

Univariate and Multivariate Time Series Modeling using a Harmonic Decomposition Methodology

A. Agustín Garibo-Morante and Fernando Ornelas-Tellez

Abstract—This paper contributes by developing an univariate and multivariate harmonic decomposition methodology to model time series. The models are stated in a state space representation derived from a Fourier series analysis to describe an arbitrary signal. The frequency values of the harmonic content are used to define state variables in order to describe a signal through a time-varying linear state space model, which serves to synthesize an optimal state observer (Kalman-Bucy filter). Once the observer converges and the states (harmonics) become constant, the observer model can be used to predict the signal, i.e., a time series forecasting can be performed. The procedure can be developed for the univariate or multivariate case of time series modeling, where in the last one, a statistical analysis is used to determine which variables should be taken into account to obtain a more accurate model. The proposed modeling approach is successfully applied for the modeling and forecast of the time series of electrical power demand and a wind/solar profile.

Index Terms—Signal modeling, Time series, Optimal observer, Kalman-Bucy filter, Forecast, Fourier series.

I. INTRODUCCIÓN

En disciplinas de ingeniería y otras ciencias, es de importancia disponer de modelos matemáticos de la respuesta dinámica de un sistema o fenómeno, mismos que pueden obtenerse a partir de las señales o mediciones disponibles. Un modelo matemático permitiría utilizar metodologías de análisis y diseño bien establecidas [1], [2], y con ello poder tomar decisiones basadas en el conocimiento del sistema, pronosticar eventos futuros, anticipándose a diferentes escenarios, o buscar optimizar recursos y evitar situaciones indeseables. Una teoría bien establecida para el modelado de señales o mediciones, es su tratamiento como series de tiempo (ST), que es un nombre dado en teoría estadística a una variable observada a intervalos de tiempo regulares (diario, semanal, semestral, anual, entre otros) [3]–[5]. Algunos ejemplos de ST son los registros de ventas anuales totales de almacenes, el valor trimestral del producto interno bruto, variables climáticas u otras variables físicas registradas durante un periodo determinado. Algunos métodos populares de modelado matemático de series de tiempo son redes neuronales artificiales (RNA) [6]–[8], modelos autorregresivos integrados de media móvil (ARIMA) [9]–[11], o mediante algoritmos genéticos (AG) [12], entre otros, donde las representaciones logradas con estos modelos generalmente son complejas para su análisis, además de ser modelos no lineales y la mayoría de tipo univariable.

Los autores están adscritos a la Facultad de Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo, Cd. Universitaria, Morelia, 58030, Michoacán, México.

A. Agustín Garibo-Morante (e-mail: angel.garibo@umich.mx)
Fernando Ornelas-Tellez (e-mail: fernando.ornelas@umich.mx)

El procedimiento común para pronosticar ST es utilizar los respectivos modelos que las representan, y evaluándolos a un determinado número de pasos o muestras hacia adelante (instantes futuros). El tiempo en que se pronostique a una señal indicará si se trata de horizontes de tiempo a corto, mediano o largo plazo [11].

En este trabajo se plantea una nueva técnica de modelado de ST [3]–[5], tanto del tipo univariable como multivariable. De forma general, la metodología propuesta se basa en modelar una señal (serie de tiempo) a través de una representación en espacio de estado (EE) de su respectivo contenido armónico, basándose para ello en un análisis de series de Fourier, lo cual puede aplicarse a señales provenientes de diversos fenómenos. Una vez obtenido el modelo de la señal, se diseña un observador óptimo (OO) en EE para actualizar los parámetros del modelo, y así lograr su respectiva convergencia hacia la señal real. Es importante mencionar que el observador óptimo de estados está basado en el filtro de Kalman-Bucy (FKB) [13], [14], que tiene la propiedad de ser robusto ante posible ruido gaussiano presente en la señal a modelar. Con un enfoque semejante en el área de señales eléctricas, inicialmente en [15] se desarrolla un estimador óptimo para determinar en línea los componentes armónicos de una señal eléctrica, con el fin de atenuarlos y mejorar la calidad de la energía en sistemas eléctricos. Por su parte, en [16] se presenta un estudio preliminar a lo expuesto en este trabajo de investigación, abordando únicamente el caso de ST univariable, y bajo la consideración de un modelo de orden reducido para el modelado de señales, lo que limita su precisión. En el presente trabajo, se extiende lo expuesto en [16] para el caso multivariable y se formulan modelos de alto orden para tener una mejor representación de una ST, lográndose resultados con errores menores de aproximación, mismos que adicionalmente se comparan con resultados de otros métodos comúnmente utilizados para modelar y pronosticar ST.

Particularmente, este trabajo se enfoca en el estudio y pronóstico de variables físicas tales como radiación solar, velocidad del viento y demanda de energía eléctrica. Es de relevancia mencionar que la mayoría de los modelos climáticos existentes son complejos, de uso comercial, o que requieren de tecnologías especializadas para medir las variables. Se ha mostrado que el pronóstico de variables climáticas, a partir de un sistema dinámico, es en general un problema complicado y sensible a variaciones en sus condiciones iniciales, fenómeno estudiado por E. Lorentz en la teoría del caos [17], [18], concluyéndose lo difícil de obtener predicciones a corto plazo con un grado razonable de precisión. En consecuencia, la precisión de las predicciones meteorológicas dependerá de los

métodos empleados para pronóstico y de los criterios usados para su medición [5], [19]; también, en muchos métodos de pronóstico, la precisión dependerá del uso de una gran cantidad de datos usados para modelar [3]. En este sentido, la metodología de modelado propuesta en este trabajo es simple de implementar y efectiva para representar señales con comportamiento periódico, como lo son las variables del clima, apoyándose en el hecho de que éstas presentan periodicidad respecto a días, estaciones del año, etc. La metodología usa como base la serie de Fourier para obtener una representación en espacio de estados de una señal, lográndose un modelo lineal variante en el tiempo, que se emplea para el diseño de un observador, donde las variables de estado están asociadas a los parámetros de amplitudes y fases de una serie de Fourier. Estos parámetros son actualizados en línea, disponiendo en consecuencia de un modelo actualizado de la señal en todo momento, lo que permite representar señales no periódicas, con contenido armónico variante en el tiempo, y donde no se requiere una gran cantidad de datos para lograr el modelado y pronóstico.

La principal aportación de este trabajo es: el diseño de un estimador óptimo univariable y multivariable, para modelar señales, ST en lo particular, y poder realizar pronósticos de precisión. El modelo obtenido tiene las siguientes características: es descrito por un sistema lineal simple, que es adaptado en línea, cuyo contenido armónico es determinado directamente en el dominio del tiempo, lo que computacionalmente es eficiente al no requerirse de transformaciones al dominio de la frecuencia, además de que el ruido presente en la señal es filtrado mediante el uso del estimador óptimo. La metodología propuesta se aplica para el pronóstico de ST reales de demanda de energía eléctrica, radiación solar y velocidad del viento. Para el caso multivariable, se realiza un análisis estadístico para determinar las variables que se usarán en la definición del modelo. El pronóstico obtenido con el método propuesto es comparado, cualitativa y cuantitativamente, contra el logrado con modelos clásicos, como RNA y ARIMA, dando resultados con menor error de predicción.

La estructura del resto del artículo es la siguiente. La Sección II describe al OO para el caso univariable. La Sección III extiende la metodología de modelado para el caso multivariable. En la Sección IV se presentan diferentes casos de estudio de modelado y pronóstico. Finalmente, las conclusiones generales se presentan en la Sección V.

II. OBSERVADOR ÓPTIMO DE ARMÓNICOS VÍA FILTRO DE KALMAN-BUCY UNIVARIABLE

El modelado consiste en tener una representación en espacio de estados, a partir de un análisis de series de Fourier, de señales variantes en el tiempo, incluyendo ST, las cuales pueden provenir de un proceso o fenómeno, que pueden presentar cierta periodicidad. Posteriormente, la señales son

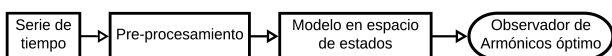


Fig. 1: Proceso de modelado del observador óptimo.

descritas en base a su contenido armónico (amplitudes y fases), mismo que es determinado a través de un estimador óptimo, conocido como FKB. Una vez que el estimador converge, éste se puede usar para realizar el pronóstico, o análisis de la señal.

La Fig. 1 muestra el proceso para modelar una señal arbitraria o serie de tiempo, donde se realiza un pre-procesamiento, que sirve para determinar el contenido armónico más representativo de la señal, lo que definirá las variables de estado y orden del observador. Finalmente, los estados (contenido armónico) son estimados con el OO.

A. Preliminares de Series de Fourier

En este trabajo, las series de Fourier son utilizadas para el modelado del OO, ya que tienen la capacidad de representar señales que presentan periodicidad [20], tal es el caso de las ST, por lo que su consideración en el proceso de modelado es directo. Para los propósitos de este trabajo, una representación conveniente a partir de la serie de Fourier para una señal periódica $s(t)$, con periodo T , es

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^N A_n \cos(w_n t + \theta_n) \quad (1)$$

donde $N \geq 1$ es un número entero (con $N \rightarrow \infty$) que representa el total de armónicos en la señal, siendo A_n la amplitud de la señal cosinusoidal, $w_n = \frac{2\pi n}{T}$ la frecuencia angular y θ_n la fase [21].

Para ejemplificar el proceso de modelado de una señal, considere el caso cuando $a_0 = 0$ y $N = 1$ en (1). Al usar la identidad trigonométrica $\cos(a+b) = \cos a \cos b - \sin a \sin b$, la señal (1) resulta en

$$\begin{aligned} s(t) &= A_1 \cos(w_1 t + \theta_1) \\ &= A_1 \cos \theta_1 \cos w_1 t - A_1 \sin \theta_1 \sin w_1 t. \end{aligned} \quad (2)$$

B. Planteamiento del observador óptimo de armónicos

Para lograr una representación en espacio de estados de (2), se definen las variables de estado como

$$x_1 = A_1 \cos \theta_1 \quad \text{y} \quad x_2 = A_1 \sin \theta_1. \quad (3)$$

Entonces (2) se puede obtener del sistema dinámico

$$\begin{aligned} \dot{x}_1 &= 0 \\ \dot{x}_2 &= 0 \\ y &= C(t)x \end{aligned} \quad (4)$$

donde $x = [x_1 \ x_2]^T \in \mathbf{R}^2$ es el vector de estados, $y \in \mathbf{R}$ es la salida, $C(t) = [\cos w_1 t \ -\sin w_1 t]$ es la matriz de salida (variante en el tiempo), y condiciones iniciales $x_1(0) = A_1 \cos \theta_1$, $x_2(0) = A_1 \sin \theta_1$. Por lo anterior, se obtiene que $s(t) = y$.

Por cuestiones prácticas, considere adicionalmente que la señal $s(t)$ está afectada por ruido Gaussiano, entonces (4) puede escribirse como el siguiente sistema lineal variante en el tiempo

$$\begin{aligned} \dot{x} &= v_1 \\ y &= C(t)x + v_2 \end{aligned} \quad (5)$$

donde $v_1 \in \mathbf{R}^2$ y $v_2 \in \mathbf{R}$ son el proceso y mediciones de ruido, respectivamente, para los cuales se supone lo siguiente:

- **A1.** v_1 y v_2 son ruidos Gaussianos de media cero, con covarianzas $E\{v_1(t)v_1^T(\tau)\} = Q_{v_1}\delta(t-\tau)$ y $E\{v_2(t)v_2^T(\tau)\} = R_{v_2}\delta(t-\tau)$, respectivamente, siendo Q_{v_1} y R_{v_2} matrices no negativas.
- **A2.** El estado inicial $x(t_0)$, los ruidos v_1 y v_2 son procesos no correlacionados.

La suposición **A1** implica que los ruidos no están correlacionados instantáneamente y que la información probabilística del ruido la contienen sus matrices de covarianza. La suposición **A2** se refiere a que los ruidos y la condición inicial son independientes.

De esta manera, para una señal $s(t)$ afectada por los ruidos v_1 y v_2 , donde la amplitud $A_1(t)$ y fase θ_1 son valores desconocidos, se puede diseñar un estimador de estado óptimo (basado en el filtro Kalman-Bucy) para estimar su contenido armónico. El estimador de estado óptimo para (5) resulta como

$$\begin{aligned}\dot{\hat{x}} &= L(s(t) - \hat{y}) \\ \hat{y} &= C(t)\hat{x}\end{aligned}\quad (6)$$

donde \hat{x} es el estimador de estados, \hat{y} es la salida estimada que representa a la señal $s(t)$, y L es la matriz de ganancia del estimador. El estimador es óptimo en el sentido de que se minimiza la matriz de covarianza del error [13], expresada como

$$E\{(x - \hat{x})(x - \hat{x})^T\}. \quad (7)$$

Se hace notar que el modelo del estimador en (6) se deriva directamente del sistema (5), más un término adicional L para garantizar la convergencia del estado del observador al estado real del sistema. En [13] se demuestra que el FKB es uniformemente asintóticamente estable, esto es, \hat{x} converge a x , cuando t tiende al infinito, a partir únicamente de la medición de $s(t)$. Es importante mencionar que dada la convergencia del observador de estados, no es necesario conocer las condiciones iniciales en (5).

De acuerdo con [22], [23], la ganancia óptima L en (6) es calculada como

$$\begin{aligned}L &= P_e C^T(t) R_{v_2}^{-1} \\ \dot{P}_e &= Q_{v_1} - P_e C^T(t) R_{v_2}^{-1} C(t) P_e\end{aligned}\quad (8)$$

con $P_e(\hat{x}(0)) = P_{e_0}$.

Una vez que el estimador ha convergido, los valores de $A_1(t)$ y θ_1 , correspondientes a la descripción armónica de la serie de tiempo, se pueden determinar a partir de (3), realizando los siguientes cálculos:

$$\frac{x_2}{x_1} = \frac{A_1(t) \cos \theta_1}{A_1(t) \sin \theta_1} \implies \theta_1 = \arctan\left(\frac{x_2}{x_1}\right) \quad (9)$$

y

$$A_1(t) = \frac{x_1}{\cos \theta_1} \quad \text{ó} \quad A_1(t) = \frac{x_2}{\sin \theta_1}. \quad (10)$$

Adicionalmente, para el caso en que una señal contiene un componente constante, dado en (1) como $a_0/2$, entonces se puede agregar un estado al sistema (5), definiéndolo como $\dot{x}_0 = 0$, con condición inicial $x_0(0) = a_0/2$.

El procedimiento anterior se puede generalizar para N armónicos contenidos en la señal $s(t)$ en (1). Notar que cada armónico requiere de la definición de dos estados (es decir, para $N = 1$ fueron necesarios declarar x_1 y x_2). Por lo tanto, la generalización de la representación en espacio de estados de $s(t)$ para $N \geq 1$, resulta en

$$\begin{aligned}\dot{x} &= v_1 \\ y &= C(t)x + v_2,\end{aligned}\quad (11)$$

donde $x \in \mathbf{R}^{2N}$ es el vector de estados definido como

$$x = [x_1 \ x_2 \ x_3 \ x_4 \ \cdots \ x_{2N-1} \ x_{2N}]^T$$

con salida $y \in \mathbf{R}$, siendo

$$C(t) = \begin{bmatrix} \cos \omega_1 t & -\sin \omega_1 t & \cos \omega_2 t & -\sin \omega_2 t \\ \cos \omega_N t & -\sin \omega_N t & & \end{bmatrix} \quad (12)$$

mientras que $v_1 \in \mathbf{R}^{2N}$ y $v_2 \in \mathbf{R}$ son los ruidos que afectan a la señal. La salida del sistema y en (11) puede ser vista de forma escalar como

$$y = \sum_{n=1,3,\dots}^{2N-1} (x_n \cos(\omega_n t) + x_{n+1} \sin(\omega_n t)) + v_2.$$

De forma semejante, para un modelo con N armónicos, se puede sintetizar un estimador de estado óptimo para (11), como el descrito por (6), de dimensión apropiada, donde \hat{y} convergerá hacia $s(t)$. La determinación de las amplitudes y fases se pueden generalizar como

$$\begin{aligned}\theta_1 &= \arctan\left(\frac{x_2}{x_1}\right), \quad \theta_2 = \arctan\left(\frac{x_4}{x_3}\right), \\ \dots, \quad \theta_N &= \arctan\left(\frac{x_{2N}}{x_{2N-1}}\right)\end{aligned}\quad (13)$$

y

$$A_1(t) = \frac{x_1}{\cos \theta_1}, \quad A_2(t) = \frac{x_2}{\cos \theta_2}, \dots, \quad A_N(t) = \frac{x_N}{\cos \theta_N}. \quad (14)$$

El proceso de modelado previo puede extenderse para el caso multivariable, al considerar que $s(t)$ proviene de la relación entre varias variables, como se describe en la siguiente sección.

III. OBSERVADOR ÓPTIMO DE ARMÓNICOS VÍA FILTRO DE KALMAN-BUCY MULTIVARIABLE

El OO de armónicos multivariable es una generalización del caso univariable. Por consiguiente, se considera que una señal es producto de la interacción de más de un fenómeno o variable, entonces puede diseñarse un modelo matemático multivariable que incluya varios fenómenos, donde la incorporación de las distintas variables puede realizarse a partir de un análisis estadístico, tomando en cuenta un análisis de correlación.

A. Modelo Multivariable

Sin pérdida de generalidad y por facilidad en la explicación del proceso de modelado, se considerará el diseño a partir de dos variables (dos series de tiempo) $s_1(t)$ y $s_2(t)$, donde cada una tiene la forma descrita por (1), y puede describirse a partir de (11). De esta forma, el modelo para una señal $s(t)$, proveniente de la interacción de dos variables $s_1(t)$ y $s_2(t)$, puede ser representado por

$$\begin{aligned} \dot{x}_{s_1} &= v_1 \\ \dot{x}_{s_2} &= v_2 \\ y &= C_{s_1}(t)x_{s_1} + C_{s_2}(t)x_{s_2} + v_3 \end{aligned} \quad (15)$$

donde

$$x_{s_i} = [x_{1s_i} \quad x_{2s_i} \quad \cdots \quad x_{2N-1s_i} \quad x_{2Ns_i}]^T$$

y con salida

$$C_{s_i}(t) = \begin{bmatrix} (\cos w_1 t)_{s_i} & (-\sin w_1 t)_{s_i} \\ \cdots & (\cos w_N t)_{s_i} & (-\sin w_N t)_{s_i} \end{bmatrix} \quad (16)$$

para $i = 1, 2$. Esto es, la señal $y = s(t)$ es la sumatoria del contenido armónico más representativo de cada variable o fenómeno $s_i(t)$. En este caso multivariable, también se considera que existen ruidos Gaussianos, v_1 , v_2 y v_3 , afectando a la señal. Puede notarse que la estructura del modelo (15) es nuevamente un modelo lineal variante en el tiempo, donde en la expresión de la salida se encuentra la interacción con las diferentes variables del fenómeno a modelar.

B. Estimador de Estado Multivariable

Para el sistema (15), el estimador de estado óptimo resulta en

$$\begin{aligned} \hat{x} &= L(s(t) - \hat{y}) \\ \hat{y} &= C_{s_1}(t)\hat{x}_{s_1} + C_{s_2}(t)\hat{x}_{s_2} \end{aligned} \quad (17)$$

donde $\hat{x} = [\hat{x}_{s_1}, \hat{x}_{s_2}]$, \hat{y} es la salida estimada y L es la matriz de ganancia del estimador. Las componentes armónicas (amplitudes y fases) se estiman a partir de los estados estimados, de manera análoga a (13) y (14).

El procedimiento de modelado puede generalizarse para tomar en cuenta cualquier cantidad de variables que puedan intervenir en un fenómeno, no obstante, el orden del modelo a obtenerse será mayor, lo que implicará un esfuerzo computacional adicional. Por lo anterior, es conveniente realizar un análisis estadístico previo, con el fin de determinar, dentro de un conjunto de posibles variables disponibles, cuales son las que más impactan sobre una variable en particular a modelar.

C. Análisis de Correlación entre Variables

Para el modelado multivariable, en este trabajo se parte del análisis estadístico de correlación de Pearson (γ), que indica las variables que están correlacionadas, y así determinar aquellas que aportan más a la dinámica de una determinada señal, con el fin de estructurar el modelo. Este análisis ha demostrado ser de utilidad para el planteamiento de modelos multivariable más precisos en diferentes trabajos [24]–[26].

Tabla I: Desviación estándar σ de todas las variables, con ST de 12 y 96 horas

Variable	σ (12 horas de muestras)	σ (96 horas de muestras)
Radiación solar	341.9587	328.9427
Temperatura	5.8948	5.3058
Presión Atmosférica	1.2858	1.1432
Velocidad del viento	2.6591	4.4661
Humedad relativa	15.7543	16.0738

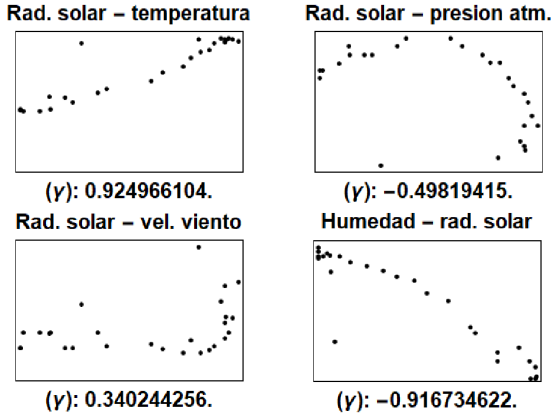
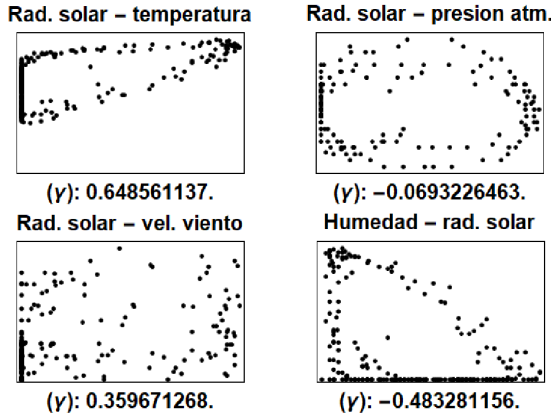
En la Fig. 2 se muestra la dispersión y correlación entre las diferentes variables del clima, indicando su valor γ , para 12 y 96 horas, respectivamente. Por su parte, la Tabla I presenta la desviación estándar σ para las variables climáticas, donde un valor bajo indica que la mayor parte de los datos tienden a estar agrupados cerca de su media, mientras que un valor alto indica que los datos se extienden sobre un rango de valores más amplio. En aplicaciones de pronóstico, una σ alta significaría que la variable a pronosticar presenta un grado alto de dispersión, lo que podría conducir a tener un mayor error de predicción.

Los datos de las variables son tomados de una estación meteorológica ubicada en la ciudad de Morelia, México. De acuerdo a [26], una correlación fuerte se encuentra en valores cercanos a la unidad, sea positiva o negativa, esto es, $\gamma \approx \pm 1$. Se puede observar que la radiación solar y la temperatura tienen una correlación fuerte (ver Fig. 2a), donde para un lapso de 12 horas se tiene $\gamma = 0.9249$ (el lapso de tiempo analizado es durante el día, cuando se tiene un índice de radiación más alto). Por otro lado, para un lapso de 96 horas, se obtiene $\gamma = 0.6485$ (ver Fig. 2b), que también es considerada una correlación alta con respecto a las otras variables. Derivado de este análisis, en la Subsección IV-B se considerarán estas variables (radiación solar y temperatura) para estructurar un modelo de estimador óptimo multivariable de radiación solar. Este modelo es de particular interés en la generación de energía eléctrica, para poder pronosticar la energía que puede estar disponible en un sistema fotovoltaico.

IV. RESULTADOS DE MODELADO Y PRONÓSTICO

En esta sección son modeladas y pronosticadas las ST normalizadas de demanda de energía eléctrica (MWh) y velocidad del viento (km/h), utilizando el modelo del OO univariable. Mientras que para el modelado de radiación solar, se considerará el caso multivariable, tomando en cuenta las ST de radiación solar (W/m^2) y temperatura ($^{\circ}C$). Adicionalmente, el método propuesto se compara, cualitativa y cuantitativamente, contra modelos clásicos de modelado de ST, como son RNA y ARIMA, donde se evalúan diferentes errores usados para medir la capacidad de aproximación [3], [11], [27], [28], entre ellos el error absoluto medio (MAE, por sus siglas del inglés *Mean Absolute Error*), el error porcentual absoluto medio (MAPE, del inglés *Mean Absolute Porcentual Error*), y la raíz cuadrada del error cuadrático medio (RMSE, del inglés *Root Mean Square Error*).

En el proceso de modelado las ST son normalizadas para tomar valores entre 0 y 1. Posteriormente, se realiza un análisis con la transformada rápida de Fourier, con la finalidad de conocer el contenido armónico que más aporta a la ST,

(a) γ en un lapso de 12 horas.(b) γ en un lapso de 96 horas.Fig. 2: γ en diferentes lapsos de tiempo, para la ST de radiación solar contra otras variables del clima.

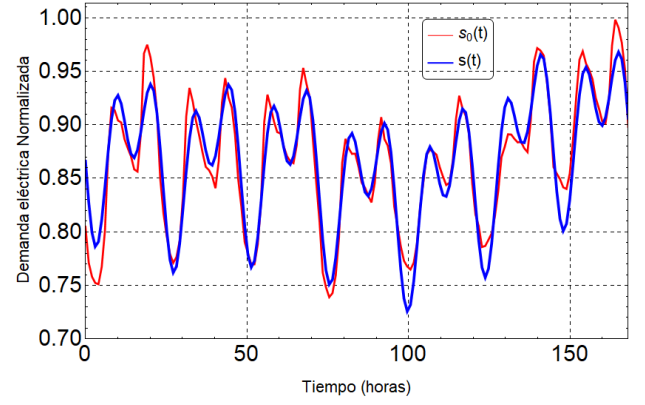
lo que permite tener un modelo en EE de orden reducido para el usarse en el diseño del OO. Se hace notar que este último análisis de Fourier puede omitirse, no obstante, es recomendable realizarlo para tener un modelo de orden bajo y que sea eficiente desde un punto de vista computacional para las simulaciones del modelo obtenido y fines de pronóstico. La implementación de los modelos y obtención de resultados se realizaron utilizando el software Matlab®.

A. Caso Univariable

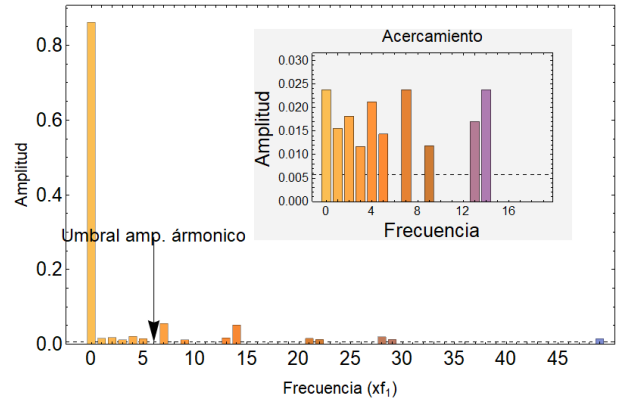
La Fig. 3a muestra la ST real de demanda de energía eléctrica normalizada $s(t)$ (muestreada en intervalos de una hora) y su respectiva aproximación $s_0(t)$, seleccionando $N = 16$ armónicos, siendo los que más aportan a la ST, de acuerdo a un valor de umbral definido por el diseñador, como se muestra en la Fig. 3b. Un valor mayor de umbral produce un modelo de menor orden, esto por considerar menos armónicos, no obstante, la capacidad de aproximación del modelo también se reduce, de forma análoga, un umbral menor genera un modelo que aproxima mejor, pero se incrementará su orden, y en consecuencia tomará un mayor tiempo computacional su ejecución.

Considerando $N = 16$, se estructura el modelo en EE como se propone en (11), con vector de estado \hat{x} dado por

$$\hat{x} = [\hat{x}_1 \quad \hat{x}_2 \quad \hat{x}_3 \quad \cdots \quad \hat{x}_{32}]^T.$$



(a) ST de demanda de energía eléctrica real y su aproximación con 16 armónicos.



(b) Primeros armónicos de la ST

Fig. 3: Selección del contenido armónico para el modelo.

y matriz de salida $C(t)$ que resulta en

$$C(t) = \begin{bmatrix} 1 & 0 & \cos(2\pi t/3) & -\sin(2\pi t/3) \\ \cos(7\pi t/12) & -\sin(7\pi t/12) & \cos(29\pi t/84) & -\sin(29\pi t/84) \\ \cos(\pi t/3) & -\sin(\pi t/3) & \cos(11\pi t/42) & -\sin(11\pi t/42) \\ \cos(\pi t/4) & -\sin(\pi t/4) & \cos(\pi t/6) & -\sin(\pi t/6) \\ \cos(13\pi t/84) & -\sin(13\pi t/84) & \cos(3\pi t/28) & -\sin(3\pi t/28) \\ \cos(\pi t/12) & -\sin(\pi t/12) & \cos(5\pi t/84) & -\sin(5\pi t/84) \\ \cos(\pi t/21) & -\sin(\pi t/21) & \cos(\pi t/28) & -\sin(\pi t/28) \\ \cos(\pi t/42) & -\sin(\pi t/42) & \cos(\pi t/84) & -\sin(\pi t/84) \end{bmatrix}.$$

Una vez definido el modelo de la ST, el OO puede diseñarse de acuerdo a (6).

La Fig. 4 muestra los resultados del modelado de la ST de demanda eléctrica con el OO, usando 168 datos para la convergencia del observador (168 horas), y realizando un pronóstico de 24 horas. La figura también presenta los resultados de los modelos RNA y ARIMA(10,1,2), obtenidos a partir de librerías especializadas disponibles en Matlab® para el pronóstico de ST. Dada la convergencia del observador, se pronostica la ST un paso adelante (una hora), en un proceso iterativo para 24 horas, esto es, cada hora se actualiza la información de la serie para pronosticar la siguiente hora, y así sucesivamente. La Fig. 5 muestra la convergencia de los primeros 8 estados del OO.

La Fig. 6 resume el modelado y pronóstico de la ST de velocidad del viento, con 17 armónicos, y usando 336 datos

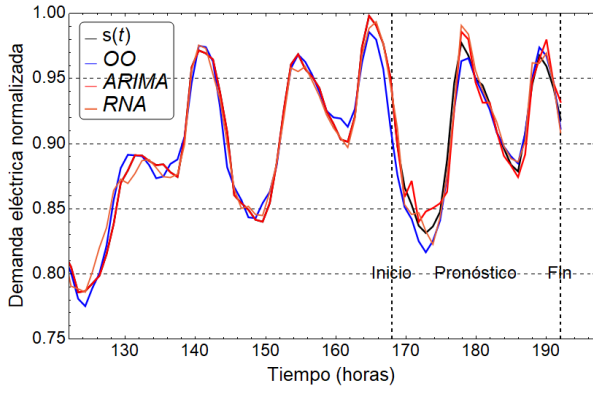
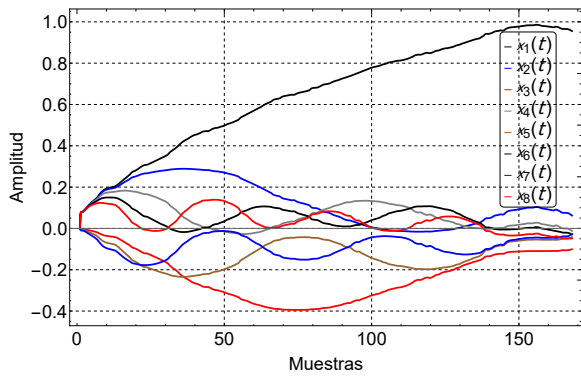


Fig. 4: Modelado y pronóstico de la ST de demanda eléctrica.


 Fig. 5: Convergencia de los primeros 8 estados ($\hat{x} \rightarrow x$).

para la convergencia del observador, tomados en intervalos de 0.5 horas y realizando pronóstico de 24 horas. Se hace notar que debido a que el estimador es actualizado en línea, no se requiere una gran cantidad de datos para modelar y pronosticar. La Tabla II muestra los errores obtenidos para los dos casos de modelado univariable y pronóstico mencionados. Se observa que el OO obtuvo un menor error en todos los casos de estudio, esto en comparación con los otros modelos analizados en cada ST. Es importante mencionar que la ST de velocidad del viento contiene altas frecuencias y variaciones no suaves, que la hacen difícil de modelar, no obstante, el OO converge, y su error de modelado y pronóstico es el menor,

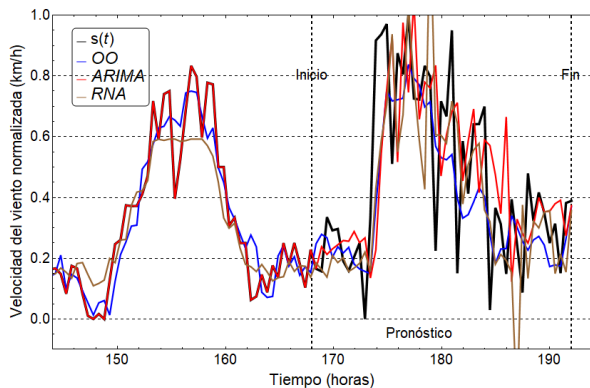


Fig. 6: Modelado y pronóstico de la velocidad del viento.

Tabla II: Error de pronóstico para diferentes ST, con modelos univariable.

Serie de tiempo	Modelo	Error del pronóstico a 24 horas		
		MAE	MAPE	RMSE
Demanda eléctrica	OO	0.007608	0.849195	0.009266
	ARIMA	0.009484	1.050806	0.011914
	RNA	0.008433	0.925215	0.009746
Velocidad del viento	OO	0.110268	30.74564	0.147818
	ARIMA	0.191413	48.37394	0.257226
	RNA	0.189544	51.16091	0.243568

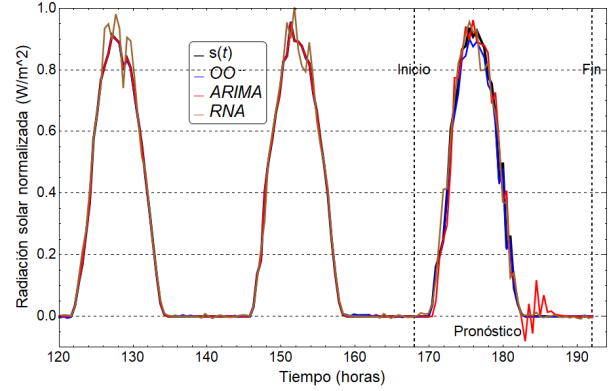


Fig. 7: Modelado multivariable y pronóstico de radiación solar.

en comparación con los modelos RNA y ARIMA.

B. Caso Multivariable

En lo que respecta al caso multivariable, se modela y pronostica una ST de radiación solar normalizada, con datos de la ciudad de Morelia, México. Para este modelo se usan dos ST, radiación solar y temperatura, seleccionadas éstas por ser las que más correlación estadística presentan.

El modelo de radiación solar está construido con 165 armónicos de la misma variable, con estado x_r , y 11 armónicos de temperatura, con estado x_t , para un total de 176 armónicos. Para realizar el modelado y pronóstico del estimador, se usaron 336 datos, tomados en intervalos de 0.5 horas y haciendo pronóstico de 24 horas.

En este caso, el estado completo del observador se conforma como $\hat{x} = [\hat{x}_r, \hat{x}_t]^T$, con

$$\hat{x}_r = [x_1 \quad \hat{x}_2 \quad \hat{x}_3 \quad \cdots \quad \hat{x}_{176}]^T$$

y

$$\hat{x}_t = [x_1 \quad \hat{x}_2 \quad \hat{x}_3 \quad \cdots \quad x_{11}]^T$$

y el OO está construido a partir del modelo dado por (17), donde $s(t)$ son los datos reales de la ST analizada.

La Fig. 7 muestra los resultados del modelo la ST de radiación solar y su comparación respecto a los modelos de RNA y ARIMA. De igual forma, se hace un pronóstico por hora, durante un periodo de 24 horas, actualizado los datos en cada iteración horaria. Esta ST tiene un comportamiento no suave cuando la radiación tiende a cero, lo que conlleva a modelarla con más armónicos para tener una mejor aproximación en todo

Tabla III: Error de pronóstico para radiación solar con un modelo multivariable.

Serie de tiempo	Modelo	Error del pronóstico a 24 horas		
		MAE	MAPE	RMSE
Radiación solar con temperatura	OO	0.010664025	48.9094782	0.016643174
	ARIMA	0.041802302	61.8083496	0.051657827
	RNA	0.030426171	56.4426692	0.053326409

su rango. Los errores de pronóstico obtenidos se muestran en la Tabla III, donde se observa que la metodología propuesta presenta ventajas para predecir la ST con mejor aproximación.

V. CONCLUSIONES

Este trabajo presenta una metodología de modelado de señales en general, con aplicación particular a series de tiempo, donde a partir de series de Fourier, se puede obtener un modelo en espacio de estados que represente su contenido armónico. A partir del modelo obtenido, se desarrolla un estimador óptimo de estados, que resulta en un sistema lineal variante en el tiempo, donde con el conocimiento de la medición de la señal, es posible estimar el contenido armónico, y en consecuencia, disponer de un modelo matemático que la describe. Una vez que el observador ha convergido, el contenido armónico permanece constante, lo que puede utilizarse para fines de pronóstico. El método de modelado puede aplicarse para el caso univariable, y extenderse al multivariable, a través de una análisis de correlación estadística. Los resultados obtenidos muestran que la metodología propuesta produce errores menores de aproximación en el pronóstico, en relación con otros métodos. Se presenta la aplicación de la técnica propuesta para el pronóstico de series de tiempo de demanda eléctrica, velocidad de viento y radiación solar.

REFERENCIAS

- [1] D. E. Kirk, *Optimal Control Theory: An Introduction*, 1st ed. Mineola, New York: Dover Publications, Inc., 2004.
- [2] A. Isidori, *Nonlinear control systems*, 1st ed. Springer-Verlag, London: Springer Science and Business Media, 2013.
- [3] M. Falk and other, *A First Course on Time Series Analysis-Examples whit SAS*. Wurzburg, Alemania: University of Wurzburg, 2012.
- [4] Y. Chen, B. Yang, J. Dong, and A. Abraham, "Time-series forecasting using flexible neural tree model," *Information Sciences*, vol. 174, pp. 219–235, 08 2005.
- [5] B. Box, G. Jenkins, G. Reinsel, and G. Ljung, *Time Series Analysis: Forecasting and Control*, 01 2016, vol. 68.
- [6] B. M. del Brio and A. S. Molina, *Redes neuronales y sistemas borrosos*, 3rd ed. Madrid, España: ALFAOMEGA GPO ED, 2015.
- [7] E. N. Sanchez, A. Y. Alanis, and J. Rico, "Electric load demand prediction using neural network trained by Kalman filtering," in *2004 IEEE International Joint Conference on Neural Networks*, vol. 4, Budapest, Hungary, July 2004, pp. 2771–2775.
- [8] R. L. Farias, J. J. Flores, and V. Puig, "Qualitative and quantitative multi-model forecasting with nonlinear noise filter applied to water demand," in *2015 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, Ixtapa, Mexico, Nov 2015, pp. 1–6.
- [9] C. Chatfield and D. Prothero, "Box-jenkins seasonal forecasting: problems in a case-study," *Journal of the Royal Statistical Society: Series A (General)*, vol. 136, no. 3, pp. 295–315, 1973.
- [10] O. D. Anderson, "The box-jenkins approach to time series analysis," *RAIRO - Operations Research*, vol. 11, no. 1, pp. 3–29, 1977.
- [11] K. H. A. McLeod, *Time Series Modelling of Water Resources and Environmental Systems, Volume 45*. P.O. Box 211, 1000, AE Amsterdam, The Netherlands: Elsevier Science, 1994.

- [12] E. de la Vega, J. J. Flores, and M. Graff, "k-nearest-neighbor by differential evolution for time series forecasting," in *Nature-Inspired Computation and Machine Learning*, A. Gelbukh, F. C. Espinoza, and S. N. Galicia-Haro, Eds., vol. 8857. Switzerland: Springer, 2014, pp. 50–60.
- [13] R. Kalman and R. Bucy, in *Journal of Basic Engineering*, vol. 83, no. 1, March 1961, pp. 95–108.
- [14] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.
- [15] F. Ornelas-Tellez, J. J. Rico-Melgoza, R. Morfin-Magaña, and S. Ramos-Paz, "Optimal dynamic harmonic extraction and suppression in power conditioning applications," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 9, pp. 7909–7918, 2020.
- [16] A. A. Garibo-Morante and F. Ornelas-Tellez, "Modeling and forecast of time series by using a harmonic decomposition approach," in *2019 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, Ixtapa, Mexico, 2019, pp. 1–6.
- [17] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [18] G. Layek, *An introduction to dynamical systems and chaos*. Springer, India, 2015.
- [19] J. C. Kniefel, D. A. Ahijevych, and K. W. Manning, "Using temporal modes of rainfall to evaluate the performance of a numerical weather prediction model," *Monthly weather review*, vol. 132, no. 12, pp. 2995–3009, 2004.
- [20] A. Zygmund, *Trigonometric series*, 3rd ed. New York, NY, USA: Cambridge university press, 2002, vol. 1.
- [21] G. P. Tolstov, *Fourier Series*, 1st ed. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1962.
- [22] B. Friedland, *Control System Design: An Introduction to State-Space Methods*, kindle ed. New York, NY, USA: McGraw-Hill, 1986.
- [23] R. F. Stengel, *Optimal Control and Estimation*, 2nd ed. New York, NY, USA: Robert F. Stengel, 1994.
- [24] J. De Leeuw, "Models and methods for the analysis of correlation coefficients," *Journal of Econometrics*, vol. 22, no. 1, pp. 113–137, 1983.
- [25] A. G. Asuero, A. Sayago, and A. G. González, "The correlation coefficient: An overview," *Critical Reviews in Analytical Chemistry*, vol. 36, no. 1, pp. 41–59, 2006.
- [26] P. Cohen, S. G. West, and L. S. Aiken, *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd ed. Mahwah, New Jersey, London: Lawrence Erlbaum Associates, 2014.
- [27] J. A. Mauricio, *Introducción al Análisis de Series Temporales*, 1st ed., Universidad Complutense de Madrid, España, 2013.
- [28] J. Tayman and D. A. Swanson, "On the validity of mape as a measure of population forecast accuracy," *Population Research and Policy Review*, vol. 18, no. 4, pp. 299–322, 1999.

A. Agustin Garibo-Morante received the B. Sc. and M.Sc. degree in electrical engineering from the Universidad Michoacana de San Nicolas de Hidalgo (UMSNH), Morelia, Mexico in 2017 and 2020, respectively. His research interests are nonlinear control, applied control, signal processing and modeling, forecasting, robotics, microcontrollers, artificial neural networks, diffusion logic and IoT.



Fernando Ornelas-Tellez received the B.Sc. degree from the Instituto Tecnológico de Morelia (ITM), Mexico, in 2005 and the M.Sc. and D.Sc. degrees in electrical engineering from the Advanced Studies and Research Center, National Polytechnic Institute (CINVESTAV-IPN), Guadalajara, Mexico, in 2008 and 2011, respectively. His research interest are dynamical system modeling, neural control, optimal control, electrical machines, and power electronics.

