



Universidad Nacional de San Juan
Facultad de Ingeniería
Instituto de Automática

Interpretabilidad e identificabilidad paramétrica en el modelado de procesos biotecnológicos

**Tesis Doctoral de
Estefanía AGUIRRE ZAPATA**

Interpretabilidad e identificabilidad paramétrica en el modelado de procesos biotecnológicos

Tesis doctoral de

Estefanía Aguirre Zapata

Aprobada por la
Facultad de Ingeniería de la Universidad Nacional de San Juan
para el otorgamiento del grado académico de

DOCTOR EN INGENIERÍA DE SISTEMAS DE CONTROL

Dirección de tesis: **Dra. Ing. Adriana N. Amicarelli**
(Instituto Automática, UNSJ)
Dr. Ing. Fernando di Sciascio
(Instituto Automática, UNSJ)
Dr. Ing. Hernán Darío Álvarez
(Universidad Nacional de Colombia, UNAL)

Jurado de tesis: **PhD. Eva Balsa-Canto**
(Instituto de Ciencias Marinas, UVigo)
PhD. Diego Alejandro Muñoz
(Universidad Nacional de Colombia)

Fecha de defensa oral: 23 Agosto 2024

ISBN 978-631-90150-7-2



Aguirre-Zapata, Estefania

Interpretabilidad e identificabilidad paramétrica en el modelado de procesos biotecnológicos / Estefania Aguirre-Zapata. - 1a ed. - San Juan : Instituto de Automática, 2024.

ISBN 978-631-90150-7-2

CDD 660.6

ESTEFANIA AGUIRRE ZAPATA

**INTERPRETABILIDAD E IDENTIFICABILIDAD PARAMÉTRICA EN EL
MODELADO DE PROCESOS BIOTECNOLÓGICOS**

Tesis de doctorado

Temas de automática

Primera edición: Agosto 2024

Impreso en Argentina

Queda prohibida la reproducción total o parcial de esta tesis, su almacenamiento o transmisión por cualquier medio, ya sea electrónico, mecánico, fotocopiado o de cualquier otra índole, sin la autorización previa y por escrito del autor o de la institución a la que pertenece.

ISBN: ISBN 978-631-90150-7-2

CCD: 660.6

©2024 ESTEFANIA AGUIRRE ZAPATA

Hecho el depósito que marca la ley.

That was a memorable day to me, for it made great changes in me. But it is the same with any life. Imagine one selected day struck out of it, and think how different its course would have been. Pause you who read this, and think for a moment of the long chain of iron or gold, of thorns or flowers, that would never have bound you, but for the formation of the first link on one memorable day.

Charles Dickens, Great expectations, 1861.

Dedicatoria

A mi tío Mario de Jesús Aguirre que depositó su confianza en mí mientras vivió, ofreciéndome la oportunidad única de formarme como ingeniera y magíster. A él que creyó en mí al emprender esta nueva aventura.

A mi madre Girlesa Zapata que lo ha sacrificado todo por mí. A ella que tuvo la fuerza en su corazón para dejarme volar alto y lejos, a 6.500km de casa, pero a menos de un milímetro de su corazón. ¡Te siento cerca siempre!

Agradecimientos

Agradezco que las tesis tengan un espacio para agradecer, porque sin lugar a dudas la experiencia de investigar e indagar sobre lo desconocido, implica la coincidencia de voluntades, recursos y personas. He sido muy afortunada por encontrar lejos de casa una familia que trasciende la consanguinidad, las culturas y las costumbres. Agradezco al INAUT, a cada integrante de este lugar que me abrigó como una casa, como una familia. Agradezco a la Dra. Adriana Amicarelli por depositar su confianza en mí, por recibirme con los brazos abiertos y dirigirme con amor y libertad durante mi doctorado. Agradezco al Dr. Fernando di Sciascio por ser siempre la palabra sincera, la mirada crítica y la palmadita en la espalda que me recordaba que lo estaba haciendo bien. Agradezco al Dr. Hernán Álvarez por su humanidad tangible, por su bondad para enseñar y guiar, por su confianza y respeto por mí, no solo como profesional, sino también como persona, y por su capacidad para escucharme y motivarme.

Agradezco a esos pequeños ángeles que me rodearon estando lejos de casa. A mis compañeros en esta carrera: Humber, Sebas y Carlos, gracias por ser mi primera familia en San Juan, mis camaradas, mis colegas, los que me recordaron que esta vez valía la pena estar lejos de casa. A mis primeros amigos: Juanchi, Fer, Chris, Vivi y Omar, gracias infinitas por alegrar mi vida con su amistad, por cada asado, cada cumpleaños, por pasar la pandemia conmigo y por cada locura juntos ¡Fue maravilloso tenerlos!. A los que vinieron después: José, Brayan y Rodri, gracias por acompañarme en la recta final. A Daniel, gracias por ser mi cómplice y mi amigo, gracias por hacerme sentir cerca de casa. Llevaré conmigo cada viaje, cada experiencia, cada recuerdo que construimos juntos, ¡Soy muy afortunada por coincidir contigo!

Agradezco a mi familia: Andre, Mariana, Iván, Jaime, Mamá y a mi viejita Titín, gracias por el sacrificio de separarnos, tienen que saber que su amor y sus oraciones siempre nos mantuvieron cerca. Agradezco a mis amigos Pini y Hugo por estar siempre y no rendirse con esta amistad en la distancia. ¡Fueron mi refugio muchas veces!

Gracias mi buen Dios por fijar sobre mi tus ojos y enseñarme el camino por el que debo andar. **Salmos 32:8**

RESUMEN

Interpretabilidad e identificabilidad paramétrica en el modelado de procesos biotecnológicos

Estefanía Aguirre-Zapata

Esta tesis doctoral aborda el desafío de la interpretabilidad en la construcción de modelos matemáticos útiles para la comprensión y optimización de sistemas complejos. La motivación central de este trabajo radica en la necesidad de establecer un marco conceptual común en todo el espectro de modelado, así como mejorar la interpretabilidad de los modelos matemáticos, especialmente en el contexto de modelos semi-físicos basados en primeros principios (MSBPP). En ella se destaca la importancia de la interpretación física de los modelos, particularmente en áreas como la biología, biotecnología y procesos agroindustriales; partiendo de la dificultad para establecer límites claros en la búsqueda de ecuaciones constitutivas durante la construcción de este tipo de modelos, a su vez que se mantiene el grado de compromiso entre la precisión del ajuste con la interpretabilidad del modelo.

En respuesta a este desafío, en este trabajo se propone un nuevo marco conceptual para abordar la interpretabilidad en el espectro de modelado matemático, y se propone un índice matemático para cuantificar la interpretabilidad en MSBPP. Además, se amplía la metodología de modelado existente al integrar la interpretabilidad como un criterio adicional en la determinación del nivel de especificación durante la búsqueda de ecuaciones constitutivas. La utilidad del índice y de la metodología propuesta se evalúa en tres aplicaciones que varían en complejidad y relevancia práctica, esto es: un reactor continuamente agitado (CSTR), un modelo de regulación de glucosa en el cuerpo humano, y un modelo de crecimiento de la polilla de la vid (*Lobesia botrana*) particularizado para la región de Cuyo, Argentina.

Las contribuciones principales incluyen el nuevo marco conceptual, el índice de interpretabilidad que acompaña la metodología para la deducción de MSBPP, y el desarrollo de un modelo específico para el proceso crecimiento y voltinismo de *Lobesia botrana* a partir del uso de la metodología propuesta.

ABSTRACT

Interpretability and Parametric Identifiability in the Modeling of Biotechnological Processes.

Estefanía Aguirre-Zapata

This doctoral thesis addresses the challenge of interpretability in the construction of mathematical models useful for the understanding and optimization of complex systems. The central motivation of this work lies in the need to establish a common conceptual framework across the spectrum of modeling and to enhance the interpretability of mathematical models, especially in the context of first-principles-based semi-physical models (FPBSM). It highlights the importance of the physical interpretation of models, particularly in areas such as biology, biotechnology, and agro-industrial processes. The difficulty in establishing clear boundaries in the search for constitutive equations during the construction of such models is emphasized, along with maintaining a balance between the precision of fitting and the interpretability of the model.

In response to this challenge, this work proposes a new conceptual framework to address interpretability in the spectrum of mathematical modeling and introduces a specific index to quantify interpretability in MSBPP models. Furthermore, it extends the existing modeling methodology by integrating interpretability as an additional criterion in determining the specification level when searching for constitutive equations. The usefulness of the index and the proposed methodology is evaluated in three applications that differ in complexity and practical relevance: a continuously stirred reactor (CSTR), a model of glucose regulation in the human body, and a model of grape berry moth (*Lobesia botrana*) growth specifically for the Cuyo region, Argentina.

The main contributions include the conceptual framework, the interpretability index, and the development of a specific model for the growth and voltinism process of *Lobesia botrana* based on the proposed methodology.

ÍNDICE GENERAL

Tabla de contenido	IV
Lista de figuras	x
Lista de tablas	xiii
Lista de acrónimos	xv
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	4
1.2.1 Objetivo general	4
1.2.2 Objetivos específicos	5
1.3 Aportes y limitaciones	5
1.4 Difusión de los aportes de investigación	6
1.4.1 Publicaciones en revistas indexadas	6
1.4.2 Participación en congresos nacionales	7
1.4.3 Proyectos	8
1.4.4 Colaboración	8
1.5 Estructura de la tesis	10
2 Modelado matemático	11
2.1 Aspectos básicos de teoría de sistemas	12
2.1.1 Proceso	12
2.1.2 Sistema	12
2.1.3 Sistema dinámico	13
2.1.4 Tipos de sistemas	13

2.1.5	Análisis de sistemas para modelado	15
2.2	Modelos matemáticos	15
2.2.1	Generalidades	15
2.2.2	Modelado matemático	16
2.2.3	Clasificación de los modelos matemáticos	16
2.3	Modelos semi-físicos basados en primeros principios . .	20
2.3.1	Método de deducción de un MSBPP	20
2.3.2	Metodología general para la deducción de modelos matemáticos	27
2.3.3	Elementos de un MSBPP	29
3	Identificabilidad y sensibilidad	33
3.1	Identificabilidad paramétrica	34
3.1.1	Identificabilidad en el modelado de procesos . . .	34
3.1.2	Identificabilidad estructural	35
3.1.3	Identificabilidad práctica	37
3.1.4	Métodos para el análisis de identificabilidad . . .	37
3.2	Sensibilidad paramétrica	40
4	Interpretabilidad de modelos	45
4.1	Estado del arte en interpretabilidad	46
4.1.1	Concepto de interpretabilidad	46
4.1.2	¿Por qué es importante la interpretabilidad? . . .	50
4.2	Un nuevo marco conceptual	51
4.2.1	Etimología y fundamento filosófico	51
4.2.2	Interpretabilidad	54
4.2.3	Interpretabilidad de parámetros estructurales . .	54
4.2.4	Interpretabilidad de parámetros funcionales . . .	55
4.3	Taxonomía de interpretabilidad	56
4.4	Índice de interpretabilidad	58
4.4.1	Escenario y grafo de modelado	59

4.4.2	Nivel de especificación	61
4.4.3	Criterio de interpretabilidad paramétrica	64
4.4.4	Coeficiente de sensibilidad	64
5	Metodología propuesta	67
5.1	Metodología propuesta	68
5.1.1	Enunciado del problema de modelado	68
5.1.2	Deducción de la estructura básica del modelo	70
5.1.3	Deducción de la estructura extendida del modelo	70
5.1.4	Reporte e interpretación de los resultados	71
5.2	Ejemplos de aplicación	75
5.2.1	Caso de estudio: CSTR	75
5.2.2	Caso de estudio: modelo mínimo de glucosa	80
5.3	Análisis de interpretabilidad	86
5.3.1	Análisis de interpretabilidad del reactor continuo de tanque agitado (CSTR)	86
5.3.2	Análisis de interpretabilidad del modelo mínimo de glucosa	99
6	Caso de estudio: <i>Lobesia botrana</i>	115
6.1	<i>Lobesia botrana</i>	116
6.1.1	Generalidades	116
6.1.2	<i>Lobesia botrana</i> en la literatura	117
6.2	Modelo de crecimiento de <i>L. botrana</i>	121
6.2.1	Etapas 1. Problema de modelado	121
6.2.2	Etapas 2. Deducción de la estructura básica	126
6.2.3	Etapas 3. Deducción de la estructura extendida	128
6.3	Voltinismo de <i>Lobesia botrana</i> : una extensión del modelo de crecimiento	158
6.3.1	Concepto de voltinismo en <i>Lobesia botrana</i>	158
6.3.2	Área de estudio	161

6.3.3	Datos fenológicos de <i>Lobesia botrana</i>	162
6.3.4	Datos meteorológicos	165
6.3.5	Extensión del modelo de crecimiento	166
6.3.6	Propiedades del modelo	171
6.3.7	Ajuste y validación	172
7	Conclusiones y trabajo futuro	183
7.1	Conclusiones	183
7.1.1	Aporte teórico	183
7.1.2	Aplicación práctica a la región	186
7.2	Trabajo futuro	187
A	Condiciones iniciales	189
A.1	Condiciones iniciales para el CSTR	189
A.2	Condiciones iniciales para el modelo mínimo de glucosa	189
A.3	Condiciones iniciales para el modelo de crecimiento y voltinismo de <i>L. botrana</i>	191
	Bibliografía	204

ÍNDICE DE FIGURAS

2.1	Espectro de modelado de Karplus y sus aplicaciones potenciales. Adaptado de (Karplus, 1977).	18
2.2	Metodología general para construir modelos de caja blanca y modelos semi-físicos basados en primeros principios (MSBPP).	21
2.3	Metodología general para construir modelos de caja negra y modelos semi-físicos de base empírica.	28
3.1	Cuadro de identificabilidad para el análisis de identificabilidad estructural. Adaptada de (Balsa-Canto et al., 2010)	38
4.1	Relación significado-significante. Adaptada de (Ogden & Richards, 1923)	52
4.2	Interpretabilidad de los modelos matemáticos. ISP: Parámetro estructural interpretable. NISP: Parámetro estructural no interpretable. GPI: Interpretabilidad paramétrica general. IPI: Interpretabilidad paramétrica heredada. PI-PI: Interpretabilidad Paramétrica Parcialmente Heredada. NIP: Parámetro funcional no interpretable.	53
4.3	Escenario de modelado en contraste con los niveles de especificación (ejemplo).	60
4.4	Grafo de modelado para el análisis de interpretabilidad de un MSBPP.	61
4.5	Representación gráfica de los niveles de especificación en MSBPP.	63
5.1	Metodología para el cálculo del índice de interpretabilidad de un MSBPP.	69
5.2	Tabla de síntesis de resultados cuantitativos de los análisis realizados para cada escenario de modelado.	71
5.3	Tabla comparativa de resultados cuantitativos para cada escenario de modelado.	72

5.4	Gráfico de radar para la evaluación del grado de influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado.	73
5.5	Gráfica de jerarquización paramétrica con base en el índice de interpretabilidad y el coeficiente de sensibilidad de cada parámetro.	74
5.6	Grafo de modelado del CSTR.	76
5.7	Grafo de modelado del modelo mínimo de glucosa.	81
5.8	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 0. .	87
5.9	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 1. .	89
5.10	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 2. .	92
5.11	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 3. .	93
5.12	Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del CSTR.	95
5.13	Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del CSTR.	97
5.14	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 0.	100
5.15	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 1.	103
5.16	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 2.	105

5.17	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 3.	107
5.18	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 4.	109
5.19	Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelos mínimos de glucosa.	111
5.20	Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado mínima de glucosa.	112
6.1	Ilustración del proceso de crecimiento de <i>Lobesia botrana</i> para una generación dada. Fuente propia (Aguirre-Zapata et al., 2022b)	121
6.2	Curvas de crecimiento de <i>Lobesia botrana</i> en laboratorio bajo condiciones controladas de laboratorio. Cada curva corresponde a una corrida experimental diferente conservando las condiciones de temperatura, humedad y foto periodo.	125
6.3	Diagrama de bloques de los sistemas de proceso considerados para modelar el proceso de crecimiento de <i>Lobesia botrana</i>	126
6.4	Grafo de modelado del proceso de crecimiento de <i>Lobesia botrana</i>	130
6.5	Perfiles de mortalidad diseñados a partir de los datos experimentales.	133
6.6	Tasas de mortalidad de <i>Lobesia botrana</i> estimadas mediante el EKF comparadas con las tasas de mortalidad construidas con las Ecuaciones 6.14-6.15.	139
6.7	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 0.	140

6.8	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 1.	142
6.9	<i>Lobesia botrana</i> curvas de crecimiento a $20^{\circ}C \pm 2^{\circ}C$ para el escenario 1.	144
6.10	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 2.	147
6.11	<i>Lobesia botrana</i> curvas de crecimiento a $20^{\circ}C \pm 2^{\circ}C$ para el escenario 2.	149
6.12	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 3.	151
6.13	Curvas de crecimiento de <i>Lobesia botrana</i> a $20^{\circ}C \pm 2^{\circ}C$ para el escenario 3.	153
6.14	Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del proceso de crecimiento de <i>Lobesia botrana</i>	154
6.15	Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado de <i>Lobesia botrana</i>	155
6.16	Comparación del voltinismo de <i>Lobesia botrana</i> en los hemisferios norte y sur. La línea roja muestra el voltinismo de <i>Lobesia botrana</i> en España, según (Martín-Vertedor et al., 2010). La línea negra representa el voltinismo de <i>Lobesia botrana</i> en la región de Cuyo, Argentina (fuente: propia).	159
6.17	Distribución de la producción de uva de mesa y de vinificación en la región de Cuyo, Argentina. Datos porcentuales del Informe Anual de Superficie Cultivada, publicado por Instituto Nacional de Vitivinicultura, 2022.	161
6.18	Trampas delta con difusor de feromona para la captura de machos de <i>Lobesia botrana</i> (Fuente propia).	163

6.19	Distribución de trampas delta en el departamento Pocito de la provincia de San Juan. Fuente: tramite.sanjuan.gov.ar/index.php/tramite/org/207	164
6.20	Datos meteorológicos de la estación meteorológica del departamento de Pocito, en San Juan - Argentina, correspondientes a la temporada 2020 – 2022.	165
6.21	Diagrama de bloque de los sistemas de procesos considerados para modelar el proceso de crecimiento y voltinismo de <i>Lobesia botrana</i>	166
6.22	Diagrama de bloque de los sistemas de procesos considerados para modelar la función de tasa de fecundidad de <i>Lobesia botrana</i>	169
6.23	Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de voltinismo de <i>Lobesia botrana</i>	171
6.24	Datos meteorológicos de la estación meteorológica de Pocito, en la región argentina de Cuyo, para la temporada 2020 – 2021.	175
6.25	Respuesta del modelo de voltinismo de <i>Lobesia botrana</i> , en contraste con los datos de captura de machos en campo en la región de Cuyo en Argentina durante la temporada 2020 – 2021.	176
6.26	Respuesta del modelo de voltinismo de <i>Lobesia botrana</i> para todos los estadios de crecimiento de la polilla (huevo, larva, pupa y adulto) en el departamento de Pocito en la región de Cuyo en Argentina durante la temporada 2020 – 2021.	177
6.27	Datos meteorológicos de la estación meteorológica de Pocito, en la región argentina de Cuyo, correspondientes a la temporada 2021 – 2022.	178
6.28	Respuesta del modelo de voltinismo de <i>Lobesia botrana</i> , en contraste con los datos de captura de machos en campo en la región de Cuyo en Argentina.	179
6.29	Respuesta del modelo de voltinismo de <i>Lobesia botrana</i> , en contraste con los datos de captura de machos en campo para el departamento de Pocito en la región de Cuyo en Argentina durante la temporada 2021 – 2022.	180

ÍNDICE DE TABLAS

4.1	Taxonomía de interpretabilidad paramétrica en el espectro de modelado.	57
5.1	Parámetros del modelo matemático del CSTR utilizados en todos los escenarios considerados.	77
5.2	Parámetros del modelo mínimo considerados en todos los escenarios analizados.	82
5.3	Análisis de interpretabilidad de un modelo del CSTR para su estructura básica.	88
5.4	Análisis de interpretabilidad de un modelo del CSTR considerando los parámetros funcionales de las nuevas ecuaciones constitutivas del escenario 1.	90
5.5	Análisis de interpretabilidad de un modelo del CSTR calculando el coeficiente global de transferencia de calor (U) mediante la ecuación presentada en (Bondy et al., 1983).	91
5.6	Análisis de interpretabilidad de un modelo del CSTR calculando el coeficiente global de transferencia de calor (U) mediante una ecuación empírica.	94
5.7	Índice de interpretabilidad $IIWS_{L,M_s}$ para cada escenario de modelado del CSRT. El cambio en el valor del índice representa el grado de mejora (+) o degradación (−) del índice en comparación con el escenario de modelado correspondiente.	96
5.8	Identificabilidad de los parámetros desconocidos del modelo matemático de un CSTR, para todos los escenarios de modelado.	98
5.9	Análisis de interpretabilidad del modelo mínimo de glucosa considerando su estructura básica.	101
5.10	Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 1.	104
5.11	Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 2.	106

5.12	Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 3.	108
5.13	Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 4.	110
5.14	Índice de interpretabilidad IIS_{L,M_s} para cada escenario planteado para el modelo mínimo de glucosa. El cambio en el valor del índice representa el grado de mejora (+) o degradación (−) del índice en comparación con el escenario de modelado correspondiente.	113
5.15	Identificabilidad de los parámetros desconocidos del modelo matemático de un modelo mínimo de glucosa, para todos los escenarios de modelado.	114
6.1	Variables y constantes del modelo de crecimiento de <i>Lobesia botrana</i> a partir de la estructura básica.	128
6.2	Parámetros del modelo matemático del proceso de crecimiento de <i>Lobesia botrana</i> utilizados en todos los escenarios considerados.	129
6.3	Valores de los parámetros identificados por el método de mínimos cuadrados no lineales para cada cinética de crecimiento. El símbolo $\otimes$ indica que este parámetro no forma parte de la cinética correspondiente, y θ corresponde al parámetro.	135
6.4	Pseudocódigo de implementación del Filtro de Kalman Extendido	137
6.5	Análisis de interpretabilidad del modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 0.	141
6.6	Análisis de interpretabilidad del modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 1.	143
6.7	Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética de Monod, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$	145
6.8	Análisis de interpretabilidad del modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 2.	148

6.9	Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética exponencial, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$	150
6.10	Análisis de interpretabilidad del modelo de crecimiento de <i>Lobesia botrana</i> en el escenario 3.	150
6.11	Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética de Hill, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$	152
6.12	Índice de interpretabilidad $IIWS_{L,M_s}$ para cada escenario de modelado <i>Lobesia botrana</i>	156
6.13	Identificabilidad de los parámetros desconocidos del modelo de crecimiento de <i>Lobesia botrana</i> , para todos los escenarios de modelado.	157
6.14	Efecto de la edad de la uva en el tiempo de desarrollo de <i>Lobesia botrana</i> . Tomado de (Gutierrez et al., 2018).	167
6.15	Parámetros añadidos a la versión extendida del modelo de crecimiento y voltinismo de <i>Lobesia botrana</i>	170
6.16	Análisis de interpretabilidad del modelo de voltinismo de <i>Lobesia botrana</i>	173
6.17	Error absoluto medio del modelo de voltinismo de <i>Lobesia botrana</i> en la fase de identificación y validación.	174
6.18	Error de detección de la estación meteorológica para cada estación.	175
6.19	Parámetros y condiciones iniciales por estación del año para el modelo de voltinismo de <i>Lobesia botrana</i> particularizado para la región de Cuyo en Argentina.	177
A.1	Condiciones iniciales para el cálculo de sensibilidades del modelo del CSTR.	190
A.2	Condiciones iniciales para el cálculo de sensibilidades del modelo del modelo mínimo de glucosa.	190
A.3	Condiciones iniciales para el cálculo de sensibilidades del modelo del modelo de crecimiento y voltinismo de <i>L. botrana</i>	191

LISTA DE ACRÓNIMOS

ADD	Grados día acumulados
CFR	Régimen fluctuante actual
CSTR	Reactor de tanque continuamente agitado
DD	Grados día
DO	Denominación de origen
DSS	Sistemas de apoyo a la toma de decisiones
FFR	Régimen fluctuante futuro
FPBSM	First principles based Semi-physical models
GL	Grados de libertad
GPI	Interpretabilidad paramétrica general
IPI	Interpretabilidad paramétrica heredada
ISP	Parámetro estructural interpretable
LHS	Latin hypercube sampling
Lb	Lobesia botrana
MSBPP	Modelos semi-físicos basados en primeros principios
NIP	Parámetro funcional no interpretable
NISP	Parámetro estructural no interpretable
PBDM	Modelos demográficos de base fisiológica
PIPI	Interpretabilidad paramétrica parcialmente heredada
dSOGMM	Modelo mínimo de glucosa oral subcutánea triangular doble

INTRODUCCIÓN

1.1	Motivación	1
1.2	Objetivos	4
1.2.1	Objetivo general	4
1.2.2	Objetivos específicos	5
1.3	Aportes y limitaciones	5
1.4	Difusión de los aportes de investigación	6
1.4.1	Publicaciones en revistas indexadas	6
1.4.2	Participación en congresos nacionales	7
1.4.3	Proyectos	8
1.4.4	Colaboración	8
1.5	Estructura de la tesis	10

(...) No puedo concebir un auténtico científico sin esa fe profunda. La situación puede expresarse con una imagen: la ciencia sin religión es coja, la religión sin ciencia es ciega. **Albert Einstein, 1930**

1.1 Motivación

En el área de ingeniería de procesos, los modelos matemáticos son una representación que busca capturar ciertas características de un sistema o proceso para un propósito previamente establecido (generación de conocimiento, estimación, control, optimización, entre otros) (Hangos & Cameron, 2001). En algunos casos, la validez de los modelos matemáticos se encuentra supeditada a una región de aplicabilidad que depende del conjunto de datos experimentales usados para su validación (Hangos &

Cameron, 2001) y su utilidad (o uso final) al tipo de modelo utilizado. De esta manera, los modelos matemáticos pueden ser clasificados de acuerdo a su naturaleza como: modelos caja blanca o basados en primeros principios, modelos caja negra o empíricos, o modelos de caja gris o semi-físicos, que son el resultado de la combinación de los dos tipos anteriores.

Los modelos de caja blanca se caracterizan por poseer una estructura completa proveniente de balances de masa, energía y cantidad de movimiento, por lo que todos sus parámetros tienen un significado físico en el contexto de aplicación del modelo. Por otra parte, los modelos de caja negra se caracterizan por producir “descripciones de datos”, que al repetir o modificar el experimento pueden producir un modelo diferente cada vez. Es decir, son modelos matemáticos que describen a los datos experimentales, pero en la mayoría de los casos sus parámetros no tienen un significado físico en el contexto de aplicación del modelo matemático. Los modelos matemáticos de caja negra, deben su nombre al hecho de que se tiene poco o nulo conocimiento sobre los mecanismos internos del sistema o proceso que se está modelando, centrándose en la relación empírica entre las entradas y las salidas del sistema.

Por su parte, los modelos de caja gris, que constituyen una combinación de los dos tipos anteriores, pueden ser semi-físicos basados en primeros principios (MSBPP), que se caracterizan por poseer una estructura básica de caja blanca (obtenida a partir de balances) y parámetros funcionales de carácter empírico, o semi-físicos de base empírica que se caracterizan por poseer una estructura básica proveniente del ajuste de datos experimentales (modelos caja negra) y algunos parámetros con un significado físico (Bohlin, 2006; Hangos & Cameron, 2001; Lema-Perez et al., 2019b). Así pues, tanto los modelos basados en primeros principios, como los modelos empíricos requieren la construcción de una estructura básica como de una estructura extendida que, dependiendo del tipo de modelo requerirá (en mayor o menor medida) la identificación de parámetros, los cuales interfieren de manera directa sobre la respuesta del modelo.

De manera particular, la construcción de modelos MSBPP implica, entre

otras cosas, la definición de un objetivo de modelado del que dependerá en gran medida el planteamiento de la hipótesis de modelado, y la búsqueda de ecuaciones constitutivas (estructura extendida) del mismo (Alvarez et al., 2009; Lema-Perez et al., 2019b). Uno de los desafíos actuales en esta gran área del modelado, tiene que ver con la dificultad que implica para el modelador establecer un punto de parada que le permita determinar en que grado profundizar en la búsqueda de ecuaciones constitutivas o de valoración, manteniendo la relación de compromiso entre la precisión y la interpretabilidad del modelo matemático (Lema-Perez et al., 2019b). Cuando se habla del ajuste de un modelo matemático, existen numerosos índices de error que permiten cuantificarlo (Dorf & Bishop, 2011; Duarte-Mermoud & Prieto, 2004; Shuaib & Ahmed, 2014). Sin embargo, no existe un consenso respecto al concepto de interpretabilidad en los modelos matemáticos, y se carece de un mecanismo formal o de una herramienta matemática para cuantificar el grado de interpretabilidad inherente a un modelo matemático dado (Lipton, 2018).

En áreas de aplicación específicas de los modelos matemáticos, los usuarios finales suelen hacer hincapié en la importancia de que tanto la estructura del modelo como los parámetros tengan un significado físico dentro del contexto de aplicación. Este énfasis es especialmente significativo en ámbitos como la biología (centrada en tratamientos médicos), la biotecnología, la bioquímica o los procesos agroindustriales. En estos dominios, el modelo no sólo predice comportamientos sino que, a través de sus parámetros y variables, facilita la extracción de información importante para el diagnóstico, seguimiento y control de los respectivos procesos (Cousins & Riondato, 2019; Erro et al., 2013; Lou et al., 2013). En consecuencia, una de las preocupaciones predominantes en el modelado matemático de procesos y sistemas es la interpretabilidad de los resultados de los modelos matemáticos a través de sus parámetros y estructura. Esta preocupación se extiende a la comprensión de la relación entre la interpretabilidad de un modelo matemático y otras propiedades intrínsecas de los modelos, tales como la precisión, la identificabilidad, la sensibilidad de los parámetros y la usabilidad general del modelo. Por ejemplo, algunas aplicaciones puntuales en las que resulta importante el nivel de interpretabilidad del modelo matemático son: el proceso de crecimiento de *Lobesia botrana* (*Lb*) o “polilla de la vid” (Burgardta et al., 2010; da Silva et al., 2013; Gonza-

lez, 2010), el proceso de homeostasis de glucosa (Corbett et al., 2020), o algunos procesos de crecimiento bacteriano como es el caso de *Bacillus thuringiensis* (Bt) (Amicarelli et al., 2010).

En esta tesis doctoral se plantea un marco conceptual para la interpretabilidad de modelos matemáticos en el espectro de modelado y se propone un índice para la cuantificación de la misma. Además, se extiende la metodología de modelado matemático propuesta en (Alvarez et al., 2009) a través de la adición de la interpretabilidad y otras propiedades del modelo matemático, como condición de parada para la búsqueda de ecuaciones constitutivas o de valoración. La utilidad del índice de interpretabilidad y la metodología propuesta es ilustrada a través de tres aplicaciones diferentes: (i) un reactor continuamente agitado (CSTR), que al ser un modelo de pocas ecuaciones diferenciales, resulta un caso simple para familiarizarse con el índice propuesto, (ii) el proceso de regulación de glucosa en el cuerpo humano (modelo mínimo de glucosa), que al ser un modelo matemático más complejo permite evidenciar la potencialidad del índice para ser aplicado en modelos más complejos y la fiabilidad de sus resultados. Finalmente, (iii) el índice es evaluado durante la deducción, parametrización y validación de un modelo de crecimiento y voltinismo de *Lobesia botrana* con aplicabilidad directa en la región de Cuyo en Argentina. Las dos primeras aplicaciones fueron utilizadas tomando como base modelos matemáticos desarrollados por otros autores para evaluar la utilidad del índice de interpretabilidad. La tercera corresponde a un modelo matemático desarrollado en esta tesis doctoral e ilustra la utilidad de la metodología de modelado completa.

1.2 Objetivos

A continuación se listan los objetivos planteados y aprobados para el desarrollo de esta tesis doctoral.

1.2.1 Objetivo general

Desarrollar y utilizar modelos originales de procesos reales, basados en primeros principios, y proponer una metodología de estimación de parámetros que permita analizar la interpretabilidad paramétrica de dichos

modelos, además de priorizar la búsqueda de los parámetros que tengan más impacto sobre la salida de los mismos.

1.2.2 Objetivos específicos

1. Realizar una revisión de los modelos semi-físicos existentes para procesos biotecnológicos y bioprocesos de interés, así como de las estrategias disponibles para la estimación de parámetros en estos modelos.
2. Desarrollar modelos semi-físicos basados en primeros principios para procesos biotecnológicos y bioprocesos de interés.
3. Validar los modelos desarrollados ya sea mediante experimentación o contraste con datos reales.
4. Proponer un criterio de interpretabilidad que permita determinar en qué nivel de detalle o especificación detener la búsqueda de ecuaciones constitutivas para el MSBPP.
5. Proponer un criterio de jerarquía que permita ordenar los parámetros de un modelo con respecto a su impacto sobre las salidas disponibles del sistema o proceso modelado, tal que este orden sea válido en todo el rango de operación del proceso.
6. Proponer un procedimiento de estimación de parámetros para MSBPP no lineales usando una aproximación basada en modelo, que permita analizar la interpretabilidad del modelo, además de priorizar la búsqueda de los parámetros que tengan más impacto sobre la salida del mismo.
7. Aplicar el procedimiento propuesto en los casos de estudio seleccionados y comparar su desempeño en términos de identificabilidad y ajuste del modelo frente a otras técnicas actualmente usadas para la identificación paramétrica.

1.3 Aportes y limitaciones

Las principales contribuciones de esta tesis doctoral son:

1. La propuesta de un nuevo marco conceptual para la interpretabilidad paramétrica y estructural en el espectro de modelado.
2. El desarrollo y propuesta de un índice matemático para la cuantifi-

cación del grado de interpretabilidad de los modelos semi-físicos basados en primeros principios.

3. La propuesta de una metodología para la construcción de modelos MSBPP que permite determinar el nivel de especificación en el que debe detenerse la búsqueda de ecuaciones constitutivas del modelo matemático.
4. El desarrollo y validación de un modelo de crecimiento y voltinismo de *Lobesia botrana*, particularizado para la región de Cuyo en Argentina.

La aplicabilidad del índice propuesto se encuentra limitada a modelos semi-físicos basados en primeros principios (MSBPP) y no ha sido verificada para modelos de caja negra.

1.4 Difusión de los aportes de investigación

A continuación se enumeran los trabajos aceptados y publicados en revistas indexadas del área (ver Subsección 1.4.1), la participación en congresos nacionales e internacionales (ver Subsección 1.4.2), y la participación en proyectos (ver Subsección 1.4.3).

1.4.1 Publicaciones en revistas indexadas

1. **Aguirre-Zapata, E.**, Alvarez, H., Kulichevsky, L. E., di Sciascio, F., & Amicarelli, A. N. (2023). Forecasting *Lobesia botrana* flight activity: A new semi-physical model. *Crop Protection Journal*, 173,106383.
2. **Aguirre-Zapata, E.**, Alvarez, H., Dagatti, C. V., di Sciascio, F., & Amicarelli, A. N. (2023). Parametric interpretability of growth kinetics equations in a process model for the life cycle of *Lobesia botrana*. *Ecological Modelling Journal*, 482, 110407.
3. **Aguirre-Zapata, E.**, Garcia-Tirado, J., Morales, H., di Sciascio, F., & Amicarelli, A. N. (2023). Methodology for the modeling and estimation of parameters of the growth process of *Lobesia botrana*. *Revista Iberoamericana de Automática e Informática Industrial*

(RIAI), 20(1), 68-79.

4. **Aguirre-Zapata, E.**, Morales, H., Dagatti, C. V., Di Sciascio, F., & Amicarelli, A. N. (2022). Semi physical growth model of *Lobesia botrana* under laboratory conditions for argentina's cuyo region. Ecological Modelling, 464, 109803.
5. **Aguirre-Zapata, E.**, Alvarez, H., Lema-Perez, L., di Sciascio, F., & Amicarelli, A. N. (2024). Interpretability in the modeling spectrum: A conceptual framework and a quantification index. Ecological Modelling, 498, 110882.
6. **Aguirre-Zapata, E.**, Lema-Perez, L., Gómez-Echavarria, L., Botero Castro, H., Maya, Juan C., Chejne, F. & Alvarez, H. (2024). Useful properties of phenomenological-based models. Ecological Modelling, 496, 110850.

1.4.2 Participación en congresos nacionales

1. **Aguirre-Zapata, E.**, Alvarez, H., Lema-Perez, L., di Sciascio, F., & Amicarelli, A. (2023, noviembre). Understanding Interpretability in Modeling Biotechnological Processes: A Critical Review. In XX Information Processing and Control Working Meeting (RPIC2023).
2. **Aguirre-Zapata, E.**, Amicarelli, A., di Sciascio, F., & Garcia-Tirado, J. (2022, September). Design of virtual sensors with physiological delay correction for glucose monitoring in patients with type I diabetes mellitus. In 2022 IEEE Biennial Congress of Argentina (ARGENCON) (pp. 1-8). IEEE.
3. **Aguirre-Zapata, E.**, Garcia-Tirado, J. di Sciascio, F., & Amicarelli, A. (2021, noviembre). Structural identifiability and parameter estimation for *Lobesia botrana* growth model. In XIX Information Processing and Control Working Meeting (RPIC2021).
4. **Aguirre-Zapata, E.**, Morales, H., di Sciascio, F., & Amicarelli, A. (2020, octubre). Optimal growth rate control strategy for δ endotoxins production. 27° Congreso Argentino de Control Automático (AADECA 2020).

1.4.3 Proyectos

- **2020 -2021**

Rol: Colaborador

Proyecto PDTs, Universidad Nacional de San Juan - Dirección de Sanidad Vegetal, Animal y Alimentos, San Juan, Argentina

Modelado del Crecimiento de *Lobesia botrana* y Diseño de Experimentos para Evaluar Efectivamente Momentos de Actuación Sobre la Plaga.

- **2023 -2024**

Rol: Colaborador

Proyecto PDTs, Universidad Nacional de San Juan - Dirección de Sanidad Vegetal, Animal y Alimentos, San Juan, Argentina.

Sistema de Soporte a la Toma de Decisiones Basado en Computación Científica Para el Manejo de la Plaga de *Lobesia botrana* en la Región de Cuyo.

- **2023 -2024**

Rol: Colaborador

Proyecto PICT, Universidad Nacional de San Juan - Dirección de Sanidad Vegetal, Animal y Alimentos, San Juan, Argentina.

Modelado del comportamiento de *Lobesia botrana* para el manejo estratégico de la plaga en los cultivos de la vid en la región de Cuyo.

1.4.4 Colaboración

En el desarrollo de este trabajo, no solo se han incorporado los aportes derivados de la investigación relacionada con el tema interpretabilidad de modelos matemáticos, sino que también se han establecido colaboraciones significativas con otros autores en el campo del modelado y estimación de procesos y bioprocesos. Estas sinergias han enriquecido la perspectiva de la tesis al integrar conocimientos diversos y experiencias especializadas. La colaboración con otros autores ha resultado fructífera, ya que se han compartido ideas innovadoras y enfoques avanzados que han fortalecido tanto la fundamentación teórica como la aplicabilidad práctica de la investigación. Como resultado de estas colaboraciones, se han publicado los siguientes artículos en revistas internacionales con referato y en

congresos nacionales:

1. Morales, H., di Sciascio, F., **Aguirre-Zapata, E.** & Amicarelli, A. N. (2024). Crystallization Process in the Sugar Industry: A Discussion On Fundamentals, Industrial Practices, Modeling, Estimation and Control. Food Engineering Reviews.
2. Morales, H., di Sciascio, F., **Aguirre-Zapata, E.** & Amicarelli, A. N. (2023). A model-based supersaturation estimator (inferential or soft-sensor) for industrial sugar crystallization process. Journal of Process Control, 129, 103065.
3. Morales, H., **Aguirre-Zapata, E.**, di Sciascio, F., & Amicarelli, A. N. (2023). Control strategies with variable reference applied to the Crystallization C process in the sugar industry. Revista Iberoamericana de Automática e Informática Industrial (RIAI), 20(1), 81-92.
4. Lema-Perez, L., **Aguirre-Zapata, E.**, Lyngvi Fougner, A., Amicarelli, A., & Alvarez, H.(2023, noviembre). Connecting sensitivity, identifiability, and interpretability of a glucose minimal model. In XX Information Processing and Control Working Meeting (RPIC2021).
5. Morales, H., **Aguirre-Zapata, E.**, di Sciascio, F., & Amicarelli, A. (2021, noviembre). Control strategies for the C Crystallization process in the sugar industry. In XIX Information Processing and Control Working Meeting (RPIC2021).
6. Morales, H., **Aguirre-Zapata, E.**, di Sciascio, F., & Amicarelli, A. (2020, octubre). Purity and Supersaturation model based estimators for sugar crystallization last stages. 27° Congreso Argentino de Control Automático (AADECA 2020).
7. Morales, H., **Aguirre-Zapata, E.**, di Sciascio, F., & Amicarelli, A. (2019, septiembre). On-line supersaturation monitoring in crystallization process. XVIII Reunión de trabajo en Procesamiento de la Información y Control (RPIC2019).

1.5 Estructura de la tesis

Para la consecución del objetivo general y objetivos específicos de esta tesis doctoral, se consideró la siguiente estructura. En el Capítulo 1 se presentan la motivación, aportes y limitaciones de la presente investigación. En el Capítulo 2 se presentan los fundamentos conceptuales necesarios para el desarrollo teórico-práctico de la investigación. En el Capítulo 3 se analizan de manera individual algunas propiedades de los modelos matemáticos como la sensibilidad e identificabilidad paramétrica, para luego en el Capítulo 4 presentar la propuesta de un nuevo marco conceptual para la interpretabilidad en el espectro de modelado, y el planteamiento de un índice para la cuantificación del grado de interpretabilidad en modelos semi-físicos basados en primeros principios (MSBPP). En ese capítulo, se evalúa la utilidad del índice propuesto considerando el modelo de un reactor continuo de tanque agitado (CSTR) y un modelo mínimo de regulación de glucosa en sangre. En el Capítulo 5 se condensa la propuesta de la sección anterior en una nueva metodología para el desarrollo y estimación paramétrica en MSBPP que considera el grado de interpretabilidad de los parámetros tanto estructurales como funcionales. Finalmente, en el Capítulo 6 se presenta el desarrollo original de un modelo semi-físico basado en primeros principios del proceso de crecimiento y voltinismo de *Lobesia botrana* mediante el uso de la metodología propuesta.

MODELADO MATEMÁTICO

2.1	Aspectos básicos de teoría de sistemas	12
2.1.1	Proceso	12
2.1.2	Sistema	12
2.1.3	Sistema dinámico	13
2.1.4	Tipos de sistemas	13
2.1.5	Análisis de sistemas para modelado	15
2.2	Modelos matemáticos	15
2.2.1	Generalidades	15
2.2.2	Modelado matemático	16
2.2.3	Clasificación de los modelos matemáticos	16
2.3	Modelos semi-físicos basados en primeros principios	20
2.3.1	Método de deducción de un MSBPP	20
2.3.2	Metodología general para la deducción de modelos matemáticos	27
2.3.3	Elementos de un MSBPP	29

La matemática es el lenguaje con el cual Dios ha escrito el universo.

Leonhard Euler

La deducción de un modelo adecuado constituye el punto de partida en el análisis de un sistema, ya que este tipo de representaciones facilita la comprensión de la respuesta del sistema y permite extraer información relevante sobre la dinámica subyacente del mismo. Los modelos matemáticos se derivan, ya sea de aplicar las leyes físicas que describen los procesos fundamentales de los sistemas o mediante el empleo de técnicas de identificación de sistemas. En el presente capítulo, se introducen los

fundamentos conceptuales relacionados con la teoría de sistemas, el modelado matemático y la deducción de modelos matemáticos. Estos conceptos son esenciales para establecer un marco conceptual común que facilite la comprensión de las herramientas propuestas en esta tesis doctoral.

2.1 Aspectos básicos de teoría de sistemas

2.1.1 Proceso

Definición 2.1.1 *Un proceso es un conjunto organizado de operaciones o transformaciones de materia, energía o información; que posee una estructura definida y cuya finalidad es obtener un producto con características específicas mediante la interacción de parámetros, ya sea a nivel abstracto o, de manera más concreta, como el resultado de la interacción entre un equipo o planta y la sustancia tratada (Alvarez, 2022).*

2.1.2 Sistema

Definición 2.1.2 *Un sistema es un conjunto de objetos normalmente interconectados que interactúan de forma coordinada para producir señales observables o salidas del sistema (DiStefano III, 2015; Kheir, 2018; Kluever, 2020).*

Los componentes o elementos que interactúan en un sistema suelen estar conectados de la forma en que lo están para una finalidad o función particular y de manera tal que es posible establecer las relaciones de causa y efecto (o entrada-salida) que suceden entre ellos (DiStefano III, 2015; Kluever, 2020). Dichas interacciones pueden verse influenciadas por señales externas que cuando son desconocidas se denominan perturbaciones y cuando son conocidas como entradas del sistema (Lennart Ljung, 1999). Los sistemas son parte del mundo real pero poseen un límite claro con su entorno. El sistema interactúa con su entorno solo a través de sus límites. Esto es, que el efecto o posible efecto del entorno sobre el sistema puede describirse mediante funciones de entrada dependientes del tiempo $u(t)$, mientras que el efecto del sistema en el entorno se describe mediante las funciones de salida $y(t)$. Así pues, el sistema se puede describir como un

operador S que mapea las entradas $u(t)$ y perturbaciones d en una salida $y(t)$, expresado como se muestra en la Ecuación 2.1 (Hangos & Cameron, 2001).

$$\mathbf{S} : u, d \rightarrow y, \quad y = \mathbf{S} [u, d] \quad (2.1)$$

2.1.3 Sistema dinámico

Definición 2.1.3 *Los sistemas dinámicos son un tipo de sistema en el que la salida actual del sistema (o variables dinámicas del sistema) dependen de las condiciones iniciales (o energía almacenada) y/o de las variables de entrada anteriores. En este tipo de sistemas, las variables dinámicas (o variables de estado) del sistema varían de manera temporal (Kluever, 2020).*

2.1.4 Tipos de sistemas

2.1.4.1. Sistemas de parámetros distribuidos o parámetros concentrados

Los sistemas de parámetros distribuidos son aquellos que requieren un número infinito de variables de estado. Este tipo de modelos incorporan la variación espacial y temporal de los estados dentro del volumen de equilibrio, lo que significa que al aplicar las ecuaciones generales de conservación se obtienen modelos que están representados por ecuaciones diferenciales parciales (EDP) en una, dos o tres dimensiones espaciales, además del tiempo (Hangos & Cameron, 2001; Kluever, 2020). De manera antagónica, un sistema de parámetros concentrados implica un número finito de variables de estado, y por lo tanto, suelen estar regidos por ecuaciones diferenciales ordinarias (EDO) (Kluever, 2020). Los modelos dinámicos de parámetros concentrados, o modelos compartimentales, obtenidos para representar este tipo de sistemas, son utilizados ampliamente para propósitos de control y diagnóstico en diversas aplicaciones, además, se utilizan con frecuencia como base para el diseño de ingeniería, así como para la evaluación de procedimientos operativos seguros (Hangos & Cameron, 2001).

2.1.4.2. Sistemas de tiempo continuo y tiempo discreto

En un sistema de tiempo continuo las variables y funciones están definidas para todos los instantes de tiempo, con el tiempo representado como un intervalo abierto en el conjunto de los reales ($T \subseteq \mathbb{R}$). Por otro lado, en un sistema de tiempo discreto, las variables se definen únicamente en puntos discretos del tiempo, los cuales están ordenados en un conjunto $T = \{\dots, t_0, t_1, t_2, \dots\}$ (Hangos & Cameron, 2001). Algunos sistemas de tiempo continuo consisten en variables que se encuentran en el dominio “analógico”, tales como la posición en un determinado instante de tiempo $x(t)$, mientras que los sistemas de tiempo discreto consisten en variables que se encuentran en el dominio “digital”, como la posición medida $x(kT_s)$ que existe solo en puntos de tiempo discreto $t = T_s$. Así mismo, los sistemas de tiempo continuo se describen a través de ecuaciones diferenciales, mientras que los sistemas de tiempo discreto se describen mediante ecuaciones en diferencias (Kluever, 2020).

2.1.4.3. Sistemas variantes e invariantes en el tiempo

Los sistemas variantes en el tiempo se caracterizan por poseer parámetros que cambian con el tiempo, mientras que en un sistema invariante en el tiempo, los parámetros se mantienen constantes. Es importante destacar, que esta clasificación no hace referencia a la variación temporal de las variables dinámicas o variables de estado del sistema (sistema dinámico o estático), sino a la variación de los parámetros que influyen en dichas variables dinámicas (Kluever, 2020). En síntesis, un sistema S es invariante en el tiempo si su respuesta a una entrada dada es invariante durante la evolución temporal del sistema. Esto quiere decir que los sistemas invariantes en el tiempo no cambian las propiedades de su sistema con el tiempo (Hangos & Cameron, 2001).

2.1.4.4. Sistemas lineales y no lineales

Definición 2.1.4 *Un sistema lineal es aquel en el que las variables involucradas en el sistema están relacionadas de manera proporcional. Además las salidas del sistema cumplen con los principios de proporcionalidad, superposición y homogeneidad.*

Suponga que tiene un sistema que se describe mediante la función $y = f(u)$ donde u es la entrada e y es la salida. Los sistemas lineales obedecen al principio de superposición que se compone de la propiedad de aditividad y de la escalabilidad del sistema: (i) $y_1 = f(u_1)$, entonces $ay_1 = f(au_1)$, donde a es una constante cualquiera. (ii) Si $y_1 = f(u_1)$ e $y_2 = f(u_2)$, entonces $y_1 + y_2 = f(u_1 + u_2)$. En el contexto de las EDO, ésta es lineal si la ecuación diferencial es lineal con respecto a las variables de estado y sus derivadas, pero puede contener términos no lineales con respecto a la variable independiente.

En segundo lugar, el principio de superposición denota que la respuesta dinámica total de un sistema lineal se puede obtener superponiendo las respuestas (o soluciones) a las funciones de entrada individuales. Finalmente, la propiedad de homogeneidad implica que si la entrada del modelo es cero, la salida también será cero (Kluever, 2020).

Definición 2.1.5 *Un sistema no lineal no cumple con al menos uno de los principios de superposición y homogeneidad (Kluever, 2020).*

2.1.5 Análisis de sistemas para modelado

El análisis de sistemas implica el uso de cálculos analíticos o herramientas de simulación numérica para determinar la respuesta del sistema con el fin de evaluar su rendimiento. El análisis repetido ayuda al proceso de diseño en el que la configuración o los parámetros del sistema se modifican, para mejorar el desempeño o cumplir con las limitaciones deseadas (Kluever, 2020).

2.2 Modelos matemáticos

2.2.1 Generalidades

Definición 2.2.1 *Un modelo matemático es una descripción o representación formal de una hipótesis planteada respecto a un sistema en particular, que permite representar las relaciones entre las variables y señales del sistema en términos matemáticos como ecuaciones, gráficos o reglas (DiStefano III, 2015; Hangos & Cameron, 2001; Kheir, 2018).*

En un modelo matemático, algunos parámetros pueden ser desconocidos, por lo que se hace necesario que estos sean determinados a través de la experimentación y la observación, lo que la mayoría de las veces conduce a relaciones empíricas. En estos casos, el modelador debe utilizar juicio experto para intercambiar la complejidad del modelo con la precisión del análisis (DiStefano III, 2015; Hangos & Cameron, 2001; Ljung & Torkel, 1994). Las no linealidades de los sistemas, suelen ignorarse en etapas preliminares de diseño con el fin de obtener aproximaciones lineales de bajo orden que permitan representar con cierto grado de precisión la dinámica del sistema, pero conservando la simplicidad de simulación y el bajo costo computacional. Los modelos no lineales, por otra parte, requieren soluciones numéricas más complejas y suelen requerir pequeños pasos de tiempo de integración para resolver con precisión los sistemas de ecuaciones diferenciales ordinarias, lo que se traduce en un aumento del costo computacional del modelo. En consecuencia, existe una relación de compromiso entre la complejidad del modelo y el tiempo de análisis (Kluever, 2020).

2.2.2 Modelado matemático

Definición 2.2.2 *El modelado matemático es el proceso de aplicación de leyes fundamentales o técnicas apropiadas para derivar ecuaciones matemáticas que describan adecuadamente la dinámica de un sistema en particular (Hangos & Cameron, 2001).*

En la bibliografía se describen habitualmente dos enfoques principales para la deducción o el desarrollo de modelos matemáticos. El primero corresponde al enfoque inductivo, en el que se observa un conjunto de datos de entrada-salida de un sistema específico y se infiere un modelo matemático, en parte o en su totalidad, a partir de estos datos. Por otro lado, el enfoque deductivo emplea el razonamiento basado en principios conocidos, como las leyes físicas o los principios de conservación, para deducir lo desconocido (Karplus, 1977).

2.2.3 Clasificación de los modelos matemáticos

Los modelos matemáticos pueden clasificarse según diversos criterios. En primer lugar, se distinguen entre modelos paramétricos y no paramé-

tricos, donde los primeros se expresan mediante funciones algebraicas, diferenciales, integrales o probabilísticas, mientras que los segundos no dependen de parámetros específicos. Estos modelos a su vez, pueden ser deterministas o estocásticos, según la naturaleza del sistema que representan. Por otro lado, los modelos pueden clasificarse según su escala temporal o espacial como modelos de tiempo discreto o continuo, así como modelos espaciales discretos o continuos, y modelos de escala global, regional o local.

Independiente de cual sea su tipo, los modelos pueden variar desde los más simples y de baja complejidad hasta otros más complejos que involucren múltiples variables y relaciones, y estos pueden ser deducidos para diferentes aplicaciones en áreas como la economía, biología, medicina, física, ingeniería, entre otras.

2.2.3.1. Modelos paramétricos y no paramétricos

Con relación al tipo de representación, los modelos matemáticos pueden ser clasificados como paramétricos o no paramétricos. El concepto de modelo no-paramétrico hace referencia a situaciones y experimentos en los que los resultados se asignan a categorías en lugar de medirse en escalas numéricas. En este tipo de modelos se incluyen los modelos de respuesta al impulso, los modelos espectrales, entre otros (Johansson, 1993). Con este tipo de modelos, los sistemas quedan perfectamente caracterizados mediante un gráfico o tablas que describen sus propiedades dinámicas (Lennart Ljung, 1999).

Por otra parte, los modelos paramétricos son aquellos que se definen mediante una estructura o forma determinada, pero dependen de un número finito de parámetros reales. En este tipo de modelos los datos siguen una distribución de probabilidad con un conjunto fijo de parámetros (Johansson, 1993). Los modelos paramétricos son usados cuando es necesario que el modelo describa las relaciones entre las variables del sistema mediante expresiones matemáticas como pueden ser ecuaciones diferenciales (Lennart Ljung, 1999).

2.2.3.2. Espectro de modelado

Los modelos paramétricos, pueden ser clasificados de acuerdo a su naturaleza en tres diferentes tipos de modelos que conforman el espectro de modelado, lo que a su vez tiene una alineación directa con las etapas del método científico, permitiendo el acceso o agrupación de información sobre el sistema modelado en cada nivel. La Figura 2.1 muestra la relación entre el espectro de modelado y las etapas científicas o de conocimiento, así como ejemplos de la aplicación de los distintos tipos de modelos.

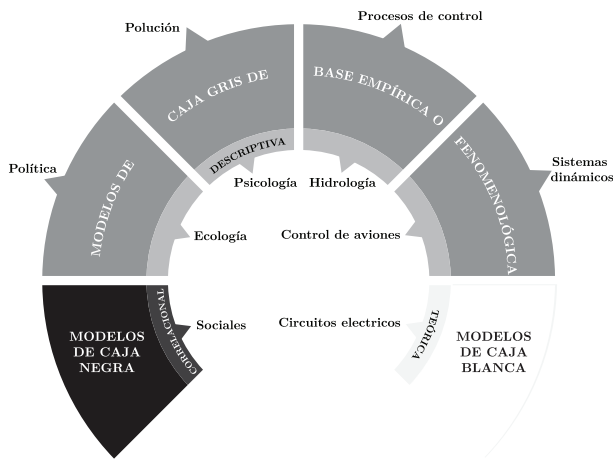


Figura 2.1: Espectro de modelado de Karplus y sus aplicaciones potenciales. Adaptado de (Karplus, 1977).

2.2.3.3. Etapa correlacional - Modelos de caja negra

Los modelos de caja negra, a menudo denominados modelos empíricos, se obtienen estimando tanto su estructura como sus parámetros a partir de los datos de entrada-salida del proceso en cuestión (Brogan, 1991; Hango & Cameron, 2001). Una ventaja de los modelos de caja negra es su capacidad para manejar los efectos aleatorios del objeto y su entorno. Sin embargo, cabe señalar que la generación de “descripciones de datos” significa que un nuevo experimento puede dar lugar a diferentes modelos

(Bohlin, 2006). Este tipo de modelo es útil para modelar sistemas sociales y económicos (Karplus, 1983).

2.2.3.4. Etapa descriptiva - Modelos de caja gris

Los modelos de caja gris se construyen mediante una combinación rigurosa de conocimientos expertos (teorías, leyes o principios) y conocimiento heurístico o datos experimentales relacionados con el sistema o proceso que se pretende modelar. El equilibrio entre el conocimiento experto y los datos experimentales depende del objetivo de la modelización, del origen de la estructura del modelo y de su aplicación final. Esta distinción da lugar a dos categorías de modelos de caja gris: los basados en primeros principios y los derivados empíricamente. Éstos son los modelos más comunes en medicina, biología, ingeniería y otras áreas del conocimiento.

2.2.3.4.1. Modelos de base empírica: estos modelos pertenecen a la subfamilia de los modelos de caja gris y su nombre proviene de una estructura fundamental derivada de un enfoque determinista, que se basa en el análisis de los datos de entrada-salida de un proceso específico (Hargos & Cameron, 2001). En este tipo de modelo, la estructura se determina mediante algoritmos de identificación del sistema y al menos uno de los parámetros se contextualiza o modela utilizando leyes o principios para impartir significado físico al modelo matemático. Los modelos de base empírica se basan menos en los primeros principios para definir la clase del modelo (Alvarez, 2022). Este tipo de modelo es especialmente útil para modelar sistemas políticos, psicológicos y ecológicos (Karplus, 1983).

2.2.3.4.2. Modelos basados en los primeros principios: los modelos basados en primeros principios o MSBPP pertenecen a la subfamilia de los modelos de caja gris y su nombre proviene de una estructura fundamental derivada de un enfoque determinista que les confiere capacidad de generalización. La estructura del modelo de caja gris (MSBPP) basado en primeros principios se esquematiza en una estructura básica y otra extendida. En la estructura básica, los parámetros se derivan de leyes y principios, mientras que al menos un parámetro en la estructura extendida se identifica o sintoniza utilizando datos experimentales (Brogan,

1991). Este tipo de modelo es particularmente valioso para el modelado de sistemas dinámicos (Karplus, 1983).

2.2.3.5. Etapa teórica - Modelos de caja blanca

Los modelos de caja blanca, a menudo denominados modelos perfectos, Estos modelos se denominan de “caja blanca” porque tanto su estructura como sus parámetros son totalmente conocidos (Alvarez, 2022). La principal ventaja de los modelos de este tipo es que permiten al usuario explotar los conocimientos previos invariantes. Su principal debilidad tiene que ver con su incapacidad para hacer frente a lo desconocido y los efectos aleatorios en el objeto y su entorno (Bohlin, 2006). Un ejemplo de modelo de caja blanca es el modelo de circuito eléctrico porque se obtienen teniendo en cuenta los primeros principios (Karplus, 1983).

2.3 Modelos semi-físicos basados en primeros principios

Como se ilustra en la Figura 2.1, los Modelos Basados en primeros principios (MSBPP) demuestran ser particularmente eficaces en la representación de sistemas dinámicos, donde es esencial definir la naturaleza de las relaciones entre las variables del sistema. Esto permite extraer conclusiones significativas sobre los procesos subyacentes que ocurren dentro del sistema. Disciplinas como biología, medicina, biotecnología y procesos agroindustriales obtienen beneficios significativos al derivar modelos de este tipo, ya que posibilita el establecimiento de relaciones precisas entre las variables y parámetros del sistema.

2.3.1 Método de deducción de un MSBPP

En la literatura se encuentran diferentes metodologías para la deducción de modelos semi-físicos basados en primeros principios (Ortega Quintana et al., 2017). Una de ellas, es la metodología propuesta por Alvarez et al., 2009, en la que los autores sintetizan el proceso de construcción de MSBPP en diez pasos que abarcan aspectos fundamentales como la definición del problema de modelado, la construcción de la estructura básica del modelo

2.3. MODELOS SEMI-FÍSICOS BASADOS EN PRIMEROS PRINCIPIOS

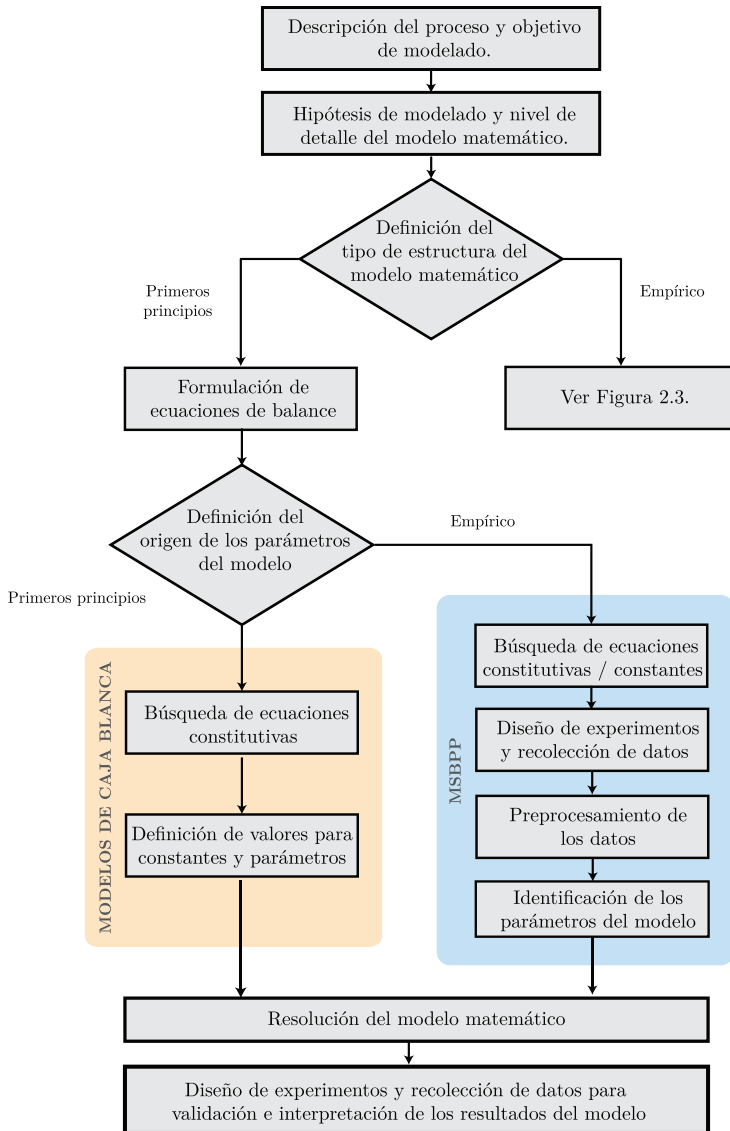


Figura 2.2: Metodología general para construir modelos de caja blanca y modelos semi-físicos basados en primeros principios (MSBPP).

a partir de los primeros principios, la solución del modelo computacional y la validación del modelo en contraste con datos experimentales (Hangos & Cameron, 2001). La metodología para la deducción de MSBPPs se ilustra en la Figura 2.2.

2.3.1.1. Paso 1. Descripción del proceso y objetivo de modelado

Los pasos 1-3 de la metodología de modelado engloban la definición del problema de modelado descrita por Hangos y Cameron, 2001. La primera etapa de modelado consiste entonces en describir de manera verbal y gráfica los procesos de interés que ocurren intrínsecamente en el sistema, evitando aspectos irrelevantes que no pretendan ser modelados. Esto implica formular al menos una pregunta que responderá el modelo cuando se resuelva. Lo anterior, enfrenta al modelador al desafío de establecer un objetivo de modelado claro y evaluar su nivel de conocimiento y comprensión respecto al proceso modelado. El desconocimiento o la falta de comprensión de los aspectos a modelar conducirá a un modelo erróneo (Alvarez, 2022).

2.3.1.2. Paso 2. Hipótesis de modelado y nivel de detalle

Este paso implica proponer una hipótesis de modelado, que es una breve descripción de la forma real o la forma en que se asume un proceso, describiendo la forma en la que interactúan varios fenómenos conocidos para explicar el comportamiento observado. Dicha hipótesis se construye relajando condiciones de fenómenos conocidos y formulando supuestos que consolidan la descripción del proceso. Cada supuesto es una declaración simple que establece una condición o valor para un objeto, operación o variable del proceso modelado. La hipótesis de modelado puede verse ligeramente modificada a lo largo de la construcción del modelo matemático por lo que se hace necesario verificar la consistencia de los supuestos en cada iteración para evitar contradicciones o inviabilidades en el proceso (Alvarez, 2022).

Por otra parte, el nivel de detalle tiene que ver con la escala en la que se realizan las diferentes particiones sobre el sistema de manera que queden expuestas las variables e interacciones de interés en el mismo. Así pues, el

nivel de detalle puede ser macroscópico en el que se ve al proceso como un todo diferenciado de su entorno, o microscópica en la que se consideran particiones más finas que dejan en evidencia relaciones más profundas entre las variables y parámetros del sistema (Alvarez, 2022).

2.3.1.3. Paso 3. Definición de los sistemas de proceso

El sistema a modelar podría ser tan amplio como toda una planta, o tan específico como una unidad operativa o elemento individual de la misma. Por lo tanto, es indispensable definir sistemas de proceso (*SdeP*) que ayuden a establecer límites, entradas y salidas, así como los procesos físico-químicos que suceden en la parte específica que se está analizando (Hangos & Cameron, 2001). Una recomendación clave para la selección de los sistemas de proceso es considerar en primera instancia las separaciones físicas del proceso, cambios de fase, o características de la masa que marquen diferenciales esenciales y de interés de acuerdo con el objetivo de modelado (Alvarez, 2022). Usualmente, los sistemas de proceso se describen mediante un diagrama de flujo que establece los límites, así como las entradas y salidas. Además, suele haber detalles sobre la estructura interna del sistema, incluyendo las unidades operativas y sus interconexiones (Hangos & Cameron, 2001).

2.3.1.4. Paso 4. Aplicación de la ley de conservación

Como su nombre lo indica, los modelos semi-físicos basados en primeros principios, derivan su estructura básica de una ley o principio que permita describir lo que ocurre al interior del proceso. La ley de conservación es una ley que puede ser aplicada a todos los procesos y que permite deducir la estructura básica de los mismos a partir de balances de materia, energía (térmica o mecánica), cantidad de movimiento o cualquier otra cantidad balanceable (Alvarez, 2022). En procesos biotecnológicos, se sugiere realizar al menos los siguientes balances para cada Sistema de Proceso (*SdeP*) definido (Alvarez, 2010):

- Balance total de masa: se aplica si hay una acumulación detectable o significativa.
- Balance por componente: se utiliza cuando al menos una sustancia experimenta un cambio en su concentración en el *SdeP*. Si hay

dos o más sustancias que cambian, se busca una relación entre ellas (ecuación constitutiva). En caso de no existir tal relación, se formula un balance para cada sustancia.

- Balance general de energía: se realiza desde al menos dos sistemas de proceso, esto es, el que entrega energía y el que la recibe. Ambas ecuaciones contienen la misma información sobre el proceso, pero se expresan con variables diferentes.
- Balance de cantidad de movimiento: el balance de cantidad de movimiento se basa en la segunda ley de Newton, que establece que la tasa de cambio del momento lineal de un sistema es igual a la suma de las fuerzas que actúan sobre él Bird, 2002.

2.3.1.5. Paso 5. Determinación de la estructura básica del modelo

Durante el proceso de deducción de la estructura básica del modelo matemático (aplicación de la ley de conservación), podrían resultar ecuaciones redundantes, innecesarias para dar respuesta a las preguntas que responderá el modelo matemático. En este paso es indispensable seleccionar aquellas ecuaciones con información valiosa que permita cumplir con el objetivo de modelado.

2.3.1.6. Paso 6. Reconocimiento y acotamiento de variables, parámetros estructurales y constantes

Este paso si bien resulta sencillo para el modelador, es fundamental considerando que se encuentra directamente relacionado con el objetivo de modelado. En este, se deben especificar las variables, parámetros y constantes del modelo, conservando las diferencias que existen entre cada uno de ellos (Alvarez, 2010, 2022). En la sección 2.3.3 se definen cada uno de estos elementos.

2.3.1.7. Paso 7. Propuesta de ecuaciones constitutivas y de valoración

La búsqueda de ecuaciones constitutivas o de valoración permite obtener la estructura extendida del modelo matemático. En este punto se trata de encontrar expresiones matemáticas que describan en primera instancia los parámetros estructurales del modelo, y en segunda instancia los pará-

metros funcionales que aparecen al abrir nuevos niveles de especificación (Alvarez, 2010, 2022).

2.3.1.8. Paso 8. Conteo de grados de libertad

Los grados de libertad (GL) se calculan considerando la diferencia entre el número de incógnitas del modelo y el número de ecuaciones. Un GL positivo implica que el modelo posee más incógnitas que ecuaciones, lo que hace necesario eliminar algunas incógnitas a través de la definición de algunos parámetros, ya sea a través de valores constantes o ecuaciones constitutivas o de valoración. En cambio, un GL negativo indica que hay más ecuaciones que incógnitas, lo que requeriría la incorporación de nuevas incógnitas o la revisión de las ecuaciones existentes. Para la reducción de los grados de libertad, el modelador debe utilizar su conocimiento experto con el objetivo de simplificar el modelo matemático. Esta tarea comúnmente implica el empleo de estrategias de identificación que facilitan la caracterización de las incógnitas restantes, lo que da paso a la creación de un submodelo generalmente empírico. Una vez concluida esta fase, el modelo es consistente desde el punto de vista computacional y puede procederse a su resolución (Alvarez, 2010).

2.3.1.9. Paso 9. Construcción del modelo computacional

El modelo computacional es un programa informático operativo que implementa el modelo de un sistema dinámico y es el encargado de resolver el sistema de ecuaciones algebraico-diferenciales inherentes al modelo matemático (Kheir, 2018). En este proceso, se transcriben todas las ecuaciones del modelo a código en un lenguaje de programación seleccionado de acuerdo con el uso final del modelo y se presenta una estructura matemática resoluble al reducir los grados de libertad a cero. Es importante destacar que un mismo MSBPP puede ser abordado mediante diversas alternativas computacionales en términos de lenguaje y lógica de programación, método numérico de solución, entre otros. La elección del tipo de modelo (ya sea con parámetros distribuidos o concentrados) determina cuál enfoque computacional resulta más apropiado (Alvarez, 2022).

2.3.1.10. Paso 10. Validación del modelo

La validación del modelo matemático conlleva la resolución del modelo computacional previamente verificado, seguido del análisis de su rendimiento en relación con datos reales extraídos del proceso real, de la literatura, o mediante la comparación del modelo con el comportamiento esperado de procesos similares basada en conocimiento previo (Kheir, 2018). En este punto se distinguen entonces dos tipos de validación que son: (i) validación cuantitativa (basada en datos numéricos) y (ii) validación cualitativa (basada en el comportamiento dinámico del sistema). Las condiciones bajo las cuales se valida el modelo matemático determinan el alcance y grado de generalización del mismo, por lo que es fundamental documentar de manera adecuada los límites de validez del modelo y las condiciones bajo las cuales fue validado (Alvarez, 2010).

En el caso de la validación cuantitativa, es importante reportar los errores máximos y promedio del modelo en comparación con los datos reales dentro del intervalo de validez establecido. En contraste, en situaciones donde los datos del proceso real no están disponibles, es posible realizar una validación del modelo en relación con la respuesta conocida o esperada (cualitativa), especialmente en procesos que aún no han sido construidos o no existen. En este contexto, es esencial que los comportamientos significativos del proceso estén modelados, o en su defecto, asegurarse de que permanezcan en valores conocidos y en condiciones de operación ideales para su interpretación (Alvarez, 2022). Los problemas con la validación de los modelos matemáticos pueden estar explicados por diversas razones (Kheir, 2018):

1. **Problemas experimentales:** las observaciones y mediciones del comportamiento del sistema contienen errores causados por problemas de instrumentación o errores de diseño experimental.
2. **Incertidumbre de modelado:** las simplificaciones del modelo matemático pueden dar como resultado la omisión de aspectos fundamentales del comportamiento del sistema.
3. **Problemas estructurales:** es posible que sea necesaria la simplificación de la estructura (básica o extendida) del modelo matemático.
4. **Linealización del modelo:** si se representa un sistema dinámico no lineal mediante ecuaciones matemáticas lineales, es posible que

esto ocasione problemas con la validación del mismo.

Todos los problemas listados pueden ser abordados *a priori* mediante el análisis de las propiedades intrínsecas del modelo matemático tales como sensibilidad paramétrica, identificabilidad estructural y práctica, e interpretabilidad como se muestra en los Capítulos 3-4.

2.3.2 Metodología general para la deducción de modelos matemáticos

En síntesis, es posible abordar el problema de modelado de sistemas dinámicos desde diferentes frentes, esto es desde el modelado de caja negra o de base empírica, o desde el modelado de caja blanca o basado en primeros principios. Las Figuras 2.2-2.3 ilustran de manera general los pasos que deben seguirse para la construcción de uno u otro tipo de modelo. Si bien el modelado de caja negra no es considerado en esta tesis doctoral, es importante a manera ilustrativa, describir las principales diferencias de este tipo de metodología de modelado con la de los MSBPP.

La Figura 2.3 ilustra los pasos generales seguidos comúnmente para desarrollar modelos de caja negra o de caja gris basados en la empiria. La construcción de estos modelos implica la disponibilidad de varios conjuntos de datos para el entrenamiento del sistema. Estos conjuntos de datos se someten a un proceso de pre-procesamiento y organización, con el objetivo de depurarlos de ruido o datos asincrónicos. A partir de los conjuntos de datos depurados, se extraen características específicas e información relevante que posibilita la predicción del comportamiento de interés. La construcción del modelo utiliza las características extraídas y la información obtenida de los datos, siendo luego evaluado con un conjunto de datos diferente que comparte las mismas características utilizadas en el entrenamiento. Las predicciones del modelo se comparan sistemáticamente con los resultados de interés. Finalmente, el algoritmo proporciona al usuario el conjunto óptimo de señales predictivas para validar y probar en un nuevo conjunto de datos independiente. En esta etapa, se realizan ajustes paramétricos si es necesario, con el objetivo de mejorar la precisión del modelo en relación con el proceso que se pretende modelar.

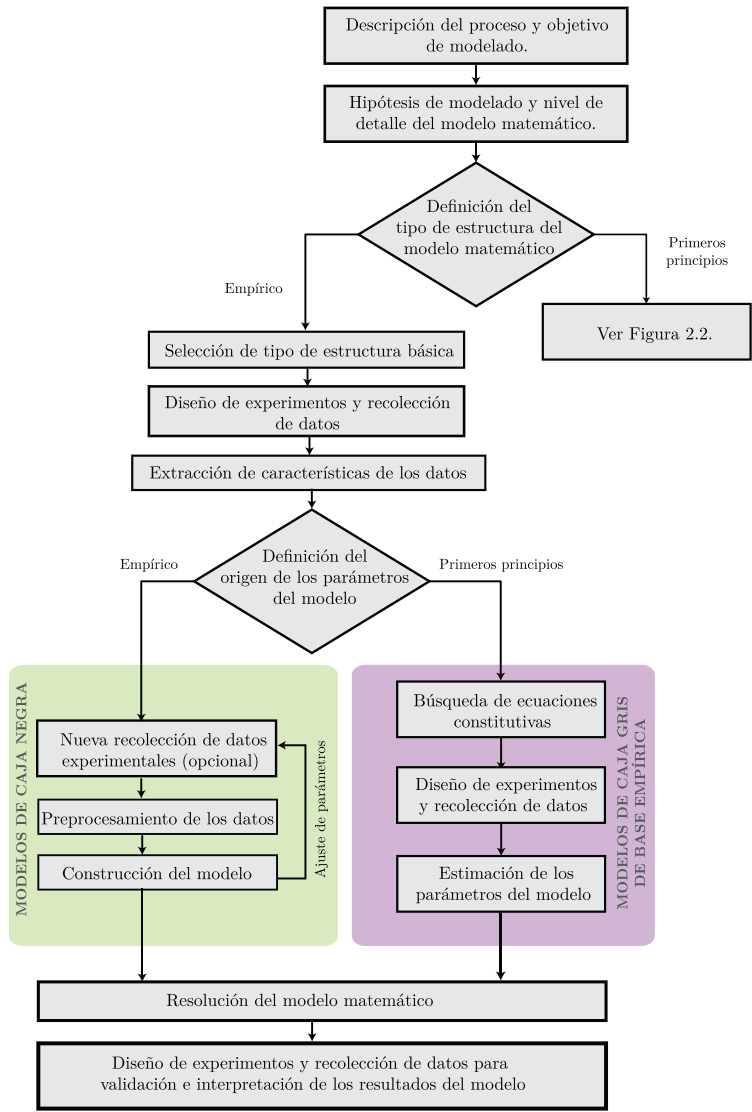


Figura 2.3: Metodología general para construir modelos de caja negra y modelos semi-físicos de base empírica.

2.3.3 Elementos de un MSBPP

2.3.3.1. Variable:

En algunos libros de matemáticas y de ingeniería de procesos, el término “variable” tiene diversas interpretaciones, por lo que en esta tesis doctoral es indispensable hacer una distinción entre los términos “variable” e “incógnita”. Las variables son las cantidades que resuelve el modelo matemático a través de las ecuaciones de balance, mientras que las incógnitas por su parte se emplean para evaluar los grados de libertad en un conjunto de ecuaciones a resolver (como se muestra en el paso 8) (Alvarez, 2017). Las variables pueden ser vistas como símbolos desde el punto de vista del modelador para la deducción del modelo matemático; o pueden ser vistas como elementos que almacenan toda la información contenida en un sistema, para aplicaciones más precisas como control o estimación (Lema-Perez et al., 2019b; Ogata, 2003).

Definición 2.3.1 *Una variable es un símbolo de la estructura básica de un modelo que solo puede conocerse mediante la solución del mismo (Lema-Perez et al., 2019b).*

Definición 2.3.2 *Las variables de estado son un conjunto mínimo de variables que determinan el estado de un sistema dinámico en un momento dado. Están diseñadas para contener toda la información necesaria para predecir el comportamiento futuro del sistema (Ogata, 2003).*

Definición 2.3.3 *Una variable de entrada es aquella que actúa sobre el sistema desde el exterior, sin verse afectadas por lo que ocurre al interior del mismo. Estas pueden ser variables manipuladas o perturbaciones al sistema (Kheir, 2018).*

Las variables del modelo forman parte de las incógnitas consideradas para la evaluación de los grados de libertad, pero no todas las incógnitas de un modelo son variables. Por otra parte, todos los parámetros del modelo son incógnitas antes de iniciar la solución del modelo, es decir que, fijar un parámetro en un valor, no lo convierte en una constante en el modelo.

2.3.3.2. Parámetro:

Un parámetro representa un valor constante o variable sujeto a una ley predefinida, que puede depender de las variaciones inherentes a los procesos que puedan tener lugar en el sistema (Alvarez, 2010). Los parámetros pueden ser estructurales o funcionales y se definen a continuación.

Definición 2.3.4 *Un parámetro es un símbolo que se atribuye al modelo en calidad de componente de configuración o de ajuste. Un parámetro se define como un valor, ya sea numérico o simbólico, que cuantifica la relación entre las diversas partes de un sistema o subsistema (Lema-Perez et al., 2019b).*

Todos los parámetros representados como símbolos en las ecuaciones de balance que conforman la estructura básica del modelo se denominan parámetros estructurales. Este tipo de parámetros tienen una injerencia significativa tanto en la salida como en la estructura del modelo. Pueden estar asociados con un valor numérico o una ecuación constitutiva o de valoración que describe cada uno de ellos.

Definición 2.3.5 *Un parámetro estructural es aquel que se encuentra dentro de la estructura básica del modelo obtenida a partir de la aplicación de la ley de conservación. Un parámetro estructural representa una cantidad que varía en el tiempo o un escalar (Lema-Perez et al., 2019b).*

Los parámetros restantes del modelo matemático se llaman parámetros funcionales, ya que se encuentran en la estructura extendida del modelo matemático, y su injerencia sobre la respuesta del sistema no es tan pronunciada como el observado con los parámetros estructurales. Al igual que los parámetros estructurales, pueden estar asociados a un valor numérico o a una ecuación constitutiva o de valoración que los describa (Lema-Perez, 2018).

Definición 2.3.6 *Un parámetro funcional es aquel que se encuentra dentro de cualquier ecuación constitutiva o de valoración. Se clasifica como parámetro acoplado, parámetro no acoplado o parámetro escalar. Estos parámetros resultan de la estructura extendida, una vez que se especifican las ecuaciones matemáticas de los parámetros estructurales (Lema-Perez et al., 2019b).*

En este sentido, se define un parámetro acoplado como aquel que depende

al menos de una variable del modelo. Por otro lado, un parámetro no acoplado se refiere a aquel asociado a una función matemática que no depende de ninguna variable del modelo, siendo independiente de las variaciones de las magnitudes representadas en el sistema. Por último, un parámetro escalar se caracteriza por poseer un valor numérico (dato) que no varía con el tiempo, y que puede ser determinado *a priori*, o mediante procesos de identificación paramétrica (Lema-Perez, 2018).

2.3.3.3. Constante:

Definición 2.3.7 *Una constante es una magnitud universalmente aceptada cuyo valor no varía en el tiempo (Alvarez, 2022; Lema-Perez, 2018).*

Conclusiones

En este capítulo se estableció un marco conceptual común alrededor de los aspectos fundamentales de esta tesis doctoral, tales como conceptos generales de teoría de sistemas, modelos matemáticos y su clasificación, y las metodologías útiles para la deducción de modelos matemáticos en todo el espectro de modelado. A partir de lo anterior, es posible concluir que esta tesis doctoral se centrará en el desarrollo de herramientas útiles para la cuantificación de la interpretabilidad en modelos semi-físicos basados en primeros principios (MSBPP) de sistemas dinámicos lineales y no lineales, de parámetros concentrados o distribuidos y variantes en el tiempo.

IDENTIFICABILIDAD Y SENSIBILIDAD

3.1	Identificabilidad paramétrica	34
3.1.1	Identificabilidad en el modelado de procesos . .	34
3.1.2	Identificabilidad estructural	35
3.1.3	Identificabilidad práctica	37
3.1.4	Métodos para el análisis de identificabilidad . .	37
3.2	Sensibilidad paramétrica	40

La abstracción es la esencia de las matemáticas, es el arte de despojar a la realidad de sus detalles y enfocarse en lo esencial.

David Hilbert

Durante el proceso de construcción de modelos de caja gris o MSBPP, se originan parámetros del sistema que tienen valores desconocidos que deben ser determinados mediante datos provenientes, ya sea de la experimentación o de fuentes bibliográficas (Ljung & Torkel, 1994). Sin embargo, el proceso de estimación de parámetros en este tipo de modelos suele presentar dificultades, que en la mayoría de los casos surgen por la falta de identificabilidad estructural (Chis et al., 2011a), y/o la falta de identificabilidad práctica que se encuentra implícitamente relacionada con la propiedad de sensibilidad paramétrica (García-Tirado et al., 2018). En este capítulo se presentan los conceptos generales relacionados con las propiedades de sensibilidad e identificabilidad paramétrica.

3.1 Identificabilidad paramétrica

3.1.1 Identificabilidad en el modelado de procesos

La identificabilidad de un modelo matemático se refiere a la posibilidad de determinar unívocamente sus parámetros desconocidos. Esta propiedad evalúa tanto el correcto condicionamiento matemático de la estructura del modelo, como la disponibilidad de datos experimentales con información relevante para la identificación de este conjunto de parámetros. El análisis de la identificabilidad de un modelo matemático se centra en dos aspectos fundamentales: (i) la evaluación de la estructura del modelo matemático (identificabilidad estructural) y (ii) el problema de la identificación de los parámetros desconocidos del modelo (identificabilidad práctica) (Balsacanto et al., 2010; Brun et al., 2001).

En los últimos años, se han publicado numerosos trabajos que exploran algoritmos y metodologías para la identificación de parámetros en modelos de caja gris. Asprion y colaboradores, abordaron el problema del modelado de caja gris en procesos químicos, destacando las bondades de los modelos de caja gris para la estimación de variables no medibles en el proceso real, el diseño de estrategias de control del sistema y la descripción teórica de bioprocesos. En su trabajo, los autores propusieron una metodología que consideraba la selección y procesamiento de los datos como tarea previa al uso de algoritmos de optimización, para la estimación de parámetros. Sin embargo, no consideraron la propiedad de identificabilidad como un requisito previo (Asprion et al., 2019). De manera similar, en el trabajo desarrollado por Baulmet y Dostál en el año 2020, si bien se proponen técnicas novedosas para la estimación de parámetros en modelos matemáticos de caja gris, los autores no consideran la propiedad de identificabilidad como requisito para la implementación de algoritmos de estimación de parámetros (Bäumelt & Dostál, 2020). En el año 2019, Brastein y colaboradores, propusieron el uso de los análisis de identificabilidad práctica e identificabilidad estructural, para determinar como el diseño óptimo de experimentos, afecta la información dinámica contenida en los datos utilizados para la estimación de parámetros de un modelo de caja gris (Brastein et al., 2019).

En el año 2021, Hotvedt y colaboradores, propusieron el uso de la regularización de parámetros, como una estrategia para conservar la interpretabilidad física de un modelo, aun cuando sus parámetros no son identificables. En su trabajo, los autores destacaron la dificultad para garantizar la identificabilidad en modelos de caja gris debido a su elevado número de parámetros, y afirmaron que el rendimiento predictivo de los modelos y su interpretabilidad, está influenciada por los datos disponibles para el proceso de ajuste (Hotvedt et al., 2021). Los hallazgos anteriores dejan en evidencia que el concepto de identificabilidad paramétrica no es considerado de manera formal en las metodologías para la estimación de parámetros utilizadas comúnmente.

3.1.2 Identificabilidad estructural

La identificabilidad estructural es una propiedad binaria de los modelos matemáticos que se refiere a la posibilidad de calcular una solución única para los parámetros de un modelo, a partir de datos de entrada- salida libres de ruido. La propiedad de identificabilidad puede ser local o global dependiendo de si la propiedad se cumple dentro de una vecindad o para todo el espacio de parámetros como se muestra en las Ecuaciones 3.1 - 3.2 (Joubert et al., 2021).

El análisis de la identificabilidad estructural se centra en la posibilidad de encontrar un conjunto único de parámetros para un modelo, teniendo en cuenta la estructura matemática del modelo. La identificabilidad estructural de un modelo puede ser local (solución única dentro de una vecindad), o global (solución única en todo el espacio de los parámetros). El software STRIKE-GOLDD fue presentado por Villaverde et al. en 2016, y es una herramienta para realizar análisis de identificabilidad estructural de modelos matemáticos complejos utilizados en biología de sistemas y otras áreas del conocimiento, empleando un enfoque de geometría diferencial. En su propuesta, los autores reformulan el problema de la identificabilidad como un problema de observabilidad aumentada, en el que el rango de la matriz de observabilidad-identificabilidad, que se construye utilizando derivadas de Lie, da información sobre la identificabilidad estructural local de los

parámetros desconocidos del modelo.

3.1.2.1. Identificabilidad estructural local

Un parámetro es estructuralmente, localmente identificable, si es posible identificar su valor único, dentro de una vecindad en el espacio de los parámetros (Chis et al., 2011b). En las definiciones que se plantean a continuación, p es el vector de parámetros desconocidos $p = (p_1, \dots, p_{n_p})$ de un modelo matemático y p^* es el vector de parámetros identificados.

Definición 3.1.1 *Un parámetro p_i es estructuralmente localmente identificable si para casi cualquier $p^* \in P$ existe un vecindario $V(p^*)$ tal que se cumple que:*

$$\text{si } p \in V(p^*) \rightarrow y(t, p) = y(t, p^*) \Rightarrow p_i = p_i^* \quad (3.1)$$

3.1.2.2. Identificabilidad estructural global

Un parámetro es estructuralmente, globalmente identificable, si es posible identificar de manera única el valor de dicho parámetro, en el espacio de los parámetros.

Definición 3.1.2 *Un parámetro p_i es estructuralmente globalmente identificable si para casi cualquier $p^* \in P$*

$$y(t, p) = y(t, p^*) \Rightarrow p_i = p_i^* \quad (3.2)$$

3.1.2.3. No identificabilidad

La no identificabilidad de un modelo matemático, implica que los algoritmos de estimación no podrán encontrar valores confiables para el conjunto de parámetros, esto es, no podrán encontrar valores únicos ni local, ni globalmente (Chis et al., 2011b).

Definición 3.1.3 *Un parámetro p_i es estructuralmente no identificable si para casi cualquier $p^* \in P$ no existe un vecindario $V(p^*)$ tal que pueda encontrarse un valor único para dicho parámetro.*

3.1.3 Identificabilidad práctica

Si bien la identificabilidad estructural permite evaluar la estructura matemática de un modelo desarrollado y garantizar que sus parámetros estructurales son identificables de forma única, en la práctica, esto puede no ser posible debido a la falta de información contenida en los datos experimentales utilizados para este fin. Aplicando un análisis práctico de identificabilidad, se puede evaluar si es posible asignar valores únicos a los parámetros desconocidos de un modelo matemático, utilizando datos experimentales o esquemas experimentales sujetos a ruido. Los datos experimentales y el ruido experimental están estrechamente relacionados con esta propiedad. La identificabilidad práctica puede determinarse de dos formas: (i) *a priori* para anticipar la calidad del esquema experimental seleccionado, o (ii) *a posteriori* para evaluar la calidad de las estimaciones de los parámetros tras la calibración o ajuste del modelo matemático (Balsa-Canto et al., 2010).

3.1.4 Métodos para el análisis de identificabilidad

En la literatura, se han presentado diferentes herramientas que buscan facilitar el cálculo y análisis de la propiedad de identificabilidad estructural, como son el software DAISY presentado en (Bellu et al., 2007), el software GenSSI presentado en (Chis et al., 2011a), el software STRIKE-GOLDD presentado en (Villaverde et al., 2016), o el software SIAN en (Hong et al., 2019). Para el análisis de identificabilidad en los casos de aplicación abordados en esta tesis doctoral, se considerará el uso de los métodos GenSSI y STRIKE-GOLD.

3.1.4.1. GenSSI

El software GenSSI se basa en el Método de Generación de Series junto con el uso de tablas de identificabilidad (Balsa-Canto et al., 2010). El método de Generación de Series, inicia mediante la generación de un sistema no lineal de ecuaciones sobre los parámetros del modelo mediante el cálculo sucesivo de las derivadas de Lie de las ecuaciones del sistema. Para el análisis de los resultados, se establece que la identificabilidad global de los parámetros se verifica si la solución del sistema de ecuacio-

nes es única. Una vez calculadas las derivadas de Lie, se construyen las tablas de identificabilidad, que no solo ayudan a construir modelos con parámetros globalmente identificables, sino que también orientan sobre la mejor manera de abordar el sistema no lineal de ecuaciones restantes. GenSSI realiza automáticamente todos los cálculos simbólicos y presenta un diagnóstico final junto con figuras que describen la identificabilidad estructural del conjunto de parámetros.

GenSSI puede proporcionar uno de los siguientes diagnósticos relacionados con la identificabilidad de un conjunto de parámetros: (i) Identificabilidad Estructural Global (SGI), que se garantiza si es posible obtener una solución única para los parámetros p a partir del conjunto de ecuaciones algebraicas generadas utilizando datos de entrada-salida. (ii) Identificabilidad Estructural Local (SLI), que se establece cuando SGI no puede demostrarse; en este caso, GenSSI intenta determinar si hay un número finito de valores de parámetros indistinguibles, lo que implicaría identificabilidad local del modelo. Finalmente, (iii) Identificabilidad Estructural No Identificable (SNI) hasta la derivada actual, que se produce cuando al menos uno de los parámetros del modelo es no identificable. La Figura 3.1 ilustra el método de cuadros de identificabilidad donde:

	θ_1	θ_2	θ_3	θ_4		θ_j
C_1						x
C_2			x			
C_3	x	x				
C_4	x	x				
C_i						

Figura 3.1: Cuadro de identificabilidad para el análisis de identificabilidad estructural. Adaptada de (Balsa-Canto et al., 2010)

una cruz en las coordenadas (i, j) indica que el coeficiente de la serie generadora depende del parámetro θ_j . Los círculos verdes representan aquellos casos en los que el parámetro correspondiente, puede ser calculado de manera única a partir de una sola ecuación del sistema (por ejemplo, el parámetro θ_3 que depende únicamente del coeficiente C_2). Los cuadros rojos representan aquellos casos en los que un conjunto de ecuaciones (dos o más), permiten identificar de manera única el parámetro correspondiente (por ejemplo, el parámetro θ_1 que depende de los coeficientes C_3 y C_4). Finalmente, la ausencia de una cruz en las coordenadas (i, j) indica que el coeficiente i no tiene dependencia del parámetro j . Un elemento único distinto de cero en una fila significa que el parámetro correspondiente es estructuralmente identificable (por ejemplo el parámetro θ_3). El rango del cuadro de identificabilidad (número de parámetros para los cuales se encontró una relación), en caso de ser completo, garantiza, al menos, la identificabilidad estructural local del modelo matemático. Es importante destacar que los coeficientes en el cuadro de identificabilidad corresponden a las sucesivas derivadas de Lie generadas. Estos coeficientes o constantes de estructura, ayudan a determinar si las derivadas sucesivas de la salida pueden proporcionar suficiente información para identificar los parámetros del sistema. Si estos coeficientes permiten que las derivadas sucesivas de h generen una base en el espacio de estados, entonces el sistema es localmente observable y, por lo tanto, identificable (Balsa-Canto et al., 2010; Chis et al., 2011b).

GenSSI realizó de manera automática todos los cálculos simbólicos para presentar finalmente un texto (describiendo el diagnóstico) y figuras (cuadros de identificabilidad) que describen la identificabilidad estructural del conjunto de parámetros. GenSSI puede producir uno de los siguientes diagnósticos con relación a la identificabilidad de un conjunto de parámetros: (i) Identificabilidad estructural global (SGI), que está garantizada si se puede obtener una solución única de los parámetros p a partir del conjunto de ecuaciones algebraicas generadas utilizando datos de entrada-salida. (ii) Identificabilidad estructural local (SLI), que se da cuando SGI no se puede demostrar; en este caso GenSSI intenta establecer si hay un número finito de valores de parámetros indistinguibles, en cuyo caso el modelo será localmente identificable. Finalmente, (iii) No identificabilidad estructural (SNI) hasta la derivada actual, que se da cuando el al menos

uno de los parámetros del modelo es no identificable.

3.1.4.2. STRIKE-GOLDD

Por otra parte, se considera también el uso del software STRIKE-GOLDD, que, utilizando también derivadas de Lie, parte de un procedimiento en el que, si la matriz de observabilidad es de rango deficiente, se eliminan las columnas correspondientes a los parámetros identificables y se obtiene una submatriz reducida. Luego, se obtiene una base para el núcleo (espacio nulo) de esta matriz. Los coeficientes del núcleo definen una o varias ecuaciones diferenciales parciales cuya solución son las combinaciones identificables. Si bien, el procedimiento tiene el potencial para poner de manifiesto combinaciones de parámetros no identificables que sean en sí mismas identificables, se debe reconocer que la complejidad computacional de calcular el núcleo limita su aplicabilidad a modelos con un número moderado de parámetros no identificables. Para abordar la complejidad computacional de analizar modelos extensos o complicados, el método propone descomponerlos en unidades más pequeñas conocidas como submodelos. Estos submodelos contienen una fracción del modelo original y pueden ser más manejables para su análisis. La selección de estos submodelos se puede realizar mediante técnicas de optimización, lo que permite identificar qué componentes del modelo original deben incluirse en cada submodelo. Esta aproximación ayuda a reducir la carga computacional al limitar la cantidad de cálculos necesarios para evaluar el modelo completo (Villaverde et al., 2016).

3.2 Sensibilidad paramétrica

Los métodos de análisis de sensibilidad permiten estudiar cómo el origen de la incertidumbre en la salida de un modelo puede distribuirse entre diferentes fuentes de incertidumbre en la entrada del modelo (Saltelli et al., 2008). Dentro de los objetivos del análisis de sensibilidad se destacan la posibilidad de detectar errores técnicos en el modelo, la identificación de regiones críticas en el espacio de las entradas, la simplificación de modelos y la posibilidad establecer prioridades de investigación a partir del análisis (Iooss & Lemaître, 2015). Por otra parte García-Tirado et al.,

2018, sugiere que un análisis de sensibilidad que permita determinar la influencia de los parámetros en la salida del modelo, puede llegar a ser esclarecedor para dos de los principales problemas asociados con la falta de identificabilidad práctica, esto es: (i) parámetros correlacionados, y (ii) parámetros sin influencia en la salida del modelo. Chis et al., 2011b, sugirieron que el uso del análisis de sensibilidad para evaluar la influencia relativa de los parámetros en la salida del modelo puede ser útil también para anticiparse a la falta de identificabilidad estructural o práctica (Chis et al., 2011b).

Considerando lo expuesto anteriormente, el coeficiente de sensibilidad paramétrica permite clasificar y jerarquizar los parámetros de un modelo en función de su influencia relativa en las predicciones (salidas) del modelo (Balsa-Canto et al., 2010). Las sensibilidades paramétricas locales pueden definirse como se muestra en la Ecuación 3.3, donde e corresponde al número de experimento, o a la salida del modelo, y θ_i a los parámetros del modelo matemático.

$$S_{\theta_{i,j}}^e = \frac{\partial y_j^{e,o}}{\partial \theta_i}(t_s^{e,o}) \quad (3.3)$$

En vista de lo anterior, Balsa-Canto y Banga, 2011 proponen el uso del método *Latin Hypercube Sampling* (LHS), y el uso de cinco factores de significancia (λ_θ), previamente descritos por Brun en (Balsa-Canto et al., 2016; Brun et al., 2001). Lo anterior como alternativa al método Monte Carlo en términos de coste computacional, para obtener una estimación más precisa de la clasificación de los parámetros de un modelo. El método LHS permite explorar la sensibilidad en un amplio espacio de parámetros, sin requerir un número excesivo de muestras. Esto se traduce en un menor costo computacional, especialmente valioso en modelos matemáticos de caja gris de bioprocesos, que suelen tener muchos parámetros. En este tipo de modelos, el problema de identificación paramétrica pasa de buscar los “valores reales” de un conjunto de parámetros desconocidos, a convertirse en un problema que busca encontrar valores de parámetros físicamente interpretables, con el fin de describir adecuadamente los datos de entrada-salida del proceso. Los propios coeficientes de sensibilidad contienen información relevante sobre la importancia de los parámetros individuales. Un coeficiente de sensibilidad grande, significará que un

cambio en el parámetro θ_n tendrá un efecto significativo en la salida del modelo y , lo que hace que el parámetro θ_n sea identificable con los datos disponibles, si todos los demás parámetros son fijos, como puede deducirse del análisis de la Ecuación 3.3.

Los factores de significancia (λ_θ) se utilizan para obtener información del cálculo de sensibilidades. La información obtenida a través de λ_θ , puede ser útil para decidir sobre la necesidad de reformulación o reestructuración del modelo matemático, o para fijar los parámetros menos relevantes para mejorar la identificabilidad estructural o práctica (Balsa-Canto et al., 2010). Los factores λ_θ^{msqr} y λ_θ^{mabs} presentados en las Ecuaciones 3.4-3.5 cuantifican que tan sensible es la salida del modelo matemático a un parámetro determinado. Sin embargo, λ_θ^{mabs} considera además las interacciones entre parámetros.

$$\lambda_{\theta_i}^{msqr} = \frac{1}{n_d} \sqrt{\sum_{mc=1}^{n_{lhs}} \sum_{e=1}^{n_e} \sum_{o=1}^{n_o} \sum_{s=1}^{n_s} \left(s_{\theta_i,j}^{e,o}(t_s^{e,o}) \right)^2} \quad (3.4)$$

$$\lambda_{\theta_i}^{mabs} = \frac{1}{N_T} \sum_{mc=1}^{n_{lhs}} \sum_{e=1}^{n_e} \sum_{o=1}^{n_o} \sum_{s=1}^{n_s} \left| \left(s_{\theta_i,j}^{e,o}(t_s^{e,o}) \right) \right| \quad (3.5)$$

Luego, λ_θ^{max} y λ_θ^{min} presentadas en las Ecuaciones 3.6-3.7, indican la presencia de valores atípicos y proporcionan información sobre el signo.

$$\lambda_{\theta_i}^{max} = \frac{1}{n_{lhs}} \sum_{mc=1}^{n_{lhs}} \left[\max_{e,o,s} s_{\theta_i,j}^{e,o}(t_s^{e,o}) \right] \quad (3.6)$$

$$\lambda_{\theta_i}^{min} = \frac{1}{n_{lhs}} \sum_{mc=1}^{n_{lhs}} \left[\min_{e,o,s} s_{\theta_i,j}^{e,o}(t_s^{e,o}) \right] \quad (3.7)$$

Finalmente, λ_θ^{mean} presentado en la Ecuación 3.8, proporciona información sobre el signo del efecto medio de un cambio en un parámetro sobre la salida del modelo.

$$\lambda_{\theta_i}^{mean} = \frac{1}{N_T} \sum_{mc=1}^{n_{lhs}} \sum_{e=1}^{n_e} \sum_{o=1}^{n_o} \sum_{s=1}^{n_s} \left(s_{\theta_i,j}^{e,o}(t_s^{e,o}) \right) \quad (3.8)$$

En las Ecuaciones 3.4-3.8, n_e es el número de experimentos, n_o el número de variables observadas (observables), n_{lhs} el número de muestras en el algoritmo LHS, n_s el tiempo de muestreo específico para cada experimento con un observable dado, $n_d = n_{lhs}n_in_on_s$, y $s_{\theta_{i,j}}^{e,o}(t_s^{e,o})$ es la sensibilidad del parámetro θ_i , con el observable j , en el experimento e (con tiempo de muestreo t_s).

Conclusiones

En este capítulo se estableció un marco conceptual común alrededor de dos propiedades claves de los modelos matemáticos tales como la sensibilidad y la identificabilidad paramétrica. Se describieron algunos métodos y herramientas útiles para la realización de los análisis asociados a cada una de ellas, así como la interpretación de sus resultados. Las propiedades fueron descritas considerando que la presente tesis doctoral parte de la hipótesis de que existe una relación directa entre la interpretabilidad de un modelo, con las demás propiedades intrínsecas del mismo (identificabilidad y sensibilidad). Los análisis descritos en este capítulo serán utilizados posteriormente para la cuantificación del grado de interpretabilidad de los MSBPP.

INTERPRETABILIDAD DE MODELOS

4.1	Estado del arte en interpretabilidad	46
4.1.1	Concepto de interpretabilidad	46
4.1.2	¿Por qué es importante la interpretabilidad? . .	50
4.2	Un nuevo marco conceptual	51
4.2.1	Etimología y fundamento filosófico	51
4.2.2	Interpretabilidad	54
4.2.3	Interpretabilidad de parámetros estructurales .	54
4.2.4	Interpretabilidad de parámetros funcionales . .	55
4.3	Taxonomía de interpretabilidad	56
4.4	Índice de interpretabilidad	58
4.4.1	Escenario y grafo de modelado	59
4.4.2	Nivel de especificación	61
4.4.3	Criterio de interpretabilidad paramétrica	64
4.4.4	Coeficiente de sensibilidad	64

El gran libro de la naturaleza está escrito en símbolos matemáticos.
Galileo Galilei

Los modelos matemáticos son una representación que busca capturar por lo menos una característica de interés de un sistema o proceso en particular para un propósito previamente establecido, como simulación dinámica, simulación estática, diseño de procesos, monitoreo, supervisión, optimización y control (Alvarez, 2022; Hangos & Cameron, 2001). En algunas áreas de aplicación del modelado matemático, el usuario final del modelo requiere en gran medida que tanto la estructura del modelo, como los parámetros de interacción usuario-modelo, tengan un significado

físico en su contexto de aplicación, debido a que a través de los mismos puede darse significado a la respuesta del modelo, y con base en ello facilitar la toma de decisiones con relación al proceso. Un claro ejemplo de estas áreas de aplicación, son los procesos biológicos (orientados a tratamientos médicos), biotecnológicos, bioquímicos o agroindustriales, para los cuales, el modelo además de permitir la predicción de ciertos comportamientos, facilita a través de sus parámetros y variables la obtención de información que se pueda utilizar para realizar diagnósticos, agrupaciones y para orientar el tratamiento o control del proceso en cuestión (Cousins & Riondato, 2019; Erro et al., 2013; Lou et al., 2013). Lo anterior, plantea dos de las preocupaciones actuales en el área de modelado de procesos y sistemas, esto es: (i) el grado de interpretabilidad que deben exhibir los modelos matemáticos, a través de sus parámetros funcionales y estructurales y (ii) la posible relación existente entre la interpretabilidad de un modelo matemático con otras propiedades importantes del mismo, como la precisión, la identificabilidad, la sensibilidad de sus parámetros y la usabilidad del modelo matemático.

4.1 Estado del arte en interpretabilidad

En los últimos años, se han presentado en la literatura diferentes trabajos que pretenden integrar los conceptos de interpretabilidad en los modelos matemáticos, proponer taxonomías para evaluarla y conferir interpretabilidad a los modelos matemáticos de caja negra o discutir la necesidad de disponer de modelos matemáticos interpretables. Un número reducido de trabajos se refiere a los modelos de caja gris, sin embargo, es necesario discutir la interpretabilidad de este tipo de modelos considerando que la estructura básica o extendida de los mismos podría no provenir de la aplicación del principio de conservación, lo que hace que el tipo y grado de interpretabilidad de sus parámetros sea desconocido.

4.1.1 Concepto de interpretabilidad

La mayoría de los trabajos sobre interpretabilidad parten de la premisa de que si un sistema puede explicar cómo toma las decisiones, entonces el

usuario final del sistema puede verificar si las decisiones son correctas con respecto a otros criterios particulares, como la seguridad, la explicabilidad y la duda técnica (Doshi-Velez & Kim, 2017).

Autores como Doshi-Velez y Kim, 2017, Miller, 2019, Molnar, 2020, Ghosh y Kandasamy, 2020, Fan et al., 2021, entre otros, han seguido la línea de pensamiento que atribuye la interpretabilidad al ser humano y no al modelo en sí, definiendo la interpretabilidad como se muestra en la Definición 4.1.1. En esta línea de pensamiento, Carvalho et al., 2019 contemplaron la interpretabilidad en el ámbito del aprendizaje automático, como una comunicación bidireccional entre dos entidades, es decir, entre el ser humano como usuario y la computadora como ejecutora del modelo. Los autores destacan que los sistemas interpretables pueden ser explicables si sus decisiones pueden ser comprendidas por un ser humano.

Definición 4.1.1 *La interpretabilidad es el grado en que un ser humano puede entender la causa de una decisión, o el grado en que un ser humano puede predecir consistentemente el resultado de un modelo (Ghosh & Kandasamy, 2020).*

- *La interpretabilidad está relacionada con la capacidad de los seres humanos para observar y razonar sobre la información proporcionada por un modelo matemático (Doshi-Velez & Kim, 2017; Molnar et al., 2020).*
- *La interpretabilidad es el grado en que un observador puede entender la causa de una decisión tomada por un modelo matemático (Miller, 2019).*

Miller, 2019 observó que la mayoría de los trabajos divulgados sobre inteligencia artificial explicable basaban su noción de interpretabilidad de modelos en la intuición de los investigadores sobre lo que representa una buena explicación, alejándose de las ciencias sociales, que son útiles para comprender cómo los humanos perciben y crean explicaciones en torno a la respuesta de un modelo matemático. Murdoch et al., 2019 afirmaron que la interpretabilidad, a grandes rasgos, es un concepto amplio y mal definido sobre el que aún no hay consenso en la literatura, y definieron la acción

de interpretar como la acción de extraer conocimiento relevante de un modelo, a partir de las relaciones contenidas en los datos que se aprenden .

En otra línea de pensamiento, algunos autores como Lipton, 2018, He et al., 2020, Li et al., 2022, entre otros, definen la interpretabilidad como una cualidad (o propiedad) del modelo matemático, y no como algo que dependa únicamente del operador humano como se muestra en la Definición 4.1.2.

Definición 4.1.2 *La interpretabilidad es un medio o requisito previo para generar confianza o seguridad en la mejor comprensión de un modelo matemático. Es la capacidad de inferir relaciones causales y transferir información útil, ya sea a través de sus resultados o de otros procedimientos computacionales (Lipton, 2018).*

- *Un modelo interpretable es una arquitectura explícita que puede utilizarse internamente para razonar y aprender sobre un sistema o proceso, y externamente para explicar los resultados del modelo a un ser humano.*

Algunas clasificaciones comunes para la interpretabilidad son “interpretabilidad local” e “interpretabilidad global”. Algunos autores como Kopitar et al., 2019, Hong et al., 2020, Azodi et al., 2020, y Oviedo et al., 2022 definen estos tipos de interpretabilidad como sigue:

Definición 4.1.3 *La interpretabilidad local consiste en proporcionar explicaciones sobre una única instancia (observación individual en un conjunto de datos). Explica la relación entre las variables y la etiqueta (o variable dependiente a predecir) para un caso particular o subconjunto de casos (Azodi et al., 2020; Hong et al., 2020).*

Definición 4.1.4 *La interpretabilidad global consiste en mostrar la estructura lógica de un modelo para explicar su funcionamiento general. Explica la relación general entre variables y etiquetas de un modelo para todas las instancias posibles (Azodi et al., 2020; Hong et al., 2020).*

Otra clasificación encontrada en la literatura tiene que ver con la interpretabilidad intrínseca, y la interpretabilidad extrínseca o interpretabilidad

post-hoc, que según autores como Lipton, 2018, Stiglic et al., 2020, y Oviedo et al., 2022 (ver Definiciones 4.1.5 y 4.1.6, respectivamente). Según Murdoch et al., 2019, la interpretabilidad *post-hoc* es especialmente difícil de lograr cuando los parámetros del modelo no muestran claramente las relaciones que el modelo ha aprendido.

Definición 4.1.5 *La interpretabilidad intrínseca, también llamada transparencia, responde a la pregunta ¿cómo funciona el modelo? Hablamos de modelos intrínsecamente interpretables, debido a la simplicidad de su estructura (Lipton, 2018; Oviedo et al., 2022).*

Definición 4.1.6 *La interpretabilidad extrínseca o post-hoc describe la aplicación de métodos que proporcionan información sobre lo que un modelo entrenado ha aprendido, y responde a la pregunta ¿qué más puede decir el modelo? (Lipton, 2018; Oviedo et al., 2022).*

Pocos autores como Hotvedt et al., 2021; Lema-Perez et al., 2019c en trabajos recientes, destacan la importancia de definir la interpretabilidad en el contexto de los modelos de caja gris y describen la importancia de la interpretabilidad de los parámetros del modelo como elementos clave para garantizar la interpretabilidad del modelo de forma más general. Hotvedt et al., 2021 dicen que un modelo es físicamente interpretable si su comportamiento de salida está en línea con los principios físicos que gobiernan el sistema y si cualquier parámetro del modelo que represente una cantidad física tiene un valor numérico físicamente factible. En este sentido, Lema-Perez et al., 2019c definen la interpretabilidad paramétrica de dos tipos (contextualizada y general) como sigue:

Definición 4.1.7 *Un parámetro tiene interpretabilidad contextualizada si su significado físico sólo es válido en un modelo matemático particular. Es decir, si su significado depende de las consideraciones y suposiciones del modelo. Luego, un parámetro tiene interpretabilidad general si su significado físico es inherente dentro de un modelo matemático en un dominio científico específico, independientemente de las consideraciones y supuestos utilizados para desarrollar el modelo (Lema-Perez et al., 2019c).*

4.1.2 ¿Por qué es importante la interpretabilidad?

Distintos autores se han aventurado a formular hipótesis sobre las situaciones o tipos de modelos en los que es fundamental garantizar la interpretabilidad. Ghosh y Kandasamy, 2020, sugirieron que la interpretabilidad es necesaria en aquellos casos en los que la predicción realizada por el modelo guía la toma de decisiones por parte de operadores humanos. Esencialmente, en aquellas aplicaciones en las que las decisiones tomadas representan un riesgo para otros humanos. Doshi-Velez y Kim, 2017; Lipton, 2018 argumentaron que la interpretabilidad es necesaria debido a la incompletitud en la formalización de un problema, es decir, la interpretabilidad es necesaria en el modelado de sistemas donde no se pueden caracterizar todas las opciones de fallo del sistema o donde se requiere predicción multiobjetivo.

Algunos autores han relacionado la interpretabilidad con conceptos como explicabilidad, simulabilidad, modularidad, entre otros, entendiendo que estos son consecuencia de la interpretabilidad del modelo (Murdoch et al., 2019). Autores como Doshi-Velez y Kim, 2017, Carvalho et al., 2019, Murdoch et al., 2019, Azodi et al., 2020, y Poursabzi-Sangdeh et al., 2021 destacaron que la interpretabilidad es útil para garantizar otras propiedades en los modelos matemáticos, tales como:

- **Equidad o imparcialidad.** Implica que la toma de decisiones no sufre ningún tipo de discriminación.
- **Fiabilidad:** Un modelo es fiable cuando el operador humano puede entender por qué se toma una determinada decisión.
- **Causalidad.** Implica que un operador humano puede prever o entender un determinado cambio en la salida del modelo ante una perturbación del sistema.
- **Usabilidad.** Los modelos utilizables proporcionan información verdadera que ayuda a los operadores humanos a realizar una tarea específica.
- **Simulabilidad.** Un modelo es simulable cuando un operador humano puede recrear o comprender internamente todo su proceso de toma de decisiones mediante la simulación y el razonamiento.
- **Generación de nuevo conocimiento.** Un modelo puede generar nuevo conocimiento, siempre que, una vez identificadas las interac-

ciones de las variables, estos resultados puedan ser interpretados.

Si bien, las investigaciones citadas hacen un importante esfuerzo en proponer taxonomías que ayuden a definir, evaluar y conferir interpretabilidad a los modelos matemáticos, un gran número de ellas, concentran sus esfuerzos en caracterizar la interpretabilidad para modelos matemáticos de caja negra, justificando que los modelos matemáticos de caja blanca son interpretables de manera intrínseca. Sin embargo, dichas investigaciones ignoran la necesidad de establecer un marco conceptual de interpretabilidad paramétrica, que abarque la totalidad del espectro de modelado (pasando por los modelos matemáticos de caja gris), y que de luz sobre las principales diferencias al hablar de interpretabilidad en el modelado empírico o en el modelado basado en primeros principios. Por esta razón, en este trabajo se propone un marco teórico para la interpretabilidad paramétrica en modelos matemáticos del espectro de modelado (caja blanca, caja gris y caja negra).

4.2 Un nuevo marco conceptual

4.2.1 Etimología y fundamento filosófico

4.2.1.1. Etimología

Interpretar es una palabra derivada del latín *interpretari* formado por el pre-fijo *inter* (entre), y el radical *pret* (comprar-vender) que se deriva a su vez del latín *pretium* (precio- valor). En su etimología original, la palabra interpretar era utilizada para denotar la acción de interceder o negociar entre dos personas que no hablan el mismo idioma (Abadi, 1986). Considerando lo anterior, la interpretabilidad podría definirse, como la capacidad de mediar entre dos entes con lenguajes diferentes para lograr una negociación por un precio o valor. La interpretabilidad se produce a través del reconocimiento de las relaciones entre el significante, es decir, las piezas de información de los modelos matemáticos (estructura y parámetros), y el significado, es decir, el contenido explícito de cada pieza de información en un contexto particular.

4.2.1.2. Fundamento filosófico

De acuerdo con Ogden y Richards, 1923, la relación entre los pensamientos, las palabras y las cosas, puede ser ejemplificada a través de un diagrama (ver Figura 4.1) donde se posicionan los tres elementos que influyen en el entendimiento (significado, significante y referente) en los vértices de un triángulo, mientras que las conexiones entre ellos se muestran a través de los lados. Dicha representación, establece relaciones causales entre un proceso cognitivo y su representación simbólica (significante). Al recibir el mensaje, los significantes evocan procesos de referencia y generan una actitud interpretativa, cuya similitud con la del emisor puede variar según el contexto y las circunstancias comunicativas.

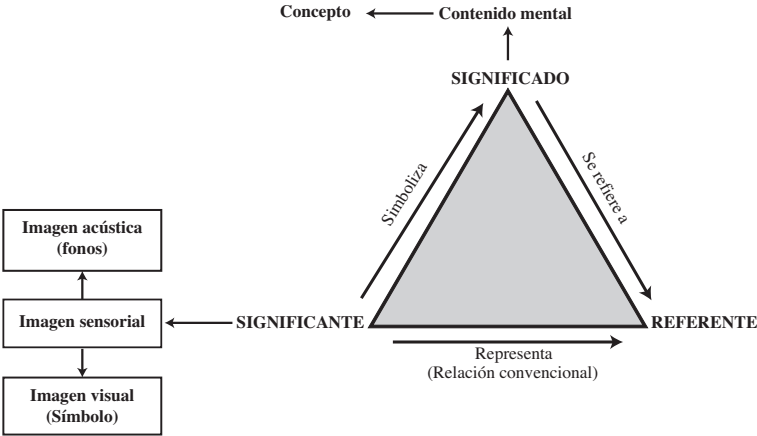


Figura 4.1: Relación significado-significante. Adaptada de (Ogden & Richards, 1923)

Por otra parte, en el proceso de vinculación entre el significado y el referente, se pueden identificar relaciones de diferente grado de directividad. Por un lado, se encuentran conexiones más inmediatas, como cuando el pensamiento se centra en una percepción directa, por ejemplo, al hacer referencia a un objeto en particular. Por otro lado, existen conexiones más indirectas, como cuando se hace referencia a un concepto o individuo

(Ogden & Richards, 1923).

Definición 4.2.1 Desde el punto de vista lingüístico, el significado es la representación mental o conceptual asociada al significante. Es la idea, concepto o referente que se evoca en la mente del individuo cuando se encuentra con el significante. Es importante destacar que el significado no está intrínsecamente contenido en el significante, sino que surge a través del proceso activo de interpretación (Ahmad, 2020).

Definición 4.2.2 Desde el punto de vista lingüístico, el significante es la forma material del signo, lo que se puede percibir a través de los sentidos, como lo que se ve, se oye, se toca, se huele o se saborea. Es la parte física o perceptible del signo que lleva consigo una carga semántica (Ahmad, 2020).

A partir de la etimología y los fundamentos filosóficos del concepto de interpretación, se presenta un nuevo marco conceptual que busca unificar la taxonomía de la interpretabilidad en el espectro de modelado. Este enfoque se detalla en la Figura 4.2, partiendo de la premisa de que la interpretabilidad de un modelo matemático está determinada por la interpretabilidad de sus parámetros.

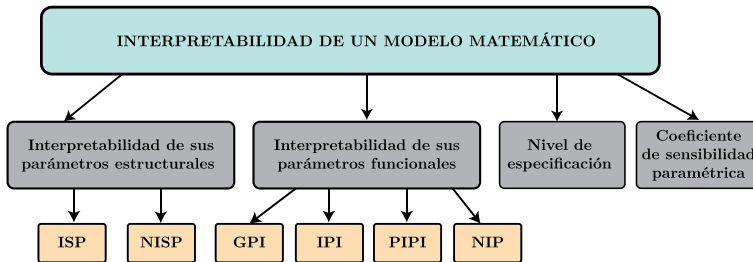


Figura 4.2: Interpretabilidad de los modelos matemáticos. ISP: Parámetro estructural interpretable. NISP: Parámetro estructural no interpretable. GPI: Interpretabilidad paramétrica general. IPI: Interpretabilidad paramétrica heredada. PIPi: Interpretabilidad Paramétrica Parcialmente Heredada. NIP: Parámetro funcional no interpretable.

4.2.2 Interpretabilidad

Definición 4.2.3 *Capacidad de los modelos matemáticos para mediar entre el conocimiento científico descrito en forma matemática o en lenguaje de máquina, y el conocimiento experto o semi-experto del usuario final del modelo matemático. Lo anterior, a través del reconocimiento de las relaciones presentes entre el significante, es decir, las piezas de información de los modelos matemáticos (estructura y parámetros), y el significado, es decir, el contenido explícito de cada pieza de información en un contexto particular.*

4.2.3 Interpretabilidad de parámetros estructurales

4.2.3.1. Parámetro estructural interpretable (ISP)

Definición 4.2.4 *Los parámetros estructurales de un modelo matemático son interpretables solamente si pueden ser explicados desde leyes o el principio de conservación aplicado al proceso de interés para la deducción de las ecuaciones de la estructura básica del modelo matemático.*

No son parámetros estructurales interpretables aquellos que hacen referencia a la configuración matemática del modelo, por ejemplo, el número de neuronas en una red neuronal.

4.2.3.2. Parámetro estructural no interpretable (NISP)

Definición 4.2.5 *Los parámetros estructurales de un modelo matemático son no interpretables cuando no pueden ser explicados mediante leyes o el principio de conservación aplicado al proceso de interés.*

La estructura de un modelo matemático puede tener hiperparámetros ¹ útiles para configurar la estructura básica de los modelos, por ejemplo, el número de capas o nodos en una red neuronal. Estos últimos son parámetros de configuración que no hacen referencia al fenómeno modelado.

¹Un hiperparámetro es un parámetro que se establece antes de que comience el proceso de aprendizaje y permanece constante durante el entrenamiento de un modelo de aprendizaje automático. No se aprende a partir de los datos, sino que es determinado por el usuario o el diseñador del modelo (Bischi et al., 2023).

4.2.4 Interpretabilidad de parámetros funcionales

Los parámetros funcionales de un modelo matemático, son aquellos que aparecen en la estructura extendida del mismo. En modelos matemáticos de caja gris, los parámetros funcionales permiten profundizar en los niveles de especificación del modelo matemático de acuerdo a las necesidades del modelador y el usuario final del modelo. En modelos matemáticos de caja negra, los parámetros funcionales son útiles para aproximar las dinámicas y comportamientos contenidos en los datos experimentales. Se resalta que, en los modelos matemáticos de caja blanca, la interpretabilidad tanto de los parámetros estructurales como funcionales, se encuentra garantizada. La interpretabilidad de los parámetros funcionales de un modelo matemático puede clasificarse de diferentes formas como sigue:

4.2.4.1. Interpretabilidad paramétrica general (GPI)

Definición 4.2.6 *Un parámetro funcional tiene interpretabilidad paramétrica general cuando su significado físico no está sujeto a la hipótesis y supuestos utilizados durante la derivación del modelo matemático.*

Lo anterior implica que, el parámetro puede interpretarse sin estar contextualizado en un modelo concreto, por lo que tanto su significado como su símbolo (relación significado-significante), pueden ser interpretados de manera general. Por ejemplo, el parámetro físico-químico de densidad asociado al símbolo ρ , usado en ingeniería de procesos.

4.2.4.2. Interpretabilidad paramétrica heredada (IPI)

Definición 4.2.7 *Un parámetro funcional tiene interpretabilidad paramétrica heredada, cuando su significado físico solo es válido en el contexto de un modelo matemático específico. Por tanto, el parámetro no puede interpretarse lejos de su aplicación. Es decir, hereda su significado de la hipótesis y supuestos utilizados para la deducción del modelo matemático.*

En este tipo de parámetros, la relación significado - significativo depende tanto de la aplicación, como de la hipótesis y supuestos de modelado.

4.2.4.3. Interpretabilidad paramétrica parcialmente heredada (PI-PI)

Definición 4.2.8 *Un parámetro funcional tiene interpretabilidad paramétrica parcialmente heredada cuando solo una parte de su significado físico puede ser heredado de la hipótesis o leyes utilizadas durante la derivación del modelo matemático. Lo anterior, en función del análisis de otras propiedades del modelo. El parámetro debe reinterpretarse, ya que su significado original no puede heredarse completamente en el nuevo contexto de aplicación del modelo.*

En este tipo de parámetros, la relación significado - significante está sujeta al contexto de aplicación del modelo flexibilizando las analogías o supuestos utilizados para la deducción del mismo.

4.2.4.4. Parámetro funcional no interpretable (NIP)

Definición 4.2.9 *Un parámetro funcional no tiene interpretabilidad paramétrica cuando no es posible establecer una relación entre su significante (símbolo) y su significado, ni en un contexto de aplicación particular, ni de manera general.*

Este tipo de parámetros suelen ser adimensionales, o tienen una combinación de unidades que surgen de un ajuste dimensional de la expresión que los contiene. Por lo tanto, tampoco sus unidades de medida aportan a su interpretabilidad.

4.3 Taxonomía de interpretabilidad

En la Tabla 4.1 se presenta la taxonomía de interpretabilidad propuesta para el espectro de modelado que va desde los modelos de caja negra, hasta los modelos de caja blanca, sin descartar los modelos que aparecen como una combinación de los extremos del espectro de modelado, esto es, los modelos matemáticos de caja gris.

En el caso de los modelos matemáticos de caja blanca, tanto sus parámetros

estructurales como funcionales tienen garantizada la interpretabilidad, ya que todos ellos pueden ser explicados por leyes o por el principio de conservación que rige el proceso que se modela, sin estar relacionados con ninguna hipótesis o supuesto específico. Por el contrario, en los modelos matemáticos de caja negra, los parámetros estructurales y funcionales no son interpretables porque este tipo de modelos se basan en datos experimentales tanto para obtener la estructura del modelo como para ajustar todos sus parámetros. En el caso de los modelos de caja gris basados en datos empíricos, los parámetros estructurales no son interpretables, ya que la estructura de dichos modelos no se origina a partir de primeros principios sino de plantillas matemáticas que se ajustan a un conjunto de datos experimentales.

En el caso de los modelos de caja gris basados en los primeros principios (MSBPP), los parámetros estructurales del modelo son interpretables, ya que las ecuaciones de estado se derivan de la aplicación de leyes físicas. Esta propiedad no se aplica directamente a los parámetros funcionales del modelo, ya que se derivan de hipótesis de modelado y supuestos considerados durante la deducción del modelo. Por lo tanto, los parámetros funcionales de un MSBPP pueden ser parámetros con interpretabilidad general, interpretabilidad heredada del contexto, interpretabilidad parcialmente heredada del contexto o pueden no ser interpretables en absoluto.

Tabla 4.1: Taxonomía de interpretabilidad paramétrica en el espectro de modelado.

	Interpretabilidad de parámetros estructurales		Interpretabilidad de parámetros funcionales			
	ISP	NISP	GIP	IPi	PIPI	NIP
Caja negra		×				×
Caja gris de base empírica		×			×	×
MSBPP	×		×	×	×	×
Caja blanca	×		×			

4.4 Índice de interpretabilidad

Después de un análisis exhaustivo del estado del arte, no se han encontrado evidencias sobre una formalización matemática que permita cuantificar el grado de interpretabilidad de un modelo matemático. Esta falta de cuantificación dificulta la toma de decisiones relacionadas con la sobre-parametrización del modelo y el nivel de especificación en el que debe detenerse la búsqueda de ecuaciones constitutivas o de valoración para el modelo. Sin embargo, sería valioso comparar modelos con dinámicas modeladas similares pero que difieren en estructura, ecuaciones constitutivas y niveles de especificación en términos de interpretabilidad. Por lo tanto, en esta sección se introduce el concepto de nivel de especificación y se propone un índice de interpretabilidad para cuantificar el grado de interpretabilidad del MSBPP, demostrando su relación con el coeficiente de sensibilidad.

El índice de interpretabilidad ponderado en función del nivel de especificación ($IIW S_L(\theta)$) presentado en las ecuaciones 4.1- 4.3 permite cuantificar el grado de interpretabilidad de un MSBPP considerando: (i) la influencia del criterio de interpretabilidad paramétrica (según la taxonomía propuesta), (ii) el coeficiente de sensibilidad de los parámetros del modelo, y (iii) el nivel de especificación para cada uno de los parámetros considerados.

$$IIW S_L(\theta) = S_{L_n}(\theta) [0,8PI_C(\theta) + 0,2S_{c_n}(\theta)] \quad (4.1)$$

$$IIW S_{L,0} = \frac{\sum_{\theta=1}^{N_\theta} IIW S_{L,0}(\theta)}{N_\theta} \quad (4.2)$$

$$IIW S_{L,M_s+1} = \frac{\frac{\sum_{\theta=1}^{N_\theta} IIW S_{L,M_s+1}(\theta)}{N_\theta}}{IIW S_{L,0}} \quad (4.3)$$

La Ecuación 4.1 corresponde al índice de interpretabilidad para cada parámetro (θ) del escenario de modelización analizado. La Ecuación 4.2 corresponde al índice de interpretabilidad para el escenario 0 (estructura básica del modelo). Este valor es importante teniendo en cuenta que la estructura básica constituye la base para la comparación de los escenarios

de modelización posteriores. La Ecuación 4.3 muestra como calcular el índice de interpretabilidad para los demás escenarios de modelización. En las Ecuaciones 4.1 - 4.3, M_s es el escenario de modelado analizado, N_θ es el número total de parámetros en el escenario de modelado analizado, $PI_C(\theta)$ es el criterio de interpretabilidad de un modelo (ver la Sección 4.4.3), $S_{c_n}(\theta)$ es el coeficiente de sensibilidad normalizado de cada parámetro (ver la Sección 4.4.4) y $S_{L_n}(\theta)$ es el nivel de especificación normalizado para cada parámetro (ver la Sección 4.4.2). En consecuencia, en la Ecuación 4.1 el criterio de interpretabilidad paramétrica se pondera con un valor de 0,8 en la función, mientras que el coeficiente de sensibilidad normalizado se pondera con un factor de 0,2. A la hora de determinar estos valores de ponderación, se partió del hecho de que al probar distintas combinaciones, esta combinación refleja mejor el efecto de las propiedades individuales sobre el índice. Esto significa que a medida que se añaden parámetros no interpretables al modelo matemático, el índice de interpretabilidad del modelo disminuye.

4.4.1 Escenario y grafo de modelado

En aquellos modelos en los que se dispone de diferentes opciones para la apertura de un parámetro, aparece lo que se denomina como escenario de modelado (M_s). Si bien cada apertura de un nuevo nivel de especificación da lugar a un nuevo escenario de modelado, cada nuevo escenario no implica un nuevo nivel de especificación. $M_s = 0$ corresponde a la estructura básica del modelo y $M_s > 0$ corresponde a los distintos escenarios de la estructura extendida del modelo. Los escenarios de modelado constituyen versiones completas del modelo, que cumplen con los grados de libertad y pueden ser resueltas computacionalmente. No debe confundirse el nivel de especificación de los parámetros con los escenarios de modelado en los que aparecen los mismos.

La Figura 4.3 presenta un ejemplo en el que se considera un modelo matemático compuesto por una ecuación de diferencial. En ella se puede observar como en la estructura básica del modelo coincide el número de escenario de modelado con el nivel de especificación de los parámetros ($M_s = S_L = 0$). Al abrir un nuevo nivel de especificación (recuadro violeta), la ecuación constituye un nuevo escenario de modelado ($M_s = 1$),

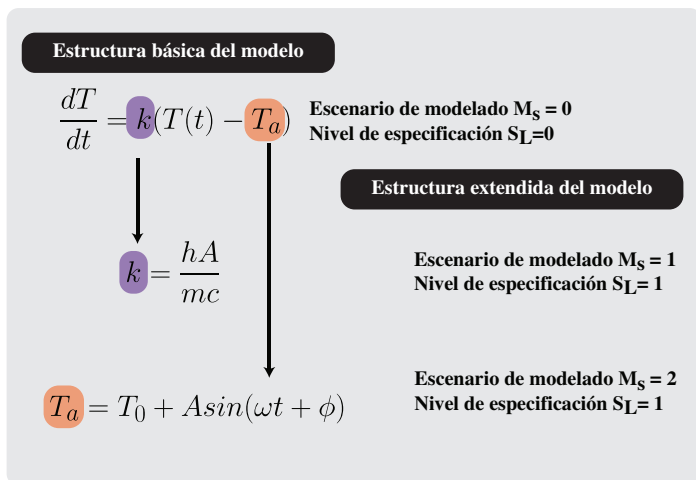


Figura 4.3: Escenario de modelado en contraste con los niveles de especificación (ejemplo).

en el que los nuevos parámetros h , A , m y c tienen nivel de especificación $S_L = 1$. Finalmente, se presenta un nuevo escenario de modelado $M_s = 2$ (recuadro naranja), en el que se describe una ecuación de valoración para el parámetro T_a (que pertenece a la estructura básica) y en el que los nuevos parámetros resultantes tienen nivel de especificación $S_L = 1$. Como se observa, si bien, el escenario de modelado es el 2, esto no implica que los parámetros se encuentren en nivel de especificación 2. Los escenarios de modelado dan lugar a un grafo dirigido que facilita la determinación del orden jerárquico necesario para establecer comparaciones en términos de ajuste y de interpretabilidad. Como se ilustra en la Figura 4.4, el grafo de modelado parte de la estructura básica del modelo, en la cual todos los parámetros son estructurales y, por lo tanto, tienen interpretabilidad garantizada. Las aristas del grafo representan los caminos que se forman al definir los parámetros del nivel de especificación anterior a través de constantes o ecuaciones constitutivas o de valoración. Dos aristas que parten de un mismo nodo (o vértice) indican la posibilidad de describir un parámetro de dos o más formas diferentes. En este contexto, el modelador, basándose en el índice de interpretabilidad y el error de ajuste, debe elegir

que pueden ser variables, parámetros, constantes, combinación de estos y operadores no lineales como exponenciales o funciones logarítmicas. Cuando se obtiene la estructura básica o nivel de especificación cero de un modelo matemático, sus parámetros deben definirse a través de otras ecuaciones que permitan representarlos con funciones matemáticas o directamente con valores numéricos.

El primer conjunto de ecuaciones utilizadas para definir los parámetros de la estructura básica del modelo (parámetros estructurales) se conoce como nivel de especificación 1. Este nivel de especificación puede generar nuevos parámetros del modelo que deben calcularse del mismo modo que los parámetros de la estructura básica, lo que significa que se necesitan ecuaciones matemáticas adicionales para definirlos. Este segundo conjunto de ecuaciones se conoce como nivel de especificación 2, que, a su vez, puede generar nuevos parámetros que deben calcularse de la misma manera. Este proceso puede continuar durante tantos niveles de especificación como sea necesario hasta que todos los parámetros del modelo se definan con ecuaciones de valoración, es decir, con valores numéricos directos, lo que garantiza que no se generen más parámetros nuevos y por definir.

Según la explicación anterior, el concepto del nivel de especificación se refiere a la posición actual o al grado de conocimiento en el cual un objeto o sus componentes están definidos dentro de un nivel de contenido de información relacionado con determinaciones matemáticas. Este nivel comienza con una cantidad mínima de información, representada por un número sin unidades descrito como $S_L = 0$, y progresa hacia una comprensión integral explicada por leyes o teorías, o ecuaciones empíricas ($S_L = n$) (Lema-Perez, 2018), como se describe en la Figura 4.5. En este contexto, el nivel representa la granularidad o resolución con la que se examina el objeto real. Al analizar el objeto en diferentes niveles, los investigadores buscan obtener perspectivas adicionales que contribuyan a una comprensión más profunda del objeto.

Es importante señalar que los niveles de especificación son una característica del parámetro que no está relacionada con el escenario de modelado

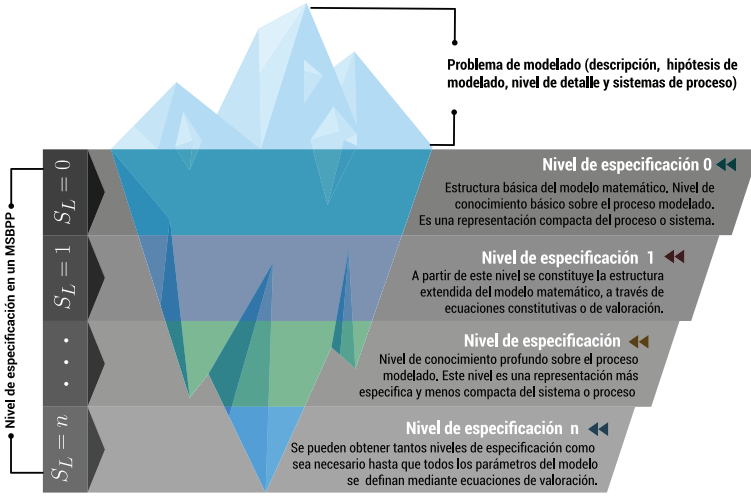


Figura 4.5: Representación gráfica de los niveles de especificación en MSBPP.

sino con la estructura del propio modelo. Esto significa, por ejemplo, que un parámetro estructural que nunca ha sido descrito por una ecuación de evaluación mantendrá su nivel de especificación $S_L = 0$ a través de las diversas representaciones del modelo matemático (escenarios de modelado) hasta que sea descrito por una ecuación de especificación de nivel $S_L = 1$, esto se encuentra ejemplificado en la Figura 4.3. En este contexto, el índice de interpretabilidad propuesto en las Ecuaciones 4.1-4.3 tiene como objetivo abordar la pregunta: “¿cuándo detener la apertura de nuevos niveles de especificación en un modelo semi-físico basado en primeros principios?” La Ecuación 4.4 representa el valor absoluto del nivel de especificación normalizado y es un componente de los parámetros del índice de interpretabilidad.

$$S_{L_n}(\theta) = \frac{S_L(\theta)}{S_{L_{max}}(\theta)} \quad (4.4)$$

donde $S_L(\theta)$ es nivel de especificación en el que aparece cada parámetro del modelo analizado, y $S_{L_{max}}(\theta)$ es el nivel de especificación máximo del modelo analizado.

4.4.3 Criterio de interpretabilidad paramétrica

En esta tesis doctoral, se propone el uso de una escala de interpretabilidad que ayuda a establecer la interpretabilidad de un parámetro con base en el conocimiento *a priori* del modelador, y de acuerdo con las definiciones presentadas en la Sección 4.2, de forma que este valor pueda ser utilizado en el índice para cuantificar la interpretabilidad del modelo matemático. Suponga que θ es un parámetro (estructural o funcional) en un MSBPP, y $PI_C(\theta)$ es el criterio de interpretabilidad de este parámetro, entonces los valores ordinales se asignan como se muestra en la Ecuación 4.5.

$$PI_C(\theta) = \begin{cases} 1 & \text{si } \theta = ISP \text{ o } GPI \\ 0,9 & \text{si } \theta = IPI \\ 0,75 & \text{si } \theta = PIPI \\ 0 & \text{si } \theta = NPI \text{ o } NISP \end{cases} \quad (4.5)$$

Dichos valores ordinales, aunque arbitrarios reflejan la degradación que sufre la interpretabilidad de un parámetro a medida que se aleja de la interpretabilidad general (valor 0).

4.4.4 Coeficiente de sensibilidad

Como se mencionó en el Capítulo 3, la sensibilidad paramétrica permite clasificar y jerarquizar los parámetros del modelo en función de su influencia relativa en las predicciones (resultados) del modelo. El coeficiente de sensibilidad obtenido a partir del análisis de sensibilidad descrito en el capítulo anterior, es utilizado para el cálculo del índice de interpretabilidad paramétrica de cada escenario de modelado. En este punto es importante destacar que cada nuevo escenario de modelado, implica un nuevo análisis de sensibilidad. Dado que existen diferentes métodos para analizar la sensibilidad de un modelo matemático, y cada método proporciona coeficientes dentro de rangos u ordenes de magnitud específicos, se hace necesaria la normalización de los coeficientes obtenidos a partir del análisis descrito en 3.3 como se muestra en la Ecuación 4.6.

$$S_{c_n}(\theta) = \frac{S_c(\theta)}{S_{c_{max}}(\theta)} \quad (4.6)$$

donde $S_c(\theta)$ es el coeficiente de sensibilidad de cada parámetro obtenido mediante el análisis de sensibilidad, $S_{c_{max}}(\theta)$ es el máximo coeficiente

de sensibilidad obtenido a partir del análisis y $S_{c_n}(\theta)$ es el coeficiente de sensibilidad normalizado.

Conclusiones

En este capítulo se abordó una brecha significativa en el campo del modelado matemático, enfatizando la necesidad crítica de modelos que no solo sean precisos, sino también interpretables, especialmente en campos aplicados como la biología, la medicina y la biotecnología. En el se introdujo un nuevo marco conceptual para la interpretabilidad en el espectro de modelado, así como un índice para la cuantificación de la interpretabilidad en MSBPP. El aporte presentado en este capítulo, no solo aborda las limitaciones encontradas en estudios previos, sino que también ofrece una solución práctica y sistemática para incorporar la interpretabilidad en todo el espectro de modelización, desde modelos de caja blanca hasta modelos de caja negra.

METODOLOGÍA PROPUESTA

5.1	Metodología propuesta	68
5.1.1	Enunciado del problema de modelado	68
5.1.2	Deducción de la estructura básica del modelo	70
5.1.3	Deducción de la estructura extendida del modelo	70
5.1.4	Reporte e interpretación de los resultados	71
5.2	Ejemplos de aplicación	75
5.2.1	Caso de estudio: CSTR	75
5.2.2	Caso de estudio: modelo mínimo de glucosa	80
5.3	Análisis de interpretabilidad	86
5.3.1	Análisis de interpretabilidad del reactor continuo de tanque agitado (CSTR)	86
5.3.2	Análisis de interpretabilidad del modelo mínimo de glucosa	99

¿Qué es la investigación, sino una cita a ciegas con el conocimiento?.

Will Harvey

En este capítulo se propone una metodología que resume y agrupa los elementos presentados en las secciones anteriores para la construcción de modelos MSBPP considerando el análisis de la interpretabilidad de los diferentes escenarios de modelado, y otras propiedades intrínsecas del modelo matemático. El objetivo principal de la metodología propuesta es que se convierta en una herramienta útil para los modeladores, que les permita utilizar la interpretabilidad como criterio de parada para determinar cuándo detener la búsqueda de ecuaciones constitutivas (Alvarez et al., 2009), evaluar la calidad de su modelo matemático en

términos de interpretabilidad paramétrica y del modelo en general y que facilite la selección de estructuras de modelos útiles para el desarrollo de sistemas de apoyo a la toma de decisiones basados en modelo (Nauck, 2003).

Es importante señalar que, si bien la interpretabilidad puede servir como criterio de parada en la búsqueda de ecuaciones constitutivas o de valoración, este criterio no es el único que debe tenerse en cuenta. En este sentido, el error absoluto del modelo respecto a los datos experimentales sigue siendo un criterio importante que va de la mano de la interpretabilidad cuando se toman decisiones sobre el nivel de especificación del modelo matemático. El modelador debe encontrar una relación de compromiso entre el error de ajuste y el índice de interpretabilidad del modelo, buscando que dicha relación entre ambos sea inversamente proporcional, es decir, que el modelo tenga el menor error posible y el mayor índice de interpretabilidad posible.

5.1 Metodología propuesta

La metodología mostrada en la Figura 5.1 comprende las tres fases que resumen los pasos de modelado propuestos en (Alvarez et al., 2009), incluyendo el análisis de interpretabilidad del modelo matemático que se realiza en la tercera fase. La primera fase corresponde a la formulación del problema de modelado, la segunda fase al desarrollo de la estructura básica del modelo matemático, y la tercera fase a la búsqueda de ecuaciones constitutivas, teniendo en cuenta la interpretabilidad del modelo matemático. Se describe una fase final adicional relacionada con el reporte e interpretación de los resultados del análisis de interpretabilidad.

5.1.1 Enunciado del problema de modelado

El planteamiento del problema de modelado abarca los pasos de modelado (i) descripción verbal y diagrama de flujo del proceso, (ii) definición del nivel de detalle del modelo y (iii) definición de los sistemas del proceso.

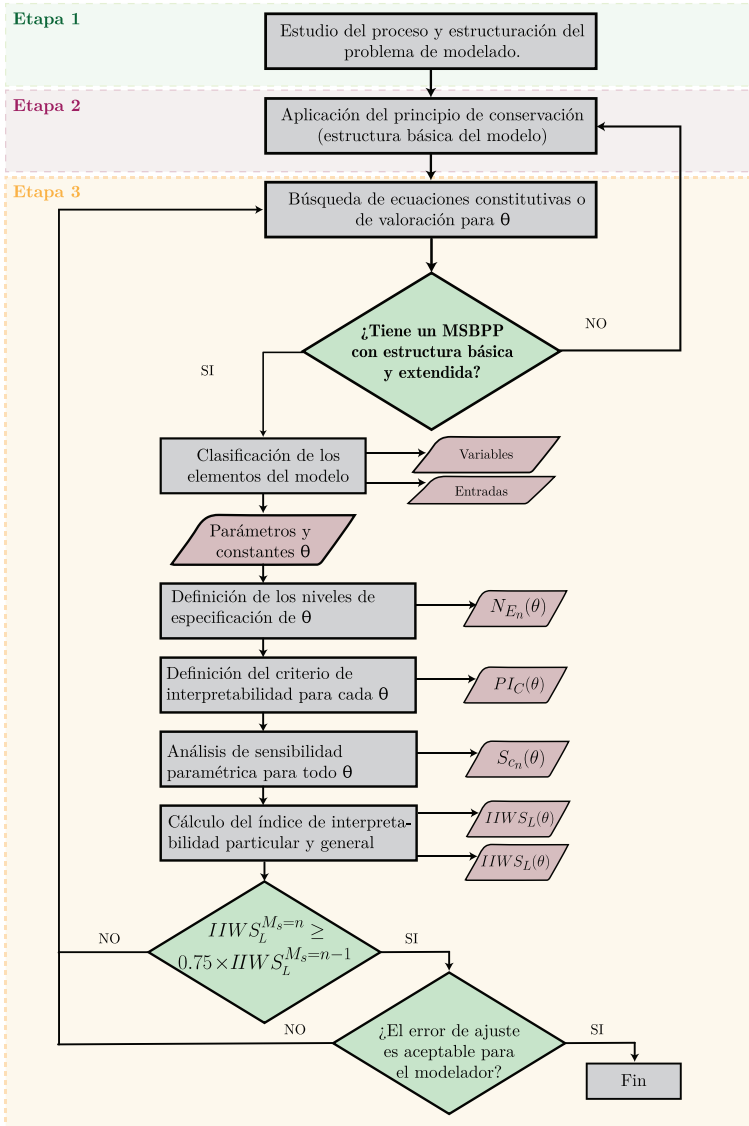


Figura 5.1: Metodología para el cálculo del índice de interpretabilidad de un MSBPP.

5.1.2 Deducción de la estructura básica del modelo

Comprende los pasos de modelado (iv) al (vi), en los que se proponen los balances dinámicos de las propiedades del sistema (masa, energía, cantidad de movimiento, etc.) y se seleccionan aquellos que aportan información valiosa para cumplir el objetivo de modelado.

5.1.3 Deducción de la estructura extendida del modelo

Comprende la etapa de modelado (paso vii), en la que se buscan ecuaciones constitutivas o de valoración para los parámetros de la estructura básica. En esta fase, es posible profundizar en los niveles de especificación del modelo matemático tanto como el modelador considere necesario. Cuando se definen parámetros funcionales mediante ecuaciones constitutivas o de valoración, es importante tener en cuenta su interpretabilidad. Esto significa que es importante analizar como afecta la inclusión de nuevas ecuaciones y parámetros las propiedades intrínsecas del modelo matemático, esto es, su sensibilidad, interpretabilidad e identificabilidad. En la metodología propuesta, se consideran como criterios de parada el índice de interpretabilidad y el error de ajuste del modelo. Se espera que de acuerdo con el objetivo de modelado, el modelador puede seguir profundizando en los niveles de especificación del modelo siempre que el índice de interpretabilidad $IIWS_L$ del escenario de modelado actual ($M_s = n$) sea mayor o igual que 75 % del índice del escenario inmediatamente anterior de acuerdo con el grafo de modelado. Es decir, la metodología propuesta acepta una degradación máxima del índice de interpretabilidad de 25 %. La evaluación del índice de interpretabilidad implica algunos pasos intermedios que comprenden el análisis de otras propiedades del modelo matemático. Los pasos intermedios son:

1. Definir el escenario de modelado a considerar de acuerdo con lo descrito en la sección 4.4.1.
2. Definir los niveles de especificación de cada parámetro, así como la normalización del mismo (ver Ecuación 4.4) para cada escenario de modelado planteado.
3. Definir el criterio de interpretabilidad de cada parámetro de acuerdo con la taxonomía propuesta en las secciones 4.2 y 4.4.3, para cada escenario de modelado planteado.

4. Realizar un análisis de sensibilidad paramétrica que permita clasificar los parámetros por orden de influencia en el resultado del modelo, así como la normalización del coeficiente de sensibilidad obtenido (ver ecuación 3.3).
5. Calcular el índice de interpretabilidad del modelo matemático en cada escenario, es decir, cada vez que aparece un nuevo nivel de especificación (ver ecuaciones 4.1-4.3).
6. Realizar el análisis de identificabilidad de los parámetros desconocidos del modelo para cada escenario (ver sección 3.1.2).

5.1.4 Reporte e interpretación de los resultados

Para el reporte e interpretación de los resultados, se propone la construcción de tablas y gráficas que faciliten el análisis de los resultados obtenidos para cada escenario de modelado.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IW S_I$
θ_1	ISP	1	0	1	1.5×10^{-2}	4×10^{-1}	2.2×10^{-1}
θ_2	PIPI	0.75	1	0.8	3.7×10^{-2}	1	1.5×10^{-1}
θ_3	NPI	0	1	0.8	3×10^{-2}	8.2×10^{-1}	3.1×10^{-2}
θ_4	IPI	0.9	3	0.3	1.3×10^{-2}	3.6×10^{-1}	5.4×10^{-2}
...	...	...	...	...	...	...	...
θ_n	GPI	1	2	0.5	n	n	n

Figura 5.2: Tabla de síntesis de resultados cuantitativos de los análisis realizados para cada escenario de modelado.

5.1.4.1. Resultados cuantitativos

En cuanto a los resultados cuantitativos, es importante documentar de manera adecuada los resultados de cada análisis realizado. La Figura 5.2 ilustra, a manera de ejemplo, como debe realizarse el reporte y síntesis de los resultados obtenidos a partir del análisis de sensibilidad, la clasificación de los parámetros de acuerdo con la taxonomía propuesta en el capítulo anterior, la definición del nivel de especificación en el que aparece cada parámetro con su respectiva normalización (Ecuación 4.4), la asignación de criterio de interpretabilidad paramétrica correspondiente a cada parámetro (ver Ecuación 4.5), el resultado del análisis de sensibilidad (ver Ecuación 3.3) con su respectiva normalización (Ecuación 4.6). La

		Escenarios de modelado			
Suma de todos los parámetros interpretables del modelo (ISP, GPI, IPI, y PIPI)		$M_s = 0$	$M_s = 1$	...	$M_s = n$
Parámetros interpretables		13	17	15	valor
Parámetros no interpretables		1	1	0	valor
$IWS_L(\theta)$		1	0.83	0.9	valor
Cambio en el valor del índice			-17%	+9%	%

% de cambio del escenario actual, respecto al escenario anterior de acuerdo con el grafo de modelado.

El índice de interpretabilidad en el escenario 0 (estructura básica) siempre es 1.

Figura 5.3: Tabla comparativa de resultados cuantitativos para cada escenario de modelado.

construcción de la tabla de síntesis, facilita el análisis de los resultados, la depuración de la metodología de búsqueda de ecuaciones constitutivas o de valoración, y el correcto reporte de resultados para la comunidad

científica. Una vez construida la tabla de síntesis y calculado el índice general del modelo matemático para todos los escenarios propuestos, se sugiere la construcción de una tabla comparativa (ver Figura 5.3) que agrupe tanto el número de parámetros interpretables y no interpretables, como el índice de interpretabilidad general para cada escenario de modelado, con el fin de poder calcular el cambio porcentual de un escenario respecto a al escenario anterior correspondiente en el grafo construido.

5.1.4.2. Resultados cualitativos

Los resultados cuantitativos reportados en la tabla de síntesis, dan lugar a una serie de gráficos que permiten visualizar de manera cualitativa la evolución del grado de interpretabilidad de cada parámetro, así como la forma como se ve modificada la jerarquización de los parámetros respecto al índice de interpretabilidad y el índice de sensibilidad. La Figura 5.4

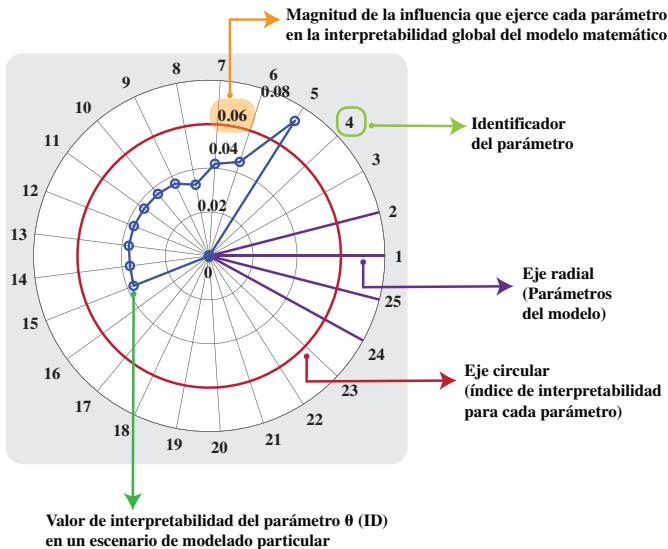


Figura 5.4: Gráfico de radar para la evaluación del grado de influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado.

ilustra, a manera de ejemplo, un gráfico de radar que permite visualizar como varía el grado de influencia de un parámetro en particular a través de las diferentes estructuras extendidas (escenarios de modelado) planteadas para un caso de aplicación particular. En este tipo de gráfico se relacionan la suma total de los parámetros presentes en todos los escenarios de modelado (identificados en el eje radial), con el grado de influencia de cada uno de ellos a la interpretabilidad general del modelo matemático (magnitud en el eje circular). Este tipo de gráfico permite evaluar de manera comparativa, como un mismo parámetro puede ser más o menos interpretable en una estructura extendida dada, para con base en ello considerar posibles variaciones estructurales en el modelo matemático. Finalmente, la Figura 5.5 ilustra a manera de ejemplo la cons-

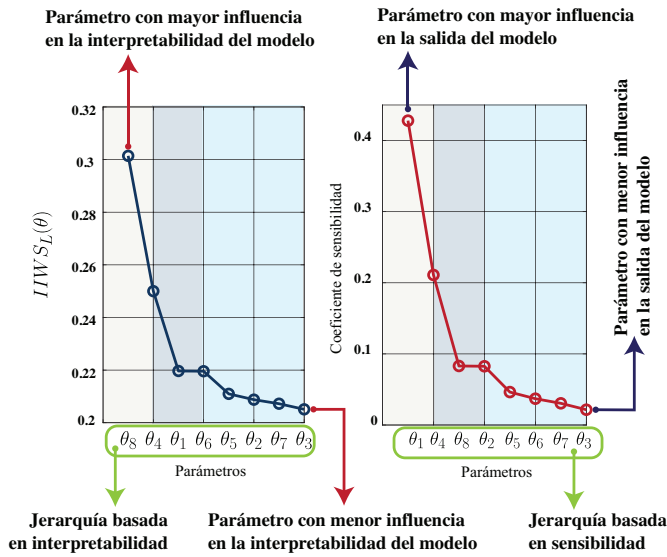


Figura 5.5: Gráfica de jerarquización paramétrica con base en el índice de interpretabilidad y el coeficiente de sensibilidad de cada parámetro.

trucción de los gráficos de jerarquización paramétrica basados tanto en el índice de interpretabilidad, como en el coeficiente de sensibilidad. Es

importante mencionar que estos gráficos son fáciles de construir a partir de la información compilada en la tabla de síntesis (ver Figura 5.2). La comparación de ambas jerarquías, permite obtener conclusiones importantes sobre la estructura del modelo. Por ejemplo, en los MSBPP, es indeseable la presencia de parámetros con un coeficiente de sensibilidad muy alto y un índice de interpretabilidad bajo, dado que este tipo de parámetros podrían dar lugar a mal condicionamiento matemático del modelo y falta de identificabilidad estructural.

5.2 Ejemplos de aplicación

Para ilustrar el funcionamiento del índice de interpretabilidad propuesto en esta tesis, se consideran dos casos de aplicación con distintos grados de complejidad. (i) En primer lugar, se considera el caso de un reactor continuo de tanque agitado (CSTR por sus siglas en inglés) para el que se proponen tres escenarios diferentes cambiando la ecuación constitutivas del coeficiente global de transferencia de calor. (ii) En segundo lugar, se aplica la metodología propuesta a una versión ampliada del modelo mínimo de Bergman (modelo mínimo de glucosa). En este caso, se proponen tres escenarios de modelado en los que se modifican las ecuaciones que definen la tasa de aparición de glucosa en el sistema portal y la tasa de producción de glucosa hepática. Esta sección incluye información general sobre los casos de aplicación planteados, así como la especificación de los escenarios de modelado considerados. El análisis de la interpretabilidad se presenta en la siguiente sección.

5.2.1 Caso de estudio: CSTR

Para el primer caso de aplicación, se utilizó el modelo matemático de un reactor continuo de tanque agitado (CSTR, por sus siglas en inglés) desarrollado en (Bequette, 1998, 2002). El ejemplo proporciona un nivel adecuado de complejidad para probar el índice propuesto y demostrar su funcionalidad en modelos MSBPP. Las Ecuaciones 5.1-5.2 muestran los balances de masa y energía resultantes después de aplicar el principio de conservación al CSTR, donde C_A es la concentración de reactivo y T es la temperatura en el reactor. El coeficiente global de transferencia de calor

$(U[\frac{J}{sm^2K}])$ se utilizó para generar los diferentes escenarios de prueba del índice de interpretabilidad propuesto, como se describe en las siguientes subsecciones. Todos los parámetros y constantes del modelo se describen en la Tabla 5.1. La Figura 5.6 ilustra el orden jerárquico de los diferentes escenarios de modelado del modelo del CSTR que se consideraron para el análisis de la interpretabilidad del modelo matemático.

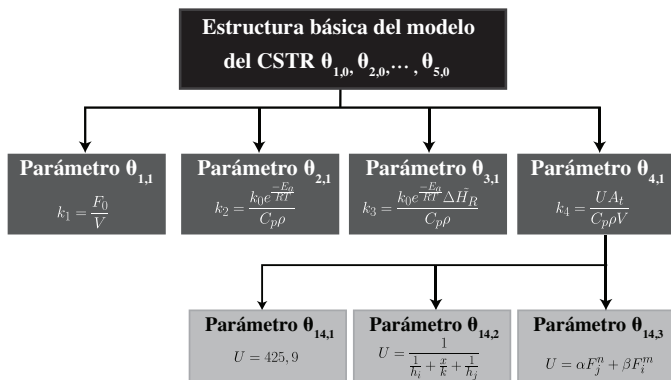


Figura 5.6: Grafo de modelado del CSTR.

5.2.1.1. Escenario 0: estructura básica

En sus trabajos de 1998 y 2002, Bequette dedujo un modelo matemático de un CSTR, basado en las hipótesis de que tanto el reactor como la camisa tenían una mezcla perfecta, el volumen del reactor era constante, los parámetros de diseño mantenían valores constantes, y se manipulaba la temperatura de la camisa (Bequette, 1998, 2002). A partir del modelo propuesto por el autor, es posible deducir la estructura básica del modelo que no considera la obtención de ecuaciones constitutivas para las tasas de transferencia k_1 , k_2 , k_3 , y k_4 como se muestra en las Ecuaciones 5.3-5.4.

$$\frac{dC_A}{dt} = k_1 C_{A_0} - k_1 C_A - k_2 C_A \quad (5.1)$$

$$\frac{dT}{dt} = k_1 T_0 - k_1 T - k_2 C_A - k_3 T + k_4 T_{0_j} + k_4 T \quad (5.2)$$

5.2. EJEMPLOS DE APLICACIÓN

Tabla 5.1: Parámetros del modelo matemático del CSTR utilizados en todos los escenarios considerados.

<i>ID</i>	θ	Descripción	Ecuación constitutiva	Unidades
1	k_1	Constante de velocidad de reacción	$\frac{F_0}{V}$	$\frac{1}{min}$
2	k_2	Constante de velocidad de reacción	$\frac{k_0 e^{-\frac{E_a}{RT}}}{C_p \rho}$	Adimensional
3	k_3	Constante de velocidad de reacción	$\frac{k_0 e^{-\frac{E_a}{RT}} \Delta \tilde{H}}{C_p \rho}$	Adimensional
4	k_4	Constante de velocidad de reacción	$\frac{U A_t}{C_p \rho V}$	Adimensional
5	T_{0_j}	Temperatura de entrada del fluido térmico a la camisa	282,96	K
6	R	Constante universal de los gases	8,32	$\frac{J}{gmol-K}$
7	E_a	Energía de activación	75361,14	$\frac{J}{gmol}$
8	k_0	Factor de frecuencia de la ley de Arrhenius	$2,83 \times 10^{11}$	$\frac{1}{min}$
9	A_t	Área de transferencia de calor	8,2	m^2
10	V	Volumen de masa reactiva en el reactor	2,41	m^3
11	F_0	Flujo de entrada del reactivo	0,16	$\frac{m^3}{min}$
12	ρ	Densidad del fluido de masa reactiva	1000	$\frac{kg}{m^3}$
13	C_p	Capacidad calorífica del fluido másico reactivo	$3,57 \times 10^3$	$\frac{J}{kg}$
14	U	Coefficiente global de transferencia de calor	Ecuaciones	$\frac{J}{s} m^2 K$
15	$\Delta \tilde{H}$	Calor de reacción del producto	$-9,0712 \times 10^{-4}$	$\frac{J}{mol}$

<i>ID</i>	θ	Descripción	Ecuación constitutiva	Unidades
16	h_i	Coeficiente de transferencia de calor en el reactor	8×10^2	$\frac{W}{m^2 K}$
17	h_j	Coeficiente de transferencia de calor en la camisa	$14,5 \times 10^2$	$\frac{W}{m^2 K}$
18	x	Espesor de la pared del reactor	0,004	m
19	k	Conductividad térmica	16,3	$\frac{W}{m K}$
20	F_j	Constante de Flujo de la camisa	0,38	$\frac{m^3}{min}$
21	F_i	Constante de flujo de reacción	0,24	$\frac{m^3}{min}$
22	α	Constante de velocidad	0,5	Adimensional
23	β	Constante de velocidad	0,1	Adimensional
24	n	Constante de velocidad	1	Adimensional
25	m	Constante de velocidad	1	Adimensional

5.2.1.2. Escenario 1: parámetro U constante

El primer escenario de modelado de la estructura extendida del CSTR consistió en definir la dinámica no modelada contenida en los parámetros k_1 , k_2 , k_3 y k_4 mediante la definición de ecuaciones constitutivas o de valoración. Para este caso, se establece el coeficiente global de transferencia de calor (contenido en el parámetro k_4) como un parámetro constante, es decir, U se define mediante una ecuación constitutiva cuyo valor es $U = 425,9 \frac{J}{sm^2 K}$. En este caso, el parámetro es interpretable, pero su ajuste debe ser tal que los resultados puedan reproducir las curvas de concentración y temperatura esperadas. Los parámetros r y $\dot{Q}_r$ se describen mediante las ecuaciones 5.5 y 5.6, respectivamente. Los demás parámetros y constantes del modelo se presentan en la Tabla 5.1. En ella se ofrece una descripción de los parámetros, junto con los valores y

unidades considerados a efectos de la simulación.

$$\frac{dC_A}{dt} = \frac{F_0}{V} C_{A_0} - \frac{F}{V} C_A - r C_A \quad (5.3)$$

$$\frac{dT}{dt} = \frac{F_0 T_0}{V} - \frac{F T}{V} - \frac{1}{C_p \rho} \dot{Q}_r - \frac{1}{C_p \rho V} (U A_t (T - T_{0j})) \quad (5.4)$$

$$r = k_0 e^{-\frac{E_a}{RT}} \quad (5.5)$$

$$\dot{Q}_r = r \Delta \tilde{H} \frac{1}{C_p} \quad (5.6)$$

5.2.1.3. Escenario 2: parámetro U descrito mediante la ecuación propuesta en (Bondy et al., 1983)

En el segundo escenario, se formuló una ecuación constitutiva para el coeficiente global de transferencia de calor, denotado como U . La relación matemática elegida es una ecuación fenomenológica derivada de primeros principios. Esta ecuación expresa explícitamente el parámetro U en función de variables clave como el coeficiente de transferencia de calor de la mezcla en el reactor h_i , el coeficiente de transferencia de calor h_j del fluido que circula por la camisa, el coeficiente de conductividad térmica del material k y el espesor de la pared del reactor, denominado x (Bondy et al., 1983).

$$U = \frac{1}{\frac{1}{h_i} + \frac{x}{k} + \frac{1}{h_j}} \quad (5.7)$$

Al utilizar la Ecuación 5.7, se obtiene un modelo caracterizado por nueve parámetros funcionales con interpretabilidad general y cuatro parámetros funcionales que poseen interpretabilidad inherente.

5.2.1.4. Escenario 3: parámetro U descrito mediante la ecuación propuesta en (Kern, 1950)

En el tercer escenario se formuló una ecuación constitutiva para el coeficiente global de transferencia de calor, denotado como U , de forma diferente, como se muestra en la Ecuación 5.8, pero siguiendo un enfoque

similar al del escenario anterior. En este caso, la relación matemática elegida es una ecuación empírica que caracteriza el parámetro U en función del flujo por la camisa F_j y del flujo interno del reactor F_i (debido a la agitación). Esta formulación adopta un supuesto no convencional en el que el interior del reactor se representa como una tubería con el fluido impulsado por una bomba que mueve a los reactantes dentro del reactor. Los parámetros m , n , α y β se introducen como coeficientes empíricos, carentes de interpretabilidad directa, pero esenciales para optimizar el ajuste del modelo a los datos de temperatura (Kern, 1950).

$$U = \alpha F_j^n + \beta F_i^m \quad (5.8)$$

Al utilizar la Ecuación 5.8 se obtuvo un modelo con ocho parámetros funcionales con interpretabilidad general, dos parámetros funcionales con interpretabilidad heredada, un parámetro funcional con interpretabilidad parcialmente heredada y cuatro parámetros no interpretables.

5.2.2 Caso de estudio: modelo mínimo de glucosa

Como segundo caso de estudio se utilizó una versión modificada del modelo mínimo de glucosa oral subcutánea triangular doble (dSOGMM) presentado en (Corbett et al., 2020). Esta versión del modelo mínimo es un modelo bicompartimental que representa las interacciones glucosa-insulina, considerando el subsistema triangular de comida e insulina. El primer compartimento representa la concentración plasmática de glucosa G variable en el tiempo y el segundo predice una tasa de crecimiento inversamente proporcional a la concentración efectiva de insulina en el compartimento remoto X . El modelo incluye también compartimentos para el consumo oral de una comida mixta, siendo Q_1 la masa de glucosa en el estómago y Q_2 la masa de glucosa en el intestino delgado.

El modelo mínimo triangular doble también considera la infusión subcutánea de insulina, representando de buena manera la farmacocinética de la insulina tras la inyección subcutánea. En este sentido, tiene un primer compartimento donde se inyecta la insulina I_{sc1} , y luego la insulina pasa a un segundo compartimento I_{sc2} antes de llegar al plasma. La insulina plasmática I se considera una variable de estado. La ingesta de comida mixta $meal$ y la señal de inyección de insulina $j(t)$ son las entradas del

modelo. El objetivo de este modelo era crear perfiles individualizados de

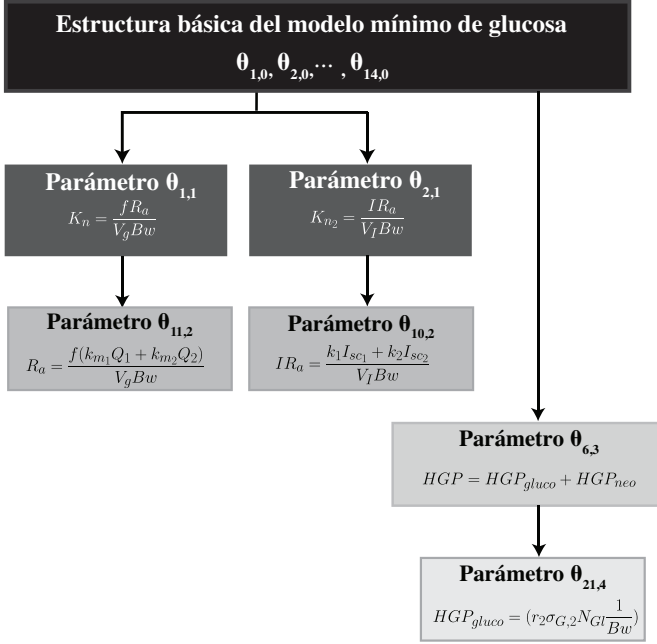


Figura 5.7: Grafo de modelado del modelo mínimo de glucosa.

alteración de las comidas que pudieran utilizarse en el contexto de un controlador predictivo basado en modelo para estabilizar los niveles de glucosa en sangre, anticipar la aparición de comidas y mejorar el control de la glucosa en sangre posprandial en personas con Diabetes Mellitus Tipo I. Sin embargo, este enfoque por compartimentos no es apropiado para proporcionar interpretabilidad directa a algunos parámetros del modelo, como se discutirá más adelante. En la Tabla 5.2 se detallan los parámetros de todos los escenarios considerados para calcular el índice de interpretabilidad. Además, a continuación se introducen las ecuaciones correspondientes a cada escenario. La Figura 5.7 ilustra el orden jerárquico de los diferentes escenarios de modelado para el caso de aplicación del modelo mínimo de glucosa.

Tabla 5.2: Parámetros del modelo mínimo considerados en todos los escenarios analizados.

<i>ID</i>	θ	Descripción	Ecuación constitutiva	Unidades
1	K_n	Tasa de aparición de glucosa en el sistema portal.	$\frac{Ra}{V_g \cdot Bw}$	$\frac{\mu mol}{kg-min}$
2	K_{n_2}	Aparición de insulina en plasma.	$\frac{IRa}{V_I \cdot Bw}$	$\frac{\mu mol}{kg-min}$
3	S_I	Sensibilidad a la insulina	1×10^{-4}	$\frac{1}{min} \text{ per } \frac{mU}{L}$
4	I_b	Insulina basal	0	mU/L
5	S_g	Efectividad fraccional de glucosa.	0,01	$\frac{1}{min}$
6	HGP	Producción hepática de glucosa	Ecuación	$\frac{\mu mol}{kg}$
7	p_2	Constante de velocidad del compartimento remoto de insulina.	0,02	$\frac{1}{min}$
8	n	Constante de velocidad.	0,178	$\frac{1}{min}$
9	f	Fracción de absorción intestinal.	0,9	Adimensional
10	IR_a	Velocidad de aparición de la insulina en el plasma.	$k_1 I_{sc1} + k_2 I_{sc2}$	mU/min
11	R_a	Velocidad de aparición de la glucosa en el sistema portal.	$k_1 I_{sc1} + k_2 I_{sc2}$	mg/min
12	V_I	Volumen de distribución de insulina.	0,118	$\frac{1}{kg-L}$
13	V_g	Volumen de distribución de glucosa.	1,7	dL/kg
14	Bw	Peso corporal.	70	kg
15	k_1	Constante de velocidad de absorción de la insulina no monomérica.	0,0018	$\frac{1}{min}$

ID	θ	Descripción	Ecuación constitutiva	Unidades
16	k_2	Constante de velocidad de absorción de la insulina no monomérica.	0,0182	$\frac{1}{min}$
17	k_d	Constante de velocidad de disociación de la insulina.	0,0164	$\frac{1}{min}$
18	k_{m_1}	Constante de velocidad del transporte subcutáneo de insulina.	0,02	$\frac{1}{min}$
19	k_{m_2}	Constante de velocidad transporte subcutáneo de insulina.	0,01	$\frac{1}{min}$
20	k_{m_d}	Constante de velocidad transporte subcutáneo de insulina.	0,015	$\frac{1}{min}$
21	HGP_{gluc}	Producción hepática de glucosa a partir de la glucogenólisis.	Ecuación	$\frac{\mu mol}{kg-min}$
22	HGP_{neo}	Producción hepática de glucosa a partir de la gluconeogénesis.	2,27	$\frac{\mu mol}{kg-min}$
23	r_2	Ritmo de producción de glucosa a partir de la glucogenólisis.	$8,3 \times 10^{-8}$	$\frac{\mu mol}{min}$
24	$\sigma_{G,2}$	Coeeficiente estequiométrico para la producción de glucosa en la reacción de glucogenólisis.	1	Adimensional
25	$N_{Glg n II}$	Moles de glucógeno en los hepatocitos.	$1,68 \times 10^5$	$mmol/L$

5.2.2.1. Escenario 0: estructura básica

El escenario 0 presenta la estructura básica del modelo mínimo utilizado en este estudio. La estructura básica se deriva considerando los balances realizados en los compartimentos descritos y se muestra en las ecuaciones 5.9 - 5.15. Es importante señalar que una de las principales dificultades en el uso de los modelos matemáticos reportados en la literatura es la falta de especificación sobre como se obtuvieron las ecuaciones matemáticas. En la estructura básica, los parámetros k_{n_1} y k_{n_2} encapsulan las dinámicas de aparición de la glucosa y de la insulina, respectivamente, tanto en el plasma como en el sistema portal.

$$\frac{dG}{dt} = -S_g G + HGP - X S_I G + K_n \quad (5.9)$$

$$\frac{dX}{dt} = -p_2 X + p_2 (I - I_b) \quad (5.10)$$

$$\frac{dQ_1}{dt} = -k_{m_1} + k_{m_d} Q_1 + meal \quad (5.11)$$

$$\frac{dQ_2}{dt} = -k_{m_2} Q_2 + k_{m_d} Q_1 \quad (5.12)$$

$$\frac{dI_{sc1}}{dt} = -(k_1 + k_d) I_{sc1} + j \quad (5.13)$$

$$\frac{dI_{sc2}}{dt} = -k_2 I_{sc2} + k_d I_{sc1} \quad (5.14)$$

$$\frac{dI}{dt} = -nI + K_{n_2} \quad (5.15)$$

5.2.2.2. Escenario 1: definición de los parámetros K_n y K_{n_2} mediante ecuaciones constitutivas

El escenario 1 consistió en expresar mediante ecuaciones los parámetros K_n y K_{n_2} relativos al volumen de distribución de la insulina y la glucosa (V_I y V_g , respectivamente), el peso corporal, la absorción intestinal de la glucosa y la aparición de insulina en el plasma. En este escenario, se describió el nivel de especificación 1, lo que da lugar a la aparición de los primeros parámetros funcionales. Las ecuaciones que representan la glucosa plasmática y la insulina plasmática se modifican sustituyendo K_n y K_{n_2} como se muestra en las ecuaciones 5.16 y 5.17. Las ecuaciones 5.10

a 5.14 se mantuvieron sin cambios.

$$\frac{dG}{dt} = -S_g G + HGP - X S_I G + \frac{f R_a}{V_g Bw} \quad (5.16)$$

$$\frac{dI}{dt} = -nI + \frac{I R_a}{V_I Bw} \quad (5.17)$$

5.2.2.3. Escenario 2: definición R_a e $I R_a$ mediante ecuaciones constitutivas

El escenario 2 consistió en definir los parámetros R_a e $I R_a$ mediante ecuaciones constitutivas. Las nuevas ecuaciones introducidas incluían factores relacionados con las tasas de transporte, disociación y absorción de la insulina. Las ecuaciones 5.16 y 5.17 cambiaron como se muestra en las ecuaciones 5.18 y 5.19.

$$\frac{dG}{dt} = -S_g G + HGP - X S_I G + \frac{f(k_{m_1} Q_1 + k_{m_2} Q_2)}{V_g Bw} \quad (5.18)$$

$$\frac{dI}{dt} = -nI + \frac{k_1 I_{sc_1} + k_2 I_{sc_2}}{V_I Bw} \quad (5.19)$$

Las ecuaciones constitutivas que definen R_a e $I R_a$ están en términos de parámetros y variables de estado conocidos, por lo que no se generan nuevos parámetros ni se clasifican como $S_L = 2$. Sin embargo, el análisis de interpretabilidad es diferente porque el número de parámetros a analizar cambia en comparación con los escenarios 0 y 1.

5.2.2.4. Escenario 3: parámetros HGP y R_a definidos mediante ecuaciones constitutivas

Los escenarios anteriores consideraban el parámetro HGP como un parámetro constante. En este escenario, se consideró definir la producción hepática de glucosa en términos de procesos bioquímicos de glucogenólisis y gluconeogénesis, generando dos nuevos parámetros en el modelo. La producción hepática de glucosa procedente de la descomposición del glucógeno en el hígado se representa por HGP_{gluco} y la producción hepática de glucosa procedente de precursores no carbohidratos se simboliza como HGP_{neo} . La única ecuación que se modifica en este caso es la Ecuación

5.18 que da lugar a la Ecuación 5.20.

$$\frac{dG}{dt} = -S_g G + HGP_{gluco} + HGP_{neo} - XS_I G + \frac{f(k_{m_1} Q_1 + k_{m_2} Q_2)}{V_g Bw} \quad (5.20)$$

5.2.2.5. Escenario 4: parámetro HGP_{gluco} definido mediante una ecuación constitutiva

El escenario 4 introdujo la definición de HGP_{gluco} con una ecuación constitutiva basada en la formulación de Arrhenius. Sustituyendo HGP_{gluco} por su ecuación constitutiva, la glucosa plasmática queda definida como se muestra en la Ecuación 5.21.

$$\frac{dG}{dt} = -S_g G + (r_2 \sigma_{G,2} N_{Gl} \frac{1}{Bw}) + HGP_{neo} - XS_I G + \frac{f(k_{m_1} Q_1 + k_{m_2} Q_2)}{V_g Bw} \quad (5.21)$$

5.3 Análisis de interpretabilidad

En esta sección se analiza la interpretabilidad de los dos casos de estudio presentados en la sección anterior y para cada uno de los escenarios planteados. Para todos los casos, se evalúa la interpretabilidad de la estructura básica del modelo, la cual sirve como punto de referencia y comparación para diferentes escenarios de estructura extendida. Se asume que las dos primeras fases de la metodología ya han sido completadas para los dos casos de aplicación mencionados, por lo tanto, solo se aborda la fase de obtención de la estructura extendida del modelo matemático.

5.3.1 Análisis de interpretabilidad del reactor continuo de tanque agitado (CSTR)

De acuerdo con la metodología planteada, se realizó el análisis de sensibilidad, la clasificación de los parámetros y el cálculo del índice de interpretabilidad. En el Anexo A.1 se especifican las condiciones de simulación utilizadas para el cálculo de sensibilidades de los parámetros, así como la solución del modelo computacional. Con base en los resultados obtenidos, se construyeron las tablas de síntesis y comparación, así como los gráficos de radar y jerarquización para cada uno de los escenarios planteados como sigue.

5.3.1.1. Escenario 0: estructura básica

La Tabla 5.3 resume los resultados del análisis de interpretabilidad para el escenario 0. En ella se presenta la clasificación de los parámetros de la estructura básica del modelo del CSTR en términos de su interpretabilidad según la taxonomía propuesta, junto con su respectivo valor ordinal PI_C , el nivel de especificación S_L en el que se encuentra cada uno de los parámetros del modelo, el coeficiente de sensibilidad S_c obtenido para cada parámetro con la estructura dada, así como el aporte individual de cada parámetro al índice de interpretabilidad general del modelo matemático. Además, se asigna a cada parámetro un número de identificación ID que

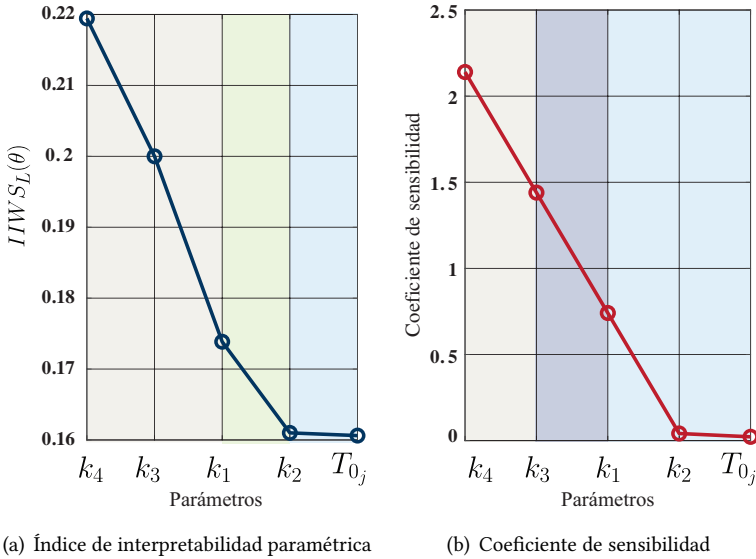


Figura 5.8: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 0.

facilita el análisis de los resultados cualitativos presentados. De acuerdo con los resultados presentados en la Tabla 5.3, el índice de interpretabilidad para la estructura básica o escenario 0 fue $IIW S_{L,0} = 1$. Este

valor actúa como base para la interpretabilidad, recordando que se calcula con los aportes de los valores reportados en la Tabla 5.3 y usando las ecuaciones 4.1 a 4.3.

Tabla 5.3: Análisis de interpretabilidad de un modelo del CSTR para su estructura básica.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_L(\theta)$
k_1	ISP	1	0	1	$7,4 \times 10^{-1}$	$3,5 \times 10^{-1}$	$1,7 \times 10^{-1}$
k_2	ISP	1	0	1	$4,1 \times 10^{-2}$	$1,9 \times 10^{-2}$	$1,6 \times 10^{-1}$
k_3	ISP	1	0	1	1,4	$6,7 \times 10^{-1}$	$1,9 \times 10^{-1}$
k_4	ISP	1	0	1	2,1	1	2×10^{-1}
$T_{0,j}$	ISP	1	0	1	$2,2 \times 10^{-2}$	$3,9 \times 10^{-3}$	$1,6 \times 10^{-1}$

La Figura 5.8 muestra la jerarquía de parámetros en el modelo del CSTR, basada tanto en el índice de interpretabilidad de cada parámetro (Figura 5.8(a)), como en la sensibilidad paramétrica de los mismos (Figura 5.8(b)). Cabe destacar que, aunque la estructura básica del modelo presenta el índice más alto de interpretabilidad general ($IIWS_L$), puede no ser la estructura óptima cuando el objetivo es representar las relaciones entre variables, parámetros y constantes dentro del sistema. En otras palabras, la estructura básica puede no ser propicia para inferir información relevante sobre los fenómenos subyacentes del sistema o proceso modelado. Sin embargo, el análisis de la estructura básica se incorpora a la metodología propuesta para que sirva como valor de base para comparar con precisión otros escenarios de modelado. En el caso del CSTR, la estructura básica carece de parámetros fundamentales necesarios para la generalización del modelo, como el volumen del reactor, la energía de activación y el coeficiente de transferencia de calor, entre otros. En consecuencia, es necesario encontrar ecuaciones constitutivas o de valoración para ampliar su estructura e incorporar estos parámetros esenciales. Este procedimiento produce varios escenarios posibles como se muestra a continuación.

5.3.1.2. Escenario 1: parámetro U constante

El análisis de interpretabilidad para el escenario 1 se muestra en la Tabla 5.4. De la misma manera que para el escenario anterior, la tabla contie-

ne los elementos necesarios para analizar el aporte individual de cada parámetro a la interpretabilidad general del modelo. En este escenario, el modelo contiene un parámetro estructural interpretable (ISP), ocho parámetros funcionales con interpretabilidad paramétrica general (GPI), dos parámetros funcionales con interpretabilidad paramétrica heredada (IPI) y cero parámetros no interpretables. De acuerdo con los resultados presentados en la Tabla 5.4, el valor del índice de interpretabilidad para el escenario 1 fue $IIWS_{L,1} = 0,51$. La Figura 5.9 ilustra la jerarquía

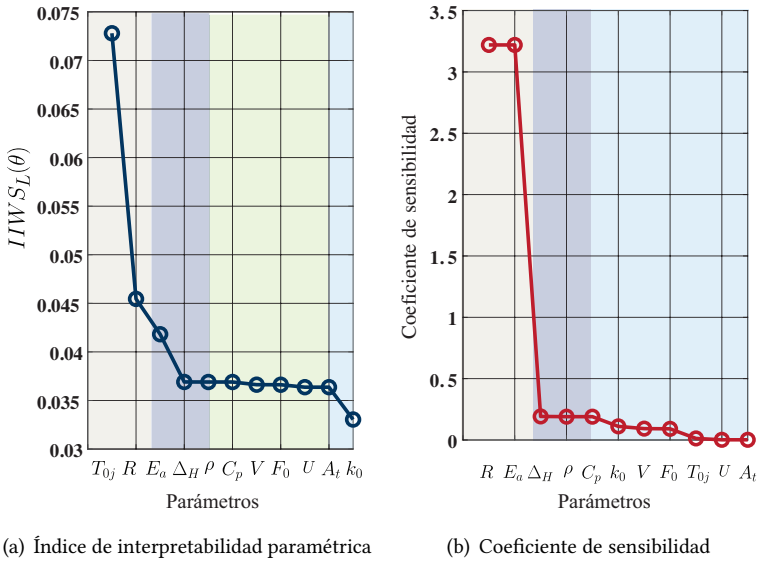


Figura 5.9: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 1.

de parámetros en el modelo del CSTR, basada tanto en el índice de interpretabilidad (Figura 5.9(a)) como en los coeficientes de sensibilidad paramétrica (Figura 5.9(b)). Como se observa en la figura, el parámetro con mayor influencia sobre la interpretabilidad del modelo general es el parámetro T_{0j} que representa la temperatura de entrada del fluido

térmico a la camisa. Es interesante que este mismo parámetro tiene una influencia menos significativa en la salida del modelo. Este tipo de

Tabla 5.4: Análisis de interpretabilidad de un modelo del CSTR considerando los parámetros funcionales de las nuevas ecuaciones constitutivas del escenario 1.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_l(\theta)$
T_{0j}	ISP	0	0	1	$1,3 \times 10^{-2}$	$3,9 \times 10^{-3}$	$7,3 \times 10^{-2}$
R	GPI	1	1	0,5	3,2	1	$4,5 \times 10^{-2}$
E_a	IPI	0,9	1	0,5	3,2	1	$4,2 \times 10^{-2}$
k_0	IPI	0,9	1	0,5	$1,1 \times 10^{-1}$	$3,4 \times 10^{-2}$	$3,3 \times 10^{-2}$
A_t	GPI	1	1	0,5	$1,6 \times 10^{-3}$	$4,9 \times 10^{-4}$	$3,6 \times 10^{-2}$
V	GPI	1	0	1	$9,3 \times 10^{-2}$	$2,9 \times 10^{-2}$	$3,7 \times 10^{-2}$
F_0	GPI	1	0	1	$9,1 \times 10^{-2}$	$2,8 \times 10^{-2}$	$3,7 \times 10^{-2}$
ρ	GPI	1	1	0,5	$1,9 \times 10^{-1}$	$5,9 \times 10^{-2}$	$3,7 \times 10^{-2}$
C_p	GPI	1	1	0,5	$1,9 \times 10^{-1}$	$5,9 \times 10^{-2}$	$3,7 \times 10^{-2}$
U	GPI	1	1	0,5	$1,6 \times 10^{-3}$	$4,9 \times 10^{-4}$	$3,6 \times 10^{-2}$
$\Delta \tilde{H}$	GPI	1	1	0,5	$1,9 \times 10^{-1}$	6×10^{-2}	$3,7 \times 10^{-2}$

relación entre interpretabilidad y sensibilidad no es deseable, ya que los parámetros de diseño deberían tener una fuerte influencia en la salida del modelo. Los resultados obtenidos sugieren la posibilidad de mejorar la estructura del modelo, de manera tal que, los parámetros con mayor influencia sobre la salida del modelo (coeficiente de sensibilidad más alto), sean aquellos con mayor grado de interpretabilidad. Esto último es deseable, considerando que los parámetros interpretables, pueden ser fácilmente acotados y correctamente identificados.

Por otra parte, el parámetro k_0 es el que tiene una menor influencia en el índice de interpretabilidad general del modelo. Esta observación coincide con lo esperado, teniendo en cuenta que k_0 es un parámetro ajustado empíricamente con interpretabilidad parcialmente heredada. En términos de sensibilidad, el parámetro constante R aparece como el parámetro con mayor influencia sobre la salida del modelo. Este resultado es deseable, dado que R representa una constante universal, cuya interpretabilidad se encuentra garantizada. Uno de los resultados destacables tiene que ver con el parámetro A_t , que a pesar de ser un parámetro con interpretabilidad

general y un parámetro importante para el diseño de reactores, tiene el coeficiente de sensibilidad más bajo. Este resultado podría indicar errores en el condicionamiento matemático del modelo o la necesidad de describir nuevos niveles de especificación para aumentar la capacidad deductiva del modelo matemático.

5.3.1.3. Escenario 2: parámetro U descrito mediante una ecuación fenomenológica

El análisis de interpretabilidad para el escenario 2 se muestra en la Tabla 5.5, de la misma forma que en los escenarios anteriores. El valor del índice de interpretabilidad para el escenario 2 utilizando los parámetros considerados fue $IIWS_{L,2} = 0,59$. La Figura 5.10 ilustra la jerarquía de

Tabla 5.5: Análisis de interpretabilidad de un modelo del CSTR calculando el coeficiente global de transferencia de calor (U) mediante la ecuación presentada en (Bondy et al., 1983).

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_l(\theta)$
$T0j$	ISP	1	0	1	$2,5 \times 10^{-2}$	$7,5 \times 10^{-3}$	$5,7 \times 10^{-2}$
R	GPI	1	1	0,7	3,4	1	$4,8 \times 10^{-2}$
E_a	IPI	0,9	1	0,7	3,4	1	$4,4 \times 10^{-2}$
$k0$	IPI	0,9	1	0,7	$1,2 \times 10^{-1}$	$3,4 \times 10^{-2}$	$3,5 \times 10^{-2}$
A_t	GPI	1	1	0,7	3×10^{-3}	9×10^{-4}	$3,8 \times 10^{-2}$
V	GPI	1	0	1	$9,8 \times 10^{-2}$	$2,9 \times 10^{-2}$	$5,8 \times 10^{-2}$
F_0	GPI	1	0	1	$9,5 \times 10^{-2}$	$2,8 \times 10^{-2}$	$5,8 \times 10^{-2}$
ρ	GI	1	1	0,7	2×10^{-1}	$5,8 \times 10^{-2}$	$3,9 \times 10^{-2}$
C_p	GPI	1	1	0,7	2×10^{-1}	$5,8 \times 10^{-2}$	$3,9 \times 10^{-2}$
$\Delta\hat{H}$	GPI	1	1	0,7	2×10^{-1}	$5,9 \times 10^{-2}$	$3,9 \times 10^{-2}$
h_i	IPI	0,9	2	0,3	$2,5 \times 10^{-3}$	$7,3 \times 10^{-4}$	$1,7 \times 10^{-2}$
h_j	IPI	0,9	2	0,3	9×10^{-7}	$2,7 \times 10^{-7}$	$1,7 \times 10^{-2}$
x	GPI	1	2	0,3	$5,7 \times 10^{-4}$	$1,7 \times 10^{-4}$	$1,9 \times 10^{-2}$
k	GPI	1	2	0,3	$5,7 \times 10^{-4}$	$1,7 \times 10^{-4}$	$1,9 \times 10^{-2}$

parámetros en el modelo del CSTR, basada en el índice de interpretabilidad (Figura 5.10(a)) y el coeficientes de sensibilidad paramétrica (Figura 5.10(b)). La figura muestra cómo al cambiar la estructura extendida del modelo matemático también cambia el grado de influencia individual de los parámetros a la interpretabilidad general del modelo. En este escena-

rio, se igualó el grado de influencia de los parámetros T_{0j} , V y F_0 . Estos parámetros tienen interpretabilidad estructural o general (ISP o GPI) y un coeficiente de sensibilidad media. Los parámetros recién introducidos asociados a la ecuación fenomenológica para el parámetro U (ver Ecuación 5.7) son los que menos impacto tienen en el índice de interpretabilidad. En términos de sensibilidad, el parámetro R mantiene su posición como el parámetro más influyente sobre la salida del modelo, mientras que los parámetros dentro de la ecuación para el parámetro U muestran los coeficientes de sensibilidad más bajos.

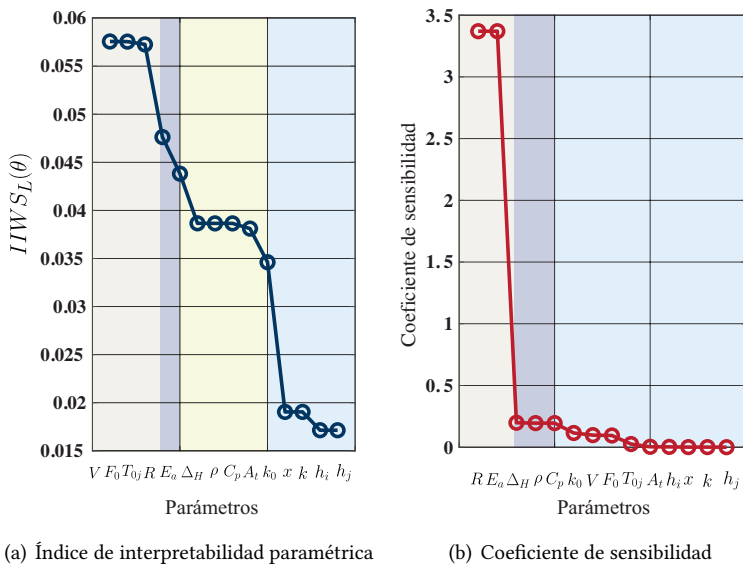


Figura 5.10: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 2.

5.3.1.4. Escenario 3: parámetro U descrito mediante una ecuación empírica

El análisis de interpretabilidad para el escenario 3 se muestra en la Tabla 5.6. El valor correspondiente del índice de interpretabilidad, calculado a partir de los elementos indicados en la tabla, fue $IIWS_{L,3} = 0,45$. La

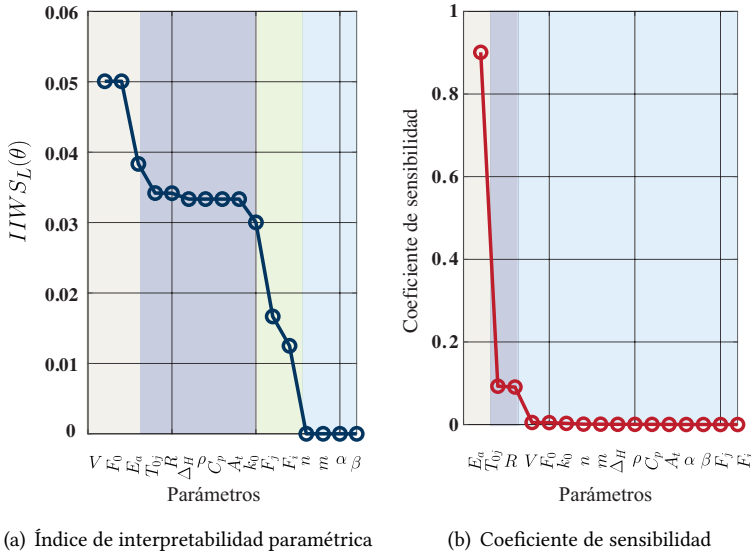


Figura 5.11: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo del CSTR en el escenario 3.

Figura 5.11 ilustra la jerarquía de los parámetros en el modelo del CSTR, basada en el índice de interpretabilidad (Figura 5.11(a)) y el coeficiente de sensibilidad paramétrica (Figura 5.11(b)). Para este escenario de modelado, el volumen (V) sigue siendo el parámetro que ejerce la mayor influencia sobre el índice de interpretabilidad del modelo. Por el contrario, los parámetros recién introducidos asociados a la ecuación empírica para el parámetro U (véase la Ecuación 5.8) son los que menos influyen

en el índice. En particular, se aprecia una alteración significativa en la clasificación de los parámetros relativos a la sensibilidad.

El parámetro R ocupa ahora la tercera posición, con un coeficiente de sensibilidad notablemente bajo. Finalmente, el parámetro E_a se convierte en el parámetro con mayor influencia sobre la salida del modelo. En este nuevo contexto, el índice permite dilucidar cómo se altera la sensibilidad paramétrica al incorporar a la estructura extendida del modelo parámetros que carecen de interpretabilidad. El índice de interpretabilidad no sólo permite evaluar las consecuencias de la sobreparametrización, sino que la penaliza explícitamente. Además, el índice también actúa como un criterio para evaluar el efecto no deseable de incluir parámetros como m y n , los cuales pueden mostrar una sensibilidad mayor que los parámetros de diseño como A_t , pero no proporcionan información relevante sobre los procesos subyacentes y la dinámica del proceso modelado.

Tabla 5.6: Análisis de interpretabilidad de un modelo del CSTR calculando el coeficiente global de transferencia de calor (U) mediante una ecuación empírica.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_I(\theta)$
$T_{0,j}$	ISP	1	0	1	$9,3 \times 10^{-2}$	1×10^{-1}	$3,4 \times 10^{-2}$
R	GPI	1	1	0,7	9×10^{-2}	1×10^{-1}	$3,4 \times 10^{-2}$
E_a	IPI	0,9	1	0,7	9×10^{-1}	1	$3,8 \times 10^{-2}$
k_0	IPI	0,9	1	0,7	3×10^{-3}	$3,3 \times 10^{-3}$	3×10^{-2}
A_t	GPI	1	1	0,7	$3,9 \times 10^{-4}$	$4,3 \times 10^{-4}$	$3,3 \times 10^{-2}$
V	GPI	1	0	1	$5,1 \times 10^{-3}$	$5,6 \times 10^{-3}$	5×10^{-2}
F_0	GPI	1	0	1	$4,8 \times 10^{-3}$	$5,4 \times 10^{-3}$	5×10^{-2}
ρ	GPI	1	1	0,7	$5,7 \times 10^{-4}$	$6,3 \times 10^{-4}$	$3,3 \times 10^{-2}$
C_p	GPI	1	1	0,7	$5,7 \times 10^{-4}$	$6,3 \times 10^{-4}$	$3,3 \times 10^{-2}$
ΔH	GPI	1	1	0,7	$6,4 \times 10^{-4}$	$7,1 \times 10^{-4}$	$3,3 \times 10^{-2}$
F_j	GPI	1	2	0,3	$1,8 \times 10^{-4}$	2×10^{-4}	$1,7 \times 10^{-2}$
F_i	PIPI	0,8	2	0,3	$1,8 \times 10^{-4}$	$1,9 \times 10^{-4}$	$1,3 \times 10^{-2}$
α	NIP	0	2	0,3	2×10^{-4}	$2,2 \times 10^{-4}$	$9,1 \times 10^{-7}$
β	NIP	0	2	0,3	$1,9 \times 10^{-4}$	$2,1 \times 10^{-4}$	$8,9 \times 10^{-7}$
n	NIP	0	2	0,3	$1,1 \times 10^{-3}$	$1,2 \times 10^{-3}$	5×10^{-6}
m	NIP	0	2	0,3	1×10^{-3}	$1,2 \times 10^{-3}$	$4,8 \times 10^{-6}$

5.3.1.5. Resultados comparativos del CSTR

Como se discute en la Sección 4.4, el índice de interpretabilidad, ponderado por el nivel de especificación, facilita la comparación de modelos MSBPP para el mismo proceso, incorporando variaciones en su estructura extendida y niveles de especificación sin introducir dinámicas modeladas adicionales. En el contexto del estudio de caso CSTR, puede realizarse un análisis comparativo de la interpretabilidad entre los modelos (como se detalla en las secciones 5.2.1.2, 5.2.1.3 y 5.2.1.4) utilizando el índice de interpretabilidad calculado. En este sentido, los gráficos de radar presentados

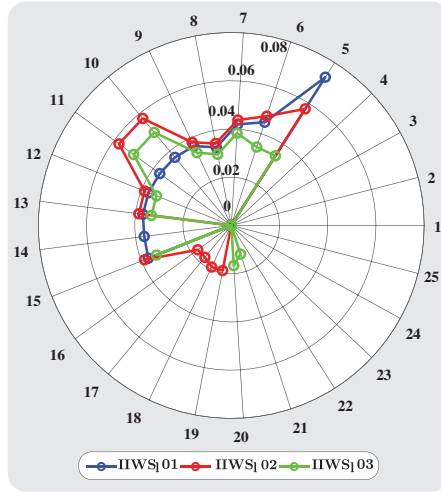


Figura 5.12: Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del CSTR.

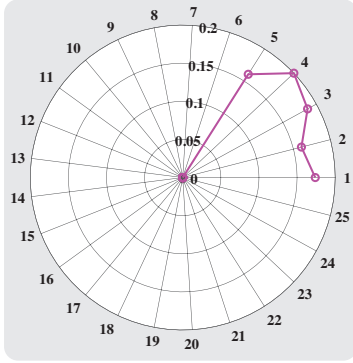
en las Figuras 5.13-5.12 ilustran la progresión del índice de interpretabilidad $IIWS_L(\theta)$ dentro de cada escenario de modelado. El ID $(1, 2, \dots, n)$ de cada parámetro en la figura corresponde al ID $(1, 2, \dots, n)$ en la Tabla 5.1. Un parámetro al que se le asigna un valor cero significa su ausencia en el escenario actual considerado. La Figura 5.13 muestra la progresión individual de la influencia de cada parámetro en el índice general de in-

terpretabilidad del modelo matemático para un CSTR en cada escenario considerado. En particular, los parámetros 1 a 4, que forman parte de la estructura básica del modelo matemático, presentan sistemáticamente el mayor índice de interpretabilidad. Tras el análisis comparativo de la Figura 5.12, se hace evidente que, aunque los parámetros 10 (V) y 11 (F_0) del escenario 1 muestran inicialmente un índice de interpretabilidad más bajo, en el segundo escenario éstos se benefician de la incorporación de nuevos parámetros con un significado físico (16 a 19). Estos parámetros adicionales aportan interpretabilidad al modelo global, mejorando en consecuencia su capacidad descriptiva. Por el contrario, el escenario 3 emerge como el menos favorable entre todos los escenarios. El índice de interpretabilidad de todos sus parámetros se ve afectado negativamente por la introducción de nuevos parámetros empíricos como F_j , F_i , α , β , n y m (20 a 25).

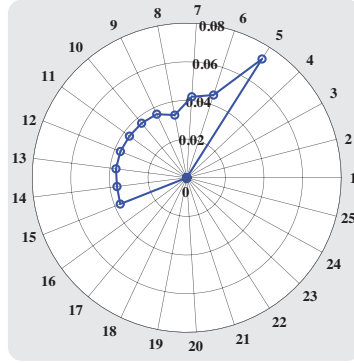
Tabla 5.7: Índice de interpretabilidad $IIWS_{L,M_s}$ para cada escenario de modelado del CSRT. El cambio en el valor del índice representa el grado de mejora (+) o degradación (−) del índice en comparación con el escenario de modelado correspondiente.

	$M_s = 0$	$M_s = 1$	$M_s = 2$	$M_s = 3$
Parámetros interpretables	5	11	14	12
Parámetros no interpretables	0	0	0	4
$IIWS_{L,M_s}$	1	0,51	0,59	0,45
Cambio en el valor del índice		−49 %	−41 %	−55 %

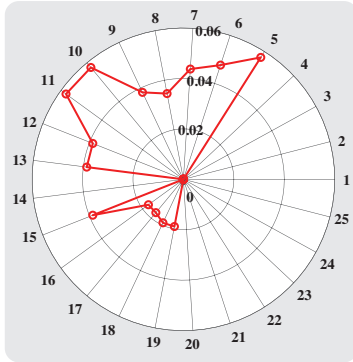
La Tabla 5.7 resume los resultados cuantitativos del análisis de interpretabilidad realizado para cada escenario de modelado. Como puede verse, en el escenario 1, el índice de interpretabilidad sufre una degradación natural de la interpretabilidad con respecto a la estructura básica del modelo. Sin embargo, como ya se ha mencionado, el escenario 0 carece de algunas relaciones que permitan la introducción de variables manipuladas o parámetros de diseño. Así pues, aunque el escenario 0 siempre es el más interpretable debido a que sus parámetros estructurales proceden de la aplicación de primeros principios, es necesario abrir nuevos niveles de especificación. En este sentido, para el caso de la estructura 2, la



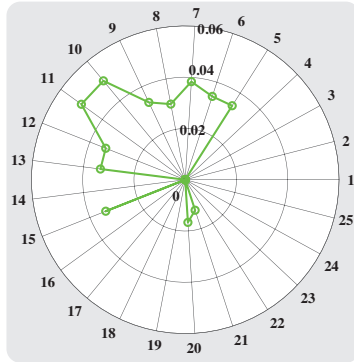
(a) $IIWS_L(\theta)$ en el escenario 0



(b) $IIWS_L(\theta)$ en el escenario 1



(c) $IIWS_L(\theta)$ en el escenario 2



(d) $IIWS_L(\theta)$ en el escenario 3

Figura 5.13: Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del CSTR.

interpretabilidad se enriquece con la inclusión de 6 nuevos parámetros interpretables. El escenario 3 resulta ser el menos favorable de todos, con una degradación del índice del 55 % respecto del escenario 0, que se explica por la adición de 4 nuevos parámetros no interpretables. Puede deducirse que la sobreparametrización no siempre empeora la calidad de los modelos matemáticos, sino que la adición de nuevos parámetros interpretables puede mejorar no sólo la estructura, sino también su ajuste. Por otra parte, el índice puede penalizar la sobreparametrización en términos de su contribución a la interpretabilidad del modelo, así como a la sensibilidad de sus parámetros.

Una vez analizada la interpretabilidad del modelo matemático, cobra una especial relevancia el análisis de identificabilidad de los escenarios planteados tal como se describió en la Sección 3.1.2 utilizando la herramienta STRIKE GOLD presentada en (Villaverde et al., 2016). La Tabla 5.8 resume el análisis de identificabilidad realizado para todos los escenarios de modelado. En primera instancia es importante recordar que el análisis de identificabilidad solo se aplica a los parámetros desconocidos del modelo que son susceptibles a ser identificados, esto es los vectores θ_{u,M_s} presentados en la tabla. Un hallazgo importante a partir de los resulta-

Tabla 5.8: Identificabilidad de los parámetros desconocidos del modelo matemático de un CSTR, para todos los escenarios de modelado.

	Vector de Parámetros desconocidos	Análisis de identificabilidad
$M_s = 0$	$\theta_{u,0} = [k_1 \quad k_2 \quad k_3 \quad k_4]$	Localmente estructuralmente identificables
$M_s = 1$	$\theta_{u,1} = [k_0 \quad E_a \quad U]$	Localmente estructuralmente identificables
$M_s = 2$	$\theta_{u,2} = [k_0 \quad E_a \quad h_i \quad h_j]$	k_0 y E_a localmente estructuralmente identificables h_i y h_j no identificables
$M_s = 3$	$\theta_{u,3} = [k_0 \quad E_a \quad F_i \quad \alpha \quad \beta \quad n \quad m]$	Estructuralmente no identificable

dos obtenidos es que la inclusión de parámetros no interpretables a la estructura extendida de un modelo, tiene una implicación directa en la identificabilidad de sus parámetros. En la tabla se observa como en el escenario 2, los parámetros h_i y h_j resultan ser no identificables, siendo estos parámetros los que exhiben una menor influencia en el índice de

interpretabilidad general del modelo (ver Figura 5.10). De manera similar, para el caso del escenario 3 la inclusión de un alto número de parámetros no interpretables (α , β , n , y m), no solo implicó su no identificabilidad, sino que además degradó la identificabilidad de los parámetros k_0 y E_a .

5.3.2 Análisis de interpretabilidad del modelo mínimo de glucosa

De acuerdo con la metodología planteada, se realizó el análisis de sensibilidad, la clasificación de los parámetros y el cálculo del índice de interpretabilidad para el modelo mínimo de glucosa. En el Anexo A.2 se especifican las condiciones de simulación utilizadas para el cálculo de sensibilidades de los parámetros, así como la solución del modelo computacional. Con base en los resultados obtenidos, se construyeron las tablas de síntesis y comparación, así como los gráficos de radar y jerarquización para cada uno de los escenarios planteados como sigue.

5.3.2.1. Escenario 0: estructura básica

El análisis de interpretabilidad para el escenario 2 se muestra en la Tabla 5.9. En ella se presenta la clasificación de los parámetros de la estructura básica del modelo mínimo de glucosa en términos de su interpretabilidad según la taxonomía propuesta, junto con su respectivo valor ordinal PI_C , el nivel de especificación S_L en el que se encuentra cada uno de los parámetros del modelo, el coeficiente de sensibilidad S_c obtenido para cada parámetro con la estructura dada, así como el aporte individual de cada parámetro al índice de interpretabilidad general del modelo. El valor correspondiente del índice de interpretabilidad, calculado a partir de los elementos indicados en la tabla, fue $IIS_{L,0} = 1$, como es de esperar. La Figura 5.14 ilustra la jerarquía de parámetros en el modelo mínimo de glucosa, tal y como se delinea mediante el índice de interpretabilidad de cada parámetro en la Figura 5.14(a), y los coeficientes de sensibilidad paramétrica en la Figura 5.14(b). La producción hepática de glucosa, HGP , afecta significativamente a las respuestas del modelo y, como era de esperar, tiene una fuerte influencia en la interpretabilidad general del modelo, considerando que es un parámetro con IPI. A pesar de su influencia en los resultados del modelo, el parámetro n carece de interpretabilidad.

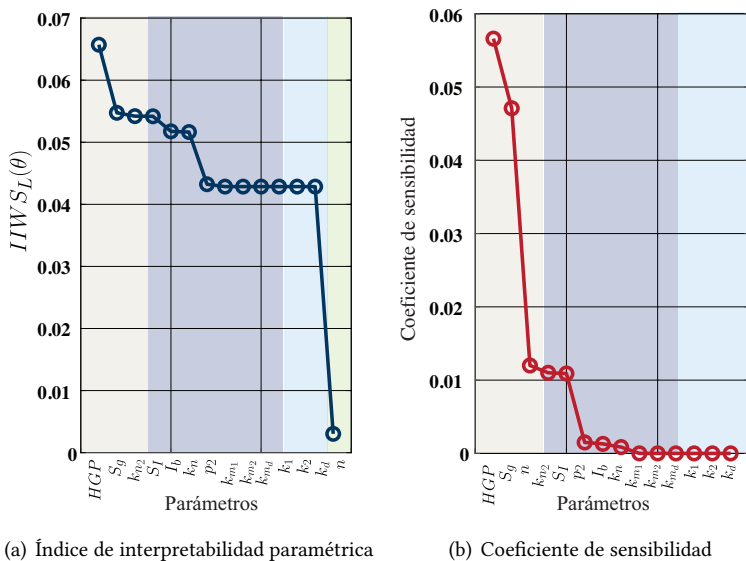


Figura 5.14: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 0.

Todas las tasas de absorción, transporte y disociación de la insulina han heredado parcialmente la interpretabilidad paramétrica, pero no influyen en las respuestas del modelo relativas a la dinámica de la glucosa, sino en la dinámica de la insulina. Este hecho sugiere que los sistemas de procesos locales se definieron aislados del sistema de procesos total, dando lugar a una dinámica desacoplada del sistema de insulina, lo que hace que la sensibilidad de estos parámetros sea nula. Posiblemente, el uso de un enfoque compartimental sea la causa de este resultado. Sin embargo, las ecuaciones diferenciales desacopladas se mantienen porque los autores del modelo consideran que estas variables dinámicas de estado representan el sistema digestivo y la farmacocinética de la insulina.

Tabla 5.9: Análisis de interpretabilidad del modelo mínimo de glucosa considerando su estructura básica.

θ	Tipo	PI_C	S_L	S_{Ln}	S_c	S_{cn}	$IW S_L(\theta)$
HGP	IPI	0,9	0	$5,7 \times 10^{-2}$	1	$6,5 \times 10^{-2}$	$5,5 \times 10^{-2}$
S_g	PIPI	0,75	0	1	$4,7 \times 10^{-2}$	$8,3 \times 10^{-1}$	
n	NPI	0	0	$1,2 \times 10^{-2}$	$2,1 \times 10^{-1}$	3×10^{-3}	$5,4 \times 10^{-2}$
K_{n2}	IPI	0,9	0	1	$1,1 \times 10^{-2}$	$1,9 \times 10^{-1}$	
S_I	IPI	0,9	0	$1,1 \times 10^{-2}$	$1,9 \times 10^{-1}$	$5,4 \times 10^{-2}$	$4,3 \times 10^{-2}$
p_2	PIPI	0,75	0	1	$1,5 \times 10^{-3}$	$2,7 \times 10^{-2}$	
I_b	IPI	0,9	0	$1,3 \times 10^{-3}$	$2,3 \times 10^{-2}$	$5,2 \times 10^{-2}$	$4,3 \times 10^{-2}$
K_n	IPI	0,9	0	$8,6 \times 10^{-4}$	$1,5 \times 10^{-2}$	$5,2 \times 10^{-2}$	
k_{m1}	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$
k_{m2}	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$
k_{md}	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$
k_1	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$
k_2	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$
k_d	PIPI	0,75	0	1	0	0	$4,3 \times 10^{-2}$

El parámetro S_g tiene un significado específico que sólo es válido en el contexto del sistema glucosa-insulina. En consecuencia, tiene un nivel relativamente alto de interpretabilidad y afecta significativamente a la dinámica de la glucosa plasmática. Aunque la sensibilidad a la insulina S_I no tiene una influencia tan fuerte como S_g , tiene el mismo nivel de peso de interpretabilidad en el modelo. En el contexto del modelo mínimo de glucosa, la estructura básica es deficiente en cuanto a disponer de parámetros medibles que permitan la generalización del modelo. En particular, están ausentes factores como el peso corporal y los volúmenes de distribución de la insulina y la glucosa. En consecuencia, es necesario formular ecuaciones constitutivas que amplíen el marco del modelo e incluyan este tipo de parámetros.

5.3.2.2. Escenario 1: definición de los parámetros K_n y K_{n2} mediante ecuaciones constitutivas

El análisis de interpretabilidad para el escenario 1 se muestra en la Tabla 5.10, de la misma forma que en los escenarios anteriores. En total, se analizaron 18 parámetros, 12 se mantuvieron del escenario 0 (K_n y K_{n2} desaparecen), y se generaron 6 nuevos parámetros en el escenario 1. Como era de esperar, el índice de interpretabilidad es inferior al del escenario 0 ($IW S_L = 0,83$) porque se sustituyeron dos parámetros clasificados como IPI en el escenario 0 para generar cinco parámetros clasificados

como PIPi, lo que proporciona un criterio de interpretabilidad inferior ($PI_C = 0,75$). La Figura 5.15 ilustra la jerarquía de parámetros en el modelo mínimo de glucosa, basada en el índice de interpretabilidad (Figura 5.15(a)) y el coeficiente de sensibilidad paramétrica (Figura 5.15(b)).

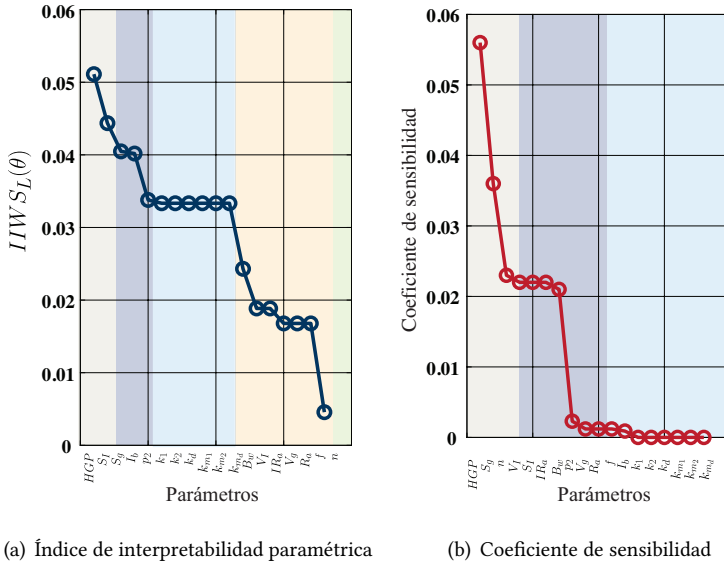


Figura 5.15: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 1.

El parámetro más interpretable y sensible en este escenario es HGP , mientras que las tasas de absorción, transporte y disociación de la insulina siguen sin influir en las respuestas del modelo, ya que R_a e IR_a son parámetros constantes, lo que significa que la dinámica de la insulina sigue desacoplada de la concentración de glucosa en plasma.

5.3.2.3. Escenario 2: definición R_a e IR_a mediante ecuaciones constitutivas

La Tabla 5.11 muestra la clasificación de los 16 parámetros analizados. El valor del índice de interpretabilidad para el escenario 2 utilizando los parámetros considerados fue $IIW S_{L,2} = 0,9$. Esto se debe a la reducción del número de parámetros analizados. En este escenario, se clasifican menos parámetros como PIPI, pero se mantiene el mismo número de

Tabla 5.10: Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 1.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_I(\theta)$
V_I	PIPI	0,75	1	0,5	$2,2 \times 10^2$	$3,9 \times 10^1$	$1,9 \times 10^2$
V_g	PIPI	0,75	1	0,5	$1,2 \times 10^{-3}$	$2,1 \times 10^{-2}$	$1,7 \times 10^{-2}$
Bw	GPI	1	1	0,5	$2,1 \times 10^{-2}$	$3,8 \times 10^{-1}$	$2,4 \times 10^{-2}$
S_I	IPI	0,9	0	1	$2,2 \times 10^{-2}$	$3,9 \times 10^{-1}$	$4,4 \times 10^{-2}$
S_g	PIPI	0,75	0	1	$3,6 \times 10^{-2}$	$6,4 \times 10^{-1}$	4×10^{-2}
p_2	PIPI	0,75	0	1	$2,3 \times 10^{-3}$	$4,1 \times 10^{-2}$	$3,4 \times 10^{-2}$
k_1	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$
k_2	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$
k_d	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$
n	NPI	0	0	1	$2,3 \times 10^{-2}$	$4,1 \times 10^1$	$4,6 \times 10^{-3}$
IR_a	PIPI	0,75	1	0,5	$2,2 \times 10^{-2}$	$3,9 \times 10^{-1}$	$1,9 \times 10^{-2}$
R_a	PIPI	0,75	1	0,5	$1,2 \times 10^{-3}$	$2,1 \times 10^{-2}$	$1,7 \times 10^{-2}$
HGP	IPI	0,9	0	1	$5,6 \times 10^{-2}$	1×10^0	$5,1 \times 10^{-2}$
I_b	IPI	0,9	0	1	9×10^{-4}	$1,6 \times 10^{-2}$	4×10^{-2}
f	PIPI	0,75	1	0,5	$1,2 \times 10^{-3}$	$2,14 \times 10^{-2}$	$1,7 \times 10^{-2}$
k_{m1}	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$
k_{m2}	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$
k_{md}	PIPI	0,75	0	1	0	0	$3,3 \times 10^{-2}$

parámetros IPI, GPI y NPI. Además, cuando R_a e IR_a se definen mediante ecuaciones constitutivas, todo el modelo está acoplado, lo que aumenta su interpretabilidad.

La Figura 5.16 ilustra la jerarquía de parámetros en el modelo mínimo de glucosa, basada por el índice de interpretabilidad (Figura 5.16(a)) y el coeficiente de sensibilidad paramétrica (Figura 5.16(b)). Como puede observarse, la clasificación paramétrica es similar a las clasificaciones anteriores. Sin embargo, en este caso, el modelo está completamente acoplado mediante la definición de los parámetros R_a e IR_a con ecuaciones constitutivas. El coeficiente de sensibilidad para las tasas de absorción, disociación y transporte de insulina refleja este acoplamiento.

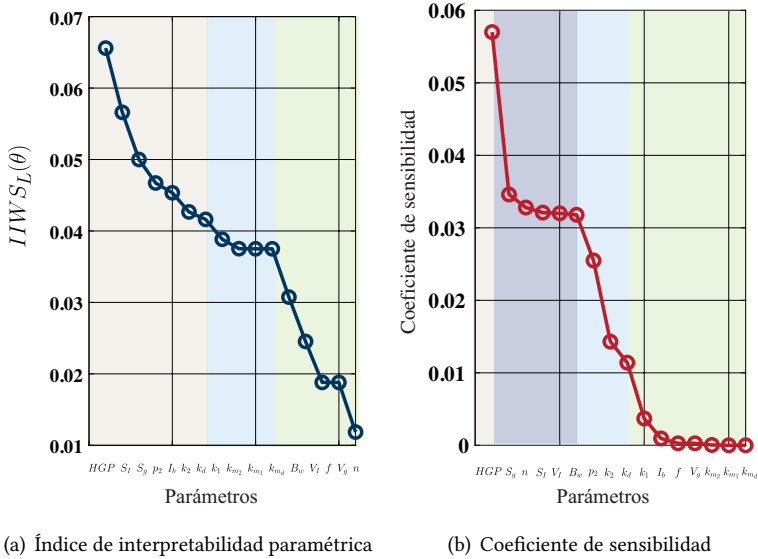


Figura 5.16: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 2.

5.3.2.4. Escenario 3: parámetros HGP y R_a definidos mediante ecuaciones constitutivas

El análisis de interpretabilidad para el escenario 3 se muestra en la Tabla 5.12. En ella se presenta el análisis de 17 parámetros, incluidos los dos nuevos parámetros HGP_{gluco} y HGP_{neo} generados tras definir HGP . Estos nuevos parámetros se clasifican como parámetros con interpretabilidad heredada (IPI) y con un nivel de especificación $S_L = 1$, por lo que en este escenario, el índice de interpretabilidad sigue siendo similar al del escenario 1, con un valor de $IIW S_{L,3} = 0,83$. Esto se debe a que el criterio de interpretabilidad no varía. Sin embargo, en comparación con el escenario 2, el índice de interpretabilidad disminuye. Esto se debe a que un único parámetro clasificado como IPI se sustituye por dos parámetros también clasificados como IPI, lo que aumenta el número total de

Tabla 5.11: Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 2.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_I(\theta)$
S_g	PIPI	0,75	0	1	$3,5 \times 10^{-2}$	$6,1 \times 10^{-1}$	$4,5 \times 10^{-2}$
HGP	IPI	0,9	0	1	$5,7 \times 10^{-2}$	$1,0 \times 10^0$	$5,8 \times 10^{-2}$
S_I	IPI	0,9	0	1	$3,2 \times 10^{-2}$	$5,6 \times 10^{-1}$	$5,2 \times 10^{-2}$
f	PIPI	0,75	1	0,5	$2,8 \times 10^{-4}$	$4,9 \times 10^{-3}$	$1,9 \times 10^{-2}$
k_{m1}	PIPI	0,75	0	1	$1,1 \times 10^{-6}$	$1,9 \times 10^{-5}$	$3,8 \times 10^{-2}$
k_{m2}	PIPI	0,75	0	1	$4,9 \times 10^{-5}$	$8,6 \times 10^{-4}$	$3,8 \times 10^{-2}$
V_g	PIPI	0,75	1	0,5	$2,8 \times 10^{-4}$	$4,9 \times 10^{-3}$	$1,9 \times 10^{-2}$
Bw	GPI	1	1	0,5	$3,2 \times 10^{-2}$	$5,6 \times 10^{-1}$	$2,9 \times 10^{-2}$
p_2	PIPI	0,75	0	1	$2,6 \times 10^{-2}$	$4,5 \times 10^{-1}$	$4,3 \times 10^{-2}$
I_b	IPI	0,9	0	1	$9,5 \times 10^{-4}$	$1,7 \times 10^{-2}$	$4,5 \times 10^{-2}$
k_{md}	PIPI	0,75	0	1	$5,5 \times 10^{-7}$	$9,6 \times 10^{-6}$	$3,8 \times 10^{-2}$
k_1	PIPI	0,75	0	1	$3,70 \times 10^{-3}$	$6,5 \times 10^{-2}$	$3,8 \times 10^{-2}$
k_2	PIPI	0,75	0	1	$1,4 \times 10^{-2}$	$2,5 \times 10^{-1}$	4×10^{-2}
k_d	PIPI	0,75	0	1	$1,1 \times 10^{-2}$	2×10^{-1}	4×10^{-2}
n	NPI	0	0	1	$3,3 \times 10^{-2}$	$5,8 \times 10^{-1}$	$7,2 \times 10^{-3}$
V_I	PIPI	0,75	1	0,5	$3,2 \times 10^{-2}$	$5,6 \times 10^{-1}$	$2,2 \times 10^{-2}$

parámetros analizados.

El efecto de la sobreparametrización se ve reflejado en el cálculo del índice. La Figura 5.11 ilustra la jerarquía de parámetros en el modelo mínimo de glucosa, basada en el índice de interpretabilidad (Figura 5.17(a)) y el coeficiente de sensibilidad paramétrica (Figura 5.17(b)).

La sensibilidad a la insulina es un parámetro importante en el control de la concentración de glucosa en sangre y su significado fisiológico es claro, como puede verse en la Figura 5.17. Desafortunadamente no existe la instrumentación necesaria para su medición, por lo que es necesario estimarla. Las tasas de absorción, disociación y transporte de insulina tienen un índice de interpretabilidad mayor en comparación con parámetros como el peso corporal y la producción hepática de glucosa. Esto se debe a que las tasas de insulina aparecen cuando se calculan los balances de masa, lo que proporciona un mayor índice de interpretabilidad, mientras que otros parámetros con un significado fisiológico más claramente definido aparecen en niveles más bajos de especificación. Este resultado no es coherente en términos de interpretación del modelo. Sin embargo, es

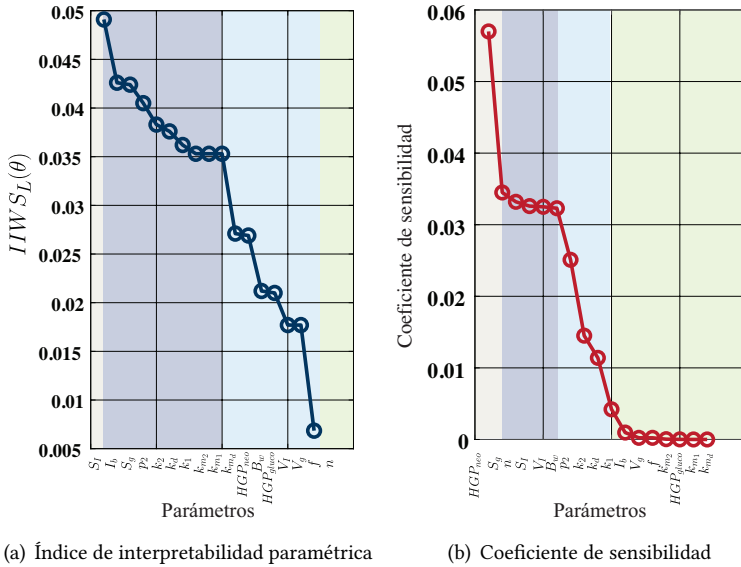


Figura 5.17: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 3.

el resultado de considerar los compartimentos como las particiones del sistema a modelar estando los compartimentos alejados de la realidad, dando lugar a mayores retos cuando, por ejemplo, el usuario del modelo quiere identificar parámetros u obtener una mejor interpretación de los resultados del modelo desde la fisiología humana.

5.3.2.5. Escenario 4: parámetro HGP_{gluco} definido mediante una ecuación constitutiva

El análisis de interpretabilidad para el escenario 4 se muestra en la Tabla 5.13, con un $IIWS_L = 0,81$. En este caso, se analizaron un total de 19 parámetros, los definidos previamente y los tres nuevos parámetros generados al definir HGP_{gluco} . Un parámetro clasificado como IPI (HGP_{gluco})

Tabla 5.12: Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 3.

θ	Tipo	PI_C	S_L	S_{Ln}	S_c	S_{cn}	$IIS_L(\theta)$
V_I	PIPI	0,75	1	0,5	$3,3 \times 10^{-2}$	$5,7 \times 10^{-1}$	$2,1 \times 10^{-2}$
V_g	PIPI	0,75	1	0,5	$2,3 \times 10^{-4}$	4×10^{-3}	$1,8 \times 10^{-2}$
Bw	GPI	1	1	0,5	$3,2 \times 10^{-2}$	$5,7 \times 10^{-1}$	$2,7 \times 10^{-2}$
S_I	IPI	0,9	0	1	$3,3 \times 10^{-2}$	$5,7 \times 10^{-1}$	$4,9 \times 10^{-2}$
S_g	PIPI	0,75	0	1	$3,5 \times 10^{-2}$	6×10^{-1}	$4,2 \times 10^{-2}$
p_2	PIPI	0,75	0	1	$2,5 \times 10^{-2}$	$4,4 \times 10^{-1}$	4×10^{-2}
k_{m1}	PIPI	0,75	0	1	$9,3 \times 10^{-7}$	$1,6 \times 10^{-5}$	$3,5 \times 10^{-2}$
k_{m2}	PIPI	0,75	0	1	4×10^{-5}	7×10^{-4}	$3,5 \times 10^{-2}$
k_{md}	PIPI	0,75	0	1	$4,5 \times 10^{-7}$	$7,8 \times 10^{-6}$	$3,5 \times 10^{-2}$
k_1	PIPI	0,75	0	1	$4,21 \times 10^{-3}$	$7,4 \times 10^{-2}$	$3,6 \times 10^{-2}$
k_2	PIPI	0,75	0	1	$1,5 \times 10^{-2}$	$2,5 \times 10^{-1}$	$3,8 \times 10^{-2}$
k_d	PIPI	0,75	0	1	$1,1 \times 10^{-2}$	2×10^{-1}	$3,8 \times 10^{-2}$
n	NPI	0	0	1	$3,32 \times 10^{-2}$	$5,8 \times 10^{-1}$	$6,8 \times 10^{-3}$
f	PIPI	0,75	1	0,5	$2,3 \times 10^{-4}$	4×10^{-3}	$1,8 \times 10^{-2}$
HGP_{gluco}	IPI	0,9	1	0,5	$1,1 \times 10^{-5}$	2×10^{-4}	$2,1 \times 10^{-2}$
HGP_{neo}	IPI	0,9	1	0,5	$5,7 \times 10^{-2}$	1	$2,7 \times 10^{-2}$
I_b	GPI	0,9	0	1	$9,63 \times 10^{-4}$	$1,7 \times 10^{-2}$	$4,3 \times 10^{-2}$

se sustituyó por dos parámetros también clasificados como IPI (r_2 y N_{GI}) y uno como GPI. El parámetro $\sigma_{G,2}$ tiene interpretabilidad general (GPI) porque procede de una reacción bioquímica, por lo que su significado físico es independiente de los supuestos utilizados para derivar el modelo matemático. Sin embargo, el cálculo de su índice de interpretabilidad es más bajo de lo esperado porque surge en un nivel de especificación más alto (2). Además, la influencia del parámetro $\sigma_{G,2}$ en la salida del modelo es pequeña como se observa en la Figura 5.18, lo que significa que este parámetro no aporta conocimiento útil al modelo, haciendo que el índice de interpretabilidad del modelo sea ligeramente menor que en el escenario anterior. En la jerarquía de parámetros en el modelo mínimo de glucosa presentado en la Figura 5.18 también se puede observar que parámetros como S_g , HGP_{neo} , S_I , y Bw tienen una alta interpretabilidad y sensibilidad, mientras que parámetros como n tiene una fuerte influencia en las respuestas del modelo aunque su interpretabilidad es pobre. Estos tipos de parámetros (como n) suelen ser útiles para ajustar adecuadamente el modelo a los datos experimentales, pero no es deseable que tengan una influencia tan significativa en la salida del modelo, considerando los problemas de identificabilidad que suelen tener.

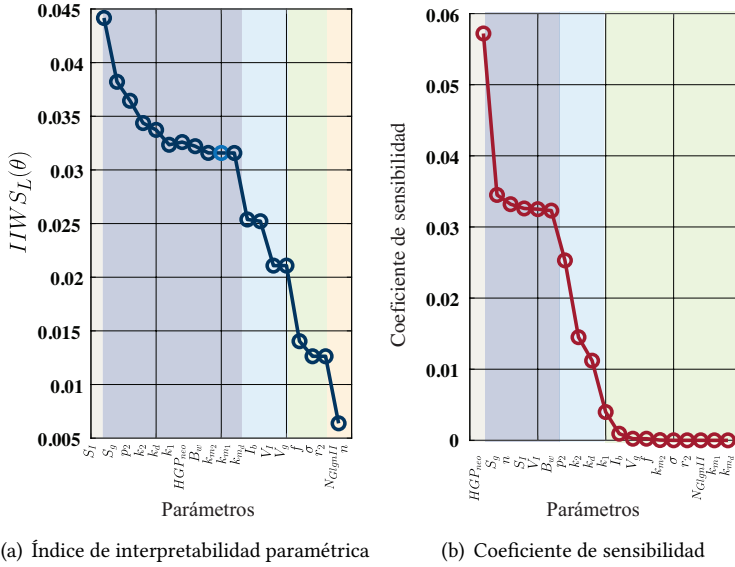


Figura 5.18: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo mínimo de glucosa en el escenario 4.

5.3.2.6. Resultados comparativos para el modelo mínimo de glucosa

Considerando los resultados obtenidos para los 4 escenarios planteados, los gráficos de radar presentados en las Figuras 5.19-5.20 ilustran la progresión del índice de interpretabilidad $IWS_L(\theta)$ dentro de cada escenario de modelado. El ID (1, 2, ..., n) de cada parámetro en la figura, corresponde al ID (1, 2, ..., n) en la Tabla 5.2. Un parámetro al que se asigna un valor cero significa su ausencia en el escenario actual considerado. La figura 5.20 muestra la progresión individual de la influencia de cada parámetro en el índice general de interpretabilidad del modelo mínimo de glucosa en cada escenario considerado. En particular, los parámetros 1-8 y 15-20, que constituyen la estructura básica del modelo

Tabla 5.13: Análisis de interpretabilidad del modelo mínimo de glucosa en el escenario 4.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IWS_I(\theta)$
V_I	PIPI	0,75	1	0,7	$3,3 \times 10^{-2}$	$5,7 \times 10^{-1}$	$2,5 \times 10^{-2}$
V_g	PIPI	0,75	1	0,7	$2,3 \times 10^{-4}$	4×10^{-3}	$2,1 \times 10^{-2}$
Bw	GPI	1	1	0,7	$3,2 \times 10^{-2}$	$5,7 \times 10^{-1}$	$3,2 \times 10^{-2}$
S_I	IPI	0,9	0	1	$3,3 \times 10^{-2}$	$5,7 \times 10^{-1}$	$4,4 \times 10^{-2}$
S_g	PIPI	0,75	0	1	$3,5 \times 10^{-2}$	6×10^{-1}	$3,8 \times 10^{-2}$
$p2$	PIPI	0,75	0	1	$2,5 \times 10^{-2}$	$4,4 \times 10^{-1}$	$3,6 \times 10^{-2}$
k_{m1}	PIPI	0,75	0	1	$9,7 \times 10^{-7}$	$1,7 \times 10^{-5}$	$3,2 \times 10^{-2}$
k_{m2}	PIPI	0,75	0	1	$3,8 \times 10^{-5}$	$6,7 \times 10^{-4}$	$3,2 \times 10^{-2}$
k_{md}	PIPI	0,75	0	1	$4,7 \times 10^{-7}$	$8,2 \times 10^{-6}$	$3,2 \times 10^{-2}$
k_1	PIPI	0,75	0	1	4×10^{-3}	7×10^{-2}	$3,2 \times 10^{-2}$
k_2	PIPI	0,75	0	1	$1,5 \times 10^{-2}$	$2,5 \times 10^{-1}$	$3,4 \times 10^{-2}$
k_d	PIPI	0,75	0	1	$1,1 \times 10^{-2}$	2×10^{-1}	$3,4 \times 10^{-2}$
n	NPI	0	0	1	$3,3 \times 10^{-2}$	$5,8 \times 10^{-1}$	$6,1 \times 10^{-3}$
f	PIPI	0,75	1	0,7	$2,3 \times 10^{-4}$	4×10^{-3}	$2,1 \times 10^{-2}$
HGP_{neo}	IPI	0,9	1	0,7	$5,7 \times 10^2$	1	$3,2 \times 10^{-2}$
$r2$	IPI	0,9	2	0,3	5×10^{-6}	$8,9 \times 10^{-5}$	$1,3 \times 10^{-2}$
$\sigma_{G,2}$	GPI	1	2	0,3	5×10^{-6}	$8,9 \times 10^{-5}$	$1,4 \times 10^{-2}$
$NGlgn_{II}$	IPI	0,9	2	0,3	5×10^{-6}	$8,9 \times 10^{-5}$	$1,3 \times 10^{-2}$

matemático, presentan sistemáticamente el índice de interpretabilidad más elevado. En este caso de aplicación concreto, un análisis de la Figura 5.19 permite identificar el escenario 4 como el menos favorable en términos de interpretabilidad. Dentro de este escenario, el modelo experimenta una notable degradación, evidente no sólo en el índice de interpretabilidad de los parámetros recién introducidos que se aproxima a cero, sino también en la disminución del índice de interpretabilidad de los parámetros preexistentes dentro del modelo matemático. Aunque los parámetros generados en el último escenario son parámetros con interpretabilidad general ($\sigma_{G,2}$) e interpretabilidad heredada ($r2$ y $NGlgn_{II}$) aparecen en un nivel de especificación superior, de manera tal que el índice penaliza la interpretabilidad del propio parámetro y deja en evidencia la sobreparametrización del modelo. Por el contrario, el escenario 2 resulta ser el más interpretable, como se indica en la Figura 5.19, gracias al acoplamiento de las variables de estado del modelo que se genera cuando R_a e IR_a se definen en términos de parámetros conocidos, haciendo que el número de parámetros analizados en este escenario sea el menor.

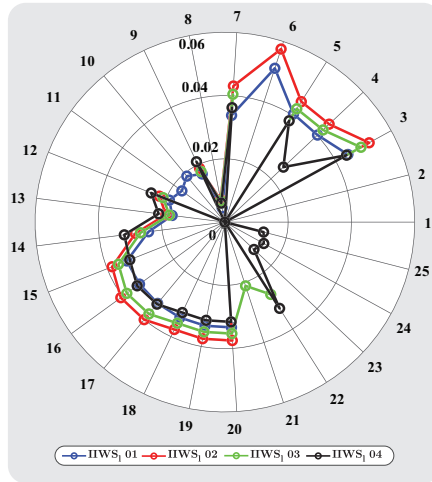
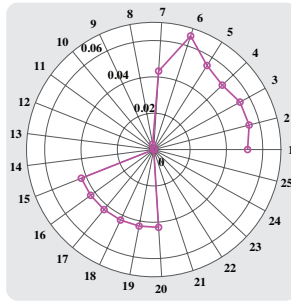
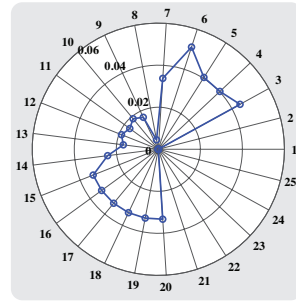


Figura 5.19: Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelos mínimos de glucosa.

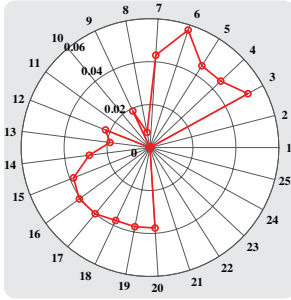
La Tabla 5.14 resume los resultados cuantitativos del análisis de interpretabilidad realizado para cada escenario de modelado. Como se puede observar, en el escenario 1, el índice de interpretabilidad sufre una degradación natural de la interpretabilidad con respecto a la estructura básica del modelo. El modelo mínimo de glucosa representa un caso de estudio interesante, teniendo en cuenta que, debido a las condiciones en las que se ha reportado el modelo, es difícil acceder a su estructura básica original. Por lo tanto, es el único caso de estudio en el que se encuentra un parámetro no interpretable en lo que en este análisis se ha descrito como la estructura básica del modelo. En este sentido, para el caso del escenario de modelado 2, el índice de interpretabilidad mejoró en 8,43 % gracias a la descripción de los parámetros del modelo 2 en términos de los parámetros existentes. En este caso, el índice fue capaz de evaluar no sólo la re-parametrización del modelo (en términos del número de parámetros) sino también la información proporcionada por las nuevas relaciones fenomenológicas, dando a este escenario el índice de interpretabilidad



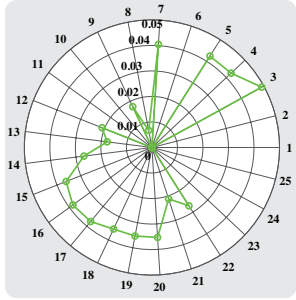
(a) $IIWS_L(\theta)$ en el escenario 0



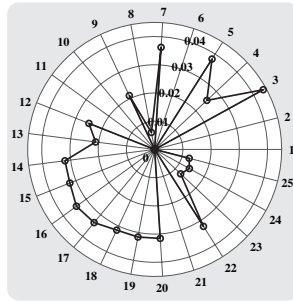
(b) $IIWS_L(\theta)$ en el escenario 1



(c) $IIWS_L(\theta)$ en el escenario 2



(d) $IIWS_L(\theta)$ en el escenario 3



(e) $IIWS_L(\theta)$ en el escenario 4

Figura 5.20: Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado mínima de glucosa.

Tabla 5.14: Índice de interpretabilidad $IIWS_{L,M_s}$ para cada escenario planteado para el modelo mínimo de glucosa. El cambio en el valor del índice representa el grado de mejora (+) o degradación (–) del índice en comparación con el escenario de modelado correspondiente.

	$M_s = 0$	$M_s = 1$	$M_s = 2$	$M_s = 3$	$M_s = 4$
Parámetros interpretables	13	17	15	16	18
Parámetros no interpretables	1	1	1	1	1
$IIWS_{L,M_s}$	1	0,83	0,9	0,83	0,81
Cambio en el valor del índice		–17 %	+8,43 %	0 %	–2,4 %

más alto. El escenario 3 es similar al escenario 1 tanto en el número de parámetros como en su respectivo índice de interpretabilidad. Por último, el escenario 4 resulta ser el menos favorable de los escenarios con una degradación del índice de 10 % respecto al mejor escenario (2).

Una vez analizada la interpretabilidad del modelo matemático, cobra una especial relevancia el análisis de identificabilidad de los escenarios planteados tal como se describió en la Sección 3.1.2. La Tabla 6.13 resume el análisis de identificabilidad realizado para todos los escenarios de modelado. Un hallazgo interesante y que ratifica los análisis realizados en los apartados anteriores, es el hecho de que todas los escenarios de modelado resultan ser estructuralmente no identificables. Cabe resaltar, que la herramienta STRIKE GOLD arrojó un resultado indeterminado para el análisis de todas las estructuras propuestas para este modelo. Como ya se mencionó antes, el modelo mínimo de glucosa presenta problemas estructurales desde la concepción de sus balances, por lo que al analizar la identificabilidad estructural el resultado es coherente.

Tabla 5.15: Identificabilidad de los parámetros desconocidos del modelo matemático de un modelo mínimo de glucosa, para todos los escenarios de modelado.

	Vector de Parámetros desconocidos	Análisis de identificabilidad
$M_s = 0$	$\theta_{u,0} = [n \quad k_n \quad k_{n_2} \quad S_g \quad S_I]$	Estructuralmente no identificables
$M_s = 1$	$\theta_{u,1} = [S_g \quad S_I \quad n \quad R_a \quad IR_a]$	Estructuralmente no identificables
$M_s = 2$	$\theta_{u,2} = [n \quad S_g \quad S_I]$	Estructuralmente no identificables
$M_s = 3$	$\theta_{u,3} = [S_g \quad S_I \quad n \quad HGP_{gluco} \quad HGP_{neo}]$	Estructuralmente no identificable
$M_s = 4$	$\theta_{u,3} = [S_g \quad S_I \quad n \quad r_2 \quad \sigma_{G,2} \quad NGI]$	Estructuralmente no identificable

Conclusiones

En este capítulo, se propuso una metodología de modelado ampliada que incorpora la interpretabilidad como criterio de parada en la búsqueda de ecuaciones constitutivas o de valoración. Lo anterior, buscando mantener el equilibrio entre la precisión y la interpretabilidad de los modelos, para facilitar la toma de decisiones en campos críticos en los que la interpretabilidad es importante. Se evaluó la utilidad del índice en la metodología propuesta a través de dos aplicaciones de diferente grado de complejidad, demostrando su versatilidad y aplicabilidad en diversos contextos, desde sistemas biológicos hasta procesos industriales. La implementación del índice propuesto reveló su utilidad como criterio de parada en la especificación de modelos, destacando su potencial para mejorar la capacidad de interacción y eficacia en la resolución de problemas complejos del mundo real en diversos ámbitos científicos y de ingeniería.

CASO DE ESTUDIO: *LOBESIA BOTRANA*

6.1	<i>Lobesia botrana</i>	116
6.1.1	Generalidades	116
6.1.2	<i>Lobesia botrana</i> en la literatura	117
6.2	Modelo de crecimiento de <i>L. botrana</i>	121
6.2.1	Etapas 1. Problema de modelado	121
6.2.2	Etapas 2. Deducción de la estructura básica	126
6.2.3	Etapas 3. Deducción de la estructura extendida	128
6.3	Voltinismo de <i>Lobesia botrana</i> : una extensión del modelo de crecimiento	158
6.3.1	Concepto de voltinismo en <i>Lobesia botrana</i>	158
6.3.2	Área de estudio	161
6.3.3	Datos fenológicos de <i>Lobesia botrana</i>	162
6.3.4	Datos meteorológicos	165
6.3.5	Extensión del modelo de crecimiento	166
6.3.6	Propiedades del modelo	171
6.3.7	Ajuste y validación	172

La ciencia más útil es aquella cuyo fruto es el más comunicable.

Leonardo Da Vinci

La industria vitivinícola es un sector económicamente importante para Argentina, que genera ingresos de alrededor de 800 millones de dólares al año, ocupando el primer lugar en producción de vino en la región y el noveno a nivel mundial (Centro de Estudios Económicos de Bodegas, 2020; CREA, 2020). Con más de 900 bodegas activas en todo el país y 17,600 productores, la producción de vino se concentra principalmente

en la región de Cuyo. Esta región está compuesta por las provincias de Mendoza, San Juan y San Luis, las cuales representan el 70,6 %, 23,73 % y 1 % de la producción total de vino del país, respectivamente, según informes del Observatorio Argentino del Vino (Observatorio Vitivinícola, 2021). Se estima que en la provincia de Mendoza (la provincia con la mayor participación en la producción de vino), casi una de cada tres hectáreas de viñedos cultivados está infestada con *Lobesia botrana* de la familia *Tortricidae: Lepidoptera*, también conocida como "la polilla de la vid" (Denis, Schiffermüller et al., 1775). Esta plaga es de importancia económica para varios países de Europa y América del Sur, y su amplia distribución se explica en gran medida por la capacidad de la misma para adaptarse a los cambios climáticos, lo que resulta en una falta de sincronización con sus depredadores naturales, causando aumentos significativos de la polilla en períodos de tiempo muy cortos (Altimira et al., 2021). Lo anterior enfatiza la necesidad del desarrollo de herramientas tendientes al control y erradicación de la plaga. En este capítulo, se aborda el modelado matemático del proceso de crecimiento y voltinismo de *Lobesia botrana* a través de la metodología propuesta en el Capítulo 5.

6.1 *Lobesia botrana*

6.1.1 Generalidades

Lobesia botrana o "polilla de la vid", se señaló por primera vez como plaga de la vid en Austria en el año 1800, veinticuatro años después de su descripción en 1776 por Denis y Schiffermüller. A partir de este momento, se extendió progresivamente por Europa durante los siglos XIX y XX (Torres Vila, 1995). En el otoño del año 2008, se registró su aparición en la zona central de Chile, y posteriormente en el año 2010 en Maipú, Mendoza (Burgardta et al., 2010; Gonzalez, 2010). La plaga de *Lobesia botrana* no ha sido eliminada en casi ninguno de los países del mundo donde ha aparecido (a excepción de los Estados Unidos), y es considerada una plaga cuarentenaria (plaga de importancia económica potencial para el área) de la Argentina. La polilla produce daños a la vid en sus diferentes estados fenológicos (Torres Vila, 1995), que se manifiestan en la destrucción de las bayas (granos) y florescencias disminuyendo la

cantidad y calidad de la cosecha y ocasionando mermas en la producción de vinos. Además, produciendo un daño indirecto tanto en uvas de mesa como de vinificación, debido a la aparición de podredumbres causadas por diversos hongos asociados a la *Lobesia botrana* que provocan mal olor y sabor al vino (Burgardta et al., 2010; da Silva et al., 2013). Por lo anterior, modelar el proceso de crecimiento de *Lobesia botrana* es considerado de interés para la región, modelo que además resulta relevante para la implementación de la metodología propuesta en los capítulos anteriores, y para la discusión sobre interpretabilidad paramétrica.

6.1.2 *Lobesia botrana* en la literatura

Con el fin de predecir el ciclo biológico de *Lobesia botrana*, diferentes autores han propuesto modelos matemáticos basados en información biológica y climática. Los estudios revelan que, la velocidad de desarrollo de la polilla se encuentra mayormente influenciada por las variables ambientales de temperatura, humedad relativa y fotoperiodo (Torres Vila, 1995). En este sentido, la temperatura es la variable meteorológica más significativa, ya que no sólo afecta su velocidad de desarrollo en los estadios iniciales sino también a la longevidad de los adultos, además de regular el desarrollo ovocitario y la producción de huevos en la hembra, alcanzando un máximo alrededor de los 22°C (Marco, 2001; Torres-Vila et al., 1996).

6.1.2.1. Modelo de Touzeau

Touzeau, 1981 propuso el primer modelo para predecir el ciclo biológico de la polilla de la vid en la región francesa de Midi Pyrénées. Este modelo se basa en el concepto de grados día (DD por sus siglas en inglés) y en la integral de la temperatura. El primero permite establecer un umbral mínimo de desarrollo de la polilla de 10°C , mientras que la integral térmica se define como el número de DD que deben acumularse (*ADD*) para que se produzca un evento determinado, como la eclosión, la muda, la pupación o la emergencia de adultos (Marco, 2001). El modelo de Touzeau fue ampliamente utilizado, identificando dos limitaciones principales: i) su dependencia de las condiciones ambientales locales y ii) la ausencia de algunas características importantes del proceso de desarrollo de la polilla, como las tasas de crecimiento, la mortalidad asociada a cada etapa y otras

variables ambientales (Marco, 2001).

Diferentes investigadores tuvieron en cuenta este modelo y lo adaptaron a condiciones específicas. Armendáriz et al., 2007 emplearon una simplificación del modelo de Touzeau comparando datos sobre la incidencia de la polilla *Lobesia botrana* en viñedos con denominación de origen Arribes de Castilla y León, España. En su trabajo, cuantificaron la población máxima de las tres generaciones de la polilla y compararon los grados-día acumulados para los picos de vuelo, con diferentes modelos reportados en la literatura. En la mayoría de los casos, hubo diferencias significativas en *ADD* entre los reportados por los modelos citados y los que se encontraron en campo, todas ellas explicadas por las condiciones climáticas de cada región. Estos hallazgos plantean la necesidad de desarrollar modelos matemáticos que sean fácilmente extrapolables a regiones del mundo que difieran de las condiciones particulares para las cuales fueron validados. Gutierrez et al., 2012 presentaron un modelo demográfico de base fisiológica (PBDM) para la distribución geográfica y cantidad relativa de la polilla en California. En su trabajo, los autores también reportaron un modelo de la vid para la evaluación de su interacción con la polilla en los diferentes estadios de crecimiento. Otro aporte interesante fue el presentado por Dagatti y Becerra, 2015 con un modelo ajustado para predecir el comportamiento de la polilla en un viñedo de la región de Mendoza en Argentina, para las temporadas 2012 a 2014.

6.1.2.2. Modelos de voltinismo

Otros estudios más recientes se centraron en el desarrollo de modelos matemáticos que no solo caracterizaran el proceso de desarrollo de la polilla, sino también su voltinismo. En esta línea, Burgardt et al., 2011 introdujeron un modelo de identificación para caracterizar el voltinismo y la dispersión local de la polilla en una finca. Los datos obtenidos fueron comparados con el modelo original de Touzeau (para Europa) para diferentes eventos como eclosión, muda, pupación o emergencia de adultos, durante el ciclo biológico de la polilla, y estableciendo las diferencias entre los grados-día necesarios para los cambios de estadio y los grados-día predichos por el modelo de Touzeau. Heit et al., 2019 introdujeron un modelo matemático para describir el voltinismo de *Lobesia botrana* en

Mendoza, Argentina. Castex et al., 2020 propusieron un modelo fenológico para describir el voltinismo de la polilla, con el fin de predecir sus fases críticas de desarrollo y, posteriormente, utilizar esta información para determinar las acciones de control y erradicación adecuadas.

6.1.2.3. Dinámica poblacional

Otro tipo de modelos desarrollados, son los de dinámica poblacional que no solo consideran los cambios de la polilla de manera temporal, sino también espacial. En este sentido, Iltis et al., 2020 propusieron una alternativa factible a los PBDMs desde el punto de vista ecológico pero basada en regímenes térmicos constantes y poco realistas. Los autores muestran una comparación para el rendimiento del estadio adulto de *Lobesia botrana* entre el régimen fluctuante actual (CFR), y el régimen fluctuante futuro (FFR). Los autores concluyen que sus hallazgos podrían contribuir a mejorar la fiabilidad de las predicciones de PBDM sobre la dinámica poblacional de *Lobesia botrana*. Amo-Salas et al., 2011, propusieron conocer con precisión los tiempos máximos de vuelo de la polilla para optimizar el uso de plaguicidas, así como generar conocimiento sobre el proceso de desarrollo de la polilla que permitiera determinar los estadios más susceptibles de la misma. Para ello, los autores propusieron el uso de un modelo empírico de transformación logística. En esta misma línea, Ortega-Lopez et al., 2014, presentaron un modelo predictivo basado también en una función logística, que consideraba la adición de humedad como un factor importante en el crecimiento de la polilla, para comparar regiones donde el voltinismo de *Lobesia botrana* presentaba similitudes. Las principales conclusiones de los autores fueron que: (i) la ecuación de Touzeau no es suficiente para predecir el voltinismo de la polilla en algunas regiones del mundo, y que (ii) el número de generaciones de la polilla en una región depende de las condiciones climáticas de la misma (Amo-Salas et al., 2011; Ortega-Lopez et al., 2014).

6.1.2.4. Modelos para el desarrollo de DSS

Con el fin de desarrollar sistemas de soporte a la toma de decisiones (DSS por sus siglas en inglés *Decision Support System*) para el control óptimo de la plaga, autores como Gilioli et al., 2016, propusieron un modelo esto-

cástico para describir la fenología de diferentes plagas, tomando *Lobesia botrana* como caso de estudio. En su trabajo proponen la estimación de las tasas de mortalidad y crecimiento utilizando datos de la literatura y datos de campo de una región de Italia. Los autores destacan que, a pesar de los importantes avances en la eficacia de los plaguicidas, el problema de determinar los tiempos óptimos de aplicación sigue sin resolverse. Por ello, sugieren que la clave para abordar esta laguna reside en el desarrollo de DSS basados en la dinámica de la plaga. En la misma línea, Carlos et al., 2018, propusieron el uso de modelos no lineales basados en ecuaciones de regresión de Boltzmann, para caracterizar la fenología de vuelo de *Lobesia botrana*, basándose en datos de 20 años de trampeo de machos en campo. Los autores expresaron la necesidad de desarrollar herramientas basadas en modelo, para generar conocimientos sobre el proceso de crecimiento de la polilla, con el fin de mejorar las estrategias de gestión integrada de plagas (GIP).

Otro planteamiento interesante fue el de Rossini et al., 2020a, que propusieron una herramienta computacional basada en un modelo de dinámica de poblaciones, desarrollado para permitir la interacción entre el usuario final y la interfaz de simulación. Los autores siguieron un enfoque de dos etapas: (i) en la primera generaron tablas de vida de la polilla en condiciones de laboratorio. (ii) Después, este modelo se simuló y validó con datos de campo. En otro trabajo, los autores contrastaron el comportamiento de dos modelos matemáticos diferentes para la descripción del ciclo de vida de *Lobesia botrana*, uno basado en la ecuación de Von Foerster (VFE) y otro basado en el modelo de retardo distribuido (DDM). En este trabajo se discuten las diferencias relacionadas con los distintos conceptos de edad (cronológica o fisiológica) entre estos dos modelos. En concreto, los autores señalan que a pesar de que los modelos DDM son más adecuados para describir el ciclo vital de estos insectos, tienen un parámetro h que carece de significado biológico, lo que dificulta el ajuste del modelo y aumenta la necesidad de datos experimentales para estimarlo (Rossini et al., 2020b). En otro trabajo, los autores destacan el potencial de desarrollar modelos matemáticos basados en primeros principios para describir los fenómenos de crecimiento de algunos insectos Rossini et al., 2021.

6.2 Modelo de crecimiento de *L. botrana*

Este trabajo de tesis doctoral se deduce un modelo semi-físico basado en primeros principios de acuerdo con la metodología propuesta en el capítulo 5, para describir cuantitativamente las diferentes etapas de crecimiento de *Lobesia botrana* en su tiempo de desarrollo normal para condiciones controladas de laboratorio y las condiciones ambientales de la región de Cuyo en Argentina. El modelo matemático, propuesto difiere principalmente de los reportados en su simplicidad conceptual, la interpretabilidad de sus parámetros y la facilidad para su cálculo y ajuste.

6.2.1 Etapa 1. Problema de modelado

6.2.1.1. Descripción del proceso y objetivo del modelo

Lobesia botrana es un insecto con metamorfosis completa, lo que significa que su ciclo de vida consta de cuatro etapas: huevo, larva, pupa y adulto (Urretabizkaya et al., 2010). En la Figura 6.1 se muestra el proceso de

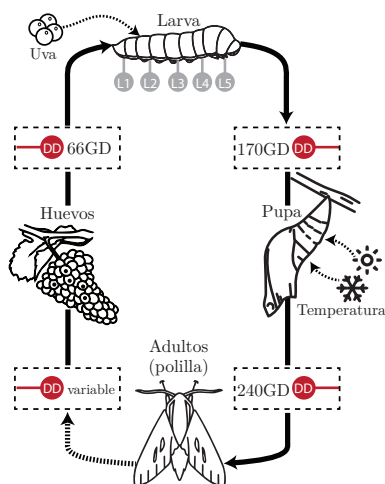


Figura 6.1: Ilustración del proceso de crecimiento de *Lobesia botrana* para una generación dada. Fuente propia (Aguirre-Zapata et al., 2022b)

crecimiento de la polilla considerando los cuatro estadios principales de crecimiento (huevo, larva, pupa, adulto) y la influencia de las variables ambientales en el mismo. La figura ilustra el proceso de crecimiento en una generación dada, más no considera el voltinismo de la polilla (ver la flecha punteada).

6.2.1.1.1. Estadio huevo: en su primer estadio, la hembra deposita sus huevos en flores, frutos o raquis de la vid. El embrión (dentro del huevo) obtiene su alimento del medio en el que se encuentra y atraviesa un proceso de desarrollo de cinco etapas hasta alcanzar el desarrollo total en este estadio (cabeza negra). El $37 \pm 4,3\%$ de los huevos depositados por la hembra mueren (Herrera et al., 2016, 2017).

6.2.1.1.2. Estadio larval: las larvas emergen después de tres a siete días de desarrollo del huevo y una vez alcanzados aproximadamente 66ADD. Perforan el fruto de la vid y se alimentan de él. Las larvas pueden consumir su peso varias veces y obtienen su alimento del medio ambiente en el que se encuentran. El estadio larval tiene cinco etapas que están limitadas por el tamaño de su cutícula, lo que resulta en cambios denominados muda. El crecimiento larval tiene lugar en temperaturas que van desde $10^{\circ}C \leq T \leq 30^{\circ}C$. Los valores fuera de este rango son desfavorables también para las otras etapas. La mortalidad de las larvas con una sola muda es de $26,34 \pm 4,6\%$, pero este valor disminuye a $1,2 \pm 0,5\%$ cuando se completa la segunda muda y se mantiene hasta pasar al siguiente estadio (Herrera et al., 2016, 2017).

6.2.1.1.3. Estadio de pupa: después de 13 a 15 días y habiendo alcanzado aproximadamente 170ADD, se completa el estadio larval y comienza el estadio de pupa o crisálida. Las pupas permanecen inmóviles en estado de diapausa y no se alimentan. El estado de diapausa se mantiene hasta que se cumplen las condiciones de humedad y temperatura adecuadas que favorecen el crecimiento de la polilla y la transición al siguiente estadio. La mortalidad es de $1,2 \pm 1,1\%$ en este estadio (Herrera et al., 2016, 2017).

6.2.1.1.4. Estadio adulto: después de 5 a 20 días y habiendo alcanzado aproximadamente 204ADD, emerge la polilla (Herrera et al., 2016). El adulto puede requerir 2 a 3 horas para realizar su primer vuelo y un día para reproducirse y así comenzar un nuevo ciclo de vida. El adulto se alimenta del néctar de las flores o del jugo de los frutos maduros y vive aproximadamente 8 días. Este ciclo se repite tres o más veces al año (voltinismo) dependiendo de las condiciones climáticas y los valores de temperatura (en ADD) que regulan las transiciones entre etapas. Teniendo en cuenta lo anterior, el modelo matemático desarrollado en este capítulo, busca describir la dinámica de crecimiento de *Lobesia botrana* para los diferentes estadios (huevo, larva, pupa y adulto) en los tiempos normales de desarrollo del insecto y bajo condiciones controladas de laboratorio.

6.2.1.2. Hipótesis de modelado y nivel de detalle

El modelo adopta una perspectiva macroscópica de cada etapa de crecimiento como parte de un conjunto más amplio, sin adentrarse en los procesos fisiológicos internos del insecto. La hipótesis de modelado parte de establecer una analogía entre el desarrollo del insecto y el crecimiento de microorganismos, incorporando tasas de crecimiento del insecto, tasas de mortalidad medidas, interacciones con parámetros ambientales y condiciones controladas de laboratorio. Los datos experimentales disponibles para la evaluación de la cinética de crecimiento provienen de un experimento de cría en laboratorio de *Lobesia botrana* en Mendoza, Argentina, bajo condiciones controladas de temperatura, humedad y fotoperíodo (Herrera et al., 2016). Las condiciones experimentales incluyeron una temperatura de $20 \pm 2^{\circ}\text{C}$, una humedad relativa de 30 – 50 %, y un fotoperíodo de 16 : 8 horas (L:O). Como medio de crecimiento de la polilla se utilizó una dieta artificial que contenía los nutrientes necesarios (Herrera et al., 2016).

En un recipiente previamente esterilizado destinado a la cría, se colocó la dieta artificial y en su interior los huevos que alcanzaron la edad adulta. Se realizaron tres réplicas de 30 huevos de 24 h de edad. Las réplicas se denominarán R_1 , R_2 y R_3 en este estudio. La Figura 6.2 muestra las curvas de crecimiento obtenidas de la cría en laboratorio. Teniendo en

cuenta que algunas observaciones experimentales estaban fuera de la tendencia de crecimiento de la polilla, fue necesario preprocesar los datos experimentales para identificar los parámetros desconocidos (Herrera et al., 2016).

1. La temperatura se controla a $15^{\circ}C$, $20^{\circ}C$ y $25^{\circ}C$.
2. La humedad se mantiene en el rango entre 30 % y 50 %.
3. La dieta artificial es similar a los nutrientes naturales.
4. El fotoperiodo se mantiene fijo en 16 : 8 (L:D).
5. Las tasas de mortalidad representan la muerte natural (se descuidan los depredadores).
6. El proceso de crecimiento es de tipo discontinuo (proceso por lotes), ya que comienza con un número conocido de huevos. Luego, no se extraen ni se agrega ningún componente durante el proceso de crecimiento.

6.2.1.3. Definición de sistemas de proceso

Se definieron los sistemas de proceso para la posterior aplicación del principio de conservación y la obtención de las ecuaciones de la estructura básica del modelo de crecimiento de *Lobesia botrana*. Para ello, se considero cada etapa de crecimiento de la polilla como un PS (sistema de proceso por sus siglas en inglés), como se muestra en la Figura 6.3. En la figura, las líneas rojas representan las pérdidas de huevos, larvas, pupas y adultos debidas a la muerte natural, mientras que las líneas verdes representan el flujo de huevos, larvas y pupas que pasan a la siguiente fase de crecimiento. La línea amarilla representa el número de adultos vivos.

En cuanto a la Figura 6.3, el primer sistema de proceso (PSI) representa el estadio de los huevos, que da como resultado el crecimiento del embrión dentro del huevo basado en el número inicial de huevos. Proporciona el número de huevos vivos que se convertirán en larvas (E_l) y el número de huevos muertos (E_d). El PSII representa el estadio de larvas. El crecimiento larval (polillas) se basa en el número de huevos vivos (E_l) como corriente de entrada de el estadio anterior. Proporciona el número de larvas vivas (L_l) y muertas (L_d) al PSII. De esta manera, el PSIII representa el estadio de pupas, que indica el desarrollo de la pupa basado en el número de larvas vivas (L_l), mientras entrega el número de pupas vivas (P_l) y muertas

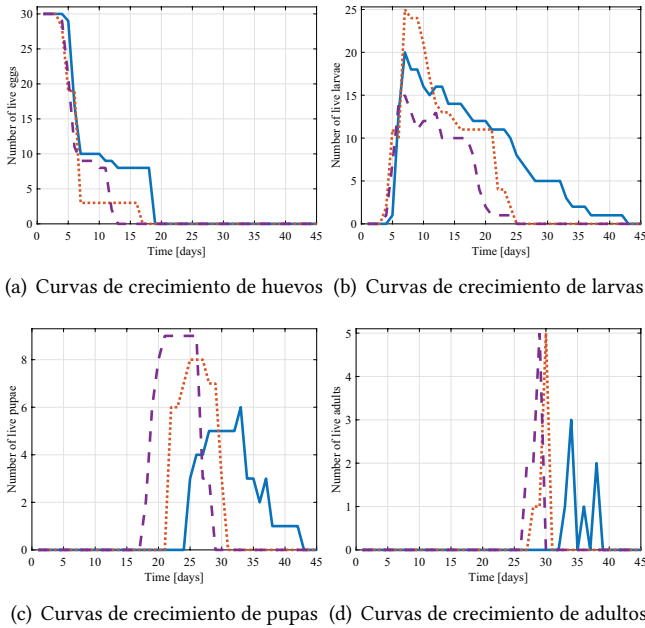


Figura 6.2: Curvas de crecimiento de *Lobesia botrana* en laboratorio bajo condiciones controladas de laboratorio. Cada curva corresponde a una corrida experimental diferente conservando las condiciones de temperatura, humedad y foto periodo.

(P_d). Finalmente, el cuarto PSIV representa el estadio de adultos, donde el adulto emerge de dentro de la pupa basado en el número de pupas vivas (P_l) y entrega el número de adultos vivos (A_l) y muertos (A_d) como corrientes de salida.

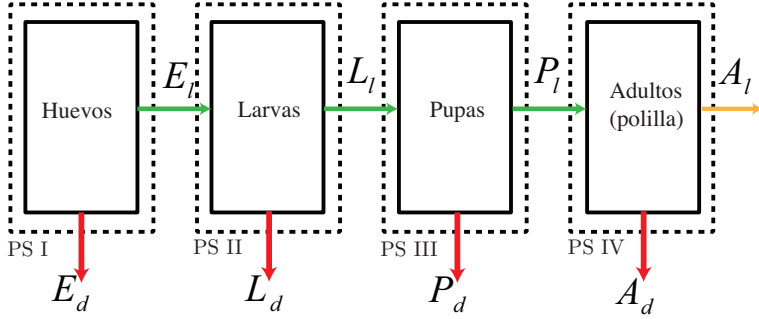


Figura 6.3: Diagrama de bloques de los sistemas de proceso considerados para modelar el proceso de crecimiento de *Lobesia botrana*.

6.2.2 Etapa 2. Deducción de la estructura básica

6.2.2.1. Aplicación del principio de conservación

Se aplicó un balance de masa total sobre el número de huevos, larvas, pupas o adultos (según corresponda) para cada sistema de proceso, de la siguiente manera:

6.2.2.1.1. *PSI*- Estadio huevos

$$\frac{dE_l(t)}{dt} = -K_E E_l(t) - K_{E_d} E_l(t) \quad (6.1)$$

dónde K_E es la tasa de crecimiento de huevos, $E_l(t)$ es el número de huevos vivos, y E_d es el número de huevos muertos.

6.2.2.1.2. *PSII*- Estadio larval

$$\frac{dL_l(t)}{dt} = K_L E_l(t) - K_{L_d} L_l(t) \quad (6.2)$$

dónde K_L es la tasa de crecimiento de larvas, $L_l(t)$ es el número de larvas vivas, y L_d es el número de larvas muertas.

6.2.2.1.3. *PSIII*- Estadio pupa

$$\frac{dP_l(t)}{dt} = K_P L_l(t) - K_{P_d} P_l(t) \quad (6.3)$$

dónde K_P es la tasa de crecimiento de pupas, $P_l(t)$ es el número de pupas vivas, y P_d es el número de pupas muertas.

6.2.2.1.4. *PSIV*- Estadio adulto

$$\frac{dA_l(t)}{dt} = K_A P_l(t) - K_{A_d} A_l(t) \quad (6.4)$$

dónde K_A es la tasa de crecimiento de adultos, $A_l(t)$ es el número de adultos vivos y A_d es el número de adultos muertos.

6.2.2.2. Estructura básica del modelo

Todas las ecuaciones obtenidas desde la Ecuación 6.1 hasta la Ecuación 6.4 son relevantes para el modelo debido a la importante información que proporcionan. Por lo tanto constituyen la estructura fundamental del MSBPP. En este sentido, la estructura básica del modelo de crecimiento queda compuesta por cuatro ecuaciones diferenciales ordinarias que representan los estadios de crecimiento: huevo, larva, pupa y adulto, respectivamente. Posee ocho parámetros estructurales que corresponden a las tasas de crecimiento K_j y las tasas de mortalidad K_{j_d} , para cada estadio de crecimiento (j).

En la estructura básica del modelo, $-K_E(t)E_l(t)$ representa la tasa a la cual los huevos vivos pasan al siguiente estadio (larvas). Este término es negativo porque la eclosión reduce el número de huevos vivos. Es decir, a medida que los huevos eclosionan, su número disminuye. En los demás estadios este término es positivo, dado que representa la tasa a la cual, por ejemplo, los huevos eclosionan en larvas, aumentando el número de larvas vivas.

6.2.2.3. Variables, parámetros estructurales y constantes:

De acuerdo con Lema-Perez et al., 2019a, los términos desconocidos y las variables deben diferenciarse porque no todas las incógnitas se constituyen como variables. Asimismo, define los parámetros estructurales como incógnitas de la estructura básica del modelo. El resumen de variables y parámetros estructurales se presentan en la Tabla 6.1.

Tabla 6.1: Variables y constantes del modelo de crecimiento de *Lobesia botrana* a partir de la estructura básica.

	Variables	Parámetros estructurales
PSI	$E_l(t)$	$K_E(t), E_d(t)$
PSII	$L_l(t)$	$K_L(t), L_d(t)$
PSIII	$P_l(t)$	$K_P(t), P_d(t)$
PSIV	$A_l(t)$	$K_A(t), A_d(t)$
Total	4	8

6.2.3 Etapa 3. Deducción de la estructura extendida

Para la deducción de la estructura extendida del modelo de crecimiento de *Lobesia botrana*, se proponen cuatro escenarios diferentes que buscan establecer una ecuación constitutiva o de valoración para la tasa de crecimiento de la polilla en todos sus estadios de crecimiento, esto es, para los parámetros estructurales K_E , K_L , K_P , y K_A . Todos los parámetros y constantes del modelo se describen en la Tabla 5.1.

La Figura 6.4 ilustra el orden jerárquico de los diferentes escenarios de modelado del modelo de crecimiento de *Lobesia botrana* que se consideraron para el análisis de la interpretabilidad del modelo matemático.

6.2.3.1. Escenarios de modelado

6.2.3.1.1. Escenario 0. Estructura básica del modelo: el escenario $M_s = 0$ corresponde a la estructura básica del modelo matemático presentada en las Ecuaciones 6.1-6.4.

Tabla 6.2: Parámetros del modelo matemático del proceso de crecimiento de *Lobesia botrana* utilizados en todos los escenarios considerados.

<i>ID</i>	θ	Descripción	Ecuación constitutiva	Unidades
1	K_E	Tasa de crecimiento-huevos	Ecuación	Adimensional
2	K_L	Tasa de crecimiento-larvas	Ecuación	Adimensional
3	K_P	Tasa de crecimiento-pupas	Ecuación	Adimensional
4	K_A	Tasa de crecimiento-adultos	Ecuación	Adimensional
5	K_{E_d}	Tasa de mortalidad-huevos	Variable	$\frac{\text{huevos}}{\text{huevos} \cdot \text{día}}$
6	K_{L_d}	Tasa de mortalidad-larvas	Variable	$\frac{\text{larvas}}{\text{larvas} \cdot \text{día}}$
7	K_{P_d}	Tasa de mortalidad-pupas	Variable	$\frac{\text{pupas}}{\text{pupas} \cdot \text{día}}$
8	K_{A_d}	Tasa de mortalidad-adultos	Variable	$\frac{\text{adultos}}{\text{adultos} \cdot \text{día}}$
9	$\mu_{E_{max}}$	Tasa máximo de crecimiento de huevos	Ecuación	$\frac{\text{huevos}}{\text{huevos} \cdot \text{día}}$
10	$\mu_{L_{max}}$	Tasa máximo de crecimiento de larvas	Ecuación	$\frac{\text{larvas}}{\text{larvas} \cdot \text{día}}$
11	$\mu_{P_{max}}$	Tasa máximo de crecimiento de pupas	Ecuación	$\frac{\text{pupas}}{\text{pupas} \cdot \text{día}}$
12	$\mu_{A_{max}}$	Tasa máximo de crecimiento de adultos	Ecuación	$\frac{\text{adultos}}{\text{adultos} \cdot \text{día}}$
13	K_{E_T}	Constante de velocidad media de crecimiento	Variable	Adimensional
14	K_{L_T}	Constante de velocidad media de crecimiento	Variable	Adimensional
15	K_{P_T}	Constante de velocidad media de crecimiento	Variable	Adimensional
16	K_{A_T}	Constante de velocidad media de crecimiento	Variable	Adimensional
17	K_{E_H}	Constante de velocidad media de crecimiento	Variable	Adimensional
18	K_{L_H}	Constante de velocidad media de crecimiento	Variable	Adimensional
19	K_{P_H}	Constante de velocidad media de crecimiento	Variable	Adimensional
20	K_{A_H}	Constante de velocidad media de crecimiento	Variable	Adimensional

ID	θ	Descripción	Ecuación constitutiva	Unidades
21	T_{Eth}	Temperatura umbral huevos	11,6	$^{\circ}C$
22	T_{Lth}	Temperatura umbral larvas	8,8	$^{\circ}C$
23	T_{Pth}	Temperatura umbral pupas	10,35	$^{\circ}C$
24	T_{Ath}	Temperatura umbral adultos	10	$^{\circ}C$
25	σ_E	Parámetro empírico	10	Adimensional
26	σ_L	Parámetro empírico	10	Adimensional
27	σ_P	Parámetro empírico	10	Adimensional
28	σ_A	Parámetro empírico	10	Adimensional
29	α	Parámetro empírico	0,99	Adimensional
30	γ	Parámetro empírico	0,99	Adimensional

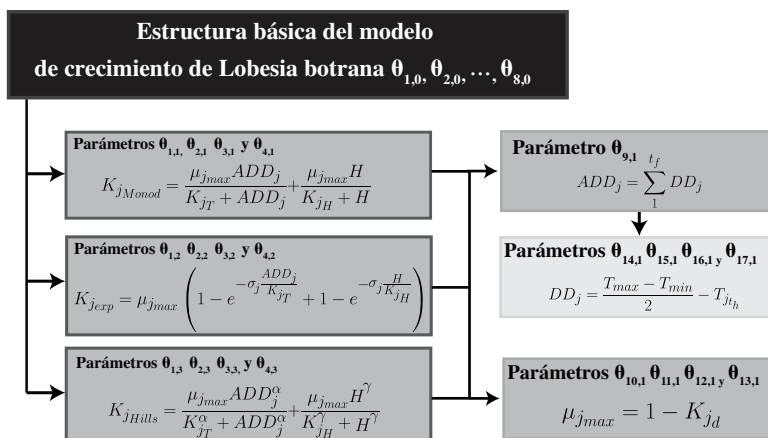


Figura 6.4: Grafo de modelado del proceso de crecimiento de *Lobesia botrana*.

6.2.3.1.2. Escenario 1. Parámetro K_j descrito mediante ecuación de Monod: en el escenario $M_s = 1$, se considera modelar las tasas de crecimiento de la polilla para cada estadio de crecimiento utilizando una ecuación de Monod como se muestra en la Ecuación 6.5. La ecuación de Monod se trata de un modelo matemático que relaciona las tasas de crecimiento microbiano en un medio acuoso con la concentración de

un nutriente limitante (Dochain, 1986; Heijnen & Romein, 1995; Monod, 1949). Para el caso particular de *Lobesia botrana* se consideran los grados-día acumulados (ADD) y la humedad del aire (H) como factores limitantes de crecimiento.

$$K_{j_{Monod}}(t) = \frac{\mu_{j_{max}}(t)ADD_j(t)}{K_{j_T} + ADD_j(t)} + \frac{\mu_{j_{max}}(t)H(t)}{K_{j_H} + H(t)} \quad (6.5)$$

K_{j_T} y K_{j_H} son las constantes de velocidad media de crecimiento para cada estadio j dependientes de la temperatura y la humedad, respectivamente. $\mu_{j_{max}}$ es la tasa de crecimiento máxima para cada estadio de crecimiento j . En el escenario 1 se describieron también los parámetros ADD , DD y $\mu_{j_{max}}$ a través de las Ecuaciones 6.6-6.8. Estas últimas se conservaron para todos los demás escenarios de modelado que contienen dichos parámetros.

$$ADD_j(t) = \sum_1^{t_f} DD_j \quad (6.6)$$

$$DD_j(t) = \frac{T_{max}(t) - T_{min}(t)}{2} - T_{j_{ti}} \quad (6.7)$$

$$\mu_{j_{max}}(t) = 1 - K_{j_d}(t) \quad (6.8)$$

6.2.3.1.3. Escenario 2. Parámetro K_j descrito mediante ecuación exponencial: para el escenario de modelado $M_s = 2$ se propone el uso de la cinética de crecimiento exponencial como se muestra en la Ecuación 6.9. Esta cinética es útil para describir la tasa de crecimiento de una población biológica en función de un sustrato limitante. Se utiliza con menos frecuencia que el modelo Monod porque el término exponencial conduce a complejidades matemáticas (Bader, 1978). Sin embargo, al poseer el parámetro σ_j que es un parámetro empírico sin significado físico, algunos modeladores la usan para ganar un grado de libertad en el ajuste del modelo matemático.

$$K_{j_{exp}}(t) = \mu_{j_{max}}(t) \left(1 - e^{-\sigma_j \frac{ADD_j(t)}{K_{j_T}}} + 1 - e^{-\sigma_j \frac{H(t)}{K_{j_H}}} \right) \quad (6.9)$$

Se asume que ambos factores limitantes del crecimiento tienen el mismo valor de σ_j . Aunque estos valores podrían ser diferentes, dado que no tienen un significado físico concreto, se puede suponer que son los mismos para ambos.

6.2.3.1.4. Escenario 3. Parámetro K_j descrito mediante ecuación de Hill: para el escenario de modelado $M_s = 3$ se propone el uso la cinética de Hill, que es una ecuación de cinética enzimática utilizada para describir el grado de cooperatividad en cinética no-Michaeliana (DiStefano III, 2015; Weiss, 1997). La forma más simple de la función de Hill de orden n escrita para el proceso de crecimiento de *Lobesia botrana* con dos factores limitantes del crecimiento, se expresa como:

$$K_{jHills}(t) = \frac{\mu_{jmax}(t)ADD(t)_j^\alpha}{K_{jT}^\alpha + ADD(t)_j^\alpha} + \frac{\mu_{jmax}(t)H(t)^\gamma}{K_{jH}^\gamma + H(t)^\gamma} \quad (6.10)$$

6.2.3.2. Identificación de parámetros

Los parámetros constantes se identificaron utilizando el método de mínimos cuadrados no lineales y los parámetros variables en el tiempo se identificaron utilizando un Filtro de Kalman Extendido (EKF). Para el proceso de identificación, los vectores de parámetros desconocidos se definieron como se muestra en las Ecuaciones 6.11 - 6.13.

$$\theta_{u_{Monod}} = [K_{jT} \quad K_{jH} \quad K_{jd}] \quad (6.11)$$

$$\theta_{u_{Exp}} = [K_{jT} \quad K_{jH} \quad K_{jd} \quad \sigma_j] \quad (6.12)$$

$$\theta_{u_{Hills}} = [K_{jT} \quad K_{jH} \quad K_{jd} \quad \alpha \quad \gamma] \quad (6.13)$$

donde $\theta_{u_{Monod}}$, $\theta_{u_{Exp}}$, y $\theta_{u_{Hills}}$ son los vectores de parámetros desconocidos considerando las diferentes cinéticas planteadas en los escenarios de modelado.

6.2.3.2.1. Parámetros constantes: Utilizando un método de optimización numérica, se ajustan los parámetros empíricos del modelo para minimizar el error entre las salidas del modelo y los datos reales. Durante la etapa de estimación de parámetros constantes, los parámetros variables en el tiempo fueron fijados en los perfiles utilizados en un trabajo previo como se muestran en la Figura 6.5. Sin embargo, este valor fijado se “mejoró” a través de la implementación de un Filtro de Kalman Extendido (EKF) como se muestra en la Figura 6.6, de manera que fue posible obtener perfiles de mortalidad más suaves que encapsulan la incertidumbre de modelado asociada a este parámetro. Para la estimación de los parámetros

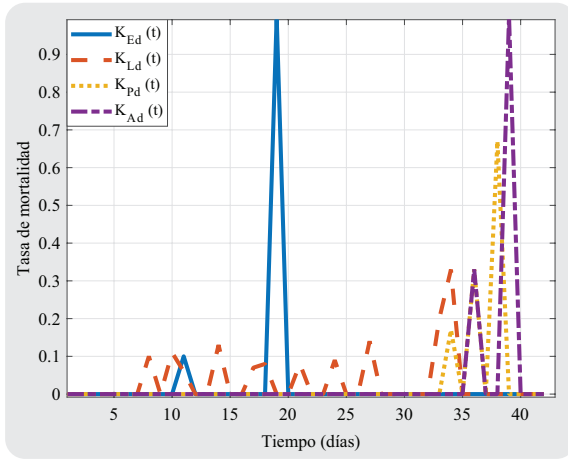


Figura 6.5: Perfiles de mortalidad diseñados a partir de los datos experimentales.

constantes se construyeron perfiles de mortalidad experimentales, tomando como base los datos de crecimiento de *Lobesia botrana* en condiciones de laboratorio ya descritas previamente. Esto se realizó tabulando los datos de crecimiento para calcular la cantidad de huevos, larvas pupas y adultos muertos por día (E_{dd} , L_{dd} , P_{dd} y A_{dd}). Lo anterior, sin considerar aquellos huevos, larvas o pupas que trascendían a la siguiente etapa de crecimiento. Las tasas de mortalidad diarias se calcularon como se muestra en las Ecuaciones 6.14-6.15.

$$K_{E_d}(1) = \frac{E_{dd}(1)}{E_{l_{max}}} \quad (6.14)$$

Para $t > 1$

$$K_{E_d} = \frac{E_{dd}(t)}{E_l(t-1)} \quad (6.15)$$

El problema de estimación de parámetros constantes se planteó como sigue:

- Considerando el vector de parámetros desconocidos

$$\theta = [K_{E_s} \quad K_{L_s} \quad K_{P_s} \quad K_{A_s}] \quad (6.16)$$

donde s es el factor limitante de crecimiento temperatura T o humedad relativa H .

- Se planteó el siguiente problema de optimización

$$\min_{\theta} f(x, \theta) = \sum_{i=1}^m |f_i(x, \theta) - y_i|^2 \quad (6.17)$$

sujeto a

$$lb_{\theta} = [0,001 \quad 0,001 \quad 0,001 \quad 0,001] \quad (6.18)$$

$$ub_{\theta} = [1 \times 10^5 \quad 1 \times 10^5 \quad 1 \times 10^5 \quad 1 \times 10^5] \quad (6.19)$$

donde lb_{θ} es el vector de límites inferiores sobre el vector θ y ub_{θ} es el vector de límites superiores sobre el vector θ . Las restricciones, tanto superiores como inferiores para el vector de parámetros, se definieron con base en el conocimiento del rango de variación de cada estado. Los parámetros $K_{E_s}, K_{L_s}, K_{P_s}, K_{A_s} > 0$ dado su significado físico y por condicionamiento matemático, debido a que se encuentran en el denominador de las ecuaciones que describen las tasas de crecimiento de la polilla en cada uno de sus estadios. Los límites superiores de los mismos parámetros fueron determinados partiendo de valores aleatorios que fueron modificados de manera iterativa hasta reducir el error en el ajuste. Para la solución del problema de optimización se utilizó el algoritmo de búsqueda local Levenberg-Marquardt.

El proceso de estimación de parámetros se realizó para el rango de temperatura de $20^{\circ}C \pm 2^{\circ}C$. Se eligieron los datos experimentales a dicha temperatura, dado que por ser una de las temperaturas ideales para el crecimiento de la polilla, permite observar comportamientos dinámicos en todos los estadios de crecimiento. Esto quiere decir que, para temperaturas muy altas o muy bajas, se inhibe el crecimiento de la polilla en el estadio adulto. Los valores de los parámetros constantes identificados para cada cinética de crecimiento se presentan en la Tabla 6.3. A continuación se presenta el análisis de identificabilidad para cada escenario, así como las curvas de respuesta del modelo ajustado con cada cinética de crecimiento.

Tabla 6.3: Valores de los parámetros identificados por el método de mínimos cuadrados no lineales para cada cinética de crecimiento. El símbolo $\otimes$ indica que este parámetro no forma parte de la cinética correspondiente, y θ corresponde al parámetro.

θ	Valores identificados		
	Monod	Exponencial	Hills
K_{E_T}	$2,1 \times 10^4$	$6,48 \times 10^2$	5×10^3
K_{L_T}	$4,1 \times 10^2$	$1,48 \times 10^2$	$2,51 \times 10^2$
K_{P_T}	$2,47 \times 10^2$	$4,95 \times 10^2$	$4,5 \times 10^2$
K_{A_T}	$9,85 \times 10^2$	$4,85 \times 10^2$	$1,45 \times 10^3$
K_{E_H}	$6,65 \times 10^4$	$2,95 \times 10^2$	$5,8 \times 10^2$
K_{L_H}	$2,16 \times 10^2$	45,95	$7,81 \times 10^2$
K_{P_H}	$2,27 \times 10^3$	12,10	$3,85 \times 10^3$
K_{A_H}	$1,08 \times 10^3$	$1,25 \times 10^2$	$0,98 \times 10^3$
σ_E	$\otimes$	0,54	$\otimes$
σ_L	$\otimes$	3,55	$\otimes$
σ_P	$\otimes$	10,5	$\otimes$
σ_A	$\otimes$	15,25	$\otimes$
α	$\otimes$	$\otimes$	0,99
γ	$\otimes$	$\otimes$	0,99

De acuerdo con los resultados presentados en la Tabla 6.3, se observa que los parámetros asociados a la cinética exponencial tienen valores considerablemente más bajos que para las otras cinéticas de crecimiento.

6.2.3.2.2. Parámetros variables en el tiempo: para el ajuste de las tasas de mortalidad, se consideró el uso del Filtro de Kalman Extendido, que es una variación del filtro de Kalman, que permite abordar el problema de la estimación de estados y parámetros variantes en el tiempo en sistemas no lineales (Simon, 2006). En todos los casos, la tasa de mortalidad de los huevos (K_{E_d}) no se consideró en el problema de identificación, ya que el análisis de identificabilidad previo presentado en (Aguirre-Zapata et al., 2022a) indicaba que este parámetro debía fijarse para garantizar la identificabilidad de los demás parámetros. Dada la interpretabilidad estructural (ISP) de este parámetro, fue posible reconstruirlo a partir de datos experimentales o de la literatura existente, por lo que K_{E_d} se fijó como parámetro conocido.

Considere un sistema no lineal de tiempo discreto que depende de forma no lineal de un vector de parámetros desconocidos θ_k , como se muestra en la Ecuación 6.20.

$$\begin{aligned}x_k &= f(x_{k-1}, \theta_k, u_{k-1}, w_{k-1}) \\y_k &= g_k(x_k, \theta_k, u_k, v_k)\end{aligned}\tag{6.20}$$

donde $f(x_{k-1}, \theta_k, u_{k-1}, w_{k-1})$ es la función de transición de estado del sistema dinámico; $g(x_k)$ es la función de medición; $x_k \in R^n$ es el vector de estados del sistema con condición inicial $x(0)$; u_k es el vector de entradas del sistema dinámico. Considerando el sistema anterior, se encuentran en la literatura diferentes aproximaciones para la estimación de estados y parámetros usando un EKF, tales como los algoritmos de estimación conjunta, estimación dual o los escenarios de estimación de parámetros cuando el estado del sistema es conocido (Simon, 2006). En este sentido, el Filtro de Kalman Extendido se convierte en un estimador de parámetros al combinar los elementos del vector de estado del sistema y el vector de parámetros θ_k , lo que se conoce como estimación conjunta o de estado aumentado. Para ello, se considera la dinámica de estado aumentado como se muestra en la Ecuación 6.21.

$$\begin{aligned}\begin{bmatrix} x_k \\ \theta_k \end{bmatrix} &= \begin{bmatrix} f_{k-1}(x_{k-1}, u_{k-1}, \theta_{k-1}, w_{k-1}) \\ \theta_{k-1} + r_{k-1} \end{bmatrix} \\y_k &= g_k(x_k, u_k, \theta_k, v_k)\end{aligned}\tag{6.21}$$

Para simplificar la notación puede escribirse el sistema como se muestra en la Ecuación 6.22.

$$\begin{aligned}X_k &= F_{k-1}(X_{k-1}, u_{k-1}, W_{k-1}) \\y_k &= g_k(X_k, u_k, v_k)\end{aligned}\tag{6.22}$$

donde X_k es el nuevo vector de estado aumentado, que comprende tanto el estado como los parámetros, W_k es el vector que comprende el ruido del proceso y el ruido de los parámetros, F es la función que combina la dinámica del estado y la dinámica de los parámetros, y r_{k-1} es un pequeño término de ruido artificial que permite al EKF cambiar su estimación de θ_k . Con el modelo aumentado, se aplican las ecuaciones del EKF como se muestra en la Tabla 6.4, cuya implementación se da en cuatro momentos: la definición del sistema dinámico no lineal, la inicialización del filtro,

la linealización punto a punto del sistema dinámico no lineal definido inicialmente, y por último la actualización de la estimación como se detalla en (Simon, 2006). donde Q y R son matrices de covarianza definidas

Tabla 6.4: Pseudocódigo de implementación del Filtro de Kalman Extendido

Calcular para $k = 1, 2, \dots, n$
Actualización de la estimación a priori
$\hat{X}_k^- = F_{k-1}(\hat{X}_{k-1}^+, u_{k-1})$
Actualización de la covarianza del error a priori
$P_k^- = A_{k-1}P_{k-1}^+A_{k-1}^T + Q_{k-1}$
Cálculo de la ganancia de Kalman
$K_k = P_k^-C_k^T(C_kP_k^-C_k^T + R_k)^{-1}$
Actualización de la estimación a posteriori
$\hat{X}_k^+ = \hat{X}_k^- + K_k[y_k - g_k(\hat{X}_k^-)]$
Actualización de la covarianza del error a posteriori
$P_k^+ = (I - K_kC_k)P_k^-$

positivas, A corresponden a la matriz de derivadas parciales del sistema F respecto al estado X y C corresponde a la matriz de derivadas parciales de la salida del sistema g respecto al estado X . Se consideró un sistema no lineal de la forma que se presentó en la Ecuación 6.22, donde se combinaron los estados del sistema dinámico y las tasas de mortalidad asociadas a cada estadio de crecimiento de la polilla (parámetros desconocidos), para la estimación de estos últimos. El vector de estados $X \in R^{+(7x1)}$ se presenta en las Ecuaciones 6.23 y 6.24.

$$X_k = [X_k^1 \ X_k^2 \ X_k^3 \ X_k^4 \ X_k^5 \ X_k^6 \ X_k^7]^T \quad (6.23)$$

$$X_k = [E_{l_k} \ L_{l_k} \ P_{l_k} \ A_{l_k} \ K_{L_{dk}} \ K_{P_{dk}} \ K_{A_{dk}}]^T \quad (6.24)$$

Considerando el sistema anterior, el EKF fue programado y simulado en MATLAB® & Simulink®, considerando las condiciones iniciales que se presentan en las Ecuaciones 6.25 y 6.26.

$$x_0 = [30 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \quad (6.25)$$

$$P_0^+ = 150I_7 \quad (6.26)$$

donde I_7 corresponde a la matriz identidad. Las matrices de covarianza Q y R fueron ajustadas empíricamente mediante simulaciones, obteniendo como resultado las matrices que se muestran en las Ecuaciones 6.27 Y 6.28.

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \quad (6.27)$$

$$R = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 20 \end{bmatrix} \quad (6.28)$$

En la Figura 6.6 se presentan los perfiles de mortalidad estimados mediante el EKF, en contraste con los perfiles de mortalidad experimentales, dónde la línea gris (-) corresponde a las tasas de mortalidad experimentales (Ecuaciones 6.14-6.15) a $20^{\circ}C$. La línea azul (-.-) corresponde a las tasas de mortalidad estimadas mediante el EKF. Se observan algunas similitudes entre los perfiles experimentales y los perfiles estimados. Sin embargo, los perfiles estimados permiten obtener una respuesta más suave del modelo como se muestra en la validación de cada escenario de modelado.

En la Figura 6.6 se observan algunas discrepancias entre las tasas de mortalidad construidas a partir de las tablas de vida presentadas por Herrera et al., 2016 y las estimadas con el EKF. Lo anterior se explica por el hecho de que las tasas de mortalidad no son constantes como lo sugieren las crías de *Lobesia botrana* en condiciones de laboratorio y no son función del tiempo, sino de la temperatura como se explicará en la Sección 6.3.5.2.

6.2.3.3. Análisis de interpretabilidad y validación

A continuación se presentan los resultados del análisis de interpretabilidad para cada uno de los escenarios de modelado planteados (Figura 6.4) considerando diferentes ecuaciones para la representación de la cinética de crecimiento de *Lobesia botrana*. Para todos los casos, el escenario de

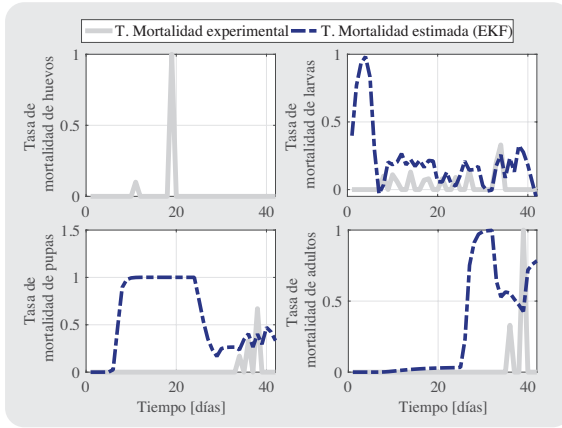


Figura 6.6: Tasas de mortalidad de *Lobesia botrana* estimadas mediante el EKF comparadas con las tasas de mortalidad construidas con las Ecuaciones 6.14-6.15.

partida que se analiza es la estructura básica del modelo. La estructura extendida se muestra en los escenarios 1 – 3. En el Anexo A.3 se especifican las condiciones de simulación utilizadas para el cálculo de sensibilidades de los parámetros, así como la solución del modelo computacional.

Además del análisis de interpretabilidad se presentan los resultados cualitativos y cuantitativos obtenidos a partir del ajuste del modelo y la comparación del mismo con datos experimentales a escala de laboratorio. Los resultados cualitativos se presentan en términos de error absoluto y del error porcentual absoluto, que se calculan como se muestra en las Ecuaciones 6.29 y 6.30. También se presentan el error absoluto máximo y mínimo para cada una de las fases de crecimiento.

$$MAE_{j_i} = \sum_{i=1}^n \frac{|y_{j_i} - x_{j_i}|}{n} \quad (6.29)$$

$$MAE_{j_i} [\%] = 100 \frac{MAE_{j_i}}{\max \{x_{j_i}\}} \quad (6.30)$$

dónde n es la duración del estadio de crecimiento, y_{ji} el número de huevos, larvas, pupas o adultos predicho por el modelo, y x_{ji} el número de huevos, larvas, pupas o adultos obtenidos de manera experimental.

6.2.3.3.1. Escenario 0. Estructura básica: el análisis de interpretabilidad para el escenario 0 se muestra en la Tabla 6.5. En ella se presenta la clasificación de los parámetros de la estructura básica del modelo de crecimiento *Lobesia botrana* en función de su interpretabilidad según la taxonomía propuesta en el apartado 4.2, con su respectivo valor ordinal según la Ecuación 4.5. En la misma tabla se indica el nivel de especi-

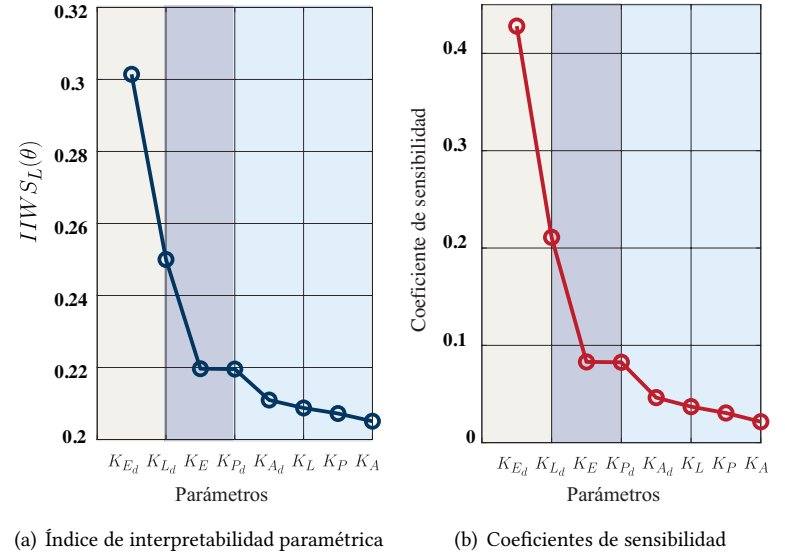


Figura 6.7: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de *Lobesia botrana* en el escenario 0.

cación en el que se encuentra cada uno de los parámetros del modelo y el resultado del análisis de sensibilidad. Considerando los elementos de la Tabla 6.5, el índice de interpretabilidad para la estructura básica o

escenario 0 fue $IIWS_{L,0} = 1$. La Figura 6.7 ilustra la jerarquía de los parámetros en el modelo de crecimiento *Lobesia botrana*, basada tanto en el índice de interpretabilidad de cada parámetro (Figura 6.7(a)), como en los coeficientes de sensibilidad paramétrica (Figura 6.7(b)). En este

Tabla 6.5: Análisis de interpretabilidad del modelo de crecimiento de *Lobesia botrana* en el escenario 0.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_l(\theta)$
K_E	ISP	1	0	1	$8,3 \times 10^{-2}$	0,19	$1,05 \times 10^{-1}$
K_L	ISP	1	0	1	$3,7 \times 10^{-2}$	0,09	$1,01 \times 10^{-1}$
K_P	ISP	1	0	1	3×10^{-2}	0,07	$1,02 \times 10^{-1}$
K_A	ISP	1	0	1	$2,2 \times 10^{-2}$	0,05	$1,01 \times 10^{-1}$
K_{E_d}	ISP	1	0	1	$4,3 \times 10^{-2}$	1	$1,3 \times 10^{-1}$
K_{L_d}	ISP	1	0	1	$2,1 \times 10^{-1}$	0,49	$1,12 \times 10^{-1}$
K_{P_d}	ISP	1	0	1	$8,3 \times 10^{-2}$	0,19	$1,05 \times 10^{-1}$
K_{A_d}	ISP	1	0	1	$4,6 \times 10^{-2}$	0,11	$1,03 \times 10^{-1}$

caso, la estructura básica del modelo carece de parámetros fundamentales necesarios para la generalización del modelo, tales como la temperatura y la humedad que regulan las transiciones de un estadio de crecimiento a otro. En consecuencia, se hace necesario derivar ecuaciones constitutivas para determinar la estructura ampliada del modelo e incorporar estos parámetros esenciales.

6.2.3.3.2. Escenario 1. Parámetro K_j descrito mediante ecuación de Monod: el análisis de interpretabilidad para el escenario 1 se muestra en la Tabla 6.6. De la misma manera que para el escenario anterior, la tabla contiene el resultado de los análisis previos al cálculo del índice de interpretabilidad. Considerando los elementos presentados en la Tabla 6.6, el índice de interpretabilidad para el escenario 1 fue $IIWS_{L,1} = 0,59$. La Figura 6.8 ilustra la jerarquía de parámetros en el modelo de crecimiento *Lobesia botrana*, basada tanto en el índice de interpretabilidad de cada parámetro, como en el coeficiente de sensibilidad de los mismos (Figuras 6.8(a) y 6.8(b)), respectivamente). Allí se muestra que, tras la incorporación de los nuevos parámetros, la tasa de mortalidad de los huevos mantiene sistemáticamente el mayor grado de influencia tanto en la interpretabilidad global como en la sensibilidad del modelo. Por el

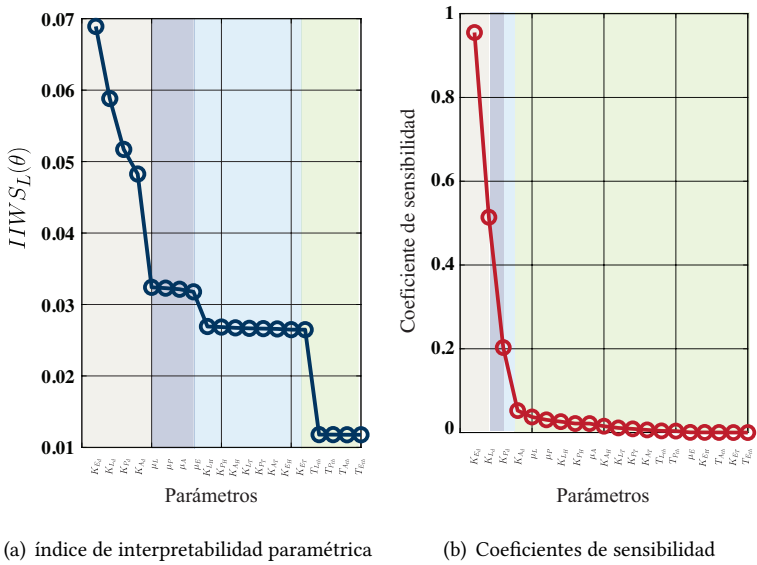


Figura 6.8: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de *Lobesia botrana* en el escenario 1.

contrario, las nuevas temperaturas umbral introducidas en este escenario tienen una menor influencia sobre la interpretabilidad general del modelo, y ocupan el último lugar con el menor coeficiente de sensibilidad. Esto último flexibiliza la identificación de estos parámetros, considerando que la polilla de la vid tiene la capacidad de adaptarse a su entorno y modificar sus mecanismos de regulación internos. Diferentes estudios en la literatura demuestran que las temperaturas umbrales varían considerablemente dependiendo de la región particular para la cual se parametriza el modelo (Carlos et al., 2018; Castex et al., 2018; Gilioli et al., 2016; Gutierrez et al., 2018).

En la Figura 6.9 se presentan las curvas de crecimiento de *Lobesia botrana* con la cinética de crecimiento de Monod, y considerando dos factores limitantes de crecimiento, en contraste con los datos experimentales de tres réplicas diferentes a $20 \pm 2^\circ C$. Los errores absolutos asociadas a

Tabla 6.6: Análisis de interpretabilidad del modelo de crecimiento de *Lobesia botrana* en el escenario 1.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_L(\theta)$
μ_E	IPI	0,90	1	0,8	$4,6 \times 10^{-4}$	$4,8 \times 10^{-4}$	$2,7 \times 10^{-2}$
μ_L	IPI	0,90	1	0,8	$3,7 \times 10^{-2}$	$3,9 \times 10^{-2}$	2,7
μ_P	IPI	0,90	1	0,8	$3,0 \times 10^{-2}$	$3,2 \times 10^{-2}$	$2,7 \times 10^{-2}$
μ_A	IPI	0,90	1	0,8	$2,2 \times 10^{-2}$	$2,3 \times 10^{-2}$	$2,7 \times 10^{-2}$
K_{E_T}	PIPI	0,75	1	0,8	$6,9 \times 10^{-5}$	$7,2 \times 10^{-5}$	$2,3 \times 10^{-2}$
K_{L_T}	PIPI	0,75	1	0,8	$1,1 \times 10^{-2}$	$1,2 \times 10^{-2}$	$2,3 \times 10^{-2}$
K_{P_T}	PIPI	0,75	1	0,8	$8,8 \times 10^{-3}$	$9,2 \times 10^{-3}$	$2,3 \times 10^{-2}$
K_{A_T}	PIPI	0,75	1	0,8	$6,3 \times 10^{-3}$	$6,6 \times 10^{-3}$	$2,3 \times 10^{-2}$
K_{E_H}	PIPI	0,75	1	0,8	$3,9 \times 10^{-4}$	$4,1 \times 10^{-4}$	$2,3 \times 10^{-2}$
K_{L_H}	PIPI	0,75	1	0,8	$2,6 \times 10^{-2}$	$2,7 \times 10^{-2}$	$2,3 \times 10^{-2}$
K_{P_H}	PIPI	0,75	1	0,8	$2,2 \times 10^{-2}$	$2,3 \times 10^{-2}$	$2,3 \times 10^{-2}$
K_{A_H}	PIPI	0,75	1	0,8	$1,5 \times 10^{-2}$	$1,6 \times 10^{-2}$	$2,3 \times 10^{-2}$
K_{E_d}	ISP	1	0	1	$9,6 \times 10^{-1}$	1	5×10^{-2}
K_{L_d}	ISP	1	0	1	$5,1 \times 10^{-1}$	$5,4 \times 10^{-1}$	$4,5 \times 10^{-2}$
K_{P_d}	ISP	1	0	1	$2,0 \times 10^{-1}$	$2,1 \times 10^{-1}$	$4,2 \times 10^{-1}$
K_{A_d}	ISP	1	0	1	$5,3 \times 10^{-2}$	$5,5 \times 10^{-2}$	$4,1 \times 10^{-2}$
$T_{E_{th}}$	GI	1	3	0,3	$4,6 \times 10^{-5}$	$4,8 \times 10^{-5}$	1×10^{-2}
$T_{L_{th}}$	GI	1	3	0,3	4×10^{-3}	$4,2 \times 10^{-3}$	1×10^{-2}
$T_{P_{th}}$	GI	1	3	0,3	$3,6 \times 10^{-3}$	$3,7 \times 10^{-3}$	1×10^{-2}
$T_{A_{th}}$	GI	1	3	0,3	$2,6 \times 10^{-4}$	$2,7 \times 10^{-4}$	1×10^{-2}

cada estadio de crecimiento respecto a los datos experimentales originales y a los datos experimentales preprocesados, se presenta en la Tabla 6.7. La Figura 6.9 muestra que el ajuste para todas las etapas de crecimiento es suave y sigue la tendencia de los datos experimentales. el estadio de huevo (E_l) muestra un comportamiento rápido que puede explicarse por incertidumbres de modelado originadas por las etapas intermedias del desarrollo del huevo que no se modelaron (Torres Vila, 1995). Sin embargo, es el estadio huevo, el que muestra el error absoluto más bajo, equivalente a 6,47 %. El estadio larvario (L_l) muestra sobreexcitación en los primeros

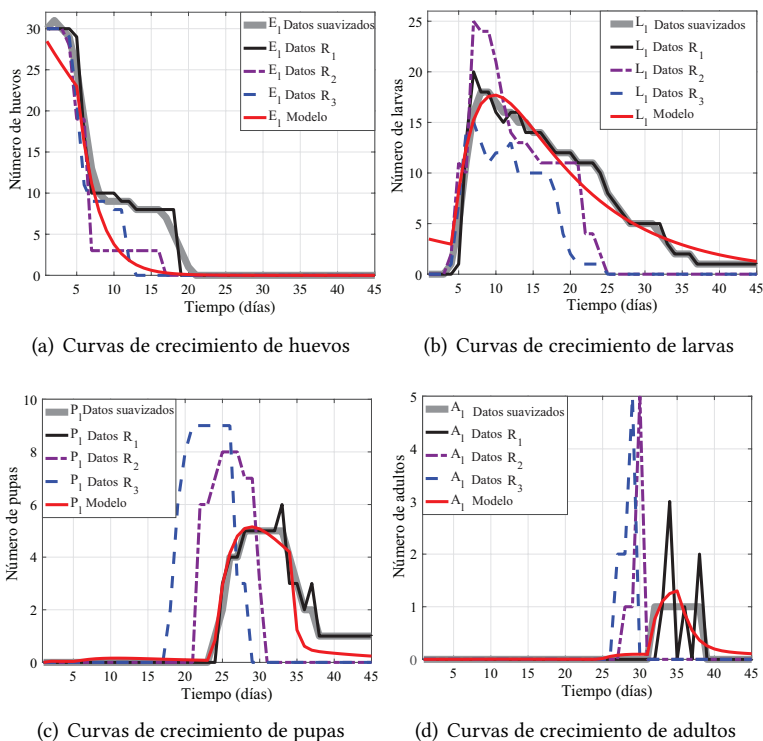


Figura 6.9: *Lobesia botrana* curvas de crecimiento a $20^{\circ}\text{C} \pm 2^{\circ}\text{C}$ para el escenario 1.

días de crecimiento, lo que le hace tener un error absoluto mínimo de 1, 1 %. Sin embargo, se ajusta rápidamente y describe correctamente la dinámica de esa etapa de crecimiento. El estadio adulto (A_l) muestra uno de los errores absolutos máximos más elevados (65 %) en torno al día 35.

Tabla 6.7: Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética de Monod, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$.

Estadio	MAE	MAE %	MAE Max %	MAE Min %
	Número de E_l, L_l, P_l or A_l			
E_l	2,0	6,47	24,43	0
L_l	1,19	6,63	19,31	1,1
P_l	0,36	7,14	35,34	0
A_l	0,08	0,18	63,72	0

6.2.3.3.3. Escenario 2. Parámetro K_j descrito mediante ecuación exponencial: el análisis de interpretabilidad para el escenario 2 se muestra en la Tabla 6.8. Considerando los análisis previos, el índice de interpretabilidad para el escenario 2 fue $IIW S_{L,1} = 0,49$. La Figura 6.8 ilustra la jerarquía de parámetros en el modelo de crecimiento *Lobesia botrana*, basada en el índice de interpretabilidad de cada parámetro (Figura 6.10(a)) y los coeficientes de sensibilidad paramétrica (Figura 6.10(b)). Allí se muestra que, al incluir cuatro parámetros empíricos no interpretables, la jerarquía basada en los coeficientes de sensibilidad paramétrica se ve afectada presentando resultados poco deseables en este tipo de modelos. Se observa que parámetros relevantes en el proceso de crecimiento como las temperaturas umbrales (por ejemplo $T_{A_{th}}$) pasan a tener una influencia casi nula sobre la salida del modelo, lo que las convierte en parámetros de sacrificio¹.

¹Un parámetro de sacrificio es un parámetro no interpretable que se ajusta para minimizar el error de ajuste del modelo. Puede tomar valores en un rango más amplio que un parámetro interpretable y su función principal es mejorar la precisión del ajuste del modelo (Lopez-

Restrepo et al., [2020](#)).

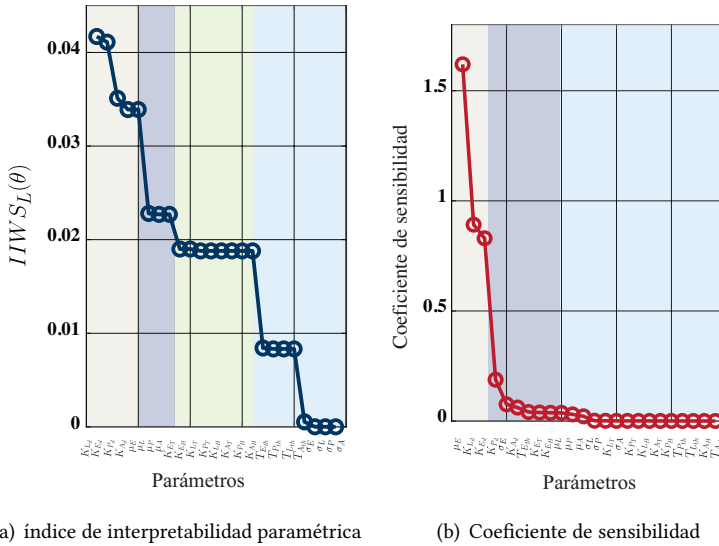


Figura 6.10: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de crecimiento de *Lobesia botrana* en el escenario 2.

Por otra parte, el parámetro σ_E no interpretable, resulta ser uno de los parámetros con mayor influencia sobre la salida del modelo. Como se ha mencionado para otros casos de aplicación, este resultado es indeseable, considerando que al ser un parámetro adimensional es difícil establecer cotas para su ajuste.

En la Figura 6.11 se presentan las curvas de crecimiento de *Lobesia botrana* con la cinética de crecimiento exponencial y considerando dos factores limitantes de crecimiento, en contraste con los datos experimentales de tres réplicas diferentes a $20 \pm 2^\circ C$. Para este caso, se observan cambios más abruptos en la dinámica, en comparación con los obtenidos con la cinética de Monod. En cuanto al estadio huevo E_l , la cinética exponencial permite describir su dinámica más lentamente en contraste con la

Tabla 6.8: Análisis de interpretabilidad del modelo de crecimiento de *Lobesia botrana* en el escenario 2.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_L(\theta)$
μ_E	IPI	0,90	1	0,8	$1,6 \times 10^{-4}$	$1,8 \times 10^{-4}$	$3,4 \times 10^{-2}$
μ_L	IPI	0,90	1	0,8	$3,7 \times 10^{-2}$	$4,1 \times 10^{-2}$	$2,3 \times 10^{-2}$
μ_P	IPI	0,90	1	0,8	3×10^{-2}	$3,4 \times 10^{-2}$ 3.18	$2,3 \times 10^{-2}$
μ_A	IPI	0,90	1	0,8	$2,2 \times 10^{-2}$	$2,4 \times 10^{-2}$	$2,3 \times 10^{-2}$
K_{E_T}	PIPI	0,75	1	0,8	$3,9 \times 10^{-5}$	$4,4 \times 10^{-5}$	$1,9 \times 10^{-2}$
K_{L_T}	PIPI	0,75	1	0,8	$8,6 \times 10^{-2}$	$9,6 \times 10^{-2}$	$1,9 \times 10^{-2}$
K_{P_T}	PIPI	0,75	1	0,8	$7,5 \times 10^{-3}$	$8,4 \times 10^{-4}$	$1,9 \times 10^{-2}$
K_{A_T}	PIPI	0,75	1	0,8	5×10^{-4}	$5,6 \times 10^{-4}$	$1,9 \times 10^{-2}$
K_{E_H}	PIPI	0,75	1	0,8	$3,8 \times 10^{-2}$	$4,2 \times 10^{-2}$	$1,9 \times 10^{-2}$
K_{L_H}	PIPI	0,75	1	0,8	$6,3 \times 10^{-4}$	$7,1 \times 10^{-4}$	$1,9 \times 10^{-2}$
K_{P_H}	PIPI	0,75	1	0,8	5×10^{-4}	$5,6 \times 10^{-4}$	$1,9 \times 10^{-2}$
K_{A_H}	PIPI	0,75	1	0,8	3×10^{-4}	$3,4 \times 10^{-4}$	$1,9 \times 10^{-2}$
K_{E_d}	ISP	1	0	1	$8,3 \times 10^{-1}$	$9,3 \times 10^{-1}$	$4,1 \times 10^{-2}$
K_{L_d}	ISP	1	0	1	$8,9 \times 10^{-1}$	1	$4,2 \times 10^{-2}$
K_{P_d}	ISP	1	0	1	$1,9 \times 10^{-1}$	$2,1 \times 10^{-1}$	$3,5 \times 10^{-2}$
K_{A_d}	ISP	1	0	1	$6,1 \times 10^{-2}$	$6,9 \times 10^{-2}$	$3,4 \times 10^{-2}$
$T_{E_{th}}$	GI	1	3	0,3	$4,1 \times 10^{-2}$	$4,6 \times 10^{-2}$	$8,4 \times 10^{-3}$
$T_{L_{th}}$	GI	1	3	0,3	$3,9 \times 10^{-4}$	$4,3 \times 10^{-4}$	$8,3 \times 10^{-3}$
$T_{P_{th}}$	GI	1	3	0,3	$4,1 \times 10^{-4}$	$4,6 \times 10^{-4}$	$8,3 \times 10^{-3}$
$T_{A_{th}}$	GI	1	3	0,3	$2,9 \times 10^{-4}$	$3,3 \times 10^{-4}$	$8,3 \times 10^{-3}$
σ_E	NIP	0	1	0,8	$7,6 \times 10^{-2}$	$8,5 \times 10^{-2}$	$5,3 \times 10^{-4}$
σ_L	NIP	0	1	0,8	$1,5 \times 10^{-3}$	$1,7 \times 10^{-3}$	1×10^{-5}
σ_P	NIP	0	1	0,8	$1,2 \times 10^{-3}$	$1,39 \times 10^{-3}$	$8,7 \times 10^{-6}$
σ_A	NIP	0	1	0,8	$8,1 \times 10^{-4}$	9×10^{-4}	$5,7 \times 10^{-6}$

cinética Monod. Sin embargo, el estadio adulto presenta un sobreimpulso indeseable, que hace que el error en este estadio sea superior al 100 %. El error MAE máximo para todos los demás estadios es significativamente superior al de la cinética analizada en el apartado anterior. El error absoluto MAE asociado a cada estadio de crecimiento, con respecto a los datos experimentales originales y a los datos experimentales preprocesados, se presenta en la Tabla 6.9.

6.2.3.3.4. Escenario 3. Parámetro K_j descrito mediante ecuación de Hill: el análisis de interpretabilidad para el escenario 3 se muestra

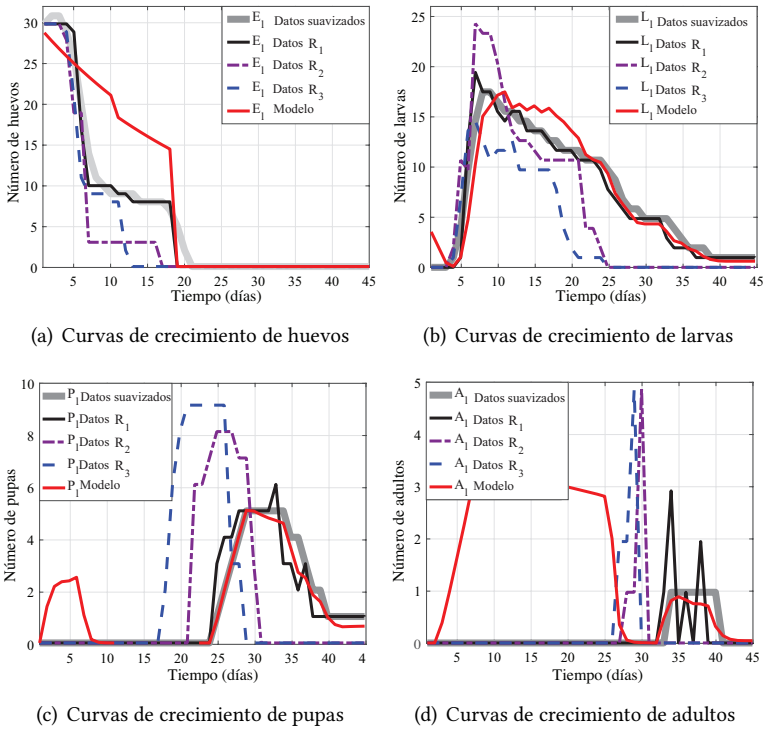


Figura 6.11: *Lobesia botrana* curvas de crecimiento a $20^{\circ}\text{C} \pm 2^{\circ}\text{C}$ para el escenario 2.

en la Tabla 6.10. Esta tabla muestra la clasificación de los parámetros de la estructura extendida del modelo de crecimiento, en términos de su interpretabilidad según la taxonomía propuesta. Considerando el análisis de las demás propiedades del modelo, el índice de interpretabilidad para el escenario 3 fue $IIWS_{L,3} = 0,54$. La Figura 6.12 ilustra la jerarquización paramétrica basada tanto en el índice de interpretabilidad como en el coeficiente de sensibilidad de los parámetros del modelo de crecimiento de *Lobesia botrana* en el escenario 3 (Figuras 6.12(a) -6.12(b)).

Tabla 6.9: Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética exponencial, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$.

Estadio	MAE Número de E_l, L_l, P_l o A_l	MAE %	MAE Max %	MAE Min %
E_l	3,05	9,83	41,7	0
L_l	1,39	7,74	33,8	0,3
P_l	0,54	10,79	49,7	0
A_l	1,59	159,6	335,9	0

Tabla 6.10: Análisis de interpretabilidad del modelo de crecimiento de *Lobesia botrana* en el escenario 3.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_L(\theta)$
μ_E	IPI	0,90	1	0,8	$3,9e^{-1}$	0,18	$2,6 \times 10^{-2}$
μ_L	IPI	0,90	1	0,8	$3,7 \times 10^{-2}$	0,02	$2,5 \times 10^{-2}$
μ_P	IPI	0,90	1	0,8	$3,0 \times 10^{-2}$	0,01	$2,5 \times 10^{-2}$
μ_A	IPI	0,90	1	0,8	$2,2 \times 10^{-2}$	0,01	$2,5 \times 10^{-2}$
K_{E_T}	PIPI	0,75	1	0,8	$1,4 \times 10^{-2}$	0,006	2×10^{-2}
K_{L_T}	PIPI	0,75	1	0,8	$5,2 \times 10^{-3}$	0,002	2×10^{-2}
K_{P_T}	PIPI	0,75	1	0,8	$4,2 \times 10^{-3}$	0,002	2×10^{-2}
K_{A_T}	PIPI	0,75	1	0,8	3×10^{-3}	0,001	2×10^{-2}
K_{E_H}	PIPI	0,75	1	0,8	$1,6 \times 10^{-2}$	0,007	$2,1 \times 10^{-2}$
K_{L_H}	PIPI	0,75	1	0,8	$7,4 \times 10^{-3}$	0,003	2×10^{-2}
K_{P_H}	PIPI	0,75	1	0,8	$6,1 \times 10^{-3}$	0,003	2×10^{-2}
K_{A_H}	PIPI	0,75	1	0,8	$4,3 \times 10^{-3}$	0,002	2×10^{-2}
K_{E_d}	ISP	1	0	1	2,1	1	$4,5 \times 10^{-2}$
K_{L_d}	ISP	1	0	1	3×10^{-1}	0,14	$3,8 \times 10^{-2}$
K_{P_d}	ISP	1	0	1	$3,1 \times 10^{-1}$	0,14	$3,8 \times 10^{-2}$
K_{A_d}	ISP	1	0	1	$4,9 \times 10^{-2}$	0,023	$3,7 \times 10^{-2}$
$T_{E_{th}}$	GI	1	3	0,3	$7,2 \times 10^{-3}$	0,004	$9,1 \times 10^{-3}$
$T_{L_{th}}$	GI	1	3	0,3	2×10^{-3}	0	$9,1 \times 10^{-3}$
$T_{P_{th}}$	GI	1	3	0,3	2×10^{-3}	0	$9,1 \times 10^{-3}$
$T_{A_{th}}$	GI	1	3	0,3	$1,3 \times 10^{-3}$	0	$9,1 \times 10^{-3}$
α	NIP	0	1	0,8	$2,3 \times 10^{-1}$	0,11	$7,3 \times 10^{-4}$
γ	NIP	0	1	0,8	$2,5 \times 10^{-1}$	0,12	$7,9 \times 10^{-4}$

trana con la cinética de crecimiento de Hill, y considerando dos factores limitantes de crecimiento en contraste con los datos experimentales de tres réplicas diferentes a $20 \pm 2^\circ C$. El MAE asociado a cada etapa de crecimiento, con respecto a los datos experimentales originales y a los datos experimentales preprocesados, se presenta en la Tabla 6.11. La cinética Hill, como se muestra en la Figura 6.13, tuvo un ajuste similar al obtenido con la cinética de Monod. La principal diferencia entre la cinética Hill y la cinética de Monod se observa en el estadio adulto, donde existe una pequeña sobreexcitación en la mitad del ciclo de crecimiento que se compensa en términos del error con la amplitud de la curva en estado estacionario. La cinética de Hills, favorece que el modelo presente un error absoluto MAE mínimo de casi 0 % en todas sus etapas de crecimiento.

Tabla 6.11: Errores absolutos medios entre el modelo matemático ajustado con tasas de crecimiento representadas por la cinética de Hill, los datos experimentales originales y los datos experimentales preprocesados para el rango de temperatura de $20 \pm 2^\circ C$.

Estadio	MAE	MAE %	MAE Max %	MAE Min %
	Número de E_l, L_l, P_l, A_l			
E_l	1,32	4,27	21,65	0
L_l	1,41	7,86	24,47	0,006
P_l	0,24	4,73	15,77	0
A_l	0,23	23,11	65,42	0

6.2.3.4. Resultados comparativos

En este contexto, las gráficas de radar representadas en las Figuras 6.14-6.15 ilustran la progresión del índice de interpretabilidad $IIWS_L(\theta)$ dentro de cada escenario de modelado. El ID $(1, 2, \dots, n)$ de cada parámetro en la figura, corresponde al ID $(1, 2, \dots, n)$ en la Tabla 6.2. Un parámetro al que se asigna un valor cero significa su ausencia en el escenario actual considerado. La Figura 6.14 muestra la progresión individual de la influencia de cada parámetro en el índice general de interpretabilidad del modelo de crecimiento *Lobesia botrana* en cada escenario considerado.

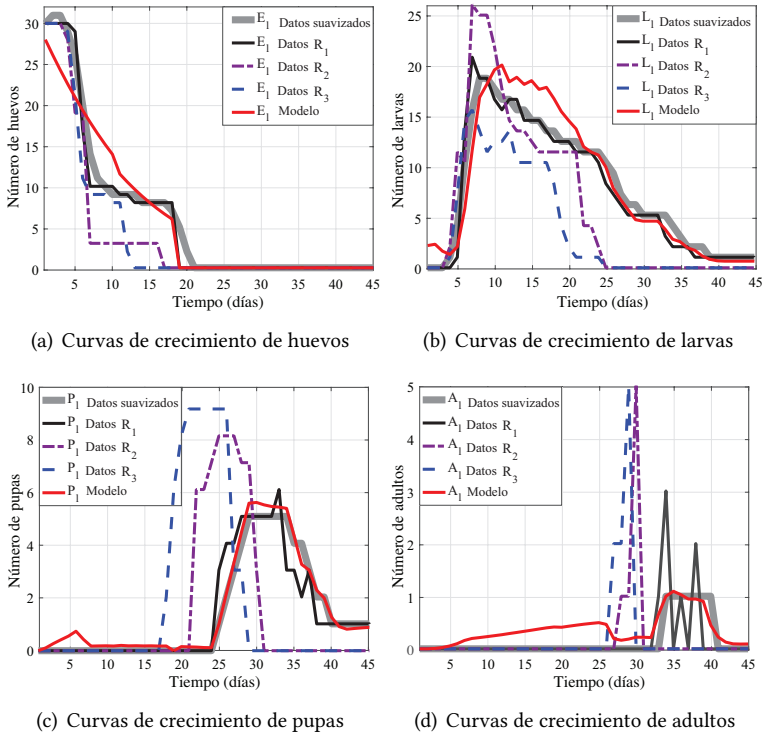


Figura 6.13: Curvas de crecimiento de *Lobesia botrana* a $20^{\circ}\text{C} \pm 2^{\circ}\text{C}$ para el escenario 3.

En particular, los parámetros 1 a 8 de la estructura básica del modelo matemático, presentan sistemáticamente el índice de interpretabilidad más elevado. En este caso concreto de aplicación, un repaso de la Figura 6.15 permite identificar el escenario 2 como el menos favorable en términos de interpretabilidad. Dentro de este escenario, el modelo experimenta una notable degradación, evidente no sólo en el índice de interpretabilidad de los parámetros recién introducidos que se aproxima a cero, sino también en la disminución del índice de interpretabilidad de los parámetros preexistentes dentro del modelo matemático.

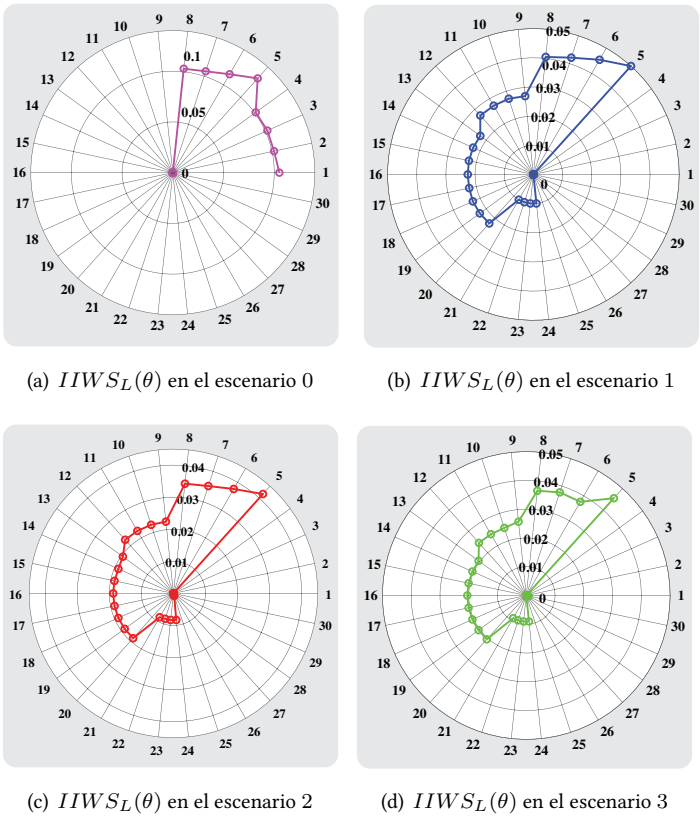


Figura 6.14: Influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado del proceso de crecimiento de *Lobesia botrana*.

Así pues, el escenario de modelado 1 se constituye como el mejor en términos de la relación entre la interpretabilidad y la sensibilidad paramétrica.

La Tabla 6.12 resume los resultados cuantitativos tanto del análisis de interpretabilidad como de la evaluación del ajuste del modelo a los datos experimentales, realizado para cada escenario de modelado. El cambio en

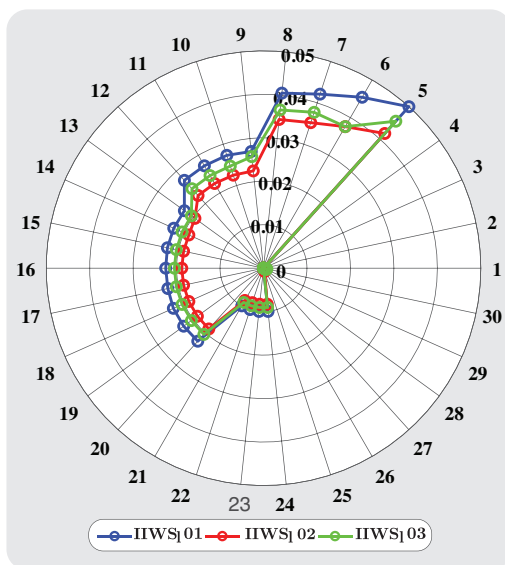


Figura 6.15: Comparación de la influencia de cada parámetro en el índice global de interpretabilidad para todos los escenarios de modelado de *Lobesia botrana*.

el índice representa el grado de mejora (+) o degradación (−) del índice de un escenario particular, en comparación con el escenario de modelado anterior (punto de partida) de acuerdo con el grafo de modelado presentado en la Figura 6.4. Como puede observarse, en el escenario 1 el índice de interpretabilidad sufre una degradación natural de la interpretabilidad relativa a la estructura básica del modelo. Sin embargo, como se ha mencionado anteriormente, el escenario 0 carece de algunas relaciones que permitan introducir parámetros fundamentales en el crecimiento de la polilla, tales como la temperatura. Por ello, aunque la estructura básica es la más interpretable al estar basada en primeros principios, es necesario abrir nuevos niveles de especificación. Para el caso de estudio particular de *Lobesia botrana*, los tres escenarios propuestos corresponden a la apertura de un mismo parámetro a través de diferentes relaciones. En este sentido, se puede inferir, que el índice de interpretabilidad no solo es

Tabla 6.12: Índice de interpretabilidad $IIWS_{L,M_s}$ para cada escenario de modelado *Lobesia botrana*.

	$M_s = 0$	$M_s = 1$	$M_s = 2$	$M_s = 3$
Parámetros interpretables	8	16	20	20
Parámetros no interpretables	0	0	4	2
$IIWS_{L,M_s}$	1	0,59	0,51	0,54
Cambio en el valor del índice		−41 %	−49 %	−46 %
MAE %		2,0	9,83	4,27
E_l, L_l, P_l, A_l		1,19	7,74	7,86
		0,36	10,79	4,73
		0,08	159,6	23,11

útil como condición de parada a la búsqueda de ecuaciones constitutivas, sino que además, es útil para la selección y comparación de diferentes ecuaciones constitutivas dentro de un modelo matemático. De acuerdo con lo reportado en la Tabla 6.12, el escenario 1 en el que se describe la dinámica de crecimiento de la polilla a través de la ecuación de Monod, no solo es la mejor opción en términos de interpretabilidad como ya se mencionó, sino que además, es la estructura de modelo que presenta un mejor ajuste a los datos experimentales disponibles. Es importante señalar que la inclusión de cuatro parámetros empíricos, adimensionales y no interpretables, empeoró la interpretabilidad y ajuste del modelo matemático de manera significativa.

6.2.3.5. Análisis de identificabilidad

Una vez analizada la interpretabilidad del modelo matemático, se analizó la identificabilidad estructural del modelo para cada uno de los escenarios planteados tal como se describió en la Sección 3.1.2. La Tabla 6.13 resume el análisis de identificabilidad realizado para todos los escenarios de modelado. Para el caso del escenario de modelado 2, el parámetro σ_E a pesar de ser un parámetro no interpretable tiene identificabilidad global. El análisis no es concluyente para los demás parámetros del modelo. Lo anterior arroja luces sobre posibles problemas estructurales al

considerar la cinética de crecimiento exponencial. Por otra parte, para el caso del escenario 3, los parámetros no interpretables α y γ resultan ser localmente identificables, mientras que la identificabilidad de los demás parámetros se degrada (en comparación con la cinética de Monod). Es posible comparar los resultados obtenidos con las cinéticas de Hill y de Monod, considerando que la única diferencia entre ambas cinéticas es la aparición de los parámetros de ajuste α y γ . Los resultados obtenidos al comparar la interpretabilidad de los parámetros con su identificabilidad es indeseable, dado que la adición de parámetros no interpretables en dos escenarios, se tradujo en la degradación de la identificabilidad de parámetros con interpretabilidad parcialmente heredada tales como K_{jT} y K_{jH} .

Tabla 6.13: Identificabilidad de los parámetros desconocidos del modelo de crecimiento de *Lobesia botrana*, para todos los escenarios de modelado.

	Vector de Parámetros desconocidos	Análisis de identificabilidad
$M_s = 0$	$\theta_{u,0} = [K_j \quad K_{j_d}]$	Localmente estructuralmente identificables
$M_s = 1$	$\theta_{u,1} = [K_{jT} \quad K_{jH}]$	Localmente estructuralmente identificables
$M_s = 2$	$\theta_{u,2} = [K_{jT} \quad K_{jH} \quad \sigma_E \quad \sigma_L \quad \sigma_P \quad \sigma_A]$	- σ_E estructuralmente globalmente identificable - Demás θ análisis no concluyente
$M_s = 3$	$\theta_{u,3} = [K_{jT} \quad K_{jH} \quad \alpha \quad \gamma]$	- K_{jT} y K_{jH} estructuralmente no identificables - α y γ estructuralmente localmente identificables

6.3 Voltinismo de *Lobesia botrana* : una extensión del modelo de crecimiento

A partir de la construcción de un modelo de crecimiento de *Lobesia botrana* validado en condiciones de laboratorio, se propone una extensión del mismo para predecir el voltinismo de la polilla. Comparado con otros modelos presentados en la literatura, el modelo desarrollado en este tesis doctoral: (i) incorpora factores bióticos (como tasas de mortalidad, tasas de fecundidad y disponibilidad de alimento) y factores abióticos (como temperatura y humedad relativa) que son importantes para describir el desarrollo y la actividad de vuelo de la polilla, (ii) es un modelo basado en primeros principios con un alto número de parámetros interpretables, (iii) es útil para el desarrollo de sistemas de apoyo a la toma de decisiones porque tiene parámetros con los que el usuario final del modelo puede interactuar fácilmente, y (iv) ha demostrado ser representativo del proceso de voltinismo de la polilla de la vid en la región de Cuyo, Argentina, para dos temporadas diferentes (2020 – 2021 y 2021 – 2022), sin necesidad de reparametrizar el modelo.

6.3.1 Concepto de voltinismo en *Lobesia botrana*

Como se mencionó antes, *Lobesia botrana* es una especie multivoltina con diapausa facultativa (estado fisiológico de inactividad) y metamorfosis completa. Su ciclo vital consta de cuatro estadios: huevo, larva (L1 a L5), pupa y adulto, y al ser multivoltina, este ciclo se repite n veces al año. El modelo de crecimiento presentado en la sección anterior, contemplaba el desarrollo de la polilla para una única generación G_1 . Sin embargo, el número de generaciones de polillas depende de la latitud, el fotoperiodo, las variables climáticas (humedad y temperatura) y el tipo de alimento disponible (Altimira et al., 2021; Rossini et al., 2015). En el caso concreto de la región de Cuyo en Argentina, el voltinismo de *Lobesia botrana* se produce entre los meses de septiembre y abril (en mayor medida) y, dependiendo de las variables climáticas, un posible cuarto pico de escape se produce entre los meses de mayo y junio, como se muestra en la Figura 6.16, dónde el eje x representa la progresión estacional (recuadro azul para el hemisferio norte y recuadro gris para el hemisferio sur), y

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

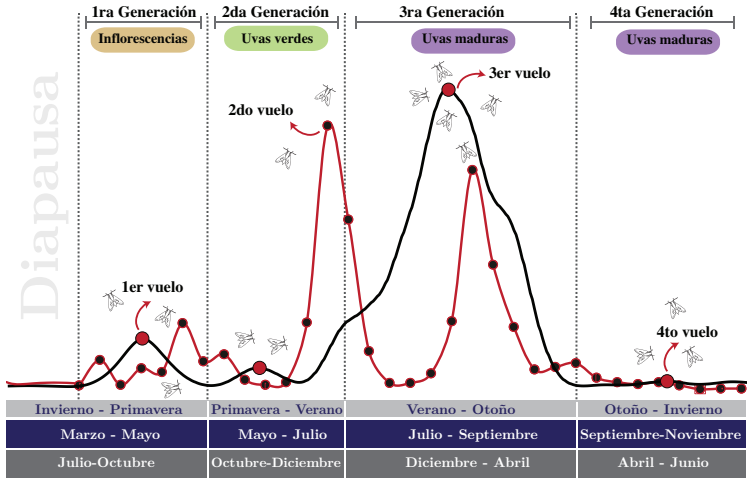


Figura 6.16: Comparación del voltinismo de *Lobesia botrana* en los hemisferios norte y sur. La línea roja muestra el voltinismo de *Lobesia botrana* en España, según (Martín-Vertedor et al., 2010). La línea negra representa el voltinismo de *Lobesia botrana* en la región de Cuyo, Argentina (fuente: propia).

el eje *y* representa el tamaño de la población de polillas (sólo machos). La figura compara las curvas típicas de voltinismo de *Lobesia botrana* en Sudamérica y Europa. Aunque la polilla es la misma especie, puede adaptarse a las diferentes condiciones climáticas de las dos regiones. Por lo tanto, es importante formular modelos matemáticos que puedan aplicarse fácilmente a cada zona para hacer predicciones precisas. La Figura 6.16 muestra en la parte superior el momento fenológico de la vid en el que se produce cada vuelo.

6.3.1.1. Generación 1 en la región de Cuyo

El primer vuelo de la polilla comienza a principios de septiembre, cuando la vid está en botón. En esta generación, las hembras ponen sus huevos en las superficies lisas de la planta, es decir, en las brácteas y botones florales, pero escasamente en el raquis de la inflorescencia, ramitas y hojas. Una

vez que el huevo eclosiona, la larva se protege con una cubierta sedosa y perfora los botones florales para destruirlos. En este punto, la larva se alimenta del primordiuelo y del gineceo. A medida que continúa el desarrollo larval (de L_1 a L_5), la larva ataca a los botones cercanos, de modo que quedan adheridos por hilos de seda secretados por la larva y forman un glomérulo o pupa fuera de la inflorescencia, preferentemente en las hojas. En la región de Cuyo, Argentina, los grados-día acumulados necesarios para que se produzca el primer pico son de aproximadamente $ADD = 150$ (Herrera et al., 2016).

6.3.1.2. Generación 2 en la región de Cuyo

El segundo vuelo de la polilla se produce principalmente en diciembre, cuando la vid está en flor. En esta generación, las hembras ponen sus huevos en bayas verdes en zonas sombreadas o en puntos de contacto entre dos bayas. Una vez que los huevos eclosionan, las larvas entran en las bayas verdes y se alimentan de la pulpa. A medida que las larvas atraviesan sus fases de desarrollo de L_1 a L_5 , perforan otras bayas y pupan en los filamentos sedosos. A menudo se pueden observar excrementos en los hoyos formados por las larvas. En la región de Cuyo, Argentina, los días-grado acumulados necesarios para que aparezca el segundo pico son de aproximadamente $ADD = 780$ (Herrera et al., 2016).

6.3.1.3. Generación 3 en la región de Cuyo

El tercer vuelo de la polilla se produce entre enero y febrero, cuando la vid está madurando. En esta generación, las hembras ponen sus huevos en bayas bien desarrolladas y maduras. El desarrollo de la polilla es el mismo que en las generaciones anteriores, pero las larvas comienzan su proceso de pupación bajo la corteza o en los postes de soporte de la planta en lugar de en las partes verdes de la planta. En esta generación, las larvas forman nidos mientras se alimentan de bayas y expulsan filamentos sedosos para la pupación. A menudo se observan bayas secas, agujeros, filamentos sedosos y excrementos. En esta generación, los daños causados por la polilla favorecen la acción de hospedadores secundarios como *Botrytis* y otros hongos causantes de la podredumbre de la fruta. En otoño, las larvas abandonan las uvas para pupar en la corteza, donde pasan el invierno en diapausa. En la región cuyana de Argentina, los días-grado acumulados

necesarios para que se produzca el tercer pico son de unos $ADD = 2150$ (Herrera et al., 2016).

6.3.1.4. Generación 4 en la región de Cuyo

Las temperaturas primaverales y estivales inusualmente altas pueden acelerar ciertos procesos, dando lugar a la aparición de una cuarta generación. En la región de Cuyo, en Argentina, los grados-día acumulados necesarios para que se produzca este pico son aproximadamente $ADD = 2900$ (Herrera et al., 2016).

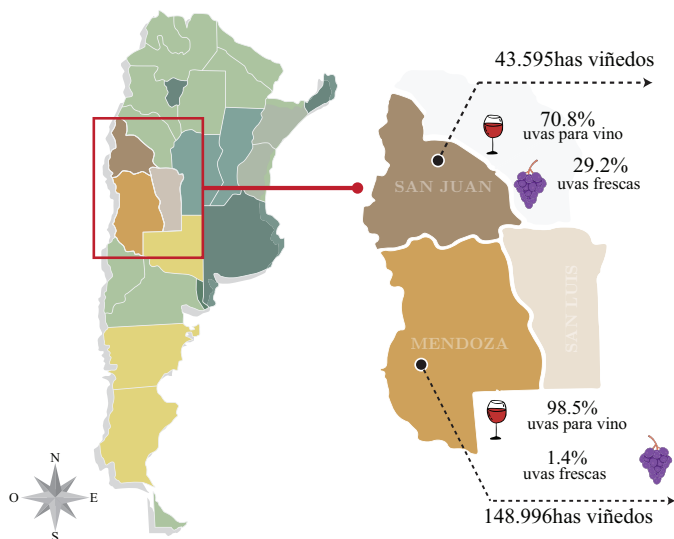


Figura 6.17: Distribución de la producción de uva de mesa y de vinificación en la región de Cuyo, Argentina. Datos porcentuales del Informe Anual de Superficie Cultivada, publicado por Instituto Nacional de Vitivinicultura, 2022.

6.3.2 Área de estudio

La región argentina de Cuyo comprende las provincias de San Juan, Mendoza y San Luis y está situada entre las latitudes 31 y 68 grados, al

pie de los Andes. Las provincias de Mendoza y San Juan son las principales zonas productoras de uva de mesa y de vino, con un 94 % de la producción total del país, seguidas de otras provincias con participaciones entre el 0,05 % y el 0,5. La Figura 6.17 muestra la distribución porcentual de la producción de uva de mesa y de vino en Mendoza y San Juan. En la provincia de Mendoza, considerada una de las grandes capitales del vino, la vitivinicultura se divide en cinco zonas: (1) Zona Norte, (2) Zona Centro, (3) Zona Este, (4) Zona Valle de Uco y (5) Zona Sur.

Cada zona tiene características diferentes, pero todas se centran principalmente en la producción de uvas (Miranda & Novello, 2012). En la provincia de San Juan, el cultivo de la vid se desarrolla en los valles de Ullum, Zonda y Tulum, entre la cordillera de los Andes y la sierra de Pie de Palo, y con menor frecuencia en los valles más altos. Una de las principales características de la provincia de San Juan es su clima cálido con escasas precipitaciones y la presencia de los vientos Zonda, que suelen ser perjudiciales para los cultivos por ser extremadamente cálidos y secos. Si bien la provincia de San Juan reserva parte de la uva cosechada para la producción vitivinícola, la provincia se ha convertido en la principal productora de uva para consumo en fresco del país, representando el 90 % del total nacional (Miranda & Novello, 2012; Sánchez, 2020).

6.3.3 Datos fenológicos de *Lobesia botrana*

Argentina cuenta con una red de monitoreo de *Lobesia botrana* distribuida por todo el país en aquellas regiones que contribuyen en mayor o menor medida a la producción de uva para vinificación o consumo en fresco. En la red de seguimiento se utilizan trampas Delta, que consisten en una trampa seca integrada por una estructura, una percha para fijarla en su lugar, un suelo adhesivo de goma removible y un dispensador de feromonas sintéticas como sustancia atrayente, específicas para *Lobesia botrana*, tal y como se muestra en la Figura 6.18. Su principio se basa en que las polillas atraídas entran en las trampas y se quedan adheridas a la superficie pegajosa del suelo. Según el Manual de procedimientos para el monitoreo de *Lobesia botrana* en Argentina, la feromona utilizada es (7E) (7E,9Z)-7,9-Dodecadien-1-il acetato (Dirección Nacional de Protección Vegetal, 2010). Las trampas Delta se instalan a una altura de 1,30m – 1,50m

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

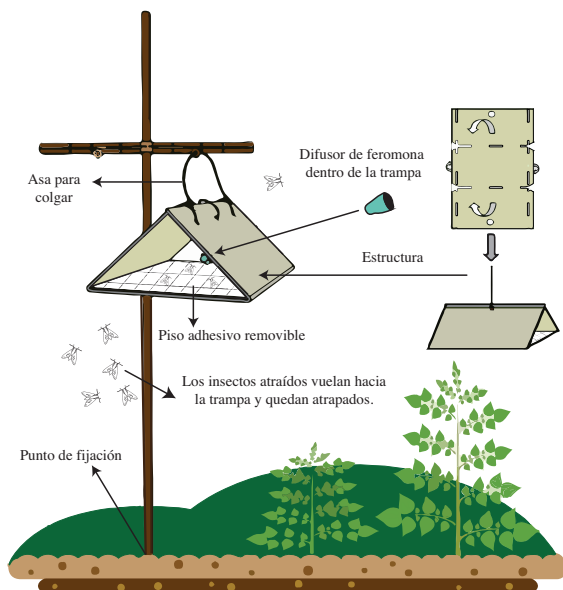


Figura 6.18: Trampas delta con difusor de feromona para la captura de machos de *Lobesia botrana* (Fuente propia).

del suelo., en cultivos con sistema de espaldera próximos a la vid. En los cultivos en espaldera, las trampas se colocan a media altura de las plantas. Los lugares de instalación de las trampas se eligen de acuerdo con el citado manual, teniendo en cuenta aspectos importantes como la accesibilidad, la existencia o no de plantación de vid y la disposición de los propietarios a permitir el acceso al personal encargado del seguimiento de la trampa. Otras consideraciones técnicas importantes son la posición de la trampa respecto a la dirección predominante del viento, la presencia de follaje que impida la difusión de la feromona o la entrada de los insectos a la planta, y la ubicación de la trampa respecto a los puntos con mayor cobertura ya que la plaga prefiere los lugares más sombríos (Dirección Nacional de Protección Vegetal, 2010). La actividad de vuelo de los machos de *Lobesia botrana* fue monitoreada en una red de viñedos en la región de San Juan, Argentina, como se muestra en la Figura 6.19, en ella, las marcas amarillas corresponden a las trampas delta georreferenciadas en

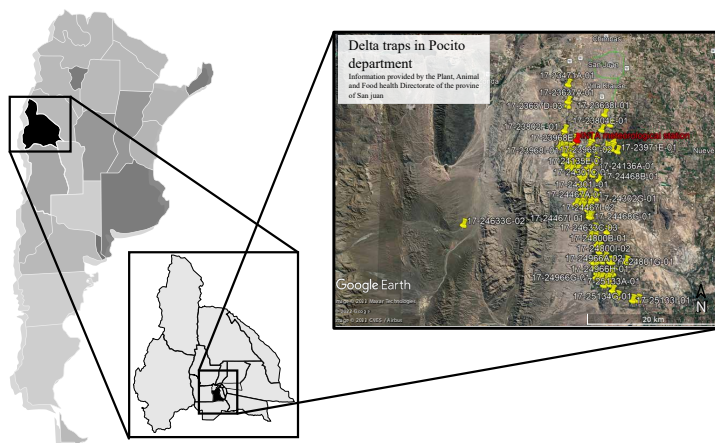


Figura 6.19: Distribución de trampas delta en el departamento Pocito de la provincia de San Juan. Fuente: tramite.sanjuan.gov.ar/index.php/tramite/org/207

un radio de 20km de la estación meteorológica del departamento Pocito. Información suministrada por la Dirección de Sanidad Vegetal, Animal y Alimentos de la provincia de San Juan.

Las trampas marcadas en amarillo corresponden a la red de monitoreo del departamento de Pocito, el cual fue seleccionado debido a que este departamento, además de contar con una extensa red de monitoreo activa, posee una estación meteorológica en las instalaciones del Instituto Nacional de Tecnología Agropecuaria (INTA) de la provincia de San Juan (marcada en rojo en la Figura 6.19). Las trampas de todo el país se identifican con un nombre compuesto por dos dígitos iniciales correspondientes al código de la provincia (17 para la provincia de San Juan), cinco dígitos posteriores correspondientes al número de marco de 1000has (23802), un elemento alfabético correspondiente a un marco de 111has (F) dentro del marco de 1000has y, por último, dos dígitos correspondientes al número de instalación de la trampa en el marco correspondiente (01).

Las trampas instaladas se revisaron semanalmente durante las tempora-

das de verano, otoño y primavera, y cada quince días durante el invierno. Durante la inspección se retiraba y volvía a colocar la banda adhesiva y se contaban las polillas que caían en la trampa para la posterior carga de los datos de captura en la base de datos oficial del SINAVIMO (Dirección Nacional de Protección Vegetal, 2010). En este trabajo se utilizaron los datos de las trampas instaladas en el departamento de Pocito. Este departamento fue seleccionado por la disponibilidad tanto de datos de trampeo como de datos meteorológicos que permitieron validar el modelo matemático propuesto. La georreferenciación de las trampas instaladas se muestra en la Figura 6.19.

6.3.4 Datos meteorológicos

La provincia de San Juan, situada en la región Cuyana de Argentina, es una de las provincias productoras de vino. Es un territorio complejo que presenta desafíos para la actividad productiva debido a las restricciones climáticas de la zona templada y semiárida de la provincia (Uñac et al., 2021). Las diferentes dinámicas geoclimáticas de la región generan una

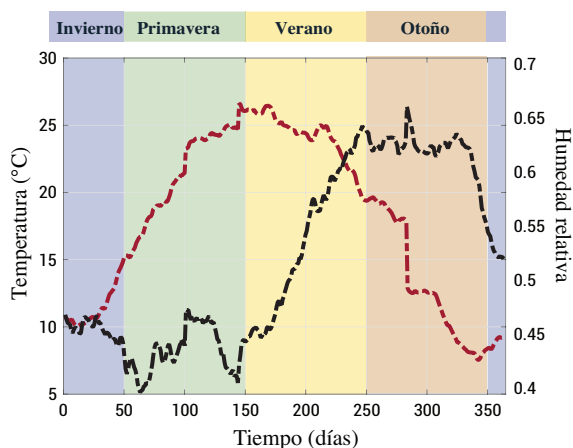


Figura 6.20: Datos meteorológicos de la estación meteorológica del departamento de Pocito, en San Juan - Argentina, correspondientes a la temporada 2020 – 2022.

serie de amenazas de origen geológico, hidrológico y climático, entre las que se destacan los vientos locales y los vientos Zonda (Uñac et al., 2021). El departamento de Pocito posee un clima seco, alcanzando temperaturas máximas de 40°C durante el verano, y mínimas bajo cero en invierno (Uñac et al., 2021). La Estación Experimental Agropecuaria de San Juan (EEA) cuenta con una Estación Meteorológica Automática instalada en las cercanías del departamento de Pocito, que registra datos diarios como temperatura mínima, máxima y media, humedad relativa, precipitaciones y viento, entre otros. La información recolectada por la estación está disponible tanto para datos en tiempo real como históricos. En este trabajo se utilizan datos de la estación entre los años 2020 a 2022.

6.3.5 Extensión del modelo de crecimiento

Para la extensión del modelo de crecimiento a todas las generaciones de una temporada completa (voltinismo) se añade al diagrama de sistemas de proceso, la conexión entre la última etapa y la primera etapa de la siguiente generación, como se muestra en la Figura 6.21. En esta figura,

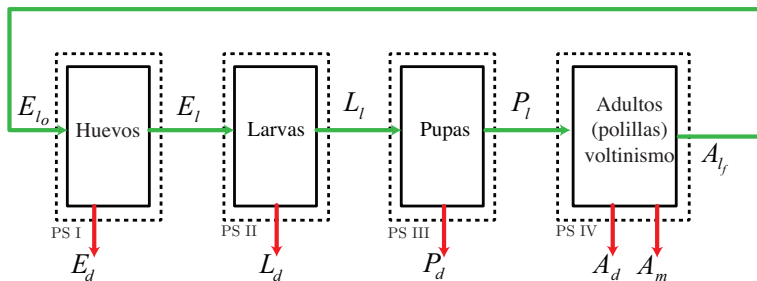


Figura 6.21: Diagrama de bloque de los sistemas de procesos considerados para modelar el proceso de crecimiento y voltinismo de *Lobesia botrana*.

se añade un nuevo flujo (E_{l_o}) para representar el proceso de oviposición por parte de las hembras adultas de la polilla. En la Figura 6.21, las líneas verdes representan las entradas a cada sistema de procesos (número de huevos E_l , larvas L_l , pupas P_l , y adultos vivos A_l) y las líneas rojas representan las salidas de los sistemas de procesos (número de huevos E_d , larvas L_d , pupas P_d , y adultos muertos A_d).

6.3.5.1. Tasa de desarrollo

La tasa de desarrollo de la polilla de la vid se modela desde una perspectiva macroscópica de cada etapa de crecimiento, sin considerar los procesos fisiológicos internos del insecto. Se hipotetizó que existe una analogía entre el proceso de desarrollo de los insectos y el proceso de crecimiento de los microorganismos. Por lo tanto, se determinó que la tasa de desarrollo de *Lobesia botrana* puede ser representada por una ecuación de Monod con tres factores de crecimiento limitantes: (i) grados día acumulados (ADD) de temperatura, (ii) humedad relativa (H) y (iii) el efecto de la edad de la uva en el tiempo de desarrollo de la polilla (Ebs_j). Los valores de Ebs_j fueron ajustados utilizando datos de la literatura, como se muestra en la Tabla 6.14. Para las etapas de huevo y adulto, Ebs_j es igual a cero.

$$K_j(t) = \frac{\mu_{j_{max}}(t)ADD_j(t)}{K_{j_T} + ADD_j(t)} + \frac{\mu_{j_{max}}(t)H(t)}{K_{j_H} + H(t)} + Ebs_j \quad (6.31)$$

Los demás parámetros de la tasa de crecimiento se mantienen tal como fueron descritos en el modelo de crecimiento inicial (Ecuaciones 6.6-6.8). El parámetro K_j describe la tasa de crecimiento de cada etapa en el ciclo

Tabla 6.14: Efecto de la edad de la uva en el tiempo de desarrollo de *Lobesia botrana*. Tomado de (Gutierrez et al., 2018).

Efecto de la edad de la uva en el tiempo de desarrollo		
	Ebs_L	Ebs_P
Inflorescencia	1,25	1,19
Uvas verdes	1,19	1,14
Uvas maduras	1	1

de vida de una polilla, incluyendo huevos, larvas, pupas y adultos. El símbolo j representa cada etapa, mientras que $\mu_{j_{max}}$ corresponde a la tasa máxima de crecimiento para esa etapa, calculada en función de las tasas de mortalidad en cada instante de tiempo, como se muestra en la Ecuación 6.8. La constante de semivelocidad para cada factor limitante de crecimiento se denota como K_{j_T} y K_{j_H} para la temperatura y la humedad relativa, respectivamente. Los grados día para cada etapa se dan por DD_j , mientras que los grados día acumulados se representan por ADD_j y se

describen mediante la ecuación de Touzeau, que es ampliamente utilizada en la literatura (Touzeau, 1981). La temperatura umbral para cada etapa de la polilla se denota como $T_{j_{tl}}$, mientras que T_{max} y T_{min} representan las temperaturas máximas y mínimas diarias, respectivamente.

6.3.5.2. Tasa de mortalidad

En este estudio, se utiliza la función de mortalidad dependiente de la temperatura propuesta por Gilioli et al., 2016 y se incorpora en las ecuaciones de balance presentadas en las ecuaciones 6.1-6.4. Los parámetros de la función fueron ajustados para cada etapa de crecimiento en diferentes estaciones del año. Solo se consideró la tasa de mortalidad intrínseca, que depende de los procesos biológicos propios de la polilla, y no se tuvo en cuenta la mortalidad relacionada con factores de control externos Gilioli et al., 2016. La ecuación 6.32 representa la función de tasa de mortalidad y es una función por partes que toma un valor específico cuando las condiciones de temperatura están dentro del rango de temperatura umbral o un valor de saturación de la tasa de mortalidad (M_{sat}) en caso contrario.

$$K_{jd}(T) = \begin{cases} \alpha_j \left(\frac{T_{av} - T_{m_j}}{T_{m_j}} \right)^2 + \varepsilon & T_{j_{tl}} \leq T_{av} \leq T_{j_{tu}} \\ M_{sat} & \text{en otro caso} \end{cases} \quad (6.32)$$

En esta función, $T_{j_{tl}}$ es la temperatura umbral inferior, $T_{j_{tu}}$ es la temperatura umbral superior, y ambas definen el rango de temperatura óptima para el proceso de crecimiento de la polilla. T_{av} corresponde a la temperatura promedio diaria, y T_{m_j} es la temperatura umbral promedio (ver Ecuación 6.33). Finalmente, α está descrito por una ecuación empírica (ver Ecuación 6.34), y ε es un parámetro empírico, ambos útiles para ajustar el modelo matemático.

$$T_{m_j} = \frac{T_{s_{tl}} + T_{s_{tu}}}{2} \quad (6.33)$$

$$\alpha_j = (M_{sat} - \varepsilon) \left(\frac{T_{j_{tu}}}{T_{m_j}} - 1 \right)^{-2} \quad (6.34)$$

6.3.5.3. Tasa de fecundidad

El modelo de crecimiento descrito en la Sección 6.2, fue desarrollado y validado para una generación de la polilla, en condiciones controladas de laboratorio y en un rango limitado de temperaturas. En esta sección, se

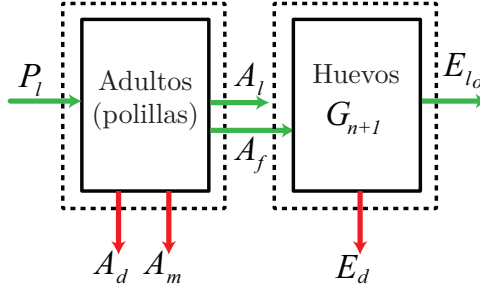


Figura 6.22: Diagrama de bloque de los sistemas de procesos considerados para modelar la función de tasa de fecundidad de *Lobesia botrana*.

considera un rango realista y amplio de temperaturas, así como la interacción entre los adultos de la generación n con la siguiente generación $n + 1$ a través de la función de tasa de fecundidad. La función de tasa de fecundidad representa la forma en que las polillas hembras llevan a cabo el proceso de oviposición. La Ecuación 6.35 corresponde a la función de tasa de fecundidad propuesta en este trabajo. La ecuación se realizó considerando los sistemas de procesos descritos en la Figura 6.22, en ella las líneas rojas representan las pérdidas de huevos y adultos, mientras que las líneas verdes representan el flujo de pupas, adultos y huevos que pasan a la siguiente etapa de crecimiento. A_d y E_d corresponden a huevos y adultos muertos, respectivamente. A_m corresponde a adultos machos ya que no realizan oviposición.

$$E_{l_0}(G) = K_{ov}A_f(G) \quad (6.35)$$

$$A_f(G) = K_{ah}A_l(G) \quad (6.36)$$

$$K_{ov}(t) = \frac{\gamma(t-1)}{\eta^{t-1}} \quad (6.37)$$

Tabla 6.15: Parámetros añadidos a la versión extendida del modelo de crecimiento y voltinismo de *Lobesia botrana*

θ	Descripción	Ecuación constitutiva	Unidades
$T_{E_{t_l}}$	Temperatura umbral inferior huevos	11,6	$^{\circ}C$
$T_{L_{t_l}}$	Temperatura umbral inferior larvas	8,8	$^{\circ}C$
$T_{P_{t_l}}$	Temperatura umbral inferior pupas	10,35	$^{\circ}C$
$T_{A_{t_l}}$	Temperatura umbral inferior adultos	10	$^{\circ}C$
$T_{E_{t_u}}$	Temperatura umbral inferior huevos	34,79	$^{\circ}C$
$T_{L_{t_u}}$	Temperatura umbral inferior larvas	34,27	$^{\circ}C$
$T_{P_{t_u}}$	Temperatura umbral inferior pupas	32,67	$^{\circ}C$
$T_{A_{t_u}}$	Temperatura umbral inferior adultos	32,67	$^{\circ}C$
M_{sat}	Mortalidad de saturación	0,9	$^{\circ}C$
K_{a_h}	Tasa de diferenciación	0,5	Adimensional
ε	Parámetro empírico	0,046	Adimensional
γ	Parámetro empírico	38,5	Adimensional
η	Parámetro empírico	1,11	Adimensional
Ebs_j	Efecto de la edad de la uva en el tiempo de desarrollo	6.14	Adimensional

Aquí, E_{l_0} es el número de huevos (condición inicial para cada generación), G corresponde al número de generación y denota que computacionalmente estas ecuaciones y parámetros se recalculan después de la finalización de una generación. A_l es el número total de adultos vivos, y A_f es el número de hembras vivas de cada generación, es decir, el número de hembras que eventualmente llevarán a cabo el proceso de oviposición. $A_m = A_f$ considerando que la tasa de diferenciación de *Lobesia botrana* es del 50 %. K_{a_h} corresponde a la tasa de diferenciación. K_{ov} es la tasa de oviposición y puede describirse mediante la Ecuación 6.37 presentada en (Gutierrez et al., 2012), donde γ y η son los parámetros empíricos. La Tabla 6.2 muestra la descripción, el valor y la referencia de los parámetros y constantes del modelo de crecimiento, los nuevos parámetros agregados al modelo de voltinismo se reportan en la Tabla 6.15.

6.3.6 Propiedades del modelo

El modelo de voltinismo desarrollado en esta sección, parte de la base del modelo de crecimiento reportado en la Sección 6.2, más la adición de una ecuación de balance que permite modelar la dinámica de voltinismo de la polilla (E_{l_o}).

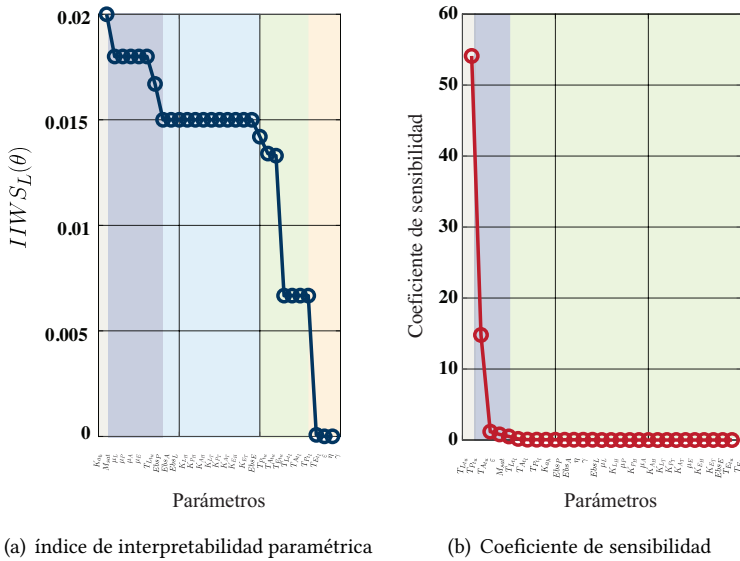


Figura 6.23: Jerarquización paramétrica según el índice de interpretabilidad de los parámetros y el coeficiente de sensibilidad paramétrica, para el modelo de voltinismo de *Lobesia botrana*.

Lo anterior dio lugar no solo a una nueva estructura básica del modelo, sino también a la adición de nuevos parámetros propios del proceso de voltinismo de la polilla. Por esta razón, el modelo de voltinismo no es comparable en términos de sus propiedades con los escenarios de modelado planteados en la Sección 6.2.3.1 y debe tratarse como un nuevo modelo. El análisis de interpretabilidad para el modelo de voltinismo se realizó con el fin de determinar como la adición de las nuevas ecuacio-

nes influye en las propiedades intrínsecas del modelo, más no con fines comparativos respecto a otros escenarios de modelado. La Tabla 6.16, resume los resultados del análisis de sensibilidad y el aporte individual de cada parámetro a la interpretabilidad general del modelo. La Figura 6.23 ilustra la jerarquización paramétrica basada tanto en el índice de interpretabilidad como en el coeficiente de sensibilidad de los parámetros del modelo de voltinismo de *Lobesia botrana* (Figuras 6.23(a) - 6.23(b)).

Como se preveía, los nuevos parámetros empíricos no interpretables ε , η y γ ocupan la última posición en la jerarquía basada en el índice de interpretabilidad más bajo, casi acercándose a cero. Sin embargo, en términos de sensibilidad, los mismos ocupan posiciones con coeficientes de sensibilidad más altos en comparación con otros parámetros importantes del modelo, tales como temperaturas umbrales, influencia de la edad de la uva, y tasas máximas de crecimiento. Este fenómeno, como ya se mencionó, se interpreta como una degradación de la calidad del modelo matemático en términos no solo de interpretabilidad e identificabilidad, sino también en términos estructurales. Sin embargo, los parámetros empíricos representan apenas un 10 % del total de parámetros del modelo, lo que, en comparación con otros modelos reportados en la literatura, permite que el modelo desarrollado tenga mayor capacidad predictiva y mayor utilidad para el desarrollo de herramientas basadas en modelo para el apoyo a la toma de decisiones.

6.3.7 Ajuste y validación

Los resultados del modelo propuesto se analizaron en dos etapas. Durante la etapa de identificación, el modelo matemático se ajustó basándose en los datos de captura en campo de las trampas descritas en la Sección 6.3.3 para la temporada 2020 – 2021. En la segunda etapa, el modelo matemático se validó en el campo utilizando las mismas trampas que en la etapa anterior pero para la temporada 2021 – 2022. A continuación se muestran los resultados cualitativos y cuantitativos obtenidos tras el ajuste y validación del modelo matemático de voltinismo (ver Tabla 6.17).

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

Tabla 6.16: Análisis de interpretabilidad del modelo de voltinismo de *Lobesia botrana*.

θ	Tipo	PI_C	S_L	S_{L_n}	S_c	S_{c_n}	$IIWS_L(\theta)$
μ_E	IPI	0,90	1	0,8	0	0	$1,8 \times 10^{-2}$
μ_L	IPI	0,90	1	0,8	$8,4 \times 10^{-5}$	$1,6 \times 10^{-6}$	$1,8 \times 10^{-2}$
μ_P	IPI	0,90	1	0,8	$1,6 \times 10^{-5}$	3×10^{-7}	$1,8 \times 10^{-2}$
μ_A	IPI	0,90	1	0,8	$9,9 \times 10^{-6}$	$1,8 \times 10^{-7}$	$1,8 \times 10^{-2}$
K_{ET}	PIPI	0,75	1	0,8	0	0	$1,5 \times 10^{-2}$
K_{LT}	PIPI	0,75	1	0,8	$4,6 \times 10^{-6}$	$8,4 \times 10^{-8}$	$1,5 \times 10^{-2}$
K_{PT}	PIPI	0,75	1	0,8	$2,1 \times 10^{-6}$	4×10^{-8}	$1,5 \times 10^{-2}$
K_{AT}	PIPI	0,75	1	0,8	1×10^{-6}	$1,9 \times 10^{-8}$	$1,5 \times 10^{-2}$
K_{EH}	PIPI	0,75	1	0,8	0	0	$1,5 \times 10^{-2}$
K_{LH}	PIPI	0,75	1	0,8	$7,7 \times 10^{-5}$	$1,4 \times 10^{-6}$	$1,5 \times 10^{-2}$
K_{PH}	PIPI	0,75	1	0,8	$1,4 \times 10^{-5}$	$2,6 \times 10^{-7}$	$1,5 \times 10^{-2}$
K_{AH}	PIPI	0,75	1	0,8	$8,9 \times 10^{-6}$	$1,6 \times 10^{-7}$	$1,5 \times 10^{-2}$
T_{Et_l}	GI	1	3	0,3	0	0	$6,7 \times 10^{-3}$
T_{Lt_l}	GI	1	3	0,3	$1,6 \times 10^{-1}$	3×10^{-3}	$6,7 \times 10^{-3}$
T_{Pt_l}	GI	1	3	0,3	$4,7 \times 10^{-2}$	$8,6 \times 10^{-4}$	$6,7 \times 10^{-3}$
T_{At_l}	GI	1	3	0,3	$5,3 \times 10^{-2}$	$9,8 \times 10^{-4}$	$6,7 \times 10^{-3}$
T_{Etu}	GI	1	3	0,3	0	0	$1,3 \times 10^{-2}$
T_{Ltu}	GI	1	3	0,3	$5,4 \times 10^{+1}$	1	$1,7 \times 10^{-2}$
T_{Ptu}	GI	1	3	0,3	$1,5 \times 10^{+1}$	$2,7 \times 10^{-1}$	$1,4 \times 10^{-3}$
$T_{A_{tu}}$	GI	1	3	0,3	$1,3 \times 10^{-3}$	0	$9,1 \times 10^{-3}$
K_{ah}	GI	1	3	0,3	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	2×10^{-2}
M_{sat}	IPI	0,9	1	0,8	5×10^{-1}	$9,3 \times 10^{-3}$	$1,8 \times 10^{-2}$
η	NIP	0	1	0,8	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	$3,3 \times 10^{-6}$
ϵ	NIP	0	1	0,8	$7,9 \times 10^{-1}$	$1,5 \times 10^{-2}$	$7,3 \times 10^{-5}$
γ	NIP	0	1	0,8	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	$3,3 \times 10^{-6}$
Ebs_E	PIPI	0,75	1	0,8	0	0	$1,5 \times 10^{-4}$
Ebs_L	PIPI	0,75	1	0,8	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	$1,5 \times 10^{-4}$
Ebs_P	PIPI	0,75	1	0,8	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	$1,5 \times 10^{-4}$
Ebs_A	PIPI	0,75	1	0,8	$3,6 \times 10^{-2}$	$6,6 \times 10^{-4}$	$1,5 \times 10^{-4}$

6.3.7.1. Primera etapa: identificación de parámetros para la Región de Cuyo

Los parámetros desconocidos del modelo matemático se identificaron mediante el método de mínimos cuadrados no lineales. El vector de parámetros desconocidos se definió como se muestra en la Ecuación 6.38, y

Tabla 6.17: Error absoluto medio del modelo de voltinismo de *Lobesia botrana* en la fase de identificación y validación.

Estadio	MAE	MAE	MAE	MAE
	Número de adultos	%	Max %	Min %
Identificación	93,79	5,9	47,5	0
Validación	215,76	10,38	59,3	0,008

el problema de optimización se definió como se muestra en la Ecuación 6.39. El proceso de identificación se llevó a cabo considerando que los parámetros desconocidos cambian en función de la temperatura, tal y como se explicó en la Sección 6.3.1.

$$\theta_{u_{nk}} = [K_{j_T} \quad K_{j_H}] \quad (6.38)$$

$$\min_{\theta_{u_{nk}}} f(x, \theta_{u_{nk}}) = \sum_{i=1}^m |f_i(x, \theta_{u_{nk}}) - y_i| \quad (6.39)$$

$$lb_{\theta_{u_{nk}}} = [0,001 \quad 0,001 \quad 0,001 \quad 0,001] \quad (6.40)$$

$$ub_{\theta_{u_{nk}}} = [1 \times 10^5 \quad 1 \times 10^5 \quad 1 \times 10^5 \quad 1 \times 10^5] \quad (6.41)$$

donde, $lb_{\theta_{u_{nk}}}$ y $ub_{\theta_{u_{nk}}}$, son los vectores restricción baja y restricción alta, respectivamente, en la identificación del problema. La restricción inferior se fijó considerando el hecho de que los parámetros aparecen en el denominador de la Ecuación de Monod y por lo tanto no pueden ser 0. Luego, la restricción superior se fijó utilizando el conocimiento disponible sobre el proceso. Se realizó un ajuste del modelo utilizando los datos correspondientes a las capturas de las trampas descritas en la Sección 6.3.3 durante la temporada 2020 – 2021, así como los datos meteorológicos correspondientes al mismo periodo. Los datos meteorológicos fueron preprocesados con el fin de filtrar el ruido y el error de medición asociado al sensor utilizado por la estación meteorológica de referencia *Vantage Pro 2*, como se muestra en la Figura 6.24. El valor del error de medida se estimó para cada estación meteorológica del año, una constante, tal y como se muestra en la Tabla 6.18. En la Figura 6.24 la línea negra representa los datos de temperatura y humedad proporcionados por la

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

Tabla 6.18: Error de detección de la estación meteorológica para cada estación.

Estación	Error de medición Δ_T
Invierno	$\pm 6,25^\circ C$
Primavera	$\pm 5^\circ C$
Verano	$\pm 2,5^\circ C$
Otoño	$\pm 2,5^\circ C$

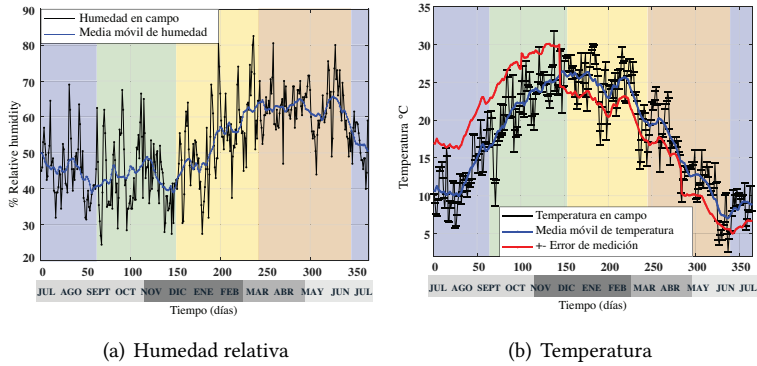


Figura 6.24: Datos meteorológicos de la estación meteorológica de Pocito, en la región argentina de Cuyo, para la temporada 2020 – 2021.

estación meteorológica de la región, la línea azul los datos de temperatura y humedad suavizados con una media móvil, y la línea roja los datos de temperatura suavizados con una media móvil y corregidos con los valores presentados en la Tabla 6.18.

La Tabla 6.19 presenta los valores de los parámetros encontrados después de resolver el problema de mínimos cuadrados no lineales. La Figura 6.25 muestra la comparación entre el modelo matemático ajustado y los datos experimentales para la temporada 2020 – 2021, con todos los valores $\times 10^6$. Allí se muestra la curva de respuesta del modelo de voltinismo

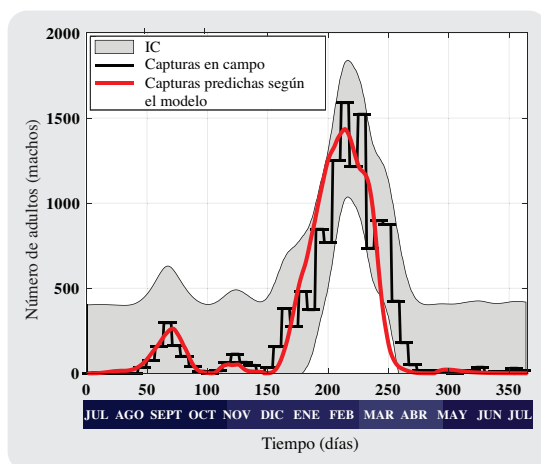


Figura 6.25: Respuesta del modelo de voltinismo de *Lobesia botrana* , en contraste con los datos de captura de machos en campo en la región de Cuyo en Argentina durante la temporada 2020 – 2021.

propuesto en este trabajo, comparada con la curva de datos obtenida del número de *Lobesia botrana* machos capturados en el campo, en Cuyo, Argentina, dónde el área gris es el intervalo de confianza (IC) de 95 %, la curva negra representa los datos discretos de captura de *Lobesia botrana* en el campo, y la curva roja representa la respuesta del modelo en la etapa de identificación. Como puede observarse, en la región de estudio, el modelo matemático propuesto es capaz de describir los picos de vuelo en tiempo y amplitud. Como se observa en la Tabla 6.17, para la etapa de ajuste, el modelo presentó un MAE de 94 adultos, lo que en términos porcentuales equivale a un error de 5,9 %. Esto es evidente en los resultados cualitativos (ver Figura 6.25) ya que la respuesta del modelo en casi todo momento se encuentra dentro del intervalo de confianza de 93 % (sombra gris).

La Figura 6.26 muestra los resultados de la predicción del modelo matemático sobre el proceso de crecimiento de la polilla en los demás estadios de crecimiento. La observación del crecimiento de la polilla en el campo es compleja, por lo que no se dispone de datos experimentales para los estadios de huevo, larva y pupa. En el estadio adulto, la curva negra representa

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

Tabla 6.19: Parámetros y condiciones iniciales por estación del año para el modelo de voltinismo de *Lobesia botrana* particularizado para la región de Cuyo en Argentina.

	E_T	L_T	P_T	A_T	E_H	L_H	P_H	A_H
Invierno	0.6	0.3	0.6	3.1	0.4	0.6	3.2	0.6
Primavera	12.6	1.8	0.006	22	1.18	0.3	5	9.5
Verano	11	5.2	9	19	3.2	0.09	1.8	1.1
Otoño	20.4	16.2	23.6	20.4	8.4	7.7	8.4	9.2

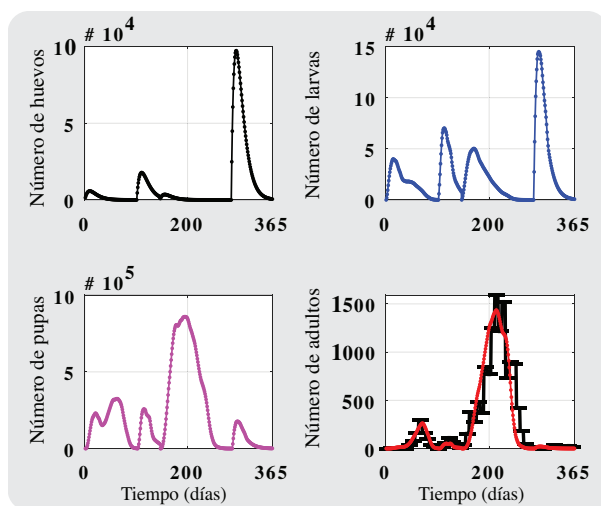


Figura 6.26: Respuesta del modelo de voltinismo de *Lobesia botrana* para todos los estadios de crecimiento de la polilla (huevo, larva, pupa y adulto) en el departamento de Pocito en la región de Cuyo en Argentina durante la temporada 2020 – 2021.

los datos discretos de captura de *Lobesia botrana* en el campo y la curva roja representa la respuesta del modelo en el estadio de identificación.

6.3.7.2. Segunda etapa. identificación de parámetros para la Región de Cuyo

La validación del modelo se realizó utilizando los datos correspondientes a las capturas de las trampas descritas en la Sección 6.3.3 durante la temporada 2021 – 2022, así como los datos meteorológicos correspondientes al mismo periodo como se muestra en la Figura 6.27. Al igual que en la sección de ajuste del modelo, los datos meteorológicos se preprocesaron para eliminar el ruido y los errores de medición asociados al sensor. Los valores utilizados de estimación del error de temperatura de los sensores (Δ_T) son los mismos ya presentados en la Tabla 6.18. La Figura 6.28 mues-

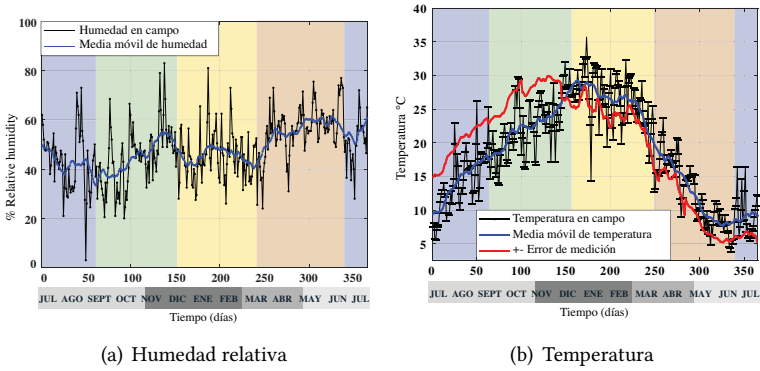


Figura 6.27: Datos meteorológicos de la estación meteorológica de Pocito, en la región argentina de Cuyo, correspondientes a la temporada 2021 – 2022.

tra la comparación entre el modelo matemático ajustado y el conjunto de datos de validación para la temporada 2021 – 2022, dónde la línea negra representa los datos de temperatura y humedad proporcionados por la estación meteorológica de la región, la línea azul los datos de temperatura y humedad suavizados con una media móvil, y la línea roja representa los datos de temperatura suavizados con una media móvil y corregidos con los valores presentados en la Tabla 6.18. En la región de estudio, el modelo matemático propuesto es capaz de describir los picos de vuelo en tiempo y amplitud bajo condiciones de operación diferentes a las utilizadas para

6.3. VOLTINISMO DE *LOBESIA BOTRANA* : UNA EXTENSIÓN DEL MODELO DE CRECIMIENTO

ajustar el modelo matemático en la etapa anterior. La Figura 6.28 muestra

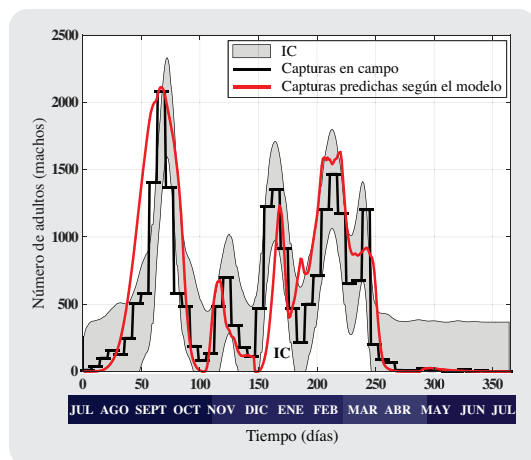


Figura 6.28: Respuesta del modelo de voltinismo de *Lobesia botrana* , en contraste con los datos de captura de machos en campo en la región de Cuyo en Argentina.

los resultados de la predicción del modelo matemático respecto al proceso de crecimiento de la polilla en las demás etapas de crecimiento, dónde el área gris es el intervalo de confianza (IC) de 95 %. La curva negra representa los datos discretos de captura de *Lobesia botrana* en el campo, y la curva roja representa la respuesta del modelo en la etapa de validación. La Figura 6.29 muestra los resultados de la predicción del modelo matemático sobre el proceso de crecimiento de la polilla en los demás estadios de crecimiento. La observación del crecimiento de la polilla en el campo es compleja, por lo que no se dispone de datos experimentales para las fases de huevo, larva y pupa. En el estadio adulto, la curva negra representa los datos discretos de captura de *Lobesia botrana* en el campo, y la curva roja representa la respuesta del modelo en la etapa de validación. Como se muestra en la Tabla 6.17 para la etapa de validación, el modelo presentó un MAE de 216 adultos, lo que en términos porcentuales equivale a un error de 10,4 %. Tal error es esperable si se considera que los parámetros constantes del modelo matemático, tales como constantes de tasa de

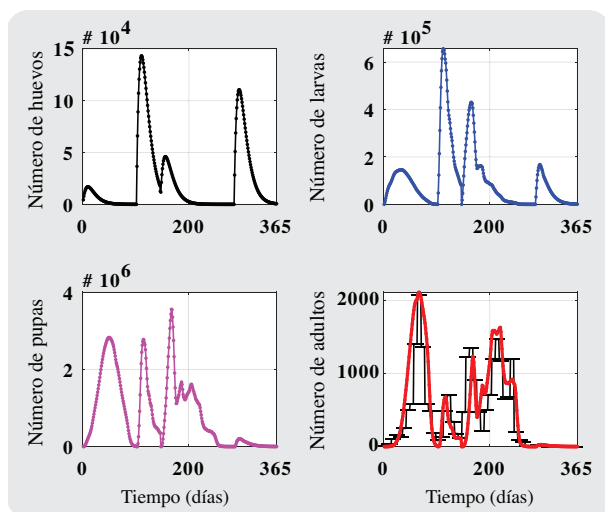


Figura 6.29: Respuesta del modelo de voltinismo de *Lobesia botrana*, en contraste con los datos de captura de machos en campo para el departamento de Pocito en la región de Cuyo en Argentina durante la temporada 2021 – 2022.

crecimiento promedio y temperaturas umbrales, entre otros, no fueron reajustados para mejorar la respuesta del modelo a un nuevo conjunto de entradas (temperatura y humedad) considerando que las condiciones de operación de la temporada 2021 – 2022 fueron diferentes a las de la temporada 2020 – 2021 (por ejemplo, las temperaturas máximas pasaron de 30°C a 35°C de la primera temporada a la segunda, respectivamente). Esto resulta evidente en los resultados cualitativos (véase la Figura 6.28), ya que la respuesta del modelo en casi todos los momentos se encuentra dentro del intervalo de confianza.

En el modelo matemático propuesto hay ocho parámetros estructurales interpretables, que corresponden a dos de los factores bióticos más influyentes en el desarrollo de la polilla (mortalidad y tasas de crecimiento). En la estructura ampliada del modelo propuesto, hay 31 parámetros que tienen significado físico cuando se aplican a una región específica y se

ajustan en función de las condiciones ambientales de la región a la que se aplica el modelo matemático, en lugar de basarse en datos experimentales. Además, cuenta con siete parámetros empíricos que, aunque deben ajustarse mediante algoritmos de optimización, tienen un coeficiente de sensibilidad bajo (es decir, influyen poco en la salida del modelo) y su valor numérico resulta ser representativo del grado de afinidad de la polilla por un determinado factor limitante del crecimiento (temperatura y humedad).

Conclusiones

En este capítulo se presentó el desarrollo de dos modelos matemáticos basados en primeros principios para el proceso de crecimiento y voltinismo de *Lobesia botrana*. Los modelos fueron validados tanto en condiciones de laboratorio (modelo de crecimiento), como en condiciones de campo (modelo de voltinismo). El modelo se desarrolló utilizando la metodología propuesta en el Capítulo 5. Un aspecto novedoso del modelo es que incorpora factores bióticos como la densidad de población, las tasas de mortalidad, las tasas de fecundidad y la influencia de la dieta en el crecimiento de la polilla, así como factores abióticos como las temperaturas máximas y mínimas, los umbrales de crecimiento y la humedad relativa. Este modelo difiere de otros en el número de parámetros empíricos que contiene. Así pues, el modelo propuesto se diferencia de los modelos de caja negra basados en datos en que la mayoría de los parámetros son interpretables, lo que facilita el ajuste y la extrapolación a otras regiones del mundo.

CONCLUSIONES Y TRABAJO FUTURO

7.1	Conclusiones	183
7.1.1	Aporte teórico	183
7.1.2	Aplicación práctica a la región	186
7.2	Trabajo futuro	187

Apunta al infinito. Aunque falles, aterrizarás entre las estrellas.
Les Brown

En esta tesis doctoral se aborda el desafío de la interpretabilidad en la construcción de modelos matemáticos útiles para la comprensión y optimización de sistemas complejos. En ella se propone un nuevo marco conceptual para abordar la interpretabilidad en el espectro de modelado matemático proponiendo un índice matemático para cuantificarla en MSBPP. Además, se amplía la metodología de modelado existente al integrar la interpretabilidad, como un criterio adicional en la determinación del nivel de especificación máximo durante la búsqueda de ecuaciones constitutivas en la estructura extendida del modelo.

7.1 Conclusiones

7.1.1 Aporte teórico

PaEn el Capítulo 4, se abordó una brecha existente en el campo del modelado matemático, que tiene que ver con la falta de un marco conceptual común alrededor del concepto de interpretabilidad de modelos en todo el espectro de modelado y la necesidad de una formalización matemática

que permita cuantificar el grado de interpretabilidad de estos modelos. En este capítulo, se enfatizó en la necesidad crítica de modelos que no solo sean precisos, sino también interpretables, especialmente en campos aplicados como la biología, la medicina y la biotecnología, en las que muchas veces es un requisito. Para ello se introdujo un nuevo marco conceptual para la interpretabilidad en el espectro de modelado, que permite clasificar los parámetros de un modelo matemático en función del tipo de relación presente entre el significante, es decir, las piezas de información de los modelos matemáticos (estructura y parámetros), y el significado, es decir, el contenido explícito de cada pieza de información en un contexto particular.

Por otra parte, para abordar la necesidad de una formalización matemática, se propuso un índice de interpretabilidad ponderado por el nivel de especificación $IIWS_L$ aplicable a MSBPP. El mismo permite evaluar de manera cuantitativa el grado de interpretabilidad de un modelo, así como la mejoría o degradación de su estructura, en términos de interpretabilidad, sensibilidad e identificabilidad, a medida que se describen nuevos niveles de especificación en la estructura extendida del modelo.

En el Capítulo 5 se amplió la metodología de modelado para la deducción de MSBPP, al incluir la interpretabilidad como criterio de parada en la búsqueda de ecuaciones constitutivas o de valoración. Con ello se proporcionó una herramienta práctica para determinar cuando detener la búsqueda de ecuaciones constitutivas para la estructura extendida de un modelo, así como para penalizar la sobreparametrización de los mismos. Con el fin de ilustrar la aplicabilidad del índice, se evaluó la interpretabilidad de dos modelos matemáticos con diferente grado de complejidad, esto es, el modelo de un CSTR y un modelo de regulación de glucosa. Lo anterior permitió extraer algunas conclusiones sobre la interpretabilidad de los modelos matemáticos. En primer lugar, los resultados demuestran que la interpretabilidad del modelo varía significativamente según los distintos escenarios de modelado que se planteen. Esto subraya la importancia de considerar diversas estructuras extendidas del modelo, de manera tal que se pueda mejorar la comprensión del sistema. En segundo lugar, se pudo establecer que el mal condicionamiento matemático de un modelo generalmente resulta en una relación poco deseable entre la interpretabilidad y la sensibilidad de los parámetros.

Un ejemplo concreto tiene que ver con algunos parámetros críticos de diseño en el caso de aplicación del CSTR, en el que parámetros clave esenciales para describir el proceso considerado (tales como parámetros de diseño), que además poseen intepretabilidad general, muestran una sensibilidad relativamente baja. Esto sugiere la necesidad de ajustar el modelo o explorar diferentes niveles de especificación. Así pues, la presencia de parámetros no interpretables plantea retos adicionales en términos de identificabilidad del modelo y definición de límites. Esto complica potencialmente la resolución del problema de identificación de parámetros. Por lo tanto, es necesario evaluar cuidadosamente los cambios en la estructura del modelo ampliado para garantizar que la interpretabilidad no se vea comprometida y que los parámetros añadidos contribuyan significativamente a la comprensión del sistema.

El Capítulo 6, además de proporcionar una solución práctica a un problema de la región de Cuyo (que se describirá más adelante), también constituye un importante aporte teórico en el ámbito del modelado de plagas. Este aporte es fundamental para el desarrollo de herramientas basadas en modelos dirigidas al control y erradicación de plagas. Un aspecto novedoso del modelo propuesto es la incorporación de factores bióticos, como la densidad de población, las tasas de mortalidad, las tasas de fecundidad y la influencia de la dieta en el crecimiento de la polilla, así como factores abióticos, como las temperaturas máximas y mínimas, los umbrales de crecimiento y la humedad relativa.

Aprovechando que la modularidad es una de las principales ventajas de los modelos MSBPP, que permite el uso de ecuaciones constitutivas o de valoración para ampliar la representación de la dinámica del proceso estudiado tras derivar su estructura básica, fue posible acoplar al modelo de esta tesis ecuaciones desarrolladas por otros autores. Por ejemplo, se incorporaron las ecuaciones desarrolladas por Gilioli et al., 2016 para describir las tasas de mortalidad, y las ecuaciones propuestas por Gutierrez et al., 2012 para describir las tasas de deposición de huevos de las hembras.

Este modelo se diferencia de otros modelos de crecimiento y voltinismo publicados en la literatura en: (i) el número de parámetros empíricos que contiene, (ii) el nivel de detalle de las dinámicas modeladas. El modelo

posee ocho parámetros estructurales interpretables, que corresponden a dos de los factores bióticos más influyentes en el desarrollo de la polilla (mortalidad y tasas de crecimiento). En la estructura extendida del modelo propuesto hay 31 parámetros que tienen significado físico cuando se aplican a una región específica, y se ajustan en función de las condiciones ambientales de la región a la que se aplica el modelo matemático, en lugar de basarse en datos experimentales. Además, cuenta con siete parámetros empíricos que, aunque deben ajustarse mediante algoritmos de optimización, tienen un coeficiente de sensibilidad bajo (es decir, influyen poco en la salida del modelo). Para el caso de las constantes de velocidad media de crecimiento, asociadas tanto a la temperatura como la humedad, si bien su interpretabilidad es parcialmente heredada, su valor numérico resulta ser representativo del grado de afinidad de la polilla por el factor limitante del crecimiento (temperatura o humedad, según corresponda). El modelo propuesto difiere de los modelos de caja negra en que la mayoría de los parámetros son interpretables, lo que facilita el ajuste y la extrapolación a otras regiones del mundo.

7.1.2 Aplicación práctica a la región

En el Capítulo 6, se ilustró la aplicabilidad de la metodología propuesta a través del desarrollo y validación de un modelo de crecimiento y voltinismo de *Lobesia botrana* parametrizado para la región de Cuyo en Argentina. La construcción del modelo de crecimiento y voltinismo de *Lobesia botrana*, involucró importantes actores de la región, tales como la Dirección de Sanidad Vegetal, Animal y Alimentos y el Instituto Nacional de Tecnología Agropecuaria (INTA). Es un modelo único en su tipo (MSBPP) para la representación del proceso de crecimiento y voltinismo de la polilla, y el único desarrollado, validado y parametrizado para esta región haciendo uso de los datos disponibles en la red de monitoreo de *Lobesia botrana* de la Nación Argentina.

El desarrollo de los modelos de crecimiento y voltinismo se enmarcó en proyectos institucionales y colaborativos con la región. Esto permitió la ejecución de dos Proyectos de Desarrollo Tecnológico y Social (PDTS) y un Proyecto de Investigación Científica y Tecnológica (PICT), todos en colaboración entre la Universidad Nacional de San Juan y la Dirección de Sanidad Vegetal, Animal y Alimentos de la provincia de San Juan,

Argentina.

Finalmente, los modelos desarrollados fueron reconocidos como la segunda mejor idea innovadora en la convocatoria CREAINnova, organizada por la Asociación Argentina de Consorcios Regionales de Experimentación Agrícola (CREA) y EmpreAR Argentina.

7.2 Trabajo futuro

Con relación al índice de interpretabilidad, se espera que su utilidad pueda extenderse para modelos de caja gris de base empírica, así como para modelos de caja negra. Este paso representaría un acercamiento a la generalización matemática de la interpretabilidad en todo el espectro de modelado. En segundo lugar, a partir de los resultados obtenidos, al comparar el índice de interpretabilidad con la identificabilidad de los parámetros de un modelo, se plantea la posibilidad de relacionar el análisis de identificabilidad con el cálculo de la interpretabilidad de un modelo matemático. Para ello se hace necesario establecer la relación matemática formal existente entre ambas propiedades.

Con relación al caso de aplicación práctica a la región, se encuentra en proceso de desarrollo una herramienta basada en el modelo de crecimiento y voltinismo desarrollado en esta tesis doctoral. Con ella se espera poder brindar soporte a la toma de decisiones por parte del personal y las instituciones gubernamentales involucradas en la contención y erradicación de la plaga. Además, se avanza en la deducción de otros modelos (como el modelo de la vid), que puedan acoplarse al modelo de *Lobesia botrana*, con el fin de generar un conocimiento más profundo sobre la polilla, su afinidad por ciertos varietales, así como la influencia de las características particulares de los mismos en el proceso de crecimiento de la polilla.

CONDICIONES INICIALES

A.1	Condiciones iniciales para el CSTR	189
A.2	Condiciones iniciales para el modelo mínimo de glucosa	189
A.3	Condiciones iniciales para el modelo de crecimiento y voltinismo de <i>L. botrana</i>	191

A.1 Condiciones iniciales para el CSTR

Para el caso de estudio del CSTR se utilizó el paquete *Cvodes* tanto para la solución del modelo computacional como para el cálculo de sensibilidades. Para el cálculo del coeficiente de sensibilidad se utilizó la medida de sensibilidad de la Diferencia Cuadrática Media Relativa (RMSD), se consideraron sensibilidades globales y los tiempos de muestreo experimentales. En la Tabla A.1 se presenta la configuración de condiciones iniciales utilizadas para el análisis de sensibilidad en el software AMIGO. θ_{in} es la condición inicial para cada parámetro y θ_{min} y θ_{max} los valores mínimos y máximos que puede tomar cada parámetro, respectivamente.

A.2 Condiciones iniciales para el modelo mínimo de glucosa

Para el caso de estudio del modelo mínimo de glucosa se utilizó el método numérico *ode15s* para la solución del modelo computacional, y el método *Sensmat* para el cálculo de sensibilidades. Para el cálculo del coeficiente de sensibilidad se utilizó la medida de sensibilidad de la Diferencia Cua-

Tabla A.1: Condiciones iniciales para el cálculo de sensibilidades del modelo del CSTR.

θ	θ_{in}	θ_{min}	θ_{max}	θ	θ_{in}	θ_{min}	θ_{max}
k_1	0,07	1×10^{-3}	1×10^3	k_2	$3,7 \times 10^{-9}$	1×10^{-3}	1×10^3
k_3	$-3,4 \times 10^{-4}$	1×10^{-3}	1×10^3	k_4	$4,9 \times 10^{-5}$	1×10^{-3}	1×10^3
T_{0j}	282,9	254,7	311,3	R	8,3	8,3	8,3
E_a	$7,5 \times 10^4$	$7,2 \times 10^4$	$7,9 \times 10^4$	k_0	$2,8 \times 10^{11}$	$2,7 \times 10^{11}$	3×10^{11}
A_t	8,17	8,17	8,17	V	2,4	2,4	2,4
F_0	0,16	0,08	0,24	ρ	1000	1000	1000
C_p	$3,5 \times 10^3$	$3,5 \times 10^3$	$3,5 \times 10^3$	U	425,9	126	600
ΔH	-9,07	-9,07	-9,07	h_i	8,0e2	7,0e2	0,0e2
h_j	14,5e2	13,5e2	15,5e6	x	0,004	0,0035	0,005
k	16,2	15	16,2	F_j	800	100	1000
F_i	500	100	800	α	0,5	$1e-3$	1e3
β	0,1	$1e-3$	1e3	n	1	0,5	2
m	1	0,5	2				

drática Media Relativa (RMSD), se consideraron sensibilidades globales y los tiempos de muestreo experimentales. En la Tabla A.2 se presenta la configuración de condiciones iniciales utilizadas para el análisis de sensibilidad en el software AMIGO. θ_{in} es la condición inicial para cada parámetro y θ_{min} y θ_{max} los valores mínimos y máximos que puede tomar cada parámetro, respectivamente.

Tabla A.2: Condiciones iniciales para el cálculo de sensibilidades del modelo del modelo mínimo de glucosa.

θ	θ_{in}	θ_{min}	θ_{max}	θ	θ_{in}	θ_{min}	θ_{max}
K_n	0,03	0	0,05	K_{n2}	0,6	0,5	2,5
S_I	6×10^{-4}	1×10^{-4}	12×10^{-4}	I_b	1	0,5	1,5
S_g	0,01	0,005	0,05	HGP	1,20	0	10
p_2	0,02	0,01	0,05	n	0,178	0,07	0,3
f	0,9	0,9	0,9	IR_a	18,8	15	20,8
R_a	4,5	0	5,5	V_I	0,06	0,05	0,07
V_g	1,6	0,7	1,2	Bw	70	56	84
k_{m1}	0,02	0,02	0,02	k_{m2}	0,01	0,01	0,01
k_{md}	0,015	0,015	0,015	k_1	0,0018	0,001	0,01
k_2	0,0182	0,01	0,1	k_d	0,0164	0,01	0,1
HGP_{neo}	2,27	2,27	2,27	$\sigma_{G,2}$	1	1	1
HGP_{gluc}	2×10^{-4}	2×10^{-4}	2×10^{-4}				
$NGlgnII$	1,68e5	$1,7 \times 10^5$	$1,7 \times 10^5$				
r_2	8×10^{-8}	8×10^{-8}	8×10^{-8}				

A.3 Condiciones iniciales para el modelo de crecimiento y voltinismo de *L. botrana*

Para el caso de estudio del modelo mínimo de glucosa se utilizó el paquete *Cvodes* tanto para la solución del modelo matemático como para el cálculo de sensibilidades. Para el cálculo del coeficiente de sensibilidad se utilizó la medida de sensibilidad de la Diferencia Cuadrática Media Relativa (RMSD), se consideraron sensibilidades globales y los tiempos de muestreo experimentales. En la Tabla A.3 se presenta la configuración de condiciones iniciales utilizadas para el análisis de sensibilidad en el software AMIGO. θ_{in} es la condición inicial para cada parámetro y θ_{min} y θ_{max} los valores mínimos y máximos que puede tomar cada parámetro, respectivamente.

Tabla A.3: Condiciones iniciales para el cálculo de sensibilidades del modelo del modelo de crecimiento y voltinismo de *L. botrana*.

θ	θ_{in}	θ_{min}	θ_{max}	θ	θ_{in}	θ_{min}	θ_{max}
K_E	0,05	0	1×10^6	K_L	0,9	0	1×10^6
K_P	0,8	0	1×10^6	K_A	43	0	1×10^6
K_{E_d}	0,05	0	1	K_{L_d}	0,05	0	1
K_{P_d}	0,05	0	1	K_{A_d}	0,05	0	1
$\mu_{E_{max}}$	0,95	0	1	$\mu_{L_{max}}$	0,95	0	1
$\mu_{P_{max}}$	0,95	0	1	$\mu_{A_{max}}$	0,95	0	1
K_{E_T}	$2,1 \times 10^4$	0	1×10^6	K_{L_T}	$4,1 \times 10^2$	0	1×10^6
K_{P_T}	$2,5 \times 10^2$	0	1×10^6	K_{A_T}	$9,9 \times 10^2$	0	1×10^6
K_{E_H}	$6,7 \times 10^4$	0	1×10^6	K_{L_H}	$2,2 \times 10^2$	0	1×10^6
K_{P_H}	$2,3 \times 10^3$	0	1×10^6	K_{A_H}	$1,1 \times 10^3$	0	1×10^6
$T_{E_{th}}$	11,6	0	12	$T_{L_{th}}$	8,8	0	10
$T_{P_{th}}$	10,35	0	11	$T_{A_{th}}$	10	0	11
σ_E	0,54	0	1×10^6	σ_L	3,55	0	1×10^6
σ_P	10,5	0	1×10^6	σ_A	15,25	0	1×10^6
α	1	0	1	γ	0,99	0	1

BIBLIOGRAFÍA

- Abadi, M. (1986). Qué es interpretar. *Revista de Psicoanálisis*, (6), 1223-1236.
- Aguirre-Zapata, E., Garcia-Tirado, J., Morales, H., di Sciascio, F., & Amicarella, A. N. (2022a). Metodología para el modelado y la estimación de parámetros del proceso de crecimiento de *Lobesia botrana* en la región de Cuyo. *Revista Iberoamericana de Automática e Informática industrial*.
- Aguirre-Zapata, E., Morales, H., Dagatti, C. V., Di Sciascio, F., & Amicarella, A. N. (2022b). Semi physical growth model of *Lobesia botrana* under laboratory conditions for Argentina's Cuyo region. *Ecological Modelling*, 464, 109803.
- Ahmad, M. (2020). Deconstructing bond of signifier & signified: a corpus-based study of variation in meaning. *International Journal of Linguistics, Literature and Culture*, 6(4), 76-87.
- Altimira, F., Vitta, N., & Tapia, E. (2021). Integrated pest management of *Lobesia botrana* with microorganism in vineyards: an alternative for clean grapes production.
- Alvarez, H. (2017). Efectos dinámicos en operaciones unitarias. *Modelado de procesos para su análisis y control, Trabajo de año sabático. Universidad Nacional de Colombia*.
- Alvarez, H., Lamanna, R., Vega, P., & Revollar, S. (2009). Methodology for obtaining phenomenological based semi-physical models applied to a sugar cane juice sulphiter (in spanish). *Iberoamerican Journal of Industrial Automation and Informatics RIAI (for its initials in Spanish)*, 6(3), 10-20.
- Alvarez, H. D. (2010). Diseño Simultáneo de Proceso y Control. Uso de efectos dinámicos en Ingeniería de Procesos. *Escuela de Procesos y Energía*.
- Alvarez, H. D. (2022). *Modelado dinámico de procesos químicos y biológicos*. Universidad Nacional de Colombia.

- Amicarelli, A., Di Sciascio, F., Toibero, J., & Alvarez, H. (2010). Including dissolved oxygen dynamics into the Bt δ -endotoxins production process model and its application to process control. *Brazilian Journal of Chemical Engineering*, 27, 41-62.
- Amo-Salas, M., Ortega-López, V., Harman, R., & Alonso-González, A. (2011). A new model for predicting the flight activity of *Lobesia botrana* (Lepidoptera: Tortricidae). *Crop Protection*, 30(12), 1586-1593. <https://doi.org/10.1016/j.cropro.2011.09.003>
- Armendáriz, I., Campillo, G., Pérez Sanz, A., Capilla, C., Juárez, S., & Miranda, L. (2007). The cluster moth *Lobesia botrana* in the Arribes DO, years 2004 to 2006 (in spanish). *Plant-Pest Health Bulletin*, 33, 477-489.
- Asprion, N., Böttcher, R., Pack, R., Stavrou, M.-E., Höller, J., Schwientek, J., & Bortz, M. (2019). Gray-Box Modeling for the Optimization of Chemical Processes. *Chemie Ingenieur Technik*, 91(3), 305-313.
- Azodi, C. B., Tang, J., & Shiu, S. H. (2020). Opening the Black Box: Interpretable Machine Learning for Geneticists. *Trends in Genetics*, 36, 442-455. <https://doi.org/10.1016/j.tig.2020.03.005>
- Bader, F. (1978). Analysis of double-substrate limited growth. *Biotechnology and bioengineering*, 20(2), 183-202.
- Balsa-Canto, E., Alonso, A. A., & Banga, J. R. (2010). An iterative identification procedure for dynamic modeling of biochemical networks. *BMC systems biology*, 4(1), 1-18.
- Balsa-Canto, E., & Banga, J. R. (2011). AMIGO, a toolbox for advanced model identification in systems biology using global optimization. *Bioinformatics*, 27(16), 2311-2313.
- Balsa-Canto, E., Henriques, D., Gábor, A., & Banga, J. R. (2016). AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics*, 32(21), 3357-3359.
- Bäumelt, T., & Dostál, J. (2020). Distributed agent-based building grey-box model identification. *Control Engineering Practice*, 101, 104427.
- Bellu, G., Saccomani, M. P., Audoly, S., & D'Angiò, L. (2007). DAISY: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine*, 88(1), 52-61.

- Bequette, B. W. (1998). Process Dynamics: Modeling, Analysis, and Simulation, 46-49.
- Bequette, B. W. Behavior of a CSTR with a recirculating jacket heat transfer system. En: *En Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)*. 4. IEEE. 2002, 3275-3280.
- Bird, R. B. (2002). Transport phenomena. *Appl. Mech. Rev.*, 55(1), R1-R4.
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., et al. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(2), e1484.
- Bohlin, T. P. (2006). *Practical grey-box process identification: theory and applications*. Springer Science & Business Media.
- Bondy, F., et al. (1983). Heat transfer in agitated vessels.
- Brastein, O. M., Sharma, R., & Skeie, N.-O. (2019). Sensor placement and parameter identifiability in grey-box models of building thermal behaviour.
- Brogan, W. L. (1991). *Modern control theory*. Pearson education india.
- Brun, R., Reichert, P., & Künsch, H. R. (2001). Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37(4), 1015-1030.
- Burgardt, M. I., da Silva, S. M., Ohanian, G., San Blas, G., Núñez McLeod, J., & Burána, M. Preliminary Data on the Study of the “grapevine moth”, *Lobesia botrana* (Lepidoptera. Tortricidae), in Mendoza (in spanish). En: *En Sixth Meeting of Engineering Researchers and Teachers*. Los Reyunos, San Rafael. Mendoza, Argentina, 2011.
- Burgardt, M. I., da Silva, S. M., Ohaniana, G., San Blas, G., McLeoda, J. N., & Burána, M. (2010). Datos Preliminares Sobre el Estudio de la “polilla de la vid”, *Lobesia botrana* (Lepidoptera. Tortricidae), en Mendoza.
- Carlos, C., Gonçalves, F., Oliveira, I., & Torres, L. (2018). Is a biofix necessary for predicting the flight phenology of *Lobesia botrana* in Douro Demarcated Region vineyards? *Crop Protection*, 110, 57-64.

- Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics (Switzerland)*, 8. <https://doi.org/10.3390/electronics8080832>
- Castex, V, Cortázar-atauri, I. G. D., Calanca, P, Beniston, M, & Moreau, J. (2020). Assembling and testing a generic phenological model to predict *Lobesia botrana* voltinism for impact studies. *Ecological Modelling*, 420(December 2019), 108946. <https://doi.org/10.1016/j.ecolmodel.2020.108946>
- Castex, V., Beniston, M., Calanca, P., Fleury, D., & Moreau, J. (2018). Pest management under climate change: The importance of understanding tritrophic relations. *Science of the Total Environment*, 616, 397-407.
- Centro de Estudios Económicos de Bodegas, A. (2020). Informe de Actividad de la industria Segundo Trimestre de 2020.
- Chis, O., Banga, J. R., & Balsa-Canto, E. (2011a). GenSSI: A software toolbox for structural identifiability analysis of biological models. *Bioinformatics*, 27(18), 2610-2611. <https://doi.org/10.1093/bioinformatics/btr431>
- Chis, O.-T., Banga, J. R., & Balsa-Canto, E. (2011b). Structural identifiability of systems biology models: a critical comparison of methods. *PloS one*, 6(11), e27755.
- Corbett, J. P., Colmegna, P., Garcia-tirado, J., & Breton, M. D. (2020). Anticipating meals with behavioral Profiles in an artificial pancreas system - An informed multistage model predictive control approach. *IFAC PapersOnLine*, 53(2), 16305-16310. <https://doi.org/10.1016/j.ifacol.2020.12.652>
- Cousins, C., & Riondato, M. (2019). CADET: interpretable parametric conditional density estimation with decision trees and forests. *Machine Learning*, 108(8), 1613-1634. <https://doi.org/10.1007/s10994-019-05820-3>
- CREA. (2020). Vitivinicultura: panorama del negocio vitivinícola.
- da Silva, S. M., Burgardta, M. I., Ohaniana, G., San Blas, G., Bergasc, V., Burána, M., & Musria, L. Aplicación del modelo de Touzeau como herramienta para el control y manejo de la “polilla de la vid” *Lobesia*

- botrana (Denis & Schiffmüller) en Mendoza. En: *En Séptimo Encuentro de Investigadores y Docentes de Ingeniería EnIDI 2013*. 2013.
- Dagatti, C. V., & Becerra, V. C. (2015). Modeling of the development rate of insects as a function of temperature. Application to integrated pest management using the degree-day method (in spanish). *INTA DIGITAL Institutional Repository - Digital Library*.
- Denis, M, Schiffermüller, I, et al. (1775). Announcement of a systematic work on the Lepidoptera of the Vienna area. *Announcement of a systematic work on the Lepidoptera of the Vienna area*.
- Dirección Nacional de Protección Vegetal, A. (2010, marzo). *Programa nacional de prevención y erradicación de Lobesia botrana: Manual de procedimientos para el monitoreo 2010-2011* (inf. téc.). Dirección Nacional de Protección Vegetal.
- DiStefano III, J. (2015). *Dynamic systems biology modeling and simulation*. Academic Press.
- Dochain, D. (1986). *On-line parameter estimation, adaptive state estimation and adaptive control of fermentation processes* [Tesis doctoral, UCL-Université Catholique de Louvain].
- Dorf, R. C., & Bishop, R. H. (2011). *Modern Control Systems* (12.^a ed.). Pearson.
- Doshi-Velez & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. <http://arxiv.org/abs/1702.08608>
- Duarte-Mermoud, M. A., & Prieto, R. A. (2004). Performance index for quality response of dynamical systems. *ISA transactions*, 43(1), 133-151.
- Erro, D., Alonso, A., Serrano, L., Navas, E., & Hernández, I. Towards physically interpretable parametric voice conversion functions. En: *En International Conference on Nonlinear Speech Processing*. Springer. 2013, 75-82.
- Fan, F. L., Xiong, J., Li, M., & Wang, G. (2021). On Interpretability of Artificial Neural Networks: A Survey. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 5, 741-760. <https://doi.org/10.1109/TRPMS.2021.3066428>
- Garcia-Tirado, J., Zuluaga-Bedoya, C., & Breton, M. D. (2018). Identifiability analysis of three control-oriented models for use in artificial

- pancreas systems. *Journal of diabetes science and technology*, 12(5), 937-952.
- Ghosh, A., & Kandasamy, D. (2020). Interpretable artificial intelligence: Why and when. *American Journal of Roentgenology*, 214, 1137-1138. <https://doi.org/10.2214/AJR.19.22145>
- Gilioli, G., Pasquali, S., & Marchesini, E. (2016). A modelling framework for pest population dynamics and management: An application to the grape berry moth. *Ecological modelling*, 320, 348-357.
- Gonzalez, M. (2010). Lobesia botrana: polilla de la uva. *Revista de Enología*, 7.
- Gutierrez, A. P., Ponti, L., Cooper, M. L., Gilioli, G., Baumg, J., & Duso, C. (2012). Prospective analysis of the invasive potential of the European grapevine moth Lobesia botrana (Den . & Schiff .) in California. *Agricultural and Forest Entomology*, 1, 1-14. <https://doi.org/10.1111/j.1461-9563.2011.00566.x>
- Gutierrez, A. P., Ponti, L., Gilioli, G., & Baumgärtner, J. (2018). Climate warming effects on grape and grapevine moth (Lobesia botrana) in the Palearctic region. *Agricultural and Forest Entomology*, 20(2), 255-271. <https://doi.org/https://doi.org/10.1111/afe.12256>
- Hangos, K., & Cameron, I. (2001). *Process modelling and model analysis. I (Process systems engineering 4.)* Acad Press.
- He, C., Ma, M., & Wang, P. (2020). Extract interpretability-accuracy balanced rules from artificial neural networks: A review. *Neurocomputing*, 387, 346-358. <https://doi.org/10.1016/j.neucom.2020.01.036>
- Heijnen, J., & Romein, B. (1995). Derivation of kinetic equations for growth on single substrates based on general properties of a simple metabolic network. *Biotechnology progress*, 11(6), 712-716.
- Heit, G. E., Sione, W. F., & Aceñolaza, P. G. (2019). Spatio-temporal risk assessment models for *Lobesia botrana* in uncolonized winegrowing areas. *Repositorio Institucional CONICET Digital*, (March 2010). <https://doi.org/10.24425/jppr.2019.129291>
- Herrera, M. E., Dagatti, C. V., & Becerra, V. C. (2016). Método práctico de cría masiva de Lobesia botrana Den. & Schiff.(Lepidoptera:

- Tortricidae) en condiciones de laboratorio. *Revista de la Sociedad Entomológica Argentina*, 75(3-4), 160-164.
- Herrera, M. E., Dagatti, C. V., & Becerra, V. C. (2017). Tabla de vida y parámetros poblacionales de *Lobesia botrana* (Den. et Schiff.) en condiciones de laboratorio. *Revista de la Sociedad Entomológica Argentina*, 76(3-4), 1-10.
- Hong, H., Ovchinnikov, A., Pogudin, G., & Yap, C. (2019). SIAN: Software for structural identifiability analysis of ODE models. *Bioinformatics*, 35, 2873-2874. <https://doi.org/10.1093/bioinformatics/bty1069>
- Hong, S. R., Hullman, J., & Bertini, E. (2020). Human Factors in Model Interpretability: Industry Practices, Challenges, and Needs. *Proceedings of the ACM on Human-Computer Interaction*, 4. <https://doi.org/10.1145/3392878>
- Hotvedt, M., Grimstad, B., & Imsland, L. (2021). Identifiability and physical interpretability of hybrid, gray-box models-a case study. *IFAC-PapersOnLine*, 54(3), 389-394.
- Iltis, C., Moreau, J., Pecharová, K., Thiéry, D., & Louâpre, P. (2020). Reproductive performance of the European grapevine moth *Lobesia botrana* (Tortricidae) is adversely affected by warming scenario. *Journal of Pest Science*. <https://doi.org/10.1007/s10340-020-01201-1>
- Instituto Nacional de Vitivinicultura, A. (2022, abril). *Informe anual de superficie 2021* (inf. téc.). Instituto Nacional de Vitivinicultura. www.inv.gob.ar
- Iooss, B., & Lemaître, P. (2015). Uncertainty Management in Simulation-Optimization of Complex Systems. *Operations Research/ Computer Science Interfaces Series*, 59, 101-122. <https://doi.org/10.1007/978-1-4899-7547-8>
- Johansson, R. (1993). *System Modeling and Identification* (Thomas Kai). Prentice-Hall International, Inc. ill.
- Joubert, D., Stigter, J. D., & Molenaar, J. (2021). Assessing the role of initial conditions in the local structural identifiability of large dynamic models. *Scientific Reports*, 11. <https://doi.org/10.1038/s41598-021-96293-9>

- Karplus, W. J. (1977). The spectrum of mathematical modeling and systems simulation. *ACM SIGSIM Simulation Digest*, 9(1), 32-38.
- Karplus, W. J. (1983). The Spectrum of Mathematical Models. *Perspectives in Computing*, 3(2), 4-13.
- Kern, D. Q. (1950). *Process heat transfer* (Vol. 871). McGraw-Hill New York.
- Kheir, N. (2018). *Systems modeling and computer simulation*. Routledge.
- Cluiever, C. A. (2020). *Dynamic systems: modeling, simulation, and control*. John Wiley & Sons.
- Kopitar, L., Cilar, L., Kocbek, P., & Stiglic, G. Local vs. Global Interpretability of Machine Learning Models in Type 2 Diabetes Mellitus Screening (M. Marcos, J. M. Juarez, R. Lenz, G. J. Nalepa, S. Nowaczyk, M. Peleg, J. Stefanowski & G. Stiglic, Eds.). En: ed. por Marcos, M., Juarez, J. M., Lenz, R., Nalepa, G. J., Nowaczyk, S., Peleg, M., Stefanowski, J., & Stiglic, G. 11979. Springer International Publishing, 2019, 108-119. ISBN: 978-3-030-37445-7. <https://doi.org/10.1007/978-3-030-37446-4>
- Lema-Perez, L. (2018). *Parameters Interpretability in Phenomenological-Based Semiphysical Models . A Human Glucose Homeostasis Model* [Doctoral Thesis]. Universidad Nacional de Colombia.
- Lema-Perez, L., Garcia-Tirado, J., Builes-Montano, C., & Álvarez, H. (2019a). Phenomenological-Based model of human stomach and its role in glucose metabolism. *Journal of theoretical biology*, 460, 88-100. <https://doi.org/10.1016/j.jtbi.2018.10.024>
- Lema-Perez, L., Muñoz-Tamayo, R., Garcia-Tirado, J., & Álvarez, H. (2019b). On parameter interpretability of phenomenological-based semiphysical models in biology. *Informatics in Medicine*, 15(November 2018). <https://doi.org/10.1016/j.imu.2019.02.002>
- Lema-Perez, L., Muñoz-Tamayo, R., Garcia-Tirado, J., & Álvarez, H. (2019c). On parameter interpretability of phenomenological-based semiphysical models in biology. *Informatics in Medicine Unlocked*, 15. <https://doi.org/10.1016/j.imu.2019.02.002>
- Lennart Ljung. (1999). *System Identification Theory for the User* (T. Kailath, Ed.). Prentice Hall information; system sciences series.
- Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J., & Dou, D. (2022). Interpretable deep learning: interpretation, interpretability,

- trustworthiness, and beyond. *Knowledge and Information Systems*, 64, 3197-3234. <https://doi.org/10.1007/s10115-022-01756-8>
- Lipton, Z. C. (2018). The Mythos of Model Interpretability. *Queue*, 16(3), 31-57.
- Ljung, L., & Torkel, G. (1994). *Modeling of dynamic systems* (K. Thomas, Ed.). Prentice Hall information; system sciences series.
- Lopez-Restrepo, S., Garcia-Tirado, J., & Álvarez, H. (2020). A Methodology for Identifying Phenomenological-Based Models using a Parameter Hierarchy. *Canadian Journal of Chemical Engineering*, 98(1), 213-224. <https://doi.org/10.1002/cjce.23500>
- Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 623-631. <https://doi.org/10.1145/2487575.2487579>
- Marco, V. (2001). Modeling of the development rate of insects as a function of temperature. Application to integrated pest management using the degree-day method. *Bulletin of the Aragonese Entomological Society*, 28, 147-150.
- Martin-Vertedor, D., Ferrero-García, J. J., & Torres-Vila, L. M. (2010). Global warming affects phenology and voltinism of *Lobesia botrana* in Spain. *Agricultural and Forest Entomology*, 12(2), 169-176.
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Miranda, O., & Novello, R. (2012, junio). *Programa Nacional Frutales Documento de la Cadena Uva de Mesa* (inf. téc.). INTA Estación Experimental Agropecuaria San Juan.
- Molnar, C. (2020). *Interpretable machine learning*. Lulu.com.
- Molnar, C., Casalicchio, G., & Bischl, B. Interpretable machine learning—a brief history, state-of-the-art and challenges. En: *En Joint European conference on machine learning and knowledge discovery in databases*. Springer. 2020, 417-431.
- Monod, J. (1949). The growth of bacterial cultures. *Annual review of microbiology*, 3(1), 371-394.

- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 22071-22080. <https://doi.org/10.1073/pnas.1900654116>
- Nauck, D. D. (2003). Measuring interpretability in rule-based classification systems. 1, 196-201. <https://doi.org/10.1109/FUZZ.2003.1209361>
- Observatorio Vitivinícola, A. (2021, noviembre). *Datos y estadísticas de la cosecha de uvas en Argentina* (inf. téc.). Observatorio Vitivinícola de Argentina.
- Ogata, K. (2003). *Ingeniería de control moderna*. Pearson Educación.
- Ogden, C. K., & Richards, I. A. (1923). The meaning of meaning: A study of the influence of thought and of the science of symbolism.
- Ortega-Lopez, V., Amo-Salas, M., Ortiz-Barredo, A., & Diez-Navajas, A. (2014). Male flight phenology of the European grapevine moth *Loebesia botrana* (Lepidoptera: Tortricidae) in different wine-growing regions in Spain. *Bulletin of entomological research*, 104(5), 566-575.
- Ortega Quintana, F. A., Álvarez, H., & Botero Castro, H. A. (2017). Enfrentando el modelado de bioprocesos: una revisión de las metodologías de modelado. *Revista ION*, 30(1), 73-90.
- Oviedo, F., Ferres, J. L., Buonassisi, T., & Butler, K. T. (2022). Interpretable and Explainable Machine Learning for Materials Science and Chemistry. *Accounts of Materials Research*, 3, 597-607. <https://doi.org/10.1021/accountsmr.1c00244>
- Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Wortman Vaughan, J. W., & Wallach, H. Manipulating and measuring model interpretability. En: *En Proceedings of the 2021 CHI conference on human factors in computing systems*. 2021, 1-52.
- Rossini, L., Rosselló, N. B., Speranza, S., & Garone, E. (2021). A general ODE-based model to describe the physiological age structure of ectotherms: Description and application to *Drosophila suzukii*. *Ecological Modelling*, 456, 109673.
- Rossini, L., Severini, M., Contarini, M., & Speranza, S. (2020a). EntoSim , a ROOT-based simulator to forecast insects ' life cycle : Description and

- application in the case of *Lobesia botrana*. *Crop Protection*, 129(March 2019), 105024. <https://doi.org/10.1016/j.cropro.2019.105024>
- Rossini, L., Speranza, S., & Contarini, M. (2020b). Distributed Delay Model and Von Foerster's equation: Different points of view to describe insects' life cycles with chronological age and physiological time. *Ecological Informatics*, 59, 101117. <https://doi.org/10.1016/j.ecoinf.2020.101117>
- Rossini, M. N., Dummel, D. M., & Agostini, J. P. (2015). *Plagas cuarentenarias de frutales de la República Argentina* (inf. téc.). Instituto Nacional de Tecnología Agropecuaria. Instituto Nacional de Tecnología Agropecuaria Centro Regional Patagonia Norte Estación Experimental Agropecuaria Alto Valle.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.
- Shuaib, A. O., & Ahmed, M. M. (2014). Robust PID Control System Design Using ITAE Performance Index (DC Motor Model). *International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)*, 3.
- Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons.
- Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(5), 1-13. <https://doi.org/10.1002/widm.1379>
- Sánchez, E. E. (2020). *Programa Nacional Frutales: superficie ocupada por plantaciones frutales en el país y cambios en su estructura productiva* (E. E. A. Balcarce, Ed.; Ediciones INTA, Vol. 1).
- Torres-Vila, L., Stockel, P., Bielza, P., & Lacasa, A. (1996). Effect of diapause and cocoon on the biotic potential of the cluster moth *Lobesia L. botrana* Den. And Schiff (in spanish). *Plant Health-Pest Bulletin*, 22, 27-36.

- Torres Vila, L. M. (1995). *Factores reguladores del potencial biótico y de la poliandria en la polilla del racimo de la vid Lobesia botrana* Den. y Schiff. [Tesis doctoral, Universidad Politécnica de Madrid].
- Touzeau, J. (1981). Modelisation de l'évolution de l'Eudémis de la vigne pour la région Midi-Pyrénées. *Bol Zool Agr Bach II*, 16, 26-8.
- Urretabizkaya, N, Vasicek, A, & Saini, E. (2010). Harmful insects of agroeconomic importance I (in spanish). *Lepidópteros*.
- Uñac, S., Uñac, M., Gattoni, R., Gattoni, G., Aubone, A. F., López, M., Producción, M. D. L., Económico, D., José, A., Cano, D., Ortiz, J. C., Salud, A. M. D., Venerando, A., Turismo, M. D., Claudia, C., Grynszpan, A., Tello, D. R., Flores, J., Gobierno, A. L. D., & Lorenzo, C. (2021, marzo). *Caracterización de la provincia de San Juan: Plan estratégico San Juan* (inf. téc.). Gobierno de San Juan.
- Villaverde, A. F., Barreiro, A., & Papachristodoulou, A. (2016). Structural identifiability of dynamic systems biology models. *PLoS computational biology*, 12(10), e1005153.
- Weiss, J. N. (1997). The Hill equation revisited: uses and misuses. *The FASEB Journal*, 11(11), 835-841.



ISBN 978-631-90150-7-2