El algoritmo Metropolis-Hastings: problemas inversos, calibración de modelos y el modelo TOMGRO

Trabajo realizado por **Carlos Nosa** ⋄
Trabajo dirigido por **Juan Galvis** ⋄ y **Rodrigo Gil** ⋆
⋄ Departamento de Matemáticas, Universidad Nacional de Colombia, Sede Bogotá
⋆ Departamento de Agronomía, Universidad Nacional de Colombia, Sede Bogotá

UNIVERSIDAD NACIONAL DE COLOMBIA

Resumen

Este trabajo investigativo aborda el estudio del algoritmo Metrópolis-Hastings (MH), perteneciente a la familia de algoritmos Markov Chain Monte Carlo (MCMC), el cual es ampliamente utilizado en la inferencia bayesiana. Se exploran sus fundamentos teóricos y aplicaciones en problemas inversos, destacando su capacidad para estimar parámetros desconocidos y capturar incertidumbres en modelos matemáticos. Además, se analizan las dificultades asociadas a su implementación, como la selección de distribuciones a priori y la evaluación de la convergencia. Este enfoque se ilustra con ejemplos prácticos que muestran la eficacia del algoritmo en el contexto de la calibración de modelos.

Preliminares

El estudio de los modelos de ecuaciones diferenciales y su estimación de parámetros requiere una base teórica sólida en problemas inversos, ecuaciones diferenciales y cadenas de Markov. A continuación, se presenta un resumen de los conceptos clave utilizados en este trabajo, incluyendo la inferencia bayesiana, el algoritmo Metropolis-Hastings y el modelo TOMGRO.

Inferencia bayesiana [2, 3]

Sean Y, X y E variables aleatorias relacionadas mediante el modelo $Y = f(X, E)$. La distribución de probabilidad conjunta entre X y Y es desconocida, pero se puede usar la información a priori $\pi_{\text{pr}}(x)$ para modelar X.
La *función de verosimilitud* $\pi(y|x)$ describe la relación entre los datos observados $y_{\text{observado}}$ y las variables desconocidas X. La distribución posterior de X, dado $y_{\text{observado}}$, se calcula con Bayes:

$$\pi_{\text{post}}(x) = \pi(x|y_{\text{observado}}) = \frac{\pi_{\text{pr}}(x)\pi(y_{\text{observado}}|x)}{\pi(y_{\text{observado}})}.$$

La estimación *máximo a posteriori* (MAP) y la *media condicional* (CM) se definen como:

$$x_{\text{MAP}} = \arg\max_x \pi(x|y_{\text{observado}}) \quad \text{y} \quad x_{\text{CM}} = \mathbb{E}(X|y_{\text{observado}}) = \int_{\mathbb{R}^n} x\pi(x|y_{\text{observado}})\,dx.$$

Algoritmo Metropolis Hastings (MH) [2]

El algoritmo de Metropolis-Hastings es una técnica de muestreo estocástico utilizada para explorar distribuciones de probabilidad. El objetivo es generar una cadena de Markov cuyo estado estacionario sea la distribución deseada.
Sea μ una medida de probabilidad en $\mathbb{R}^n$ y $\{X_j\}$ una cadena de Markov homogénea en el tiempo con núcleo de transición P. Si μ es una medida invariante de P, y P es irreducible y aperiódica, entonces:

$$\lim_{N\to\infty} P^{(N)}(x, B) = \mu(B) \quad \text{para todo} \quad B \in \mathcal{B}.$$

Y para $f \in L^1(\mu(dx))$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} f(X_j) = \int_{\mathbb{R}^n} f(x)\mu(dx) \quad \text{casi seguramente.}$$

Algoritmo. El objetivo es construir un núcleo de transición $P(x, B)$ tal que μ sea su medida invariante. Esto se logra mediante las siguientes etapas:

1. Proponer un nuevo punto $y \in \mathbb{R}^n$ usando una distribución propuesta $q(x, y)$.
2. Calcular la tasa de aceptación $\alpha(x, y)$ con la fórmula:
$$\alpha(x, y) = \min\left(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}\right)$$
donde $\pi(x)$ es la distribución objetivo.
3. Aceptar o rechazar el movimiento:
   - Si $\alpha(x, y) \geq t$ (con $t \in [0, 1]$ aleatorio), aceptar y, es decir, $x_{k+1} = y$.
   - Si $\alpha(x, y) < t$, rechazar el movimiento y mantener $x_{k+1} = x_k$.

Este proceso se repite hasta que se alcance el tamaño deseado de la muestra.

Modelo TOMGRO [1]

TOMGRO es un modelo dinámico que representa el crecimiento y desarrollo de plantas de tomate en invernadero. Se basa en relaciones fuente-depósito e incluye una representación estructurada por edad de las variables de estado. El objetivo es predecir el crecimiento y rendimiento de las plantas en respuesta a condiciones ambientales como temperatura, radiación solar y concentración de CO_2. El modelo usa siete variables de estado: número de tallos (N_S), número de hojas (N_L), número de frutos (N_F), peso de tallos (W_S), peso de hojas (W_L), área foliar (A_L) y peso de frutos (W_F). El crecimiento se describe a través de estos componentes en diferentes clases de edad.

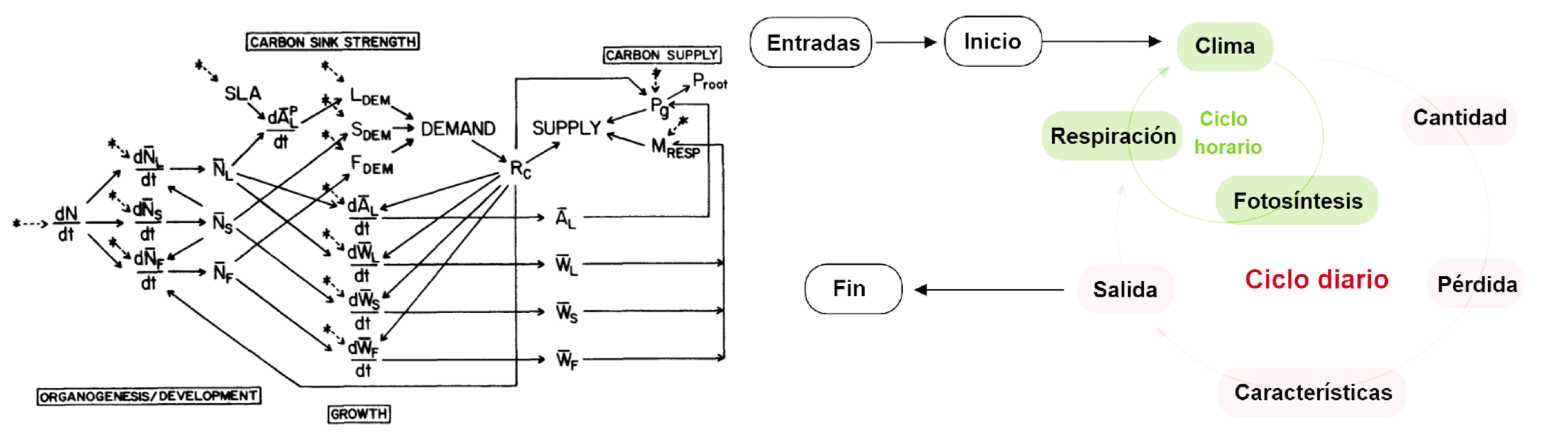

Figura 1: Izquierda: Esquema del desarrollo general de las variables de estado de la planta de tomate en TOMGRO. Figura tomada de [1]. Derecha: Ciclos de funcionamiento del modelo a nivel computacional

Ecuaciones del Modelo

Las dinámicas de cada órgano O en m clases de edad están descritas por:

$$(O)\begin{cases} \frac{dN_{O,1}}{dt} = o_{1,1}N_{O,1} - o_{1,2}N_{O,1}^2 & \\ \frac{dN_{O,k}}{dt} = o_{k-1,2}N_{O,k-1}^2 + o_{k,1}N_{O,k} - o_{k,2}N_{O,k}^2 & \text{para } k = 2, \ldots, m \\ \frac{dC_{O,1}}{dt} = c_{o,1,1}N_{O,1}C_{O,1} - c_{o,1,2}N_{O,1}C_{O,1}^2 & \\ \frac{dC_{O,k}}{dt} = c_{o,k,0}N_{O,k-1}C_{O,k-1}^2 + c_{o,k,1}N_{O,k}C_{O,k} - c_{o,k,2}N_{O,k}C_{O,k}^2 & \text{para } k = 2, \ldots, m \end{cases} \quad (1)$$

Donde los coeficientes dependen del tiempo y de otras variables de estado.

Problema

Objetivo principal
Aplicar técnicas de inferencia bayesiana y métodos de Markov Chain Monte Carlo para la estimación de parámetros en modelos de ecuaciones diferenciales, como TOMGRO, garantizando una caracterización probabilística de la incertidumbre en los parámetros.

Pregunta de investigación
¿Cómo estimar los parámetros de modelos de ecuaciones diferenciales, como TOMGRO, utilizando un enfoque bayesiano basado en métodos de Markov Chain Monte Carlo para mejorar su precisión y capacidad predictiva?

Resultados

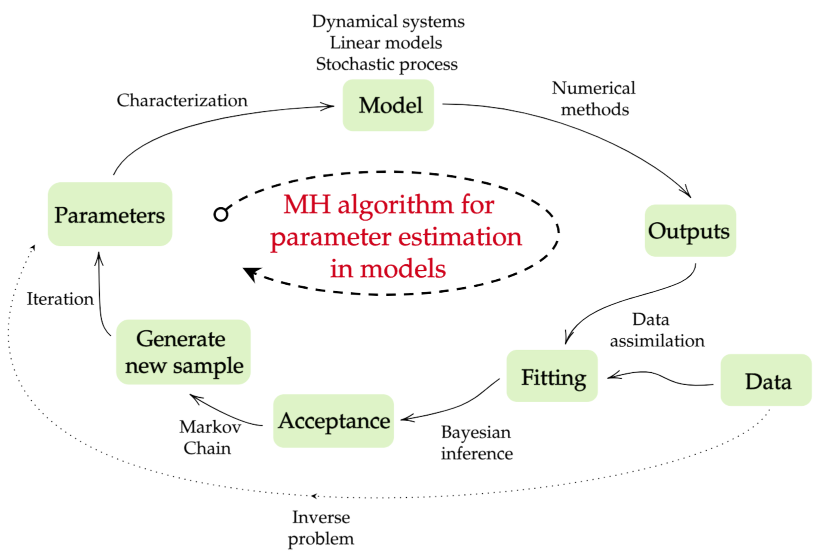

Figura 2: Uso del enfoque bayesiano para la estimación de parámetros de modelos dinámicos.

Se generaron 11×4 datos simulados siguiendo la dinámica descrita por la ecuación 1, utilizando el método de Euler para resolver las ecuaciones diferenciales acopladas y la relación recursiva entre dos clases de edad. El resultado del método numérico se perturbó con un ruido multiplicativo proveniente de una distribución Gamma $\Gamma(\alpha = 100, \beta = 1/100)$. La función de verosimilitud se modeló como una distribución normal multivariada (asumiendo ruido aditivo) ($\Sigma = \gamma^2 I, \gamma = 1/100$), y los parámetros se asumieron uniformemente distribuidos en $(-2.5, 2.5)$. Se ejecutó el algoritmo MH con 50000 iteraciones, utilizando un núcleo Gaussiano ($\sigma^2 = 0.01$) y punto inicial en el origen.

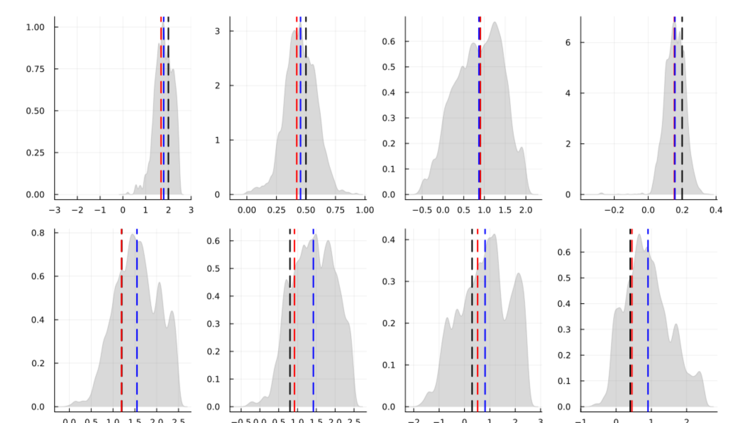

Figura 3: Histogramas obtenidos del proceso de calibración (negro: real, azul: CM, rojo: MAP).

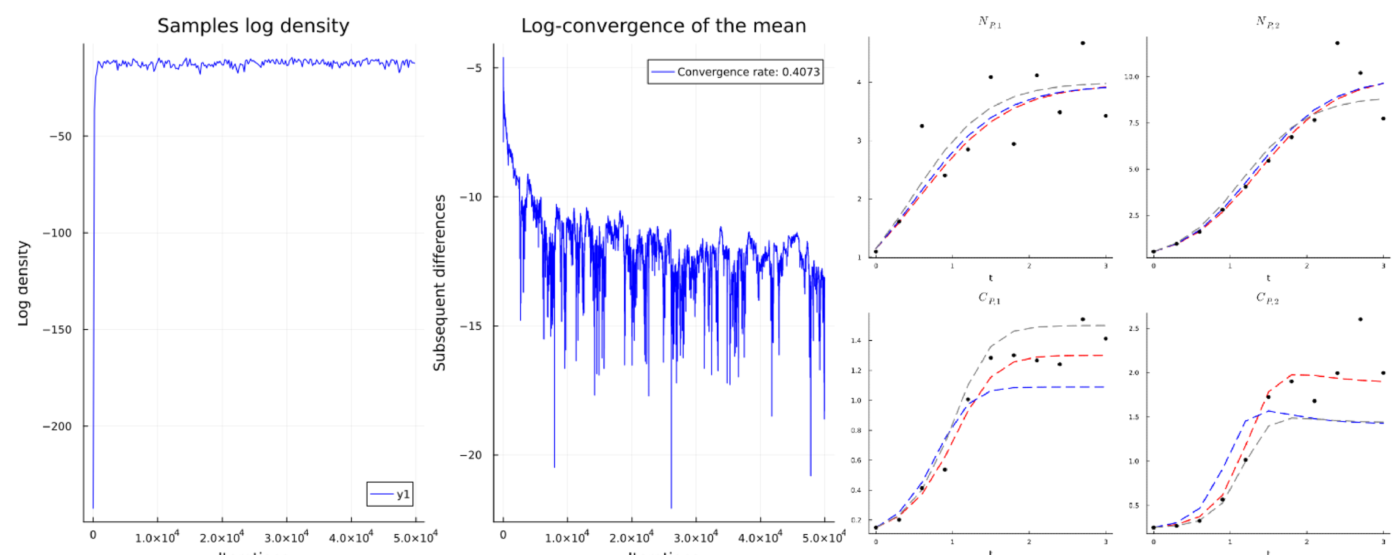

Figura 4: A la derecha, medidas de convergencia; a la izquierda, resultado de calibración en el modelo (negro: real, azul: CM, rojo: MAP).

Las figuras muestran que las estimaciones puntuales concuerdan con los valores iniciales, lo que indica una correcta ejecución del algoritmo. La tasa de aceptación de muestras fue del 70.29 %, y la tasa de convergencia de los parámetros fue de 0.4096. El modelo ajustó adecuadamente los datos, con valores de verosimilitud obtenidos de:

$$\frac{\pi_{\text{post}}(x_{\text{MAP}})}{\pi_{\text{post}}(x_{\text{Real}})} \approx 8.1575 \quad \text{y} \quad \frac{\pi_{\text{post}}(x_{\text{CM}})}{\pi_{\text{post}}(x_{\text{Real}})} \approx 3.1915$$

Conclusiones

Evaluación del Algoritmo: El algoritmo Metrópolis-Hastings es efectivo en la estimación de parámetros del modelo TOMGRO, con una tasa de aceptación adecuada. El enfoque bayesiano con MH estima con precisión los parámetros, logrando valores de verosimilitud cercanos a los reales. Aunque los resultados son prometedores, la correcta elección de distribuciones a priori y la eficiencia en la implementación del algoritmo son cruciales para optimizar el proceso de calibración y estimación.
Aplicaciones Futuras: La metodología puede extenderse a otros modelos de ecuaciones diferenciales, mejorando la predicción y caracterización de la incertidumbre. Se proyecta hacer la calibración del modelo con una mayor cantidad de datos reales.

Referencias

[1] J.W. Jones et al. "A dynamic tomato growth and yield model (TOMGRO)." English. En: *Transactions of the ASAE* 34.2 (1991), págs. 663-672. ISSN: 0001-2351.

[2] Jari Kaipio y E. Somersalo. *Statistical and computational inverse problems.* en. Springer Science & Business Media, mar. de 2006. ISBN: 9780387271323.

[3] Luke Tierney. "Markov chains for exploring posterior distributions". En: *The Annals of Statistics* 22.4 (dic. de 1994). ISSN: 0090-5364. DOI: 10.1214/aos/1176325750. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-22/issue-4/Markov-Chains-for-Exploring-Posterior-Distributions/10.1214/aos/1176325750.full (visitado 08-01-2025).