

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



**TESIS**

**PRONÓSTICO DE SERIES DE TIEMPO PARA EL MERCADO DE ENERGÍA  
MEXICANO**

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

Presenta:

**MCC. Lemuel Rodríguez Moya**  
**D09070427**  
**No. CVU de CONACyT 663496**

Director de Tesis  
**Dr. Juan Frausto Solís**  
**No. CVU de CONACyT 31308**

Co-Director de Tesis  
**Dr. Juan Javier González Barbosa**

Ciudad Madero, Tamaulipas, **27/abril/2022**

**OFICIO No. : U.056/22**  
**ASUNTO: AUTORIZACIÓN DE**  
**IMPRESIÓN DE TESIS**

**C. LEMUEL RODRÍGUEZ MOYA**  
**No. DE CONTROL D09070427**  
**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:

**"PRONÓSTICO DE SERIES DE TIEMPO PARA EL MERCADO DE ENERGÍA MEXICANO"**

El Jurado está integrado por los siguientes catedráticos:

PRESIDENTE:	DR.	JUAN FRAUSTO SOLÍS
SECRETARIO:	DR.	JUAN JAVIER GONZÁLEZ BARBOSA
PRIMER VOCAL:	DRA.	GUADALUPE CASTILLA VALDEZ
SEGUNDO VOCAL:	DR.	RUBÉN SALAS CABRERA
TERCER VOCAL:	DR.	LUCIANO AGUILERA VÁZQUEZ
SUPLENTE	DR.	ULISES PÁRAMO GARCÍA
DIRECTOR DE TESIS:	DR.	JUAN FRAUSTO SOLÍS
CO-DIRECTOR:	DR.	JUAN JAVIER GONZÁLEZ BARBOSA

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.

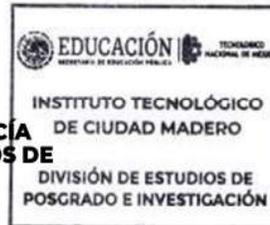
**ATENTAMENTE**

*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



Av. 1ª de Mayo y Sor Juana I. de la Cruz S/N Col. Los Mangos C.P. 89440 Cd.  
Madero, Tam.  
Tel. 01 (833) 357 48 20, ext. 3110, e-mail: [depi\\_cdmadero@tecnm.mx](mailto:depi_cdmadero@tecnm.mx)  
[tecnm.mx](http://tecnm.mx) | [cdmadero.tecnm.mx](http://cdmadero.tecnm.mx)



**2022** Ricardo  
**Flores**  
Año de  
**Magón**  
PROCLAMACIÓN DE LA REVOLUCIÓN MEXICANA

## **DECLARACIONES DE ORIGINALIDAD Y PROPIEDAD INTELLECTUAL**

Declaro que la investigación es original y este documento de tesis son producto de mi directa contribución intelectual ya que sus contenidos no infringen los derechos de terceros, tales como derechos de publicación, derechos de autor y similares. Todos los datos y las referencias a materiales ya publicados están debidamente identificados con su respectivo crédito e incluidos en las notas bibliográficas y en las citas que se destacan como tal.

Por lo tanto, me hago responsable de cualquier infracción o reclamación relacionada con derechos de propiedad intelectual, exonerando de responsabilidad a mi director de tesis codirectores y al Tecnológico nacional de México, así como a al Instituto tecnológico de Ciudad de México.

A handwritten signature in black ink, appearing to read 'Lemuel Rodriguez Moya', written over a horizontal line.

Lemuel Rodriguez Moya

## AGRADECIMIENTOS

- Le doy gracias a Dios por ser la fuerza, el sustento y la guía en mi vida.
- A mi esposa por su apoyo emocional a lo largo de todo este estudio.
- Le agradezco a mis padres por su apoyo incondicional brindado en estos años.
- A mis directores de proyecto, Dr. Juan Frausto Solís y Dr. Juan Javier González por todo el apoyo durante el desarrollo de esta investigación además de los valiosos conocimientos y la formación brindados a lo largo del desarrollo de este trabajo.
- A los miembros de mi comité tutorial, por sus puntuales aportaciones y contribuciones a la mejora continua de este trabajo
- A mi compañera Mirna Patricia por su amistad y apoyo en estos años.
- 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 los doctores que forman parte del núcleo académico del posgrado, ya que cada uno me brindo conocimientos que me permitieron crecer como profesionalista, además de su apoyo en múltiples trámites permitiendo la obtención de este grado.
- Al Consejo Nacional de Ciencia y Tecnología, por el apoyo ofrecido con de la beca de manutención, con la que pude solventar los gastos personales y de los estudios en el tiempo de duración del estudio.
- Le agradezco a todas las personas que de una manera u otra apoyaron para que este logro fuera posible, a todos los que directa e indirectamente me impulsaron a emprender esta etapa.

Gracias.

# **Pronóstico de Series de Tiempo para el mercado de energía mexicano**

Lemuel Rodriguez Moya

## **Resumen**

La combinación de métodos de pronóstico es una técnica muy extendida que ha sido útil para diseñar métodos híbridos en varias áreas. Sin embargo, en los tiempos modernos se requieren cada vez más aplicaciones más precisas, y nuevas técnicas de hibridar. Muchos enfoques han surgido para resolver este problema. La principal dificultad de esta técnica es encontrar la metodología más adecuada para combinar métodos de pronóstico. Este trabajo presenta una nueva metodología denominada FCTA (Método de combinación de pronósticos con Threshold Accepting). Esta metodología ensambla un conjunto de métodos de pronóstico y utiliza un algoritmo Threshold Accepting para optimizar la ponderación de cada pronóstico en la combinación. FCTA parte de una ponderación inicial y tiene como objetivo encontrar la mejor ponderación de cada método con el objetivo de mejorar la precisión del pronóstico global. Para probar FCTA se usan conjuntos de datos tomados de la competencia M4-Makridakis además FCTA se compara con los mejores métodos de pronóstico individuales y otros exitosos métodos de la actualidad. Los Resultados de la experimentación muestran que FCTA supera a los mejores métodos individuales además es equivalente o mejor que las mejores metodologías del área de pronóstico.

# **Pronóstico de Series de Tiempo para el mercado de energía mexicano**

Lemuel Rodriguez Moya

## **Abstract**

Combining forecasting methods is a widespread technique that has been useful for designing hybrid methods in several areas. However, in modern times, more and more precise applications and new hybridization techniques are required. Many approaches have emerged to solve this problem. The main difficulty with this technique is finding the most suitable methodology for combining forecasting methods. This paper presents a new methodology called FCTA (Method of Combining Forecasts with Threshold Accepting). This methodology assembles a set of forecasting methods and uses a Threshold Accepting algorithm to optimize the weighting of each forecast in the combination. FCTA starts with an initial weighting and aims to find the best weighting for each method with the aim of improving overall forecast accuracy. To test FCTA, datasets taken from the M4-Makridakis competition are used. In addition, FCTA is compared with the best individual forecasting methods and other successful methods of today. The results of the experimentation show that FCTA outperforms the best individual methods and is equivalent to or better than the best methodologies in the forecasting area.

# Índice General

Resumen .....	V
Abstract.....	VI
Índice Tablas.....	IX
Índice de Figuras .....	XI
Índice de Ecuaciones .....	XII
Nomenclatura.....	XIII
<b>1 Introducción .....</b>	<b>1</b>
1.1 Planteamiento del problema.....	4
1.2 Hipótesis .....	6
1.3 Objetivos.....	6
1.3.1 Objetivo General.....	6
1.3.2 Objetivos Específicos .....	6
1.4 Justificación del estudio.....	6
1.5 Alcances y Limitaciones.....	7
1.5.1 Alcances.....	7
1.5.2 Limitaciones .....	8
1.6 Organización de la tesis .....	8
<b>2 Antecedentes .....</b>	<b>9</b>
2.1 Marco Teórico.....	9
2.1.1 Series de Tiempo .....	9
2.1.2 Métodos de Pronóstico .....	10

2.1.3	Métodos Híbridos .....	13
2.1.4	Combinación de pronósticos .....	13
2.1.5	Medidas de error de Pronóstico .....	15
2.1.6	Algoritmo Treshold Accepting.....	17
2.1.7	Pronóstico de Energía.....	19
2.2	Estado del Arte.....	21
<b>3</b>	<b>Metodología.....</b>	<b>28</b>
3.1	Combinación de Pronóstico FCTA.....	29
3.1.1	Arquitectura de FCTA .....	31
3.1.2	Complejidad de TAE .....	44
<b>4</b>	<b>Análisis y Resultados.....</b>	<b>45</b>
4.1	Equipo y Material .....	45
4.1.1	Equipo de Cómputo .....	45
4.1.2	Software Utilizado.....	46
4.1.3	Datasets.....	46
4.2.1	Pronóstico con FCTA para la M4 de Makridakis.....	48
4.2.2	Comparativa entre FCTA y otros ensambles.....	64
4.2.3	Pronóstico de Energía con FCTA.....	68
<b>5</b>	<b>Conclusiones y trabajos futuros.....</b>	<b>73</b>
5.1	Conclusiones.....	73
5.2	Trabajos Futuros .....	74
5.3	Publicaciones derivadas del presente trabajo.....	74
	Bibliografía.....	75

# Índice Tablas

Tabla 2.1: Estado del Arte de Combinación de Pronósticos .....	27
Tabla 4.1: Series Seleccionadas de M4 competition .....	47
Tabla 4.2: Conjunto de métodos de pronóstico para probar FCTA.....	51
Tabla 4.3: Métodos de referencia usados para la comparativa con FCTA.....	51
Tabla 4.4: Resultados de errores de pronóstico del promedio sMAPE para todos los métodos individuales.....	52
Tabla 4.5: Resultados Friedman y Wilcoxon con distribución de peso normal .....	55
Tabla 4.6: Resultados Friedman y Wilcoxon con distribución de peso uniforme.....	56
Tabla 4.7: Prueba Wilcoxon con rangos para el subconjunto <i>Yearly</i> para la distribución de peso normal estándar .....	57
Tabla 4.8: Prueba Wilcoxon con rangos para el subconjunto <i>Quarterly</i> para la distribución de peso normal estándar .....	57
Tabla 4.9: Prueba Wilcoxon con rangos para el subconjunto <i>Monthly</i> para la distribución de peso normal estándar .....	57
Tabla 4.10: Prueba Wilcoxon con rangos para el subconjunto <i>Weekly</i> para la distribución de peso normal estándar .....	58
Tabla 4.11: Prueba Wilcoxon con rangos para el subconjunto <i>Daily</i> para la distribución de peso normal estándar .....	58
Tabla 4.12: Resultados de la prueba estadística W de Wilcoxon para todos los subconjuntos para la distribución de peso normal estandar.....	58
Tabla 4.13: Prueba Wilcoxon con rangos para el subconjunto <i>Yearly</i> para la distribución de peso uniforme .....	59
Tabla 4.14: Prueba Wilcoxon con rangos para el subconjunto <i>Quarterly</i> para la distribución de peso uniforme .....	59
Tabla 4.15: Prueba Wilcoxon con rangos para el subconjunto <i>Monthly</i> para la distribución de peso uniforme .....	60
Tabla 4.16: Prueba Wilcoxon con rangos para el subconjunto <i>Weekly</i> para la distribución de peso uniforme .....	60

Tabla 4.17: Prueba Wilcoxon con rangos para el subconjunto <i>Daily</i> para la distribución de peso uniforme .....	60
Tabla 4.18: Resultados de la prueba estadística W de Wilcoxon para todos los subconjuntos para la distribución de peso uniforme .....	61
Tabla 4.19: Resultados de errores promedio de pronóstico sMAPE y tiempos de ejecución para todos los conjuntos .....	65
Tabla 4.20: Friedman y Wilcoxon para los resultados del horizonte H1 y Hn .....	66
Tabla 4.21: Prueba W de Wilcoxon con rangos para el horizonte H1. ....	67
Tabla 4.22: Prueba W de Wilcoxon con rangos para el horizonte Hn .....	67
Tabla 4.23: Resultados estadísticos W de Wilcoxon para todos los subconjuntos .....	67
Tabla 4.24: Comparativa de errores sMAPE promedio de pronóstico horizonte 10.....	70
Tabla 4.25: Comparativa de errores sMAPE promedio de pronóstico horizonte 1.....	71

# Índice de Figuras

Figura 2.1: Serie de Demanda Energética .....	10
Figura 2.2: Métodos de Híbridos Pronóstico.....	14
Figura 2.3: Combinación de Pronósticos.....	14
Figura 2.4: Curva de demanda de energía Eléctrica en México.....	20
Figura 2.5: Zonas de Potencia del Sistema Eléctrico Nacional.....	20
Figura 2.6: Zonas y sub-zonas de Potencia del Sistema Eléctrico Nacional.....	21
Figura 3.1: Metodología general para el pronóstico.....	28
Figura 3.2: Combinación de pronóstico .....	30
Figura 3.3: Metodología FCTA.....	32
Figura 3.4: Diagrama de Secuencia de la metodología FCTA.....	33
Figura 3.5: Proceso de Ponderación de FCTA .....	35
Figura 3.6: Ejemplo de distribución normal estándar .....	36
Figura 3.7: Ejemplo de disminución de quantum de los pesos .....	40
Figura 3.8: Disminución del espacio de búsqueda .....	41
Figura 3.9: Golden Ratio (GR) scheme .....	42
Figura 3.10: Estrategia de Reheat.....	43
Figura 4.1: Arquitectura de FCTA Ampliada.....	49
Figura 4.2: División de la Serie de Tiempo.....	50
Figura 4.3: Comparativa de FCTAU y FCTAN versus métodos de referencia .....	53
Figura 4.4: Comparativa entre los métodos propuestos .....	54
Figura 4.5: Comparación de los métodos propuestos frente a los tres mejores métodos individuales.....	54
Figura 4.6: Comparativa entre FCTA y el mejor método para el subconjunto Yearly .....	62
Figura 4.7: Comparativa entre FCTA y el mejor método para el subconjunto Quarterly....	62
Figura 4.8: Comparativa entre FCTA y el mejor método para el subconjunto Monthly .....	63
Figura 4.9: Comparativa entre FCTA y el mejor método para el subconjunto Weekly.....	63
Figura 4.10: Comparativa entre FCTA y el mejor método para el subconjunto Daily .....	64
Figura 4.11: Relación de error del FCTAN propuesto en comparación con otros algoritmos de pronóstico de conjuntos .....	68

Figura 4.12: División de las series de Tiempo experimento energía .....	69
Figura 4.13: Comparativa entre FCTAU y FCTAN versus métodos individuales .....	69
Figura 4.14: Comparativa FCTA Contra CENACE.....	72

## Índice de Ecuaciones

Ecuación	
Ecuación 2.1 .....	11
Ecuación 2.2 .....	11
Ecuación 2.3 .....	11
Ecuación 2.4 .....	11
Ecuación 2.5 .....	12
Ecuación 2.6 .....	15
Ecuación 2.7 .....	15
Ecuación 2.8 .....	16
Ecuación 2.9 .....	16
Ecuación 2.10 .....	16
Ecuación 2.11 .....	16
Ecuación 2.12 .....	17
Ecuación 2.13 .....	17
Ecuación 3.1 .....	30
Ecuación 3.2 .....	34
Ecuación 3.3 .....	35
Ecuación 3.4 .....	41
Ecuación 4.1 .....	68

# Nomenclatura

En esta sección se contienen los símbolos o términos que se emplean en ecuaciones, diagramas o figuras que representan conceptos que se emplean en el documento presentado. Si no hay símbolos o términos que listar, no es necesario incluir esta sección.

---

---

## Introducción

La estimación de un evento futuro ha cautivado constantemente a los seres humanos dado que en su naturaleza siempre buscan entender que ocurrirá con relación a muchos aspectos de la vida cotidiana. Debido a esto el conocer algún aspecto del futuro es importante para la planificación de eventos. A lo largo de los años muchas técnicas de pronóstico se han desarrollado para diversas ramas científicas y económicas. El pronóstico es un tema que se presenta regularmente en diferentes organizaciones de todo el mundo a la hora de tomar decisiones, por ejemplo, cuando se planifica dónde se requiere invertir capital o se desea anticipar el comportamiento del mercado antes de una inversión. El éxito del pronóstico de valores futuros consiste en la relación que existe entre el valor actual de una variable y los valores pasados de la misma. El propósito común de los pronósticos es auxiliar en la toma de decisiones. Actualmente en la era de la tecnología y la información, la mayoría de las actividades generan cantidades masivas de información que son almacenados en bases de datos, donde se pueden manipular y explotar para obtener información útil y relevante. Dentro del área financiera el interés de los inversores siempre ha radicado en el pronóstico de los activos financieros para uno o más periodos en el futuro, a partir de los precios que han tenido en el pasado. Por eso las bolsas de valores alrededor del mundo son de gran importancia para las economías de cada país.

Un gran número de las ganancias económicas de las empresas de una nación provienen de transacciones efectuadas en las diferentes bolsas del mundo. Con esto los inversionistas que trabajan activamente cotizando acciones en la bolsa de valores buscan siempre obtener la mejor información respecto a los diferentes activos cotizados, siempre buscando invertir en activos donde puedan obtener mayores ganancias con el menor riesgo posible [1].

Para el pronóstico se han desarrollado algunos métodos que por su buen desempeño siguen vigentes y son considerados clásicos, entre ellos están: el Suavizamiento Exponencial [2] y la metodología de Box Jenkins [3] con los Métodos *ARIMA*. Entre los métodos de pronósticos de la categoría de aprendizaje automático (Machine Learning), los más desatacados son Redes Neuronales (*NN*) y *SVR* (Support Vector Regression) [4]. Algunas mejoras en los métodos anteriores se han centrado en combinaciones de los mismos buscando obtener mejores resultados. A pesar de todo el trabajo realizado, la experiencia de pronóstico indica que la mayoría de los métodos de pronóstico individuales pueden tener un buen rendimiento para algún tipo específico de series temporales pero pueden no ser tan eficientes para la generalidad de estas [5], [6].

El pronóstico es un problema complejo, y durante mucho tiempo, la comunidad de pronósticos ha centrado su investigación en encontrar mejores técnicas minimizando el error del pronóstico.

Más recientemente, se han propuesto métodos híbridos que combinan diferentes métodos de pronóstico [6]–[14] y han presentado un buen rendimiento en muchas áreas. Este rendimiento se debe a que las técnicas de hibridación le permiten a los métodos individuales compensarse entre sí sus debilidades y con esto explorar mejor la información histórica de las series temporales [15].

De entre las hibridaciones más recientes para el pronóstico se encuentran combinaciones de los métodos tradicionales como *ARIMA*, *SES*, Box & Jenkins y técnicas de aprendizaje automático.

Métodos ANN (Artificial Neural Networks) y *ARIMA*[16],[17],[18]

Métodos ANN (Artificial Neural Networks) y Métodos Box & Jenkins [19], [20]

*ARIMA* y *SVR* [21], [22],[22][23]

ANN y *SES* Suavizamiento Exponencial [24],[25]

Metodologías comunes aplican filtros y agrupan los mejores métodos individuales en muchas áreas[26]–[32]; entre ellas, las dos con mejor desempeño son las metodologías *PP-MA* [26]y

LA [27]. El primero utiliza una etapa de filtrado y el Criterio de Información Bayesiana (BIC) para seleccionar los mejores modelos de pronóstico para cada ventana de la serie temporal[26]; el segundo, utiliza un nuevo enfoque de autómatas de aprendizaje que deriva dos algoritmos de conjunto principales llamados LA-SW y LA-MW [27]. Otros métodos de conjunto son EMD-PE-ANN y LSTM-KF aplicados a la velocidad del viento y al consumo de energía, respectivamente [29], [30].

Dentro de las hibridaciones más recientes se encuentra la combinación de pronósticos. La motivación de la combinación de pronósticos surge del hecho de que todos los modelos de un proceso de generación de datos del mundo real están mal especificados [15], por eso elegir solo un método de pronóstico supone un riesgo si el rendimiento y los datos del modelo varían a lo largo del tiempo [12]. El combinar varios métodos de pronóstico tiene como objetivo disminuir el error y mejorar la precisión. Para realizar una combinación de métodos de pronóstico se necesita principalmente seleccionar los métodos de pronóstico que se van a combinar y la estrategia de combinación. La combinación de múltiples métodos de pronóstico implica seleccionarlos y determinar su mejor distribución de pesos [12], [33]–[37]. La selección de un solo método de pronóstico puede verse como un riesgo cuando el rendimiento de los modelos y los datos varían con el tiempo [12], y es mejor utilizar dos o más métodos individuales en una combinación. La metodología de combinación de Bates-Granger (B&G) tiene bases sólidas [7], y ha inspirado enfoques modernos exitosos [37], [38] incluida la metodología propuesta en este documento.

Una de las técnicas utilizadas para combinar métodos consiste en el uso del meta-aprendizaje, algunas obras más recientes como [12], [33], [35], [36] han combinado el aprendizaje automático con el cálculo de las características de las series, y han logrado buenos resultados. Recientemente *FFORMA* (Selección de modelo de pronóstico basada en características) propuso utilizar características de series temporales asociadas con meta aprendizaje para seleccionar el mejor método de pronóstico.

Finalmente, el enorme desafío en el pronóstico es desarrollar un método con altos requisitos que tenga un amplio espectro, buen rendimiento en múltiples escenarios, conjuntos de datos de varias áreas y para diversos horizontes de pronóstico. La competencia M4 permite probar nuevas metodologías de pronóstico en diferentes escenarios, como se propone en este documento. La competencia M4 se ha convertido en una importante referencia en el área de

los pronósticos ya que incluye los resultados de pronóstico para cien mil series, con los mejores métodos de pronóstico en la actualidad [38].

Por su parte dentro del mercado de Energía Mexicano en la actualidad la electricidad (potencia y energía) es una mercancía susceptible de ser comprada, vendida y comercializada. Este mercado funciona de manera análoga al mercado financiero en el cual se buscan efectuar las compras, a través de ofertas de compra y venta en operaciones a corto plazo, generalmente en forma de permutas. En este mercado se aplican los principios de la oferta y la demanda para fijar el precio. Las operaciones a largo plazo son contratos similares a los acuerdos de compra de energía y generalmente se consideran las transacciones privadas bilaterales entre las empresas. Dado que las series de tiempo de demanda y precios de electricidad son no lineales y presentan una periodicidad diaria y semanal por lo ende, un modelo desarrollado para el mercado financiero también puede ser útil para predecir valores futuros como los precios de la energía en series de tiempo del mercado de energía mexicano. Pronosticar es un reto dada la complejidad que pueden presentar las series de tiempo, por lo tanto, se espera que los algoritmos desarrollados para el mercado financiero encuentren soluciones de alta calidad con un consumo de tiempo que compita con métodos tradicionales de pronóstico de series de tiempo de energía actuales.

En el proceso de pronóstico, existen una serie de requisitos recomendadas para obtener un buen resultado. Primero se debe pre procesar la serie temporal para eliminar datos incorrectos y detectar valores atípicos [39], [40]. Otro requisito muy recomendable a la hora de pronosticar es utilizar no solo el mejor método de pronóstico, sino también colocar los mejores métodos individuales para el tipo de serie con la que se está trabajando. Otro buen requisito a tener en cuenta cuando se utiliza un método híbrido para pronóstico es determinar la correcta ponderación de cada método de pronóstico en una hibridación. Una ponderación incorrecta puede impactar negativamente el resultado final e invalidar la ventaja propia de una hibridación obteniéndose errores mayores.

## 1.1 Planteamiento del problema

Las series de tiempo son conjuntos de datos registrados u observados en intervalos periódicos de tiempo de un fenómeno variable. Es decir, una serie de tiempo registra valores consecutivos en el tiempo y estos registros pueden estar bajo periodos de tiempo, por

ejemplo, cada hora, diarios, semanales, mensuales, trimestrales o anuales. Considerando siempre que, si el primer valor es por ejemplo mensual, los demás datos en la serie tienen que ser mensuales. La idea más básica para obtener un pronóstico de las series de tiempo es que al dirigir una mirada al pasado a través de una serie de tiempo e interpretar esos datos históricos se puede planificar y pronosticar valores futuros [41]. Los datos organizados en una serie de tiempo son conocidos como variables de la serie de tiempo. Lo relevante en una serie de tiempo para la planificación es el comportamiento de la serie en su conjunto, no los datos como valores individuales. No obstante, el problema del pronóstico es un tema más complicado, porque en la actualidad no existe una solución explícita para el mismo, por lo que existe una gran área de oportunidad la cual se puede abordar con propuestas de heurísticas que obtengan una buena solución o la solución óptima del problema. Las heurísticas son algoritmos que buscan obtener soluciones aproximadas a problemas en los cuales no existe un método de solución, o bien donde los métodos de solución existentes requieren de un tiempo extremadamente grande. Por lo tanto, las heurísticas con meta aprendizaje pueden ser herramientas eficientes para lograr resultados de alta calidad en menos tiempo y con menos costo computacional. Además de que pueden aplicarse fácilmente a pronósticos de series de tiempo para muchas áreas como Finanzas, Economía, Meteorología etc. Un buen método de pronóstico es imprescindible en el área de energía tanto en el área de producción y comercialización. El pronosticar valores futuros de la demanda energética, precios de la energía permiten a las grandes empresas de energía determinar la estrategia de mercado que van utilizar tanto para producir como para distribuir la energía evitando gastos innecesarios, fallas del sistema y problemas como sobreproducción o escasez de energía.

En el área de energía se han desarrollado diversos modelos de pronósticos de series de tiempo que pueden ser aplicados a múltiples problemas dependiendo de las necesidades, pero no todos los problemas relativos al pronóstico de series de tiempo han sido resueltos. Inicialmente los pronósticos de series de tiempo se realizaban con modelos independientes unos de otros.

Hasta el momento actual no existe un modelo generalizado de aplicación universal de pronóstico dado que los estudios realizados no se han validado lo suficiente y se han desarrollado bajo circunstancias particulares. Especialmente en el mercado de energía mexicano, dado que es un mercado de energía emergente, no se han estudiado ampliamente las nuevas técnicas para el pronóstico de variables

## 1.2 Hipótesis

En función de la metodología usada actualmente para pronóstico de energía, se plantea la siguiente Hipótesis.

H: La metodología de pronóstico usada para mercados financieros con optimización heurística es aplicable al mercado de energía con un desempeño similar o superior a las metodologías en uso actualmente.

## 1.3 Objetivos

### 1.3.1 Objetivo General

Desarrollar un método híbrido basado en meta-aprendizaje, métodos clásicos y computación inteligente y meta heurísticas para el pronóstico en el mercado de energía con un desempeño similar o mejor a los métodos del estado del arte.

### 1.3.2 Objetivos Específicos

- Desarrollo de métodos de distribución de pesos (combinación de pronósticos)
- Pronóstico de series financieras, demanda de energía y otras series de tiempo
- Analizar técnicas de ensamble de métodos de pronóstico
- Diseño de una metodología nueva de ensamble para pronóstico.
- Aplicación de algoritmos heurísticos en pronóstico.
- Comparar resultados (literatura).

## 1.4 Justificación del estudio

El pronóstico de series de tiempo es un gran desafío dado que en la actualidad en las empresas y organizaciones se busca siempre tomar la mejor decisión. La importancia de conocer los valores futuros aumenta, por ejemplo, cuando se trabaja con series de tiempo financieras dado a la volatilidad propia del mercado financiero y los bruscos cambios que este puede sufrir. De manera análoga, cuando se trabaja con series del mercado de energía la estabilidad

del sistema eléctrico depende del correcto pronóstico de la demanda energética y parámetros relacionados con la producción energética de un país.

El propósito de desarrollar un método híbrido para realizar el pronóstico de series de tiempo consiste en demostrar que estos métodos poseen una eficiencia superior con resultados de alta calidad que los métodos clásicos usados ampliamente a lo largo de la historia. Diversos estudios apuntan a que los métodos híbridos combinando alguno de los métodos de pronóstico clásicos como *ARIMA*, *SES* y *SVR* tienen mejor desempeño.

Se plantea desarrollar un método híbrido de computación inteligente que haga uso de métodos clásicos y técnicas heurísticas de bajo costo computacional, sencillo de implementar, que presente resultados precisos y que sea de aplicación general para la mayoría de las variables de pronóstico. Dado que los estudios realizados no se han validado lo suficiente y se han desarrollado bajo circunstancias particulares.

Muchos estudios reportados en el Estado del Arte indican que métodos híbridos combinando métodos de pronóstico clásicos con meta aprendizaje y computación inteligente obtienen los mejores resultados y los primeros lugares en las competencias de pronóstico como la M4 [25], [37], [42], [43].

## 1.5 Alcances y Limitaciones

### 1.5.1 Alcances

- El método propuesto será un método de combinación de pronóstico que usará meta aprendizaje y optimización con técnicas meta heurísticas.
- El método de combinación podrá usar métodos clásicos y otras hibridaciones como base para realizar la combinación.
- El método de combinación podrá usar series de cualquier tipo financieras, energía u otra variable de interés.
- El método de combinación podrá compararse con cualquier otro método de pronóstico del estado del arte.

## 1.5.2 Limitaciones

- La comparación del método de combinación propuesto se desarrollará con aquellos trabajos del estado del arte cuyos desarrollos estén disponibles (ejecutables) o se tengan los algoritmos para ser reproducidos.
- Se utilizarán Series de competencias de pronósticos como M4, y series del Mercado de energía mexicano.

## 1.6 Organización de la tesis

El trabajo desarrollado en esta tesis, se encuentra organizado en 6 capítulos.

- **Capítulo 1.-** En este capítulo se presenta a una breve introducción del problema a resolver, se incluyen los objetivos propuestos, así como su justificación, alcances y limitaciones.
- **Capítulo 2.-** Se presenta la parte teórica del problema del pronóstico y el estado del Arte donde se revisan los trabajos de investigación más relevantes y recientes del área de pronósticos.
- **Capítulo 3.-** Se detalla la metodología propuesta para el pronóstico.
- **Capítulo 4.-** En este Capítulo se detallan, los equipos y conjuntos de datos que se utilizaron durante para la experimentación de la metodología propuesta
- **Capítulo 5.-** En este Capítulo se detallan los resultados obtenidos a partir de los experimentos realizados con un subconjunto de datos de las series de la competencia M4 y del área de energía de México.
- **Capítulo 6.-** En este Capítulo se especifican las conclusiones finales y los trabajos futuros para esta investigación

---

## Antecedentes

### 2.1 Marco Teórico

#### 2.1.1 Series de Tiempo

Una serie de tiempo es un conjunto de observaciones medidas secuencialmente a través del tiempo [44] de un fenómeno cambiante. Las observaciones que se denotan como  $Y_t$  donde el subíndice  $t$  corresponde a un dato almacenado en el tiempo  $t$  [45]. El análisis de las series de tiempo financiera se refiere a la teoría y la práctica de la valoración de activos a lo largo del tiempo [46] o a cualquier variable que conste de datos reunidos, registrados u observados sobre incrementos sucesivos de tiempo [41]. La idea subyacente del pronóstico de valores futuros es que existe una relación entre el valor actual de una variable y los valores pasados de la misma. Dado que se expresa la secuencia de las observaciones de una serie de tiempo por  $X_{(t)} = (X_1, X_2, \dots, X_m)$  entonces la expresión  $X(t+h)$  denota el horizonte de pronóstico para el período  $h$  en el tiempo  $t$ , donde  $h = 1$  significa un paso adelante y  $h > 1$  significa múltiples pasos en el pronóstico por delante. En la Figura 2.1 se presenta un ejemplo de una serie de histórica demanda energética (azul) contra el pronóstico 10 valores en el futuro (naranja) realizado con un método de pronóstico.



Figura 2.1: Serie de Demanda Energética

### 2.1.2 Métodos de Pronóstico

Los métodos de pronóstico son utilizados con el objetivo de estimar una variable en el futuro. Los métodos de pronóstico de las series de tiempo pueden ser clasificados en 3 tipos: critico, univariados y multivariado [47].

El pronóstico crítico está basado en la interpretación de cualquier información relevante. En los métodos univariados el pronóstico depende solo en valores actuales o pasados de una sola serie de tiempo. En los métodos multivariados el pronóstico para una variable depende, al menos parcialmente, de los valores de las variables de otras series de tiempo. En [48] los métodos univariados y multivariados se organizan en dos subclases:

- Métodos estadísticos tradicionales (Suavizamiento exponencial, ARIMA, etc.
- Métodos de computo inteligente como redes neuronales, máquinas de soporte vectorial y algoritmos heurísticos,

Existen hibridaciones la cuales pueden incluir por ejemplo el uso de redes neuronales con métodos clásicos de pronóstico ARIMA, ANN, o SVR con ARIMA, SVR con algoritmos genéticos.

### 2.1.2.1 Modelos ARMA

Uno de los modelos más importantes y más popularizados para series de tiempo es llamado ARMA, que es la combinación del modelo auto regresivo (AR) y de medias móviles (MA). El modelo AR es representado en la ecuación 2.1.

$$Y_t = c + \sum_{i=1}^p \Phi_i X_{t-i} + \varepsilon_t \quad 2.1$$

Donde  $Y_t$  es una variable estimada en el periodo  $t$  en términos del primer dato  $p$  en la serie de tiempo;  $\Phi_1, \dots, \Phi_p$  son los parámetros del modelo de regresión simple,  $c$  es una constante y  $\varepsilon_t$  es un término de error. Muchos autores omiten el término constante, para fines de simplificación.

Este modelo es denotado como  $AR(p)$  porque los datos  $p$  son tomados en cuenta. En el modelo MA,  $Y_t$  es estimado cerca del promedio  $\mu$  de la información de la serie de tiempo; esto es hecho a través de la ponderación de los errores  $\varepsilon_t$  en varios periodos pasados  $q$  al periodo  $t$  (2.2).

$$Y_t = \mu_t - \Phi_1 \varepsilon_{t-1} - \Phi_2 \varepsilon_{t-2} - \dots - \Phi_q \varepsilon_{t-q} \quad 2.2$$

Dado que número de errores en la ecuación es  $q$ , este modelo es representado como  $MA(q)$ .

### 2.1.2.2 ARIMA

El modelo ARMA es la combinación del modelo AR y MA, mientras que el modelo ARIMA incluye el componente estacionario al modelo ARMA. Cuando una serie de tiempo no es estacionaria, puede ser integrada definiendo una nueva variable de la siguiente manera (2.3).

$$Z_t = y_t - y_{t-1} \quad 2.3$$

donde:  $t = (2; 3; \dots; n)$  Las segundas diferencias son determinadas por (2.4).

$$Z_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \quad 2.4$$

Para  $t = (3; 4; \dots; n)$  El modelo resultante, integrando estacionalidad es conocido como ARIMA (Auto-Regresive Integrated Moving Average) el cual es un modelo ARMA con problemas no estacionales. Para poder integrar el componente estacional, ARMA es adaptado y se genera el modelo SARMA. Al combinar el modelo ARIMA y SARMA es generado el modelo SARIMA (Seasonal AutoRegresive Integrated Moving Average).

Los retrasos en los procesos y las perturbaciones aleatorias pueden ser representados por una forma periódica en cada patrón estacional. Los retrasos estacionales ocurren dado a una dependencia en periodos similares de años sucesivos. Otro modelo es capaz de representar la tendencia correctamente, estacionalidad y componentes no estacionales de una serie de tiempo. Este modelo es conocido como el modelo  $ARIMA(p; d; q) \times SARIMA(P; D; Q)$ , dónde:

$p$ : orden del modelo  $AR$ .

$d$ : diferenciación del orden en las partes regulares y no regulares de una serie estacionaria.

$q$ : orden del promedio móvil del  $MA$ .

$P$ : orden del modelo  $SAR$ .

$D$ : diferenciación del orden en las partes estacionales de las series.

$Q$ : orden del promedio móvil estacional del  $SMA$ .

### 2.1.2.3 Suavizamiento Exponencial

El Suavizamiento exponencial es un método que puede pronosticar en donde los datos no tienen una tendencia predecible, este usa un promedio móvil con peso exponencial de los valores observados con anterioridad [49]. El Suavizamiento exponencial suaviza (promedia) los valores pasados de una serie de forma exponencialmente decreciente. Donde la observación más reciente tiene un peso  $\alpha$  ( $0 < \alpha < 1$ ), la siguiente recibe menos peso ( $1 - \alpha$ ), la observación de dos periodos anteriores tendría  $\alpha(1 - \alpha)$  y así de manera sucesiva. La forma de realizar el nuevo pronóstico se muestra en la ecuación **¡Error! No se encuentra el origen de la referencia..**

$$Y_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t \quad 2.5$$

donde período

$Y_{t+1}$  = es el valor del pronóstico para el periodo  $t + 1$

$\alpha$  = es la constante de Suavizamiento ( $0 < \alpha < 1$ )

$\hat{Y}_t$  = es la nueva observación

$Y_t$  = es el antiguo valor suavizado o pronóstico para el período  $t$

### **2.1.3 Métodos Híbridos**

Los métodos híbridos buscan revolucionar el pronóstico de series temporales combinando lo mejor de los métodos estadísticos y de aprendizaje automático. La idea fundamental es que la hibridación compensa las limitaciones de un enfoque con las fortalezas del otro. Por ejemplo, la efectividad de los métodos estadísticos con disponibilidad limitada de datos puede compensar los amplios requerimientos de datos de métodos de aprendizaje automático. Aparte de eso, la consideración del conocimiento a priori puede simplificar la tarea de pronóstico esperada y disminuir el esfuerzo computacional. Además, los métodos híbridos pueden incorporar el aprendizaje cruzado que es la capacidad de los modelos de aprender a partir de múltiples series cómo pronosticar con una mayor precisión los valores de una serie individual, una capacidad de la que carecen muchos métodos estadísticos.

Finalmente, los métodos híbridos proporcionan una solución al dilema de la asunción de linealidad. Como las series temporales del mundo real pueden ser puramente lineales, puramente no lineales, o a menudo contienen una combinación de esos dos patrones, los métodos híbridos pueden ser efectivos cuando los enfoques tradicionales alcanzan sus límites. Dado lo anterior el objetivo principal de los métodos híbridos es compensar deficiencias individuales de cada método y aprovechar sus ventajas. En la Figura 2.2 se muestra algunos de los métodos que comúnmente se utilizan en hibridaciones.

### **2.1.4 Combinación de pronósticos**

La combinación de métodos de pronóstico independientes se realiza con el objetivo de obtener un método compuesto con una mejor precisión (menor error) que cada uno de los métodos individuales (Figura 2.3). La calidad de combinación de pronósticos depende en parte que cada uno de los métodos individuales posea un buen rendimiento. La idea de combinar pronósticos surge del conocimiento de que el pronóstico obtenido a partir de cada método puede no cubrir todos los aspectos relevantes de la serie y que pueden existir métodos que evalúan mejor una serie que otros o sea pueden capturar características de la serie que otros no pueden. Los modelos individuales proporcionan valores de pronóstico de un conjunto de series que se combinan en un pronóstico final. Como resultado, un método combinado puede producir buenos resultados [11].



Figura 2.2: Métodos de Híbridos Pronóstico

Un tema clave en la literatura es cuales son los métodos más adecuados que deben usarse dentro de una combinación. Generalmente se usan aquellos que tengan el menor error de pronóstico y una amplia aplicabilidad para el caso requerido. En otras palabras, la mejor práctica es combinar varios métodos en función de su rendimiento [11],[50]. Formalmente podemos describir la combinación de pronósticos como un problema de optimización donde el objetivo es elegir la mejor distribución de pesos  $\omega_i$  que minimicen el MAPE de la combinación.

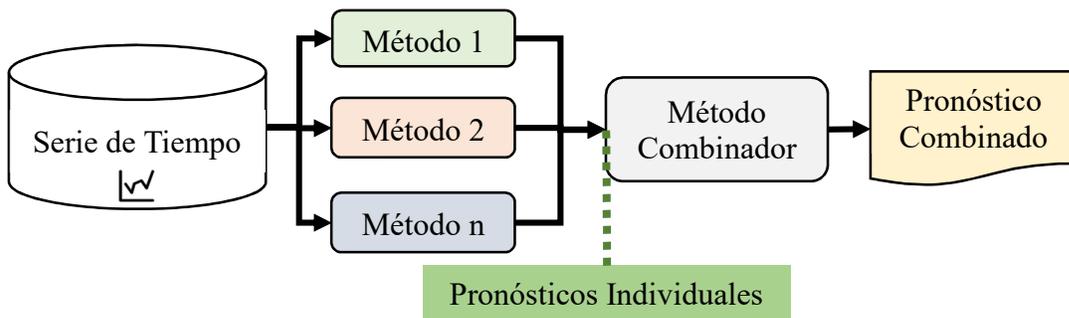


Figura 2.3: Combinación de Pronósticos

### 2.1.5 Medidas de error de Pronóstico

Si queremos evaluar un pronóstico y estimar el funcionamiento de un modelo se requieren comparar los valores pronosticados contra los valores reales observados. Para esto se usan diferentes medidas de precisión del pronóstico (comúnmente llamadas medidas de error). Con estas técnicas se mide la confiabilidad de una metodología de pronóstico permitiendo diferenciarla de otras de acuerdo a su rendimiento [51]. Básicamente el error de pronóstico se define como se muestra en la ecuación (2.6).

$$E_t = Y_t - \hat{Y}_t \quad 2.6$$

donde

$E_t$  = Error de pronóstico para el período t

$Y_t$  = Valor real de la serie en el período t (observación actual)

$\hat{Y}_t$  = Valor del pronóstico en el período t

#### 2.1.5.1 MAD (Desviación Media Absoluta)

La desviación media Absoluta promedia los valores absolutos de los errores de pronóstico. Es una medida de exactitud, su principal ventaja es que utiliza las mismas unidades que la serie original para medir el pronóstico [49]. La formulación matemática de esta medida se muestra en la ecuación 2.7.

$$MAD = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t| \quad 2.7$$

donde

$n$  = el número de punto de datos

#### 2.1.5.2 MSE (Error Cuadrático Medio)

El error cuadrático medio (*MSE*), consiste en elevar al cuadrado cada error o residuo sumarlos y dividirlos por el número de observaciones. La característica más notable de esta medida de error es que penaliza grandes errores de pronóstico ya que eleva al cuadrado el error [49]. La formulación matemática de esta medida se muestra en la ecuación 2.8.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad 2.8$$

### 2.1.5.3 *RMSE* (Raíz cuadrada del Error Cuadrático Medio)

Al igual que *MSE*, La raíz cuadrada del error cuadrado medio (*RMSE*), esta medida de error es que penaliza grandes errores de pronóstico. Esta medida es de más fácil interpretación y también es muy comúnmente usada para evaluar métodos de pronóstico [52]. La formulación matemática de esta medida se muestra en la ecuación 2.9.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad 2.9$$

### 2.1.5.4 *MAPE* (Error Porcentual Absoluto Medio)

El error porcentual absoluto medio (*MAPE*), consiste en obtener el promedio de los valores de error absolutos. el resultado de estos se multiplica por 100 y se expresa como un porcentaje de los valores reales. Se puede decir que el *MAPE* no se expresa en unidades si no que es un porcentaje. Esta medida es muy usada en el área de pronósticos y muy útil para cálculo de precisión de una técnica o más de una en diferentes series de tiempo [53]. La formulación matemática de esta medida se muestra en la ecuación 2.10.

$$MAPE = \left[ \frac{1}{n} * \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{|Y_i|} \right] * 100(\%) \quad 2.10$$

### 2.1.5.5 *sMAPE* (Error Porcentual Absoluto Medio Simétrico)

El error *sMAPE* [54]. La formulación matemática de esta medida se muestra en la ecuación 2.11.

$$sMAPE = \left[ \frac{2}{h} * \sum_{t=n+1}^{n+h} \frac{|Y_t - \hat{Y}_t|}{(|Y_t| + |\hat{Y}_t|)} \right] * 100(\%) \quad 2.11$$

$h$  = el horizonte de pronóstico

### 2.1.5.6 MASE

El error *MASE* [54]. La formulación matemática de esta medida se muestra en la ecuación 2.12.

$$MASE = \frac{1}{h} \frac{\sum_{t=n+1}^{n+h} |Y_t - \hat{Y}_t|}{\frac{1}{n-m} \sum_{t=m+1}^n |Y_t - Y_{t-m}|} \quad 2.12$$

donde

$m$  = el intervalo de tiempo entre observaciones sucesivas consideradas

### 2.1.5.7 Evaluación de la calidad de la clasificación

Los investigadores utilizan la matriz de confusión para evaluar la calidad de la clasificación. La matriz de confusión contiene las clases de falsos positivos y negativos, comparando con la mejor selección real para cada serie, la siguiente medida utilizada en la ecuación 2.13.

$$Acc = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad 2.13$$

Dónde

TP: Verdadero Positivo

TN: Verdadero Negativo

FP: Falso positivo

FN: Falso negativo

## 2.1.6 Algoritmo Treshold Accepting

El algoritmo Threshold Accepting (TA) propuesto por Dueck-Scheuer [55] se ha aplicado para resolver muchos problemas de optimización NP-hard de diferentes áreas. TA es una variación del recocido simulado. De la misma manera que en SA, el algoritmo TA representa una analogía de un proceso termodinámico donde un metal se calienta y se enfría lentamente. Este proceso permite que el metal varíe sus propiedades físicas. En el algoritmo TA, la función objetivo es encontrar una mejor solución que mejore la actual. El algoritmo es una búsqueda local clásica que comienza con una solución aleatoria factible y luego busca aleatoriamente una nueva solución en el espacio de búsqueda, aceptándola cuando esta mejora la función objetivo. TA tiene los siguientes elementos (Algoritmo 2.1):

- a) La función objetivo. Eso es  $f(x)$ , el error de pronóstico en el Algoritmo 2.1.
- b) La estructura de vecindad. Define cómo pasar de una solución actual a una nueva. Esta estructura es sencilla en TA, actuando como la función *Generar* en la línea seis.
- c) El criterio de deterioro: este define cuando aceptar una mala solución. El recocido simulado utiliza el criterio de Boltzmann. En cambio, TA simplifica la forma de aceptar soluciones incorrectas utilizando un rango de deterioro de la función objetivo. Este criterio se ejecuta en la línea siete de este algoritmo.
- d) El criterio de parada. Se define idealmente con la diferencia "delta" de los errores de pronóstico para las dos soluciones vecinas. Está representado por  $\Delta$  en las líneas seis y siete. y la temperatura actual. Se representa como la regla en la línea 11.

En el algoritmo TA, la función objetivo propone encontrar la solución más cercana al óptimo. TA busca aleatoriamente una nueva solución en el espacio de búsqueda, aceptándola cuando se mejora la función objetivo. Este algoritmo consta de dos ciclos:

- El ciclo externo: está controlado por el parámetro de temperatura que varía desde la temperatura inicial hasta la temperatura final. La primera es demasiado alta (por ejemplo, 100, o 1000), y la segunda es demasiado pequeña (cerca a cero).
- El ciclo interno: este ciclo comúnmente aumenta el número de iteraciones con una tasa fija o variable, llamada beta.

---

**Algoritmo 2.1** Threshold Accepting clásico

---

- (1) **Inicialización:**  $n_{Steps}$ ,  $\alpha$ , convergence=falso;  $k=1$
  - (2) Calcular la secuencia threshold  $T_r$ , Calcular  $f(x^{old}) = f(x^r)$
  - (3) Generar aleatoriamente la solución actual  $x^{new} \in X$
  - (4) **while** convergence= falso **do**
  - (5)     **for**  $i = 1 : n_{Steps}$  **do**
  - (6)         Generar  $x^n \in N(x^c)$  y Calcular  $\Delta = f(x^{new}) - f(x^{old})$
  - (7)         **if**  $\Delta < T_k$  **then**  $x^{old} = x^{new}$ ;
  - (8)     **end for**
  - (9)      $k = k + 1$ ;
  - (10)      $T_k = \alpha T_k$
  - (11)     **if**  $\Delta \leq \epsilon$ ,  $0 k \geq N_{max}$ , convergence=verdadero.
  - (12)     **end while**
  - (13)  $x^{sol} = x^{old}$
-

## **2.1.7 Pronóstico de Energía**

El pronóstico de energía eléctrica en un sistema eléctrico de potencia, es de suma importancia para el sector energético para la planificación y distribución de la energía en los diferentes entornos ya sean el Industrial, el comercial o el doméstico. Uno de los objetivos básicos de un buen pronóstico en el sector eléctrico es mantener el sistema estable evitando los riesgos que conllevan a una sobreproducción o baja capacidad de producción de energía porque cualesquiera de estos extremos representan para el sector energético situaciones que propician pérdidas económicas. El pronóstico en el sector energético es de suma importancia ya que a partir de este se puede garantizar que no existan cortes en la distribución de la energía por falta de producción además de proporcionar una idea clara de donde y cuando se deben para llevar a cabo programas de mantenimiento, de la infraestructura actual como programas de actualización y ampliación de la infraestructura existente con el objetivo de satisfacer la creciente necesidad de energía eléctrica en cualquier sistema eléctrico de potencia. El principal coeficiente, por lo tanto, que se debe pronosticar dentro de un sistema eléctrico de potencia es la demanda.

La información histórica de la demanda eléctrica de un sistema eléctrico es la base para el éxito del pronóstico de energía. Además de la información histórica existen algunas otras variables relevantes como la temperatura diaria, variables socioeconómicas como el producto interno bruto (PIB), la población del área, el crecimiento económico y el número de consumidores que permiten conocer más sobre el comportamiento de la demanda energética de un país. El pronóstico de la demanda de energía de un país es la base para que se pueda lograr el objetivo de gestionar eficientemente la energía, operando el sistema eléctrico de manera segura y sostenible.

### **2.1.7.1 Pronóstico de la demanda Energética de México**

La demanda energética es un indicador de crecimiento y desarrollo de un país por eso es que se requiere conocer de la manera más segura el comportamiento de la demanda de manera que su pronóstico se aproxime lo más posible a la realidad para evitar afectaciones en el suministro de energía.

La demanda energética por su naturaleza es cíclica como vemos en la Figura 2.4 donde se ve que la curva presenta un patrón que se repite a cada año. Así que el pronóstico de la demanda se radica en encontrar la tendencia de esta serie.

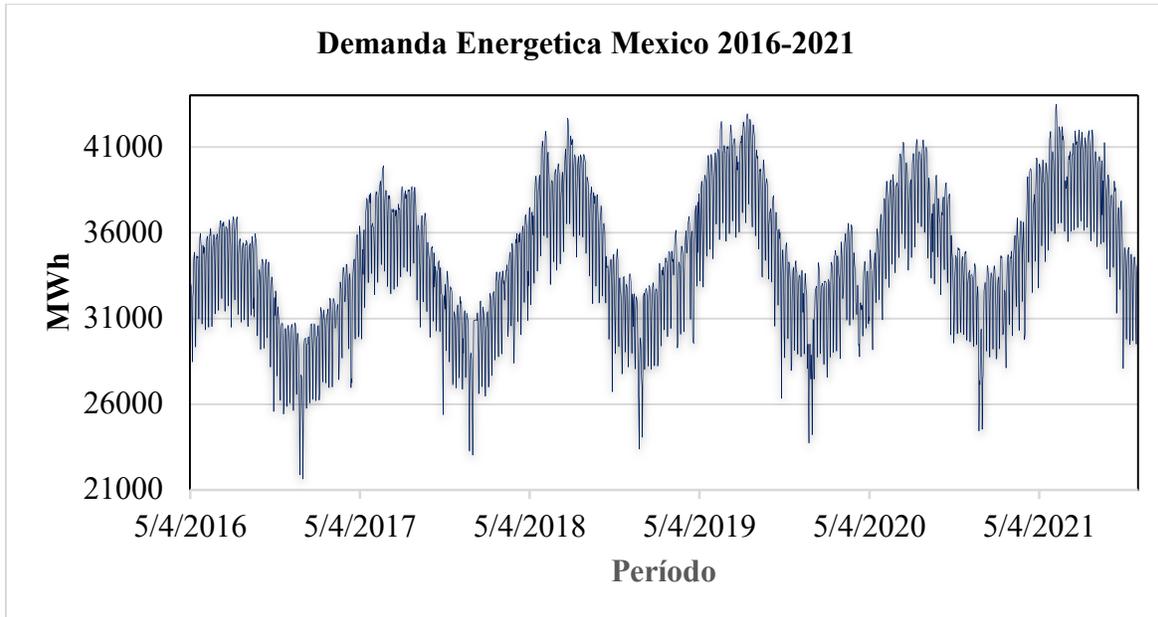


Figura 2.4: Curva de demanda de energía Eléctrica en México

### 2.1.7.2 Zonas de Potencia En México

En México existen Zonas de Potencia que están Interconectadas entre ellas, para el Mercado para el Balance de Potencia 2020, las Zonas de Potencia definidas son (Figura 2.5):

- I. Sistema Interconectado Nacional (SIN).
- II. Sistema Interconectado Baja California (BCA).
- III. Sistema Interconectado Baja California Sur (BCS).



Figura 2.5: Zonas de Potencia del Sistema Eléctrico Nacional

La zona de Potencia (SIN) a su vez está compuesta por 7 áreas interconectadas las cuales son (central oriental Occidental Noroeste, Norte Noreste y Peninsular) (Figura 2.6).



Figura 2.6: Zonas y sub-zonas de Potencia del Sistema Eléctrico Nacional.

## 2.2 Estado del Arte

En el área de pronósticos de series de tiempo, muchos trabajos se han desarrollado dado que ha sido un área donde año a año se han desarrollado muy buenas técnicas, muchas de las cuales se siguen usando a las fechas actuales a pesar de su antigüedad.

A continuación, se presentan algunos de los trabajos más relevantes relacionados con la investigación que se está desarrollando en el área de combinación de pronósticos.

- La combinación de pronósticos utilizando pesos fue propuesta Bates-Granger en [7], donde los pesos se determinan de acuerdo con la varianza de los errores de los métodos individuales. B&G determina los pesos utilizando la correlación entre la varianza de los errores de los métodos de pronóstico individuales en la combinación. Este método de combinación se puede aplicar usando otros métodos de pronóstico. Cuando solo hay dos métodos de pronóstico, los pesos  $\omega_k$ , y  $1 - \omega_k$  deben asignarse al primer y segundo método, respectivamente. El método combinado con esta estrategia no necesariamente produce un mejor pronóstico que el mejor método individual. Sin embargo, estos pesos

deben modificarse para mejorar la previsión combinada. B&G estableció que la mejor combinación de métodos individuales se obtiene cuando [7]:

- La combinación produce el error más bajo,
- La varianza total del error es lo más baja posible.
- Winkler [8] muestra extensos resultados empíricos de la precisión de las pronósticos combinados obtenidas a través de promedios ponderados. Este desarrollo se aplica a diferentes horizontes. Los autores afirman que la precisión de los promedios ponderados es superior a la media simple y promueve la metodología utilizada. Explican que, si se combinan métodos de pronóstico más diversos, el pronóstico final tiene una mayor precisión.
- Aunque la tendencia común mostrada en los primeros el trabajos de pronósticos combinados usando pesos en Granger [56] considera tres enfoques alternativos para obtener combinaciones lineales. En el trabajo se demuestra que la mejor técnica consiste en añadir un término constante y no limitar los pesos a la unidad. Este trabajo utiliza tres técnicas de combinación que se pueden realizar en función de los costes. Este trabajo propone abandonar la práctica habitual de obtener una media ponderada de métodos de pronóstico en favor de una combinación lineal sin restricciones que incluya un marco temporal constante. Este trabajo concluye afirmando que siempre habrá dificultades para combinar. A pesar de esto, si los pronósticos se combinan, no necesariamente tendrán los mismos errores que los métodos individuales.
- Clemen [57], proporciona una revisión y bibliografía de esta literatura, bajo la premisa de que la precisión del pronóstico se mejora sustancialmente mediante la combinación de múltiples pronósticos individuales. Clemen expresa que el pronosticador no puede crear un modelo específico correctamente al combinar pronósticos. Desarrollar más y más métodos de combinación puede ser una forma complicada de resolver el problema. Sin embargo, los investigadores creen que el estudio de la combinación de pronósticos puede ser una pieza crucial para ayudar a los pronosticadores a especificar los procesos subyacentes de manera más apropiada y, por lo tanto, construir mejores modelos individuales.

- West [58] propone un sistema basado en pesos y se utiliza un conjunto de datos con 4275 pronósticos de modelos econométricos y combina siete estrategias de pronóstico. El objetivo es combinar los pronósticos con una precisión media alta y un bajo riesgo de metodología de inexactitud. Los resultados indican que una de las estrategias, más específicamente la estrategia de estimación de pesos sin restricciones superó las otras. Ninguno de las otras seis estrategias evaluadas es más preciso relacionado con la estrategia vencedora. Los resultados obtenidos se contrastan con [59], quien indicó una guía para determinar los métodos a combinar.
- Deustch [10] abunda más en la teoría propuesta inicialmente por [7] en un método con pesos variables a lo largo del tiempo. Los resultados muestran que los esquemas de combinación más sencillos propuestos en este trabajo pueden generar pronósticos más altos que los otros modelos de pronóstico más sofisticados.
- Smitch [60] utilizó un estudio de Montecarlo y una reevaluación para apoyar la explicación propuesta para concluir que se encontraron combinaciones de pronóstico simples para cargar repetidamente combinaciones ponderadas sofisticadas de carga externa en aplicaciones empíricas. Una conclusión de este trabajo es ignorar los errores de covarianza en los pronósticos al calcular las ponderaciones. Si los pesos de la combinación de pronóstico son iguales o cercanos a la igualdad, se espera que un promedio simple sea más preciso que una combinación de pronóstico con ponderación. Cuando el número de pronósticos es grande, de modo que se obtienen pesos diminutos, un promedio simple puede ganar eficiencia intercambiando un pequeño sesgo contra una varianza de estimación más considerable.
- Aiolfi [13] describe los factores que determinan las ventajas de combinar pronósticos. Se analizan diferentes técnicas para combinar métodos de pronóstico. El trabajo concluyó que la investigación estudiada muestra que los peores modelos han producido buenos resultados combinando pronósticos con diferentes pesos. Para conjuntos de datos grandes, se requieren parámetros de estimación.
- Lemke [12] realizó experimentos combinando métodos de pronóstico clásicos usando las competencias NN3 y NN5 para determinar si los métodos de combinación de pronóstico eran mejores que los métodos individuales, probándolos en diferentes escenarios. Se

experimentaron cinco tipos de combinaciones. Se utilizaron árboles de decisión para la exploración de datos. Se comparan tres algoritmos de meta-aprendizaje (NN, árboles de decisión y *SVM*) y realizan algunas combinaciones entre ellos, y también aplican el algoritmo k-means a un grupo y clasifican más tarde. Como las principales contribuciones, se construyó un gran conjunto de características diversas. Con un árbol de decisión, determinan qué características son esenciales para la selección de métodos. Sólo la diversidad no es la clave para la combinación correcta de métodos. El uso de métodos de pronóstico individuales es útil cuando tiene series temporales cortas mientras se utiliza el enfoque de métodos de combinación significa que tiene una diversidad de métodos de precisión aceptables. La combinación de métodos basados en clasificación es superior a un enfoque de selección de modelo simple.

- Adhikari [14] desarrollo cinco métodos de pronósticos individuales y diez combinaciones de métodos en 9 series temporales (puntos de referencia). Como literatura, algunos de los modelos de combinación de peso se incluyen en los pronósticos (más conocidos). Se realizó la experimentación con puntos de referencia de series temporales para determinar si una combinación de métodos de pronóstico con pesos es mejor que los métodos de pronósticos individuales. También se trata de determinar cuáles de las 15 opciones (individuales y combinadas) de los métodos de pronóstico son mejores y por qué). Las medidas de error de los pronósticos son *MSE* y *sMAPE*. Se encontró que los métodos de peso combinado eran mejores que los métodos individuales. Este trabajo concluyó que la combinación de métodos basados en el error *sMAPE* fue la que tuvo el mejor desempeño, mientras que la combinación *LSR* (regresión menos cuadrada). El promedio simple está mejor posicionado que los tres métodos estadísticos (participantes). Así que esta investigación concluye que es mejor utilizar una combinación lineal de pronósticos que los métodos de pronóstico individuales.
- Chan [61], realizó una comparación entre las diferentes técnicas con el objetivo de encontrar la mejor distribución de peso en una combinación de métodos para el pronóstico bancario. La idea fue combinar diferentes variaciones del método de pronóstico individual Holt and Winters. Las técnicas de combinación que se compararon fueron Promedio simple (*SA*), Pesos fijos (*FW*) y Ponderación de ventana variable (*RW*).

El artículo concluye que la combinación con pesos fijos bien determinados obtuvo mejores resultados en contra de lo que la literatura sugiere. La literatura indica que es posible variar los pesos de la combinación en busca de un mejor resultado.

- Chan [62] presenta una metodología para encontrar los pesos óptimos en una combinación de pronósticos y establecer un método para controlar la calidad de esos pesos. Algunas técnicas utilizadas en estos trabajos son la técnica de Pesos controlados (*CW*), pesos fijos (*FW*), suma acumulativa (*CUSUM*) y longitud de ejecución media (*ARL*). Los autores concluyen que muchos métodos de pronóstico con combinación se han desarrollado, pero se desarrolla poco trabajo en los criterios para actualizar los pesos del pronóstico. Algunos utilizan pesos fijos y otros utilizan ponderaciones variables que se actualizan para cada período pronosticado. La técnica *CUSUM* surge como una solución para certificar la calidad de los pesos. La metodología *CUSUM* y el tiempo correcto para actualizar usando el método *CW* mejoraron el rendimiento ya que realiza una actualización de peso inteligente en lugar de reducir la frecuencia de las actualizaciones. La metodología propuesta tiene resultados sencillos y satisfactorios.
- Feng [63] combina tres métodos de pronóstico individuales: BP (Back Propagation), *ARMA* (Autoregression and Moving Average ) y ES (Exponential Smoothing) para pronosticar los índices bursátiles de Shanghai. Este trabajo utiliza métodos para desarrollar la combinación de pronóstico agregado a través de la ponderación exponencial (*AFTER*) y la optimización de enjambre de partículas (*PSO*). Estos métodos utilizados en conjunto funcionan bien para diferentes datos.
- Fan [64] se propone un método de pronóstico combinado con ponderaciones variables basadas en el error de pronóstico inicial (*VWCFFE*). La combinación de pronósticos con un peso fijo no puede obtener un resultado aceptable porque el pronóstico de cada modelo varía según el tiempo. La gran ventaja de variar los pesos es que el modelo se puede generalizar más de una vez mediante la combinación de pronósticos, lo que mejora significativamente la precisión del pronóstico.
- Chen [65] propone un método de combinación de pronósticos utilizando la distribución de errores históricos de los métodos *PM*, *ARMA* y *NWP* para construir la función de Scoring para el pronóstico final. Las técnicas utilizadas son la función de densidad de

probabilidad (*PDF*) y la regla de Scoring adecuada (PSR). El uso del método de Scoring se aplica para medir el rendimiento del modelo mediante una distribución probabilística. La combinación de pronósticos en este trabajo varía con el tiempo en comparación con las combinaciones tradicionales basadas en el peso. Los resultados muestran la superioridad del modelo sobre los métodos tradicionales de pronóstico y los autores concluyen que la combinación de modelos de pronóstico es favorable para mejorar la precisión.

La metodología B&G [7] se aplicó muchas veces para el horizonte de pronóstico de un día [27], [30] y más amplio [6], [8], [10], [14], [29], [58], [62]. Algunos métodos son relativamente fáciles de implementar, pero la mayoría de ellos requieren ajustar sus parámetros para diferentes aplicaciones [26], [27]. Además, una combinación lineal no necesariamente supera el pronóstico del mejor método en la combinación [56]. Como resultado, la comunidad de pronósticos propuso muchas alternativas [43]:

- Determinar los pesos utilizando los métodos de Monte Carlo con una densidad de distribución probablemente definida por la matriz de varianzas de la combinación,
- utilizar solo los valores de varianzas diagonales como factor para obtener pesos mejorados.

Algunos autores han utilizado estas últimas ideas para grandes conjuntos de datos. De acuerdo con los experimentos con un peso mejorado, un grupo de métodos obtuvo mejores resultados que los métodos individuales [12]. Además, los autores aplican la agrupación y clasificación, y utilizan un gran conjunto de características en un árbol de decisión. Determinan qué características son esenciales para seleccionar métodos en la combinación. Observamos que la mejor combinación generalmente se obtiene cuando se utilizan los mejores métodos individuales y cuando la combinación tiene los enfoques más diversos [27], [28]. Por ejemplo, Jaganathan y Prakash proponen dos métodos de enfoques de combinación de pronósticos basados en [42]:

- evidencia histórica de métodos bien evaluados inspirados en la distribución uniforme del peso [66].
- optimización de pesos [42].

Jaganathan y Prakash tomaron 24 métodos de pronóstico de varias fuentes para diseñar una combinación que superó a la mayoría de los métodos M4-Competition. La Tabla 2.1 muestra algunos de los trabajos relacionados más relevantes con respecto a la técnica utilizada, las métricas de rendimiento, el punto de referencia, las ventajas y las desventajas del trabajo.

Tabla 2.1: Estado del Arte de Combinación de Pronósticos

Trabajo	Técnica	Métricas	Benchmark	Ventajas	Desventajas
2018 [26]	PP-MA	MAE, RMSE, MAPE, MASE	PlanetLab	Fácil de implementar	Carece de optimización para mejorar la combinación
2018 [27]	LA	RMSD, Error Ratio, Absolute Error	PlanetLab	Fácil de implementar	Carece de optimización para mejorar la combinación
2020 [29]	EMD-PE-ANN	MSE, RMSE MAE	Datos de velocidad del viento	Alto rendimiento en horizontes más amplios. Se puede utilizar en tiempo real	Se requiere reajustes de parámetros para el entrenamiento
2021 [30]	LSTM-Filtros de Kalman	MAPE, MSE, RMSE, and R2 score	consumo eléctrico Residencial	Pronostico de sistemas que cambian dinámicamente	Pronostica solo un período por delante. Se requiere ajuste para nuevas aplicaciones
2018 [37]	FFORMA	sMAPE, OWA, MASE	M4 Competition	Alto rendimiento	No permite seleccionar métodos individuales específicos
2004 [62]	CUSUM	SRMSE	Datos de demanda mensual para diez artículos	Detecta si los pesos deben ajustarse para nuevas aplicaciones	La precisión depende de su parámetro ARL que debe ajustarse.
1969 [7]	B&G	MSE	Datos de pasajeros de aerolíneas	Buen rendimiento en horizontes pequeños	Rendimiento deficiente para grandes horizontes
2020 [42]	Jagannathan	sMAPE, OWA, MASE	M4 Competition	Mejores pronósticos puntuales, mejores intervalos de confianza.	No permite seleccionar métodos individuales específicos
Propuesto	FCTA	sMAPE, ErrorRatio	M4 Competition	Buen rendimiento para pronosticar en múltiples escenarios y grandes horizontes	Su proceso de ajuste requiere conocimientos especializados

## Metodología

Dado que un pronóstico debe de ser fidedigno es necesario seguir una secuencia de pasos para ser justificable independientemente de los elementos o métodos que se usen para pronosticar [42]. Según [67] y [68] los pasos del proceso de pronósticos son los siguientes y se muestran en la Figura 3.1.

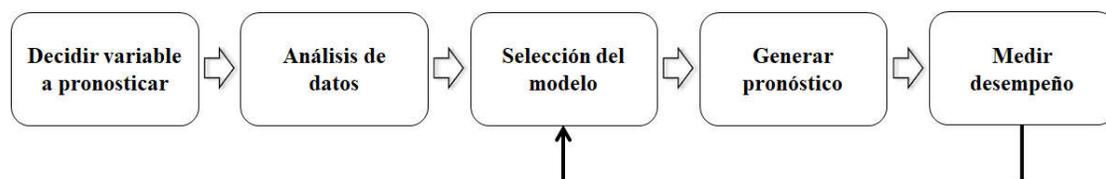


Figura 3.1: Metodología general para el pronóstico

- **Decidir la variable a pronosticar:** En este paso se requiere conocer el tipo de serie que se desea pronosticar, que horizonte de esta serie se requiere pronosticar y a partir de que intervalo histórico de la serie se desea pronosticar y que precisión se desea obtener de este pronóstico. Además, existen prácticas bien recomendadas como pre procesar las series temporales para detectar y corregir valores atípicos [40], que pueden ser necesarias o no dependiendo de la serie de tiempo.
- **Análisis de datos:** Es un paso importante antes de proceder al pronóstico de cualquier variable. En este paso se verifican aspectos relevantes de la serie de tiempo en cuestión como son la tendencia los ciclos la estacionalidad y la aleatoriedad. Además, se pueden verificar aspectos como la media, mediana y la desviación estándar de los datos.

- **Selección del modelo:** En este paso de acuerdo a las particularidades de la serie se selecciona el modelo de pronóstico más adecuado para el pronóstico de valores futuros de esa serie.
- **Generar pronóstico:** En este paso se ejecuta el modelo y se obtienen los resultados.
- **Medir Desempeño:** En este paso se evalúa el modelo de pronóstico y se determina el desempeño ya sea a partir de métricas de precisión del pronóstico como las métricas de error. A partir de estos resultados se determina si el modelo necesitara algún ajuste.

### 3.1 Combinación de Pronóstico FCTA

La metodología que se propone, FCTA (Forecasting Combined method with Threshold Accepting), combina varios métodos de pronóstico, sus elementos significativos son su arquitectura y el algoritmo TAE (Threshold Accepting-Enhanced). TAE es un método de optimización innovador que utiliza una heurística de reheat y golden ratio para mejorar la exploración del espacio de búsqueda. Los experimentos utilizan casi veinte mil veces series y métodos de la M4 muestran que FCTA obtiene mejores resultados de pronóstico que todos los métodos individuales en la combinación. Además, FCTA tiene un rendimiento estadísticamente equivalente o mejor que otras buenas metodologías del área.

El proceso de combinación de FCTA se muestra en la Figura 3.2 donde  $\langle x_t^k \rangle$  es la serie temporal  $k$  donde ocurre el último elemento en el período  $t$ . Por ejemplo, el conjunto de datos  $\{\langle x_t^1 \rangle, \langle x_t^2 \rangle, \dots, \langle x_t^6 \rangle\}$  contiene las siguientes seis series temporales  $\{x_1^1, x_2^1 \dots x_t^1\}, \{x_1^2, x_2^2 \dots x_t^2\}, \dots, \{x_1^6, x_2^6 \dots x_t^6\}$ . En otras palabras, el conjunto  $\langle x_1^k | x_2^k | \dots x_t^k \rangle$  representa los datos del período  $1, 2, \dots, t$  para el  $k$ -ésimo elemento de este conjunto de datos.

FCTA requiere seleccionar los mejores métodos  $k$  de pronóstico de entre los más utilizados en el área para cualquier conjunto de datos tomados. Los métodos individuales seleccionados se nombran en la Figura 3.2. como  $M1, M2 \dots Mk$ , y  $f_1, f_2, \dots, f_k$  representan los pronósticos obtenidos por estos métodos. Luego FCTA obtiene  $k$  estimaciones en períodos  $h$  en el futuro que están representados por  $\hat{Y}_{t+h} (M1), \hat{Y}_{t+h} (M2), \dots, \hat{Y}_{t+h} (Mk)$ .

Explicamos el proceso de estimación de peso en la sección **¡Error! No se encuentra el origen de la referencia.** Este proceso obtiene un peso  $\omega_j$  para cada método. Estos pesos se promedian en un proceso de combinación de la figura para producir el pronóstico inicial para cada serie temporal. Sin embargo, este pronóstico generalmente no es mejor que los métodos individuales. Así, que para solucionar esto se aplica un algoritmo refinamiento, denominado TAE (Threshold Accepted Enhanced) basado en el algoritmo de Threshold Accepting [55], que se presenta en la sección 3.1.1.4. El pronóstico combinado  $\hat{F}_{t+h}$  se puede obtener en función de los  $k$  pronosticos individuales de los mejores métodos, donde se asigna una ponderación inicial  $\overline{\omega}_k$ . Sin embargo, obtener esta función no es, una tarea fácil, aunque comúnmente se utiliza una función simple que pondera los pronósticos individuales. Por el contrario, como se mencionó anteriormente, en FCTA se refina el pronóstico  $\hat{F}_{t+h}$  por medio de TAE.

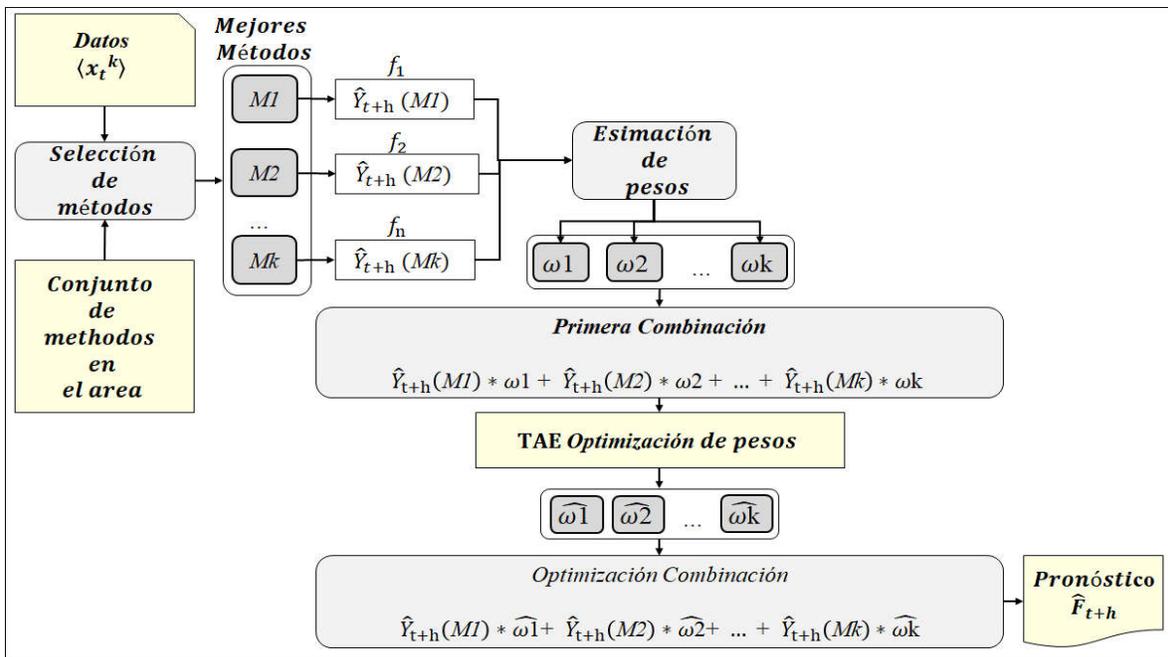


Figura 3.2: Combinación de pronóstico

Matemáticamente el problema de la combinación de pronóstico se puede describir de la siguiente manera:

$$\min_{\omega_i} \hat{\varepsilon} = \sum_{i=1}^M \omega_i \hat{Y}_i \tag{3.1}$$

Subject to  $\sum_{m=1}^M \omega_i \leq 1; \forall \omega_i \geq 0; i=1, \dots, M$

Donde  $\hat{\epsilon}$  es el error estimado para los métodos combinados de pronóstico;  $\omega_i = (\omega_1, \dots, \omega_M)$ , y  $\hat{Y}_i = f(M_i, \hat{Y}_{t+h,i})$  da el pronóstico para el período  $t + h$  del  $M_i$  método individual. Todos los métodos individuales pronostican  $h$  períodos en el horizonte.

### 3.1.1 Arquitectura de FCTA

La arquitectura de la metodología propuesta FCTA (Forecasting Combined method con Threshold Accepting) se describe en la Figura 3.3. Esta arquitectura fue diseñada para probar FCTA para cualquier área de aplicación en particular (en nuestro caso, la competencia M4). Por lo tanto, los métodos de pronóstico se ejecutan previamente para todas las series temporales, y sus errores se almacenan en un conjunto de datos denominado SFM (Set of Forecasting Methods).

En primer lugar, tenemos dos bases de datos. a) Los métodos de pronóstico para el área de aplicación y b) el conjunto de datos con las series temporales que se requieren pronosticar. Como es típico, cada serie temporal se divide en tres partes: Entrenamiento, Validación y Prueba. Para cada serie temporal, el proceso de Selección determina la calidad de los métodos que lo clasifican de mejor a peor. Esta clasificación se almacena en Selección como un atributo de cada serie temporal. Por lo tanto, se determinan los mejores métodos para cada serie temporal. FCTA determina el error promedio y la desviación estándar para cada método en este proceso. La ponderación es un proceso que calcula un peso inicial para cada uno de los métodos  $k$ . Este proceso utiliza dos estrategias manejando un proceso estocástico con distribuciones Uniforme y Normal Estándar. Estos pesos se almacenan en el  $X_o$ , una solución inicial para el algoritmo TAE que determina la mejor ponderación para cada método. Tanto la ponderación como el TAE se explican en dos secciones posteriores.

Finalmente, TAE obtiene la distribución óptima del peso para determinar el Pronóstico Combinado de los mejores métodos seleccionados. Este pronóstico final se calcula utilizando la parte de pruebas de la serie temporal. Luego se calcula la métrica de error para el pronóstico final y finaliza el proceso.

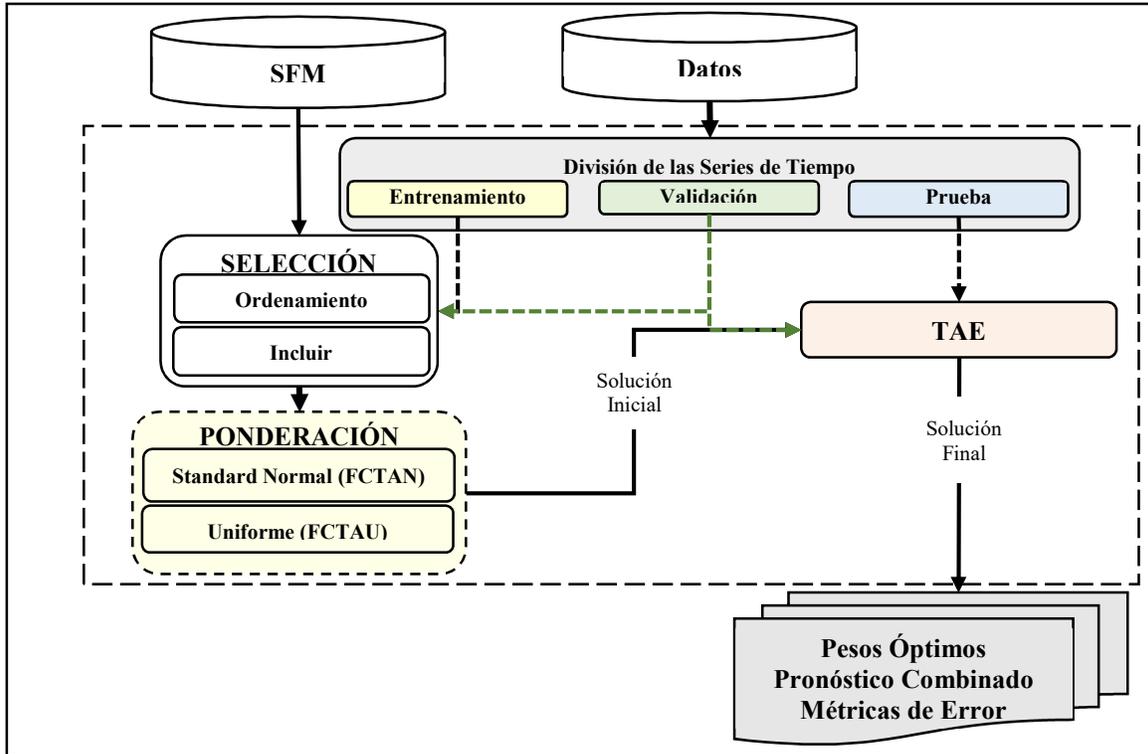


Figura 3.3: Metodología FCTA

La Figura 3.4 muestra el diagrama secuencial para FCTA, que ordena las acciones como se describe a continuación:

- El actor le pide a "Selección" que inicie el proceso con algunas opciones. A continuación, "Selección" solicita a "DataSet" el subconjunto de series temporales (subdivididas en datos de entrenamiento, validación y prueba), que se le devuelven; "Selección" también solicita a "SFM" la métrica de error.
- "Selección" ordena los mejores métodos  $k$  y los envía en un vector con sus pesos iniciales. "Ponderación" utiliza una de las dos distribuciones (Normal o Uniforme) para obtener un vector de ponderación inicial llamado  $X_0$  y lo envía a TAE;  $X_0$  tiene la primera ponderación de los métodos.
- TAE obtiene los datos de validación del "Dataset", con  $X_0$  y el proceso de optimización de TAE mejora la ponderación. A continuación, TAE obtiene los valores del pronóstico y las métricas de error para la parte de prueba. Finalmente, TAE envía los resultados al actor.

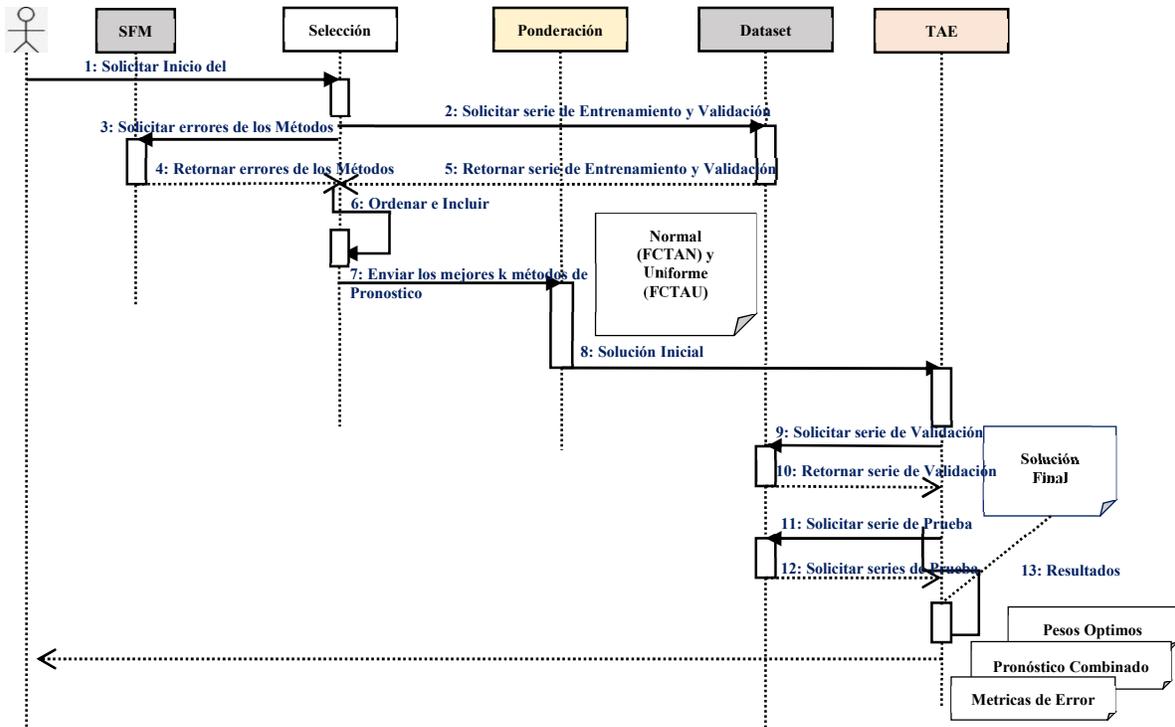


Figura 3.4: Diagrama de Secuencia de la metodología FCTA

### 3.1.1.1 Selección de los mejores métodos

FCTA selecciona los mejores métodos para cada serie temporal con el proceso *Selección* mostrado en la Figura 3.3. Este proceso funciona de la siguiente manera:

- La lectura de rendimiento de cada método en la parte de validación se almacena previamente en SFM.
- Los métodos se clasifican de acuerdo con sus métricas de error para cada serie temporal. El orden se almacena en SFM en un vector de clasificación. Los métodos de este vector se denominan  $BestM_1, BestM_2, \dots, BestM_k, \dots, BestM_N$ .
- El proceso selecciona los mejores métodos  $2 \leq k \leq N$  con el error más bajo.

Para incluir varios métodos en el pronóstico combinado, el proceso *Selección* utiliza el proceso de inclusión que se muestra en el Algoritmo 3.1. En primer lugar, este proceso determina una combinación inicial con los dos mejores métodos:  $BestM_1; BestM_2$ , el proceso *Incluir* decide si se debe agregar otro mejor método posterior a la combinación. El proceso continúa, decidiendo para cada método posterior hasta que se aplica un criterio de

parada basado en la calidad del pronóstico. Este criterio depende del umbral  $\nu$  definido en el proceso de afinación.

---

Algoritmo 3.1: Proceso de Inclusion

---

```

(1) Ranking vector = ( $BestM_1, BestM_2, \dots, BestM_k \dots BestM_N$ )
(2) Select ( $BestM(1), BestM(2)$ ) // Seleccionar los dos mejores métodos
(3)  $BestM(current) = BestM(2)$  ;  $BestM(next) == BestM(k)$  ;
(4)  $include = true$ 
(5) while  $include = true$  // Selección de los siguientes mejores métodos
(6)     if  $BestM(current) - BestM(next) \leq \nu$  , then
(7)          $add BestM(next)$ ;  $k = k + 1$ :
(8)          $BestM(current) = BestM(k)$ ;  $BestM(next) == BestM(k + 1)$  ;
(9)     else  $include = false$ 

```

---

Para encontrar los mejores pesos para el pronóstico combinado, definimos un vector de peso  $X_o = (\omega_1, \omega_2, \dots, \omega_i \dots \omega_j \dots \omega_m)$ . Este vector es una solución inicial para ponderar los métodos de pronóstico (1, 2, ..., m) en el pronóstico combinado. El proceso de ponderación en la Figura 3.5 utiliza las distribuciones *Uniforme* y *Normal Estándar* para definir el vector de pesos  $X_o$ . Para la distribución uniforme, el proceso simplemente asigna el mismo peso a cada método. La Figura 3.5 muestra que el proceso de *Ponderación* recibe para cada serie temporal la siguiente información previamente generada: los pronósticos y las métricas de error, Además, este proceso lee el promedio y la desviación estándar de estas métricas asociadas con cada grupo de series temporales.

Para la distribución *Normal Estándar* se utiliza una población obtenida de los errores de pronóstico para cada serie en cada grupo, donde se toman varias muestras de esos errores. Todas las muestras se caracterizan a partir de su error promedio. Como es bien sabido, cuando el número de muestras es lo suficientemente grande, siguen la distribución *Normal Estándar*. A continuación, la ponderación determina la  $\mu$  media de estos errores medios y su desviación estándar  $\sigma$ . En este punto, utiliza la distribución *Normal Estándar* que utiliza valores  $Z$  para cada uno aplicando la relación popular 3.2

$$Z_i = \frac{\bar{e}_i - \mu}{\sigma} \tag{3.2}$$

Donde  $\bar{e}_i$  representa el promedio del error de pronóstico del método  $M_i$  que se determina utilizando los errores del grupo al que pertenece la serie temporal. En otras palabras,  $\bar{e}_i$  es el

promedio de estos errores. El parámetro  $\mu$  se define como la media de todos los errores promedio de los métodos en todo el conjunto de datos.

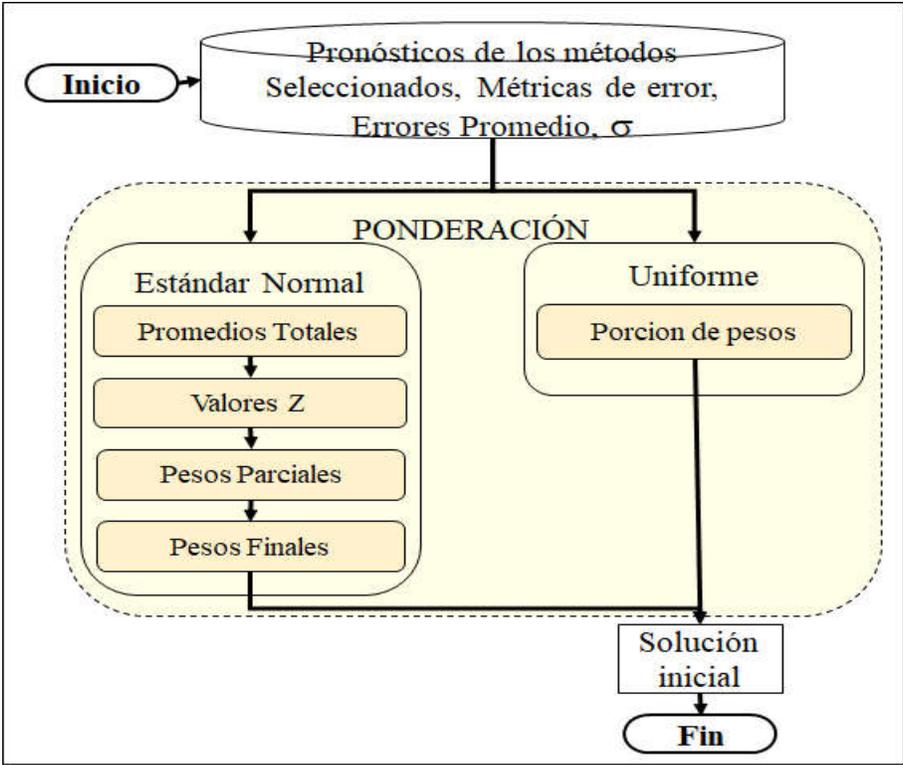


Figura 3.5: Proceso de Ponderación de FCTA

### 3.1.1.2 Proceso de Ponderación

Para estos valores Z, se puede calcular el área bajo la curva de la distribución. El siguiente proceso en la secuencia conocida como pesos parciales consiste en obtener pesos que no son definitivos porque no satisfacen la restricción de la suma igual a uno.

Por definición, el área total bajo la curva estándar representa alrededor del 100% de la probabilidad relacionada con la variable de estudio, como se muestra en la

Figura 3.6. Por lo tanto, en nuestro caso, necesitamos averiguar qué porcentaje de la probabilidad representa el área bajo la curva para cada error promedio con el fin de obtener el peso parcial. Así, aplicamos la fórmula 3.3 para la obtención de los pesos de las piezas para cada método  $Pw_l$

$$Pw_l = 1 - (A_l/SA) \tag{3.3}$$

Donde  $A_I$  representa el área bajo la curva para cada método en la previsión combinada, y  $SA$  representa la suma de todas las áreas bajo la curva para estos métodos. En la fórmula 3.3 el número uno en el lado derecho representa el área total bajo la curva estándar. Por ejemplo, en la Figura 3.6, damos un ejemplo hipotético, donde el área bajo la curva para un  $\sigma$ , corresponde al valor  $z$  del mejor método entre los métodos individuales seleccionados. En esta figura, el 99,7% de los valores de datos están dentro de  $\mu \pm 3\sigma$ . El proceso final consiste en determinar los pesos finales  $Fw_I$ . Primero, teniendo los pesos parciales, los normalizamos igualando su suma a uno. Por lo tanto, cada peso parcial se divide por la suma de todo el peso parcial. En otras palabras, cada peso final de cada método se obtiene como  $Fw_I = Pw_I / \sum Pw$ .

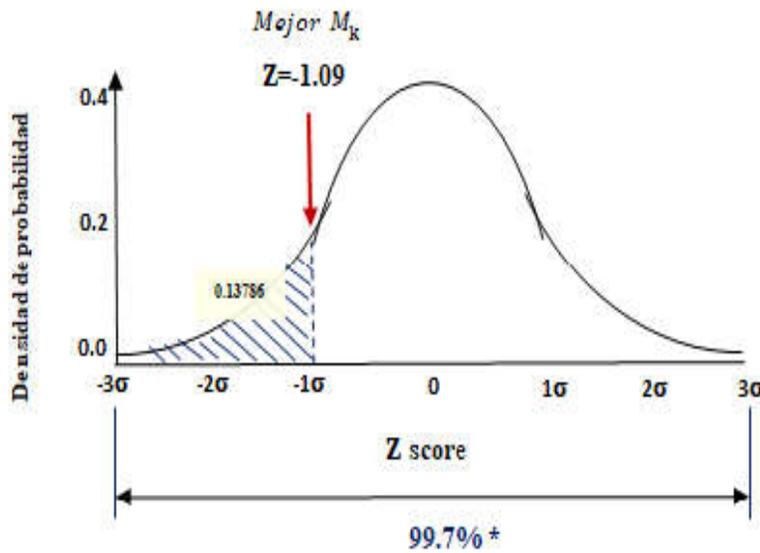


Figura 3.6: Ejemplo de distribución normal estándar

La segunda forma es asumir que los pesos siguen una distribución uniforme. En este caso, se asigna el mismo peso a todos los métodos. Luego, para satisfacer que la suma de peso es igual a uno, determinamos la proporción de peso para cada error de pronóstico utilizando  $Fw_I = 1/n$  donde  $n$  representa el número de métodos. Luego, la ponderación proporciona una solución inicial para el algoritmo TA.

### 3.1.1.3 *Threshold Accepting* para FCTA

El resultado del proceso de ponderación se mejora con TAE en el siguiente paso. TA es una poderosa heurística de optimización que se ha aplicado de manera eficiente a diferentes

problemas de optimización en muchas áreas. En nuestro caso, TAE resuelve un problema de minimización donde la función objetivo es encontrar la distribución óptima del peso para reducir el error de pronóstico combinado. Además, TAE tiene características especiales, como se describe en esta sección.

### 3.1.1.4 Algoritmo TAE

El algoritmo TAE propuesto se muestra en el Algoritmo 3.2, y algunas de las estrategias incorporadas en este algoritmo son ruleta, memoria, esquema de enfriamiento, recalentamiento y el equilibrio térmico que se explicarán en secciones posteriores.

TAE utiliza los parámetros:  $T_{initial}$ ,  $T_{final}$ ,  $\alpha$ ,  $L_k$ ,  $\beta$ ,  $X_o = (\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_j, \dots, \omega_m)$ ,  $E_0$ ,  $F_{t0}$ . La variable  $X_o$ , contiene los pesos del método en la combinación y es proporcionada por el proceso de ponderación (línea 1). Primero, los siguientes parámetros se inicializan  $T_k$ ,  $X_{old}$ ,  $\alpha$ ,  $\Phi$  (líneas 2 -3). El parámetro de temperatura controla el ciclo externo (líneas 4 – 33), mientras que el ciclo interno (líneas 6 – 28) es controlado por un número de iteraciones (que en nuestro caso es variable); el parámetro  $L_k$  determina el número de iteraciones. El primer procedimiento es generar una nueva solución factible utilizando la solución inicial  $X_o$  almacenada en  $X_{old}$ . Un número aleatorio es utilizado por los métodos de ruleta y memoria (líneas 8 y 10). La nueva solución  $X_{new}$  se determina aplicando una función de perturbación a  $X_{old}$ . Luego, se realiza un cálculo de energía (línea 12). TAE obtiene la energía  $E_k$  con el sMAPE,  $E_0$  con  $X_o$  y la posterior con  $X_k$ . El proceso resultante en la secuencia es el cálculo  $\Delta E$  (línea 13), que es la diferencia entre la energía de la solución actual y la solución candidata. Por lo tanto, restamos el error del pronóstico combinado actual del nuevo (solución candidata). En este momento, si  $\Delta E < 0$  (línea 15) (la combinación MAPE disminuye), esta nueva solución es aceptada  $X_{old} = X_{new}$  (línea 16). Además, el proceso de generación de una nueva solución comenzó de nuevo, de lo contrario si  $\Delta E > 0$  y si  $\Delta E \leq Tol_k$  (línea 21) esta nueva solución también se acepta  $X_{old} = X_{new}$  (línea 22).

Luego, el proceso de generar una nueva solución comienza nuevamente. Si  $\Delta E > Tol_k$  el proceso continúa, si  $k < L_k$ ,  $k$  se incrementa (línea 27), generando una nueva solución, comenzando de nuevo e iterando una búsqueda local. Por otro lado, si  $k \geq L_k$ , finaliza el ciclo interno y se actualizan los parámetros: la velocidad de enfriamiento, (parámetro  $\alpha$ ) en

la línea 29 y una primera temperatura final con el parámetro  $\Phi$ ; en la línea 30, la temperatura con  $T_{k+1} = \alpha * T_k$ , la tolerancia umbral  $Tol_k$ , y el  $L_k$  parámetro con el valor  $\beta$  (línea 30).

---

Algoritmo 3.2 Threshold Accepting Enhanced (TAE) propuesto para FCTA

---

```

(1) Parametros TAE: ( $T_{initial}, T_{final}, \alpha, L_k, \beta, X_0, E_0, F_{t0}, NonImprove, elitism, slope$ )
(2) Inicialización Parámetros:  $T_k = T_{initial}, T_{min} = 0.1, X_{old} = X_0, \alpha = 0.7, \Phi = 0.618,$ 
(3)  $MaxY = 10, MaxE = 0.8, Min\zeta = -0.001$ 
(4) while  $T_k \geq T_{final}$  do
(5)   if ( $NonImprove == MaxY$ ) then Reheat ( $T_k$ ) end if
(6)   while  $k < L_k$ 
(7)     if  $elitism < MaxE$  then
(8)        $X_{new} = \mathbf{GenerateNewSolutionRou}(X_{old})$  // Roulette Method
(9)     else
(10)       $X_{new} = \mathbf{GenerateNewSolutionMem}(X_{old})$  // Memory Method
(11)    end if
(12)     $E_{new} = \mathbf{EnergyCalc}(X_{new}, F_t)$ 
(13)     $\Delta E = E(X_{new}) - E(X_{old})$ 
(14)     $memory = \mathbf{UpdateMemory}(\Delta E, X_{old}, X_{new})$ 
(15)    if  $\Delta E < 0$  then
(16)       $X_{old} = X_{new}$ 
(17)      if  $E_{new} < E_{best}$  then
(18)         $Best = (E_{new}, X_{new})$ 
(19)      end if
(20)    Else
(21)      if  $\Delta E \leq Tol_k$  then
(22)         $X_{old} = X_{new}$ 
(23)      if  $E_{new} < E_{best}$  then
(24)         $Best = (E_{new}, X_{new})$ 
(25)      end if
(26)    end if
(27)    end if;  $k = k + 1$ 
(28)  end while
(29)   $Tf = T_k * \Phi$ ; Update CoolingSpeed( $\alpha, Tf$ );
(30)   $T_{k+1} = \alpha * T_k$ ;  $Tol_{k+1} = \gamma * Tol_k$ ;  $L_{k+1} = \beta * L_k$ ;
(31)   $slope = \mathbf{SlopeCalc}(T_k, T_{k+1})$ 
(32)  if ( $T_k < T_{min}$  OR  $slope > Min\zeta$ ) then stop end if
(33) end while
(34) return ( $Best$ )
(35) end Threshold Accepting

```

---

Las variables y parámetros en el algoritmo TAE son los siguientes:

- **Parámetros Iniciales:**  $T_{initial}$ : Temperatura inicial,  $T_{final}$ : Temperatura final,  $\alpha$ : Tasa de enfriamiento,  $L_k$ : Número de iteraciones del ciclo interno,  $\beta$ : Tasa de aumento de búsqueda,  $X_0$ : Solución inicial (pesos),  $E_0$ : Energía inicial (combinación sMAPE)

- **Parámetros de Control:**  $T_k$ : Temperatura actual,  $T_{min}$ : Temperatura mínima para detener el algoritmo,  $Tol_k$ : Tolerancia de temperatura actual,  $\gamma$ : Tasa de tolerancia,  $k$ : Iteración actual,  $\Phi$ : Número Golden ratio,  $\Delta E$ : Diferencia entre la energía actual y la anterior,  $NonImprove$ : Número de iteraciones sin mejora,  $MaxY$ : Número Máximo de iteraciones sin mejora,  $elitism$ : valor de elitismo,  $MaxE$ : Valor de elitismo máximo,  $Min\zeta$ : Valor mínimo de la pendiente de la curva
- **Variables de la Solución:**  $X_{old}$ : Solución actual,  $X_{new}$ : Nueva solución,  $E_{old}$ : Energía actual,  $E_{new}$ : Nueva energía,  $Tf$ : Ventana de temperatura,  $slope$ : Pendiente de curva,  $F_{t0}$ : Valores de pronóstico de los métodos seleccionados

La estrategia de recalentamiento se aplica cuando se detecta estancamiento. En el ciclo externo, la pendiente se calcula para activar el criterio de parada se activa en la línea 32. Por último, el proceso finaliza cuando se cumple el criterio de parada. Los principales procesos de TAE se describen en la siguiente subsección.

### 3.1.1.5 Generación de una nueva Solución

La nueva solución se genera siguiendo uno de los dos métodos: ruleta o memoria. Que se explican en los siguientes párrafos.

Método de la ruleta. Teniendo una configuración de vector o peso  $X_k$ , el vector posterior  $X_{k+1}$  se determina mediante el método de Monte Carlo de la siguiente manera:

En primer lugar, el proceso comienza con la solución inicial  $X_0$  que es un vector de los pesos en la combinación  $X_0 = (\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_j, \dots, \omega_m)$ . TAE determina los intervalos para los pesos en la solución inicial  $X_0$  y las soluciones posteriores. Por ejemplo, para cinco métodos  $X_0 = (0.30, 0.15, 0.20, 0.25, 0.10)$ , los cinco intervalos se definen con los rangos:

$$0 < w_i \leq 0.30; 0.30 < w_i \leq 0.45; 0.45 < w_i \leq 0.65; 0.65 < w_i \leq 0.90; 0.90 < w_i < 1.0$$

En segundo lugar, el siguiente proceso consiste en generar una nueva solución  $X_{k+1}$ ; debido a la condición de que la suma de los pesos debe ser uno el nuevo peso debe satisfacer esta condición. Por lo tanto, si el algoritmo elimina una porción (cuántica) de peso de un método, debería compensar esta porción en otro peso. Se generan dos números aleatorios en el rango

[0,1]; el primero identifica el método para agregar un valor, mientras que el segundo corresponde al método para restar este valor (Figura 3.7).

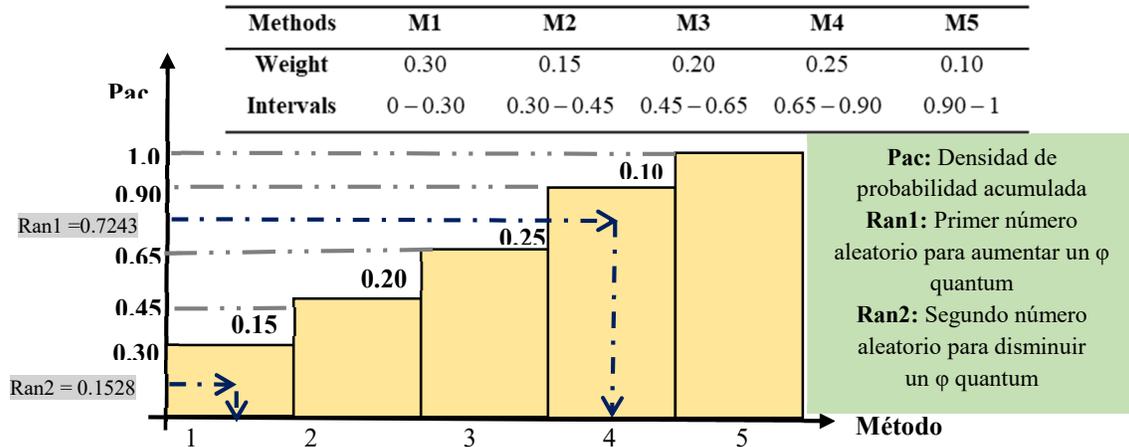


Figura 3.7: Ejemplo de disminución de quantum de los pesos

Método de memoria. En este caso, el algoritmo tiene una estructura llamada "memoria" que almacena los últimos cinco movimientos de iteraciones del algoritmo.

Esta estructura permite verificar en los últimos movimientos aquellos métodos que han mejorado la métrica de error para la solución óptima. Así como en el método de la ruleta, el proceso comienza con una solución  $X_o = (\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_j, \dots, \omega_m)$ .

En primer lugar, se calcula un promedio de los últimos cinco movimientos y se almacena en la memoria. En segundo lugar, con el promedio calculado previamente, se evalúa qué método es beneficioso para agregar un quantum de peso y qué método es conveniente para disminuir un valor cuántico de peso.

### 3.1.1.6 Cálculo de Energía

El algoritmo TAE es análogo al clásico recocido simulado de proporción áurea (GRSA), que minimiza una variable de energía a cada temperatura [69]. Por lo general, el algoritmo determina esta energía con la solución de estado actual, y al comienzo del proceso, comúnmente tiene un alto valor. En cada paso, con una nueva solución, se determina su nueva energía. El valor energético de una solución determina su probabilidad de ser aceptada como una solución actual. Cuando la temperatura es alta, el algoritmo acepta casi todas las soluciones nuevas, incluso si tienen un error alto. Cuando la temperatura es baja, el algoritmo

acepta solo unas pocas soluciones en las que la energía no se ha reducido. Al igual que en los algoritmos TA, la energía define el espacio de búsqueda de la siguiente iteración. A medida que se reduce la variable de energía, también se reduce el espacio de búsqueda (Figura 3.8). Un nuevo valor de energía se evalúa en TAE con la métrica de error sMAPE. En este caso, el valor de pronóstico combinado viene dado por la ecuación 3.4.

$$F_{t0} = X_{new} * (f_{t1}, f_{t2}, \dots, f_{tn}) \tag{3.4}$$

donde  $X_{new} = (\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_j, \dots, \omega_m)$  tiene los nuevos pesos para ponderar el pronóstico individual  $f_{t1}, f_{t2}, \dots, f_{tn}$  para los n métodos elegidos.

En TAE, la diferencia entre la energía actual y la nueva permite la evaluación del algoritmo si la iteración actual produjo un movimiento bueno o malo en el proceso de optimización.

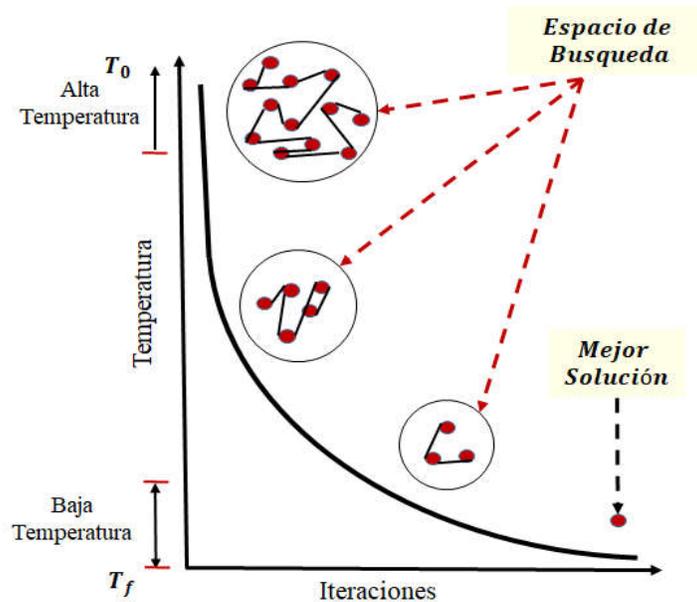


Figura 3.8: Disminución del espacio de búsqueda

### 3.1.1.7 Esquema de enfriamiento

TAE se puede mejorar utilizando la técnica de proporción áurea (GR o  $\Phi$ ). En un proceso de búsqueda, GR es un parámetro utilizado para dividir el espacio de búsqueda en varias secciones de GR; en recocido simulado y heurística de umbral, GR es un parámetro para dividir el ciclo de temperatura; cada sección tiene diferentes temperaturas iniciales y finales [35]. La estrategia de búsqueda utilizada por GRSA realiza búsquedas más rápidas en cada sección de GR. Esta constante  $\Phi$  (o parámetro de aura) divide el intervalo de temperatura en

segmentos, y de esta manera, el espacio de búsqueda se explora más rápido. Utilizamos el esquema de enfriamiento geométrico,  $T_{k+1} = \alpha * T_k$ , donde  $T_k$  y  $T_{k+1}$  son las temperaturas actuales y nuevas, respectivamente. El parámetro  $\alpha$  se conoce como el factor de enfriamiento, y está en el rango de  $0.7 \leq \alpha < 1$ . En el primer GR,  $\alpha$  es pequeño ( $\alpha = 0.7$ ), luego la temperatura disminuye rápidamente. Por el contrario, en los posteriores GRs,  $\alpha$  está creciendo, y las exploraciones son cada vez más lentas; la última sección GR tiene un valor de  $\alpha$  cercano a uno, y el criterio de aceptación se vuelve muy estricto. Esta variación del factor de enfriamiento permite disminuir la temperatura más lentamente cuando la temperatura está más cerca de la temperatura final. La temperatura  $T_\Phi$  se modifica, utilizando el número áureo  $\Phi$ , el espacio de búsqueda para encontrar la solución óptima se reduce cuando se utiliza la estrategia GR para definir el esquema de enfriamiento. La Figura 3.9 muestra las secciones donde las tasas de disminución de temperatura cambian, disminuyendo así el número de iteraciones del algoritmo para llegar a la solución óptima.

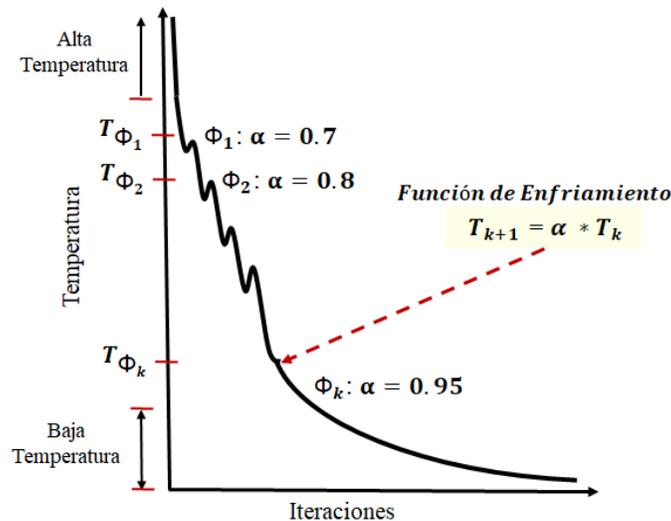


Figura 3.9: Golden Ratio (GR) scheme

### 3.1.1.8 Estrategia de Reheat

Otra estrategia utilizada en TAE es Reheat (RH) [70]; en RH, la temperatura actual  $T_k$  se reinicia a un valor anterior si el algoritmo no mejora su calidad durante un número determinado de ciclos de metrópolis. Los algoritmos reinician la temperatura para explorar más soluciones desde el espacio de búsqueda. En TAE, el RH se aplica en la última sección

de GR, y cuando detecta condición de estancamiento; esto se realiza cuando para algunas iteraciones (diez es común), no se logra una mejora significativa ( $MaxY == 10$  en las líneas 3 y 5 del Algoritmo 3.2). El proceso de RH se inicia cuando detecta condición de estancamiento (línea 5 del Algoritmo 3.2). Luego, el proceso de RH se repite un número máximo de veces (comúnmente tres veces). Después de este número, TAE normalmente se ejecuta hasta que se alcanza la condición de parada. La estrategia RH permite que el algoritmo escape del óptimo local, explore más soluciones y evite la convergencia prematura. La Figura 3.10 muestra algunos elementos de la estrategia de RH.

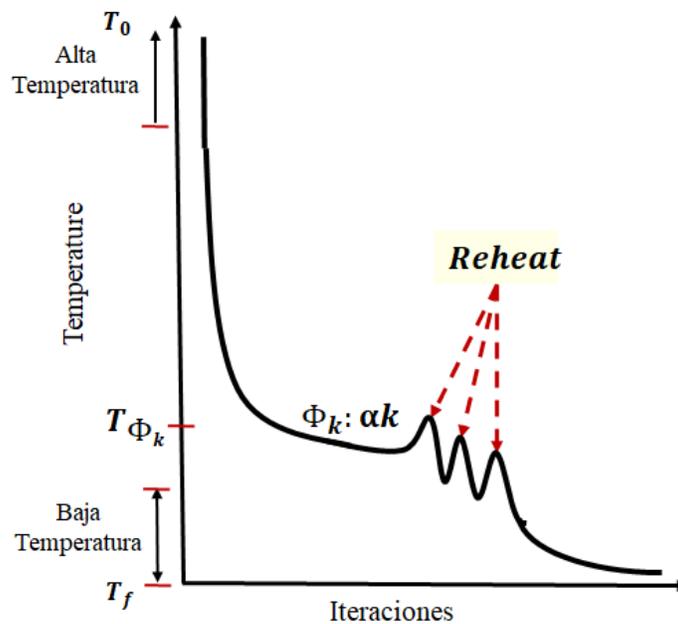


Figura 3.10: Estrategia de Reheat

### 3.1.1.9 Equilibrio Térmico

El equilibrio térmico de un algoritmo TA es la métrica que expresa cuando la temperatura actual ( $T_k$ ) ha alcanzado un valor muy cercano a cero, lo que indica que la búsqueda ha finalizado. El equilibrio térmico en el algoritmo TAE se logra cuando se cumplen dos condiciones:  $T_k < T_{min}$  o cuando  $slope < Min\zeta$ . Por lo tanto, el algoritmo calcula la pendiente de la curva en cada iteración. El valor mínimo de la curva de pendiente es  $Min\zeta$  lo que indica que se ha alcanzado el equilibrio térmico. En este punto, se determina el final de la búsqueda.

### 3.1.2 Complejidad de TAE

Como se muestra en la Figura 3.3, el tiempo de ejecución del FCTA depende esencialmente de los procesos “*Selección*”, “*Ponderación*” y TAE, cuyas complejidades para el peor de los casos son las siguientes:

- Para la selección:  $O(k \log k)$  se debe a ordenar los  $k$  mejores métodos, y  $O(k)$  para seleccionar los mejores métodos que participarán en la combinación.
- Para ponderación:  $O(k)$  para ponderar los  $k$  métodos, y  $O(1)$  para almacenar los pesos en el vector prioritario y enviarlo a TAE. Por lo tanto, la complejidad de la ponderación es  $O(k)$ .
- Para TAE:  $O((n^2 + n) \log n)$ , donde  $n$ , es el número de temperaturas, que es la clase de complejidad para el algoritmo de aceptación de umbral, incluso para el caso multiobjetivo [37], [38].

FCTA se basa en métodos de ponderación como B&G, que realiza tareas similares, excepto que este último no tiene una etapa de mejora como TAE. En otras palabras, TAE provoca una sobrecarga cuya complejidad es después de la reducción en términos de Big-O igual a  $O(n^2 \log n)$  que es la complejidad final del algoritmo.

---

## Análisis y Resultados

### 4.1 Equipo y Material

Los equipos, materiales y datos utilizados de datos utilizados durante el desarrollo de esta investigación se describen a continuación.

#### 4.1.1 Equipo de Cómputo

La experimentación de la metodología propuestas se desarrolló en los siguientes equipos con las siguientes características:

Equipo: Dell G3 (2020)

- Procesador: Intel Core Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz - 2.59 GHz  
Memoria: 16 GB

Equipo: Dell 5559 (2016)

- Procesador: Intel Core Intel(R) Core(TM) i7-6500u CPU @ 2.59 GHz  
Memoria: 12 GB

Equipo: Genérico (2015)

- Procesador: Intel Core Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz  
Memoria: 12 GB

## 4.1.2 Software Utilizado

Para el desarrollo de esta investigación se utilizó el software comercial de entre los cuales se encuentran, MINITAB, R, C Excel, SPSS y herramientas gráficas entre otros.

## 4.1.3 Datasets

Respecto a los datos usados para evaluar la investigación propuesta se determinó desarrollar la experimentación usando las series de las competencias de Makridakis, ya que este conjunto posee una gran cantidad de serie de diversas áreas, además de que se tienen en este conjunto elementos como la serie histórica y una porción de prueba, que son importantes dentro de la ejecución de nuestro experimento. Además de estas series se obtuvieron series del mercado de energía mexicano importantes para probar la metodología propuesta también en esta área.

### 4.1.3.1 M4-Makridakis

El conjunto de datos utilizado corresponde a algunas de las series de competición M4. El Concurso Makridakis M4 contiene cien mil series temporales compiladas continuamente por la Universidad Técnica Nacional de Atenas (NTUA). Se construyó a partir de diferentes fuentes de dominio público, como industrias, servicios, turismo, importaciones y exportaciones, demografía, educación, trabajo y salario, datos gubernamentales, hogares, bonos, acciones, seguros, préstamos, bienes raíces, transporte, productos naturales, recursos y medio ambiente [43]. Las series temporales se dividen en seis grupos con diferentes frecuencias: Anual (Y), Trimestral (Q), Mensual (M), Semanal (W), Diaria (D) y Horaria (H). Debido a la extensa longitud de este conjunto de datos, las series temporales también se pueden dividir en seis tipos: Demográficas, Financieras, Industria, Macro, Micro y Otros. El criterio para seleccionar un subconjunto representativo de estos datos para las pruebas fue de que las series tuvieran entre doscientas y mil observaciones.

Como se muestra en la Tabla 4.1, en cada frecuencia, las series son de diferentes tipos, debido a eso las series del grupo horario no se seleccionaron porque solo contiene series del tipo “Otro”. En la Tabla 4.1, hay dos columnas denominadas Tamaño y Porcentaje (%). La primera hace referencia al número de series seleccionadas en cada grupo de frecuencias. La

última columna es el porcentaje correspondiente de series temporales seleccionadas en relación con el total.

Tabla 4.1: Series Seleccionadas de M4 competition

Frecuencia	Demográficas	Financieras	Industria	Macro	Micro	Otras	Total	Series Seleccionadas	
								Tamaño	%
<b>Y</b>	1,088	6,519	3,716	3,903	6,538	1,236	23,000	56	0.24
<b>Q</b>	1,858	5,305	4,637	5,315	6,020	865	24,000	256	1.07
<b>M</b>	5,728	10,987	10,017	10,016	10,975	277	48,000	18,360	38.25
<b>W</b>	24	164	6	41	112	12	359	125	34.81
<b>D</b>	10	1,559	422	127	1476	633	4,227	941	22.26
<b>H</b>	0	0	0	0	0	414	414	0	0
<b>Total</b>	<b>8,708</b>	<b>24,534</b>	<b>18,798</b>	<b>19,402</b>	<b>25,121</b>	<b>3,437</b>	<b>100,000</b>	<b>19,738</b>	<b>20.152</b>

#### 4.1.3.2 Series de Demanda de Energética de México

Se obtuvieron series de demanda energética de México el conjunto de prueba contiene 19 de entre las cuales se tienen series de energía diarias con los datos del consumo real energético, la cantidad de energía generada y el estimado propuesto por el CENACE. El periodo de las mismas es de aproximadamente 5.5 años desde el 5 de abril del 2016 al 30 de noviembre 2021 para las tres zonas principales de carga (BCA, BCS, SIN) y para cada una de las 7 subregiones del SIN Sistema Interconectado Nacional (CEN, NES, NOR, NTE, OCC, ORI, PEN). [71].

## 4.2 Experimentación

Para evaluar la metodología FCTA se llevaron a cabo una serie de experimentos en diferentes escenarios.

El primer experimento se enfocó en probar que la metodología FCTA era superior que los métodos clásicos de pronóstico y que algunas de las mejores hibridaciones de la competencia de pronóstico M4. En este experimento se contrastaron los pronósticos de algunos métodos clásicos y dos hibridaciones propuestas en M4 contra FCTA. Se usó un conjunto significativo de series de tiempo para desarrollar el experimento. Los resultados obtenidos se evaluaron usando promedios, gráficas y pruebas estadísticas.

El segundo experimento buscó comparar la metodología propuesta con otras hibridaciones de pronóstico exitosas desarrolladas para otras áreas de aplicación y que no participaron en la competencia de pronóstico M4. Los resultados obtenidos con estas técnicas también se evaluaron usando promedios, gráficas y pruebas estadísticas.

Finalmente, en el tercer experimento se evaluó el desempeño de la metodología FCTA para pronosticar valores futuros de series de tiempo de demanda energética de México. Los resultados se contrastaron contra los resultados de los métodos individuales además se evaluó el pronóstico de FCTA contra el pronóstico realizado por el CENACE.

### 4.2.1 Pronóstico con FCTA para la M4 de Makridakis

El experimento consistió en aplicar la metodología FCTA propuesta y pronosticar los valores futuros de las series temporales seleccionadas de la competencia M4. La metodología se aplica tal y como se muestra en la Figura 4.1

Se definieron dos escenarios:

- Escenario A. Métodos de pronóstico individuales comparados contra FCTA.
- Escenario B. Rendimiento de FCTA para diferentes frecuencias del conjunto de datos (Y, Q, M, W, D)

A partir de las series de la M4 y realizando un filtrado como se muestra en la Tabla 4.1 se obtuvieron dos conjuntos de series uno con 19738 series y otro con 150 series, estas últimas que corresponden a las primeras 150 series de la M4 que cumplen con el filtro.

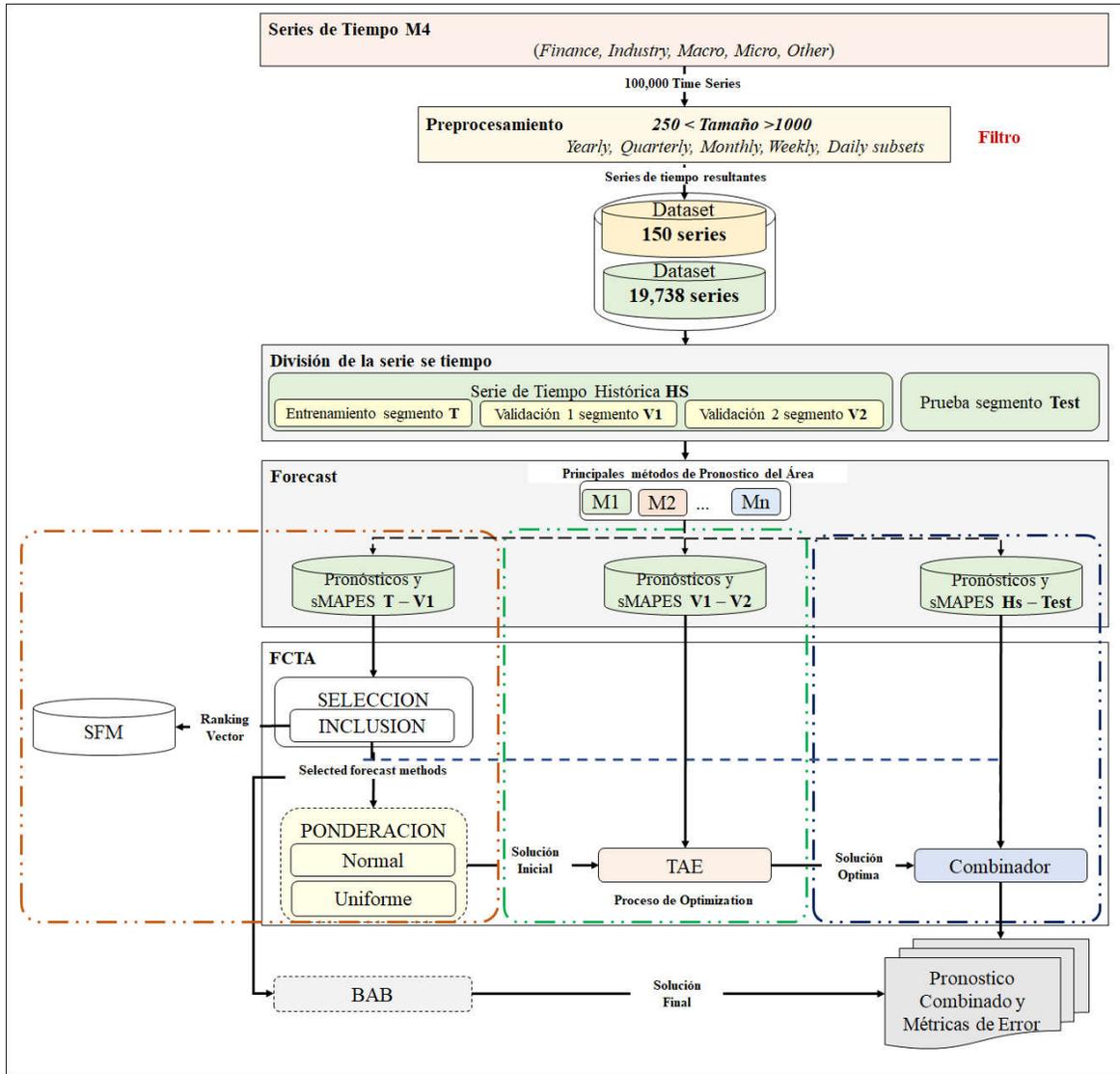


Figura 4.1: Arquitectura de FCTA Ampliada

El experimento realizado se ejecuta de la siguiente manera:

- Como se muestra en la Figura 4.1 a partir del set de datos (Dataset) seleccionados se procede a la división de cada una de las series de tiempo del Dataset. Cada serie temporal de los subconjuntos seleccionados de la M4 se divide en tres partes: Entrenamiento, *Validación1* y *Validación2*, como se muestra en la Figura 4.2. La parte de *Prueba* descrita

en la figura es desconocida para el algoritmo, ya que se usa para corroborar el pronóstico final.

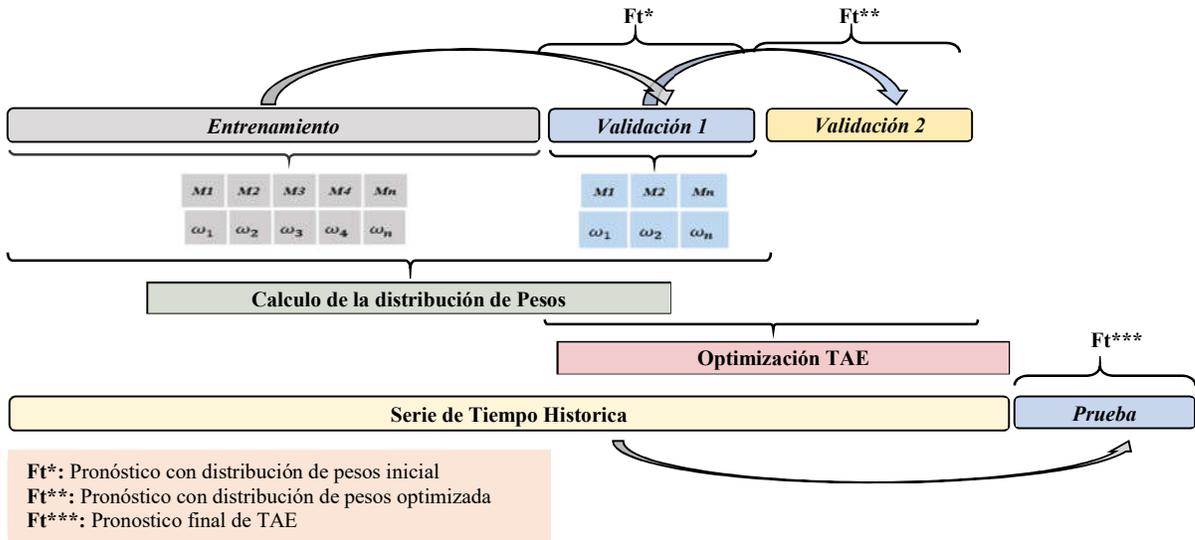


Figura 4.2: División de la Serie de Tiempo

- El experimento combina los mejores métodos de pronóstico del área para obtener un pronóstico de los valores futuros de las series utilizando una de las dos técnicas de distribución de peso (*Uniforme* o *Normal Estándar*) (Figura 4.1), que será optimizada con el algoritmo TAE.
- En primer lugar, la parte de *Entrenamiento* de cada serie se utiliza para pronosticar la parte de *Validación 1*, con todos los métodos de pronóstico individuales disponibles. Con los resultados obtenidos, se determinan los mejores métodos para pronosticar la porción de *Validación 1*. En este punto también definen la distribución de pesos para la combinación de estos mejores métodos.
- A continuación, con la porción de *Validación 1*, se pronostica *Validation 2*, y se obtienen los mejores métodos, con la distribución de peso como se calculó previamente. El resultado se almacena y el proceso de optimización con TAE comienza a utilizar la parte de *Validation 1* para pronosticar *Validation 2*.
- Los pesos se ajustan hasta que no haya más mejoras posibles.
- Finalmente, la mejor combinación de ponderaciones y los mejores métodos de pronóstico se aplican utilizando la serie de tiempo histórica para pronosticar la porción de prueba.

Los catorce métodos utilizados en el experimento se describen en la Tabla 4.2.

Tabla 4.2: Conjunto de métodos de pronóstico para probar FCTA

No.	Método	Acrónimo	Descripción
1	Auto-Arima	AA	AA determina el mejor método ARIMA
2	ETS	ET	ETS (Suavizamiento exponencial de estado de espacio) [72].
3	NNetar	NN	Método de pronóstico con redes neuronales.[73].
4	TBATS	TB	Suavizamiento exponencial de estado de espacio con Box-Cox-ARMA [74].
5	STLM-AR	ST	STLM (Aprendizaje Profundo) con método AR [73].
6	RW-Drift	RW	Pronósticos Naive y Random Walk con drift activado[73].
7	Theta method	TH	Suavizamiento exponencial [75], [76]
8	Naïve Method	NA	Método Naive Clásico [73]
9	ARFIMA	AR	Media móvil integrada fraccionalmente autorregresiva [77]
10	Bagged	BA	Método bootstrapped [78]
11	Spline	SP	Método SplinesCubic [79]
12	Holt	HO	Método Holt clásico Suavizado exponencial [80]
13	FFORMA	FF	Modelo de pronóstico usando promedio basado en características [37]
14	Jaganathan	JA	Método de Jaganathan [42]

Los primeros doce métodos están disponibles en [39] mientras que los dos últimos métodos, FFORMA [13] y Jaganathan [33] están en M4-Competition. Otros métodos utilizados en esta comparación se muestran en la Tabla 4.3.

Tabla 4.3: Métodos de referencia usados para la comparativa con FCTA

No.	Método	Acrónimo	Descripción
1	Smyl	SM	Método híbrido de Suavizamiento exponencial y redes neuronales recurrentes [25]
2	Pawlikowski	PA	Ensamble ponderado de modelos estadísticos (Naïve, ETS, Theta, ARIMA y Regresión lineal) [81]
3	Fiorucci	FI	Pronóstico Combinado usando (Naïve, SES, Damped, ETS, y ARIMA) [82]
4	Petropoulos	PE	Combinación de modelos univariados (SCUM)usando(ETS, CES, ARIMA, DOTM) [83]
5	Bates	BAC	B&G [7] clásico: Usando (AA) y (HO)

El proceso de optimización utiliza el algoritmo TAE propuesto, que observamos en el Algoritmo 3.2. Como se muestra en la Tabla 4.4, los resultados de esta experimentación presentan los 14 métodos individuales, incluyendo [34], [35], utilizados para el conjunto. Los resultados de los métodos anteriores se contrastan con los resultados de los métodos propuestos, que son: FCTA (WD optimizado con TAE) y Bates Best (BAB). Diseñamos este último basándonos en los métodos clásicos de B&G. BAB básicamente toma los primeros 2 mejores métodos del método de Selección de la metodología FCTA (Figura 5.1) y realiza su combinación. Este no utiliza la optimización de pesos propuesta por TAE.

Los resultados indican que, en promedio, el FCTA tiene un error promedio más bajo (a pesar de la distribución de peso elegida) que el mejor método individual para cada serie en todos los subconjuntos utilizados. Sin embargo, en el subconjunto mensual, la diferencia entre el error promedio del mejor individuo y FCTA es muy pequeña. Debido a esto, para validar los resultados obtenidos, se les aplicaron dos pruebas estadísticas: las pruebas de Friedman y Wilcoxon.

Tabla 4.4: Resultados de errores de pronóstico del promedio sMAPE para todos los métodos individuales

Subset	Horizonte	sMAPE (Promedio de errores de pronóstico Métodos ensamblados)														Métodos de referencia					Propuestas		Distribución
		AA	ET	NN	TB	ST	RW	TH	NA	AR	BA	SP	HO	FF	JA	SM	PA	FI	PE	BAC	BAB	FCTA	
Y	6	6.67	6.84	12.6	6.65	18.0	7.17	6.85	7.12	8.46	5.61	<u>5.52</u>	6.91	6.45	6.70	<u>6.11</u>	6.68	6.79	6.79	8.39	3.49	<b>3.86</b>	U*
																						<b>3.87</b>	N
Q	8	<u>4.32</u>	4.72	6.45	4.38	51.4	4.71	4.80	5.27	12.0	4.75	5.97	4.61	4.48	4.45	4.29	4.40	4.33	<u>4.28</u>	4.54	3.59	<b>4.05</b>	U
																						<b>4.07</b>	N
M	18	7.24	8.18	8.43	7.19	19.7	8.31	7.69	8.31	10.2	7.93	17.9	8.63	<u>7.17</u>	7.23	6.37	<u>6.32</u>	6.38	6.38	15.2	13.5	<b>7.15</b>	U
																						<b>7.16</b>	N
W	13	5.97	5.89	7.88	5.56	29.1	6.35	5.91	6.20	10.0	5.93	8.94	6.26	<u>5.41</u>	5.61	5.98	5.52	5.65	<u>5.48</u>	7.53	4.43	<b>4.83</b>	U
																						<b>4.84</b>	N
D	14	3.66	3.47	3.85	3.56	11.6	4.03	3.43	<u>3.39</u>	4.62	3.64	5.19	3.56	3.51	3.52	3.46	<u>3.13</u>	3.48	3.48	3.67	3.57	<b>3.13</b>	U
																						<b>3.14</b>	N

\* Distribución Normal (N), Distribución Uniforme (U)

Se han desarrollado tres gráficos para ver los resultados desde otra perspectiva. La primera gráfica (Figura 4.3) muestra una comparación entre FCTA utilizando una distribución de peso de forma uniforme (FCTAU), FCTA utilizando una distribución de peso normal estándar (FCTAN) y los métodos de referencia.

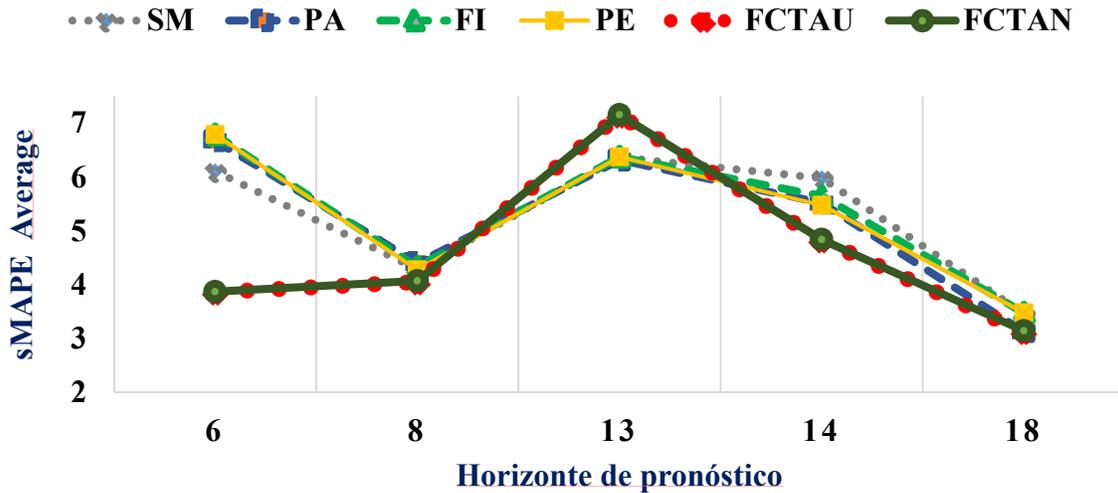


Figura 4.3: Comparativa de FCTAU y FCTAN versus métodos de referencia

Luego, en la Figura 4.4, se han comparado los tres métodos propuestos (BAB, FCTAU, FCTAN). Finalmente, en la Figura 4.5 hemos seleccionado los tres mejores métodos de los catorce para compararlos con FCTAN, FCTAU y BAB.

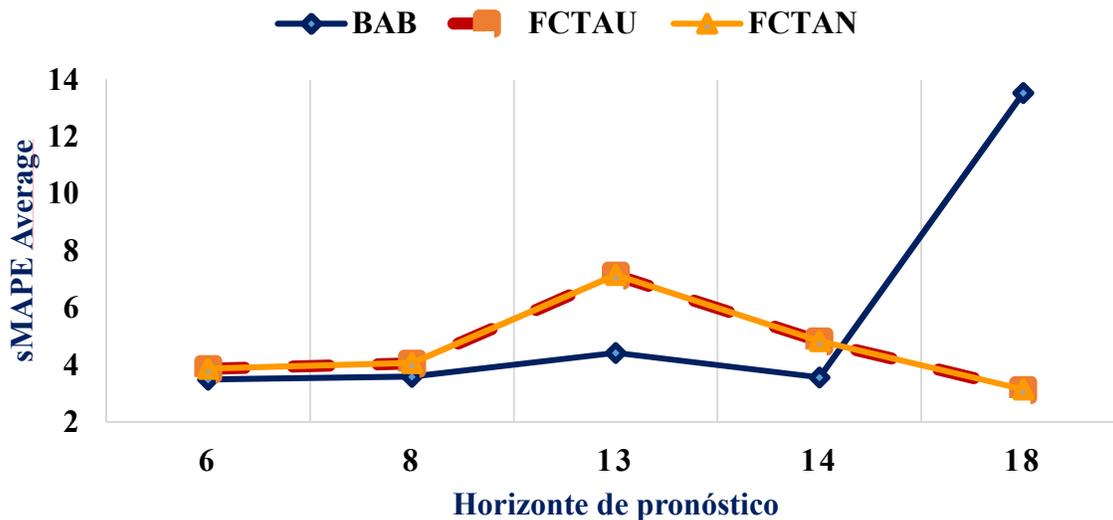


Figura 4.4: Comparativa entre los métodos propuestos

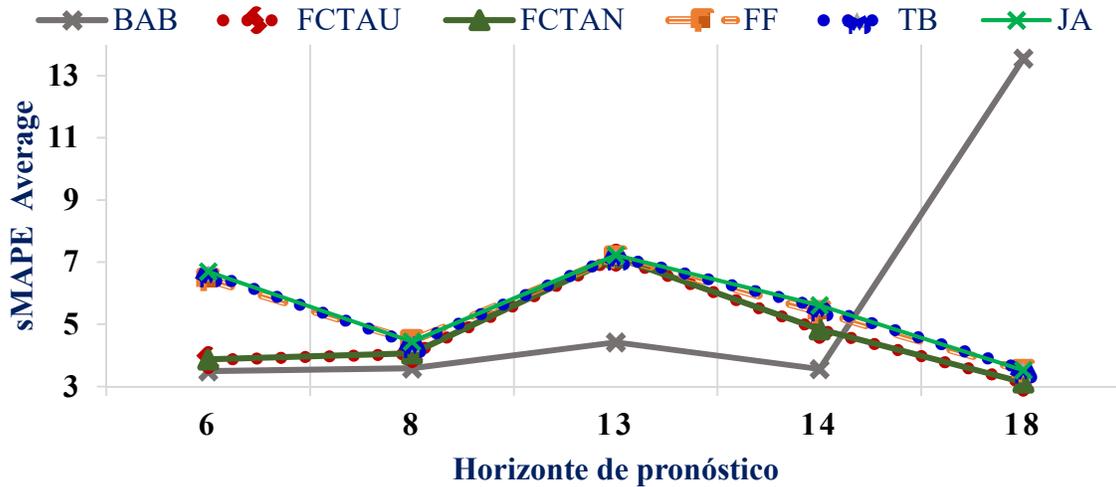


Figura 4.5: Comparación de los métodos propuestos frente a los tres mejores métodos individuales

Observamos en las Figura 4.4 4.4 y Figura 4.5 que el método BAB aumenta su error cuando el horizonte es superior a catorce. Por lo tanto, se debe usar el método BAB para horizontes de pronóstico pequeños. Por el contrario, FCTAU y FCTAN obtienen consistentemente un buen desempeño en diferentes horizontes.

Los análisis de FCTA en los dos escenarios produjeron las siguientes observaciones:

- FCTA produce mejores resultados que los métodos de pronóstico en la combinación.
- FCTA produce excelentes resultados de pronóstico para todas las frecuencias en conjuntos de datos (Y, Q, W, D) excepto para las series mensuales.

#### 4.2.1.1 Pruebas Estadísticas

La prueba de hipótesis no paramétrica de Friedman busca probar las diferencias entre grupos (tres o más grupos pareados) cuando el mismo parámetro se ha medido bajo diferentes condiciones en el mismo sujeto. Wilcoxon, también una prueba no paramétrica, busca encontrar diferencias significativas entre dos grupos específicos. Dado que la prueba de Friedman no da el rango promedio para cada uno de los grupos, la idea es ejecutar la prueba de Friedman utilizando todos los métodos individuales y propuestos, y luego ejecutar la

prueba de Wilcoxon con las dos frecuencias superiores. Los resultados de las pruebas estadísticas que utilizan la distribución normal estándar se muestran en la Tabla 4.5.

Tabla 4.5: Resultados Friedman y Wilcoxon con distribución de peso normal

Subset	Friedman			Wilcoxon		Conclusión	FCTA Posición
	1st y 2nd Mejores	% Certeza	FCTA Position	FCTA con Ganador	% Certeza		
<b>Y</b>	FCTAN - Bagged	99.9 %	1st	FCTAN - Bagged	73.3 %	FCTAN ≈ Bagged	<b>1st ,1st</b>
<b>Q</b>	FCTAN - Holt	99.9 %	1st	FCTAN	99.9 %	<b>FCTAN</b>	<b>1st</b>
<b>M</b>	FCTAN - FFORMA	99.9 %	1st	FCTAN	99.9 %	<b>FCTAN</b>	<b>1st</b>
<b>W</b>	FCTAN - FFORMA	99.9 %	1st	FCTAN	99.9%	<b>FCTAN</b>	<b>1st</b>
<b>D</b>	FCTAN - Jaganathan	99.9 %	1st	FCTAN	99.9%	<b>FCTAN</b>	1st

De acuerdo con la Tabla 4.5, la prueba de Friedman indica que con un 99,9% de certeza, el FCTAN es el mejor método en comparación con todos los individuos en los subconjuntos Y, Q, M, W y D. Para verificar qué tan significativo es, la diferencia entre el mejor y el segundo mejor método de pronóstico, proceda a la prueba de Wilcoxon. Los resultados indican que el FCTAN es estadísticamente equivalente a Bagged para el conjunto Y en esta segunda prueba. Sin embargo, en los otros subconjuntos, el FCTAN es estadísticamente superior al segundo con una certeza del 99,9%.

Tabla 4.6: Resultados Friedman y Wilcoxon con distribución de peso uniforme

Subset	Friedman			Wilcoxon		Conclusión	FCTA Posición
	1st y 2nd Mejores	% Certeza	FCTA Position	FCTA con Ganador	% Certeza		
<b>Y</b>	FCTAU - Bagged	99.9 %	1st	FCTAU - Bagged	95.2%	<b>FCTAU</b>	<b>1st</b>
<b>Q</b>	FCTAU - Holt	99.9 %	1st	FCTAU	99.9 %	<b>FCTAU</b>	<b>1st</b>
<b>M</b>	FCTAU - FFORMA	99.9 %	1st	FCTAU	99.9 %	<b>FCTAU</b>	<b>1st</b>
<b>W</b>	FCTAU - FFORMA	99.9 %	1st	FCTAU	99.9%	<b>FCTAU</b>	1st
<b>D</b>	FCTAU - Jaganathan	99.9 %	1st	FCTAU	99.9%	<b>FCTAU</b>	1st

En la Tabla 4.6, podemos observar los resultados de las pruebas estadísticas cuando se utiliza la distribución uniforme del peso. La prueba de Friedman indica que con un 99,9% de certeza, el FCTAU es el mejor método en comparación con todos los individuos en los subconjuntos Y, Q, M, W y D. La prueba de Wilcoxon se ejecutó para verificar la diferencia entre el mejor y el segundo mejor pronóstico. Los resultados de Wilcoxon indican que el FCTAU es estadísticamente superior al segundo, con un 95,2% de certeza en el conjunto Y además con un 99,9% de certeza en los subconjuntos Q, M, W y D. Los resultados de Wilcoxon para todos los subconjuntos de la distribución de peso normal estándar se pueden verificar en las Tablas 4.7 a la 4.12.

La Tabla 4.7 muestra los resultados de la prueba de Wilcoxon aplicados al FCTAN y al método de pronóstico Bagged. La suma de rangos negativos es mayor que la suma de rangos positivos. Por lo tanto, el error sMAPE del método FCTAN es menor que el método Bagged.

Por otro lado, el valor p es igual a 0,267, (Tabla 4.12) indicando que ambos métodos son estadísticamente equivalentes.

Tabla 4.7: Prueba Wilcoxon con rangos para el subconjunto *Yearly* para la distribución de peso normal estándar

<b>FCTAN - Bagged</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	32a	29.19	934.00
Rangos positivos	24b	27.58	662.00
Vínculos (Empates)	0c		
Total	56		

a. FCTA < bagged    b. FCTA > bagged    c. FCTA = bagged

Tabla 4.8: Prueba Wilcoxon con rangos para el subconjunto *Quarterly* para la distribución de peso normal estándar

<b>FCTAN - Holt</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	157a	143.37	22509.00
Rangos positivos	99b	104.92	10387.00
Vínculos (Empates)	0c		
Total	256		

a. FCTA < holt                      b. FCTA > holt                      c. FCTA = holt

Tabla 4.9: Prueba Wilcoxon con rangos para el subconjunto *Monthly* para la distribución de peso normal estándar

<b>FCTAN - FFORMA</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	6782a	9497.00	64408632.00
Rangos positivos	11563b	8982.97	103870053.0
Vínculos (Empates)	0c		
Total	18345		

a. FFORMA < FCTA                      b. FFORMA > FCTA                      c. FFORMA = FCTA

Tabla 4.10: Prueba Wilcoxon con rangos para el subconjunto *Weekly* para la distribución de peso normal estándar

<b>FCTAN - FFORMA</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	86a	69.37	5966.00
Rangos positivos	39b	48.95	1909.00
Vínculos (Empates)	0c		
Total	125		

a. FFORMA < FCTA      b. FFORMA > FCTA      c. FFORMA = FCTA

Tabla 4.11: Prueba Wilcoxon con rangos para el subconjunto *Daily* para la distribución de peso normal estándar

<b>FCTAN - Jaganathan</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	611a	504.10	308005.00
Rangos positivos	329b	408.10	134265.00
Vínculos (Empates)	0c		
Total	940		

a. FCTA < Jaganathan      b. FCTA > Jaganathan      c. FCTA = Jaganathan

Tabla 4.12: Resultados de la prueba estadística W de Wilcoxon para todos los subconjuntos para la distribución de peso normal estándar

<b>Subset</b>	<b>1st y 2nd Best</b>	<b>Z</b>	<b>p-Value</b>
<b>Y</b>	FCTAN - Bagged	-1.109b	0.267
<b>Q</b>	FCTAN - Holt	-5.111b	< 0.001
<b>M</b>	FCTAN - FFORMA	-27.507c	< 0.001
<b>W</b>	FCTAN - FFORMA	-4.998b	< 0.001
<b>D</b>	FCTAN - Jaganathan	-10.433b	< 0.001

a. Prueba de rango firmado de Wilcoxon      b. Se basa en rangos positivos.

c. Se basa en rangos negativos.

Mientras tanto, las Tablas 4.8 a 4.11, para frecuencias de Q, M, W y D, muestran que la suma de rangos negativos es mayor que la suma de rangos positivos con valores de p inferiores a 0.001, lo que significa que los métodos comparados tienen una diferencia estadísticamente significativa. Por lo tanto, en general, según la prueba de Wilcoxon, FCTAN es el mejor método.

Los resultados de Wilcoxon para todo el subconjunto con distribución uniforme del peso se pueden verificar en las Tablas 4.13 a la 4.18.

Tabla 4.13: Prueba Wilcoxon con rangos para el subconjunto *Yearly* para la distribución de peso uniforme

<b>FCTAU - bagged</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	35a	29.71	1040.00
Rangos positivos	21b	26.48	556.00
Vínculos (Empates)	0c		
Total	56		

a. FCTA < bagged    b. FCTA > bagged    c. FCTA = bagged

Tabla 4.14: Prueba Wilcoxon con rangos para el subconjunto *Quarterly* para la distribución de peso uniforme

<b>FCTAU - Holt</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	149a	150.44	22415.00
Rangos positivos	107b	97.95	10481.00
Tie	0c		
Total	256		

a. FCTA < holt                                    b. FCTA > holt                                    c. FCTA = holt

Tabla 4.15: Prueba Wilcoxon con rangos para el subconjunto *Monthly* para la distribución de peso uniforme

<b>FCTAU - FForma</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	11448 a	8955.96	102527846.0
Rangos positivos	6897b	9533.25	65750839.00
Vínculos (Empates)	0c		
Total	18345		

a. FFORMA < FCTA      b. FFORMA > FCTA      c. FFORMA = FCTA

Tabla 4.16: Prueba Wilcoxon con rangos para el subconjunto *Weekly* para la distribución de peso uniforme

<b>FCTAU - FFORMA</b>	<b>N</b>	<b>Mean range</b>	<b>Suma de Rangos</b>
Rangos negativos	87a	68.06	592100
Rangos positivos	38b	51.42	1954.00
Vínculos (Empates)	0c		
Total	125		

a. FFORMA < FCTA      b. FFORMA > FCTA      c. FFORMA = FCTA

Tabla 4.17: Prueba Wilcoxon con rangos para el subconjunto *Daily* para la distribución de peso uniforme

<b>FCTAU - Jaganathan</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	617a	492.18	303673.00
Rangos positivos	323b	429.09	138597.00
Vínculos (Empates)	0c		
Total	940		

a. FFORMA < Jaganathan      b. FFORMA > Jaganathan      c. FFORMA = Jaganathan

Tabla 4.18: Resultados de la prueba estadística W de Wilcoxon para todos los subconjuntos para la distribución de peso uniforme

Subset	1st y 2nd Best	Z	p-Value
Y	FCTAU - Bagged	-1.974b	0.048
Q	FCTAU- Holt	-5.032b	< 0.001
M	FCTAU - FFORMA	-25.636c	< 0.001
W	FCTAU - FFORMA	-4.887b	< 0.001
D	FCTAU - Jaganathan	-9.913b	< 0.001

- a. Prueba de rango firmado de Wilcoxon    b. Se basa en rangos positivos.  
 c. Se basa en rangos negativos.

La Tabla 4.13 muestra los resultados de la prueba de Wilcoxon aplicados a los métodos de pronóstico FCTAU y Bagged. La suma de rangos negativos es mayor que la suma de rangos positivos. Por lo tanto, el error sMAPE del método FCTAU es menor que el método Bagged. Por otro lado, el valor p es igual a 0,048 (Tabla 4.18), lo que indica que los métodos comparados tienen una diferencia estadísticamente significativa. Mientras tanto, las Tablas 4.14 a 4.17 para las frecuencias Y, Q, W y D muestran que la suma de rangos negativos es mayor que la suma de rangos positivos con valores de p inferiores a 0.001, lo que significa que los métodos comparados tienen una diferencia estadísticamente significativa. Por lo tanto, en general, según la prueba de Wilcoxon, FCTAU es el mejor método.

Los resultados se han trazado en diagramas de caja para visualizar la diferencia entre los métodos 1° y 2° en todos los subconjuntos. Por ejemplo, las Figuras 4.6 a 4.10 muestran la variabilidad de la medida sMAPE por frecuencia de series temporales para los dos mejores métodos de pronóstico, donde las gráficas izquierdas (a) son la distribución de peso normal estándar y las gráficas derechas (b) corresponden a la distribución de peso uniforme.

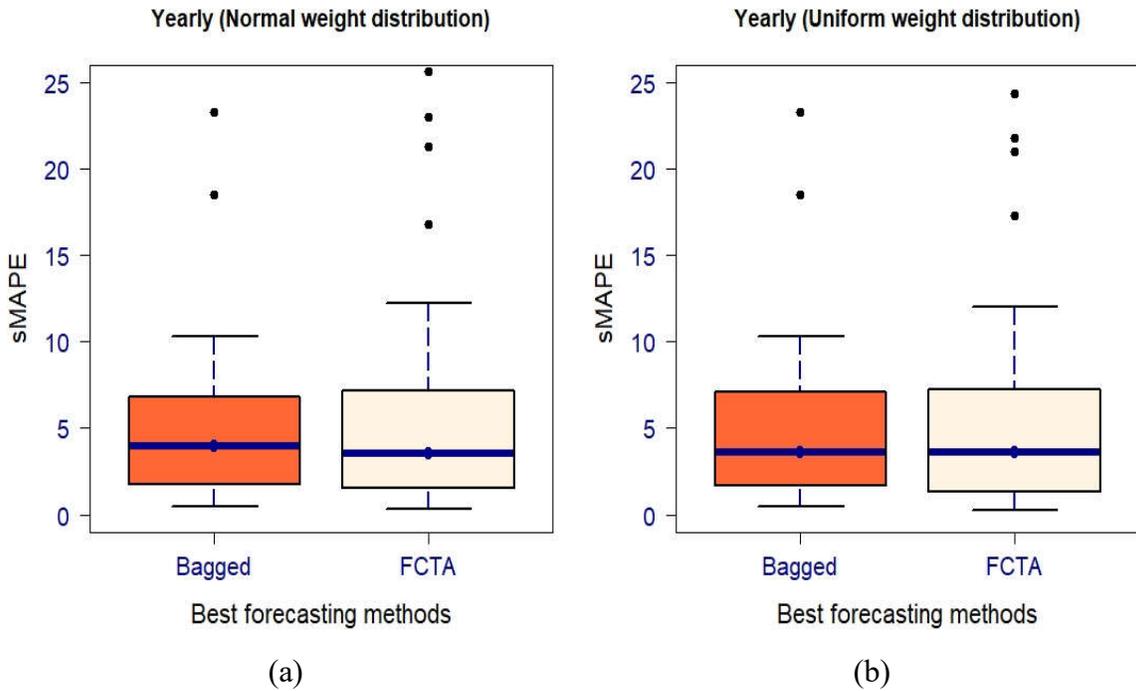


Figura 4.6: Comparativa entre FCTA y el mejor método para el subconjunto Yearly

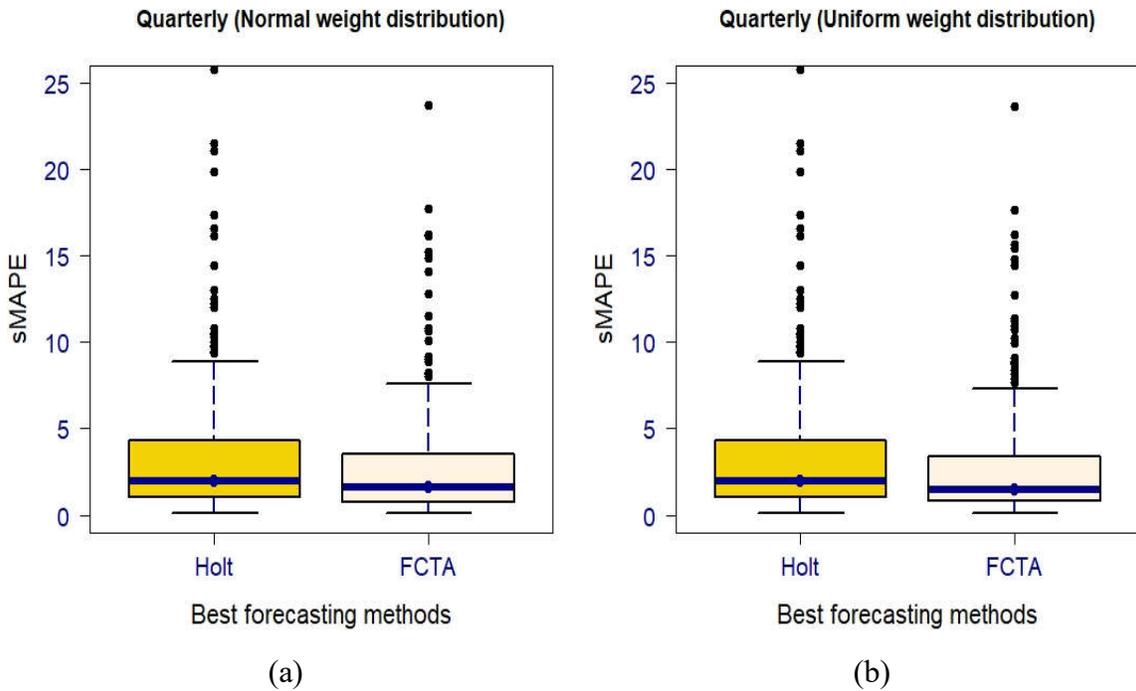


Figura 4.7: Comparativa entre FCTA y el mejor método para el subconjunto Quarterly

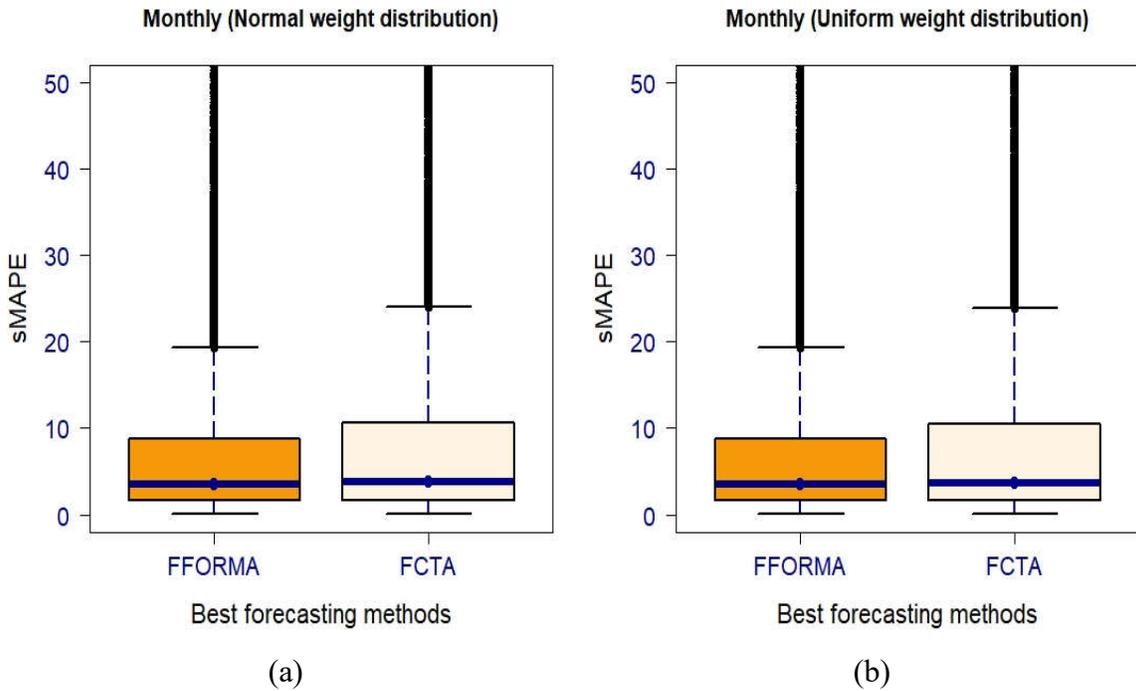


Figura 4.8: Comparativa entre FCTA y el mejor método para el subconjunto Monthly

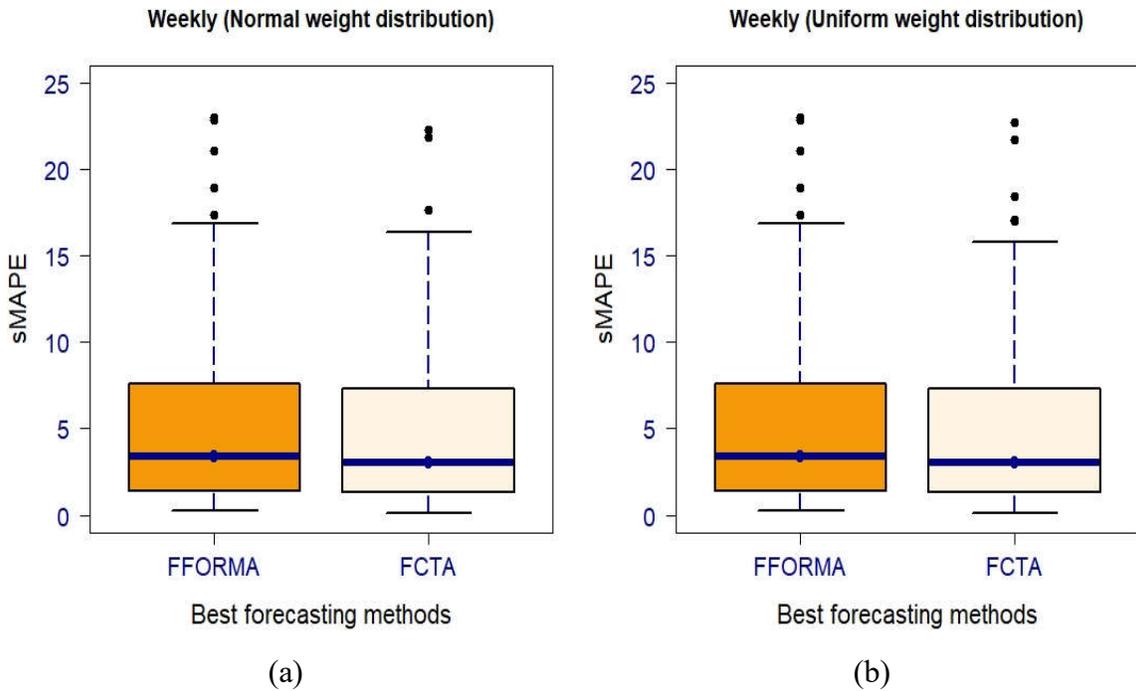


Figura 4.9: Comparativa entre FCTA y el mejor método para el subconjunto Weekly

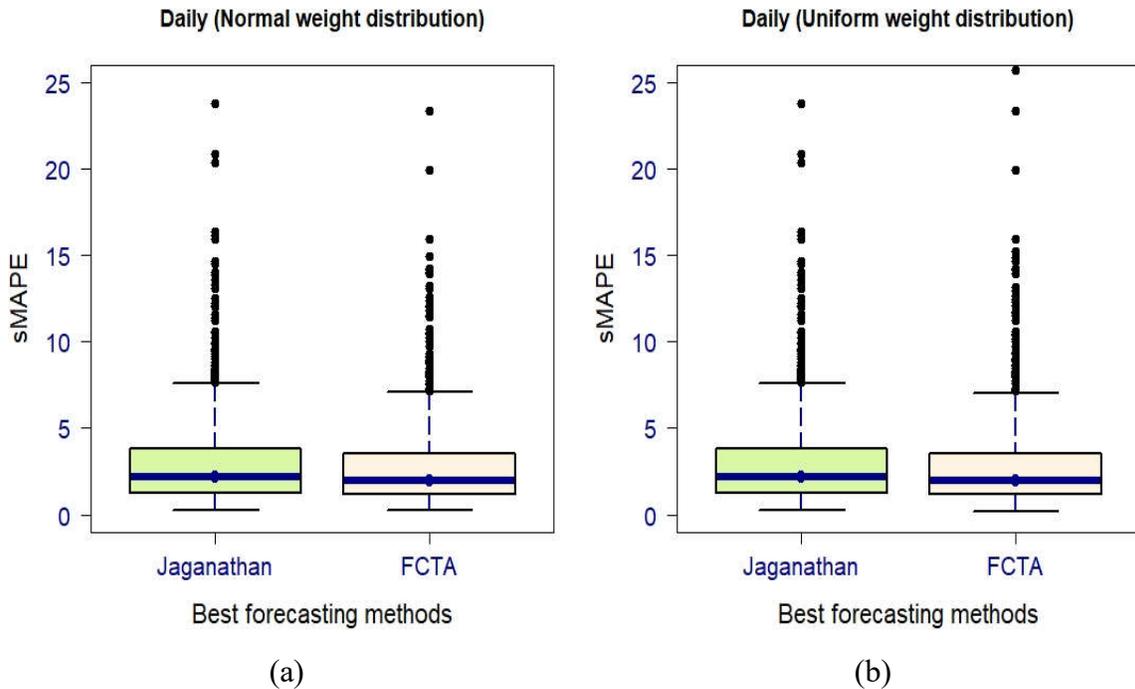


Figura 4.10: Comparativa entre FCTA y el mejor método para el subconjunto Daily

Las Figuras 4.7 y 4.10 (frecuencia Q y D) muestran una menor variabilidad del error sMAPE del método FCTA. Mientras tanto, la Figura 4.9 (W) muestra una variabilidad de error equivalente entre los dos métodos. Finalmente, las Figuras 4.6 y 4.8 ilustran claramente la mayor variabilidad de error con FCTA.

### 4.2.2 Comparativa entre FCTA y otros ensambles

En este experimento contrasta el FCTA con los métodos actuales de vanguardia para combinar pronósticos; cinco de los 14 métodos individuales se utilizan para combinar los pronósticos AA, ET, HO, TH, BA (explicados en la Tabla 4). Las metodologías de comparación son FCTA (con FCTAU, FCTAN); LA (con LA-SW, LA-MW) [3]; PP-MA [2]; y BIC [84], [85] (con BICU, BICN). BICU y BICN son métodos que utilizan el método de *Selección* de FCTA con la única diferencia que la selección de los métodos individuales que se combinarán se realiza según la medida del BIC al contrario de FCTA que lo hace según la medida de error. BICU y BICN optimizan los pesos asignados a cada uno de estos métodos utilizando el algoritmo Threshold Accepting Enhanced (TAE) propuesto. En esta experimentación, los algoritmos LA y PP-MA no utilizan la optimización propuesta por TAE

para mejorar sus pesos. El experimento utiliza las primeras 30 series temporales de cada grupo obtenidas por el proceso de filtrado mostrado en la Tabla 4.1. Este experimento evalúa los algoritmos de combinación de pronóstico en términos de la calidad de la solución final y el tiempo para obtenerla. Cada metodología de combinación se ejecuta utilizando el mismo conjunto de métodos individuales y el mismo conjunto de series temporales. Las diferencias consisten en cómo las metodologías exploran las series temporales y combinan los métodos individuales. El experimento consistió en probar para cada metodología su pronóstico de un día por delante y el pronóstico de n días por delante (determinado con los mismos métodos individuales). Los resultados del promedio de error sMAPE para las 150 series temporales utilizadas en el análisis se muestran en la Tabla 4.19. La primera fila presenta los resultados para el horizonte de un día por delante, mientras que la segunda fila presenta los resultados para el horizonte de n (determinado por la frecuencia de la serie de la M4) días por delante. Finalmente, la fila final muestra el tiempo promedio de ejecución en segundos para procesar la serie temporal 150.

Los resultados presentados en la Tabla 5.19 indican que, en promedio, FCTAN tiene un error promedio más bajo que los otros algoritmos de pronóstico de conjunto en ambos horizontes de pronóstico (H1, Hn).

Tabla 4.19: Resultados de errores promedio de pronóstico sMAPE y tiempos de ejecución para todos los conjuntos

Horizonte	FCTAU	FCTAN	LA-SW	LA-MW	PP-MA	BICU	BICN
<b>H1 sMAPE</b> <b>prom</b>	3.38	<b>3.35</b>	3.84	3.50	3.46	3.46	3.46
<b>Hn sMAPE</b> <b>prom</b>	5.06	<b>5.00</b>	5.64	5.46	5.64	5.11	5.09
<b>Tiempo</b> <b>segundos</b>	9.38	10.28	611.03	608.96	607.34	9.71	9.88

En cuanto a los tiempos de ejecución, las diferentes variantes FCTA tienen un tiempo más corto que las variantes LAs y PP-MA. Para validar los resultados, se aplicaron dos pruebas estadísticas: las pruebas de Friedman y Wilcoxon mostradas en la Tabla 4.20.

La Tabla 4.20 muestra las pruebas de Friedman y Wilcoxon, donde la primera columna muestra el horizonte, que puede ser 1 o n. Las columnas segunda y tercera muestran el resultado de la prueba de Friedman, los métodos ordenados de mejor a peor, y el porcentaje de certeza de que hay una diferencia estadística entre esos métodos. Finalmente, las columnas cuarta y quinta muestran el resultado de la prueba de Wilcoxon entre los métodos más relevantes identificados por la prueba de Friedman y el porcentaje de certeza de la diferencia estadística.

Según la Tabla 4.20, FCTAN muestra el mejor desempeño con respecto a la prueba de Friedman para ambos horizontes. Además, el porcentaje de certeza de que existe una diferencia estadística entre los métodos es del 99,4% y el 100% para el horizonte 1 (H1) y n (HN), respectivamente. Por lo tanto, para H1, ejecutamos una prueba de Wilcoxon comparando FCTAN con BICN; esta prueba muestra que FCTAN tiene estadísticamente el mismo rendimiento que BICN. Además, la prueba de Wilcoxon para HN comparó los resultados de FCTAN con BICN, mostrando una diferencia estadística con un porcentaje de certeza del 99,8%.

Los resultados de Wilcoxon para el horizonte H1 y Hn se pueden verificar en las Tablas 4.21 a la 4.23. La Tabla 4.21 muestra la prueba de Wilcoxon para el FCTAN y el BICN.

Tabla 4.20: Friedman y Wilcoxon para los resultados del horizonte H1 y Hn

Horizonte	Friedman		Wilcoxon	
	Métodos	% Certeza	Método con Ganador	% Certeza
1	1.- FCTAN 2.- BICN 3.- FCTAU 4.- BICU	5.- LA-MW 6.- LA-SW 7.- PP-MA	99.4 %	FCTAN – BICN 90.2 %
n	1.- FCTAN 2.- FCTAU 3.- BICN 4.- BICU	5.- PP-MA 6.- LA-MW 7.- LA-SW	100 %	FCTAN - BICN 99.8%

La suma de rangos negativos es mayor que la suma de rangos positivos; el error sMAPE del método FCTA es menor que el método BICN. Por otro lado, el valor p es igual a 0,198 (véase la Tabla 4.23), lo que demuestra que ambos métodos tienen el mismo rendimiento estadístico. La Tabla 4.22 muestra los resultados de la prueba de Wilcoxon aplicados a los métodos de pronóstico FCTAN y BICN. La suma de rangos negativos es mayor que la suma de rangos positivos. Por lo tanto, el error sMAPE del método FCTAN es menor que el método BICN. Finalmente, el valor p es igual a 0,002 (ver Tabla 4.23), lo que indica que ambos métodos tienen una diferencia estadísticamente significativa. Por lo tanto, según las pruebas de Wilcoxon, no hay diferencias significativas entre FCTAN y BICN en H1.

Tabla 4.21: Prueba W de Wilcoxon con rangos para el horizonte H1.

<b>FCTAN – BICN</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	60a	54.07	3244.00
Rangos positivos	46b	52.76	2427.00
Vínculos (Empates)	44c		
Total	150		

a. FCTAN < BICN    b. FCTAN > BICN    c. FCTAN = BICN

Tabla 4.22: Prueba W de Wilcoxon con rangos para el horizonte Hn

<b>FCTAN – BICN</b>	<b>N</b>	<b>Rango Promedio</b>	<b>Suma de Rangos</b>
Rangos negativos	71a	53.45	3795.00
Rangos positivos	35b	53.60	1876.00
Vínculos (Empates)	44c		
Total	150		

a. FCTAN < BICN    b. FCTAN > BICN    c. FCTAN = BICN

Tabla 4.23: Resultados estadísticos W de Wilcoxon para todos los subconjuntos

<b>Subset</b>	<b>Métodos</b>	<b>Z</b>	<b>p-Value</b>
<b>H1</b>	<b>FCTAN – BICN</b>	-1.29	0.198
<b>Hn</b>	<b>FCTAN - BICN</b>	-3.02	0.002

En Hn FCTAN supera a los otros métodos. Finalmente, para verificar la calidad del FCTA, se usa el error ratio [27], pero con la métrica de error sMAPE (ecuación 4.1).

$$ErrorRatio = \frac{sMAPE_{Evaluating\ Algorithm}}{sMAPE_{Baseline\ Algorithm}} \quad 4.1$$

La Figura 4.11 muestra la relación de error de la metodología de combinación mejor propuesta FCTAN en comparación con otros algoritmos de pronóstico de conjuntos; de acuerdo con esta métrica, FCTAN tiene el mejor rendimiento que todas las demás metodologías de la comparación.

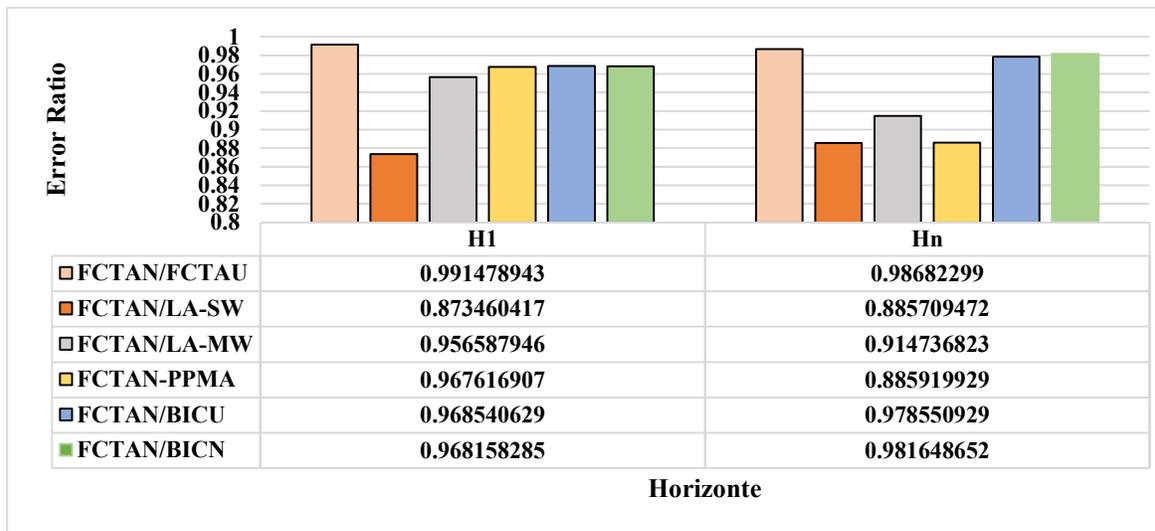


Figura 4.11: Relación de error del FCTAN propuesto en comparación con otros algoritmos de pronóstico de conjuntos

### 4.2.3 Pronóstico de Energía con FCTA

En este experimento para el pronóstico de la demanda y consumo de energía en México se aplicó la metodología FCTA al conjunto de series de la sección 4.1.3.2. usando los métodos de la Tabla 4.2.

El experimento aplica las dos técnicas de distribución de peso (Uniforme o Normal Estándar) propuestas en FCTA y la optimización con el algoritmo TAE. Cada serie del conjunto de energía se divide en cuatro porciones: Entrenamiento (70% de la serie), Validación1 (10%), Validación2 (10%), y Prueba (10%) como se muestra en la Figura 4.12.

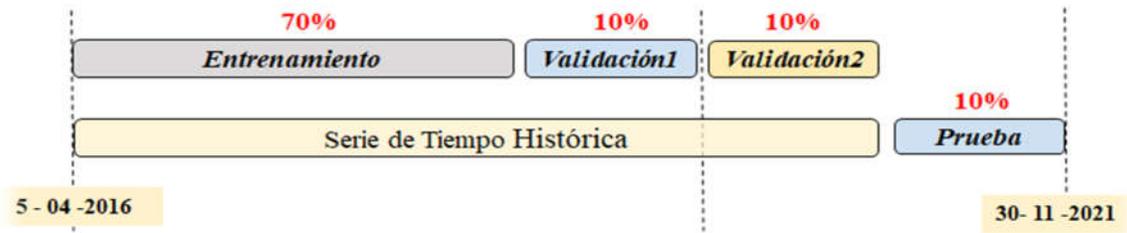


Figura 4.12: División de las series de Tiempo experimento energía

Siguiendo la misma secuencia de pasos que se explica en la sección 4.2.1 se obtiene el pronóstico en la porción de la Prueba y se compara contra el mismo pronóstico en la porción de la Prueba obtenido de cada uno de los métodos individuales. Los resultados de esta comparativa para un horizonte de 10 valores futuros se muestran en la Tabla 4.24.

Los resultados indican que, en promedio, FCTA tanto utilizando una distribución de peso uniforme (FCTAU) como utilizando una distribución de peso normal estándar (FCTAN) presentan un error promedio más bajo que los mejores métodos individuales. En el gráfico de la Figura 4.13 se puede visualizar los resultados. En este grafico se comparan los métodos propuestos (FCTAU, FCTAN) contra los 5 mejores métodos individuales para estas series.

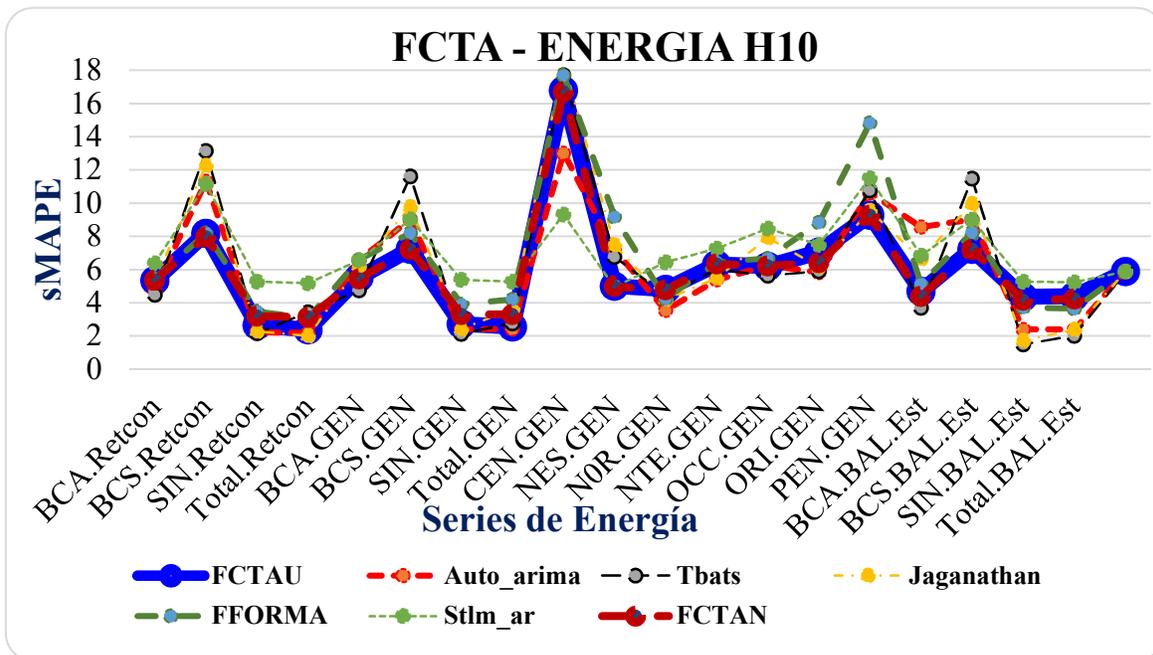


Figura 4.13: Comparativa entre FCTAU y FCTAN versus métodos individuales

Tabla 4.24: Comparativa de errores sMAPE promedio de pronóstico horizonte 10

Subset	sMAPE (Promedio de errores de pronóstico Métodos individuales)														Propuestas	
	AA	ET	NN	TB	ST	RW	TH	NA	AR	BA	SP	HO	FF	JA	FCTAN	FCTAU
<b>BCA. GEN</b>	6.6	8.1	6.7	4.8	6.6	8.1	8.1	8.1	3.9	11.6	6.9	8.0	6.6	6.2	5.4	5.5
<b>BCS. GEN</b>	9.0	8.2	9.1	11.6	9.0	7.2	7.2	7.2	11.3	8.4	9.5	7.2	8.2	9.8	7.2	7.2
<b>SIN. GEN</b>	2.4	4.4	6.7	2.1	5.4	4.2	4.7	4.2	4.3	5.1	4.6	4.6	3.9	2.4	3.3	2.7
<b>Total. GEN</b>	2.4	4.5	6.5	2.8	5.3	4.4	4.9	4.4	4.5	5.3	4.6	4.9	4.2	3.7	3.3	2.6
<b>CEN. GEN</b>	13.0	24.3	16.0	17.7	9.3	21.9	24.4	21.7	8.9	32.6	20.9	24.3	17.7	18.5	16.7	16.8
<b>NES. GEN</b>	7.2	12.4	5.2	6.8	5.0	12.5	12.4	12.4	5.4	10.4	12.0	12.4	9.2	7.5	5.0	5.0
<b>NOR. GEN</b>	3.6	4.6	4.6	4.5	6.5	4.3	4.6	4.3	5.1	10.5	10.1	4.5	4.2	4.2	4.8	4.8
<b>NTE. GEN</b>	5.4	6.2	9.7	6.0	7.3	6.0	6.2	6.0	8.3	8.0	6.8	6.2	6.5	5.5	6.3	6.3
<b>OCC. GEN</b>	6.2	9.5	6.0	5.7	8.5	9.5	9.6	9.5	9.3	9.7	9.4	9.6	6.7	7.9	6.2	6.2
<b>ORI. GEN</b>	5.8	9.3	5.4	5.9	7.5	12.1	11.9	12.1	9.7	8.3	10.6	11.7	8.8	6.2	6.4	7.1
<b>PEN. GEN</b>	10.6	9.3	22.3	10.8	11.5	10.2	9.3	10.2	12.3	8.9	9.5	9.3	14.8	9.6	9.3	9.3
<b>BCA. Est</b>	8.6	8.1	5.1	3.7	6.9	8.1	8.1	8.1	9.4	12.5	6.6	8.1	5.2	6.7	4.4	4.7
<b>BCS. Est</b>	9.0	8.2	9.1	11.5	9.0	7.2	7.2	7.2	11.0	8.4	9.5	7.2	8.2	10.0	7.2	7.2
<b>SIN. Est</b>	2.4	4.4	6.7	1.5	5.3	4.2	4.6	4.2	7.5	5.0	4.5	4.6	3.7	1.7	4.2	4.4
<b>Total. Est</b>	2.4	4.6	6.3	2.0	5.2	4.3	4.8	4.3	4.4	5.2	4.5	4.8	3.7	2.4	4.2	4.4
<b>PROM</b>	6.1	8.2	8.0	6.3	7.2	7.9	8.2	7.9	7.5	9.6	8.3	8.2	6.9	6.5	<b>6.0</b>	<b>5.9</b>

Observamos en la Figura 4.13 que FCTAU y FCTAN obtienen un buen desempeño consistente pronosticando para un horizonte de 10 días por delante y en algunas series se presentan un pronóstico con un error más bajo.

El mismo experimento se realizó para el pronóstico usando solo las series de demanda energética (BCA.Retcon, BCS.Retcon, SIN.Retcon, Total.Retcon). En el experimento se realizó un pronóstico del día siguiente de la demanda, los resultados de la comparativa entre los métodos propuestos y los individuales se muestran en la Tabla 4.25.

Tabla 4.25: Comparativa de errores sMAPE promedio de pronóstico horizonte 1

Subset	sMAPE (Promedio de errores de pronóstico Métodos individuales)														Propuestos	
	AA	ET	NN	TB	ST	RW	TH	NA	AR	BA	SP	HO	FF	JA	FCTAN	FCTAU
<b>BCA. GEN</b>	1.6	6.8	3.6	1.6	1.8	6.8	6.8	6.8	2.9	10.4	4.1	6.8	3.3	3.2	2.8	2.8
<b>BCS. GEN</b>	9.0	10.9	4.8	10.5	9.7	6.0	8.9	6.0	10.4	10.3	10.9	8.9	7.1	10.7	5.5	5.6
<b>SIN. GEN</b>	2.2	6.0	5.2	0.6	3.9	5.6	6.0	5.7	3.3	6.9	6.0	6.2	4.2	2.9	2.7	2.5
<b>Total. GEN</b>	2.2	6.0	5.4	1.4	3.8	5.7	6.0	5.7	3.2	6.9	5.9	6.1	3.9	2.6	2.6	2.4
<b>CEN. GEN</b>	2.4	6.5	6.6	3.9	2.4	6.5	6.5	6.5	2.6	11.6	6.4	6.5	6.3	4.3	5.4	5.5
<b>NES. GEN</b>	6.1	7.1	5.0	6.7	6.1	5.8	5.8	5.8	6.6	7.4	10.6	5.8	6.1	6.9	5.6	5.7
<b>NOR. GEN</b>	2.4	6.1	6.1	0.6	4.2	5.8	6.5	5.8	3.4	7.3	6.3	6.5	5.1	3.0	3.2	2.9
<b>NTE. GEN</b>	2.6	6.1	6.5	1.1	4.0	5.8	6.7	5.8	3.5	7.4	6.3	6.7	5.8	4.6	3.5	3.3
<b>OCC. GEN</b>	8.0	11.0	0.8	0.2	7.7	8.5	11.0	8.4	6.7	19.5	0.9	11.0	1.5	1.8	0.8	1.0
<b>ORI. GEN</b>	4.7	2.7	4.6	0.6	5.4	2.6	2.7	2.6	6.3	5.2	9.8	2.7	3.7	1.8	2.6	2.3
<b>PEN. GEN</b>	1.7	3.5	1.4	0.1	3.0	2.3	3.5	2.3	2.1	11.0	7.7	3.4	1.7	1.3	1.3	1.3
<b>BCA. Est</b>	1.6	3.9	0.4	3.0	3.9	3.4	3.9	3.4	4.0	6.9	5.5	3.9	1.7	2.0	0.9	1.1
<b>BCS. Est</b>	1.5	5.4	1.2	0.3	2.4	5.6	5.9	5.6	4.4	6.2	2.7	5.9	1.5	2.7	2.6	2.9
<b>SIN. Est</b>	4.8	8.0	6.2	6.3	4.9	11.5	11.3	11.5	7.9	6.8	6.7	11.3	8.2	6.0	5.7	5.6
<b>Total. Est</b>	10.6	9.1	13.6	10.8	11.1	10.7	9.1	10.7	11.7	4.3	1.0	9.1	13.3	10.0	10.1	10.1
<b>PROM</b>	5.2	6.8	4.5	<b>2.1</b>	<b>2.7</b>	6.8	6.8	6.8	4.4	12.5	5.5	6.8	4.9	3.9	<b>3.8</b>	<b>3.9</b>

Los resultados indican que, para este horizonte, (FCTAU, FCTAN quedan por detrás de TB y ST los cuales presentan un rendimiento promedio mejor en este horizonte. Si se grafican los resultados obtenidos para las series relacionadas con el consumo energético (Figura 4.14) podemos ver que FCTAU y FCTAN presentan un buen rendimiento que es comparable con el pronóstico del CENACE para el mismo horizonte.

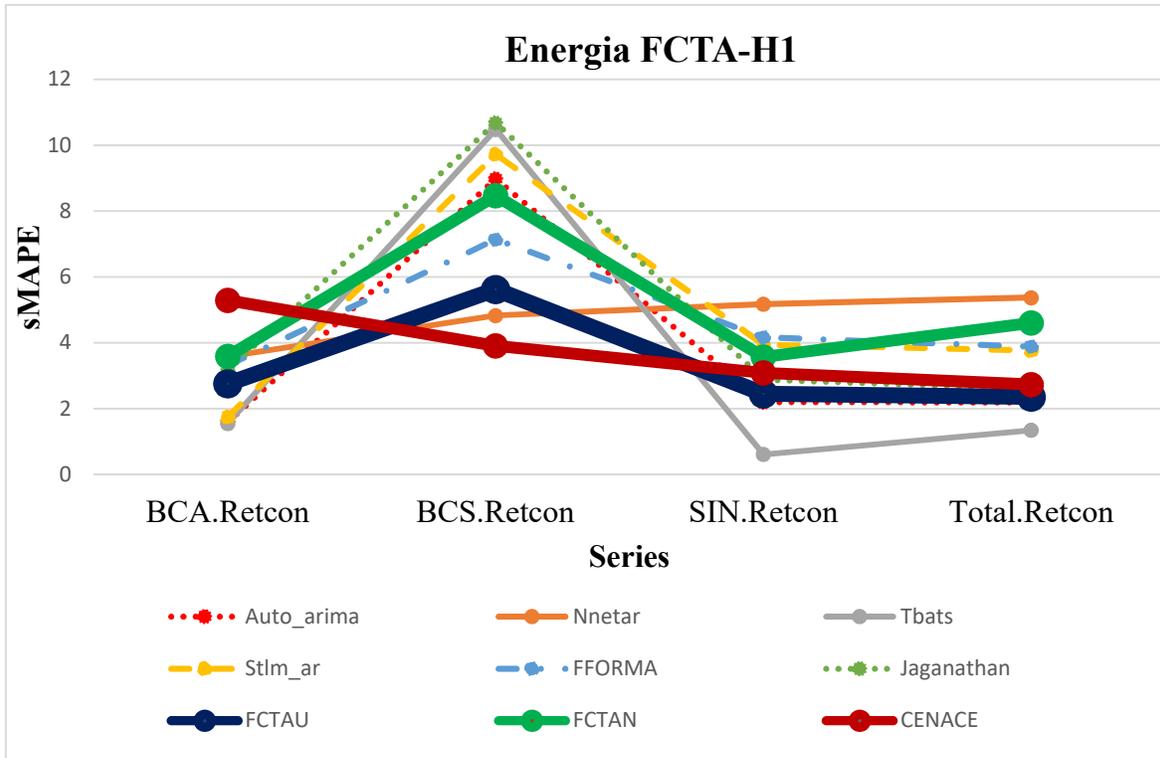


Figura 4.14: Comparativa FCTA Contra CENACE

A partir de los resultados mostrados por la Figura 4.14 podemos ver que el pronóstico usando la metodología FCTA es comparable al pronóstico del CENACE. FCTAN es mejor para 3 de las series evaluadas.

En base a los anteriores experimentos y los resultados obtenidos en cada uno podemos concluir que FCTA en sus dos variantes produce resultados consistentemente buenos para diferentes horizontes de pronósticos, en diferentes escenarios y pronosticando series de diferentes áreas, por lo tanto, podemos concluir que el uso de la metodología FCTA puede aplicarse al pronóstico de series del mercado de energía mexicano con buenos resultados.

---

---

## Conclusiones y trabajos futuros

### 5.1 Conclusiones

En el área de pronóstico, los grandes retos siempre se han enfocado en la mejora de la capacidad de obtener buenos pronósticos para 1 y n periodos hacia adelante en múltiples escenarios y con datos de diversas áreas incluso energía. Este trabajo presenta la metodología de pronóstico FCTA para series temporales de diferentes escenarios.

Las variantes de FCTAU y FCTAN se compararon en un primer experimento contra algunos de los métodos clásicos más exitosos y los mejores métodos de referencia de la competencia M4. En ambos períodos 1 y n hacia adelante, FCTAN fue estadísticamente superior a todos estos métodos.

En un segundo experimento, las variantes de FCTA se compararon con otros métodos exitosos del estado del arte que no participaron en la competencia M4. Los resultados del experimento para 1 y n periodos hacia adelante, muestran que FCTAN obtiene los mejores resultados seguidos por FCTAU. Además, hibridaciones con FCTA (BICU y BICN) obtuvieron mejores resultados que los métodos de última generación.

Finalmente, los resultados del tercer experimento realizado con series de tiempo del área de energía de México indican que FCTA obtiene resultados consistentes también en esta área. Sus resultados compiten con el pronóstico provisto por el CENACE. En consecuencia, FCTA es una buena metodología y se puede hibridar con otros métodos con la condición de estos

produzcan buenos resultados. Otros métodos también pueden ser hibridados con FCTA. Por lo tanto, FCTA es una contribución significativa al área de combinación de pronósticos.

## 5.2 Trabajos Futuros

Existen diversas áreas de oportunidad que se han encontrado durante la realización de esta investigación para la mejoría de la metodología FCTA, a continuación, se enumeran las más relevantes:

- Optimizar el método de selección de los mejores métodos usando otro tipo de que no sean las de error, además de que se puede implementar un método de Selección usando técnicas de optimización.
- Ampliar el estudio realizado probando otras distribuciones de los pesos iniciales
- Implementar un método que permita aplicar Filtros en las series de tiempo antes de realizar la optimización de pesos propuesta en FCTA.
- Ampliar y probar la metodología FCTA para series difíciles.
- Trabajar en mejoras al algoritmo de TA dentro de la metodología FCTA para que esta pueda tener un desempeño superior en un número mayor de los conjuntos estudiados.

## 5.3 Publicaciones derivadas del presente trabajo

Artículo **FCTA: A Forecasting Combined methodology with a Threshold Accepting Approach** aceptado y publicado en la revista Hindawi Mathematical Problems in Engineering en Marzo 2022. <https://doi.org/10.1155/2022/6206037>

## Bibliografía

- [1] G. Santamaría-Bonfil, J. Frausto-Solís, and M. Chi-chim, “Una metodología basada en Máquinas de Vectores de Soporte para el pronóstico de la volatilidad del Índice de Precios y Cotizaciones,” *Komput. Sapiens*, vol. III, no. 4, pp. 12–17, 2010.
- [2] R. G. Brown, ““Exponential Smoothing for Predicting Demand,”” *Cambridge, Massachusetts Arthur D. Little Inc.*, vol. 5, no. 1, p. 145, 1956, [Online]. Available: <https://www.industrydocuments.ucsf.edu/tobacco/docs/#id=jzlc0130>.
- [3] O. D. Anderson, “Box-Jenkins Approach To Time Series Analysis.,” *RAIRO Rech. Oper.*, vol. 11, no. 1, pp. 3–29, 1977, doi: 10.1051/ro/1977110100031.
- [4] N. Sapankevych and R. Sankar, “Time series prediction using support vector machines: A survey,” *IEEE Comput. Intell. Mag.*, vol. 4, no. 2, pp. 24–38, 2009, doi: 10.1109/MCI.2009.932254.
- [5] R. Fildes and E. J. Lusk, “The choice of a forecasting model,” *Omega*, vol. 12, no. 5, pp. 427–435, 1984, doi: 10.1016/0305-0483(84)90042-2.
- [6] P. Newbold and C. W. Granger, “Experience with Forecasting Univariate Time Series and the Combination of Forecasts,” *J. R. Stat. Soc.*, vol. 137, no. 2, pp. 131–165, 1974, [Online]. Available: <http://www.jstor.org/stable/2344546> .
- [7] J. M. Bates ; C. W. Granger, “Combination of Forecast,” *Oper. Res. Q.*, vol. 20, no. 4, pp. 451–468, 1969, doi: 10.2307/3008764.
- [8] R. L. Winkler and S. Makridakis, “The Combination of Forecasts,” *J. R. Stat. Soc.*, vol. 146, no. 2, pp. 150–157, 1983, doi: 10.2307/2982011.
- [9] D. W. Bunn, “Statistical efficiency in the linear combination of forecasts,” *Int. J. Forecast.*, vol. 1, no. 2, pp. 151–163, 1985, doi: 10.1016/0169-2070(85)90020-2.
- [10] M. Deutsch, C. W. J. Granger, and T. Teräsvirta, “The combination of forecasts using changing weights,” *Int. J. Forecast.*, vol. 10, no. 1, pp. 47–57, 1994, doi: 10.1016/0169-2070(94)90049-3.

- [11] J. S. Armstrong, “Combining Forecasts,” in *Principles of forecasting: a handbook for researchers and practitioners*, 1st ed., USA: Springer, Boston, MA, 2001, pp. 417–439.
- [12] C. Lemke and B. Gabrys, “Meta-learning for time series forecasting and forecast combination,” *Neurocomputing*, vol. 73, no. 10–12, pp. 2006–2016, 2010, doi: 10.1016/j.neucom.2009.09.020.
- [13] M. Aiolfi, C. Capistrán, and A. G. Timmermann, “Forecast Combinations,” *SSRN Electron. J.*, 2012, doi: 10.2139/ssrn.1609530.
- [14] R. Adhikari and R. K. Agrawal, “Performance evaluation of weights selection schemes for linear combination of multiple forecasts,” *Artif. Intell. Rev.*, vol. 42, no. 4, pp. 529–548, 2014, doi: 10.1007/s10462-012-9361-z.
- [15] M. Cuartas and L. López Menéndez, Álvarez, “Comparación y combinación de predicciones: aplicación a las series temporales,” pp. 1–22, 2000.
- [16] G. P. Zhang, “Time series forecasting using a hybrid ARIMA and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003, doi: 10.1016/S0925-2312(01)00702-0.
- [17] M. Khashei and M. Bijari, “An artificial neural network (p, d, q) model for timeseries forecasting,” *Expert Syst. Appl.*, vol. 37, no. 1, pp. 479–489, 2010, doi: 10.1016/j.eswa.2009.05.044.
- [18] I. Khandelwal, R. Adhikari, and G. Verma, “Time series forecasting using hybrid arima and ann models based on DWT Decomposition,” *Procedia Comput. Sci.*, vol. 48, no. C, pp. 173–179, 2015, doi: 10.1016/j.procs.2015.04.167.
- [19] S. BuHamra, N. Smaoui, and M. Gabr, “The Box-Jenkins analysis and neural networks: Prediction and time series modelling,” *Appl. Math. Model.*, vol. 27, no. 10, pp. 805–815, 2003, doi: 10.1016/S0307-904X(03)00079-9.
- [20] A. Singh and G. C. Mishra, “Application of box-jenkins method and artificial neural network procedure for time series forecasting of prices,” *Stat. Transit.*, vol. 16, no. 1,

- pp. 83–96, 2015, doi: 10.21307/stattrans-2015-005.
- [21] Y. S. Kao, K. Nawata, and C. Y. Huang, “Predicting primary energy consumption using hybrid arima and ga-svr based on eemd decomposition,” *Mathematics*, vol. 8, no. 10, pp. 1–19, 2020, doi: 10.3390/math8101722.
- [22] Y. Zhang, L. Luo, J. Yang, D. Liu, R. Kong, and Y. Feng, “A hybrid ARIMA-SVR approach for forecasting emergency patient flow,” *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 8, pp. 3315–3323, 2019, doi: 10.1007/s12652-018-1059-x.
- [23] J. Pokora, “Hybrid arima and support vector regression in short-term electricity price forecasting,” *Acta Univ. Agric. Silvic. Mendelianae Brun.*, vol. 65, no. 2, pp. 699–708, 2017, doi: 10.11118/actaun201765020699.
- [24] O. Basheer Shukur, N. Salem Fadhil, M. Hisyam Lee, and M. Hura Ahmad, “Electricity load forecasting using hybrid of multiplicative double seasonal exponential smoothing model with artificial neural network,” *J. Teknol. (Sciences Eng.*, vol. 69, no. 2, pp. 65–70, 2014, doi: 10.11113/jt.v69.3109.
- [25] S. Smyl, “A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 75–85, 2020, doi: 10.1016/j.ijforecast.2019.03.017.
- [26] S. Tofighy, A. A. Rahmanian, and M. Ghobaei-Arani, “An ensemble CPU load prediction algorithm using a Bayesian information criterion and smooth filters in a cloud computing environment,” *Softw. - Pract. Exp.*, vol. 48, no. 12, pp. 2257–2277, 2018, doi: 10.1002/spe.2641.
- [27] A. A. Rahmanian, M. Ghobaei-Arani, and S. Tofighy, “A learning automata-based ensemble resource usage prediction algorithm for cloud computing environment,” *Futur. Gener. Comput. Syst.*, vol. 79, pp. 54–71, 2018, doi: 10.1016/j.future.2017.09.049.
- [28] S. F. Crone, M. Hibon, and K. Nikolopoulos, “Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction,”

- Int. J. Forecast.*, vol. 27, no. 3, pp. 635–660, 2011, doi: 10.1016/j.ijforecast.2011.04.001.
- [29] J. J. Ruiz-Aguilar, I. Turias, J. González-Enrique, D. Urda, and D. Elizondo, “A permutation entropy-based EMD–ANN forecasting ensemble approach for wind speed prediction,” *Neural Comput. Appl.*, vol. 33, no. 7, pp. 2369–2391, 2020, doi: 10.1007/s00521-020-05141-w.
- [30] A. N. Khan, N. Iqbal, R. Ahmad, and D. H. Kim, “Ensemble prediction approach based on learning to statistical model for efficient building energy consumption management,” *Symmetry (Basel)*, vol. 13, no. 3, pp. 1–26, 2021, doi: 10.3390/sym13030405.
- [31] T. McAndrew, N. Wattanachit, G. C. Gibson, and N. G. Reich, “Aggregating predictions from experts: A review of statistical methods, experiments, and applications,” *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 13, no. 2, pp. 1–26, 2020, doi: 10.1002/wics.1514.
- [32] M. Etemadi, M. Ghobaei-Arani, and A. Shahidinejad, “Resource provisioning for IoT services in the fog computing environment: An autonomic approach,” *Comput. Commun.*, vol. 161, no. March, pp. 109–131, 2020, doi: 10.1016/j.comcom.2020.07.028.
- [33] R. Prudêncio and T. Ludermir, “Using machine learning techniques to combine forecasting methods,” in *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 2004, vol. 3339, pp. 1122–1127, doi: 10.1007/978-3-540-30549-1\_111.
- [34] R. B. C. Prudêncio and T. B. Ludermir, “Meta-learning approaches to selecting time series models,” *Neurocomputing*, vol. 61, no. 1–4, pp. 121–137, 2004, doi: 10.1016/j.neucom.2004.03.008.
- [35] K. Mirko, C. Sven F., and F. Michael, “Meta-Learning with Neural Networks and Landmarking for Forecasting Model Selection,” *Proc. Int. Jt. Conf. Neural Networks*,

- pp. 1499–1506, 2016, doi: 10.1109/IJCNN.2016.7727376.
- [36] Y. Kang, R. J. Hyndman, and K. Smith-Miles, “Visualising forecasting algorithm performance using time series instance spaces,” *Int. J. Forecast.*, vol. 33, no. 2, pp. 345–358, 2017, doi: 10.1016/j.ijforecast.2016.09.004.
- [37] P. Montero-Manso, G. Athanasopoulos, R. J. Hyndman, and T. S. Talagala, “FFORMA: Feature-based Forecast Model Averaging FFORMA: Feature-based Forecast Model Averaging FFORMA: Feature-based Forecast Model Averaging,” no. December, 2018, [Online]. Available: <http://business.monash.edu/econometrics-and-business-statistics/research/publications>.
- [38] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 Competition: 100,000 time series and 61 forecasting methods,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 54–74, 2020, doi: 10.1016/j.ijforecast.2019.04.014.
- [39] H. J. Motulsky and R. E. Brown, “Detecting outliers when fitting data with nonlinear regression - A new method based on robust nonlinear regression and the false discovery rate,” *BMC Bioinformatics*, vol. 7, pp. 1–20, 2006, doi: 10.1186/1471-2105-7-123.
- [40] P. Raña, G. Aneiros, and J. M. Vilar, “Detection of outliers in functional time series,” *Environmetrics*, vol. 26, no. 3, pp. 178–191, 2015, doi: 10.1002/env.2327.
- [41] I. Z. Batyrshin and L. B. Sheremetov, “Perception-based approach to time series data mining,” vol. 8, pp. 1211–1221, 2008, doi: 10.1016/j.asoc.2007.02.020.
- [42] S. Jaganathan and P. K. S. Prakash, “A combination-based forecasting method for the M4-competition,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 98–104, 2020, doi: 10.1016/j.ijforecast.2019.03.030.
- [43] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 Competition: Results, findings, conclusion and way forward,” *Int. J. Forecast.*, no. June, 2018, doi: 10.1016/j.ijforecast.2018.06.001.
- [44] S. Makridakis *et al.*, “The M2-competition: A real-time judgmentally based

- forecasting study,” *Int. J. Forecast.*, vol. 9, no. 1, pp. 5–22, 1993, doi: 10.1016/0169-2070(93)90044-N.
- [45] P. J. Brockwell and R. a Davis, *Introduction to Time Series and Forecasting*, *Second Edition*. 2002.
- [46] R. S. Tsay, *Analysis of Financial Time Series*, Third Edit. John Wiley & Sons, Inc., 2010.
- [47] C. Chatfield, *Time-Series Forecasting*, 1st ed. Chapman and Hall, 2001.
- [48] 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, doi: 10.1016/j.neucom.2008.10.017.
- [49] A. M. Hanke, J. E., Wichern, D. W., & Zepeda, *Pronósticos en los negocios*. Pearson Educación. 2006.
- [50] A. C. B. Mancuso and L. Werner, “Review of Combining Forecasts Approaches,” *Indep. J. Manag. Prod.*, vol. 4, no. 1, pp. 248–277, 2013, doi: 10.14807/ijmp.v4i1.59.
- [51] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.
- [52] R. G. Pontius, O. Thontteh, and H. Chen, “Components of information for multiple resolution comparison between maps that share a real variable,” *Environ. Ecol. Stat.*, vol. 15, no. 2, pp. 111–142, 2008, doi: 10.1007/s10651-007-0043-y.
- [53] R. S. Tsay, *Analysis of Financial Time Series*. John Wiley & SONS, 2005.
- [54] P. Goodwin and R. Lawton, “On the asymmetry of the symmetric MAPE,” *Int. J. Forecast.*, vol. 15, no. 4, pp. 405–408, 1999, doi: 10.1016/S0169-2070(99)00007-2.
- [55] G. Dueck and T. Scheuer, “Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing,” *J. Comput. Phys.*, vol. 90, no. 1, pp. 161–175, 1990, doi: 10.1016/0021-9991(90)90201-B.

- [56] C. W. J. Granger and R. A. M. U. Ramanathan, "Forecasts," vol. 3, no. July 1982, pp. 197–204, 1984.
- [57] R. T. Clemen, "Combining forecasts: A review and annotated bibliography," *Int. J. Forecast.*, vol. 5, no. 4, pp. 559–583, 1989, doi: 10.1016/0169-2070(89)90012-5.
- [58] C. T. West, "System-based weights versus series-specific weights in the combination of forecasts," *J. Forecast.*, vol. 15, no. 5, pp. 369–383, 1996, doi: 10.1002/(SICI)1099-131X(199609)15:5<369::AID-FOR629>3.0.CO;2-W.
- [59] C. W. Bischoff, "The combination of macroeconomic forecasts," *J. Forecast.*, vol. 8, no. 3, pp. 293–314, 1989, doi: 10.1002/for.3980080312.
- [60] J. Smith and K. F. Wallis, "A simple explanation of the forecast combination puzzle," *Oxf. Bull. Econ. Stat.*, vol. 71, no. 3, pp. 331–355, 2009, doi: 10.1111/j.1468-0084.2008.00541.x.
- [61] C. K. Chan, B. G. Kingsman, and H. Wong, "Value of combining forecasts in inventory management - a case study in banking," *Eur. J. Oper. Res.*, vol. 117, no. 2, pp. 199–210, 1999, doi: 10.1016/S0377-2217(98)00277-X.
- [62] C. K. Chan, B. G. Kingsman, and H. Wong, "Determining when to update the weights in combined forecasts for product demand - An application of the CUSUM technique," *Eur. J. Oper. Res.*, vol. 153 153, no. 3 3 SPEC. ISS., pp. 757–768, 2004, doi: 10.1016/S0377-2217(02)00528-3.
- [63] X. Y. Feng, L. M. Wan, Y. C. Liang, Y. F. Sun, H. P. Lee, and Y. Wang, "A modified particle swarm optimization for combining forecasting," *Proc. 2004 Int. Conf. Mach. Learn. Cybern.*, vol. 4, no. August, pp. 2384–2389, 2004, doi: 10.1109/icmlc.2004.1382201.
- [64] W. G. Fan and F. Q. Deng, "Variable weight combining forecasts based on forecasting error," *2007 IEEE Int. Conf. Control Autom. ICCA*, vol. 2, pp. 1610–1613, 2007, doi: 10.1109/ICCA.2007.4376632.
- [65] X. Chen, Y. Jiang, K. Yu, Y. Liao, J. Xie, and Q. Wu, "Combined time-varying

- forecast based on the proper scoring approach for wind power generation,” *J. Eng.*, vol. 2017, no. 14, pp. 2655–2659, 2017, doi: 10.1049/joe.2017.0843.
- [66] J. S. Armstrong, *Principles of forecasting: a handbook for researchers and practitioners*, 1st ed. Springer Science & Business Media, 2001.
- [67] M. K. Starr and S. K. Gupta, “The routledge companion to production and operations management,” *Routledge Companion to Prod. Oper. Manag.*, pp. 1–712, 2017, doi: 10.4324/9781315687803.
- [68] Douglas C. Montgomery; Cheryl L. Jennings; Murat Kulahci, *Introduction to time Series Analysis and Forecasting*. Hoboken, New Jersey: John Wiley & Sons, 2007.
- [69] J. Frausto-Solis and F. Martinez-Rios, “Golden ratio annealing for satisfiability problems using dynamically cooling schemes,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4994 LNAI, pp. 215–224, 2008, doi: 10.1007/978-3-540-68123-6\_24.
- [70] S. Abdullah, L. Golafshan, and M. Z. A. Nazri, “Re-heat simulated annealing algorithm for rough set attribute reduction,” *Int. J. Phys. Sci.*, vol. 6, no. 8, pp. 2083–2089, 2011, doi: 10.5897/IJPS11.218.
- [71] CENACE, “Datos de Demanda.” <https://www.cenace.gob.mx/Paginas/SIM/Reportes/EstimacionDemandaReal.aspx>.
- [72] Rob J. Hyndman, Anne B. Koehler, Ralph D. Snyder, Simone Grose, “A state space framework for automatic forecasting using exponential smoothing methods,” *Int. J. Forecast.*, vol. 18, no. 4, pp. 439–454, 2002, doi: [https://doi.org/10.1016/S0169-2070\(01\)00110-8](https://doi.org/10.1016/S0169-2070(01)00110-8).
- [73] E. W. & F. Y. Hyndman, R. J. G Athanasopoulos, C Bergmeir, G Caceres, L Chhay, M O’Hara-Wild, F Petropoulos, S Razbash, “forecast: Forecasting functions for time series and linear models,” *R package version 8.12*, 2018. <http://pkg.robjhyndman.com/forecast>.
- [74] A. M. de Livera, R. J. Hyndman, and R. D. Snyder, “Forecasting time series with

- complex seasonal patterns using exponential smoothing,” *J. Am. Stat. Assoc.*, vol. 106, no. 496, pp. 1513–1527, 2011, doi: 10.1198/jasa.2011.tm09771.
- [75] V. Assimakopoulos and K. Nikolopoulos, “The theta model: A decomposition approach to forecasting,” *Int. J. Forecast.*, vol. 16, no. 4, pp. 521–530, 2000, doi: 10.1016/S0169-2070(00)00066-2.
- [76] R. J. Hyndman and B. Billah, “Unmasking the Theta method,” *Int. J. Forecast.*, vol. 19, no. 2, pp. 287–290, 2003, doi: 10.1016/S0169-2070(01)00143-1.
- [77] Y. Hyndman, R. J., & Khandakar, “Forecast, Automatic Time Series Forecasting: The forecast Package for R,” *J. Stat. Softw.*, vol. 27, no. 3, pp. 1–22., 2008, doi: <https://doi.org/10.18637/jss.v027.i03>.
- [78] C. Bergmeir, R. J. Hyndman, and J. M. Benítez, “Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation,” *Int. J. Forecast.*, vol. 32, no. 2, pp. 303–312, 2016, doi: 10.1016/j.ijforecast.2015.07.002.
- [79] R. J. Hyndman, M. L. King, I. Pitrun, and B. Billah, “Local linear forecasts using cubic smoothing splines,” vol. 47, no. 1, pp. 87–99, 2005, doi: <https://doi.org/10.1111/j.1467-842X.2005.00374.x>.
- [80] P. R. Hyndman, P. A. Koehler, P. K. Ord, and A. P. R. Snyder, *Forecasting with Exponential Smoothing*. 2008.
- [81] M. Pawlikowski and A. Chorowska, “Weighted ensemble of statistical models,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 93–97, 2020, doi: 10.1016/j.ijforecast.2019.03.019.
- [82] J. A. Fiorucci, T. R. Pellegrini, F. Louzada, F. Petropoulos, and A. B. Koehler, “Models for optimising the theta method and their relationship to state space models,” *Int. J. Forecast.*, vol. 32, no. 4, pp. 1151–1161, 2016, doi: 10.1016/j.ijforecast.2016.02.005.
- [83] F. Petropoulos and I. Svetunkov, “A simple combination of univariate models,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 110–115, 2020, doi: 10.1016/j.ijforecast.2019.01.006.
- [84] G. Schwarz, “Estimating the Dimension of a Model,” *Ann. Stat.*, vol. 6, no. 2, pp. 461–

464, 1978.

- [85] A. E. Raftery, "Bayesian Model Selection in Social Research," *Sociol. Methodol.*, vol. 25, no. May, p. 111, 1995, doi: 10.2307/271063.