

Calibración del modelo TOMGRO aplicado a condiciones de la ciudad de Bogotá

Carlos Enrique Nosa Guzman

Tercera conferencia colombiana de Matemáticas aplicadas e industriales - MAPI 3

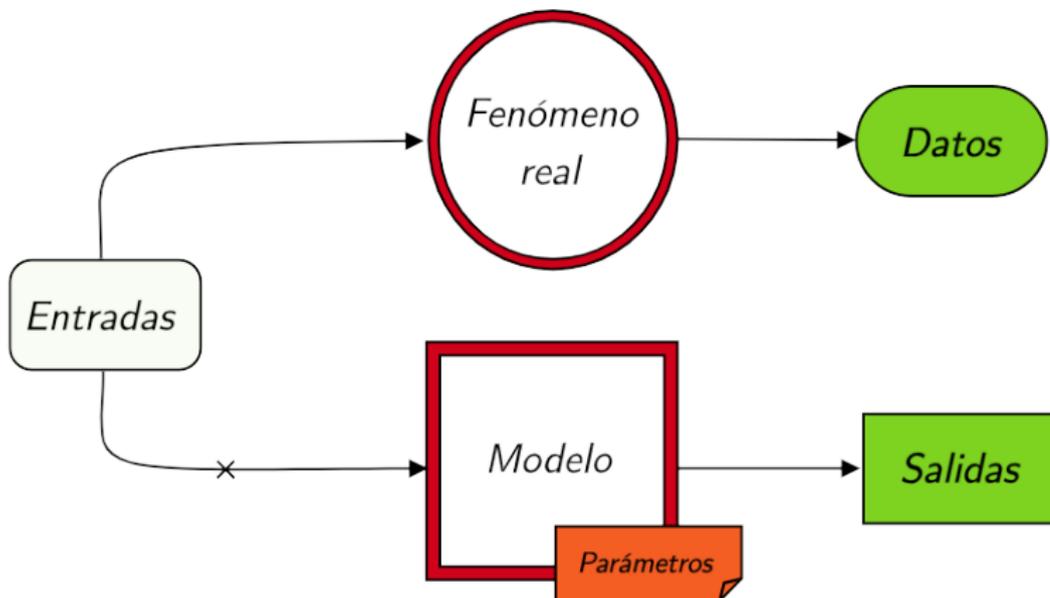
13 de junio del 2024

Contenidos

1 Introducción
2 Modelo

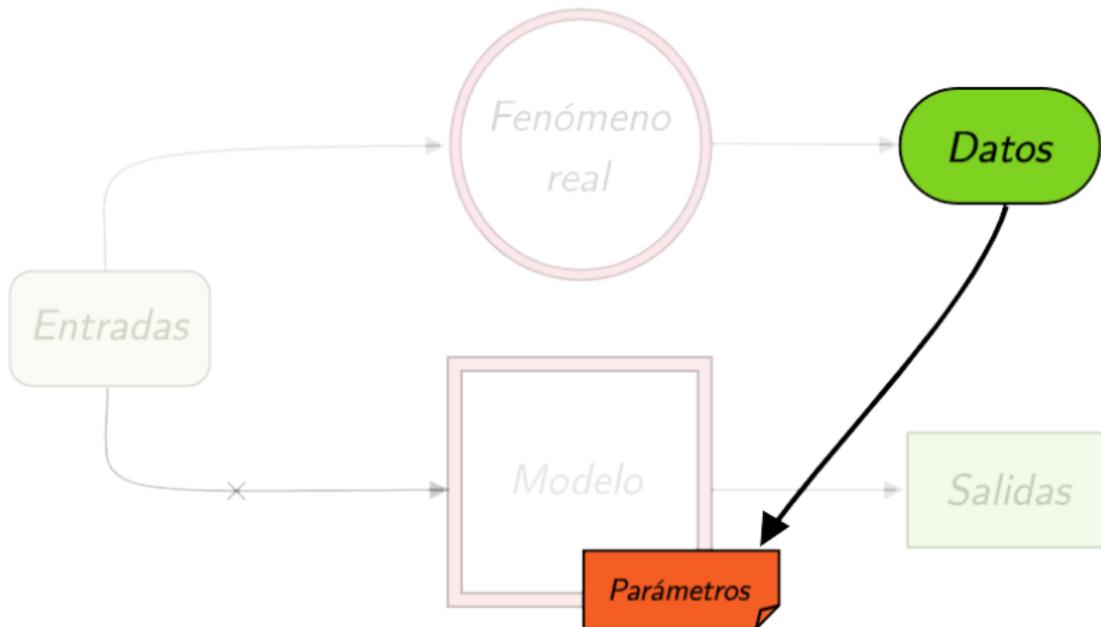
3 Calibración
4 Avances

Introducción



Problema inverso / Calibración

¿Cómo conocer el valor de los parámetros?



Modelo TOMGRO



Figura: Camas de siembra. Invernadero en el campus de la UN en Bogotá. Mayo 17.

- Matemáticamente, el modelo es un sistema dinámico de la forma:

$$\frac{dy}{dt} = F(C, \theta, y), \quad y(t_0) = y_0.$$

- El modelo utiliza un enfoque oferta-demanda para repartir los carbohidratos en el crecimiento de diferentes órganos.
- Es usado para crear estrategias óptimas de cultivo.

Estructura del modelo

- Como variables de estado se tienen la cantidad de órganos (tallos, hojas y frutos), su peso en seco y el área foliar.
- Los parámetros describen propiedades que **identifican** a cada variedad o híbrido de la planta dadas unas condiciones ambientales.



Figura: Tomates en maduración. Invernadero en el campus de la UN en Bogotá. Mayo 24.

Estructura del modelo

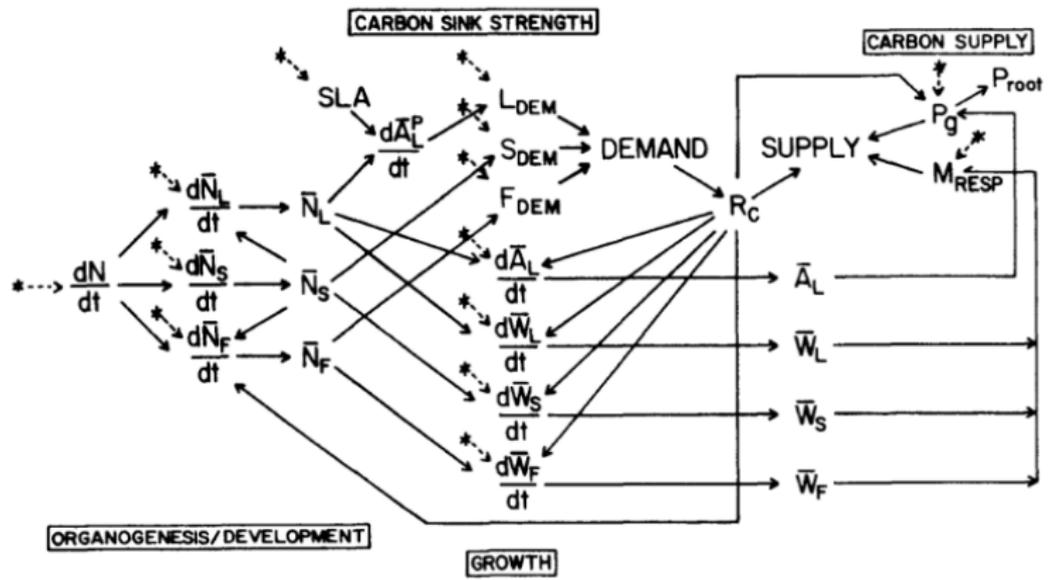


Figura: Relación entre variables del modelo. Tomado de [3].

Estructura del modelo

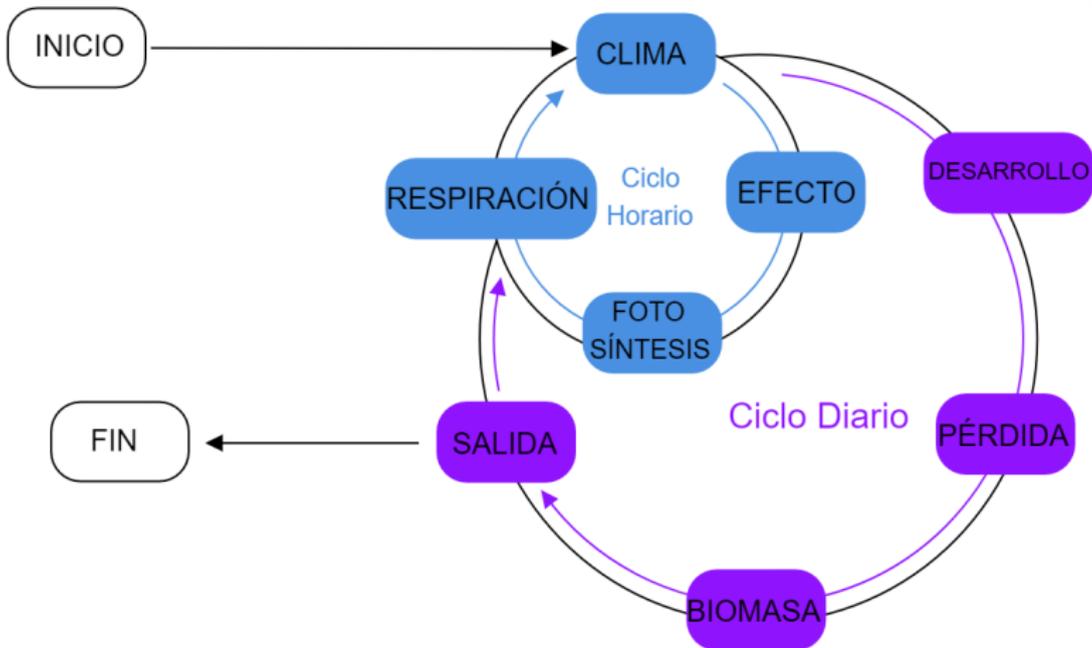


Figura: Ciclos usados para el cambio de información. Realización propia.

Datos de entrada

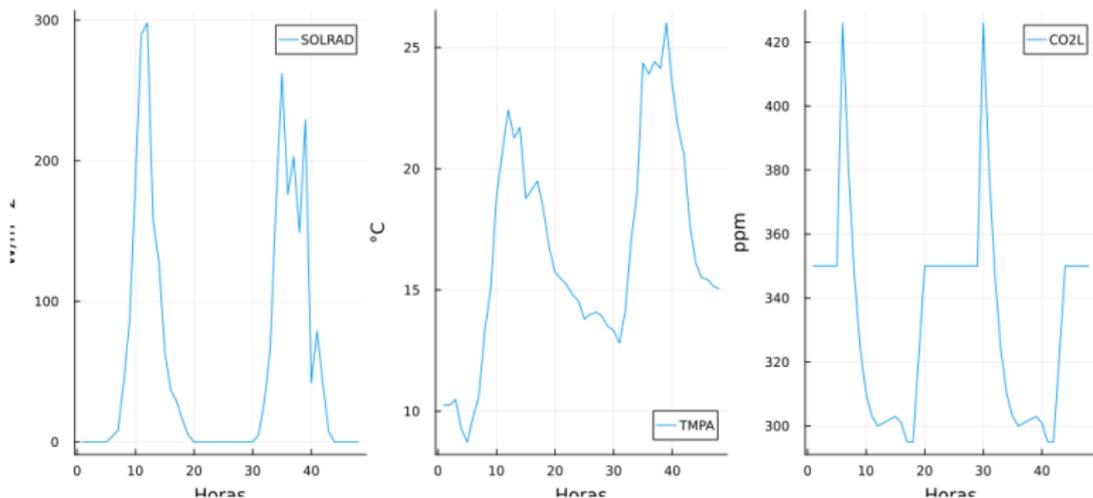


Figura: Curvas de variables climáticas que se toman como datos de entrada del modelo TOMGRO.

Algunos resultados del TOMGRO

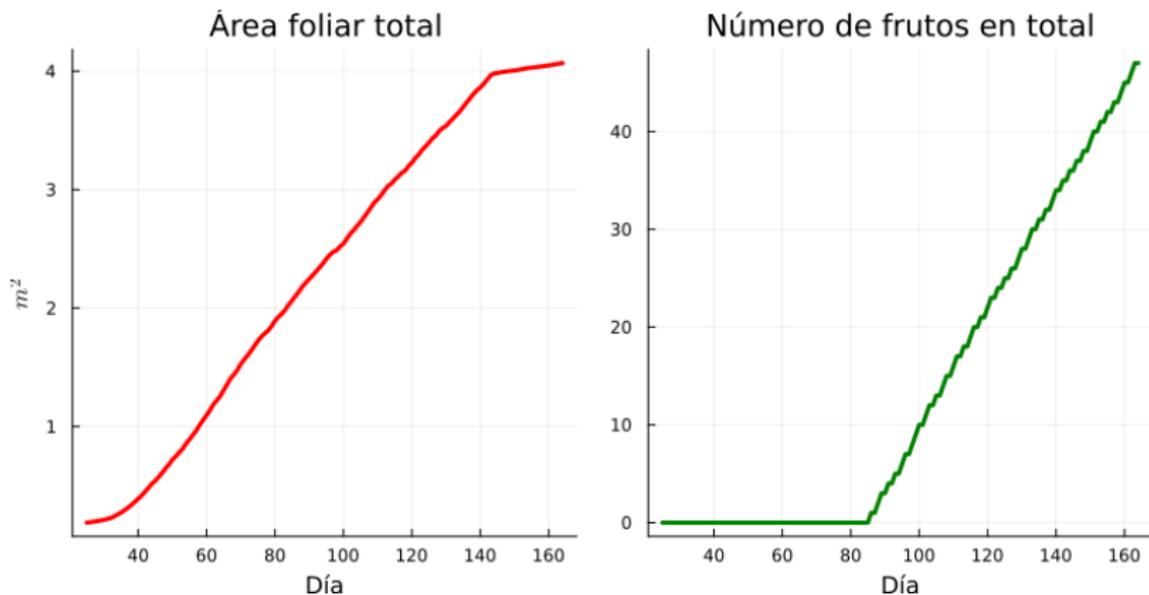


Figura: Simulación de área foliar total y número de frutos. Realización propia.

Ejemplo

Para ilustrar el proceso de calibración se recurre a un modelo ejemplo que busca estudiar el comportamiento del índice de área foliar (LAI).

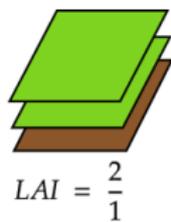


Figura: Definición LAI.

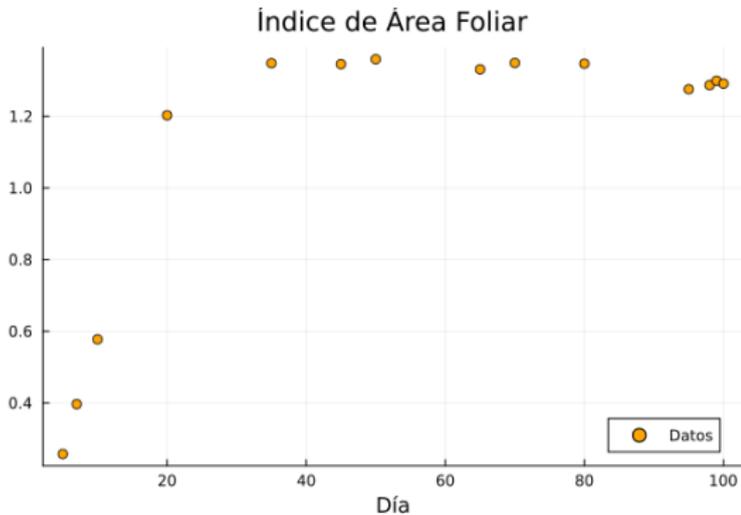


Figura: Datos simulados del LAI.

Ejemplo de modelo para el índice de área foliar

Sea y la variable que representa el índice de área foliar y y_0 el valor de éste índice en el tiempo inicial.

$$\frac{dy}{dt} = ry \left(1 - \frac{y}{K} \right), \quad y(0) = y_0,$$

donde r representa la razón de crecimiento y K la capacidad máxima que tiene la planta para modificar tal índice.

- r y K son los parámetros del modelo. En principio se pueden asumir como **constantes**, sin embargo, se supone que estos parámetros sean **funciones** de condiciones ambientales.

Modelación del problema desde el enfoque bayesiano

Sea $X \in \mathbb{R}^m$ el vector aleatorio de parámetros del modelo, $Y \in \mathbb{R}^n$ los datos y $\pi_{X,Y}(x, y)$ la densidad de probabilidad conjunta.

- Distribución **a priori** del vector de parámetros X .

$$\pi_X(x) = \int_{\mathbb{R}^n} \pi_{X,Y}(x, y) dy.$$

- La distribución de probabilidad condicional o función de **verosimilitud**.

$$\pi_{Y|X}(y|x) = \frac{\pi_{X,Y}(x, y)}{\pi_X(x)}.$$

- Si se tiene que $Y = y_{obs}$, la distribución de probabilidad condicional

$$\pi_{X|Y}(x|y_{obs}),$$

se llama la distribución **a posteriori** de X .

Modelación del problema desde el enfoque bayesiano

- Por el Teorema de Bayes,

$$\pi_{X|Y}(X|y_{obs}) = \frac{\pi_{Y|X}(y_{obs}|X)\pi_X(X)}{\pi_Y(y_{obs})}.$$

- Algunas de las siguientes estadísticas son de relevancia:

$$\underbrace{x_{MAP} = \max_{x \in \mathbb{R}^m} \{\pi_{X|Y}(x|y_{obs})\}}_{\text{Máximo a posteriori}} \quad \text{y} \quad \underbrace{x_{CM} = E[\pi_{X|Y}(x|y_{obs})]}_{\text{Media condicional}}.$$

- Se **explora** la distribución a posteriori $\pi_{X|Y}(x|y_{obs})$.

Algoritmo Metropolis Hastings Random Walk (MHRW)

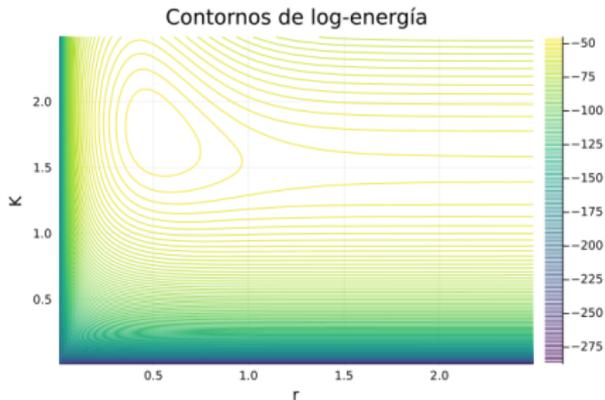


Figura: Contornos de la función de logenergía. Realización propia.

- El algoritmo MHRW es del tipo **Markov Chain Monte Carlo (MCMC)**.
- Tiene como objetivo muestrear la distribución a posteriori usando el Teorema de Bayes:

$$\underbrace{\pi_{X|Y}(X|y_{obs})}_{A \text{ posteriori}} \propto \underbrace{\pi_{Y|X}(y_{obs}|X)}_{\text{Verosimilitud}} \underbrace{\pi_X(X)}_{A \text{ priori}}.$$

Energía

Algoritmo MHRW

Datos: Punto inicial x_0 , número de muestras $N \in \mathbb{N}$ y función de energía E_n .

Resultado: Muestras $\{x_0, \dots, x_N\}$ de la distribución a posteriori.

Proceso: Para $i = 0, 1, \dots, N - 1$:

- Hacer una perturbación a la muestra $x_i + \epsilon$. Construcción del **camino aleatorio**.
- Comparar la energía de la muestra x_i con la muestra perturbada $x_i + \epsilon$:

$$\alpha = \min \left\{ 1, \frac{E_n(x_i + \epsilon)}{E_n(x_i)} \right\}.$$

- Aceptar la muestra perturbada con una probabilidad de α .

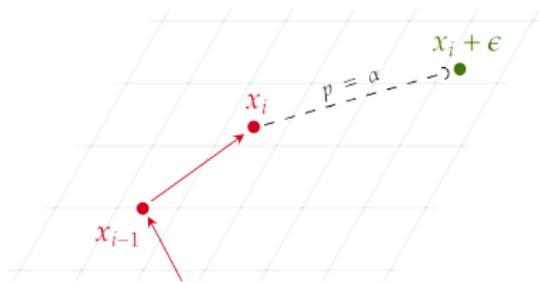


Figura: Camino aleatorio.

Algoritmo MHRW

Resultados

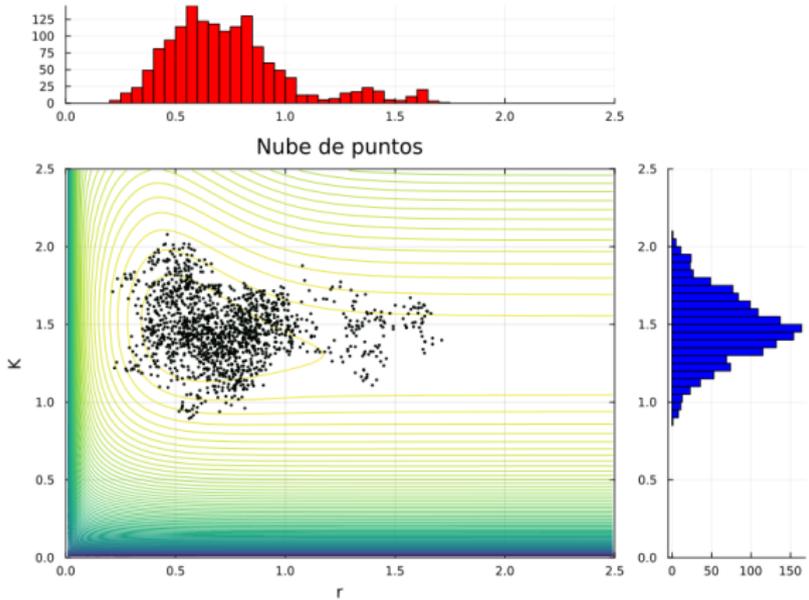


Figura: Resultados gráficos de la calibración de parámetros.

Algoritmo MHRW

Resultados

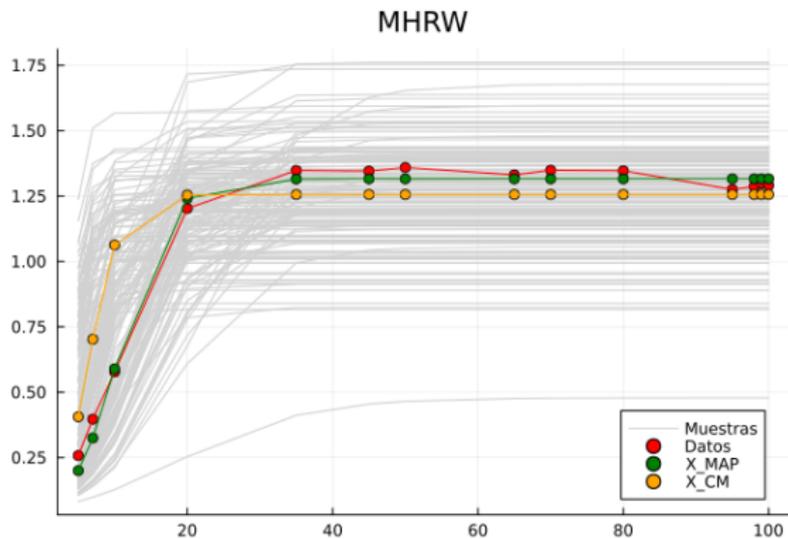


Figura: Resultados gráficos de la calibración de parámetros.

Datos

Desde el transplante hasta el momento

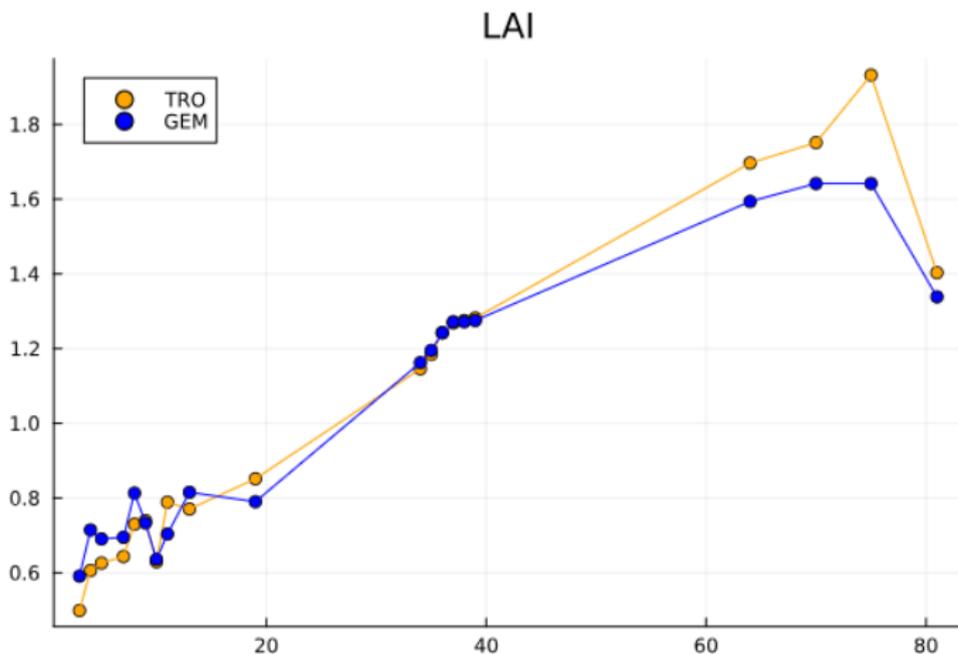


Figura: Índice de área foliar para dos híbridos.

Datos

Desde el trasplante hasta el momento

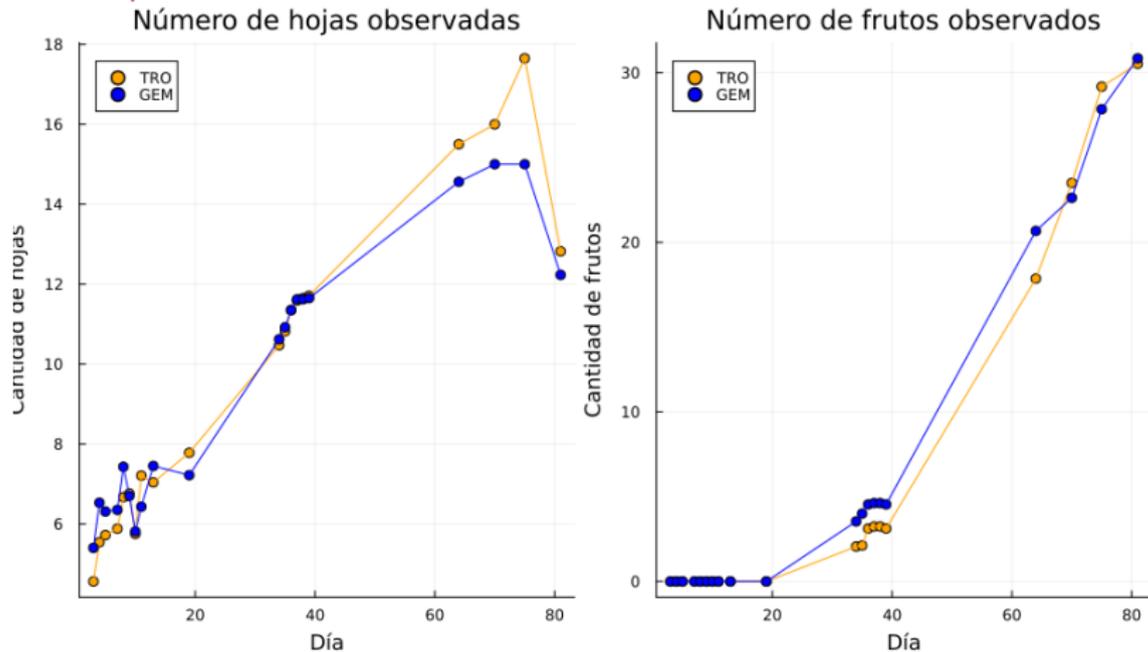


Figura: Datos recogidos del número de hojas y del número de frutos para diferentes variedades.

Problemas de calibración

- **Elección** de parámetros de interés para realizar la calibración.
- Análisis de **sensibilidad** de los resultados del modelo para los parámetros de interés.
- **Escasez** de datos de entrada que requiere el modelo.
- **Ausencia** de datos (series de tiempo) de las diferentes variables de estado.
- El periodo de recolección de datos no ha sido finalizado.

Posible trabajo futuro

- Representación de manera digital del desarrollo y crecimiento de una planta de tomate (**gemelo digital**).
- Una vez finalizada la calibración del modelo TOMGRO para diferentes variedades o híbridos de la planta, se puede buscar cómo **optimizar** variables referentes a la planta por medio del **control** de condiciones ambientales (*temperatura, radiación solar, concentración de dióxido de carbono*).

Referencias

- [1] Cooman, A. (2002). Feasibility of protected tomato cropping in the high altitude tropics using statistical and system dynamic models for plant growth and development [Doctoral thesis]. Katholieke Universiteit Leuven.
- [2] Galvis, J., Capistrán, M. (2021). Beyond Research: Asimilación de datos en sistemas complejos [Curso].
- [3] J. W. Jones, E. Dayan, L. H. Allen, H. Van Keulen, & H. Challa. (1991). A dynamic tomato growth and yield model(Tomgro). Transactions of the ASAE, 34(2), 0663-0672.
<https://doi.org/10.13031/2013.31715>
- [4] Kaipio, J., Somersalo, E. (2006). Statistical and computational inverse problems. Springer Science Business Media.