



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

**INSTITUTO TECNOLÓGICO DE CIUDAD MADERO**  
**DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN**  
**DOCTORADO EN CIENCIAS DE LA INGENIERÍA**



**TESIS**

**MÉTODOS HÍBRIDOS PARA PRONÓSTICO DE SERIES FINANCIERAS BASADAS  
EN SUAVIZAMIENTO EXPONENCIAL**

Que para obtener el Grado de  
**Doctora en Ciencias de la Ingeniería**

Presenta  
**MC. Mirna Patricia Ponce Flores**  
**D94070293**

Director de Tesis  
**Dr. Juan Frausto Solís**

Co-director de Tesis  
**Dra. Guadalupe Castilla Valdez**

Cd. Madero, Tamaulipas

Junio 2021



Instituto Tecnológico de Ciudad Madero  
Subdirección Académica  
División de Estudios de Posgrado e Investigación

Cd. Madero, Tam. **27 de mayo de 2021**

OFICIO No. : U.034/21  
ASUNTO: AUTORIZACIÓN DE  
IMPRESIÓN DE TESIS

**C. MIRNA PATRICIA PONCE FLORES**  
**No. DE CONTROL D94070293**  
**P R E S E N T E**

Me es grato comunicarle que después de la revisión realizada por el Jurado designado para su Examen de Grado de Doctorado en Ciencias de la Ingeniería, se acordó autorizar la impresión de su tesis titulada:

**“MÉTODOS HÍBRIDOS PARA PRONÓSTICO DE SERIES FINANCIERAS BASADAS EN SUAVIZAMIENTO EXPONENCIAL”**

El Jurado está integrado por los siguientes catedráticos:

PRESIDENTE:	DR. JUAN FRAUSTO SOLÍS
SECRETARIO:	DR. PEDRO MARTÍN GARCÍA VITE
VOCAL 1:	DRA. GUADALUPE CASTILLA VALDEZ
VOCAL 2:	DR. JUAN JAVIER GONZALÈZ BARBOSA
VOCAL 3:	DRA. SIVIA BEATRIZ BRACHETTI SIBAJA
DIRECTOR DE TESIS:	DR. JUAN FRAUSTO SOLÍS
CO-DIRECTOR DE TESIS:	DRA. GUADALUPE CASTILLA VALDEZ

Es muy satisfactorio para la División de Estudios de Posgrado e Investigación compartir con usted el logro de esta meta. Espero que continúe con éxito su desarrollo profesional y dedique su experiencia e inteligencia en beneficio de México.

**ATENTAMENTE**

*Excelencia en Educación Tecnológica®*

*"Por mi patria y por mi bien"®*

**MARCO ANTONIO CORONEL GARCÍA**  
**JEFE DE LA DIVISIÓN DE ESTUDIOS DE**  
**POSGRADO E INVESTIGACIÓN**



c.c.p.- Archivo  
MACG 'mdcoa'



# Agradecimientos

A mis directores de proyecto, Dr. Juan Frausto Solís y Dra. Guadalupe Castilla por todo el apoyo brindado para la realización de esta investigación y su invaluable transferencia de conocimiento a lo largo de los estudios de doctorado.

Al Consejo Nacional de Ciencia y Tecnología, por el apoyo brindado a través de la beca de manutención 737188, con la que fue posible solventar los gastos personales y los estudios de posgrado realizados .

Al Tecnológico Nacional de México/Instituto Tecnológico de Ciudad Madero, por las facilidades otorgadas, así como la gestión en cada una de las etapas requeridas para la realización de este proyecto.

A Dios, por acomodar las piezas del rompecabezas en el momento indicado, permitiendo un giro completo en mis paradigmas de vida y abriéndome nuevas puertas hacia lo que tanto me gusta realizar.

A mi familia:

A mi esposo, por estar a mi lado en todo momento, brindándome no solo apoyo emocional sino también académico, sin dudarle, este logro es de ambos.

A mis hijas, ustedes han sido el aliento que necesito día a día para dejar el cansancio de lado y seguir esforzándome para lograr los objetivos, mil gracias mis pequeñas, este camino sin ustedes, no hubiese tenido el mérito que hoy tiene el haberlo recorrido juntas, esas sonrisas,

## *Agradecimientos*

llantos y gritos forman parte de los recuerdos que atesoro.

A mi mamá, quién sin dudarle me externo su apoyo y ayuda incondicional, además de alentarme a seguir adelante. Mamá, mis hijas no pudieron estar en mejores manos que las tuyas.

A mi papá, hermanos, hermana y sobrina, todos aportaron su granito de arena para motivarme de una u otra manera para seguir adelante.

A mis compañeros de posgrado, ya que juntos nos apoyamos y alentamos en cada una de las entregas, nos brindamos ánimo cuando fue necesario.

A los doctores que forman parte del núcleo académico del posgrado, ya que cada uno me brindo conocimientos que me permitieron crecer como profesionista, además de su apoyo en múltiples trámites permitiendo la obtención de este grado.

Y a todos los que directa e indirectamente me impulsaron a emprender esta etapa, gracias.



# Índice general

Índice de tablas . . . . .	VI
Índice de figuras . . . . .	VIII
<b>1 Introducción</b>	<b>1</b>
1.1 Planteamiento del problema . . . . .	4
1.2 Objetivos . . . . .	5
1.2.1 Objetivo general . . . . .	5
1.2.2 Objetivos específicos . . . . .	5
1.3 Justificación del estudio . . . . .	6
1.4 Alcances . . . . .	6
1.5 Limitaciones . . . . .	7
1.6 Organización de la tesis . . . . .	7
<b>2 Marco teórico</b>	<b>9</b>
2.1 Serie de tiempo . . . . .	10
2.2 Autocorrelación . . . . .	11
2.2.1 Correlograma . . . . .	12
2.2.2 Patrones en los datos . . . . .	13
2.3 Pronóstico de series de tiempo . . . . .	14
2.4 Taxonomía de los métodos de pronósticos . . . . .	14
2.5 Notación en pronósticos y medidas de error . . . . .	14
2.5.1 Desviación media absoluta . . . . .	15
2.5.2 Error cuadrático medio . . . . .	16

2.5.3	Raíz cuadrada del error cuadrático medio . . . . .	16
2.5.4	Error porcentual absoluto medio . . . . .	17
2.5.5	Error porcentual absoluto medio simétrico . . . . .	17
2.5.6	Error porcentual medio . . . . .	17
2.6	Ventajas y desventajas de los métodos de pronósticos . . . . .	18
2.7	Selección de técnicas para pronosticar . . . . .	18
2.8	Método de suavizamiento exponencial . . . . .	19
2.9	Metodología Box-Jenkins . . . . .	22
2.10	Métodos híbridos basados en suavizamiento exponencial . . . . .	23
<b>3</b>	<b>Estado del arte</b>	<b>25</b>
<b>4</b>	<b>Metodología</b>	<b>32</b>
4.1	Análisis de la relación Hurst-MAPE y sMAPE . . . . .	33
4.1.1	Exponente de Hurst . . . . .	35
4.1.2	Algoritmo k-means . . . . .	37
4.2	Complejidad y desempeño en el pronóstico de series de tiempo . . . . .	38
4.3	Desarrollo de la metodología MetaLSS . . . . .	40
4.3.1	Fase de entrenamiento . . . . .	41
4.3.2	Fase de prueba . . . . .	44
<b>5</b>	<b>Equipos y conjuntos de datos</b>	<b>58</b>
5.1	Equipo de cómputo . . . . .	59
5.2	Software utilizado . . . . .	59
5.3	Conjuntos de datos . . . . .	59
<b>6</b>	<b>Análisis y resultados</b>	<b>62</b>
6.1	Pruebas preliminares . . . . .	62
6.1.1	Pronósticos con métodos clásicos . . . . .	62
6.1.2	Identificación de características de las series de tiempo . . . . .	63
6.1.3	Relación entre el exponente de Hurst vs. MAPE y sMAPE . . . . .	67

6.1.4	Resultados de complejidad en series de tiempo . . . . .	69
6.2	Metodología MetaLSS . . . . .	71
6.2.1	Resultados con el subconjunto de datos diario . . . . .	72
6.2.2	Resultados con el subconjunto de datos mensual . . . . .	75
<b>7</b>	<b>Conclusiones y trabajos futuros</b>	<b>79</b>
7.1	Conclusiones . . . . .	79
7.2	Trabajos futuros . . . . .	80
7.3	Publicaciones derivadas del presente trabajo . . . . .	81
	<b>Glosario</b>	<b>82</b>
	<b>Bibliografía</b>	<b>83</b>

# Índice de tablas

Tabla 2.1	Serie de Tiempo del tipo de cambio del dólar frente al peso. . . . .	10
Tabla 2.2	Ejemplo de una serie de tiempo con 1, 2 y 3 retrasos en el tiempo. . . .	12
Tabla 2.3	Autocorrelación del tipo de cambio del dólar-peso en enero 2021. . . . .	12
Tabla 2.4	Ventajas y desventajas de algunos métodos de pronósticos. . . . .	19
Tabla 2.5	Selección de métodos de pronósticos de acuerdo con los patrones de los datos. . . . .	20
Tabla 2.6	Peso asignado de acuerdo con el número de observación. . . . .	20
Tabla 3.1	Estado del arte de aprendizaje y metaaprendizaje. . . . .	29
Tabla 4.1	Interpretación de los valores del exponente de Hurst. . . . .	35
Tabla 4.2	Métodos de pronósticos seleccionados de las competencias M4. . . . .	43
Tabla 4.3	Casos potenciales de los clasificadores. . . . .	46
Tabla 4.4	Métodos híbridos desarrollados. . . . .	49
Tabla 5.1	Series de tiempo de la competencia M3. . . . .	60
Tabla 5.2	Series de tiempo de la competencia M4. . . . .	60
Tabla 6.1	Muestra de errores sMAPE del subconjunto de datos diario. . . . .	73
Tabla 6.2	Prueba de Friedman al subconjunto de datos diario. . . . .	74
Tabla 6.3	Prueba W de Wilcoxon con rangos al subconjunto de datos diario. . . .	75
Tabla 6.4	Prueba estadística W de Wilcoxon al subconjunto de datos diario. . . .	75
Tabla 6.5	Muestra de errores sMAPE del subconjunto de datos mensual. . . . .	76
Tabla 6.6	Prueba de Friedman al subconjunto de datos mensual. . . . .	77
Tabla 6.7	Prueba W de Wilcoxon con rangos al subconjunto de datos mensual. . .	78

Tabla 6.8 Prueba estadística de Wilcoxon al subconjunto de datos mensual. . . . 78

# Índice de figuras

Figura 2.1	Serie de Tiempo del tipo de cambio del dólar frente al peso. . . . .	11
Figura 2.2	Autocorrelación del tipo de cambio del dólar-peso en enero 2021. . . . .	13
Figura 2.3	Taxonomía general de los métodos de pronósticos [1]. . . . .	15
Figura 2.4	Métodos híbridos. . . . .	16
Figura 2.5	Diagrama de flujo del modelo Box-Jenkins [2]. . . . .	23
Figura 4.1	El proceso del pronóstico [3]. . . . .	33
Figura 4.2	Primera etapa del análisis realizado [4]. . . . .	34
Figura 4.3	Segunda y tercera etapa del análisis realizado [4]. . . . .	35
Figura 4.4	Paso uno para el cálculo del exponente de Hurst. . . . .	36
Figura 4.5	Cálculo del rango reescalado de cada partición. . . . .	36
Figura 4.6	Posibles caracterizaciones de un sistema dinámico [5]. . . . .	39
Figura 4.7	Metodología del análisis de la relación entre el pronóstico y la complejidad de las series de tiempo [5]. . . . .	40
Figura 4.8	Esquema general, segmentación y fases de la metodología MetaLSS. . . . .	42
Figura 4.9	Selección de subconjuntos de datos de M4-Competition. . . . .	43
Figura 4.10	Fase de entrenamiento de MetaLSS. . . . .	45
Figura 4.11	Fase de prueba de MetaLSS. . . . .	47
Figura 4.12	Parámetros de entrada de MetaLSS. . . . .	48
Figura 4.13	Método híbrido: MetaComb. . . . .	50
Figura 4.14	Método híbrido: MetaComb-S. . . . .	51
Figura 4.15	Método híbrido: MetaComb-SwA. . . . .	52
Figura 4.16	Método híbrido: MetaSmoothing. . . . .	54

Figura 4.17 Método híbrido: MetaSmoothing-S. . . . .	55
Figura 4.18 Método híbrido: MetaSmoothing-SwA. . . . .	56
Figura 6.1 Ejemplo de la gráfica de una serie de tiempo. . . . .	63
Figura 6.2 Relación Hurst vs MAPE y sMAPE usando R/S. . . . .	68
Figura 6.3 Medidas de complejidad basadas en entropía y análisis de componentes principales. . . . .	69
Figura 6.4 Series sintéticas y medidas de entropía (ESC). . . . .	70
Figura 6.5 Análisis de las series de tiempo con medidas de entropía y los pronósti- cos de ARIMA, Smyl, Theta y ETS en M4-Competition. . . . .	70
Figura 6.6 Medida de error-complejidad en el primer cuartil. . . . .	71
Figura 6.7 Medida de error-complejidad en el cuarto cuartil. . . . .	72
Figura 6.8 Histograma de los métodos de pronósticos generados con MetaLSS en el subconjunto de datos diario. . . . .	73
Figura 6.9 Histograma de los métodos de pronósticos generados con MetaLSS en el subconjunto de datos mensual. . . . .	76

---

# Introducción

Los pronósticos, de manera general, consisten en realizar estimaciones de eventos futuros a partir de la ocurrencia de eventos en el pasado. Desde hace mucho tiempo se han realizado pronósticos en diferentes áreas y cada vez se han vuelto más cotidianos. La mayoría de las personas realizamos estas estimaciones de manera natural. Existen áreas que se especializan en realizar pronósticos para la toma de decisiones. Particularmente, en el área financiera, se estudian y pronostican series de tiempo de activos financieros que son cotizados en mercados estandarizados o bolsas de valores internacionales.

Existen dos problemas que son de gran interés por su aplicabilidad:

- **Pronósticos de activos:** se trata de determinar los valores de activos para uno o más periodos en el futuro a partir de valores que han tenido en el pasado esos mismos activos.
- **Pronósticos de la volatilidad:** se trata de determinar la tendencia o variabilidad existente para uno o más periodos(instantes en el tiempo) en el futuro, a partir de los precios que han tenido en el pasado esos mismos activos.



Los métodos más comúnmente utilizados para el pronóstico de series financieras son:

1. Métodos de suavizamiento exponencial: acorde con [6, 7], el método de suavizamiento exponencial simple, tuvo sus orígenes en el año de 1944, por el investigador Robert G. Brown, pero fue hasta el año de 1959, cuando se presentó de manera formal su primer libro [8].
2. Método doblemente exponencial: como se puede leer en [6, 7], el trabajo original de Charles C. Holt, fue documentado en el año 1957, sin embargo la publicación oficial inédita se realizó en [9].
3. Métodos ARMA y ARIMA o métodos de Box & Jenkins [2].
4. Métodos GARCH para volatilidad [10–12].
5. Métodos de redes neuronales (ANN, artificial Neural Networks) [13–15].
6. Métodos de máquinas de soporte vectorial [16, 17].
7. Métodos heurísticos [18].

Este trabajo de investigación, aborda específicamente el problema de pronósticos de activos. Los métodos en general, tienen ventajas y desventajas, pero lo más importante a destacar es que para las series financieras ninguno de ellos puede garantizar un pronóstico aceptable, incluso con un error permisible moderado. Por esta razón, una alternativa es utilizar métodos híbridos a partir de algunos de ellos así como alguna estrategia de hibridación. Cabe destacar que los métodos de suavizamiento exponencial aun cuando se originaron en los años 1950's con el trabajo de Brown [8] han sido recientemente utilizados como base de hibridación [19–23].

Algunas de las publicaciones de los métodos híbridos se basan en metaaprendizaje, los cuales:

1. A partir de las características de las series de tiempo obtienen patrones para determinar:

- El o los métodos más adecuados para realizar pronósticos en función del patrón de la serie.
  - Cuál combinación de métodos es el método híbrido más adecuado para la serie de tiempo cuyas características de las series resulte el más adecuado.
2. Realizan una puntualización o scoring (por su nombre en inglés) con la cual cada patrón de series de tiempo es clasificado, por lo regular como simple, regular, difícil y muy difícil [24].
- Este tipo de métodos determina para una serie de tiempo la combinación de métodos más adecuada, no ha sido aplicado en el caso de series financieras. Por otro lado, la mejor forma de hibridar estos métodos es un tema abierto de investigación.
3. El uso de funciones de regresión para pronóstico de series de tiempo utilizando plantillas de funciones de regresión es un tema que se ha abordado de manera incipiente en el área de pronóstico. En esta área, [10] propone para el pronóstico de la volatilidad, una función general de la forma:

$$\varepsilon_t \mid \Psi_{t-1} \sim N(0, h_t) \quad (1.1)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (1.2)$$

$$= \alpha_0 + A(L)\varepsilon_t^2 + B(L)h_t \quad (1.3)$$

donde:

$\varepsilon_t$ : denota un valor real del tiempo discreto en el proceso estocástico.

$\Psi_t$ : el conjunto de información de toda la información a través del tiempo  $t$ .

$$\begin{aligned} p &\geq 0, & q &> 0 \\ \alpha_0 &> 0, & \alpha_i &\geq 0, & i &= 1, \dots, q, \\ \beta_i &\geq 0, & i &= 1, \dots, p. \end{aligned}$$

Es decir, dependiendo de los valores de los coeficientes de regresión encontrados, solo son aceptadas: funciones de regresiones cuadráticas, lineales y constantes. Sin embargo, una plantilla que acepte cualquier tipo de función es la siguiente [24]:

$$f(t) = \alpha e^{\beta t} + \phi(t) \quad (1.4)$$

donde:

$t$ , variable dependiente del tiempo

$\alpha, \beta$ , define una componente exponencial

$\phi(t)$ , es un polinomio de grado  $n$  que se define como sigue:

$$\phi(t) = \sum_{i=0}^n a_i t^i \quad (1.5)$$

Como se puede observar no solo funciones cuadráticas lineales y constantes son admitidas por la plantilla de función de regresión anterior. En este trabajo de investigación se propone incluir al menos uno de los dos enfoques anteriores.

## 1.1. Planteamiento del problema

Los pronósticos de series de tiempo son de gran importancia en muchas aplicaciones de la ciencia y cada disciplina tiene sus propias características y necesidades [25]. Este problema pretende encontrar valores que no se han producido o calculado en la actualidad, estos puedan servir para hacer estimaciones futuras evitando tomar decisiones sin información. Algunos ejemplos de las aplicaciones de pronósticos de series de tiempo son: manejo de operaciones, mercadotecnia, finanzas y manejo de riesgos, economía, control de proceso industrial y demografía [3].

Existen diversos modelos de pronósticos de series de tiempo que pueden ser aplicados a múltiples problemas dependiendo de las necesidades. Sin embargo, sería un error pensar que ya se han resuelto todos los problemas relativos al pronóstico de series de tiempo.

En sus inicios, la investigación del pronóstico de series de tiempo fue realizado con modelos independientes unos de otros. Actualmente los investigadores han optado por diseñar e implementar modelos híbridos que tienen cómo propósito compensar sus mutuas deficiencias. Estos modelos híbridos tienden a estar compuestos por modelos de regresión, suavizamiento exponencial, técnicas heurísticas, entre otros.

## **1.2. Objetivos**

Esta investigación busca desarrollar un método híbrido que realice pronósticos de series de tiempo financieras, para lo cual se proponen los siguientes objetivos:

### **1.2.1. Objetivo general**

Desarrollar un método híbrido basado en metaaprendizaje y suavizamiento exponencial para pronóstico de series financieras con un desempeño similar a los métodos tradicionales y actuales en el estado del arte.

### **1.2.2. Objetivos específicos**

1. Revisar el estado del arte de métodos híbridos de suavizamiento exponencial e identificar implementaciones exitosas mediante experimentación usando casos de prueba comúnmente utilizadas en la literatura especializada.
2. Estudiar las técnicas híbridas para seleccionar y diseñar un nuevo algoritmo híbrido que utilice suavizamiento exponencial como núcleo e incluya componentes que fortalezcan las debilidades del suavizamiento exponencial.
3. Seleccionar y caracterizar las series de tiempo.

### 1.3. Justificación del estudio

Existen varios estudios que sugieren modelos híbridos combinando alguno de los métodos de pronóstico mencionados con anterioridad, tales como:

- Métodos ANN (Artificial Neural Networks) y métodos Box & Jenkins [26]. En particular son muy comunes en esta área, los métodos NARX (combinación de ANN y ARIMA) también conocidos como ARIMA (acrónimo del inglés AutoRegressive Integrated Moving Average)-BPNN (acrónimo del inglés Back Propagation Neural Network) [27, 28].
- Métodos GMDH (Group Method of Data Handling) y ANN [29].
- GMDH y evolución diferencial [30].
- GMDH y LSSVM (Least Squares Support Vector Machine) [31].
- Métodos ANN y suavizamiento exponencial [20].
- Métodos SVM y doble suavizamiento exponencial [23].
- Métodos ANN y ESSS(Espacio de estado de Suavizamiento Exponencial) [21].
- Métodos ANN, RBF (Radial Basis Function) NN, ARIMA y suavizamiento exponencial [22].

Se decide el uso del método de suavizamiento exponencial, dado que es un método robusto [6], de bajo costo, sencillo de implementar, con buena precisión a corto plazo [1] y que además ha ido evolucionando con el paso de los años [32].

### 1.4. Alcances

- El método propuesto será un método híbrido que incluya metaaprendizaje y el método clásico de suavizamiento exponencial.
- Se puede utilizar cualquier método del estado del arte para realizar comparaciones con el método propuesto.

- Se puede seleccionar cualquier serie de tiempo financiera para realizar pruebas de pronósticos con el método propuesto.
- De los métodos clásicos de pronóstico para las series de tiempo, se seleccionan suavizamiento exponencial, ARIMA y theta.

## 1.5. Limitaciones

- Los métodos con los que se compara el método híbrido propuesto, son de software libre y ejecutable en los equipos disponibles en el Instituto Tecnológico de Ciudad Madero.
- Las series de tiempo utilizadas en este proyecto de tesis, están reguladas por los mercados estandarizados tal como la bolsa mexicana de valores, las cuales son regidas por las leyes nacionales e internacionales.

## 1.6. Organización de la tesis

Esta tesis se divide en siete capítulos, de los cuales el primero corresponde a una breve introducción del problema a resolver, se incluyen los objetivos propuestos, así como su justificación, alcances y limitaciones.

El capítulo dos, corresponde al marco teórico, donde se definen elementos de la teoría fundamental de los pronósticos a corto, mediano y largo plazo.

En el capítulo tres, se describe una revisión del estado del arte con los trabajos de investigación más relevantes y recientes en el área de pronósticos relacionados con el tema de investigación.

Posteriormente, en el capítulo cuatro, se detalla la relación encontrada entre el exponente Hurst y el error MAPE y sMAPE, así mismo se presenta el estudio de la medida de complejidad aplicada a los pronósticos de las series de tiempo y se especifica la metodología propuesta, a la cual se le denominó **MetaLSS** que dio origen al desarrollo de seis métodos híbridos.

Por otro lado, el capítulo cinco, describe los equipos y conjuntos de datos que se utilizaron durante el desarrollo de la investigación.

El capítulo seis corresponde a los análisis y resultados obtenidos durante la experimentación realizada con dos subconjuntos de datos de las series de la competencia M4.

Por último, en el capítulo siete, se puntualizan las conclusiones finales que surgieron durante el desarrollo de la investigación.

---

## Marco teórico

Acorde con algunos datos históricos [33], los pronósticos se han realizado de manera formal a partir del siglo XVII. Sin embargo, no es hasta el siglo XX que se reportan los primeros métodos para pronosticar basados en datos como lo son: los métodos de análisis de regresión, descomposición, suavizamiento y promedios móviles autorregresivos los cuales han demostrado ser altamente eficientes [1]. Dado que su uso se ha extendido en diversas áreas de aplicación [3, 34], se han desarrollado una diversidad de métodos que aplican técnicas de aprendizaje y metaaprendizaje, así como diferentes combinaciones de ellos a los que regularmente se les denomina métodos híbridos, entre estos prevalece el uso de las redes neuronales (Artificial Neural Network) y máquinas de soporte vectorial (Support Vector Machines).

Actualmente, debido a la era tecnológica, se pueden encontrar en diversos programas computacionales (MiniTab 17, SPSS, R, entre otros) que incorporan implementaciones de los métodos mencionados y de otros métodos que fueron desarrollados en los 90's, tales como: suavizamiento exponencial, ARMA, ARIMA y sus variantes. Además de lo anterior, actualmente existen competencias de métodos de pronósticos [35–38] que proporcionan series de tiempo real de diversas áreas y que además comparten algunos de los códigos



fuentes de los métodos ganadores.

## 2.1. Serie de tiempo

La serie de tiempo es un conjunto de observaciones que se denota como  $Y_t$  donde el subíndice  $t$  corresponde a un dato almacenado en el tiempo  $t$  (en la terminología de series de tiempo financieras, también se le conoce como periodo  $t$ ) [39]; la Tabla 2.1 muestra una serie de tiempo, que corresponde a los tipos de cambio del dólar frente al peso del 01-enero-2021 al 31-marzo-2021 [40], mientras que la Figura 2.1 muestra los mismos datos de forma gráfica.

**Tabla 2.1** Serie de Tiempo del tipo de cambio del dólar frente al peso.

Día	Enero	Febrero	Marzo	Día	Enero	Febrero	Marzo
1			20.939	17		20.2023	20.5753
2		20.2248	20.6123	18	19.7875	20.2257	20.6828
3		20.1722	20.6903	19	19.757	20.4165	20.4415
4	19.9087	20.1953	20.8957	20	19.6788		
5	19.8457	20.4435	20.8498	21	19.5793		
6	19.9437			22	19.7393	20.4222	20.4823
7	19.725			23		20.6783	20.5788
8	19.91	20.1388	21.2635	24		20.5833	20.6408
9		20.1227	21.4177	25	19.9833	20.4752	20.7882
10		20.0913	21.2575	26	20.1598	20.8523	20.7662
11	19.9705	20.0588	20.9853	27	20.0152		
12	20.113	19.9535	20.7133	28	20.2168		
13	19.9373			29	20.2697		20.5677
14	19.7935			30			20.6047
15	19.7793	19.9622		31			20.6025
16		19.9518	20.7637				

Las series de tiempo se clasifican en deterministas y estocásticas (aleatorias). Acorde con [1], existen diferentes tipos de patrones de datos, estos son: horizontales, con tendencia, estacionales y cíclicos, cuando los datos fluctúan alrededor de un nivel o una media constante se les conoce como serie estacionaria en su media; la tendencia es el componente de largo plazo que representa el crecimiento o descenso en la serie de tiempo, durante un periodo extenso. El componente cíclico es la oscilación alrededor de la tendencia y el componente estacional, esto es; un patrón de cambio que se repite año tras año.

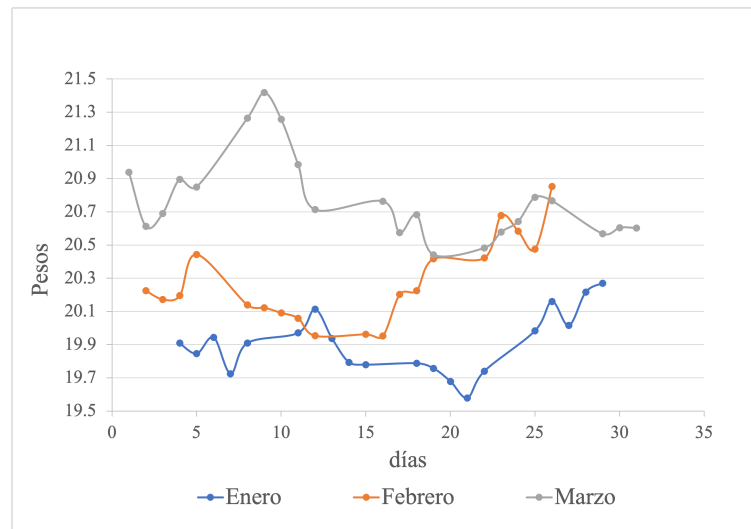


Figura 2.1 Serie de Tiempo del tipo de cambio del dólar frente al peso.

## 2.2. Autocorrelación

Al medir una variable a lo largo del tiempo, las observaciones (valores) en diferentes periodos de tiempo  $t$ , generalmente están relacionadas entre sí. A esta relación se le denomina correlación y se mide con el coeficiente de autocorrelación. En otras palabras, la autocorrelación es la relación que existe entre una variable retrasada uno o más periodos (en el tiempo) consigo misma. En la Tabla 2.2, se puede visualizar una serie de tiempo con observaciones originales ( $Y_t$ ) y 3 desfases en el tiempo ( $Y_{t-1}$ ,  $Y_{t-2}$ ,  $Y_{t-3}$ ), aunque técnicamente, como se puede ver, son los mismos datos pero con corrimientos, llamados desfases. A través de estos desfases, se calcula el coeficiente de autorrelación.

La fórmula del coeficiente de autocorrelación se define de la siguiente manera:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad k = 0, 1, 2, \dots \quad (2.2.1)$$

donde

$r_k$  = coeficiente de autocorrelación para un retraso de  $k$  periodos (instantes)

$\bar{Y}_t$  = media de los valores de la serie

**Tabla 2.2** Ejemplo de una serie de tiempo con 1, 2 y 3 retrasos en el tiempo.

Tiempo $t$	Meses	Observaciones			
		$Y_t$	$Y_{t-1}$	$Y_{t-2}$	$Y_{t-3}$
1	Mes1	110			
2	Mes2	140	110		
3	Mes3	132	140	110	
4	Mes4	140	132	140	110
5	Mes5	134	140	132	140
6	Mes6	142	134	140	132
7	Mes7	140	142	134	140
8	Mes8	124	140	142	134
9	Mes9	145	124	140	142
10	Mes10	156	145	124	140
11	Mes11	147	156	145	124
12	Mes12	187	147	156	145

$Y_t$  = observación en el periodo (instante)  $t$

$Y_{t-k}$  = observación de  $k$  periodos (instantes) anteriores a partir del tiempo  $t$

### 2.2.1. Correlograma

Un correlograma es una gráfica de las autocorrelaciones para varios retrasos de una serie de tiempo. La Tabla 2.3 muestra un ejemplo de una serie de tiempo del tipo de cambio del peso-dólar durante el mes de enero de 2021 y la columna *ACF1* corresponde al coeficiente de autocorrelación calculados con diferentes desfases en el tiempo.

**Tabla 2.3** Autocorrelación del tipo de cambio del dólar-peso en enero 2021.

Día	Tipo cambio	ACF1
4	19.9087	1.0000
5	19.8457	0.1821
6	19.9437	-0.2048
7	19.7250	-0.6066
8	19.9100	-0.0901
11	19.9705	0.0695
12	20.1130	0.1364
13	19.9373	-0.0037
14	19.7935	0.0331
15	19.7793	-0.0161

Por otra parte, en la Figura 2.2 se puede visualizar y analizar estos datos para identificar la existencia de patrones.

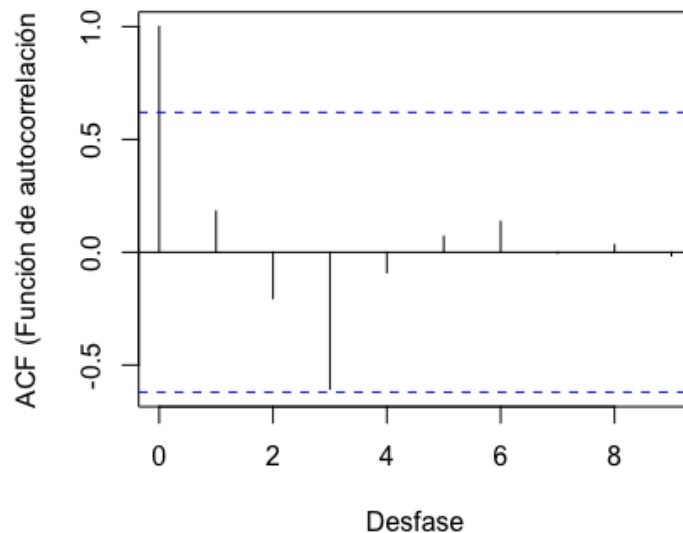


Figura 2.2 Autocorrelación del tipo de cambio del dólar-peso en enero 2021.

### 2.2.2. Patrones en los datos

Con el coeficiente de autocorrelación o correlograma, es posible determinar los patrones de los datos, los cuales pueden ser:

- *Aleatoriedad*: las autocorrelaciones entre  $Y_t$  y  $Y_t - k$  para cualquier retraso de tiempo  $k$  son cercanas a cero. Los valores sucesivos de una serie de tiempo no están relacionados entre sí.
- *Tendencia*: las observaciones sucesivas están altamente correlacionadas; los coeficientes de correlación son diferentes de cero para los primeros retrasos de tiempo, y de forma gradual tienden a cero conforme se incrementa el número de retrasos.
- *Estacional*: presenta un coeficiente de autocorrelación significativo en el retraso de tiempo estacional o en los múltiplos del retraso estacional. El retraso estacional se considera de tamaño 4 para datos trimestrales y tamaño 12 para datos mensuales [1].
- *Estacionaria*: es una serie de tiempo donde la media y la varianza, permanecen constantes en el tiempo. Una serie que contiene una tendencia es no estacionaria. Los coeficientes de autocorrelación de una serie estacionaria decrecen hacia cero bastante

rápidamente, por lo común después del segundo o tercer retraso de tiempo.

Cuando las autocorrelaciones de las series de tiempo no estacionarias, permanecen muy grandes durante varios periodos, se recomienda eliminar la tendencia antes de aplicar cualquier modelo de pronóstico [1].

### **2.3. Pronóstico de series de tiempo**

Un pronóstico consiste en la estimación futura de un dato u observación  $Y_{t+1}$ . Utilizando diferentes herramientas como los datos históricos  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$  y algunos métodos de predicción que sean los más adecuados acorde a la serie de tiempo que se está analizando y se desea pronosticar [41].

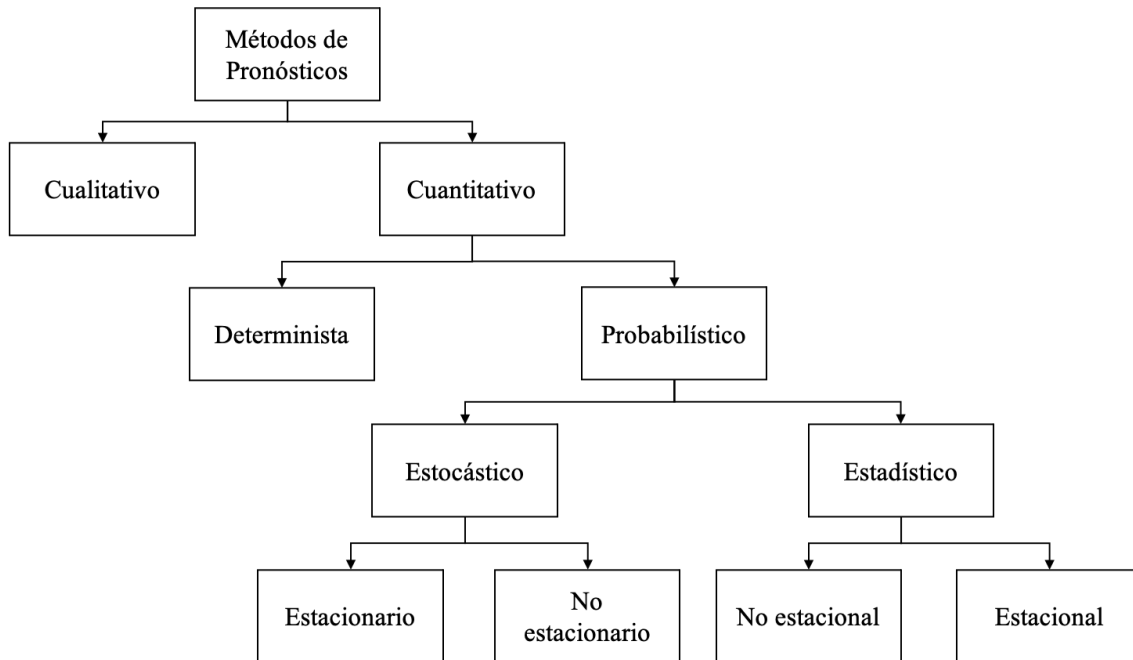
### **2.4. Taxonomía de los métodos de pronósticos**

Para realizar un pronóstico de una serie de tiempo, se puede aplicar uno o varios métodos, actualmente existe una amplia clasificación de estos métodos. La Figura 2.3 muestra, de manera general, la clasificación acorde con [1]. Las series de tiempo generalmente están clasificadas como estocásticas y estadísticas, por lo que las investigaciones de estas, se enfocan en métodos probabilísticos.

Como se ha mencionado previamente, en la actualidad hay una diversidad de métodos combinados que permiten minimizar las debilidades propias de cada uno volviéndolos robustos, la Figura 2.4 muestra algunos de los métodos más utilizados en el desarrollo de métodos híbridos.

### **2.5. Notación en pronósticos y medidas de error**

Cuando se trata del área de pronósticos, se utiliza una notación básica estándar, ésta se describe a continuación:



**Figura 2.3** Taxonomía general de los métodos de pronósticos [1].

$Y_t$  = valor de una serie de tiempo en el periodo (instante)  $t$

$\hat{Y}_t$  = valor pronosticado de  $Y_t$

$e_t = Y_t - \hat{Y}_t$  = residuo o error del pronóstico

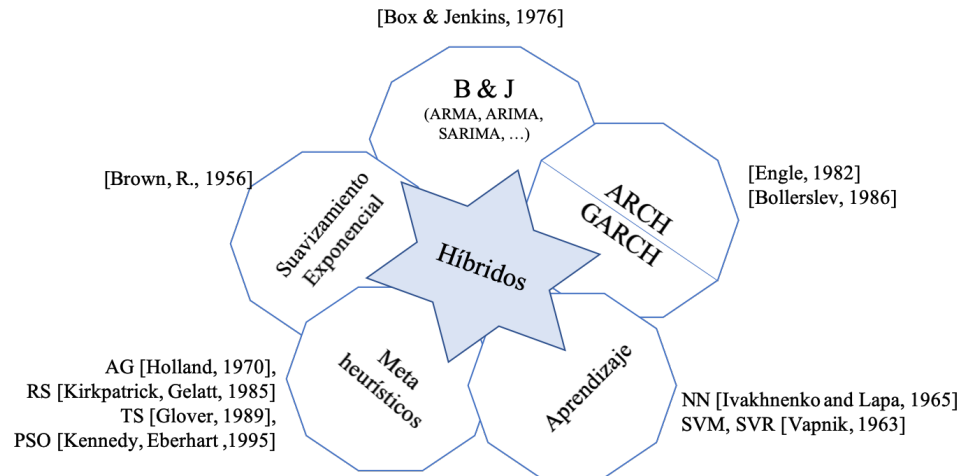
Después de realizar un pronóstico, es importante cuantificar el error obtenido, por lo que estos se evalúan mediante diversas medidas de precisión [1] y son usadas para:

- Comparar la exactitud de dos (o más) técnicas diferentes.
- Medir la utilidad o confiabilidad de una técnica en particular.
- Ayudar en la búsqueda de una técnica óptima.

Las medidas más comúnmente utilizadas son las siguientes:

### 2.5.1. Desviación media absoluta

La desviación media absoluta (MAD, por sus siglas en inglés) mide la exactitud del pronóstico, promediando las magnitudes de los errores del pronóstico (los valores absolutos de los



**Figura 2.4 Métodos híbridos.**

errores). La MAD está en las mismas unidades que la serie original, y proporciona un tamaño promedio de los errores sin importar la dirección. La ecuación 2.5.1 corresponde al cálculo de esta medida.

$$MAD = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t| \quad (2.5.1)$$

### 2.5.2. Error cuadrático medio

En el error cuadrático medio (MSE, por sus siglas en inglés), cada error o residuo se eleva al cuadrado; posteriormente éstos se suman y se dividen entre el número de observaciones o periodos. En este enfoque los errores grandes obtenidos al pronosticar, impactan mayormente ya que los errores están elevados al cuadrado. La ecuación 2.5.2 corresponde al cálculo de esta medida.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (2.5.2)$$

### 2.5.3. Raíz cuadrada del error cuadrático medio

La raíz cuadrada del MSE, o la raíz cuadrada del error cuadrado medio (RMSE, por sus siglas en inglés), se utiliza para evaluar los métodos generadores de pronósticos. Tanto la

RMSE como la MSE, castigan los errores grandes, pero tienen las mismas unidades de la serie que se está pronosticando, de modo que su magnitud se interpreta con mayor facilidad. La ecuación 2.5.3 corresponde al cálculo de esta medida.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (2.5.3)$$

#### 2.5.4. Error porcentual absoluto medio

El error porcentual absoluto medio (MAPE, por sus siglas en inglés) se calcula obteniendo el error absoluto de cada periodo, dividiendo éste entre el valor real observado en ese periodo y promediando estos errores porcentuales absolutos. El resultado final se multiplica después por 100 y se expresa como porcentaje. El MAPE es especialmente útil cuando los valores  $Y_t$  son grandes. El MAPE no tiene unidades de medición (es un porcentaje) y sirve para comparar la exactitud de la misma técnica o de otras técnicas en dos series completamente diferentes. La ecuación 2.5.4 corresponde al cálculo de esta medida.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{|Y_t|} \quad (2.5.4)$$

#### 2.5.5. Error porcentual absoluto medio simétrico

Similar al error MAPE, el sMAPE es un error independiente de la escala, el valor absoluto en el denominador puede evitar valores negativos en el error sMAPE. Fue propuesto en [36], y ha tenido diferentes adaptaciones [42], [43]. La ecuación 2.5.5 corresponde a una de los cálculos propuestos para esta medida de error.

$$sMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|\hat{Y}_t - Y_t|}{(Y_t + \hat{Y}_t)/2} \quad (2.5.5)$$

#### 2.5.6. Error porcentual medio

Algunas veces es necesario determinar si el método para pronosticar está sesgado (con pronósticos consistentemente altos o bajos). En estos casos, se usa el error porcentual



medio (MPE, por sus siglas en inglés), el cual se calcula obteniendo el error en cada periodo, dividiendo éste entre el valor real de ese periodo y luego promediando estos errores porcentuales. El resultado usualmente se multiplica por 100 y se expresa como un porcentaje. Si el enfoque del pronóstico no tiene sesgo, el MPE producirá un resultado que esté cercano a cero. Si el resultado es un porcentaje negativo grande, el método de elaboración del pronóstico está sobreestimando consistentemente. Si el resultado es un porcentaje positivo grande, el método de elaboración del pronóstico está subestimando consistentemente. La ecuación 2.5.6 corresponde al cálculo de esta medida.

$$MPE = \frac{1}{n} \sum_{t=1}^n \frac{(Y_t - \hat{Y}_t)}{Y_t} \quad (2.5.6)$$

## 2.6. Ventajas y desventajas de los métodos de pronósticos

Como se indica en diversas fuentes de la literatura [6, 44–48], cada método de pronóstico tiene sus respectivas ventajas y desventajas, en la Tabla 2.4 se recopilan algunos métodos más utilizados y se indican algunas de ellas. Como se puede observar, algunos métodos son complementarios de otros; razón por la cual, surge la combinación entre algunos de ellos, tratando de minimizar las deficiencias que puedan tener. A partir de ese enfoque, surgen los métodos híbridos, los cuales son ampliamente utilizados en los últimos años.

## 2.7. Selección de técnicas para pronosticar

Un factor importante que influye en la selección de la técnica de elaboración del pronóstico, es la identificación y comprensión de patrones históricos en los datos. La Tabla 2.5 muestra los métodos que se pueden utilizar de acuerdo a los patrones que se encuentren en la serie de tiempo que se desea pronosticar [1].

**Tabla 2.4** Ventajas y desventajas de algunos métodos de pronósticos.

<b>Método</b>	<b>Ventajas</b>	<b>Desventajas</b>
Suavizamiento exponencial	Efectivo [44]. Ampliamente usado en finanzas y negocios [44]. Robusto [6, 45] . Capaz de encontrar modelos lineales [44]. Sencillo y de bajo costo [1]. Fácil adaptación a los cambios de los mercados [46].	No es capaz de encontrar patrones de no linealidad [44].
Redes neuronales artificiales	Capacidad para encontrar modelos no lineales [44]. Explora relaciones no estructuradas de los datos [47]. Gran número de parámetros [49, 50].	No es capaz de encontrar modelos lineales [44]. Presenta problemas con la regresión lineal múltiple [48].
ARIMA	Pueden representar series de tiempo autoregresivas, de promedios móviles y una combinación de ambas [51].	No captura patrones no lineales [51].
Soporte vectorial de regresión	Presenta una mejor precisión en la predicción debido a el principio de minimización de estructura de riesgo [52].	Determinación de sus parámetros [52].
Algoritmos genéticos	Optimización de parámetros [52]. Eficiente en un numero complejo de restricciones [53].	

## 2.8. Método de suavizamiento exponencial

El método de suavizamiento exponencial, es un método de pronóstico que tuvo sus orígenes en los años 1950's-1960's con el trabajo de Brown [8], Holt [9] y Winters [54].

El método analiza continuamente un valor estimado, dando mayor importancia a las experiencias más recientes. Se basa en suavizar valores pasados de una serie, asignando un peso de manera exponencial, de tal manera que, la observación más reciente recibe el

**Tabla 2.5** Selección de métodos de pronósticos de acuerdo con los patrones de los datos.

Métodos	Patrones de los datos				Plazo		
	Estacionarios	Tendencia	Estacionales	Serie Cíclicas	C	M	L
Informales	x						
Promedio simple	x						
Promedio móviles	x	x			x	x	
ARMA (Box & Jenkins)	x				x	x	
ARIMA (Box & Jenkins)	x	x	x	x	x	x	
Exponenciales		x					
Regresión simple							
Regresión Múltiple			x	x	x	x	x
Suavizamiento Exponencial lineal Holt		x			x		
Suavizamiento Exponencial de Winter			x		x		

**C – Corto plazo, M – Mediano plazo, L – Largo plazo**

peso mas grande, igual a  $\alpha$ , la siguiente observación recibe el peso igual a  $\alpha(1 - \alpha)$ , a la tercer observación se le asigna el peso igual a  $\alpha(1 - \alpha)^2$  y así sucesivamente con las demás observaciones. La Tabla 2.6 muestra la asignación del valor de  $\alpha$  que corresponde al peso que se asigna a cada una de las observaciones, de acuerdo al periodo (instante) de tiempo  $t$ .

**Tabla 2.6** Peso asignado de acuerdo con el número de observación.

Observación	Peso asignado
$Y_t$	$\alpha = (0 < \alpha < 1)$
$Y_{t-1}$	$\alpha(1 - \alpha)$
$Y_{t-2}$	$\alpha(1 - \alpha)^2$

La ecuación de Suavizamiento Exponencial es la siguiente:

$$\hat{Y}_{t-1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t \tag{2.8.1}$$

Y puede quedar de la siguiente manera:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t = \alpha Y_t + \hat{Y}_t - \alpha \hat{Y}_t \tag{2.8.2}$$

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t) \tag{2.8.3}$$

donde

$\hat{Y}_{t+1}$  = nuevo valor suavizado o el valor del pronóstico para el siguiente periodo

$\alpha$  = constante de suavización ( $0 < \alpha < 1$ )

$Y_t$  = nueva observación o el valor real de la serie en el periodo  $t$

$\hat{Y}_t$  = último valor suavizado o el pronóstico del periodo  $t$

Los dos modelos más sencillos que relacionan los valores observados ( $Y_t$ ) de una serie de tiempo, con los patrones (componentes) de tendencia ( $T_t$ ), estacionalidad ( $S_t$ ) e irregulares o aleatorios ( $I_t$ ) son el modelo de componentes aditivo y el multiplicativo, los cuales se describen a continuación:

- Modelo de componente aditivo (suavizamiento exponencial), su resultado es mejor cuando la serie de tiempo se analiza y tiene la misma variabilidad a lo largo de toda la serie, es decir, todos los valores de la serie caen dentro de un intervalo de valores constante centrados en la tendencia.

$$\bar{Y}_t = T_t + S_t + I_t \quad (2.8.4)$$

- Modelo de componente multiplicativo, funciona mejor cuando la variabilidad de la serie de tiempo aumenta con el tiempo, es decir, los valores de la serie se dispersan conforme la tendencia aumenta, y el conjunto de observaciones tiene la apariencia de un megáfono o embudo.

$$Y_t = T_t + S_t + I_t \quad (2.8.5)$$

Acorde con [7], existen 15 métodos de suavizamiento exponencial, cada uno de ellos genera buenos resultados acorde con ciertos patrones de datos, a continuación algunos de los que destacan:

- Simple exponencial simple (Simple Exponential Smoothing): para series de tiempo donde no hay tendencia y no son estacionarias.
- Suavizamiento exponencial lineal de Holt ( Holt's Linear): para series de tiempo con tendencia aditiva.

- Suavizamiento exponencial de Holt-Winters (Holt-Winters additive method ): para series de tiempo con tendencia aditiva y estacionalidad aditiva.
- Suavizamiento exponencial de Holt-Winters (Holt-Winters multiplicative method): para series de tiempo con tendencia aditiva y estacionalidad multiplicativa.

Debido a que el método de suavizamiento exponencial es simple, robusto [6] y con buenas predicciones a corto plazo [1], sigue vigente en la actualidad [32] y ha sido utilizado como parte de las hibridaciones exitosas entre métodos [47].

## 2.9. Metodología Box-Jenkins

En el siglo XIX, los estudios de las series de tiempo eran en su mayoría deterministas, sin embargo, en [55], Yule propuso que las series de tiempo podrían ser estimadas como procesos estocásticos. A partir de este momento, se desarrollaron métodos de pronósticos autorregresivos (AR) y de promedios móviles (MA).

En [2], Box-Jenkins integraron el conocimiento existente y desarrollaron una metodología iterativa de tres etapas para la identificación, estimación y verificación de series de tiempo. A este enfoque se le dio el nombre de *metodología Box & Jenkins* la cual es diferente de la mayoría de los métodos porque no supone ningún patrón particular en los datos históricos de las series que se van a pronosticar.

Se basa en un enfoque iterativo para identificar un posible modelo, el modelo seleccionado se coteja con los datos históricos para ver si describe la serie con exactitud. Se dice que un modelo de pronóstico está bien ajustado si los residuos son generalmente pequeños, están distribuidos aleatoriamente y no contienen información útil. Por otra parte, si el modelo producido no está bien ajustado, la metodología repetirá el proceso usando un nuevo modelo con el objetivo de encontrar uno que mejore al anterior. Este procedimiento iterativo continúa hasta que se encuentra un modelo bien ajustado. La Figura 2.5 muestra este enfoque. Con el uso de las computadoras, se popularizó el uso de los modelos autorregresivos integrados de promedios móviles (ARIMA, por sus siglas en inglés) y sus extensiones, éstos

son una clase de modelos lineales que son capaces de representar tanto series de tiempo estacionarias como no estacionarias. Los modelos ARIMA dependen mucho de los patrones de autocorrelación en los datos [1].

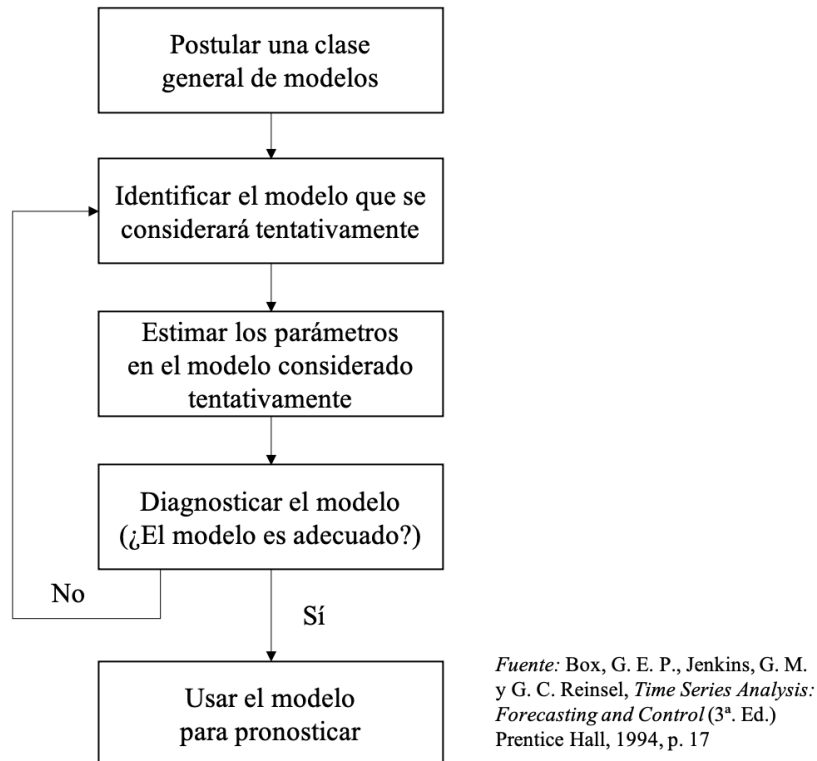


Figura 2.5 Diagrama de flujo del modelo Box-Jenkins [2].

## 2.10. Métodos híbridos basados en suavizamiento exponencial

Los métodos híbridos se han vuelto populares en los últimos años, ya que tratan de complementar las ventajas de los diferentes métodos que se combinan, algunas de estas combinaciones que incluyen el método de Suavizamiento exponencial son:

- SVM-ES (máquinas de soporte vectorial con suavizamiento exponencial)
- ES-NN (redes neuronales con suavizamiento exponencial)
- ES-ARIMA-RNN (suavizamiento exponencial con ARIMA y redes neuronales)

En [38], se reportan los resultados de las competencias de Makridakis llamadas M4-Competition realizadas en el año 2018, donde participaron 61 métodos de pronósticos para las 100,000 series de tiempo de la vida real que fueron seleccionadas y proporcionadas como conjunto de datos. De estas competencias, resultó ganador un método híbrido que combina suavizado exponencial con redes neuronales recurrentes (ES-RNN, por sus siglas en inglés), del cual se proporciona el código fuente que se encuentra en los lenguajes c++ y python. Más detalles de la implementación de este método ganador se encuentra en [47].

---

## Estado del arte

Dado que los pronósticos de series de tiempo financieras han sido un nicho de oportunidades de continua mejora en las investigaciones, en la actualidad existe una variedad de trabajos que lo abordan.

La revisión de los trabajos relacionados en esta investigación, se ha dividido en dos partes, la primera incluye trabajos generales de una diversidad de métodos de pronósticos aplicados en series de tiempo de diferentes sectores, los cuales se enlistan y describen a continuación.

- Hadavandi et al. [56], realizan una investigación con el objetivo de utilizar una cantidad mínima de datos de entrada y un modelo de predicción que no sea complejo. Con ese objetivos se presentó un enfoque integrado, basado en un sistema genético difuso (Genetic Fuzzy Systems) y una red neuronal artificial (ANN) en el que realizaron agrupamiento usando el método SOM (Self-Organizing MAP). Se aplicó el enfoque en series de tiempo del sector de aerolíneas y se emplearon los indicadores MAPE (Mean Absolute Percentage Error) and RMSE (Root Mean Square Error) para evaluar los resultados obtenidos. Dichos resultados mostraron que el enfoque propuesto, mejora todos los métodos previos.
- Wang et al. [16] presentan una predicción de los precios de las acciones o índices con



datos no estructurados, se aplica una transformación de ondas (wavelet transform) para descomponer la serie de tiempo en 6 capas de baja y alta frecuencia y se utiliza para la eliminación del ruido existente. El modelo es basado en una red neuronal con back propagation (WBDP), la experimentación se realizó con datos del Shanghai Composite Index y se comparó contra una BPNN (Back Propagation Neural Network), mostrando que la capa 4, resultado de la descomposición de la serie de tiempo, obtiene errores mas pequeños.

- Guresen et al. [46], evalúan la efectividad de algunos modelos de redes neuronales como Multi Layer Perceptron (MLP), dynamic artificial neural network (DAN2) y una red neuronal híbrida que utiliza el método GARCH (Generalized Autoregressive Conditional Heteroscedasticity) aplicándolos en predicciones de mercados. Se realiza una comparativa con las medidas de error MSE (Mean Square Error) y MAD (Mean Absolute Deviation) utilizando índices de la bolsa NASDAQ, los resultados señalan que el modelo de red neuronal MLP es el modelo que obtuvo mejores resultados para el índice de mercado de prueba.
- Chang et al. [14] realizan una predicción de la tendencia de los precios de las acciones usando técnicas de indicadores como entrada. El método propuesto es llamado: Evolving Partially Connected Neural Networks (EPCNNs). En el estudio se realiza una comparación del rendimiento contra otros modelos como: BPN (Back Propagation Neural Network), sistema difuso TSK (Tak-agi–Sugeno–Kang) y el método MARS (Multivariate Adaptive Regression Splines), con la experimentación realizada, el método propuesto logró un 94% de predicción en los pronósticos realizados, probando ser efectivo en la predicción de las tendencias de los índices de mercado.
- Wang et al. [57], desarrollaron un modelo híbrido de pronósticos, combinando ESM (Exponential Smoothing Model), ARIMA (Autoregressive Integrated Moving Average (ARIMA),) y BPNN (Back Propagation Neural Network), se determina el peso del modelo a través de un algoritmo genético. El modelo se aplica para pronosticar los precios de cierre y apertura de los índices Shenzhen Integrated y Dow Jones Industrial

Average, finalmente se realizó una comparativa con métodos de pronóstico individuales donde se observa un mejor desempeño del método propuesto.

- Ticknor [58] emplea una Red Neuronal Artificial para pronosticar el comportamiento del mercado financiero. Su red neuronal emplea regularización bayesiana y experimentó con activos de Microsoft y el grupo Goldman Sachs para determinar la efectividad del modelo. Los resultados indicaron que el modelo propuesto se ejecuta tan bien como los modelos más avanzados, sin necesidad de preprocesamiento de datos, pruebas de estacionalidad o análisis de ciclos.
- Oliveira et al. [15], experimentaron con la predicción del comportamiento de futuros precios de cierre de activos (a corto plazo) en la bolsa de valores de Brazil (BM & FBO-VESPA), a través de una Red Neuronal implementada en Encog(Heaton, 2013), cuyo nombre completo es “Encog Artificial Intelligence Framework” el cual es un ambiente comúnmente utilizado para desarrollar modelos de diversa índole, incluyendo modelos económicos y financieros (Heaton, 2015). Encog permite combinar técnicas de análisis fundamental de series de tiempo.
- Por otra parte, Palivonaite et al. [19] identificaron un indicador algebraico con suavizamiento para pronósticos de series de tiempo financieras de corto plazo; el método propuesto es llamado Algebraic Prediction Mixed Smoothing (APMS). La experimentación se llevo a cabo con series de tiempo reales para mostrar la efectividad del algoritmo propuesto.
- Finamore et al. [59] realizaron una investigación aplicada a pronósticos del viento desarrollando una activación directa de una Red Neuronal. En este trabajo se utilizaron series de tiempo reales con datos meteorológicos y una frecuencia en escala de horas en el Sur de Italia. En su investigación concluyen que se obtuvo un modelo de pronóstico persistente con buenos resultados y emplearon las mediciones de errores APE (Absolute Percentage Error), MAPE y MSE.
- Tratar [25], utilizó series de tiempo que contienen observaciones iguales a cero y

con gran ruido, el autor desarrolló un método de pronóstico aditivo y multiplicativo. Al método clásico Holt-Winters se le introdujo un parámetro de suavizamiento exponencial en el nivel de recurrencia de la ecuación y se emplearon las series de tiempo de M3-Competition.

- Pan et al. [60] proponen un nuevo método de entrenamiento de una red neuronal el cual combina los métodos PSO (Particle Swarm Optimización) con SSO (Simplified Swarm Optimization) denominado PSOSS, el cual mejora la habilidad para entrenar los pesos de la red neuronal. El método fue empleado para entrenar un modelo de pronóstico de una red neuronal para el pronóstico del índice bursátil HSI (por sus siglas en inglés, Hang Seng index) el cual pertenece a la bolsa de valores de Hong Kong.
- Adicionalmente, Wu et al. [61] emplean el método PSO junto con el método de suavizamiento exponencial. Los autores emplean el PSO con un enfoque de generación de parvadas que busca lograr un efecto de suavizamiento y agregar un peso adicional a los cambios recientes; el cual es aplicado al pronóstico del precio del cerdo en China. Los resultados han demostrado que el método propuesto supera al tradicional método de doble exponencial en problemas de pronósticos.

En la Tabla 3.1, se muestran algunos trabajos de investigación más recientes, que abordan el tema del aprendizaje (learning) y metaaprendizaje (metalearning) para series de tiempo no financieras y financieras, la columna cinco y seis corresponden a datos relacionados con técnicas de aprendizaje(A) y Metaaprendizaje(MA).

**Tabla 3.1** Estado del arte de aprendizaje y metaaprendizaje.

<b>Año</b>	<b>Título</b>	<b>Método propuesto</b>	<b>Serie de tiempo financiera</b>	<b>A</b>	<b>MA</b>
2015 [24]	Forecasting Oil Production Time Series with a Population-Based Simulated Annealing Method.	Redes Neuronales con Recocido Simulado / ARIMA.	NO	SI	SI
2015 [62]	A novel hybrid model using teaching-learning based optimization and a support vector machine for commodity futures index forecasting.	SVM (Máquina de Soporte Vectorial) y Enseñanza-aprendizaje basada en optimización (TLBO).	SI	SI	SI
2017 [34]	A hybrid ETS-ANN model for time series forecasting.	ETS (Suavizamiento exponencial de espacios de estado) y Redes Neuronales.	NO	SI	NO
2018 [63]	Cross-domain Meta-learning for time series forecasting.	Suavizamiento exponencial, Prome-dios móviles, ARI-MA, Redes neuro-nales, entre otros.	SI	SI	SI

<b>Tabla 3.1 – continuación de la página anterior</b>					
<b>Año</b>	<b>Título</b>	<b>Método propuesto</b>	<b>Serie de tiempo financiera</b>	<b>A</b>	<b>MA</b>
2018 [64]	A hybrid forecasting model for prediction of stock value of tata steel using support vector regression and particle swarm optimization.	SVR (Regresión de Soporte Vectorial) y PSO (Optimización por enjambre de partículas).	SI	SI	SI
2018 [52]	Forecasting leading industry stock prices based on a hybrid time-series forecast model	SVM (Máquina de Soporte Vectorial) y Algoritmo genético.	SI	SI	SI
2020 [47]	A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting	Suavizamiento exponencial, Redes Neuronales Recurrentes.	NO	SI	SI
2020 [65]	A combination-based forecasting method for the M4-competition	Naïve/snaïve, Suavizamiento exponencial, Bagging, Multi-aggregation prediction algorithm, ARIMA, Theta, etc.	NO	SI	SI

<b>Tabla 3.1 – continuación de la página anterior</b>					
<b>Año</b>	<b>Título</b>	<b>Método propuesto</b>	<b>Serie de tiempo financiera</b>	<b>A</b>	<b>MA</b>
2020 [66]	FFORMA: Feature-based forecast model averaging	Naïve, Rwf, Theta, ARIMA, Redes neuronales, etc	NO	SI	SI
2021 [67]	A new financial data forecasting model using genetic algorithm and long short-term memory network	Algoritmo Genético, Memoria de largo plazo, Redes neuronales	SI	SI	SI

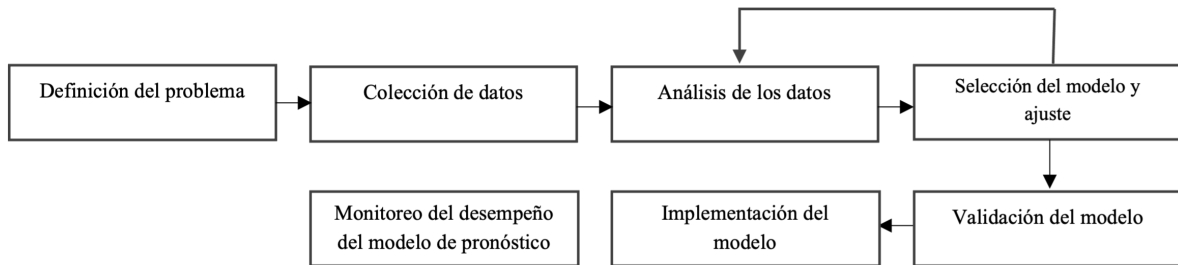
Como se puede visualizar en los trabajos mencionados previamente, en la actualidad el uso de métodos híbridos que utilizan aprendizaje y metaaprendizaje son comúnmente estudiados y propuestos con diversas variantes, sus resultados han demostrado que dichos métodos han obtenido un mejor desempeño que algunos métodos tradicionales.

---

## Metodología

El proceso de pronósticos general, esta compuesto por siete pasos, los cuales se muestran en la Figura 4.1, se definen en [3] y se resumen a continuación:

1. Definición del problema: Este paso consiste en entender cómo será utilizado el pronóstico con las expectativas del cliente. Se deben considerar algunas de las siguientes variables:
  - a) el horizonte de proyección o el tiempo de espera
  - b) el intervalo de tiempo  $t$
  - c) el nivel de precisión del pronóstico
2. Colección de datos: consiste en obtener datos históricos relevantes de las variables que se desea pronosticar. A menudo es necesario tratar con valores faltantes de algunas variables, existen diversidad de técnicas para tratar estos valores y dependerá del tipo de serie de tiempo con la que se esta trabajando [1, 3].
3. Análisis de los datos: es un importante paso preliminar a la selección del modelo de pronóstico que se utilizará. Los datos de las series de tiempo deben ser construidas



**Figura 4.1** El proceso del pronóstico [3].

y visualmente inspeccionadas para el reconocimiento de patrones, tales como los componentes de tendencia, estación y cíclicos.

4. Selección del modelo y ajuste: consiste en seleccionar, uno o más modelos de pronósticos, así como el ajuste del modelo a los datos.
5. Validación del modelo: consiste en la evaluación del modelo de pronósticos para determinar el desempeño para una determinada aplicación.
6. Implementación del modelo de pronóstico: implica la ejecución del modelo y los resultados del pronóstico usados por el cliente.
7. Monitoreo del desempeño del modelo de pronóstico: debe ser una actividad que se realice cuando el modelo ha sido utilizado, para asegurarse que se está ejecutando de manera satisfactoria, la naturaleza del pronóstico es que las condiciones cambian sobre el tiempo, y el modelo que se ejecutó bien en el pasado, podría alterar su desempeño.

Como parte de la caracterización de las series, una de las preguntas de investigación que surgió en la etapa de experimentación, fue si el exponente de Hurst era realmente una característica importante, por lo que se profundizó en este tema y se planteó una metodología sencilla, la cual se define en la siguiente sección.

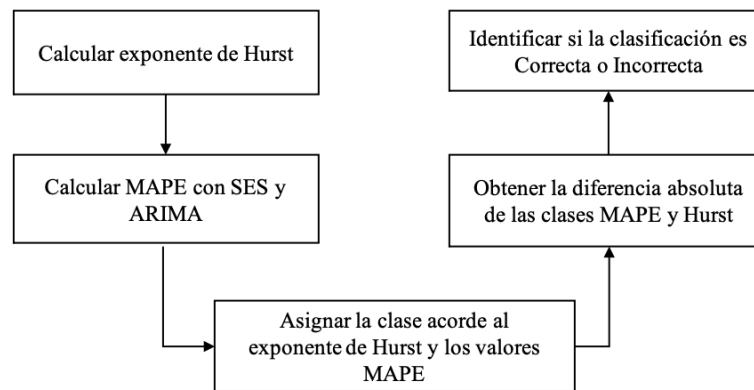
#### 4.1. Análisis de la relación Hurst-MAPE y sMAPE

Considerando que el exponente Hurst permite determinar la persistencia de las series de tiempo, se realizó el cálculo con las series de tiempo del conjunto de M3-Competition.



Posteriormente se realizaron los pronósticos con los métodos clásicos SES (suavizamiento exponencial simple) y ARIMA y se calcularon los valores de las medidas de error MAPE y sMAPE. A través del análisis realizado, se encontró la relación entre el exponente de Hurst y el error de pronóstico de cada serie. En síntesis, el análisis propuesto, consta de tres etapas que aparecen en las Figuras 4.2 y 4.3, las cuales se describen a continuación:

- Primera etapa: Calcular el exponente de Hurst para cada serie y asociarlo a una categoría de complejidad. Esto se realiza calculando la medida de error de cada uno de los métodos de pronóstico aplicados y asociarlo a la medida de complejidad establecida.
- Segunda etapa: Aplicar un método de agrupamiento para identificar la relación entre el error de pronóstico MAPE y sMAPE de cada método y el exponente de Hurst.
- Tercera etapa: Realizar un análisis gráfico para determinar si existe tendencia para las series analizadas. En este caso se consideraron cuatro tipos de series de tiempo: anual, cuatrimestral, mensual y otros (Year, Quarter, Month y Other, en inglés).



**Figura 4.2** Primera etapa del análisis realizado [4].

Como dato de confirmación, en [68], se documentó que la alta persistencia de una serie es un indicador de la alta facilidad de realizar un buen pronóstico, por el contrario, la baja persistencia debiera estar asociada con la complejidad de dichas series.

El resto de esta sección contiene los conceptos y métodos relacionados con el análisis propuesto anteriormente.

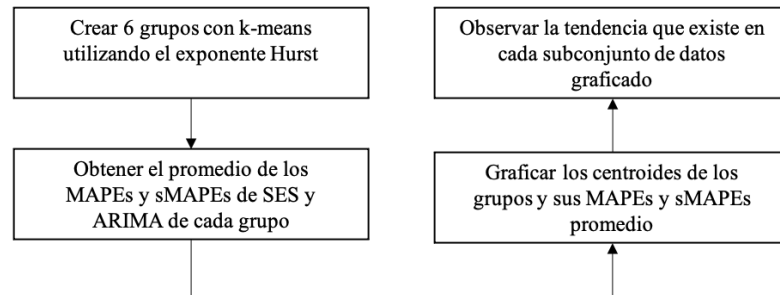


Figura 4.3 Segunda y tercera etapa del análisis realizado [4].

#### 4.1.1. Exponente de Hurst

Existe una variedad de métodos que permiten calcular el exponente de Hurst ( $H$ ), se revisaron, entre otros, el método de Rango Reescalado (R/S) [69] y el proceso ARFIMA (Autoregressive Fractionally Integrated Moving Average) desarrollado por Granger & Joyeux [70], establecida por Hosking [71] y utilizado por Hyndman [72].

Las características que se atribuyen a cada serie, dependiendo del valor de  $H$ , son las que se señalan en la Tabla 4.1. Se hace notar que conforme  $H$  es cercano a la unidad, las series son más persistentes, de forma que debieran ser menos complejas [69].

Tabla 4.1 Interpretación de los valores del exponente de Hurst.

Intervalo	Descripción
$0.0 < H < 0.5$	Series anti-persistentes, tienen una tendencia de reversión a la media.
$H = 0.5$	Series de caminata aleatoria, los eventos son aleatorios y no hay correlación.
$0.5 < H < 1.0$	Series de tiempo persistentes.
$H = 1.0$	Series deterministas.
Otros valores.	Es posible encontrar otros valores, pero no se han estudiado este tipo de series.

Como se señaló anteriormente, el valor de  $H$  permite evaluar la persistencia de las series de tiempo [73]. En esta experimentación se utiliza el método Rango Reescalado (R/S) de Peters [69], a través de los siguientes pasos:

- Paso uno: dividir la serie de tiempo en intervalos conformando particiones como se muestra en la Figura 4.4.
- Paso dos: obtener del Rango Reescalado de cada partición en el paso uno tal como se

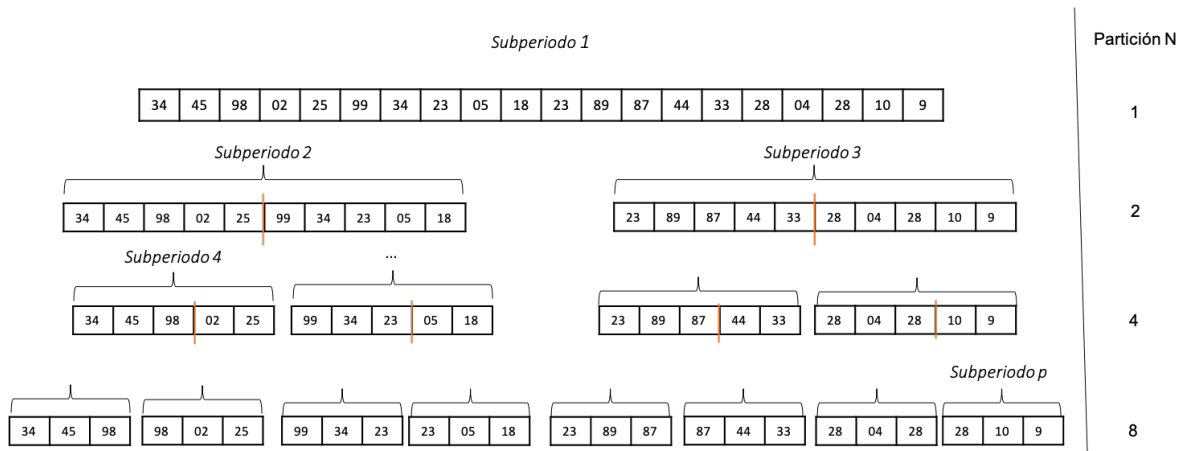


Figura 4.4 Paso uno para el cálculo del exponente de Hurst.

muestra en la Figura 4.5.

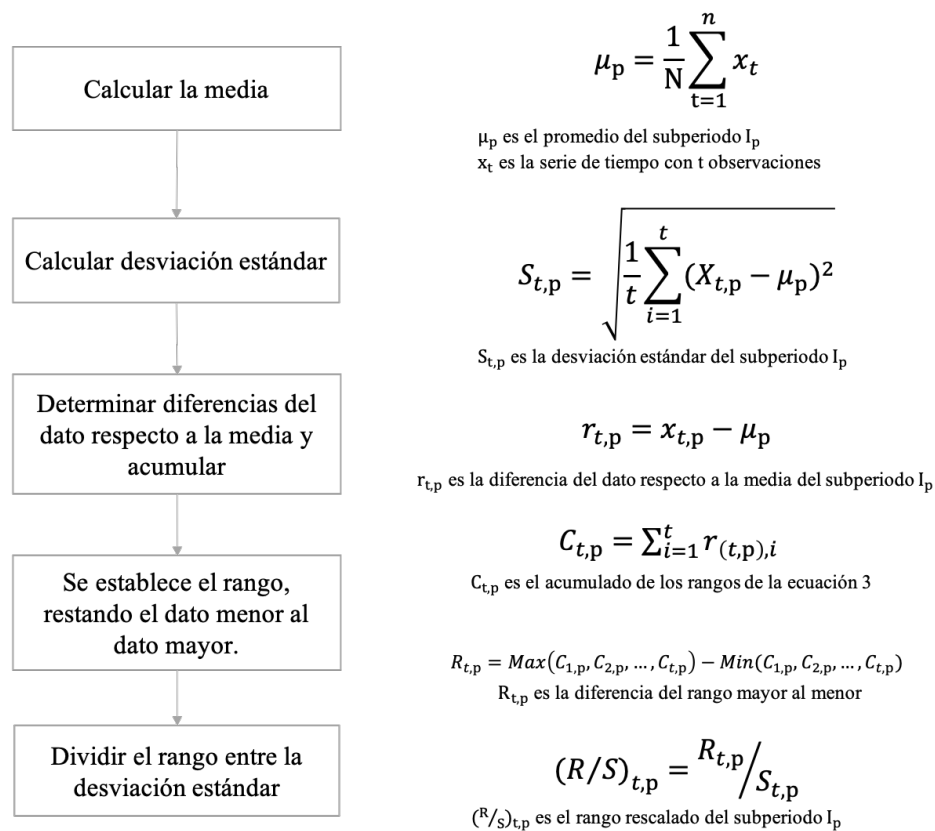


Figura 4.5 Cálculo del rango reescalado de cada partición.

Se calcula el Rango Reescalado con el promedio de todos los rangos para cada valor

de  $N$ .

Posteriormente, el exponente de Hurst se calcula despejando  $H$  de la ecuación 4.1.1 [74].

$$(R/S)_t = (\alpha * N)^H \quad (4.1.1)$$

donde:

$(R/S)_t$ ,	rango reescalado de toda la serie
$N$	número de observaciones
$\alpha$	constante ( $0 \leq \alpha \leq 1$ )
$H$	exponente de Hurst

Tomando logaritmos de 4.1.1; se obtiene la ecuación 4.1.2 donde  $H$  corresponde a la pendiente. Finalmente, la regresión con el mejor ajuste será el exponente de Hurst estimado.

$$\log_{10}(R/S)_N = H \log_{10}(N) + \log_{10}(\alpha) \quad (4.1.2)$$

#### 4.1.2. Algoritmo k-means

Dentro de los algoritmos de agrupamiento, existe una amplia categorización, sin embargo, en la clasificación de métodos de agrupamiento por particionamiento [75], el algoritmo k-means es uno de los más usados [76, 77] debido a su simpleza y eficiencia [78]. El algoritmo k-means intenta encontrar una partición de  $N$  observaciones en  $k$  grupos, de manera que cada observación pertenezca al grupo cuyo centro geométrico este más cercano. El algoritmo trata de minimizar la varianza total del conjunto de observaciones. Para esta experimentación, se aplica k-means como sigue:

- a) Determinar los valores del exponente de Hurst para los conjuntos de series M3 de Makridakis publicados [37].

- b) Obtener el mejor número  $k$  de grupos con el método elbow [79] y por experimentación.
- c) Identificar las series de tiempo de cada grupo, y calcular las medidas MAPE y sMAPE promedio para cada grupo.
- d) Graficar los puntos que corresponden a los valores promedio de MAPE y sMAPE calculados con los centroides del exponente de Hurst de cada grupo.

Las gráficas obtenidas por este procedimiento se utilizan para realizar un análisis cualitativo de la relación entre MAPE, sMAPE y Hurst, mismas que se presentan en la Sección 5: Análisis y Resultados.

## 4.2. Complejidad y desempeño en el pronóstico de series de tiempo

La entropía de Shannon en las series de tiempo, es una característica que ha sido ampliamente utilizada. En [80] se presenta un paquete para calcular las medidas de emergencia, auto-organización y complejidad (Emergence, Self-organization y Complexity, mejor conocido como ESC) basado en la entropía de Shannon. Tomando como base estas medidas, se realizó un estudio que permitiera determinar si la medida de complejidad aplicada a las series de M4-Competition influía en el pronóstico de las series de tiempo.

El conjunto de datos utilizado en este estudio, se divide en dos partes, la primera está compuesta por series sintéticas subdivididas en a) senos, b) mapa logístico y c) series generadas con la herramienta GRATIS [81]. Algunas de las series del inciso a, fueron corrompidas con una señal de ruido gaussiano. La segunda parte de los datos son las series de las competencias M4 de Makridakis, el conjunto completo es de 100,000 series, sin embargo, se consideraron solo series de tiempo cuyo número de observaciones fuese mayor o igual a 250.

Por otro lado, existe una amplia diversidad en las medidas de entropía, en este caso, se consideran solo cuatro de ellas: distribución, espectral, dos regímenes y permutación, aunque en el área de computación, la más utilizada es la espectral.

Una vez definidos los cuatro tipos de entropía y las series de tiempo a utilizar, fue necesario obtener cada una de las series en sus cuatro posibles caracterizaciones, por lo que en la Figura 4.6 se muestra de manera gráfica, la conversión de una series de tiempo a las cuatro medidas de entropía y con ello, posteriormente se determinan las medidas de auto-organización y complejidad (Self-Organization y Complexity) que pertenecen al framework ESC, obteniendo así un total de doce medidas o características de las series de tiempo.

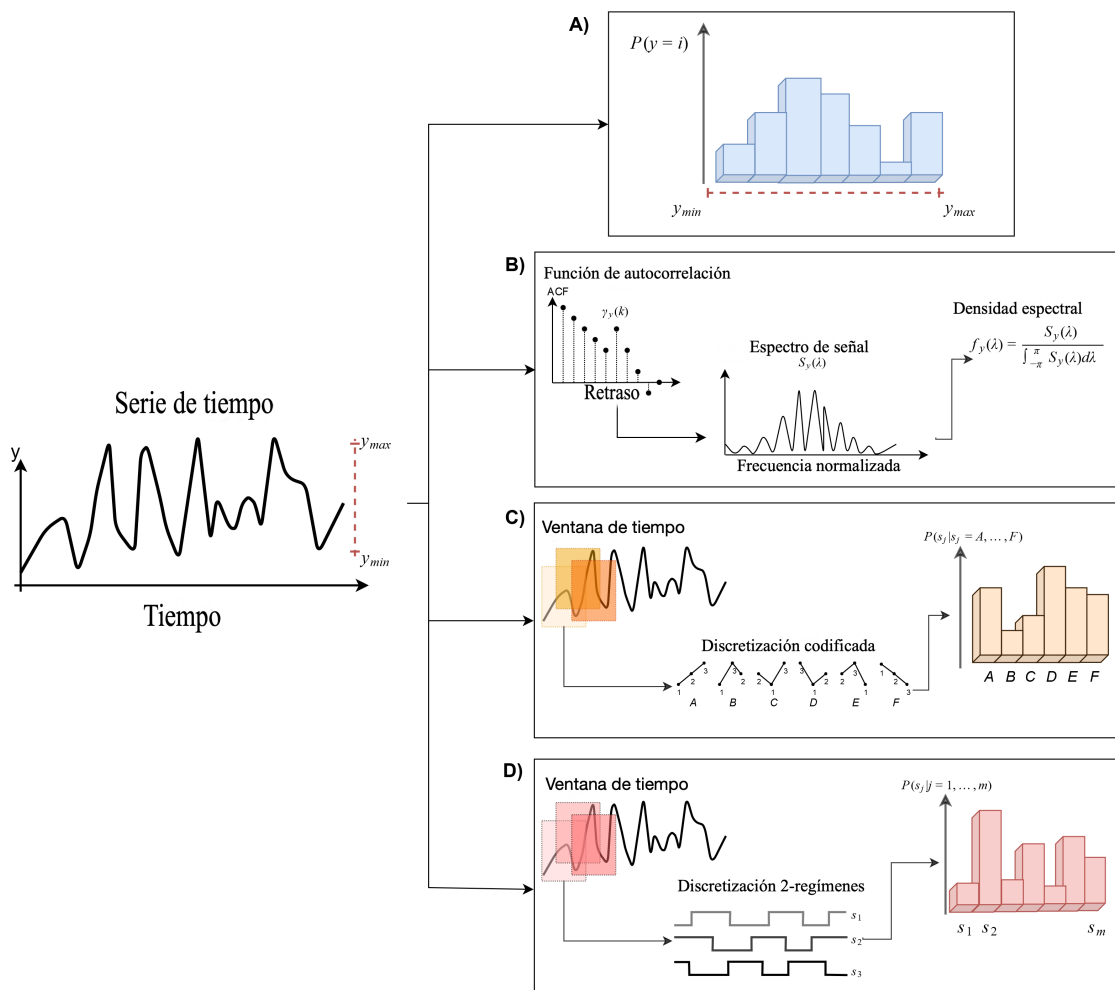
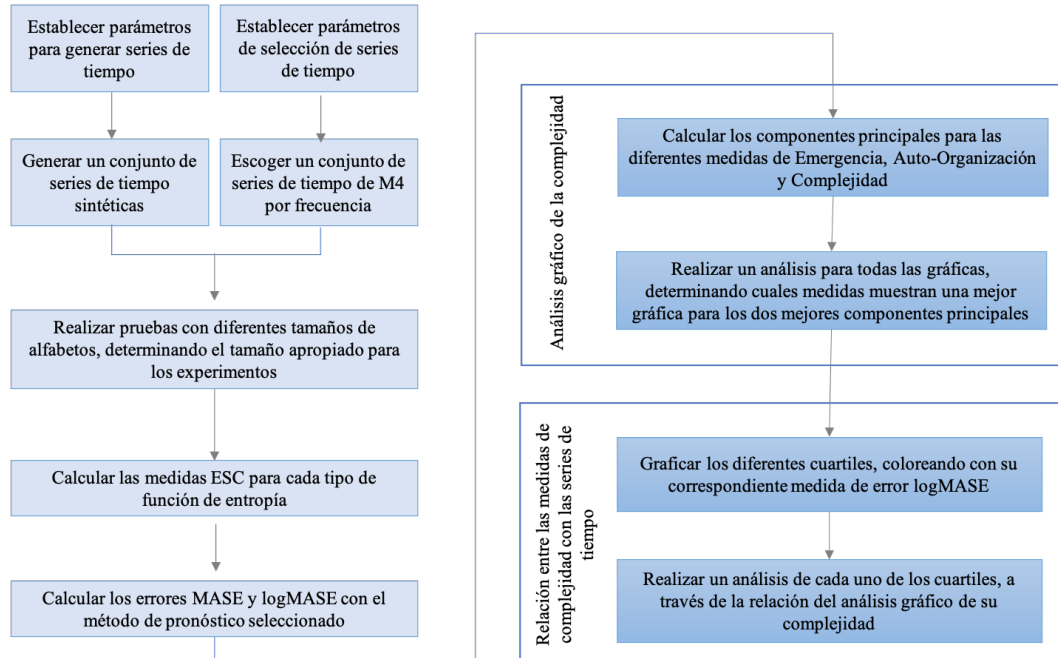


Figura 4.6 Posibles caracterizaciones de un sistema dinámico [5].

También se seleccionaron cuatro métodos de pronósticos utilizados en las competencias M4, los cuales fueron: Smyl, Theta, ARIMA y ETS (suavizamiento exponencial), a partir de los cuales se obtuvo la medida sMAPE para cada una de las series de tiempo del conjunto de

datos. La Figura 4.7 muestra de manera global, los pasos realizados para construir el espacio de características de complejidad (12 medidas) para analizar la relación entre el pronóstico de las series de tiempo y las medidas de complejidad calculadas en este estudio.



**Figura 4.7** Metodología del análisis de la relación entre el pronóstico y la complejidad de las series de tiempo [5].

Los resultados obtenidos aplicando este procedimiento, son presentados y explicados en la Sección 5: Análisis y Resultados.

### 4.3. Desarrollo de la metodología MetaLSS

Este trabajo de investigación propone el desarrollo de un método híbrido que permita pronosticar series financieras utilizando como base el método de suavizamiento exponencial. A esta metodología se le ha asignado el nombre de MetaLSS (MetaLearning Smoothing Scoring, en inglés). En el inicio de este trabajo de investigación, se experimentó con el conjunto de datos de las competencias de M3 [37], posteriormente se incluyó el conjunto de datos de las competencias de M4 [38]. La metodología que se presenta a continuación, muestra el uso específico en dos de los subconjuntos de datos de M4: mensual y diario (monthly y daily en

inglés), sin embargo, esta metodología se puede reproducir en cualquier conjunto de series de tiempo.

La metodología MetaLSS se muestra de manera general en la Figura 4.8 , el inciso *a* corresponde a la segmentación de las series de tiempo y se indica en cuales de las fases son utilizadas, mientras tanto, en el inciso *b* se pueden observar las 3 fases de la implementación del método híbrido propuesto.

La selección de las series de tiempo que serán utilizadas durante el proceso, es una parte fundamental de la metodología MetaLSS, la Figura 4.9 muestra los subconjuntos utilizados y su segmentación en porcentajes de 80-20 para entrenamiento y prueba, respectivamente.

Es importante mencionar que, la fase de prueba contiene implícitas dos etapas que se llevan a cabo para cada una de las series de tiempo que pertenecen al 20% del total de las series de tiempo.

1. Etapa de Preprueba: Series de tiempo que incluyen la división de entrenamiento y preprueba.
2. Etapa de Prueba: Series de tiempo que incluyen la división de entrenamiento, preprueba y prueba.

Las cuales se revisan con mayor detalle durante la descripción de cada una de las fases.

#### **4.3.1. Fase de entrenamiento**

En la fase de entrenamiento del modelo, participan  $n$  métodos de pronósticos utilizados en las competencias de M4 [38], algunos de ellos son métodos clásicos y comúnmente utilizados en el área de pronósticos, mientras que otros se seleccionaron por haber obtenido los primeros lugares en la competencia, además de tener acceso al código y hacer fácil su reproducción. Es importante mencionar que cualquier método que cumpla con los requisitos de accesibilidad y reproducción, puede ser agregado a este subconjunto de métodos de pronósticos.

Los métodos que componen el conjunto de  $n$  métodos, se presentan en la Tabla 4.2.



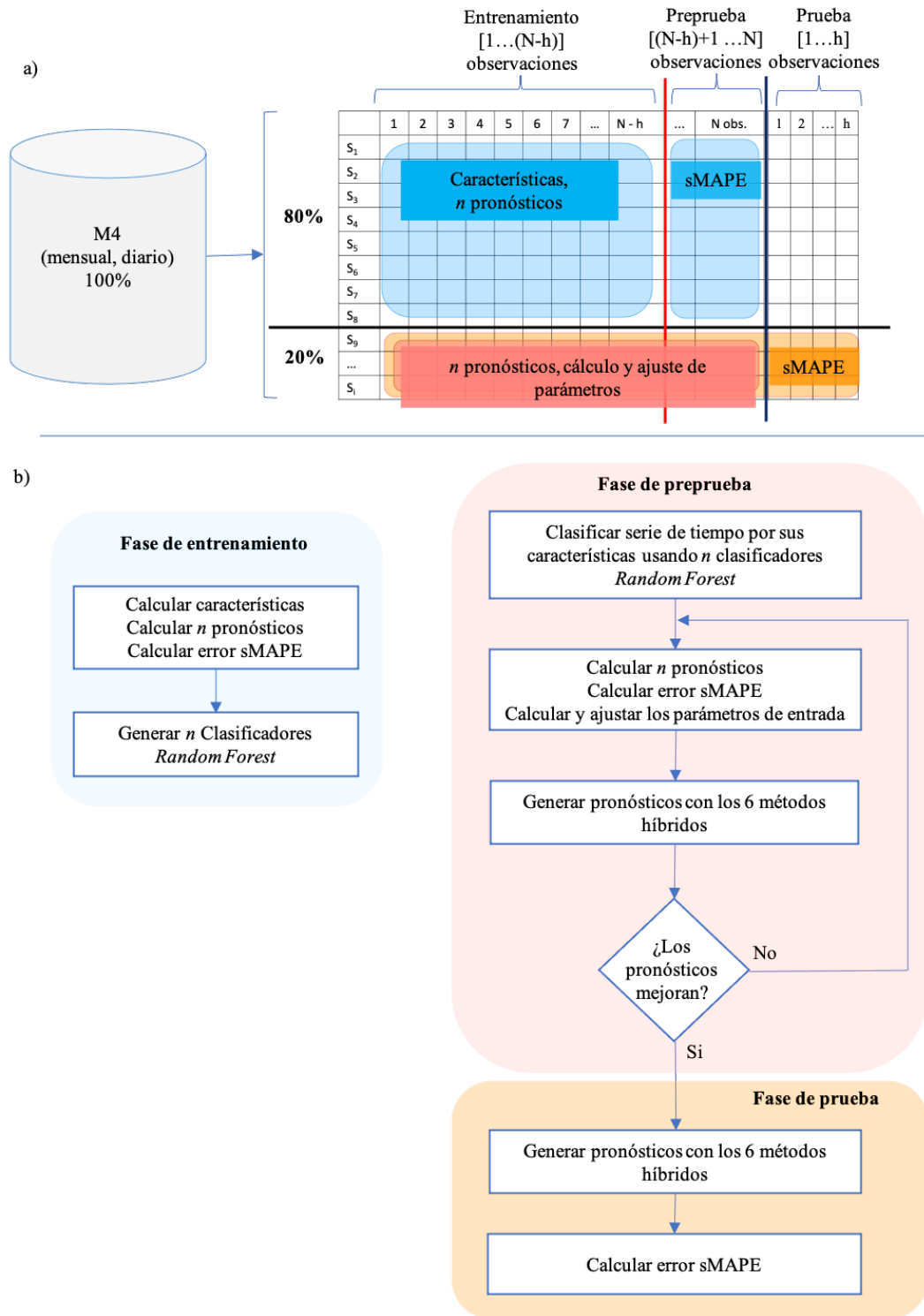
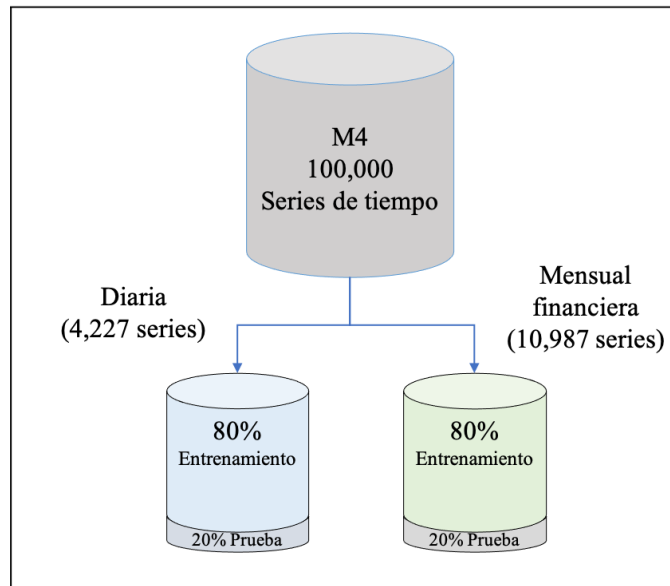


Figura 4.8 Esquema general, segmentación y fases de la metodología MetaLSS.



**Figura 4.9** Selección de subconjuntos de datos de M4-Competition.

**Tabla 4.2** Métodos de pronósticos seleccionados de las competencias M4.

Identificador	Tipo	Autor(es)	Afiliación	Posición
245	Combination	Montero-Manso, et al.	University of A Coruña & Monash University	2
72	Combination	Jaganathan & Prakash	Individual	4
Theta	Statistical	Theta - Benchmark		18
Com	Combination	Comb - Benchmark		19
Damped	Statistical	Damped - Benchmark		22
Holt	Statistical	Holt - Benchmark		33
SES	Statistical	SES - Benchmark		34
Naive2	Statistical	Naive2 - Benchmark		37
Naive	Statistical	Naïve - Benchmark		41
sNaive	Statistical	sNaive - Benchmark		43

Las series de tiempo de M4, en su versión original, están divididas en un subconjunto de entrenamiento, compuesto por  $N$  observaciones y un subconjunto de prueba de tamaño  $h$ , que corresponde al horizonte para cada una de las frecuencias (mensual, diaria, etc) de los datos que se van a pronosticar. En el método desarrollado, cada una de las series de tiempo en su subconjunto de entrenamiento, se subdividen en *entrenamiento* y *preprueba*, de tal manera que, el tamaño de la serie en su subconjunto de entrenamiento es  $[1, \dots, N-h]$  observaciones, la preprueba es  $[N-h+1, \dots, N]$  observaciones mientras que, el subconjunto de *prueba* es  $[1, \dots, h]$  observaciones, el cual se mantiene igual que la serie original.

Una vez subdivididas las series de tiempo:

- Se extraen sus características o atributos [66, 82].
- Se pronostica con los  $n$  métodos participantes.
- Se obtiene el sMAPE de cada uno y se ordenan de menor a mayor valor de sMAPE.
- Para cada una de las series de tiempo, se mantiene el orden de los métodos que obtuvieron los primeros tres lugares.
- A partir del cuarto lugar se asigna un mismo orden ( $n$ ) a los métodos.
- Con las características y el orden que obtuvieron cada uno de los  $n$  métodos, se generan  $n$  modelos clasificadores, aplicando la técnica de aprendizaje automático *Random Forest*.

La Figura 4.10 muestra un diagrama de flujo con los pasos que se realizan en la fase de entrenamiento y la generación de los modelos clasificadores utilizando el 80% de las series de tiempo de la frecuencia de los datos seleccionada.

#### 4.3.2. Fase de prueba

La fase de prueba consiste, en utilizar el 20% de las series de tiempo, cuyas características no son conocidas por los modelos generados con el método *Random Forest*. Como se mencionó en la sección previa, se generaron  $n$  modelos clasificadores, los cuales a partir de las características de las series de tiempo del subconjunto de prueba, retornarán un vector resultante de tamaño  $n$ , donde cada elemento corresponde a la posición que obtuvo el método ( $m_i$ ) para un tipo de serie en específico, los posibles valores son  $1, 2, 3$  y  $n$ , donde:

- Uno (1) indica que el método de pronóstico obtuvo el primer lugar.
- Dos (2) el segundo lugar.
- Tres (3) el tercer lugar.

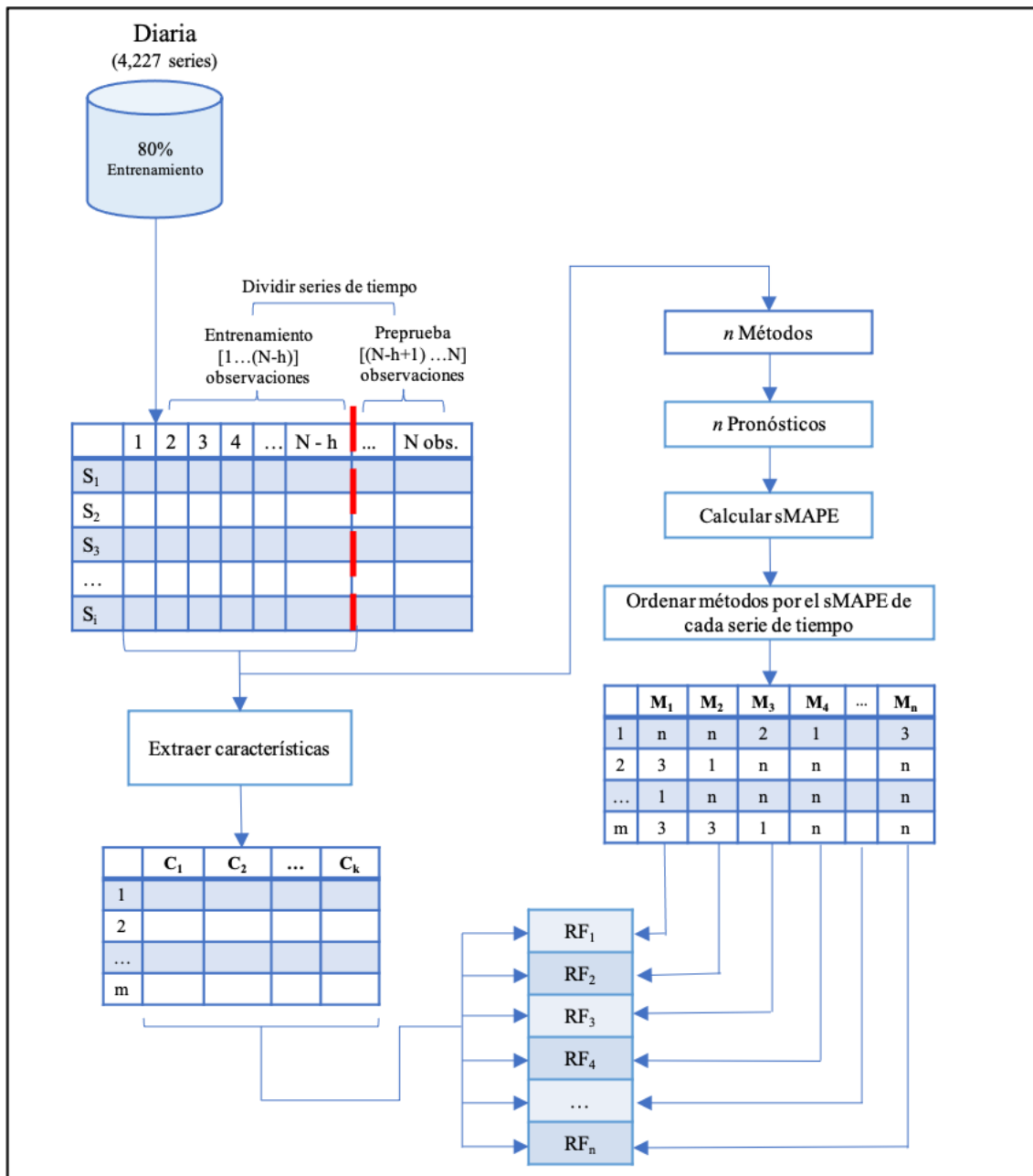


Figura 4.10 Fase de entrenamiento de MetaLSS.

- n (variable que se sustituye por el número total de métodos de pronósticos participantes), indica que el método no es adecuado para resolver la serie de tiempo que se está analizando.

La Tabla 4.3 muestra seis posibles casos que se pueden presentar, los cuales se describen a

continuación:

- Caso 1: existe más de un método de pronóstico en primero, segundo o tercer lugar.
- Caso 2: solo existe un método en primero, uno en segundo y uno en tercer lugar.
- Caso 3: existen únicamente primeros lugares.
- Caso 4: existen solamente segundos lugares.
- Caso 5: existen solo terceros lugares.
- Caso 6: no existen métodos adecuados para pronosticar la serie de tiempo.

**Tabla 4.3** Casos potenciales de los clasificadores.

n Métodos						
Caso	$m_1$	$m_2$	$m_3$	$m_4$	...	$m_n$
1	1	n	1	1	...	n
2	1	n	3	2	...	n
3	1	n	n	1	...	n
4	n	2	n	2	...	2
5	3	3	3	n	...	n
6	n	n	n	n	...	n

En la Figura 4.11 se muestra un diagrama de flujo con los pasos iniciales que se llevan a cabo con las series de tiempo del subconjunto de prueba, los cuales forman parte de la metodología MetaLSS. Mientras que en la Figura 4.12 se visualizan dos de los principales *parámetros de entrada* que se obtienen en el proceso mostrado previamente en la Figura 4.11, estos dos parámetros se obtienen de manera individual para cada una de las series de tiempo que pertenecen al conjunto de prueba:

- Inciso **a)**, las series de tiempo se subdividen en *entrenamiento, preprueba y prueba*.
- Inciso **b)**, se asigna un porcentaje de participación en la elaboración del pronóstico durante la etapa de preprueba a los métodos que resultaron ser los más adecuados para pronosticar las series en su fase de entrenamiento.

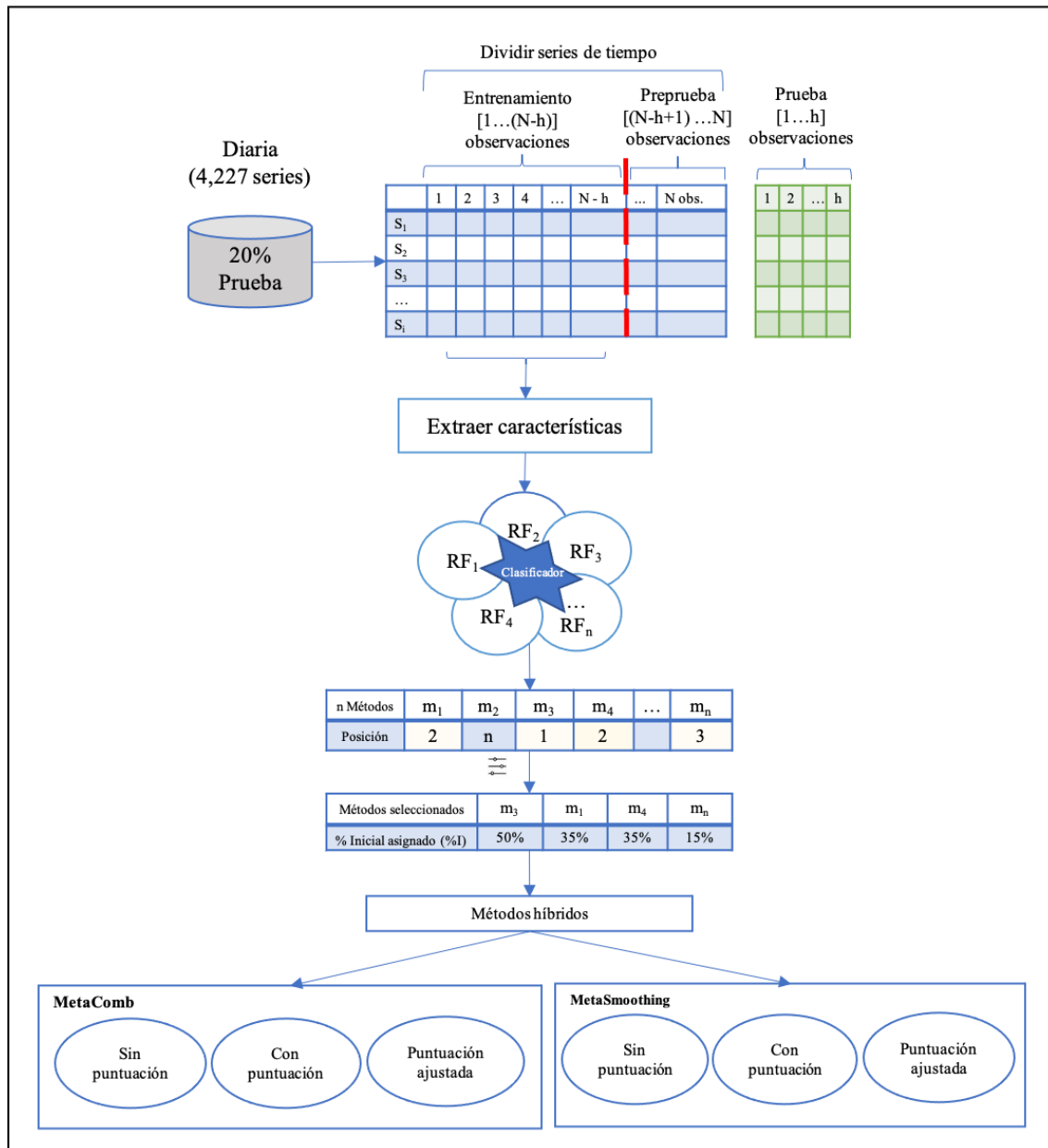


Figura 4.11 Fase de prueba de MetaLSS.

Inicialmente se asigna de manera arbitraria un 50% a los métodos que obtuvieron el primer lugar, un 35% a los métodos que obtuvieron el segundo lugar y un 15% a los métodos que obtuvieron el tercer lugar. En el caso de que la suma de los porcentajes exceda el 100%, estos serán ajustados de manera proporcional con los porcentajes asignados de manera inicial utilizando la ecuación 4.3.1.

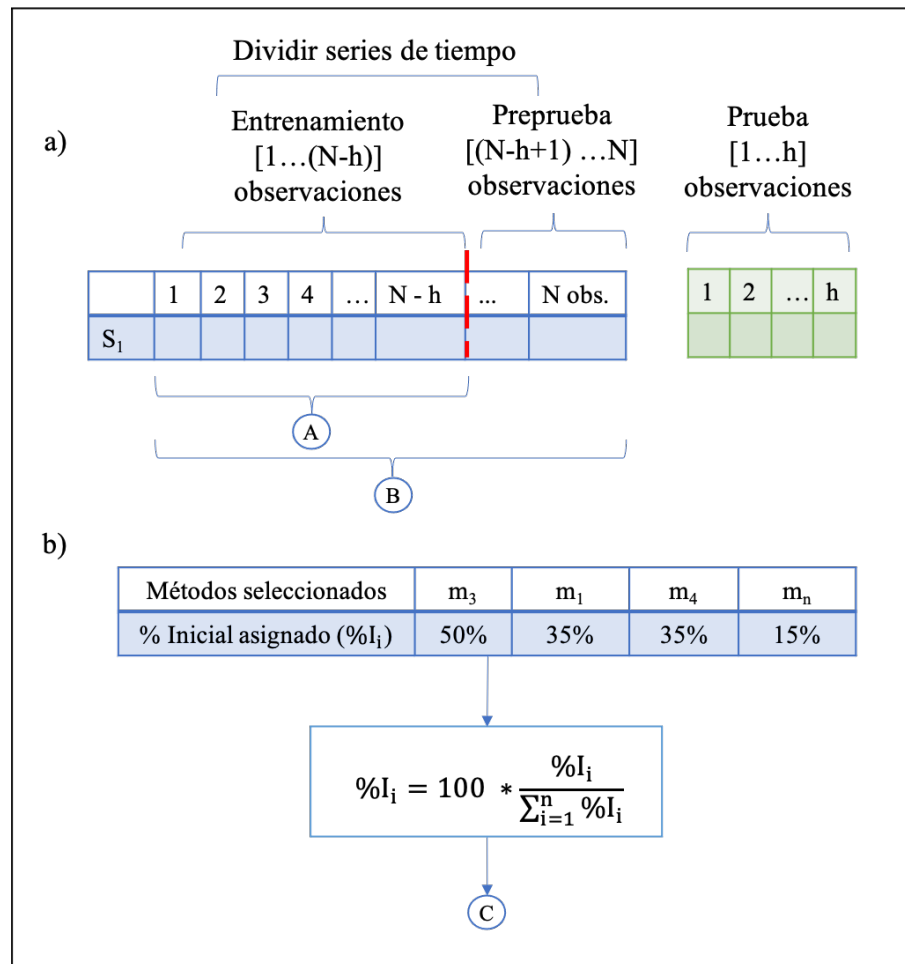


Figura 4.12 Parámetros de entrada de MetaLSS.

Como se mencionó anteriormente, es posible que no se encuentren métodos adecuados para determinadas series de tiempo, en ese caso, se asigna de manera inicial un 5%, el cual será ajustado por los primeros tres métodos que tengan un sMAPE más pequeño en el entrenamiento y un pronóstico de la preprueba.

$$\%I_i = 100 * \left( \%I_i / \sum_{i=1}^n \%I_i \right) \tag{4.3.1}$$

Durante el desarrollo de la metodología MetaLSS, surgieron varias alternativas en el ensamble de los métodos de pronóstico, dando como resultados seis métodos híbridos: *MetaComb*, *MetaComb-S*, *MetaComb-SwA*, *MetaSmoothing*, *MetaSmoothing-S*, *MetaSmoothing-SwA*, los cuales se presentan en el orden de desarrollo y que van desde lo más básico a los más

complejo en su implementación. La Tabla 4.4 muestra un resumen de los seis métodos desarrollados.

**Tabla 4.4** Métodos híbridos desarrollados.

	Ensamble de métodos	Puntuación	Puntuación ajustada	Suavizamiento	Suavizamiento ajustado
MetaComb	✓				
MetaComb-S	✓	✓			
MetaComb-SwA	✓	✓	✓		
MetaSmoothing	✓			✓	
MetaSmoothing-S	✓	✓		✓	
MetaSmoothing-SwA	✓		✓	✓	✓

La ejecución de estos seis métodos desarrollados, se realiza en dos ocasiones para cada una de las series de tiempo que conforman el 20% del subconjunto de fase de prueba:

1. Series de tiempo que comprenden el entrenamiento y la preprueba.
2. Series de tiempo que comprenden el entrenamiento, preprueba y prueba.

#### **Método híbrido: MetaComb**

El método MetaComb surge de una combinación de los mejores métodos de pronósticos que son seleccionados con base al uso de  $n$  modelos clasificadores generados con el método Random Forest. La combinación se lleva a cabo con base a porcentajes asignados de manera permanente, el (los) método (s) de pronósticos que hayan alcanzado el primer lugar, obtendrán un 50% de participación, mientras que el (los) método (s) que hayan quedado en segundo lugar, tendrá (n) asignado un 35% y por último los que hayan obtenido el tercer lugar, se les asignará un 15%. Para los casos donde no haya métodos adecuados para una determinada serie de tiempo, se consideran los primeros tres métodos que tengan una menor medida de error sMAPE y se les asigna el porcentaje acorde con esta medida. La Figura 4.13 muestra el proceso que se realiza para generar un nuevo pronóstico utilizando este enfoque, los conectores B y C del diagrama de flujo, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12.



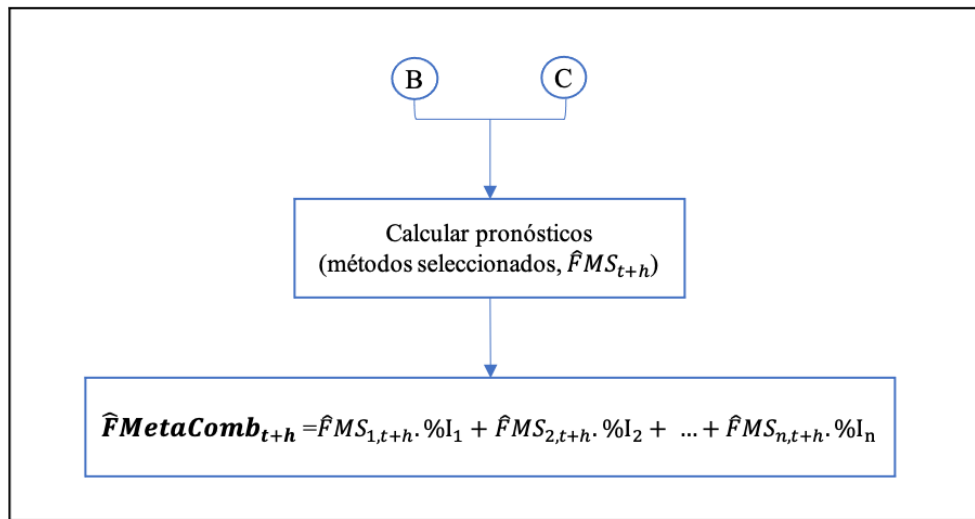


Figura 4.13 Método híbrido: MetaComb.

### Método híbrido: MetaComb-S

Este método es una variante del método anterior, la diferencia consiste en utilizar un Scoring (puntuación), el cual es llamado bat-scoring, esto debido a que los porcentajes iniciales (50%, 35%, 15% y 5%), serán modificados hacia un punto en el espacio de porcentajes; siempre y cuando el pronóstico sea mejorado, de tal manera que se establecen porcentajes promedios asignados de manera equitativa entre el número de métodos de pronósticos que participarán en la combinación final de método. La Figura 4.14 muestra el diagrama de flujo con los pasos desarrollados para este enfoque, los conectores A, B y C, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12.

### Método híbrido: MetaComb-SwA

MetaComb-SwA se diferencia del método anterior en la actualización de los porcentajes de combinación de los métodos de pronósticos que participan en tal combinación. El procedimiento a seguir es similar al del método híbrido MetaComb-S, descrito en la subsección anterior, pero los pesos promedio iniciales se asignan basándose en la medida de error sMAPE de cada uno de los métodos de pronósticos que se utilizarán. De tal manera que el método de pronóstico que tiene un error de sMAPE mayor tendrá un porcentaje promedio

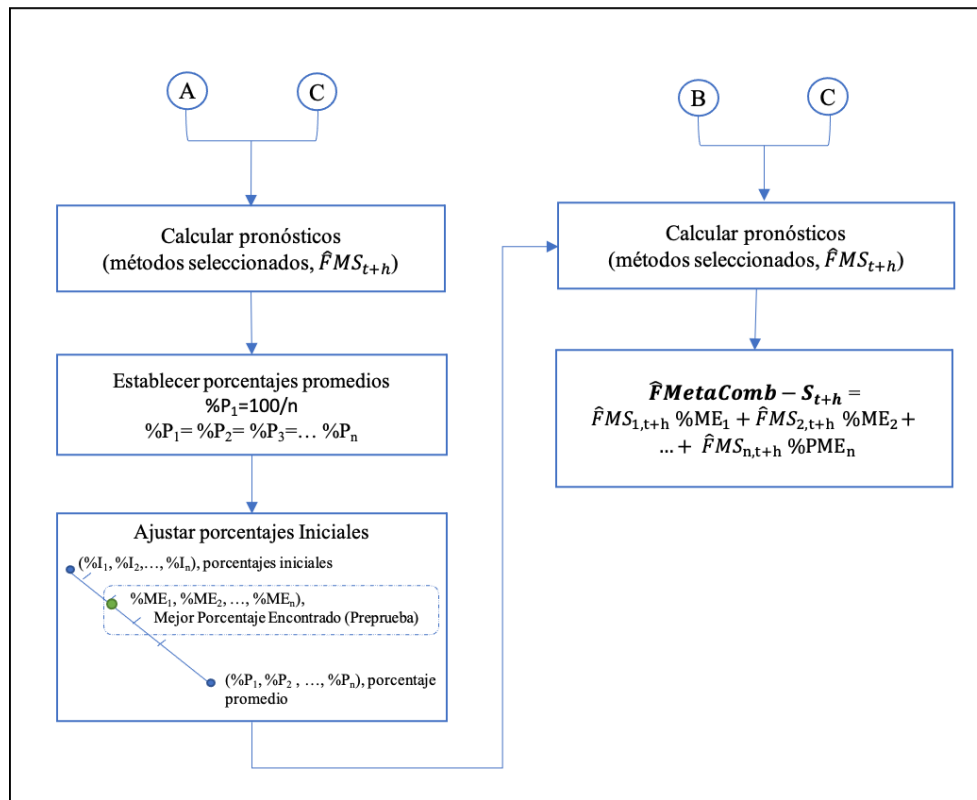


Figura 4.14 Método híbrido: MetaComb-S.

inicial más pequeño y el que tenga el error sMAPE más pequeño, tendrá un porcentaje inicial promedio más grande. Estos cálculos se pueden realizar con la Ecuación 4.3.2, 4.3.3 y 4.3.4.

$$w_i = sMAPE_{max} + sMAPE_{min} - X_i \quad i \leq i \leq n \quad (4.3.2)$$

donde,

$sMAPE_{max}$  y  $sMAPE_{min}$  son el máximo y mínimo error sMAPE de los métodos de pronósticos seleccionados, respectivamente,  $X_i$  es el error sMAPE de  $i$  método de pronóstico.

$$T_w = \sum_{i=1}^n w_i \quad (4.3.3)$$

$$\%A_i = w_i / T_w \quad (4.3.4)$$

Los porcentajes iniciales promedio calculados son solo una referencia, no son los definitivos,

ya que hay porcentajes iniciales de 50%, 35% y 15% que se estarán ajustando hacia los porcentajes iniciales promedios y a través del error sMAPE con los pesos ajustados, se definirá finalmente cuáles serán los valores de los porcentajes finales.

La Figura 4.15 muestra un diagrama de flujo con los pasos desarrollados para este método, los conectores A, B y C, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12..

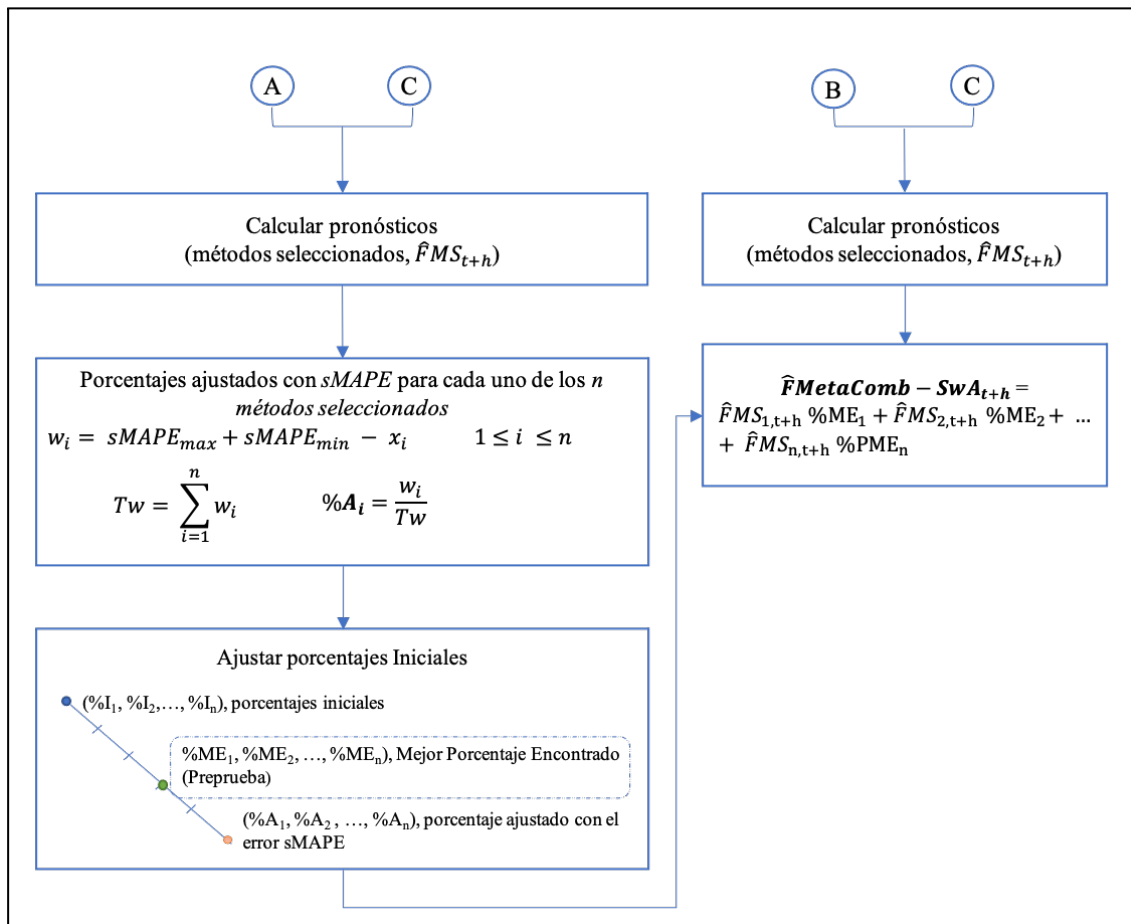


Figura 4.15 Método híbrido: MetaComb-SwA.

Los tres métodos descritos (MetaComb, MetaComb-S y MetaComb-SwA) realizan ensambles con varios métodos de pronóstico, y su principal objetivo es determinar cuál es el porcentaje de participación de cada uno de los métodos seleccionados como los mejores para la serie de tiempo que se pronosticará.

Los siguientes tres métodos híbridos incluyen, además, un método de suavizamiento expo-

nencial, en este caso, el método Holt [83] el cual se utiliza como parte complementaria en el ensamble de métodos. Las siguientes tres subsecciones describen cada uno de ellos con mayor detalle.

### **Método híbrido: MetaSmoothing**

El método MetaSmoothing surge con la idea de utilizar una variable de suavizamiento para el ensamble de los métodos. El método clásico de suavizamiento exponencial, descrito en la sección 2: Antecedentes, hace uso de una variable de suavizamiento  $\alpha$ , cuyo valor esta en el rango de  $0 \leq \alpha \leq 1$ . La Ecuación 4.3.5 corresponde al método de suavizamiento exponencial simple.

$$\hat{Y}_{t-1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t \quad (4.3.5)$$

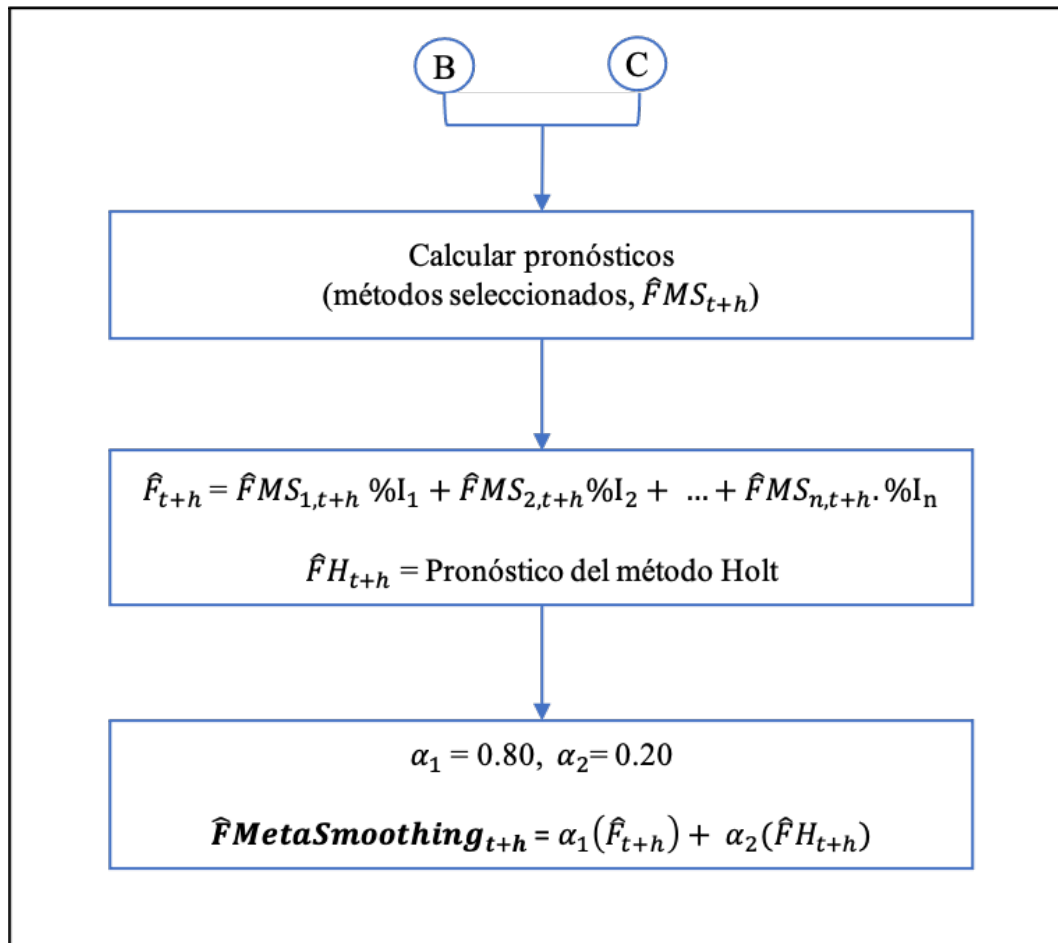
El método implementado asigna de manera arbitraria el valor de 0.80 para la variable de suavizamiento  $\alpha_1$  y un valor complemento de 0.20, el cual se denominará de aquí en adelante  $\alpha_2$ , la suma de las dos variables ( $\alpha_1, \alpha_2$ ) es igual a la unidad.

La combinación de los métodos de pronósticos seleccionados tendrá asignado un peso del valor de  $\alpha_1$ , mientras que el método de pronóstico Holt tendrá asignado un peso del valor de  $\alpha_2$ .

La Figura 4.16 muestra un diagrama de flujo con los pasos desarrollados para este método, los conectores B y C, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12.

### **Método híbrido: MetaSmoothing-S**

El método Meta-Smoothing-S es una variante del método anterior. Se agrega un scoring (puntuación), el cual consiste en determinar los porcentajes de los pesos que se le asignarán a cada uno de los métodos de pronósticos seleccionados, inicialmente los porcentajes asignados son del 50%, 35% y 15% acorde con el resultado de los modelos de clasificación Random Forest utilizados, por otro lado, los pesos promedio hacia donde se realizará la



**Figura 4.16** Método híbrido: *MetaSmoothing*.

búsqueda de mejores porcentajes, será un porcentaje fijo que está en función del número de métodos de pronósticos seleccionados (ejemplo:  $100\%/5 = 20\%$ ).

Se hace uso de la variable de suavizamiento  $\alpha$ , la cual como se mencionó en el método anterior, se subdivide en  $\alpha_1$  y  $\alpha_2$ , donde la suma ambos valores serán el valor de la unidad.

La Figura 4.17 corresponde al diagrama de flujo del método híbrido desarrollado, incluyendo puntuación y el método de pronóstico Holt, los conectores A, B y C, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12.

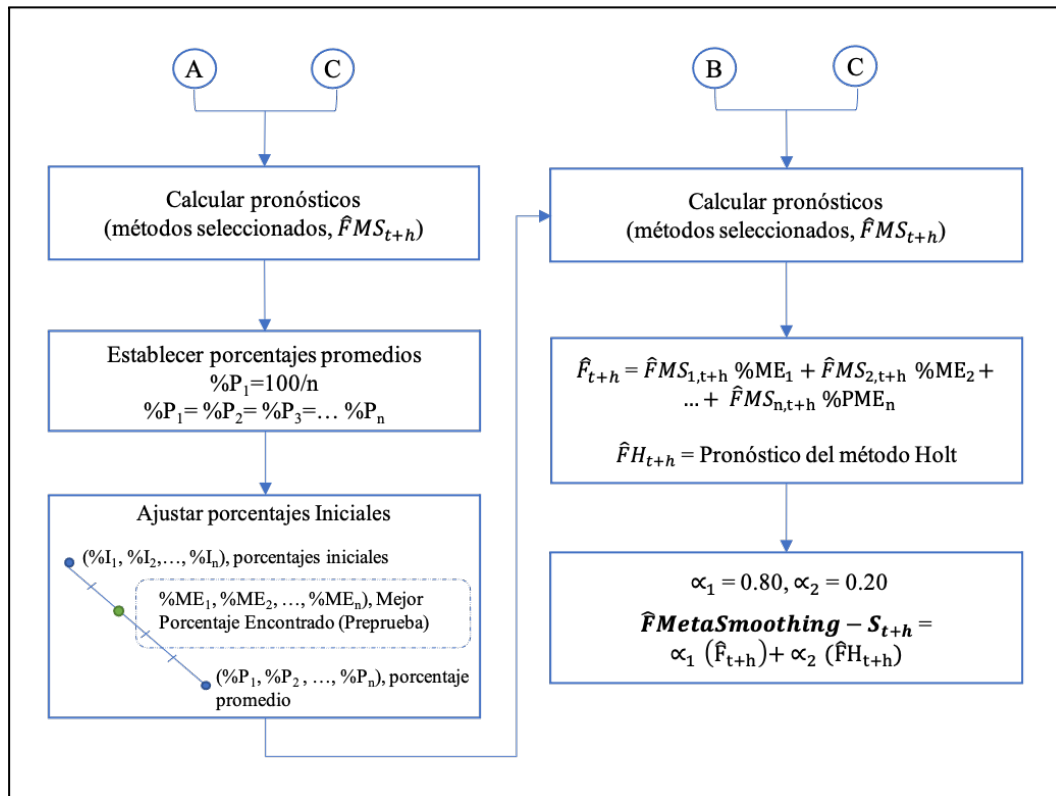


Figura 4.17 Método híbrido: MetaSmoothing-S.

### Método híbrido: MetaSmoothing-SwA

El método MetaSmoothing-SwA es el resultado final del ensamble de los métodos seleccionados como los más adecuados para las características de la serie de tiempo que se desea pronosticar, este método incluye una puntuación (scoring) con pesos adaptados, utilizando unos pesos promedio asignados acorde con los errores SMAPE de cada uno de estos métodos de pronósticos, donde los mayores pesos se le asignan a los pronósticos que tienen un error más pequeño y los menores pesos se asignan a los pronósticos que tienen un error más grande. Además, se incluye una variable de suavizamiento  $\alpha$ , subdividida en  $\alpha_1$  y  $\alpha_2$ , donde la suma de ambos valores es igual a la unidad. Estos valores de  $\alpha_1$  y  $\alpha_2$  también son ajustados tratando de buscar la mejor combinación de valores para obtener un buen pronóstico. Los valores iniciales asignados son  $\alpha_1 = 0.80$  y  $\alpha_2 = 0.20$ . La variable  $\alpha_1$  suaviza la combinación de pronósticos de los métodos seleccionados, mientras que la variable  $\alpha_2$  suaviza el pronóstico de un método seleccionado de manera arbitraria, en este caso se

seleccionó el método Holt, pero es posible utilizar el método con el menor error sMAPE de todos los métodos seleccionados por los modelos clasificadores *Random Forest*.

La Figura 4.18 corresponde al diagrama de flujo con los pasos desarrollados para el método MetaSmoothing-SwA, el cual incluye una puntuación ajustada a los errores sMAPE de los métodos de pronósticos seleccionados, con dos variables de suavizamiento de los pronósticos que componen el pronóstico final, los conectores A, B y C, corresponden a las variables que se obtienen a partir de los parámetros de entrada de la Figura 4.12..

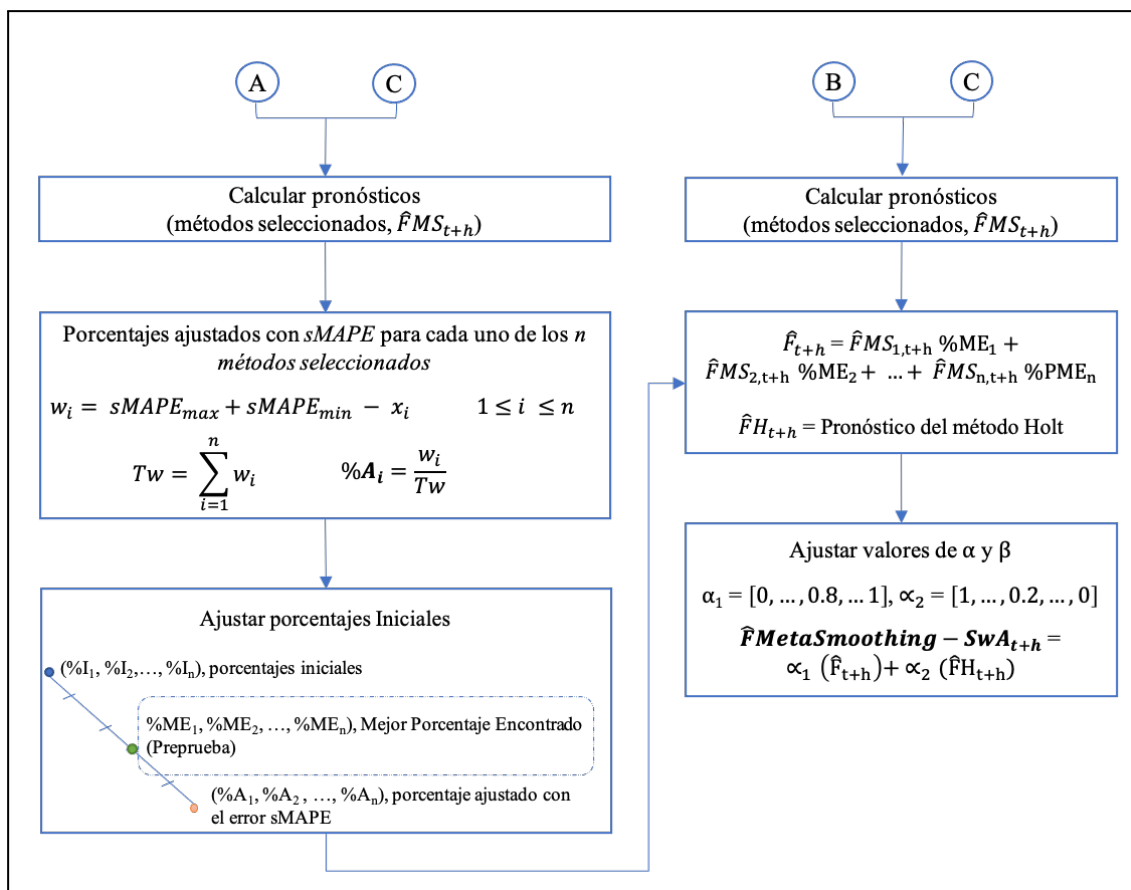


Figura 4.18 Método híbrido: MetaSmoothing-SwA.

Es importante mencionar que la fase de prueba, en su etapa de preprueba, permitió ajustar parámetros tales como los porcentajes de participación de los métodos de pronósticos que intervienen en el pronóstico final, de tal manera que, una vez teniéndolos, se utiliza la serie de tiempo identificada como entrenamiento para ejecutar los seis métodos de pronósticos

nuevamente y obtener su medida de error sMAPE, la cual es comparada con las medidas de error sMAPE de los  $n$  métodos de pronósticos iniciales participantes.

Los resultados de la experimentación realizada, se muestran en la Sección 5: Análisis y Resultados.



---

## Equipos y conjuntos de datos

En esta Sección, se incluye una descripción de los equipos y conjuntos de datos utilizados durante la experimentación realizada con el método híbrido propuesto. Este trabajo de investigación se enfoca en la generación de pronósticos de buena calidad y con un error SMAPE pequeño, por lo que el equipo de cómputo y sus características no juegan un papel relevante, sin embargo, se debieron cumplir algunos requisitos como reproducir el entorno de desarrollo en el lenguaje R con las librerías necesarias para la ejecución de los métodos de pronósticos utilizados. Por otro lado, el conjunto de datos es de total relevancia ya que existe una diversidad de series de tiempo, aunque no todas están reguladas o evaluadas, por lo que, las series de tiempo de las competencias de Makridakis M4, se seleccionaron debido a que proporcionan series de tiempo revisadas y evaluadas, clasificadas como benchmarks para realizar pruebas de métodos de pronósticos, así mismo proporcionan los pronósticos de los 25 mejores métodos participantes en la competencia, lo cual permite evaluar de manera confiable y precisa el o los métodos propuestos. El detalle del equipo, software y conjuntos

de datos utilizado durante la experimentación se detalla en la siguientes subsecciones.

## 5.1. Equipo de cómputo

La experimentación final de el método híbrido se realizó en un equipo de escritorio con las siguientes características.

Equipo: iMac (Retina 4K, 21.5-inch, 2019)  
Procesador: 3.2 GHz 6-Core Intel Core i7  
Memoria: 8 GB 2667 MHz DDR4  
Gráficos: Radeon Pro 560X 4 GB

y en una computadora portátil con las siguientes características:

Equipo: MacBook Air (Retina, 13-inch, 2018)  
Procesador: 1.6 GHz Dual-Core Intel Core i5  
Memoria: 8 GB 2133 MHz LPDDR3  
Gráficos: Intel UHD Graphics 617 1536 MB

## 5.2. Software utilizado

Software comercial como: MiniTab, C, R, Excel, PSPP (versión Mac) y herramientas gráficas entre otros.

## 5.3. Conjuntos de datos

Se decidió realizar experimentación con las series de las competencias de Makridakis, debido a que se proporcionan las series de tiempo de entrenamiento, de prueba y los pronósticos de los métodos participantes, en el caso del conjunto de M4 de los 61 métodos participantes, se incluyen únicamente los primeros 25 métodos de pronósticos acorde al lugar que obtuvieron. Los conjuntos de datos corresponden a series de tiempo de las competencias M3 y M4, donde el primer conjunto de series de tiempo, esta compuesto por 3,003 series de tiempo de la vida real, divididas en frecuencias de datos: anuales, cuarto de año, mensuales y otros, y el segundo conjunto de series lo integran 100,000 series de tiempo de la vida real divididos

en frecuencias de datos: anual, cuarto de año, mensual, diario, semanal y por hora. La Tabla 5.1 y 5.2 muestran la distribución de cada conjunto.

**Tabla 5.1** *Series de tiempo de la competencia M3.*

Frecuencia	Demográfica	Finanzas	Industria	Macro	Micro	Otra	Total
<b>Anual</b>	146	102	83	58	245	11	645
<b>Cuarto de año</b>	204	83	336	76	57	0	756
<b>Mensual</b>	474	334	312	145	111	52	1,428
<b>Otras</b>	4	0	0	29	0	141	174
<b>Total</b>	828	519	731	308	413	204	3,003

**Tabla 5.2** *Series de tiempo de la competencia M4.*

Frecuencia	Demográfica	Finanzas	Industria	Macro	Micro	Otra	Total
<b>Anual</b>	1,088	6,519	3,716	3,903	6,538	1,236	23,000
<b>Cuarto de año</b>	1,858	5,305	4,637	5,315	6,020	865	24,000
<b>Mensual</b>	5,728	10,987	10,017	10,016	10,975	277	48,000
<b>Semanal</b>	24	164	6	41	112	12	359
<b>Diaria</b>	10	1,559	422	127	1,476	633	4,227
<b>Hora</b>	0	0	0	0	0	414	414
<b>Total</b>	8,708	24,534	18,798	19,402	25,121	3,437	100,000

Una de las marcadas diferencias entre los conjuntos de datos de las competencias, es el número de series de tiempo, el segundo conjunto es más amplio y las observaciones de cada una de las series de tiempo, es mayor comparado con el primer conjunto de series de tiempo estudiado. En el caso de las competencias de M4, se tienen algunos métodos que se pueden reproducir debido a que los códigos utilizados por los participantes, son compartidos en el sitio web GitHub.

Debido al gran número de series de tiempo que conforman el conjunto de datos de las competencias de M4, se seleccionaron solo las frecuencias de datos *mensual* y *diario*, en el caso de la frecuencia *mensual* se consideraron las de tipo financieras, sumando un total de 10,987 series de tiempo, mientras que en las de la frecuencia *diaria* se incluyeron los 6 diferentes tipos con un total de 4,227 series de tiempo.

Es importante mencionar que para algunas series de tiempo seleccionadas, al intentar extraer sus características, no fue posible hacerlo, por lo que las series con características cuyos valores fueron iguales a NA, se excluyeron del subconjunto de datos final.

Finalmente, el número de series de tiempo del subconjunto de datos diario fue de 4,218 series de tiempo.

---

# Análisis y resultados

## 6.1. Pruebas preliminares

Las series de tiempo financieras pueden ser pronosticadas por diversidad de métodos que han sido propuestos a lo largo de la historia, este trabajo de investigación propone el desarrollo de la metodología MetaLSS que utiliza técnicas de metaaprendizaje y es basada en suavizamiento exponencial.

Unas de las pruebas preliminares consistieron en pronosticar series de tiempo financieras con el método ARIMA (Autoregression Integrated Moving Average), con la finalidad de obtener conocimiento de los métodos clásicos utilizados en el área de la econometría, así como el cálculo de algunas medidas de error en los pronósticos generados.

### 6.1.1. Pronósticos con métodos clásicos

La siguiente prueba se realizó utilizando solo una serie de tiempo del conjunto de M3 Competition, en la Figura 6.1 se muestra una gráfica de la serie de tiempo seleccionada de manera aleatoria y pronosticada a través del método ARIMA.

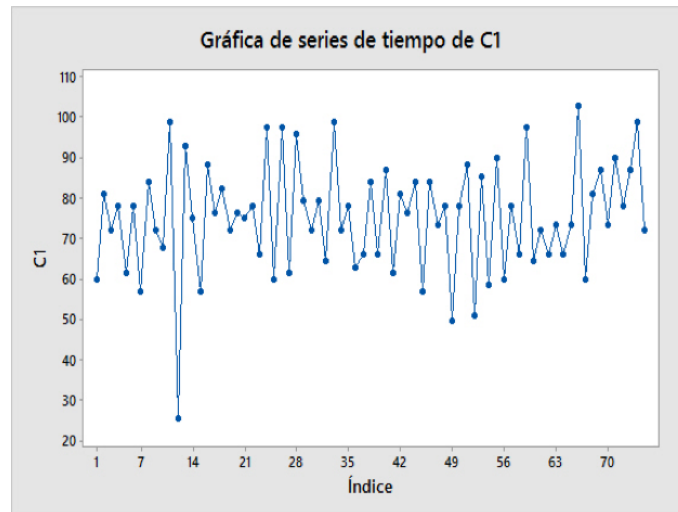


Figura 6.1 Ejemplo de la gráfica de una serie de tiempo.

El pronóstico fue realizado con la ejecución de un modelo ARIMA (1,0,0), aplicado a un horizonte de 5 observaciones.

### 6.1.2. Identificación de características de las series de tiempo

Como se mencionó anteriormente, las series de tiempo tienen algunas características que pueden medirse, algunas veces se les llama parámetros o atributos, existen algunos estudios en los cuales han realizado algún tipo de aprendizaje relacionado con dichos parámetros y los métodos de pronósticos para determinar que método se ajusta más acorde con los parámetros.

En [84] se mencionan las metacaracterísticas consideradas por la metodología MPro-oil para las series de tiempo de producción mensual de aceite y gas, las cuales son:

1. Exponentes de Hurst
2. Sesgo o Asimetría
3. Curtosis
4. Autocorrelación
5. Tendencia

6. Autocorrelación parcial
7. Porcentaje de cambio de paso o step change
8. Prueba de rachas
9. Porcentaje de cambios de dirección
10. No linealidad
11. Desviación estándar (SD)
12. Cociente de SD
13. Número de segmentos
14. Radio de segmentación

Mientras que en [82] y en [72] mencionan las características de las series de tiempo que consideraron en su investigación, las cuales son:

1. Tendencia
2. Estacionalidad
3. Periodicidad
4. Correlación
5. Sesgo o Asimetría
6. Curtosis
7. No linealidad
8. Autosemejanza (Exponente de Hurst)
9. Caos (Exponent Lyapunov)

En este trabajo de investigación, como fase preliminar, se inició con la caracterización de las series de tiempo financieras, considerando algunas de las características que mencionan los trabajos mencionados anteriormente. Las características que se consideraron son:

1. Exponente de Hurst
2. Sesgo o Asimetría
3. Curtosis
4. AutoCorrelación
5. Autocorrelación parcial
6. Tendencia
7. Desviación estándar

El método de aprendizaje supervisado utilizado es un árbol de decisión llamado J48, también conocido como ID3, C4.5 ó C5.0, estos métodos se encuentran ya implementados en software como Weka, lenguaje R, entre otros; para efectos de familiarización y sencillez con el método, se decidió utilizar Weka versión 3.9.

Las series de tiempo utilizadas fueron tomadas del conjunto de datos de las competencias de Makridakis: M-Competitions [35], M2-Competitions [36], M3-Competitions [37], NN3-Competitions [85] y NN5-Competitions [86].

### **Primera prueba de categorización, utilizando el árbol de decisión J48**

Para realizar la primera prueba de la caracterización de las series de tiempo, se llevaron a cabo los siguientes pasos:

1. Escoger aleatoriamente 42 series de todas las competencias mencionadas, estas series se utilizarán para entrenar y probar el método J48 en el software Weka 3.9.
2. Calcular los 7 parámetros o características de las 42 series seleccionadas en el punto 1, normalizar los datos de sesgo, curtosis y desviación estándar.



3. Categorizar las series, en Fácil, Regular y Difícil, acorde con los errores reportados en las competencias.
4. Generar los archivos de entrenamiento y de prueba con el formato que utiliza el software Weka, es decir, archivos con extensión .arff.
5. Entrenar, probar y categorizar las series.
6. Comparar los resultados obtenidos contra los deseados.
7. Establecer conclusiones: En esta primera prueba, se obtuvo un 100% de error en la categorización, la posible razón es que se construyó el árbol de decisión con series de datos de diversas competencias y 4 de las series de las que se obtuvo la categorización formaban parte de un conjunto con el que no se construyó el árbol de decisión; por lo que la siguiente prueba la conformaron solo series de una sola competencia.

### **Segunda prueba de categorización, utilizando el árbol de decisión J48**

Debido a los resultados obtenidos en la primera prueba preliminar, se planteó utilizar solo series de tiempo del conjunto de datos de la competencia M3-Competitions de Makridakis, esto debido a que es la competencia que tiene un detalle más completo de los resultados y se encuentra estandarizada en cuanto al manejo de estos.

Para realizar la segunda prueba de la caracterización de las series de tiempo, se llevaron a cabo los siguientes pasos:

1. Escoger aleatoriamente 31 series de tiempo del conjunto de M3-Competitions, estas series se utilizarán para entrenar y probar el árbol de decisión J48 en el software Weka 3.9.
2. Calcular los 7 parámetros o características de las 31 series seleccionadas en el punto 1, normalizar los datos de sesgo, curtosis y desviación estándar.
3. Categorizar las series, en Fácil, Regular y Difícil, acorde con los errores reportados en las competencias *Mean Absolute Percentage Error (MAPE)*, las series se dividen en

los periodos de: Yearly, Quarter, Monthly y Other.

4. Generar los archivos de entrenamiento y de prueba con el formato que utiliza el software Weka, es decir, archivos con extensión .arff.
5. Entrenar, probar y categorizar las series.
6. Comparación de los resultados obtenidos contra los deseados.
7. Establecer conclusiones: En esta segunda prueba, se obtuvo un 50% de error en la categorización, la posible razón es que se está entrenando con 25 series de un conjunto de 3003 series, por lo que se plantea, seguir realizando pruebas con un mayor número de series y además utilizar otros métodos de clasificación que acorde con la literatura, prometen resultados más favorables.

### **6.1.3. Relación entre el exponente de Hurst vs. MAPE y sMAPE**

De los experimentos realizados, en uno de ellos se hizo uso de la técnica de agrupamiento tal cual como se mostró en la subsección 3.1. Este experimento toma cada subconjunto de datos anual, cuarto de año, mensual y otros (Year, Quarter, Month y Other) del conjunto de series de M3 y en cada uno de ellos realiza una agrupación utilizando el algoritmo k-means con el valor de H (método R/S y proceso H-ARFIMA) de cada serie. En cada subconjunto se obtiene el promedio del valor MAPE y sMAPE de cada grupo y para cada método de pronóstico. De acuerdo con el método elbow [79] el número de grupos adecuado fue entre tres y seis y finalmente el utilizado fue de seis. La Figura 6.2 muestra los resultados del algoritmo k-means para los 6 grupos creados. En estas gráficas se puede observar la relación entre el MAPE y sMAPE con el exponente de Hurst (método R/S) para las series estudiadas correspondientes a los cuatro subconjuntos de datos.

Como resultado, se observa que se muestra una tendencia negativa en la relación entre el exponente de Hurst (R/S) y el valor MAPE y sMAPE para los métodos SES y ARIMA (Figura 6.2), también es posible observar que cerca del valor 0.9 del exponente Hurst se eleva el error promedio de MAPE y sMAPE en tres de los cuatro subconjuntos. Otra característica

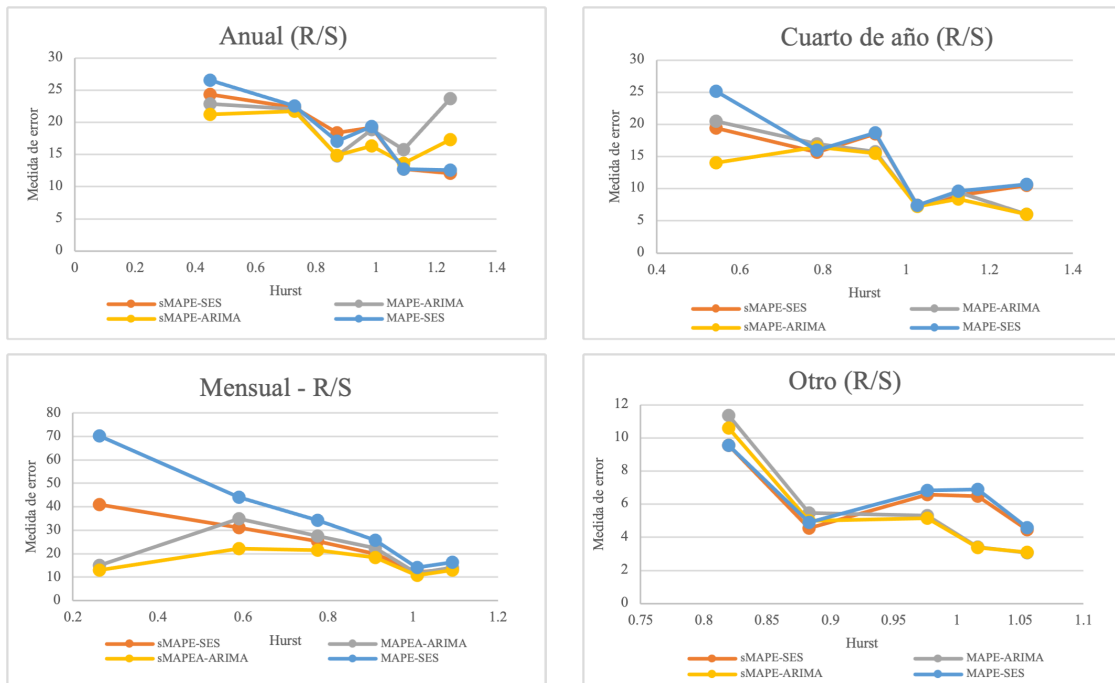


Figura 6.2 Relación Hurst vs MAPE y sMAPE usando R/S.

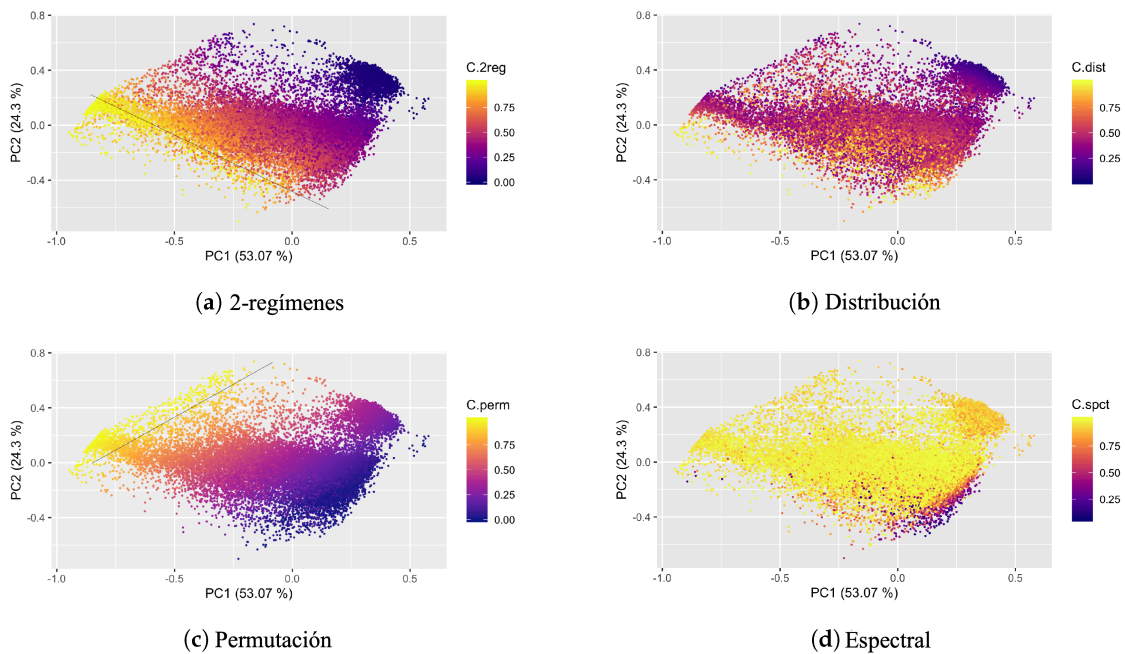
que se puede observar es una tendencia descendente para el MAPE, desde  $H$  mayor a cero hasta la unidad y dicha tendencia del MAPE se invierte cuando el valor del exponente de Hurst es mayor a uno. Este patrón es notorio para los subconjuntos de datos anual, cuarto de año, y mensual (Yearly, Quarterly y Monthly), mientras que para el subconjunto otros (Other) tal comportamiento no se presenta.

Las dos medidas de error que se consideraron muestran patrones similares, aunque se puede observar que el sMAPE se mantiene más estable durante la experimentación.

Se puede observar también que conforme el valor del exponente Hurst es cercano a cero, el MAPE es mayor para los dos métodos de pronóstico. De tal forma que se presenta una fuerte correlación inversa entre  $H$  y MAPE, para los valores de  $H$  menores a 1. En contraste, para valores de  $H$  mayores a uno, esa situación no se presenta. Los detalles de estos resultados pueden ser consultados en [4].

### 6.1.4. Resultados de complejidad en series de tiempo

En la experimentación realizada, se hizo uso del análisis de los componentes principales (PCA, Principal Component Analysis), donde fue posible identificar en dos dimensiones, cuáles de las doce medidas del espacio de características proporciona mayor información, por lo que el estudio se enfocó en las medidas de complejidad calculadas para todos los periodos de las series de tiempo de M4-Competitions. La Figura 6.3 muestra el PCA de las cuatro medidas de entropía. En esta figura fue posible observar que las medidas de complejidad de permutación (C.perm) y 2-regímenes (C.2Reg) son las medidas que aportan más información del sistema y específicamente para las series sintéticas permiten observar una clara agrupación acorde con el error de pronóstico y su medida de complejidad, esto se puede observar en la Figura 6.4.



**Figura 6.3** Medidas de complejidad basadas en entropía y análisis de componentes principales.

Por otro lado, para las series de M4, se encontró que de los cuatro métodos de pronósticos utilizados, no existe alguno que sea el ganador acorde a las frecuencias de las series (anual, cuarto de año, mensual, semanal y diario, en inglés: yearly, quarterly, monthly, weekly y daily), por lo que el teorema de No Free Lunch [87] sigue vigente. En la Figura 6.5 se puede

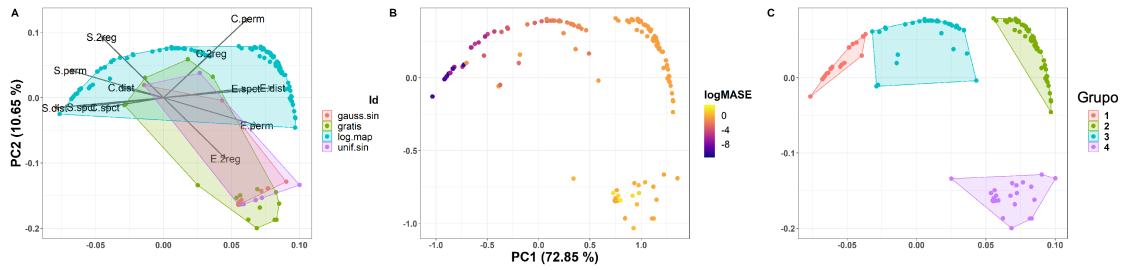


Figura 6.4 Series sintéticas y medidas de entropía (ESC).

observar este comportamiento.

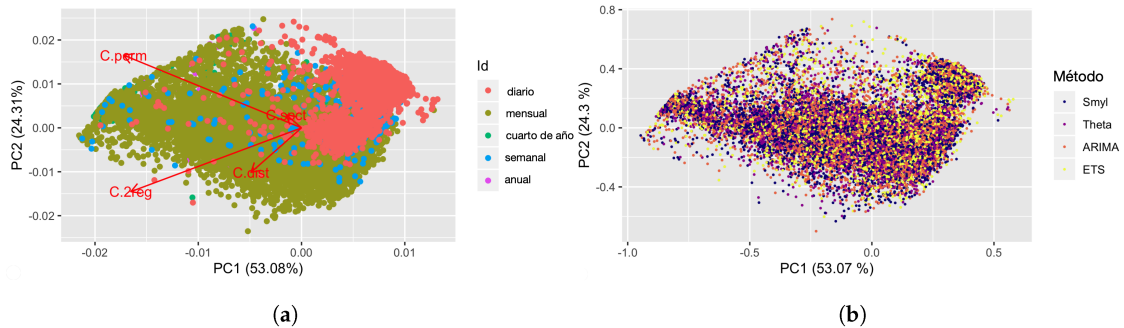


Figura 6.5 Análisis de las series de tiempo con medidas de entropía y los pronósticos de ARIMA, Smyl, Theta y ETS en M4-Competition.

Algo más que fue posible observar es que cuando se grafica el conjunto de series de M4 por cuartiles, los errores de pronósticos más pequeños se concentran en el primer componente principal. En las Figuras 6.6 y 6.7 se pueden observar cada una de las series de tiempo del conjunto de M4 y su relación con la medida de complejidad y la medida de error sMAPE (log(sMAPE) por funcionalidad).

A partir de estas figuras se puede determinar que cuando la medida de complejidad es alta, el valor del log(MAPE) es más alto también, y considerando las frecuencias de las series de tiempo (anual, cuarto de año, mensual, semanal y diario, en inglés: yearly, quarterly, monthly, weekly y daily), la mayor parte de las series de tiempo corresponden a la frecuencia diaria. Por último, se puede determinar que cuando la medida de complejidad es cercana a cero, ésta corresponde a un error de pronóstico pequeño y cuando la medida de complejidad es grande, corresponde a un error de pronóstico alto.

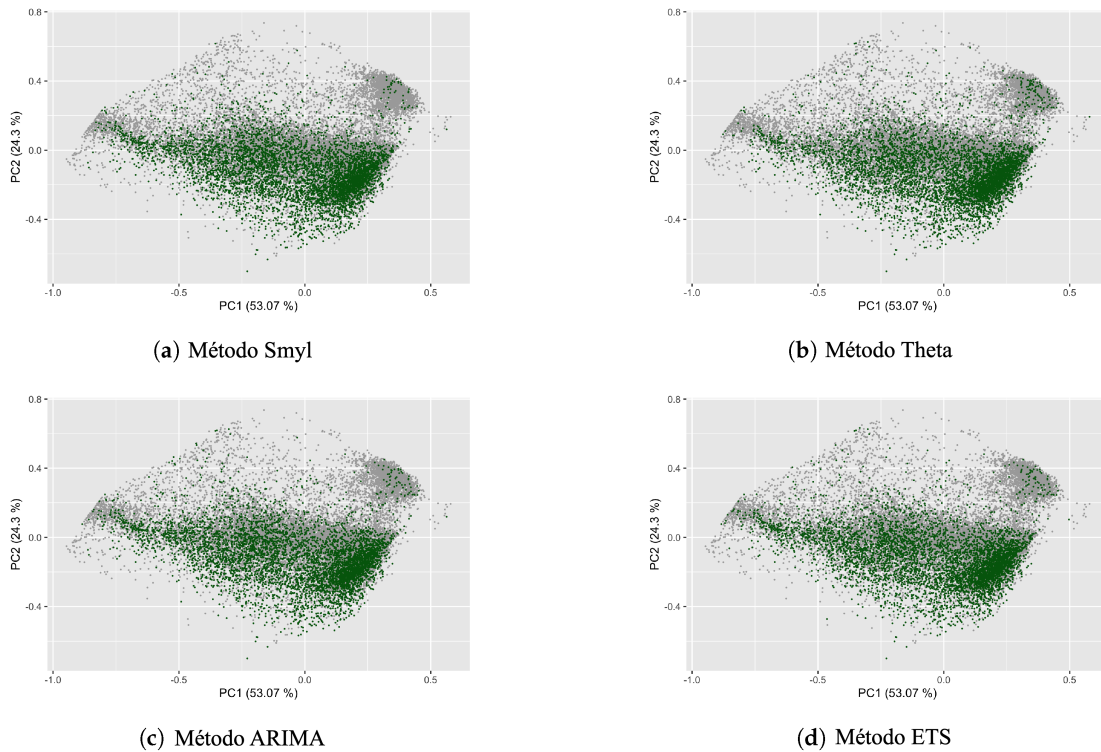


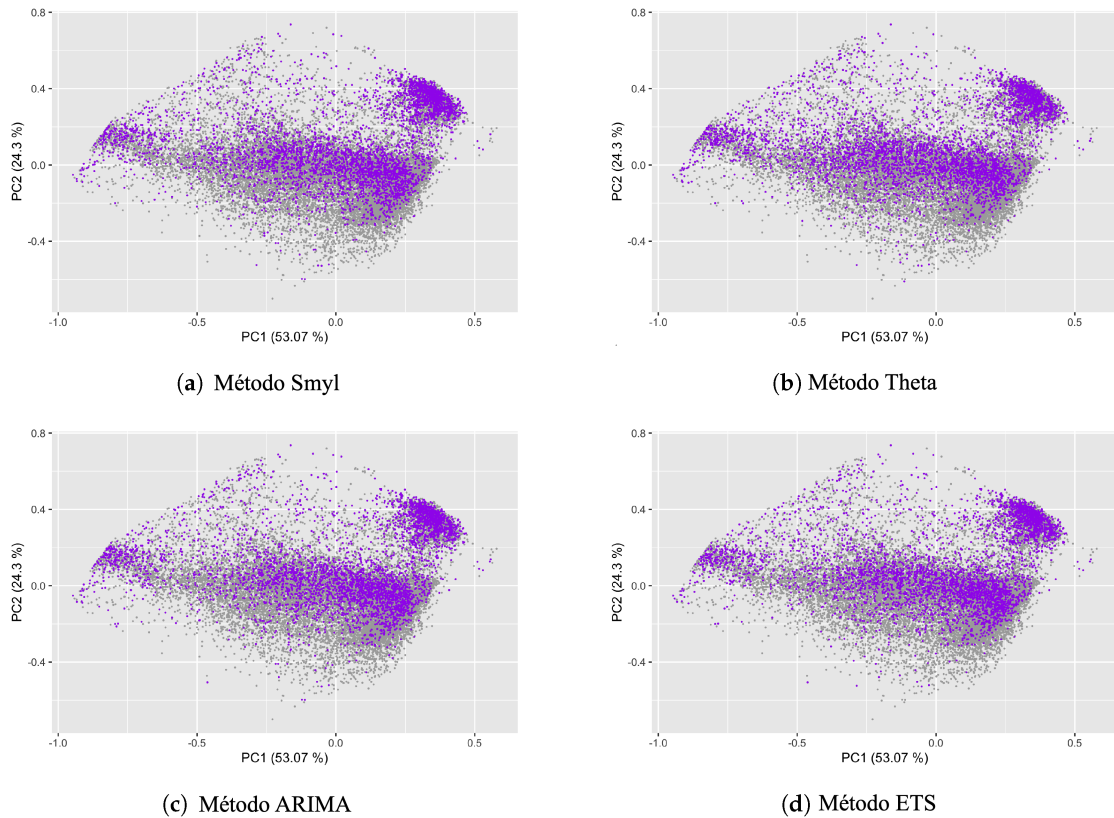
Figura 6.6 Medida de error-complejidad en el primer cuartil.

## 6.2. Metodología MetaLSS

Como se mencionó previamente en la sección 3: Metodología, con el desarrollo del método híbrido, se generaron seis métodos de pronósticos.

- MetaComb
- MetaComb-S
- MetaComb-SwA
- MetaSmoothing
- MetaSmoothing-S
- MetaSmoothing-SwA

Originalmente se planteó el desarrollo de un solo método (método híbrido basado en suavizamiento exponencial). Con el avance de la investigación se observó que se obtenían



**Figura 6.7** Medida de error-complejidad en el cuarto cuartil.

resultados diferentes cada que se le adicionaba alguna característica al método base y se decidió considerarlos como métodos independientes con la finalidad de medir los errores de los pronósticos que cada uno de ellos generaba.

Se utilizaron dos subconjuntos de datos de las competencias M4 de Makridakis: mensual y diario.

### 6.2.1. Resultados con el subconjunto de datos diario

La Figura 6.8 muestra el histograma de algunos de los métodos de pronósticos participantes incluyendo los seis métodos de pronósticos implementados, donde las columnas corresponden a los métodos y las filas al numero de series de tiempo que alcanzaron una determinada posición (Ej., en el método Holt, 91 de 844 series de tiempo quedaron en primer lugar, 102 de 844 series de tiempo quedaron en segundo lugar y así sucesivamente, por lo que, acorde

con esta figura, el método Jaganathan es el método que alcanzó colocar la mayor cantidad de series de tiempo en primer lugar, con un total de 334 series de tiempo).

	Holt	Damped	Theta	FFORMA	Joganathan	MetaComb	MetaComb-S	MetaComb-SwA	MetaSmoothing	MetaSmoothing-S	MetaSmoothing-SwA
1	91	23	27	39	334	33	3	33	34	35	136
2	102	37	47	42	70	95	31	52	31	96	118
3	55	61	73	41	10	125	98	57	35	109	25
4	20	38	52	34	10	123	131	85	61	109	14
5	7	42	46	41	7	71	128	110	136	88	35
6	37	36	37	28	29	94	68	84	171	77	62
7	66	44	44	49	53	70	90	52	145	56	63
8	105	60	77	61	16	84	68	87	68	48	21
9	64	213	213	174	19	128	180	183	137	147	37
10	44	98	74	88	12	12	32	53	14	61	20
11	92	111	92	103	44	3	13	33	7	17	69
12	109	62	43	97	170	2	2	15	4	1	178
13	36	19	17	40	45	4	0	0	1	0	56
14	16	0	2	7	25	0	0	0	0	0	10
Total	844	844	844	844	844	844	844	844	844	844	844

Figura 6.8 Histograma de los métodos de pronósticos generados con MetaLSS en el subconjunto de datos diario.

Observar de esta manera el número de series que obtuvieron una posición determinada no permite identificar cuál de estos métodos es el mejor, además de que en el caso de las series de la frecuencia diaria, estas son realmente fáciles y los métodos de pronósticos generalmente obtienen errores sMAPE promedios por debajo del 5%.

La Tabla 6.1 contiene los errores sMAPE de algunos de los métodos de pronósticos participantes, las celdas en color verde indican el método que presenta el menor error sMAPE para una determinada serie de tiempo.

Tabla 6.1 Muestra de errores sMAPE del subconjunto de datos diario.

Holt	Damped	Theta	FFORMA	Jaganathan	MetaComb	MetaComb-S	MetaComb-SwA	MetaSmoothing	MetaSmoothing-S	MetaSmoothing-SwA
1.39	1.93	1.60	2.62	1.31	1.67	1.67	1.62	1.62	1.67	3.46
2.20	2.26	2.53	2.64	2.40	2.55	2.55	2.55	2.47	2.55	1.89
2.74	2.04	2.10	1.19	2.53	0.95	0.95	1.02	1.14	0.95	1.78
2.55	2.27	2.35	1.52	2.70	1.13	1.13	1.09	1.27	1.13	1.18
1.27	1.26	1.22	1.29	1.42	1.02	1.05	1.05	1.08	1.05	1.05
2.72	1.96	2.01	1.87	2.40	1.14	1.14	1.09	1.34	1.14	2.05
1.41	1.43	1.43	1.52	0.92	1.06	1.12	1.12	1.17	1.12	0.97
1.23	1.19	1.22	1.41	1.33	0.97	0.97	0.97	0.94	0.97	1.06
1.33	1.22	1.27	0.86	0.91	0.85	0.85	0.83	0.86	0.85	0.90
0.82	0.87	0.89	0.95	1.71	0.67	0.67	0.67	0.68	0.67	0.75
0.62	0.61	0.62	0.63	1.66	0.86	0.86	0.94	0.80	0.86	1.63
0.66	0.65	0.64	0.88	1.59	0.79	0.79	0.79	0.76	0.79	0.74
0.69	0.67	0.67	0.66	0.93	0.67	0.67	0.67	0.67	0.67	0.63
1.25	1.28	1.28	1.29	1.67	1.41	1.41	1.41	1.37	1.41	2.59
0.52	0.52	0.52	0.51	0.58	0.53	0.53	0.53	0.53	0.53	0.53
0.68	0.73	0.71	0.84	0.87	0.72	0.72	0.72	0.71	0.70	0.64
0.56	1.15	1.12	2.20	1.04	1.50	1.50	1.50	1.30	1.50	2.37
1.24	1.19	1.20	1.95	0.92	1.42	1.42	1.42	1.38	1.42	2.41

Al analizar estos porcentajes, no es posible visualizar a simple vista cual de los métodos es el mejor para este subconjunto de series de tiempo. Lo que es posible apreciar, es que generalmente los  $n$  métodos de pronósticos que se analizan, tienen errores pequeños y en



algunos casos son similares, por lo que se puede decir, que este subconjunto de series de tiempo, en general es realmente fácil de pronosticar y la competencia entre métodos de pronósticos se vuelve todo un reto.

Con los errores sMAPE calculados, se realizaron algunas pruebas estadísticas que permitieron determinar el comportamiento general de cada uno de los métodos de pronósticos participantes para el subconjunto completo de las series de tiempo. En este caso, el total de series de tiempo del subconjunto de prueba equivale a 844 series de tiempo. La Tabla 6.2 muestra los resultados obtenidos al aplicar la prueba de Friedman, donde se destaca la participación de los seis métodos de pronósticos híbridos desarrollados en esta investigación, siendo el método *MetaComb-SwA* el mejor posicionado, seguido del método *MetaSmoothing-S*. El nivel de significancia es igual a 0.000 con 13 grados de libertad, concluyendo así que los métodos son diferentes entre si.

**Tabla 6.2** Prueba de Friedman al subconjunto de datos diario.

	<b>Rango medio</b>	<b>Posición</b>
Naive2	8.82	12
SES	8.83	13
Holt	7.88	8
Damped	8.77	11
Theta	8.12	9
Com	8.2	10
FFORMA	9.05	14
Jaganathan	6.5	6
MetaComb	6.25	4
MetaComb_S	6.22	3
MetaComb_SwA	6.12	1
MetaSmoothing	6.38	5
MetaSmoothing_S	6.14	2
MetaSmoothing_SwA	7.73	7
<b>Pruebas Estadísticas</b>		
N		844
Chi-cuadrado		875.06
df		13
Sign. Asint.		0.000

La prueba estadística de Wilcoxon, la cual se muestra en la Tabla 6.3 permitió realizar una comparación de pares de métodos de pronósticos, en este caso, entre el método que obtuvo un menor rango en la prueba de Friedman y los demás métodos participantes. A partir de la

suma de rangos positivos y negativos, se observó que el método *Jaganathan* es el que tiene una menor diferencia entre valores de rangos positivos y negativos y al revisar el valor del *p-value*, el cual se observa en la Tabla 6.4, este es igual a 0.915, muy cercano a la unidad, indicando que ambas muestras tienen un comportamiento estadísticamente equivalente. El valor de rango positivo es ligeramente mayor al rango negativo mostrando que el método híbrido MetaComb-SwA tiene un mejor desempeño de manera general comparado con el método Jaganathan, recordando que este método alcanzó el cuarto lugar general en las competencias de M4 [65].

**Tabla 6.3** Prueba *W* de Wilcoxon con rangos al subconjunto de datos diario.

		<b>N</b>	<b>Rango medio</b>	<b>Suma de rangos</b>
Jaganathan - MetaComb-SwA	Rangos negativos	431	410.95	177118
	Rangos positivos	412	433.56	178628
	Vínculos (empate)	1		
	Total	844		

**Tabla 6.4** Prueba estadística *W* de Wilcoxon al subconjunto de datos diario.

	<b>Z</b>	<b>Sig. asint. (2 colas)</b>
Jaganathan - MetaComb_SwA	-0.11	0.915
MetaSmoothing_S - MetaComb_SwA	-2.61	0.009
MetaComb_S - MetaComb_SwA	-3.95	0.000

### 6.2.2. Resultados con el subconjunto de datos mensual

El subconjunto de series de tiempo de frecuencia mensual se encuentra conformado por un total de 48,000 series de tiempo, las cuales se categorizan por su tipo en: demographic, finance, industry, macro, micro y other. Para los experimentos realizados, se consideraron solo las series de tiempo financieras (finance), con un total de 10,987 series de tiempo. Después de realizar la extracción de características, se excluyeron aquellas series de tiempo que contenían valores NAs en algunas de sus características, por lo que finalmente se consideraron en total 10,450 series de tiempo, este subconjunto se subdividió en entrenamiento y prueba utilizando un 80%-20% respectivamente.

En la Figura 6.9 se muestra un histograma con las posiciones de algunos de los métodos

de pronósticos participantes, donde las celdas contienen la cantidad de series de tiempo que obtuvieron una determinada posición; analizando el método de pronóstico Holt, el cual se encuentra en la columna número 2, se puede observar que de un total de 2,090 series de tiempo, en 128 de ellas, el método Holt obtuvo la primera posición, mientras que en 102 series de tiempo obtuvo la segunda posición y así sucesivamente. De esta manera, a simple vista no se puede determinar con precisión cuáles métodos son los ganadores.

	Holt	Damped	Theta	FFORMA	Jaganathan	MetaComb	MetaComb-S	MetaComb-SwAA	MetaSmoothing	MetaSmoothing-S	MetaSmoothing-SwAA
1	128	398	128	58	218	152	23	7	106	63	9
2	102	118	106	86	122	347	106	56	132	184	94
3	105	125	114	92	143	200	211	194	200	148	237
4	113	81	123	108	153	179	230	271	204	185	223
5	90	41	116	123	93	284	238	241	193	247	216
6	83	17	85	146	106	167	396	222	220	253	208
7	101	47	104	117	125	136	231	312	253	243	206
8	108	116	157	154	180	134	167	176	215	176	171
9	112	64	149	298	160	130	134	138	201	140	181
10	392	203	441	538	430	176	229	274	256	284	320
11	212	123	221	163	160	45	64	83	60	100	105
12	311	413	183	108	117	43	46	84	44	36	87
13	103	60	69	75	49	51	12	28	5	28	18
14	100	117	68	24	31	45	3	4	1	3	15
15	30	167	26	0	3	1	0	0	0	0	0
Total	2090	2090	2090	2090	2090	2090	2090	2090	2090	2090	2090

Figura 6.9 Histograma de los métodos de pronósticos generados con MetaLSS en el subconjunto de datos mensual.

La Tabla 6.5 contiene los errores sMAPE de algunos de los métodos de pronósticos participantes para un subconjunto de 18 series de tiempo seleccionadas de manera aleatoria, las celdas con fondo color verde indican el método que presenta el menor error sMAPE para una determinada serie de tiempo. Es posible observar que la medida de error sMAPE en esta selección de series se encuentra en el rango  $10 \leq sMAPE \leq 32$  lo que nos indica que las series de tiempo no son tan fáciles de pronosticar.

Tabla 6.5 Muestra de errores sMAPE del subconjunto de datos mensual.

Holt	Damped	Theta	FFORMA	Jaganathan	MetaComb	MetaComb-S	MetaComb-SwA	MetaSmoothing	MetaSmoothing-S	MetaSmoothing-SwA
22.92	31.83	24.30	23.30	21.10	24.08	23.79	23.67	21.58	20.95	21.16
17.81	21.89	19.91	19.87	16.77	18.16	17.94	18.04	18.72	19.90	19.66
26.76	32.84	22.59	27.20	24.43	21.08	21.08	21.08	23.31	22.47	22.47
14.78	12.53	14.06	13.57	12.27	15.26	14.20	14.67	13.60	14.20	13.26
15.40	10.29	15.37	12.58	17.84	10.32	10.74	10.66	10.47	10.74	10.66
11.66	23.96	19.57	18.58	13.10	11.10	11.00	10.90	13.83	14.17	14.02
20.48	20.95	20.45	20.36	20.78	20.34	20.34	20.34	20.30	20.40	20.40
10.00	10.04	14.93	9.45	10.31	8.26	8.26	8.26	7.34	7.61	7.61
24.17	21.92	24.16	23.39	28.06	22.68	22.98	22.95	22.76	22.98	22.95
28.09	23.36	26.75	26.03	21.25	19.14	19.24	19.30	19.81	19.24	19.30
25.85	22.30	22.36	22.36	25.15	22.18	22.20	22.20	22.15	22.20	22.20
9.62	16.78	9.89	11.37	10.27	9.60	9.60	9.60	10.55	9.60	9.60
10.01	16.62	9.98	10.35	10.15	10.01	10.05	10.04	10.16	10.05	10.04
24.45	22.94	22.88	23.45	24.43	23.56	23.45	23.47	23.35	23.45	23.47
14.37	12.40	14.26	13.63	12.94	11.78	11.78	11.77	11.63	11.78	11.77
31.90	44.16	31.88	35.20	31.57	31.55	31.55	31.55	33.57	35.15	35.15
33.59	48.25	34.99	35.12	34.60	33.42	33.71	33.70	32.37	35.17	35.17

A partir de los pronósticos generados para cada uno de los métodos de pronósticos participantes, se calcularon los errores sMAPE, posteriormente se realizaron algunas pruebas estadísticas, las cuales permitieron determinar el comportamiento general de cada uno de los métodos de pronósticos participantes para el subconjunto completo de las series de tiempo. En este caso, el total de series de tiempo del subconjunto de prueba, equivale a 2,090 series de tiempo. La Figura 6.6 muestra los resultados obtenidos al aplicar la prueba de Friedman, donde se destaca la participación de los seis métodos de pronósticos híbridos desarrollados en esta investigación, siendo el método *MetaComb* el mejor rankeado, seguido del método *MetaComb-SwA*. El nivel de significancia es igual a 0.000 con 14 grados de libertad, concluyendo así que los métodos son diferentes entre sí. Es importante destacar que para este subconjunto de pruebas, los seis métodos implementados, obtuvieron las seis primeras posiciones.

**Tabla 6.6** Prueba de Friedman al subconjunto de datos mensual.

	<b>Rango medio</b>	<b>Posición</b>
sNaive	9.52	13
Naive2	11.70	15
SES	9.76	14
Holt	9.08	12
Damped	8.73	11
Theta	8.48	10
Com	7.62	8
FFORMA	8.40	9
Jaganathan	7.44	7
<b>MetaComb</b>	<b>6.33</b>	<b>1</b>
MetaComb_S	6.41	3
MetaComb_SwA	6.39	2
MetaSmoothing	6.46	4
MetaSmoothing_S	6.87	6
MetaSmoothing_SwA	6.79	5
<b>Pruebas estadísticas</b>		
N		2090
Chi-cuadrado		3686.61
df		14
Sign. Asint.		0.000

Otra prueba estadística realizada, es la prueba de Wilcoxon, mostrada en la Tabla 6.7, la cual permitió realizar una comparación de pares de métodos de pronósticos, en este caso, entre el método que obtuvo un menor rango en la prueba de Friedman y los demás métodos

participantes. A partir de la suma de rangos positivos y negativos, se observó que los métodos MetaComb-S y MetaComb-SwA tienen sumas de rangos negativos mayores a los positivos y el valor *p-value* es igual a 0.000 y 0.031 respectivamente, valores que se pueden ver en la Tabla 6.8, por lo que la hipótesis nula se rechaza, esto indica que existe una diferencia significativa entre los métodos de pronósticos, otra observación a destacar es que el método MetaSmoothing presenta un valor de *p-value* igual a 0.061, indicando que la hipótesis nula es aceptada, esto significa que ambos métodos son estadísticamente equivalentes.

**Tabla 6.7** Prueba *W* de Wilcoxon con rangos al subconjunto de datos mensual.

		<b>N</b>	<b>Rango medio</b>	<b>Suma de rangos</b>
MetaComb_S - MetaComb	Rangos negativos	386	462.96	178704.00
	Rangos positivos	404	331.04	133741.00
	Vínculos(empate)	1300		
	Total	2090		
MetaComb_SwA - MetaComb	Rangos negativos	379	492.030	186480.00
	Rangos positivos	449	349.06	156726.00
	Vínculos(empate)	1262		
	Total	2090		

**Tabla 6.8** Prueba estadística de Wilcoxon al subconjunto de datos mensual.

	<b>Z</b>	<b>Sig. Asint. (2colas)</b>
MetaComb_S - MetaComb	-3.500	0.000
MetaComb_SwA - MetaComb	-2.160	0.031
MetaSmoothing - MetaComb	-1.860	0.062

Debido a los resultados obtenidos en la suma de rangos, se calculó la prueba de Wilcoxon, esta vez realizando una prueba emparejada del método MetaComb-S contra los demás trece métodos de pronósticos participantes, se elige este método debido a que de los dos métodos que presentaron un mayor número en la suma de rangos negativos, este es el método que tiene una mayor diferencia entre valores de rangos positivos y negativos. Los resultados muestran que se presenta una suma de rangos positivos mayores a la suma de rangos negativos en todos los casos y los respectivos valores *p-value* calculados, son menores a el valor de 0.05, lo cual indica que se rechaza la hipótesis nula, indicando que todos los métodos son estadísticamente diferentes. De esta manera, podemos concluir que para el subconjunto de series de tiempo mensual, el método MetaComb-S presenta al mejor desempeño.

---

# Conclusiones y trabajos futuros

## 7.1. Conclusiones

Las conclusiones encontradas durante el desarrollo de la investigación, son las siguientes: Los métodos híbridos han demostrado ser exitosos. Prueba de ello son los mejores métodos reportados en las pruebas de Makridakis.

A partir de la característica del exponente de Hurst, el cual es utilizado para medir la persistencia en las series, se encontró una fuerte relación entre el error generado por los métodos de pronósticos de suavizamiento exponencial y ARIMA con dicho exponente. De tal forma que un valor del exponente de Hurst cercano a cero tendrá relación con un promedio de error alto. Mientras que un valor del exponente de Hurst cercano a uno tendrá relación con un menor promedio de error.

Trabajando con las series de tiempo de M3-Competition, se encontró que el error del pronóstico depende del método que lo resuelve, es decir, para un método la serie puede ser difícil, mientras que para otro método puede ser bastante fácil. Esto soporta el teorema de *No Free lunch* [87].

Se analizaron cuatro medidas de complejidad en las series de tiempo de M4-Competition

utilizando la entropía de Shannon. De las cuatro medidas de complejidad, sobresale la complejidad de permutación y 2-regímenes, ya que permiten visualizar patrones relacionados con los componentes principales y el error del pronóstico.

Por otra parte, se desarrollaron seis algoritmos cuyos resultados fueron estadísticamente similares y superiores a los que se reportan en el estado del arte. Se comprobó comparándolos con series de tiempo de M4-Competitions para las frecuencias de datos mensual y diario, posteriormente se realizaron pruebas estadísticas de Friedman y Wicolxon en ambos subconjuntos de datos, dichas pruebas muestran la equivalencia o superioridad de los métodos propuestos.

El objetivo general de este trabajo de investigación consistió en *el desarrolló de un método híbrido basado en metaaprendizaje y suavizamiento exponencial para realizar pronósticos de series de tiempo financieras con un desempeño similar a los métodos del estado del arte*, por lo que, se puede decir que se concluyó satisfactoriamente el objetivo general así como todos los objetivos específicos de esta tesis. De manera complementaria, se implementó un marco de trabajo (framework) que permite realizar pronósticos de calidad en series de tiempo, a corto, mediano y largo plazo.

## 7.2. Trabajos futuros

Existen diversas áreas de oportunidad que se encontraron durante la realización de esta investigación, a continuación, se enumeran las más relevantes:

- Aplicar diversas características al método híbrido propuesto en otros conjuntos de datos.
- Estandarizar y ampliar la metodología propuesta para otros tipos de series con diferentes características.

### 7.3. Publicaciones derivadas del presente trabajo

- Artículo aceptado y publicado en la revista IEEE Latin American Transactions, con fecha de noviembre de 2019.

*Hurst Exponent with ARIMA and Simple Exponential Smoothing for Measuring Persistence of M3- Competition Series [4].*

- Artículo aceptado y publicado en la revista MDPI Special Issue: Entropy Application for Forecasting, con fecha de enero de 2020.

*Time Series Complexities and Their Relationship to Forecasting Performance [5].*



# Glosario

**Autocorrelación** Relación existente entre la misma variable desfasada en el tiempo.

**Componente cíclico** Es la fluctuación de los datos en el tiempo, alrededor del patrón de tendencia.

**Componente estacional** Es un patrón de cambio que se repite año tras año, con base en la frecuencia (mensual, bimestral, trimestral o cuatrimestral ) de los datos.

**Componente estacionario** Es cuando los datos de una serie de tiempo fluctúan cerca de su media.

**Componente de tendencia** Representa el crecimiento o el descenso en la serie de tiempo, durante un periodo extenso. Cuando los datos crecen o descienden en varios periodos.

**Correlación** Asociación lineal que existe entre dos variables numéricas.

**Datos de entrenamiento** Se le conoce así, al subconjunto de datos (usualmente un 80%) con el que se construye un modelo de predicción, es decir, los datos que son conocidos.

**Datos de prueba** Se le determina de esta manera, al subconjunto de datos (usualmente un 20%) con el que se realiza

la prueba del modelo construido, para determinar la del modelo.

**Frecuencia** Número de veces en un tiempo determinado, que sucede un evento o suceso, las frecuencias más comunes son: anual, mensual y diaria.

**Horizonte** Se le conoce así al número de datos que se pronosticaran en un corto, mediano o largo plazo, generalmente se le denota con la letra  $h$ .

**Método de pronóstico** Técnica con una metodología específica, la cual permite realizar predicciones de series de tiempo para un corto, mediano o largo plazo.

**Metodología Box-Jenkins** Conjunto de procedimientos que permiten reconocer, ajustar y verificar modelos ARIMA (Autoregressive Integrated Moving Average, del inglés) con los datos de la serie de tiempo que se analiza.

**Modelo Autorregresivo (Autoregression, AR)** Modelo de pronóstico que hace uso de valores previos de la serie de tiempo para pronosticar.

**Observación** Se le conoce así a cada uno de los datos que componen a una serie de tiempo, sin importar la frecuencia de estos.

**Promedios Móviles (Moving Average, MA)**

Método de pronóstico para un dato futuro, a partir de la media aritmética de los valores de la serie de tiempo para un intervalo de datos  $k$  más recientes.

**Patrones de datos**

Comportamiento de los datos numéricos que son recurrentes, de los cuales los más comunes son: tendencia, estacionalidad, estacionariedad y cíclicos. En series de tiempo, a los patrones se les conocen como componentes.

**Periodo**

Se le llama así, a cada uno de los datos u observaciones de una serie de tiempo, donde cada dato que compren-

de a la serie de tiempo, es un valor obtenido que corresponde a la frecuencia de la serie de tiempo.

**Regresión lineal simple**

Es un modelo de regresión que utiliza la ecuación de la recta o pendiente para determinar valores en el futuro, a partir de una variable que se conoce como variable independiente.

**Residual**

Es la diferencia que existe, entre un valor real de la serie de tiempo y su valor de pronóstico correspondiente.

**Serie de tiempo**

Son los datos que se registran u observan a través del tiempo.

# Bibliografía

- [1] J. E. Hanke and D. W. Wichern, *Pronósticos en los negocios*, 9th ed., R. F. Rivera, Ed. Pearson Educación, 2006, 2010.
- [2] G. E. Box, G. M. Jenkins, and G. C. Reinsel, "Time Series Analysis - Forecasting and Control," *Prentice Hall New Jersey 1994*, vol. 10, no. Chapter 5, pp. 837–900, 1994. [Online]. Available: [http://books.google.com.au/books?id=B8\\_1UBmqVUoC](http://books.google.com.au/books?id=B8_1UBmqVUoC)
- [3] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, D. J. Balding, N. Cressie, G. Fitzmaurice, and I. Johnstobe, Eds. WILEY Series in Probability And Statistics, 2008.
- [4] M. Ponce, J. Frausto, G. Castilla, J. González, J. Pérez, and J. Téran, "Hurst Exponent with ARIMA and Simple Exponential Smoothing for Measuring Persistency of M3-Competition Series," *IEEE Latin America Transactions*, vol. 17, no. 5, pp. 815–822, 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8891950>
- [5] M. Ponce-Flores, J. Frausto-Solís, G. Santamaría-Bonfil, J. Pérez-Ortega, and J. J. González-Barbosa, "Time series complexities and their relationship to forecasting performance," *Entropy*, vol. 22, no. 1, p. 89, 2020. [Online]. Available: <https://doi.org/10.3390/e22010089>
- [6] E. Gardner, "Exponential smoothing: The state of the art Part II," *International Journal of Forecasting*, vol. 22, pp. 637–666, 2006. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0169207006000392>
- [7] J. G. De Gooijer, R. J. Hyndman, J. G. D. Gooijer, and R. J. Hyndman, "25 Years of Time Series Forecasting," p. 2006, 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0169207006000021>
- [8] R. G. Brown, "Statistical forecasting for inventory control," *McGraw-Hill, New York*, p. 1959, 1959.
- [9] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.
- [10] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

- [11] M. Sheraz and V. Preda, "Implied Volatility in Black-scholes Model with Garch Volatility," *Procedia Economics and Finance*, vol. 8, no. 14, pp. 658–663, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2212567114001415>
- [12] D. M. AL-Najjar, "Modelling and Estimation of Volatility Using ARCH/GARCH Models in Jordan's Stock Market," *Asian Journal of Finance & Accounting*, vol. 8, no. 1, p. 152, 2016. [Online]. Available: <http://www.macrothink.org/journal/index.php/ajfa/article/view/9129>
- [13] C. Hamzaçebi, "Improving artificial neural networks' performance in seasonal time series forecasting," *Information Sciences*, vol. 178, no. 23, pp. 4550–4559, 2008. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0020025508002958>
- [14] P. Chang, D. Wang, and C. Zhou, "A novel model by evolving partially connected neural network for stock price trend forecasting," *Expert Systems With Applications*, vol. 39, no. 1, pp. 611–620, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2011.07.051>
- [15] F. A. De Oliveira, C. N. Nobre, and L. E. Zárata, "Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index - Case study of PETR4, Petrobras, Brazil," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7596–7606, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2013.06.071>
- [16] J. Z. Wang, J. J. Wang, Z. G. Zhang, and S. P. Guo, "Forecasting stock indices with back propagation neural network," *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 346–14 355, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2011.04.222>
- [17] M. Boži and M. Stojanovi, "Application of SVM Methods for Mid-Term Load Forecasting," *Serbian Journal of Electrical Engineering*, vol. 8, no. 1, pp. 73–83, 2011. [Online]. Available: [http://www.journal.ftn.kg.ac.rs/Vol\\_8-1/08-Bozic-Stojanovic.pdf](http://www.journal.ftn.kg.ac.rs/Vol_8-1/08-Bozic-Stojanovic.pdf)
- [18] C. Yang, G. N. Odvody, C. J. Fernandez, J. A. Landivar, R. R. Minzenmayer, and R. L. Nichols, "Evaluating unsupervised and supervised image classification methods for mapping cotton root rot," *Precision Agriculture*, vol. 16, no. 2, pp. 201–215, 2015. [Online]. Available: <https://link.springer.com/article/10.1007/s11119-014-9370-9>
- [19] R. Palivonaite, K. Lukoseviciute, and M. Ragulskis, "Short-term time series algebraic forecasting with mixed smoothing," *Neurocomputing*, vol. 171, pp. 854–865, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.neucom.2015.07.018>
- [20] O. B. Shukur, N. Salem Fadhil, M. Hisyam Lee, and M. H. Ahmad, "Electricity Load Forecasting using Hybrid of Multiplicative Double Seasonal Exponential Smoothing Model with Artificial Neural Network," *Jurnal Teknologi*, vol. 69, no. 2, pp. 65–70, 2014. [Online]. Available: <https://journals.utm.my/jurnalteknologi/article/view/3109>
- [21] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, "Satellite image analysis and a hybrid ESSS/ANN model to forecast solar irradiance in the tropics," *Energy Conversion and Management*, no. March, pp. 66–73. [Online]. Available: <https://doi.org/10.1016/j.enconman.2013.11.043>
- [22] R. Shan, H. Dai, J. Zhao, and W. Liu, "Forecasting Study of Shanghai's and Shenzhen's Stock Markets Using a Hybrid Forecast Method," *Communications in Statistics -*

- Simulation and Computation*, vol. 44, no. 4, pp. 1066–1077, 2014. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.1080/03610918.2013.804554>
- [23] L. Tang, L. Yu, F. Liu, and W. Xu, “An Integrated Data Characteristic Testing Scheme for Complex Time Series Data Exploration,” *International Journal of Information Technology & Decision Making*, vol. 12, no. 03, pp. 491–521, 2013. [Online]. Available: <http://www.worldscientific.com/doi/abs/10.1142/S0219622013500193>
- [24] J. Frausto-Solís, M. Chi-Chim, and L. Sheremetov, “Forecasting oil production time series with a population-based simulated annealing method,” *Arabian Journal for Science and Engineering*, no. 4, pp. 1081–1096. [Online]. Available: <https://link.springer.com/article/10.1007/s13369-015-1587-z>
- [25] L. F. Tratar, B. Mojškerc, and T. Aleš, “Demand forecasting with four-parameter exponential smoothing,” *Int. J. Production Economics journal*, vol. 181, pp. 162–173, 2016. [Online]. Available: <https://doi.org/10.1016/j.ijpe.2016.08.004>
- [26] S. BuHamra, N. Smaoui, and M. Gabr, “The Box-Jenkins analysis and neural networks: Prediction and time series modelling,” *Applied Mathematical Modelling*, no. 10, pp. 805–815. [Online]. Available: [https://doi.org/10.1016/S0307-904X\(03\)00079-9](https://doi.org/10.1016/S0307-904X(03)00079-9)
- [27] Y. Xiao, J. Xiao, and S. Wang, “A hybrid model for time series forecasting,” *Human Systems Management*, vol. 31, no. March, pp. 133–143, 2012.
- [28] M. E. Banihabib, A. Ahmadian, and F. S. Jamali, “Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data,” *GeoResJ*, vol. 13, pp. 9–16, jun 2017. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S2214242816300584>
- [29] A. Shabri and R. Samsudin, “A hybrid GMDH and Box-Jenkins models in time series forecasting,” *Applied Mathematical Sciences*, vol. 8, no. 61-64, pp. 3051–3062, 2014. [Online]. Available: <http://dx.doi.org/10.12988/ams.2014.44270>
- [30] G. C. Onwubolu, “Design of hybrid differential evolution and group method of data handling networks for modeling and prediction,” *Information Sciences*, vol. 178, no. 18, pp. 3616–3634, 2008. [Online]. Available: <https://doi.org/10.1016/j.ins.2008.05.013>
- [31] R. Samsudin, P. Saad, and A. c. Shabri, “A hybrid GMDH and least squares support vector machines in time series forecasting,” *Neural Network World*, vol. 21, no. 3, pp. 251–268, 2011.
- [32] A. F. Osman and M. L. King, “A new approach to forecasting based on exponential smoothing with independent regressors,” in *Department of Econometrics and Business Statistics*. Monash University, 2015, no. 02/15, ch. 02, pp. 1–19. [Online]. Available: <http://econpapers.repec.org/paper/mshebswps/2015-2.htm>
- [33] P. L. Bernstein, *Against the gods : the remarkable story of risk*. John Wiley & Sons, 1996.

- [34] S. Panigrahi and H. S. Behera, "A hybrid ETS–ANN model for time series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 66, no. June, pp. 49–59, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.engappai.2017.07.007>
- [35] S. Makridakis, M. Hibon, E. Lusk, and M. Belhadjali, "Confidence intervals. An empirical investigation of the series in the M-competition," *International Journal of Forecasting*, vol. 3, no. 3-4, pp. 489–508, 1987. [Online]. Available: [https://doi.org/10.1016/0169-2070\(87\)90045-8](https://doi.org/10.1016/0169-2070(87)90045-8)
- [36] S. Makridakis, "Accuracy measures: theoretical and practical concerns," *International Journal of Forecasting*, vol. 9, no. 4, pp. 527–529, dec 1993. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/0169207093900793>
- [37] S. Makridakis and M. Hibon, "The M3-Competition: results, conclusions and implications," *International Journal of Forecasting*, vol. 16, no. 4, pp. 451–476, oct 2000. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0169207000000571>
- [38] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition : Results , findings , conclusion and way forward The M4 Competition : Results , findings , conclusion and way forward," *International Journal of Forecasting*, no. June, 2018. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2018.06.001>
- [39] P. Brockwell and R. Davis, *Introduction to Time Series and Forecasting*, 2002. [Online]. Available: <https://link.springer.com/book/10.1007%2F978-3-319-29854-2>
- [40] S. de Gobernación, "DOF - Diario Oficial de la Federación," 2021. [Online]. Available: <http://diariooficial.gob.mx/estadisticas.php>
- [41] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed., S. David J. Balding, Noel A. C. Cressie, Garrett M. Fitzmaurice, Geof H. Givens, Harvey Goldstein, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, Ruey S. Tsay, Ed. WILEY Series in Probability And Statistics, 2016.
- [42] B. E. Flores, "A pragmatic view of accuracy measurement in forecasting," *Omega*, vol. 14, no. 2, pp. 93–98, jan 1986. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/0305048386900137>
- [43] P. Goodwin and R. Lawton, "On the asymmetry of the symmetric MAPE," *International Journal of Forecasting*, vol. 15, no. 4, pp. 405–408, 1999. [Online]. Available: [https://doi.org/10.1016/S0169-2070\(99\)00007-2](https://doi.org/10.1016/S0169-2070(99)00007-2)
- [44] K. K. Lai, L. Yu, S. Wang, and W. Huang, "Hybridizing exponential smoothing and neural network for financial time series predication," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3994 LNCS, pp. 493–500, 2006. [Online]. Available: [https://link.springer.com/content/pdf/10.1007%2F11758549\\_69.pdf](https://link.springer.com/content/pdf/10.1007%2F11758549_69.pdf)
- [45] E. L. D. Faria, M. P. Albuquerque, J. L. Gonzalez, J. T. P. Cavalcante, and M. P. Albuquerque, "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods," *Expert Systems With Applications*, vol. 36, no. 10, pp. 12 506–12 509, 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2009.04.032>

- [46] E. Guresen, G. Kayakutlu, and T. U. Daim, "Using artificial neural network models in stock market index prediction," *Expert Systems With Applications*, vol. 38, no. 8, pp. 10 389–10 397, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2011.02.068>
- [47] S. Smyl, "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting," *International Journal of Forecasting*, vol. 36, no. 1, pp. 75–85, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2019.03.017>
- [48] A. Bansal, R. J. Kauffman, and R. R. Weitz, "Comparing the modeling performance of regression and neural networks as data quality varies: A business value approach," *Journal of Management Information Systems*, vol. 10, no. 1, pp. 11–32, 1993. [Online]. Available: <https://doi.org/10.1080/07421222.1993.11517988>
- [49] S. Haykin, *Neural networks and learning machines*, 3rd ed., M. J. Horton, A. Dworkin, D. Mars, and W. Opaluch, Eds. Pearson Hall, Prentice, 1999.
- [50] D. F. Spencht, "A General Regression Neural Network," in *IEEE International Conference on Neural Networks - Conference Proceedings*, vol. 2, no. 6, 1991, pp. 1959–1964.
- [51] P. G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, no. February, pp. 159–175, 2003. [Online]. Available: [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)
- [52] M. C. Tsai, C. H. Cheng, M. I. Tsai, and H. Y. Shiu, "Forecasting leading industry stock prices based on a hybrid time-series forecast model," *PLoS ONE*, vol. 13, no. 12, pp. 1–24, 2018. [Online]. Available: <https://doi.org/10.1371/journal.pone.0209922>
- [53] C. T. Cheng, Z. K. Feng, W. J. Niu, and S. L. Liao, "Heuristic methods for reservoir monthly inflow forecasting: A case study of xinfengjiang reservoir in pearl river, China," *Water (Switzerland)*, vol. 7, no. 8, pp. 4477–4495, 2015. [Online]. Available: <https://doi.org/10.3390/w7084477>
- [54] P. R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," in *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, 1960, vol. 132, pp. 324–342. [Online]. Available: [https://doi.org/10.1007/978-3-642-51565-1\\_{ }116](https://doi.org/10.1007/978-3-642-51565-1_{ }116)
- [55] G. U. Yule, "On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers," in *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 1927, vol. 226, no. 636-646, pp. 267–298.
- [56] E. Hadavandi, H. Shavandi, and A. Ghanbari, "Knowledge-Based Systems Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 800–808, 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.knosys.2010.05.004>
- [57] J.-j. Wang, J.-z. Wang, Z.-g. Zhang, and S.-p. Guo, "Stock index forecasting based on a hybrid model," *Omega*, vol. 40, no. 6, pp. 758–766, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.omega.2011.07.008>

- [58] J. L. Ticknor, "A Bayesian regularized artificial neural network for stock market forecasting," *Expert Systems With Applications*, vol. 40, no. 14, pp. 5501–5506, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.eswa.2013.04.013>
- [59] A. R. Finamore, V. Galdi, V. Calderaro, A. Piccolo, G. Conio, and S. Grasso, "Artificial Neural Network Application in Wind Forecasting : an One-Hour-Ahead Wind Speed Prediction," in *The 5th IET International Conference on Renewable. 5th IET International Conference on Renewable Power Generation (RPG) 2016*, 2016, pp. 1–6. [Online]. Available: <https://ieeexplore.ieee.org/document/7835492>
- [60] J. Pan, Y. Y. Tang, Y. Wang, P. Shen, and P. Wang, "A hybrid swarm optimization for neural network training with application in stock price forecasting," in *IEEE International Conference on Systems, Man, and Cybernetics. 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 4450–4453. [Online]. Available: <https://ieeexplore.ieee.org/document/7844932>
- [61] L. Wu, S. Liu, and Y. Yang, "Grey double exponential smoothing model and its application on pig price forecasting in China," *Applied Soft Computing Journal*, vol. 39, pp. 117–123, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.asoc.2015.09.054>
- [62] S. P. Das and S. Padhy, "A novel hybrid model using teaching–learning based optimization and a support vector machine for commodity futures index forecasting," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 1, pp. 97–111, 2015. [Online]. Available: <https://link.springer.com/article/10.1007/s13042-015-0359-0>
- [63] A. R. Ali, B. Gabrys, and M. Budka, "Cross-domain Meta-learning for Time-series Forecasting," *Procedia Computer Science*, vol. 126, pp. 9–18, 2018. [Online]. Available: <https://doi.org/10.1016/j.procs.2018.07.204>
- [64] M. Siddique, S. Mohanty, and D. Panda, "a Hybrid Forecasting Model for Prediction of Stock Value of Tata Steel Using," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 14, pp. 1719–1727, 2018. [Online]. Available: <https://acadpubl.eu/hub/2018-119-14/articles/3/46.pdf>
- [65] S. Jaganathan and P. K. Prakash, "A combination-based forecasting method for the M4-competition," *International Journal of Forecasting*, vol. 36, no. 1, pp. 98–104, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2019.03.030>
- [66] P. Montero-Manso, G. Athanasopoulos, R. J. Hyndman, and T. S. Talagala, "FFORMA: Feature-based forecast model averaging," *International Journal of Forecasting*, vol. 36, no. 1, pp. 86–92, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2019.02.011>
- [67] Y. Huang, Y. Gao, Y. Gan, and M. Ye, "A new financial data forecasting model using genetic algorithm and long short-term memory network," *Neurocomputing*, vol. 425, no. xxxx, pp. 207–218, 2021. [Online]. Available: <https://doi.org/10.1016/j.neucom.2020.04.086>



- [68] L. Tang, H. Lv, F. Yang, and L. Yu, “Complexity testing techniques for time series data: A comprehensive literature review,” *Chaos, Solitons and Fractals*, vol. 81, pp. 117–135, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.chaos.2015.09.002>
- [69] E. E. Peters, *Chaos and order in the capital markets : a new view of cycles, prices, and market volatility*. Wiley, 1996.
- [70] C. W. J. Granger and R. Joyeux, “An Introduction to Long-Memry Time Series Models and Fractional Differencing,” *Journal of Time Series Analysis*, vol. 1, no. 1, pp. 15–29, 1980. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-9892.1980.tb00297.x>
- [71] J. R. Hosking, “Fractional Differencing,” *Biometrika*, vol. 68, no. 1, pp. 165–176, 1981.
- [72] X. Wang, K. Smith-Miles, and R. Hyndman, “Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series,” *Neurocomputing*, vol. 72, no. 10-12, pp. 2581–2594, 2009. [Online]. Available: <https://doi.org/10.1016/j.neucom.2008.10.017>
- [73] E. E. Hurst, “Long-Term Storage Capacity of Reservoirs,” *Transactions of The American Society of civil Engineers*, vol. 116, pp. 770–799, 1951.
- [74] B. B. Mandelbrot and J. R. Wallis, “Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence,” *Water Resources Research*, vol. 5, no. 5, pp. 967–988, oct 1969. [Online]. Available: <http://doi.wiley.com/10.1029/WR005i005p00967>
- [75] S. Bhardwaj, “Data Mining Clustering Techniques – A Review,” *International Journal of Computer Science and Mobile Computing*, vol. 6, no. 5, pp. 183–186, 2017.
- [76] L. A. Albores and M. López, “Classification of User Interactive Interfaces with K-modes,” *IEEE Latin America Transactions*, vol. 13, no. 7, pp. 2308–2313, 2015. [Online]. Available: <https://ieeexplore.ieee.org/document/7273792>
- [77] B. Liu, *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Springer, 2011.
- [78] A. K. Jain, “Data clustering: 50 years beyond K-means,” *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.patrec.2009.09.011>
- [79] R. S. Tsay, *Analysis of Financial Time Series: Third Edition*, third. ed. ed. Wiley a John Wiley & Sons Inc. Publication, 2010.
- [80] G. Santamaría-bonfil, “A Package for Measuring emergence , Self-organization , and Complexity Based on Shannon entropy,” vol. 4, no. March, pp. 1–12, 2017. [Online]. Available: <https://doi.org/10.3389/frobt.2017.00010>
- [81] Y. Kang, R. J. Hyndman, and F. Li, “GRATIS: GenerAting Time Series with diverse and controllable characteristics,” *Statistical Analysis and Data Mining*, mar 2020. [Online]. Available: <https://doi.org/10.1002/sam.11461>

- [82] X. Wang, K. Smith, and R. Hyndman, "Characteristic-based clustering for time series data," *Data Mining and Knowledge Discovery*, vol. 13, no. 3, pp. 335–364, 2006. [Online]. Available: <https://doi.org/10.1007/s10618-005-0039-x>
- [83] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *International Journal of Forecasting*, vol. 36, no. 1, pp. 54–74, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2019.04.014>
- [84] M. Chi, "Pronóstico de series de tiempo en la ingeniería de yacimientos de petróleo con estrategias híbridas basadas en metaheurísticas," Ph.D. dissertation, INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY, 2015.
- [85] S. F. Crone, M. Hibon, and K. Nikolopoulos, "Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction," *International Journal of Forecasting*, vol. 27, no. 3, pp. 635–660, 2011. [Online]. Available: <https://doi.org/10.1016/j.ijforecast.2011.04.001>
- [86] R. R. Andrawis, A. F. Atiya, and H. El-Shishiny, "Forecast combinations of computational intelligence and linear models for the NN5 time series forecasting competition," *International Journal of Forecasting*, vol. 27, no. 3, pp. 672–688, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.ijforecast.2010.09.005>
- [87] D. Wolpert, "No free lunch theorems for search," *Technical Report SFI-TR-95-02-010*, pp. 1–38, 1995.