

NUEVOS ENFOQUES METODOLÓGICOS PARA EL TRATAMIENTO DE DATOS IMPRECISOS. INTEGRACIÓN DE LÓGICA DIFUSA Y R-SHINY*

Matilde Inés Cesari**
Santiago Cristóbal Pérez**

Recibido: 9/11/2024 • Aceptado: 09/05/2025

<https://doi.org/10.22395/rium.v25n47a2>

RESUMEN

Este estudio propone una metodología basada en lógica difusa para el análisis de datos biotecnológicos con incertidumbre, aplicada a la caracterización y clasificación de cepas de hongos filamentosos. La metodología busca optimizar la representación de la variabilidad en datos experimentales, y permitir una mejor interpretación en comparación con enfoques estadísticos tradicionales.

Se presenta un caso de estudio que se centra en el crecimiento micelial de distintas cepas de hongos, considerando parámetros como velocidad de expansión, densidad estructural y dimensiones morfométricas. Para modelar esta información se utilizó la partición óptima de Fisher para definir los rangos de clasificación y posteriormente estimar los conjuntos difusos. Una vez establecidas las variables difusas, los datos fueron transformados en números borrosos, representados en tablas de contingencia para su análisis estructurado.

Para evaluar la relación entre las cepas y sus categorías de crecimiento se empleó la prueba exacta de Fisher, que permite determinar la significancia estadística de las asociaciones entre las variables difusas obtenidas. Esta estrategia mejora la caracterización del crecimiento micelial al capturar transiciones graduales y reducir la pérdida de información asociada a segmentaciones rígidas. Como aporte innovador, se desarrolló una aplicación web en R-Shiny que facilita el análisis interactivo de datos, sin requerir conocimientos avanzados en programación. Los

* Tesis en proceso de evaluación. Doctorado en Ingeniería mención en Computación. Universidad Tecnológica Nacional, Mendoza, Argentina.

** UTN Regional Mendoza - Sistemas de Información, GIDeCo. Mendoza, Argentina. Correo electrónico: matilde.cesari@frm.utn.edu.ar Orcid: <https://orcid.org/0000-0003-4786-7886>

*** UTN Regional Mendoza - Sistemas de Información, GIDeCo. Mendoza, Argentina. Correo electrónico: santiagoocp@frm.utn.edu.ar Orcid: <https://orcid.org/0000-0002-7241-3694>

resultados muestran que la modelización difusa mejora la interpretación y clasificación de cepas de hongos, y proporciona un marco metodológico replicable en biotecnología y otras áreas con incertidumbre en los datos.

Palabras clave: lógica difusa, hongos filamentosos, crecimiento micelial, biotecnología, análisis de datos, modelización matemática, inteligencia artificial, clasificación de datos, sistemas difusos, *software* científico.

NEW METHODOLOGICAL APPROACHES FOR THE TREATMENT OF IMPRECISE DATA: INTEGRATION OF FUZZY LOGIC AND R-SHINY

ABSTRACT

This study proposes a methodology based on fuzzy logic for the analysis of biotechnological data with uncertainty, applied to the characterization and classification of filamentous fungi strains. The methodology aims to optimize the representation of variability in experimental data and enable better interpretation compared to traditional statistical approaches.

A case study is presented focusing on the mycelial growth of different fungal strains, considering parameters such as expansion rate, structural density, and morphometric dimensions. Fisher's optimal partition was used to model this information, defining classification ranges and subsequently estimating the fuzzy sets. Once the fuzzy variables were established, the data were transformed into fuzzy numbers and represented in contingency tables for structured analysis.

To assess the relationship between the strains and their growth categories, Fisher's exact test was employed, allowing the determination of statistical significance of the associations among the fuzzy variables obtained. This strategy improves the characterization of mycelial growth by capturing gradual transitions and reducing information loss associated with rigid segmentation.

As an innovative contribution, an R-Shiny web application was developed to facilitate interactive data analysis without requiring advanced programming skills. The results show that fuzzy modeling enhances the interpretation and classification of fungal strains and provides a replicable methodological framework in biotechnology and other fields with data uncertainty.

Keywords: fuzzy logic, filamentous fungi, mycelial growth, biotechnology, data analysis, mathematical modeling, artificial intelligence, data classification, fuzzy systems, scientific software.

INTRODUCCIÓN

En un mundo impulsado por la información, la gestión adecuada de datos imprecisos es esencial para la toma de decisiones en diversos campos de la ciencia, la investigación y la tecnología. La lógica difusa se ha consolidado como una herramienta clave para abordar esta incertidumbre, pues permite asignar grados de pertenencia en lugar de categorizaciones binarias. Se le ha reconocido impacto relevante en áreas diversas, como la inteligencia artificial, el análisis de tráfico y el transporte de carga, entre otras [1].

En ese sentido, ha demostrado ser un instrumento efectivo en múltiples aplicaciones que necesitan la identificación de patrones en datos inciertos, con una mayor flexibilidad en la representación de la imprecisión [2;3]. En estudios previos –como el de Zurek, Romero, Blanco y Gutiérrez, [4],– se ha utilizado un sistema de inferencia difusa para la identificación de medicamentos de venta libre, y demostrado su eficacia en la clasificación de datos no estructurados. En [5] se adopta un enfoque similar aplicando lógica difusa para el análisis y clasificación de datos imprecisos en un contexto biotecnológico, con el objetivo de optimizar la representación de la incertidumbre y mejorar la toma de decisiones.

El desarrollo reciente de la lógica difusa también ha permitido avances significativos en su aplicación en sistemas inteligentes [6]. Además, el análisis de datos ha cobrado relevancia en otras áreas del conocimiento. Se ha aplicado en sectores como la ingeniería y la gestión de procesos industriales. Estudios han explorado el uso de modelos difusos para la toma de decisiones en estos entornos con incertidumbre [7]. En el ámbito educativo, la lógica difusa se ha utilizado para evaluar actitudes y desempeños estudiantiles de manera más precisa. Por ejemplo, Arroyo y Antolínez, [8], desarrollaron un modelo basado en lógica difusa para la valoración actitudinal en estudiantes universitarios, con lo que se logró una evaluación más objetiva y adaptada a las realidades complejas del entorno académico. En el sector salud, la lógica difusa se ha aplicado en el diagnóstico médico para manejar la incertidumbre inherente a los síntomas y signos clínicos. Por ejemplo, Ameen, Balachandran y Theodoridis, [9], propusieron un enfoque que combina reglas difusas y conocimiento experto para mejorar la clasificación de enfermedades hematológicas, y demostraron una mayor precisión en los diagnósticos al incorporar grados de certeza en la evaluación de parámetros clínicos. Hay también varios antecedentes de su uso en la agricultura. Martínez Rojas y Andrade Sosa, [10], integraron la lógica difusa con la dinámica de sistemas para seleccionar terrenos de cultivo agrícola, y así abordaron la complejidad y la incertidumbre inherentes a estos procesos.

Sin embargo, y a pesar de lo expresado, su adopción ha estado limitada por la necesidad de conocimientos técnicos avanzados. Este artículo propone una

metodología innovadora que combina la lógica difusa con el desarrollo de aplicaciones web interactivas en R-Shiny, que facilitan su uso por investigadores y profesionales sin experiencia en programación [11]. Esta integración permite realizar análisis de datos imprecisos con mayor exactitud y accesibilidad. Esto se ejemplifica con un caso de estudio sobre la modelización de procesos biotecnológicos, y permite manejar la incertidumbre y la complejidad asociadas al crecimiento micelial en diferentes cepas de hongos [12]. Usando la metodología y la herramienta web desarrollada, se evidencia una representación más realista de la variabilidad inherente a los datos biológicos y un mayor nivel de personalización en el análisis.

Los principales aportes de este trabajo son los siguientes: i) desarrollo e implementación de una metodología que integra lógica difusa con herramientas web interactivas de análisis de datos en R-Shiny, con lo que se logra una mayor accesibilidad y usabilidad en el tratamiento de datos imprecisos; ii) validación experimental de la metodología en un caso de estudio en el análisis del crecimiento de micelio en tres cepas de hongos; se alcanzaron resultados cuantitativos detallados que demuestran la utilidad del enfoque; iii) comparación y optimización de funciones de pertenencia para la modelización de datos inciertos, destacando cómo la selección de la función afecta la interpretación de los resultados; y iv) propuesta de un marco metodológico replicable que puede ser extendido a otros ámbitos científicos y tecnológicos en los que se requiera manejar incertidumbre en los datos.

El artículo se organiza de la siguiente manera: la sección 1 describe la metodología propuesta, y abarca la transformación de datos imprecisos y su modelización; la sección 2 detalla el caso de estudio, los experimentos realizados y los resultados obtenidos; la sección 3 analiza los hallazgos y sus implicaciones en distintos campos de aplicación; y finalmente, la sección 4 presenta las conclusiones más relevantes junto con posibles líneas de investigación futura.

1. METODOLOGÍA

El análisis se basa en la transformación de datos experimentales en variables difusas mediante un enfoque estructurado, que incluye: a) la codificación y estandarización de datos, b) la discretización de datos mediante la partición óptima de Fisher para definir clases ajustadas a la variabilidad de los datos y estimar los parámetros de los conjuntos difusos; y c) la transformación de datos a variables difusas. Este enfoque permite representar la incertidumbre sin imponer límites arbitrarios, y asegura una modelización precisa de los conjuntos difusos.

a) Codificación y estandarización de datos

En la primera etapa, los datos recolectados se organizan en una matriz en la cual las variables cuantitativas se mantienen en su forma numérica, mientras que las variables cualitativas se convierten en categorías discretas. Este proceso es esencial para estructurar los datos y prepararlos para su transformación en valores difusos. Opcionalmente, se aplican técnicas de normalización y escalado para garantizar comparabilidad entre variables con diferentes unidades de medición y variabilidad; así, se reducen sesgos en los cálculos posteriores.

Para este proceso, el paquete *caret* de R [13], proporciona herramientas avanzadas para la preparación de datos, incluyendo centrado y escalado de variables. Estas técnicas son útiles para evitar que variables con rangos muy diferentes influyan desproporcionadamente en la modelización difusa.

b) Discretización de datos mediante la partición óptima de Fisher

Para definir los límites de los conjuntos difusos se utilizó la partición óptima de Fisher, un método que segmenta las variables sin introducir ruido innecesario, ajustando la discretización a la variabilidad real de los datos en lugar de establecer rangos arbitrarios [14]. Este método divide los valores en clases óptimas basándose en la inercia interna de los datos; asegura que la segmentación refleja la estructura natural de las observaciones.

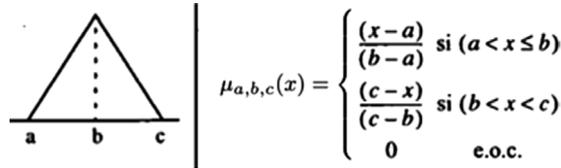
El uso de la partición de Fisher es clave en la construcción de conjuntos difusos porque permite definir intervalos que se ajustan dinámicamente a la distribución de los datos; de esta forma se evita la necesidad de establecer límites fijos que podrían no representar adecuadamente la variabilidad observada [5]. En este estudio, la implementación de este método se realizó mediante la función *classIntervals* del paquete *classInt* en R, lo que facilitó la generación automática de los intervalos óptimos [15].

c) Transformación de datos a variables difusas

La transformación de datos a variables difusas, conocida como borrosificación, convierte valores numéricos en números difusos mediante la asignación de funciones de pertenencia, lo que permite representar la incertidumbre y la variabilidad inherente en los datos. Para cada variable se definen conjuntos difusos a partir de descriptores lingüísticos que reflejan distintos niveles de la variable medida, con lo que se garantiza una mejor interpretación de los datos. Las funciones de pertenencia más utilizadas son la triangular, que permite representar transiciones más abruptas entre categorías (figura 1), y la gaussiana, que proporciona una transición más suave y facilita el modelado de datos con distribuciones continuas e incertidumbre estructural (figura 2) [16]. La selección de la función de pertenencia adecuada depende de la naturaleza de los datos y del

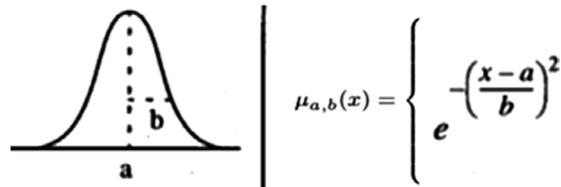
propósito del análisis, lo que influye en la precisión de la representación difusa [17]. El uso de estimaciones por intervalos en lógica difusa ha sido explorado en diversas aplicaciones, y proporciona una representación más robusta de los datos inciertos [18].

Figura 1. Representación gráfica y parámetros función de membresía triangular



Fuente: elaboración propia

Figura 2. Representación gráfica y parámetros función de membresía gaussiana



Fuente: elaboración propia

Una vez transformados, los datos borrosos pueden organizarse en estructuras tabulares para su análisis, con lo que se facilita su interpretación y manipulación mediante esquemas de representación específicos (figura 3). Para su implementación eficiente, el paquete FuzzyR de R permite generar funciones de pertenencia de manera automatizada mediante la función *genmf()*; su uso se optimiza en estudios de datos difusos y se asegura una clasificación más robusta y adaptable [19].

Figura 3. Esquema representación genérico de datos borrosos en tabla.

evaluator	Muestra	otra info...	Descriptor A				Descriptor ...				Descriptor ...					
			c0	c1	...	ci	c0	c1	...	ci	c0	c1	...	ci		
J1	M1															
...	M1															
Ji	M1															
...														
J1	Mx															
...	Mx															
Ji	Mx															

Fuente: elaboración propia

CASO DE ESTUDIO, EXPERIMENTACIÓN Y RESULTADOS

Este estudio se basa en un conjunto de datos experimentales sobre cepas de hongos, registrados en un archivo de Excel en el que se recopiló información sobre cada cepa, número de repetición, velocidad de crecimiento inicial y final (cm/h), índice de densidad, diámetro máximo de Feret, diámetro mínimo de Feret y perímetro. Por ejemplo, para la cepa E47 en su primera repetición, se observó una velocidad de crecimiento inicial de 0.386944 cm/h y final de 0.397083 cm/h, un índice de densidad de 8537.13, un diámetro máximo de Feret de 8.0 cm, un diámetro mínimo de Feret de 8.0 cm y un perímetro de 58.18 cm. Estos parámetros morfométricos permiten caracterizar el crecimiento y la estructura del micelio, lo que facilita la comparación entre cepas.

a) Particionamiento óptimo de los datos y estimación de parámetros difusos

Para analizar los datos mediante lógica difusa se realizó una partición óptima de los valores experimentales utilizando el método de Fisher, con el fin de definir los intervalos más adecuados para cada categoría de crecimiento. A partir de esta partición se establecieron tres conjuntos difusos correspondientes a las categorías de crecimiento bajo, moderado y alto; se aseguró así que cada intervalo reflejara la distribución real de los datos.

A diferencia de una segmentación rígida en rangos predefinidos, la partición óptima de Fisher ajusta la discretización de las variables según su variabilidad, con lo que se evita introducir ruido innecesario y se proporciona una base más precisa para la estimación de los conjuntos difusos. Esto permitió que los límites de las categorías emergentes reflejaran la estructura real de los datos; de esta manera se optimizó la clasificación de las cepas según sus características de crecimiento. Una vez definidos los conjuntos difusos se estimaron los parámetros de las funciones de pertenencia mediante técnicas estadísticas, asignando valores de posibilidad a cada observación. Se emplearon dos modelos principales: las funciones triangulares, que generan transiciones lineales entre categorías y ofrecen una interpretación más simple; y las funciones gaussianas, que permiten transiciones más progresivas y capturan mejor la variabilidad en los datos. Las particiones resultantes se presentan en tablas que resumen las principales variables analizadas, con lo cual se proporciona una visión estructurada del comportamiento de los datos bajo este enfoque.

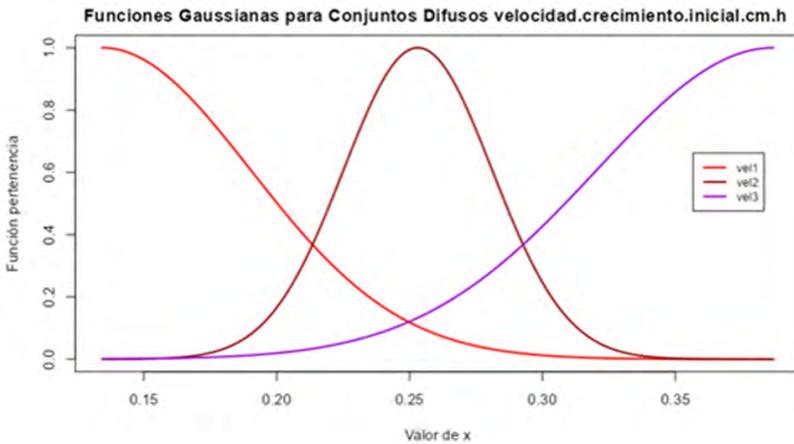
Los resultados obtenidos tras la partición de los datos mediante el método de Fisher se presentan en las tablas 1 a 6.

Tabla 1. Velocidad de crecimiento inicial (cm/h)

Conjunto	Mínimo	Máximo	a	b
vel1[0.1, 0.2]	0.13	0.21	0.13	0.08
vel2[0.2, 0.3]	0.21	0.29	0.25	0.04
vel3[0.4, 0.4]	0.29	0.39	0.39	0.09

Fuente: elaboración propia

Figura 4. Funciones de pertenencia gaussianas para los conjuntos difusos de la velocidad de crecimiento inicial (cm/h).



Fuente: elaboración propia

Tabla 2. Velocidad de crecimiento final (cm/h)

Conjunto	Mínimo	Máximo	a	b
vel1[0.1, 0.2]	0.11	0.20	0.11	0.09
vel2[0.2, 0.4]	0.20	0.38	0.29	0.09
vel3[0.4, 0.5]	0.38	0.51	0.45	0.06
vel4[0.7, 0.7]	0.51	0.66	0.66	0.15

Fuente: elaboración propia

Tabla 3. Índice de densidad

Conjunto	Mínimo	Máximo	a	b
Ind1[1850.7, 4250.6]	1850.69	4250.65	1850.69	2399.96
Ind2[4250.6, 6335.5]	4250.65	6335.49	5293.07	1042.42
Ind3[6335.5, 7992.5]	6335.49	7992.50	7163.99	828.50
Ind4[9032.5, 9032.5]	7992.50	9032.47	9032.47	1039.98

Fuente: elaboración propia

Tabla 4. Diámetro máximo Feret

Conjunto	Mínimo	Máximo	a	b
Max1[5, 6.2]	5.04	6.24	5.04	1.20
Max2[6.2, 7.3]	6.24	7.27	6.76	0.52
Max3[7.3, 7.7]	7.27	7.73	7.50	0.23
Max4[8, 8]	7.73	8.00	8.00	0.27

Fuente: elaboración propia

Tabla 5. Diámetro mínimo Feret

Conjunto	Mínimo	Máximo	a	b
Min1[4.8, 5.8]	4.78	5.77	4.78	0.99
Min2[5.8, 6.9]	5.77	6.91	6.34	0.57
Min3[6.9, 7.7]	6.91	7.73	7.32	0.41
Min4[8, 8]	7.73	8.00	8.00	0.27

Fuente: elaboración propia

Tabla 6. Perímetro

Conjunto	Mínimo	Máximo	a	b
Per1[28.5, 41.9]	28.50	41.90	28.50	13.40
Per2[41.9, 61.7]	41.90	61.71	51.81	9.91
Per3[61.7, 81.9]	61.71	81.90	71.81	10.10
Per4[91.4, 91.4]	81.90	91.43	91.43	9.53

Fuente: elaboración propia

Cada una de las tablas 1 a 6 muestra la distribución de los valores experimentales y su asignación a los conjuntos difusos, lo que permite una representación detallada de la variabilidad en los datos. En la tabla 1 se presentan los rangos de clasificación de la velocidad de crecimiento inicial (cm/h) obtenidos mediante la partición óptima de Fisher, la cual definió los límites de cada conjunto difuso sin imponer segmentaciones rígidas. La asignación de valores de posibilidad en un rango de 0 a 100 permitió representar la incertidumbre en la pertenencia de cada medición a una categoría específica.

Las demás tablas siguen el mismo enfoque de análisis aplicado a otras variables clave. La tabla 3, que representa el índice de densidad, permite evaluar la compactación del micelio en diferentes intervalos, y proporciona información sobre la estructura y la distribución de las colonias. Del mismo modo, los diámetros de Feret en las tablas 4 y 5 describen la variabilidad en la expansión del micelio a partir de sus dimensiones máximas y mínimas. Finalmente, la tabla 6, que presenta los valores de perímetro, complementa el análisis morfométrico al reflejar los distintos niveles de extensión del micelio en función de la segmentación difusa.

Los gráficos asociados a cada variable (figuras 4) muestran las funciones de pertenencia gaussianas utilizadas para definir los conjuntos difusos correspondientes a cada medición. Estas representaciones visuales permiten observar la transición progresiva entre categorías, y evidencian la flexibilidad del modelo difuso en la clasificación de los datos experimentales. En particular, la gráfica de la variable difusa (figura 4) destaca cómo la distribución de los valores influye en la pertenencia a cada conjunto difuso, y proporciona una interpretación más detallada de la incertidumbre en cada variable analizada.

b) Análisis de asociación y evaluación de crecimiento

Para evaluar la relación entre las cepas y las categorías de crecimiento se construyeron tablas de frecuencia a partir de los conjuntos difusos y se aplicó la prueba exacta de Fisher, que permite determinar la significancia estadística de la asociación. Además, se calcularon valores de test, en los cuales aquellos superiores a 2 indican una asociación fuerte entre la cepa y su categoría de crecimiento.

Los resultados obtenidos, presentados en la tabla 7, indican diferencias significativas entre las cepas analizadas en términos de velocidad de crecimiento, índice de densidad y morfología del micelio. La prueba exacta de Fisher permitió determinar asociaciones significativas entre las cepas y sus categorías de crecimiento, respaldadas por valores de test elevados.

Tabla 7. Resultados de discretización y valores de test para cada variable y cepa

Variable	Cepa	Conjunto Difuso	Rango	Valor de Test	Significancia
Velocidad crecimiento inicial	E47	vel3	[0.4, 0.4]	9.928	*** $p < 0.01$
	ECO	vel2	[0.2, 0.3]	6.497	*** $p < 0.01$
	FLO	vel1	[0.1, 0.2]	13.515	*** $p < 0.01$
Velocidad crecimiento final	E47	vel4	[0.7, 0.7]	9.151	*** $p < 0.01$
	ECO	vel2	[0.2, 0.4]	9.546	*** $p < 0.01$
	FLO	vel1	[0.1, 0.2]	5.043	*** $p < 0.01$
	FLO	vel3	[0.4, 0.5]	1.834	* $p < 0.10$
Índice de densidad	E47	Ind4	[9032.5, 9032.5]	12.081	*** $p < 0.01$
	E47	Ind3	[6335.5, 7992.5]	10.100	*** $p < 0.01$
	ECO	Ind2	[4250.6, 6335.5]	8.493	*** $p < 0.01$
	FLO	Ind1	[1850.7, 4250.6]	7.400	*** $p < 0.01$
	FLO	Ind2	[4250.6, 6335.5]	1.862	* $p < 0.10$

Variable	Cepa	Conjunto Difuso	Rango	Valor de Test	Significancia
Diámetro máximo de Feret	E47	Max4	[8, 8]	11.894	*** $p < 0.01$
	ECO	Max3	[7.3, 7.7]	5.425	*** $p < 0.01$
	ECO	Max2	[6.2, 7.3]	5.171	*** $p < 0.01$
	FLO	Max1	[5, 6.2]	4.732	*** $p < 0.01$
Diámetro mínimo de Feret	E47	Min4	[8, 8]	15.523	*** $p < 0.01$
	ECO	Min3	[6.9, 7.7]	4.528	*** $p < 0.01$
	ECO	Min2	[5.8, 6.9]	3.484	** $p < 0.05$
	FLO	Min1	[4.8, 5.8]	5.981	*** $p < 0.01$
	FLO	Min2	[5.8, 6.9]	4.570	*** $p < 0.01$
Perímetro	E47	Per4	[91.4, 91.4]	4.659	** $p < 0.05$
	ECO	Per2	[41.9, 61.7]	3.956	** $p < 0.05$

Fuente: elaboración propia

La cepa FLO presentó un crecimiento inicial significativamente mayor que las demás; se situó en la clase vel1 [0.1, 0.2] con un valor de test de 13.515 ($p < 0.01$). En contraste, la cepa E47 mostró la mayor velocidad de crecimiento final y se ubicó en la clase vel4 [0.7, 0.7] con un valor de test de 9.151 ($p < 0.01$). En términos de densidad del micelio, la cepa E47 alcanzó los valores más altos (Ind4 [9032.5, 9032.5], test = 12.081, $p < 0.01$), mientras que FLO presentó la menor densidad (Ind1 [1850.7, 4250.6], test = 7.4, $p < 0.01$).

El análisis de los diámetros de Feret reveló que E47 mantiene una estructura morfológica homogénea con poca variabilidad, y presentó los valores de test más altos en diámetro máximo y mínimo (Max4 [8, 8], test = 11.894, $p < 0.01$; Min4 [8, 8], test = 15.523, $p < 0.01$). En cuanto al perímetro, E47 mostró un micelio con mayor extensión (Per4 [91.4, 91.4], test = 4.659, $p < 0.05$), mientras que ECO presentó dimensiones más reducidas (Per2 [41.9, 61.7], test = 3.956, $p < 0.05$).

Estos resultados evidencian diferencias estructurales y de crecimiento entre las cepas, lo que sugiere posibles aplicaciones en la optimización del crecimiento micelial para diversas industrias.

c) Aplicación del modelo en R y visualización interactiva

Para facilitar la interpretación de los resultados, los datos transformados en valores difusos se integraron en una aplicación interactiva desarrollada en Shiny, una plataforma de R que permite la visualización dinámica de datos. Shiny proporciona una interfaz intuitiva para explorar cómo las diferentes funciones de pertenencia afectan la clasificación de los datos, y permite ajustar parámetros y evaluar distintos enfoques en tiempo real [4; 21].

d) Comparación de la metodología de la lógica difusa con la metodología tradicional

El análisis del crecimiento micelial mediante lógica difusa permitió una clasificación más detallada de las cepas en función de sus patrones de crecimiento, debido al empleo de la discretización óptima de Fisher y a la asignación de valores de pertenencia a conjuntos difusos. Esta metodología capturó transiciones graduales entre categorías y permitió evaluar la incertidumbre inherente a los datos experimentales.

Se confirmaron diferencias significativas entre cepas en múltiples variables, entre el análisis tradicional basado en el ANOVA y la prueba de Fisher. Los resultados del ANOVA indicaron diferencias altamente significativas en la velocidad de crecimiento inicial ($p < 0.01$) con la cepa E47, y mostraron los valores más altos en comparación con ECO y FLO. Para la velocidad de crecimiento final se observaron diferencias igualmente significativas ($p < 0.01$): E47 alcanzó la mayor velocidad final, mientras que FLO presentó los valores más bajos. En cuanto al índice de densidad, se identificó una diferencia altamente significativa ($p < 0.001$): E47 presentó la mayor compactación del micelio, seguido de ECO y FLO, en concordancia con la clasificación obtenida mediante lógica difusa.

El análisis de correlación lineal permitió evaluar relaciones entre variables, y mostró que la velocidad de crecimiento inicial y el índice de densidad presentan una correlación significativa de 0.850 ($p < 0.001$), lo que indica que una mayor velocidad inicial tiende a estar asociada con una mayor densidad del micelio. Sin embargo, al analizar la velocidad final y el diámetro de Feret, la correlación fue 0.577 ($p < 0.01$), lo que sugiere una relación moderada, pero con mayor variabilidad. La correlación entre diámetro máximo y mínimo de Feret fue 0.835 ($p < 0.001$), lo que refleja la uniformidad del crecimiento en ciertas cepas.

Si bien el ANOVA y las correlaciones permiten identificar diferencias y relaciones entre variables, la lógica difusa, en combinación con la partición óptima de Fisher, ofrece una ventaja en la interpretación de los datos, ya que no depende de la suposición de distribuciones normales o de rangos arbitrarios para segmentar las mediciones. En este sentido, la discretización basada en Fisher permite una estimación más precisa de los conjuntos difusos, pues refleja transiciones graduales entre categorías y captura mejor la incertidumbre en los datos experimentales.

3. DISCUSIÓN DE LOS RESULTADOS

Los resultados obtenidos en este estudio son consistentes con investigaciones recientes sobre la aplicación de lógica difusa en biotecnología y análisis de datos con

incertidumbre. Macnamara, [20], aplicó este enfoque para modelar el crecimiento de tumores, y demostró que la lógica difusa permite representar con mayor precisión la variabilidad en datos biomédicos y mejorar la interpretación de patrones complejos. De manera similar, Houssein, Mahmoud, Hassanien y Emam, [21], destacaron la utilidad de estos modelos en el análisis de datos clínicos y reforzaron su aplicabilidad en diferentes áreas biotecnológicas. En este sentido, Patel, [22], demostró la utilidad de este enfoque para analizar desequilibrios hormonales: resaltó que la lógica difusa mejora la clasificación de datos biomédicos al capturar la incertidumbre inherente en los procesos fisiológicos.

Desde una perspectiva más amplia, Talpur, Das, Kumar y Singh, [23], han señalado que los sistemas neurodifusos pueden mejorar la interpretación de datos complejos en múltiples dominios, lo que respalda la elección de este enfoque para analizar el crecimiento de micelio en este estudio. En comparación con métodos tradicionales como el ANOVA y la correlación lineal, la lógica difusa no solo confirmó las diferencias entre cepas, sino que también permitió modelar la variabilidad con mayor flexibilidad. La aplicación de este enfoque es particularmente relevante en estudios en los cuales los datos presentan variaciones graduales en lugar de cambios abruptos entre categorías discretas.

En el ámbito industrial, Al-Ghamdi, [24], mostró que los modelos de lógica difusa mejoran la eficiencia en procesos de manufactura, mientras que IEEE [25] los aplicó en rehabilitación con exoesqueletos y logró una mejor evaluación psicofisiológica de los pacientes. Arango Serna, Serna Urán y Gómez Builes, [26], desarrollaron un modelo de inferencia difusa para la toma de decisiones en condiciones de incertidumbre, aplicado al diseño de productos bajo un esquema de ingeniería concurrente.

Estas aplicaciones demuestran que la lógica difusa es una herramienta versátil que puede ser utilizada en distintos escenarios, desde la optimización de procesos industriales hasta el análisis de datos biomédicos y ecológicos. Finalmente, la revisión de Jabiyeva, [27], sobre el uso de lógica difusa en biomedicina confirma que estas metodologías son clave en la optimización del procesamiento de datos con incertidumbre, y consolida su importancia en la investigación y análisis de sistemas complejos.

Si bien los resultados obtenidos confirman la utilidad de la lógica difusa en la clasificación de datos con incertidumbre, es importante destacar algunas limitaciones de la metodología utilizada. En primer lugar, la precisión de la modelización depende en gran medida de la selección de la función de pertenencia adecuada, lo que puede afectar la representación de los conjuntos difusos. Además, la estimación de los conjuntos difusos podría beneficiarse de técnicas de optimización adicionales que permitan una segmentación más precisa de los datos. Futuras investigaciones podrían explorar

enfoques híbridos que integren aprendizaje automático con lógica difusa para mejorar la robustez del modelo.

En el presente estudio, la aplicación de la lógica difusa permitió una clasificación más detallada del crecimiento micelial, al optimizar la segmentación de los datos mediante la partición óptima de Fisher. Esta estrategia evitó la introducción de ruido debido a la imposición de rangos predefinidos, como ocurre en algunos enfoques tradicionales. La capacidad de la lógica difusa para representar transiciones graduales entre categorías permitió una diferenciación más precisa de las cepas, lo que se tradujo en una mejor evaluación de sus patrones de crecimiento.

4. CONCLUSIONES

Este estudio demuestra la utilidad de la lógica difusa en el análisis de datos imprecisos, pues proporciona una caracterización más flexible del crecimiento micelial en comparación con métodos tradicionales. La aplicación de la partición óptima de Fisher permitió definir conjuntos difusos adaptados a la variabilidad de los datos sin imponer rangos arbitrarios, y optimizó la representación de incertidumbre. Mientras que el ANOVA y la correlación lineal confirmaron diferencias significativas entre cepas, la lógica difusa ofreció una interpretación más detallada de la dinámica del crecimiento.

Los resultados sugieren que este enfoque puede aplicarse en otros ámbitos en los que los datos presentan transiciones graduales en lugar de cortes discretos. Además, la integración con R-Shiny, resumida aquí por brevedad, facilita su implementación en entornos científicos y tecnológicos. Como futuras líneas de investigación, se recomienda explorar la combinación de lógica difusa con técnicas de aprendizaje automático para mejorar la precisión en la clasificación de datos imprecisos.

REFERENCIAS

- [1] R. Krishnapuram, *Fuzzy Data Analysis — Methods and Industrial Applications*. Elsevier, 2024. doi: 10.1016/0165-0114(94)90280-1.
- [2] C. Kahraman, B. Oztaysi, I. Otay, y S. C. Onar, “Extensions of ordinary fuzzy sets: a comparative literature review,” en *Intelligent and Fuzzy Techniques: Smart and Innovative Solutions: Proceedings of the INFUS 2020 Conference*, Springer, 2020, pp. 1655–1665. doi: 10.1007/978-3-030-51156-2_193.
- [3] S. M. Baas, *Feasible Algorithms for Approximate Reasoning Using Fuzzy Logic*. Elsevier, 2024. doi: 10.1016/0165-0114(80)90022-6.
- [4] E. E. Zurek, D. J. Romero, H. D. Blanco, y C. A. Gutiérrez, “Fuzzy inference system for the identification of over-the-counter (OTC) drugs,” *Rev. Ing. Univ. Medellín*, vol. 12, no. 22, pp. 109–116, 2014. doi: 10.22395/rium.v12n22a9.

- [5] M. I. Césari, “Soluciones metodológicas para el análisis de datos imprecisos: Lógica difusa y R-Shiny,” *Actas Jorn. Eventos Académicos UTN*, 2024. [En línea]. Disponible en: <https://rtyc.utn.edu.ar/index.php/ajea/article/download/1738/1565/8194>.
- [6] C. Swathi, J. J. Ebienezar, M. Swathi, y S. Suruthipriya, “Fuzzy Logic,” *Int. J. Innov. Res. Inf. Secur.*, vol. 9, no. 3, pp. 147–152, 2023. doi: 10.26562/ijiris.2023.v0903.19.
- [7] M. D. Arango Serna, C. A. Serna Urán y A. C. Gómez Builes, “Inferencia difusa aplicada a la ingeniería concurrente para el diseño de productos de manufactura en condiciones de incertidumbre,” *Rev. Ing. Univ. Medellín*, vol. 11, no. 21, pp. 127–138, 2012. [En línea]. Disponible en: https://www.scielo.org.co/scielo.php?pid=S1692-33242012000200011&script=sci_arttext.
- [8] M. Arroyo y J. Antolínez, “Aplicación de lógica difusa para la valoración actitudinal en estudiantes universitarios,” *Rev. Electr. Investig. Educ.*, vol. 17, no. 1, pp. 120–136, 2015. [En línea]. Disponible en: <https://www.redalyc.org/pdf/4677/467746222001.pdf>.
- [9] S. Ameen, R. Balachandran, y T. Theodoridis, “Deriving Hematological Disease Classes Using Fuzzy Logic and Expert Knowledge,” *arXiv preprint*, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2406.13015>.
- [10] N. Martínez Rojas y H. H. Andrade Sosa, “Integración de la lógica difusa a la dinámica de sistemas para la selección de terrenos de cultivos agrícolas,” *Rev. Investig. Des. Innov.*, vol. 7, no. 1, pp. 45–56, 2016. [En línea]. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/5676771.pdf>.
- [11] R. Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2023. [En línea]. Disponible en: <https://www.R-project.org/>.
- [12] D. Dubois, “On statistical inference for fuzzy data with applications to descriptive statistics,” *Fuzzy Sets Syst.*, 2024. doi: 10.1016/0165-0114(92)90213-N.
- [13] M. Kuhn, “caret: Classification and Regression Training,” 2021. [En línea]. Disponible en: <https://cran.r-project.org/web/packages/caret/index.html>.
- [14] W. D. Fisher, “On grouping for maximum homogeneity,” *J. Am. Stat. Assoc.*, vol. 53, no. 284, pp. 789–798, 1958. doi: 10.1080/01621459.1958.10501479.
- [15] R. Bivand et al., “classInt: Choose Univariate Class Intervals (Version 0.4-11) [Paquete de R],” CRAN, 2025. [En línea]. Disponible en: <https://cran.r-project.org/web/packages/classInt/classInt.pdf>.
- [16] L. A. Zadeh, “Calibration with imprecise signals and concentrations based on fuzzy theory,” *Chemom. Intell. Lab. Syst.*, 2024. doi: 10.1016/0169-7439(86)80027-2.
- [17] K. Al-Ghamdi, “Fuzzy data-driven scenario-based robust data envelopment analysis for prediction and optimization,” *Expert Syst. Appl.*, vol. 116, p. 419, 2022. doi: 10.1016/j.eswa.2021.116419.

- [18] M. A. Gil, “A note on interval estimation with fuzzy data,” *Fuzzy Sets Syst.*, 2024. doi: 10.1016/0165-0114(88)90201-1.
- [19] D. Zhang, y S. Chen, “FuzzyR: Fuzzy logic toolkit in R (Version 2.2-3) [Manual de software],” CRAN, 2023. Disponible en: <https://cran.r-project.org/web/packages/FuzzyR/FuzzyR.pdf>.
- [20] C. K. Macnamara, “A fuzzy mathematical model for tumor growth pattern using generalized Hukuhara derivative,” *Appl. Soft Comput.*, vol. 122, 2022. doi: 10.1016/j.asoc.2022.108467.
- [21] E. H. Houssein, A. K. Mahmoud, A. E. Hassanien, y M. K. Emam, “Soft computing techniques for biomedical data analysis,” *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 2599–2649, 2023. doi: 10.1007/s10462-023-10585-2.
- [22] S. Patel, “Fuzzy machine learning logic utilization on hormonal imbalance dataset,” *Comput. Biol. Med.*, vol. 151, 2024. doi: 10.1016/j.compbiomed.2024.108429.
- [23] N. Talpur, S. K. Das, S. Kumar, y J. Singh, “Deep Neuro-Fuzzy System application trends and challenges,” *Artif. Intell. Rev.*, vol. 56, pp. 865–913, 2023. doi: 10.1007/s10462-022-10188-3.
- [24] K. Al-Ghamdi, “Fuzzy data-driven scenario-based robust data envelopment analysis for prediction and optimization of an electrical discharge machine’s parameters,” *Expert Syst. Appl.*, vol. 116, 2022. doi: 10.1016/j.eswa.2021.116419.
- [25] IEEE, “A Data-Driven Fuzzy Logic Method for Psychophysiological Assessment,” *IEEE Trans. Med. Robot. Bionics*, 2025. doi: 10.1109/TMRB.2024.3377453.
- [26] M. D. Arango Serna, C. A. Serna Urán y A. C. Gómez Builes, “Inferencia difusa aplicada a la ingeniería concurrente para el diseño de productos de manufactura en condiciones de incertidumbre,” *Rev. Ing. Univ. Medellín*, vol. 11, no. 21, pp. 127–138, 2012. [En línea]. Disponible en: https://www.scielo.org/co/scielo.php?pid=S1692-33242012000200011&script=sci_arttext.
- [27] A. Jabiyeva y M. Khudaverdiyeva, “Application of Fuzzy Logic in Computer Systems of Medical Diagnosis,” *Socio World Social Research & Behavioral Sciences*, vol. 12, no. 02, pp. 17–25, dic. 2023, doi: 10.36962/SWD12022023-17. Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4667205