

Universidad Central “Marta Abreu” de Las Villas  
Facultad de Ingeniería Eléctrica  
Departamento de Automática y Sistemas Computacionales



## TESIS DE MAESTRÍA

# Modelado de riego de *Solanum lycopersicum* L. en cultivo protegido basado en mapas cognitivos difusos

Tesis presentada en opción al grado de  
Máster en Automática y Sistemas Informáticos

Autor: Ing. Alexandra Madruga Peláez

Tutor: Drs.C. Iván Santana Ching y Yelenys Alvarado Capó  
MSc. Richar Sosa López

Santa Clara

2019

“Año 61 de la Revolución”

Universidad Central “Marta Abreu” de Las Villas  
Facultad de Ingeniería Eléctrica  
Departamento de Automática y Sistemas Computacionales



## TESIS DE MAESTRÍA

# Modelado de riego de *Solanum lycopersicum* L. en cultivo protegido basado en mapas cognitivos difusos

Tesis presentada en opción al grado de  
Máster en Automática y Sistemas Informáticos

Autor: Ing. Alexandra Madruga Peláez  
alexandra@ibp.co.cu

Tutor: Drs.C. Iván Santana Ching y Yelenys Alvarado Capó Prof. Titular  
Dpto. de Automática e Instituto de Biotecnología de las Plantas  
email: ching@uclv.edu.cu y yelenys@ibp.co.cu

MSc. Richar Sosa López Prof. Asistente  
Dpto. de Automática, Facultad de Ing. Eléctrica, UCLV  
richar@uclv.cu

Santa Clara

2019

“Año 61 de la Revolución”



Hago constar que la presente Tesis de Maestría fue realizada en la Universidad Central “Marta Abreu” de Las Villas como parte de la culminación de estudios de la maestra en Automática y Sistemas Informáticas, autorizando a que el mismo sea utilizado por la Institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos, ni publicados sin autorización de la Universidad.

---

Ing. Alexandra Madruga Peláez  
Autor

---

Fecha

Los abajo firmantes certificamos que el presente trabajo ha sido realizado según acuerdo de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.

---

Ing. Alexandra Madruga Peláez  
Autor

---

Fecha

---

, Dr.C  
Jefe del Departamento

---

Fecha

---

Responsable ICT o J' de Carrera, (Dr.C., M.Sc. o Ing.)  
Responsable de Información Científico-Técnica

---

Fecha

## PENSAMIENTO

*“Vale más saber alguna cosa de todo, que saber todo de una sola cosa.”*

*Blaise Pascual*

## DEDICATORIA

A mis padres y mi novio,  
*por el amor, el apoyo y la formación que me han brindado.*

A toda mi familia y amistades,  
*por apoyarme en todo momento.*

A la Revolución.

A todos los profesores que me han formado.

## AGRADECIMIENTOS

A mi tutora Yelenys por su disposición y entrega en todo momento.

A mis tutores Richar y Ching por su apoyo incondicional en esta investigación. Por convertir este tiempo en un continuo aprendizaje como la inteligencia artificial, a ellos, mi eterno agradecimiento.

A Ariel, Novi, Mayda y Mesa que de una forma u otra aportaron su granito de arena en este proyecto.

A mis amigos y compañeros de trabajo, por compartir un momento muy importante de mi vida profesional.

En general quisiera agradecer a todo aquel que ayudó a la realización de esta tesis de maestría.

Santa Clara, Cuba, 2019

## RESUMEN

La automatización en el interior de las casas de cultivo posee una gran importancia, ya que garantiza precisión y fiabilidad. Una supervisión, medición y análisis de los aspectos físicos y fenómenos en dichas casas; mejora la gestión de los cultivos y la toma de decisiones. Con el propósito de lograr una correcta toma de decisiones, se puede realizar un modelado del riego en cultivo protegido para comprender mejor los cambios y condiciones climáticas que se produce dentro de ellas. El objetivo de este trabajo es diseñar un modelo de riego para el cultivo protegido del *Solanum lycopersicum* L., basado en la ingeniería del conocimiento y los mapas cognitivos difusos, que permita una estimación de variables de impacto productivo. Inicialmente se determinaron las variables y se confeccionó un modelo inicial aplicando la ingeniería del conocimiento. Luego, se evaluó dicho modelo, el cual se decidió ajustar mediante el proceso de aprendizaje; utilizando el algoritmo del gradiente descendente. Por último, se valoró el modelo obtenido por los expertos y se realizó un análisis estadístico para la comparación entre el modelo inicial brindado por la ingeniería del conocimiento, y el obtenido por la biblioteca de *software*. Los resultados fueron diez variables involucradas en el riego del cultivo protegido de *S. lycopersicum*, que demostraron la efectividad de la ingeniería del conocimiento. La biblioteca de *software* logró un modelo de riego para el cultivo protegido de *S. lycopersicum* con un error cuadrático medio de 0,95%. Los expertos valoraron el modelo de aceptable; ya que logró la estimación de variables de impacto productivo para ayudar en la toma de decisiones en la producción del *S. lycopersicum*. El análisis estadístico corroboró los excelentes resultados del modelo obtenido, pues presentó una buena correlación y concordancia los valores estimados con respecto a los reales de producción.

## TABLA DE CONTENIDO

	<u>Página</u>
PENSAMIENTO . . . . .	I
DEDICATORIA . . . . .	II
AGRADECIMIENTOS . . . . .	III
RESUMEN . . . . .	IV
ÍNDICE DE TABLAS . . . . .	VIII
ÍNDICE DE FIGURAS . . . . .	X
INTRODUCCIÓN . . . . .	1
1. MARCO TEÓRICO PARA MODELADO DEL RIEGO EN CULTIVO PRO- TEGIDO DE TOMATE . . . . .	7
1.1. Introducción . . . . .	7
1.2. Mapas Cognitivos Difusos . . . . .	7
1.2.1. La función de transferencia . . . . .	10
1.2.2. Algoritmos de aprendizaje . . . . .	11
1.2.3. Herramientas de <i>software</i> . . . . .	14
1.2.4. Aplicaciones de los mapas cognitivos difusos en la agricultura . . . . .	14
1.3. Producción del <i>Solanum lycopersicum</i> L. en casas de cultivo . . . . .	17
1.3.1. Modelos de predicción en tomate . . . . .	19
1.4. Metodologías en el modelado basado en la teoría de los mapas cognitivos difusos . . . . .	22
1.5. Ingeniería del conocimiento . . . . .	22
1.6. Consideraciones finales del capítulo . . . . .	25
2. MATERIALES Y MÉTODOS . . . . .	27
2.1. Introducción . . . . .	27

2.2.	Metodología para el modelado mediante la ingeniería del conocimiento y los mapas cognitivos difusos . . . . .	27
2.3.	Determinación de variables aplicando la ingeniería del conocimiento . . .	29
2.3.1.	Selección de los expertos . . . . .	29
2.3.2.	Elección de la metodología . . . . .	30
2.3.3.	Ejecución de la metodología . . . . .	31
2.3.4.	Procesamiento de la información . . . . .	31
2.4.	Confección del modelo inicial . . . . .	32
2.5.	Recopilación de datos históricos . . . . .	33
2.6.	Evaluación, ajuste y validación del modelo . . . . .	35
2.7.	Análisis del modelo . . . . .	37
2.7.1.	Análisis estadístico . . . . .	37
2.8.	Consideraciones finales del capítulo . . . . .	39
3.	RESULTADOS Y DISCUSIÓN . . . . .	40
3.1.	Introducción . . . . .	40
3.2.	Determinación de variables aplicando la ingeniería del conocimiento . . .	40
3.2.1.	Selección de los expertos . . . . .	40
3.2.2.	Elección de la metodología . . . . .	41
3.2.3.	Procesamiento de la información . . . . .	41
3.3.	Confección del modelo inicial . . . . .	44
3.4.	Evaluación del modelo inicial . . . . .	46
3.5.	Ajuste del modelo . . . . .	48
3.6.	Validación del modelo . . . . .	49
3.7.	Análisis del modelo . . . . .	51
3.7.1.	Análisis estadístico . . . . .	51
3.8.	Consideraciones finales del capítulo . . . . .	54
	CONCLUSIONES . . . . .	56
	RECOMENDACIONES . . . . .	57
	REFERENCIAS BIBLIOGRÁFICAS . . . . .	68
A.	ENCUESTAS PARA EL MODELADO DEL RIEGO EN CULTIVO PROTEGIDO DEL TOMATE . . . . .	69

A.1.	Encuesta de selección de expertos . . . . .	69
A.2.	Encuesta para determinar los principales factores del crecimiento y desarrollo de hortalizas en cultivo protegido . . . . .	69
A.3.	Encuesta para determinar las principales variables involucradas en el riego del tomate en cultivo protegido . . . . .	70
A.4.	Relaciones causales entre las variables involucradas en riego del cultivo protegido del tomate . . . . .	70
B.	ANÁLISIS ESTADÍSTICOS . . . . .	72
B.1.	Histogramas . . . . .	72

## ÍNDICE DE TABLAS

<u>Tabla</u>	<u>Página</u>
1-1. Comparación de las herramientas de <i>software</i> . . . . .	14
1-2. Fases fenológicas del tomate, según Moreno (2007) . . . . .	18
2-1. Metodología para el modelado con MCD . . . . .	27
2-2. Escala tipo Linkert para el grado de influencia de las fuentes de argumentación	29
2-3. Representación de los términos lingüísticos . . . . .	32
2-4. División del conjunto de datos para las fases fenológicas del tomate . . . . .	35
3-1. Prueba de W de Kendall de la primera encuesta . . . . .	41
3-2. Prueba de W de Kendall de la segunda encuesta . . . . .	42
3-3. Rango de trabajo de las entradas y los conceptos para el modelado de riego	44
3-4. Relaciones causales entre las variables . . . . .	45
3-5. Prueba de normalidad de Shapiro Wilk . . . . .	52
3-6. Correlación de Pearson . . . . .	52
3-7. Correlación de Spearman . . . . .	53
A-1. Escala tipo Linkert para el grado de influencia de las fuentes de argumentación	69
A-2. Escala tipo Linkert para el grado de influencia de las fuentes de argumentación	69
A-3. Encuesta determinar las variables en el crecimiento y desarrollo de hortalizas en cultivo protegido . . . . .	70
A-4. Encuesta para determinar las principales variables involucradas en el riego del tomate en cultivo protegido . . . . .	71
A-5. Términos lingüísticos para las relaciones entre variables . . . . .	71

A-6. Relaciones causales entre las variables en el riego del tomate en cultivo protegido . . . . .	71
--	----

## ÍNDICE DE FIGURAS

<u>Figura</u>	<u>Página</u>
1-1. Ejemplo modificado del mapa cognitivo difuso de la investigación de Mour- hir <i>et al.</i> (2017) . . . . .	8
1-2. Método de gradiente descendente . . . . .	13
2-1. Diagrama de flujo de la metodología para modelado MCD . . . . .	28
2-2. Diagrama de flujo del proceso de inferencia . . . . .	36
2-3. Diagrama de flujo del proceso de aprendizaje . . . . .	37
3-1. Coeficiente de competencia de los candidatos . . . . .	40
3-2. Frecuencia de los expertos en la primera encuesta . . . . .	41
3-3. Frecuencia de los expertos en la segunda encuesta . . . . .	43
3-4. Representación del mapa cognitivo difuso para el modelado del riesgo del tomate en cultivo protegido, según la ingeniería del conocimiento . . . . .	46
3-5. Representación de los conceptos, según la ingeniería del conocimiento . . . . .	47
3-6. Representación de los conceptos, según la ingeniería del conocimiento . . . . .	47
3-7. Error cuadrático medio del modelo inicial . . . . .	47
3-8. Representación de los conceptos, resultado del proceso de aprendizaje . . . . .	48
3-9. Representación de los conceptos resultado del proceso de aprendizaje . . . . .	48
3-10. Evolución del error en el proceso de aprendizaje . . . . .	49
3-11. Representación de los conceptos en la etapa de validación . . . . .	49
3-12. Representación de los conceptos en la etapa de validación . . . . .	50
3-13. Error cuadrático medio del modelo final . . . . .	50

3-14.Dispersión de puntos del volumen de riego y la conductividad eléctrica de la solución lixiviada del modelo final . . . . .	53
3-15.Dispersión de puntos de la conductividad eléctrica de la solución nutritiva y el rendimiento acumulado del modelo final . . . . .	54
3-16.Representación del mapa cognitivo difuso para el modelado del riego del tomate en cultivo protegido . . . . .	55
B-1. Representación de los histogramas de los conceptos de los datos históricos .	72
B-2. Representación de los histogramas de los conceptos del modelo obtenido . .	73

## INTRODUCCIÓN

La automatización en la agricultura todavía está en fase de investigación debido a la naturaleza compleja y ambiente realista. A medida que la complejidad y la incertidumbre del medio ambiente aumentan, el modelo matemático del entorno no es fácil de establecer. La información multisensorial combinada con otras tecnologías como el procesamiento de señales, la teoría de control, probabilidad, estadísticas e inteligencia artificial juntas pueden proporcionar una solución técnica para trabajar en un entorno dinámico y desconocido, como la agricultura (Pingzeng and Shusheng, 2011).

Los sistemas de riego automatizados se utilizan con el objetivo de aplicar el agua en el momento y cantidad adecuado, según la evolución de las necesidades del cultivo. Con el uso eficiente de dicho recurso, la energía y los fertilizantes, permitiría reducir los costos de producción (Tajuelo, 2005). Conocer las dinámicas y relaciones causales entre la mayor cantidad posible de los elementos que componen este sistema, es de vital importancia si se pretende alcanzar un determinado resultado. En este caso, las relaciones que existen entre las distintas variables no se pueden obtener mediante la combinación de leyes concretas conocidas, tales como fórmulas matemáticas. Por esta razón, es necesario recurrir a otras técnicas de modelado, como puede ser la inteligencia artificial.

La inteligencia artificial (IA) es un logro poderoso en varios campos, ya que puede complementar al ser humano en diversas actividades, tanto fáciles de realizar, complejas e incluso imposibles de hacer; además de llevar a cabo una actividad sin interrupción por largos períodos de tiempo y donde se requiera conocimiento (Berdel, 1989).

En el modelado de sistemas de riego se han utilizado técnicas de IA como las redes neuronales artificiales (RNA) (Tsang and Jim, 2016), la lógica difusa (Rode *et al.*, 2010) y algoritmos genéticos (González *et al.*, 2016). Muestra de ello, lo demuestran Rode *et al.* (2010) por describir las fases de análisis, diseño e implementación de un invernadero inteligente, a partir de la información acerca de las condiciones ambientales y nutricionales requeridas para el cultivo de tomate, y con el uso de Lógica Difusa. La implementación del control inteligente, el sistema hidropónico y las energías alternativas ofrecen una significativa reducción de costos energéticos y de operación, y la protección al medio ambiente minimizando el uso ineficiente de recursos como el agua y el suelo. También Sadati *et al.* (2014), diseñan un modelo de optimización de programación no lineal con un balance integrado de suelo/agua para determinar un patrón óptimo en los cultivos de trigo (*Triticum*

L.), maíz (*Zea mays* ssp.), remolacha (*Beta vulgaris* L.) y cebada (*Hordeum vulgare* L.). Se concluyó que el enfoque de optimización se aplicó con éxito en la región de la presa de Doroudzan. Así, los responsables de la toma de decisiones y las autoridades pueden usarlo como una herramienta eficaz para problemas complejos de planificación del riego. [Hassan-Esfahani et al. \(2015\)](#) presentaron un enfoque de modelado para la asignación óptima de agua en relación con la maximización de la uniformidad del riego, y la minimización de la reducción del rendimiento. Sobre la base de los resultados de simulación y optimización obtenidos del modelo, el irrigador del área de estudio podría usar hasta un 20 % menos de agua (cantidad ahorrada sobre la cantidad total de agua) durante la temporada de crecimiento, en comparación con los procedimientos operativos tradicionales, sin reducir los beneficios. El estudio de [Tsang and Jim \(2016\)](#) utilizan las RNA y lógica difusa, en la simulación de los cambios en la humedad del suelo para desarrollar una estrategia de riego óptima. La irrigación con inteligencia artificial mantuvo una humedad adecuada del suelo de 0,13 a 0,22  $m^3/m^3$  y redujo el 20 % del uso del agua con una cobertura mejorada de las plantas.

Otras aplicaciones de las técnicas de inteligencia artificial lo constituyó el proyecto de [González et al. \(2016\)](#) los cuales desarrollaron una metodología híbrida que combina RNA y algoritmos genéticos multi-objetivo, para la predicción a corto plazo de la demanda de agua cuando la disponibilidad de datos es escasa. El modelo fue desarrollado utilizando datos de series temporales del Sector VII de la Zona Regable Bembézar en Córdoba, España. El modelo desarrollado fue capaz de explicar el 95 % de la varianza total de los datos observados con un error estándar de predicción del 9,38 % en el período de prueba. En la investigación desarrollada por [López et al. \(2016\)](#), propusieron el modelado de un sistema de control de riego para lograr un uso eficiente de agua en invernaderos de plantas heterogéneas. Los resultados obtenidos, después de los seis días de experimentación, fueron satisfactorios ya que los tiempos de riego y espera permitieron mejorar el consumo de agua con respecto al riego convencional en cada uno de los grupos definidos, sin sacrificar el buen estado y crecimiento favorable de las plantas. Con el objetivo de contribuir al uso racional del recurso hídrico en la Sabana de Bogotá para cultivos de mini-clavel, [Gamboa et al. \(2017\)](#) han generado y simulado un modelo basado en técnicas de inteligencia artificial que predice el valor para el siguiente día de la evapotranspiración (Etc), para un cultivo de mini-clavel variedad Rony.

Dentro de las principales técnicas de IA que se han desarrollado hasta la actualidad, se encuentran los modelos conexionistas ([Granados, 2017](#)), los cuales imitan de un modo bastante básico la estructura biológica del cerebro animal. Dentro de los distintos tipos de modelos existentes se encuentran los mapas cognitivo difusos (MCDs), que, aunque son una representación directa de los sistemas, poseen una estructura neurosináptica, similar a la del cerebro. Dentro de sus ventajas se encuentran que son altamente multi-variables;

y que los elementos que componen el modelo tienen un significado concreto dentro del mundo real, lo cual hace mucho más entendibles e interpretables los resultados (Felix *et al.*, 2017).

Los MCDs se convirtieron en un miembro importante de la familia *Soft Computing* ya que son capaces de resolver numerosos problemas en una variedad de dominios como soporte de decisiones, control de sistemas, pronóstico de series de tiempo, clasificación de patrones y muchos otros (Tettamanzi, 2013). Impulsada por el progreso tangible, la investigación sobre los MCDs continúa atrayendo un gran interés de muchos investigadores de todo el mundo. La revista *Neurocomputing* dedicó un número especial a los avances teóricos del mapeo cognitivo difuso, que comprende nuevas contribuciones a este campo (Froelich, 2014). Sin embargo, la mayoría de esas revisiones son bastante parciales y a menudo se centran en el problema específico abordado en el documento. Sin embargo, la teoría actual de los MCDs es rígida en cuanto a una serie de condiciones y consideraciones que son tomadas a la hora de su desarrollo matemático, las cuales, al ser implementadas en la práctica, ofrecen un conjunto de dificultades. Por otro lado, no se cuenta en la actualidad con bibliotecas de *software* que encapsulen la teoría de los MCDs de modo que pueda ser aplicada con facilidad en el desarrollo de aplicaciones prácticas. Incluso las principales plataformas para el trabajo con minería de datos y aprendizaje automático (ML, por su siglas en inglés) como WEKA (Holmes *et al.*, 1994) y Orange (Demšar *et al.*, 2013), no poseen módulos para el trabajo con esta modalidad de IA.

Los MCDs están contruidos principalmente por el conocimiento de expertos a través de un gráfico (Felix *et al.*, 2017). La ingeniería del conocimiento constituye un método de alto nivel científico que permite el desarrollo de un criterio grupal, a partir de aproximaciones cualitativas procedentes de la experticia y el conocimiento (Stylios and Groumos, 2004). En este sentido, se destaca el método Delphi como técnica de investigación que ha ganado gran popularidad en las investigaciones cubanas de las ciencias agrícolas, donde se emplean el método Delphi para la validación de indicadores, variables, modelos de gestión, procedimientos, propuestas metodológicas y estrategias en el sector agropecuario (González Viera *et al.*, 2018).

En las últimas décadas se han incrementado las investigaciones relacionadas con la teoría de los MCDs en muchas esferas tales como ingeniería, negocios y gestión, medio ambiente, medicina y telecomunicaciones. Ejemplos dentro de la gestión ambiental, se encuentra Ozesmi (2004) donde se examinaron las percepciones entre diferentes partes interesadas en un conflicto ambiental de una presa, para el desarrollo de planes de gestión ambiental y resaneamiento de personas desplazadas. También, Hobbs *et al.* (2002) presentó los pasos involucrados en la construcción de un ecosistema construido para la cuenca hidrográfica del Lago Erie, a través del MCD usando estadísticas multivariadas.

En el caso de la esfera agrícola, Papageorgiou *et al.* (2010) investigaron el rendimiento y sus variaciones en el complejo proceso de gestión del cultivo del algodón (*Gossypium arboreum* L.), donde influyen aspectos como el suelo, el propio cultivo y determinados factores climáticos, todo ello usando los MCDs para modelar y representar el conocimiento de los expertos. Luego, Papageorgiou *et al.* (2011) utilizaron los MCDs para la predicción del rendimiento de ese cultivo, en el cual los algoritmos de aprendizaje automático comparativos propusieron decisiones que coinciden con los datos medidos.

Otros ejemplos utilizando una metodología del MCD lo constituyó la investigación referida por Papageorgiou (2012), donde los MCDs se diseñaron y desarrollaron para representar el conocimiento de los expertos en predicción del rendimiento del cultivo de la manzana (*Malus domestica* B.) utilizando un algoritmo de aprendizaje Hebbian no lineal, que mostró superioridad en la predicción, comparado con la herramienta MCD convencional y los algoritmos de aprendizaje automático de referencia. En la investigación de Halbrecht *et al.* (2014) se comprendió la relación entre las tendencias en el razonamiento de los expertos y los agricultores rurales y comparar las diferencias entre los modelos mentales con las condiciones ambientales locales y los resultados agrícolas, en términos de producción. Christen *et al.* (2015) presentaron la metodología del MCD sobre la afectación a los agricultores y las prácticas agrícolas de la regulación ambiental y la búsqueda de los factores del control de la regulación. También se ha utilizado una metodología mejorada del MCD para el modelado del nivel de producción de coco (*Cocos nucifera* L.) (Jayashree *et al.*, 2015) y la caña de azúcar (*Saccharum officinarum* L.) (Natarajan *et al.*, 2016). En (Mourhir *et al.*, 2017) se propuso una variante del MCD basado en reglas dinámicas para simulaciones en la identificación de los efectos del aumento o disminución de los fertilizantes, donde los resultados revelaron que la disminución a la mitad de tres nutrientes en el fertilizante no reduce el rendimiento en más del 10%.

La automatización en la agricultura es de extrema importancia ya que garantiza precisión y fiabilidad (Reca *et al.*, 2018). En Cuba, el interior de las casas de cultivo presenta un nivel de automatización bajo, de allí, que la mayoría de los procesos agrícolas se realizan gracias a la experiencia de sus trabajadores. Con el monitoreo, medición y análisis de los aspectos físicos y fenómenos, dentro de ellas, se mejorará la gestión de los cultivos y la toma de decisiones. Para lograr todo lo anterior, es necesario el modelado del riego el cual permitirá comprender mejor los procesos que allí se producen.

En la empresa de cultivos varios "Valle del Yabú", principal polo productivo de la provincia de Villa Clara, se encuentra la unidad empresarial de base (UEB) de cultivos protegidos y semiprotegidos. La UEB cuenta con un sistema de casas de cultivo destinadas a la producción de hortalizas como el tomate (*Solanum lycopersicum* L.), con un sistema de riego por goteo y fertirriego automatizado. Sin embargo, este centro aún requiere de la automatización en el interior de las casas de cultivo, y por eso se han venido desarrollando

programas de inversiones para la fabricación y adquisición de equipamientos modernos que integran las nuevas tecnologías de automatización, control y comunicaciones, con el objetivo de desarrollar una agricultura sostenible y productiva (Espinosa, 2017).

**Problema científico:** No se dispone de un modelo de riego en cultivo protegido, que permita ayudar a los expertos en la toma de decisiones de la producción de *S. lycopersicum*.

Una vez realizada la revisión bibliográfica y confeccionado el marco teórico que se describe en el **Capítulo I**, se plantea la siguiente hipótesis:

**Hipótesis:** *Un modelo de riego para el cultivo protegido, basado en la ingeniería del conocimiento y los mapas cognitivos difusos, ayudará a una correcta toma de decisiones por parte de los expertos en la producción de S. lycopersicum.*

Con esta investigación se pretende cumplir los siguientes objetivos:

**Objetivo general:** Diseñar un modelo de riego para el cultivo protegido de *S. lycopersicum*, basado en la ingeniería del conocimiento y los mapas cognitivos difusos, que permita una estimación de variables de impacto productivo para una correcta toma de decisiones.

**Objetivos específicos:**

1. Identificar las principales variables involucradas en el modelado del riego de *S. lycopersicum* en casas de cultivo, mediante la ingeniería del conocimiento.
2. Diseñar el modelo de riego a través de una herramienta de *software* y de la ingeniería del conocimiento, para el cultivo protegido de *S. lycopersicum*.
3. Analizar el modelo de riego obtenido para el cultivo protegido de *S. lycopersicum*, mediante el criterio de los expertos y análisis estadísticos.

El informe está formado por la introducción, el desarrollo organizado en tres capítulos, las conclusiones, seguido de las recomendaciones relativas a investigaciones futuras, y finalmente, las referencias bibliográficas y anexos.

**Capítulo 1:** se describen los principales aspectos de los mapas cognitivos difusos. También, se definen los factores a tener en cuenta para la producción de *S. lycopersicum* en casas de cultivo. Además, se profundiza en las metodologías de diferentes aplicaciones para el modelado utilizando la teoría de los mapas cognitivos difusos. Por último, se explica todo lo relacionado con la ingeniería del conocimiento.

**Capítulo 2:** se presenta la forma de determinar de las variables involucradas en el modelo de riego del cultivo protegido de *S. lycopersicum*. Así como, se describe el procedimiento para la obtención del modelo inicial a través los mapas cognitivos difusos y de la ingeniería del conocimiento. A su vez, se explica el desarrollo de la colecta de los datos históricos para utilizar en el modelado. También, se establece la herramienta de *software* para las etapas de evaluación, ajuste y validación de la metodología. Finalmente, se procede a explicar el proceso de análisis del modelo.

**Capítulo 3:** se determinan las variables involucradas en el modelo de riego del cultivo protegido de *S. lycopersicum*. Luego, se obtiene su modelo inicial a través de los mapas cognitivos difusos y de la ingeniería del conocimiento. Además, se presenta la evaluación el modelo inicial de riego para el cultivo protegido de *S. lycopersicum* utilizando una herramienta de *software* MCD. Para después, ajustar el modelo utilizando el método del gradiente descendente y, por consiguiente, validar el modelo obtenido. Por último, que se produce una valoración del modelo por parte de los expertos; y se realiza el análisis estadístico del mismo con respecto al modelo inicial ofrecido por los expertos.

# CAPÍTULO 1

## MARCO TEÓRICO PARA MODELADO DEL RIEGO EN CULTIVO PROTEGIDO DE TOMATE

### 1.1. Introducción

En este capítulo se incursiona dentro del campo de la inteligencia artificial con los principales aspectos de los mapas cognitivos difusos. Se aborda el marco teórico, enfocado a concebir el planteamiento de la hipótesis de la tesis sobre la base de una actualizada revisión bibliográfica. También, se realiza una descripción de los principales factores y los modelos de predicción encontrados en la literatura científica para la producción de *S. lycopersicum* en casas de cultivo. Además, se profundiza en las metodologías para el modelado utilizando la teoría de los mapas cognitivos difusos de diferentes aplicaciones en la esfera de la agricultura. Por último, se describe la metodología de la ingeniería del conocimiento que será aplicada en el modelado de riego en el cultivo protegido de *S. lycopersicum*.

### 1.2. Mapas Cognitivos Difusos

Una de las variantes más modernas de las RNA lo constituyen los mapas cognitivos difusos, que fueron propuestos por Kosko (1986) como un método de representación del conocimiento basado en un gráfico. Esta forma de representar, describe un conjunto de conceptos en un dominio de interés que están conectados por relaciones de causa y efecto entre ellos. Desde la perspectiva estructural, un MCD es un dígrafo cognitivo que describe el comportamiento de un sistema físico en términos de nodos y aristas que los conectan (figura 1-1). Los conceptos (es decir, los nodos del gráfico) se pueden entender como conjuntos difusos que describen las variables, objetos o entidades del sistema bajo investigación. Las aristas indicadas y equilibradas del gráfico representan las relaciones causales entre los conceptos. Al caracterizar la interacción entre conjuntos difusos en varias iteraciones, los MCDs pueden representar escenarios difusos y complejos (Felix *et al.*, 2017).

La semántica detrás de un MCD estándar se puede definir por 4 tuplas de  $(C, W, A, f)$ , donde  $C = C_1, C_2, \dots, C_M$  es la familia de  $M$  conceptos modelados después de los conjuntos difusos,  $W : C \times C \rightarrow [-1, 1]$  es la matriz que contiene el peso  $w_{ij} \in [-1, 1]$  asignado a cada par de conceptos  $(C_i, C_j)$ . El valor de  $w_{ij}$  determina el signo y la magnitud que conecta la causa del concepto  $C_i$  con el efecto del concepto  $C_j$ . La función  $A : C \rightarrow A_i^t$  calcula el grado de activación  $A_i \in \mathfrak{R}$  de cada concepto  $C_i$  en el paso de tiempo discreto

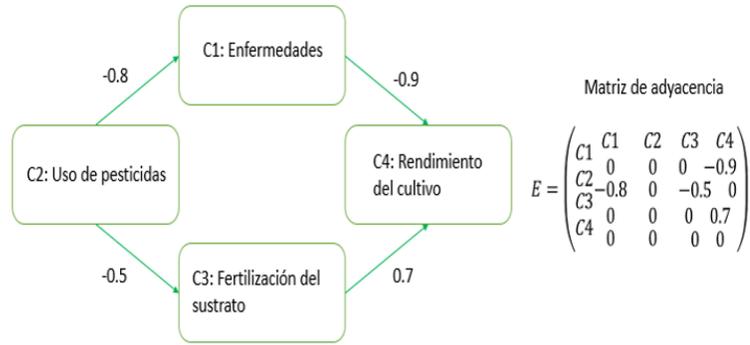


Figura 1–1: Ejemplo modificado del mapa cognitivo difuso de la investigación de Mourhir *et al.* (2017)

$t = 1, 2, \dots, T$ . Finalmente, la función de transferencia  $f : \mathfrak{R} \rightarrow I$  agrega el impacto de eventos causales múltiples sobre el concepto objetivo y fija el resultado al intervalo de activación predefinido. La interpretación del peso causal  $w_{ij}$  entre dos conceptos  $C_i$  y  $C_j$  según Zhi-Qiang (2001) es:

- Si  $w_{ij} > 0$ , entonces un incremento (decremento) en el concepto  $C_i$  producirá un incremento (decremento) en el concepto  $C_j$  con intensidad  $|w_{ij}|$ .
- Si  $w_{ij} < 0$ , entonces un incremento (decremento) en el concepto  $C_i$  producirá una decremento (incremento) en el concepto  $C_j$  con intensidad  $|w_{ij}|$ .
- Si  $w_{ij} = 0$  (o muy cerca de 0), esto denota la ausencia de una relación causal de  $C_i$  sobre  $C_j$ , entonces no hay relación causal en el gráfico.

La selección de la regla de activación requiere una comprensión sólida del sistema físico en cuestión (Felix *et al.*, 2017). La regla de activación de Kosko (1986) para los MCDs se muestra en la ecuación 1.1, siendo  $A^0$  el vector de activación,  $w_{ij}$  el valor de la relación causal que conecta el concepto  $C_j$  con concepto  $C_i$  mientras que  $A_i^t$  denota el valor de activación del concepto  $C_i$  en el  $t$ -ésimo paso de tiempo. Esta regla de activación se repite iterativamente hasta que se cumple una condición de detención. Observe que un MCD produce un vector de estado en cada paso de tiempo discreto que comprende el grado de activación de todos conceptos.

$$A_i^{(t+1)} = f \left( \sum_{j=1, j \neq i}^M w_{ji} A_j^{(t)} + A_i^{(t)} \right) \quad (1.1)$$

Una de las características más relevantes de los MCDs es la interpretabilidad de su topología. De hecho, los MCDs pueden definirse como redes neuronales recurrentes interpretables que incluyen elementos lógicos difusos durante la fase de ingeniería del conocimiento (Felix *et al.*, 2017). Más explícitamente, un MCD explota una vector de activación mediante el uso de una regla similar al modelo *McCulloch-Pitts* estándar (McCulloch, 1988), donde los conceptos pueden ser considerados como entidades de procesamiento neuronal.

Esto implica que el grado de activación de cada neurona del mapa viene dado por el valor de la sumatoria transformada de los pesos recibidos de las neuronas conectadas en la red causal.

Algunos artículos en la literatura científica como (Tsadiras, 1999, 2008; Papakostas *et al.*, 2012; Kreinovich and Stylios, 2015; Nápoles *et al.*, 2016, 2017a) conciben a los modelos basados en los MCDs como redes neuronales artificiales, incluso cuando algunos de los métodos teóricos de los MCDs, por ejemplo los algoritmos de aprendizaje de Hebb, tienen un claro significado neuronal. Esto es debido a sus diferencias, ya que las RNA clásicas se presentan regularmente como cajas negras donde las neuronas y conexiones ocultas no tienen ningún significado claro para el problema en sí mismo (Nápoles *et al.*, 2016). Sin embargo, las neuronas del MCD y sus conexiones tienen una interpretación precisa para el sistema en estudio. Además, los MCDs no involucran neuronas ocultas, ya que tales entidades, no podrían ser interpretadas, ni ayudar a explicar si el resultado es una solución adecuada para un problema determinado. Esto sugiere que la capacidad de representación de los sistemas MCDs es superior al de los modelos basados en las RNA.

El limitado horizonte de predicción del espacio de activación de las neuronas y la ausencia de un algoritmo de aprendizaje intrínseco son limitantes que presentan la red neuronal MCD. Por ello, han surgido algunas variantes para solucionar estos problemas, destacando los modelos Short-Term Cognitive Networks (STCNs, por sus siglas en inglés) y Long-Term Cognitive Networks (LTCNs, por sus siglas en inglés) como se puede apreciar en (Nápoles *et al.*, 2019b) y (Nápoles *et al.*, 2019a) respectivamente. Sin embargo, ninguna de las variantes anteriores presentan nodos que constituyan puras entradas, debido a que son sistemas cerrados donde todas sus neuronas son inferidas en el estado siguiente, lo cual no posee significado físico para una entrada.

En la investigación desarrollada por (Sosa *et al.*, 2019) fueron solucionados estos problemas, partiendo del enfoque LTCN, al incluir un nuevo tipo de neurona llamada entrada, lo cual introduce una nueva variante de MCD llamada Long-Term Cognitive Networks with Inputs (LTCNIs, por sus siglas en inglés). Por tanto, considerando el enfoque LTCNIs, la ecuación de activación quedó representada por la ecuación 1.2.

$$A^{(k+1)} = f(WA^{(k)} + HU^{(k)}) \quad (1.2)$$

Donde  $U^k$  constituye el vector que contendrá los valores de activación de las entradas en el instante  $k$  y que no pueden ser inferidos en un instante  $k + 1$  y,  $H$  una matriz dada por la ecuación (1.5), no necesariamente cuadrada, que almacenará los pesos causales entre las entradas y conceptos similar a como lo hace  $W$  (ecuación 1.4).

$$U^k = [u_1^k, u_2^k, \dots, u_M^k] \quad (1.3)$$

Siendo:  $M$  la cantidad de entradas del mapa

$$W^k = \begin{pmatrix} w_{11}^k & w_{12}^k & \cdots & w_{1N}^k \\ w_{21}^k & w_{22}^k & \cdots & w_{2N}^k \\ \vdots & \vdots & \cdots & \vdots \\ w_{N1}^k & w_{N2}^k & \cdots & w_{NN}^k \end{pmatrix} \quad (1.4)$$

Donde  $w_{ij}$  es el efecto causal que tiene la activación del concepto causa  $j$  sobre el concepto efecto  $i$ . Se toma esta consideración para establecer un acuerdo con la ecuación 1.5.

$$H^k = \begin{pmatrix} h_{11}^k & h_{12}^k & \cdots & h_{1M}^k \\ h_{21}^k & h_{22}^k & \cdots & h_{2M}^k \\ \vdots & \vdots & \cdots & \vdots \\ h_{N1}^k & h_{N2}^k & \cdots & h_{NM}^k \end{pmatrix} \quad (1.5)$$

Donde  $h_{ij}$  es el efecto causal que tiene la activación de la entrada  $j$  sobre el concepto  $i$ . Se toma esta consideración para poder realizar las operaciones entre las matrices que componen a la ecuación 1.2.

### 1.2.1. La función de transferencia

La selección de la función de transferencia está frecuentemente condicionada por los requisitos del sistema, es decir, por el papel que juega cada concepto al modelar el sistema bajo consideración (Bueno, 2009). La función  $f : R \rightarrow I$  denota una forma monótona, no decreciente para fijar el valor de activación de cada concepto al intervalo  $I$  deseado, donde  $I = [0, 1]$  o  $I = [-1, 1]$  según el dominio del problema. De acuerdo con la cardinalidad del espacio de estado, las funciones de la transferencia pueden ser discretas o continuas.

Las funciones de transferencia más utilizadas son la tangente bivalente, trivalente e hiperbólica; y la función sigmoidea. En las funciones discretas bivalente y trivalente, la primera solo produce respuestas binarias que conducen a un número finito de estados y la segunda produce un número finito de diferentes estados (Tsadiras, 2008). La función hiperbólica, es una función de transferencia continua que produce estados infinitos que se pueden distribuir dentro del hipercubo  $[-1, 1]^M$ . La función sigmoidea, es una función continua que produce un infinito número de estados diferentes que se distribuyen libremente dentro del hipercubo  $[0, 1]^M$  (Bueno, 2009).

Para este trabajo la función de transferencia que ha sido utilizada, se muestra en la ecuación 1.6, que no es más, que la ecuación de saturación acotada en los valores mínimo y máximo de cada variable. Su elección se encontró justificada por ser una función sencilla que puede funcionar más fácilmente, ante un modelado de ambiente complejo con una

numerosa cantidad de variables. Además, de que se facilitó su uso con la ayuda de los expertos que ofrecieron los rangos admisibles y de trabajo de cada variable.

$$f(x) = \begin{pmatrix} \min, & x < \min \\ x, & \min \leq x \leq \max \\ \max, & x > \max \end{pmatrix} \quad (1.6)$$

### 1.2.2. Algoritmos de aprendizaje

La fase de aprendizaje es la esencia de la flexibilidad de los modelos conexionistas, y cómo el conocimiento se almacena en los pesos de las conexiones sinápticas; pues el proceso de aprendizaje justamente se basa en el ajuste de esos pesos hasta conseguir que la red se comporte del modo deseado ante los estímulos del medio (Nápoles Ruiz, 2014; Palacios Corrella, 2012). El principal objetivo detrás del aprendizaje del MCD es derivar la matriz de peso basada en la intervención de los expertos, datos históricos disponibles o ambos. La mayoría de los enfoques de aprendizaje existentes suponen que el conjunto de etiquetas conceptuales se proporciona a priori por el experto (Papageorgiou, 2012); y solo la matriz de peso se aprende de datos o con la ayuda de su criterio. Por ello, es la capacidad de evolucionar mediante su entrenamiento o aprendizaje de las RNA; el cual puede ser de dos tipos, supervisado o no supervisado.

Los métodos no supervisados se basan en la idea de que el agente adquiere el conocimiento constantemente durante su interacción con el medio ambiente. La vía más usada ha sido la regla de Hebb (Benítez *et al.*, 2014), la cual tiene una base netamente biológica. En el año 1949, el neurofisiólogo Donald Hebb descubrió que, si dos neuronas se encontraban excitadas simultáneamente, la fuerza de la conexión sináptica entre estas aumentaba, y en caso contrario, la misma disminuía. De este modo la red aprende paulatinamente de la experiencia cotidiana. El aprendizaje no es supervisado cuando la neuro-red avanza dependiendo de los datos de entrenamiento que haya recibido y; finalmente arroja un resultado que puede ser satisfactorio o no (Serrano and Gallizo, 1996). Es decir, se presentan a la red multitud de patrones, pero sin anexar la respuesta deseada; y es entonces que el sistema, a partir de una regla de aprendizaje, debe reconocer regularidades, extraer rasgos, estimar densidades de probabilidad o agruparlos según su similitud.

Por otro lado, los métodos supervisados utilizan, básicamente, un conjunto de datos con la información del comportamiento deseado (Tang *et al.*, 2015). De este modo, la red ajusta sus pesos sinápticos en un proceso iterativo, con el fin de acercarse lo más posible al comportamiento que produce esos valores. Este proceso se desarrolla en una fase de entrenamiento previa a la utilización práctica del programa. En la fase de entrenamiento, un agente externo o supervisor es quien comprueba que dicha salida de la red sea la deseada a través de una comparación y, si existen diferencias, ajusta iterativamente los pesos hasta

conseguirlo, utilizando para ello información detallada del error cometido en cada paso (del Brío and Cinca, 1995; Gamboa *et al.*, 2017). Una gran ventaja es, que sin necesidad de proponer un punto de partida, este modelo es capaz de identificar y estimar relaciones suficientemente correctas. Normalmente se utiliza un porcentaje de los datos en la etapa de aprendizaje y el resto se destina a la validación del resultado final. Según (Sosa *et al.*, 2019), se utiliza el 70 % de los datos recopilados, en la etapa de aprendizaje y los restantes 30 % de los datos se destina para la validación del modelo ajustado.

El tipo de aprendizaje más común es el supervisado y, aunque suele ser más complejo a la hora de su desarrollo computacional es el que ofrece resultados más exactos. No obstante, ambos tipos de aprendizaje destacan por sus propias aplicaciones (Nápoles *et al.*, 2014; Tang *et al.*, 2015).

Se ha propuesto una cantidad diversa de algoritmos de aprendizaje en la literatura científica; ellos se basan principalmente en principios provenientes del campo de las redes neuronales artificiales. Los algoritmos más importantes para el aprendizaje de MCD se pueden clasificar en tres tipos en base a su paradigma de aprendizaje subyacente: aprendizaje basado en Hebbian, basado en errores y algoritmos híbridos. Además, existen otros esquemas alternativos de aprendizaje, que no caen dentro de los tres grupos antes mencionados, que combinan otras funciones matemáticas y teorías computacionales para construir un MCD (Felix *et al.*, 2017).

Los métodos de aprendizaje basados en errores apuntan a generar matrices de peso que minimicen una función de error basado en la diferencia entre las respuestas esperadas y las salidas inferidas del mapa. Estos algoritmos son técnicas de optimización más costosas dado que intentan adaptar un modelo para un conjunto de observaciones históricas. Por otro lado, requieren la definición de la función objetivo que se optimizará, que es el núcleo del procedimiento de aprendizaje (Felix *et al.*, 2017).

El método del gradiente descendente (Flores Villarreal, 2005) es uno de los más sencillos y, por tanto, utilizados en aprendizaje de modelos de IA. Se basa en encontrar un juego de parámetros del modelo que optimicen una función objetivo. En el caso de estudio de esta investigación, los parámetros que componen el modelo son los pesos sinápticos entre conceptos y, la función objetivo es el error cuadrático medio entre la salida del modelo y los datos experimentales. El vector  $w$  de la ecuación 1.7 contiene estos parámetros:

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_P \end{pmatrix} \quad (1.7)$$

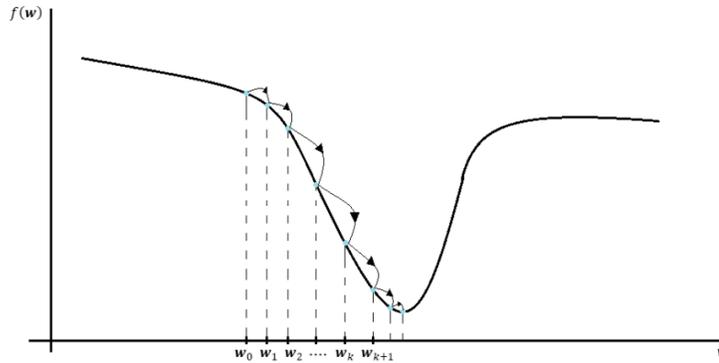


Figura 1-2: Método de gradiente descendente

Siendo:  $P$  el número total de parámetros

Suponga que se tiene una función  $f(w)$  de la cual se quiere obtener el valor de  $w$  que minimiza su valor. Primero se definirá el gradiente de  $f(w)$  como el vector de las derivadas parciales de  $f$  respecto a cada uno de los parámetros de la ecuación 1.8

$$\nabla f(w) = \begin{pmatrix} \frac{\delta f}{\delta w_1} \\ \frac{\delta f}{\delta w_2} \\ \vdots \\ \frac{\delta f}{\delta w_N} \end{pmatrix} \quad (1.8)$$

Teniendo en cuenta, se puede definir el algoritmo del gradiente descendente por pasos de la siguiente forma:

- Paso 1: Se elige un vector  $w_0$  (este valor inicial puede responder a algún criterio en específico o puede ser elegido al azar). Se define  $k$  como el número de la iteración del proceso, siendo inicialmente  $k = 0$ .
- Paso 2: Se calcula el gradiente para el vector a partir de la siguiente ecuación:

$$w_{k+1} = w_k - \alpha \nabla f w_k \quad (1.9)$$

- Paso 3: Se calcula a través de la ecuación del paso 2 para  $k = k + 1$ . Así, iterativamente hasta que se alcance un número máximo de iteraciones, previamente especificado.

Supóngase el caso más sencillo posible donde  $w$  posee un solo componente para observar gráficamente en la figura 1-2 como ocurre el proceso. Este método solo hace uso del gradiente de la función objetivo, por lo que es considerado un método de primer orden. Existen métodos de orden superior que utilizan, además, la matriz Hessiana.

### 1.2.3. Herramientas de *software*

En los trabajos revisados en la literatura científica relacionados con el desarrollo de estas herramientas para un MCD, suelen tener presentes métodos teóricos o aplicaciones prácticas, pero raramente son respaldados por implementaciones de *software*. Dentro de estas bibliotecas (tabla 1–1) se encontraron: FCM Modeler (Mohr, 1997), FCM Designer (Aguilar and Contreras, 2010), *Tool* (León *et al.*, 2011), Mental Modeler Gray *et al.* (2013), JFCM (De Franciscis, 2014), ISEMK (Poczqueta *et al.*, 2015; Papageorgiou *et al.*, 2016), FCM Expert (Nápoles *et al.*, 2017b). Las implementaciones FCM Designer, Mental Modeler y FCM Tool, JFCM carecen de experiencia y no permite el manejo de problemas de aprendizaje automático, lo que obstaculiza su utilización frente a situaciones del mundo real. Por el contrario, las herramientas FCM Expert e ISEMK son más convenientes para el desarrollo de sistemas dedicado a escenarios de simulación, clasificación de patrones y previsión de series de tiempo. A pesar de esto, dichos *software* no logran traducir los avances teóricos que se han realizado en el campo del MCD en implementaciones prácticas.

Tabla 1–1: Comparación de las herramientas de *software*

Herramientas	Año	Opciones de simulaciones	Algoritmo de aprendizaje	Gráficas
FCM Modeler	1997	ninguna	solo uno	pobre
FCM Designer	2010	limitado	ninguna	adecuada
FCM Tool	2011	varias	solo uno	avanzada
JFCM	2013	para programadores	ninguna	ninguna
Mental Modeler	2014	limitado	ninguna	adecuada
ISEMK	2015	varias	varios	adecuadas
FCM Expert	2017	varias	varios	adecuadas

### 1.2.4. Aplicaciones de los mapas cognitivos difusos en la agricultura

En las últimas décadas se han incrementado las investigaciones relacionadas con la teoría de los MCDs en varios ámbitos económicos, sociales y ambientales. Los MCDs son capaces de manejar situaciones que incluyen descripciones inciertas con procedimientos similares, al razonamiento humano (Felix *et al.*, 2017). Eso es un enfoque desafiante para la toma de decisiones, especialmente en entornos de procesamiento complejos como es el caso de la agricultura. Aunque en la actualidad no se cuenta con muchas aplicaciones en dicha esfera, debido al fuerte carácter estocástico que presenta (Arriaga, 2015).

Muchos tipos de problemas ecológicos o ambientales se beneficiarían de modelos basados en el conocimiento de las personas. Para crear modelos ecológicos con conocimiento tanto de los expertos como de las personas locales, Ozesmi (2004) propuso un enfoque basado en MCDs en varios pasos. En esta investigación se examinan las percepciones de diferentes partes interesadas en un conflicto ambiental, con el objetivo de facilitar el desarrollo de planes de gestión ambiental y determinar los deseos de resaneamiento de personas desplazadas por un proyecto de presa a gran escala. También, Hobbs *et al.* (2002) presentaron

los pasos involucrados en la construcción de un MCD de un ecosistema y la interpretación de la salida del MCD usando estadísticas multivariadas. Para ilustrar estas ideas, se creó un modelo de ecosistema complejo, de más de 160 variables, construido para la cuenca hidrográfica del Lago Erie.

Por otra parte, [Papageorgiou \*et al.\* \(2010\)](#) investigaron el rendimiento y sus variaciones en el complejo proceso de gestión del cultivo del algodón, donde influyen aspectos como el suelo, el propio cultivo y determinados factores climáticos, todo esto usando los MCDs para modelar y representar el conocimiento de los expertos. Luego, ([Papageorgiou \*et al.\*, 2011](#)) utilizaron los MCDs para predecir el rendimiento del cultivo del algodón. El modelo MCD desarrollado consistió en nodos unidos por bordes dirigidos, donde los nodos representaron los principales factores que afectan la producción del algodón, tales como textura, materia orgánica, pH, K, P, Mg, N, Ca, Na, y los bordes dirigidos mostraron las relaciones causa-efecto entre las propiedades del suelo y el rendimiento de algodón. En esta investigación se usó la teoría de los mapas ya no solo para modelar las condiciones del sistema en cuestión, también para inferir el posible desarrollo del mismo. Revelaron su ventaja comparativa con respecto a los algoritmos de aprendizaje automático comparativos probados para un conjunto de datos; al proporcionar decisiones que coinciden mejor con los reales medidos. La principal ventaja de este enfoque radicó en su estructura y flexibilidad simples, que representaron el conocimiento de forma visual y más descriptiva. Aquí se demostró las potencialidades de esta herramienta a la hora de interpretar datos recopilados en ambientes con dinámicas altamente complejas.

Otro ejemplo de una metodología del MCD para representar el conocimiento de los expertos, la desarrollaron [Papageorgiou \*et al.\* \(2013\)](#) con una predicción de rendimiento y manejo de cultivos. El modelo MCD desarrollado consistió en nodos unidos, que representaron los principales factores del suelo que afectan el rendimiento, [como textura del suelo, conductividad eléctrica del suelo, potasio, fósforo, contenido de materia orgánica, calcio y zinc], y los bordes dirigidos mostraron las relaciones causa-efecto entre las propiedades del suelo y el rendimiento. Se clasificó el rendimiento del cultivo de la manzana utilizando un algoritmo de aprendizaje MCD eficiente, el aprendizaje Hebbian no lineal, que comparado con la herramienta MCD convencional y los algoritmos de aprendizaje automático de referencia, demostró superioridad del enfoque de aprendizaje de MCD en la predicción del rendimiento.

Un estudio desarrollado por [Halbrendt \*et al.\* \(2014\)](#) buscó comprender la relación entre las tendencias en el razonamiento de los expertos y los agricultores rurales; y las predicciones sobre los resultados asociados con la tecnología de desarrollo basada en estas creencias. Además, comparó estas diferencias basadas en modelos mentales con las condiciones ambientales locales (usando mediciones del suelo) y los resultados agrícolas en términos de

producción(es decir, rendimiento). Mientras que los modelos mentales de los investigadores predijeron que la labranza mínima mejoraría el rendimiento, los modelos mentales de dos de las tres aldeas predijeron que el rendimiento disminuiría. Las mediciones locales de suelo y rendimiento apoyaron las predicciones del modelo mental de los agricultores. Los resultados indicaron que las técnicas de agricultura de conservación no deberían aplicarse universalmente; los profesionales del desarrollo deben participar en un aprendizaje bidireccional con las comunidades locales para beneficiarse del conocimiento local.

Otra investigación diseñada por [Christen et al. \(2015\)](#) presentó una metodología MCD sobre la afectación a los agricultores y las prácticas agrícolas de la regulación ambiental y la búsqueda de los factores más importantes en el cumplimiento o el incumplimiento de esta regulación. El estudio comparó las opiniones de dos grupos (agricultores y no agricultores), utilizando visualizaciones de la red de MCD que fueron validadas mediante entrevistas y una sesión de taller. Entre los dos grupos, el proceso de MCD reveló una percepción muy diferente de la importancia y la interacción de los factores; y sugiere firmemente que el problema radicó en un fracaso institucional en lugar de en una simple falta de voluntad de los agricultores para obedecer las reglas. MCD permitió un proceso estructurado de identificación de áreas de percepciones conflictivas, pero también áreas donde grupos de partes interesadas muy diferentes pueden ganar terreno común. De esta manera, el MCD ayudó en la identificación de puntos para el desarrollo de políticas específicas y tuvo el potencial de convertirse en una herramienta útil en el diseño y la comunicación de políticas agrícolas. Los resultados mostraron la utilidad del MCD al señalar cómo la regulación ambiental podría modificarse para aumentar el cumplimiento de las normas y dónde podrían buscarse los motivos del fracaso institucional identificado.

En el trabajo de [Jayashree et al. \(2015\)](#) categorizaron el nivel de producción de *C. nucifera* para un conjunto dado de condiciones agroclimáticas utilizando la metodología del MCD mejorada por sus capacidades de aprendizaje. Los datos de campo de cocoteros reales de diferentes temporadas para el período 2009-2013 de la región de Malabar del estado de Kerala se utilizaron para la capacitación y evaluación del MCD. Se demostró las capacidades de clasificación y predicción de MCD para la aplicación de agricultura de precisión descrita, con los dos enfoques de aprendizaje de MCD más conocidos y eficientes: Hebbian no lineal (NHL) y Hebbian no lineal basado en datos (DDNHL). El DDNHL-FCM ofreció una precisión de clasificación general del 96 %, y su capacidad para inferir nuevos conocimientos utilizando NHL. También se desarrolló una herramienta de *software* con una interfaz gráfica de usuario (GUI) para agrónomos y agricultores que proporciona la estimación del nivel de rendimiento de coco para el conjunto dado de parámetros ambientales y del suelo.

Un enfoque híbrido que utilizó algoritmos de aprendizaje de MCD presentaron [Natarajan et al. \(2016\)](#) para la caña de azúcar. Se propuso una clasificación de rendimiento,

combinando los aspectos clave del algoritmo DDNHL y algoritmo genético (GA) llamado FCM-DDNHL-GA. El modelo MCD desarrollado incluye varios parámetros de suelo y clima que influyen en la agricultura de precisión. Las precisiones de clasificación y las capacidades de inferencia del algoritmo de aprendizaje híbrido de los MCDs se analizaron y se compararon con algoritmos de aprendizaje para la aplicación de monitoreo del rendimiento de la caña de azúcar. Los resultados experimentales mostraron la superioridad del aprendizaje híbrido que proporcionaron una precisión de clasificación significativamente mayor.

Se propuso una metodología de MCD basado en reglas dinámicas para realizar simulaciones donde la novedad reside en un algoritmo de inferencia avanzada, mejorado con un razonamiento que se caracteriza por magnitudes de causa y efecto (Mourhir *et al.*, 2017). El método propuesto aprovechó el conocimiento de los expertos para proporcionar una estimación del rendimiento de los cultivos y, por lo tanto, puede permitir a los agricultores obtener información sobre la variación del rendimiento en un campo, para que puedan determinar cómo adaptar la aplicación de fertilizantes en consecuencia. Permitió también, producir simulaciones que pueden ser utilizadas por los administradores para identificar los efectos del aumento o la disminución de los fertilizantes en el rendimiento, y por lo tanto, puede facilitar la adopción de regulaciones de agricultura de precisión por parte de los agricultores. Se presentó un ejemplo para predecir el cambio en el rendimiento del algodón, como respuesta a las opciones de manejo estimuladas utilizando escenarios proactivos, basados en la disminución de fósforo, potasio y nitrógeno. Los resultados revelaron que disminuir los tres nutrientes a la mitad no disminuye el rendimiento en más de un 10 %.

### 1.3. Producción del *Solanum lycopersicum* L. en casas de cultivo

El tomate (*Solanum lycopersicum* L.) pertenece a la familia *Solanaceae*, es una de las hortalizas mundialmente más cultivadas en condiciones protegidas. Justifican la inversión bajo ambiente protegido, la alta demanda en los mercados y su valor comercial, además de la posibilidad de sembrar en épocas con condiciones climáticas desfavorables. La siembra de este cultivo es una de las más globalizadas, avanzadas e innovadoras de la industria hortícola (Cook and Calvin, 2005; Costa and Heuvelink, 2007).

Uno de los vegetales más populares, nutritivos y demandados en el mundo es el tomate, tanto para consumo fresco como a nivel industrial (Martin, 2013; Sturm and An, 2014; Padayachee *et al.*, 2017). Contribuye a suplir los requerimientos de nutrientes para la alimentación y la salud humana. Además, el tomate es rico en licopenos para mejorar la calidad de vida en pacientes con enfermedades cardiovasculares, reducir el riesgo de cáncer de próstata y otros tipos de cáncer (Zhang *et al.*, 2015; Ilahy *et al.*, 2016). Estadísticas de la FAO refieren que el tomate es la hortaliza más cultivada e importante a nivel mundial. En el año 2017 se sembraron 4 848 382 ha y se obtuvieron 182 301 395 t (37,6 t/ha)

de las cuales correspondió a continente americano el 18,1%. Los mayores productores de dicho vegetal son China, Estados Unidos de América e India (FAOSTAT, 2018). En Cuba, también es una de las hortalizas de mayor demanda, por tanto, predomina su cultivo. En el año 2017 se sembraron 48 713 ha que representaron el 25,13% del total destinadas a hortalizas. La producción alcanzó 584 072 t (12 t/ha) (ONE, 2018).

En el cultivo protegido de hortalizas, las fechas de plantación y la duración de los ciclos de cultivo varía considerablemente en respuesta a los precios del mercado, las condiciones ambientales y las estrategias de manejo de los productores (FAO, 2013). En tomate, la duración del ciclo de cultivo varía según las características del cultivar, las condiciones climáticas y el manejo agronómico (Redmond *et al.*, 2018). Algunos autores lo sitúan en 130 días después del trasplante (ddt) (Antúnez and Felmer, 2017) y otros en menos tiempo, por ejemplo entre 65 y 100 días (Redmond *et al.*, 2018). En este trabajo de tesis se utilizarán las descripciones de las fases fenológicas presentadas por (Moreno, 2007) como se muestra en el cuadro 1-2.

Tabla 1-2: Fases fenológicas del tomate, según Moreno (2007)

Fases	Objetivos
I	Trasplante a emisión del 1er racimo floral
II	Emisión del 1er racimo floral al cuaje del 3er
III	Cuaje del 3er racimo a inicio de cosecha
IV	Inicio de cosecha a cosecha completa del antepenúltimo racimo
V	Antepenúltimo racimo cosechado hasta el final

La duración de cada fase puede variar de acuerdo con el método de cultivo, las características propias de cada cultivar y las condiciones climáticas (Mollinedo and Tapia, 2008). Aunque se produce en una amplia gama de condiciones de clima y suelo, el tomate prospera mejor en climas secos con temperaturas moderadas. Su rusticidad asociada a nuevos cultivares permite su cultivo en condiciones adversas. No obstante, el tomate es una especie de estación cálida, su temperatura óptima de desarrollo varía entre 18°C y 30°C. Respecto a la humedad relativa, el desarrollo del tomate requiere que ésta oscile entre 60% y 80%. Humedades relativas muy elevadas favorecen el desarrollo de enfermedades (fúngicas y bacterianas) y dificultan la fecundación, debido a que el polen se compacta y abortan parte de las flores. La luminosidad en el cultivo de tomate cumple un rol importante, más allá del crecimiento vegetativo de la planta, ya que el tomate requiere al menos seis horas diarias de luz directa para la etapa de floración. Atendiendo a ello, se debe garantizar la limpieza de los techos de las casas de cultivo para una buena iluminación (Moreno, 2007; Antúnez and Felmer, 2017).

Para el cultivo de tomate se requiere agua suficiente con el objetivo de reponer la humedad perdida por evapotranspiración (Etc). El riego también servirá para enfriar el cultivo por medio de la transpiración, especialmente en días muy calurosos, además de permitir la

lixiviación de sales que se acumulan en la zona de raíces. La cantidad de agua que requiere el tomate dependerá de las condiciones meteorológicas durante el ciclo de cultivo, de las propiedades físicas de retención de agua en el suelo y de las prácticas de riego ([Antúnez and Felmer, 2017](#)).

Con una mejor estimación de la evapotranspiración se lograría optimizar el manejo del riego en cultivo protegido de tomate. Para el cálculo de la evapotranspiración del cultivo (mm/día) se emplean generalmente métodos indirectos para estimarla e implican análisis del balance de energía, transferencia de masas, balance de agua en el suelo o a través de fórmulas empíricas o semi-empíricas ([Allen et al., 1998](#)). Sin embargo, debido a la dificultad de obtener mediciones de campo precisas, la Etc se calcula comúnmente con datos meteorológicos ([Antúnez and Felmer, 2017](#)).

Un manejo adecuado del riego depende de la demanda hídrica del cultivo en relación con las condiciones ambientales y la fase fenológica. Además, es importante la determinación correcta del volumen de agua que se requiere aplicar según el tipo de suelo que se dispone y el sistema radical de las plantas para reducir las pérdidas de agua y no afectar los rendimientos ([FAO, 2013](#)).

El riego por goteo es uno de los sistemas de riego que más se emplean en cultivo protegido ([Intriago et al., 2019](#)). El riego por goteo ofrece ventajas para obtener elevadas producciones con gran eficiencia en el uso del agua. En este sentido, se destacan: amplia y exacta distribución uniforme del agua, se puede emplear la fertilización localizada junto al riego (fertirriego), permite flexibilidad en los horarios de riego, los volúmenes de descarga son bajos lo que se traduce en una economía del bombeo, se puede utilizar en todos los cultivos en hilera, no es afectado por el viento, facilita la automatización, entre otras ([FAO, 2013](#)).

### **1.3.1. Modelos de predicción en tomate**

Los modelos matemáticos de cultivo simulan algunas características como el crecimiento a través de varios procesos fisiológicos de las plantas tales como la fotosíntesis, respiración, transpiración y desarrollo de cultivo. Los modelos permiten probar diferentes hipótesis del comportamiento de los cultivos sin la necesidad de realizar experimentos, lo cual implica reducir los costos. Por lo tanto, a partir de los modelos matemáticos de los cultivos es posible obtener buenas aproximaciones de la realidad cuando se hace una adecuada calibración y evaluación de estos ([Bouman et al., 1994](#)).

Para el cultivo de tomate se han desarrollado varios modelos de crecimiento a lo largo de los años. Un modelo determinístico de crecimiento de tomate de invernadero escrito en lenguaje Pascal fue desarrollado basado en un modelo de asimilación de la hoja, más una teoría de respiración y una del rango de fotosíntesis, lo cual es controlado por las condiciones ambientales y los contenidos de carbohidratos en las hojas. El modelo fue aplicado para estimar el efecto del  $CO_2$  sobre la producción de frutos de tomate ([KANO](#)

and VAN GAVEL, 1988). En cuanto a la distribución de materia seca entre hojas, tallo, raíz y frutos de las plantas; Heuvelink (1989) desarrollaron un modelo dinámico. Los resultados de simulación de la materia seca entre las hojas, tallos y frutos de tomate correspondieron razonablemente bien con los datos medidos en el experimento.

Jones *et al.* (1991) implementaron el complejo modelo TOMGRO, inicialmente, propuso 69 variables de estado para el desarrollo y producción de plantas de tomate. El modelo tuvo un vector de entrada de clima y las variables: temperatura del aire, radiación solar (PAR) y concentración  $CO_2$ . El vector de salida consistió en siete principales grupos de variables de estado: número de hojas, número de nuevos tallos, número de frutos, masa seca de las hojas, peso seco de nuevos tallos, peso seco de frutos y foliar zona. La última versión de TOMGRO (v3,0) consta de 574 variables de estado, simulando con gran detalle el desarrollo de fruta. El desarrollo de la fruta está modelado separado del crecimiento de la biomasa, que permite que el tamaño pueda ser manejado como una variable (?). Luego, Jones *et al.* (1999) propusieron un modelo TOMGRO más simple para cultivos de tomate que se comporta de manera similar, pero con la ventaja de tener solo cinco variables de estado (Ramírez-Arias, 2005). Esta versión simplificada fue evaluada con datos de varios experimentos, incluidos los datos recogidos en un invernadero comercial y los resultados mostraron que el modelo puede describir con precisión el crecimiento y el rendimiento de tomates en diferentes lugares y marcos temporales. Los resultados mostraron que el modelo es inadecuado para describir el balance de  $CO_2$  del agrosistema del invernadero, sin embargo, se determinó que puede ser usado en un modelo para describir el crecimiento y el desarrollo del cultivo de tomate (Zekki *et al.*, 1999). Dicho modelo ha sido estudiado y calibrado bajo diferentes condiciones climáticas que demuestran que es robusto (Medina-Ruíz *et al.*, 2011). Dentro de los ejemplos se encuentra, de Koning (1994) que desarrolló un modelo de 300 variables para predecir la distribución de materia seca con tomates cultivados en invernaderos. El número de frutos fueron estimados a partir de predecir la iniciación, aborto y recolección de individuos. El modelo desarrollado logró predecir razonablemente la formación de racimos de fruta, la tasa del crecimiento del fruto, y la distribución de la materia seca. La predicción del número de frutos por racimo; no se encontró aceptablemente exacto.

Tomsim es otro modelo desarrollado para tomate con una estructura modular, que predice su crecimiento y desarrollo (Ramírez-Arias, 2005) y representó un modelo simplificado del TOMGRO. Dimokas *et al.* (2008), propusieron un nuevo modelo modificado del modelo original TOMGRO bajo las condiciones del este de Grecia. El modelo fue evaluado y calibrado en el cultivo de tomate bajo condiciones de invernadero, ya que hubo una buena relación entre los datos medidos y simulados tomando en cuenta el desarrollo del cultivo, la biomasa y la producción de frutos. Ruíz-García *et al.* (2015) utilizaron el método de

asimilación de datos basado en el *Kalman Unscented Filter* (UKF) para mejorar la predicción de tres variables de estado, incorporando mediciones del cultivo. Sobre la base de los resultados de las pruebas estadísticas, se concluyó que UKF mejora con éxito la predicción de las tres variables de estado del modelo TOMGRO. Por lo tanto, el filtro Kalman es un método de asimilación de datos eficiente para modelos dinámicos no lineales de crecimiento de cultivos bajo invernaderos.

Existen otros modelos para el cultivo protegido del tomate como los informados por [Abreu et al. \(1998\)](#) y [Vaca et al. \(2015\)](#) para predecir la producción. Mientras que [Dai et al. \(2006\)](#), desarrollaron un modelo relativamente simple para predecir la producción de biomasa del dosel; además de la cosecha de frutos para tres cultivos, incluido el tomate. Los resultados mostraron que el modelo predice satisfactoriamente la producción de biomasa del dosel y la cosecha de los frutos para los tres cultivos. También, se realizó un análisis de crecimiento de tomate por [Juárez-Maldonado et al. \(2015\)](#) donde se presentaron una etapa de crecimiento exponencial y otra lineal. Se demostró que existe una fuerte correlación entre las masas frescas y secas de los diferentes órganos de la planta, y en las diferentes etapas de crecimiento. La generación y acumulación de biomasa por las plantas de tomate se afectó por las condiciones climáticas internas de la casa.

En los últimos años se encontraron otras aplicaciones como es el caso de [Gutiérrez and Arias \(2017\)](#) que realizaron un control óptimo para regular la temperatura en el interior de un invernadero, a partir de un modelo matemático. El modelo consideró cinco estados: relativo con el cultivo, el número de nodos, la masa seca del dosel y la masa seca de la raíz; con respecto al invernadero, la temperatura y la humedad relativa. La propuesta tuvo como finalidad aumentar la ganancia del agricultor y disminuir el consumo de energía. También, se diseñó un sistema neurodifuso para un cultivo, que controla la humedad del suelo mediante un riego automatizado. Este sistema utilizó una red neuronal tipo *feedforward* con algoritmo *backpropagation* (topología multicapa) y el sistema difuso tipo Takagi-Sugeno. La red neuronal emplea los valores de entradas de los sensores de humedad y el sistema difuso los valores de salida para el control de humedad del cultivo. Un sistema neurodifuso fue simulado en Matlab por [de la Piedra et al. \(2016\)](#), e implementado en una placa Arduino. [Silva \(2016\)](#) calibró y probó el modelo CROPGRO-Tomato con datos de cinco experimentos de campo con procesamiento de tomate bajo diferentes concentraciones de fósforo del suelo. Por otra parte, ([Moreno-Ramos et al., 2017](#)) propusieron bajo ciertas condiciones una trayectoria subóptima de la temperatura, para con ciclos estables, maximizar la producción de peso y número de fruto cosechado de tomate. Por último, se realizó un modelo matemático del sistema de riego equilibrando la humedad del suelo, la presión y flujo de agua. Basado en la entrada/salida, y usando las tres mediciones de estos tres parámetros, y utilizando la caja de herramientas "Ident" de MatLab se obtuvo

un modelo estimado lo más cercano posible al sistema de riego real (Sahbani and Ferjani, 2018).

#### 1.4. Metodologías en el modelado basado en la teoría de los mapas cognitivos difusos

En las metodologías de modelado MCD encontradas en la literatura científica, la mayoría inician identificando las variables que mejor describen el modelo del sistema con la ayuda de la ingeniería del conocimiento, ya que los expertos conocen los factores, funcionamiento y comportamiento del sistema (Morone *et al.*, 2019). Luego, dichos expertos son los encargados de describir las relaciones entre los conceptos, aplicando las reglas *IF-THEN* o a través del empleo de técnicas *fuzzy*, para justificar sus sugerencias de causa y efecto entre los conceptos. Así de una forma u otra, se logra la obtención de un peso difuso entre cada una de las variables, que se traduce a un peso numérico el cual se utiliza en el programa de razonamiento MCD impuesto (Papageorgiou *et al.*, 2011, 2016). En tales casos, su desarrollo podría verse afectado significativamente por los limitados conocimientos y habilidades del ingeniero del conocimiento. Por lo tanto, es esencial utilizar algoritmos de aprendizaje para lograr buen modelado MCD: ejemplos son las investigaciones por Papageorgiou *et al.* (2013) que aplicaron NHL y Natarajan *et al.* (2016) que implementaron el DDNHL. Una comparación entre ambos algoritmos realizaron Jayashree *et al.* (2015) los cuales demostraron mejores resultados con NHL en la categorización de la producción de *C. nucifera*. En todas las aplicaciones revisadas se constató que las implementaciones utilizando la teoría MCD no presentan un orden establecido para el modelado en diferentes casos de estudio.

#### 1.5. Ingeniería del conocimiento

La ingeniería del conocimiento es una disciplina tecnológica que se centra en la aplicación de una aproximación sistemática, disciplinada y cuantificable; al desarrollo, funcionamiento y mantenimiento de sistemas basados en el conocimiento. Esta ingeniería surgió de la necesidad de establecer principios metodológicos y científicos que permitan desarrollar esos sistemas, a partir de la inteligencia computacional. Además, esta disciplina desarrolla la adquisición del conocimiento así como su especificación, verificación, validación, diseño e implementación en sistemas informáticos o lenguajes apropiados para la construcción de bases de conocimiento para la toma de decisiones (González *et al.*, 2017).

Para la adquisición del conocimiento se utilizó la metodología del criterio de expertos como una alternativa para el modelado bajo condiciones de incertidumbre. Estos han alcanzado popularidad en las Ciencias Sociales (Manzanet *et al.*, 2018), Económicas (Abreu *et al.*, 2016), Médicas (Pérez *et al.*, 2018) y Técnicas (González Robaina *et al.*, 2015), entre otras.

Los métodos basados en criterios de expertos emplean la consulta a personas que tienen grandes conocimientos sobre el objeto de estudio. La síntesis, el consenso y la estabilidad

del juicio colectivo pueden ofrecer una visión verosímil del futuro, combinando la imaginación y el talento individual. Por este motivo, estos métodos son básicamente subjetivos (Cruz, 2009). Se tiene en cuenta como experto, tanto al individuo en sí, como a un grupo de personas u organizaciones capaces de ofrecer valoraciones conclusivas sobre un problema y de hacer recomendaciones respecto a sus posibles soluciones, con un máximo de competencia (Ramírez and Toledo, 1999).

### **Metodología de gestión de la información que proporciona el criterio de expertos:**

1. Elaboración del objetivo: la formulación de este debe hacerse en función de la validación del resultado investigativo que se propone. (Matos and Cruz, 2011; Cruz, 2009; Ramírez and Toledo, 1999; Gorina, 2010; de Arquer, 2011; Schery and Sánchez, 2013).
2. Selección de los expertos: para que sea certera esta selección hay que tener en cuenta que un experto altamente calificado debe poseer cualidades como ética profesional, maestría, imparcialidad, intuición, amplitud de enfoques, independencia de juicios. En consecuencia, para su selección se utilizan criterios relativos a su competencia, creatividad, disposición a participar en la encuesta, conformidad, capacidad de análisis, espíritu colectivista y autocrítico y efectividad de su actividad profesional. De ahí, que para determinar la competencia del candidato a experto se emplee una autoevaluación del nivel de conocimiento sobre el tema, experiencia de orden empírico (práctica profesional), experiencia en el desarrollo de las investigaciones teóricas, capacidad de análisis y conocimiento del estado actual del problema y comprensión del mismo (Gorina, 2010; Cruz, 2009; Ramírez and Toledo, 1999; Schery and Sánchez, 2013).
3. Elección de la metodología o método: existen varias metodologías y métodos para la evaluación a través del criterio de expertos también se pueden combinar para superar las debilidades de las mismas.
  - Preferencia: es más empleada por su exactitud, objetividad y rapidez. Según la misma, los expertos ubican los aspectos evaluados siguiendo una encuesta, la que generalmente ha sido conformada por grupos de ítems que siguen un orden descendente en cuanto al nivel de calidad que reflejen, es decir, que manifieste el resultado de las investigaciones objeto de análisis. Así el lugar que ocupa un ítem está determinado por la cantidad de puntos acumulados por este y mientras mayor sea el total de esos puntos, mayor será la calidad del resultado valorado y viceversa (Ramírez and Toledo, 1999).
  - Método de agregados individuales: requiere individualmente de cada experto una estimación directa de la probabilidad de éxito o fracaso en cada uno de los ítems descritos. Después se tratan estadísticamente los datos recogidos. Lo habitual es calcular la media aritmética del conjunto de estimaciones obtenidas individualmente para cada ítem, lo que se hace presuponiendo que el conjunto de los posibles datos

tengan una distribución simétrica y, por tanto, la media aritmética es un buen índice de tendencia central. Este método es considerado económico porque, al igual que el método Delphi, no exige que se reúna a los expertos en un lugar determinado. Puede parecer limitado ya que los expertos no pueden intercambiar sus opiniones, puntos de vista y experiencias, las que se les solicitan de manera individual; no obstante, esta limitación es precisamente lo que se busca para evitar sesgos de los datos que pudieran ser ocasionados por conflictos interpersonales, presiones entre los expertos, etc. (de Arquer, 2011).

- Comparación por pares: cada experto rellena una tabla de contingencia, en la cual los aspectos a evaluar se encuentran ubicados tanto en sentido horizontal como vertical. Cada celda de la citada tabla guarda relación con los aspectos comparados vertical y horizontalmente y en ella se coloca aquel número, perteneciente a una escala predeterminada, que a juicio del experto mejor refleja el resultado objeto de evaluación (Ramírez and Toledo, 1999).
- Metodología Delphi o Delfos: es la más completa y compleja, motivo por el cual su aplicación requiere más tiempo para obtener los resultados de la evaluación de los expertos. Esta metodología, a decir de los especialistas, consiste en la utilización sistemática del juicio intuitivo de un grupo de expertos para obtener un consenso de opiniones informadas. Su esencia está dada por la organización de un dialogo anónimo entre expertos consultados individualmente. Dicha consulta se lleva a cabo mediante cuestionarios, con vistas a obtener un consenso general o, al menos, los motivos de una cierta discrepancia (Gorina, 2010; Cruz, 2009; Kayo and Securato, 1997).
- Método del consenso grupal: consiste en reunir a los expertos y indicarles que se deben lograr una estimación de la probabilidad de éxito o de fracaso para cada ítem, que sea satisfactoria para todos los expertos. Con estas instrucciones se maximizan los intercambios de información y opiniones dentro del grupo de expertos. Si el grupo no logra un consenso, puede intentarse un consenso artificial en que se recogen las estimaciones individuales y se sintetizan estadísticamente. Este método, como el anterior, también precisa que el grupo de expertos sea pequeño, se fomente la libre expresión y se eviten las discusiones tensas y los sistemas de votación (de Arquer, 2011).
- Taller de Socialización: permite el debate y la discusión grupal como vía para la valoración científica. Debe elaborarse un informe de investigación para estos efectos, contentivo de los principales aportes construidos, de modo que los expertos seleccionados puedan realizar un estudio exhaustivo de ellos, previo a la realización del taller, para garantizar la profundidad de las valoraciones a realizar. Al momento de la realización del taller, el investigador, realizará una breve presentación oral de sus

resultados y propondrá criterios de análisis como una propuesta de guía para la valoración científica a realizar. Posteriormente, los especialistas procederán a debatir acerca de sus criterios y juicios críticos, los que serán registrados por el investigador para posteriormente proceder a su valoración cualitativa, con vistas al perfeccionamiento y enriquecimiento de su investigación (Matos and Cruz, 2011; Schery and Sánchez, 2013).

4. Ejecución de la metodología: se desarrolla en correspondencia con los criterios establecidos por la metodología o método seleccionado. En general debe iniciarse con una cuidadosa elaboración de la guía de aspectos a someter a la valoración de los expertos. Estos deberán expresar sus ideas y criterios sobre las bondades, deficiencias o insuficiencias que presenta el resultado científico valorado, así como ordenar los aspectos que con mayor calidad reflejen dicho resultado. Debe velarse porque los expertos comprendan adecuadamente la forma en que se les pide que evalúen los indicadores y porque todos los indicadores sean evaluados (Gorina, 2010; Cruz, 2009; Kayo and Securato, 1997; Ramírez and Toledo, 1999; de Arquer, 2011; Schery and Sánchez, 2013).
5. Procesamiento de la información: se procesa toda la información almacenada en la base de datos, lo que está condicionado a la metodología seleccionada. Si se asume una metodología cuantitativa, se deben calcular las medidas de interés para la toma de decisiones, en especial las relativas al nivel de asociación entre variables o expertos, dando cuenta de la confiabilidad de estas últimas. También será conveniente realizar pruebas de hipótesis para poder hacer valoraciones o tomar decisiones a partir del criterio de probabilidad, pudiéndose incluso concluir que se deben aplicar nuevamente algunos de los pasos anteriores para poder tomar una decisión final por parte del investigador. En el caso que la metodología asumida sea cualitativa, debido a que la información disponible sea textual, deberá realizarse el correspondiente análisis del contenido, sobre la base de los indicadores previstos. Será beneficioso realizar recuentos simples de respuestas similares para elaborar una estructura de codificación a partir de temas recurrentes, categorías analíticas y explicaciones teóricas, todo lo cual facilitará arribar a conclusiones sobre los objetivos específicos definidos para la consulta a expertos. (Gorina, 2010; Cruz, 2009; Kayo and Securato, 1997; Ramírez and Toledo, 1999; de Arquer, 2011; Schery and Sánchez, 2013). Finalmente, cabe precisar que puede elegirse trabajar con una metodología mixta (cualitativa y cuantitativa).

### 1.6. Consideraciones finales del capítulo

En el modelado de los mapas cognitivos difusos es muy importante la elección de la ecuación de activación, la función de transferencia y el método de aprendizaje. Para ello, se elige la ecuación de activación y el algoritmo de aprendizaje implementado en la herramienta de *software* LTCNIs. La ecuación de activación muestra el grado de activación de los conceptos, teniendo en cuenta su estado anterior y el de las entradas. El algoritmo de

aprendizaje implementado es el gradiente descendente por ser uno de los más sencillos para simplificar el modelado en cuestión. La función de transferencia es la de saturación para un modelado de naturaleza tan compleja con numerosa cantidad de variables. Dentro del modelados en la agricultura utilizando MCD, a pesar de la existencia de varios ejemplos, no logran cerrar la brecha existente entre los avances teóricos en el campo del MCD y el desarrollo de aplicaciones prácticas.

El cultivo de *S. lycopersicum* en cultivo protegido necesita agua y el método más adecuado para irrigar este cultivo es el riego por goteo ya que aumenta el rendimiento con una menor entrada de este preciado recurso. Los principales factores que influyen en la producción del tomate son los climáticos, condiciones del suelo y fenología del cultivo.

Las metodologías encontradas en la literatura científica no presentan un orden establecido para el modelado utilizando la teoría MCD en diferentes casos de estudio.

La metodología en la ingeniería del conocimiento para el modelado del riego de *S. lycopersicum* utilizando mapas cognitivos difusos cuenta con las etapas: elaboración del objetivo, selección de los expertos, elección del método, ejecución y procesamiento de la información.

# CAPÍTULO 2

## MATERIALES Y MÉTODOS

### 2.1. Introducción

En este capítulo se describe la metodología para el modelado basado en los mapas cognitivos difusos y la ingeniería del conocimiento; de la cual inicialmente se describe la forma de determinar las principales variables. Luego, se explica los métodos a utilizar en la obtención del modelo inicial. Además, se define la vía de registro y almacenamiento de las mediciones reales de las variables, previamente definidas, en la UEB de cultivos varios “Valle del Yabú”. También, se describe la biblioteca de *software* a utilizar en la evaluación, ajuste y validación del modelo de riego del tomate en casas de cultivo. Por último, se plantea la forma de valoración por los expertos y definición de las pruebas estadísticas a las que somete el modelo obtenido.

### 2.2. Metodología para el modelado mediante la ingeniería del conocimiento y los mapas cognitivos difusos

Las investigaciones científicas que realizan modelos MCDs en una aplicación práctica no presentan un orden metodológico establecido (Papageorgiou *et al.*, 2010). Por ello, se empleó la metodología del cuadro 2–1 utilizando etapas encontradas en diversas aplicaciones (Papageorgiou *et al.*, 2013; Jayashree *et al.*, 2015; Morone *et al.*, 2019). Esta metodología ha sido representada en la figura 2–1 como un diagrama de flujo.

Tabla 2–1: Metodología para el modelado con MCD

No	Acción	Recursos involucrados
1	Determinación de variables	Expertos
2	Confección del modelo inicial	Expertos y MCD
3	Recopilación de datos históricos	Sistema de adquisición de datos
4	Evaluación del modelo inicial	MCD
5	Ajuste del modelo	MCD
6	Validación del modelo final	MCD

En la primera acción, el experto determina las variables del sistema (Papageorgiou *et al.*, 2013; Morone *et al.*, 2019). Dentro de esta lista de variables según Sosa *et al.* (2019), el experto debe clasificarlas en conceptos internos del sistema y entradas del mismo. Además de

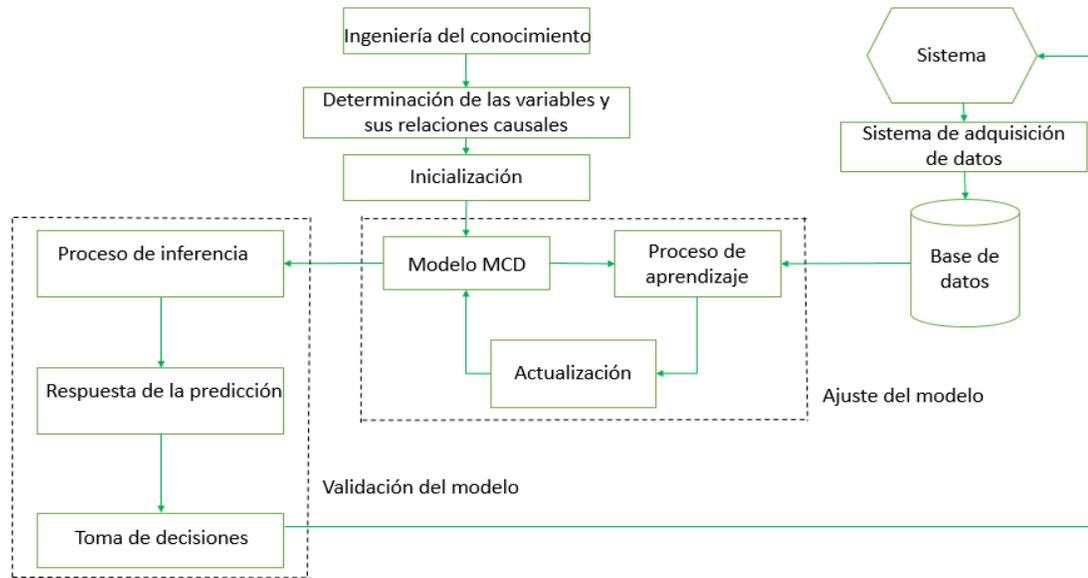


Figura 2–1: Diagrama de flujo de la metodología para modelado MCD

que son los encargados de establecer para estos conceptos las unidades de medida, los rangos de valores admisibles y los rangos adecuados de cuyos valores depende el rendimiento y el éxito del proceso.

Para la confección del modelo inicial, el experto aporta sus conocimientos sobre cuáles deben ser las relaciones causales entre los distintos conceptos (Stylios and Groumos, 2004). Para ello, se pretende que al menos sea capaz de decidir sobre una de las cuatro posibles opciones (positivo, negativo, cero o desconocido), en cuanto al peso sináptico entre el concepto A y B (dirigido de A hacia B). Si el experto refiere una relación positiva, un cambio en el valor de A influye en un cambio en el valor de B en el mismo sentido, el negativo refleja que un cambio en el valor de A influye en un cambio en el valor de B en sentido contrario y el cero simboliza que un cambio en el valor de A no influye en el valor de B (Zhi-Qiang, 2001). El desconocido es para cuando el experto desconoce cuál es el tipo de relación directa entre A y B.

Luego de establecido este paso, se pasa a la próxima fase, durante la cual se procede a recopilar datos históricos de las variables. Esta etapa requiere un período de muestreo, que no es más que el espacio de tiempo que transcurre entre una medición y otra.

La evaluación de este modelo inicial se basa en dos acciones. Primero se establecen las condiciones iniciales y se realiza el proceso de inferencia. Luego, los resultados de este proceso se comparan con los datos históricos mediante el cálculo y el graficado del error (Jayashree *et al.*, 2015).

En la etapa de ajuste del modelo se utiliza el método de aprendizaje, que en este caso es el gradiente descendente, para acercar la respuesta el máximo posible a los datos recopilados (Nápoles Ruiz, 2014). En este proceso se debe utilizar el 70 % de estos valores que fueron

previamente almacenados en la etapa anterior (Sosa *et al.*, 2019). Aquí es posible que se tenga que variar el factor de aprendizaje del mapa hasta obtener los mejores resultados posibles.

Por último, se procede a la validación del resultado final, para lo cual se compara la respuesta del modelo ajustado con el 30 % restante de los datos medidos (Sosa *et al.*, 2019), a través del cálculo y graficado del error.

### 2.3. Determinación de variables aplicando la ingeniería del conocimiento

La metodología de la ingeniería del conocimiento se utilizó para determinar las principales variables teniendo en cuenta las condiciones del suelo, climáticas y fenológicas involucradas en el riego del *S. lycopersicum* en cultivo protegido.

#### 2.3.1. Selección de los expertos

Los criterios para seleccionar los expertos fueron: nivel de competencia, creatividad, disposición a participar en la encuesta, capacidad de análisis y de pensamiento, espíritu colectivista y autocrítico. Para determinar el nivel de competencia de un candidato a experto, fue implementada la metodología de Evlanov and Kutuzov (1978), ampliamente utilizada en diversas aplicaciones (Zambrano *et al.*, 2015; Almenara and Cejudo, 2015; Crespo *et al.*, 2018). En esta metodología, la competitividad del experto se determina por un coeficiente  $k = \frac{1}{2}(kc + ka)$ , donde  $kc$  representa una medida del nivel de conocimientos sobre el tema investigado y  $ka$  una medida de las fuentes de argumentación.

Tabla 2–2: Escala tipo Linkert para el grado de influencia de las fuentes de argumentación

Fuentes de argumentación	ALTO(A)	MEDIO(M)	BAJO(B)
1. Análisis de las investigaciones teóricas y/o experimentales	0,3	0,2	0,1
2.Experiencia obtenida en la actividad profesional	0,5	0,4	0,2
3.Análisis de la literatura especializada y publicaciones de autores nacionales	0,05	0,05	0,05
4.Análisis de la literatura especializada y publicaciones de autores extranjeros	0,05	0,05	0,05
5.Conocimiento del estado actual de la problemática en el país y en el extranjero	0,05	0,05	0,05
6.Intuición	0,05	0,05	0,05

El cálculo de  $kc$  requiere de la autoevaluación del candidato en una escala de 0 a 10, donde el valor seleccionado se divide entre 10 para lograr cierta normalización. Respecto al cálculo de  $ka$ , también es necesario que el encuestado se autoevalúe, pero atendiendo a seis posibles fuentes de argumentación en una escala tipo Likert. Para ello debe completar marcando con  $x$  en cada fila del tabla 2–2, donde los números constituyen los pesos asignados a cada fuente y aparecen ocultos en el instrumento. Con los valores de la tabla patrón para cada

una de las celdas marcadas por el experto, se calcula  $ka$  que no es más que el número de puntos obtenidos en total (Zambrano *et al.*, 2015).

El número de expertos a utilizar en un caso de estudio no se encuentra determinado (Williams and Webb, 1994). No obstante, existen en la literatura algunas propuestas, entre las que se puede señalar a (Martínez *et al.*, 2018), que utilizó nueve candidatos; mientras que (Zambrano *et al.*, 2015) aplicó la metodología a 25; Witkin and Altschuld (1995) no indican un número concreto, pero sí que debe ser menor que 50, reconociendo que en algunos casos se puede recurrir a un número mayor en función de los objetivos que se persigan. Aunque también se puede obtener una cantidad estimada a través de la ecuación 2.1, según Sarache-Castro *et al.* (2015):

$$n = \frac{P(1 - P)K}{i^2} \quad (2.1)$$

Donde:

n: número de expertos

i: nivel de precisión deseado ( $\pm 10$ )

P: porcentaje de error tolerado

K: constante calculada a partir del nivel de confianza

El instrumento del anexo A.1, se aplicó a 17 candidatos con conocimientos sobre los procesos agronómicos que se desarrollan dentro del cultivo protegido. Entre los candidatos se encuentran profesores de la Facultad de Ciencias Agropecuarias de la Universidad Central "Marta Abreu" de las Villas con más de 10 años de experiencia en la docencia de la carrera de ingeniería en agronomía. Además, fueron encuestados ingenieros agrónomos y obreros agrícolas de la UEB de cultivos protegidos y semiprotegidos del "Valle del Yabú" con vasta experiencia en el manejo agronómico en casas de cultivo. Del total de candidatos fueron seleccionados como expertos los que cumplieron con el criterio de competitividad establecido, y por tanto, los que participaron en la ingeniería del conocimiento para el modelado del riego en el cultivo protegido del tomate.

### 2.3.2. Elección de la metodología

Dentro de los métodos de consulta a expertos, se elige la metodología Delphi donde a través de dos encuestas con preguntas cerradas se definieron las principales variables involucradas en el riego del tomate en cultivo protegido. La etapa preparatoria de este método se procedió a seleccionar los expertos, disposición en el ejercicio y explicar a los expertos el método; con el objetivo de lograr resultados más confiables, que en este caso ya estuvo anteriormente descrita. En la fase exploratoria el grupo coordinador (responsables de la investigación) sometieron a los expertos a un análisis y discusión de una primera encuesta, resultado de una previa y extensa revisión sobre el tema aplicando el método bibliográfico,

y con la fundamentación teórica originada del mismo. Luego, se sometió al grupo de expertos a una segunda ronda para la validación con la intención de profundizar en unos criterios cuantitativos y cualitativos más estables. La etapa final quedó diseñado y validado el cuestionario. Todo este proceso fue considerado por los investigadores del diseño de encuestas como validación de contenido, justamente por evaluar la concordancia entre los criterios de los expertos.

La programación del riego se encuentra dado por procedimientos técnicos desarrollados para predecir la forma y cantidad de riego. Generalmente, se basa en la medición directa o en cálculos de balance de agua, se tienen en cuenta la influencia y variabilidad de las condiciones climáticas, especialmente la temperatura, humedad relativa y luminosidad, asociados todos al desarrollo fenológico de las plantas (Dorais *et al.*, 2016). Por ello, en la primera encuesta estuvieron reflejados los factores generales que pueden influir en el crecimiento y desarrollo de hortalizas en cultivo protegido, tales como: factores climáticos, condiciones del suelo y fenología del cultivo.

Se inició la primera encuesta con preguntas cerradas, a manera de enfoque preinvestigativo del problema, para que en el subsiguiente cuestionario se puedan concretar las preguntas e ir cerrando el entorno de las respuestas a los puntos neurálgicos para definir las variables involucradas en el riego del cultivo protegido del tomate. Ello posibilitó una secuencia lógica de elementos que permitieron concretar las preguntas en el posterior cuestionario.

### 2.3.3. Ejecución de la metodología

#### Primera encuesta

La primera encuesta que se ilustra en el anexo A.2 propuso, a través de una escala de relevancia, determinar las variables involucradas en el cultivo protegido de hortalizas. Los especialistas seleccionados valoraron cada pregunta de la encuesta según las categorías: extremadamente relevante (*ER*), muy relevante (*MR*), relevante (*R*), poco relevante (*PR*) y no relevante (*NR*).

#### Segunda encuesta

A partir de la primera encuesta fue ajustada una segunda encuesta, expuesta en el anexo A.3, que tuvo como objetivo identificar las principales variables involucradas en el modelado del riego en cultivo protegido de tomate. Para ello, el experto valoró la posibilidad de mantener u omitir algunas variables del primer cuestionario. Además, se le brindó la opción de adicionar variables de forma justificada que no se encontraran en la encuesta inicial.

### 2.3.4. Procesamiento de la información

Las respuestas de los cuestionarios fueron procesadas estadísticamente en el *software* IBM SPSS Statistics 23 a través de la prueba de W de Kendall y las tablas de frecuencia.

La prueba W de Kendall se utilizó para conocer el grado de asociación entre un conjunto de rangos (Siegel *et al.*, 1972), por lo cual es particularmente útil cuando se les solicita a los expertos asignarle rangos a las preguntas (Escobar-Pérez and Cuervo-Martínez, 2008). Para el caso de la primera encuesta el rango es de 1 a 5, debido a las categorías de relevancia por las que debieron exponer su criterio los expertos. En la segunda encuesta el rango sería del 1 a 3 pues los expertos eligieron entre una omisión, mantenimiento de las variables del primer cuestionario, o agregación de una nueva variable.

La tabla de frecuencia representó el porcentaje de expertos que optó por uno de los cinco grados de relevancia para el primer cuestionario; y entre las tres opciones para el caso del segundo.

#### 2.4. Confección del modelo inicial

Para la confección del modelo inicial se aplicó una tercera encuesta como se muestra en el anexo A.4. En dicho cuestionario, se determinó las relaciones causales entre las variables obtenidas en la etapa anterior. Las relaciones entre las variable fueron establecidas por cada experto mediante un conjunto de términos lingüísticos difusos  $S$  de la ecuación 2.2, propuestos por (Pérez, 2014) expuesto en la tabla 2–3. Este conjunto agrupa el dominio de pesos causales entre  $(-1; 1)$ , donde los valores positivos expresan la implicación directa y los negativos la implicación inversa (Mar Cornelio *et al.*, 2015).

$$S_i = \{s_0, \dots, s_k\} \quad (2.2)$$

Tabla 2–3: Representación de los términos lingüísticos

Variable	Término lingüístico	Valor
$S_0$	Negativamente muy fuerte	-1
$S_1$	Negativamente fuerte	-0.75
$S_2$	Negativamente media	-0.5
$S_3$	Negativamente débil	-0.25
$S_4$	Cero	0
$S_5$	Positivamente débil	0.25
$S_6$	Positivamente media	0.5
$S_7$	Positivamente fuerte	0.75
$S_8$	Positivamente muy fuerte	1

La agregación o fusión de distintos modelos es relativamente fácil en los MCDs (Mar Cornelio *et al.*, 2015). Esta agregación de conocimiento posibilita un mejor modelo final, ya que no se encuentra sujeto únicamente a un experto sino que permite integrar conocimientos de diferentes expertos en un solo modelo MCD (Pérez-Teruel *et al.*, 2014). El proceso de agregación de las relaciones se estableció a partir de una función promedio de las matrices que representa el conocimiento causal de los expertos, tal como se muestra en la ecuación 2.3. Los pesos resultantes de la agregación para el modelado del riesgo en el

cultivo protegido del tomate fueron agrupados mediante una matriz de adyacencia ([White and Mazlack, 2011](#)). Esta matriz conformó las relaciones entre las variables del sistema a modelar y, a su vez, el modelo MCD inicial.

$$VA_{ij} = \frac{\sum_{i=1}^n w_{ij}}{E} \quad (2.3)$$

Donde:

$VA_{ij}$ : representa el valor agregado

E: Cantidad de expertos que participan en el proceso

$w_{ij}$ : Vector de correlación expresado por los expertos

## 2.5. Recopilación de datos históricos

La colecta de datos se desarrolló del 8 agosto del 2016 al 30 de marzo del 2017, la unidad empresarial de base de cultivos protegidos y semiprotegidos. Dicha UEB, cuenta con instalaciones (casas de cultivo) modelo Granma 1 de la empresa cubano española EMBA-MSA A-12, diseñado por la empresa cubano-española CARISOMBRA, de  $540 \text{ m}^2$  (12 m de ancho y 45 m de largo), con una altura a la cumbre de 4,4 m, cerramiento superior con rafia plastificada (Ginegar,  $140 \mu\text{m}$ , 80% luz difusa), ventana cenital abierta y malla sombreadora (35%) por los laterales y el frente. Tres casas de cultivo con modelo A-12 integran un módulo ([Mesa, 2019](#)).

En la tesis de maestría [Mesa \(2019\)](#), propuso un programa de riego teniendo en cuenta las características de las instalaciones, el sistema de fertirriego automatizado disponible y el cultivo. Se aplicó en un módulo estableciendo de frecuencia, turno y tiempo de riego, que permite una estabilidad en los valores de conductividad eléctrica de la solución nutritiva, de la solución lixiviada y de la fracción de lavado fijada, a lo largo del ciclo de cultivo. El programa diseñado ofreció, a partir de la fase inicio de cosecha hasta el final del ciclo una excelente respuesta agronómica de plantas de tomate en condiciones de producción. Por ello, algunas de los datos históricos de las variables utilizados en el modelado fueron obtenidos de la evaluación en las casas de cultivo de dicha investigación. Los datos de las variables fueron registrados de forma manual; y luego almacenados en bases de datos dispuestas para ello.

El diseño y planificación del riego está determinado por la estimación de la Etc, la cual se obtuvo bajo condiciones estándar utilizando el coeficiente de cultivo ( $Kc$ ) y la evapotranspiración de referencia ( $Eto$ ), mediante la ecuación propuesta por [Doorenbos and Pruitt \(1977\)](#) ( $Etc = Kc * Eto$ ) ([FAO, 2006](#)). El valores de  $Kc$  fueron establecidos según [Zamora et al. \(2014\)](#), empleando un valor de 0,61 inicial, 1,17 medio y 0,77 final (promedios de los valores referidos para tres cultivares de tomate). La estimación de la  $Eto$  se realizó mediante el modelo propuesto por [Jensen and Haise \(1963\)](#).

El modelo de [Jensen and Haise \(1963\)](#), mostrado en la ecuación 2.4, requiere datos de temperatura media, número de horas reales de sol, temperatura media mínima y máxima del periodo de experimentación. Todos estos datos fueron adquiridos de la estación Meteorológica 78343 del Ministerio de Ciencia Tecnología y Medio Ambiente (CITMA), ubicada en la Empresa Agropecuaria "Valle del Yabú". Además que dicho modelo necesitó de los valores de la radiación solar si no existiera atmósfera y el número máximo de horas de sol dado por [Allen et al. \(1998\)](#).

$$Eto = C_t(T - T_x) * Rs \quad (2.4)$$

Siendo:

$C_t$ : coeficiente de temperatura del aire constante para un área dada

T: temperatura media del día

$T_x$ : constante regional que representa la intercepción de la ecuación lineal con el eje de la temperatura

Rs: Radiación solar incidente sobre la superficie, dado por la ecuación 2.5

$$Rs = Ro(0,18 + 0,55 \frac{n}{N}) \quad (2.5)$$

Siendo:

Ro: radiación solar si no existiera atmósfera

n: número de horas de sol reales

N: número máximo de horas de sol

Para el cálculo de la Eto se necesita obtener los valores de los coeficientes  $C_t$  y  $T_x$  dados por las ecuaciones 2.6 y 2.7 respectivamente.

$$C_t = \frac{1}{38 - \frac{h}{152,5} + \frac{380}{e_2 - e_1}} \quad (2.6)$$

$$T_x = -2,5 - 0,14(e_2 - e_1) - \frac{h}{550} \quad (2.7)$$

Estos a su vez requieren de:

$e_2$ : Presión de vapor a saturación (*mbar*) correspondiente a la temperatura media máxima ( $T_{max}$ , °C)(ecuación 2.8)

$e_1$ : Presión de vapor a saturación (*mbar*) correspondiente a la temperatura media mínima ( $T_{min}$ , °C) (ecuación 2.9)

$$e_2 = 6,108e^{\frac{17,27T_{max}}{T_{max}+237,3}} \quad (2.8)$$

$$e_1 = 6,108e^{\frac{17,27T_{max}}{T_{min}+237,3}} \quad (2.9)$$

Con la aplicación del programa de riego en el cultivo protegido de tomate desde el primer día después del trasplante (ddt) hasta los 234 ddt. La cantidad de datos por cada variable con un período de muestreo cada 3 días, se muestra en el cuadro 2-4 según las fases fenológicas del tomate (tabla 1-2). De este conjunto de datos solo se utilizan los de inicio de cosecha a fin de producción (etapa reproductiva) para lograr la estimación de variables de impacto productivo.

Tabla 2-4: División del conjunto de datos para las fases fenológicas del tomate

Fases	Total de datos	Datos para el ajuste	Datos para la validación
I, II y II	35	25	10
IV y V	43	30	13

Inicialmente se normalizan los datos, con el objetivo de reducir a una misma escala todas las variables del modelo. El rango seleccionado estuvo de [0, 1] por ser el rango de activación que poseen los conceptos del modelo y, además, es el rango de evaluar en la función de activación. Este proceso se llevó a cabo mediante la ecuación 2.10:

$$X_{normalizado} = \frac{X - X_{min}(d_2 - d_1)}{X_{max} - X_{min}} + d_1 \quad (2.10)$$

Donde:

X: valor a normalizar

$X_{min}$  y  $X_{max}$ : rango de óptimo de X

$d_1$  y  $d_2$ : rango de normalización de X (0,1)

## 2.6. Evaluación, ajuste y validación del modelo

Para la evaluación del modelo inicial, su ajuste y validación ha sido necesario utilizar una biblioteca de *software*. La herramienta empleada fue diseñada por Sosa *et al.* (2019), la cual consiste en un paquete para Python que abarca la teoría de mapas cognitivos difusos en cuatro clases fundamentales: FCM, Concept, Boundary e Input. Esta biblioteca ofreció como ventaja principal que su lenguaje de programación es multiplataforma y multiparadigma (Lynch, 2014); además, de ser *software* libre y de código abierto. La versión de Python utilizada fue la 3.6.5 por ser la más actualizada. Debido a su característica de ser multiplataforma, se eligió la suite Anaconda como gestor de paquetes para el desarrollo de esta investigación ya que recoge una serie de aplicaciones, bibliotecas y conceptos diseñados

para el desarrollo de la Ciencia de datos con Python (Kadiyala and Kumar, 2018), y que se encuentra disponible para ejecutarse en *Microsoft Windows*. Además, de que integra varios IDEs de desarrollo como Spyder, el cual fue seleccionado debido a que permite la depuración y la comprobación de los códigos generados a través de una terminal de IPython.

La clase FCM representa el mapa cognitivo difuso, el cual posee como atributos una lista de conceptos, instancias de la clase *Concept*, y una lista de entradas, instancias de la clase *Input*. La clase *Boundary* solo presenta métodos estáticos, y por lo tanto, solo se encarga de proporcionar las funciones de acotación y sus derivadas al mapa, durante los procesos de inferencia y de aprendizaje.

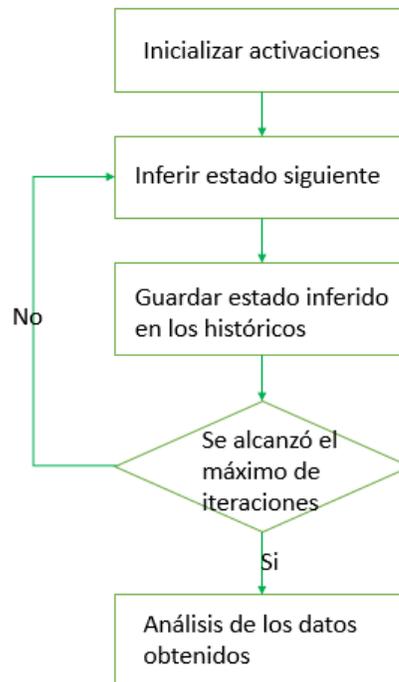


Figura 2–2: Diagrama de flujo del proceso de inferencia

El paquete proporciona dos funcionalidades principales: inferencia y aprendizaje. En el proceso de inferencia, una vez inicializado el mapa, este método calcula y asigna los valores de activación de los distintos conceptos, según la ecuación 1.1, guardando en el historial los valores de cada uno. Mediante esta función realizó la predicción del comportamiento del sistema en el tiempo, conocidas las condiciones iniciales del mismo, así como sus dinámicas internas, representadas por las matrices causales. En la figura 2–2 se muestra el flujo del proceso de inferencia del mapa.

El proceso de aprendizaje se realizó con los datos históricos del comportamiento del sistema. Este método representa un algoritmo iterativo que en cada paso ajusta las conexiones sinápticas del mapa con el objetivo de disminuir cada vez más una función de error del

modo que pueda semejarse lo más posible a los datos medidos. En la figura 2-3 se muestra el flujo del proceso de aprendizaje del mapa.

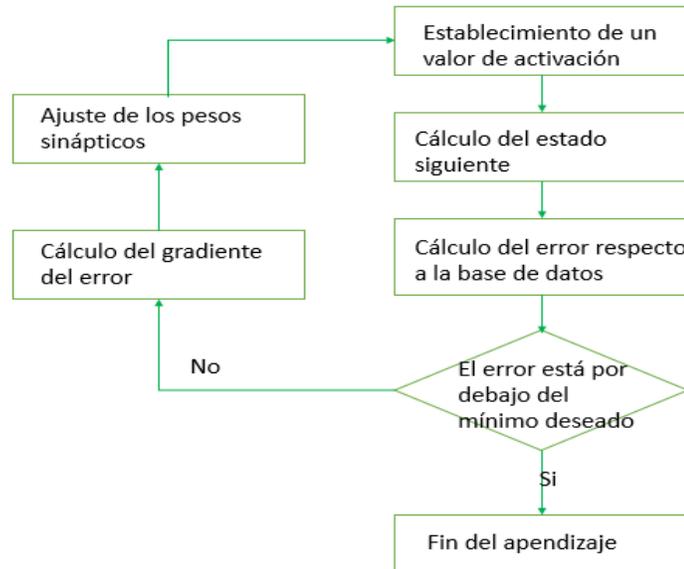


Figura 2-3: Diagrama de flujo del proceso de aprendizaje

## 2.7. Análisis del modelo

Como etapa cumbre de la investigación, los expertos apreciaron las respuestas del modelo obtenido con respecto al conjunto de datos de referencia; y de ello valoraron si los resultados son o no aceptables para un modelo de riesgo de *S. lycopersicum* en casas de cultivo. El criterio de los expertos también estuvo apoyado con los resultados obtenidos del error cuadrático medio ofrecido por el modelo.

### 2.7.1. Análisis estadístico

Se analizó el modelo estadísticamente en el IBM SPSS Statistics 23 a través de pruebas de correlación entre los conceptos y la dispersión de puntos. La correlación determinó el grado de asociación entre los conceptos del modelo propuesto y el conjunto de datos. La dispersión de puntos permitió que en torno a una línea recta se compruebe el sentido de la relación de los conceptos, en términos de aumento o disminución.

Existen tres tipos de correlaciones W de Kendall, Pearson y Spearman (Puth *et al.*, 2015). Es importante resaltar que la correlación no implica la existencia de causalidad, sino que existe alguna asociación con una variable intermedia entre ambas (Garmendia and Monserrat, 2018). Antes de calcular el coeficiente de correlación se realizó la prueba de normalidad de los datos; debido a que la correlación de W de Kendall y Spearman son no paramétrica, es decir, se utilizan en variables que no presentan una distribución normal (Puth *et al.*, 2015). Mientras que la correlación de Pearson, determina el grado de asociación lineal entre las variables cuantitativas que posean normalidad (Dagnino, 2014).

Se aplicó la prueba Shapiro-Wilk de normalidad de los datos para un nivel de significación de 0.05(nivel de confianza de 95 %). Esta prueba considera como hipótesis nula,  $H_0$ , que la distribución de las variables provienen de una distribución normal. La selección de este test estuvo condicionada a que el tamaño de muestra de los datos fue menor de 30. Para la prueba de Shapiro-Wilk si la significación ( $p - valor$ )  $> 0,05$  aceptamos  $H_0$  y la variable presenta una distribución normal, de lo contrario, si ( $p - valor$ )  $< 0,05$  posee una distribución no normal.

La correlación de Pearson determina si están relacionadas o no las variables a través de la regresión lineal y, de estarlo, una medida de la fuerza con que están relacionadas es el coeficiente de correlación lineal ( $r$ ). En esencia,  $r$  mide la dispersión de los puntos en torno a una tendencia lineal (Dagnino, 2014).

Dagnino (2014), define como coeficiente de correlación lineal simple o coeficiente de correlación de Pearson a la raíz cuadrada de la siguiente expresión:

$$r = \frac{XY}{\sqrt{X^2Y^2}} \quad (2.11)$$

Donde: X, Y representan variables.

El coeficiente de correlación de Spearman ( $\rho$ ) es una medida de asociación lineal que utiliza los rangos, números de orden, de cada grupo de sujetos y compara dichos rangos (Torres *et al.*, 2017). El coeficiente de correlación de Spearman viene dado por la expresión:

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)} \quad (2.12)$$

D: es la diferencia entre los rangos de las variables X-Y.

N: es el número de parejas de datos.

La escala de interpretación de los coeficientes de correlación de Pearson y Spearman es la misma, en este caso se utilizó la realizada por Garmendia and Monserrat (2018) que indica:

1. Buena positivamente:  $0,6 < r$  ó  $\rho \leq 1$
2. Débil positivamente:  $0,4 < r$  ó  $\rho \leq 0,6$
3. Mala:  $0 < r$  ó  $\rho \leq 0,4$
4. Débil negativamente:  $-0,4 > r$  ó  $\rho \geq -0,6$
5. Buena negativamente:  $-0,6 > r$  ó  $\rho \geq -1$

## 2.8. Consideraciones finales del capítulo

La metodología propuesta basado en la ingeniería del conocimiento y los mapas cognitivos difusos se presenta como una herramienta para la obtención del modelo de riego en el cultivo protegido de *S. lycopersicum*.

La metodología de la ingeniería del conocimiento se utiliza como un instrumento en la determinación y confección del modelo inicial del riego en el cultivo protegido de *S. lycopersicum*.

La utilización de plataformas libres y de código abierto en el desarrollo de la biblioteca de *software* MCD, convierten a la misma en un proyecto portable y fácilmente actualizable.

# CAPÍTULO 3

## RESULTADOS Y DISCUSIÓN

### 3.1. Introducción

Inicialmente se determinan las variables y se confecciona el modelo inicial aplicando la ingeniería del conocimiento. Luego, se realiza la evaluación del modelo ofrecido por los expertos. De dicha etapa se decide ajustar el modelo mediante el proceso de aprendizaje, como se refirió en el sexto paso de la metodología propuesta. También se valida el modelo ajustado para comprobar los resultados obtenidos. Por último, se valora el modelo obtenido por los ingenieros del conocimiento; y se realiza un análisis estadístico para comparar el modelo inicial brindado por los expertos y el final obtenido por la biblioteca de *software*. Como resultado se establece el diseño del modelo de riego para ayudar a los expertos en la toma de decisiones en la producción de *S. lycopersicum* en cultivo protegido.

### 3.2. Determinación de variables aplicando la ingeniería del conocimiento

#### 3.2.1. Selección de los expertos

Para la selección de los expertos se estableció como criterio de selección un  $k \geq 0,8$  pues, según [Feria et al. \(2018\)](#), representa un coeficiente de competencia del experto alto. Los resultados muestran en la figura 3-1, que solo nueve expertos cumplieron con el criterio de conocimiento establecido. Según lo que plantea [Zambrano et al. \(2015\)](#), a partir de diez candidatos en adelante se reduce el porcentaje de errores de los resultados a alcanzar, por lo que se demostró que la cantidad de expertos seleccionados fue correcta.

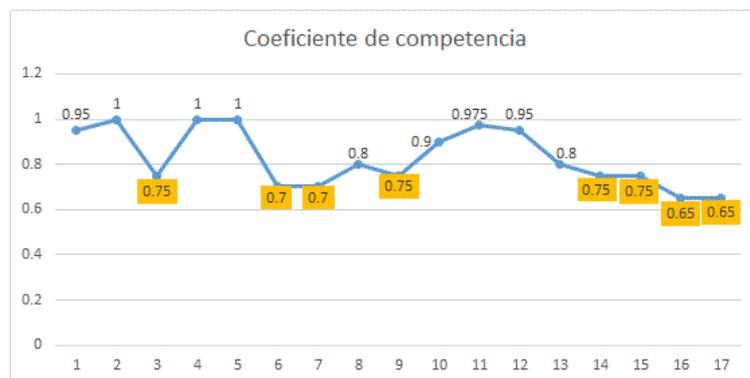


Figura 3-1: Coeficiente de competencia de los candidatos

### 3.2.2. Elección de la metodología

La aplicación de la metodología Delphi donde a través de las encuestas se definieron las principales variables involucradas en el riego del tomate en cultivo protegido. Aunque, esta metodología tuvo como limitante que se consumió aproximadamente tres meses para que se completara el intercambio de los documentos de las ronda de encuestas, esto afectó la motivación y fluidez de la investigación. Esta sentencia se encuentra detallada en la revisiones realizadas como una de las principales desventajas del la metodología en cuestión, pues según las investigaciones de Varela-Ruiz *et al.* (2012) y Powell (2003) el proceso se extiende a un promedio de 45 días.

### 3.2.3. Procesamiento de la información

#### Primera encuesta

En los resultados de la primera encuesta que se muestran en la tabla 3-1, la prueba de W de Kendall destacó una correlación significativa entre los criterios de los expertos, con un grado de coincidencia entre las respuestas de 0.259. Este resultado implica en la escala de Landis and Koch (1977) un acuerdo justo entre el juicio emitido por cada experto. Un valor justo del coeficiente W de Kendall puede interpretarse como un reflejo de que los nueve expertos están aplicando los mismos estándares al asignar rangos a las preguntas.

Tabla 3-1: Prueba de W de Kendall de la primera encuesta

Expertos	W de Kendall	Sig. asintótica
9	0.259	0.003

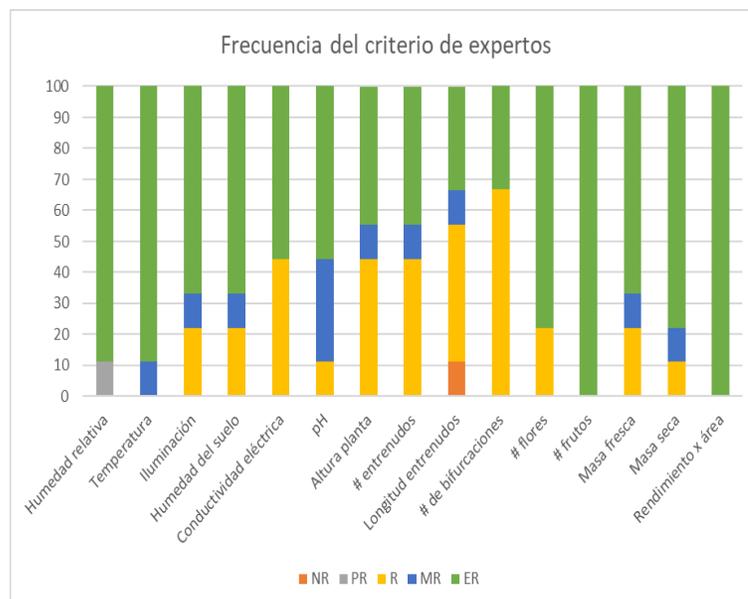


Figura 3-2: Frecuencia de los expertos en la primera encuesta

El gráfico 3-2 expone que de las 19 variables, las que fueron catalogadas de relevantes por los expertos. De la tabla de frecuencia de esta figura se tradujo de extremadamente

relevantes en el cultivo protegido de hortalizas con más del 50 % de los expertos: los factores climáticos, la humedad relativa (C1), temperatura (C2), iluminación (C3), humedad del suelo (C4); en cuanto a las condiciones del suelo, la conductividad eléctrica (C5) y pH (C6); por parte de la fenología de las hortalizas, el número de flores y frutos (C10 y C11), masa fresca (C12), masa seca (C13), rendimiento (C14). Mientras que, la altura de la planta (C7), número y longitud de los entrenudos (C8 y C9), aunque no cumplieron la condición anterior, los porcentajes se encontraron dentro de los niveles de relevancia. Otro aspecto importante a destacar, dentro del análisis de la encuesta, se manifestó en el número de frutos y el rendimiento que son variables que los expertos señalaron con un 100 % de extrema relevancia por ser fundamentales en un proceso productivo.

### Segunda encuesta

La prueba de W de Kendall igualmente presentó una correlación significativa entre los criterios de los expertos, con un grado de coincidencia entre las respuestas de 0.630. Este resultado implica en la escala de [Landis and Koch \(1977\)](#) un acuerdo sustancial entre el juicio emitido por cada experto. Este valor reflejó una mayor coincidencia de criterios entre los nueve expertos.

Tabla 3-2: Prueba de W de Kendall de la segunda encuesta

Expertos	W de Kendall	Sig. asintótica
9	0.630	0.000

La segunda encuesta estuvo dirigida hacia las variables involucradas en el riego del cultivo protegido del tomate. Los resultados de la tabla de frecuencia se exponen en la figura 3-3. En la cual, se produjo el mantenimiento de las variables de conductividad eléctrica (55,6 %) y rendimiento (66,7 %). El 100 % de los expertos eliminó la altura de la planta y el número de bifurcaciones. Los nueve expertos sustituyeron la humedad relativa, temperatura, iluminación, humedad del suelo y el pH, en su defecto, agregaron la evapotranspiración. Con más del 50 % del criterio de los expertos, se agregó la fracción de lavado, demanda hídrica, volumen, frecuencia, turno y tiempo de riego; por las variables número de entrenudos, de flores y frutos, longitud de entrenudos, masa seca y fresca. Debido a que adquieren mayor relevancia para la planificación del riego en el cultivo protegido del tomate ([Mesa, 2019](#)).

La Etc se adicionó porque causa el movimiento de agua, nutrientes y minerales desde las raíces hasta el resto de los órganos de la planta y juega un papel importante en el crecimiento y la productividad. Depende de las condiciones climáticas, por tanto, se omiten las variables climáticas (humedad relativa, temperatura e iluminación) y las características del suelo (humedad del suelo y pH) que se expusieron en el primer cuestionario. Además la Etc depende del tipo de cultivo, la calidad del agua, el sistema de cultivo y las prácticas agrícolas ([Dorais et al., 2016](#); [Antúnez and Felmer, 2017](#); [Qiu et al., 2017](#)). Comparados

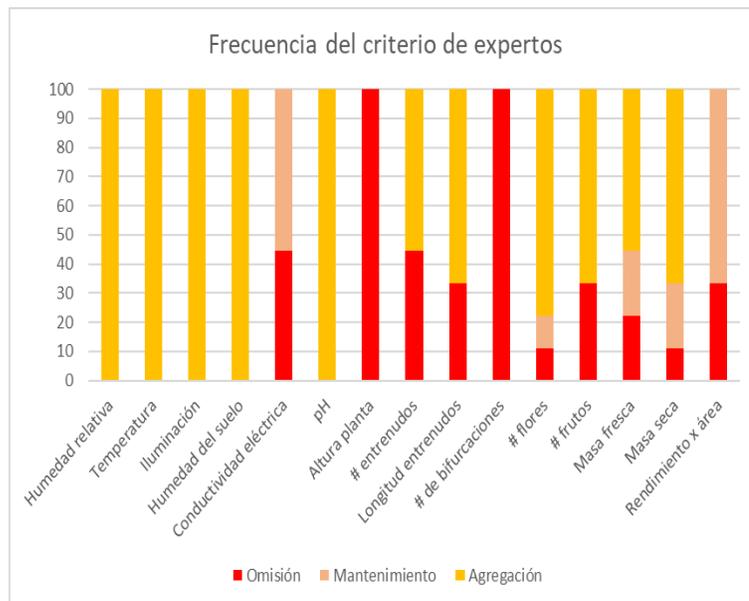


Figura 3-3: Frecuencia de los expertos en la segunda encuesta

con los cultivos que crecen a campo abierto, la Etc en los cultivos protegidos es considerablemente menor (León and Cun, 2001; León *et al.*, 2005; Fernández *et al.*, 2010).

Se incluyó la fracción de lavado ya que depende de diferentes factores que intervienen en el balance de agua del cultivo (entre los que se mencionan el manejo del riego, la calidad del agua, el clima, el suelo, la evapotranspiración) y de la sensibilidad del cultivo a la salinidad (Letey *et al.*, 2011). También son variables determinantes en el riego la frecuencia, tiempo y turnos de riego pues dependen de la calidad del agua, del clima, del suelo y de la sensibilidad del cultivo a la salinidad.

La demanda hídrica constituyó una variable esencial en el riego del cultivo, pues guarda relación con las condiciones ambientales y la fase fenológica en que se encuentre. Además, es importante la determinación correcta del volumen de agua que se requiere aplicar según el tipo de suelo que se dispone y el sistema radical de las plantas para reducir las pérdidas de agua y no afectar los rendimientos (FAO, 2013).

El establecimiento del volumen de riego como concepto se deriva principalmente de la evapotranspiración del cultivo, la tensión del agua en el suelo, tipo de suelo (capacidad de campo, porcentaje de saturación), eficacia de riego (uniformidad de caudal de los emisores) y calidad del agua de riego (a peor calidad, mayores son los volúmenes de agua, ya que es necesario desplazar el frente de sales del bulbo de humedad)(Mollinedo and Tapia, 2008).

La conductividad eléctrica referida como extremadamente relevante es importante para una óptima producción. Sin embargo, estuvo representada por la conductividad eléctrica de la solución nutritiva (Cesn) y de la solución lixiviada (Cesl). La primera porque indica el potencial osmótico y está determinado por la concentración de solutos o sustancias osmóticamente activas para la planta, mientras que la segunda brinda una información

rápida de lo que acontece con la solución nutritiva una vez que interactúa con el suelo (Hernández Díaz *et al.*, 2014; Sonneveld and Voogt, 2009).

El rendimiento acumulado representó el número de frutos cosechados en cada período de muestreo, lo cual determina los índices productivo en el modelo de riego en cultivo protegido de tomate. Este depende principalmente de factores biológicos y de condiciones ambientales (dEsposito *et al.*, 2017; Raza *et al.*, 2017). Entre estas últimas la disponibilidad de agua y nutrientes son fundamentales y tienen influencia en las fases vegetativa y reproductiva. Por tanto, una planificación efectiva del riego con tiempos específicos y estrategias de fertilización pueden incrementar significativamente los rendimientos (Hernández Díaz *et al.*, 2014; Wang *et al.*, 2018).

Con la aplicación de la segunda encuesta el total de variables se redujo a diez, expuestas en la tabla 3-3, con sus rangos de trabajo y unidad de medida que fueron brindados por los expertos.

Tabla 3-3: Rango de trabajo de las entradas y los conceptos para el modelado de riego

Variables	Nomenclatura	Rango	Unidad de medida	Clasificación
Fracción de lavado	Fl	0-20	%	Entrada
Demanda hídrica	Dac	0-3/3.8	L/planta/día	Entrada
Evapotranspiración	Etc	0-3	mm/día	Entrada
Frecuencia de riego	Fr	1-3	días de intervalo	Entrada
Cantidad de turnos de riego	Ctr	1-4	turnos	Entrada
Tiempo de riego	Tr	10-45	min	Entrada
Volumen de agua de riego	Vr	0-8668.8	L/día	Concepto
Conductividad eléctrica de la solución nutritiva	CEsn	0-3	dS/m	Concepto
Conductividad eléctrica de la solución lixiviada	Cesl	0-3.2	dS/m	Concepto
Rendimiento acumulado	Ra	0-11810	Kg	Concepto

### 3.3. Confección del modelo inicial

Los resultados de la agregación de los pesos causales de cada experto, se exponen en la tabla 3-4. De la confección del modelo inicial, se puede apreciar que para el volumen de riego fueron influyentes la demanda hídrica y la Etc. Este criterio de los expertos se encontró, posiblemente dado por lo referido en la literatura científica, que al aumentar la demanda de agua de los cultivos para cubrir las pérdidas por Etc, se incrementa el volumen de riego (FAO, 2006).

Tabla 3–4: Relaciones causales entre las variables

Relación	Vr	Cesl	Cesn	Ra
Fl	0.5	0.25	0	0.75
Dac	0.75	0.5	0.25	0.25
Etc	0.75	0.5	0.25	0.25
Fr	-0.5	0.75	0.25	0.25
Ctr	0.25	-0.75	-0.25	0.25
Tr	0	-0.75	0.25	0.25
Vr	0	0.25	-0.75	0.25
CEsl	0	0	0	0.75
Cesn	0.25	0.5	0	0.75
Ra	0	0	0	0

La frecuencia de riego insidió en el volumen de riego negativamente, este resultado se encuentra sustentado en la investigación de (Mesa, 2019) ya que demostró que la aplicación de un programa de riego con un determinado volumen de riego y un aumento de la frecuencia produjo incrementos en los rendimientos.

En el concepto perteneciente a la conductividad eléctrica de la solución lixiviada, la cantidad de turnos y el tiempo de riego influye negativamente fuerte. De la conductividad eléctrica nutritiva solo se destacó la influencia negativamente fuerte del volumen de riego. Ambas sentencias, se producen debido a que para contribuir a una mejor planificación del riego que cubra la demanda de agua del cultivo, es preciso lavar las sales que pueden acumularse en el suelo a consecuencia de la salinidad del agua o el aporte de fertilizantes con la solución nutritiva en el fertirriego y minimice el impacto negativo de las pérdidas por la eficiencia del sistema (Mesa, 2019).

El rendimiento acumulado representó la variable objetivo del proceso productivo del modelo. Esta importante variable no influye sobre ningún otro concepto porque constituyó el resultado final del modelo como se puede observar en la tabla 3–4. A su vez, sobre él inside la fracción de lavado y las conductividades eléctricas de la solución lixiviada y nutritiva positivamente fuerte. Esta afirmación sucede porque para determinar la dosis de riego necesaria para lavar las sales en el suelo a los niveles adecuados se utiliza la fracción de lavado mínima para mantener la salinidad del suelo por debajo del umbral de sensibilidad del cultivo y que no se afecte el rendimiento (Ayer and Westcot, 1985; FAO, 2006; Letey *et al.*, 2011).

Con los pesos agregados se representan las matrices de adyacencias 3.1 y 3.2 correspondiente a la ecuación de activación 1.2. La representación del modelo inicial a través de la ingeniería del conocimiento se muestra en la figura 3–4.

$$W^{inicial} = \begin{pmatrix} 0 & 0,25 & -0,75 & 0,25 \\ 0 & 0 & 0 & 0,75 \\ 0,25 & 0,5 & 0 & 0,75 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.1)$$

$$H^{inicial} = \begin{pmatrix} 0,5 & 0,75 & 0,75 & -0,5 & 0,25 & 0 \\ 0,25 & 0,5 & 0,5 & 0,75 & -0,75 & -0,75 \\ 0 & 0,25 & 0,25 & 0,25 & -0,25 & 0,25 \\ 0,75 & 0,25 & 0,25 & 0,25 & 0,25 & 0,25 \end{pmatrix} \quad (3.2)$$

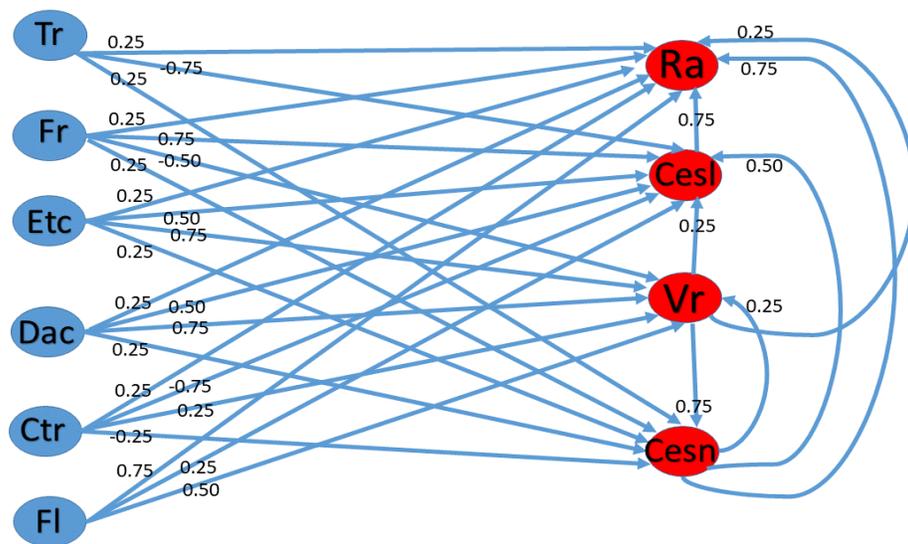
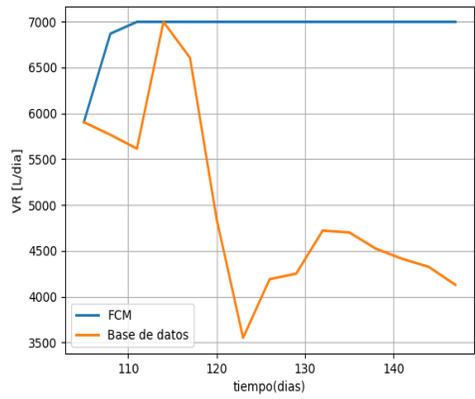


Figura 3-4: Representación del mapa cognitivo difuso para el modelado del riesgo del tomate en cultivo protegido, según la ingeniería del conocimiento

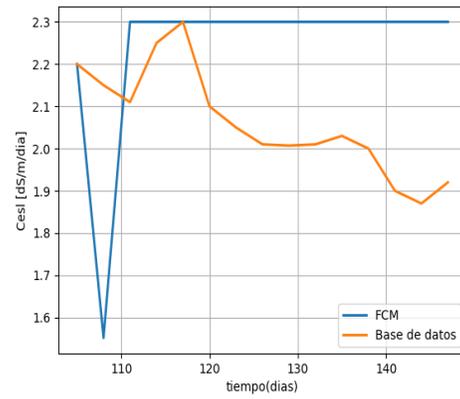
### 3.4. Evaluación del modelo inicial

La evaluación del modelo inicial se muestra en las figuras 3-5 y 3-6. Dicho etapa, se encuentra representada por las respuestas del proceso de inferencia, descrito en 2-2, de los cuatro conceptos con respecto al 30% de las mediciones reales almacenadas en la base de datos. Los resultados se representan de forma desnormalizada para facilitar la interpretación en sus medidas reales. Como se puede observar, el modelo presentó una mala estimación en términos de seguimiento de los valores de activación de los conceptos del conjunto de evaluación.

El error cuadrático medio del modelo fue de 5,27%, como se aprecia en la figura 3-7, proporcionando una exactitud media de 94,73%. Según los expertos, estos resultados no ofrecieron un modelo MCD correcto en correspondencia con el conjunto de datos reales adquiridos.

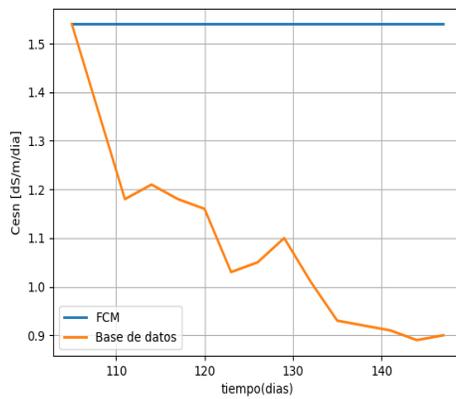


(a) Volumen de riego

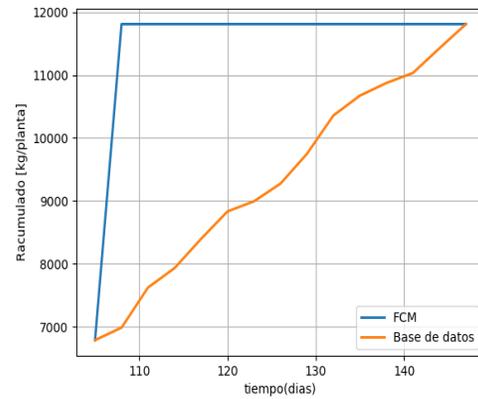


(b) Conductividad eléctrica de la solución lixiviada

Figura 3-5: Representación de los conceptos, según la ingeniería del conocimiento



(a) Conductividad eléctrica de la solución nutritiva



(b) Rendimiento acumulado

Figura 3-6: Representación de los conceptos, según la ingeniería del conocimiento

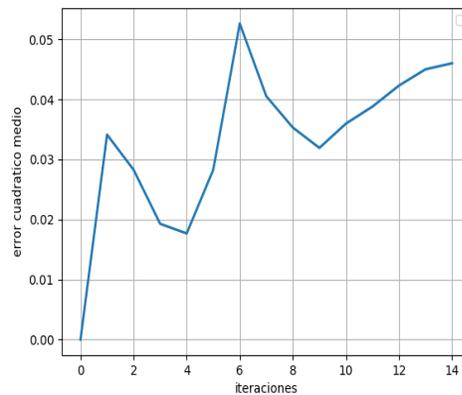
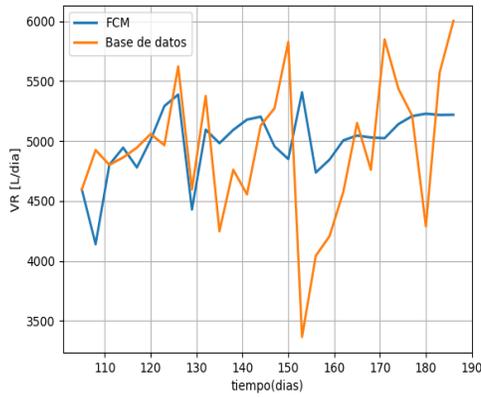


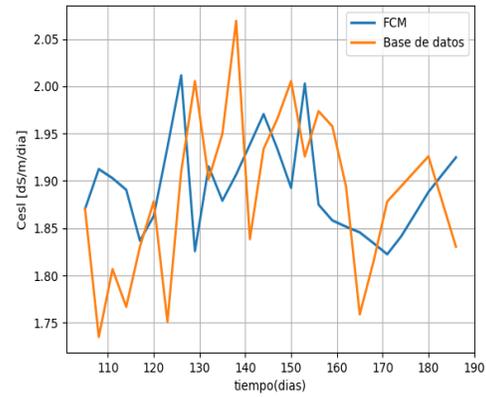
Figura 3-7: Error cuadrático medio del modelo inicial

### 3.5. Ajuste del modelo

El proceso de aprendizaje descrito en la figura 2–3 se produjo durante 400 iteraciones, buscando el error máximo cada diez pasos iterativos con un factor de aprendizaje de 0,1.

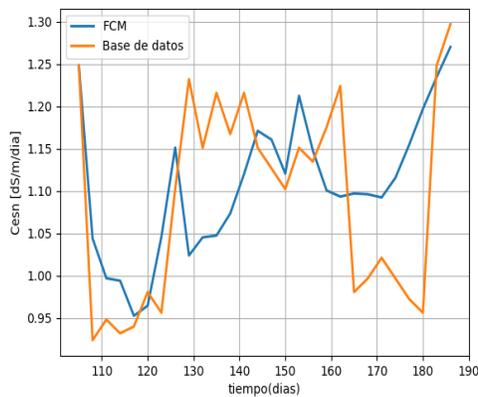


(a) Volumen de riego

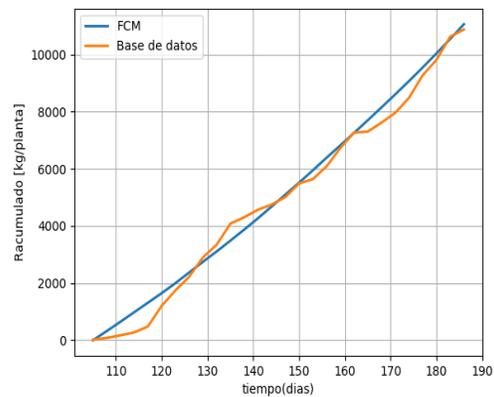


(b) Conductividad eléctrica de la solución lixiviada

Figura 3–8: Representación de los conceptos, resultado del proceso de aprendizaje



(a) Conductividad eléctrica de la solución nutritiva



(b) Rendimiento acumulado

Figura 3–9: Representación de los conceptos resultado del proceso de aprendizaje

La respuesta del error cuadrático medio se muestra en la figura 3–10, donde cumplió con el objetivo del algoritmo del gradiente descendente, de disminuir iterativamente el error cuadrático medio entre el modelo respecto a los datos de referencia. Al mismo tiempo, se apreció la rápida convergencia que presentó este método. Así, como resultado de este proceso de aprendizaje se obtuvieron nuevamente las gráficas 3–8 y 3–9, que representan las respuestas desnormalizadas del modelo con respecto a los datos históricos de referencia.

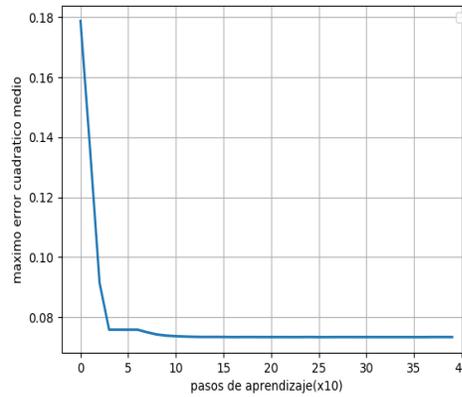
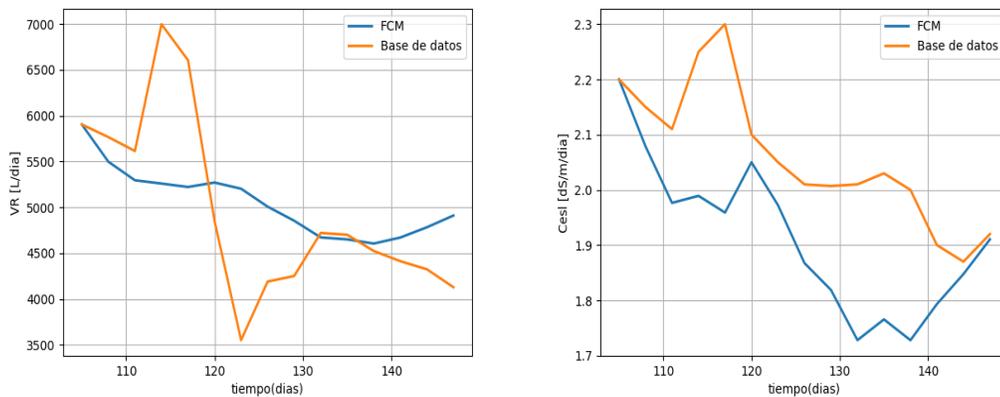


Figura 3-10: Evolución del error en el proceso de aprendizaje

### 3.6. Validación del modelo

La etapa de validación del modelo se muestran en las gráficas desnormalizadas 3-11 y 3-12, en ellas se encuentran las respuestas del modelo ajustado con respecto a los restantes 30 % de los datos adquiridos. Si se comparan estos resultados con los obtenidos del modelo inicial de la figura 3-5 y 3-6, se aprecia una mejor estimación que el modelo inicial obtenido por la ingeniería del conocimiento.

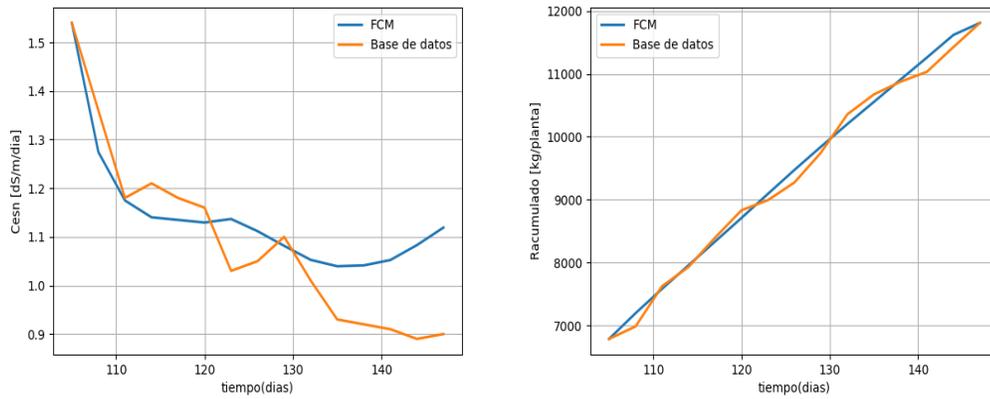


(a) Volumen de riego

(b) Conductividad eléctrica de la solución lixiviada

Figura 3-11: Representación de los conceptos en la etapa de validación

El error cuadrático medio corroboró los excelentes resultados del modelo obtenido, pues un error de 5,27 % disminuyó a 0.95 %, como se presenta en la figura 3-13. Esto trae consigo, una exactitud media de 99.05 %, superando en gran medida a los resultados obtenidos en el proceso de inferencia a partir del modelo inicial brindado por la ingeniería del conocimiento (figura 3-7).



(a) Conductividad eléctrica de la solución nutritiva

(b) Rendimiento acumulado

Figura 3-12: Representación de los conceptos en la etapa de validación

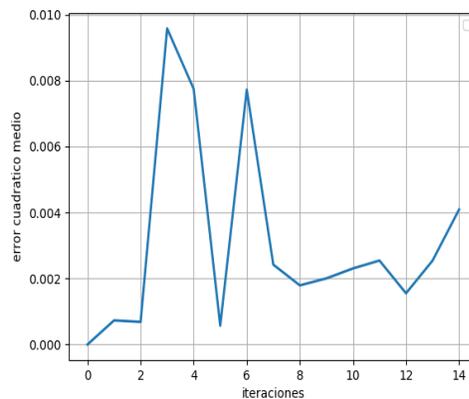


Figura 3-13: Error cuadrático medio del modelo final

### 3.7. Análisis del modelo

Los expertos valoraron que las respuestas del modelo obtenido con respecto al conjunto de datos de referencia cumple con los objetivos propuestos. También expresaron que en él, pudiera ocurrir mejoras, pero en general los resultados del modelo de riego de *S. lycopersicum* en casas de cultivo lo definieron como aceptable. Se planteó que, para ser el primer modelado de naturaleza compleja, no lineal y de dinámicas con un nivel de dificultad alto; se consiguió representar un modelo lineal aproximado. A pesar, de que la base de datos utilizada presentó poca cantidad de datos para las pruebas y simulaciones, fue un gran paso de avance haber obtenido dichos resultados.

En cuanto, a los resultados de la conductividad eléctrica de la solución lixiviada, los expertos opinaron que se mantuvo menor al valor umbral de conductividad eléctrica referido para el cultivo ( $2,5 \text{ dS/m}$ ), como en el caso de la investigación de Mesa (2019), en las fases desde el desarrollo de los frutos hasta la cosecha.

Según los expertos, se cumplió fielmente el criterio de que  $CE_{sl} > CE_{sn}$  (Mesa, 2019). Esto permite que el suministro de nutrientes sea el adecuado, limita el riesgo de salinidad del suelo que puede afectar la disponibilidad de agua y nutrientes para la planta y posibilita hacer las correcciones necesarias (Hernández Díaz *et al.*, 2014).

En relación al volumen de riego valoraron que aunque no siguió una buena estimación con respecto a los datos medidos, posiblemente este resultado estuvo influenciado por la poca cantidad de datos en las simulaciones realizadas.

Los expertos argumentaron que después de 234 ddt el rendimiento acumulado no mostró decrecimiento, lo que ofreció resultados similares a los obtenidos por (Mesa, 2019), lo cual indicó las potencialidades del modelo de estimar las variables de impacto productivo.

Fueron aceptables los resultados del modelo de riego obtenido en cada uno de los conceptos, según el grupo de expertos, y aseguraron que se cumplió el objetivo de estimar las variables de impacto productivo, para ayudar en la toma de decisiones en el interior de las casas de cultivo.

#### 3.7.1. Análisis estadístico

En la prueba de Shapiro Wilk, que se muestra en la tabla 3-5, los datos históricos de los cuatro conceptos presentaron un  $p - \text{valor} > 0,05$ , por tanto, se aceptó la hipótesis nula porque presentaron normalidad las cuatro variables. Similares resultados se obtuvieron con los datos inferidos del modelo MCD obtenido, en cuanto a la normalidad de sus variables; excepto en el caso de la conductividad eléctrica de la solución nutritiva, que  $p - \text{valor} < 0,05$  es decir, este concepto no presentó una distribución normal. Mientras que, el modelo inicial rechazó la  $H_0$  de sus variables porque el  $p - \text{valor} < 0,05$ . Del análisis de normalidad de los conceptos se pudo concluir que para los datos históricos y el

modelo MCD final se utilizó la correlación de Pearson. Para el caso del modelo inicial y la Cesn del modelo final fue necesario recurrir a las pruebas no paramétricas, específicamente la correlación de Spearman. Para corroborar las pruebas de normalidad, se aprecia en el anexo B.1, los histogramas de las variables que tuvieron una distribución normal.

Tabla 3–5: Prueba de normalidad de Shapiro Wilk

Conceptos	Estadístico	No muestra	Sig.
Vr	0.913	15	0.151
Cesl	0.96	15	0.695
Cesn	0.899	15	0.091
Ra	0.954	15	0.596
Vr-inicial	0.33	15	0
Cesl-inicial	0.335	15	0
Cesn-inicial	-	-	-
Ra-inicial	0.284	15	0
Vr-final	0.922	15	0.204
Cesl-final	0.959	15	0.683
Cesn-final	0.7	15	0
Ra-final	0.957	15	0.649

La correlación de Pearson (tabla 3–6) entre las mediciones reales y el modelo obtenido de los conceptos Vr, Cesl y Ra se encuentran en sentido positivo y bien relacionados. Estos resultados se encontraron respaldados con un coeficiente entre  $0,6 < r \leq 1$ , según Garmendia and Monserrat (2018), de forma significativa a 0.05 para el volumen de riego y la conductividad eléctrica de la solución lixiviada; y a 0.01 para el rendimiento acumulado.

Tabla 3–6: Correlación de Pearson

Correlaciones	Correlación de Pearson	Sig.(bilateral)	N
Vr-Vrfinal	0,558*	0.03	15
Cesl-Ceslfinal	0,635*	0.011	15
Ra-Rafinal	0,997**	0.00	15

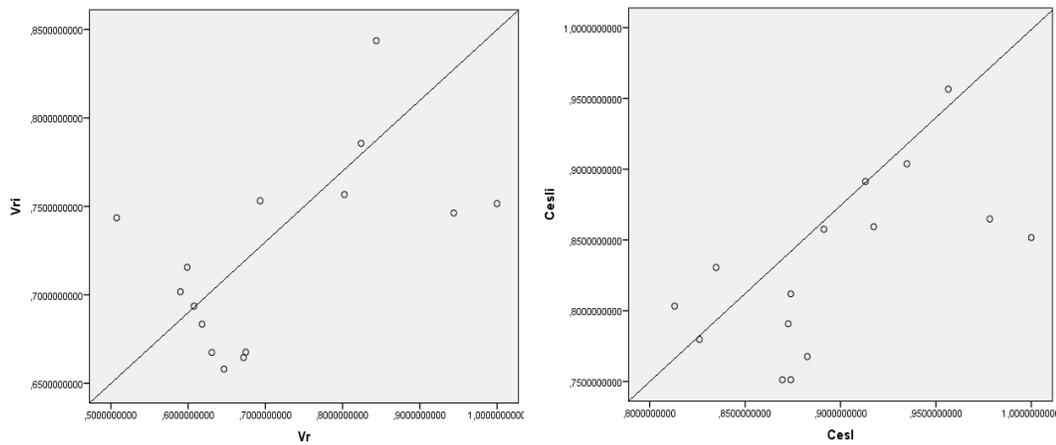
La conductividad eléctrica de solución nutritiva, a través de la correlación de Spearman, logró una relación positiva y buena con respecto a los conjunto de datos medidos a un nivel de significación de 0.01, como se aprecia en la tabla 3–7. En dicha tabla, también se puede apreciar que el modelo inicial tuvo una correlación débil en ambos sentidos con los datos históricos, pues el coeficiente osciló entre los intervalos de  $0,4 < \rho \leq 0,6$  y  $-0,4 > \rho \geq -0,6$ , según Garmendia and Monserrat (2018). Por tanto, el modelo final presentó mejor asociación conjunta entre las mediciones reales que el modelo ofrecido por los expertos. Por ello, el modelo MCD obtenido realizó una aceptable estimación con respecto a las mediciones registradas en el proceso productivo del cultivo de *S. lycopersicum*.

En las figuras 3–14 y 3–15 se aprecian la dispersión de los puntos del modelo MCD obtenido entre los datos históricos, porque presentó los mejores valores de correlación. El volumen de

Tabla 3–7: Correlación de Spearman

Correlaciones	Correlación de Spearman	Sig.(bilateral)	N
Cesn-Cesnfinal	0,777**	0.001	15
Vr-Vrinicial	-0.411	0.129	15
Cesl-Ceslinicial	-0.405	0.134	15
Cesn-Cesninitial	-	-	-
Ra-Rainicial	0.433	0.107	15

riego y la conductividad eléctrica de la solución lixiviada expusieron una mayor dispersión de los puntos. Aunque se corroboró que el volumen de riego obtuvo peor correlación que la conductividad eléctrica de la solución lixiviada, porque el coeficiente  $r < 0,6$ . Para el caso de la conductividad eléctrica de la solución nutritiva y el rendimiento acumulando los puntos se encontraron más próximos entre sí ya que era más visible la línea de tendencia entre los ellos, que proporciona el sentido de ambas variables. Esto último, se traduce en una buena correlación y concordancia. Pero de manera general, los cuatro conceptos del modelo final exhibieron poca dispersión de los datos.

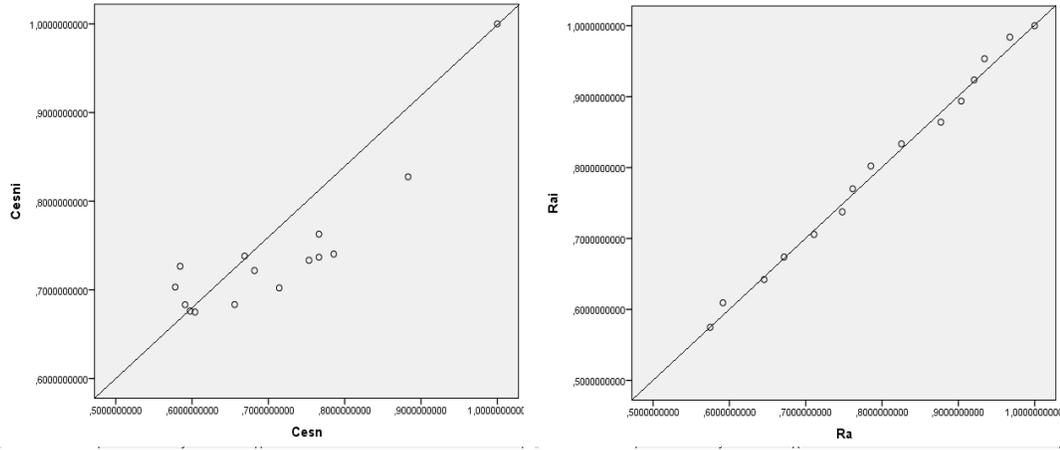


(a) Volumen de riego

(b) Conductividad eléctrica de la solución lixiviada

Figura 3–14: Dispersión de puntos del volumen de riego y la conductividad eléctrica de la solución lixiviada del modelo final

El trabajo presentó novedad científica ya que se obtuvo un modelo de riego, basado en la ingeniería del conocimiento y los mapas cognitivos difusos, para el cultivo protegido del tomate. Este resultado es novedoso debido a que los trabajos encontrados en la literatura científica, por los autores [Gutiérrez and Arias \(2017\)](#); [de la Piedra et al. \(2016\)](#); [Moreno-Ramos et al. \(2017\)](#), solo realizaron modelos de riego en el cultivo del tomate; utilizando la técnica neurodifusa de la inteligencia artificial. Además, de que no hubo evidencia de aplicaciones de los MCDs en el riego para ningún otro cultivo, aunque si fueron revisadas investigaciones en la predicción del rendimiento ([Papageorgiou et al., 2011, 2013](#); [Jayashree et al., 2015](#); [Natarajan et al., 2016](#); [Mourhir et al., 2017](#)). Por ello, este modelo



(a) Conductividad eléctrica de la solución nutritiva

(b) Rendimiento acumulado

Figura 3–15: Dispersión de puntos de la conductividad eléctrica de la solución nutritiva y el rendimiento acumulado del modelo final

representó uno de los primeros de naturaleza compleja, no lineal y de dinámica compleja en la cual a través de técnicas de inteligencia artificial se haya llegado a un modelo lineal aproximado. Todo esto acompañado, de un modelo obtenido con excelentes resultados con respecto a mediciones reales en la producción del tomate en las casas de cultivo. Con los análisis y discusiones realizadas, se demostró la validez del modelo de riego diseñado para el cultivo protegido de tomate que se muestra en la figura 3–16. Para lograr un mejor comprensión del resultado de las relaciones causales obtenidas entre las variables, se aprecia en las matrices de causalidad resultantes 3.3 y 3.4.

$$W^{final} = \begin{pmatrix} 0,0198 & 0,2433 & -0,7074 & 0,1926 \\ 0,0008 & 0,0631 & 0,0012 & 0,0317 \\ 0,2096 & 0,2712 & 0,2683 & 0,1337 \\ 0,0289 & 0,0069 & 0,00002 & 1,0054 \end{pmatrix} \quad (3.3)$$

$$H^{final} = \begin{pmatrix} 0,309656 & 0,6958 & 0,6603 & -0,5897 & 0,1603 & 0,0378 \\ 0,251925 & 0,5706 & 0,6170 & 0,8671 & -0,6329 & -0,6329 \\ 0,090397 & 0,00002 & 0,000003 & 0,0034 & -0,0453 & 0,0097 \\ 0,000025 & 0,0002 & 0,0002 & 0,00002 & 0,00002 & 0,00002 \end{pmatrix} \quad (3.4)$$

### 3.8. Consideraciones finales del capítulo

La elección de la cantidad de candidatos se consideró correcta, ya que reduce el porcentaje de errores de los resultados a alcanzar.

El método Delphi obtuvo, a través de tres encuestas a los expertos, las variables y sus relaciones causales para la representación del modelo inicial.

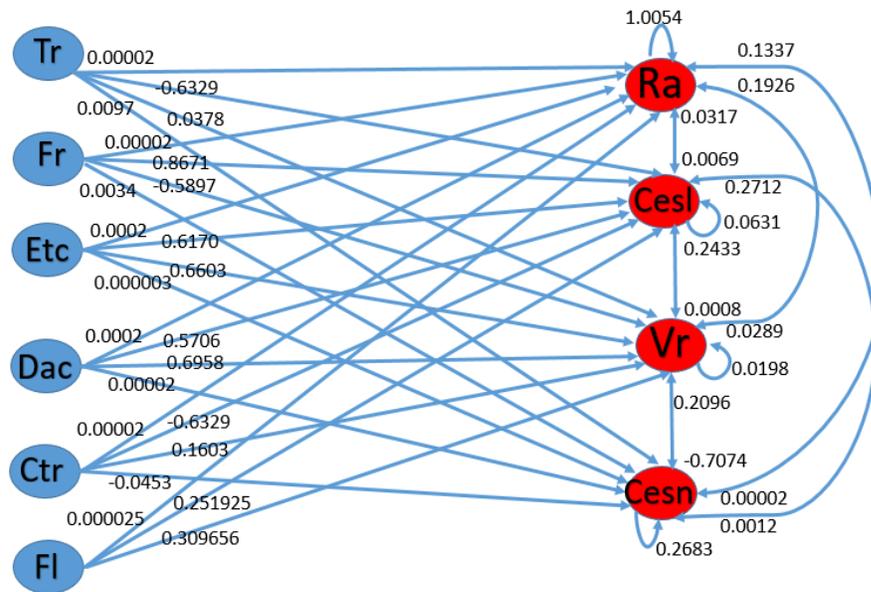


Figura 3–16: Representación del mapa cognitivo difuso para el modelado del riego del tomate en cultivo protegido

La evaluación del modelo inicial obtenidos por la ingeniería del conocimiento no ofrece buenos resultados por eso se procedió a realizar un ajuste del modelo a través del proceso de aprendizaje.

La validación del modelo obtenido presentó un mejor seguimiento con respecto a los datos históricos, que el modelo ofrecido por los expertos.

El modelo obtenido fue valorado por lo expertos de aceptable, en cada uno de los conceptos. Con este modelo se pudo lograr la estimación de variables de impacto productivo para ayudarles en la toma de decisiones.

En el análisis estadístico se corroboró la buena correlación y concordancia entre las mediciones reales y el modelo obtenido. Demostrando los excelentes resultados del modelo de riego diseñado para el cultivo protegido de *S. lycopersicum*.

## CONCLUSIONES

Como resultado final arrojado por esta investigación, se diseña un modelo de riego para el cultivo protegido del tomate para la estimación de variables de impacto productivo, lo cual queda demostrado mediante simulación y en pruebas experimentales. A partir de estos resultados, se plantean las conclusiones generales siguientes:

- Las diez variables involucradas en el riego del cultivo protegido de *S. lycopersicum*, identificadas mediante la ingeniería del conocimiento, demuestran la efectividad de la metodología en determinar la primera etapa del modelado mediante los mapas cognitivos difusos.
- El paquete de *software* logra un modelo de riego para el cultivo protegido de *S. lycopersicum* con una estabilización del error cuadrático medio de 0,95 %; mediante un algoritmo de aprendizaje basado en el gradiente descendente.
- Los expertos valoran el modelo de aceptable ya que se logra la estimación de variables de impacto productivo para ayuda en la toma de decisiones en la producción de *S. lycopersicum*, lo que es sustentado con un análisis estadístico.

## RECOMENDACIONES

Para establecer la necesaria continuidad que debe tener este trabajo se recomienda lo siguiente:

- ★ Continuar con la metodología propuesta en esta investigación, para realizar otras pruebas y simulaciones con otros métodos de aprendizaje.
- ★ Utilizar una mayor cantidad de datos experimentales para cada una de las variables del modelo para contar con una mejor base de datos para el proceso de aprendizaje.

## REFERENCIAS BIBLIOGRÁFICAS

- Abreu, Omar, Miguel E Naranjo, Bertha S Rhea and Mónica C Gallegos (2016). Modelo didáctico para la facultad de ciencias administrativas y económicas de la universidad técnica del norte en ecuador. *Formación universitaria* **9**(4), 3–10.
- Abreu, P, JF Meneses and C Gary (1998). Tompousse, a model of yield prediction for tomato crops: calibration study for unheated plastic greenhouses. In: *XXV International Horticultural Congress, Part 9: Computers and Automation, Electronic Information in Horticulture 519*. pp. 141–150.
- Aguilar, J. and J. Contreras (2010). *Cognitive maps: advances in theory, methodologies, tools and applications*. Chap. The FCM designer tool, pp. 71–87. Vol. 247. Berlin.
- Allen, Richard G, Luis S Pereira, Dirk Raes, Martin Smith et al. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *FAO* **300**(9), 5–109.
- Almenara, Julio Cabero and María del Carmen Llorente Cejudo (2015). Entornos personales de aprendizaje (ple): valoración educativa a través de expertos. *Areté: Revista Digital del Doctorado en Educación de la Universidad Central de Venezuela* **1**(1), 7–19.
- Antúnez, A. and S. Felmer (2017). Manejo del riego en tomate. In: *Manual de cultivo del tomate bajo invernadero*. pp. 66–79. Number 12. Santiago de Chile, Chile.
- Arriaga, J. (2015). Modelado, simulación y control del uso del agua en la agricultura. Tesis doctoral. Universidad de Sevilla: Escuela Técnica Superior de Ingeniería. España.
- Ayer, RS and DW Westcot (1985). Water quality for irrigation. *FAO irrigation and drainage paper*.
- Benítez, Raúl, Gerard Escudero, Samir Kanaan and David Masip Rodó (2014). *Inteligencia artificial avanzada*. Editorial UOC. Barcelona, España.
- Berdel, J. (1989). La agricultura artificial en la agricultura: Perspectivas de los sistemas expertos. *Revista de Estudios Agro-Sociales* (149), 62–77.
- Bouman, Bas AM, MCS Wopereis and JJM Riethoven (1994). *The use of crop growth models in agro-ecological zonation of rice*. AB-DLO.
- Bueno, S.; Salmeron, JL. (2009). Benchmarking main activation functions in fuzzy cognitive maps. **36**(3), 5221–5229.
- Christen, Benjamin, Chris Kjeldsen, Tommy Dalgaard and Julia Martin-Ortega (2015). Can fuzzy cognitive mapping help in agricultural policy design and communication?. *Land Use Policy* **45**, 64–75.
- Cook, Roberta L and Linda Calvin (2005). Greenhouse tomatoes change the dynamics of the north american fresh tomato industry. Technical report.

- Costa, JM and E Heuvelink (2007). Today's worldwide tomato production. *Fruit & Veg tech* **2007**, 14–16.
- Crespo, María Eugenia Ramos, Maricela González Pérez, Inés Margarita Torres Rivero and Reynier Fernández de la Torre (2018). Validación de indicadores para la gestión pública de la seguridad alimentaria y nutricional. *Cooperativismo y Desarrollo: COODES* **6**(2), 169–178.
- Cruz, M. (2009). El método delphi en las investigaciones educacionales. *La Habana: Editorial Academia* pp. 1–48.
- Dagnino, Jorge (2014). Coeficiente de correlacion lineal de pearson. *Chil Anest* **43**, 150–153.
- Dai, Jianfeng, Weihong Luo, Yongxiu Li, Changmei Yuan, Yongshan Chen and Jiheng Ni (2006). A simple model for prediction of biomass production and yield of three greenhouse crops. In: *III International Symposium on Models for Plant Growth, Environmental Control and Farm Management in Protected Cultivation 718*. pp. 81–88.
- de Arquer, M.I. (2011). *NTP 401: Fiabilidad humana: métodos de cuantificación, juicio de expertos*. Centro nacional de condiciones de trabajo. España.
- De Franciscis, Dimitri (2014). Jfcm: A java library for fuzzycognitive maps. In: *Fuzzy Cognitive Maps for Applied Sciences and Engineering*. pp. 199–220. Springer.
- de Koning, Adrianus Noël Maria (1994). *Development and dry matter distribution in glass-house tomato: a quantitative approach*. De Koning.
- de la Piedra, Mario Culebro, Aída Guillermina Cossío Martínez, IT de Tuxtla Gutiérrez, Joaquín Adolfo López Molina, Néstor Antonio Morales Navarro and Jorge Octavio Guzmán Sánchez (2016). Modelo neurodifuso para el control de humedad del suelo en cultivo hidropónico para la planta de tomate. *Revista Tecnología Digital Vol* **6**(1), 43–56.
- del Brío, Bonifacio Martín and Carlos Serrano Cinca (1995). Fundamentos de redes neuronales artificiales: hardware y software. *Scire: representación y organización del conocimiento* **1**(1), 103–125.
- Demšar, Janez, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinovič, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič et al. (2013). Orange: data mining toolbox in python. *The Journal of Machine Learning Research* **14**(1), 2349–2353.
- dEsposito, Daniela, Francesca Ferriello, Alessandra Dal Molin, Gianfranco Diretto, Adriana Sacco, Andrea Minio, Amalia Barone, Rossella Di Monaco, Silvana Cavella, Luca Tardella et al. (2017). Unraveling the complexity of transcriptomic, metabolomic and quality environmental response of tomato fruit. *BMC plant biology* **17**(1), 66–85.
- Dimokas, G, C Kittas and M Tchamitchian (2008). Validation of a tomato crop simulator for mediterranean greenhouses. In: *International Workshop on Greenhouse Environmental Control and Crop Production in Semi-Arid Regions 797*. pp. 247–252.

- Doorenbos, Jan and WO Pruitt (1977). Crop water requirements. *FAO irrigation and drainage paper*.
- Dorais, Martine, Beatrix Alsanius, Wim Voogt, Steeve Pepin, Hakki Tüzel, Yüksel Tüzel and Kurt Möller (2016). *Impact of water quality and irrigation management on organic greenhouse horticulture*.
- Escobar-Pérez, Jazmine and Ángela Cuervo-Martínez (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medicina* **6**(1), 27–36.
- Espinosa, ERJ (2017). Parámetros de explotación y uniformidad de riego en la máquina de pivote central otech-irrimec. *Revista Ingeniería Agrícola* **1**(1), 7–12.
- Evlanov, L. G. and V. A. Kutuzov (1978). *Valoraciones de expertos en la dirección*. Moscú: Ehkonomika.
- FAO (2006). Evapotranspiración del cultivo: guías para la determinación de los requerimientos de agua de los cultivos. riego y drenaje. Technical report. FAO.
- FAO (2013). Good agricultural practices for greenhouse vegetable crops. principles for mediterranean climate areas. Technical Report 217. FAO. Roma, Italy.
- FAOSTAT (2018). Estadísticas de la fao. Technical report. FAOSTAT.
- Felix, Gerardo, Gonzalo Nápoles, Rafael Falcon, Wojciech Froelich, Koen Vanhoof and Rafael Bello (2017). A review on methods and software for fuzzy cognitive maps. *Artificial Intelligence Review* pp. 1–31.
- Feria, Ulises Pacheco, Rosa Marina Castellanos Dorado and Milagros Morales Pérez (2018). Evaluación prospectiva de la eficiencia económica de la producción de frijol en la provincia santiago de cuba. *TERRA: Revista de Desarrollo Local* (4), 71–97.
- Fernández, MD, S Bonachela, F Orgaz, R Thompson, JC López, MR Granados, M Gallardo and E Fereres (2010). Measurement and estimation of plastic greenhouse reference evapotranspiration in a mediterranean climate. *Irrigation Science* **28**(6), 497–509.
- Flores Villarreal, Humberto Javier (2005). Operación eficiente de sistemas de transporte de gas natural mediante el método de gradiente reducido generalizado. PhD thesis. Universidad Autónoma de Nuevo León. México.
- Froelich, W.; Salmeron, J.L. (2014). Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series. **55**(6), 1319–1335.
- Gamboa, Sativa, Juan Sebastián, Ingrid Nathaly Villamor Medina et al. (2017). Generación y Simulación de un Modelo para Mejorar los Criterios de Riego en los Cultivos de Mini-Clavel de la Sabana de Bogotá Mediante Técnicas de Inteligencia Artificial. tesis de diploma. Universidad Distrital Francisco José de Caldas. Bogotá, Colombia.
- Garmendia, JL García and F Maroto Monserrat (2018). Interpretación de resultados estadísticos. *Medicina Intensiva* **42**(6), 370–379.
- González, R, E Camacho Poyato, P Montesinos and JA Rodríguez Díaz (2016). Optimización de la predicción de demanda de agua mediante algoritmos neuro-genéticos para un conjunto de datos reducido. In: *XXXIV Congreso Nacional de Riegos, Sevilla 2016*.

- Escuela Técnica Superior de Ingeniería Agronómica.
- González Robaina, Felicita, Teresa López Seijas and Julián Herrera Puebla (2015). Indicadores de productividad del agua por cultivos y técnicas de riego en cuba. *Revista Ciencias Técnicas Agropecuarias* **24**(4), 57–63.
- González Viera, Deborah, Miguel Socorro Quesada, Idaris Gómez Ravelo, Héctor R de las Cuevas Milán and Astrid Fernández de Castro Fabr  (2018). Indicadores para la evaluaci n de impacto de la gesti n tecnol gica en el cultivo del arroz. *Revista Ciencias T cnicas Agropecuarias* **27**(1), 99–109.
- Gonz lez, Yunia Reyes, Natalia Mart nez S nchez, Adolfo D az Sardi nas and Maidelis Milan s Luque (2017). De la extracci n al modelado del conocimiento en un sistema basado en el conocimiento. un enfoque desde el agrupamiento conceptual l gico combinatorio (from the extraction to knowledge modeling in a knowledge based system. a logical combinatorial conceptual grouping approach). *GECONTEC: Revista Internacional de gesti n del Conocimiento y la Tecnolog a* **5**(2), 41–57.
- Gorina, A. (2010). Din mica del procesamiento de la informaci n en las investigaciones sociales. Tesis doctoral. Universidad de Oriente. Cuba.
- Granados, L.F.M. (2017). L nea de inteligencia artificial y procesos de razonamiento. *TED Tecne Episteme Didaxis*.
- Gray, Steven A, Stefan Gray, Linda J Cox and Sarah Henly-Shepard (2013). Mental modeler: a fuzzy-logic cognitive mapping modeling tool for adaptive environmental management. In: *2013 46th Hawaii International Conference on System Sciences*. IEEE. pp. 965–973.
- Guti rrez, EJ C rdoba and JM Guti rrez Arias (2017). Control  ptimo para el control de temperatura en el cultivo de tomate en nvernadero basado en un sistema din mico. *Pistas Educativas*.
- Halbrendt, Jacqueline, Steven A Gray, Susan Crow, Theodore Radovich, Aya H Kimura and Bir Bahadur Tamang (2014). Differences in farmer and expert beliefs and the perceived impacts of conservation agriculture. *Global Environmental Change* **28**, 50–62.
- Hassan-Esfahani, Leila, Alfonso Torres-Rua and Mac McKee (2015). Assessment of optimal irrigation water allocation for pressurized irrigation system using water balance approach, learning machines, and remotely sensed data. *Agricultural Water Management* **153**, 42–50.
- Hern ndez D az, Mar a Isabel, Marisa Chailloux Laffta, V ctor Moreno Placeres, Alberto Igarza S nchez and Anselma Ojeda Veloz (2014). Niveles referenciales de nutrientes en la soluci n del suelo para el diagn stico nutricional en el cultivo protegido del tomate. *Idesia (Arica)* **32**(2), 79–88.
- Heuvelink, E; Marcelis, LFM (1989). Dry matter distribution in tomato and cucumber. In: *International Symposium on Growth and Yield Control in Vegetable Production 260*. pp. 149–180.

- Hobbs, Benjamin F, Stuart A Ludsin, Roger L Knight, Phil A Ryan, Johann Biberhofer and Jan JH Ciborowski (2002). Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems. **12**(5), 1548–1565.
- Holmes, Geoffrey, Andrew Donkin and Ian H Witten (1994). *Weka: A machine learning workbench*. Hamilton, New Zealand: Universidad de Waikato.
- Ilahy, Riadh, Gabriella Piro, Imen Tlili, Anissa Riahi, Rabaoui Sihem, Imen Ouerghi, Chafik Hdidier and Marcello Salvatore Lenucci (2016). Fractionate analysis of the phytochemical composition and antioxidant activities in advanced breeding lines of high-lycopene tomatoes. *Food & function* **7**(1), 574–583.
- Intriago, Dídimo Alexander Mendoza, Fanny Rodríguez Jarama, Edison Cuenca Cuenca, Rubén Rivera Fernández, Bernardo Navarrete Cedeño and Ernesto Cañarte Bermúdez (2019). Efecto de cuatro láminas de riego y meloidogyne spp. sobre características agronómicas y de rendimiento de tomate de mesa solanum lycopersicum l.. *Revista Universidad de Zulia* **10**(26), 28–47.
- Jayashree, LS, Nidhil Palakkal, Elpiniki I Papageorgiou and Konstantinos Papageorgiou (2015). Application of fuzzy cognitive maps in precision agriculture: a case study on coconut yield management of southern india malabar region. *Neural Computing and Applications* **26**(8), 1963–1978.
- Jensen, Marvin E and Howard R Haise (1963). Estimating evapotranspiration from solar radiation. *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division* **89**, 15–41.
- Jones, James W, Ehud Dayan, LH Allen, Herman Van Keulen and Hugo Challa (1991). A dynamic tomato growth and yield model (tomgro). *Transactions of the ASAE* **34**(2), 663–6672.
- Jones, J.W., A. Kenig and C.E. Vallejos (1999). Reduced states variable tomato growth model. *Transactions of the ASAE* **42**(1), 255.
- Juárez-Maldonado, Antonio, Karim de Alba Romenus, Alejandro Zermeño González, Homero Ramírez and Adalberto Benavides Mendoza (2015). Análisis de crecimiento del cultivo de tomate en invernadero. *Revista mexicana de ciencias agrícolas* **6**(5), 943–954.
- Kadiyala, Akhil and Ashok Kumar (2018). Applications of python to evaluate the performance of bagging methods. *Environmental Progress & Sustainable Energy* **37**(5), 1555–1559.
- KANO, Atsushi and Cornelius HM VAN GAVEL (1988). Design and test of a simulation model of tomato growth and yield in a greenhouse. *Journal of the Japanese Society for Horticultural Science* **56**(4), 408–416.
- Kayo, Eduardo Kazuo and José Roberto Securato (1997). Método delphi: fundamentos, críticas e vieses. *Cadernos de pesquisa em administração* **1**(4), 51–61.
- Kosko, Bart (1986). Fuzzy cognitive maps. *Int. J. Man-Mach. Stud* **24**, 65–75.

- Kreinovich, V. and C. Stylios (2015). Why fuzzy cognitive maps are efficient. **10**(5), 825–833.
- Landis, J Richard and Gary G Koch (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**(1), 159–174.
- León, Maikel, Gonzalo Nápoles, Ciro Rodriguez, María M García, Rafael Bello and Koen Vanhoof (2011). A fuzzy cognitive maps modeling, learning and simulation framework for studying complex system. In: *International Work-Conference on the Interplay Between Natural and Artificial Computation*. Springer. pp. 243–256.
- León, María and R Cun (2001). Necesidades hídricas del tomate protegido en las condiciones de cuba. *Revista Ciencias Técnicas Agropecuarias* **10**(3), 67–71.
- León, María, R Cun, Yohima Chaterlán and R Rodríguez (2005). Uso eficiente del agua en el cultivo del tomate protegido. resultados obtenidos en cuba. *Revista Ciencias Técnicas Agropecuarias*.
- Letey, John, Glenn J Hoffman, Jan W Hopmans, Stephen R Grattan, D Suarez, Dennis L Corwin, Jim D Oster, Laosheng Wu and Christopher Amrhein (2011). Evaluation of soil salinity leaching requirement guidelines. *Agricultural water management* **98**(4), 502–506.
- López, Verónica del Rocío Ochoa, Carlos Lino Ramírez, María del Rosario Baltazar-Flores, Miguel Ángel Casillas Araiza, Víctor Manuel Zamudio Rodríguez, Sandra Jacqueline López Cervera and Guillermo Eduardo Méndez Zamora (2016). Propuesta de un sistema para optimizar el riego en invernaderos de plantas heterogéneas usando wns y algoritmos evolutivos.. *Research in Computing Science* **113**, 131–145.
- Lynch, Stephen (2014). *Dynamical Systems with Applications Using Python*. 2nd ed.. Springer.
- Manzanet, Alfonso, José Enrique, Roberto Zayas Mujica, Alberto Juan Dorta-Contreras and José Luis Cadenas Freixas (2018). Propuesta para la evaluación de la calidad y el funcionamiento de revistas científicas en ciencias de la salud. *Revista Habanera de Ciencias Médicas* **17**(2), 325–334.
- Mar Cornelio, Omar, Maykel Y Leyva Vázquez and Ivan Santana Ching (2015). Modelo multicriterio multiexperto utilizando mapa cognitivo difuso para la evaluación de competencias. *Ciencias de la Información* **46**(2), 17–22.
- Martin, Cathie (2013). The interface between plant metabolic engineering and human health. *Current Opinion in Biotechnology* **24**(2), 344–353.
- Martínez, Ricardo Echeverri, Juan Carlos Osorio Gómez, Diego Echeverri Ibarra and Carlos Arturo Lozano Moncada (2018). Selección de una infraestructura de medición inteligente de energía usando una técnica de decisión multicriterio. *Scientia et Technica* **23**(2), 136–142.
- Matos, Eneida and Lorna Cruz (2011). La práctica investigativa, una experiencia en la formación doctoral en ciencia pedagógica. *Santiago de Cuba: Ediciones UO*.

- McCulloch, WS.;Pitts, W. (1988). A logical calculus of the ideas immanent in nervous activity. pp. 15–27.
- Medina-Ruíz, CA, IA Mercado-Luna, GM Soto-Zarazúa, I Torres-Pacheco and E Rico-García (2011). Mathematical modeling on tomato plants: A review. *African Journal of Agricultural Research* **6**(33), 6745–6749.
- Mesa, Rolando (2019). Incremento de la producción de *Solanum lycopersicum* L. cv. Aegean en cultivo protegido mediante alternativa de manejo del riego. Tesis de maestría. Universidad Central “Marta Abreu” de Las Villas, Facultad de Ciencias Agropecuarias.
- Mohr, Stephen (1997). Software design for a fuzzy cognitive map modeling tool. *Tensselaer Polytechnic Institute*.
- Mollinedo, V. and A. Tapia (2008). Fertirriego del cultivo de tomate. In: *Fertilización de cultivos y posturas*. 2da ampliada y actualizada ed. Hemisferio Sur. Argentina.
- Moreno-Ramos, Eduardo, América Morales-Díaz, Antonio Juárez-Maldonado, Karim de Alba-Romenus and Adalberto Benavides-Mendoza (2017). Obtención de trayectorias subóptimas de temperatura con el método de monte carlo para la maximización de la producción de tomate (*Solanum lycopersicum* l.) en invernadero. *Agrociencia* **51**(8), 849–866.
- Moreno, V. (2007). *Procedimientos para el manejo de la nutrición y el control de la fertirrigación en las casas de cultivo*. 2da ed.. Grupo empresarial frutícola. La Habana, Cuba.
- Morone, Piergiuseppe, Pasquale Marcello Falcone and Antonio Lopolito (2019). How to promote a new and sustainable food consumption model: A fuzzy cognitive map study. *Journal of cleaner production* **208**, 563–574.
- Mourhir, Asmaa, Elpiniki Papageorgiou, Konstantinos Kokkinos and Tajjeeddine Rachidi (2017). Exploring precision farming scenarios using fuzzy cognitive maps. *Sustainability* **9**(7), 1241.
- Nápoles, Gonzalo, Elpiniki Papageorgiou, Rafael Bello and Koen Vanhoof (2016). On the convergence of sigmoid fuzzy cognitive maps. *Information Sciences* **349**, 154–171.
- Nápoles, Gonzalo, Elpiniki Papageorgiou, Rafael Bello and Koen Vanhoof (2017a). Learning and convergence of fuzzy cognitive maps used in pattern recognition. *Neural Processing Letters* **45**(2), 431–444.
- Nápoles, Gonzalo, Maikel Leon, Isel Grau and Koen Vanhoof (2017b). Fuzzy cognitive maps tool for scenario analysis and pattern classification. In: *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE. pp. 644–651.
- Nápoles, Gonzalo, Rafael Bello and Koen Vanhoof (2014). How to improve the convergence on sigmoid fuzzy cognitive maps?. *Intelligent Data Analysis* **18**(65), 577–588.
- Nápoles Ruiz, Gonzalo R (2014). Algoritmo para mejorar la convergencia en Mapas Cognitivos Difusos Sigmoidales. PhD thesis. Universidad Central Marta Abreu de Las Villas. Facultad de Matemática.

- Natarajan, Rajathi, Jayashree Subramanian and Elpiniki I Papageorgiou (2016). Hybrid learning of fuzzy cognitive maps for sugarcane yield classification. *Computers and Electronics in Agriculture* **127**, 147–157.
- Nápoles, G., F. Vanhoenshoven and K. Vanhoof (2019a). Nonsynaptic error backpropagation in long-term cognitive networks. *IEEE Transactions on Neural Networks and Learning Systems* pp. 1–11.
- Nápoles, G., F. Vanhoenshoven and K. Vanhoof (2019b). Short-term cognitive networks, flexible reasoning and nonsynaptic learning. *Neural Networks* **115**, 72–81.
- ONE (2018). Capítulo 9: Agricultura, ganadería, silvicultura y pesca. In: *Anuario Estadístico de Cuba 2017*. 2018 ed. La Habana.
- Ozesmi, U.; Ozesmi, SL. (2004). Ecological models based on peoples knowledge: a multistep fuzzy cognitive mapping approach. **176**, 43–64.
- Padayachee, A, L Day, K Howell and MJ Gidley (2017). Complexity and health functionality of plant cell wall fibers from fruits and vegetables. *Critical reviews in food science and nutrition* **57**(1), 59–81.
- Palacios Corella, Adrián (2012). Implementación de un módulo para el entrenamiento y evaluación de redes neuronales mediante GPUs. PhD thesis.
- Papageorgiou, EI, AT Markinos and TA Gemtos (2010). Soft computing technique of fuzzy cognitive maps to connect yield defining parameters with yield in cotton crop production in central greece as a basis for a decision support system for precision agriculture application. In: *Fuzzy Cognitive Maps*. pp. 325–362. Springer.
- Papageorgiou, E.I., K. Poczketta and C. Laspidou (2016). Hybrid model for water demand prediction based on fuzzy cognitive maps and artificial neural networks. In: *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE. pp. 1523–1530.
- Papageorgiou, EI., KD. Aggelopoulou, TA. Gemtos and GD. Nanos (2013). Yield prediction in apples using fuzzy cognitive map learning approach. **91**, 19–29.
- Papageorgiou, EI. and Kannappan, A. (2012). Fuzzy cognitive map ensemble learning paradigm to solve classification problems: application to autism identification. **12**(12), 3798–3809.
- Papageorgiou, Elpiniki I, Athanasios T Markinos and Theofanis A Gemtos (2011). Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. **11**(4), 3643–3657.
- Papakostas, George A, Dimitris E Koulouriotis, Athanasios S Polydoros and Vassilios D Tourassis (2012). Towards hebbian learning of fuzzy cognitive maps in pattern classification problems. *Expert Systems with Applications* **39**(12), 10620–10629.
- Pérez, Agustin Mulet, Agustín M Mulet Gámez and Jorge González Ferrer (2018). Síndrome. empleo inadecuado del término en la discusión diagnóstica. semejanzas y divergencias en el criterio de expertos. *Correo Científico Médico*.

- Pérez, K (2014). Modelo de proceso de logro de consenso en mapas cognitivos difusos para la toma de decisiones en grupo. PhD thesis. Universidad de las Ciencias Informáticas, la Habana. Cuba.
- Pérez-Teruel, Karina, Maikel Leyva-Vázquez, Macarena Espinilla and Vivian Estrada-Sentí (2014). Computación con palabras en la toma de decisiones mediante mapas cognitivos difusos. *Revista Cubana de Ciencias Informáticas* **8**(2), 19–34.
- Pingzeng, L and B Shusheng (2011). Obstacle avoidance system for agricultural robots based on multi-sensor information fusion. pp. 1181–1185.
- Poczketka, K., A. Yastrebov and E.I. Papageorgiou (2015). Learning fuzzy cognitive maps using structure optimization genetic algorithm. In: *2015 federated conference on computer science and information systems (FedCSIS)*. IEEE. pp. 547–554.
- Powell, Catherine (2003). The delphi technique: myths and realities. *Journal of advanced nursing* **41**(4), 376–382.
- Puth, Marie-Therese, Markus Neuhäuser and Graeme D Ruxton (2015). Effective use of spearman's and kendall's correlation coefficients for association between two measured traits. *Animal Behaviour* **102**, 77–84.
- Qiu, Rangjian, Chunwei Liu, Zhenchang Wang, Zaiqiang Yang and Yuanshu Jing (2017). Effects of irrigation water salinity on evapotranspiration modified by leaching fractions in hot pepper plants. *Scientific reports* **7**(1), 7231.
- Ramírez, L. and A. Toledo (1999). Algunas consideraciones acerca del método de evaluación utilizando el criterio de expertos. In: *Conferencia dictada en Santa Fé de Bogotá, DC Colombia*.
- Ramírez-Arias, JA. (2005). Hierarchical multiobjective control for greenhouse crops. Phd. Universidad de Almería. España.
- Raza, Waseem, Ning Ling, Ruifu Zhang, Qiwei Huang, Yangchun Xu and Qirong Shen (2017). Success evaluation of the biological control of fusarium wilts of cucumber, banana, and tomato since 2000 and future research strategies. *Critical reviews in biotechnology* **37**(2), 202–212.
- Reca, J, C Trillo, JA Sánchez, J Martínez and D Valera (2018). Optimization model for on-farm irrigation management of mediterranean greenhouse crops using desalinated and saline water from different sources. *Agricultural Systems* **166**, 173–183.
- Redmond, Ramin Shamshiri, James W Jones, Kelly R Thorp, Ahmad Desa, Hasfalina Che Man and Sima Taheri (2018). Review of optimum temperature, humidity, and vapour pressure deficit for microclimate evaluation and control in greenhouse cultivation of tomato: a review. *International Agrophysics* **32**(2), 287.
- Rode, PF Cepeda, RJ Rocha Gamarra, HE Ponce Espinosa, AI García-Ravizé Guizar, DC Romero Díaz, P Ponce Cruz and A Molina Gutiérrez (2010). Invernadero inteligente basado en un enfoque sustentable para la agricultura mexicana. In: *VIII Congreso Internacional sobre Innovación y Desarrollo Tecnológico*. p. 8.

- Ruíz-García, A, JA Ramírez-Arias, RD Peña-Moreno, JC Torres-Monsivais and IL López-Cruz (2015). Data assimilation to improve states estimation of a dynamic greenhouse tomatoes crop growth model. In: *International Symposium on New Technologies and Management for Greenhouses-GreenSys2015 1170*. pp. 433–440.
- Sadati, Somayeh, Stijn Speelman, Mahmood Sabouhi, Mohsen Gitizadeh and Bijan Ghahraman (2014). Optimal irrigation water allocation using a genetic algorithm under various weather conditions. *Water* **6**(10), 3068–3084.
- Sahbani, Farhat and Ezzidine Ferjani (2018). Identification and modelling of drop-by-drop irrigation system for tomato plants under greenhouse conditions. *Irrigation and Drainage* **67**(4), 550–558.
- Sarache-Castro, William Ariel, Yasel José Costa-Salas and Jhully Paulin Martínez-Giraldo (2015). Environmental performance evaluation under a green supply chain approach. *Dyna* **82**(189), 207–215.
- Schery, José Angel Álvarez and Alexander Gorina Sánchez (2013). Plataforma computacional gplipce: Gestión y procesamiento en línea de información proporcionada por el criterio de expertos.
- Serrano, C and JL Gallizo (1996). Las redes neuronales artificiales en el tratamiento de la información financiera. *Biblioteca Electrónica Ciberconta*.
- Siegel, Sidney, Javier Aguilar Villalobos, Lisha Jacobo Ruseil and Ricardo Vinos Cruz-López (1972). *Estadística no paramétrica aplicada a las ciencias de la conducta*. Vol. 4. Trillas México.
- Silva, Juliana Aparecida dos Santos da (2016). Models for estimation growth, yield and nutrients content of processing tomato.
- Sonneveld, C. and W. Voogt (2009). *Plant nutrition of greenhouse crops*. Springer.
- Sosa, Richar, Alejandro Alfonso and Gonzalo Nápoles (2019). Synaptic learning of long-term cognitive networks with inputs. In: *IJCNN 2019: The International Joint Conference on Neural Networks*.
- Sturm, Roland and Ruopeng An (2014). Obesity and economic environments. *CA: a cancer journal for clinicians* **64**(5), 337–350.
- Stylios, CD and PP Groumpos (2004). Modeling complex systems using fuzzy cognitive maps. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* (34), 155–162.
- Tajuelo, J.M. (2005). *El riego por aspersión y sus tecnologías*. Mundi-Prensa.
- Tang, Jiexiong, Chenwei Deng and Guang-Bin Huang (2015). Extreme learning machine for multilayer perceptron. *IEEE transactions on neural networks and learning systems* **27**(4), 809–821.
- Tettamanzi, A.G.; Tomassini, M. (2013). *Soft computing: integrating evolutionary, neural, and fuzzy systems*. Springer. Berlin, Alemania.

- Torres, Martha Constanza Daza, Hugo Stiven Meneses Carvajal, Aldemar Reyes Trujillo and Norberto Urrutia Cobo (2017). Necesidades hídricas de estevia calculadas con el coeficiente del cultivo. *Agronomía Mesoamericana* pp. 509–521.
- Tsadiras, A.K. (2008). Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. **178**(20), 3880–3894.
- Tsadiras, A.K.; Margaritis, K.G. (1999). An experimental study of the dynamics of the certainty neuron fuzzy cognitive maps. **24**, 95–116.
- Tsang, SW and Chi Yung Jim (2016). Applying artificial intelligence modeling to optimize green roof irrigation. *Energy and Buildings* **127**, 360–369.
- Vaca, Luz Elva Marín, Martha Lilia Domínguez Patiño, Nadia Lara Ruiz and Miguel Aguilar Cortes (2015). Simulación processes of a mathematical model to determine the growth of tomato under plastic cover. *Agricultural Sciences* **6**(12), 1532.
- Varela-Ruiz, Margarita, Laura Díaz-Bravo and Rocío García-Durán (2012). Descripción y usos del método delphi en investigaciones del área de la salud. *Investigación en educación médica* **1**(2), 90–95.
- Wang, Xiukang, Jia Yun, Peng Shi, Zhanbin Li, Peng Li and Yingying Xing (2018). Root growth, fruit yield and water use efficiency of greenhouse grown tomato under different irrigation regimes and nitrogen levels. *Journal of Plant Growth Regulation* pp. 1–16.
- White, Ethan and Lawrence J Mazlack (2011). Discerning suicide notes causality using fuzzy cognitive maps. In: *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*. IEEE. pp. 2940–2947.
- Williams, Patricia L and Christine Webb (1994). The delphi technique: a methodological discussion. *Journal of advanced nursing* **19**(1), 180–186.
- Witkin, B.R. and J.W. Altschuld (1995). *Planning and conducting needs assessment: A practical guide*. Thousand Oaks: Sage.
- Zambrano, Diana Corina, Ruth Rebeca Bonilla Buitrago, Laura Avellaneda Rivera and Gregorio Zambrano (2015). Análisis prospectivo de los bioinsumos agrícolas en Colombia: una consulta a expertos. *Revista colombiana de Biotecnología* **17**(2), 103–117.
- Zamora, E.I., C.E. Duarte and R. Cun (2014). Coeficientes de cultivos (kc) en Cuba. *Revista Ingeniería Agrícola* **4**(3), 16–22.
- Zekki, H, C Gary, A Gosselin and L Gauthier (1999). Validation of a photosynthesis model through the use of the CO<sub>2</sub> balance of a greenhouse tomato canopy. *Annals of botany* **84**(5), 591–598.
- Zhang, Yu-Jie, Ren-You Gan, Sha Li, Yue Zhou, An-Na Li, Dong-Ping Xu and Hua-Bin Li (2015). Antioxidant phytochemicals for the prevention and treatment of chronic diseases. *Molecules* **20**(12), 21138–21156.
- Zhi-Qiang, L.I.U (2001). Causation, bayesian networks, and cognitive maps. *Acta automatica sinica* **27**(4), 552–566.

# ANEXO A

## ENCUESTAS PARA EL MODELADO DEL RIEGO EN CULTIVO PROTEGIDO DEL TOMATE

### A.1. Encuesta de selección de expertos

¿Valore su nivel de conocimiento sobre las variables que influyen en el crecimiento y desarrollo de hortalizas en el cultivo protegido en un rango del 0-10?

Tabla A-1: Escala tipo Linkert para el grado de influencia de las fuentes de argumentación

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

Marque con una X su nivel de conocimiento acerca del problema que se está resolviendo y las fuentes bibliográficas.

Tabla A-2: Escala tipo Linkert para el grado de influencia de las fuentes de argumentación

Fuentes de argumentación	ALTO(A)	MEDIO(M)	BAJO(B)
1. Análisis de las investigaciones teóricas y/o experimentales			
2.Experiencia obtenida en la actividad profesional			
3.Análisis de la literatura especializada y publicaciones de autores nacionales			
4.Análisis de la literatura especializada y publicaciones de autores extranjeros			
5.Conocimiento del estado actual de la problemática en el país y en el extranjero			
6.Intuición			

### A.2. Encuesta para determinar los principales factores del crecimiento y desarrollo de hortalizas en cultivo protegido

Clasifique su respuesta de acuerdo con MR (muy relevante), BR (bastante relevante), R (relevante), PR (poco relevante) y NR (no relevante).

Tabla A-3: Encuesta determinar las variables en el crecimiento y desarrollo de hortalizas en cultivo protegido

VARIABLES CLIMÁTICAS	VARIABLES DE LAS CONDICIONES DEL SUELO	VARIABLES FENOTÍPICAS DE LA PLANTA
Humedad relativa	Humedad del suelo	Altura de la planta
Temperatura	Conductividad eléctrica	Número de entrenudos
Iluminación	pH	Longitud de entrenudos
	Tipo de suelo	Número de bifurcaciones
		Número de flores
		Número de frutos
		Masa fresca de los frutos
		Masa seca de los frutos
		Rendimiento/área
		Tipo de suelo
		Tamaño de la planta
		Caída de flores
		Caída de frutos pequeños
		Tipo de lóculos

### A.3. Encuesta para determinar las principales variables involucradas en el riego del tomate en cultivo protegido

De las variables propuestas como relevantes para el desarrollo y crecimiento de las hortalizas en casas de cultivos, marque con una X mediante omisión, mantenimiento o agregación las principales variables involucradas en el riego del tomate en cultivo protegido, cuadro A-4. En el caso de agregación especifique por cuál la cambiaría y justifique su respuesta.

### A.4. Relaciones causales entre las variables involucradas en riego del cultivo protegido del tomate

Las variables que usted propuso como relevantes en el riego del tomate en cultivo protegido. Proponga una relación una con otra, catalogando dicha influencia como:

Tabla A-4: Encuesta para determinar las principales variables involucradas en el riego del tomate en cultivo protegido

Variables	Omisión	Mantenimiento	Agregación
Humedad relativa			
Temperatura			
Iluminación			
Humedad del suelo			
Conductividad eléctrica			
pH			
Número de entrenudos			
Longitud de entrenudos			
Número de bifurcaciones			
Número de flores			
Número de frutos			
Masa fresca de los frutos			
Masa seca de los frutos			
Rendimiento/área			

Tabla A-5: Términos lingüísticos para las relaciones entre variables

$S_0$	Negativamente muy fuerte
$S_1$	Negativamente fuerte
$S_2$	Negativamente media
$S_3$	Negativamente débil
$S_4$	Cero
$S_5$	Positivamente débil
$S_6$	Positivamente media
$S_7$	Positivamente fuerte
$S_8$	Positivamente muy fuerte

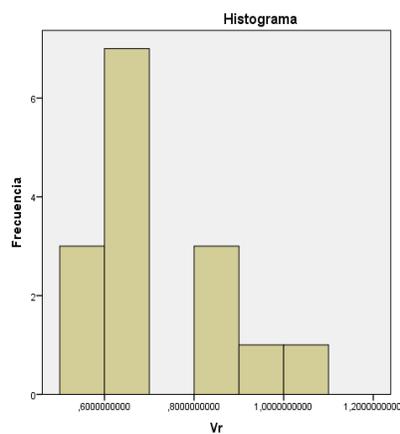
Tabla A-6: Relaciones causales entre las variables en el riego del tomate en cultivo protegido

Relación	Vr	Cesl	Cesn	Ra
F1				
Dac				
Etc				
Fr				
Ctr				
Tr				
Vr				
CEsl				
Cesn				
Ra				

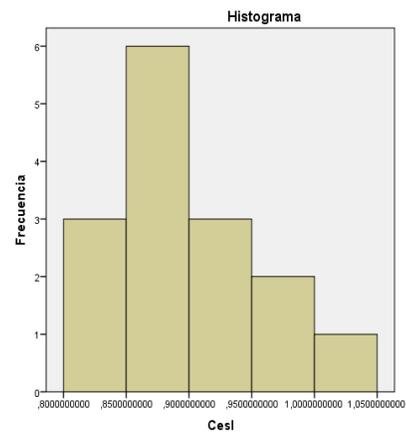
# ANEXO B

## ANÁLISIS ESTADÍSTICOS

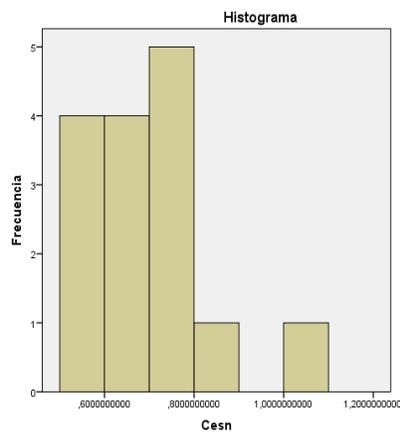
### B.1. Histogramas



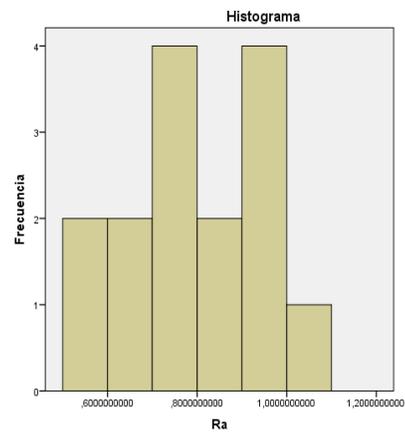
(a) Volumen de riego



(b) Conductividad eléctrica de la solución lixiviada

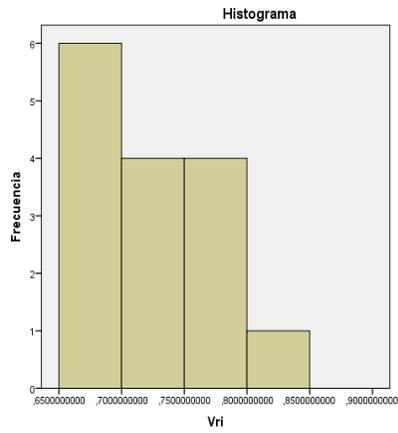


(c) Conductividad eléctrica de la solución nutritiva

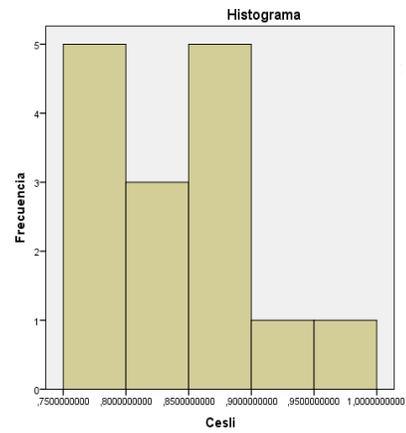


(d) Rendimiento acumulado

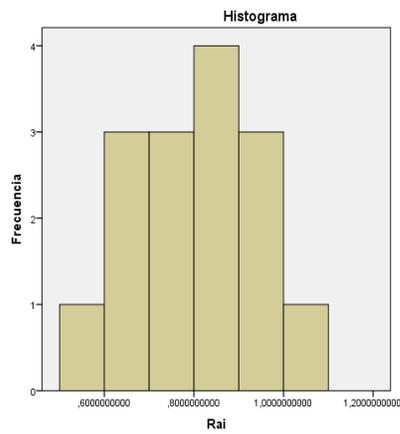
Figura B-1: Representación de los histogramas de los conceptos de los datos históricos



(a) Volumen de riego



(b) Conductividad eléctrica de lixiviada



(c) Rendimiento acumulado

Figura B-2: Representación de los histogramas de los conceptos del modelo obtenido