

Estimación de la rentabilidad en la producción de café sostenible a través de un modelo cualitativo jerárquico



Leidy Katherine Casilimas López

Tesis de Maestría en Ingeniería Telemática

Director:

Dr. David Camilo Corrales Muñoz, PhD

Co-Director:

Dr. Juan Carlos Corrales Muñoz, PhD

Universidad Del Cauca
Facultad de Ingeniería Electrónica y telecomunicaciones
Departamento de Telemática
e- @mbiente
Popayán, agosto de 2021

Leidy Katherine Casilimas López

Estimación de la rentabilidad en la producción de café sostenible a través de un modelo cualitativo jerárquico

Tesis presentada a la
Facultad de Ingeniería Electrónica y Telecomunicaciones de la
Universidad del Cauca, Colombia
para otorgar el grado académico de

Magíster en:
Ingeniería Telemática

Director:

Dr. David Camilo Corrales Muñoz, PhD

Co-Director:

Dr. Juan Carlos Corrales Muñoz, PhD

Popayán
2021

Agradecimientos

Dedico este trabajo de grado a mis queridos padres: Adiela Patricia López y Pedro Luis Casilimas, les agradezco cada palabra de aliento y cada ayuda brindada durante el desarrollo de mi carrera profesional y de mi formación personal. Agradezco a mi hermana Jessica Casilimas por ser una amiga incondicional en todo momento. Agradezco a mis demás familiares, porque de todos ellos recibo siempre las mejores energías para seguir adelante con mis proyectos. A mis tutores David Camilo Corrales y Juan Carlos Corrales, personas que me brindaron su conocimiento desde sus diferentes áreas de especialidad para cumplir satisfactoriamente con este proyecto.

Un agradecimiento especial al grupo de investigación de Ingeniería Telemática de la Universidad del Cauca y a la finca la Sultana y su equipo de trabajo, gracias a su apoyo la propuesta de Maestría presentada en este documento es una realidad. Finalmente, agradezco a Dios por todas las oportunidades que me brinda día a día para ser una mejor persona en este mundo.

Resumen estructurado

Antecedentes: la producción de café es una actividad económica fundamental en muchos países en vía de desarrollo, donde el 90% es producido por pequeños productores. Pese a esto, la mayor parte de ellos, tienen grandes dificultades para ganarse la vida dignamente debido a diversos factores económicos, sociales y climáticos entre otros, que giran entorno a los cultivos e impactan su rentabilidad. Por lo tanto, conocer la rentabilidad y los factores determinantes relacionados a las pequeñas producciones de café, es un factor habilitador para la toma de decisiones que permitan desarrollar sistemas agrícolas económicamente más sostenibles.

Objetivos: proponer un modelo cualitativo jerárquico basado en conocimiento experto para estimar la rentabilidad en la producción de café sostenible a pequeña escala.

Métodos: se propone en primer lugar, construir una base de conocimiento basado en la operación de la finca “La Sultana” con el fin de caracterizar las producciones de café sostenible a pequeña escala. Seguidamente, se propone un modelo cualitativo jerárquico utilizando la base de conocimiento resultante, para estimar diferentes niveles de rentabilidad en pequeñas producciones de café sostenible. Finalmente, el modelo es evaluado con datos reales del Instituto de Café de Costa Rica.

Resultados: el presente trabajo entrega como resultado una base de conocimiento (i.e., Modelo Jerárquico Conceptual) que caracteriza las producciones de café sostenible a pequeña escala. Adicionalmente, un modelo cualitativo jerárquico basado en conocimiento experto que estima de manera precisa diferentes niveles de rentabilidad en pequeñas producciones de café sostenible. Y finalmente, un conjunto de datos organizado, que contiene información de rentabilidad de producciones de café sostenible.

Conclusiones: la creación de un modelo cualitativo jerárquico difuso basado en conocimiento experto permitió aportar una solución no impulsada por datos, de fácil interpretación y con un nivel de precisión; perfectamente adaptable al contexto de la estimación de la rentabilidad en pequeñas producciones de café sostenible donde: se evidenciaba una limitación en el acceso los datos, la necesidad de conocer la interacción entre variables y de obtener información de rentabilidad confiable.

Palabras clave: Métodos Jerárquicos Multicriterio, Lógica difusa, Reglas-Mandammi, Topsis, Saw, Rentabilidad, Pequeñas producciones de café, Sostenibilidad, Agricultura.

Abstract

Background: coffee production is a fundamental economic activity in many developing countries, where smallholders produce around 90%. Despite this, most of them have great difficulties earning a decent living due to various economic, social, and climatic factors, among others, that revolve around crops and influence the profitability. Therefore, knowing the profitability and the determining factors related to small coffee productions is an enabling factor for making decisions that allow the development of more economically sustainable agricultural systems. However, knowing the profitability at this level represents a research challenge because small farms in developing countries do not have precise information and data systems on their agricultural microeconomic activities.

Objectives: propose a hierarchical qualitative model based on expert knowledge to estimate profitability in small-scale sustainable coffee production.

Methods: first of all, we propose to build a knowledge base based on the operation of the “La Sultana” farm to characterize sustainable small-scale coffee productions. Next, we propose a hierarchical qualitative model using the resulting knowledge base to estimate different levels of profitability in small sustainable coffee productions. Finally, we evaluate the model with actual data from the Costa Rican Coffee Institute.

Results: the present work results in a knowledge base (i.e., Conceptual Hierarchical Model) that characterizes small-scale sustainable coffee productions. Additionally, a hierarchical qualitative model based on expert knowledge precisely estimates different levels of profitability in small sustainable coffee productions. Finally, an organized data set containing profitability information for sustainable coffee productions.

Conclusions: the creation of a fuzzy hierarchical qualitative model based on expert knowledge made it possible to provide a solution no data-driven, easy to interpret, and precise; perfectly adaptable to the context of profitability in small sustainable coffee farming.

Keywords: Hierarchical Multicriteria Methods, Fuzzy Logic, Mandammi-Rules, Topsis, Saw, Profitability, Small coffee productions, Sustainability, Agriculture.

Contenido

1. Introducción.....	1
1.1. Definición del Problema	1
1.2. Hipótesis	2
1.3. Objetivos	3
1.3.1. Objetivo general	3
1.3.2. Objetivos específicos	3
1.4. Metodología	3
1.5. Actividades generales y específicas.....	4
1.6. Contribuciones	5
2. Estado del Arte	7
2.1. Conceptualización	7
2.1.1. Desde el dominio de la Agricultura.....	7
2.1.2. Desde el dominio de investigación de la ingeniería	7
2.2. Trabajos Relacionados.....	9
2.2.1. Modelos jerárquicos relacionados a rentabilidad en el sector agrícola	9
2.2.2. Modelos difusos en agricultura.....	10
2.2.3. Métodos multicriterio difusos en agricultura	10
2.2.4. Resumen.....	12
3. Modelo jerárquico multicriterio de rentabilidad.....	14
3.1. Área de estudio y panel de expertos	14
3.2. Construcción del modelo HMP-Coffee	15
3.2.1. Fase de conocimiento del contexto: Análisis multinivel de granularidad	16
3.2.2. Fase Jerárquica Multicriterio Convencional.....	24
3.2.3. Resumen.....	29
4. Sistema de Lógica Difusa en el HMM	31
4.1. Implementación de un Sistema de Lógica difusa en el HMM.....	31
4.1.1. Fuzzificación de atributos.....	32
4.1.2. Método Multicriterio Difuso	37
4.1.3. Resumen.....	47
4.2. Conjunto de datos para la evaluación de los enfoques multicriterio	49
4.3. Métricas de desempeño	50
4.4. Resultados	51
4.4.1. Matrices de confusión por clase.....	51
4.4.2. Análisis de métricas de desempeño.....	55

4.4.3.	Discusión.....	56
4.4.4.	Resumen.....	57
5.	Conclusiones y trabajos futuros.....	57
5.1.	Conclusiones.....	57
5.2.	Trabajos futuros	60
6.	Referencias	61

Lista de Tablas

Tabla 1. Resumen análisis de trabajos relacionados	12
Tabla 2 Equipo involucrado en la concepción de HMP-Coffee.....	15
Tabla 3. Lista de conceptos de HMP-Coffee.	19
Tabla 4. Rejilla de resumen de HMP-Coffee. La rejilla 1 analiza el nivel de relación. La rejilla 2 analiza el nivel de generalidad.....	21
Tabla 5. Definición de escalas para el conjunto de atributos básicos de HPM-Coffee. Celdas grises: escalas cualitativas definidas por el panel de expertos de HPM-Coffee. Celdas blancas: rangos por valor cualitativo (determinados por un grupo de cafetaleros locales con su unidad respectiva).....	25
Tabla 6. Pesos definidos para el conjunto de atributos HPM-Coffee. La columna atributo contiene los atributos básicos, los atributos agregados y la variable de salida. La columna peso incluye cinco niveles de ponderación para los atributos del modelo de café HMP.....	27
Tabla 7. Valores máximos, mínimos y promedios del atributo básico	34
Tabla 8. Valores de los mínimos y los máximos extendidos un valor delta para cada estado Costoso, Promedio y Económico.	35
Tabla 9. Valores normalizados de cada función triangular Económico, Promedio y Costoso...	36
Tabla 10. Atributos del conjunto de datos para la evaluación de los enfoques multicriterio.....	49
Tabla 11. Matriz de confusión de múltiples clases.....	50
Tabla 12. Matriz de confusión para el modelo HMP-Coffee Convencional.....	51
Tabla 13. Matriz de confusión para el modelo HMP-Coffee Reglas de Mandammi.....	52
Tabla 14. Matriz de confusión para el modelo HMP-Coffee Topsis.	52
Tabla 15. Matriz de confusión para el modelo HMP-Coffee Saw.	52
Tabla 16 Resultados de los criterios estadísticos (Precisión, Recuperación, Puntuación-F1) de HMP-Coffee convencional, , Reglas de Mandammi , Topsis y Saw.	55

Lista de Figuras

Figura 1. Estructura genérica de un HMM	8
Figura 2. Ubicación de la finca de café de Colombia La Sultana (2 ° 22'28.51 "N, 76 ° 43'31.89" W). Fuente : www.proyectomamundi.com	14
Figura 3. Proceso conceptual para la construcción del modelo HMP-Coffee. Fase 1: Conocimiento del contexto; Fase 2: multicriterio jerárquico. Fuente: autoría propia.....	16
Figura 4. Documentos recuperados para las tres cadenas de consulta: Producción de café a pequeña escala (Small Coffee Production-SCP), Rentabilidad económica del café (Economic Coffee Profitability-ECP) y, Rentabilidad de Café a pequeña escala (Small Coffee Production Profitability-SCPR). Fuente: autoría propia.	17
Figura 5. Proceso de análisis de texto. 1. Actividades genéricas del análisis textual (diagrama de flujo color blanco) 2. Actividades instanciadas para el modelo HMP-Coffee (diagrama de flujo color verde). Fuente: autoría propia.....	18
Figura 6. Diagrama de Venn de los conceptos frecuentes. Los colores azul, gris y verde representan las regiones por cadenas de consulta: Producción de café a pequeña escala (SCP), Rentabilidad económica del café (ECP) y Rentabilidad del café a pequeña escala (SCPR). Las 24 palabras ubicadas en las regiones de superposición se seleccionan como lluvia de conceptos de HMP-Coffee (palabras en negrita).....	19
Figura 7. Modelo conceptual de rentabilidad HMP-Coffee . Fuente: autoría propia.	24
Figura 8. Tabla DEXI para el atributo agregado "labores del cultivo" (captura de pantalla del software Dexi). Contiene los valores cualitativos por atributo: Para Y= "labores del cultivo", Alto, Promedio o Bajo; para X1 = "Tiempo de ejecución en actividades de cuidado del cultivo", Costoso, Promedio o Económico y para X2 = "Precio de la mano de obra", Alto, Normal o Bajo. Además, presenta las nueve combinaciones posibles entre atributos.	29
Figura 9. Proceso conceptual para la construcción del modelo HMP-Coffee. Fase 1: Conocimiento del contexto; Fase 2: Multicriterio jerárquico difuso. Fuente: autoría propia.	32
Figura 10. Pasos por cuadrantes para la construcción de las funciones de membresía. I. Identificación de mínimos (min), promedios (prom) y máximos (max). II: Construcción funciones triangulares. III: Calculo Delta. IV: Extensión de máximos y mínimos. Fuente: autoría propia.	33

Figura 11. Funciones de membresía triangulares del atributo básico "Precio de mano de obra" con escalas cualitativas representadas por colores: costoso en rojo, promedio en negro y económico en verde	34
Figura 12. Función de membresía resultante correspondiente al atributo básico "Precio de mano de obra" con escalas cualitativas representadas por colores: Costoso en rojo, Promedio en negro y Económico en verde.	35
Figura 13. Proceso de normalización de las funciones de membresía. Fuente: autoría propia	36
Figura 14. Función de membresía normalizada del atributo básico "Precio de mano de obra" representadas por colores: Costoso en rojo, Promedio en negro y Económico en verde.....	37
Figura 15. Proceso conceptual para implementar el modelo HMP-Coffee basado en Reglas-Mandammi. Paso 1: Fuzzificación; Paso 2: Método basado en Reglas-Mandammi. Fuente: autoría propia.....	38
Figura 16. Funcionamiento de una regla difusa para valores de entrada de “Precio de Precio de mano de obra” = 1,75 y “Tiempo de ejecución en actividades del cultivo” =1,2.....	39
Figura 17. Proceso conceptual para la implementación del modelo HMP-Coffee basado en Topsis. Paso 1: Fuzzificación; Paso 2: Método basado en Topsis. Fuente: autoría propia.	40
Figura 18. Proceso conceptual para la implementación del modelo HMP-Coffee basado en Saw. Paso 1: Fuzzificación; Paso 2: Método basado en Saw. Fuente: autoría propia.	44
Figura 19. Implementación de la fase multicriterio difusa en cada nivel de modelo Jerárquico. Fuente: autoría propia.	47
Figura 20 Comparación grafica de precisión. El eje x incluye los cuatro enfoques implementados: Reglas de Mandammi, Topsis, Convencional y Saw. El eje y corresponde al porcentaje de precisión general alcanzados por cada enfoque. Fuente: autoría propia.	53
Figura 21 Curvas ROC de HMP-Coffee.(a) Reglas de Mandammi, (b) Topsis, (c) convencional y (d) Saw. El eje x corresponde a la tasa de FP's/ especificidad y el eje y a la tasa de TP's/sensibilidad. El recuadro contiene el valor de ROC/AUC a analizar. Fuente: autoría propia.	54

Capítulo 1

En este capítulo presenta el planteamiento del problema del proyecto, la hipótesis, el objetivo general, los objetivos específicos, la metodología, las actividades específicas que guían el presente trabajo de maestría, y, finalmente, las contribuciones realizadas a partir de su desarrollo.

1.Introducción

1.1. Definición del Problema

Según el observatorio de economía compleja del Instituto Tecnológico de Massachusetts [1], en 2017, Colombia recibió cerca de 39.1 mil millones de dólares por exportaciones, donde sus principales productos fueron petróleo crudo o minerales bituminosos y café. En particular, el café genera un promedio de 2.7 mil millones de dólares en ingresos anuales [1], lo cual tiene un alto impacto en la economía interna del país. En esa dinámica, dos perspectivas de la economía cafetera son manifiestas: en el ámbito nacional, los gobiernos impulsan políticas orientadas a mejorar la producción por cantidad [2], [3]; desde lo local, las cooperativas y los centros de acopio incentivan la compra de café de alta calidad. Los dos enfoques mencionados tienen un factor diferenciador: el proceso de producción [4]. La producción por cantidad o tradicional, tiene como fin aumentar la cantidad de café producido por hectárea haciendo uso de productos químicos [5]. Estos ayudan a regular el crecimiento de las plantas y de esta forma, garantizan un volumen producción por cosecha. Mientras que, la producción sostenible, incorpora el uso de prácticas con independencia parcial o total de productos químicos, con el fin de mejorar la calidad del café [6].

Aunque el enfoque tradicional puede aumentar la producción a corto plazo, el uso de productos químicos como fungicidas y pesticidas deteriora la calidad del café [1], generando problemas de salud en los consumidores y el medio ambiente [8]. Por el contrario, la producción sostenible trabaja en el beneficio ambiental, social y económico. Desde la perspectiva ambiental y social, diversos trabajos evidencian el impacto positivo de las alternativas de producción sostenible en el mejoramiento de la calidad del café, de los suelos y las condiciones sanitarias de los trabajadores [9]–[12]. Desde la perspectiva, económica otras investigaciones, mencionan la rentabilidad de los cultivos de producción sostenible, como el resultado de múltiples acciones relacionadas con prácticas agronómicas adecuadas y oportunas, además de una gestión eficiente de los agro insumos y la mano de obra empleada en la producción [13]. Las propuestas en esta dirección mencionan aspectos que afectan la rentabilidad en el proceso de producción sostenible [13]–[15], sin embargo, no llegan a modelar su impacto en la rentabilidad.

Debido a la importancia de la agricultura a nivel mundial, varios esfuerzos de investigación desde la Inteligencia Artificial (Artificial Intelligence - AI) han propuesto modelos basados en

variables numéricas para optimizar la producción de diversos cultivos, y como consecuencia, mejorar la rentabilidad. Por ejemplo, los trabajos propuestos en [2], [3], [16] predicen la producción en cultivos de plátano, arroz, frijol, entre otros, basados en variables meteorológicas y agronómicas con el fin de aplicar mecanismos para incrementar el volumen de producción por hectárea. Sin embargo, estos trabajos hacen uso de modelos de caja negra. Los modelos de caja negra son creados directamente de un conjunto de datos a partir de funciones complejas (generando planos multidimensionales) [17]. Esto dificulta la interpretación del modelo, llegando incluso a no entender como las variables se relacionan entre sí para alcanzar la predicción final [18]

Particularmente, para modelar la rentabilidad en la producción de café sostenible a pequeña escala, es necesario considerar algoritmos que permitan interpretar la interacción entre variables ya que el proceso de producción de café involucra a su vez subprocesos (*siembra, cuidado del cultivo, producción, beneficio, transformación*) y actividades [19]. La interacción entre actividades al interior de cada subproceso determina que tan sostenible y rentable es la producción, además de brindar información relevante para la toma de decisiones basadas en los intereses de los pequeños productores. Por esta razón, los modelos de caja negra mencionados, no resultan útiles para los caficultores y agrónomos que desean conocer la interacción entre variables asociadas a los procesos y subprocesos de producción sostenible [8]. Además, diversos autores afirman que alrededor del 78 % de los pequeños caficultores no acostumbran a tener registros históricos almacenados [20] . Esto dificulta la utilización de modelos basados en datos para estimar la rentabilidad en la agricultura a pequeña escala debido a la escasez de los mismos.

Teniendo en cuenta las anteriores consideraciones, es necesario enfocar los esfuerzos en proponer un modelo para estimar la rentabilidad en la producción de café sostenible que explique el comportamiento de sus variables, involucre aspectos cualitativos, cuantitativos y sea independiente de los conjuntos de datos para su creación. A partir de lo mencionado se plantea la siguiente pregunta de investigación:

¿Cómo proponer un modelo que explique el comportamiento de sus variables (cualitativas y cuantitativas) para la estimación de la rentabilidad de producción de café sostenible teniendo en cuenta la escasez de datos?

1.2. Hipótesis

La incorporación de lógica difusa en el tratamiento de variables cualitativas para un modelo jerárquico basado en conocimiento experto mejorará la estimación de la rentabilidad en la producción de café sostenible.

1.3. Objetivos

1.3.1. Objetivo general

Proponer un modelo cualitativo jerárquico basado en conocimiento experto para estimar la rentabilidad en la producción de café sostenible.

1.3.2. Objetivos específicos

- Construir una base de conocimiento a partir de información de rentabilidad por cosecha en la producción de café sostenible.
- Diseñar un modelo cualitativo jerárquico basado en conocimiento experto que permita representar la rentabilidad por cosecha en la producción de café sostenible.
- Implementar un sistema de lógica difusa en el modelo cualitativo jerárquico propuesto, con el fin de garantizar un grado de pertenencia a los valores cualitativos definidos.
- Evaluar el desempeño del modelo cualitativo jerárquico propuesto para un estudio de caso.

1.4. Metodología

El desarrollo del proyecto estará basado en dos componentes: el Modelo de Investigación Documental [21] y la metodología de BUCHANAN [22]. El Modelo de Investigación documental ofrece cualidades de gran adaptación para el propósito del presente proyecto propuesto. El modelo presenta cuatro fases de actividad:

A continuación, se describen el conjunto de componentes a desarrollar

- **Fase Preparatoria:** en esta fase será realizada una revisión del estado actual del conocimiento definiendo las etapas de investigación.
- **Fase Descriptiva:** esta fase identificará los lineamientos más importantes respecto a los métodos cualitativos jerárquicos.
- **Fase de Construcción:** esta fase seguirá la metodología BUCHANAN para el desarrollo del proyecto en el contexto de los métodos cualitativos jerárquicos a partir de 5 pasos. **1. Identificación** de participantes, recursos y fuentes; **2. Conceptualización** del área de estudio; **3. Formalización** o mapeo de la información conceptual; **4 Implementación:** Formalización del conocimiento en un ambiente de programación y **5. Evaluación:** observación del comportamiento del prototipo.
- **Fase de Extensión y Publicación:** en esta fase se realizará la divulgación de los resultados obtenidos en el desarrollo del proyecto en revistas indexadas.

El Modelo de Investigación documental permite la construcción del estado actual del conocimiento. En la construcción del modelo basado en conocimiento experto será utilizada la metodología BUCHANAN la cual consta de las siguientes fases de trabajo: identificación, conceptualización, formalización, implementación y validación.

1.5. Actividades generales y específicas

A continuación, se muestra una descripción detallada del conjunto de actividades que se desarrollarán para cumplir con el paquete de trabajo en relación a la metodología BUCHANAN (WP- Work Package)

WP1 Identificación y Conceptualización: Capítulo 1, 2 y 3 - Construcción de base de conocimiento relacionado con la producción de café sostenible.

- **Estado del arte:** recopilar información relacionada con la pregunta de investigación.
- **Recolección de información:** recopilar y analizar información sobre los parámetros relevantes basados en la producción de café sostenible que influye en el análisis de la rentabilidad por cosecha.
- **Análisis de alternativas:** definir criterios de evaluación que identifiquen los parámetros más relevantes que influyen en la estimación de la rentabilidad.
- **Selección de parámetros:** seleccionar los parámetros que trabajarán como entradas en la construcción del modelo de estimación de costos para determinar la rentabilidad por cosecha.
- **Construcción de la base de conocimiento experto:** condensar la información en una base de conocimiento experto.

WP2 Formalización e Implementación: Capítulo 4 - Diseño e implementación del modelo cualitativo jerárquico difuso

- **Diseño general:** diseñar el modelo cualitativo jerárquico difuso.
- **Implementación:** desarrollar la arquitectura propuesta.
- **Prototipo Final:** obtener un prototipo integrado y funcional de la propuesta-solución.

WP3 Evaluación: Capítulo 5 – Caso de Estudio y Evaluación

- **Selección del contexto:** seleccionar el estudio de caso bajo el cual ejecutar el experimento.
- **Definición:** definir el escenario de prueba a realizar en el estudio de caso.
- **Selección de Métricas:** seleccionar las métricas bajo las cuales evaluar el estudio de caso.

- **Diseño:** diseñar el escenario de prueba en el estudio de caso seleccionado.
- **Evaluación:** analizar y validar los resultados obtenidos.

WP4: Entregas

- **Documento final:** preparar el documento final del trabajo de grado.
- **Artículo de investigación:** preparar el artículo de investigación
- **Otras entregas:** entregar información completa del trabajo de grado, que consiste en el documento final y un repositorio digital.
- **Presentación final:** exponer la presentación final del trabajo de grado.

1.6. Contribuciones

1. El aporte investigativo del presente proyecto está centrado en la línea de investigación en e-@ambiente del Grupo de Ingeniería Telemática de la Universidad del Cauca y potencialmente beneficiará a la comunidad de pequeños y medianos agricultores, brindando soporte en la estimación de la rentabilidad en cultivos de café con producción sostenible a través de un modelo cualitativo jerárquico difuso.
2. Un modelo interpretable e independiente de grandes conjuntos de datos que permite estimar el nivel de rentabilidad en las producciones de café a pequeña escala en los países en desarrollo considerando los tres factores de rentabilidad del café (costo de producción, volumen de producción y mercado internacional) definidos por la Organización Internacional del Café (OIC) [17,18]. La independencia de grandes conjuntos de datos, la alcanza a través de la fase de conocimiento del contexto basada en el conocimiento experto, que permite crear un modelo conceptual sobre la rentabilidad, mientras que, la interpretabilidad, la logra a en la fase jerárquica-multicriterio que traduce el modelo conceptual en un modelo cualitativo jerárquico.
3. Una base de conocimiento resultante de la fase de conocimiento del contexto y adquirida a partir de la operación de la finca la Sultana y el panel de expertos.
4. La implementación de un sistema de lógica difusa en un modelo jerárquico cualitativo multicriterio para estimar el nivel de rentabilidad en las producciones de café a pequeña escala en los países.
5. La formalización de un conjunto de datos que contiene información de rentabilidad por cosecha en producciones de café sostenible.
6. Artículo científico titulado: "HMP-Coffee: A hierarchical multicriteria model to estimate the profitability for small coffee farming in Colombia" publicado en mes de julio, en la revista científica "Applied Sciences", homologada en la categoría Q1 de JSR y Publindex, en la categoría A1

Capítulo 2

2. Estado del Arte

Este capítulo presenta en la sección 2.1 los principales conceptos relacionados con la propuesta. Primero, una descripción de los conceptos relacionados al dominio de la agricultura: rentabilidad y sostenibilidad en producciones de café; segundo, los conceptos relacionados al área de investigación de la ingeniería: modelos jerárquicos multicriterio y lógica difusa. Posteriormente, la sección 2.2 describe los trabajos relacionados bajo dos enfoques: en primer lugar, las propuestas que utilizan métodos jerárquicos multicriterio para caracterizar la rentabilidad en cultivos agrícolas y, en segundo lugar, los enfoques que aplican modelos difusos en el sector agrícola.

2.1. Conceptualización

2.1.1. Desde el dominio de la Agricultura

a. Rentabilidad en la producción de café

La rentabilidad en el café hace referencia al beneficio obtenido por su producción, venta o comercialización [25], [26]. Según [27], la rentabilidad es el resultado de la relación entre factores internos y externos que intervienen en la producción agrícola. A nivel interno, las decisiones en materia de producción están compuestas por factores financieros (costos, inversiones, préstamos), físicos (construcciones y edificaciones), naturales (tierra, clima, biodiversidad) y sociales (comunidad, agrupamiento, parentesco). Por otro lado, los mercados, las políticas institucionales, los bienes públicos, los recursos naturales y las tecnologías hacen parte de los factores externos [26]. La adecuada relación entre estos factores, combinada con buenas prácticas agronómicas impactan positivamente la rentabilidad.

b. Producción de café sostenible

La producción de café sostenible [28] es aquella, que, de acuerdo con los estándares de calidad nacional, produce y procesa el café en un sistema alineado con el beneficio ambiental y social, y económico. Este tipo de producciones buscan mejorar el medio ambiente (p. Ej., Reducción de las emisiones de gases de efecto invernadero), la sociedad (p. Ej., Mejora de las condiciones de los trabajadores) y la economía (p. Ej., Mayor rentabilidad para los agricultores) como estrategias fundamentales para el cumplimiento de los objetivos de sostenibilidad [29]. En particular, desde la perspectiva económica, permitir mejoras en la rentabilidad general de los cultivos de café es un factor habilitador esencial para generar producciones más sostenibles.

2.1.2. Desde el dominio de investigación de la ingeniería

c. Modelos Jerárquicos Multicriterio (Hierarchical Multicriteria Model - HMM)

HMM es un método de análisis de decisiones que evalúa opciones determinando "la opción más apropiada". Un HMM se puede representar como la descomposición de un problema de decisión en sub-problemas más pequeños y menos complejos. Está compuesto por atributos y funciones de utilidad [30]. En un HMM, los atributos representan variables que toman valores de una escala cualitativa correspondiente (p. Ej., **Favorable**, **Promedio** y **Desfavorable**). Hay tres tipos: atributos básicos (i.e., variables externas), atributos agregados (i.e., variables internas) y atributos objetivos (i.e., variables de estudio o interés); están organizados jerárquicamente en un árbol de atributos. Las funciones de utilidad definen un conjunto de reglas mediante la combinación de las escalas cualitativas de los atributos básicos. Los HMM son modelos interpretables debido al uso de reglas lingüísticas comprensibles para los no especialistas [31], [32]. Las reglas a menudo se derivan del conocimiento de expertos, útiles en contextos con un alto nivel de experiencia en campo [33]. La **Figura 1** muestra la estructura genérica de un HM.

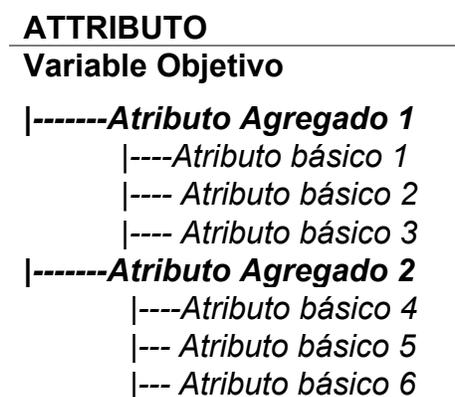


Figura 1. Estructura genérica de un HMM

d. Lógica difusa

La lógica difusa es una herramienta útil en el manejo de problemas con ambigüedades o alto grado de incertidumbre [31] e introduce el concepto de verdad parcial, donde el valor de verdad puede oscilar entre completamente verdadero y completamente falso. Al contrario, de la teoría clásica de conjuntos donde, el valor de verdad puede ser únicamente un número entero: "0" o "1" [32]. Estos modelos tienen la capacidad de reconocer, representar, manipular e interpretar información imprecisa y que carecen de certeza [33]. En el contexto particular de esta propuesta de maestría, la lógica difusa es útil para representar las percepciones de los calificadores plasmadas en escalas cualitativas (p. Ej., **Costos**, **Promedio** y **Económico**) con un grado de certeza para cada escala. Por ejemplo, **Costos** puede tener un grado de certeza de las 0,2 mientras que **Promedio** 0,7 y **Económico** 0,1. Esto significa que, para una entrada dada, la probabilidad de ser **Promedio** es mayor con un grado de pertenencia de 0,7. Estos valores son obtenidos de lo que en los sistemas de lógica difusa se denominan como funciones de membresía o pertenencia.

e. Modelos Multicriterio Difusos (Fuzzy Multicriteria Models - FMM)

Son modelos de análisis de decisiones que evalúa opciones determinando "la opción más apropiada" basados en la teoría difusa. Estos enfoques pueden ser vistos como métodos alternativos que incorporan información difusa representada por funciones de membresía y valores de pertenencia para definir el conjunto de posibilidades que determinan "la opción más apropiada" (e.i., el nivel de rentabilidad). En el contexto de esta propuesta de maestría, se busca utilizar los grados de pertenencia de las percepciones de los caficultores como base en la construcción del conjunto de posibilidades (ej. reglas difusas, matrices) que determinan el nivel de rentabilidad como **Favorable**, **Promedio** o **Desfavorable**.

2.2. Trabajos Relacionados

2.2.1. Modelos jerárquicos relacionados a rentabilidad en el sector agrícola

Varios enfoques jerárquicos han caracterizado la rentabilidad de diversos cultivos. De Salvo et al. [34] propusieron un enfoque jerárquico conceptual para mejorar la industria del vino mediante la evaluación del impacto de los criterios relacionados con el clima, las características de los viñedos, las características de los viticultores y las prácticas de gestión en la rentabilidad de este gremio en Europa del Este. Hawes et al. [35] definieron un enfoque jerárquico basado en árboles de decisión para estimar la rentabilidad del cultivo agroecológico mediante la combinación de los criterios de riesgo de producción (p. Ej., clima y riesgo de plagas y enfermedades) y rentabilidad potencial (p. Ej., márgenes brutos y subsidios directos).

Rising et al. [36] presentaron un enfoque jerárquico basado en redes bayesianas para evaluar el impacto de los criterios climáticos, como la estacionalidad de la temperatura, la precipitación anual, la estacionalidad de la precipitación y la fracción de riego, en el potencial de rentabilidad de varios cultivos, incluyendo cebada, soja, trigo, algodón, maíz y arroz. Shakoore et al. [37] propusieron un enfoque jerárquico de base numérica, que utiliza criterios relacionados con el clima y las variables de rendimiento para predecir cuales cultivos (entre, arroz aus, arroz aman, arroz boro, papa, yute y trigo) logran el máximo beneficio en algunas regiones de Bangladesh. Haaster et al. [38] introdujeron un enfoque jerárquico basado en la técnica de árboles de decisión, condiciones climáticas y actividades del mercado para predecir el momento oportuno de venta de los cultivos de flores en los Países Bajos.

Milne et al. [39] presentaron un enfoque jerárquico basado en la optimización de objetivos múltiples para simular el efecto de la aplicación de fertilizantes en la rentabilidad de algunos cultivos del Reino Unido. En este caso, la rentabilidad fue calculada como la diferencia entre los ingresos por rendimiento y los costos asociados con el fertilizante y su aplicación en el cultivo mismo; los autores probaron su enfoque en 20 cultivos, incluidas el cultivo de papa, la remolacha y cebolla. Cardozo et al. [40] desarrollaron un enfoque jerárquico basado en datos para el análisis multi-variable del clima (p. Ej., lluvia y región) y criterios de rendimiento (p. Ej., toneladas de caña por hectárea y azúcares recuperables totales) para determinar la rentabilidad en las producciones de caña de azúcar de Brasil [41].

Aubertot et al. [42] [43] presentaron un enfoque jerárquico basado en conocimiento experto para la predicción de infestación de la mancha del trigo marrón y óxido, con el fin de tomar acciones que disminuyan el impacto dichas enfermedades en la rentabilidad del cultivos de trigo.

El análisis de trabajos relacionados de la **Sección 2.2.1** revela diversos hechos:

De Salvo et al. [34] y Hawes et al. [35] propusieron enfoques multicriterios jerárquicos conceptuales que no ofrecen una implementación orientada a la rentabilidad para probarlos. Rising et al. [36], Shakoor et al. [37] y Haaster et al. [38] desarrollaron soluciones basadas en estructuras jerárquicas que no consideran datos económicos, como el costo de producción, el volumen de producción y el mercado internacional, para las estimaciones de rentabilidad. Además, estas soluciones no se centran en los cultivos de café. Milne et al. [39] y Cardozo et al. [40] propusieron enfoques basados en datos para estimar la rentabilidad en diversos cultivos considerando variables climáticas y de rendimiento. No obstante, la escasa adopción de herramientas modernas de gestión de la información en la caficultura a pequeña escala, es una barrera grave que dificulta la adopción de soluciones económicas agrícolas basadas en datos en los países en desarrollo [44]. Aubertot et al. [42][43] presentaron dos enfoques jerárquicos basado en conocimiento experto que no superan el 60% de precisión en la predicción.

2.2.2. Modelos difusos en agricultura

En esta subsección son presentados los trabajos que hacen uso de modelos difusos en el sector agrícola:

En la literatura existen diferentes trabajos enfocados en modelar el comportamiento de variables asociadas a los cultivos agrícolas a partir del concepto de lógica difusa. Por ejemplo, un sistema de evaluación difuso (Fuzzy Evaluation System - FES) ha sido utilizado para medir el nivel de sostenibilidad en la producción agrícola [45]. Una red neuronal difusa (Fuzzy Neural Network - FNN) ha sido utilizada para la notificación y control de riego de cultivos agrícolas, en función del tipo de producción, características del suelo y condiciones meteorológicas (temperatura, precipitación, etc.) [46]. Los mapas cognitivos difusos (Fuzzy Cognitive Maps - FCMs) han sido propuestos para estimar la variabilidad del rendimiento en cultivos de algodón y de coco en la región de malabar, al sur de India [47]. En Irán, un sistema de inferencia adaptativo neuro-difuso (Adaptive-Network-based Fuzzy Inference System - ANFIS) y un modelo de Regresión Lineal Difuso Jerárquico (Fuzzy Logic Regression - FLR) han sido empleados en el pronóstico de los ingresos por exportaciones de productos agrícolas y la demanda de energía en este sector [48], [49].

El análisis de trabajos relacionados de la **Sección 2.2.2** revela que:

Estas propuestas son **modelos de caja negra basados en datos** (construidos partir de funciones complejas). Por ejemplo, [48] define la ecuación de regresión lineal difusa utilizando datos históricos relacionados con la demanda de energía agrícola en Irán. Como lo mencionamos anteriormente, la escasa adopción de herramientas modernas de gestión de la información en la caficultura a pequeña escala es una barrera grave que dificulta la adopción de soluciones económicas agrícolas basadas en datos en los países en desarrollo[50].

2.2.3. Métodos multicriterio difusos en agricultura

En la literatura existen diferentes trabajos enfocados en modelar el comportamiento de variables asociadas a los cultivos agrícolas a partir de métodos multicriterio difusos. Por ejemplo, Prakash et al. [51] propusieron un enfoque multicriterio espacial dentro de un ambiente difusos para identificar la eficiencia de los cultivos de Kharif al sur de la India. El estudio considera características del terreno y el ambiente, incluida la topografía, las propiedades del suelo, la vegetación, el clima y los factores socioeconómicos. Anagnostopoulos et al. [52] aplicaron un enfoque multicriterio difuso de costo-beneficio para la selección del proyecto óptimo en la gestión racional de recursos hídricos del río Nestos en Grecia. La selección del proyecto óptimo se hizo basada en información climática como las precipitaciones promedio y máximas anuales. Elaalem et al. [53] plantearon la uso de un método de evaluación multicriterio en un ambiente difuso con el propósito de evaluar índices de calidad de los suelos para la producción de aceitunas en la llanura de Jeffara en Libia. Esta evaluación considero factores del suelo incluidos la capacidad de retención de agua disponible, profundidad, infiltración, textura, salinidad y reacción. Ozkan et al. [54] propusieron un método multicriterio basado en GIS en un ambiente difuso utilizando datos del satélite Landsat 8 para identificar áreas potenciales adecuadas para el desarrollo de actividades agrícolas en el ecosistema terrestre semiárido de la región de Anatolia central. Elleuch et al. [55] han desarrollado un enfoque combinado basado en métodos difusos multicriterio y programación de optimización matemática para la gestión de la agricultura de regadío basados en criterios de fiabilidad, disponibilidad y adaptabilidad del recurso hídrico. Chen et al. [56] propusieron un modelo multicriterio de promediado ponderado lingüístico difuso (FLOWA- Fuzzy Linguistic Weighted Averaging) que utiliza criterios de topografía, paisaje y características de suelo con el propósito de determinar la idoneidad de tierras de regadío para agricultura en el monte de Gambier en Australia.

El análisis de trabajos relacionados de la **Sección 2.2.3** revela que:

Estas investigaciones proponen el uso de lógica difusa sobre modelos de multicriterio de un “único” nivel de profundidad. La **profundidad** hace referencia al número de capas o niveles que posee un modelo en su estructura interna: a mayor profundidad se generan relaciones más específicas [57]. Cuando los modelos con bajo nivel de profundidad, aumentan significativamente el número de variables de entrada, las relaciones internas crecen exponencialmente. Como resultado, se obtienen modelos de alta complejidad. En particular, la producción de café sostenible está compuesta por varios procesos (donde cada proceso tiene a su vez subprocesos con sus propias entradas) [21]. Este alto número de entradas aplicado a un modelo con un “único” nivel de profundidad, produce soluciones de alta complejidad que impide la fácil interpretación del comportamiento de las variables al interior del modelo [21].

A continuación, la **Tabla 1** condensa los hechos revelados en la sección 2.2.1, 2.2.2 y 2.2.3 para finalmente presentar el enfoque de la propuesta del presente trabajo de maestría.

- **Conceptualización e Implementación:** presentación conceptual y ejecución del enfoque propuesto
- **Modelo no dirigido por datos:** enfoque modelado parcial o totalmente independientes de grandes conjuntos de datos.
- **Interpretabilidad:** nivel de entendimiento de un proceso de decisión particular.

- **Precisión:** métrica de desempeño que representa el funcionamiento general del modelo de estimación.
- **Profundidad:** número de capas o niveles que posee un modelo en su estructura interna.

Tabla 1. Resumen análisis de trabajos relacionados

	Conceptualización e Implementación	No dirigido por datos	Interpretabilidad	Precisión	Profundidad
De Salvo et al. [34] Hawes et al. [35]	*	*	✓	*	*
Rising et al. [36] Shakoor et al. [37] Haaster et al. [38]	*	*	✓	*	✓
Milne et al. [39] Cardozo et al. [40]	*	*	✓	*	✓
Aubertot et al. [42] Robín et al. [43]	*	✓	✓	*	✓
Cornelissen et al. [45] Giusti et al. [58] Mohapatra et al. [46] Jayashree et al. [47] Mohaddes et al. [49] Taghizadeh et al. [48]	*	*	*	✓	*
Prakash et al. [51] Anagnosto et al. [52] Elaalem et al. [53] Ozkan et al. [54] Elleuch et al. [55] Chen et al. [56]	*	*	✓	✓	*
Propuesta	✓	✓	✓	✓	✓

La **Tabla 1** evidencia que, hasta el momento, no se ha desarrollado un modelo conceptual con su respectiva implementación, que permita la estimación de la rentabilidad cumpliendo las premisas relevantes para el caso de pequeñas producciones de café: un enfoque no dirigido por datos, con alto nivel de interpretabilidad (también relacionado con la profundidad) y garantizando un nivel de precisión en la estimación. En particular, las propuestas de Aubertot et al. [42] [43] aportaron los enfoques con el mayor número de características relevantes para este trabajo (i.e., no dirigido por datos, interpretables y con varios niveles de profundidad), sin embargo, son modelos que no superan el 60% de precisión en sus resultados.

2.2.4. Resumen

Este capítulo presentó los principales conceptos teóricos de la investigación, tales como: rentabilidad en la producción de café, producción de café sostenible, modelos jerárquicos multicriterio, lógica difusa y modelos multicriterio difusos en agricultura. Luego, fueron expuestos los trabajos de investigación en tres líneas: Modelos jerárquicos relacionados a rentabilidad en el sector agrícola, modelos difusos en agricultura y modelos multicriterio en agricultura. Finalmente, se consensó en la **Tabla 1** los hechos revelados en las tres líneas de

investigación, donde se definió finalmente el enfoque de la propuesta. Se concluyó que, hasta el momento, no se había desarrollado un modelo conceptual con su respectiva implementación, que permitiera la estimación de la rentabilidad cumpliendo tres premisas relevantes para el caso de pequeñas producciones de café: un enfoque no dirigido por datos, con alto nivel de interpretabilidad (también relacionado con la profundidad) y garantizando un nivel de precisión en la estimación. Además, las propuestas de Aubertot et al. [42] [43] aportaron los enfoques con el mayor número de características relevantes para este trabajo (i.e., no dirigido por datos, interpretables y con varios niveles de profundidad), sin embargo, son soluciones que no superan el 60% de precisión en sus resultados.

Capítulo 3

3. Modelo jerárquico multicriterio de rentabilidad

En este capítulo presenta la construcción del modelo jerárquico multicriterio de rentabilidad HMP-Coffee, un enfoque útil para estimar el nivel de rentabilidad en pequeñas producciones de café.

3.1. Área de estudio y panel de expertos

El modelo HMP-Coffee se basa en el funcionamiento de la finca cafetera "La Sultana", perteneciente a la Universidad del Cauca. La finca está ubicada en Timbío, Cauca, Colombia ($2^{\circ} 22'28.51''\text{N}$, $76^{\circ} 43'31.89''\text{W}$), a 1700 m.s.n.m. con una superficie total de 13,5 hectáreas, de las cuales 4,9 hectáreas son de producción de café [59]–[61]. En 2006, La Sultana migró hacia un cultivo de producción de café sostenible a través de un sistema de procesamiento ecológico y manejo de subproductos. Además, la producción, manejo y transformación del café se desarrollan con buenas prácticas para producir café de calidad bajo el concepto

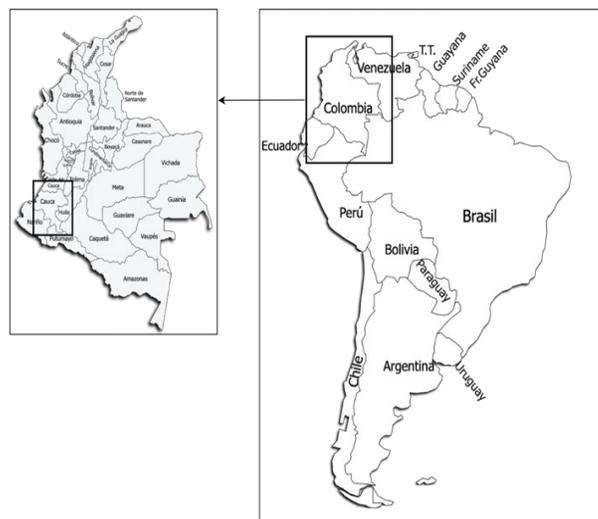


Figura 2. Ubicación de la finca de café de Colombia La Sultana ($2^{\circ} 22'28.51''\text{N}$, $76^{\circ} 43'31.89''\text{W}$). Fuente : www.proyectomamundi.com.

de sostenibilidad ambiental, social y económica [62]. La Sultana reúne tres características esenciales que aborda el modelo HMP-Coffee: (i) producción de café a pequeña escala, (ii) incorporación de buenas prácticas, principalmente para mejorar la sostenibilidad económica; y (iii) ubicado en un país en desarrollo. (Colombia; ver **Figura 2**).

Diferentes expertos estuvieron involucrados en la construcción del modelo HMP-Coffee. Estos expertos contribuyeron con su conocimiento en la definición de los atributos de entrada,

agregación, la salida del modelo y los valores cualitativos correspondientes. El equipo estuvo compuesto por seis expertos en agricultura, con más de 10 años de experiencia en el sector económico cafetero. El rol de cada experto se detalla en la **Tabla 2**.

Tabla 2 Equipo involucrado en la concepción de HMP-Coffee.

Experto	Área de especialización	Años de experiencia	Experiencia	Organización
1	Problemas económicos agroindustriales	28	La transformación agroindustrial y la economía de la actividad cafetera.	Universidad del Cauca
2	Agroindustria	10	Estrategias para la producción y comercialización de café orgánico en países en desarrollo	Finca de café experimental La Sultana
3	Gestión Agrícola	20	Implementación de estrategias de sustentabilidad en la finca experimental "La Sultana" y certificaciones de café sustentable (p. Ej., Rainforest Alliance)	Finca de café experimental La Sultana
4	Agronomía, suelos y agua	30	Estimación de indicadores técnicos y económicos sobre el beneficio del café, caracterización de fincas integrales y productivas y producción de café orgánico	Universidad del Cauca
5	Ingeniería agronómica	15	Optimización de los procesos de transformación del café, cosecha y pos cosecha de café de calidad.	Tecnicafe
6	Gestión Agrícola	18	Prácticas para el manejo y administración de cultivos de café en pequeños cultivos de café.	Federación Nacional de Cafeteros

3.2. Construcción del modelo HMP-Coffee

A continuación, se presenta el modelo HMP-Coffee para estimar el nivel de rentabilidad en pequeñas producciones de café. El modelo incluye dos fases: i) **Fase de conocimiento del contexto** para crear un modelo conceptual sobre la rentabilidad en la producción de café a pequeña escala compuesta por dos pasos: granularidad gruesa y granularidad fina y ii) la **Fase jerárquica-multicriterio convencional** para convertir el modelo conceptual en un modelo cualitativo jerárquico comprensible y ejecutable capaz de estimar el nivel de rentabilidad en pequeñas producciones de café; desarrollada en tres pasos: la definición de escalas, pesos y reglas [63]. HMP-Coffee involucra a las siguientes partes interesadas: el experto en economía agrícola, el ingeniero del conocimiento y el tomador de decisiones o pequeño productor. El experto en economía agrícola interviene en la granularidad fina (e.i., el análisis detallado del dominio de aplicación), la definición de escalas, pesos y reglas para estructurar el modelo conceptual. El ingeniero del conocimiento ejecuta el paso de granularidad gruesa (e.i., el análisis general del dominio de aplicación) y la implementación del modelo multicriterio jerárquico. Finalmente, el tomador de decisiones (p. Ej., el pequeño productor) configura HMP-Coffee proporcionando los datos de entrada del modelo (p. Ej., inversión en fertilizantes, tiempo dedicado a las actividades de cuidado de cultivos y volumen de producción de café) y obtiene de este, la estimación del

nivel de rentabilidad (i.e., **Favorable**, **Promedio**, **Desfavorable**). Con base en esta estimación, el pequeño productor podrá tomar decisiones sobre el cultivo variando el comportamiento de los datos de entrada para mejorar el nivel de rentabilidad del café. Las dos fases se presentan en la **Figura 3**.

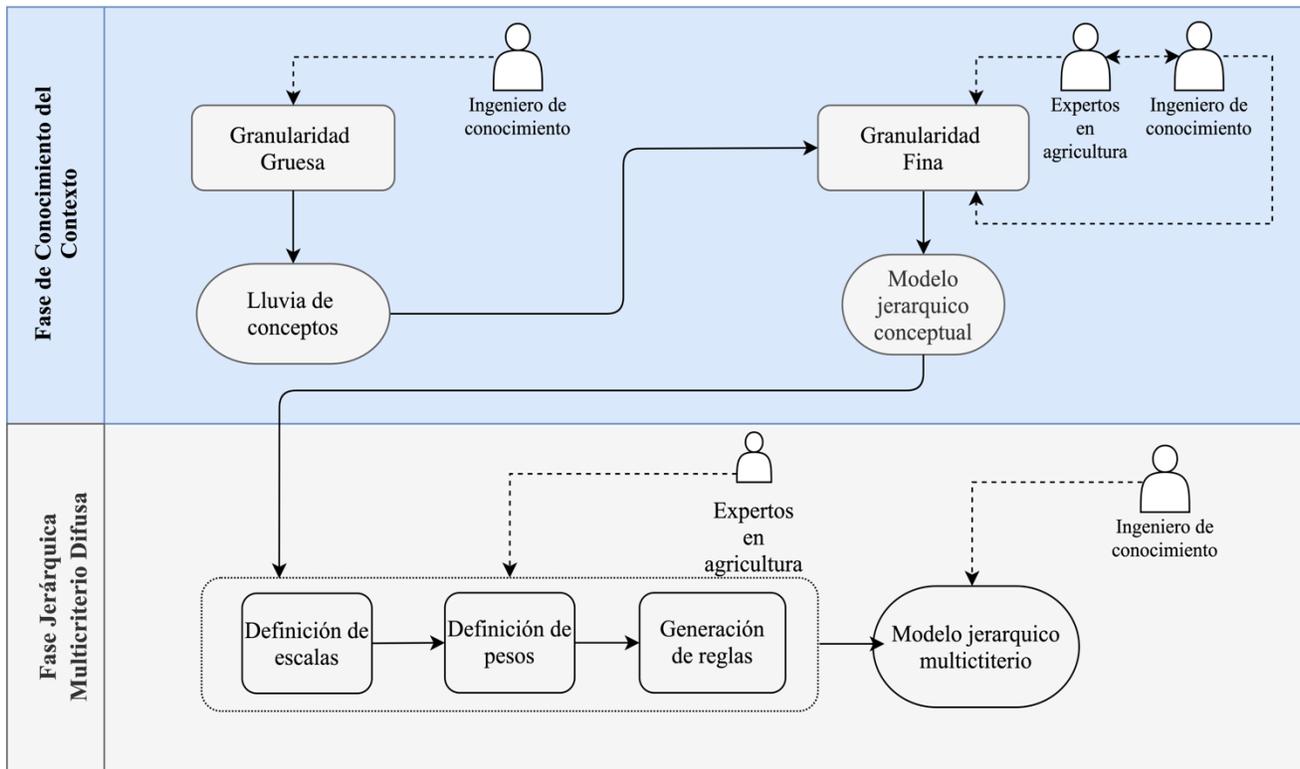


Figura 3. Proceso conceptual para la construcción del modelo HMP-Coffee. Fase 1: Conocimiento del contexto; Fase 2: multicriterio jerárquico. Fuente: autoría propia.

3.2.1. Fase de conocimiento del contexto: Análisis multinivel de granularidad

Los aspectos clave relacionados con la rentabilidad en los cultivos de café a pequeña escala se recopilan de diferentes fuentes de conocimiento (p. Ej., artículos científicos, informes técnicos y expertos). A partir de esos aspectos, se propone un modelo conceptual jerárquico. Con el fin de construir el modelo conceptual, es utilizado un análisis de granularidad multinivel útil para el descubrimiento de conocimientos desde una perspectiva global a una específica [64]. El análisis de granularidad multinivel implica dos pasos, la granularidad gruesa y la granularidad fina.

a. Granularidad Gruesa

La **Granularidad Gruesa** permite la construcción de una lluvia de conceptos relacionados con la “rentabilidad en pequeñas producciones de café”. Este paso incluyó tres actividades: *selección documental* para identificar la documentación disponible sobre la rentabilidad en la pequeña producción de café, *análisis textual* para extraer los conceptos más frecuentes de la

documentación y la *conceptualización iterativa* para identificar los conceptos más importantes de la información analizada [65].

Selección documental: para cubrir los aspectos relevantes de la estimación de la rentabilidad en pequeñas producciones de café, se definieron tres cadenas de consulta: producción de café a pequeña escala (Small Coffee Production-SCP), rentabilidad económica del café (Economic Coffee Profitability-ECP) y, rentabilidad de Café a pequeña escala (Small Coffee Production Profitability-SCPR). Las cadenas de consulta fueron ejecutadas sobre bases de datos científicas (i.e., SCOPUS, Science Direct y Google Scholar) y páginas oficiales sobre la actividad cafetera, como ICO [23], Instituto del Café de Costa Rica (ICAFFE) [66], y Caravela Coffee [67].

La **Figura 4** muestra los 45 documentos resultantes de la actividad de selección documental, incluyendo 19 informes técnicos, 19 artículos científicos y 7 documentos de literatura gris. El SCP se centró en documentos de producción de café a pequeña escala, con 20 resultados: 8 informes técnicos, 10 artículos científicos y 2 documentos de literatura gris. ECP se centró en los documentos relacionados con las variables de rentabilidad económica para la producción de café, con 15 resultados: 7 informes técnicos, 5 artículos científicos y 3 documentos de literatura gris. Finalmente, el SCPR se centró en documentos relacionados con la rentabilidad y los cultivos de café en pequeña escala, con 10 resultados: 4 informes técnicos, 4 artículos científicos y 2 documentos de literatura gris.

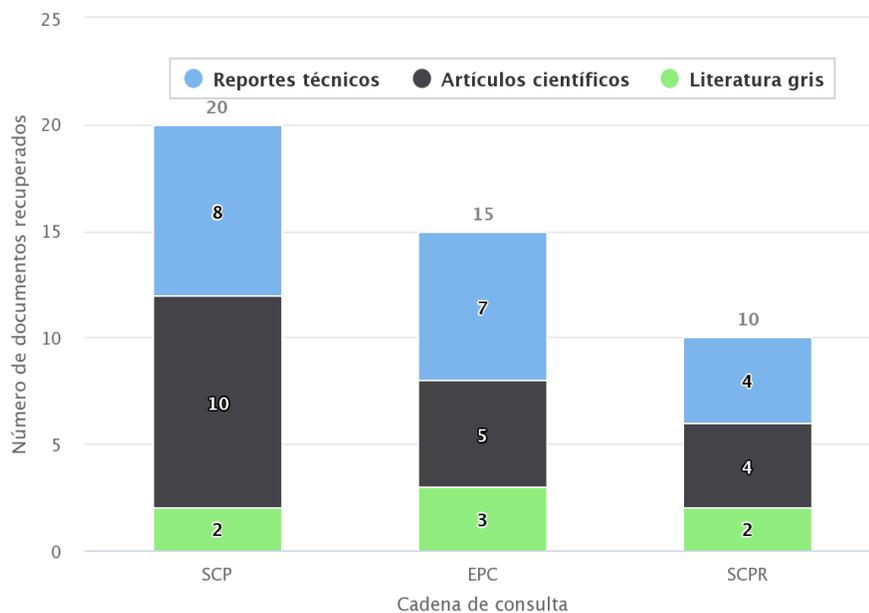


Figura 4. Documentos recuperados para las tres cadenas de consulta: Producción de café a pequeña escala (Small Coffee Production-SCP), Rentabilidad económica del café (Economic Coffee Profitability-ECP) y, Rentabilidad de Café a pequeña escala (Small Coffee Production Profitability-SCPR). Fuente: autoría propia.

científicos y 2 documentos de literatura gris.

Análisis textual: el método estadístico " Frecuencia de palabras " se utiliza para la ejecución de la actividad de análisis textual [68]. El método identifica los conceptos más frecuentes en una

colección de documentos [65]. La **Figura 5** describe el proceso realizado en este análisis utilizando la herramienta estadística, R versión 1.2.50 [69]. Esta técnica implica la identificación de los conceptos iterativos sobre un grupo de documentos (i.e., los 45 documentos recopilados), la eliminación de palabras redundantes o "palabras vacías" (p. Ej., artículos, pronombres y preposiciones) y su selección. Como resultado, la actividad proporciona alrededor de 40 términos.

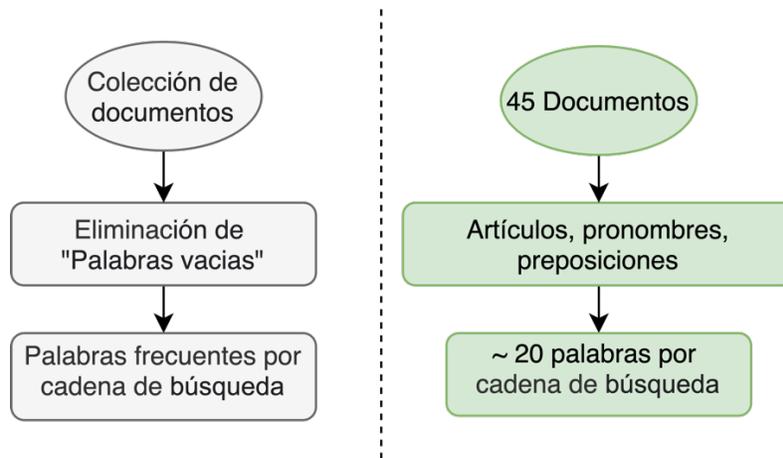


Figura 5. Proceso de análisis de texto. 1. Actividades genéricas del análisis textual (diagrama de flujo color blanco) 2. Actividades instanciadas para el modelo HMP-Coffee (diagrama de flujo color verde). Fuente: autoría propia.

Conceptualización iterativa: los conceptos más frecuentes se mapearon por cadena de consulta en un diagrama de Venn como se muestra en la **Figura 6**. Posteriormente, se seleccionaron las palabras ubicadas en las regiones superpuestas (palabras en negrita representadas en la **Figura 6**). Posteriormente, se identificaron 24 términos, incluidos conceptos como **mano de obra** perteneciente a cadenas de consulta de SCP y EPC; **labores** del cultivo perteneciente a EPC y SCPR; fertilizantes **químicos** a SCPR y SCP, y **cosecha** como intersección de las tres cadenas de búsqueda.

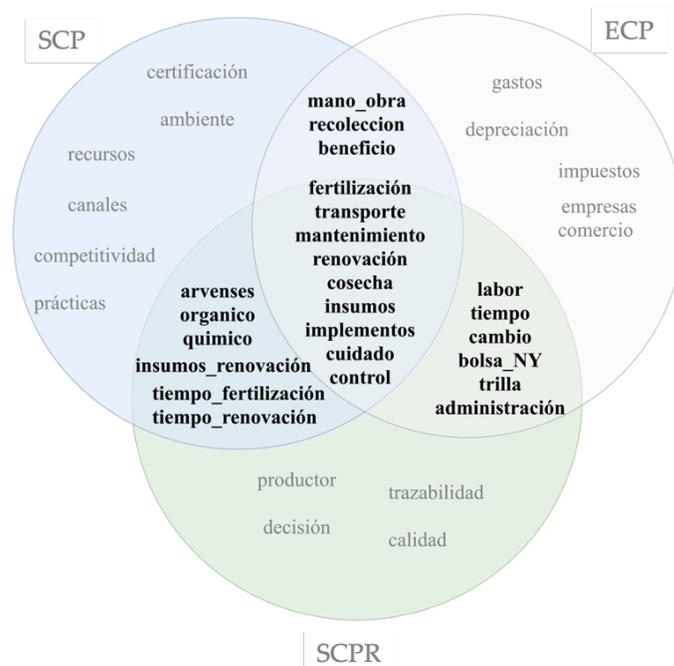


Figura 6. Diagrama de Venn de los conceptos frecuentes. Los colores azul, gris y verde representan las regiones por cadenas de consulta: Producción de café a pequeña escala (SCP), Rentabilidad económica del café (ECP) y Rentabilidad del café a pequeña escala (SCPR). Las 24 palabras ubicadas en las regiones de superposición se seleccionan como lluvia de conceptos de HMP-Coffee (palabras en negrita)

Adicionalmente, la **Tabla 3** muestra la lista de la lluvia de conceptos final. Para cada concepto se define: id, nombre, descripción y frecuencia.

Tabla 3.Lista de conceptos de HMP-Coffee.

ID	CONCEPTO	DESCRIPCIÓN	FRECUENCIA
T1	Tiempo en Fertilización	Tiempo empleado en fertilizar el cultivo.	80
T2	Tiempo en Renovación	Tiempo invertido en la renovación del cultivo.	70
T3	Tiempo en Control (Arvenses/Plagas & enfermedades)	Tiempo dedicado al control de malezas / plagas y enfermedades en el cultivo.	34
T4	Mano de Obra	Costo del salario	190
T5	Fertilizantes químicos	Insumos para fertilización química.	220
T6	Fertilizantes orgánicos	Insumos para la fertilización orgánica.	153
T7	Insumos arvenses	Entradas para control químico mediante parches con el selector de malezas, control con machete o guadaña.	40
T8	Insumos para renovación	Insumos para renovación del cultivo.	32
T9	Mantenimiento	Gastos de mantenimiento de instalaciones.	124

T10	Trasporte	Transporte de suministros y café.	27
T11	Recolección	Pago de la cosecha de café por kilogramos, bushel, entre otras medidas.	87
T12	Implementos	Provisión de recolectores de café (p. Ej., canasta, guantes, etc.)	13
T13	Beneficio	Proceso de beneficio del café.	34
T14	Trilla	Proceso de trilla del café	18
T15	Tasa de Cambio	La relación entre una moneda (p. Ej., COP, CRC) y otra (p. Ej., USD).	45
T16	Bolsa de NY	Precio del café en las acciones de Nueva York.	16
T17	Volumen	El volumen de la cosecha en @ / ha.	72
T18	Tiempo	Término general que se refiere a la inversión de tiempo en el cuidado del cultivo.	28
T19	Labores	Todos los aspectos relacionados con el trabajo del cultivo, incluido el tiempo de trabajo, el precio por salario, etc.	54
T20	Insumos	Todos los suministros utilizados en labores de cuidado de cultivos.	99
T21	Gestión	Costos de manejo de la finca, incluyendo pagos por servicios, mantenimiento, entre otros.	42
T22	Cosecha	Todos los aspectos relacionados con la cosecha, incluida la recolección, provisión de recolectores de café, entre otros.	90
T23	Cuidado	Todos los aspectos relacionados con el cuidado de los cultivos, incluidos los insumos, el pago de salarios, plagas y control de enfermedades, entre otros.	19
T24	Fertilización	Actividad de fertilización, incluida la fertilización orgánica y química.	99

b. Granularidad Fina

En el paso de granularidad fina, hace uso de los términos obtenidos de la lluvia de conceptos para crear el modelo jerárquico conceptual de HMP-Coffee siguiendo dos actividades: Análisis de relaciones para identificar relaciones entre conceptos y Estructuración del modelo para organizar la información en una estructura jerárquica.

Análisis de relaciones: para definir los niveles de relación y generalidad entre la lluvia de conceptos y la rentabilidad, se aplicó la técnica de Rejilla (Repertory Grid - RG) [70]. Esta técnica ubica un grupo de términos en las filas y otro en las columnas. Los expertos asignan una puntuación a las intersecciones que corresponde a un nivel de relación y generalidad entre conceptos (dados desde su experticia, ver **Anexo 1**). Para construir HMP-Coffee, creamos dos rejillas de la siguiente manera:

La primera rejilla determina el nivel de relación entre la variable objetivo (i.e., la rentabilidad) y la lluvia de conceptos. Las filas contienen los 24 términos y las columnas, tres factores que caracterizan la rentabilidad en producciones de café: el costo de producción, el volumen de

producción y el mercado (renombradas como variables globales - gv1, gv2, gv3). Estas variables globales son seleccionadas adoptando la definición formal de rentabilidad, propuesta por la ICO y la Federación Nacional de Cafeteros de Colombia (FNC) [24], [71] que la presenta como la diferencia entre los ingresos (asociados al volumen de producción y al mercado) y la inversión (asociada al costo de producción) [24]. Cada celda de la Rejilla 1 (ver **Tabla 4**) contiene el puntaje promedio asignado por el panel de expertos para relacionar dichos conceptos. Para esto, cada experto proporcionó puntajes en una escala de 1 a 4, donde (1) Sin relación directa, (2) Relación baja, (3) Relación media y (4) Relación alta. En la Rejilla 1, las celdas verdes resaltan la variable global más relacionada para cada término (p. Ej., t1 tiene mayor relación con la variable global costo de producción-gv1).

La segunda rejilla determina el nivel de generalidad de la lluvia de conceptos. Las filas representan los 24 términos y las columnas, los cuatro niveles de generalidad: muy alta generalidad (vhg), alta generalidad (hg), media generalidad (mg) y específica (s). Cada celda de la Rejilla 2 (ver **Tabla 4**) contiene el puntaje de generalidad promedio asignado por el panel de expertos (**Tabla 2**). Cada experto proporcionó puntajes en una escala de 1 a 4, donde (1) Sin relación directa, (2) Relación baja, (3) Relación media y (4) Relación alta. En la Rejilla 2, las celdas azules resaltan el nivel de generalidad más ponderado para cada término de la lluvia de conceptos (p. Ej., t1 logra la generalidad más baja: Nivel 6_Específico. La **Tabla 4** mapea los resultados del nivel de relación (ver Rejilla 1) y el nivel de generalidad (ver Rejilla 2). Adicionalmente, con el fin de mejorar la comprensión de los términos de HMP-Coffee, propusimos nombres descriptivos para los 24 conceptos. Desde este punto, los utilizaremos para referirnos a cada uno de ellos.

Tabla 4.Rejilla de resumen de HMP-Coffee. La rejilla 1 analiza el nivel de relación. La rejilla 2 analiza el nivel de generalidad.

ID	Concepto	Nombres descriptivos	Rejilla 1 Nivel de Relación			Rejilla 2 Nivel de Generalidad			
			gv1	gv2	gv3	vhg	hg	mg	s
t1	Tiempo en Fertilización	Tiempo invertido en labores de Fertilización	3,4	2,8	1,6	3,4	2,8	2,8	3,8
t2	Tiempo en Renovación	Tiempo invertido en labores de Renovación	3,2	2,6	2	2,6	3,4	3,4	4
t3	Tiempo en Control (Arvenses/Plagas & enfermedades)	Tiempo invertido en labores de Control (Arvenses/Plagas & enfermedades)	3,2	2	2,4	2	2,6	2,6	4
t4	Mano de Obra	Precio de mano de obra	3,8	1,6	2	1,6	3	3,2	3,4
t5	Fertilizantes químicos	Inversión en fertilizantes químicos	3,2	1,6	3	2	2,8	3	3,4
t6	Fertilizantes orgánicos	Inversión en fertilizantes orgánicos	3,4	2	1,6	1,6	2,6	2,8	3,2
t7	Insumos arvenses	Inversión en insumos para el control de arvenses	3,2	2,8	2	3,4	2	3,4	2,8
t8	Insumos para renovación	Inversión en insumos renovación del cultivo	3,4	3,4	2	2,6	2,8	2,6	3,4
t9	Mantenimiento	Inversión en mantenimiento	3,2	2,6	2,8	2	3,4	3,6	3,8
t10	Trasporte	Gastos de transporte	3,8	2	3,4	1,6	2,6	2,8	3,2
t11	Recolección	Pago por recolección de café	3,4	1,6	2,4	3	2	3,4	3

t12	Implementos	Implementos para recolectores de café	2,6	2	2	2	3,4	2,6	2,8
t13	Beneficio	Inversión en proceso de beneficio de café	2,6	1,6	3	2	2,6	2	2,4
t14	Trilla	Inversión en proceso de trilla de café	3	2,6	1,2	1,6	2	1,6	2,6
t15	Tasa de Cambio	Variación del Precio del Dólar	2	2,6	3,8	2	1,6	2	2,6
t16	Bolsa de NY	Precio del Café en la bolsa de valores de NY	3	2,6	3,4	3,2	2,6	2	3,4
t17	Volumen	Volumen de producción de café	1,6	4	2	2	1,6	2,6	2,8
t18	Tiempo	Tiempo de ejecución en actividades del cultivo	1,6	2	1,6	1,6	2,6	3,4	2
t19	Labores	Labores del Cultivo	2	2	2	2	3,8	3	2,6
t20	Insumos	Inversión en insumos	2,8	1,6	1,6	3	3,4	2,8	2
t21	Gestión	Gastos de gestión	3,4	2	2,6	3,4	2,8	3	2,6
t22	Cosecha	Inversión en tiempo de cosecha	2,6	2,8	2,6	3,8	3,4	2,6	2
t23	Cuidado	Inversión en actividades de cuidado de cultivos	2	2,6	2	3,8	2,6	2	3,4
t24	Fertilización	Inversión en actividades de fertilización	3	2,8	1,6	2,8	3	3,4	2

Estructuración del modelo: una vez identificada la relación (celdas verdes en la **Tabla 4**) y la generalidad (celdas azules en la **Tabla 4**), estructuramos el modelo jerárquico mostrado en la **Figura 7**.

Por niveles (generalidad):

- **El nivel 1 (variable de salida)** incluye el nivel de rentabilidad (vo).
- **El nivel 2 (variables globales)** contiene el costo de producción (gv1), el volumen de producción (gv2) y el mercado (gv3).
- **El nivel 3 (Generalidad muy alta)** contiene gastos de gestión (t21), inversión en tiempo de cosecha (t22) e inversión en actividades de cuidado de cultivos (t23)
- **El nivel 4 (generalidad alta)** contiene las labores del Cultivo (t19) y la inversión en insumos (t20).
- **El nivel 5 (Generalidad Media)** contiene tiempo de ejecución en actividades de cultivo (t18), inversión en actividades de fertilización (24).
- **El nivel 6 (específico)** contiene el tiempo dedicado a la fertilización (t1), el tiempo dedicado a la renovación (t2), el tiempo dedicado al control de arvenses / plagas y enfermedades (t3), el precio de la mano de obra (t4), la inversión en fertilizantes químicos (t5), inversión en fertilizantes orgánicos (t6), inversión en insumos para el control de arvenses (t7), inversión en insumos de renovación (t8), inversión en mantenimiento (t9), gastos de transporte (t10), pago por recolección de café (t11), implementos para recolectores de café (12), inversión en proceso de beneficio de café (t13), inversión en proceso de trilla de café (t14), variación del precio del dólar (t15) y precio del café en la bolsa de valores de nueva york (t16), y volumen de producción de café (t17).

Por relacionamiento entre conceptos:

- **La rentabilidad** está relacionada con el costo de producción (gv1), el volumen de producción (gv2) y el mercado (gv3).
- **El costo de producción** está relacionado con 20 conceptos: tiempo dedicado a la fertilización (t1) / nivel 6, tiempo dedicado a la renovación (t2) / nivel 6, tiempo dedicado al control de malezas / plagas y enfermedades (t3) / nivel 6, precio de mano de obra (t4) / nivel 6, inversión en fertilizantes químicos (t5) / nivel 6, inversión en suministros orgánicos (t6) / nivel 6, inversión en suministros de deshierbe (t7) / nivel 6, inversión en insumos de renovación (t8) / nivel 6, inversión en mantenimiento (t9) / nivel 6, gastos de transporte (t10) / nivel 6, pago por recolección de café (t11) / nivel 6, implementos para recolectores de café (t12) / nivel 6, inversión en proceso de beneficio de café (t13) / nivel 6, inversión en el proceso de trilla del café (t14) / nivel 6, tiempo de ejecución en actividades de cultivo (t18) / nivel 5, labores del cultivo (t19) / nivel 4, inversión en insumos (t20) / nivel 4, gastos de gestión (t21) / nivel 3, inversión en tiempo de cosecha (t22) / nivel 3, inversión en actividades de cuidado de cultivos (t23) / nivel 3 e inversión en actividades de fertilización (t24) / nivel 5.
- **El volumen de producción** tiene relación con el concepto volumen de producción de café (t17) / nivel 6.
- **Mercado** tiene asociación con dos conceptos: variación del precio del dólar (t15) / nivel 6 y precio del café en la bolsa de valores de Nueva York (t16) / nivel 6.

La **Figura 7** presenta el modelo jerárquico conceptual HMP-Coffee obtenido por la fase de conocimiento del contexto.

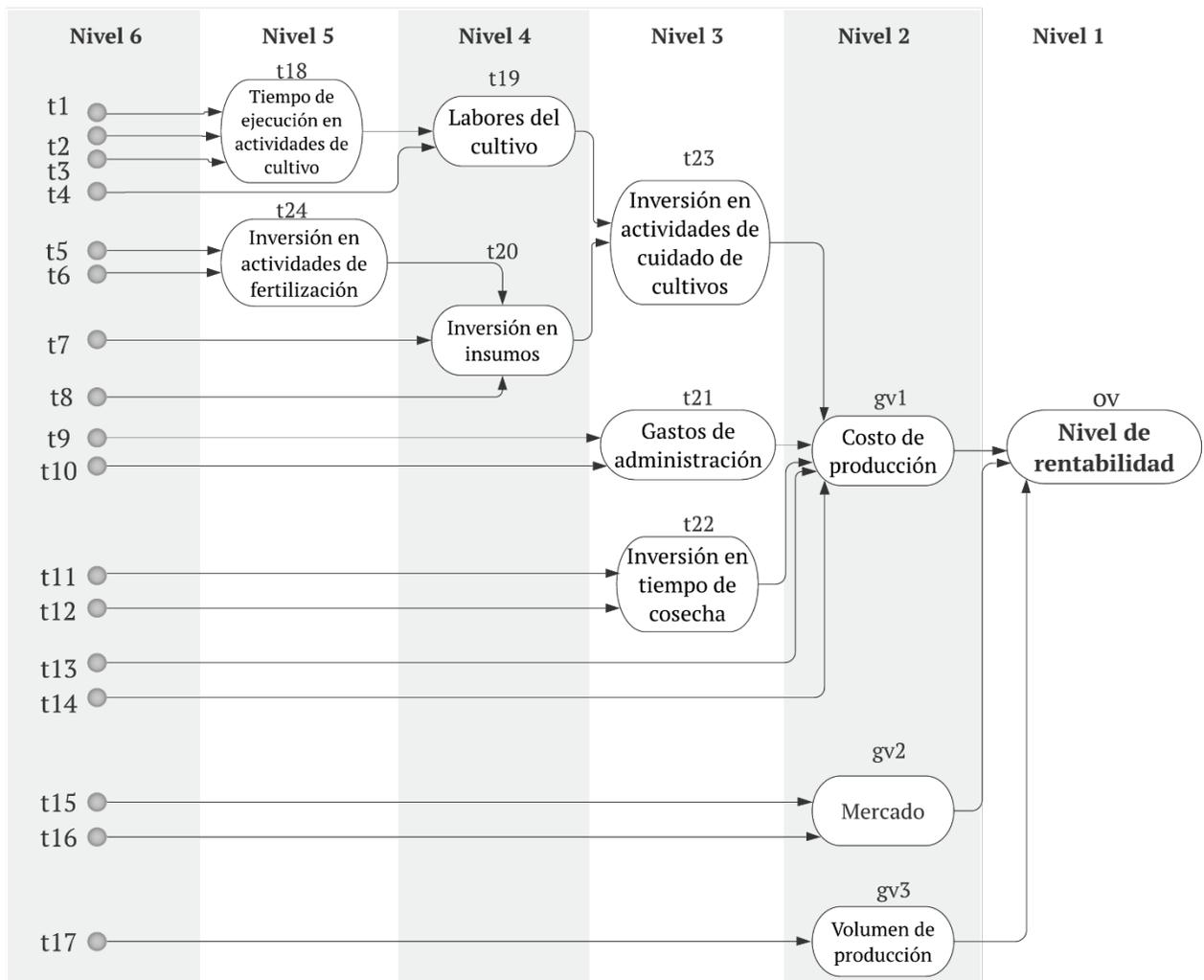


Figura 7. Modelo conceptual de rentabilidad HMP-Coffee . Fuente: autoría propia.

3.2.2. Fase Jerárquica Multicriterio Convencional

Para convertir el modelo conceptual en un modelo de estimación de rentabilidad, es utilizada la herramienta para la toma de análisis de decisiones denominada DEX (Decision EXpert) [72]. DEX divide un problema de decisión complejo en sub-problemas más pequeños y menos complicados y los organiza jerárquicamente en un árbol de decisión. En consecuencia, DEX permite de manera simple interpretar el modelo. La construcción de la fase jerárquica multicriterio basada en DEX comprende la definición de escalas (a), definición de pesos (b) y generación de reglas (c) [72].

a. Definición de escalas

La definición de escalas representa los diferentes niveles de impacto para cada atributo básico del modelo HMP-Coffee. Para la definición de escalas, asignamos valores ordinales a cada atributo del modelo según el conocimiento de los expertos. La asignación de valores cualitativos refleja el efecto del atributo en relación al nivel de rentabilidad. Por ejemplo, "Inversión en tiempo de cosecha" se representa con valores de tres escalas (**Alto**, **Promedio** y **Bajo**) y "Rentabilidad" (**Favorable**, **Promedio** y **Desfavorable**). Los valores cualitativos rojos tienen una connotación negativa y los verdes tienen una connotación positiva en la rentabilidad. En este sentido, una "Inversión en tiempo de cosecha" **Alta** contribuye a una rentabilidad **Desfavorable**, y una "Inversión en tiempo de cosecha" **Baja** contribuye a una rentabilidad **Favorable**. Es importante señalar que los valores de la escala de atributos son independientes del contexto local. Por ejemplo, HMP-Coffee define tres escalas para el atributo básico "Transporte" (**Alto**, **Promedio** y **Bajo**) en cualquier parte del mundo. Por el contrario, el significado de los valores o rangos de la escala puede ser diferente entre países e incluso regiones del mismo país. Por ejemplo, 10 dólares pueden representar un valor **Alto** para "Transporte" en Colombia, mientras que en Brasil puede representar un valor **Promedio**.

La **Tabla 5** presenta las escalas cualitativas definidas por el panel de expertos para HPM-Coffee (celdas grises) y el significado por valor cualitativo (i.e., rangos) determinado por un grupo de cafetaleros locales (celdas blancas). En este caso, los cafetaleros trabajan para la Asociación Cafambiente[73]. En particular, para definir los niveles de rentabilidad (i.e., la variable objetivo), los expertos consideraron un Pr , dado por la Ecuación 1 [74], y la pirámide de necesidades de Maslow [75]. Acordaron clasificar la rentabilidad de la siguiente manera. **Desfavorable** ($Pr \leq 1$), es decir, el pequeño agricultor no puede satisfacer sus necesidades fisiológicas. En **Promedio** ($1 < Pr \leq 1.4$), el pequeño productor obtiene suficientes ganancias para satisfacer sus necesidades fisiológicas. **Favorable** ($Pr > 1.4$), la pequeña producción de café genera ingresos que permiten satisfacer las necesidades de seguridad. Las columnas de la tabla contienen el Id, nombres de los atributos básicos, nombres de las escalas cualitativas (p. Ej., tiempo bajo, tiempo medio y tiempo alto) y la unidad de medida de los rangos cuantitativos (p. Ej., horas, Costa Rica Colon (CRC), USD, Bushel por hectárea (Bsh / ha)).

$$Pr = \frac{Ingresos}{Inversion} = \frac{Ingresos\ por\ venta\ de\ Cafe}{Costos\ de\ produccion} \quad (1)$$

Tabla 5. Definición de escalas para el conjunto de atributos básicos de HPM-Coffee. Celdas grises: escalas cualitativas definidas por el panel de expertos de HPM-Coffee. Celdas blancas: rangos por valor cualitativo (determinados por un grupo de cafetaleros locales con su unidad respectiva).

Id	Atributo básico	Escala Cualitativa	Unidad
----	-----------------	--------------------	--------

		Tiempo bajo	Tiempo Promedio	Bastante tiempo	
t1	Tiempo invertido en labores de Fertilización	<98,7	98,7-99,25	99,25-99,8	
t2	Tiempo invertido en labores de Renovación	<6,8	6,8-8,6	8,6-10,4	
t3	Tiempo invertido en labores de Control (Arvenses/Plagas & enfermedades)	<129,4	129,4-143,2	143,2-157	Hora
		Económico	Promedio	Costoso	
t4	Precio de mano de obra	680.24-895.93	895.93-1111,62	1111,62-1327,31	CRC
		Bajo costo	Moderado	Alto	
t5	Inversión en fertilizantes químicos	4597.43-6500	6500-9500	8817,09-11630.21	
t6	Inversión en fertilizantes orgánicos	1321.56-1793.25	1793,25-2736.63	2736.63-3308.32	CRC
t7	Inversión en insumos para el control de arvenses	1089-1569,87	1569,87-2050,14	2050,14-2530,71	
t8	Inversión en insumos para renovación del cultivo	377.917-844.197	844.197-1310.47	1310.47-1776.746	
		Baja	Moderada	Alta	
t9	Inversión en mantenimiento	5060.26-6976.353	6976.353-8892.447	8892.447-10808.54	CRC
t10	Gastos de transporte	459.916-1063.037	1063.037-2269.279	2269.279-2872.4	
		Económico	Promedio	Costoso	
t11	Pago por recolección de café	13609.39-16802.9	16802.9-19996.41	19996.41-23189.92	CRC
		Baja inversión	inversión Moderada	Alta inversión	
t12	Implementos para recolectores de café	65,25-<86	86-87	>87-107.049	
t13	Inversión en proceso de beneficio de café	5512.327-6895.7711	6895.7711-8279.05	8279.05-9662.389	CRC
t14	Inversión en proceso de trilla de café	2000.828-2446.986	2446.986-2893.145	2893.145-3339.303	
		Favorable	Promedio	Desfavorable	
t15	Variación del Precio del Dólar	557.303-585.65	528.957-557.303	500.61-528.957	USD
t16	Precio del Café en la bolsa de valores de NY	183.07-255.16	147.025-183.07	110.98-147.025	
		Alta Producción	Producción promedio	Baja Producción	
t17	Volumen de producción de café	<=24.4	>24.4 - <37.4	>=37.4	Bsh /ha
		Favorable	Promedio	Desfavorable	
vo	Nivel de rentabilidad	>=1.4	1<= Pr <1.4	<=1	

b. Definición de pesos

La definición de pesos determina la relevancia de los atributos con respecto a la variable de salida (i.e., nivel de rentabilidad). Los pesos se definieron con base en valores reales obtenidos de la finca "La Sultana", incluyendo costo de insumos, actividades de manejo de cultivos,

cosecha, gastos administrativos, procesamiento de café, proceso de transformación y utilidad. Cabe destacar que "La Sultana" engloba el modelo económico de diez pequeñas fincas cafetaleras de la Asociación Cafiambiente [73]. La asociación opera bajo los Principios y Criterios del Estándar de Agricultura Sostenible RAS [76], promoviendo prácticas saludables para lograr la sostenibilidad en pequeños cultivos de café.

En particular, el peso para cada atributo básico y agregado es definido (en total 24 atributos) calculando la relevancia promedio correspondiente (i.e., un porcentaje parcial) con respecto a la rentabilidad total (i.e., porcentaje total del 100%). Por ejemplo, en el período 2010 - 2019, la inversión promedio en pago por la recolección de café (i.e., atributo básico t11) fue de USD 4.250 por año. Este valor representó el 99% de la inversión total en la cosecha (es decir, el atributo agregado t22) y el 33% del costo total de producción (es decir, una variable global gv1). El costo de producción impactó el 60% de la rentabilidad total (es decir, la variable objetivo vo). La **Tabla 6** resume los pesos de todos los atributos básicos y agregados. Es importante tener en cuenta que la suma de las ponderaciones dentro de cada atributo agregado es del 100%. La suma de las ponderaciones de los atributos agregados es 100% para obtener la variable objetivo.

Tabla 6. Pesos definidos para el conjunto de atributos HPM-Coffee. La columna atributo contiene los atributos básicos, los atributos agregados y la variable de salida. La columna peso incluye cinco niveles de ponderación para los atributos del modelo de café HMP.

Atributo	Pesos basados en la operación de la finca "La Sultana"				
	Peso 1	Peso 2	Peso 3	Peso 4	Peso 5
vo: Nivel de rentabilidad					=100%
gv1: Costos de producción					40%
t19: Labores del Cultivo				42,60%	
t23: Inversión en actividades de cuidado de cultivos			56%		
t18 Tiempo de ejecución en actividades del cultivo		70%			
t1: Tiempo invertido en labores de Fertilización	25%				
t2: Tiempo invertido en labores de Renovación	56%				
t3: Tiempo invertido en labores de Control (Arvenses/Plagas & enfermedades)	18%				
t4: Precio de mano de obra		30%			
t20: inversión en insumos			44%		
t24: inversión en actividades de fertilización		79,83%			
t5: Inversión en fertilizantes químicos	21%				
t6: Inversión en fertilizantes orgánicos	79%				
t7: Inversión en insumos para el control de arvenses		13%			
t8: Inversión en insumos para renovación del cultivo		7,13%			
t21: Gastos de gestión				15,80%	
t9: Inversión en mantenimiento					71%

t10: Gastos de Transporte	29%
t22: Inversión en tiempo de cosecha	32,84%
t11: Pago por recolección de café	99%
t12: Implementos para recolectores de café	1%
t13: Inversión en proceso de beneficio de café	5,03%
t14: Inversión en proceso de trilla de café	3,05%
gv2: Volumen de producción	40%
t17: Volumen de producción de café	
gv3: Mercado	20%
t15: Variación del Precio del Dólar	50%
t16: Precio del Café en la bolsa de valores de NY	50%

c. Generación de Reglas

En el paso de generación de la regla, usamos los conceptos de funciones de utilidad y pesos proporcionados por DEXI [30]. Las funciones de utilidad f definen reglas de toma de decisiones combinando valores de escala de atributos básicos X_1, X_2, \dots, X_n para obtener el valor cualitativo de los atributos agregados (Ver Ecuación 2). Los pesos w_1, w_2, \dots, w_n (i.e., la salida del paso de definición de peso) permiten definir el valor de las funciones de utilidad f estableciendo la contribución del atributo correspondiente al resultado final f (Ver Ecuación 3).

$$f: X_1 = value_1 \text{ and } X_2 = value_2 \text{ and } \dots \text{ and } X_n = Value_n \text{ then } Y = Value_n \quad (2)$$

$$f(X_1, X_2, \dots, X_n) = (w_1 X_1) \text{ and } (w_2 X_2) \text{ and } (w_n X_n) \quad (3)$$

DEXI representa las funciones de utilidad en una tabla denominada "Tabla de atributos" donde cada fila indica una función f para cada una de las combinaciones resultantes de entre todos los estados de las escalas cualitativas [72]. En este paso de generación de reglas para el modelo HMP-Coffee, se obtienen un total de 10 tablas de atributos y 93 reglas SI-ENTONCES. La **Figura 8** muestra un ejemplo para la generación de reglas del atributo agregado $Y =$ "labores del cultivo", el atributo básico $X_1 =$ "Tiempo de ejecución en actividades de cuidado del cultivo" debe ser llenado en la tabla como **Alto**, **Promedio** o **Bajo** y $X_2 =$ "Precio de la mano de obra" como **Costoso**, **Promedio** o **Económico**. A su vez, el atributo agregado $Y =$ "Labore del cultivo" tiene como salida la escala cualitativa **Alto**, **Normal** o **Bajo**. Una vez son realizadas todas las posibles combinaciones (por los expertos), el modelo puede estimar el nivel de rentabilidad. Todas las tablas desarrolladas para HPM-Coffee están disponibles en el **Anexo 2**.

	Investment	Investment	Labor of the crop
1	High	High	High
2	High	Average	High
3	High	Low	Average
4	Normal	High	Average
5	Normal	Average	Average
6	Normal	Low	Average
7	Low	High	Average
8	Low	Average	Low
9	Low	Low	Low

Rules: 2/9 (22.22%), determined: 100.00% [High:2,Average:5,Low:2]

Figura 8. Tabla DEXI para el atributo agregado “labores del cultivo” (captura de pantalla del software Dexi). Contiene los valores cualitativos por atributo: Para Y= “labores del cultivo”, Alto, Promedio o Bajo; para X1 = “Tiempo de ejecución en actividades de cuidado del cultivo”, Costoso, Promedio o Económico y para X2 = “Precio de la mano de obra”, Alto, Normal o Bajo. Además, presenta las nueve combinaciones posibles entre atributos.

3.2.3. Resumen

Este capítulo presentó y desarrolló la metodología para la construcción del modelo jerárquico multicriterio convencional llamado HMP-Coffee. En primer lugar, abordó una descripción del área de estudio (e.i., la finca la Sultana) y el panel de expertos que participó durante la construcción del modelo. En segundo lugar, se desarrolló el proceso de construcción de HMP-Coffee el cual fue desarrollado en dos fases: i) **Fase de conocimiento del contexto** para crear un modelo conceptual sobre la rentabilidad en la producción de café a pequeña escala compuesta por dos pasos: granularidad gruesa y granularidad fina y ii) la **Fase jerárquica-multicriterio convencional** para convertir el modelo conceptual en un modelo cualitativo jerárquico comprensible y ejecutable capaz de estimar el nivel de rentabilidad en pequeñas producciones de café; desarrollada en tres pasos: la definición de escalas, pesos y reglas.

Capítulo 4

4. Sistema de Lógica Difusa en el HMM

Este capítulo presenta el sistema de lógica difusa desarrollado para HMM, específicamente, en la fase jerárquica multicriterio para mejorar el desempeño del modelo convencional (**sección 3.2.2**). Para este propósito, son implementados/probados tres enfoques multicriterio difusos: basado en reglas de Mandammi, Topsis y Saw. Para cada enfoque contiene la conceptualización, implementación y un ejemplo práctico que facilita su comprensión.

4.1. Implementación de un Sistema de Lógica difusa en el HMM

El uso de modelos jerárquicos cualitativos es útil para garantizar la creación de modelos sencillos y de fácil interpretación, pero poco flexibles y susceptibles a errores. Esto sucede debido a que su funcionamiento está basado en la teoría clásica de conjuntos donde un elemento pertenece o no a un conjunto, más no existen situaciones intermedias [48]. Los autores en [49] afirman que se debe mantener dos premisas para el buen funcionamiento de estos modelos: garantizar cierto nivel de homogeneidad interna por estado y un grado significativo de diferenciación entre ellos. Si estas dos condiciones no se mantienen, la respuesta de los modelos presenta graves problemas de precisión. Su adaptación hacia enfoques como la teoría difusa, permitirá considerar situaciones intermedias, y lograr así, modelos más flexibles manteniendo el bajo número de escalas cualitativas por atributo.

A continuación, la **Figura 9** presenta el proceso conceptual para la construcción del modelo HMP-Coffee difuso para estimar el nivel de rentabilidad en pequeñas producciones de café.

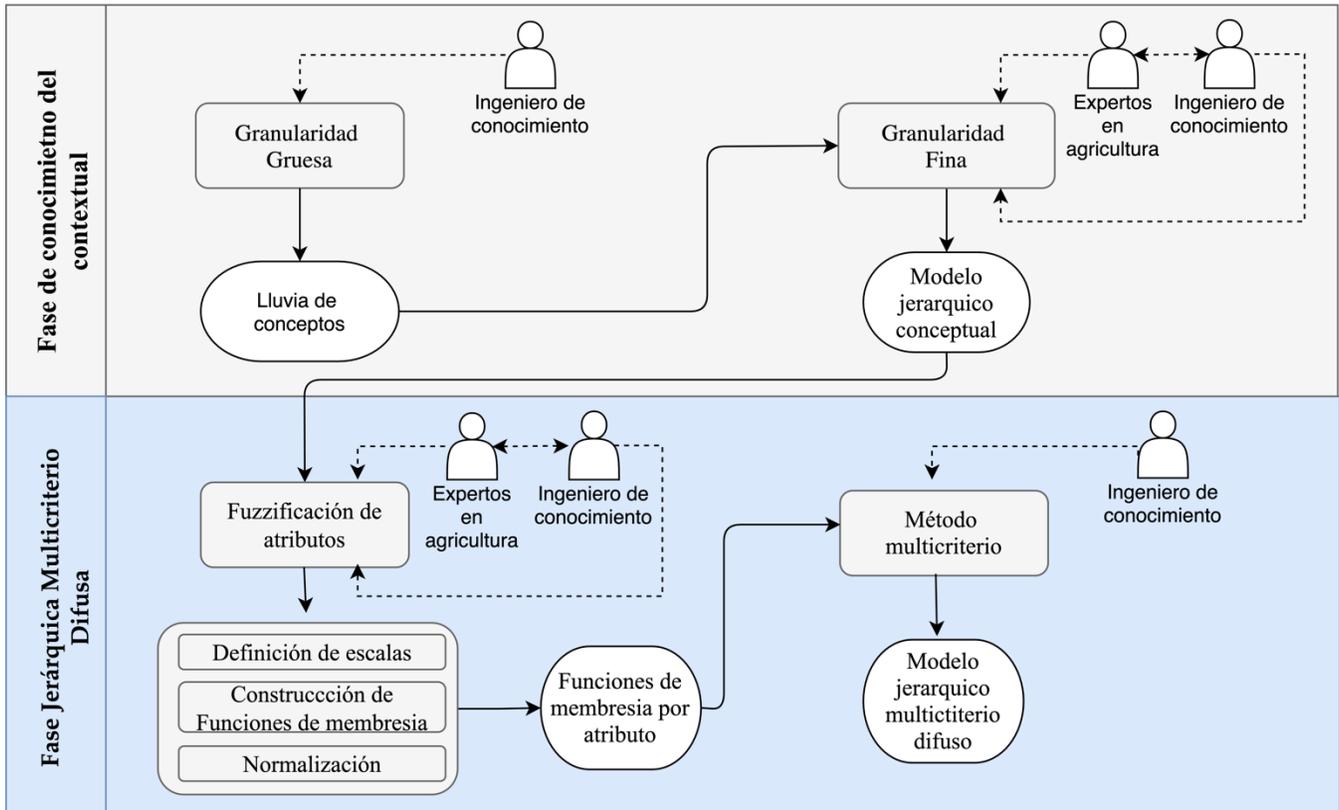


Figura 9. Proceso conceptual para la construcción del modelo HMP-Coffee. Fase 1: Conocimiento del contexto; Fase 2: Multicriterio jerárquico difuso. Fuente: autoría propia.

La adaptación del modelo convencional (**sección 3.2.2**) se desarrolla en la fase 2, renombrada fase jerárquica multicriterio difusa la cual incluye dos pasos: la fuzzificación de atributos y el método multicriterio difuso.

El paso de fuzzificación de atributos tiene como fin representar cada atributo del modelo a través de funciones de membresía. Para este propósito la Fuzzificación de Atributos desarrolla tres actividades: la definición de escalas, construcción de las funciones y normalización. Como salida de este paso, obtenemos el conjunto de funciones membresía de todos los atributos del modelo (i.e., básicos y agregados). **El paso de método multicriterio difuso** tiene como fin la aplicación de un método multicriterio que tome decisiones basado en valores difusos. Para este fin, utiliza las salidas de las funciones de membresía (i.e., los valores de pertenencia) como criterios de decisión. Finalmente, la salida de este paso, presenta la decisión final del modelo jerárquico multicriterio difuso.

4.1.1. Fuzzificación de atributos

Este paso tiene como propósito cambiar el uso de escalas cualitativas tradicionales, basada en la teoría clásica de conjuntos (p.Ej., Alto, Medio y Bajo) en funciones de pertenencia basadas en la teoría difusa. Las funciones de membresía son la forma de representar gráficamente el grado de pertenencia que un valor de entrada posee con respecto a cada uno de los estados de la escala cualitativa. Para este propósito se procede a definir las escalas, construir las funciones de membresía y normalizarlas:

- a. La definición de una escala es determinada por el panel de expertos (**sección 3.1, Tabla 2**) y el significado por valor cualitativo (i.e., rangos) es definido por un grupo de cafetaleros locales (**Tabla 5**). El desarrollo de este paso es explicado en detalle en la **sección 3.2.2**.
- b. Para la construcción de las funciones aplicamos la metodología propuesta por Chen et al. [77] explicada en la **Figura 10**.

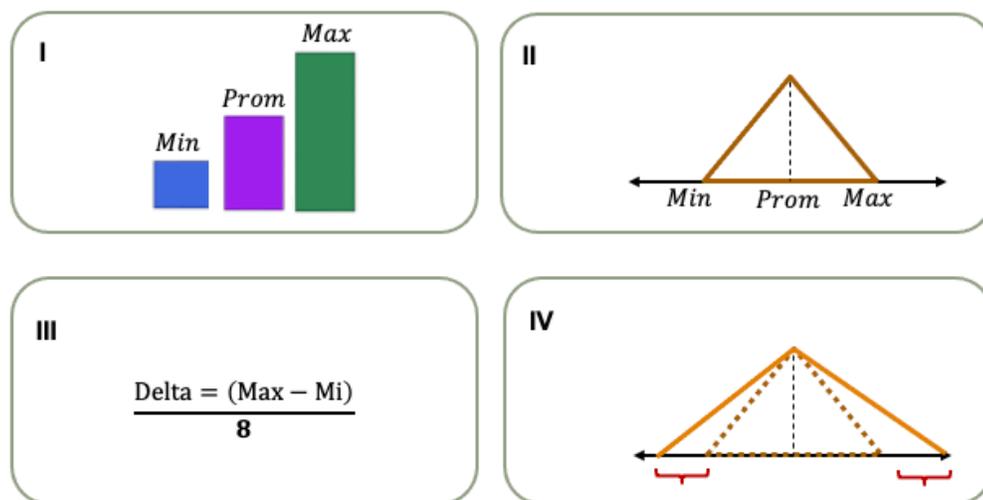


Figura 10. Pasos por cuadrantes para la construcción de las funciones de membresía. I. Identificación de mínimos (min), promedios (prom) y máximos (max). II: Construcción funciones triangulares. III: Calculo Delta. IV: Extensión de máximos y mínimos. Fuente: autoría propia.

La **Figura 10** ejemplifica los cuadrantes para la construcción de las funciones de membresía donde:

- I. Son identificados los valores mínimos (min), promedios (prom) y máximos (max).
- II. de cada estado (x) de la escala cualitativa. En HMP-Coffee, esta información es proporcionada por el grupo de cafeteros locales (p.Ej., Los productores de la asociación Cafiambiente del Cauca).
- III. Se construyen las funciones de pertenencia triangular utilizando los valores **$min(x)$** , **$prom(x)$** y **$max(x)$** para cada estado (x).

IV. Se determina el valor **delta (x)** de cada estado (x), donde **delta (x)** es dado por la siguiente expresión:

$$\mathit{delta}(x) = \frac{\mathit{max}(x) - \mathit{min}(x)}{8} \tag{4}$$

V. Se extiende el valor de min(x) a **min_d (x)** donde es **min_d (x) = min(x)-delta(x)** y el valor de max(x) a **max_d (x)** donde **max_d (x) = max(x)+delta(x)**

Para ejemplificar este proceso, se presenta la construcción de la función de membresía del atributo básico "Precio de mano de obra" con escalas **Costoso (x)**, **Promedio (y)** y **Económico (z)**. Los productores definen los valores máximos, mínimos y promedios mostrados en la **Tabla 7** (Cuadrante I):

Tabla 7. Valores máximos, mínimos y promedios del atributo básico

	<i>Min</i>	<i>Max</i>	<i>Prom</i>
Económico	680,24	895,93	788,085
Promedio	895,93	1111,62	1003,775
Costoso	1111,62	1327,31	1219,465

Para los estados cualitativos **Costoso(x)**, **Promedio (y)**, o **Económico (z)**, son construidas las funciones de membresía triangular utilizando los valores calculados en la **Tabla 7** como se muestra a continuación en la **Figura 11** (Cuadrante II).

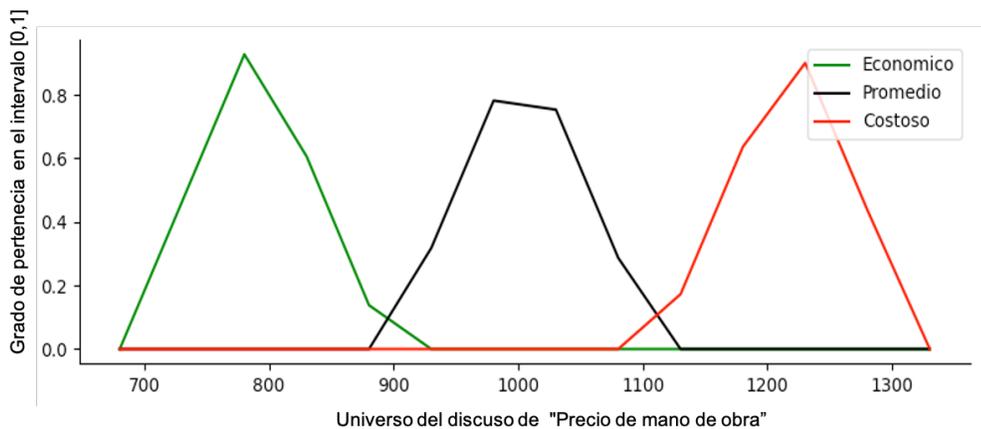


Figura 11. Funciones de membresía triangulares del atributo básico "Precio de mano de obra" con escalas cualitativas representadas por colores: costoso en rojo, promedio en negro y económico en verde

Posteriormente, calculamos el valor **delta** para cada estado **Costoso** (ver ecuación 5), **Promedio** (Ver ecuación 6), y **Económico** (Ver ecuación 7) (Cuadrante III)

$$\text{delta}(x) = \frac{895,93-680,24}{8} = 26,9 \quad (5)$$

$$\text{delta}(y) = \frac{1111,62-895,93}{8} = 26,9 \quad (6)$$

$$\text{delta}(z) = \frac{1327,31-1111,62}{8} = 26,9 \quad (7)$$

Seguidamente, la **Tabla 8** presenta los valores de los mínimos y los máximos extendidos (un valor **delta**) tanto a la derecha **max_d** como a la izquierda **min_d** (Cuadrante IV) para cada estado **Costoso**, **Promedio** y **Económico**.

Tabla 8.Valores de los mínimos y los máximos extendidos un valor delta para cada estado Costoso, Promedio y Económico.

	min_d	max_d
Económico	653,27875	922,89125
Promedio	868,96875	1138,58125
Costoso	1084,65875	1354,27125

Finalmente, la **Figura 12** muestra la función de membresía resultante correspondiente al atributo básico "Precio de mano de obra".

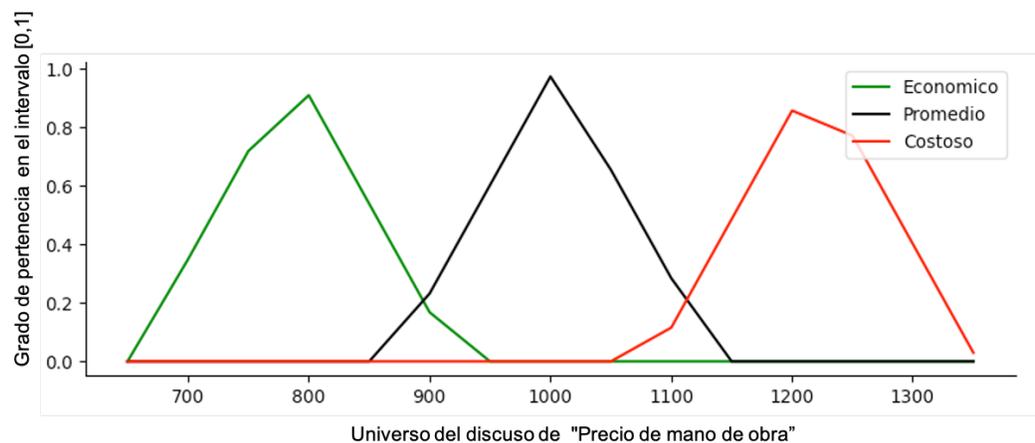


Figura 12. Función de membresía resultante correspondiente al atributo básico "Precio de mano de obra" con escalas cualitativas representadas por colores: Costoso en rojo, Promedio en negro y Económico en verde.

Es necesario repetir el método de Chen et al. [77] para cada atributo del modelo. Una vez finalizada la construcción de funciones de membresía, el modelo habrá pasado de la dimensión clásica a la difusa. Todas las funciones de membresía están en el **Anexo 3**.

c. La normalización es el proceso matemático que permite ajustar las funciones de membresía en diferentes escalas respecto a una escala común. Cada función de membresía es normalizada basado en la técnica de Re-escalado, aplicando la siguiente expresión para los tres valores de la función triangular $min_d(x)$, $prom(x)$ y $max_d(x)$ de cada estado de la escala cualitativa

$$X_{normalizada} = \frac{X - Min}{Max - Min} \quad (8)$$

La **Figura 12** ejemplifica el proceso de normalización para un valor $X \in$ Funcion triangular.

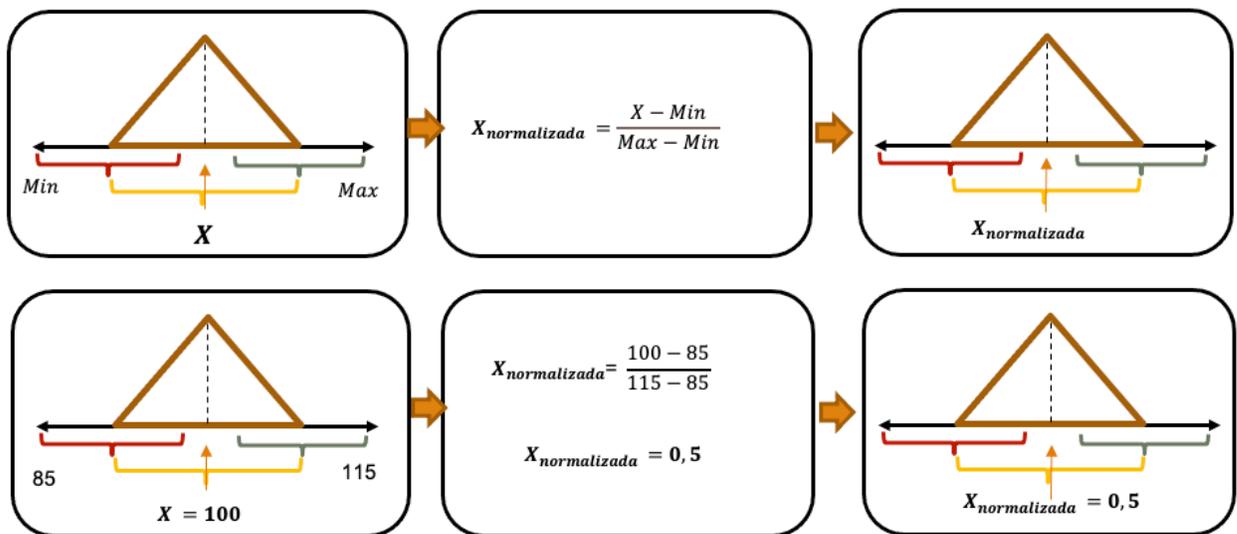


Figura 13.Proceso de normalización de las funciones de membresía. Fuente: autoría propia

Continuando con el ejemplo del paso anterior, para la función de membresía del atributo "Precio de mano de obra" la **Tabla 9** presenta los valores normalizados de cada función triangular por estados **Económico**, **Promedio** y **Costoso**. Para el caso de HMP-Coffee difuso, normalizamos las funciones de membresía y multiplicamos por un factor 3 con el fin de manejar las mismas escalas de HMP-Coffee convencional donde la salida varía entre 1 a 3 (p.Ej., **1: Favorable**, **2: Promedio** y **3: Desfavorable**) y facilitar la comparación.

Tabla 9.Valores normalizados de cada función triangular Económico, Promedio y Costoso

	Normalizadas*3		
	min_d	prom	max_d
Económico	0,01280697	0,5826324	1,15241557
Promedio	0,92446849	1,49429393	2,0640771
Costoso	1,83613001	2,40595545	2,97573862

Finalmente, la **Figura 14** muestra la función de membresía normalizada correspondiente al atributo básico "Precio de mano de obra".

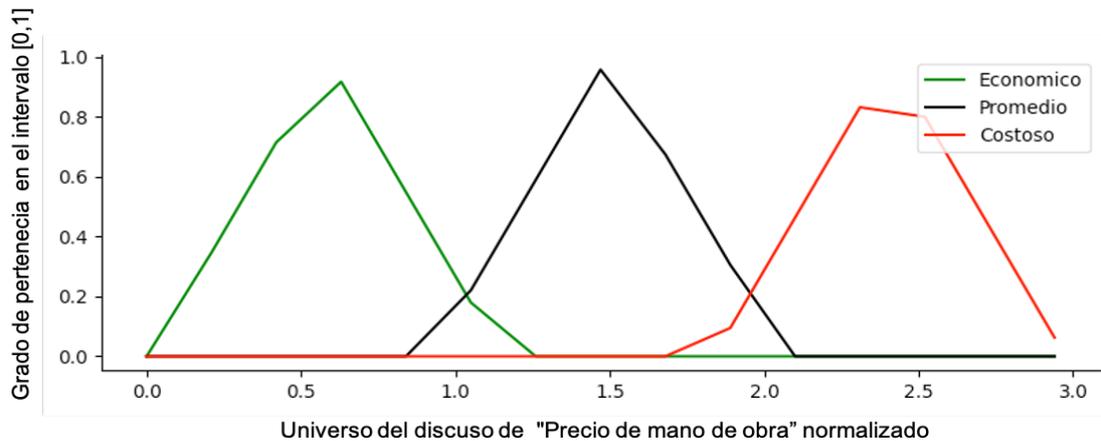


Figura 14. Función de membresía normalizada del atributo básico "Precio de mano de obra" representadas por colores: Costoso en rojo, Promedio en negro y Económico en verde.

4.1.2. Método Multicriterio Difuso

Este paso pretende incorporar los valores de pertenencia de las funciones de membresía, como alternativas del método multicriterio (implementado de forma jerárquica) que conlleven a estimar el estado de la variable de salida/interés (i.e., la rentabilidad en pequeñas producciones de café). Para este propósito, el método multicriterio detecta el mayor valor de pertenencia entre el conjunto alternativas (i.e., de escalas cualitativas posibles), que condicionado por una serie de disposiciones (i.e., pesos y/o reglas) selecciona el "mejor camino". Como salida de este paso, el "mejor camino" estima el estado de la variable de interés (i.e., rentabilidad **Favorable**, **Promedio** o **Desfavorable**).

Para la selección de los métodos multicriterio a implementar, la presente propuesta de investigación mantiene como premisa principal la interpretabilidad del modelo. En este sentido, son seleccionados tres métodos multicriterio: El método basado en reglas difuso llamado reglas de Mandammi y los métodos matemáticos Topsis y Saw los cuales, según la revisión literaria de Velásquez et al., mantienen un alto nivel de simplicidad e interpretabilidad, en comparación a otros métodos multicriterio. A continuación, es presentado en mayor detalle, el funcionamiento de cada método y su interacción con los valores de pertenencia arrojados por las funciones de membresía.

a. HMP-Coffee basado en reglas Difusas de Mandammi

La **Figura 15** presenta la arquitectura para implementar el método multicriterio basado en reglas difusas de Mandammi. Este método estima la variable objetivo (i.e., el nivel de rentabilidad) a partir de reglas difusas formadas con los operadores AND y OR. Toma la salida

de la Fuzzificación (i.e., valores de pertenencia) como antecedentes o condiciones de una sentencia tipo IF-THEN para producir las reglas.

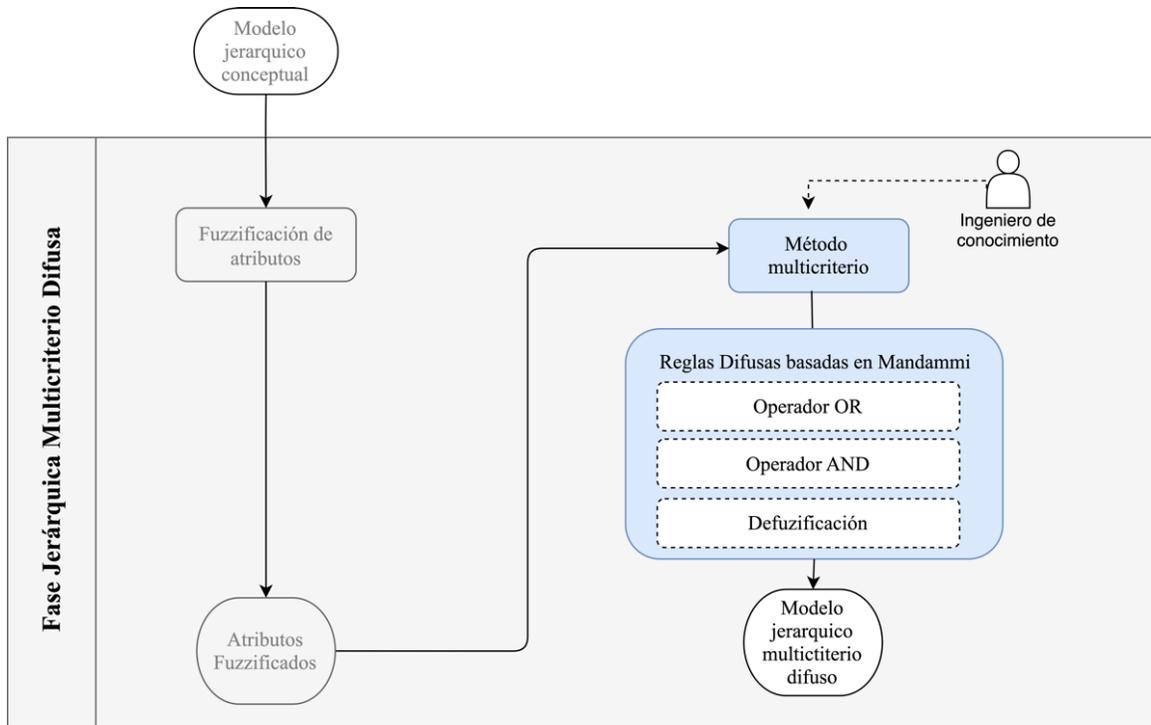


Figura 15. Proceso conceptual para implementar el modelo HMP-Coffee basado en Reglas-Mandammi. Paso 1: Fuzzificación; Paso 2: Método basado en Reglas-Mandammi. Fuente: autoría propia.

En general, las reglas difusas son proposiciones asociadas a adverbios que pueden transformar los conjuntos difusos. Estas reglas emplean las sentencias del tipo:

IF < antecedente o condición> THEN < consecuente o conclusión>

El < antecedente o condición> y < consecuente o conclusión> son proposiciones difusas que pueden formarse usando conjunciones (AND) o disyunciones (OR). Las formas más comunes para implementar las reglas difusas son el método de Mandammi [79] y el de Takagi-Sugeno [80], sin embargo, Mandammi genera un conjunto de reglas más interpretable [81], razón por la cual es el método seleccionado para aplicar en el desarrollo de esta propuesta. Mandammi utiliza el operador AND u OR para obtener una única función que representa el resultado de la evaluación de la regla. La función resultante es defuzzificada para obtener un único valor de salida [82]. En el dominio difuso el operado AND o unión está definido por la función T- Conorma y el operador OR o la intersección por la función Norma Triangular T-Norma [83].

Ejemplo de HMP-Coffee difuso basado en reglas de Mandammi

La **Figura 16** ejemplifica el funcionamiento de una regla difusa para valores de entrada de “Precio de mano de obra” = 1,75 y “Tiempo de ejecución en actividades del cultivo” =1,2. Para la implementación de las reglas difusas se utilizó la librería matemática de python 3.7, scikit-fuzzy (skfuzzy).

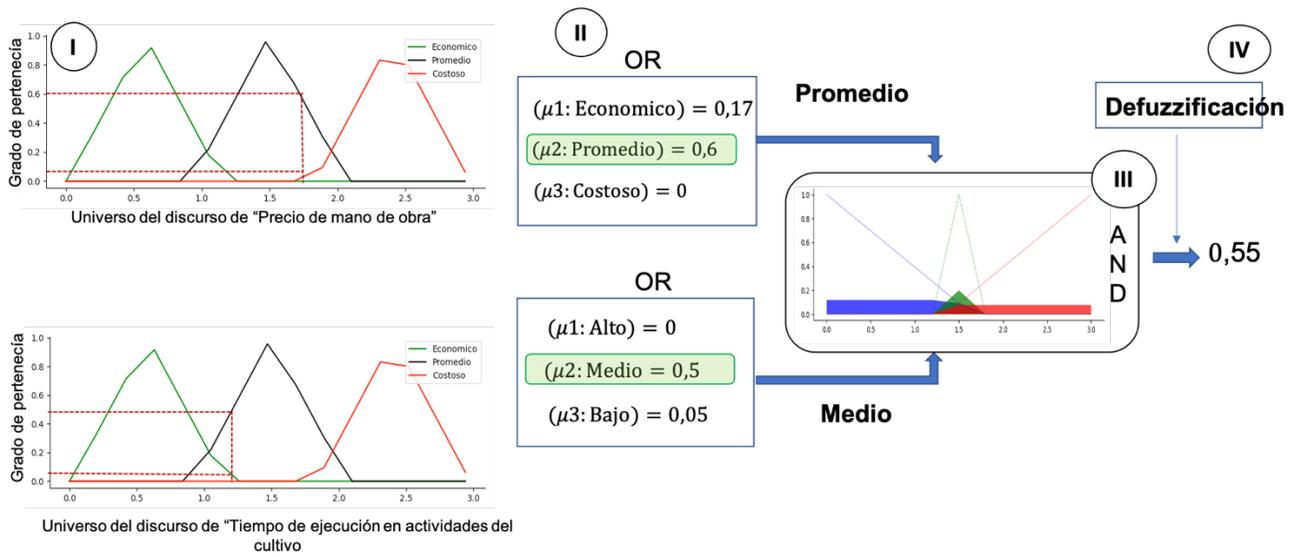


Figura 16. Funcionamiento de una regla difusa para valores de entrada de “Precio de mano de obra” = 1,75 y “Tiempo de ejecución en actividades del cultivo” =1,2.

- I. Para cada dato de entrada es asignado un valor de pertenencia $\mu_1, \mu_2 \dots \mu_n$ extraídos de las funciones de membresía triangulares. “Precio de mano de obra” tiene valores de pertenencia $\mu_{x1} = 0,17, \mu_{x2} = 0,6$ y $\mu_{x3} = 0$. Por su parte “Tiempo de ejecución en actividades del cultivo” tiene valores de pertenencia $\mu_{y1} = 0, \mu_{y2} = 0,5$ y $\mu_{y3} = 0,05$.
- II. Estas funciones pasan por el operador OR o la intersección identificando el mayor valor de pertenencia de cada una de ellas. En el caso de “Precio de mano de obra”, el mayor valor de pertenencia es $\mu_{x2} = 0,6$. Por su parte, para “Tiempo de ejecución en actividades del cultivo” el mayor valor de pertenencia es $\mu_{y2} = 0,5$.
- III. Los valores de pertenencia identificados ejecutan una regla utilizando el operador AND o unión. Esta operación genera una única función que representa el resultado de la evaluación de la regla (ver apartado III, **Figura 16**).
- IV. Esta función es defuzzificada generando un único valor de salida como resultado. Para este ejemplo en particular es 0,55.

b. HMP-Coffee difuso basado en Topsis

Topsis permite estimar el estado de la variable objetivo (i.e., el nivel de rentabilidad) a partir del análisis de múltiples alternativas. El principio básico es que la alternativa elegida (i.e., rentabilidad **Favorable**, **Promedio** o **Desfavorable**) debe tener la distancia más corta de la solución ideal positiva y la distancia más lejana de la solución ideal negativa.

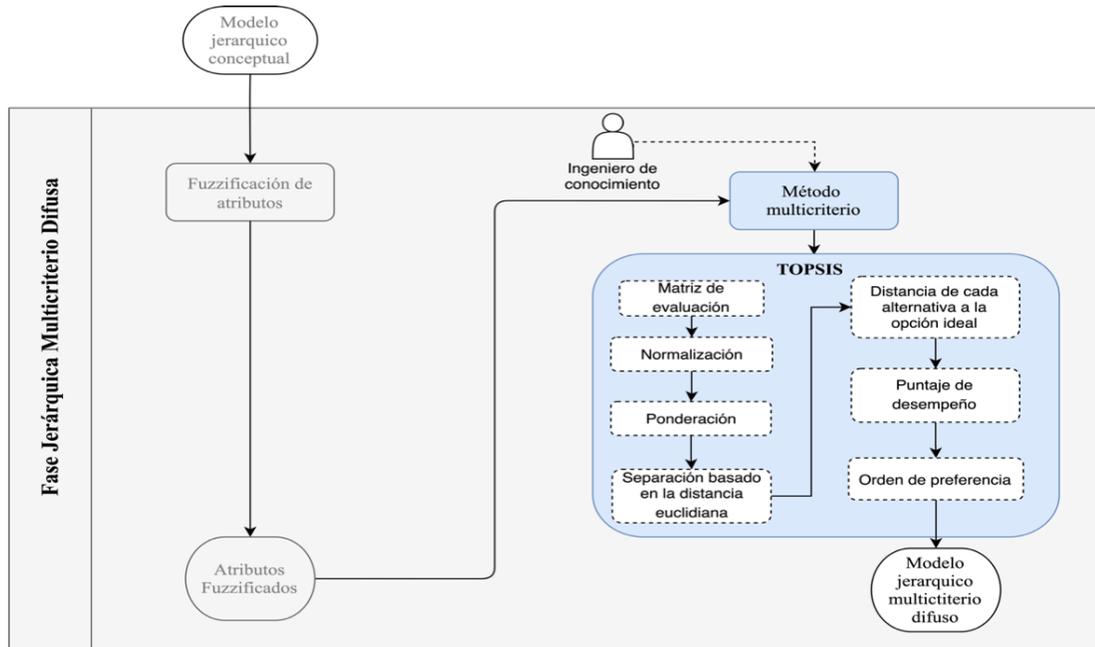


Figura 17. Proceso conceptual para la implementación del modelo HMP-Coffee basado en Topsis. Paso 1: Fuzzificación; Paso 2: Método basado en Topsis. Fuente: autoría propia.

La **Figura 17** presenta el proceso conceptual para la implementación de la fase Jerárquica Multicriterio Difusa basada en Topsis. El modelo jerárquico conceptual (**sección 3.2.1, Figura 7**) atraviesa la Fuzzificación, generando una función de membresía y un grupo de valores de pertenencia para cada atributo del modelo. Esa información, entra al método multicriterio basado en Topsis el cual ejecuta los siguientes 8 pasos.

Paso 1: identificación de alternativas a través de una matriz de evaluación. Esta matriz consta de m alternativas y n criterios, la intersección de cada alternativa y los criterios dados como $x_{i,j}$, donde, obtenemos una matriz $(x_{i,j})_{m \times n}$. Para el enfoque propuesto, la matriz de evaluación esta compuesta por los valores de pertenencia calculados para cada dato de entrada al modelo jerárquico (en cada una de las relaciones por nivel y en cada nivel del modelo). En este sentido la matriz se representa como:

$$X_{i,j} = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,n} \\ X_{2,1} & X_{2,2} & \dots & X_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{m,1} & X_{m,2} & \dots & X_{m,n} \end{bmatrix} \quad (9)$$

Paso 2: normalización de la matriz de decisión. El valor normalizado $(\bar{X}_{ij})_{m \times n}$ se calcula como:

$$\bar{X}_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^n x_{ij}^2}}, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n \quad (10)$$

Paso 3: ponderación de la matriz de decisión normalizada como:

$$V_{ij} = \bar{X}_{ij} \times W_j, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n, \quad (11)$$

donde W_j son las ponderaciones por atributo del modelo

Paso 4: determinación de la solución ideal positiva y la solución ideal negativa representada con la siguiente expresión

$$V^+ = \{(max_i V_{ij} | j \in J^+), \{(min_i V_{ij} | j \in J^-) | i = 1, 2, \dots, m\} \quad (12)$$

$$V^- = \{(min_i V_{ij} | j \in J^+), \{(max_i V_{ij} | j \in J^-) | i = 1, 2, \dots, m\} \quad (13)$$

donde ,

$$J^+ = 1, 2, \dots, n | j \text{ asociado al criterio de impacto positivo}$$

$$J^- = 1, 2, \dots, n | j \text{ asociado al criterio de impacto negativo}$$

Paso 5: cálculo de la separación de cada alternativa a la solución ideal dada por la siguiente expresión:

$$S_i^+ = [\sum_{j=1}^m (V_{ij} - V_{j^+})^2]^{0.5}, \quad i = 1, 2, \dots, m \quad (14)$$

Paso 6: cálculo de la separación de cada alternativa a la solución más lejana dada por la siguiente expresión

$$S_i^- = [\sum_{j=1}^m (V_{ij} - V_{j^-})^2]^{0.5}, \quad i = 1, 2, \dots, m \quad (15)$$

Paso 7: cálculo de del puntaje de desempeño dada por la siguiente expresión

$$P_i = \frac{S_i^-}{S_i^+ + S_i^-} \quad (16)$$

Paso 8: organización de las alternativas en orden creciente

Ejemplo de HMP-Coffee difuso basado en Topsis

En este apartado, ejemplificamos el funcionamiento de Topsis mostrando la respuesta de cada paso para facilitar su comprensión. En este punto, el modelo ya ha atravesado la etapa de Fuzzificación. Para el valor de entrada “Tiempo de ejecución en labores del cultivo” = 1,7 y valores de pertenecía $\mu_{x1} = 0,17$, $\mu_{x2} = 0,6$ y $\mu_{x3} = 0$ y Precio de mano de obra = 1,5 y valores de pertenecía $\mu_{y1} = 0$, $\mu_{y2} = 0,5$ y $\mu_{y3} = 0,05$.

Paso 1: se identifica la matriz de evaluación compuesta por los valores de pertenencia de cada entrada (i.e., Precio de mano de obra y Tiempo de ejecución en labores del cultivo) como se presenta a continuación:

$$X_{i,j} = \begin{bmatrix} \mu_{1_{alto}} & \mu_{1_{costoso}} \\ \mu_{1_{promedio}} & \mu_{1_{normal}} \\ \mu_{1_{bajo}} & \mu_{1_{economico}} \end{bmatrix}$$

$$X_{i,j} = \begin{bmatrix} 0,29_{Alto} & 0,22_{costoso} \\ 0,71_{Promedio} & 0,85_{normal} \\ 0_{Bajo} & 0,5_{economico} \end{bmatrix}$$

Paso 2: se normaliza la matriz de evaluación aplicando la Ecuación 10. La matriz resultante es presentada a continuación.

$$\bar{X}_{ij} = \begin{bmatrix} 0,469_{Alto} & 0,21_{costoso} \\ 1,150_{Promedio} & 0,83_{normal} \\ 0,275_{Bajo} & 0,48_{economico} \end{bmatrix}$$

Paso 3: se pondera la matriz de decisión normalizada (Ver ecuación 11), con valores de ponderación o pesos de 0,64 para Precio de mano de obra y 0,36 para Tiempo de ejecución en actividades del cultivo (pesos tomados de la **Tabla 6**, Sección 3.2.2)

$$V_{i,j} = \bar{X}_{ij} \times W_j, \quad i = 1,2, \dots, m, \quad j = 1,2, \dots, n$$

$$V_{ij} = \begin{bmatrix} 0,469_{Alto} & 0,21_{costoso} \\ 1,150_{Promedio} & 0,83_{normal} \\ 0,275_{Bajo} & 0,48_{economico} \end{bmatrix} \times (0,64, 0,36)$$

$$V_{ij} = \begin{bmatrix} 0,174_{Alto} & 0,086_{costoso} \\ 0,426_{Promedio} & 0,333_{normal} \\ 0,102_{Bajo} & 0,293_{economico} \end{bmatrix}$$

Paso 4: se determina la solución ideal positiva y la solución ideal negativa (Ver Ecuación 12 y 13). Para “Precio de mano de obra” la $V^+ = 0.102$ y $V^- = 0.426$. Para “Tiempo de ejecución en actividades del cultivo” $V^+ = 0.426$ y $V^- = 0.333$ como se muestra a continuación.

$$V^+ = \left\{ \min \begin{pmatrix} 0,174_{Alto} \\ 0,426_{Promedio} \\ 0,102_{Bajo} \end{pmatrix}, \min \begin{pmatrix} 0,086_{costoso} \\ 0,333_{normal} \\ 0,293_{economico} \end{pmatrix} \right\}$$

$$V^+ = \{0.102, 0.086\}$$

$$V^- = \left\{ \max \begin{pmatrix} 0,174_{Alto} \\ 0,426_{Promedio} \\ 0,102_{Bajo} \end{pmatrix}, \max \begin{pmatrix} 0,086_{costoso} \\ 0,333_{normal} \\ 0,293_{economico} \end{pmatrix} \right\}$$

$$V^- = \{0.426, 0.333\}$$

Paso 5: se calcula la distancia euclidiana desde la solución ideal desarrollando la ecuación 14 como se muestra a continuación.

$$S_{i1}^+ = (((0,174_{Alto} - 0.102)^2 + (0,086_{costoso} - 0.086)^2)^{0.5}) = 0.072_{Desfavorable}$$

$$S_{i2}^+ = (((0,426_{Promedio} - 0.102)^2 + (0,333_{normal} - 0.086)^2)^{0.5}) = 0.407_{promedio}$$

$$S_{i3}^+ = (((0,102_{Bajo} - 0.102)^2 + (0,293_{economico} - 0.086)^2)^{0.5}) = 0.207_{Favorable}$$

Paso 6: se calcula la distancia euclidiana desde la peor solución desarrollando la ecuación 15 como se muestra a continuación

$$S_{i1}^- = (((0,174_{Alto} - 0.426)^2 + (0,086_{costoso} - 0,333)^2)^{0.5}) = 0.352_{Desfavorable}$$

$$S_{i2}^- = (((0,426_{Promedio} - 0.426)^2 + (0,333_{normal} - 0,333)^2)^{0.5}) = 0_{Promedio}$$

$$S_{i3}^- = (((0,102_{Bajo} - 0.426)^2 + (0,293_{economico} - 0,333)^2)^{0.5}) = 0.327_{Favorable}$$

Paso 7: se calcula el puntaje de desempeño desarrollando la ecuación 16 como se muestra a continuación.

$$P_{i1} = \frac{0.352_{Desfavorable}}{0.072_{Desfavorable} + 0.352_{Desfavorable}} = 5.25_{Desfavorable}$$

$$P_{i2} = \frac{0_{Promedio}}{0.407_{promedio} + 0_{Promedio}} = 0_{promedio}$$

$$P_{i3} = \frac{0.327_{Favorable}}{0.207_{Favorable} + 0.327_{Favorable}} = 3.55_{favorable}$$

Paso 8: se organizan las alternativas por orden de preferencia. Para este caso, el modelo basado en Topsis estimó la variable de salida como **Desfavorable**.

$$Opcion\ 3 = 0_{promedio}$$

$$Opcion\ 2 = 1.89_{favorable}$$

$$Opcion\ 1 = 5.25_{Desfavorable}$$

c. Modelo basado en Saw

Saw permite estimar el estado de la variable objetivo (i.e., el nivel de rentabilidad) a partir del cálculo del desempeño ponderado para cada alternativa (i.e., rentabilidad **Favorable**, **Promedio** o **Desfavorable**) y las organiza desde la mejor hasta la de menor desempeño [84]. La alternativa de mejor desempeño será el estado estimado de la variable objetivo (i.e., el nivel de rentabilidad).

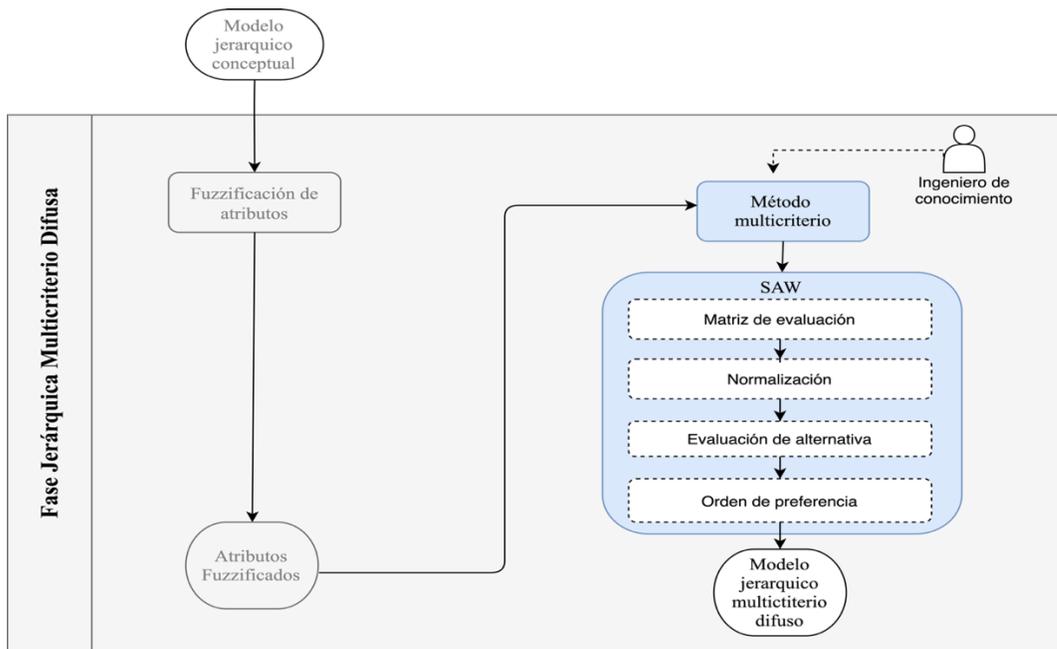


Figura 18. Proceso conceptual para la implementación del modelo HMP-Coffee basado en Saw. Paso 1: Fuzzificación; Paso 2: Método basado en Saw. Fuente: autoría propia.

La **Figura 18** presenta el proceso conceptual para la implementación de la fase Jerárquica Multicriterio Difusa basada en Saw. El modelo jerárquico conceptual (**sección 3.2.1, Figura 7**) atraviesa la Fuzzificación, generando una función de membresía y un grupo de valores de pertenencia por cada atributo del modelo. Esa información, entra al método multicriterio basado en Saw el cual desarrolla los siguientes 4 pasos.

Paso 1: identificación de alternativas a través de una matriz de evaluación. Esta matriz consta de m alternativas y n criterios, con la intersección de cada alternativa y los criterios dados como $x_{i,j}$, donde, obtenemos una matriz $(x_{i,j})_{m \times n}$. Para este enfoque, la matriz de evaluación está compuesta por los valores de pertenencia calculados para cada dato de entrada al modelo jerárquico. En este sentido la matriz se representa como:

$$X_{i,j} = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,n} \\ X_{2,1} & X_{2,2} & \dots & X_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{m,1} & X_{m,2} & \dots & X_{m,n} \end{bmatrix} \quad (17)$$

Paso 2: Normalización de la matriz $X_{i,j}$ usando la siguiente expresión

$$\bar{X}_{ij} = \frac{x_{ij}}{\text{Max}(x_{ij})} \quad (18)$$

Donde $\text{Max}(x_{ij})$ es el valor máximo de cada fila y columna.

Paso 3: Evaluación de cada alternativa, dada por la siguiente expresión

$$V_{ij} = \bar{X}_{ij} \times W_j, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n \quad (19)$$

Donde X_{ij} es la puntuación de la i -ésima alternativa con respecto al j -ésimo criterio, y W_j son las ponderaciones por atributo del modelo

Paso 4: Orden de preferencia.

Ejemplo de HMP-Coffee difuso basado en Saw

En este apartado, ejemplificamos el funcionamiento de Saw mostrando su respuesta en cada paso. Para valores de entrada donde “Tiempo de ejecución en actividades del cultivo” = 1,7 (“Promedio”) y “Precio de mano de obra” 1,5 (“Promedio”) se ejecutan los siguientes 4 pasos:

Paso 1: se identifica la matriz de evaluación compuesta por los valores de pertenencia de cada entrada (i.e., precio de mano de obra y tiempo de ejecución en actividades del cultivo) como se presenta a continuación:

$$X_{i,j} = \begin{bmatrix} \mu_{1_{alto}} & \mu_{1_{costoso}} \\ \mu_{1_{promedio}} & \mu_{1_{normal}} \\ \mu_{1_{bajo}} & \mu_{1_{economico}} \end{bmatrix}$$

$$X_{i,j} = \begin{bmatrix} 0,29_{Alto} & 0,22_{costoso} \\ 0,71_{promedio} & 0,85_{normal} \\ 0_{Bajo} & 0,5_{economico} \end{bmatrix}$$

Paso 2: se normaliza la matriz de evaluación. (Ver Ecuación 18)

$$X_{i,j} = \begin{bmatrix} \frac{0,29_{Alto}}{0,71_{promedio}} & \frac{0,22_{costoso}}{0,85_{normal}} \\ \frac{0,71_{promedio}}{0,71_{promedio}} & \frac{0,85_{normal}}{0,85_{normal}} \\ \frac{0_{Bajo}}{0,71_{promedio}} & \frac{0,5_{economico}}{0,85_{normal}} \end{bmatrix}$$

$$X_{i,j} = \begin{bmatrix} 0,4084507 & 0,2588235 \\ 1 & 1 \\ 0,23943662 & 0,588235 \end{bmatrix}$$

Paso 3. se evalúa cada alternativa, denotada V_i desarrollando la ecuación 19 como se muestra a continuación:

$$V_{i,j} = \bar{X}_{ij} \times W_j, \quad i = 1,2, \dots, m, \quad j = 1,2, \dots, n$$

$$V_{i,j} = X_{i,j} = \begin{bmatrix} 0,4084507 & 0,2588235 \\ 1 & 1 \\ 0,23943662 & 0,588235 \end{bmatrix} \times (0,6, 0,4)$$

$$V_{i,j} = \begin{bmatrix} 0,34859983 \\ 1 \\ 0,37895 \end{bmatrix}$$

Paso 4: se organizan las alternativas por orden de preferencia. Para este caso, el modelo basado en Saw estimo la variable de salida como Promedio.

Opcion 1 = 1_{promedio}

Opcion 2 = 0,37_{favorable}

Opcion 3 = 0,34_{Desfavorable}

d. Implementación de la fase jerárquica difusa

Finalmente, la **Figura 19** esquematiza como se implementó la fase multicriterio dentro del modelo jerárquico. El ingeniero de conocimiento ejecutó el modelo variando el método multicriterio. El círculo amarillo representa las reglas de Mandammi; el círculo verde, Topsis y el círculo morado, Saw. Como lo ejemplifica la **Figura 19**, la fase multicriterio difusa se ejecuta en cada una de las relaciones entre atributos básicos (p.Ej., t1, t2 y t3) y atributos agregados (p.Ej., t19 y t20) para cada nivel del modelo de forma jerárquica hasta estimar la variable objetivo (i.e., El nivel de rentabilidad).

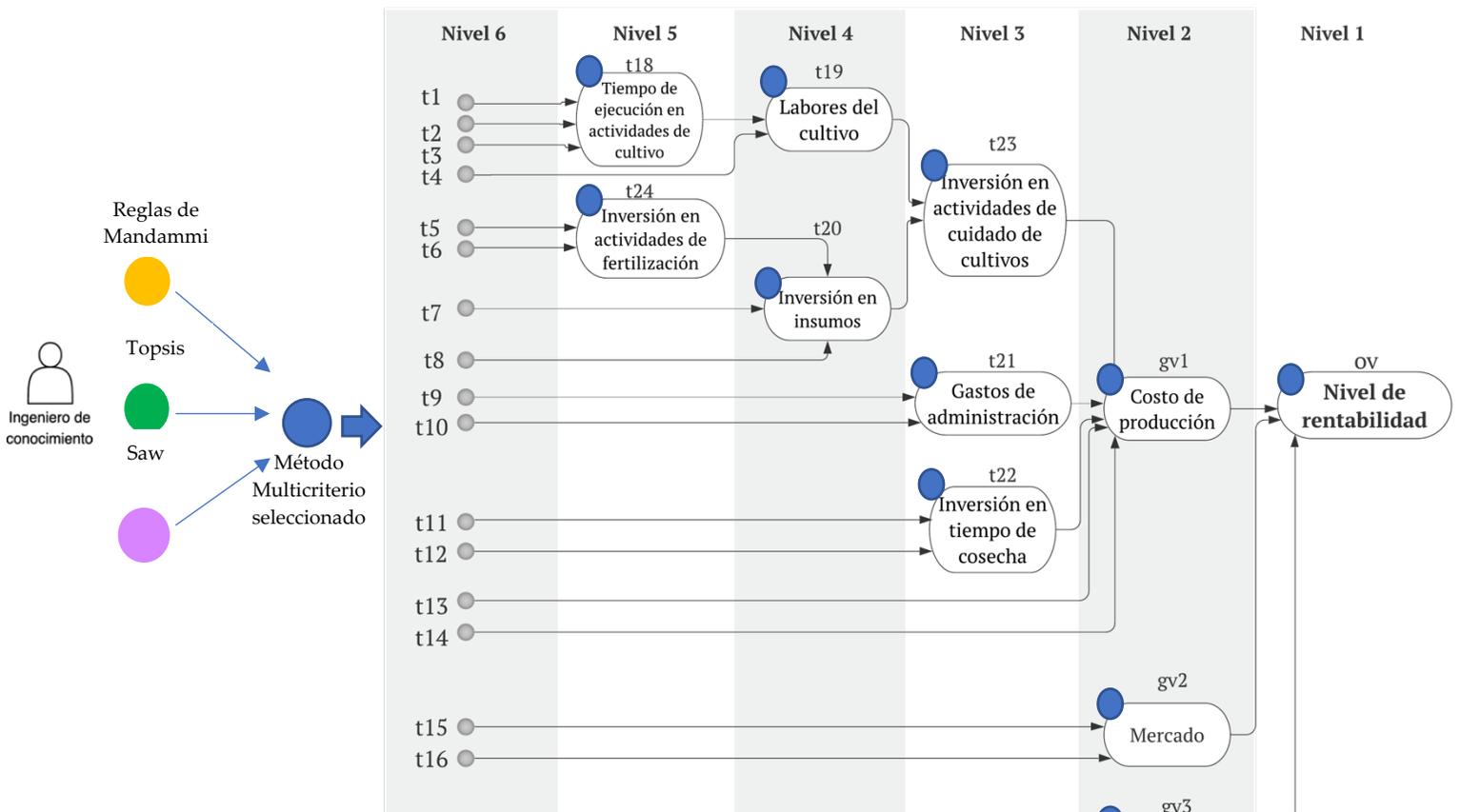


Figura 19. Implementación de la fase multicriterio difusa en cada nivel de modelo Jerárquico. Fuente: autoría propia.

4.1.3. Resumen

Este capítulo presentó el proceso de implementación de un sistema de lógica difusa dentro de un HMM. En primer lugar, el capítulo abordó las razones por las cuales fue importante incorporar este sistema dentro del HMM. En segundo lugar, presentó el proceso de adaptación de la fase jerárquica-multicriterio convencional a una fase jerárquica-multicriterio difusa. Este proceso fue desarrollado en dos pasos: fuzzificación de atributos y método multicriterio difuso. El paso de

Fuzzificación de atributos tuvo como fin representar cada atributo del modelo a través de funciones de membresía, realizado en tres actividades: la definición de escalas, construcción de las funciones y normalización. El paso de método multicriterio difuso tuvo como fin la aplicación de un método multicriterio basado en valores difusos (i.e., reglas de Mandammi, Topsis y Saw). Para este fin, utilizó las salidas de las funciones de membresía (i.e., los valores de pertenecía) como criterios de decisión. Finalmente, la salida de este paso, arrojó la decisión final del modelo jerárquico multicriterio difuso, es decir, el nivel de rentabilidad.

Capítulo 5

En este capítulo se evalúa el modelo HMP-Coffee para los 4 enfoques, el modelo convencional presentado en la **sección 3.2.2** y los tres enfoques multicriterio difusos (i.e., basado en reglas de Mandammi, Topsis y Saw). Para esta evaluación se analizan cuatro métricas de desempeño y la curva ROC, las cuales son comparadas y discutidas en los resultados.

4.2. Conjunto de datos para la evaluación de los enfoques multicriterio

El modelo HMP-Coffee es evaluado a través de un conjunto de datos proporcionado por el Instituto del Café de Costa Rica ICAFE [66]. Usamos este conjunto de Datos ya que, según la Organización Internacional del Café (International Coffee Organization -ICO), la estructura económica cafetera es similar en Costa Rica y Colombia. Por ejemplo, en 2017, el costo de producción de café en Costa Rica fue de 1956,36 [USD / ha], mientras que, en Colombia, de 1874,85 [USD / ha]. Además, en Costa Rica y Colombia, alrededor del 90% de la actividad cafetera corresponde a pequeños productores [85]. El conjunto de datos comprende 40 instancias con 17 atributos (15 relacionados a costo de producción, dos relacionado al mercado internacional y uno al volumen de producción). El conjunto de datos ICAFE está disponible en el repositorio <https://github.com/Leidy9428/HMP-Coffee/tree/main/ICAFF%20dataset%20test>. (**Anexo 4**, adicionalmente se realiza evaluación con conjunto de datos sintético)

La **Tabla 10**, presenta los 18 atributos que componen el conjunto de Datos ICAFE con su respectivo Id, unidad y descripción.

Tabla 10. Atributos del conjunto de datos para la evaluación de los enfoques multicriterio

Id	Atributos básicos	Unidad
t1	Tiempo invertido en fertilización	Hora
t2	Tiempo invertido en renovación	
t3	Tiempo invertido en el control de arvenses/plagas y enfermedades	
t4	Precio de mano de obra	CRC
t5	Fertilizantes Químicos	
t6	Fertilizantes Orgánicos	
t7	Insumos para el manejo de arvenses	
t8	Insumos para renovación del cultivo	
t9	Mantenimiento	
t10	Gastos en transporte	
t11	Pago por recolección de café	
t12	Implementos para recolectores de café	

t13	Beneficio de café	
t14	Trilla de café	
t15	Precio del Dólar	USD
t16	Precio del Café en la bolsa de NY	
t17	Volumen de Producción	Bsh /ha
vo	Nivel de Rentabilidad	

4.3. Métricas de desempeño

Para visualizar el desempeño de los cuatro métodos propuestos, es empleada la matriz de confusión. Esta matriz resume el desempeño de la clasificación de clases de la variable objetivo y proporciona una representación visual de las clasificaciones de clases reales frente a las previstas. En la matriz de confusión, cada columna representa el número de predicciones por clase mientras que cada fila representa las observaciones. La Diagonal principal indica los casos correctamente clasificados, mientras que las otras celdas muestran el número de casos clasificados de forma incorrecta. Además, a partir de la matriz de confusión se pueden calcular todas las métricas de desempeño: precisión general (Ecuación 17), precisión por clase (Ecuación 18), recuperación (Ecuación 19) y puntuación-F1 [86][87]. En la **Tabla 11** se muestra una estructura simple de una matriz de confusión que tiene tres clases.

Tabla 11.Matriz de confusión de múltiples clases

Predicción Real	Clase 1	Clase 2	Clase 3
Clase 1	a	b	c
Clase 2	d	e	f
Clase 3	g	h	i

Las métricas de desempeño fueron calculadas a partir de tasas de verdaderos positivos (TP), falsos positivos (FP), falsos negativos (FN) y verdaderos negativos (TN).

- Precisión general mide el porcentaje de casos clasificados correctamente como Rentabilidad **Favorable**, **Promedio** y **Desfavorable** por el modelo HMP-Coffee entre todos los casos posibles. Es expresada formalmente por la Ecuación 17.

$$\begin{aligned}
 \text{Precision General} &= \frac{\# \text{ Casos clasificados correctamente}}{\# \text{ Casos totales}} \\
 &= \frac{a + e + i}{a + b + c + d + e + f + g + h + i} \quad (17)
 \end{aligned}$$

- La Precisión mide el porcentaje de casos clasificados correctamente como “Rentabilidad **Promedio**” por el modelo HMP-Coffee entre todos los casos que realmente pertenecen a esta clase. Es expresada formalmente por la Ecuación 18.

$$P = TP / (TP + FP) \quad (18)$$

- La Recuperación hace referencia al porcentaje total de casos clasificados correctamente como “rentabilidad **Promedio**” por el modelo HMP-Coffee. Es expresada formalmente por la Ecuación 19.

$$\text{Recuperación} = TP / (TP + FN) \quad (19)$$

- La Puntuación-F1 es la media armónica de la Precisión y la Recuperación. Esta métrica alcanza su valor ideal en 1 (Precisión y Recuperación perfectos) y el peor caso es 0 [21]. La ecuación 20 muestra el cálculo de puntuación -F1:

$$F1 = 2 * ((P * \text{Recuperación}) / (P + \text{Recuperación})) \quad (20)$$

4.4. Resultados

La presente sección reporta los resultados del modelo HMP-Coffee para la variable de rentabilidad (clases: **Favorable**, **Promedio** y **Desfavorable**). El modelo es evaluado empleando el conjunto de datos de ICAFE [66] para cada uno de los enfoques implementados: convencional, Mandammi, Topsis y Saw.

4.4.1. Matrices de confusión por clase

En esta sección, es presentada la matriz de confusión, la precisión general y la curva ROC por cada enfoque. La diagonal principal (i.e., celdas verdes) representa los casos clasificados correctamente, mientras que las otras celdas indican el número de clasificaciones erróneas.

Tabla 12. Matriz de confusión para el modelo HMP-Coffee Convencional.

Real \ Predicción	Clase 1: Favorable	Clase 2: Promedio	Clase 3: Desfavorable
Clase 1: Favorable	4	0	0
Clase 2: Promedio	9	7	8
Clase 3: Desfavorable	0	1	11

Para HMP-Coffee Convencional, la **Tabla 12** evidencia los siguientes resultados: La clase 1 cuenta con 4 casos, todos correctamente clasificados. La clase 2 con 24 casos: 7 clasificados con éxito y 17 incorrectos: nueve como clase 1 y 8 como clase 3. Finalmente, la clase 3 cuenta con 12 casos totales: 11 clasificados con éxito y 1 caso clasificado incorrectamente como clase 2. En general, HMP-Coffee Convencional logró una precisión igual a 55%.

Tabla 13.Matriz de confusión para el modelo HMP-Coffee Reglas de Mandammi.

Predicción \ Real	Clase 1: Favorable	Clase 2: Promedio	Clase 3: Desfavorable
Clase 1: Favorable	4	0	0
Clase 2: Promedio	0	17	7
Clase 3: Desfavorable	0	0	12

Para HMP-Coffee basado en Reglas de Mandammi, la **Tabla 13** evidencia que para la clase 1 y 3 todos los casos fueron correctamente clasificados. Por su parte, la clase 2 obtuvo 17 casos clasificados con éxito y 7 casos clasificados incorrectamente como clase 3. HMP-Coffee basado en Reglas de Mandammi logró una precisión igual al 82,5%.

Tabla 14.Matriz de confusión para el modelo HMP-Coffee Topsis.

Predicción \ Real	Clase 1: Favorable	Clase 2: Promedio	Clase 3: Desfavorable
Clase 1: Favorable	2	2	0
Clase 2: Promedio	0	11	13
Clase 3: Desfavorable	0	1	11

Para HMP-Coffee difuso basado en Topsis, la **Tabla 14** muestra que para la clase 1, dos casos se clasificaron erróneamente como Clase 2. La clase 2 obtuvo 11 casos clasificados con éxito y 15 casos clasificados incorrectamente como clase 3. Finalmente, la clase 3 clasificó un caso incorrectamente como clase 2 y los 11 restantes, correctamente. En general, HMP-Coffee logró una precisión igual al 60%.

Tabla 15.Matriz de confusión para el modelo HMP-Coffee Saw.

Predicción \ Real	Clase 1: Favorable	Clase 2: Promedio	Clase 3: Desfavorable
Clase 1: Favorable	4	0	0
Clase 2: Promedio	10	2	12
Clase 3: Desfavorable	0	0	12

Para HMP-Coffee difuso basado en Saw, la **Tabla 15** evidencia que en la clase 1 y 3 todos los casos fueron correctamente clasificados. Por su parte, para la clase 2, dos casos fueron clasificados con éxito y 22 incorrectos, 10 como clase 1 y 12 como clase 3. En general, HMP-Coffee Saw logró una precisión igual al 45%.

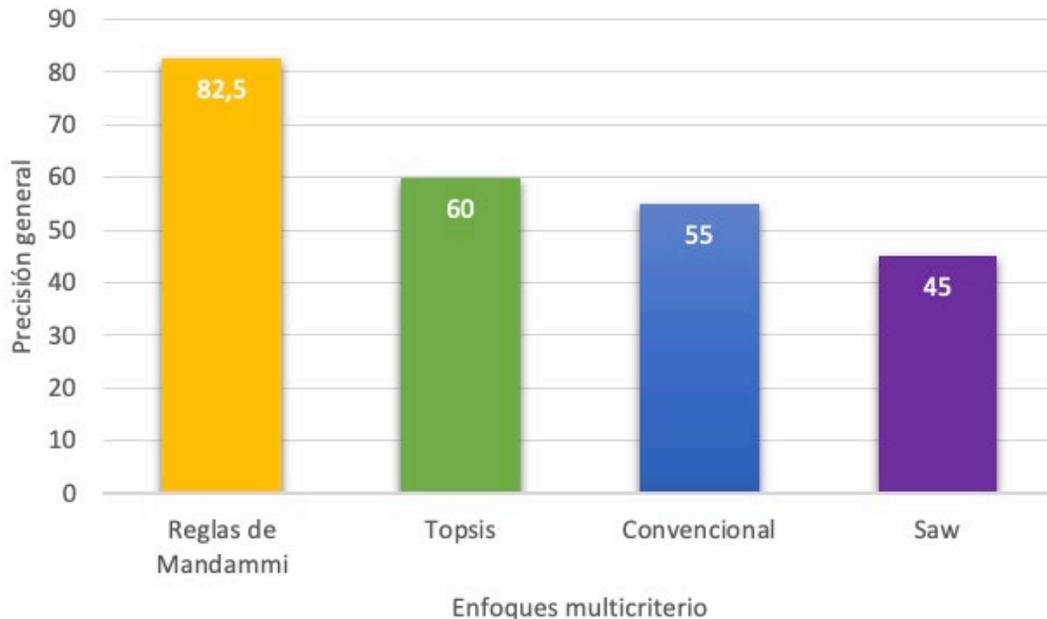


Figura 20 Comparación grafica de precisión. El eje x incluye los cuatro enfoques implementados: Reglas de Mandammi, Topsis, Convencional y Saw. El eje y corresponde al porcentaje de precisión general alcanzados por cada enfoque. Fuente: autoría propia.

La **Figura 20** para la precisión general de todos los modelos donde se observa que Reglas de Mandammi obtiene el mejor valor de precisión general con un 82,5% seguido por Topsis con 60%, Convencional con 55% y Saw con 45%. Adicionalmente, se evidencia que la precisión general del modelo convencional (i.e., alrededor del 55%) es similar al de implementaciones multicriterio convencionales desarrolladas en agricultura para la estimación de otras variables [42], [43]. De donde concluimos que, hasta el momento, independientemente del dominio de aplicación, el enfoque convencional mantiene sus niveles de precisión. Por otro lado, reglas de Mandammi logró superar en un 22,7 % a Topsis y en alrededor de un 30% al método convencional y Saw. Las matrices de confusión (Ver **Tabla 13, 14, 15 y 16**), muestran que, a diferencia de los otros enfoques, Reglas de Mandammi mejoró notablemente la estimación de la rentabilidad de las clases extremas (0% de error) y **Promedio** (la mejor estimación en relación a los otros enfoques), incrementando así, su precisión general. Para profundizar en estos resultados, en el siguiente apartado analizaremos con mayor detalle el comportamiento de los modelos para cada una de las clases (i.e., **Favorable**, **Promedio** y **Desfavorable**).

Adicionalmente, la **Figura 21** presentan las Curvas ROC para cada uno de los enfoques (e.i., la relación entre la tasa de FP/ especificidad y la TP/sensibilidad)

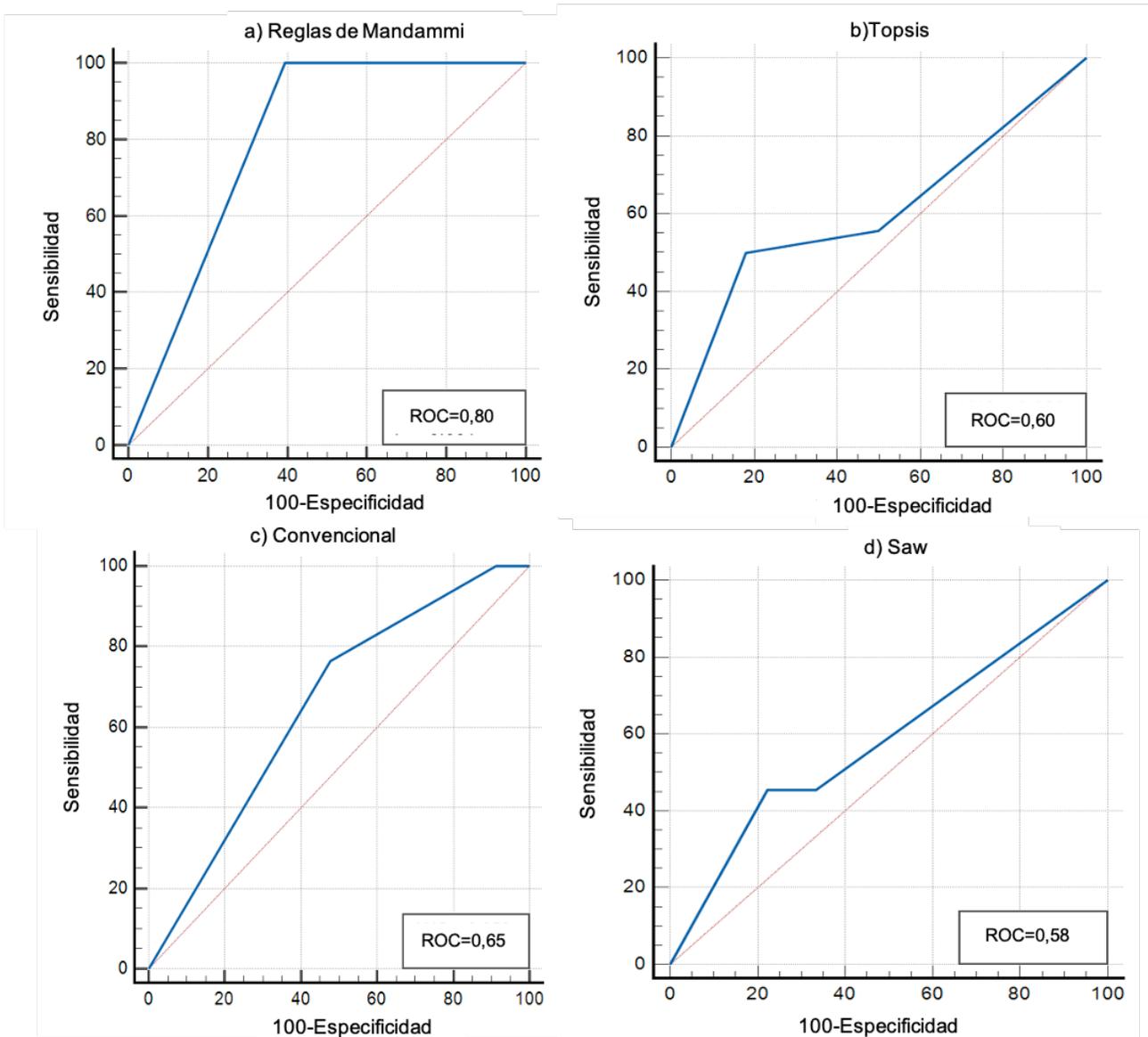


Figura 21 Curvas ROC de HMP-Coffee.(a) Reglas de Mandammi, (b) Topsis, (c) convencional y (d) Saw. El eje x corresponde a la tasa de FP's/ especificidad y el eje y a la tasa de TP's/sensibilidad. El recuadro contiene el valor de ROC/AUC a analizar. Fuente: autoría propia.

Las curvas ROC (Area Under the Curve- AUC) son una técnica de visualización del desempeño de un modelo basado en la clasificación. Generalmente, de ellas se obtiene el valor de ROC el cual varía entre 0 y 1, donde un mayor resultado hace referencia a una precisión de clasificación mas alta. En este sentido, el enfoque ideal arroja un ROC de 1. La **Figura 21** presentan las Curvas ROC para cada uno de los enfoques donde se evidencia que mejor valor de ROC esta dado por Reglas de Mandammi con un 0,83 seguido por el enfoque convencional con 0,65 , Topsis con 0,60 y Saw con 0,58. De los resultados obtenidos hasta el momento, se

obtiene a Reglas de Mandammi como el enfoque con mejor desempeño en la precisión de la clasificación, sin embargo, para analizar estos resultados más a fondo, se realiza el análisis de métricas de desempeño por clase como se desarrolla en la siguiente sección.

4.4.2. Análisis de métricas de desempeño

En esta sección se presenta el análisis de las métricas: precisión, recuperación, y puntuación-F1 [87]. A continuación, se calculan los valores de precisión, la recuperación y puntuación-F1 para cada una de las clases (i.e., **Favorable**, **Promedio** y **Desfavorable**) y enfoques (i.e., Convencional, Reglas de Mandammi, Topsis y Saw)

La **Tabla 16** presentan la Precisión, la Recuperación, y la Puntuación-F1 lograda por HMP-Coffee para las clases **Favorable**, **Promedio** y **Desfavorable** de los 4 enfoques implementados: Convencional, Reglas de Mandammi, Topsis y Saw.

Tabla 16 Resultados de los criterios estadísticos (Precisión, Recuperación, Puntuación-F1) de HMP-Coffee convencional, Reglas de Mandammi, Topsis y Saw.

		1: Favorable	2: Promedio	3: Desfavorable
Convencional	<i>Precisión</i>	0,307	0,875	0,578
	<i>Recuperación</i>	1,000	0,291	0,916
	<i>Puntuación -F1</i>	0,470	0,437	0,709
Reglas de Mandammi	<i>Precisión</i>	1,000	1,000	0,631
	<i>Recuperación</i>	1,000	0,708	1,000
	<i>Puntuación -F1</i>	1,000	0,829	0,774
Topsis	<i>Precisión</i>	1,000	0,785	0,458
	<i>Recuperación</i>	0,500	0,458	0,916
	<i>Puntuación-F1</i>	0,666	0,578	0,611
Saw	<i>Precisión</i>	0,285	1,000	0,500
	<i>Recuperación</i>	1,000	0,083	1,000
	<i>Puntuación-F1</i>	0,444	0,153	0,666

HMP-Coffee convencional (ver **Tabla 16**, color azul) clasificó con mayor precisión la clase **Promedio** (0,875), que las clases **Favorable** y **Desfavorable** (0,307 y 0,578 respectivamente). Es decir que este método funciona mejor para las escalas cualitativas intermedias que las extremas. En particular, HPM-Coffee convencional tiene problemas de detección de FP al identificar las clases **Favorable** y **Desfavorable**. El problema se produjo debido al uso de escalas cualitativas basadas en la teoría clásica de conjuntos. Por ejemplo, cuando los valores de entrada de “Precio de la mano de obra” son 10 y 19,9 dólares en una escala donde **Económico** oscila entre 0-9,9 dólares, Promedio entre 10-19,9 dólares y **Costoso** es superior a 20 dólares, se clasifican como Promedio; sin embargo, su cercanía al límite tiene una clara implicación en la precisión general del modelo, en particular en la detección de FP’s. Adicionalmente, el método arroja valores altos de *recuperación* para las clases **Favorable** y **Desfavorable** (1 y 0.916), lo que indica una apropiada detección de FN’s. Por su parte, **Favorable** y **Desfavorable** tienen una Puntuación-F1 en promedio del 58% (0.470 y 0.709) lo

que evidencia un equilibrio medio entre las métricas de Precisión y Recuperación de las clases extremas.

HMP-Coffee Reglas de Mandammi (ver **Tabla 16, color amarillo**) clasificó con la mayor precisión las clases **Favorable** y **Promedio** (1 y 1 respectivamente), mientras que la clase **Desfavorable** superó el 60% (i.e., 0,631). Este método flexibilizó las escalas cualitativas al reemplazarlas por funciones de membresía, mejorando así, la clasificación tanto de las clases extremas como las intermedias. Estos resultados reflejan que este enfoque supera el problema de detección de FP del método convencional. Como se esperaba, logró una alta recuperación para las tres clases simultáneamente (1, 0.708 y 1, respectivamente). Además de una Puntuación-F1 alta para los tres casos (1, 0.829 y 0.774 respectivamente) superando al método convencional en un 50%,40 y 10% respectivamente. Esto indica que existe un buen equilibrio entre la Precisión y la Recuperación de las tres clases.

Por su parte, HMP-Coffee Topsis (ver **Tabla 16, color verde**) clasificó con la mayor precisión la clase **Favorable** (1); sin embargo, la clase **Promedio** y **Desfavorable** no lograron a superar al método Reglas de Mandammi ni el convencional (0,785 y 0,458). Adicionalmente, HMP-Coffee Topsis alcanzó valores de Puntuación-F1 superiores a 60%. Aunque no logro igualar el desempeño de Reglas de Mandammi, superó al método convencional. En general, la flexibilización de las escalas cualitativas tuvo un efecto positivo en el funcionamiento general del modelo logrando mejorar la detección de los FP.

Finalmente, para HMP-Coffee Saw (ver **Tabla 16, color morado**) los valores de precisión no superaron el 50% en las clases extremas. La Recuperación y la Puntuación-F1 se vieron altamente afectadas. A comparación de los otros dos enfoques difusos (i.e., Reglas de Mandammi y Topsis), Saw contiene el algoritmo de toma de decisiones más simple de los métodos matemáticos multicriterio [84]. En este sentido, la flexibilización que proporcionó la lógica difusa no fue suficiente para mejorar la estimación del modelo manteniendo el problema de detección de FP's al identificar las clases extremas (obtenido Precisiones bajas de 0,285 y 0,500 respectivamente) y de FN's, para la clase media (obtenido una recuperación bajo de 0,153).

4.4.3. Discusión

A partir de los resultados de precisión general condensado en la **Figura 20** , se observó que Reglas de Mandammi presentó el mayor valor de precisión con un 82,5% seguido por Topsis con 60%, Convencional con 55% y Saw con 45%. Esto evidenció que, a diferencia de los otros enfoques, Reglas de Mandammi mejora notablemente la estimación de la rentabilidad de las clases extremas (0% de error) y **Promedio** (la mejor estimación en relación a los otros enfoques), incrementando así, su precisión general. En este sentido, es posible afirmar que, en relación a esta métrica general, el modelo multicriterio de reglas difusas de Mandammi en HMP-Coffee representa la mejor opción asegurando el 82,5% de precisión en la clasificación de la rentabilidad como **Favorable, Promedio** y **Desfavorable**.

Adicionalmente, desde los resultados obtenidos del análisis de métricas de desempeño condensados en la **Tabla 17** , se observó que este mismo método multicriterio clasificó con la

mayor precisión las clases **Favorable** y **Promedio** (1 y 1 respectivamente), mientras que la clase **Desfavorable** superó el 60% (i.e., 0,631). En este sentido es posible concluir que, reglas de Mandammi flexibilizó las escalas cualitativas al reemplazarlas por funciones de membresía, mejorando así, la clasificación tanto de las clases extremas como las intermedias. Estos resultados reflejaron que esta propuesta supera el problema de detección de FP del método convencional. También logró una alta recuperación para las tres clases simultáneamente (1, 0.708 y 1, respectivamente) y Puntuación-F1 alta para los tres casos (1, 0.829 y 0.774 respectivamente) superando al método convencional en un 50%,40 y 10% respectivamente. Esto indica que existe un buen equilibrio entre la Precisión y la Recuperación de las tres clases.

4.4.4. Resumen

Este capítulo evaluó el modelo HMP-Coffee para los cuatro enfoques, el modelo convencional presentado en la **sección 3.2.2** y los tres enfoques multicriterio difusos (i.e., basado en reglas de Mandammi, Topsis y Saw) con el fin de analizar y comparar su desempeño. Esta evaluación hizo uso de la matriz de confusión multi-clase para el cálculo de las métricas de desempeño: precisión general, curva ROC, precisión por clase, recuperación y puntuación-F1. En el análisis general entre modelos dado por la precisión general, se evidenció que las reglas de Mandammi arrojó el mejor resultado con un 82,5%, seguido por Topsis con 60%, Convencional con 55% y Saw con 45%. En el análisis por clases los resultados mostraron que, HMP-Coffee Reglas de Mandammi presentó el mejor desempeño entre los cuatro modelos, clasificando con la más alta precisión las clases **Favorable** y **Promedio** (1 y 1 respectivamente), y con un valor superior al 60% (i.e., 0,631) la clase **Desfavorable**. Este comportamiento también se vio reflejado en los altos valores de recuperación (e.i., 1, 0.708 y 1,) y puntuación (e.i., 1, 0.829 y 0.774) resultantes para las tres clases simultáneamente, superando así, los otros tres enfoques.

5. Conclusiones y trabajos futuros

5.1. Conclusiones

En el desarrollo de un modelo cualitativo jerárquico basado en conocimiento experto para la estimación precisa de la rentabilidad en producciones de café sostenible a pequeña escala, se obtienen las siguientes conclusiones:

- En general, HMP-Coffee basado en Reglas de Mandammi (i.e., El enfoque con mejor desempeño) puede estimar el nivel de rentabilidad de pequeñas producciones de café con una precisión de aproximadamente el 82,5%. Es de destacar que, a diferencia de los trabajos relacionados en la sección 2.2.1, el enfoque propuesto ofrece las siguientes características en conjunto: no dirigido por datos, es fácil de interpretar y considera los tres factores de rentabilidad del café. A diferencia de HMP-Coffee, Milne et al. [39] estimó la rentabilidad con un enfoque basado en datos; este enfoque es difícil de adoptar en los países en desarrollo donde la adopción de herramientas modernas de gestión de la información en las pequeñas fincas de café es escasa [41]. A diferencia de Rising et al. [36] y Haaster et al. [38] que estima la rentabilidad con base en variables climáticas en diversos tipos de cultivo, HMP-Coffee realiza estimaciones considerando

las particularidades de la rentabilidad del cultivo de café definidas por ICO. Shakoor et al [37] implementó los enfoques de inteligencia artificial aplicando los algoritmos K-Neighbors Neighbors para seis cultivos de arroz Aus, arroz Aman, arroz Boro, papa, yute y trigo; sin embargo, los modelos de salida son caja negra. Finalmente, a diferencia de Aubertot et al. [42], [43] que presentaron enfoques jerárquicos basado en conocimiento experto (i.e., Enfoque convencional) con resultados de precisión inferiores al 60%, HMP-Coffee basado en Reglas de Mandammi arrojó resultados de precisión superiores al 80%. Es de resaltar que las características en conjunto que ofrece el enfoque propuesto, están alineadas con las premisas identificadas desde las particularidades de las pequeñas producciones de café sostenible en países en vía de desarrollo.

Adicionalmente es de notar que, a diferencia de los trabajos relacionado de la sección 2.2.2, donde se proponía la utilización de lógica difusa sobre modelos de predicción, pronóstico o estimación de un “único” nivel de profundidad y caja negra con alto grado de complejidad, HMP-Coffee basado en Reglas de Mandammi incorporó este concepto de forma desagregada a lo largo de sus 6 niveles, garantizando un modelo con alto nivel de interpretabilidad, a la vez que, proporcionó altos niveles de precisión superior al 80% (i.e., El modelo basado en reglas de Mandammi), superando así, enfoques convencionales como los propuestos por Aubertot et al. [42] [43].

- El enfoque HMP-Coffee convencional clasificó con mayor precisión la clase Promedio (0,875), que las clases **Favorable** y **Desfavorable** (0,307 y 0,578 respectivamente). Es decir que este método funciona mejor para las escalas cualitativas intermedias que las extremas. En particular, HPM-Coffee convencional sufre de FP al identificar las clases **Favorable** y **Desfavorable**. El problema de detección de FP's, se produjo debido al uso de escalas cualitativas basadas en la teoría clásica de conjuntos. Este problema conlleva a que la precisión de estos enfoques no supere el 60%.
- El enfoque HPM-Coffee basado en Reglas de Mandammi para la estimación de la rentabilidad en producciones de café sostenible a pequeña escala mejora considerablemente la precisión de la estimación en relación al enfoque convencional. Este método flexibilizó las escalas cualitativas al reemplazarlas por funciones de membresía, mejorando así, la clasificación tanto de las clases extremas como las intermedias. Estos resultados reflejan que este enfoque supera el problema de detección de FP del método convencional. Como se esperaba, logró una alta recuperación para las tres clases simultáneamente (1, 0,708 y 1, respectivamente). Además de una Puntuación-F1 alta para los tres casos (1, 0,829 y 0,774 respectivamente) superando al método convencional en un 50%,40% y 10% respectivamente. Esto indica que existe un buen equilibrio entre la Precisión y la Recuperación de las tres clases, y consecuentemente, el mejor desempeño entre los 4 enfoques evaluados.

- El enfoque HMP-Coffee Topsis clasificó con la mayor precisión la clase **Favorable** (1); sin embargo, la clase **Promedio** y **Desfavorable** no lograron a superar al método Reglas de Mandammi ni el convencional (0,785 y 0,458). Adicionalmente, HMP-Coffee Topsis alcanzó valores de Puntuación-F1 superiores a 60%. Aunque no logró igualar el desempeño de Reglas de Mandammi, superó al método convencional. En general, la flexibilización de las escalas cualitativas tuvo un efecto positivo en el funcionamiento general del modelo logrando mejorar la detección de los FP.
- En el enfoque HMP-Coffee Saw, los valores de precisión no superaron el 50% en las clases extremas. Como es de esperarse, la Recuperación y la Puntuación-F1 se vieron altamente afectadas. A comparación de los otros dos enfoques difusos (i.e., Reglas de Mandammi y Topsis), Saw contiene el algoritmo de toma de decisiones más simple de los métodos matemáticos multicriterio [84]. En este sentido, la flexibilización que proporcionó la lógica difusa no fue suficiente para mejorar la estimación del modelo manteniendo el problema de detección de FP's al identificar las clases extremas (obtenido precisiones bajas de 0,285 y 0,500 respectivamente) y de FN's, para la clase media (obtenido una baja recuperación de 0,153).
- En general, el enfoque HMP-Coffee está constituido por un modelo conceptual, su implementación y evaluación correspondiente. Cabe destacar que, enfoque propuesto ofrece las siguientes características en conjunto: no dirigido por datos, es fácil de interpretación (aportando múltiples niveles de profundidad), considera los tres factores de rentabilidad del café y garantiza altos niveles de precisión. Para lograr esto, el modelo se concibe con base en el conocimiento de expertos, superando así, la falta de datos económicos sobre pequeños cultivos de café en los países en desarrollo; mantiene una naturaleza jerárquica cualitativa adecuada para obtener modelos interpretables e incorpora la Lógica difusa para flexibilizar las escalas cualitativas y garantizar altos niveles de precisión en la estimación.
- HMP-Coffee es un enfoque pensado para estimar el nivel de rentabilidad en pequeñas producciones de café donde existe un escaso control de registros escritos relacionados a la economía del cultivo. En este sentido, esta propuesta se presenta como una solución útil en países en vía de desarrollo, donde alrededor del 90% de los pequeños productores no registra los movimientos contables asociados a su producción. Es de resaltar que el modelo no está enfocado en grandes producciones de café y/o tecnificadas donde cada actividad es monitoreada, situación frecuente en países desarrollados productores de café como Vietnam.
- Para la evaluación se utilizó el conjunto de datos de ICAFE, el cual proporcionó todas las características de entrada requeridas por el modelo. Los datos de evaluación fueron seleccionados debido a que, según la Organización Internacional del Café, la estructura económica cafetera es similar en Costa Rica y Colombia. En este sentido, es posible resaltar que la evaluación desarrollada, es válida en contextos cafeteros similares al de

estos dos países. Para validar el modelo en otros contextos, basta ajustar los límites de las condiciones de entrada de cada región (con ayuda de un panel de expertos).

5.2. Trabajos futuros

1. Actualmente, la construcción de las funciones de membresía y el establecimiento de los límites de los estados cualitativos en los sistemas de lógica difusa son definidos por modelos manuales o basados en conocimiento experto. La introducción del Aprendizaje Automático (ML- Machine Learning) es una alternativa factible para definir de forma autónoma límites que se adapten a las condiciones de los diferentes contextos cafeteros (e.i., de producciones de café en diferentes países o regiones). De esta manera, el modelo se entrenaría con datos de diversos lugares para definir los límites de las escalas cualitativas dentro de las funciones de membresía, proporcionando así, mayor autonomía y capacidad de adaptación para estimar la rentabilidad en las pequeñas producciones de café.
2. Actualmente, la fase de conocimiento del contexto genera el modelo jerárquico conceptual basado en técnicas de extracción de conocimiento experto comúnmente aplicadas en la ingeniería de conocimiento. Estas técnicas, en su gran mayoría tienen una alta interacción con los profesionales en el área y gran parte de los resultados dependen de sus experiencias. La minería de texto combinadas con técnicas de inteligencia artificial proporcionarían la posibilidad de analizar mayor cantidad de información para la construcción autónoma de uno o muchos modelos conceptuales relacionados a un mismo dominio de aplicación (p.Ej., la rentabilidad en pequeñas producciones). De esta manera, sería posible obtener modelos jerárquicos conceptuales con información escrita o experiencial diversificada que incorporen la mayor cantidad de características independiente del país o la región.
3. Ampliar la base de conocimiento actual (basada en la operación de la finca “La Sultana”), con información de pequeñas producciones cafeteras ubicadas en distintas ciudades de Colombia o países de Centro y Sur América. Esto puede contribuir a la generación de un modelo más completo, al considerar un mayor número de características de los diferentes territorios en el contexto de la agricultura a pequeña escala.
4. Actualmente, HMP-Coffee considera el volumen de producción, como una variable conocida por el caficultor utilizando algún método propio (p. Ej. métodos manuales de conteo). Sin embargo, incorporar en el enfoque propuesto, un sistema propio de predicción del volumen de producción, que, a su vez, considere el comportamiento económico del cultivo para estimar la rentabilidad, le proporcionarían mayor autonomía y confiabilidad al modelo.

6. Referencias

- [1] "OEC - The Observatory of Economic Complexity." <https://oec.world/> (accessed Dec. 30, 2020).
- [2] P. Shah, D. Hiremath, and S. Chaudhary, "Big Data Analytics Architecture for Agro Advisory System," in *2016 IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW)*, Dec. 2016, pp. 43–49. doi: 10.1109/HiPCW.2016.015.
- [3] P. Mohan and K. K. Patil, "Crop production rate estimation using parallel layer regression with deep belief network," in *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, Dec. 2017, pp. 168–173. doi: 10.1109/ICEECCOT.2017.8284659.
- [4] K. B. Ramí-rez and W. D. L. Bohorquez, "Producción agroecológica vs producción convencional: el problema de la creación de valor y el mercado. discusión en perspectiva marxista.," *Rev. Kavilando*, vol. 10, no. 1, Art. no. 1, Feb. 2018.
- [5] A. Rojas Sanchez, P. U. Javeriana, K. Hartman Ulloa, P. U. Javeriana, R. A. Marques, and U. E. de Colombia, "El impacto de la producción de café sobre la biodiversidad, la transformación del paisaje y las especies exóticas invasoras," *Ambiente Desarro. Vol 16 Núm 30 2012 Morfol. Ecotechos Biodivers. 93-104*, Jun. 2012, Accessed: May 24, 2021. [Online]. Available: <http://repository.javeriana.edu.co/handle/10554/22838>
- [6] V. Ml. C. Granados, "AGROECOLOGÍA DEL AGROECOSISTEMA CAFÉ ('COFFEA ARABICA') Y SU RELACIÓN CON LA ERODABILIDAD DE LADERAS EN EL VALLE DE OROSI, CARTAGO, COSTA RICA," *Anu. Estud. Centroam.*, vol. 37, pp. 271–305, 2011.
- [7] A. M. del Puerto Rodríguez, S. Suárez Tamayo, and D. E. Palacio Estrada, "Efectos de los plaguicidas sobre el ambiente y la salud," *Rev. Cuba. Hig. Epidemiol.*, vol. 52, no. 3, pp. 372–387, Dec. 2014.
- [8] V. Kumar, R. Kumar, J. Singh, and P. Kumar, *Contaminants in Agriculture and Environment: Health Risks and Remediation*. Agro Environ Media, Publication Cell of AESA, Agriculture and Environmental Science Academy, 2019.
- [9] A. Yong-Chou, A. Crespo-Morales, B. Benítez-Fernández, M. I. Pavón-Rosales, and G. R. Almenares-Garlobo, "Uso y manejo de prácticas agroecológicas en fincas de la localidad de san andrés, municipio La Palma," *Cultiv. Trop.*, vol. 37, no. 3, pp. 15–21, Sep. 2016, doi: 10.13140/RG.2.1.2756.3761.
- [10] "Caracterización de nueve agroecosistemas de café de la cuenca del río Porce, Colombia, con un enfoque agroecológico." https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-34292015000100008 (accessed May 24, 2021).
- [11] A. Hruska, *Memorias de Talleres de Agroecología y Caficultura en Mesoamérica y República Dominicana*. 2017.
- [12] K. Padmavathy and P. Gopalsamy, "Alternative Farming Techniques for Sustainable Food Production," vol. 7, 2011, pp. 367–424. doi: 10.1007/978-94-007-1521-9_13.
- [13] R. V. Vallejo, "La rentabilidad del caficultor, un compromiso de todos.," 2016. [Online]. Available: https://federaciondecafeteros.org/static/files/Periodico_2016Baja2.pdf

- [14] “6 Ways to Increase Your Coffee Shop Profits | National Coffee Association Blog.” <https://nationalcoffee.blog/2016/10/28/6-ways-to-increase-your-coffee-shop-profits/> (accessed May 24, 2021).
- [15] “World Food Preservation Center® LLC | International Society for Horticultural Science.” http://www.actahort.org/books/1194/1194_2.htm (accessed May 24, 2021).
- [16] A. Nabavi-Pelesaraei, S. Rafiee, S. S. Mohtasebi, H. Hosseinzadeh-Bandbafha, and K. Chau, “Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production,” *Sci. Total Environ.*, vol. 631–632, pp. 1279–1294, Aug. 2018, doi: 10.1016/j.scitotenv.2018.03.088.
- [17] W. Samek, T. Wiegand, and K.-R. Müller, “Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models,” *ArXiv170808296 Cs Stat*, Aug. 2017, Accessed: May 24, 2021. [Online]. Available: <http://arxiv.org/abs/1708.08296>
- [18] “Identificación con estimación para sistemas tipo caja negra.” http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-62302014000300004 (accessed May 24, 2021).
- [19] R. G. García Cáceres and É. S. Olaya Escobar, “Caracterización de las cadenas de valor y abastecimiento del sector agroindustrial del café,” *Cuad. Adm.*, vol. 19, no. 31, pp. 197–217, Jun. 2006.
- [20] “Los pequeños productores y su participación en el proceso de comercialización agrícola.” http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0378-78182004000200007 (accessed May 24, 2021).
- [21] C. Tancara Q, “LA INVESTIGACION DOCUMENTAL,” *Temas Soc.*, no. 17, pp. 91–106, Dec. 1993.
- [22] L. Delgado Montenegro, A. Cortez Vásquez, and E. Ibáñez Prentice, “Aplicación de metodología Buchanan para la construcción de un sistema experto con redes bayesianas para apoyo al diagnóstico de la tetralogía de Fallot en el Perú,” *Ind. Data*, vol. 18, no. 1, p. 135, Jun. 2016, doi: 10.15381/idata.v18i1.12076.
- [23] “International Coffee Organization.” <http://www.ico.org/> (accessed Feb. 17, 2021).
- [24] “Eje Económico FNC,” *Federación Nacional de Cafeteros*. <https://federaciondefcafeteros.org/wp/sostenibilidad/eje-economico/> (accessed Dec. 16, 2020).
- [25] “Administración y rentabilidad en las fincas cafeteras | Agronomía Colombiana.” <https://revistas.unal.edu.co/index.php/agrocol/article/view/21137> (accessed May 24, 2021).
- [26] B. De La Hoz Suárez, M. A. Ferrer, and A. De La Hoz Suárez, “Indicadores de rentabilidad: herramientas para la toma decisiones financieras en hoteles de categoría media ubicados en Maracaibo,” *Rev. Cienc. Soc.*, vol. 14, no. 1, pp. 88–109, Apr. 2008.
- [27] J. Dixon, D. P. Gibbon, A. Gulliver, and M. Hall, *Farming systems and poverty: improving farmers' livelihoods in a changing world*. Rome : Washington, D.C: FAO ; World Bank, 2001.
- [28] F. Farfán-Valencia and P. M. Sánchez-Arciniegas, “CERTIFICACIÓN DE FINCAS DE PRODUCCIÓN DE CAFÉ ORGÁNICO,” p. 8.
- [29] “Agricultura sostenible | Objetivos de Desarrollo Sostenible | Organización de las Naciones Unidas para la Alimentación y la Agricultura.” <http://www.fao.org/sustainable-development-goals/overview/fao-and-post-2015/sustainable-agriculture/es/> (accessed May 25, 2021).
- [30] M. Bohanec, “DEXi: A Program for Multi-Attribute Decision Making,” *DEXi: A Program for Multi-Attribute Decision Making*, 2020. <http://kt.ijs.si/MarkoBohanec/dexi.html> (accessed Dec. 03, 2020).

- [31] C. Bockstaller, S. Beauchet, V. Manneville, B. Amiaud, and R. Botreau, "A tool to design fuzzy decision trees for sustainability assessment," *Environ. Model. Softw.*, vol. 97, pp. 130–144, Nov. 2017, doi: 10.1016/j.envsoft.2017.07.011.
- [32] L. Andriantiatsaholiniana, V. Kouikoglou, and Y. Phillis, "Evaluating strategies for sustainable development: Fuzzy logic reasoning and sensitivity analysis," *Ecol. Econ.*, vol. 48, pp. 149–172, Feb. 2004, doi: 10.1016/j.ecolecon.2003.08.009.
- [33] B. J. Wielinga and B. Bredeweg, "Knowledge and Expertise in Expert Systems," in *Human-Computer Interaction*, Berlin, Heidelberg, 1988, pp. 290–297. doi: 10.1007/978-3-642-73402-1_18.
- [34] M. De Salvo, D. Begalli, R. Capitello, L. Agnoli, and E. Tabouratzi, "Determinants of winegrowers' profitability: evidence from an Eastern Europe wine region," *EuroMed J. Bus.*, vol. 12, no. 3, pp. 300–315, Jan. 2017, doi: 10.1108/EMJB-12-2016-0043.
- [35] Hawes *et al.*, "Whole-Systems Analysis of Environmental and Economic Sustainability in Arable Cropping Systems: A Case Study," *Agronomy*, vol. 9, no. 8, p. 438, Aug. 2019, doi: 10.3390/agronomy9080438.
- [36] J. Rising and N. Devineni, "Crop switching reduces agricultural losses from climate change in the United States by half," *Nat. Commun.*, vol. 11, Oct. 2020, doi: 10.1038/s41467-020-18725-w.
- [37] M. T. Shakoor, K. Rahman, S. N. Rayta, and A. Chakrabarty, "Agricultural production output prediction using Supervised Machine Learning techniques," in *2017 1st International Conference on Next Generation Computing Applications (NextComp)*, Jul. 2017, pp. 182–187. doi: 10.1109/NEXTCOMP.2017.8016196.
- [38] F. Haaster and T. Kleynhans, "Decision tree Modelling to support investment decisions on flower production for the northern Hemisphere," *Agrekon*, vol. 41, pp. 75–96, Mar. 2002, doi: 10.1080/03031853.2002.9523587.
- [39] A. E. Milne, K. Coleman, L. C. Todman, and A. P. Whitmore, "Model-based optimisation of agricultural profitability and nutrient management: a practical approach for dealing with issues of scale," *Environ. Monit. Assess.*, vol. 192, no. 11, p. 730, Oct. 2020, doi: 10.1007/s10661-020-08699-z.
- [40] N. P. Cardozo, R. de O. Bordonal, A. R. Panosso, and C. A. C. Crusciol, "A Multivariate Approach to Determine the Economic Profitability of Sugarcane Production Under Diverse Climatic Conditions in Brazil," *Sugar Tech*, vol. 22, no. 6, pp. 954–966, Dec. 2020, doi: 10.1007/s12355-020-00854-7.
- [41] A. Stratigea, "ICTs for rural development: potential applications and barriers involved," *Netcom Réseaux Commun. Territ.*, no. 25–3/4, Art. no. 25–3/4, Dec. 2011, doi: 10.4000/netcom.144.
- [42] J.-N. Aubertot and M.-H. Robin, "Injury Profile SIMulator, a qualitative aggregative modelling framework to predict crop injury profile as a function of cropping practices, and the abiotic and biotic environment. I. Conceptual bases," *PloS One*, vol. 8, no. 9, p. e73202, 2013, doi: <https://doi.org/10.1371/journal.pone.0073202>.g001.
- [43] M.-H. Robin *et al.*, "IPSIM-Web, An Online Resource for Promoting Qualitative Aggregative Hierarchical Network Models to Predict Plant Disease Risk: Application to Brown Rust on Wheat," *Plant Dis.*, vol. 102, no. 3, pp. 488–499, Mar. 2018, doi: 10.1094/PDIS-12-16-1816-SR.

- [44] “Principal barriers to the adoption of ICTs in agriculture and in rural areas | Repositorio Digital | Comisión Económica para América Latina y el Caribe.” <https://repositorio.cepal.org/handle/11362/35440> (accessed Dec. 03, 2020).
- [45] A. M. G. Cornelissen, J. van den Berg, W. J. Koops, M. Grossman, and H. M. J. Udo, “Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory,” *Agric. Ecosyst. Environ.*, vol. 86, no. 2, pp. 173–185, Aug. 2001, doi: 10.1016/S0167-8809(00)00272-3.
- [46] A. Mohapatra and S. Lenka, “Hybrid decision support system using PLSR-fuzzy model for GSM-based site-specific irrigation notification and control in precision agriculture,” *Int. J. Intell. Syst. Technol. Appl.*, vol. 15, p. 4, Jan. 2016, doi: 10.1504/IJISTA.2016.076101.
- [47] J. Subramanian, N. Palakkal, E. Papageorgiou, and K. Papageorgiou, “Application of fuzzy cognitive maps in precision agriculture: a case study on coconut yield management of southern India’s Malabar region,” *Neural Comput. Appl.*, vol. 26, Nov. 2015, doi: 10.1007/s00521-015-1864-5.
- [48] M. R. Taghizadeh Yazdi, H. Shakouri G., E. Asgharizadeh, and M. Sakawa, “Design of a multi-level fuzzy linear regression model for forecasting: A case study of Iran,” *Sci. Res. Essays*, vol. 6, pp. 5072–5081, Oct. 2011.
- [49] F. Agin, R. Khosravanian, M. Karimifard, and A. Jahanshahi, “Application of adaptive neuro-fuzzy inference system and data mining approach to predict lost circulation using DOE technique (case study: Maroon oilfield),” *Petroleum*, vol. 6, no. 4, pp. 423–437, Dec. 2020, doi: 10.1016/j.petlm.2018.07.005.
- [50] “A data portrait of smallholder farmers | FAO.” <http://www.fao.org/family-farming/detail/es/c/385074/> (accessed Dec. 03, 2020).
- [51] M. Pramanik, “Site suitability analysis for agricultural land use of Darjeeling district using AHP and GIS techniques,” *Model. Earth Syst. Environ.*, 2016, doi: 10.1007/s40808-016-0116-8.
- [52] “(3) A fuzzy multicriteria benefit-cost approach for irrigation projects evaluation | Request PDF.” https://www.researchgate.net/publication/227411441_A_fuzzy_multicriteria_benefit-cost_approach_for_irrigation_projects_evaluation (accessed Aug. 03, 2021).
- [53] “A Comparison of Parametric and Fuzzy Multi-Criteria Methods for Evaluating Land Suitability for Olive in Jeffara Plain of Libya | Semantic Scholar.” <https://www.semanticscholar.org/paper/A-Comparison-of-Parametric-and-Fuzzy-Multi-Criteria-Elaalem/29f9d846e4937fd2a39f6ec1da4d86ee66ab6879> (accessed Aug. 03, 2021).
- [54] “Site suitability analysis for potential agricultural land with spatial fuzzy multi-criteria decision analysis in regional scale under semi-arid terrestrial ecosystem | Scientific Reports.” <https://www.nature.com/articles/s41598-020-79105-4> (accessed Aug. 03, 2021).
- [55] M. A. Elleuch, M. Anane, J. Euch, and A. Frikha, “Hybrid fuzzy multi-criteria decision making to solve the irrigation water allocation problem in the Tunisian case,” *Agric. Syst.*, vol. 176, p. 102644, Nov. 2019, doi: 10.1016/j.agry.2019.102644.
- [56] “Evaluation of potential irrigation expansion using a spatial fuzzy multi-criteria decision framework,” *Land Portal*, Jul. 27, 2017. <https://landportal.org/es/library/resources/agrisus201500200282/evaluation-potential-irrigation-expansion-using-spatial-fuzzy> (accessed Aug. 02, 2021).

- [57] G. R. Dattatreya, "Decision Trees," in *Artificial Intelligence Methods in the Environmental Sciences*, S. E. Haupt, A. Pasini, and C. Marzban, Eds. Dordrecht: Springer Netherlands, 2009, pp. 77–101. doi: 10.1007/978-1-4020-9119-3_4.
- [58] E. Giusti and S. Marsili-Libelli, "A Fuzzy Decision Support System for irrigation and water conservation in agriculture," *Environ. Model. Softw.*, vol. 63, pp. 73–86, Jan. 2015, doi: 10.1016/j.envsoft.2014.09.020.
- [59] S. Morales-Velasco, N. J. Vivas-Quila, and V. Gomez-Medina, "Variable biophysics of twelve systems agrosilvopastoriles in the plateau of popayán - cauca," *Biotechnol. En El Sect. Agropecu. Agroindustrial*, vol. 13, no. 2, pp. 11–21, Dec. 2015, doi: 10.18684/BSAA(13)11-21.
- [60] J. M. Muñoz C, J. A. Muñoz P, and C. Montes R, "Evaluation of organic manures in lettuce and cabbage plants at Popayán, Cauca," *Biotechnol. En El Sect. Agropecu. Agroindustrial*, vol. 13, no. 1, pp. 73–82, Jun. 2015.
- [61] A. D. P. Campo-Martínez, R. L. Acosta-Sanchez, S. Morales-Velasco, and F. A. Prado, "Evaluation of microorganisms of mountain (mm) in the production of chard on the plateau of Popayán," *Biotechnol. En El Sect. Agropecu. Agroindustrial*, vol. 12, no. 1, pp. 79–87, Jun. 2014.
- [62] A. Ordonez, "La Sultana Farm," *Universidad del Cauca-Unidad experimental*. <https://faca.unicauca.edu.co/cienciasagrarias/infraestructura> (accessed Dec. 03, 2020).
- [63] H. Adeli, *Expert Systems in Construction and Structural Engineering*. CRC Press, 2003.
- [64] J. Qi, L. Wei, and Q. Wan, "Multi-level granularity in formal concept analysis," *Granul. Comput.*, vol. 4, no. 3, pp. 351–362, Jul. 2019, doi: 10.1007/s41066-018-0112-7.
- [65] H. Ahonen-Myka, "Discovery of Frequent Word Sequences in Text," vol. 2447, Oct. 2002, doi: 10.1007/3-540-45728-3_14.
- [66] ICAFE, "Costo de la Actividad Cafetalera," *ICAFE*. <http://www.icafe.cr/sector-cafetalero/informacion-de-mercado/costo-de-la-actividad-cafetalera/> (accessed Dec. 15, 2020).
- [67] "Caravela Coffee." <https://caravela.coffee/> (accessed Feb. 17, 2021).
- [68] L. Guo, F. Shi, and J. Tu, "Textual analysis and machine learning: Crack unstructured data in finance and accounting," *J. Finance Data Sci.*, vol. 2, no. 3, pp. 153–170, Sep. 2016, doi: 10.1016/j.jfds.2017.02.001.
- [69] U. Grömping, "Using R and RStudio for Data Management, Statistical Analysis and Graphics (2nd Edition)," *J. Stat. Softw.*, vol. 68, Nov. 2015, doi: 10.18637/jss.v068.b04.
- [70] N. Shadbolt and P. R. Smart, "Knowledge Elicitation: Methods, Tools and Techniques," in *Evaluation of Human Work*, J. R. Wilson and S. Sharples, Eds. CRC Press, 2015, pp. 163–200. Accessed: Dec. 03, 2020. [Online]. Available: <https://eprints.soton.ac.uk/359638/>
- [71] M. Gonzalez-Perez and S. Gutierrez-Viana, "Cooperation in coffee markets: the case of Vietnam and Colombia," *J. Agribus. Dev. Emerg. Econ.*, vol. 2, no. 1, pp. 57–73, Jan. 2012, doi: 10.1108/20440831211219237.
- [72] M. Bohanec, M. Žnidaršič, V. Rajkovic, I. Bratko, and B. Zupan, "DEX Methodology: Three Decades of Qualitative Multi-Attribute Modeling," *Informatika*, vol. 37, pp. 49–54, Mar. 2013.
- [73] "Asociacion cafiambiente del Cauca." <http://colombiapymes.com/ficha/asociacion-cafiambiente-cauca-208962> (accessed Feb. 17, 2021).
- [74] C. J. O'Donnell, "Measuring and decomposing agricultural productivity and profitability change*," *Aust. J. Agric. Resour. Econ.*, vol. 54, no. 4, pp. 527–560, 2010, doi: <https://doi.org/10.1111/j.1467-8489.2010.00512.x>.

- [75] S. A. Drakopoulos and K. Grimani, "Maslow's Needs Hierarchy and the Effect of Income on Happiness Levels," Oct. 2013. <https://mpr.ub.uni-muenchen.de/50987/> (accessed Apr. 27, 2021).
- [76] L. F. G. Pinto, T. Gardner, C. L. McDermott, and K. O. L. Ayub, "Group certification supports an increase in the diversity of sustainable agriculture network–rainforest alliance certified coffee producers in Brazil," *Ecol. Econ.*, vol. 107, pp. 59–64, Nov. 2014, doi: 10.1016/j.ecolecon.2014.08.006.
- [77] S.-M. CHEN and C.-H. CHANG, "A New Method to Construct Membership Functions and Generate Weighted Fuzzy Rules from Training Instances," *Cybern. Syst.*, vol. 36, no. 4, pp. 397–414, Jun. 2005, doi: 10.1080/01969720490929562.
- [78] M. Velasquez and P. T. Hester, "An Analysis of Multi-Criteria Decision Making Methods." 2013.
- [79] S. S. Izquierdo and L. R. Izquierdo, "Mamdani Fuzzy Systems for Modelling and Simulation: A Critical Assessment," *J. Artif. Soc. Soc. Simul.*, vol. 21, no. 3, p. 2, 2018.
- [80] F. Cavallaro, "A Takagi-Sugeno Fuzzy Inference System for Developing a Sustainability Index of Biomass," *Sustainability*, vol. 7, no. 9, Art. no. 9, Sep. 2015, doi: 10.3390/su70912359.
- [81] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man-Mach. Stud.*, vol. 7, no. 1, pp. 1–13, Jan. 1975, doi: 10.1016/S0020-7373(75)80002-2.
- [82] P. Zhang, "CHAPTER 7 - Industrial intelligent controllers," in *Advanced Industrial Control Technology*, P. Zhang, Ed. Oxford: William Andrew Publishing, 2010, pp. 257–305. doi: 10.1016/B978-1-4377-7807-6.10007-5.
- [83] I. P. Vania Peneva, "Fuzzy Logic Operators in Decision-Making," *Cybern. Syst.*, vol. 30, no. 8, pp. 725–745, Nov. 1999, doi: 10.1080/019697299124966.
- [84] A. Salehi and M. Izadikhah, "A novel method to extend SAW for decision-making problems with interval data," *Decis. Sci. Lett.*, vol. 3, no. 2, pp. 225–236, 2014.
- [85] J. Berdegué and R. Fuentealba, "Latin America: The State of Smallholders in Agriculture," Mar. 2014, doi: 10.1093/acprof:oso/9780199689347.003.0005.
- [86] M. Ali, D.-H. Son, S.-H. Kang, and S.-R. Nam, "An Accurate CT Saturation Classification Using a Deep Learning Approach Based on Unsupervised Feature Extraction and Supervised Fine-Tuning Strategy," *Energies*, vol. 10, no. 11, Art. no. 11, Nov. 2017, doi: 10.3390/en10111830.
- [87] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," in *Advances in Information Retrieval*, Berlin, Heidelberg, 2005, pp. 345–359. doi: 10.1007/978-3-540-31865-1_25.

