional supera al modelo durante el brote intenso. En 2024 el estadístico vuelve a ser negativo (-2.08) con $p \approx 0.021$, sugiriendo una mejora modesta pero estadísticamente significativa frente a *snaive*. Finalmente, la prueba de Ljung-Box aplicada a los residuos del ajuste SARIMA (en la escala transformada de Box-Cox) no detecta autocorrelación remanente significativa en ninguno de los años evaluados: los p-valores se sitúan entre 0.47 y 0.58, de modo que no se rechaza la hipótesis nula de independencia serial de los residuos. Esto es consistente con un ajuste interno adecuado del modelo a los datos de entrenamiento, aun cuando la capacidad predictiva fuera de muestra se degrade en años de comportamiento epidémico extremo.

Modelo B

Comportamiento gráfico observados vs. Predichos

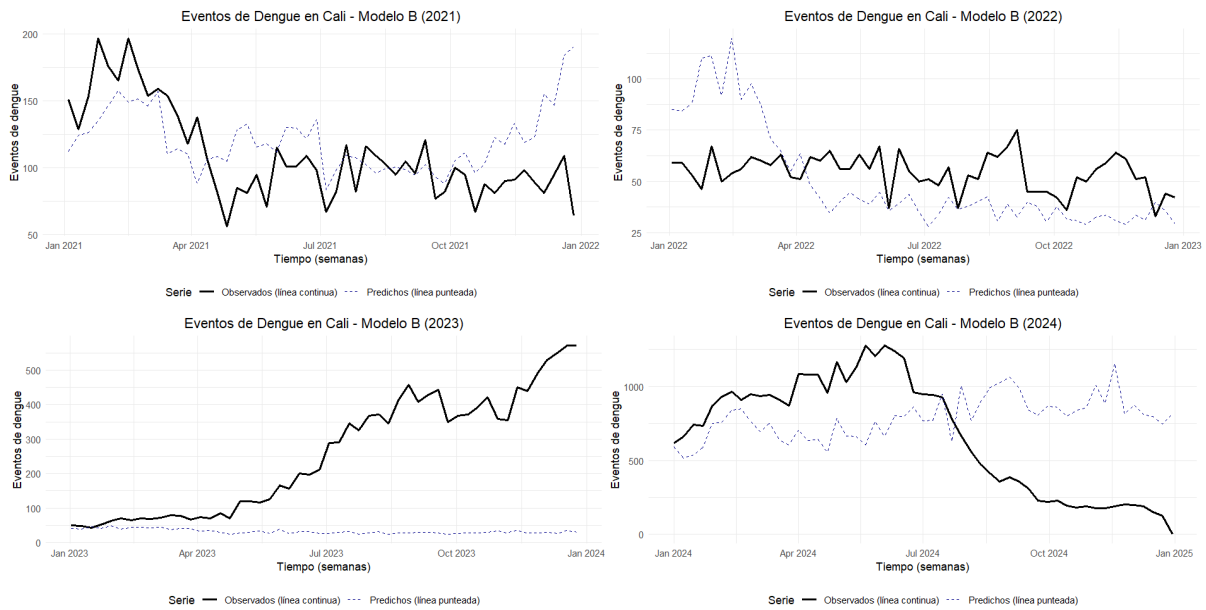
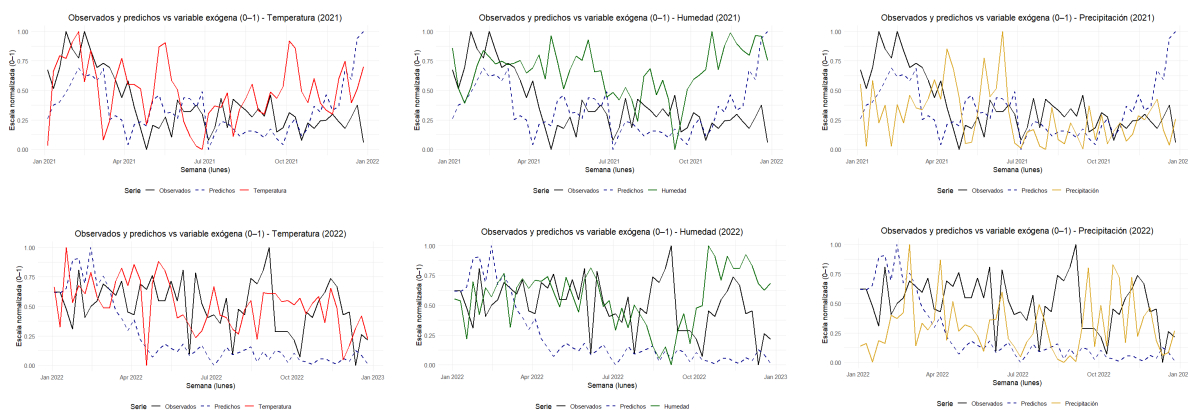


Figura 5-8. Casos semanales de dengue en Cali y pronóstico fuera de muestra del Modelo B para 2021-2024. La línea negra continua muestra los casos observados y la línea azul punteada los valores pronosticados.

En 2021 (Fig. 5) el modelo reproduce parcialmente la disminución de casos observada después del primer trimestre, pero las trayectorias pronosticadas tienden a sobreestimar la carga de dengue a partir de mitad de año. La serie predicha (línea punteada) muestra un aumento progresivo hacia finales de 2021 que no se observa con la misma intensidad en los datos reales, lo que sugiere un sesgo alcista en el segundo semestre. En 2022 (Fig. 6) los pronósticos tienden a sobreestimar la incidencia durante las primeras semanas, con valores predichos por encima de los casos observados, y luego pasan a subestimar de manera moderada la variabilidad de mitad y final de año. Aunque el patrón estacional general es compatible con la serie real, la amplitud de las oscilaciones es menor en la serie predicha. Para 2023 y 2024 el desempeño gráfico es claramente deficiente. En 2023 (Fig. 7), mientras la serie observada presenta un incremento sostenido y marcado de la incidencia a partir de mitad de año, el modelo mantiene valores predichos relativamente bajos y poco variables, sin capturar el brote de gran magnitud. De forma similar, en 2024 (Fig. 8) el modelo subestima el pico epidémico del primer semestre y, posteriormente, tiende a mantener niveles altos cuando la serie observada ya muestra un descenso acusado, evidenciando un desacople entre la dinámica real y la dinámica generada por el modelo.



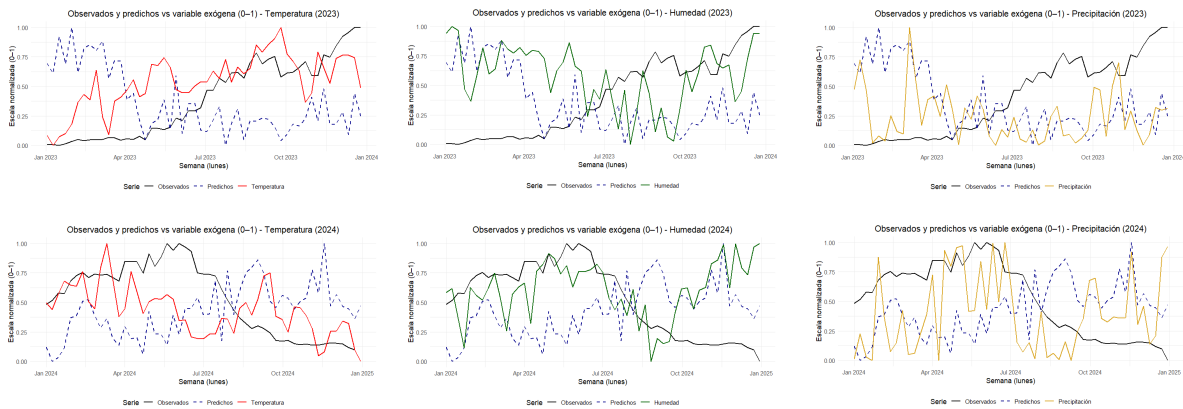


Figura 9-20. Casos semanales de dengue en Cali, pronóstico fuera de muestra del Modelo B, y variables meteorológicas para 2021-2024. La línea negra continua muestra los casos observados, la línea azul punteada los valores pronosticados, la línea roja continua la temperatura, la línea verde la humedad, y la línea amarilla la precipitación. Los datos se transformaron para un rango de 0 a 1 para facilitar la visualización

En 2021 (Fig. 9–11) los casos observados de dengue alcanzan valores altos en el primer trimestre y descienden a niveles bajos-moderados el resto del año. Las predicciones del Modelo B siguen la fase inicial de mayor incidencia, pero muestran una variabilidad menor en la segunda mitad del año. La humedad y la temperatura normalizadas se mantienen en rangos medios-altos con máximos a comienzos y hacia el final del año, mientras que la precipitación aparece concentrada en pulsos episódicos, sobre todo en el primer semestre. Para 2022 (Fig. 12–14) los eventos observados se mantienen en niveles moderados durante buena parte del año, con picos marcados hacia el tercer trimestre, mientras que las predicciones del Modelo B tienden a concentrarse en valores más bajos a partir del segundo trimestre. La humedad relativa permanece en niveles medios, con incrementos notorios a finales de año; la precipitación muestra picos frecuentes distribuidos a lo largo del año, y la temperatura presenta valores altos en el primer trimestre y un descenso progresivo hacia el último trimestre. En 2023 (Fig. 15–17) 2023 la serie observada inicia en valores muy bajos y aumenta casi de forma continua hasta alcanzar el máximo al final del año, mientras que las predicciones del Modelo B se mantienen en niveles bajos-medios con menor pendiente ascendente. La humedad comienza en valores altos y desciende a niveles intermedios en la mitad del año, con un nuevo aumento hacia el cierre del periodo. La precipitación se concentra en pulsos intensos durante los primeros meses y se mantiene en valores más bajos y aislados después de mitad de año. La temperatura muestra un incremento gradual desde valores bajos al inicio hasta niveles altos en el tercer trimestre, estabilizándose posteriormente en rangos medios-altos. Para 2024 (Fig. 18–20) la humedad se mantiene en rangos medios-altos al inicio del año, cae a finales del tercer trimestre y vuelve a niveles altos hacia el cierre del periodo. La precipitación muestra una serie de picos breves sobre un nivel de base bajo a moderado, sin tramos prolongados de lluvia intensa. La temperatura alcanza sus valores más altos en el primer semestre y desciende gradualmente durante la segunda mitad del año.

Métricas de precisión por año

Año	2021	2022	2023	2024
MAE	27.66	21.12	222.31	425.71
RMSE	36.17	25.34	282.41	486.43
MAPE (%)	29.2	37.77	72.84	1.54E+11

MASE	0.19	0.146	1.597	2.975
Cov80 (%)	100	92.3	32.7	98.1
Cov95 (%)	100	100	73.1	98.1
IS80	404.06	177.31	1438.77	3190.09
IS95	925.21	430.7	957.7	7613.83
DM_stat	-3.978	-5.83	7.74	-3.014
DM_pvalue	0.0002	0	0	0.004
LB_stat	89.25	63.03	87.056	87.715
LB_pvalue	0.0003	0.0591	0.0003	0.0003

Tabla 2. Métricas de evaluación y resultados de las pruebas para el modelo B.

En 2021 el error absoluto medio (MAE) fue de aproximadamente 28 casos semanales y la raíz del error cuadrático medio (RMSE) de 36 casos, mientras que el MAPE se situó en torno al 29 %. El error absoluto medio escalado (MASE = 0.19) indica que, en promedio, el modelo cometió un error sustancialmente menor que el pronóstico ingenuo estacional. En 2022 los errores absolutos se redujeron aún más (MAE \approx 21 casos y RMSE \approx 25 casos), aunque el MAPE aumentó hasta cerca de 38 %, lo que sugiere errores porcentuales más altos en un contexto de menores conteos; aun así, el MASE se mantuvo claramente por debajo de 1 (\approx 0.15), por lo que el modelo siguió superando al baseline estacional en términos de error medio. En 2023 se observa un deterioro marcado del desempeño: el MAE ascendió a unos 222 casos y el RMSE a 282 casos, con un MAPE cercano al 73 % y un MASE de alrededor de 1.60, lo que indica que el error absoluto medio del modelo fue mayor que el del pronóstico estacional ingenuo. En 2024 los errores crecieron todavía más (MAE \approx 426 y RMSE \approx 486 casos), con un MASE cercano a 3, lo que refleja errores casi tres veces superiores a los de la línea base. El MAPE de 2024 alcanza un valor extremadamente elevado ($\sim 1.5 \times 10^{11}$ %), numéricamente inestable debido a semanas con conteos muy pequeños en el denominador, por lo que solo se interpreta de manera descriptiva y se privilegia la lectura del MASE para valorar el desempeño relativo frente al modelo ingenuo.

Cobertura de intervalos e Interval score

Las bandas de predicción obtenidas por simulación bootstrap muestran coberturas muy altas en algunos años, pero con un comportamiento heterogéneo. En 2021 las coberturas empíricas fueron de 100 % tanto para el intervalo del 80 % como para el del 95 % (Cov80 y Cov95), es decir, todas las observaciones se ubicaron dentro de los intervalos. En 2022 las coberturas siguieron siendo elevadas (Cov80 \approx 92.3 % y Cov95 = 100 %), cercanas o algo superiores a los niveles nominales. En 2023, en cambio, la cobertura del intervalo del 80 % cayó a \sim 32.7 % y la del 95 % a \sim 73.1 %, evidenciando que los intervalos resultaron demasiado estrechos o mal centrados durante el año epidémico. Para 2024 las coberturas volvieron a ser muy altas (Cov80 \approx 98.1 % y Cov95 \approx 98.1 %). El interval score complementa esta lectura: en 2021 y 2022 los valores fueron relativamente moderados (IS80 \approx 404 y 177; IS95 \approx 925 y 431, respectivamente), mientras que en 2023 y 2024 se observaron scores mucho más elevados (IS80 \approx 1,439 y 3 190; IS95 \approx 958 y 7,614). Esto indica que, aunque en algunos años las bandas fueron “muy seguras”, esa seguridad se logró a costa de intervalos muy anchos o—en el caso de 2023—de una subcobertura importante.

Comparación con la línea base estacional y diagnóstico de residuos

La comparación formal frente a la línea base estacional tipo snave se realizó mediante la prueba de Diebold-Mariano. En 2021 y 2022 el estadístico DM fue negativo y de gran magnitud (DM \approx -3.98 y -5.83, con valores de $p \approx$ 0.0002 y $<$ 0.001, respectivamente), lo que indica que el Modelo B logró errores cuadráticos medios significativamente menores que el pronóstico estacional ingenuo. En 2024 también se observó un DM negativo (\approx -3.01, $p \approx$ 0.004), mientras que en 2023 el estadístico fue positivo y elevado (DM \approx 7.74, $p <$ 0.001), evidenciando que ese año la línea base estacional superó de forma clara al modelo con covariables

meteorológicas. El diagnóstico de autocorrelación residual mediante la prueba de Ljung-Box mostró un ajuste aceptable solo en 2022 ($LB \approx 63.0$, $p \approx 0.059$), mientras que, en 2021, 2023 y 2024 los valores de p fueron muy pequeños ($p \approx 0.0003$), señalando autocorrelación remanente en los residuos y, por tanto, una captación incompleta de la estructura temporal de la serie incluso tras incorporar las variables climáticas.

Modelo C

Comportamiento gráfico observados vs. Predichos

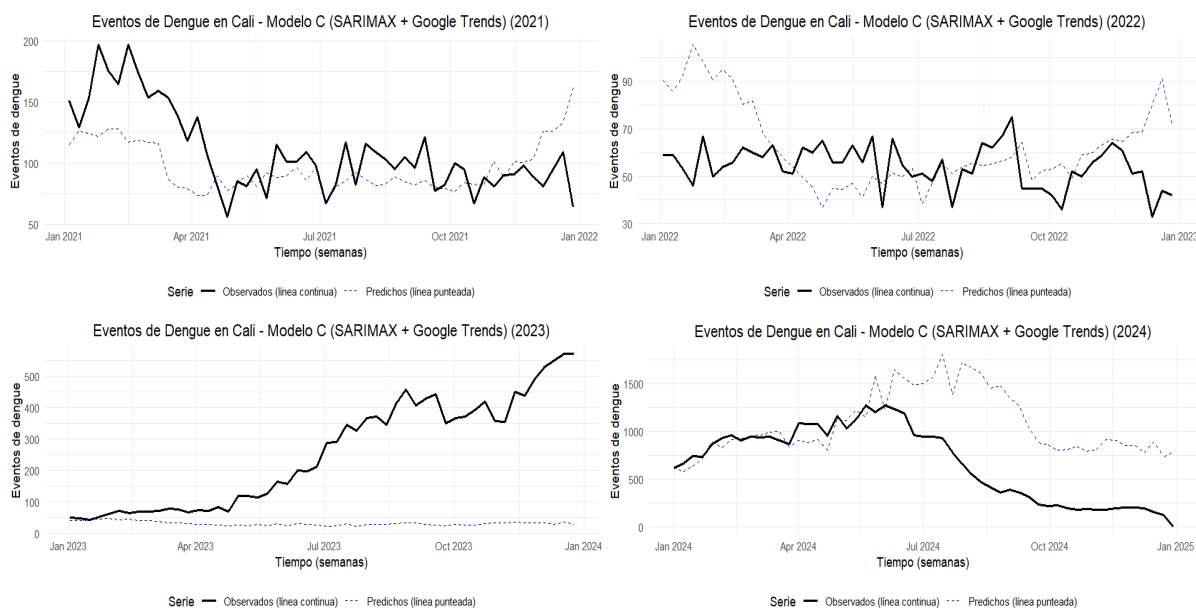


Figura 21-24. Casos semanales de dengue en Cali y pronóstico fuera de muestra del Modelo C para 2021-2024. La línea negra continua muestra los casos observados y la línea azul punteada los valores pronosticados.

En 2021 (Fig. 21) el Modelo C reproduce adecuadamente el nivel general de la serie y la variación intra-anual de los casos. Durante el primer trimestre las trayectorias observada y predicha son similares, con ligera subestimación en algunas semanas. A partir de abril se observa la caída de casos registrada en la vigilancia, que también es capturada por el modelo, aunque con una pendiente algo más suave; hacia el segundo semestre las predicciones se mantienen cercanas a los valores notificados, sin desfases visuales evidentes. En 2022 (Fig. 22) los casos de dengue se mantienen en un rango relativamente estrecho y el modelo produce una trayectoria pronosticada de amplitud comparable. La serie predicha sigue de forma razonable las oscilaciones semanales, con diferencias puntuales en algunos picos aislados; no se observan desfases marcados entre las curvas, y la magnitud de los errores aparenta ser moderada a lo largo del año. En contraste, en 2023 (Fig. 23) se presenta un incremento pronunciado de los casos notificados a partir de mediados de año, con un patrón epidémico sostenido hasta diciembre. El Modelo C, en cambio, genera una trayectoria suavizada de baja amplitud, que permanece muy por debajo de la serie observada durante casi todo el periodo epidémico. La discrepancia entre ambas curvas es especialmente visible a partir del tercer trimestre, donde los picos observados no son acompañados por aumentos equivalentes en las predicciones. Finalmente, en 2024 (Fig. 24) el modelo acompaña la fase inicial de aumento de casos durante el primer semestre, pero a partir de julio las predicciones se mantienen en niveles elevados mientras que la serie observada muestra una disminución progresiva hasta valores cercanos a cero al final del año. Esto se traduce en una sobreestimación sostenida de la incidencia durante el segundo semestre, con divergencias crecientes entre ambas curvas.

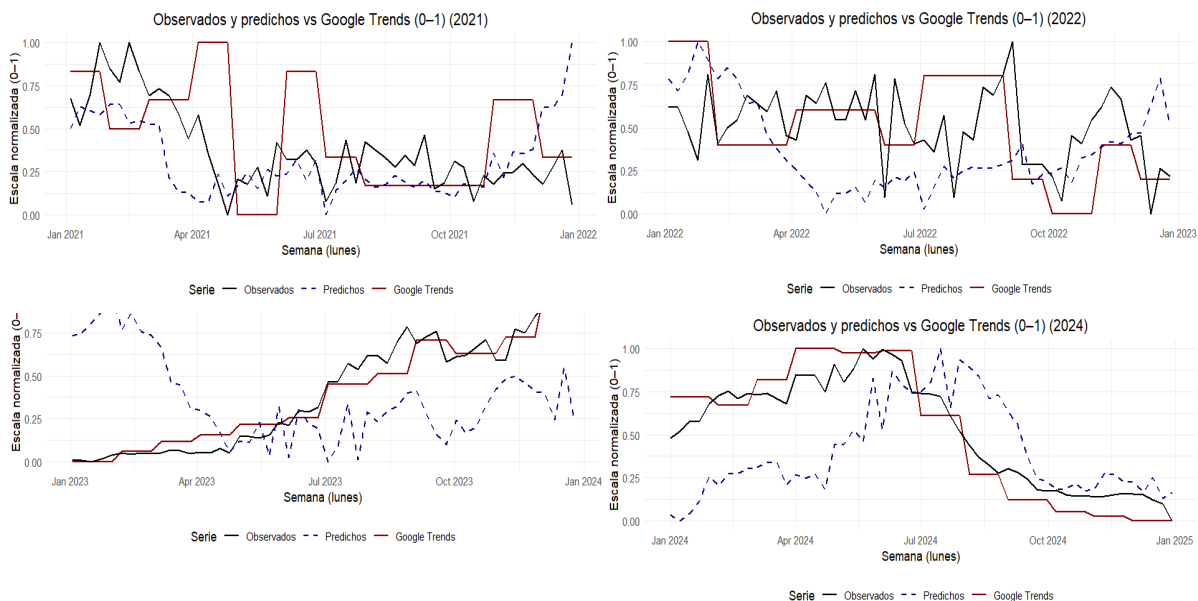


Figura 25-28. Casos semanales de dengue en Cali y pronóstico fuera de muestra del Modelo C para 2021-2024. La línea negra continua muestra los casos observados, la línea azul punteada los valores pronosticados, y la línea roja denota al índice de búsqueda de Google tren. Los datos se transformaron para un rango de 0 a 1 para facilitar la visualización

En 2021 (Fig. 25) los casos observados descienden desde valores altos a comienzos de año hacia niveles bajos-moderados; el índice de GT también baja desde valores elevados y se mantiene en niveles intermedios, mientras que el Modelo C reproduce el descenso general, pero con oscilaciones más suaves. Para 2022 (Fig. 26) los casos se mantienen en niveles moderados con un aumento en el tercer trimestre; GT pasa de valores altos a intermedios y vuelve a subir a mitad de año, mientras que las predicciones del modelo permanecen en valores bajos y casi planos, por debajo de los observados. En 2023 (Fig. 27) los casos de dengue aumentan de forma casi continua desde valores muy bajos hasta un máximo al final del año; el índice de GT muestra un patrón creciente similar, pero el Modelo C predice valores altos al inicio, luego bajos y solo incrementos moderados en la segunda mitad del año. En 2024 (Fig. 28) los casos aumentan desde niveles medios hasta un pico hacia mitad de año y luego caen de forma marcada; GT también sube hasta mediados de año y después desciende, mientras que el modelo sigue el aumento inicial y la caída posterior, aunque se mantiene en niveles intermedios cuando los casos observados ya son muy bajos.

Métricas de precisión por año

Año	2021	2022	2023	2024
MAE	26.01	16.26	223.29	453.49
RMSE	34.32	21.35	282.19	577.07
MAPE (%)	22.96	31.97	74.04	1.47E+11
MASE	0.179	0.112	1.604	3.169
Cov80 (%)	100	94.2	30.8	98.1
Cov95 (%)	100	100	78.8	98.1
IS80	382.53	238.99	1435.94	4120.97
IS95	973.25	581.99	836.15	9806.41
DM_stat	-4.03488	-5.96597	8.349041	-1.2458
DM_pvalue	9.17E-05	1.16E-07	1	0.109212

LB_stat	82.56447	80.47859	82.14829	83.23406
LB_pvalue	8.7E-05	0.000155	9.76E-05	7.21E-05

Tabla 3. Métricas de evaluación y resultados de las pruebas para el modelo C.

En 2021 el error absoluto medio (MAE) fue de aproximadamente 26 casos por semana y la raíz del error cuadrático medio (RMSE) de 34 casos, con un MAPE cercano al 23 %. El MASE fue 0.18, lo que indica que el error absoluto medio del modelo fue claramente inferior al de la aproximación estacional naive. En 2022 los errores absolutos disminuyeron aún más (MAE \approx 16 casos y RMSE \approx 21 casos), aunque el MAPE aumentó a alrededor de 32 %, coherente con un año de menor incidencia en el que errores moderados representan un porcentaje mayor. El MASE se redujo a 0.11, confirmando que el modelo siguió superando cómodamente a la línea base estacional. En 2023 las métricas se deterioran de forma marcada: el MAE asciende a unos 223 casos semanales y el RMSE a 282 casos, con un MAPE cercano al 73 % y un MASE de 1.60, lo que implica que el error medio del modelo es ya superior al del pronóstico ingenuo. Para 2024 el MAE aumenta hasta \sim 453 casos y el RMSE a \sim 577 casos; el MAPE toma un valor extremadamente grande (del orden de 10^{11}) debido a semanas con denominadores muy pequeños, por lo que solo se interpreta como indicio de inestabilidad numérica, mientras que el MASE (\approx 3.17) indica que el error absoluto medio del modelo triplica aproximadamente al de la referencia estacional. En conjunto, las métricas puntuales sugieren un buen desempeño relativo en 2021-2022 y errores mucho mayores en 2023-2024, especialmente durante los brotes intensos.

Cobertura de intervalos e Interval score

Los intervalos de predicción presentan coberturas muy altas en 2021 y 2022. En 2021, la cobertura empírica de los intervalos del 80 % y 95 % es del 100 % en ambos casos (Cov80 y Cov95), con valores de interval score de aproximadamente 383 (IS80) y 973 (IS95). En 2022 la cobertura se mantiene elevada, con Cov80 \approx 94.2 % y Cov95 = 100 %, y scores de alrededor de 239 y 582 para los intervalos del 80 % y 95 %, respectivamente. En 2023 la cobertura disminuye de forma notable: solo \sim 30.8 % de las observaciones caen dentro del intervalo del 80 % y \sim 78.8 % dentro del del 95 %, acompañadas de interval scores elevados (IS80 \approx 1,436; IS95 \approx 836), lo que refleja intervalos poco calibrados frente al brote epidémico. En 2024 la cobertura vuelve a ser muy alta (\approx 98.1 % para ambos niveles), pero asociada a scores muy grandes (IS80 \approx 4,121; IS95 \approx 9,806), lo que sugiere que se logra contener casi todas las observaciones a costa de intervalos excesivamente anchos y—por tanto—de utilidad limitada para la toma de decisiones.

Comparación con la línea base estacional y diagnóstico de residuos

La comparación formal con la línea base estacional tipo snaive mediante la prueba de Diebold-Mariano también se resume en la Tabla X-C1. En 2021 y 2022 la estadística DM es negativa y de gran magnitud (DM \approx -4.03 y -5.97, con valores de p de aproximadamente 0.00009 y 1.2×10^{-7}), lo que indica que el Modelo C presenta pérdidas cuadráticas medias significativamente menores que la referencia estacional. En 2023 la estadística DM se vuelve positiva (\approx 8.35) con $p \approx 1$ bajo la formulación unilateral utilizada, lo que es consistente con un peor desempeño del modelo frente al snaive durante el año epidémico. En 2024 el DM es nuevamente negativo (\approx -1.25) pero con $p \approx 0.11$, de modo que no hay evidencia estadística suficiente para afirmar una mejora clara frente a la línea base. Finalmente, los resultados de la prueba de Ljung-Box muestran valores de estadístico altos y p-valores muy pequeños en los cuatro años (entre $\sim 8.7 \times 10^{-5}$ y $\sim 1.6 \times 10^{-4}$), lo que señala la presencia de autocorrelación remanente en los residuos del Modelo C e indica que, pese a la inclusión de la covariable de GT, aún queda estructura temporal sin capturar completamente.

Modelo D

Comportamiento gráfico observados vs. Predichos

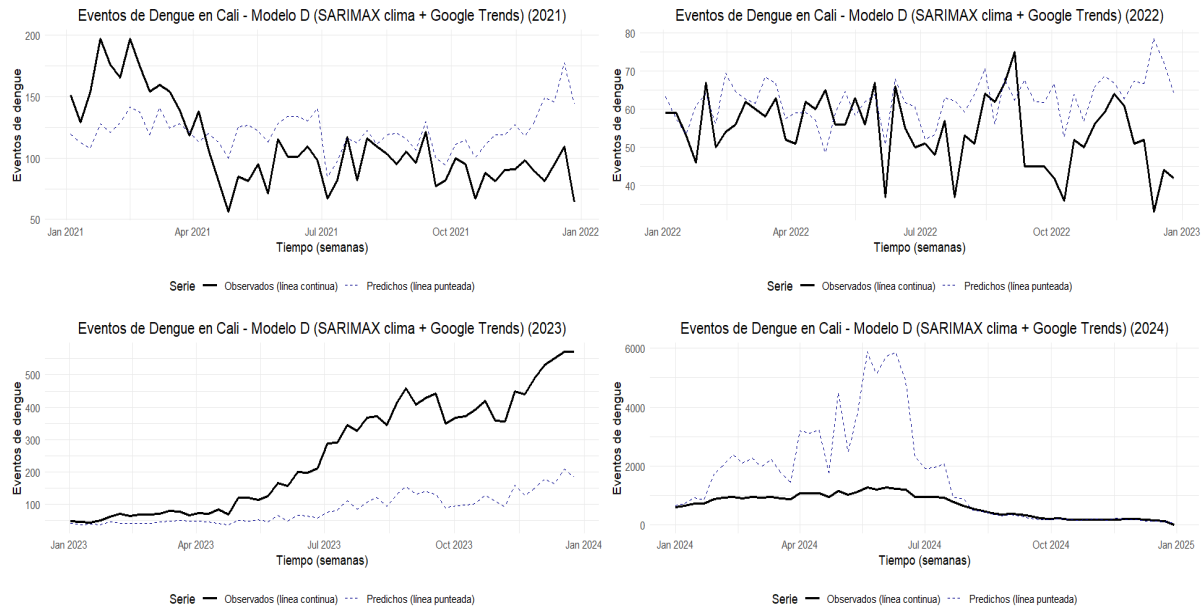


Figura 28-31. Casos semanales de dengue en Cali y pronóstico fuera de muestra del Modelo D para 2021-2024. La línea negra continua muestra los casos observados, la línea azul punteada los valores pronosticados.

Métricas de precisión por año

Año	2021	2022	2023	2024
MAE	29.82	9.93	168.07	970.87
RMSE	34.9	13.17	210.42	1635.28
MAPE (%)	30.08	21.2	56.72	1.56E+10
MASE	0.203	0.068	1.197	6.731
Cov80 (%)	98.1	100	59.6	66
Cov95 (%)	100	100	100	98.1
IS80	351.27	177.21	355.62	4738.21
IS95	774.68	396.84	521.43	6600.77
DM_stat	-4.015	-5.79	-5.091	3.07
DM_pvalue	0.0001	0	0	0.9983
LB_stat	101.212	103.958	116.34	112.49
LB_pvalue	0	0	0	0

Tabla 4. Métricas de evaluación y resultados de las pruebas para el modelo D.

Cobertura de intervalos e Interval score

En 2021 la cobertura empírica de los intervalos de predicción es elevada: aproximadamente 98.1 % de las observaciones se ubican dentro del intervalo del 80 % (Cov80) y el 100 % dentro del intervalo del 95 %, con interval scores moderados (IS80 \approx 351 y IS95 \approx 775), lo que sugiere bandas relativamente informativas. En 2022 la cobertura es aún más alta (100 % tanto para el 80 % como para el 95 %), con interval scores de alrededor de 177 y 397, respectivamente, coherentes con intervalos bien calibrados en un contexto de baja variabilidad. En 2023 la cobertura del 80 % desciende a \sim 59.6 %, mientras que la del 95 % se mantiene en 100 %; los interval scores (IS80 \approx 356; IS95 \approx 521) indican que, aunque el modelo logra conservar casi todas las observaciones dentro del intervalo amplio, la calibración en el nivel del 80 % es deficiente durante el brote,

con bandas que resultan estrechas o mal centradas para parte del año. En 2024, finalmente, las coberturas se sitúan en 66 % para el 80 % y 98.1 % para el 95 %, pero acompañadas de interval scores muy altos (IS80 \approx 4 738; IS95 \approx 6,601), lo que refleja intervalos excesivamente anchos y de utilidad limitada para la toma de decisiones, pese a la alta proporción de observaciones contenidas en ellos.

Comparación con la línea base estacional y diagnóstico de residuos

La prueba de Diebold-Mariano frente a la línea base estacional snave muestra que en 2021 y 2022 el Modelo D ofrece mejoras claras en términos de error cuadrático medio: los estadísticos DM son negativos (\approx -4.02 y -5.79, respectivamente) con valores de p muy pequeños ($p \approx 0.0001$ y $p \approx 0$), lo que indica que la combinación de covariables climáticas y de GT reduce significativamente el MSE respecto al pronóstico ingenuo en esos años. En 2023 el DM sigue siendo negativo (\approx -5.09, $p \approx 0$), lo que sugiere que, a pesar de los errores grandes observados, el Modelo D aún supera estadísticamente a la referencia estacional en términos de pérdida cuadrática promedio. En 2024 la situación se invierte: el estadístico DM es positivo (\approx 3.07) con un valor de p cercano a 1 ($p \approx 0.998$), lo que implica que, bajo la formulación unilateral utilizada, la línea base snave resulta preferible al modelo con dos bloques de covariables. Por otro lado, los resultados de la prueba de Ljung-Box indican en todos los años valores de estadístico elevados y p-valores prácticamente nulos, evidenciando autocorrelación remanente en los residuos del Modelo D. Esto sugiere que, aunque en varios años el modelo mejora la precisión respecto a snave, todavía deja sin capturar parte de la estructura temporal de la serie, especialmente en los periodos de mayor variabilidad epidémica.

Comparación entre modelos

Comparación	Año	DM_stat	DM_p value
B vs A	2021	2.5135	0.9924
B vs A	2022	0.1347	0.5533
B vs A	2023	-6.6616	0.0000
B vs A	2024	-3.4693	0.0005
C vs A	2021	2.4866	0.9919
C vs A	2022	-2.7338	0.0043
C vs A	2023	-6.0054	0.0000
C vs A	2024	0.8632	0.8040
C vs B	2021	-0.5141	0.3047
C vs B	2022	-1.8917	0.0321
C vs B	2023	-0.5562	0.2902
C vs B	2024	1.9496	0.9717
D vs A	2021	2.0059	0.9749
D vs A	2022	-3.2501	0.0010
D vs A	2023	-6.1608	0.0000
D vs A	2024	3.1019	0.9984
D vs B	2021	-0.4297	0.3346
D vs B	2022	-3.4654	0.0005
D vs B	2023	-6.0654	0.0000
D vs B	2024	3.1930	0.9988
D vs C	2021	0.1864	0.5736
D vs C	2022	-3.1487	0.0014
D vs C	2023	-6.1441	0.0000

D vs C	2024	2.9970	0.9979
--------	------	--------	--------

Tabla 5. Resultados pruebas DM por pares, comparando los modelos.

Con base en las pruebas de Diebold-Mariano por pares (hipótesis alternativa: el primer modelo tiene menor MSE que el segundo), se observa que:

- B vs A: en 2021-2022 no hay diferencias significativas; en 2023 ($DM < 0$, $p \ll 0.001$) y 2024 ($DM < 0$, $p \approx 0.001$) el Modelo B supera al univariado A.
- C vs A: en 2021 y 2024 no hay evidencia de mejora; en 2022 ($DM < 0$, $p < 0.01$) y 2023 ($DM < 0$, $p \ll 0.001$) el Modelo C supera claramente a A.
- D vs A: en 2021 y 2024 tampoco se detectan diferencias significativas ($DM > 0$ con p alta, compatible incluso con A ligeramente mejor); en 2022 y 2023 el Modelo D mejora de forma marcada al univariado ($DM < 0$, $p \leq 0.001$).
- C vs B (H_1 : C mejor que B): en 2021 y 2023 no hay diferencias; en 2022 C es mejor que B ($DM < 0$, $p < 0.05$); en 2024 la señal se invierte ($DM > 0$ con p muy alta), lo que sugiere que B es preferible.
- D vs B (H_1 : D mejor que B): en 2021 no hay diferencias; en 2022 y 2023 D supera significativamente a B ($DM < 0$, $p \leq 0.001$); en 2024 ocurre lo contrario ($DM > 0$, $p \approx 1$), favoreciendo a B.
- D vs C (H_1 : D mejor que C): en 2021 no se observan diferencias; en 2022 y 2023 el Modelo D presenta menor MSE que C ($DM < 0$, $p \leq 0.001$); en 2024 la evidencia apunta a que C es mejor ($DM > 0$, $p \approx 1$).

En conjunto, las mejoras frente al modelo univariante A se concentran en 2022-2023 para los tres modelos con covariables, con el Modelo D mostrando la reducción de error más consistente frente a A, B y C en esos años, mientras que en 2024 los modelos con clima (B) y con GT (C) resultan más favorables que el modelo combinado D.

Discusión

Los resultados muestran que los modelos con información exógena (B, C y D) superan de forma sistemática al enfoque univariado cuando la dinámica del sistema permanece dentro de su régimen “típico”. En 2021-2022, los cuatro modelos logran errores absolutos moderados, pero los SARIMAX presentan valores de MASE claramente inferiores a 1, lo que indica que mejoran al pronóstico *snaive*. Por ejemplo, en el Modelo C, el MAE se mantiene por debajo de 30 casos semanales en 2021-2022, con MASE entre ~ 0.18 y 0.11 ; en el Modelo B, el MAE es similar o ligeramente mayor, con MASE entre ~ 0.19 y 0.15 ; y en el Modelo D el MAE de 2022 desciende a ~ 10 casos por semana, con un MASE de 0.07 . Estas cifras, junto con valores de RMSE relativamente bajos y MAPEs moderados, reflejan una buena capacidad para reproducir el nivel medio y las variaciones semanales en años de incidencia relativamente estable. En contraste, en 2023-2024 todos los modelos experimentan un deterioro marcado de la precisión puntual, coherente con la aparición de un brote de gran magnitud. Los valores de MAE y RMSE se incrementan fuertemente, con MASE por encima de 1 en prácticamente todas las combinaciones modelo-año, lo que indica errores absolutos medios superiores a los de la línea base estacional. En el Modelo B, el MAE supera los 220 casos en 2023 y ~ 426 en 2024, con MASE de ~ 1.60 y ~ 2.98 , respectivamente; el Modelo C muestra un comportamiento similar, con MAE por encima de 220 casos en 2023 y ~ 453 en 2024 y MASE en torno a 1.60 y 3.17 . El Modelo D, aunque reduce el MAE en 2023 (~ 168 casos, $MASE \approx 1.20$), colapsa en 2024 con un MAE cercano a 971 casos y MASE superior a 6, lo que sugiere una incapacidad para anticipar la magnitud del brote pese a incorporar dos bloques de covariables. Además, el MAPE de 2024 se vuelve numéricamente inestable en todos los modelos, debido a semanas con

denominadores muy pequeños; en consecuencia, se interpreta solo de forma descriptiva y se privilegia la lectura de MAE, RMSE y MASE como indicadores más robustos del error.

Las métricas de cobertura de intervalos y el interval score complementan esta lectura, proporcionando información sobre la calidad probabilística de los pronósticos. En 2021-2022 los modelos con variables exógenas producen intervalos con coberturas empíricas (Cov80 y Cov95) cercanas o iguales a los niveles nominales. Por ejemplo, el Modelo C alcanza coberturas del 80 % y 95 % del 100 %—en ambos casos— en 2021, y del 94 % y 100 % en 2022; el Modelo D presenta Cov80 y Cov95 cercanas al 100 % en ambos años, con valores de IS80 e IS95 relativamente contenidos. Esto indica intervalos razonablemente calibrados en años de incidencia moderada, aunque algo conservadores. En 2023, sin embargo, las coberturas del 80 % caen de forma notable en los modelos B y C (≈ 31 -33 %), y aunque el Modelo D mantiene Cov95 del 100 %, lo hace a costa de un incremento marcado de IS80 e IS95. En 2024 la cobertura vuelve a ser muy alta en B, C y D (Cov95 ≈ 98 -100 %), pero asociada a valores de IS80 e IS95 elevados, especialmente en el Modelo D, lo que indica que los intervalos deben ensancharse de forma considerable para abarcar las observaciones en un entorno de alta incertidumbre. En términos prácticos, las exógenas aportan buena calibración en años “normales”, pero frente a brotes extremos el sistema responde abriendo fuertemente las bandas, algo consistente con la idea de que, ante un aumento de varianza residual, los modelos deben sacrificar precisión puntual para mantener cobertura (Gneiting & Raftery, 2007). La comparación formal con la línea base estacional mediante la prueba de Diebold-Mariano refuerza esta interpretación.

Frente a *snaive*, los modelos B, C y D presentan estadísticos DM negativos y p-valores muy pequeños ($p \ll 0.01$) en 2021-2022, lo que evidencia reducciones significativas del error cuadrático medio. En el caso del Modelo D, la prueba DM es favorable en 2022, coherente con su MAE y MASE particularmente bajos. En 2023 los resultados son más matizados: el Modelo C deja de superar a la línea base (DM positiva o no significativa), mientras que el Modelo D aún mantiene DM negativa frente a *snaive*, aunque con un MASE > 1 que indica que—aunque reduce algo el MSE—lo hace con errores absolutos aún grandes en términos operativos. En 2024, por el contrario, los DM se vuelven positivos o no significativos para el Modelo D, indicando que la combinación clima + búsquedas no garantiza mejoras frente al pronóstico estacional sencillo en un contexto de comportamiento epidémico extremo.

Los contrastes por pares entre modelos muestran que las mejoras frente al univariado A se concentran en 2022-2023 (especialmente para C y D), y que en 2022 el Modelo D tiende a superar tanto al modelo con clima solo (B), como al modelo con búsquedas solo (C), mientras que en 2024 las diferencias entre modelos se difuminan o incluso favorecen en algunos casos a las formulaciones más simples. El diagnóstico de residuos mediante la prueba de Ljung-Box indica que el ajuste interno de los modelos no es homogéneo. En el univariado A, los p-valores tienden a ser moderados en años “tranquilos”, lo que sugiere una captura razonable de la estructura temporal básica. En el Modelo B, solo 2022 muestra un p cercano al umbral de significancia, mientras que 2021, 2023 y 2024 presentan LB_pvalue muy pequeños, señalando autocorrelación remanente aun después de incorporar clima. Para el Modelo C, los p-valores de Ljung-Box son muy bajos en todos los años, lo que indica que la inclusión de GT no elimina completamente dependencias temporales no modeladas. El Modelo D tampoco corrige este problema: sus LB_pvalue son cercanos a 0 en 2021-2024, evidenciando que, aunque la combinación de covariables puede mejorar MAE, RMSE o incluso DM en ciertos años, persiste estructura serial en los residuos.

En conjunto, estos resultados sugieren que la familia SARIMA/SARIMAX, tal como se especificó (lineal en las exógenas y con rezagos fijos), captura buena parte de la estacionalidad y de la señal forzada por clima y búsquedas en años típicos, pero deja sin modelar componentes adicionales (no linealidades, cambios de

régimen, microclimas, factores sociales y demográficos) que se vuelven críticos en episodios extremos (Colón-González et al., 2013; Restrepo et al., 2014; Nova et al., 2020).

Desde el punto de vista conceptual, la mejora observada en 2021-2022 es coherente con el marco que vincula clima, conducta y transmisión. La temperatura, la humedad y la precipitación influyen de forma directa en el desarrollo de *Aedes*, en la supervivencia del vector y en la duración de la incubación extrínseca del virus, modulando la capacidad vectorial y el calendario de transmisión (Morin et al., 2013; Ebi & Nealon, 2016). Estudios en distintos contextos climáticos han mostrado que la variabilidad de temperatura y lluvia, así como el cambio climático, modifican tanto la intensidad como la estacionalidad y la distribución geográfica del dengue (Banu et al., 2011; Xu et al., 2016; Fujita et al., 2023), y que estos efectos se encuentran mediados por la disponibilidad de susceptibles y la dinámica interna del sistema (Pongsumpun et al., 2008; Nova et al., 2020). Cuando la estacionalidad es relativamente estable, las covariables climáticas añaden información sobre el “timing” y la magnitud de los picos que no está contenida únicamente en la inercia de la serie, y ayudan a explicar por qué grandes ciudades cálidas y densamente pobladas pueden sostener transmisión persistente y actuar como núcleos de expansión regional del dengue, tal como se ha documentado en Brasil (Barcellos & Lowe, 2014) y en patrones espaciales de Colombia (Restrepo et al., 2014).

De forma análoga, el índice de búsquedas de “dengue” refleja la demanda de información ante síntomas y noticias, aportando una señal rápida que mejora el nowcasting en contextos endémicos (Guo et al., 2017; Yang et al., 2017; Cheong et al., 2025). El Modelo C capitaliza esta dimensión conductual, mientras que el Modelo D sintetiza ambas fuentes: clima como forzante físico-biológico y GT como proxy de la respuesta conductual. No es casual que, en 2022, el Modelo D muestre simultáneamente MAE y MASE muy bajos, DM claramente favorables frente a *snaive* y buena cobertura de intervalos con IS moderado. El deterioro observado en 2023-2024 es consistente con un escenario climático y epidemiológico anómalo que desestructura la estacionalidad local y desplaza el sistema fuera del régimen en el que fue entrenado. En presencia de forzantes interanuales, como ENSO, y anomalías marcadas de temperatura y lluvia, la coherencia entre ciclo anual y variabilidad interanual se altera, desplazando el calendario de transmisión y erosionando la habilidad de modelos lineales ajustados sobre un régimen previo (Colón-González et al., 2013; Nova et al., 2020; Muñoz et al., 2021). La señal digital puede desacoplarse de la curva clínica—por ruido mediático, cambios en el comportamiento de búsqueda o fatiga informativa—y perder su ventaja para anticipar picos (Gluskin et al., 2014; Cheong et al., 2025). Además, el clima urbano introduce heterogeneidad espacial adicional: islas de calor, diferencias en la amplitud térmica diaria, cobertura arbórea y uso del suelo pueden generar variaciones de varios grados entre barrios, modificando la capacidad vectorial a escalas que no se reflejan en promedios meteorológicos de ciudad (Banu et al., 2011; Misslin et al., 2016; Fujita et al., 2023). Factores sociales y entomológicos, como la intermitencia del suministro de agua, el almacenamiento doméstico, la densidad de susceptibles o los índices de infestación por *Aedes* spp., también condicionan la traducción del forzante climático en casos y, de no ser incorporados, limitan la “tracción” predictiva de los modelos (Nova et al., 2020; Ordoñez-Sierra et al., 2021). En este contexto, la expansión y consolidación de áreas de transmisión observada en grandes núcleos urbanos de Brasil y otros países (Banu et al., 2011; Barcellos & Lowe, 2014; Restrepo et al., 2014) ilustra cómo la combinación de clima adecuado, tamaño poblacional y conectividad puede generar regímenes de transmisión más persistentes y difíciles de modelar con especificaciones lineales simples.

Asimismo, estudios sobre desastres climáticos y vulnerabilidad social en Colombia muestran que los extremos hidrometeorológicos reconfiguran la exposición y los patrones de movilidad (Gelvez & Bauer, 2025), lo que sugiere que choques climáticos intensos pueden alterar el contexto estructural en el que operan los modelos de pronóstico de dengue. Metodológicamente, los hallazgos avalan la elección de modelos SARIMA/SARIMAX y la comparación rigurosa con *snaive* mediante DM, así como la evaluación probabilística

mediante coberturas e interval score (Diebold & Mariano, 1995; Tashman, 2000; Gneiting & Raftery, 2007; Hyndman & Athanasopoulos, 2021), pero también señalan líneas claras de mejora. Los p-valores bajos de Ljung-Box en los modelos con exógenas indican la necesidad de introducir no linealidades, umbrales—por ejemplo, temperaturas mínimas a partir de las cuales se activa la transmisión—interacciones clima-estación y mecanismos para detectar cambios de régimen (Colón-González et al., 2013; Xu et al., 2016; Nova et al., 2020). Del lado de las covariables, incorporar forzantes de gran escala (ENSO), desagregar el clima a escalas espaciotemporales más finas y enriquecer la dimensión social/entomológica podría reducir el error estructural (Banu et al., 2011; Restrepo et al., 2014; Muñoz et al., 2021).

En términos de señal digital, enfoques tipo ARGO con selección de términos específicos, regularización y reentrenamiento periódico pueden mitigar el ruido, la deriva de plataforma y el desacople temporal entre búsquedas y casos (Gluskin et al., 2014; Yang et al., 2017; Cheong et al., 2025). Operativamente, los resultados apoyan el desarrollo de sistemas de alerta temprana que combinen clima observado y pronosticado con indicadores digitales, evaluados de manera continua out-of-sample y con umbrales definidos en términos de MAE, RMSE, MASE y coberturas, más allá de la vigilancia puramente retrospectiva (Banu et al., 2011; Lowe et al., 2012; Ebi & Nealon, 2016; Fujita et al., 2023).

En síntesis, la evidencia empírica y la literatura convergen en que las variables exógenas mejoran la precisión de los modelos ARIMA para dengue, pero su aporte es heterogéneo y dependiente del régimen. En años con estacionalidad estable, clima y búsquedas añaden valor claro: reducen MAE y RMSE, mantienen MASE por debajo de 1, y logran coberturas cercanas a las nominales con interval scores moderados, en línea con lo descrito para otras regiones y contextos urbanos (Banu et al., 2011; Barcellos & Lowe, 2014; Restrepo et al., 2014; Fujita et al., 2023). En años anómalos, las ganancias se preservan más en términos de calibración que de precisión puntual, y se concentran en el Modelo D, que aprovecha la sinergia entre clima y conducta, aunque a costa de intervalos muy anchos y cierta autocorrelación residual. Esta lectura no solo explica por qué los modelos B, C y D superan al univariado A y a la línea base *snaive* en buena parte de los horizontes, sino que orienta mejoras concretas para robustecer el desempeño frente a shocks como el de 2023, avanzando hacia un sistema de predicción y alerta temprana más flexible, resiliente y sensible a los cambios de régimen climáticos y sociales.

Conclusion

Para uso operativo se requieren modelos adaptativos, centrados en calibración bajo años anómalos. En periodos con quiebres (p. ej., 2023), la señal digital puede desacoplarse de la curva clínica y el sistema responde con intervalos más anchos para sostener cobertura (como en 2024), por lo que la prioridad pasa de la precisión puntual a la gestión del riesgo. Esto exige incorporar no linealidades y detección de cambios de régimen, además de integrar forzantes de gran escala (p. ej., ENSO) dentro de un sistema de alerta temprana que combine clima y datos digitales.

Agradecimientos

La carrera de biología ha sido—para mi—una de las mejores decisiones que he podido tomar en mi vida, y es un camino que decidí tomar en mi vida desde pequeño, fantaseando en convertirme en un científico de laboratorio, sin embargo, he podido aprender que la biología va más allá de eso, más allá del laboratorio, del campo, de las reservas o parques naturales, del animal, de la planta, etc., es una ciencia tan extensa y maravillosa, y hay tantas cosas por hacer, y conocer, que nunca deja de sorprenderme, ni de cambiar mi forma de ver al mundo, por más insignificante que sea el nuevo conocimiento adquirido. Por ello, quiero agradecer a mis padres y mi hermano—mi familia en general—por su amor, paciencia, y apoyo incondicional durante estos

5 años de vida universitaria. A mis amigos, aquellos que vienen desde el colegio y a los que tuve la fortuna de conocer en la universidad, porque con el paso del tiempo, de las risas y de las clases, mas que amigos se han vuelto mi familia. También quiero agradecer a mis profesores, y mi director—David—que me acompañaron durante todo este proceso universitario, y que me permitieron conocer poco a poco lo valioso y vasto que es el mundo de la biología y las ciencias naturales. Y gracias a las empresas municipales de Cali (EMCALI), por darme la oportunidad de atender la universidad al financiar mi matrícula.

Referencias

- Alam, S., Islam, S. M. S., & Khan, M. A. (2025). Temporal trends, SARIMA forecasting of Dengue, and the influence of Dengue-related meteorological factors in Bangladesh: a time series analysis. *Environmental Research*, 241, 117565. <https://doi.org/10.1101/2025.04.09.25325511>
- Banu, S., Hu, W., Hurst, C., & Tong, S. (2011). Dengue transmission in the Asia-Pacific region: impact of climate change and socio-environmental factors. *Tropical Medicine and International Health*, 16(5), 598-607. <https://doi.org/10.1111/j.1365-3156.2011.02734.x>
- Barcellos, C., & Lowe, R. (2013). Expansion of the dengue transmission area in Brazil: the role of climate and cities. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 108(2), 63-71. <https://doi.org/10.1111/tmi.12227>
- Bengtsson, H. (2025) *Future: Unified Parallel and Distributed Processing in R for Everyone*. Package R version 1.67.0
- Bhatt, S., Gething, P., Brady, O. et al., (2013). The global distribution and burden of dengue. *Nature*, 496, 504-507. <https://doi.org/10.1038/nature12060>
- Chagas, G. C. L., Rangel, A. R., Noronha, L. M., et al., (2022). Risk factors for mortality in dengue: a systematic review and meta-analysis. *Trop Med Int Health*, 27(8): 656-668. <https://doi.org/10.1111/tmi.13797>
- Cheng, X. & Moraga, P. (2024). Assessing dengue forecasting methods: A comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil. *Trop Med Health* 53, 52. <https://doi.org/10.1186/s41182-025-00723-7>
- Cheong, K. H., Li, K., Yu, D., Zhao, X. (2025). Forecasting dengue cases through time-series modeling with Google Trends and deep neural networks. *Chaos, Solitons and Fractals*, 201, 117290. <https://doi.org/10.1016/j.chaos.2025.117290>
- Colón-González F. J., Fezzi C., Lake I. R., Hunter P. R. (2013) The Effects of Weather and Climate Change on Dengue. *PLoS Negl Trop Dis* 7(11): e2503. <https://doi.org/10.1371/journal.pntd.0002503>
- Dicker, Richard C. et al., (2006). *Principles of epidemiology in public health practice; an introduction to applied epidemiology and biostatistics*. 3rd ed. <https://stacks.cdc.gov/view/cdc/6914>
- Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- Ebi, K. L., & Nealon, J. (2016). Dengue in a changing climate. *Environmental Research*, 151, 115-123. <https://doi.org/10.1016/j.envres.2016.07.026>
- Fujita, D. M., Salvador, F. S., Nali, L. H. S., et al., (2023). Dengue and climate changes: Increase of DENV-1 in Sao Paulo/Brazil. *Travel Medicine and Infectious Disease*, 56, 102668. <https://doi.org/10.1016/j.tmaid.2023.102668>
- Gelvez, J. D., & Bauer, L. (2025). Climate Disasters, Armed Conflict, and Forced Displacement: A 23-Year Spatial Analysis in Colombia. *Social Science Quarterly*, 106, e70053. <https://doi.org/10.1111/ssqu.70053>
- Gluskin, R. T., Johansson, M. A., Santillana, M., & Brownstein, J. S. (2014). Evaluation of Internet-based dengue query data. *PLOS Neglected Tropical Diseases*, 8(2), e2713. <https://doi.org/10.1371/journal.pntd.0002713>
- Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477), 359-378. <https://doi.org/10.1198/016214506000001437>
- Gómez-Millán G., Cucunubá, Z. M., Huguett-Aragón, C., et al., (2024). *sivirep: Data Wrangling and Automated Reports from 'SIVIGILA' Source*. R package version 1.0.1.

- Guo, P., Wang, L., Zhang, Y. et al., (2017). Can internet search queries be used for dengue fever surveillance in China? *International Journal of Infectious Diseases*, 63, 74-76. <http://dx.doi.org/10.1016/j.ijid.2017.08.001>
- Hamner, B., Frasco, M., LeDell, E. (2018). *Metrics: Evaluation Metrics for Machine Learning*. R package version 0.1.4.
- Hasan, P., Khan, T. D., Alam, I., & Haque, M. E (2024). Dengue in Tomorrow: Predictive Insights from ARIMA and SARIMA Models in Bangladesh: A Time Series Analysis. *Health Science Reports*, 7: e70276. <https://doi.org/10.1002/hsr2.70276>
<https://doi.org/10.1111/nyas.13084>
- Hussain, S. S. A., & Dhiman, R. C. (2022). Distribution expansion of dengue vectors and climate change in India. *GeoHealth*, 6, e2021GH000477. <https://doi.org/10.1029/2021GH000477>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.).
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of forecasting*, 22, 679-688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Hyndman, R. J., Athanasopoulos, G., Bergmeir, C. et al., (2025). *forecast: Forecasting Functions for Time Series and Linear Models*. R package version 8.24.0.
- Instituto Nacional de Salud, Colombia. (2024). *Protocolo de vigilancia en salud pública: Dengue*. Versión 7.
- Jain, R., Sontisirikit, S., Iamsirithaworn, S., & Prendinger, H. (2019). Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data. *BMC Infectious Diseases*, 19(1), 272. <https://doi.org/10.1186/s12879-019-3874-x>
- Khairul, M. S. B. M., Dom, N. C., Abdullah, S., Abdullah, A. M. (2021). Correlation analysis between dengue incidence and Google Trends data. *Teikyo Medical Journal*, 44(5).
- Ljung, G. M., & Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65(2), 297-303. <https://doi.org/10.2307/2335207>
- Lowe, R., Bailey, T. C., Stephenson, D. B., et al., (2012) The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. *Statist. Med.*, 32, 864-883. <https://doi.org/10.1002/sim.5549>
- Mislin, R., Telle, O., Daudé, E., et al., (2016). Urban climate versus global climate change: what makes the difference for dengue? *Annals of the New York Academy of Sciences*, 1382(1), 56-72.
- Morin, C. W., Comrie, A. C., & Ernst, K. (2013). Climate and dengue transmission: evidence and implications. *Environmental Health Perspectives*, 121(11-12), 1264-1272 <https://doi.org/10.1289/ehp.1306556>.
- Muñoz, E., Poveda, G., Arbelaéz, M. P., & Vélez, I. D. (2021). Spatiotemporal dynamics of dengue in Colombia in relation to local climate and ENSO. *Acta Tropica*, 224, 106136. <https://doi.org/10.1016/j.actatropica.2021.106136>
- Weir, L., & Mykhalovskiy, E. (2009). Global Public Health Vigilance: Creating a World on Alert.
- NASA Langley Research Center. (2023). *NASA POWER Project: Prediction Of Worldwide Energy Resources*.
- Nova, N., Deyle, E. R., Shocket, M. S., et al., (2020). Susceptible host availability modulates climate effects on dengue dynamics. *Ecology letters*, 24(3), 415-425. <https://doi.org/10.1111/ele.13652>
- Ooms J (2025b). *writexl: Export Data Frames to Excel 'xlsx' Format*. R package version 1.5.4.
- Ooms, J., Lang, D. T., Hilaiel, L. (2025). *jsonlite: A Simple and Robust JSON Parser and Generator for R*. R package version 2.0.0.
- Ordoñez-Sierra, R., Sarmiento-Senior, D., Gomez, J. F. J. et al., (2021). Multilevel analysis of dengue determinants in Colombia. *One Health*, 12, 100234. <https://doi.org/10.1016/j.onehlt.2021.100234>
- Pongsumpun, P., Lopez, D. G., Favier, C., et al., (2008). Dynamics of dengue epidemics in urban contexts. *Tropical Medicine and International Health*, 13(9), 1180-1187. <https://doi.org/10.1111/j.1365-3156.2008.02124.x>
- R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reich, N. G., Lauer, S. A., Sakrejda, K., et al., (2016). Challenges in real-time prediction of infectious disease: A case study of dengue in Thailand. *PLoS Neglected Tropical Diseases*, 10(6), e0004761. <https://doi.org/10.1371/journal.pntd.0004761>

- Restrepo, A. C., Baker, P., & Clemens, A. C. A. (2014). National spatial and temporal patterns of notified dengue cases, Colombia 2007-2010. *Tropical Medicine and International Health*, 19(7), 863-871. <https://doi.org/10.1111/tmi.12325>
- Sebastianelli, A., Spiller, D., Carmo, R. et al., (2024). A reproducible ensemble machine learning approach to forecast dengue outbreaks. *Nature Sci. Rep.*, 14, 3807. <https://doi.org/10.1038/s41598-024-52796-9>
- Spinu, V., Grolemond, G., Wickham, H. et al., (2024). *lubridate: Make Working with Dates in R Easier*. R package versión 1.9.4.
- Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy as a framework for model selection. *International Journal of Forecasting*, 16, 437-450. [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)
- Tuan, D. A. & Dang, T. N. (2024). Leveraging Climate Data for Dengue Forecasting in Ba Ria Vung Tau Province, Vietnam: An Advanced Machine Learning Approach. *Trop. Med. Infect. Dis.*, 9, 250. <https://doi.org/10.3390/tropicalmed9100250>
- Vaughan, D. (2022) *furrr: Apply Mapping Functions in Parallel using Futures*. Package R version 0.3.1
- Waggoner, J. J., Gresh, L., Vargas, M. J., et al., (2016). Viremia and Clinical Presentation in Nicaraguan Patients Infected with Zika Virus, Chikungunya Virus, and Dengue Virus. *Clinical Infectious Diseases*, 63(12), 1584-1590. <https://doi.org/10.1093/cid/ciw589>
- Wickham, H. (2023). *httr: Tools for Working with URLs and HTTP*. R package versión 1.4.7.
- Wickham, H. (2023b). *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 2.0.0.
- Wickham, H., Chang, W., Henry, L. et al., (2025). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. R package version 3.5.2.
- World Health Organization. (2009). *Dengue: Guidelines for diagnosis, treatment, prevention, and control*. New edition. Geneva: World Health Organization.
- Xiao, Y., Soares, G., Bastos, L. et al., (2024). Nowcasting dengue in Brazil using meteorological and digital data. *MedRxiv*. <https://doi.org/10.1101/2024.09.02.24312934>
- Xu, L., Stige, L. C., Chan, K. S., et al., (2016). Climate variation drives dengue dynamics. *PNAS*, 114(1), 113-118. <https://doi.org/10.1073/pnas.161855811>
- Yang, S., Kou, S. C., Lu, F., et al., (2017). Advances in using Internet searches to track dengue. *PLOS Computational Biology*, 13(7), e1005607. <https://doi.org/10.1371/journal.pcbi.1005607>
- Zhang, Z. (2019). Predictive analytics in the era of big data: Concepts, models, and applications. *Annals of Translational Medicine*, 8(4). <http://dx.doi.org/10.21037/atm.2019.10.97>

Anexo 1. Repositorio con los códigos de R usados para los modelos.

https://github.com/Juan-David-Castillo/Codigo_Tesis_Biolog-a_Javeriana

Anexo 2. Tablas de métricas completas Modelo A.

Año	2021	2022	2023	2024
MAE	22.28	18.03	226.54	470.91
RMSE	29.72	25.08	288.79	538.87
sMAPE	20.06	27.36	129.79	70
WMAPE	20.31	33.2	89.07	70.32
MdAPE	17.63	21.61	90.98	42.92
MAPE	21.54	34.99	72.86	1.87E+11
RMSLE	0.2642	0.3605	2.2683	1.3451
MASE	0.153	0.124	1.627	3.291
Cov80	100	90.4	30.8	75.5
Cov95	100	98.1	42.3	98.1
IS80	322.63	218.56	1673.98	2107.47
IS95	632.64	407.37	3246.54	3615.19
DM_stat	-4.03672	-6.22548	8.07488	-2.0789
DM_pvalue	9.11E-05	4.54E-08	1	0.021288
LB_stat	47.14168	44.9831	44.53397	47.1809
LB_pvalue	0.466762	0.556461	0.57527	0.46516

Anexo 3. Tablas de métricas completas Modelo B.

Año	2021	2022	2023	2024
MAE	27.66	21.12	222.31	425.71
RMSE	36.17	25.34	282.41	486.43
sMAPE	24.12	38.8	126.3	66.76
WMAPE	25.22	38.88	87.41	63.57
MdAPE	21.54	33.39	88.82	40.94
MAPE	29.2	37.77	72.84	1.54E+11
RMSLE	0.315	0.4411	2.0106	1.2832
MASE	0.19	0.146	1.597	2.975
Cov80	100	92.3	32.7	98.1
Cov95	100	100	73.1	98.1
IS80	404.06	177.31	1438.77	3190.09
IS95	925.21	430.7	957.7	7613.83
DM_stat	-3.978	-5.83	7.74	-3.014
DM_pvalue	0.0002	0	0	0.004
LB_stat	89.25	63.03	87.056	87.715
LB_pvalue	0.0003	0.0591	0.0003	0.0003

Anexo 4. Tablas de métricas completas Modelo C.

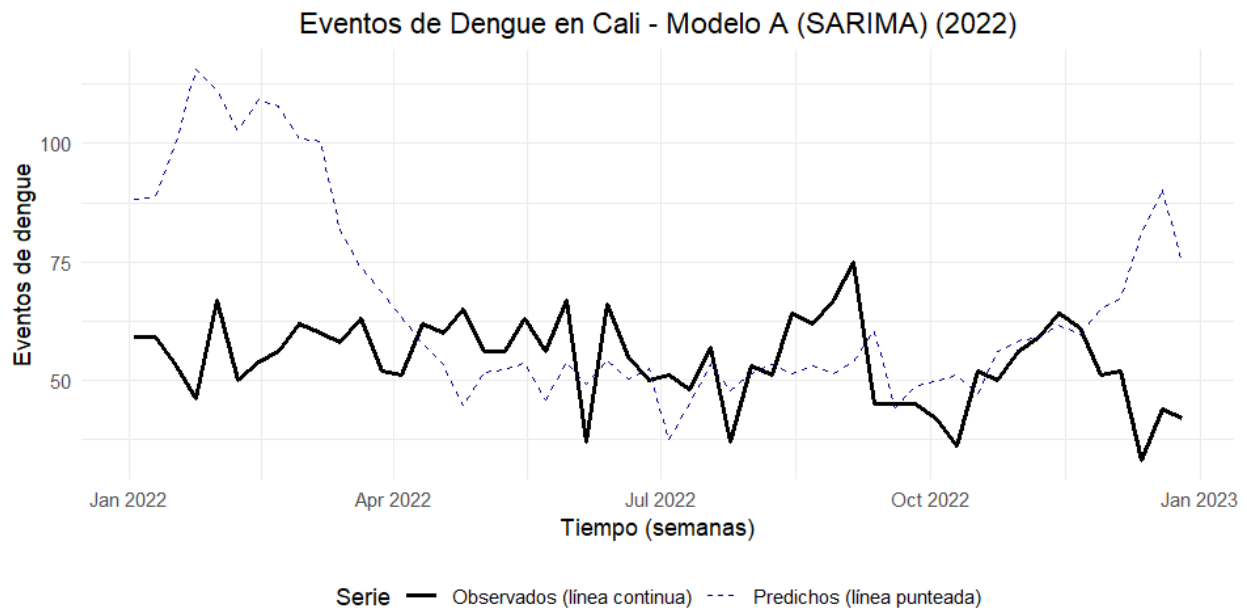
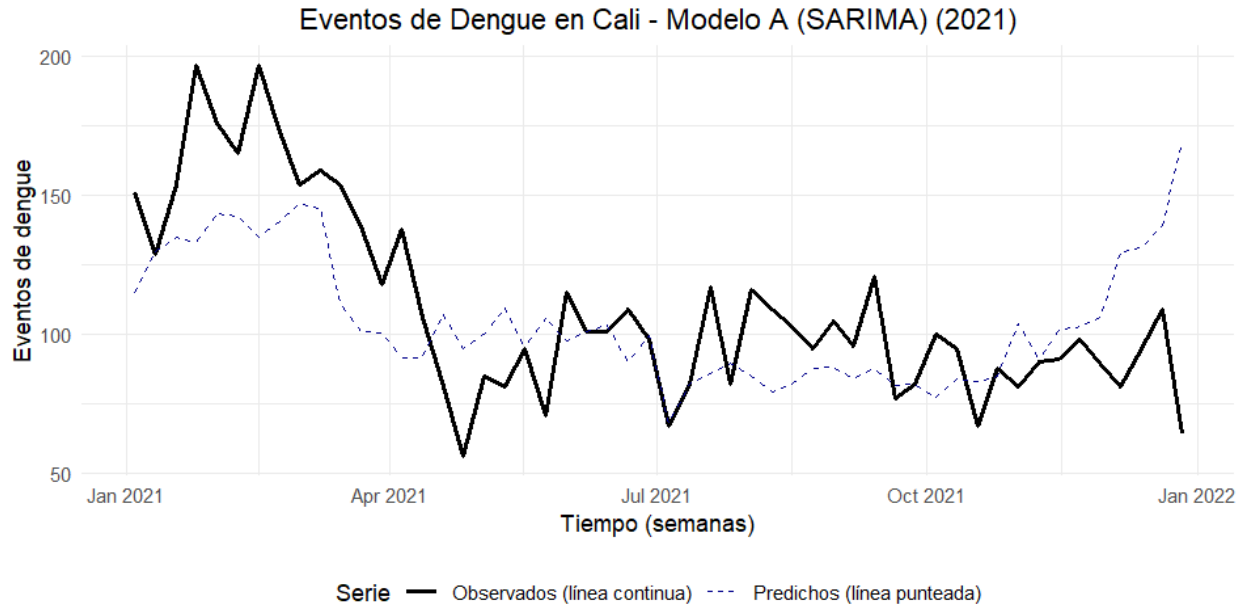
Año	2021	2022	2023	2024
MAE	26.01	16.26	223.29	453.49
RMSE	34.32	21.35	282.19	577.07
sMAPE	23.43	26.51	129.06	63.32
WMAPE	23.72	29.92	87.8	67.72
MdAPE	22.26	23.42	88.91	58.3
MAPE	22.96	31.97	74.04	1.47E+11

RMSLE	0.2979	0.3357	2.023	1.2997
MASE	0.179	0.112	1.604	3.169
Cov80	100	94.2	30.8	98.1
Cov95	100	100	78.8	98.1
IS80	382.53	238.99	1435.94	4120.97
IS95	973.25	581.99	836.15	9806.41
DM_stat	-4.03488	-5.96597	8.349041	-1.2458
DM_pvalue	9.17E-05	1.16E-07	1	0.109212
LB_stat	82.56447	80.47859	82.14829	83.23406
LB_pvalue	8.7E-05	0.000155	9.76E-05	7.21E-05

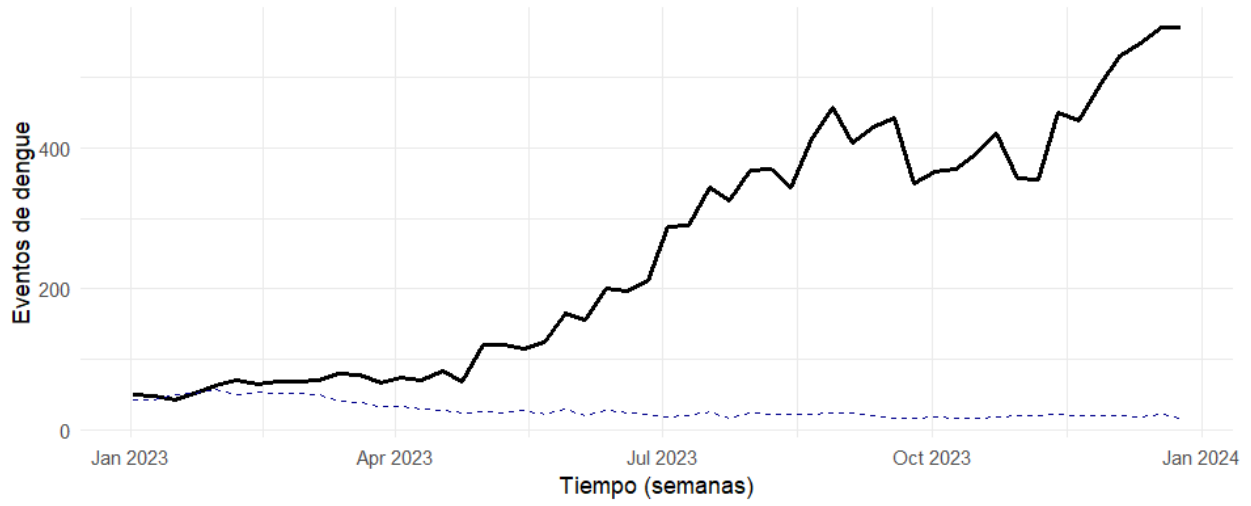
Anexo 5. Tablas de métricas completas Modelo D.

Anio	2021	2022	2023	2024
MAE	29.82	9.93	168.07	970.87
RMSE	34.9	13.17	210.42	1635.28
sMAPE	26.01	17.71	83.23	52.84
WMAPE	27.19	18.29	66.09	144.97
MdAPE	25.41	13.05	66.42	34.62
MAPE	30.08	21.2	56.72	1.56E+10
RMSLE	0.3052	0.2398	0.9836	0.9416
MASE	0.203	0.068	1.197	6.731
Cov80	98.1	100	59.6	66
Cov95	100	100	100	98.1
IS80	351.27	177.21	355.62	4738.21
IS95	774.68	396.84	521.43	6600.77
DM_stat	-4.015	-5.79	-5.091	3.07
DM_pvalue	0.0001	0	0	0.9983
LB_stat	101.212	103.958	116.34	112.49
LB_pvalue	0	0	0	0

Anexo 6. Figuras Modelo A. La línea negra continua representa los eventos de dengue observados, la línea azul discontinua representa los eventos que fueron predichos por el modelo A.

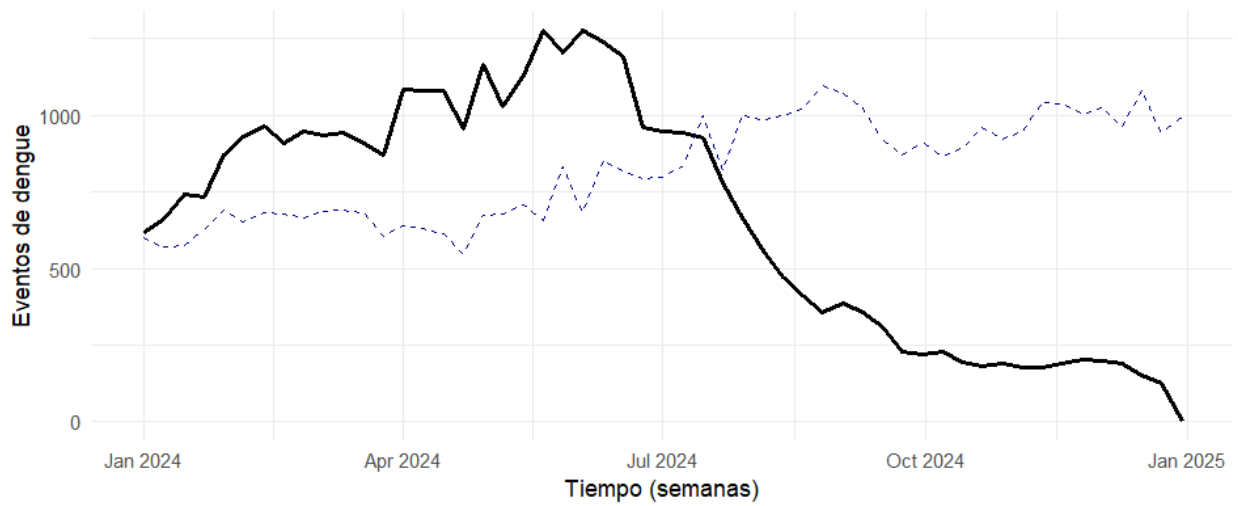


Eventos de Dengue en Cali - Modelo A (SARIMA) (2023)



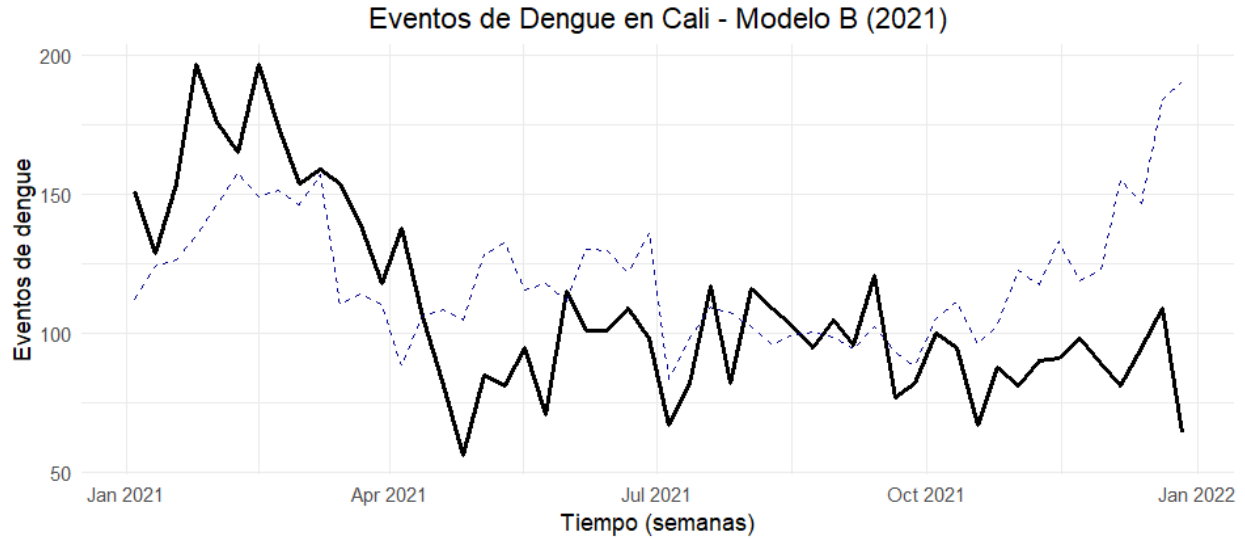
Serie — Observados (línea continua) - - - Predichos (línea punteada)

Eventos de Dengue en Cali - Modelo A (SARIMA) (2024)

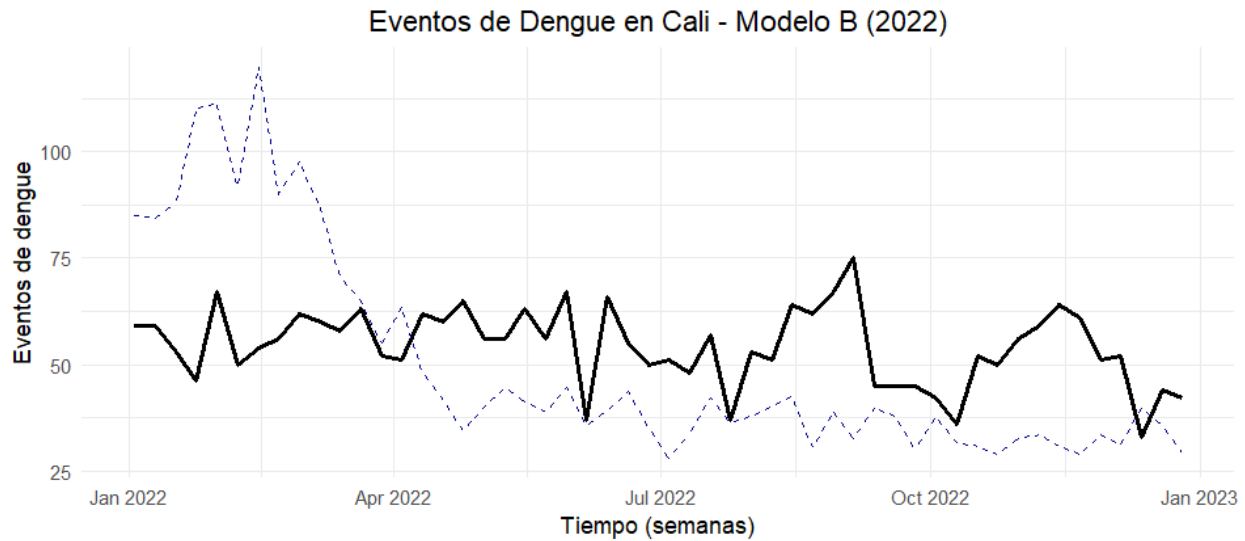


Serie — Observados (línea continua) - - - Predichos (línea punteada)

Anexo 7. Figuras Modelo B. La línea negra continua representa los eventos de dengue observados, la línea azul discontinua representa los eventos que fueron predichos por el modelo B. La línea roja representa la temperatura, la línea verde representa la humedad, la línea amarilla representa la precipitación. La información en las gráficas con variables exógenas se encuentra transformada en una escala (0-1) para facilitar la visualización.

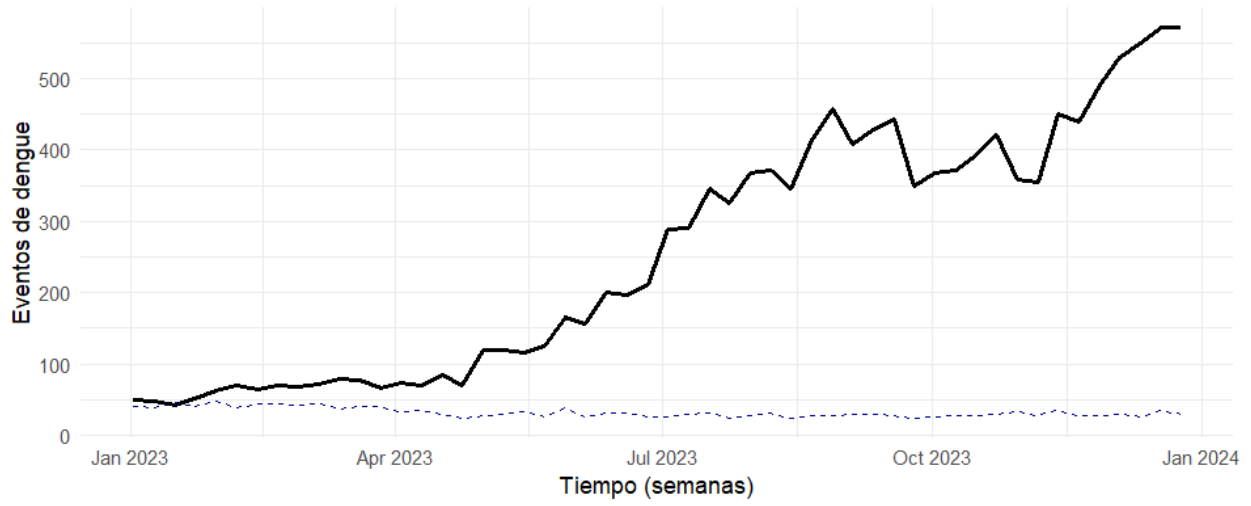


Serie — Observados (línea continua) - - - Predichos (línea punteada)



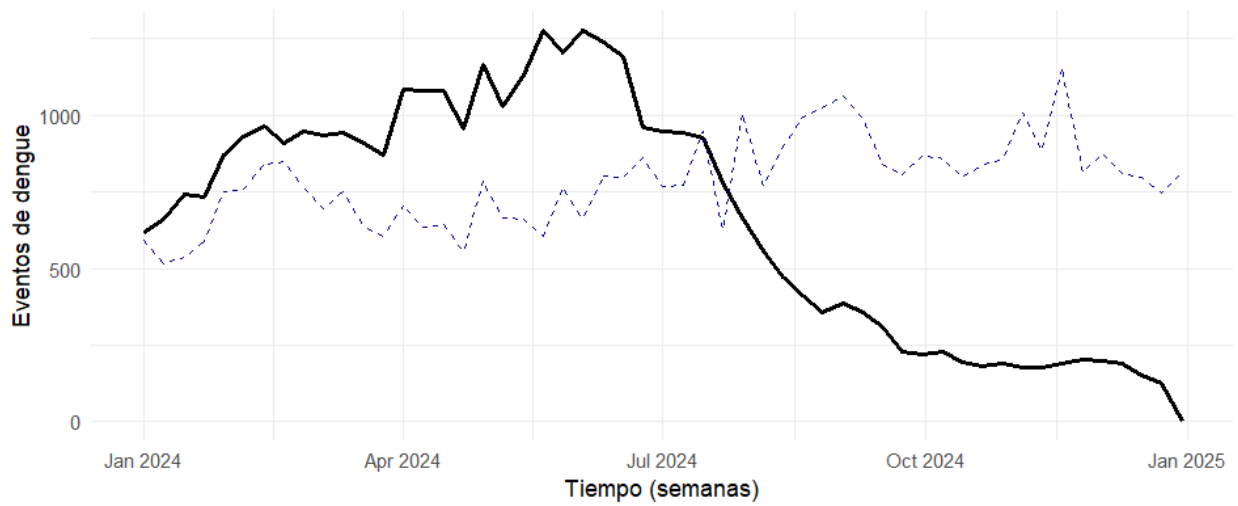
Serie — Observados (línea continua) - - - Predichos (línea punteada)

Eventos de Dengue en Cali - Modelo B (2023)



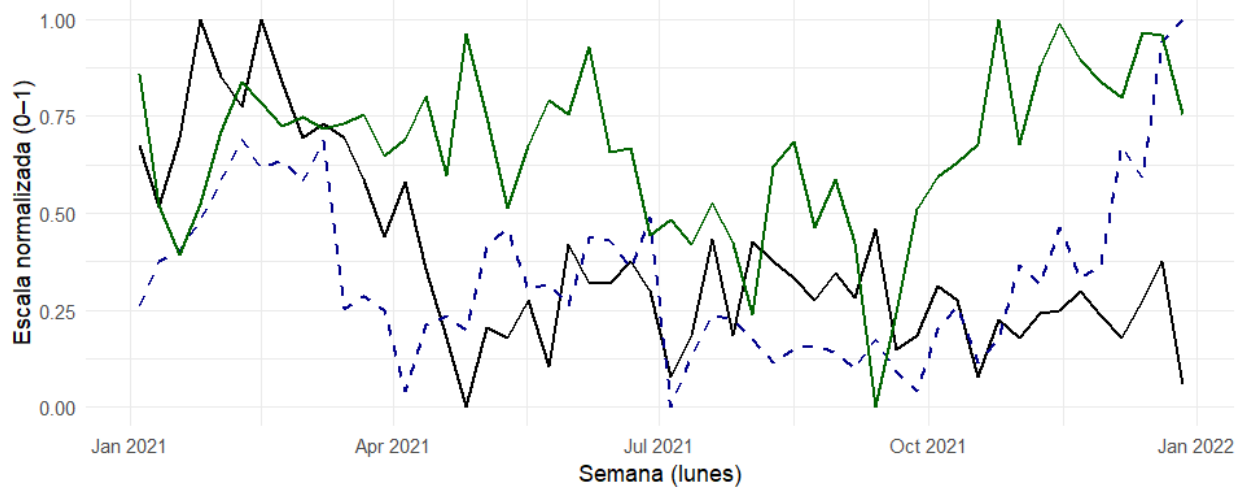
Serie — Observados (línea continua) - - - Predichos (línea punteada)

Eventos de Dengue en Cali - Modelo B (2024)



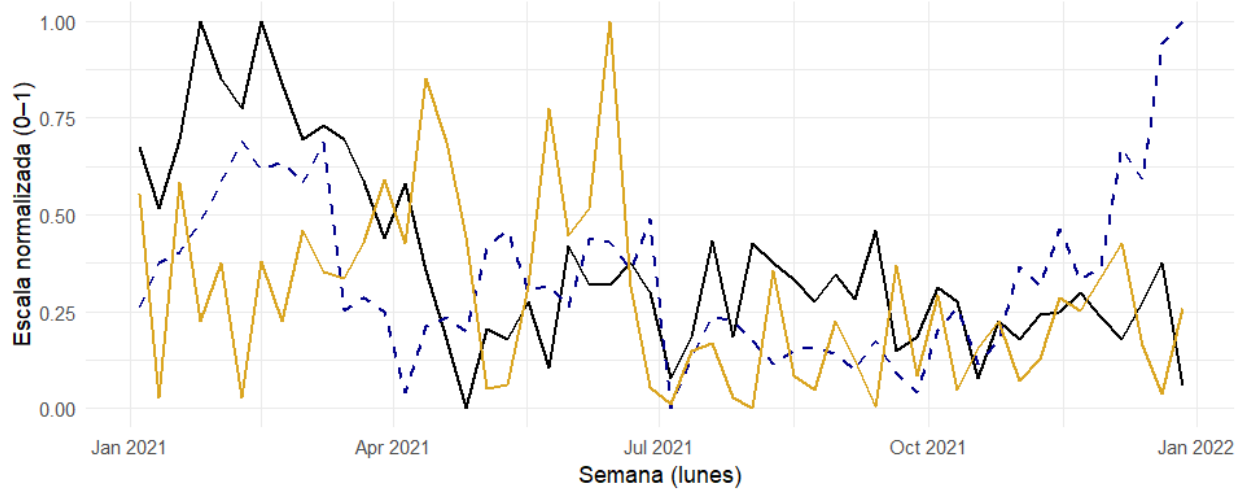
Serie — Observados (línea continua) - - - Predichos (línea punteada)

Observados y predichos vs variable exógena (0-1) - Humedad (2021)



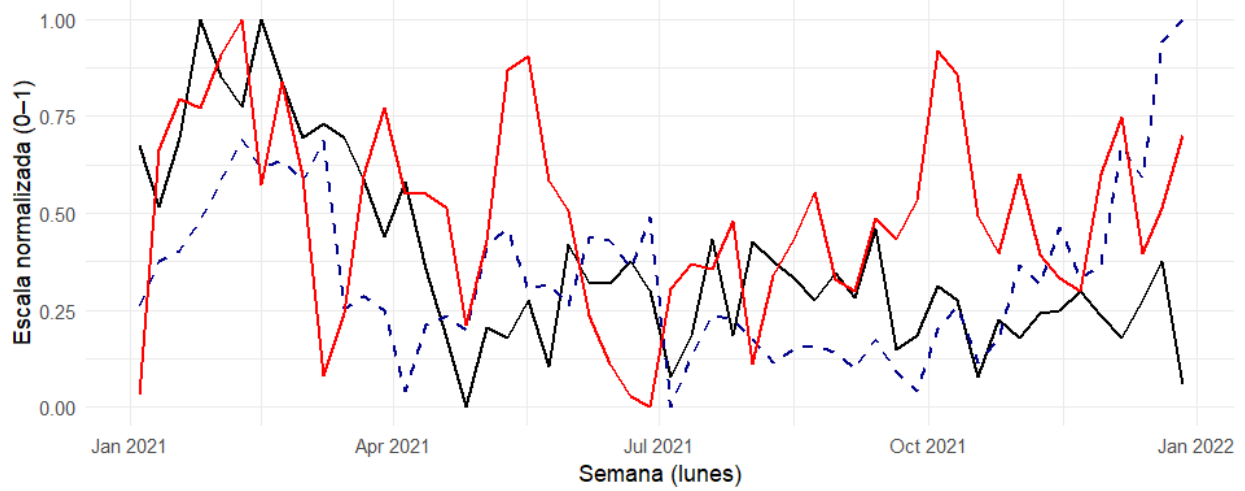
Serie — Observados - - Predichos — Humedad

Observados y predichos vs variable exógena (0-1) - Precipitación (2021)



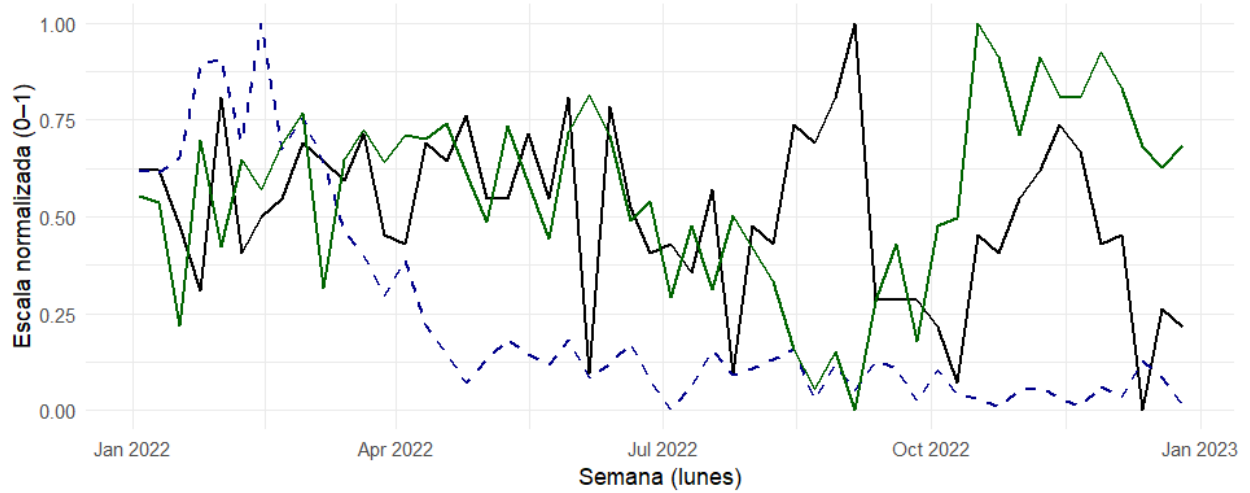
Serie — Observados - - Predichos — Precipitación

Observados y predichos vs variable exógena (0-1) - Temperatura (2021)



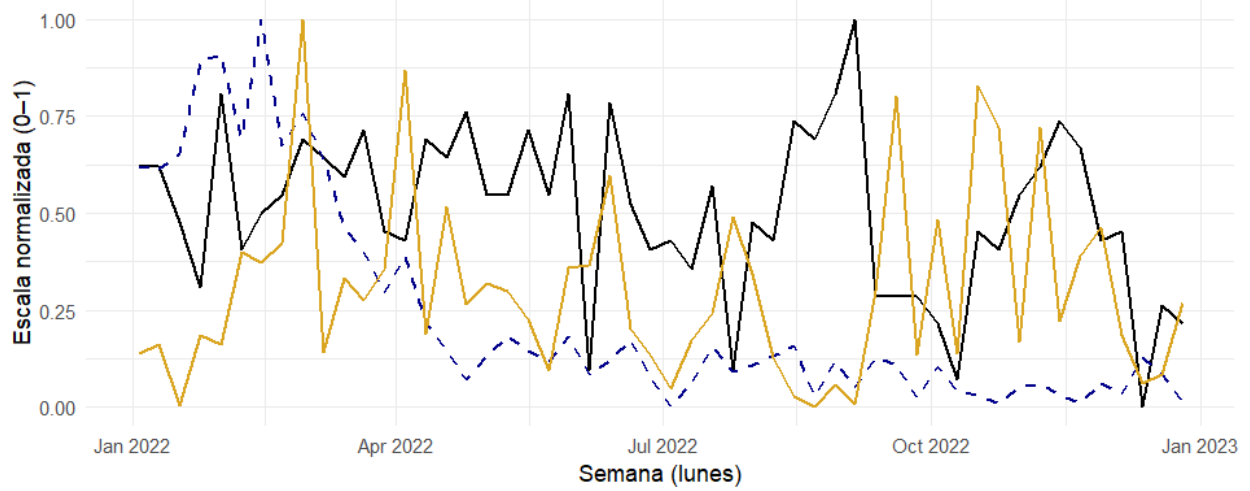
Serie — Observados - - Predichos — Temperatura

Observados y predichos vs variable exógena (0-1) - Humedad (2022)



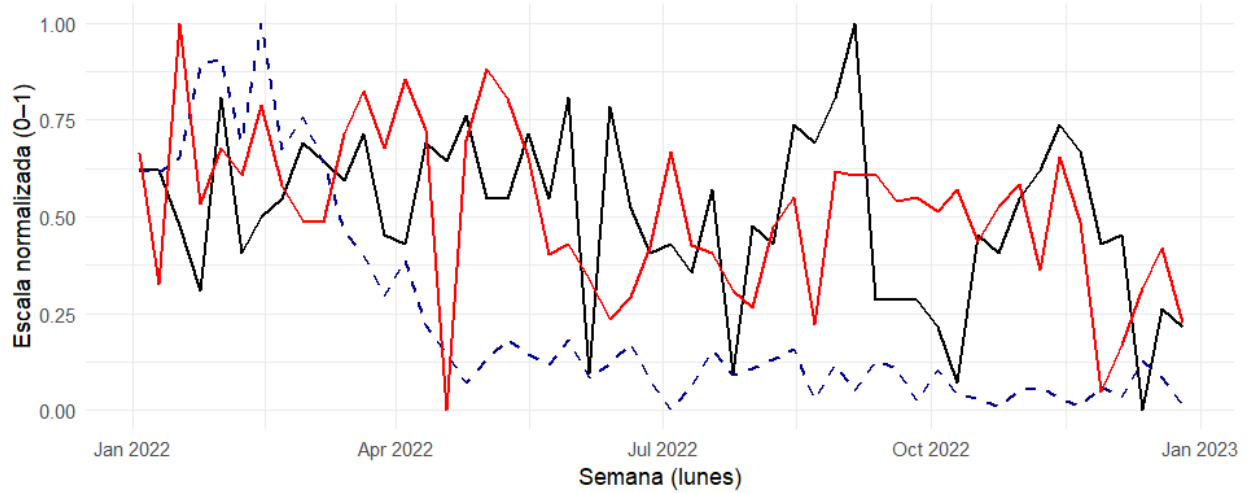
Serie — Observados - - Predichos — Humedad

Observados y predichos vs variable exógena (0-1) - Precipitación (2022)



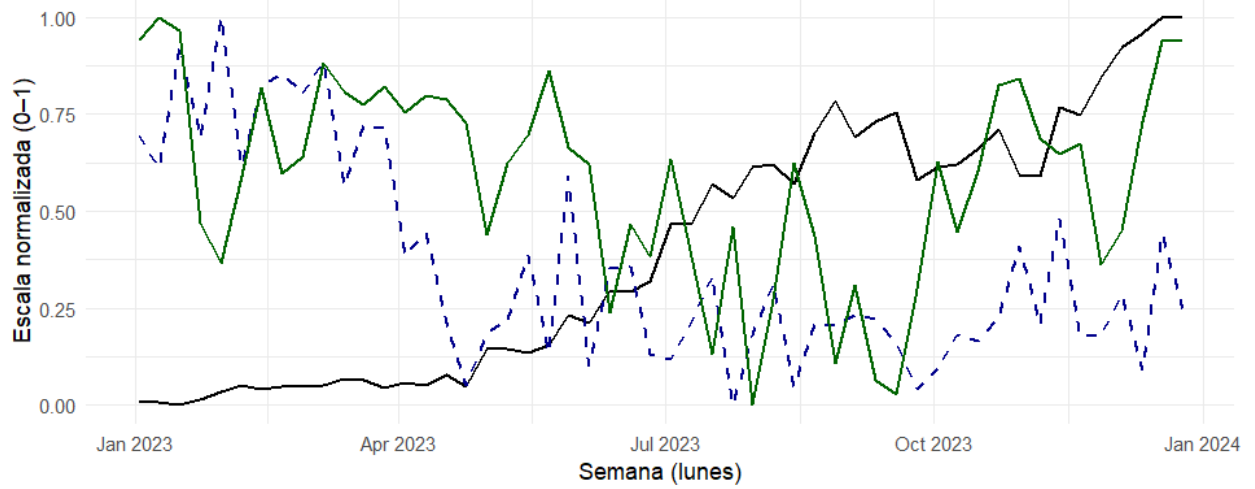
Serie — Observados - - Predichos — Precipitación

Observados y predichos vs variable exógena (0-1) - Temperatura (2022)



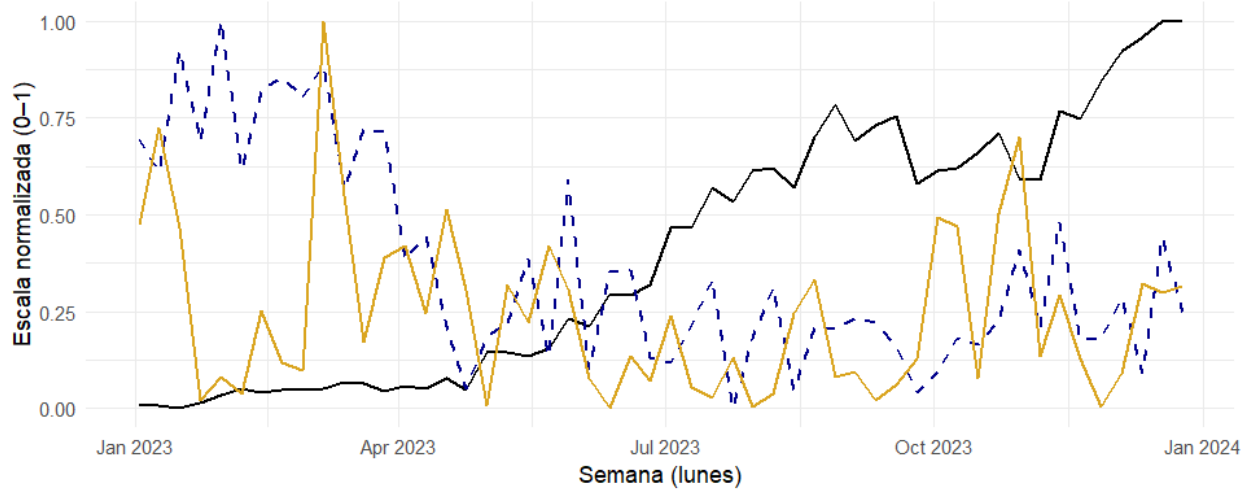
Serie — Observados - - Predichos — Temperatura

Observados y predichos vs variable exógena (0–1) - Humedad (2023)



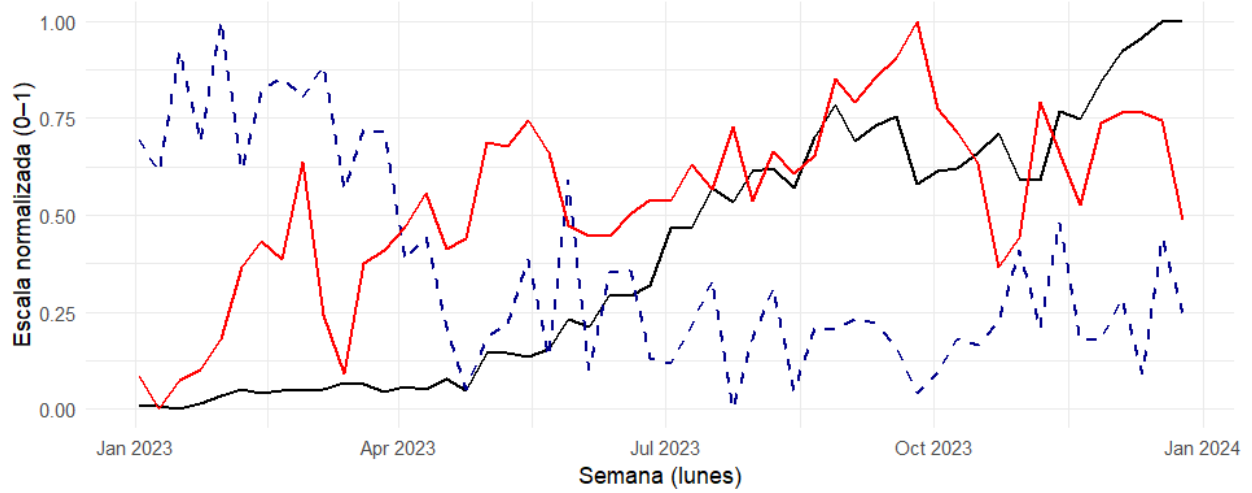
Serie — Observados - - Predichos — Humedad

Observados y predichos vs variable exógena (0–1) - Precipitación (2023)



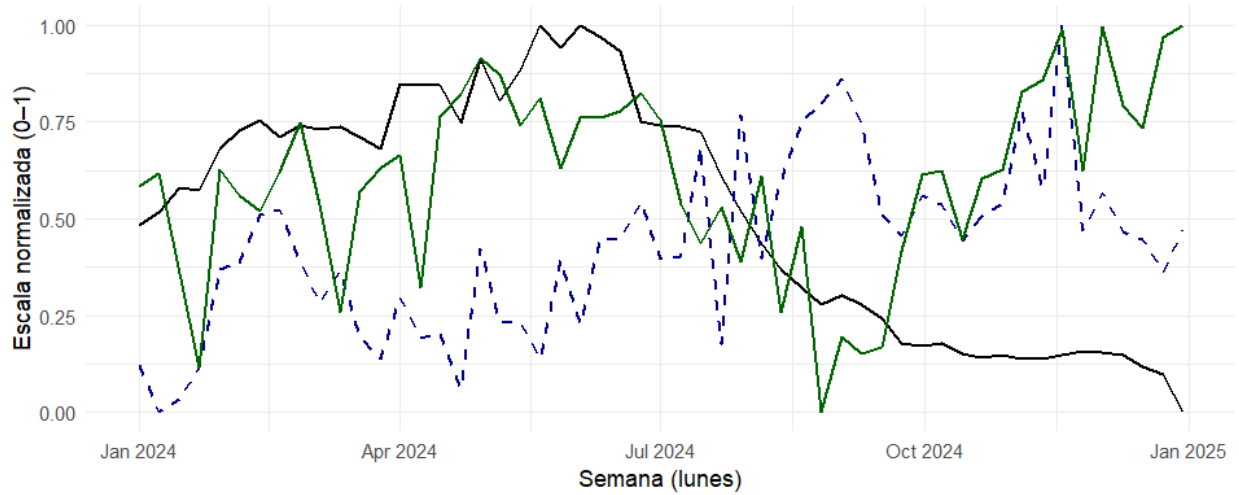
Serie — Observados - - Predichos — Precipitación

Observados y predichos vs variable exógena (0-1) - Temperatura (2023)



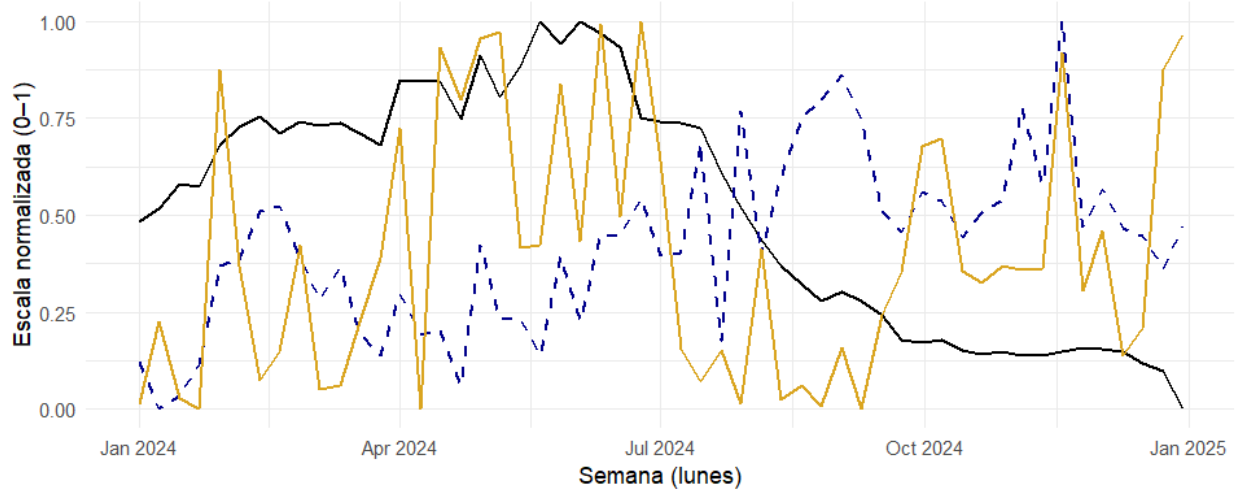
Serie — Observados - - Predichos — Temperatura

Observados y predichos vs variable exógena (0-1) - Humedad (2024)



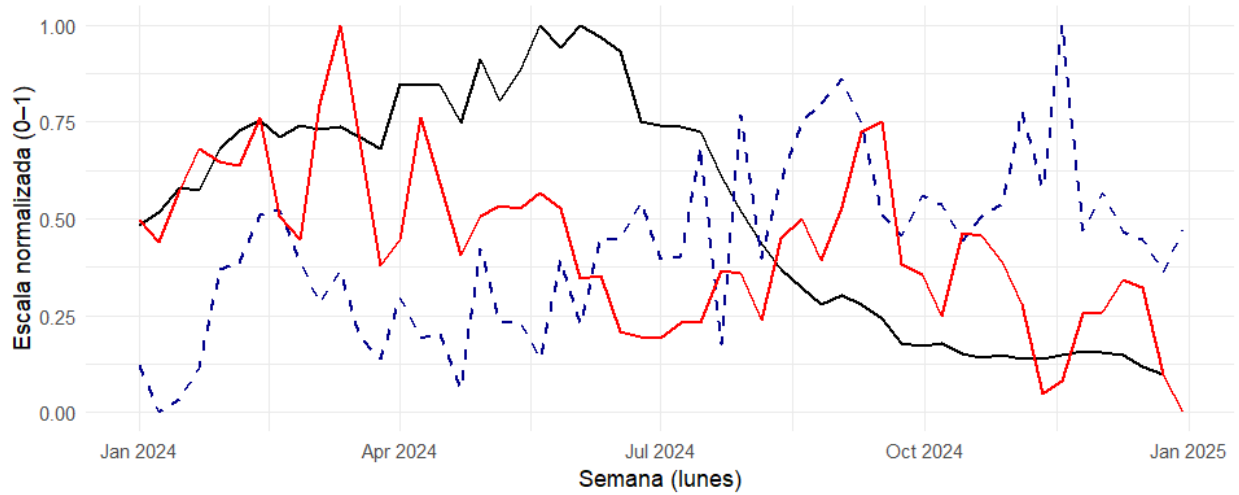
Serie — Observados - - Predichos — Humedad

Observados y predichos vs variable exógena (0-1) - Precipitación (2024)



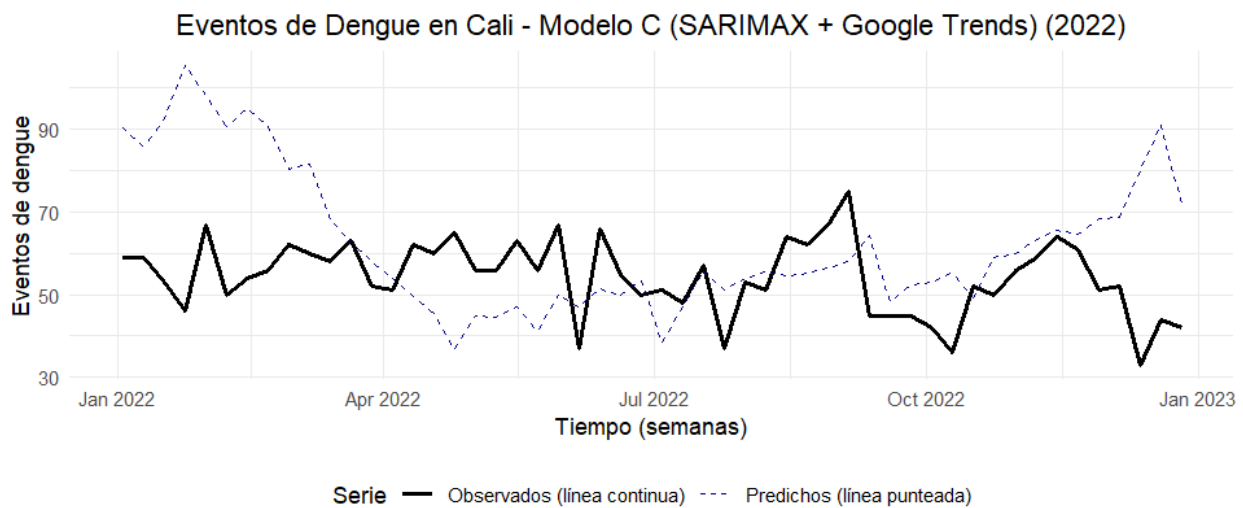
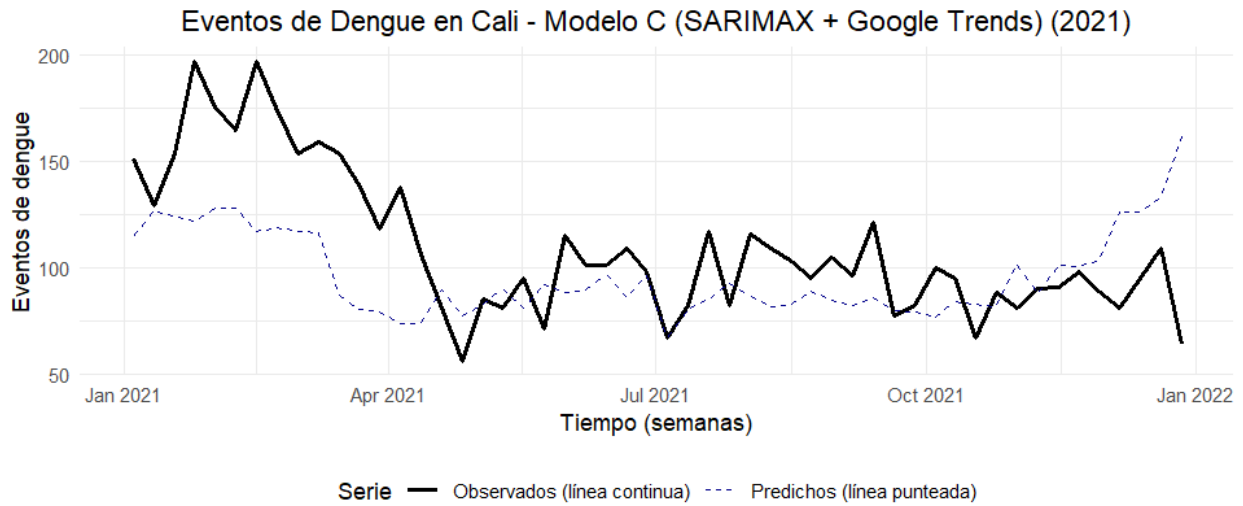
Serie — Observados - - Predichos — Precipitación

Observados y predichos vs variable exógena (0-1) - Temperatura (2024)

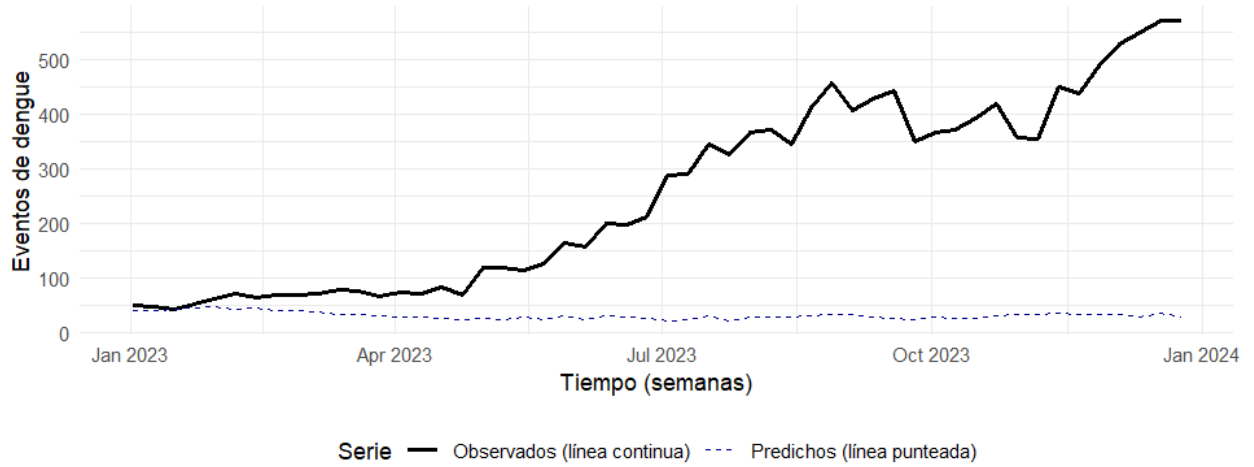


Serie — Observados - - Predichos — Temperatura

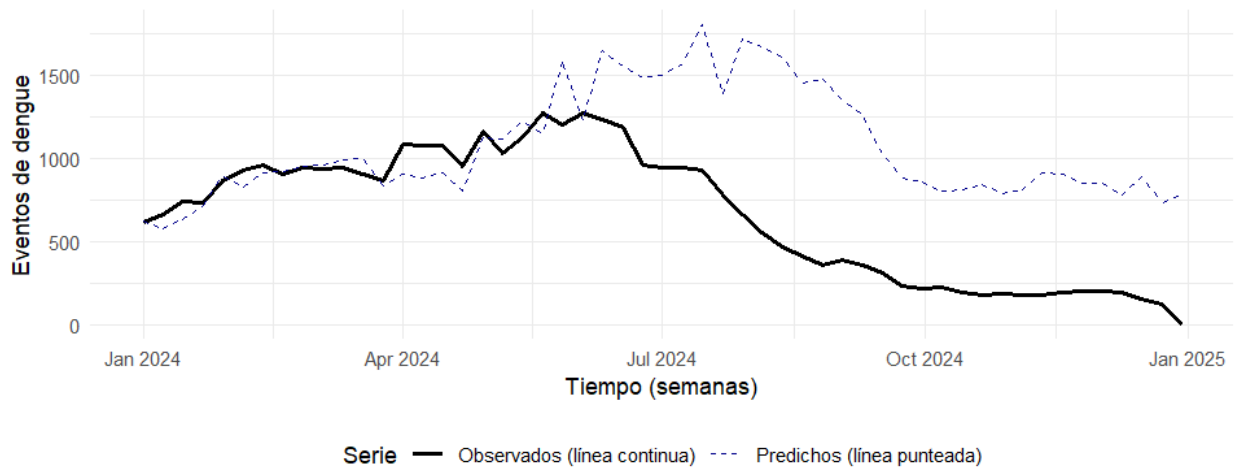
Anexo 8. Figuras Modelo C. La línea negra continua representa los eventos de dengue observados, la línea azul discontinua representa los eventos que fueron predichos por el modelo C. La línea roja representa el valor del índice de Google Trends. La información en las gráficas con variables exógenas se encuentra transformada en una escala (0-1) para facilitar la visualización.

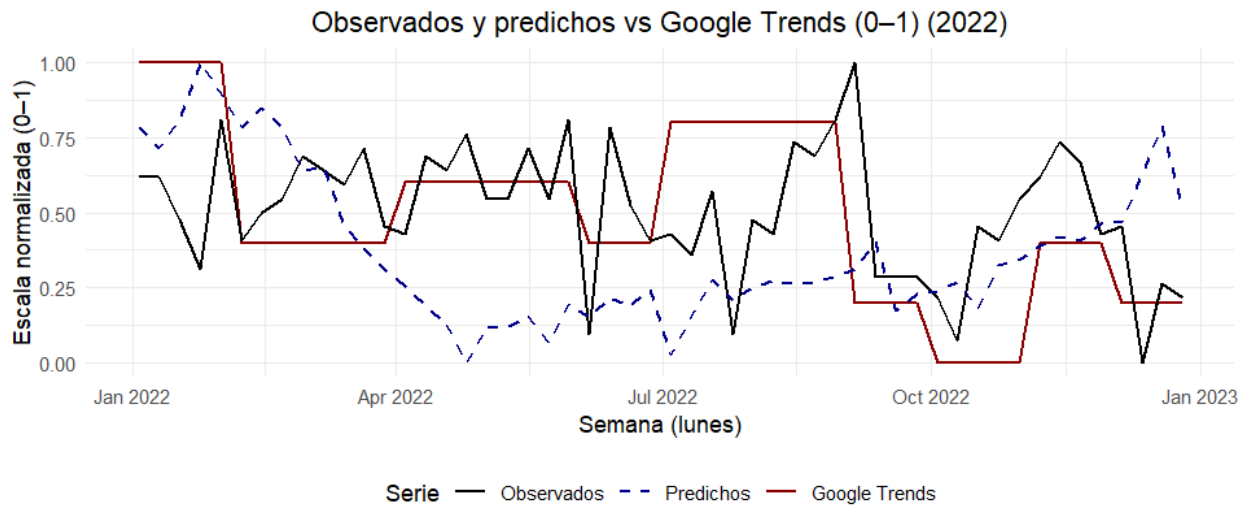
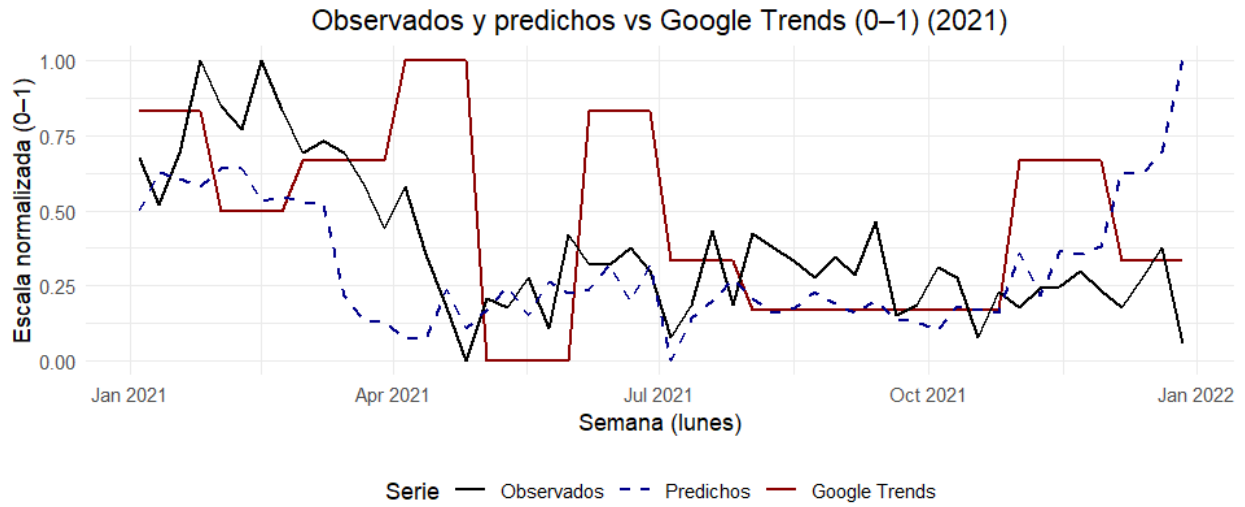


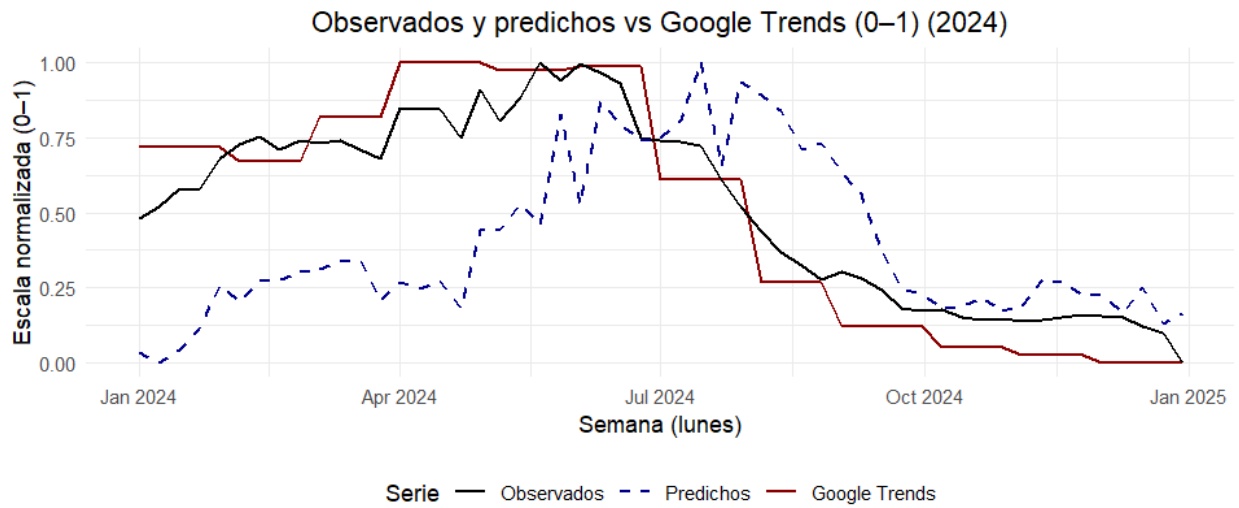
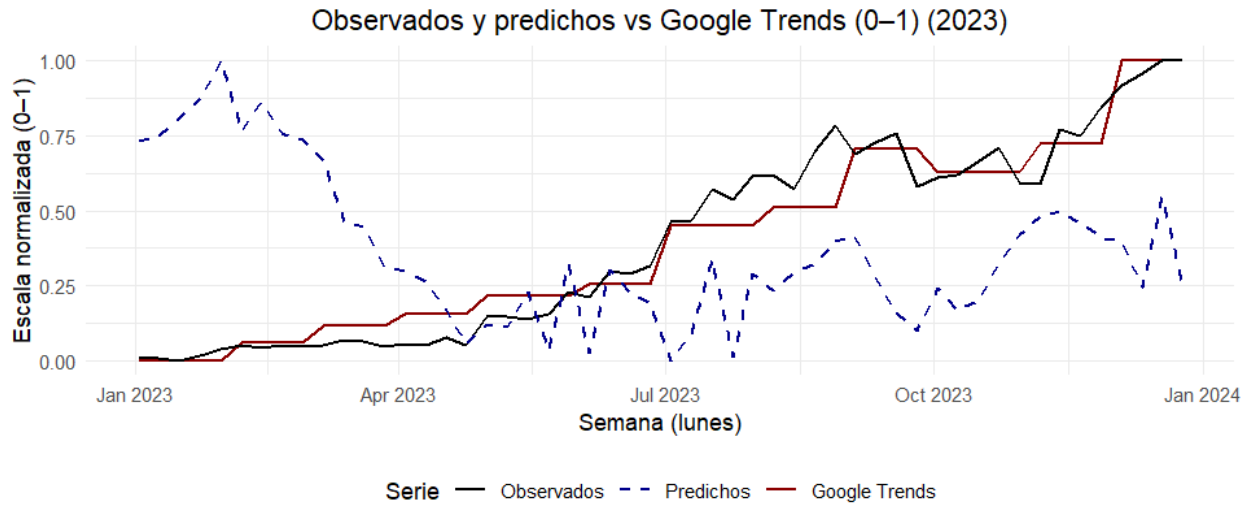
Eventos de Dengue en Cali - Modelo C (SARIMAX + Google Trends) (2023)



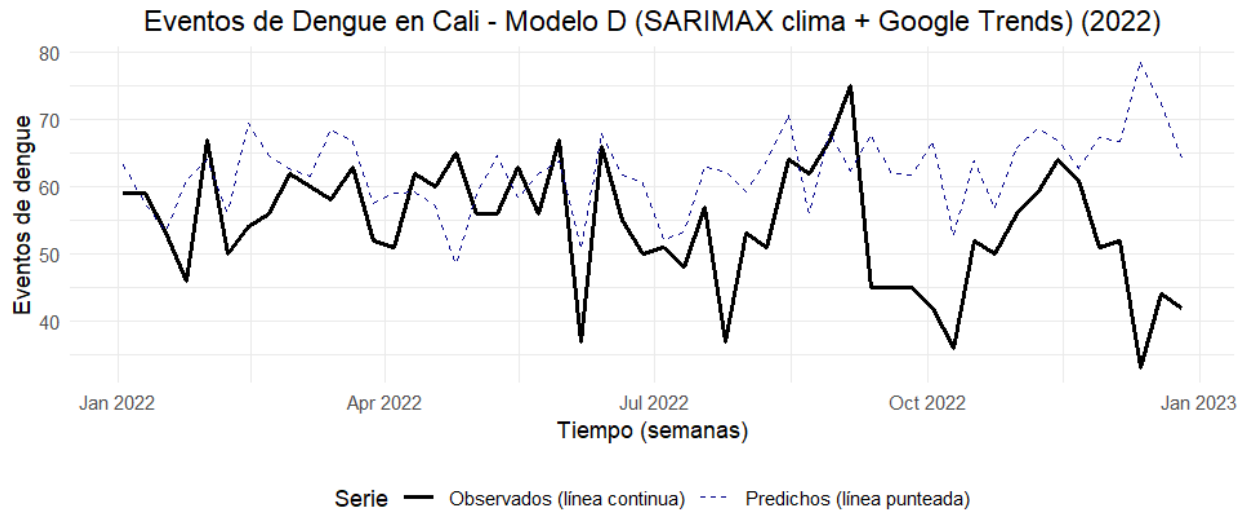
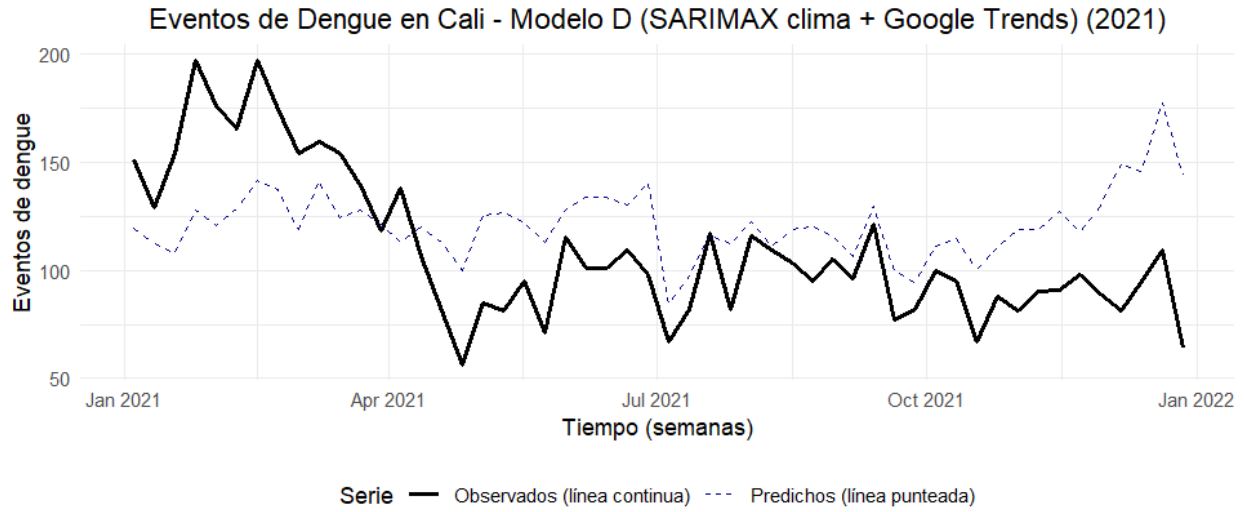
Eventos de Dengue en Cali - Modelo C (SARIMAX + Google Trends) (2024)



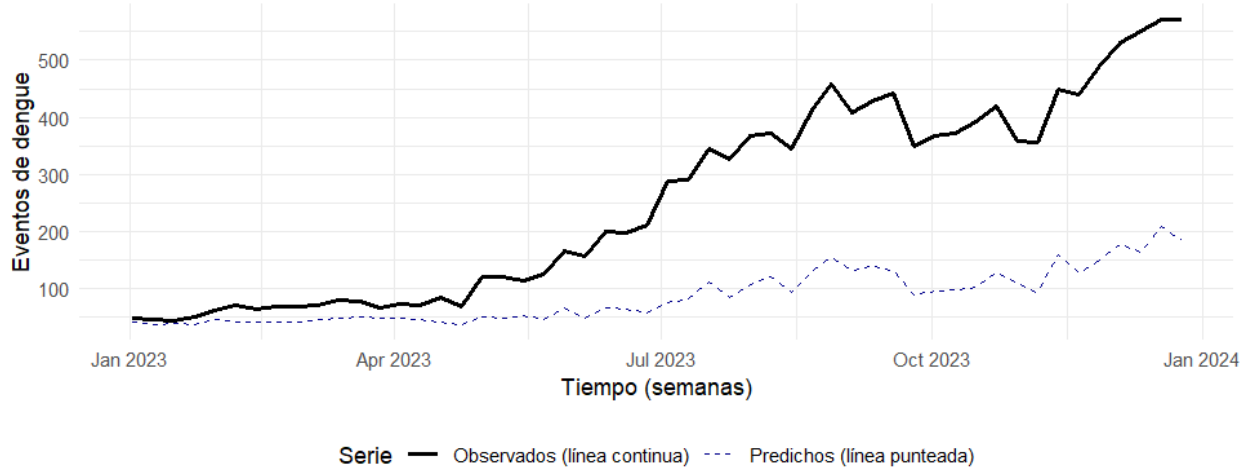




Anexo 9. Figuras Modelo D. La línea negra continua representa los eventos de dengue observados, la línea azul discontinua representa los eventos que fueron predichos por el modelo D.



Eventos de Dengue en Cali - Modelo D (SARIMAX clima + Google Trends) (2023)



Eventos de Dengue en Cali - Modelo D (SARIMAX clima + Google Trends) (2024)

