

DOI: 10.24850/j-tyca-2022-04-09

Notas

Predicción de la máxima demanda en un sector de la red de distribución de agua potable de la ciudad de Guanajuato, México

Prediction of the maximum demand in a DMA of the water distribution network of Guanajuato City, Mexico

Laurent Reyes¹, ORCID: <https://orcid.org/0000-0003-1419-1680>

Lizeth Torres², ORCID: <https://orcid.org/0000-0002-4937-4586>

Martín Jiménez³, ORCID: <https://orcid.org/0000-0003-1034-5996>

Carlos Ríos⁴, ORCID: <https://orcid.org/0000-0001-9555-3289>

¹Tecnológico Nacional de México, Campus Tuxtla Gutiérrez, Tuxtla Gutiérrez, Chiapas, México, laurentreyesc@gmail.com

²Instituto de Ingeniería, Universidad Nacional Autónoma de México, Ciudad de México, México, ftorreso@iingen.unam.mx

³Facultad de Estudios Superiores Aragón, Universidad Nacional Autónoma de México, Nezahualcóyotl, Estado de México, México, mr.jimenez@comunidad.unam.mx



⁴Tecnológico Nacional de México, Campus Tuxtla Gutiérrez, Tuxtla Gutiérrez, Chiapas, México, crios@ittg.edu.mx

Autora para correspondencia: Lizeth Torres, ftorreso@iingen.unam.mx

Resumen

En este artículo se presenta un modelo para simular y predecir la máxima demanda semanal de un sector de la red de distribución de agua de la ciudad de Guanajuato. El modelo está compuesto de dos submodelos, uno determinístico y uno estocástico, los cuales se calibran con datos históricos de la demanda. El submodelo determinístico está descrito por la serie de Fourier, mientras que el submodelo estocástico involucra la función cuantil de la distribución log-logística. Con el fin de probar la capacidad del modelo para describir el comportamiento de la demanda, se utilizaron datos para calibrar y validar el modelo, los cuales fueron proporcionados por un sistema holístico de telemetría. Dicho sistema consta de una red de estaciones para la medición y registro de datos, así como de un sistema de almacenamiento y procesamiento de la información.

Palabras clave: demanda de agua, modelado, redes de distribución de agua, series de Fourier, distribuciones probabilísticas.

Abstract



This article presents a model to simulate and predict the maximum weekly water demand of a DMA belonging to the distribution network of Guanajuato city. The model is made up of two sub-models, one deterministic and one stochastic, which can be calibrated with historical demand data. The Fourier series describes the deterministic sub-model, while the stochastic sub-model involves the quantile function of the log-logistic distribution. Data from a holistic telemetry system were used to calibrate and validate the model to test the model's goodness in describing the demand behavior. Such a system consists of a network of stations for the measurement and recording of data and a system for storing and processing the information.

Keywords: Water demand, modeling, water distribution networks, Fourier series, probability distributions.

Recibido: 15/06/2020

Aceptado: 12/05/2021

Introducción



Debido a la naturaleza pseudoaleatoria de la demanda de agua en una red de distribución se ha propuesto en diversas publicaciones científicas modelar la demanda de agua considerando exclusivamente un enfoque probabilístico. Sin embargo, es bien sabido que existen patrones, concretamente periodicidades diarias, semanales y estacionales, que están ligadas con las actividades rutinarias de los consumidores (Zhou, McMahon, Walton, & Lewis, 2002; Alvisi, Franchini, & Marinelli, 2007). Por esta razón, el modelado de la demanda de agua también se ha formulado utilizando ecuaciones deterministas, como la serie de Fourier, los polinomios Chebyshev y regresiones (Brentan, Luvizotto Jr., Izquierdo, & Pérez-García, 2016).

El objetivo de este artículo es presentar un modelo para simular y predecir la máxima demanda semanal en un sector de la red de distribución de agua de la ciudad de Guanajuato utilizando ambos enfoques: el aleatorio y el determinista, y proponer un modelo híbrido compuesto de dos módulos: un submodelo estocástico y uno periódico. El primero expresado por la función cuantil de una distribución de probabilidad y el segundo por la serie de Fourier. La unificación de enfoques no es una idea original, de hecho, fue en los años de 1980 que Maidment y Parzen (1984) propusieron por primera vez un modelo híbrido, al que se llamó *modelo en cascada*, y el que se puede enunciar de la siguiente manera:

$$D(t) = D_l(t) + D_s(t)$$



donde $D(t)$ representa la demanda mensual promedio, la cual se expresa como la suma de una componente de memoria larga o determinista, denotada por $D_l(t)$, más una componente de memoria corta o estocástica denotada por $D_s(t)$.

Una serie de tiempo de memoria larga es aquella que puede contener cualquiera de las siguientes características: tendencia, periodicidad, correlación con decaimiento lento, ciclos de muy baja frecuencia, o densidades espectrales con un rango dinámico muy grande (Box, Jenkins, Reinsel, & Ljung, 2015).

Las series temporales de memoria corta tienen un comportamiento aleatorio debido a la naturaleza incierta de agentes exógenos como la temperatura, el clima y fenómenos socioeconómicos, entre otros, cuyo origen no tiene que ver con la cotidianeidad de los usuarios. Estas series deben modelarse con herramientas estadísticas.

Se han propuesto varias distribuciones de probabilidad para describir la componente de memoria corta o estocástica, por ejemplo, la distribución de Poisson, que se ha utilizado para modelar la demanda de sectores poco habitados o la demanda residencial (Tzatchkov, Alcocer-Yamanaka, Arreguín-Cortés, & Feliciano-García, 2003; Tzatchkov & Alcocer-Yamanaka, 2016). En general, los modelos que se basan en distribuciones de probabilidad discretas, como la de Poisson, describen la demanda como la superposición de varios pulsos rectangulares que se

generan al activar los dispositivos hidrosanitarios (grifos, regaderas, excusados) de las viviendas (Alcocer-Yamanaka, Tzatchkov, & Arreguín-Cortés, 2012). Sin embargo, conforme el número de usuarios aumenta, la tendencia de la demanda de agua durante el día se suaviza y los pulsos individuales ya no son distinguibles (Gargano, Tricarico, Granata, Santopietro, & De-Marinis, 2017). En estas condiciones, la demanda de agua puede representarse mediante una única variable aleatoria que es continua y positiva.

De-Marinis, Gargano y Tricarico (2006) demostraron mediante un análisis estadístico que, para una red de agua con 1 200 usuarios, las distribuciones de Gumbel y log-normal representan mejor la demanda máxima de agua. Demostraron también que la media de la demanda tiene una distribución normal y que la demanda nocturna está bien caracterizada con el modelo de Poisson. Otros autores han utilizado la distribución normal para modelar la componente de memoria corta sin importar la hora del día. Sin embargo, hay algunos estudios que han demostrado que los datos no se distribuyen normalmente, sino que tienen una distribución asimétrica positiva de cola larga (Surendran & Tota-Maharaj, 2015; Surendran & Tota-Maharaj, 2018). Por esta razón, Tanyimboh, Surendran, Naga, Topping y Soares (2004) propusieron modelar la componente aleatoria con la distribución log-logística que se ha utilizado en hidrología para modelar flujo de corrientes y precipitaciones; en economía, como un modelo simple de distribución de

riqueza o ingreso; y en redes de comunicación, para modelar los tiempos de transmisión de datos considerando tanto la red como el *software*.

Es importante mencionar que la distribución log-logística se parece en forma a la distribución log-normal, la cual es flexible y conveniente para modelar. Además, es una de las pocas distribuciones con una función de probabilidad de densidad, una función de distribución acumulada, y una función cuantil en forma cerrada y simple (Al-Shomrani, Shawky, Arif, & Aslam, 2016). Este hecho es una gran ventaja, ya que una función cuantil simple se puede usar para "construir" una variable aleatoria con una determinada distribución sin necesidad de métodos numéricos, por lo que puede servir, con la ayuda de un generador de números aleatorios, para simular, modelar y predecir el "muestreo" de una distribución arbitraria.

Así pues, en este artículo se pretende sacar ventaja de la función cuantil de la distribución log-logística proponiéndola para el modelado de la componente aleatoria de la demanda semanal máxima que, junto con la serie de Fourier, conforman un modelo integral híbrido. Para probar la capacidad del modelo se utilizan las mediciones de gasto registradas durante 12 semanas de un sector de la red de agua de la ciudad de Guanajuato, que está dotado de cuatro estaciones de telemetría distribuidas en el territorio de manera estratégica. Estas estaciones forman parte de un sistema holístico de monitoreo y diagnóstico de la red de agua.

De manera general, la ciudad de Guanajuato se abastece de dos tipos de fuente: una superficial, agua proveniente de presas; y otra subterránea, de pozos al sur de la ciudad; de los pozos se extrae 60 %, mientras que las presas suministran 40 % del agua que se consume en la ciudad. Bajo situaciones especiales es deseable estimar con cierta anticipación cómo será el comportamiento del consumo en la ciudad, esto permitiría establecer las políticas de operación de presas de la manera más adecuada y determinar los volúmenes máximos permisibles de extracción.

El artículo está organizado de la siguiente manera. La sección "Descripción del caso de estudio" detalla las características de la red de distribución de Guanajuato, en particular, del sector *Los Alcaldes*, así como la arquitectura del sistema holístico de monitoreo y diagnóstico de la red de distribución. En la sección "Descripción del modelo" se presentan las ecuaciones propuestas para describir la demanda de agua en un sector. En la sección "Resultados" se describe el proceso de calibración del modelo con los datos provenientes de *Los Alcaldes*, así como una evaluación de la predicción resultante del modelo calibrado.

Descripción del caso de estudio



Descripción de la red de agua de la ciudad Guanajuato

El 40 % del agua demandada por los habitantes de la ciudad de Guanajuato proviene de las plantas de tratamiento *Los Filtros* e *Ing. Salvador Yañez Castro*. La primera planta es abastecida por las presas *La Soledad* y *La Esperanza*, mientras que la segunda lo es por la *Presa de Mata*; 60 % del agua potable restante proviene de pozos. La red de agua de la ciudad (RDA) de Guanajuato tiene una longitud de aproximadamente 655.382 km y tuberías con diámetros que van de 4 a 16 pulgadas. Más del 72 % es de PVC. La red se divide en 26 macrosectores, de los cuales uno se llama *3 Estrellas*, el cual incluye al sector *Los Alcaldes*: área designada para la implementación de un sistema holístico de telemetría. El sector *Los Alcaldes* abarca un vecindario de clase media, un hospital público que tiene una cisterna de 35 m³ y una estación de autobuses. En *Los Alcaldes* se entrega agua a más de mil clientes fijos, pero hay clientes flotantes (viajeros y pacientes) que cambian la demanda. El trazo del sector hecho en EPANET se muestra en la Figura 1. *Los Alcaldes* es abastecido por un tanque superficial con un volumen de 1 000 m³ llamado *3 Estrellas*, que a su vez se abastece de agua enviada por bombeo desde la estación *Cárcamo Yerbabuena* que recibe agua de los 20 pozos de la Batería del Sistema Puenteillas

localizada al sur de la ciudad. En Cano-Rodríguez *et al.* (2000), y Cano-Rodríguez, Gómez-Vallejo, Aguilera-Alvarado, De-la-Rosa y Gardea-Torresdey (2004) se puede encontrar información adicional sobre algunos pozos de la batería, como ubicación, nivel estático y gasto.



Figura 1. Trazo del sector *Los Alcaldes*.

Descripción del sistema holístico de telemetría *Los Alcaldes*

La estación localizada en la entrada del sector *Los Alcaldes* es un espacio subterráneo rectangular con paredes de concreto a la que se accede mediante una escalera marina. La línea principal (una tubería de 3"), que entrega agua al sector pasa a través de esta estación, donde además se localiza una válvula reductora de presión (VRP) instalada en la línea principal. En la estación también se encuentran instalados dos transductores de presión: uno aguas arriba y otro aguas abajo de la VRP. Adicionalmente, un bypass está montado en la línea principal donde se colocaron sensores de pH, temperatura y cloro. La principal ventaja del bypass es que los sensores se pueden aislar fácilmente para mantenimiento sin interrumpir el suministro de agua. Un medidor de gasto está instalado aguas abajo de la VRP para medir el consumo de los usuarios del sector.

La estación fue diseñada para ser autónoma, cuenta con un sistema de suministro de energía de origen fotovoltaico con una capacidad de almacenamiento para operar hasta 30 días sin sol (Figura 2).





Figura 2. Estación de registro y envío de datos.

La estación está equipada con dos dispositivos de telemetría que envían las mediciones tomadas por los sensores hacia una plataforma web. Un dispositivo envía las mediciones proporcionadas por los transductores de presión y el otro envía las mediciones proporcionadas por el medidor de gasto, pH, temperatura y cloro. Ambos dispositivos utilizan diferentes protocolos de comunicación para trabajar con los

diversos sensores que están conectados a ellos, incluidos los protocolos analógicos y digitales 4-20 mA, HART, RS-232/485, UART, SPI, I2C. Además, los dispositivos tienen un módulo de comunicación de radiofrecuencia que es compatible con las redes 2G y 3G de cualquier operador de telefonía celular global, lo que permite la comunicación entre los dispositivos y la plataforma web, donde se reciben, muestran y almacenan los datos para su análisis.

La plataforma web, que administra el análisis y el almacenamiento de datos, utiliza los servicios de computación en la nube de Microsoft Azure®, que incluyen herramientas como IoT Hub y Stream Analytics, para manejar los dispositivos de telemetría y las bases de datos de SQL Server, para almacenar la información. La plataforma web fue diseñada para ser escalable y robusta, así como para recibir y procesar grandes cantidades de datos. Además, se concibió para permitir el uso de algoritmos complejos que son esenciales para un sistema de abastecimiento inteligente.

El envío de datos desde ambos dispositivos se realiza cada 10 minutos. Los datos almacenados se pueden descargar en los siguientes archivos de datos .csv, .pdf, .xls; están etiquetados con los nombres de las variables, y tienen marcas de fecha y hora asociadas que indican el envío de las mediciones.

Descripción del modelo

El modelo está compuesto de dos submodelos: uno determinístico y uno estocástico, los cuales sirven para representar, respectivamente, la componente periódica y la componente aleatoria de la demanda de agua.

El modelo se puede escribir entonces de la siguiente manera:

$$D(t) = \zeta(t) + \lambda\eta(h(t))$$

donde $D(t)$ denota la demanda; $\zeta(t)$ representa el submodelo de la componente determinística; $\eta(h(t))$ es la función cuantil de la distribución log-logística; $h(t)$ es un variable aleatoria distribuida uniformemente en el intervalo $(0,1)$, y $\lambda \in (0,1)$ expresa un peso de intensidad aleatoria.

El submodelo de la componente determinística se puede expresar como una serie de Fourier, como:

$$\zeta(t) = a_0 + \sum_{k=1}^N (a_k \cos(k\omega t) + b_k \sin(k\omega t))$$

donde a_0 , a_k y b_k son los coeficientes de Fourier, que pueden estimarse utilizando información histórica y algún método de estimación; ω denota la frecuencia fundamental de la demanda.

En esta contribución se propone que la función elemental del submodelo estocástico sea la función cuantil de la distribución log-logística:

$$\eta(h(t)) = \alpha \left(\frac{h(t)}{1-h(t)} \right)^{1/\beta}$$

El parámetro $\alpha > 0$ es el de escala y la mediana de la distribución. Cuanto más grande α , más amplia la distribución. El parámetro $\beta > 0$ es el parámetro de forma. La distribución es unimodal cuando $\beta > 1$ y su dispersión disminuye a medida que aumenta β . Ambos parámetros deben estimarse utilizando datos reales y un método de estimación como el método de los momentos.

La distribución log-logística (también conocida como distribución de Fisk) es una distribución probabilística continua para variables aleatorias no negativas y se expresa de la siguiente manera:

$$f(x) = \frac{\left(\beta/\alpha\right) (x/\alpha)^{\beta-1}}{\left(1 + (x/\alpha)^\beta\right)^2}$$

donde x representa la componente aleatoria de los valores de la demanda.

Resultados

Procesamiento de los datos

Para calibrar el modelo se descargaron de la plataforma web los registros de gasto adquiridos en la estación de la entrada del sector *Los Alcaldes*. En concreto, se descargó el historial de tres meses: julio, agosto y septiembre de 2019. Los datos se segmentaron en semanas, así que se obtuvieron 12 conjuntos de datos.

Con el fin de verificar que la función de probabilidad log-logística representa convenientemente la distribución de la demanda semanal se obtuvieron 12 histogramas, cada uno correspondiente a cada una de las semanas. Estos histogramas se muestran en la Figura 3. Posteriormente, se ajustó la función log-logística a cada una de las distribuciones

semanales. Con el fin de lograrlo, se estimaron los parámetros α y β involucrados en la función de probabilidad a partir de los datos de cada semana.

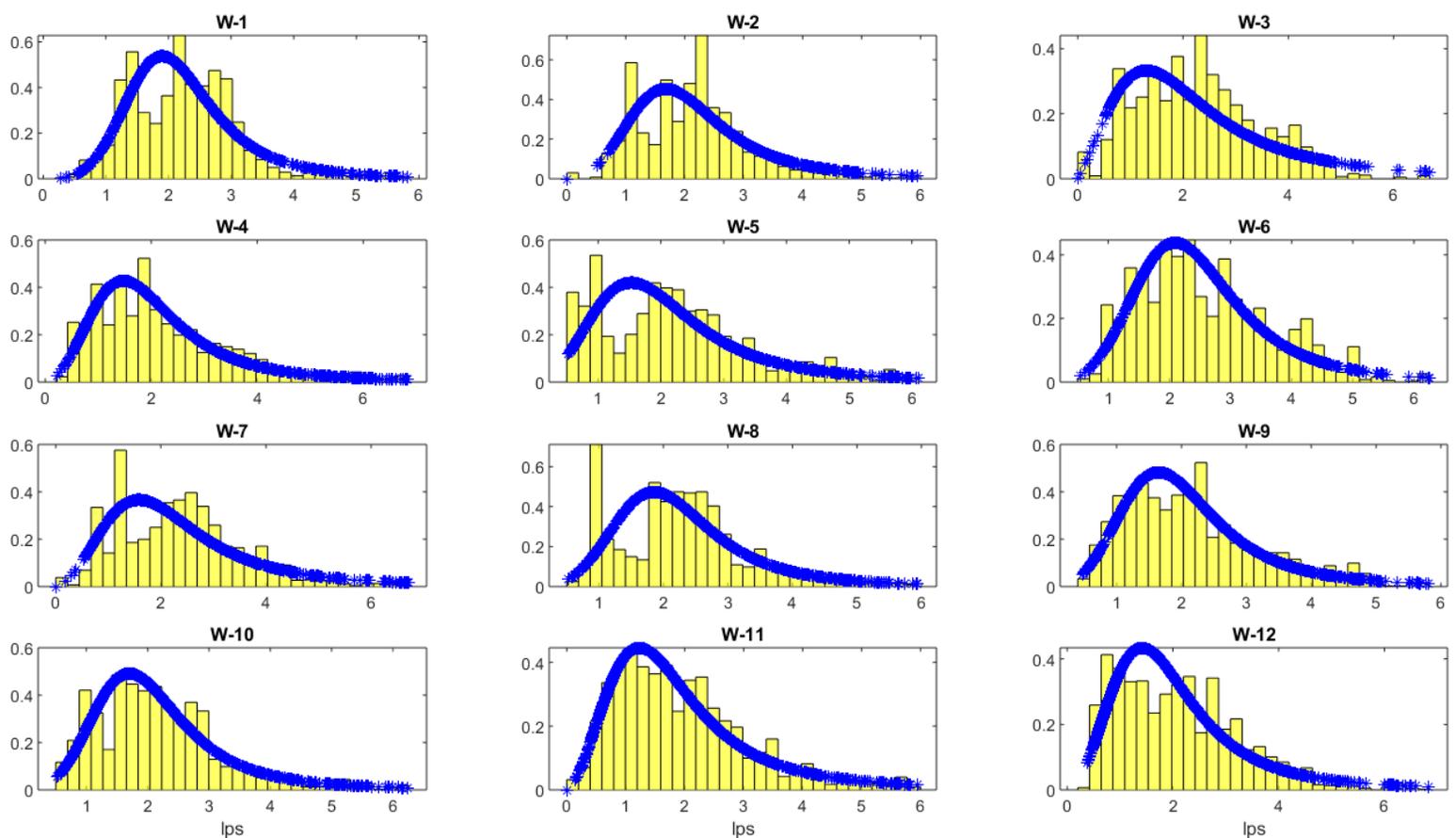


Figura 3. Histogramas de la demanda de agua semanal y funciones de probabilidad ajustadas.

Los parámetros se estimaron utilizando el método de los momentos y se validaron, i.e. se volvieron a estimar, utilizando máxima verisimilitud. La Tabla 1 resume el valor de estos parámetros para cada una de las semanas. La última columna muestra la media de estos parámetros. La desviación estándar para α fue $\sigma_{\alpha} = 0.0765$, que representa aproximadamente 10 % de la media calculada para las 12 semanas; para β fue $\sigma_{\beta} = 0.0435$, que representa un 5 % de la media. Esto significa que durante estas 12 semanas hubo poca variación de los parámetros y, por ende, escasa variación en la distribución.

Tabla 1. Parámetros de la función log-logística y de la serie de Fourier para cada una de las distribuciones semanales.

	W-1	W-2	W-3	W-4	W-5	W-6	W-7	W-8	W-9	W-10	W-11	W-12	Media
α	0.7505	0.7016	0.7137	0.6418	0.6699	0.8705	0.7492	0.7600	0.6691	0.6870	0.5238	0.6118	0.7256
β	0.2320	0.2996	0.4572	0.3483	0.3440	0.2553	0.3720	0.2657	0.2888	0.2768	0.3875	0.3554	0.2770
a_0	2.2230	2.1780	2.3180	2.2110	2.2030	2.5810	2.3560	2.5810	2.1640	2.1720	1.9530	2.1030	2.2381
a_1	-0.6145	-0.7154	-0.9320	-1.0180	-0.7012	-0.5944	-0.8994	-0.5944	-0.7205	-0.6756	0.0259	-0.1395	-0.5928
b_1	-0.3325	-0.4412	-0.5065	-0.5473	-0.6265	-0.2803	-0.7271	-0.2803	-0.4659	-0.5080	-0.6712	-0.6144	-0.5455
a_2	0.2998	0.2475	0.0772	0.2477	0.4712	0.4005	0.2526	0.4005	0.2552	0.4468	0.3200	0.5040	0.3479
b_2	-0.3896	-0.2220	-0.2745	-0.3870	-0.5087	-0.5200	-0.0375	-0.5200	-0.3524	-0.2930	-0.0699	-0.5040	-0.2257
a_3	0.0948	0.1584	0.0934	0.0697	0.2406	0.3392	0.2968	0.3392	0.0956	0.1296	-0.1357	0.3431	0.0766
b_3	0.1043	-0.0317	0.0596	0.3421	-0.0157	-0.0710	0.0781	-0.0710	-0.1261	-0.2666	-0.2916	-0.1180	-0.0291
ω	0.0044	0.0044	0.0045	0.0044	0.0044	0.0043	0.0044	0.0043	0.0043	0.0044	0.0044	0.0045	0.0044

En la última fila de la Tabla 1 se muestra la frecuencia angular para cada semana, la cual, para todos los casos, es $\omega = 0.0044$ (rad/min). Esto corresponde a un periodo de $T = 1\ 440$ (min) a un periodo de 24 horas.

Con el fin de verificar que la serie de Fourier representa de modo adecuado la componente determinística de la demanda semanal, se estimaron los coeficientes de la serie de Fourier que aproxima la serie temporal de cada semana. La estimación se llevó a cabo utilizando el observador espectral propuesto por Torres, Jiménez-Cabas, Gómez-Aguilar y Pérez-Alcázar (2018), que tiene la ventaja de estimar los coeficientes en cada instante de tiempo, de tal suerte que este algoritmo puede utilizarse para la predicción de la demanda de agua en tiempo real.

En la Tabla 1 se listan los coeficientes de Fourier que aproximan la componente periódica de cada semana. Nótese que se calcularon los coeficientes correspondientes al primer, segundo y tercer armónico, ya que a partir del cuarto armónico las amplitudes de las componentes frecuenciales son casi cero. Véase también que la última columna muestra la media de cada coeficiente calculada para las 12 semanas.

Calibración del modelo



Cuando se habla de la calibración de un modelo, se habla de la estimación de los parámetros que lo caracterizan utilizando mediciones (muestras) de variables recolectadas en campo, las cuales contienen información del comportamiento del sistema, proceso o fenómeno que se está modelando. Entre más información significativa se utilice para la calibración se obtendrá un modelo que describirá con mayor precisión la realidad. Sin embargo, esto no siempre es posible, por lo que hay que lidiar con la incertidumbre.

Una primera calibración del modelo propuesto en esta contribución se puede hacer con los datos de una semana. Si se utilizan los datos medidos de la semana 1, entonces los parámetros listados en la primera columna de la Tabla 1 son los parámetros del modelo, el cual se expresaría de la siguiente forma:

$$D(t) = 2.223 - 0.6145 \cos(0.0044t) - 0.3325 \sin(0.0044t) + 0.2998 \cos(0.0088t) - 0.3896 \sin(0.0088t) + 0.0948 \cos(0.0132t) + 0.1043 \sin(0.0132t) + \lambda \left(0.7505 \left(\frac{h(t)}{1-h(t)} \right)^{1/0.232} \right)$$

El único parámetro por definir en el modelo es λ . Este parámetro permite regular el grado de aleatoriedad que se le quiere otorgar al modelo y se puede elegir con base en cuánta información se utilizó para la calibración. Dado que los datos de una semana son muy pocos para modelar la máxima demanda semanal de un sector, se puede elegir $\lambda = 1$.

La Figura 4 muestra la respuesta del modelo (en gris claro) y de la demanda medida durante las 12 semanas (en diversos colores). En esta figura, aparentemente, se ve que los registros medidos de la demanda están por debajo de la respuesta del modelo, que es lo que se desea. Sin embargo, para verificar esta hipótesis, se calculó el error relativo porcentual entre el modelo y los datos de la semana 5 en los instantes en que los datos sobrepasaron la respuesta del modelo. La gráfica del error se muestra en la Figura 5, donde se puede apreciar que durante varios instantes de tiempo los datos de campo sobrepasaron el modelo de la demanda. Lo mismo sucedió con los datos de las demás semanas.

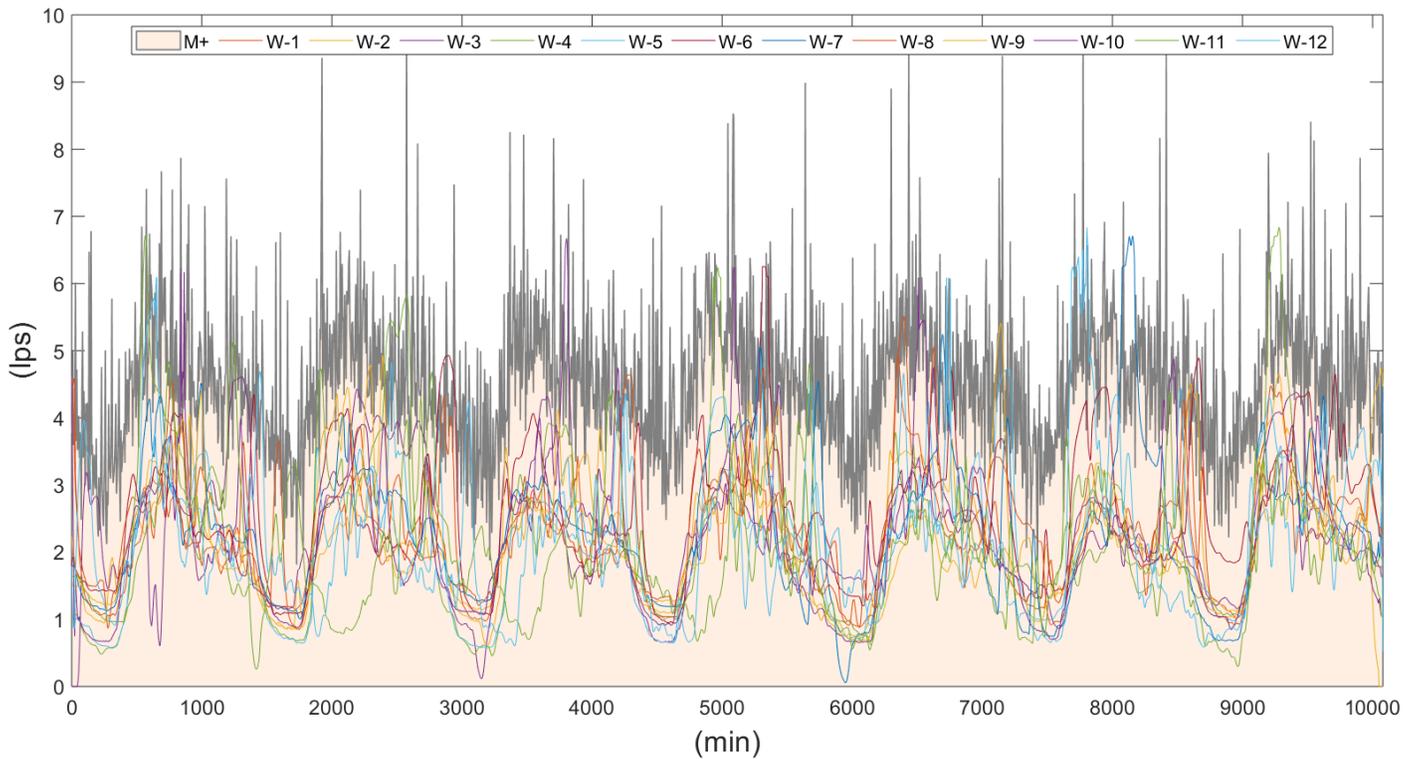


Figura 4. Modelo calibrado con los datos de la semana 1 vs. datos recolectados en campo durante 12 semanas.

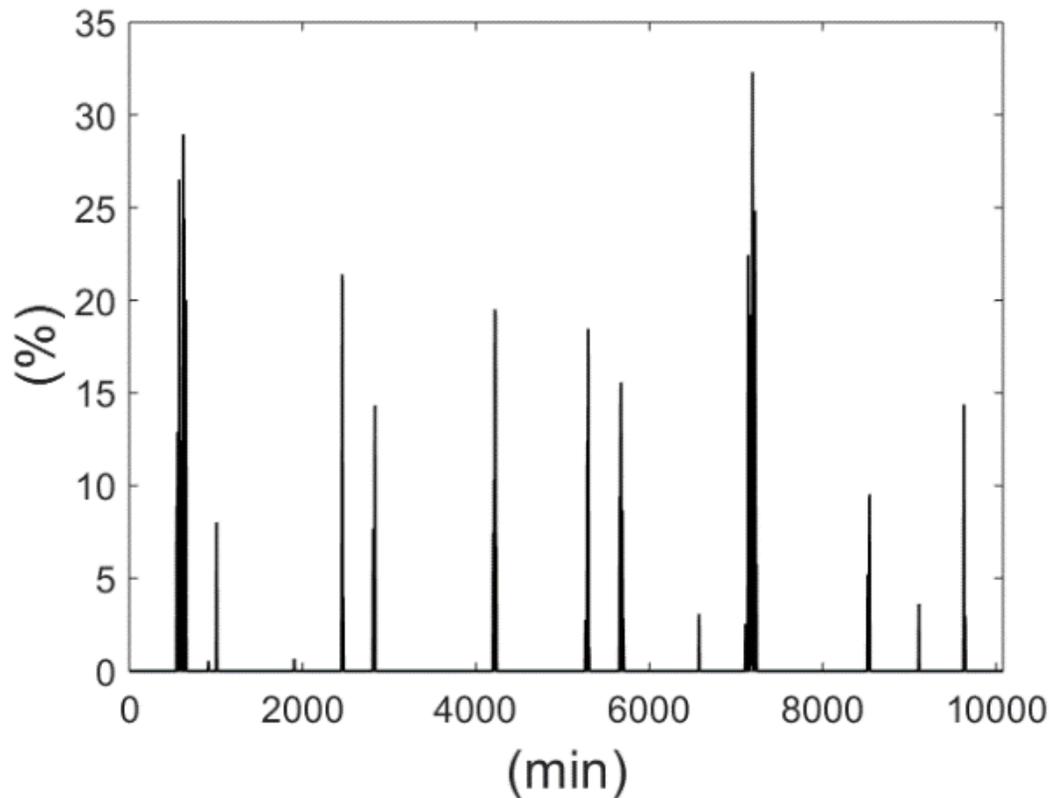


Figura 5. Error relativo porcentual del modelado.

Con el fin de tener un mejor modelo, se hace la calibración utilizando la información de tres semanas, para ello se usan los valores máximos de cada instante de tiempo de las tres primeras semanas:

$$W_{max}[k] = \max\{W_1[k], W_2[k], W_3[k]\}; \forall k = 1, 2, 3, \dots, N$$

donde $W_1(k)$, $W_2(k)$ y $W_3(k)$ denotan, respectivamente, las k -ésimas muestras de las series temporales de la primera, segunda y tercera semana; k es el índice de la muestra y N es el número total de muestras que conforman la serie temporal.

De modo que se crea un nuevo vector con la combinación de la información de tres semanas, por ejemplo, para el valor $W_{max}(1)$ se comparan los datos de las tres semanas en el instante 1 y se toma el máximo de ellos.

Generalizando el procedimiento para M semanas, se tiene:

$$W_{max} = \max\{W_1[k], W_2[k], \dots, W_M[k]\}; \forall k = 1, 2, 3, \dots, N$$

Una vez que se calcula $W_{max}[k]$, se utiliza para calibrar el modelo. El resultado es:

$$D(t) = 2.66 - 0.6284 \cos(0.0044t) - 0.7801 \sin(0.0044t) + 0.3778 \cos(0.0088t) - 0.0684 \sin(0.0088t) + 0.0155 \cos(0.0132t) - 0.0076 \sin(0.0132t) + \lambda \left(0.9198 \left(\frac{h(t)}{1-h(t)} \right)^{1/0.2357} \right)$$

La Figura 6 muestra el modelo calibrado (en negro) con los datos de tres semanas y de la demanda medida durante las 12 semanas (en diversos colores). Por otro lado, la Figura 7 muestra los instantes de tiempo en los que los datos de la semana 5 sobrepasaron la respuesta del

modelo con $\lambda = 1$. Aunque el error de modelado disminuye cuando el modelo se calibra con datos de tres semanas, aún sigue existiendo en algunos instantes de tiempo.

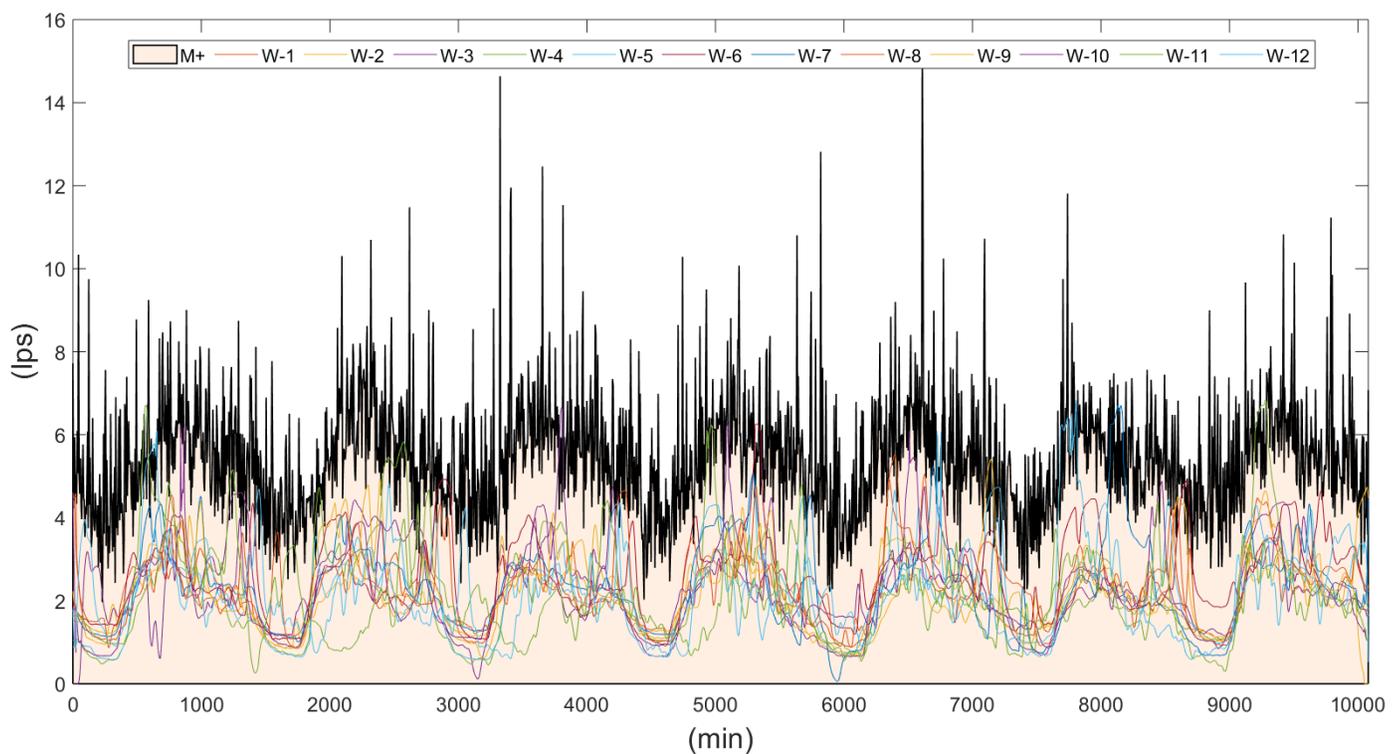


Figura 6. Modelo calibrado con los datos de las semanas 1, 2 y 3 vs. datos recolectados en campo durante 12 semanas.

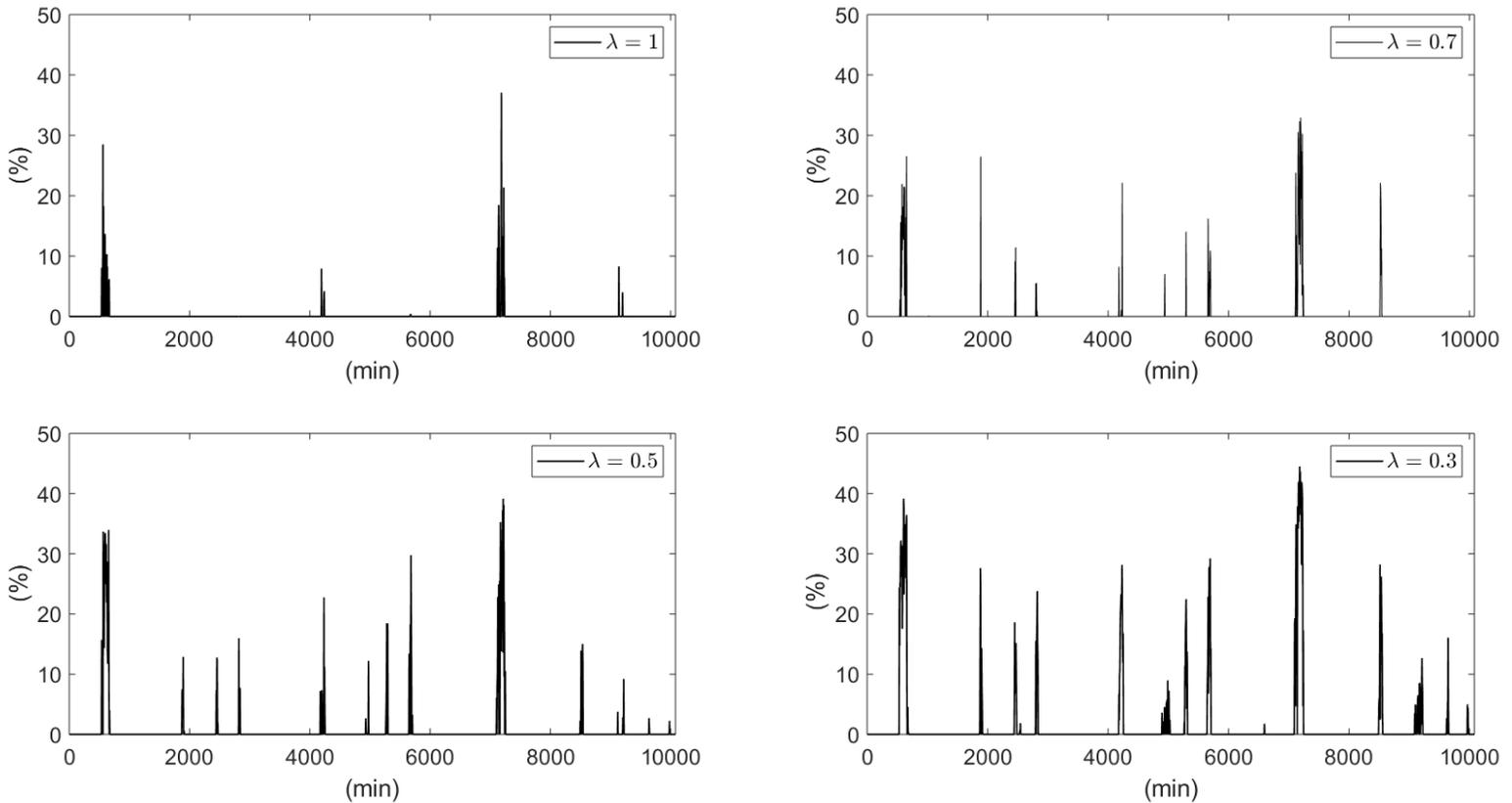


Figura 7. Error relativo porcentual del modelado.

En la Figura 7 se puede observar también el efecto de reducir el valor de λ . Entre más pequeño es su valor, mayor es el error de modelado. Al aumentar o disminuir el valor de λ se le da un peso a la componente aleatoria, es decir, se expande o se contrae la incertidumbre.

Si el modelo es calibrado con datos de las 12 semanas y se calcula el error de modelado con datos de cualquiera de esas 12 semanas, el error en cualquier instante de tiempo será cero, pues la aleatoriedad se

reduce. Esta reducción permite que el valor de λ sea menor. Por ejemplo, si $\lambda = 0.7$, el modelo calibrado con datos de las 12 semanas está dado por

$$D(t) = 3.704 - 0.5935 \cos(0.0044t) - 0.8478 \sin(0.0044t) + 0.7168 \cos(0.0088t) - 0.2567 \sin(0.0088t) + 0.2974 \cos(0.0132t) + 0.2358 \sin(0.0132t) + 0.7 \left(1.2834 \left(\frac{h(t)}{1-h(t)} \right)^{1/0.2074} \right)$$

La Figura 8 muestra la respuesta del modelo junto con los datos medidos.

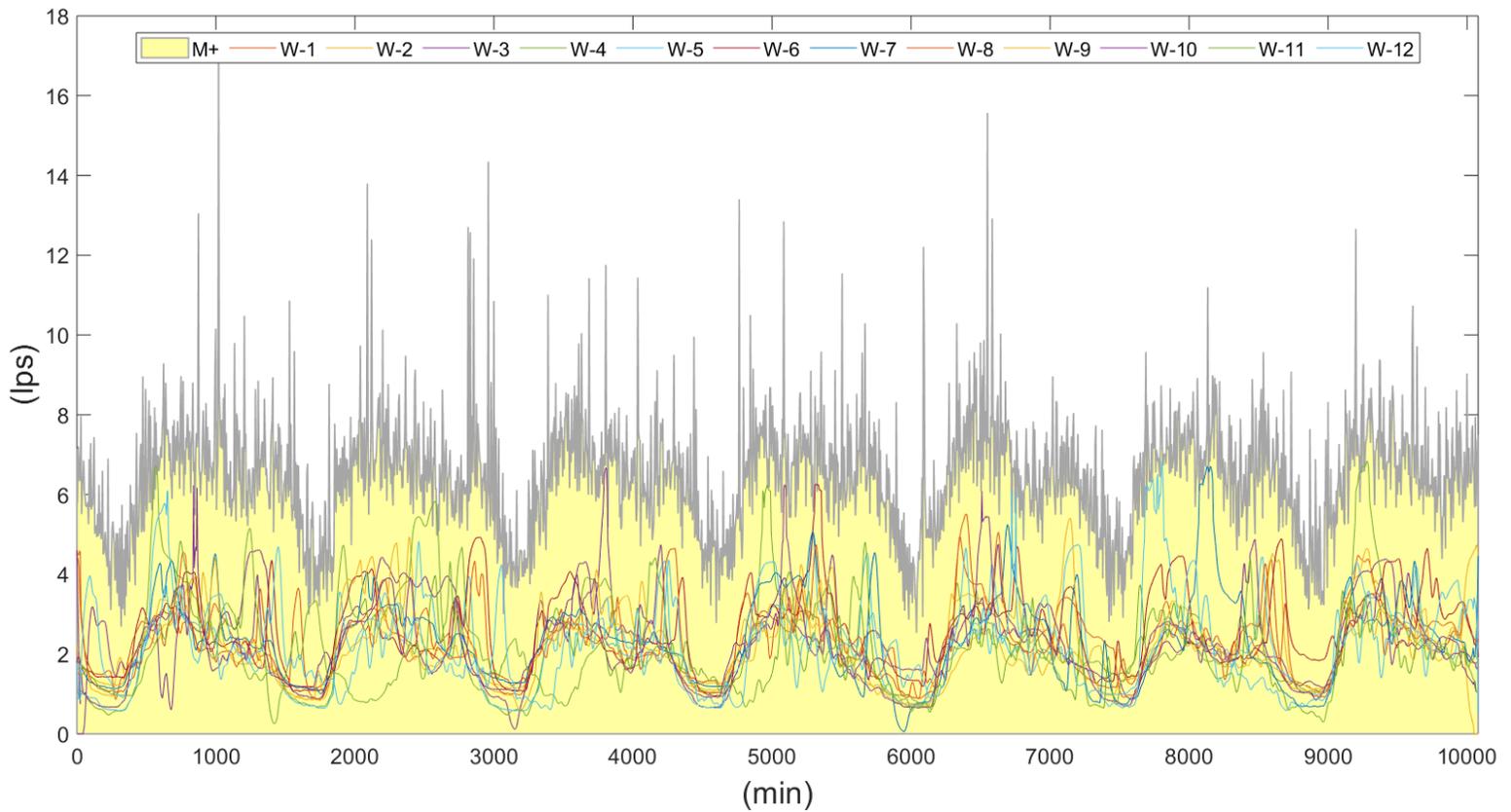


Figura 8. Modelo calibrado con los datos de las 12 semanas vs. datos recolectados en campo durante 12 semanas.

Conclusiones



En este artículo se propuso un modelo para estimar la máxima demanda semanal de un sector de la red de distribución de agua potable de la ciudad de Guanajuato, México. El modelo está compuesto de un submodelo determinístico y un modelo estocástico. El primero se usa para describir la periodicidad de la demanda, mientras que el segundo para representar la componente aleatoria, es decir, la incertidumbre debida a agentes exógenos no cuantificables.

Se mostró que la componente aleatoria de la demanda semanal de agua puede representarse con una función de probabilidad log-logística, la cual tiene una función cuantil simple que puede usarse para generar datos con una distribución aproximada por la función log-logística.

Para regular la incertidumbre o la componente aleatoria del modelo se propuso un parámetro de peso λ que afecta el submodelo aleatorio. Este parámetro debe elegirse o estimarse con base en cuánta información se está utilizando para calibrar el modelo; si la información es escasa, λ debe ser 1 o cercano a 1; si la información es abundante o significativa, λ puede tomar valores menores a 1.

El modelo propuesto puede emplearse en situaciones en las que, por eventos no previstos, no se cuente con una lectura de la demanda, la cual resulta imprescindible para la toma de decisiones, por ejemplo, en la ampliación de la red o en la agregación de nuevas tomas. El modelo también puede utilizarse para representar o simular numéricamente la demanda de otros sectores con características similares, como tamaño de población, actividades económicas preponderantes, nivel socioeconómico.

Agradecimientos

Este trabajo se llevó a cabo gracias al apoyo otorgado por Conacyt (Atención a Problemas Nacionales, Convocatoria 2017, Proyecto 4730), así como al apoyo proporcionado por el Sistema de Agua Potable y Alcantarillado de Guanajuato (SIMAPAG) y por la Universidad de Guanajuato (UG). En particular, los autores agradecen toda la ayuda brindada por la maestra Josefina Ortiz (UG), el ingeniero Carlos Navarro (SIMAPAG) y el ingeniero Miguel Molina (SIMAPAG).

Referencias

- Al-Shomrani, A. A., Shawky, A. I., Arif, O. H., & Aslam, M. (2016). Log-logistic distribution for survival data analysis using MCMC. *SpringerPlus*, 5(1), 1774.
- Alcocer-Yamanaka, V. H., Tzatchkov, V. G., & Arreguín-Cortés, F. I. (2012). Modeling of drinking water distribution networks using stochastic demand. *Water Resources Management*, 26(7), 1779-1792.
- Alvisi, S., Franchini, M., & Marinelli, A. (2007). A short-term, pattern-based model for water-demand forecasting. *Journal of Hydroinformatics*, 9(1), 39-50.

- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Nueva Jersey, Estados Unidos de América: John Wiley & Sons.
- Brentan, B., Luvizotto Jr. E., Izquierdo, J., & Pérez-García, R. (2016). Fourier series and Chebyshev polynomials applied to real-time water demand forecasting. *Acta Universitaria Multidisciplinary Scientific Journal*, 26, 74-81.
- Cano-Rodríguez, I., Gómez-Vallejo, F., Aguilera-Alvarado, A. F., De-la-Rosa, G., & Gardea-Torresdey, J. (2004). Transporte y destino final de fósforo en el acuífero de Puentecillas de Guanajuato. *Revista Mexicana de Ingeniería Química*, 3(2), 201-208.
- Cano-Rodríguez, I., Gómez-Vallejo, F., Mendez, V. R., Barbosa, P. M., Rodríguez, E. R., & Alvarado, A. A. (2000). *Determinación de contaminantes en la presa La Purísima y su efecto en el sistema de pozos puentecillas de Guanajuato*. En: Scott, C. A.; Wester, P.; Maranon-Pimental, B. (Eds.) 2000. Asignación, productividad y manejo de recursos hídricos en cuencas (in Spanish) (pp. 123-133). Mexico, DF., Mexico: IWMI. xix, 234p. (IWMI Serie Latinoamericana 020) doi: <http://dx.doi.org/10.3910/2009.411>
- De-Marinis, G., Gargano, R., & Tricarico, C. (2006). Water demand models for a small number of users. In: *8th Annual Water Distribution Systems Analysis Symposium* (pp. 1-14). Cincinnati, Ohio, USA, August 27-30.

- Gargano, R., Tricarico, C., Granata, F., Santopietro, S., & De-Marinis, G. (2017). Probabilistic models for the peak residential water demand. *Water*, 9(6), 417.
- Maidment, D. R., & Parzen, E. (1984). Cascade model of monthly municipal water use. *Water Resources Research*, 20(1), 15-23.
- Surendran, S., & Tota-Maharaj, K. (2018). Effectiveness of log-logistic distribution to model water-consumption data. *Journal of Water Supply: Research and Technology-AQUA*, 67(4), 375-383.
- Surendran, S., & Tota-Maharaj, K. (2015). Log logistic distribution to model water demand data. *Procedia Engineering*, 119, 798-802.
- Tanyimboh, T., Surendran, S., Naga, D., Topping, B. H. V., & Soares, C. M. (2004). Loglogistic modelling of municipal water demands. In: *Proceedings of the Fourth International Conference on Engineering Computational Technology*. Stirling, Escocia: Civil-Comp Press.
- Torres, L., Jiménez-Cabas, J., Gómez-Aguilar, J. F., & Pérez-Alcázar, P. (2018). A simple spectral observer. *Mathematical and Computational Applications*, 23(2), 23.
- Tzatchkov, V., Alcocer-Yamanaka, V., Arreguín-Cortés, F., & Feliciano-García, D. (2003). Medición y caracterización estocástica de la demanda instantánea de agua potable. *Ingeniería Hidráulica en México*, 20(1), 67-76.

Tzatchkov, V. G., & Alcocer-Yamanaka, V. H. (2016). Modelación de la variación del consumo de agua potable con métodos estocásticos. *Tecnología y ciencias del agua*, 7(3), 115-133.

Zhou, S. L., McMahon, T. A., Walton, A., & Lewis, J. (2002). Forecasting operational demand for an urban water supply zone. *Journal of Hydrology*, 259(1-4), 189-202.