

Modelos matemáticos para el entendimiento del dengue

María Eugenia

Puerta Yepes

Editora académica



MATEMÁTICAS

COLECCIÓN ACADÉMICA

Modelos matemáticos para el entendimiento del dengue

María Eugenia Puerta Yepes

Editora académica



Modelos matemáticos para el entendimiento del dengue / Andrea Arévalo-Cortés...
[et al.] ; María Eugenia Puerta Yepes, editora académica. -- Medellín : Editorial EAFIT, 2025.

254 p. ; il. ; 24 cm. -- (Académica).

ISBN: 978-958-720-986-0

ISBN: 978-958-720-987-7 (versión EPUB)

ISBN: 978-958-720-988-4 (versión PDF)

1. Dengue – Diagnóstico. 2. Dengue – Prevención y control. 3. Dengue – Modelos matemáticos. 4. Dengue – Investigaciones – Colombia. 5. Epidemiología – Modelos matemáticos. 6. Enfermedades transmitidas por vectores. I. Puerta Yepes, María Eugenia, edit. II. Tít. III. Serie.

614.58852 cd 23 ed.

M689

Universidad EAFIT - Centro Cultural Biblioteca Luis Echavarría Villegas

Modelos matemáticos para el entendimiento del dengue

Primera edición: julio de 2025

© Editorial EAFIT

Carrera 49 No. 7 sur – 50. Medellín, Antioquia

<http://www.eafit.edu.co/editorial>

Correo electrónico: obraseditorial@eafit.edu.co

ISBN: 978-958-720-986-0

ISBN: 978-958-720-987-7 (versión EPUB)

ISBN: 978-958-720-988-4 (versión PDF)

DOI: <https://doi.org/10.17230/978-958-720-986-0>

Coordinación editorial: Heiner Mercado Percia

Corrección de textos: Christian Martínez y Heiner Mercado Percia

Diseño y diagramación: Daniel Felipe Loaiza

Imagen de carátula: www.freepik.es

Universidad EAFIT | Vigilada Mineducación. Reconocimiento como Universidad: Decreto Número 759, del 6 de mayo de 1971, de la Presidencia de la República de Colombia. Reconocimiento personería jurídica: Número 75, del 28 de junio de 1960, expedida por la Gobernación de Antioquia. Acreditada institucionalmente por el Ministerio de Educación Nacional hasta el 2026, mediante Resolución 2158 emitida el 13 de febrero de 2018.

Prohibida la reproducción total o parcial, por cualquier medio o con cualquier propósito, sin la autorización escrita de la editorial

Editado en Medellín, Colombia

Capítulo 5

Estimación de parámetros y control

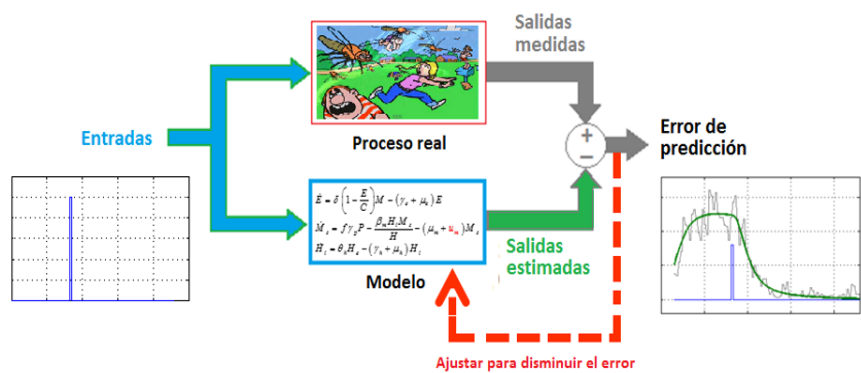
DOI:10.17230/978-958-720-986-0ch5

*Carlos M. Vélez S., Alexandra Cataño-López y
Daniel Rojas Díaz*

Introducción

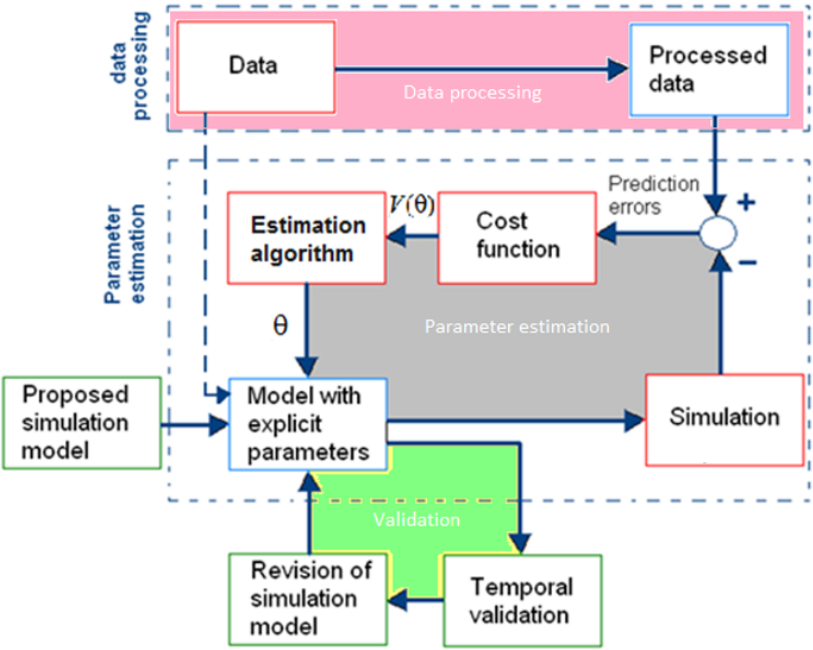
La estimación de parámetros es el proceso de obtención de los valores nominales de parámetros de un modelo matemático a partir de datos experimentales de las entradas y salidas (ver figuras 17 y 18).

Figura 17: Idea general del proceso de estimación de parámetros con un modelo dado y un conjunto de datos experimentales.



Fuente: Elaboración propia.

Figura 18: Procedimiento para la estimación de parámetros considerando todas las etapas: recolección de datos, propuesta del modelo, simulación, selección del método de optimización y criterio de estimación, validación.



Fuente: Elaboración propia.

Adicionalmente, si no se conoce la estructura del modelo matemático, entonces se habla de identificación de sistemas dinámicos, siendo el problema inverso a la simulación matemática. La estimación y simulación en sistemas dinámicos es ampliamente aplicada a diferentes fenómenos físicos, químicos y biológicos. En este último caso, se encuentran aplicaciones a la transmisión de diferentes enfermedades infecciosas como dengue, zika y fiebre amarilla.

En este capítulo se aborda el problema de estimación de parámetros de un modelo matemático de infección del dengue por vectores (mosquitos), buscando obtener unos parámetros del modelo que describan la dinámica de infección de los humanos y que tengan significado biológico. Se utilizan datos del municipio de Bello en el departamento de Antioquia en Colombia, en varios años endémicos y epidémicos. Dado que se cuenta con una sola variable medible (salida) correspondiente al número de humanos infectados en cada instante del tiempo $H_i(t)$, la obtención de diferentes conjuntos de parámetros (mínimos locales) para un valor dado de la función de coste es posible, por lo que se utilizan los rangos biológicos obtenidos en laboratorio para acotar el conjunto solución. Además, no se cuenta con la opción de manipular alguna entrada del sistema de manera que se puedan obtener conjuntos de datos informativos [164], ya que se tiene un registro de datos de varios años donde se han aplicado ciertas acciones de control (químico por fumigación, en este trabajo) generalmente desconocidas o poco precisas. Sin embargo, se pueden incluir en el modelo acciones de control en forma de pulsos que se pueden estimar y que, conocidas, permiten considerar los datos experimentales como un experimento con una señal de entrada establecida y, si se tienen diferentes pulsos en varios años, la variable puede ser entendida como informativa (persistentemente excitada). La variable de fumigación se modela en forma de un pulso con tres valores asociados (tiempo inicial de aplicación, duración o efecto residual e intensidad). Para dar más

sentido a los resultados, el análisis se realiza teniendo en cuenta los rangos de incertidumbre de los parámetros tanto del modelo como de las entradas de fumigación, por cuanto que existen muchas fuentes de incertidumbre (subregistro o sobrerregistro) de pacientes, falsos positivos, falsos negativos, registro de pacientes que no corresponden a la región de estudio, pacientes que no indican su lugar correcto de residencia, aislamientos naturales o artificiales de propagación de la enfermedad en la región, entre otros.

Con base en el problema planteado en el párrafo anterior, en este capítulo se aplican varios métodos que apuntan a los siguientes objetivos: a) dado un modelo conocido en la literatura científica [165], revisar su estructura y realizar ajustes para un mejor modelado de las condiciones conocidas, tales como que la picadura solo la realizan los mosquitos hembra, la mortalidad de los mosquitos infecciosos y expuestos al virus del dengue es mayor que la mortalidad de los mosquitos libres del virus, y la fumigación se aplica en momentos concretos con diferente intensidad y duración; b) estimar los parámetros a partir de sus rangos biológicos (facilitando los cálculos y descartando soluciones no viables al problema de optimización inherente a la estimación matemática) y modelando el cambio de la tasa de mortalidad por medio de una entrada tipo pulso (con una media que corresponde al valor obtenido en el laboratorio y un incremento dado por una acción de fumigación); c) obtener intervalos de confianza de los parámetros, de manera que el análisis tenga sentido a pesar

de las incertidumbres mencionadas; d) con base en el modelo, analizar estrategias de control de la propagación de la enfermedad a partir de la simulación; e) ilustrar los cálculos detalladamente con la aplicación de un entorno de programación basado en Matlab.

Para empezar, es importante resaltar que los parámetros a estimar de los modelos matemáticos deben ser identificables [164], en el sentido de que sus parámetros pueden ser obtenidos única y directamente de los datos experimentales y con una estructura del modelo correcta (ver figura 18). Aunque existen pruebas de identificabilidad complejas, hay criterios que a simple vista descartan ciertos parámetros, como es el caso del modelo [166] en el cual los parámetros k y b están multiplicados por valores que no permiten su estimación directa. En general, la estimación de parámetros depende de las condiciones del sistema real, de la estructura que se seleccione del modelo, del método utilizado para el cálculo de los parámetros y de las condiciones experimentales, en especial del número de datos experimentales y de la variación de la señal de entrada (grado de excitación persistente). Por lo tanto, si se selecciona un modelo de orden pequeño se obtendrán valores de los parámetros muy diferentes de los esperados, por cuanto se debe compensar con valores de los parámetros lo que falta de variables (variables de estado explicativas). Si el modelo, por el contrario, es de un orden muy grande, el ajuste a los datos experimentales es más fácil, pero se pueden obtener muchos conjuntos solución.

Para el modelado matemático de la infección del dengue se han propuesto diversos modelos, cada uno de los cuales tiene en cuenta consideraciones diversas. Una introducción histórica al modelado de enfermedades de transmisión se puede revisar en la literatura existente [167]. Por otro lado, hay modelos más complejos, como el de [168], el cual es base de este estudio con sus 10 variables de estado. Ese modelo considera explícitamente que la población de humanos es constante (suposición aceptable en un lapso de unos pocos meses o años), mientras que la población de mosquitos no lo es. A diferencia de ese modelo, el presentado en este capítulo incluye el parámetro f que representa la proporción de mosquitos hembra, las cuales propagan el virus; dicho parámetro es completamente identificable y su valor (cercano a 0,5) debe ayudar a dar confianza a los otros parámetros estimados. También se introduce el parámetro α , el cual representa el incremento o decremento de la tasa de mortalidad de los mosquitos infectados por el virus del dengue (en laboratorio se determinó que esa mortalidad es mayor y, por lo tanto, $\alpha \geq 1$). La condición inicial de los humanos susceptibles (H_s) igualmente aporta a la confiabilidad del modelo, dado que un valor muy pequeño o que supere la población de Bello es un indicio de que la estimación no es correcta; lo normal es esperar un valor cercano y menor al censo poblacional del municipio, dado que no toda la población es susceptible (puede haber grupos muy alejados de los focos de infección).

Existen muchos tipos de modelos que consideran diferentes restricciones y suposiciones de desarrollo y propagación de la enfermedad. Un análisis de los artículos sobre enfermedades de transmisión por mosquitos de 1970 a 2010, en [127] se hace referencia a 325 artículos con modelos, 53 % de ellos entre 2005 y 2010, de los cuales 77 se relacionan con el dengue. Esta información es fundamental para entender las diferentes formas de abordar el problema de modelado de la transmisión de enfermedades con mosquitos. Otro texto como [128] complementa ese diagnóstico y se concentra solo en el dengue. En temas de control [128] se identifica en la literatura científica tres acciones principales (fumigación de mosquitos adultos, larvicidas y control mecánico por remoción de contenedores de crías), además de las ventajas de la aplicación periódica de insecticidas en determinados meses. También se afirma que es escasa la investigación en el control teniendo en cuenta los cuatro serotipos del dengue (el modelado de los cuatro serotipos es un tema complejo). En este capítulo se considera solo un comportamiento general de la enfermedad y se aplica un análisis de incertidumbre que permite dar más confianza a las conclusiones, pero es un tema de interés analizar el valor de los parámetros estimados del modelo en función de diferentes escenarios de propagación de la enfermedad, y con un análisis de incertidumbre, de manera que pueda dar indirectamente indicios de ciertas características específicas del entorno (por ejemplo, mayor control mecánico o prevalencia de un serotipo, entre otras). En estos modelos es

necesario incluir todas o algunas de las fases acuáticas (huevos, larvas o pupas) si se desea estimar el efecto del control mecánico (limpieza de criaderos de mosquitos, por ejemplo) en la propagación de la enfermedad. Finalmente, conviene resaltar la importancia de la matemática en ciencias de la vida, pero sin perder de vista sus limitaciones, al menos, en dos casos [169]: a) aplicación de métodos y modelos matemáticos en circunstancias que no cumplen con los requisitos que los soportan y b) obtención de modelos con mucho detalle cuando hay factores poco conocidos.

Con base en lo anterior, este capítulo está dividido en varias secciones que apuntan a los objetivos mencionados. Inicialmente, se presentan los métodos conocidos y nuevos utilizados para obtener los resultados. Los métodos se describen brevemente, se referencian para profundizar en ellos y se explica con más detalles su implementación con Matlab. Luego, en los resultados y la discusión se muestran gráficas y tablas relevantes, y se describen los aspectos más sobresalientes. Por último, se exponen las conclusiones del estudio.

Métodos

En la presente sección se abordará un conjunto de métodos implementados en el análisis de un modelo de transmisión del dengue, el cual es aplicado a una zona de estudio altamente afectada. En primer lugar, se hace referencia a los aspectos experimentales relacionados con la zona, por ejemplo, cómo se obtienen

los datos reales utilizados (casos de dengue por semana) y estrategias de control aplicadas en la zona de estudio. Luego, se describe el modelo matemático junto con sus supuestos y parámetros, incluyendo una entrada tipo pulso que simula el control por fumigación. Posteriormente, se presentan los métodos de estimación de parámetros y de sus correspondientes intervalos de confianza, incluyendo los parámetros de la entrada de tipo pulso. Al final, se describen los métodos de análisis de incertidumbre y sensibilidad que se utilizan para evaluar los intervalos de confianza obtenidos y encontrar los parámetros más significativos dentro del modelo, respectivamente.

Aspectos experimentales: zona de estudio y casos de dengue

De acuerdo con el Área Metropolitana del Valle de Aburrá, Bello (Antioquia) es un municipio localizado a 1.450 msnm y su precipitación media anual es de 153 mm con una temperatura de 21,7°C. Estas condiciones climáticas son las apropiadas para la presencia y proliferación del *Aedes aegypti* y, por consiguiente, de los casos endémicos o epidémicos de dengue [170, 171]. Los datos a estudiar son el número de personas infectadas durante 93 semanas epidemiológicas en Bello, comenzando desde la semana epidemiológica 29 del 2009 y finalizando en la 37 del 2011, donde el brote epidémico ocurrió en los años 2009-2010.

Independientemente de la presencia o ausencia de brotes epidémicos, los casos de dengue en Colombia son recolectados

por los hospitales y reportados semanalmente por el Sistema de Vigilancia en Salud Pública (SIVIGILA). Los protocolos para el reporte de casos son preestablecidos con el fin de realizar monitoreos sistemáticos en cada municipio y departamento; esto es aplicado a los casos de dengue clásico y grave que se presentan. Finalmente, dicha información es utilizada para identificar factores de riesgo y realizar acciones de prevención o control [172].

Las acciones de control comúnmente realizadas por las secretarías de salud en Colombia son las jornadas de limpieza de criaderos y las fumigaciones. Para el caso de Bello, y debido al alto número de casos reportados en 2010, la Secretaría de Salud de Bello realizó una fumigación durante la semana epidemiológica 29 de dicho año. El insecticida utilizado en esta intervención fue el malatión. Este es un compuesto organofosforado que tiene un efecto residual de máximo 12 semanas sobre una población de mosquitos adultos [173]. Dicha información sobre el control es de importancia para la separación de efectos exógenos en la estimación de parámetros del modelo, como será explicada en las siguientes secciones.

Modelo matemático y simulación de la infección del dengue

A continuación, se presenta el modelo matemático de propagación del dengue (22), basado en el modelo de [165], con algunos ajustes importantes y con una explicación detallada de los supuestos de aplicación, las variables y los parámetros.

$$\left\{ \begin{array}{l} \dot{E} = \delta \left(1 - \frac{E}{C}\right) M - (\gamma_e + \mu_e) E \\ \dot{L} = \gamma_e E - (\gamma_l + \mu_l + u_l) L \\ \dot{P} = \gamma_l L - (\gamma_p + \mu_p + u_p) P \\ \dot{M}_s = f \gamma_p P - \frac{\beta_m H_i M_s}{H} - (\mu_m + u_m) M_s \\ \dot{M}_e = \frac{\beta_m H_i M_s}{H} - (\theta_m + \alpha \mu_m + u_m) M_e \\ \dot{M}_i = \theta_m M_e - (\alpha \mu_m + u_m) M_i \\ \dot{H}_s = \mu_h H - \frac{\beta_h M_i H_s}{M} - \mu_h H_s - u_v H_s \\ \dot{H}_e = \frac{\beta_h M_i H_s}{M} - (\theta_h + \mu_h) H_e \\ \dot{H}_i = \theta_h H_e - (\gamma_h + \mu_h) H_i \\ \dot{H}_r = \gamma_h H_i - \mu_h H_r + u_v H_s \end{array} \right. \quad (22)$$

En la tabla 16 se definen las variables y parámetros del modelo.

Tabla 16: Definición de variables y parámetros del modelo matemático de propagación del dengue.

Vble	Descripción	Parámetro	Descripción
E	Número de huevos	δ	Tasa de oviposición por mosquito
L	Número de larvas	C	Capacidad total de huevos
P	Número de pupas	γ_e	Tasa de transformación de huevo a larva (indicativo del efecto del control mecánico)
M_s	Número de mosquitos susceptibles de transmitir el virus	μ_e	Tasa de mortalidad de huevos
M_e	Número de mosquitos expuestos (infectados pero no infecciosos)	γ_l	Tasa de transformación de larva a pupa
M_i	Número de mosquitos infecciosos	μ_l	Tasa de mortalidad de larvas
H_s	Número de humanos susceptibles de contraer el virus	γ_p	Tasa de eclosión de pupa a mosquito
H_e	Número de humanos expuestos (infectados pero no infecciosos)	μ_p	Tasa de mortalidad de pupas
H_i	Número de humanos infectados e infecciosos	f	Fracción de hembras que emergen de todos los huevos
H_r	Número de humanos recuperados	β_m	Tasa de contacto efectiva (o coeficiente de transmisión) humano-mosquito
		μ_m	Tasa de mortalidad de mosquitos

Vble	Descripción	Parámetro	Descripción
u_l	Variable (tasa) de control de larvas con larvicidas	α	Proporción de cambio de la tasa de mortalidad de los mosquitos debido a la infección con el virus
u_p	Variable (tasa) de control de pupas con larvicidas	θ_m	Tasa de incubación extrínseca (en el tiempo θ_m^{-1} se vuelve infeccioso)
u_m	Variable (tasa) de control de mosquitos con insecticidas	μ_h	Tasa de mortalidad de humanos
u_v	Variable de control de vacunación (tasa de vacunados por porcentaje de eficacia)	β_h	Tasa de contacto efectiva (o coeficiente de transmisión) mosquito-humano
H	Número total de humanos ($H = H_s + H_e + H_i + H_r$)	θ_h	Tasa de incubación intrínseca (en el tiempo θ_h^{-1} se evidencian los síntomas de la enfermedad)
M	Número total de mosquitos ($M = M_s + M_e + M_i$)	γ_h	Tasa de recuperación en humanos
		A_{ci}	Amplitud del pulso i de control por fumigación
		t_{0ci}	Tiempo inicial del pulso i de control por fumigación
		Δt_{ci}	Ancho del pulso i de control por fumigación

Nota: Definición de 10 variables de estado, 4 variables de entrada, 17 parámetros del sistema y 3 parámetros de la entrada de fumigación, correspondientes al modelo matemático de propagación del dengue en consideración.

Es importante aclarar que cada tasa en el modelo indica la relación entre la variable de salida (variable derivada) y la variable de entrada (variable de cada término) por unidad de tiempo. Por ejemplo, θ_h en el término $\dot{H}_i = \theta_h H_s$ representa la proporción de personas que pasan de un estado expuesto a un estado infectado durante la unidad de tiempo considerada (semanas en este artículo). El parámetro $\lambda_m = \beta_m H_i / H$ en el término $\dot{M}_e = \lambda_m M_s$ es la tasa de transición de mosquitos susceptibles a expuestos (tasa de contacto efectiva per cápita de todos los humanos infecciosos con mosquitos), la cual depende del parámetro β_m (tasa de contacto efectiva o coeficiente de transmisión humano-mosquito)

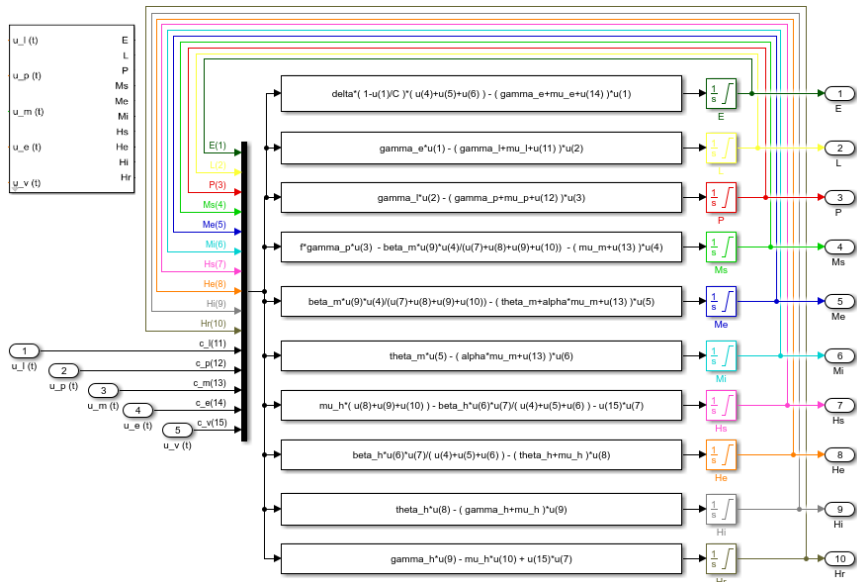
que determina el número de mosquitos por unidad de tiempo que un humano infectado está en capacidad de infectar (depende del número total de picaduras, que en este caso está entre 1 y 4 picaduras, y de la probabilidad de infección), y de H_i/H , la proporción de humanos infectados.

El modelo se basa en las siguientes suposiciones: a) la población total de humanos H es constante ($\dot{H} = \dot{H}_s + \dot{H}_e + \dot{H}_i + \dot{H}_r = 0$), es decir, la tasa de nacimientos y muertes de humanos es igual a μ_h , lo cual es admisible para intervalos cortos de estudio; b) la población total de mosquitos $M = M_s + M_e + M_i$ es variable; c) no se hace una separación entre los distintos serotipos del dengue, es decir, se asume que el modelo representa un promedio del comportamiento de propagación; d) hay una mezcla homogénea o uniforme de la población de infecciosos y susceptibles, por lo que todas las tasas son las mismas en cualquier punto geográfico (esto es equivalente a decir que los parámetros son el promedio de toda la población y el modelo representa el promedio del comportamiento en distintos puntos geográficos); e) no hay transmisión vertical del virus; f) el número de humanos infectados corresponde a los nuevos casos: no se lleva un registro del estado de los anteriores humanos infectados al final de cada semana, por lo que el valor de H_i puede ser menor o mayor a los casos registrados y se decidió tomar los nuevos casos como el valor más probable e incluir los posibles errores en el análisis de incertidumbre (el valor puede ser mayor si hay enfermos remanentes de la semana anterior, pero puede ser menor si un

enfermo reportado al inicio de la semana –que pudo contagiarse la semana anterior– logró recuperarse totalmente).

El modelo (22) se diferencia del modelo de [165] en lo siguiente: a) se incluye el parámetro f correspondiente a la fracción de hembras que emergen de todos los huevos, el cual se incluye en una ecuación que lo convierte en un parámetro identificable y que sirve de parámetro de control, debido a que su valor debe ser cercano a 0,5 [174]; b) se introduce el parámetro α que modela un aumento de la tasa de mortalidad de los mosquitos infectados con el virus y cuyo valor debe ser mayor que 1 si está bien estimado [175]; c) se modela una acción de control por medio de un pulso positivo (fumigación), pero se podrían incluir más pulsos para más acciones de control y pulsos negativos (condiciones favorables para el crecimiento de la población de mosquitos o la disminución de su tasa de mortalidad); d) se trabaja con variables sin escalar con el fin de observar la magnitud de todas las variables involucradas; e) se adiciona la acción de control de vacunación (u_v) para efectos de simulación, también como una entrada tipo pulso, con lo cual se modela la tasa de humanos susceptibles que pasan directamente a la fase de humanos recuperados sin pasar por la fase de exposición o infección en un intervalo determinado de tiempo (se podría incluir una nueva fase de humanos vacunados, pero se considera aquí innecesario). La implementación del modelo en el entorno de Matlab/Simulink se muestra en la figura 19.

Figura 19: Diagrama de simulación del modelo (22), implementado en Matlab/Simulink.



Fuente: Elaboración propia

Estimación de parámetros y cálculo de los intervalos de confianza

La estimación de parámetros de un modelo matemático es el proceso de obtención de valores óptimos generando familias factibles de valores de parámetros y simulando con ellos el modelo, de manera que los resultados producidos se ajusten a los datos experimentales (ver figura 18). Para el presente estudio se utilizan casos reales de humanos infectados con el virus del dengue (H_i) para ajustar la salida del modelo, aplicando los siguientes métodos y herramientas sobre el modelo de dengue: a) el modelo matemático es de tiempo continuo y no lineal, y se implemen-

ta en Matlab/Simulink (ver figura 19); b) se resolvieron las ecuaciones de estado utilizando el método de Dormand-Prince de paso variable implementado en la función `ode45`; c) las ecuaciones se resolvieron utilizando matemática simbólica y el método numérico de paso variable Dormand-Prince implementado en la función `ode45` de Matlab; d) se realizó la estimación de parámetros utilizando el algoritmo `trust-region-reflective` y el criterio de mínimos cuadrados no lineales (Mean Squared Error, MSE, por sus siglas en inglés) con un parámetro y valores de tolerancia funcional iguales a 10^{-7} (se utilizó la herramienta *Simulink Design Optimization* de Matlab); e) se utilizó el *Uncertainty Confidence Interval (GSUA-UCI) Toolbox* de Matlab como herramienta de estimación y obtención de los intervalos de confianza; f) todo se integró en funciones y *scripts* de Matlab para facilitar los cálculos en paralelo en un entorno multiprocesador; g) las características del equipo utilizado para dichas tareas son las siguientes: Linux 4.15.0-58-generic con arquitectura de 64 bits y 8 procesadores Intel(R) Core(TM) i7-4700HQ CPU@ 2.40GHz.

Para la estimación de parámetros del modelo de dengue en Bello, se realizó una serie de estimaciones para un caso: un brote epidémico con entrada durante 93 semanas epidemiológicas. Ya se han hecho pruebas con más entradas tipo pulso (positivos y negativos) para simular múltiples brotes epidémicos consecutivos, pero se mostrarán en otras publicaciones futuras. De esta manera, se estimaron 15 parámetros (2 se dejaron fijos), 8 condiciones iniciales (2 se dejaron fijas) y se estimaron tres paráme-

tros del pulso de control, relacionado con la tasa de mortalidad de mosquitos. Las condiciones iniciales fijas fueron $H_r(0) = 0$, $H_i(0) = 8$, dado que no se consideran humanos recuperados al principio y se sabe que hubo 8 humanos infectados al inicio de la epidemia. Se adicionan restricciones en los límites inferiores y superiores de cada parámetro con base en estudios de laboratorio conocidos y ampliando los rangos en aquellos parámetros cuyo valor podría cambiar en condiciones silvestres reales. Además, se realizaron múltiples estimaciones debido a que algunos parámetros presentaban una tendencia a anclarse en el límite de los rangos de estimación, bloqueando el proceso posterior de estimación, fijando dichos parámetros en un extremo del intervalo y afectando la estimación de los demás parámetros. Dada esta situación, se expandieron los intervalos y se usaron los parámetros nominales obtenidos como puntos iniciales para verificar que los valores fueran iguales o cercanos a los nominales.

El principal problema al estimar los parámetros del modelo es la disponibilidad y confiabilidad de los datos, debido a que solo se tiene un conjunto de datos de salida para ajustar los parámetros, por eso es probable que diferentes conjuntos de parámetros den un ajuste igual de bueno y en estos casos el acotamiento del conjunto solución a partir de información biológica es una buena estrategia. Además, se puede reducir la probabilidad de obtención de varios conjuntos de solución si se realizan diferentes estimaciones y se seleccionan solo aquellas que no se diferencien más de determinado porcentaje del valor de

la función de coste, con el fin de calcular los valores nominales de los parámetros utilizando la mediana estadística, dado que la mediana es una medida robusta de la centralidad de los datos. Calculada la mediana, es posible obtener los intervalos de confianza de los parámetros utilizando la desviación estándar σ_{θ_i} , N estimaciones y la siguiente expresión:

$$\theta_i = \text{mediana}(\theta_i) \pm 1.96 \sqrt{\frac{\pi}{2}} \frac{\sigma_{\theta_i}}{\sqrt{N}} \quad (23)$$

Sin embargo, la expresión anterior puede entregar (y en el presente trabajo entrega) soluciones no factibles al realizar un análisis de incertidumbre, es decir, al simular por el método de Montecarlo modelos con diferentes valores de parámetros incluidos en los intervalos de confianza estimados. Para reducir los conjuntos no factibles de solución se utiliza un método ya descrito [176], cuya idea básica es la siguiente: a) se realizan múltiples estimaciones de parámetros hasta dar con una curva nominal y se estiman los intervalos de confianza con la expresión (23; b) se construyen unos intervalos iniciales promisorios aplicando una simulación de Montecarlo variando un solo parámetro a la vez (método de uno a la vez u OAT, (One-At-a-Time); c) en un proceso iterativo se aplica el método de Montecarlo variando todos los parámetros a la vez, descartando las soluciones no factibles y acotando los intervalos de los parámetros.

Dicho método pretende definir una caja (dada por unos intervalos de confianza) dentro del contorno de confianza que contiene a los parámetros nominales. El contorno de confianza es el

lugar de los valores de los parámetros cuyo ajuste con respecto a los datos reales es bueno (en el sentido de mínimos cuadrados). Así, los intervalos de confianza calculados por este método reciben el nombre de intervalos de la caja del subcontorno de confianza (CSB, por sus siglas en inglés). El método numérico expuesto para calcular los intervalos CSB es muy efectivo y garantiza que el 95 % o más de las soluciones simuladas estén dentro de un rango preestablecido alrededor de una salida nominal [176]. El método se implementó en Matlab en una herramienta de fácil uso y que se puede descargar libremente [177].

Análisis de incertidumbre y sensibilidad

El análisis de incertidumbre determina cómo una incertidumbre en las entradas o parámetros de un modelo se refleja en una incertidumbre en las salidas. El análisis de sensibilidad clasifica los parámetros de los más a los menos sensibles (importantes). Se dice que una solución (decisión o análisis) es sensible si ante un cambio pequeño en el modelo o el entorno la solución cambia radicalmente y pierde, por lo tanto, su sentido; una solución poco sensible se dice que es robusta. El análisis de sensibilidad e incertidumbre se aplica en este trabajo para a) comprobar la robustez de las predicciones del modelo y dar mayor credibilidad a la solución, dado que hay muchas incertidumbres en los datos (humanos infectados): subregistro de pacientes (pacientes que se infectaron y sus síntomas no ameritaron una consulta médica), falsos positivos (casos que no correspondían y fueron re-

gistrados como casos de dengue) o falsos negativos, casos que proceden de otras regiones, pacientes que consultaron en servicios médicos por fuera de la ciudad de estudio (Bellos), otros casos; b) identificar los parámetros más sensibles (importantes) y cuyos valores son más confiables para concluir algo; c) estimar los intervalos de confianza de los parámetros al descartar soluciones no muy cercanas a la solución nominal (ver sección anterior).

El análisis de sensibilidad e incertidumbre puede ser local o global [178]. En el primer caso, el análisis se realiza alrededor de un punto de operación en el espacio de variables de estado y parámetros, es un análisis lineal a partir de un modelo nominal, no considera la interacción entre parámetros y no es aplicable a sistemas altamente no lineales; el método más utilizado en este caso es el método de Morris OAT. De otro lado, el análisis global es un análisis no lineal, considera todos los parámetros a la vez y la interacción entre ellos, y no utiliza un modelo nominal. En este trabajo se realiza un análisis global de sensibilidad e incertidumbre.

Existen diversos métodos globales de sensibilidad e incertidumbre, entre los cuales se encuentran los siguientes:

- Métodos estadísticos de sensibilidad (costosos computacionalmente, la sensibilidad puede ser estimada, pero no calculada exactamente): basados en la regresión lineal (método lineal, usa los coeficientes de regresión como medidas de la sensibilidad, el coeficiente de determinación permi-

te confirmar la linealidad del modelo), métodos basados en la varianza, método de la matriz hessiana, método de descomposición de la varianza, método de la superficie de respuesta, expansión del caos polinomial, método de Karhunen–Loève, método bayesiano.

- Métodos basados en el muestreo (simulación): método de Montecarlo (uso de distribuciones de probabilidad para la generación aleatoria de gran cantidad de combinaciones de valores de ciertas variables de entrada o parámetros, con los cuales se simula el modelo y se realiza un análisis probabilístico posterior de las variables de salida para tener una idea del proceso físico), método de efectos elementales, muestreo adaptativo, muestreo de importancia, muestreo del hipercubo latino.
- Métodos deterministas (eficiencia computacional, desventaja por el cálculo de las derivadas): métodos de expansión local (series de Taylor, método de la perturbación), análisis de intervalos, teoría borrosa, método de la función de Green, derivadas de Frechet, *Adjoint Sensitivity Analysis Procedure* (ASAP), funciones adjuntas, Forward Sensitivity Analysis Procedure (FSAP).

En este capítulo, el análisis de sensibilidad e incertidumbre se enfoca en los métodos globales basados en la varianza con utilización del método de Montecarlo. Entre dichos métodos están los siguientes: fuerza bruta, método de Sobol, método de Jansen,

método de Saltelli, Fourier Amplitude Sensitivity Test (FAST), emuladores, HDMR, entre otros. Estos métodos tienen en común las siguientes características: el análisis aplica a modelos no lineales, se captura la influencia de todo el rango de variación de un parámetro, permiten el análisis de la interacción entre parámetros y el costo computacional es alto. El procedimiento general a seguir en estos métodos se enumera a continuación.

1. Plantear el modelo matemático a analizar
2. Determinar los intervalos de incertidumbre de cada parámetro
3. Simular el modelo varias veces para una muestra significativa de combinaciones de todos los valores de parámetros y entradas, de acuerdo con cierta distribución (simulación Montecarlo)
4. Graficar las soluciones anteriores para cualificar el efecto de la incertidumbre de los parámetros y entradas en las salidas (análisis de incertidumbre)
5. Calcular los índices de sensibilidad a partir de las soluciones anteriores para cuantificar el efecto de cada parámetro en la incertidumbre de una salida definida (análisis de sensibilidad global), la cual en este trabajo es la función de coste y corresponde a la función escalar del error cuadrático medio, es decir, la sumatoria de la diferencia en cada instante del tiempo y elevada al cuadrado entre la curva experimental y la curva simulada

6. Graficar los índices de sensibilidad por diferentes métodos

Los métodos basados en la varianza parten de la siguiente descomposición ANOVA-HDMR [179] para parámetros independientes, donde la varianza observada en una variable particular (función de coste en el presente trabajo) se divide en componentes atribuibles a diferentes fuentes de variación (variación individual de cada uno de los N_p parámetros y de las 2^{N_p-1} combinaciones de 2, 3 y más varianzas):

$$V(Y) = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < k} V_{ijk} + \dots + V_{12\dots N_p} \quad (24)$$

Donde,

$$V_i = V_{X_i} [E_{\mathbf{X}_{\sim i}}(Y|X_i)], V_{ij} = V_{X_{i,j}} [E_{X_{\sim i,j}}(Y|X_i, X_j)] - V_i - V_j, \dots \quad (25)$$

Desde el punto de vista computacional, en las expresiones anteriores se generan N familias o combinaciones de parámetros por simulación Montecarlo, cada uno dentro de sus intervalos de confianza o interés, formando una matriz de N filas y N_p columnas:

$$\mathbf{M} = \begin{bmatrix} x_1^{(1)} & \dots & x_i^{(1)} & \dots & x_{N_p}^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(N)} & \dots & x_i^{(N)} & \dots & x_{N_p}^{(N)} \end{bmatrix} \quad (26)$$

Por lo tanto, $V_{X_i} [E_{\mathbf{X}_{\sim i}}(Y|X_i)]$ indica la varianza de la esperanza matemática de la salida (función de coste) haciendo $X_i = \text{const}$ y variando todos los demás parámetros excepto el parámetro i ($X_{\sim i}$), por lo que se obtiene la siguiente matriz para el parámetro

i (se generan muestras aleatorias de todos los parámetros y se forman bloques, dejando fijo el factor X_i en cada columna del bloque):

$$\mathbf{M}_i = \left[\begin{array}{ccccc} x_1^{(1)} & \cdots & x_i^{(1)} & \cdots & x_{N_p}^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(N)} & \cdots & x_i^{(1)} & \cdots & x_{N_p}^{(N)} \\ \hline x_1^{(1)} & \cdots & x_i^{(2)} & \cdots & x_{N_p}^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(N)} & \cdots & x_i^{(2)} & \cdots & x_{N_p}^{(N)} \\ \hline \vdots & & \vdots & & \vdots \end{array} \right] \quad (27)$$

De otro lado, $V_{X_{\sim i}} [E_{\mathbf{X}_i}(Y|X_{\sim i})]$ indica la varianza de la esperanza matemática de la salida (función de coste) haciendo todos los $X_i = \text{const}$ excepto X_i ($X_{\sim i}$), por lo cual se obtiene la siguiente matriz para el parámetro i (se generan muestras aleatorias de todos los parámetros y se forman bloques, dejando fijos todos los factores, menos X_i en cada columna):

$$\mathbf{M}_{T_i} = \left[\begin{array}{ccccc} x_1^{(1)} & \cdots & x_i^{(1)} & \cdots & x_{N_p}^{(1)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(1)} & \cdots & x_i^{(2)} & \cdots & x_{N_p}^{(1)} \\ \hline x_1^{(2)} & \cdots & x_i^{(1)} & \cdots & x_{N_p}^{(2)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(2)} & \cdots & x_i^{(2)} & \cdots & x_{N_p}^{(2)} \\ \hline \vdots & & \vdots & & \vdots \end{array} \right] \quad (28)$$

Las expresiones anteriores están relacionadas de la siguiente manera:

$$V(Y) = V_{X_i} [E_{\mathbf{X}_{\sim i}}(Y|X_i)] + E_{X_i} [V_{\mathbf{X}_{\sim i}}(Y|X_i)] \quad (29)$$

$$= V_{\mathbf{X}_{\sim i}} [E_{X_i}(Y|\mathbf{X}_{\sim i})] + E_{\mathbf{X}_{\sim i}} [V_{X_i}(Y|\mathbf{X}_{\sim i})] \quad (30)$$

Teniendo en cuenta las expresiones anteriores, y dividiendo por la varianza total, se pueden obtener dos tipos de índices de sensibilidad:

- El índice S_i es el índice de sensibilidad fraccional de primer orden del factor X_i , el cual indica en cuanto se reduce en promedio la varianza de la salida si se fija X_i :

$$S_i = \frac{V_{X_i}[E_{\mathbf{X}_{\sim i}}(Y|X_i)]}{V} = \frac{V_i}{V} \dots S_{ijk\dots} = \frac{V_{ijk\dots}}{V} \quad (31)$$

$$\sum_i S_i + \sum_{i < j} S_{ij} + \sum_{i < j < k} S_{ijk} + \dots + S_{12\dots N_p} = 1, \quad \sum_i S_i < 1$$

- El índice S_{T_i} es el índice de sensibilidad total (efecto total) del factor X_i , el cual indica en cuanto se reduce en promedio la varianza de la salida si se fijan todos los factores menos X_i :

$$S_{T_i} = \frac{E_{\mathbf{X}_{\sim i}}[V_{X_i}(Y|\mathbf{X}_{\sim i})]}{V} = 1 - \frac{V_{\mathbf{X}_{\sim i}}[E_{X_i}(Y|\mathbf{X}_{\sim i})]}{V}$$

$$S_{T_i} = S_i + \sum_j S_{ij} + \sum_{j,k} S_{ijk} + \dots, \quad \sum_i S_{T_i} \geq 1 \quad (32)$$

$$S_{T_i} > S_i$$

A partir de las definiciones anteriores se pueden calcular los índices utilizando diferentes métodos, todos ellos computacionalmente exigentes. En primer lugar, se puede recurrir al método de fuerza bruta (debe su nombre al hecho que recurre solamente a la capacidad computacional sin ningún tipo de consideración matemática que mejore la eficiencia de cómputo). A partir de la matriz (26) se construyen las matrices M_i (27) y M_{T_i} (28). Para cada muestra de parámetros (fila de cada matriz) se simula el sistema y se calcula la salida Y_i (función de coste) que conforma el vector-columna Y . Con los resultados anteriores se obtienen V_i y V_{T_i} , y de ahí los índices de sensibilidad (31) y (32), respectivamente. Se requieren $(N^2 \times N_p)$ simulaciones, lo cual implica un costo computacional considerable. Por lo tanto, y recurriendo a propiedades matemáticas, es posible simplificar las operaciones del método de fuerza bruta, lo cual lleva a un conjunto de métodos cuyo principal exponente es el método de Sobol, y que ha llevado a mejoras que se reflejan en otros métodos equivalentes como el método de Jansen y el de Saltelli. El número de operaciones es mucho menor e igual a $N \times (N_p + 2)$, lo cual permite ir aumentando N para analizar la convergencia. En estos métodos se definen las siguientes matrices, obtenidas a partir de la matriz (26), la cual se divide en dos matrices A y B :

$$\mathbf{A} = \begin{bmatrix} x_1^{(1)} & \cdots & x_i^{(1)} & \cdots & x_{N_p}^{(1)} \\ x_1^{(2)} & \cdots & x_i^{(2)} & \cdots & x_{N_p}^{(2)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(N)} & \cdots & x_i^{(N)} & \cdots & x_{N_p}^{(N)} \end{bmatrix} \quad (33)$$

$$\mathbf{B} = \begin{bmatrix} x_1^{(N+1)} & \dots & x_i^{(N+1)} & \dots & x_{N_p}^{(N+1)} \\ x_1^{(N+2)} & \dots & x_i^{(N+2)} & \dots & x_{N_p}^{(N+2)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(2N)} & \dots & x_i^{(2N)} & \dots & x_{N_p}^{(2N)} \end{bmatrix} \quad (34)$$

Se obtienen las siguientes matrices a partir de las anteriores, intercambiando solo la columna i :

$$\mathbf{A}_{B_i} = \begin{bmatrix} x_1^{(1)} & \dots & x_i^{(N+1)} & \dots & x_{N_p}^{(1)} \\ x_1^{(2)} & \dots & x_i^{(N+2)} & \dots & x_{N_p}^{(2)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(N)} & \dots & x_i^{(2N)} & \dots & x_{N_p}^{(N)} \end{bmatrix} \quad (35)$$

$$\mathbf{B}_{A_i} = \begin{bmatrix} x_1^{(N+1)} & \dots & x_i^{(1)} & \dots & x_{N_p}^{(N+1)} \\ x_1^{(N+2)} & \dots & x_i^{(2)} & \dots & x_{N_p}^{(N+2)} \\ \vdots & & \vdots & & \vdots \\ x_1^{(2N)} & \dots & x_i^{(N)} & \dots & x_{N_p}^{(2N)} \end{bmatrix} \quad (36)$$

Se simula el sistema y se calcula la salida Y (función de coste) con cada conjunto de parámetros (fila) de cada una de las matrices anteriores (por ejemplo, $Y_{A_{B_i}}$ es el vector-columna con los valores de la función de coste a partir de los parámetros dados en la matriz A_{B_i}):

$$\mathbf{Y}_A, \mathbf{Y}_B, \mathbf{Y}_{A_{B_i}}, \mathbf{Y}_{B_{A_i}} \quad (37)$$

El método de Sobol establece los siguientes estimadores para la varianza total:

$$V(Y) \simeq \frac{1}{N} \sum_{k=1}^N [\mathbf{Y}(k)]^2 - f_0^2 \quad f_0^2 = \left[\frac{1}{N} \sum_{k=1}^N \mathbf{Y}(k) \right]^2 \quad (38)$$

Finalmente, se obtienen las siguientes expresiones para el cálculo de los dos tipos de índices de sensibilidad para cada método:

- Método de Sobol (1993 y 2007):

$$\begin{aligned} S_i &\simeq \frac{\frac{1}{N} \sum_{k=1}^N \mathbf{Y}_A(k) \mathbf{Y}_{B_{A_i}}(k) - f_0^2}{V(Y)} \\ S_{T_i} &\simeq \frac{\frac{1}{N} \sum_{k=1}^N \mathbf{Y}_A(k) \left[\mathbf{Y}_A(k) - \mathbf{Y}_{A_{B_i}}(k) \right]}{V(Y)} \end{aligned} \quad (39)$$

- Método de Jansen (1999):

$$\begin{aligned} S_i &\simeq 1 - \frac{\frac{1}{2N} \sum_{k=1}^N \left[\mathbf{Y}_B(k) - \mathbf{Y}_{A_{B_i}}(k) \right]^2}{V(Y)} \\ S_{T_i} &\simeq \frac{\frac{1}{2N} \sum_{k=1}^N \left[\mathbf{Y}_A(k) - \mathbf{Y}_{A_{B_i}}(k) \right]^2}{V(Y)} \end{aligned} \quad (40)$$

- Método de Saltelli (2010):

$$\begin{aligned} S_i &\simeq \frac{\frac{1}{N} \sum_{k=1}^N \mathbf{Y}_B(k) \left[\mathbf{Y}_{A_{B_i}}(k) - \mathbf{Y}_A(k) \right]}{V(Y)} \\ S_{T_i} &\simeq \frac{\frac{1}{2N} \sum_{k=1}^N \left[\mathbf{Y}_A(k) - \mathbf{Y}_{A_{B_i}}(k) \right]^2}{V(Y)} \end{aligned} \quad (41)$$

Los valores de los índices de sensibilidad permiten hacer diferentes tipos de inferencias:

- La medida del impacto de las interacciones la da la siguiente diferencia: $1 - \sum_i S_i$

- Las interacciones son consideradas pequeñas si: $1 - \sum_i S_i < 0.35$. Si ese valor es significativo, hay interacciones importantes del parámetro
- Para un modelo aditivo puro $1 - \sum_i S_i = 1$ y $S_{T_i} = S_i$
- Si $S_{T_i} = 0$ el factor X_i no influye en nada
- Si $S_{T_i} = 0$ el factor $X_i \simeq 0$ puede fijarse en un valor

Respecto al análisis de incertidumbre, se realizan simulaciones con los N conjuntos o familias de parámetros, con lo cual se puede evaluar visualmente, y con algunas características cuantitativas, el efecto de la incertidumbre de los parámetros en la cercanía de los modelos a la salida experimental (humanos infectados H_i en el caso del dengue).

El análisis de sensibilidad e incertidumbre depende del número de muestras con el que se explora el espacio de los parámetros: un mayor número de muestras permite un mayor entendimiento del comportamiento del modelo y de la influencia de los parámetros dentro de sus respectivos intervalos. Para determinar un número de muestras apropiado, es recomendable proceder con un análisis de convergencia de los índices utilizando métodos que garanticen una rápida convergencia (diseño del hipercubo latino, por ejemplo); sin embargo, el costo computacional de dichos análisis puede ser muy alto. Para el caso de los modelos epidemiológicos (y muchos otros) es común que algunos índices de primer orden (S_i) tiendan a 0, por lo cual es

posible obtener valores negativos con estimadores como el proporcionado por el método de Saltelli, lo cual, por definición, no tiene sentido. De esta manera, se puede plantear un indicador de confianza basado en la sumatoria de S_i respecto a la sumatoria de valores absolutos de S_i : cuando la diferencia entre ambos tienda a 0, entonces consideraremos que el tamaño de la muestra fue apropiado.

A continuación, se muestran diferentes tipos de gráficas para representar tanto el análisis de sensibilidad como el análisis de incertidumbre, dando además detalles de los cálculos utilizando Matlab.

Resultados

En este punto, se presentan los resultados y la discusión de la estimación de parámetros e intervalos de confianza para el modelo (22), junto con su correspondiente validación (análisis de sensibilidad e incertidumbre). Posteriormente, se muestran algunos escenarios de control químico y vacunación, en los cuales se utilizan los valores nominales obtenidos de la estimación de parámetros. Además, se explica paso a paso la implementación de cada uno de los métodos descritos en la sección anterior utilizando las funciones de la herramienta GSUA-UCI de Matlab [177], la cual conecta con una guía de usuario que explica con detalle las distintas funciones utilizadas.

Estimación de parámetros

Inicialmente, se define el modelo (22) con la herramienta de matemáticas simbólicas que ofrece Matlab. Como se muestra en el código 1, se declaran los estados, seguidos de los parámetros del modelo y, luego, la entrada de tipo pulso junto a las ecuaciones diferenciales del modelo.

Código 1: Definición del modelo (22) en matemática simbólica.

```
1 %Declaracion de estados del sistema
2 syms E(t) L(t) P(t) Ms(t) Me(t) Mi(t) Hs(t) He(t) Hi(t) Hr(t)
3 %Declaracion de parametros del sistema
4 syms alpha delta C gamma_e mu_e gamma_l mu_l gamma_p mu_p f...
5     beta_m mu_m theta_m mu_h beta_h theta_h gamma_h...
6     A_c1 t0_c1 Delta_tc1 Pul(t);
7 %Definicion del pulso de control
8 Pul(t) = piecewise(t0_c1<=t<=t0_c1+Delta_tc1,A_c1,0);
9 %Definicion de las ecuaciones del modelo con matematica simbolica
10 ode1 = diff(E) == delta*(1-E/C)*(Ms+Me+Mi) - (gamma_e+mu_e)*E;
11 ode2 = diff(L) == gamma_e*E - (gamma_l+mu_l)*L;
12 ode3 = diff(P) == gamma_l*L - (gamma_p+mu_p)*P;
13 ode4 = diff(Ms) == f*gamma_p*P - beta_m*Hi*Ms/(Hs+He+Hi+Hr) - (mu_m+Pul(t))*Ms;
14 ode5 = diff(Me) == beta_m*Hi*Ms/(Hs+He+Hi+Hr) - (theta_m+alpha*mu_m+Pul(t))*Me;
15 ode6 = diff(Mi) == theta_m*Me - (alpha*mu_m+Pul(t))*Mi;
16 ode7 = diff(Hs) == -beta_h*Mi*Hs/(Ms+Me+Mi) + (He+Hi+Hr)*mu_h;
17 ode8 = diff(He) == beta_h*Mi*Hs/(Ms+Me+Mi) - (theta_h+mu_h)*He;
18 ode9 = diff(Hi) == theta_h*He - (gamma_h+mu_h)*Hi;
19 ode10 = diff(Hr) == gamma_h*Hi-mu_h*Hr;\
```

Aunque en esta guía se utiliza matemática simbólica, el usuario es libre de construir el modelo en bloques que se muestra en la figura 19 siguiendo la documentación del *GSUA-UCI Toolbox* [177] para la configuración inicial de modelos en Simulink.

Tras definir el modelo de interés, se crea un objeto tipo tabla que agrupa sus diferentes características, tales como los parámetros fijos, el valor nominal, los intervalos de los parámetros, el método numérico para resolver el sistema de ecuaciones dife-

renciales, las salidas de interés del sistema, entre otros (ver código 2). En caso de que se desconozca el orden de ingreso de los parámetros y sus intervalos, se ejecuta una primera vez la función *gsua_dpmat*, la cual indica el orden de ingreso de dichos valores. Posteriormente, se crea un arreglo (*ranges*) con las condiciones iniciales fijas ($H_i(0) = [8, 8]$ y $H_r(0) = [0, 0]$) y los intervalos de los parámetros en el orden requerido (los valores utilizados se presentan en la tabla 17). Nuevamente se ejecuta la función *gsua_dpmat*, pero esta vez ingresando el arreglo *ranges* y anexando la opción *ñominal',[]*, la cual indica que el valor nominal a implementar es la media de los intervalos definidos en el vector *ranges*.

Código 2: Definición de los intervalos de estimación de parámetros para el modelo (22).

```

1 %Orden de ingreso de rangos y parametros biologicos
2 gsua_dpmat(odes, vars, [0 93], 'Modelo1P', 'output', 1);
3 %Arreglo de rangos de estimacion
4 Ranges = [8 8; 0 0; 0 100000; 0 100000; 0 100000; 10000 10000000; 0 2000;
5           0 1000; 50000 407000; 0 1000; 0 1; 60000 340000; 1 12; 1 2; 0 4;
6           0 4; 20 120; 0 0.7; 0 2.3; 0 2; 0 1.6; 0 1.7; 0 1.3;
7           1.00E-05 0.001; 0 3.22; 0 0.9; 0 1.4; 30 40; 0.4 1.8; 0.4 1];
8 %Agrupacion de las ecuaciones del sistema
9 odes = [ode1; ode2; ode3; ode4; ode5; ode6; ode7; ode8; ode9; ode10];
10 %Agrupacion de los estados del modelo
11 vars = [Hi(t) Hr(t) E(t) L(t) P(t) Ms(t) Me(t) Mi(t) Hs(t) He(t)];
12 %Creacion del objeto tipo tabla con intervalos iniciales de estimacion
13 %Metodo numerico por defecto: ode45
14 [T, Sol] = gsua_dpmat(odes, vars, [0 93], 'Modelo1P', 'range', Ranges, 'output', 1,
15                       'nominal', [])

```

Tabla 17: Intervalos: de estimación, de confianza y valores nominales.

Factor	Intervalos de estimación	CSB	Valor nominal
$E(0)$	0 -100.000	20.000-80.000	50.000
$L(0)$	0 -100.000	9.100 -95.000	61.000
$P(0)$	0 -100.000	20.000- 90.000	50.000
$M_s(0)$	10.000-10.000.000	4.500.000- 7.900.000	4.800.000
$M_e(0)$	0 -2.000	380 -770	520
$M_i(0)$	0 -1.000	120- 250	210
$H_s(0)$	50.000-407.000	93.000 -140.000	120.000
$H_e(0)$	0 -1.000	26- 35	27
δ	20- 120	29 -160	31
C	60.000-340.000	190.000- 230.000	220.000
γ_e	0 -2,3	1,07-1,42	1,17
μ_e	0 -1,3	0,25 -1,6	1
γ_l	0 -1,6	0,78 -1,2	0,82
μ_l	0 -3,22	1,102-1,581	1,58
γ_p	0 -1,7	0,72 -1,5	0,78
μ_p	0 -1,4	0,4 -0,9	0,6
f	0 -0,7	0,28- 0,367	0,352
β_m	0 -4	1,77 -1,825	1,817
μ_m	0 -0,9	0,382076- 0,38721	0,387208
α	1 -2	1,61- 1,64	1,63
θ_m	0,4-1	0,545 -0,573	0,56
μ_h	1,00E-05 -0,001	0,00013- 0,00061	0,00044
β_h	0 -4	1,16 -1,19	1,18
θ_h	0,4- 1,8	0,567 -0,647	0,574
γ_h	0 -2	1,759- 1,801	1,797
A_{c1}	0 -1	0,68- 2,5	0,69
t_{0c1}	30- 40	36,02- 37,77	36,04
Δt_{c1}	1 -12	4 -9	7

Nota: Intervalos de estimación, intervalos de confianza (utilizando método UCI) y valores nominales para 8 condiciones iniciales, 17 parámetros del modelo y 3 parámetros de entrada tipo pulso.

Como se observa en el código 3, se carga la curva de datos reales a los que el modelo 22 se ajustará, los cuales se almacenan en la variable *ydata* y corresponden a los casos de dengue en 93 semanas epidemiológicas y al vector de tiempo *xdata*. La función que realiza la estimación de parámetros es *gsua_pe*, e implementa los datos reales, los rangos definidos en la tabla *T*, una serie de algoritmos de optimización y un solucionador del modelo definido como un *function handle*. Para dicha función de estimación se deben definir algunas opciones utilizando *optimoptions*, las cuales para el caso de estudio corresponden a la siguiente configuración: número de estimaciones de parámetros ($N = 300$), número máximo de evaluaciones del modelo por estimación (*MaxFunctionEvaluations* = 5,000), número máximo de iteraciones por estimación (*MaxIterations* = 300) y la tolerancia en el paso (*StepTolerance* = $1e^{-7}$). Los algoritmos que se utilizan en el proceso de estimación son *'lsqcurvefit'* y *trust-region-reflective* (este último es implementado por defecto). El primer algoritmo soluciona problemas de ajuste de curvas no lineales y el segundo es un algoritmo de optimización para el primero.

Código 3: Estimación de parámetros para el modelo 22.

```

1 %Lectura de datos reales
2 load('Data_Bello_2010_2014.mat')
3 ydata = yBello(1:93)';
4 %Vector tiempo
5 xdata = linspace(0,93,93);
6 %Definicion de opciones de estimacion usados en la funcion gsua_pe()
7 opt = optimoptions('lsqcurvefit','UseParallel',1,'MaxFunctionEvaluations',...
8     5000,'MaxIterations',300,'StepTolerance',1e-7);
9 solver = 'lsq';
10 %Funcion de estimacion de parametros para el modelo de interes
11 [T,res]=gsua_pe(T,xdata,ydata,'solver',solver,'opt',opt,'N',300,'Show','on');
```

Al finalizar las estimaciones predefinidas en el código 3, la función *gsua_pe* encargada de la estimación guarda los resultados obtenidos de dos maneras: a) anexando a la tabla T una variable tipo columna con el valor de los parámetros o b) guardando en un archivo *.mat* un vector con el valor de la función de coste (x) y una tabla con el valor de los parámetros (res). En este caso se trabaja con el archivo denominado *Estimations.mat* y, como se muestra en el código 4, se obtiene el gráfico de todas las estimaciones junto con los datos reales para determinar visualmente la calidad de los resultados.

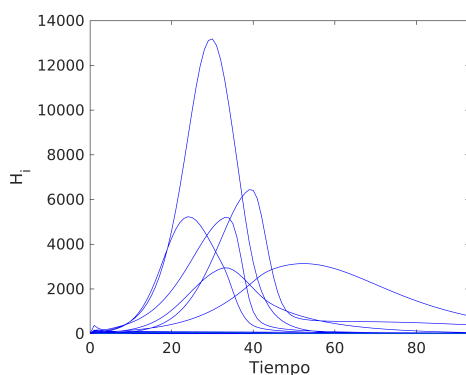
Código 4: Gráficos de todas las estimaciones realizadas para el modelo (22).

```
1 %Cargar datos de estimaciones
2 load('Estimations.mat')
3 figure
4 %Simulacion de todas las estimaciones
5 gsua_eval(x',T); %Funcion que simula y grafica
6 hold on
7 %Comparacion de datos reales y simulados
8 plot(xdata,ydata)
9 ylabel('H_i')
10 xlabel('Tiempo')
```

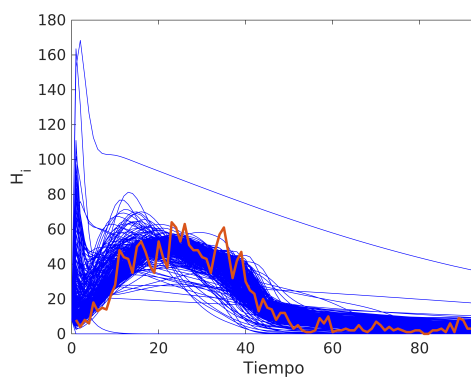
En la figura 20(a) se presentan las curvas correspondientes a las 300 estimaciones realizadas, de las cuales algunas curvas son atípicas y no permiten observar el comportamiento de los datos reales. Por lo tanto, son removidas, obteniendo un gráfico con mayor visibilidad del comportamiento epidémico y en el cual se observa la tendencia de la mayoría de las estimaciones realizadas (ver figura 20(b)). Aun así, se presenta una gran cantidad de curvas que no se ajustan a los datos reales, a pesar de haber pasado por el proceso de optimización de parámetros. Esto res-

ponde principalmente a mínimos locales o a estimaciones que no convergen debido al cumplimiento de alguna de las condiciones impuestas en *optimoptions* en el código 3.

Figura 20: Simulaciones de los parámetros estimados (curvas azules) junto a los datos reales de casos de dengue (curva roja).



(a) Simulaciones de las 300 estimaciones de parámetros realizadas.



(b) Simulaciones retirando las estimaciones de parámetros más lejanas.

Fuente: Elaboración propia.

A causa de la gran cantidad de curvas atípicas obtenidas, como se muestra en las figuras 20(a) y 20(b), es necesario escoger aquellas estimaciones que converjan y se ajusten a los datos reales. Esto se realiza mediante el filtrado por valor de función de coste (ver sección anterior), donde se seleccionan aquellas estimaciones cuya función de coste no supere en un 30 % la función de coste mínima obtenida (ver código 5). Con esto se busca eliminar los resultados atípicos, debido a que estos no representan el comportamiento del modelo de interés y, a su vez, afectan considerablemente la estimación de los parámetros nominales para el modelo de interés.

Código 5: Selección de las 300 estimaciones de parámetros realizadas para el modelo (22).

```

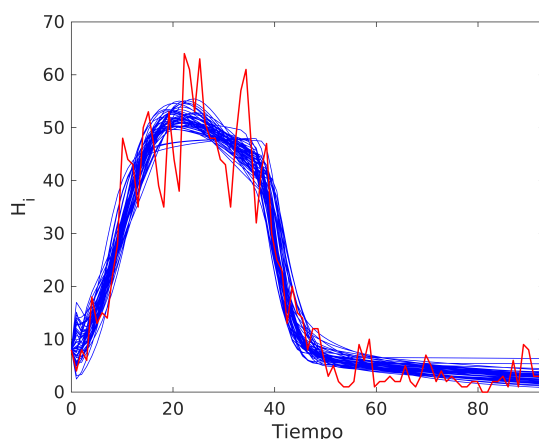
1 %Filtrar de curvas atipicas
2 Minfunc = min(res); %Funcion de coste minima obtenida
3 %Seleccion de curvas que cumplen criterio de filtrado
4 FilteredRes = find(res <= Minfunc*1.3);
5 FilterPar = Par(FilteredRes,:);
6 %Grafico de estimaciones seleccionadas
7 figure
8 gsua_eval(FilterPar,T,xdata);
9 hold on
10 %Comparacion de datos reales y simulados
11 plot(xdata,ydata)
12 ylabel('H_i')
13 xlabel('Tiempo')

```

Al realizar el filtrado anterior, el 16 % de las estimaciones fueron aceptadas por el criterio de selección; las demás fueron descartadas. En la figura 21 se presentan las estimaciones seleccionadas, donde se observa una franja de curvas que rodean los datos reales sin presentar curvas atípicas. A partir de estas estimaciones se obtiene una curva nominal, es decir, una familia

de parámetros que representa todas las estimaciones filtradas. Como se muestra en el código 6, se utilizan dos estadísticos de centralidad (la media y la mediana) para estimar dicha curva. Dependiendo del ajuste obtenido para cada uno de los estadísticos, se elige uno para obtener intervalos de confianza y simular diferentes escenarios epidémicos.

Figura 21: Datos reales (curva en rojo) y estimaciones de parámetros seleccionado por criterio de filtrado (curvas en azul).



Fuente: Elaboración propia.

Cabe resaltar que el gran número de estimaciones fallidas se pudo producir al tamaño de los intervalos de estimación, debido a que si dichos intervalos son muy amplios puede existir una mayor cantidad de mínimos locales que hacen que las estimaciones adecuadas o cercanas al mínimo global se reduzcan, y de esa manera sea necesario realizar más estimaciones. Por lo tanto, es importante definir intervalos de estimación adecuados

que comprendan valores biológicos y restrinjan valores atípicos, siendo necesario realizar diferentes estimaciones de parámetros variando los intervalos iniciales con el fin de encontrar intervalos en los que la mayoría de estimaciones puedan converger a un valor óptimo. Asimismo, se deben escoger intervalos en los que los parámetros no tiendan a uno de sus extremos, ya que se estaría forzando al algoritmo a compensar dicha limitación con los demás parámetros.

Código 6: Cálculo del valor nominal utilizando la media y mediana de los parámetros estimados.

```

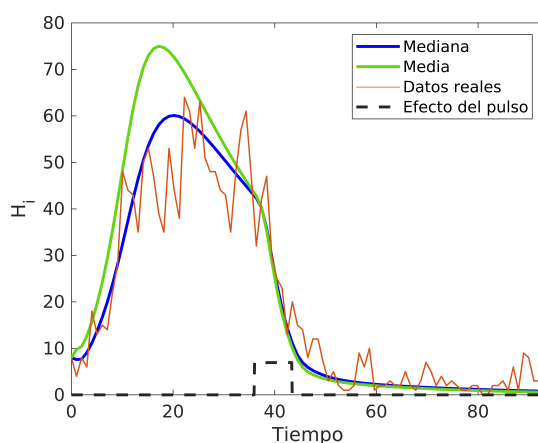
1  %Estadísticos de centralidad de estimaciones filtradas
2  Median = median(FilterPar ');
3  Mean = mean(FilterPar ');
4  %Grafico de curva nominal usando estadísticos de centralidad
5  figure(3)
6  gsua_eval([Median,Mean],T,xdata);
7  hold on
8  %Comparacion de datos reales y simulados
9  plot(xdata,ydata)
10 ylabel('H_i')
11 xlabel('Tiempo')
12 %Definicion de parametros de entrada tipo pulso
13 Delta_tc1=table2array(T('Delta_tc1','Median')); %Ancho de pulso
14 t0_c1=table2array(T('t0_c1','Median'));%Tiempo inicial
15 A_c1=table2array(T('A_c1','Median')); %Amplitud
16 %Funcion que representa al pulso
17 f = @(t)(A_c1*10).*((t0_c1 <= t & t <= Delta_tc1 + t0_c1));
18 x=linspace(0,93,93);
19 %Grafico del pulso
20 plot(x, f(x))

```

En la figura 22 se observan dos curvas calculadas con los estadísticos de centralidad, donde la curva en el primer caso se aleja de los datos reales, generando así un pico más pronunciado. Por otro lado, la mediana conserva la tendencia del pico epidémico y parte de la zona endémica, pero la enfermedad tiende a des-

aparecer con el tiempo. Por lo tanto, se escoge la mediana como estimador de la curva nominal, en tanto que esta logra captar la dinámica epidémica y endémica de la enfermedad y es una medida robusta que no se ve afectada por los datos atípicos. En la tabla 17 se presentan los resultados nominales para cada uno de los parámetros estimados, junto con sus intervalos de confianza y de estimación (ver siguiente sección).

Figura 22: Curvas nominales calculadas con estadísticos de centralidad (media y mediana), junto al efecto del control químico modelado como una entrada de tipo pulso.



Fuente: Elaboración propia.

Cálculo de los intervalos de confianza

A continuación, se presentan dos métodos para obtener los intervalos de confianza de los parámetros nominales (intervalos paramétricos usando la mediana e intervalos CSB), los cuales se describen en la sección anterior. En primer lugar, para obtener

los intervalos paramétricos se utiliza la ecuación (23), la cual es aplicada en el código 7 junto a un análisis de incertidumbre. Dicho análisis ayuda a la validación gráfica de los intervalos obtenidos, ya que genera 1,000 combinaciones de parámetros diferentes que son simulados utilizando las funciones *gsua_dmatrix* y *gsua_ua*. Dependiendo del comportamiento de las curvas simuladas, respecto a la curva nominal, se decide si los intervalos son adecuados para describir la incertidumbre asociada al modelo.

Código 7: Cálculo de Intervalos de confianza paramétricos utilizando la mediana y la ecuación (23).

```

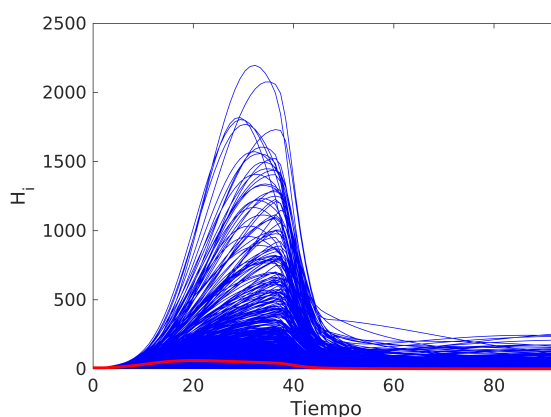
1 %% Calculo de intervalos con la mediana
2 [NumParams, c]=size(FilterPar); %c: numero de estimaciones filtradas
3 standarDesv = std(FilterPar ');
4 %Determina limites superior e inferior de intervalos de confianza
5 DistM=1.96*sqrt(pi)*standarDesv/(sqrt(c-1));
6 intervalinf=Median - DistM';
7 intervalsup=Median + DistM';
8 %Se anexan los intervalos al objeto T
9 T.Range=[intervalinf, intervalsup]
10 %Se anexan la mediana como valor nominal al objeto T
11 T.Nominal=Median;
12 %Grafico del analisis incertidumbre para intervalos calculados con la mediana
13 [M,T]=gsua_dmatrix(T,1000);
14 gsua_ua(M,T, 'xdata', xdata);

```

En la figura 23 se observa el resultado del análisis de incertidumbre, el cual muestra que las familias de parámetros generadas son curvas atípicas que sobrepasan en gran medida a los datos reales. Por lo tanto, dichos intervalos no son informativos por la gran incertidumbre que se presenta. Si se quisiera reducir el número de curvas atípicas y, por consiguiente, los intervalos de confianza, sería necesario realizar N estimaciones, con un número suficientemente grande. Por lo tanto, dicho méto-

do paramétrico no entrega buenos resultados a menos de que se realice un gran número de estimaciones, siendo esto muy costoso a nivel computacional, y es necesario utilizar otra estrategia para estimar intervalos coherentes para cierto nivel de incertidumbre. Se sugiere utilizar el método de los CSB explicados en la sección anterior, lo cual garantiza intervalos de confianza que generan curvas con comportamientos similares a la curva nominal. Dicho método también se explica en la guía de usuario del *GSUA-UCI Toolbox* y en [177]. El CSB utiliza los fundamentos del análisis de sensibilidad para reducir o ampliar los intervalos de confianza.

Figura 23: Análisis de incertidumbre utilizando los intervalos obtenidos a partir de la distribución de la mediana.



Nota: La curva nominal se representa en rojo, mientras que en azul se muestran las diferentes simulaciones obtenidas de los intervalos de confianza paramétricos.

Como se muestra en el código 8, antes de utilizar el método CSB se debe llamar la función *gsua_oatr*, la cual crea un intervalo inicial para cada parámetro a partir de la ampliación del intervalo de cada parámetro, dejando los demás parámetros en su valor nominal. Posteriormente, para la función *gsua_ucis* del método CSB se utilizan los rangos obtenidos anteriormente por el *gsua_oatr* (objeto tabla nombrado como *RangesOAT*) y se define el número de simulaciones por iteración (en este caso son 1,000), así como el número máximo de iteraciones (400) y si se desea utilizar computación en paralelo. El nivel de confianza predefinido para los intervalos es de aproximadamente 95 %, es decir, al menos el 95 % de las curvas obtenidas de los intervalos CSB cumplen con un criterio de similitud respecto a la curva nominal.

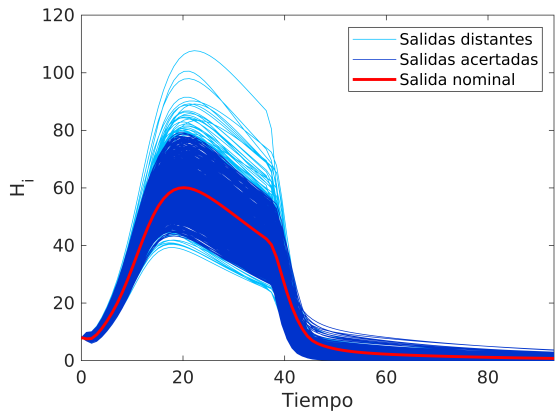
Código 8: Cálculo de Intervalos de confianza utilizando el método CSB.

```
1 %Calculo de intervalos OAT para funcion gsua_ucis()
2 RangesOAT=gsua_oatr(T, 'parallel',true);
3 %Calculo de Confidence Subcontour Box (\textsc{CSB})
4 T=gsua_ucis(RangesOAT,1,000,'reps',400,'parallel',true)
```

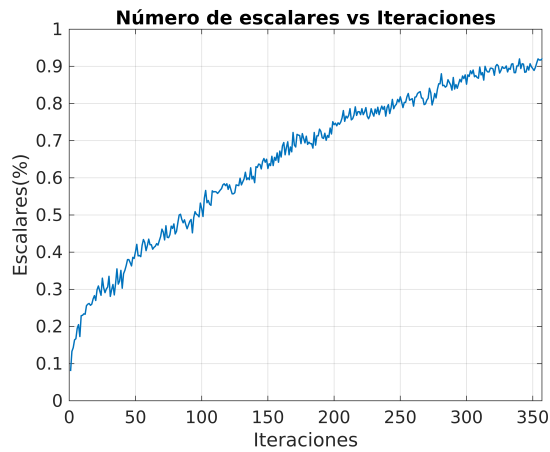
En la tabla 17 se presentan los intervalos obtenidos por el método CSB para cada uno de los parámetros y, como se puede ver en la figura 24(a), estos intervalos no generan curvas atípicas como las presentadas en la figura 23. Las curvas más internas mantienen la misma tendencia y forman una franja sobre la curva nominal (en rojo). De manera similar, las curvas atípicas (en azul claro) representan aproximadamente el 5 % del total de las curvas generadas y, aunque el criterio las catalogue como dis-

tantes, siguen la misma tendencia de la curva nominal y de las curvas internas. En la figura 24(b) se observa la convergencia del método, el cual tomó cerca de 357 iteraciones y 11 horas de cómputo (en la plataforma computacional descrita en la sección *Estimación de parámetros*) en lograr que al menos el 95 % de las curvas generadas por los intervalos de confianza fueran catalogadas como acertadas o adecuadas. No obstante, aparte del gráfico de UA que cumple con el requerimiento de pocos o ningún valor atípico, es necesario aclarar otros aspectos como la necesidad de calcular los intervalos para la mediana y no para el valor medio. Igualmente, el método exigió mucho tiempo computacional debido, posiblemente, a que el espacio de solución puede presentar diferentes zonas con soluciones atípicas; para resolver este problema se pueden realizar más estimaciones que mejoren el valor de la mediana como curva nominal y produzcan un espacio de solución más sencillo. Estos dos últimos asuntos serán abordados en futuros estudios, pero hasta el momento, en las pruebas realizadas, el método entrega excelentes resultados.

Figura 24: Resultados gráficos de la estimación de intervalos de confianza utilizando el método CSB.



(a) Análisis de incertidumbre para intervalos de confianza estimados con CSB, separando los tipos de salida por distantes y acertadas.



(b) Convergencia del método CSB para le modelo en (22) y los parámetros nominales obtenidos con la mediana.

Fuente: Elaboración propia.

Tras obtener los valores nominales y sus respectivos intervalos de confianza (ver tabla 17), se realiza ahora una breve discusión de los resultados obtenidos a la luz de diversos estudios científicos que han reportado con anterioridad el valor de diferentes parámetros relacionados con el desarrollo del vector, la población humana y la transmisión del dengue.

En el caso del modelo 22, se tiene información reportada para 14 de sus 20 parámetros, de los cuales Δt_{c1} , t_{0c1} , α , f , θ_m , μ_p , γ_p , μu_l , γ_e y θ_h presentan valores consistentes con los evidenciados en la literatura [174, 175, 180–182]. Esto es un indicativo de que la estimación de parámetros e intervalos de confianza ha logrado captar información de la dinámica de la enfermedad a través del modelo matemático, entregando valores de parámetros con sentido biológico. Asimismo, cabe resaltar que estimar parámetros incorporando la entrada de tipo pulso ayuda a separar el efecto de fumigación del valor estimado de los parámetros del modelo.

A pesar de haber obtenido parámetros con valores encontrados en la literatura, también se presentaron algunos que no coinciden con dichos reportes, entre ellos están μ_m , γ_l , δ y γ_h [182]. Aun así, cada uno de estos parámetros tiene una explicación enmarcada en la biología del vector o del humano. En el caso de la tasa de mortalidad es intuitivo determinar que su valor será mucho mayor en condiciones de campo que en condiciones ideales de laboratorio, ya que situaciones de depredación, cambios en el clima y en la alimentación hacen que la mortalidad del mosquito

aumente [131]. En el caso de los valores estimados para γ_l (superior al reportado por [180]) y δ (inferior al reportado por [182]), estos pueden ser alterados por el efecto de condiciones extrínsecas como la limpieza de criaderos de mosquitos. Si se quisiera obtener un valor aproximado de dichos parámetros sería necesario anexar una entrada de tipo pulso a las condiciones acuáticas, la cual separaría el efecto de la limpieza de criaderos y el valor de los parámetros de estadios acuáticos. Finalmente, el parámetro relacionado con la población humana (γ_h) puede presentar una variación de su valor respecto al presentado por la literatura, ya que diferentes condiciones ambientales y genéticas pueden hacer variar el valor nominal de un parámetro de población en población.

Gráficas del análisis de sensibilidad global

En este caso, el análisis de sensibilidad (SA) permite identificar los parámetros más importantes para la curva solución y su comportamiento en una vecindad definida por el 1 % de incertidumbre en cada parámetro, con lo que se genera una franja similar a la obtenida en el análisis de incertidumbre de los CSB (comparar figuras 24(a) y 25). Es aconsejable realizar el SA a fin de determinar aquellos parámetros en los que se debe enfocar la estimación y el control, para afectar de manera más radical el comportamiento de la salida del modelo. La razón por la que todos los parámetros tienen el mismo nivel de incertidumbre es para compararlos en la misma escala: no estamos tratando de en-

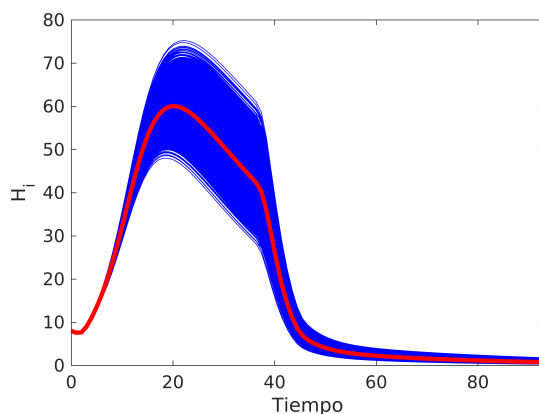
contrar los parámetros que determinan el comportamiento del modelo dentro de los rangos de estimación (lo cual es útil antes de la estimación de parámetros), sino que queremos identificar los parámetros con el mayor potencial para modificar la curva nominal. En ese orden de ideas, el SA adquiere gran similitud con el análisis basado en derivadas parciales sobre una familia de parámetros específica, pero las diferencias fundamentales radican en que se obtiene información de todas las interacciones entre los parámetros y, además, se penaliza la distancia, ya que se evalúa con respecto a la MSE. En el código 9 se generan intervalos a partir de los valores nominales ampliados en 1 % con la función *gsua_dataprep*.

Código 9: Análisis de incertidumbre de los intervalos creados a partir de ampliar el valor de la mediana en 1 %.

```

1 %Creacion de tabla con rangos de porcentajes de la curva nominal
2 Percent = [8;0;Median];
3 Percent = cat(2,Percent,[0;0;ones(28,1)*1])
4 [T,sol] = gsua_dataprep(odes,vars,[0 93], 'SIR', 'range',Percent, 'rMethod', 'percent', '
    output',[1]);
5 % Analisis de incertidumbre
6 [M,T]=gsua_dmatrix(T,6000);
7 gsua_ua(M,T, 'xdata',xdata, 'show', 'on');
```

Figura 25: Análisis de incertidumbre utilizando intervalos creados a partir del 1 % de los valores nominales.



Fuente: Elaboración propia.

La matriz M es una matriz de diseño creada con el método de hipercubo latino, el cual trata de lograr la máxima cobertura del hiperespacio definido por los intervalos de los parámetros y, de esa forma, es posible estimar con mayor confianza la contribución de los parámetros. La matriz M , creada en el código 9 con un tamaño de 6,000 familias de parámetros, se utilizó tanto para el análisis de incertidumbre (con *gsua_au*) como para el de sensibilidad (con *gsua_sa*); este último se llevó a cabo implementando el método de Saltelli, descrito en la sección *Análisis de incertidumbre y sensibilidad*, el cual entrega índices de sensibilidad escalar (STi y vectorial (STi_vec). Al utilizar la función *gsua_plot* y los índices de sensibilidad, se pueden obtener gráficos que muestran la importancia de los parámetros en porcentaje (*'Pie'*) y en el tiempo (*'TotalSensitivityArea'*).

Código 10: Análisis de sensibilidad de los intervalos creados a partir de la ampliación del 1 % de la mediana.

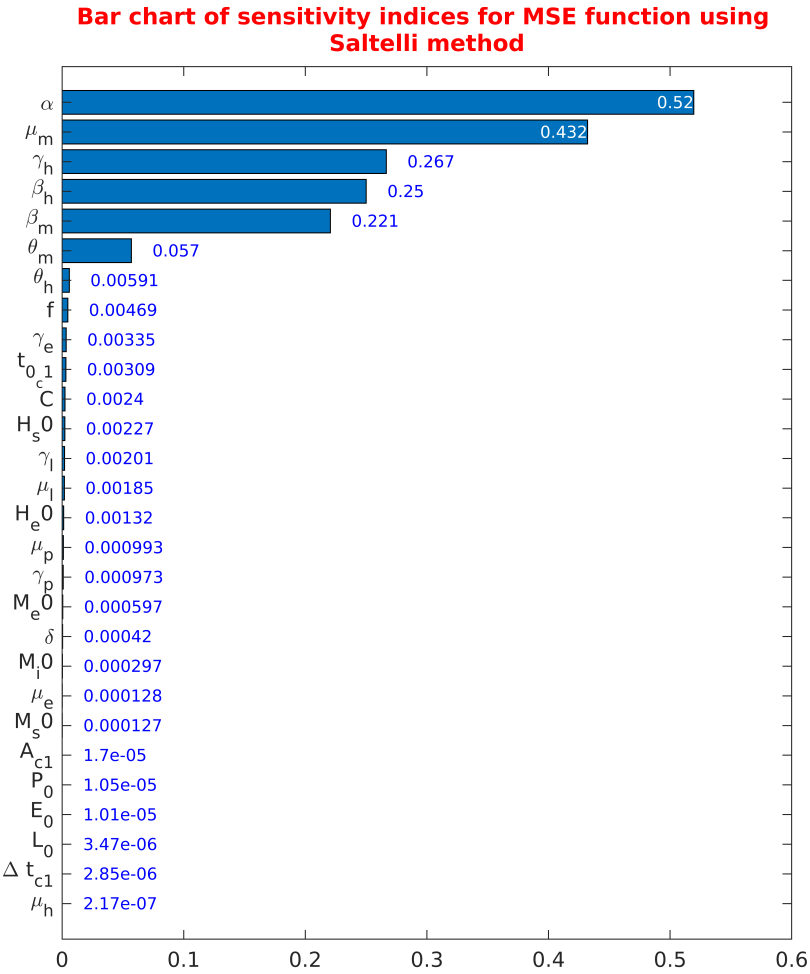
```

1  % Analisis de sensibilidad
2  [T,J,Y]=gsua_sa(M,T, 'SensMethod', 'Saltelli');
3  %Torta de indices de sensibilidad global (STi)
4  figure
5  gsua_plot( 'Pie',T,T.STi)
6  %Grafico de barras del valor de los STi
7  figure
8  gsua_plot( 'Bar',T,T.STi)
9  %Grafico de indices STi vectoriales
10 figure
11 gsua_plot( 'TotalSensitivityArea',T,T.STi_vec,xdata);

```

Como se observa en la figura 26(a), los parámetros más importantes en el enfoque escalar son α , μ_m , γ_h , β_h y β_m , mientras que los demás parámetros ($\Sigma(23\text{factores})$) no son importantes en el modelo. La figura 26(b) muestra que cerca del 50 % del modelo es explicado por parámetros relacionados con la mortalidad del mosquito (α y μ_m). En la figura 26(c), que representa la importancia de los parámetros a lo largo del tiempo, se observa que la clasificación de los índices escalares se mantiene y, adicionalmente, aparece un parámetro extra en la semana 36. Dicho parámetro corresponde a la fecha inicial de la acción de control químico efectuada en el municipio ($t_{0_{c1}}$), y su importancia se mantiene durante el número de semanas en que persiste el efecto residual (Δt_{c1}).

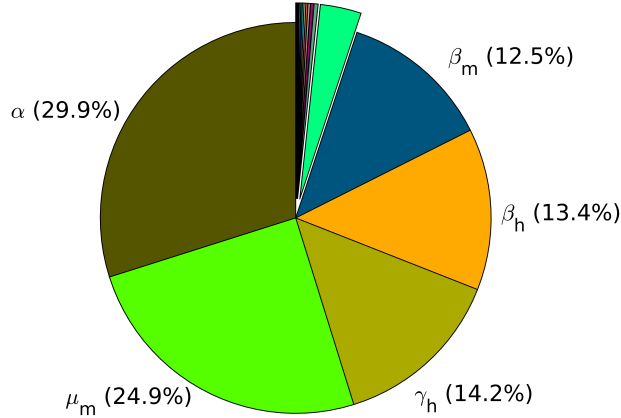
Figura 26: Análisis de sensibilidad global para intervalos de confianza obtenidos a partir del 1 % de la curva nominal estimada.



(a) Gráfica de barras del valor de los índices escalares para cada parámetro.

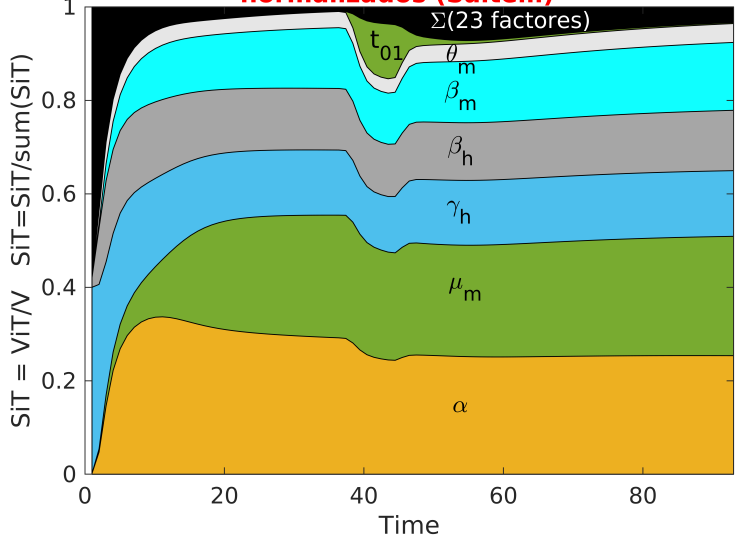
Indices de sensibilidad escalares (Saltelli)

Σ (23 Factores) = 5.07%



(b) Gráfica de torta del porcentaje para índices escalares.

Indices de sensibilidad vectoriales normalizados (Saltelli)



(c) Gráfica de área vectorial para índices vectoriales normalizados.

El análisis de sensibilidad vectorial revela que el comportamiento a largo plazo del sistema depende casi exclusivamente de los valores de 6 de los 29 parámetros, pero durante los procesos de perturbación (dados por el tiempo inicial de la simulación o el tiempo de efecto del pulso de control) los 23 parámetros con poca relevancia juegan un papel más importante.

Estos resultados sugieren que el modelo podría ser reducido o que varios parámetros podrían ser fijados sin afectar considerablemente el ajuste del modelo. Parece ser que la dinámica de los estadios inmaduros del vector no aporta mayor información a la dinámica de la epidemia (excepto para su inicio), lo cual concuerda con la interpretación del R_0 y su cálculo por medio del enfoque de la matriz de la siguiente generación [183]. En ese orden de ideas, parece que el proceso de la epidemia parte de un punto alejado del equilibrio del sistema donde casi todos los parámetros desempeñan un papel y, luego, el sistema es llevado a un punto de equilibrio mayoritariamente determinado por 6 de los parámetros, donde el papel más importante corresponde a la mortalidad del vector y al aumento de la mortalidad en los vectores expuestos al virus. Lo anterior es bastante relevante y concuerda con lo que puede esperarse de las enfermedades transmitidas por vectores: lo que mayoritariamente determina la dinámica de la epidemia es la longevidad del vector, su capacidad de transmitir e infectarse por el virus, y la duración de la infección en el hospedero.

Resumiendo: los brotes epidémicos parecen estar determinados por cambios drásticos en la dinámica del vector y, una vez ha empezado el brote, lo fundamental es aumentar la mortalidad del vector y reducir la exposición entre el mismo y los hospederos de la enfermedad. De esta manera, el mejor método para controlar los brotes epidémicos consistiría en estrategias de prevención basadas en controlar la dinámica del vector (limpieza de criaderos, etc.), fumigaciones y cuarentena, una vez se ha declarado el brote epidémico.

Simulación de diferentes escenarios de control

En esta sección se simulan dos escenarios epidémicos utilizando los parámetros nominales obtenidos previamente. En primer lugar, se evalúa el efecto de la fumigación al variar el momento inicial de la aplicación de dicho control ($t_{0_{c_i}}$). En segundo lugar, se simula una entrada tipo pulso que representa el efecto de la vacunación sobre una población (u_v), con la que se remueven humanos susceptibles y se transfieren a la fase de humanos recuperados.

Como se muestra en el código 11, para simular la fumigación tardía y anticipada como un análisis de incertidumbre se fijan todos los parámetros y se generan familias aleatorias de $t_{0_{c_i}}$ a partir del aumento o disminución en un 30 % respecto a sus valores originales. Como se ve en las figuras 27(a) y 27(b), dichas familias son simuladas y contrastadas con los datos reales. Para el primer caso (27(a)), se puede observar que al realizar una fu-

migación temprana la incidencia del dengue disminuye en gran medida; por otro lado, al realizar una fumigación tardía (27(b)) el número de humanos infectados aumenta respecto al número de casos originales reportados. Por lo tanto, si la campaña de fumigación se hubiera realizado con anterioridad, siguiendo el criterio de canal endémico [182], la epidemia hubiera sido más corta y el número de afectados menor.

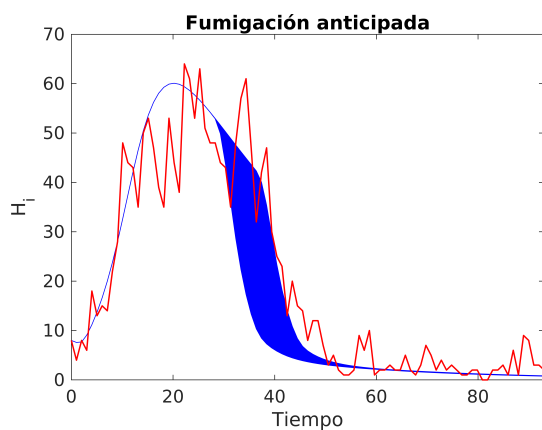
Código 11: Simulación de escenario de control: fumigación anticipada o tardía respecto a la fecha de fumigación estimada.

```

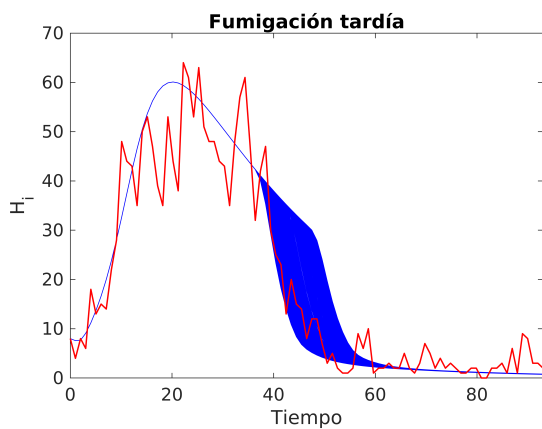
1 N=1000; %Numero de simulaciones
2 %Creacion de matriz de dimension N*NumParams con valores de mediana
3 Parameters= repmat (Median,N);
4 %Porcentaje minimo y maximo de inicio de fumigacion
5 Mint0=[36/1.3 36];
6 Maxt0=[36 36*1.3];
7 %Simulacion de escenarios de fumigacion
8 for i=1:2
9     %Momento de fumigacion aleatorio entre Mint0 y Maxt0
10    t0 = Mint0(i) + (Maxt0(i)-Mint0(i)).*rand(N,1);
11    Parameters(26,:)=t0;
12    figure
13    gsua_eval(Parameters,T,xdata);
14    hold on
15    plot(xdata,ydata, 'r')
16    xlabel('Tiempo')
17    ylabel('H_i')
18 end

```

Figura 27: Escenario epidémico: variación en el inicio del efecto químico de la fumigación.



(a) Fumigación temprana, el parámetro $t_{0_{c_1}}$ se mueve entre la semana 28 y 36.



(b) Fumigación tardía, el parámetro $t_{0_{c_1}}$ se mueve entre la semana 36 y 47.

Fuente: Elaboración propia.

Con lo anterior, se puede observar que la fumigación es un método efectivo para reducir un brote epidémico en proceso. Esto se debe a que la entrada de tipo pulso que lo simula afecta directamente la tasa de mortalidad del mosquito, la cual es catalogada como uno de los parámetros más influyentes sobre la salida del sistema (ver sección anterior). Por ende, con dichas simulaciones y teniendo en cuenta el análisis de sensibilidad, es posible suponer que si al aumentar la mortalidad del vector se puede finalizar una epidemia, entonces, al disminuir la mortalidad del mismo (debido a factores externos como cambios favorables en la temperatura y humedad) puede haber un aumento en la población de mosquitos y en el número de casos de dengue, siendo un desencadenante de epidemias.

Finalmente, se presenta el escenario de vacunación, el cual se simula como un efecto de tipo pulso en el que los humanos susceptibles son vacunados en determinado momento y el efecto de transferencia directa a una fase de recuperación se presenta en semanas específicas. En este caso, el efecto de inmunidad será efectivo en la semana 36 (misma que el efecto de control químico) y dura 7 semanas aproximadamente (mismo período del efecto residual del control químico); la intensidad de la vacunación (porcentaje de personas inmunizadas) se evalúa con diferentes valores. Para su implementación en Matlab, en primer lugar se crea un modelo con matemática simbólica en el que se define un pulso de vacunación ($vac(t)$) y se excluye el pulso de fumigación (código [12](#)).

Código 12: Definición del modelo (22) en matemática simbólica, utilizando una entrada tipo pulso para simular vacunación

```

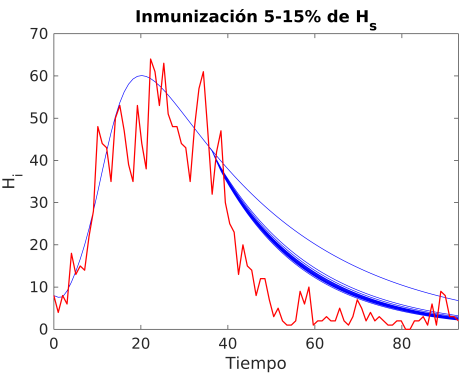
1 %Definicion de estados del sistema
2 syms E(t) L(t) P(t) Ms(t) Me(t) Mi(t) Hs(t) He(t) Hi(t) Hr(t)
3 %Definicion de parametros del sistema
4 syms alpha delta C gamma_e mu_e gamma_l mu_l gamma_p mu_p f...
5     beta_m mu_m theta_m mu_h beta_h theta_h gamma_h...
6     A_cv t0_cv Delta_tcv vac(t);
7 %Definicion del pulso de vacunacion
8 vac(t) = piecewise(t0_cv<=t<=t0_cv+Delta_tcv,A_cv,0);
9 %Modelo incluyendo el pulso de vacunacion
10 ode1 = diff(E) == delta*(1-E/C)*(Ms+Me+Mi) - (gamma_e+mu_e)*E;
11 ode2 = diff(L) == gamma_e*E - (gamma_l+mu_l)*L;
12 ode3 = diff(P) == gamma_l*L - (gamma_p+mu_p)*P;
13 ode4 = diff(Ms) == f*gamma_p*P - beta_m*Hi*Ms/(Hs+He+Hi+Hr) - (mu_m)*Ms;
14 ode5 = diff(Me) == beta_m*Hi*Ms/(Hs+He+Hi+Hr) - (theta_m+alpha*mu_m)*Me;
15 ode6 = diff(Mi) == theta_m*Me - (alpha*mu_m)*Mi;
16 ode7 = diff(Hs) == -beta_h*Mi*Hs/(Ms+Me+Mi) + (He+Hi+Hr)*mu_h-vac(t)*Hs;
17 ode8 = diff(He) == beta_h*Mi*Hs/(Ms+Me+Mi) - (theta_h+mu_h)*He;
18 ode9 = diff(Hi) == theta_h*He - (gamma_h+mu_h)*Hi;
19 ode10 = diff(Hr) == gamma_h*Hi-mu_h*Hr+vac(t)*Hs;

```

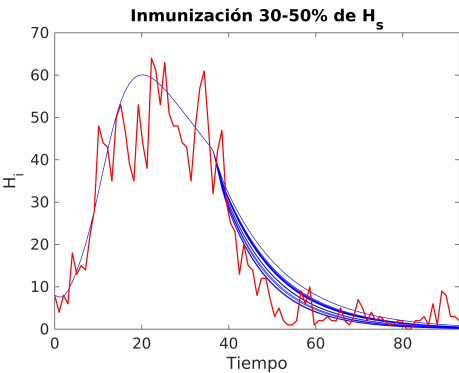
Similar a lo realizado para la simulación del efecto químico en el código 11, el código 13 muestra cómo se obtuvieron los porcentajes de inmunización (A_{cv}) de la población susceptible; dichos valores oscilan entre el 5-15 % para un primer caso (figura 28(a)), y 30-50 % para un segundo caso (figura 28(b)). Comparando ambas figuras se puede observar que entre mayor sea la inmunización de la población de humanos susceptibles, menor será el impacto de la enfermedad sobre la población. Aunque la vacunación sea un método efectivo a largo plazo comparado con la fumigación, es una estrategia bastante costosa, ya que se tendrían que vacunar más de 40,000 personas durante siete semanas de campaña de vacunación. Por lo tanto, sería conveniente encontrar un punto de equilibrio entre diferentes acciones de control como la vacunación (en caso de que se apruebe

una vacuna totalmente efectiva), control químico, limpieza de criaderos, utilización de repelentes y toldillos, entre otros. Dicha combinación de métodos de control puede ser simulada a partir de entradas tipo pulso en el modelo (22).

Figura 28: Efecto de inmunización por vacunación sobre epidemia de dengue.



(a)



(b)

Fuente: Elaboración propia.

Código 13: Simulación de escenario de control: efecto de vacunación sobre cierto porcentaje de la población susceptible

```
1 %Numero de simulaciones por escenario
2 N=1000;
3 Parameters= repmat(Median,N);
4 %porcentaje minimo y maximo de personas inmunes
5 MinVac=[0.05 0.3];
6 MaxVac=[0.15 0.5];
7 %Agrupacion de las ecuaciones del sistema
8 odes=[ode1; ode2 ;ode3; ode4; ode5 ;ode6; ode7; ode8; ode9; ode10];
9 %Agrupacion de los estados del sistema
10 vars=[Hi(t) Hr(t) E(t) L(t) P(t) Ms(t) Me(t) Mi(t) Hs(t) He(t)];
11 %Creacion del objeto tipo tabla
12 [T, Sol]=gsua_dpmt(odes,vars,[0 93], 'SIR', 'range', Ranges, 'output', 1, 'nominal', [8;0;
    Median])
13 %Simulacion de escenarios de vacunacion
14 for i=1:2
15     %Amplitud de vacunacion aleatoria entre MinVac y MaxVac
16     Av = MinVac(i) + (MaxVac(i)-MinVac(i)).*rand(N,1);
17     Parameters(11,:)=Av;
18     figure
19     gsua_eval(Parameters,T,xdata);
20     hold on
21     plot(xdata,ydata, 'r')
22     xlabel('Tiempo')
23     ylabel('H_i')
24 end
```

Conclusiones

En este capítulo se utilizó un modelo matemático que describe la transmisión del dengue en una población de humanos, utilizando como caso de estudio la epidemia ocurrida en el 2010 en el municipio de Bello (Colombia). Al implementar dicho modelo y diferentes métodos, fue posible realizar una guía paso a paso de la estimación de parámetros e intervalos de confianza (específicamente, los intervalos *CSB*), junto a los análisis de sensibilidad e incertidumbre que validaron los parámetros más importantes y

los intervalos de confianza, respectivamente. Con esto se pudieron obtener valores de parámetros e intervalos con significado biológico donde los más significativos fueron aquellos relacionados con la mortalidad del vector (α y μ_m) y las probabilidades de infección (β_h). Se mostró que fue posible obtener dichos valores de parámetros al agregar una entrada de tipo pulso que modela el efecto de la fumigación, ya que el pulso separa el efecto del control del valor medio de los parámetros. También se realizaron algunas simulaciones que explican el beneficio de las acciones de control químico y por vacunación, pero empleando un modelo con parámetros más fieles a la realidad.

Los resultados obtenidos con dicha metodología abren paso a otro tipo de preguntas como: a) estimación de parámetros para diferentes brotes epidémicos de un mismo lugar; b) estimación de parámetros e intervalos de confianza para otras zonas endémicas y su comparación con el fin de determinar si son similares; c) utilización de entradas de tipo pulso para otras estrategias de control como la limpieza de criaderos y repelentes; d) evaluación de otras enfermedades transmitidas por vectores como el zika o chikungunya; e) uso de pulsos negativos para modelar las condiciones que favorecen la proliferación de mosquitos, y buscar su correlación con condiciones ambientales históricas conocidas.