



Modelo para el pronóstico de la demanda real de agua potable en Quito

Model for the forecast of the real demand of drinking water in the city of Quito



* **Mat. Danilo Gortaire J., Ph.D.** (UCE) / Facultad de Ingeniería, Ciencias Físicas y Matemática. Docente titular de la Carrera de Ingeniería Matemática. / danilo_gortaire@yahoo.es

****Ing. Edgar Ayabaca C., Ph.D.** (EPMAPS, UCE), edgarjoseac@gmail.com, edgar.ayabaca@aguaquito.gob.ec /

Ing. Felipe Borja B., M.Sc. (UCE), / lborja@uce.edu.ec

Mat. Benjamín Valarezo P., M.Sc. (UCE) / benjavalarezo@gmail.com

Resumen

Se propone desarrollar, calibrar y validar un modelo matemático que pronostique la demanda real de agua potable en Quito con horizontes de corto, mediano y largo plazos, considerando las principales variables que actúan en el suministro de agua. Se supone que la demanda de agua potable en Quito puede ser pronosticada ya que su fluctuación y crecimiento están definidos en gran parte por variables meteorológicas, urbanísticas o de ordenamiento territorial, demográficas, económicas, sociales y la autocorrelación propia de la demanda. El modelo de pronóstico de la demanda de agua corresponde al estudio y análisis de una serie temporal o vector X_1 de 2557 datos de caudales medios diarios (litros/seg.) entregados por la Planta de Tratamiento de Agua Potable de Bellavista, PTAP de Quito, desde enero 2007 hasta diciembre 2013. Para el análisis de la serie temporal o vector X_1 se utilizan varias herramientas estadísticas y la transformada rápida de Fourier que nos ayuda a determinar las periodicidades del fenómeno. Se concluye obteniendo explícitamente la función de pronóstico de agua potable buscada. La validación de la función de demanda obtenida se la hará en la segunda etapa de la investigación, usando los datos que van desde el inicio de 2014 hasta el año 2016.

Palabras clave: Serie temporal de demanda de agua; modelo aditivo; tendencia; componente periódica; componente aleatoria; transformada de Fourier; frecuencias; método de mínimos cuadrados; coeficiente de determinación; polinomio trigonométrico de ajuste; función final de demanda

Abstract

*It is proposed to develop, calibrate and validate a mathematical model that predicts the real demand for drinking water in Quito horizons short, medium and long term, considering the main variables that act in the water supply. It is assumed that the demand for drinking water in Quito can be predicted as their fluctuation and growth are defined largely by weather, urban planning or land use, demographic, economic, social variables and autocorrelation own demand. The model forecast water demand corresponds to the study and analysis of a time series or vector X_1 2557 data average daily flow (liters / sec.) Delivered by Treatment Plant Drinking Water Bellavista, POWPA of Quito, from January 2007 to December 2013. for the analysis of the time series or vector X_1 several statisticians tools and fast Fourier transform that helps us determine the periodicities of the phenomenon are used. It is concluded explicitly obtaining sought forecast function potable water. The validation of the demand function obtained is the will in the second stage of the research, using data ranging from the beginning of 2014 until 2016. The highest number of resistance (4 antibiotics) was exhibited by *Brevundimonas diminuta* strains. All isolates showed resistance to Cephalothin. It is demonstrated that these waters possess a characteristic microbiota and the presence of environmental resistomas.*

Keywords: time series of water demand; additive model; trend; periodic component; random component; Fourier transform; frequency; method of least squares; coefficient of determination; trigonometric polynomial adjustment; final demand function

1. Introducción

El suministro de agua por parte del Sistema Papallacta-Bellavista al Centro y Norte de Quito ha alcanzado un promedio anual de 2,4 m³/s, lo cual implica que en cerca de 2 o 3 años más, la ciudad estará utilizando permanentemente toda la capacidad instalada de conducción de agua cruda y potabilización en la Planta de Tratamiento de Bellavista [3; 12]. En estas condiciones se producirán horas y/o días con déficit en el suministro, razón por la cual la EPMAPS, a más de construir nuevas tuberías de conducción y de ampliar la planta de tratamiento, ha previsto gestionar la demanda de agua mediante el uso de un “Sistema de Pronóstico de Demanda de Agua Potable” que permita conocer con anticipación los requerimientos de agua y programar la forma más eficiente de satisfacerlos.

Los pronósticos sobre la demanda de agua que requiere el suministro de agua potable de Quito abarcan el corto, mediano y largo plazos. Así, la operación de las plantas de tratamiento exige pronósticos a nivel horario, diario y semanal [14; 19]; mientras que la operación y manejo de los embalses demandan pronósticos semanales, mensuales, estacionales y anuales [9; 18]. Esta contribución de los modelos predictivos a la operación y gestión es particularmente importante cuando la capacidad de producción de agua está en su punto límite, por cuanto al prever con suficiente antelación las demandas (máximas, medias y mínimas), se puede: reducir los fallos en el suministro, programar las entregas de agua por los embalses, optimizar el bombeo y el consumo de energía, planificar la adquisición de químicos y otros insumos, reducir las pérdidas y fugas, evitar riesgos, etc.

Los pronósticos de largo plazo encuentran especial aplicación en la planificación y diseño de las obras ya que la capacidad de los futuros proyectos debe seleccionarse considerando las demandas reales y otras condiciones que se producirán durante su vida útil. La magnitud de los beneficios, que tales pronósticos pueden proporcionar, puede superar varios millones de dólares al año para el caso de Quito.

Esta estimación se sustenta en las mediciones de los últimos años que la EPMAPS realiza cada 15 minutos de los caudales entregados por la Planta de Tratamiento de Bellavista; esas mediciones demuestran que el caudal máximo diario de la demanda supera sólo en el 12%, al caudal medio anual. Este porcentaje es significativamente inferior al 25% que establecen diversas normas de diseño e inclusive el Plan Maestro de Agua Potable de Quito [10; 11; 12; 13; 16; 17]. Si la investigación aquí propuesta, confirma que ese porcentaje del 12% puede aplicarse en el diseño y construcción de las obras de agua potable, ello implicará un ahorro en inversiones de cerca de 4 millones de USD al año en Quito.

Estos beneficios se explican por el hecho de que las obras de captación de agua, conducciones, tuberías, plantas de tratamiento, estaciones de bombeo, equipos y otras instalaciones de agua potable se están diseñando con un coeficiente de mayoración igual al 125%. Reducir

ese coeficiente al 112% equivale a reducir aproximadamente en un mismo porcentaje el tamaño y las inversiones anuales de las obras que requiere la ciudad. Es evidente que los beneficios aquí descritos serán mucho mayores si los pronósticos de la demanda de agua potable se aplican en otras ciudades y lugares del país; por lo tanto, esta investigación aplicada, pionera en nuestro medio, contribuirá a la gestión eficiente del agua.

Por lo expuesto, esta propuesta de investigación plantea aplicar a las condiciones de Quito la teoría de los modelos de predicción de demanda máxima de agua potable con horizontes de cortísimo, corto, mediano y largo plazos, para lo cual se aprovecharán las mediciones realizadas por la EPMAPS durante los últimos cuatro años en la Planta de Bellavista.

La investigación aplicada para desarrollar el modelo de pronóstico se apoyará en métodos clásicos (análisis de Fourier, modelos estocásticos) y posiblemente otras propuestas matemáticas (redes neuronales, análisis fuzzy, etc.) para prever la demanda de agua potable en el corto, mediano y largo plazos en Quito. Para la calibración y validación de los modelos se cuenta con cuatro años de mediciones, realizadas cada 15 minutos en Bellavista, la mayor planta de tratamiento de agua potable de Quito, así como mediciones complementarias en otras plantas menores.

Se estudiarán y cuantificarán las relaciones de la demanda de agua potable con factores meteorológicos (temperatura del aire, radiación solar, número de días consecutivos secos, etc.) y aspectos sociales, como los patrones de conducta de la población durante las festividades. Los resultados de la investigación contribuirán a optimizar la operación y costos de la producción de agua potable, a la vez que coadyuvarán a elevar la eficacia en la gestión de las aguas urbanas en el País. Particular atención se brindará al estudio de los caudales máximos diarios y su real proporción respecto a los caudales medios anuales, ya que en base a esa proporción se definen importantes dimensiones y, por ende, los costos de conducciones, estaciones de bombeo, plantas de tratamiento, equipos y otros componentes de la infraestructura de agua potable. Esta investigación creará la base para la formación de jóvenes investigadores en la Universidad Central en gestión de los recursos hídricos.

2. Marco Teórico

El presente proyecto de investigación se enmarca en la aplicación de modelos estocásticos y de inteligencia artificial al pronóstico de la demanda de agua potable que consideran la importancia de los factores que condicionan la demanda, el horizonte del pronóstico requerido, así como la calidad de los datos utilizados en la predicción. La demanda de agua está influenciada por variables que son controlables (restricciones físicas del sistema, precio del agua, etc.) y por factores no controlables (hábitos de la población, desarrollo económico, estado del tiempo, vacaciones, etc.), cuya incidencia es



variable en dependencia de cuál sea el horizonte del pronóstico deseado.

La mayoría de los métodos desarrollados para pronosticar la demanda de agua potable requieren datos históricos de la demanda y asumen que ese comportamiento observado se reproducirá en el futuro [1; 2; 4; 7]. Así es como la modelización y pronóstico de series temporales de demanda de agua, desde sus inicios, se apoyó en el análisis de Fourier, ya que con su ayuda se pueden identificar los componentes periódicos de la demanda de agua (a nivel estacional, semanal, horario, etc.) y sus resultados permiten identificar los armónicos predominantes. Junto con este análisis es habitual estudiar la correlación que suele existir entre la demanda de agua de un período de tiempo con la demanda de períodos previos [6; 8; 14; 15], así como la correlación con factores externos (temperatura ambiental) y sociales (hábitos de la población en feriados, vacaciones, paros, huelgas, etc.).

Un avance en la elaboración de los pronósticos de demanda de agua potable constituye la utilización de modelos matemáticos de series temporales ARIMA [5; 19] con términos estacionales y no estacionales de órdenes p , q , P y Q , según la metodología de Box y Jenkins. Estos modelos estocásticos pueden ser calibrados y validados con las mediciones de la demanda de agua que se han recopilado durante varios años; su desempeño se evalúa cuantificando los errores y el sesgo de los pronósticos que proporcionan. Se seleccionan aquellos modelos que mejor satisfacen los requerimientos del usuario del pronóstico, en función del horizonte predicción y de la disponibilidad oportuna de mediciones sobre las variables exógenas (pronósticos de temperatura, etc.) que considera el modelo. Para mejorar la predictividad de estos modelos se ha planteado la posibilidad de formularlos para cada época o estación del año, así como combinarlos con modelos basados en redes neuronales, ya que ellas permiten considerar la no linealidad de las series temporales [1; 2; 9].

En complemento a lo indicado cabe agregar que en la elaboración de pronósticos de demanda, últimamente se están aplicando modelos basados en la teoría de la lógica difusa, los algoritmos genéticos, el razonamiento inductivo fuzzy, sistemas expertos, etc. que permiten superar limitaciones tales como la incertidumbre o insuficiencia de la información disponible, a la vez que abren la posibilidad de incluir el criterio de expertos.

3. Preguntas de Investigación

Los planteamientos y lineamientos fundamentales de esta investigación aplicada pueden resumirse en las siguientes preguntas:

- ¿Cuáles son las variables endógenas y exógenas (climáticas, sociales, etc.) que explican las fluctuaciones y el desarrollo de la demanda de agua potable en Quito a

cortísimo, corto, mediano y largo plazos?

- ¿Cuáles son los modelos estocásticos o de otro tipo que proporcionan los mejores pronósticos a cortísimo, corto, mediano y largo plazos de la demanda real de agua potable en Quito?
- ¿Cuál es la calidad de los pronósticos proporcionados por los modelos de pronósticos de demanda real de agua potable en Quito?
- ¿Cuál es el coeficiente de mayoración que se recomienda utilizar en la determinación del caudal máximo diario de agua potable que se utiliza en los diseños de la EPMAPS?
- ¿Qué mediciones adicionales deben realizarse en Quito para mejorar los pronósticos de demanda de agua potable a cortísimo, corto, mediano y largo plazos?
- ¿Qué otras recomendaciones serían factibles formularlas basándonos en el modelo obtenido?

4. Justificación

La revisión bibliográfica de los métodos que se han venido desarrollando para pronosticar la demanda de agua potable sustenta la hipótesis de que esa demanda en Quito depende de variables endógenas (inercia de la demanda o su autocorrelación) y exógenas (meteorológicas, sociales, económicas, urbanísticas).

Estas variables en nuestro medio físico geográfico tendrán una importancia y composición diferentes ya que en nuestro clima ecuatorial de montaña no existen las fuertes variaciones mensuales de temperatura ambiental que se observan en otras latitudes (en contraposición a ello, aquí tenemos una gran amplitud en la variación de la temperatura en el transcurso del día) y por ello será necesario investigar cuáles son las variables meteorológicas que deben utilizarse en nuestros modelos de pronóstico.

La identificación de cuáles son los factores endógenos y exógenos que resultan válidos para pronosticar la demanda en nuestro medio Ecuatorial, constituye ya un aporte al acervo científico en modelación matemática y en gestión de los recursos hídricos. Mayor será la contribución al conocimiento el momento en que la investigación demuestre la bondad de utilizar modelos diferenciados para las diversas épocas del año, así como las posibilidades que surgen al aplicar la teoría de la lógica difusa.

5. Metodología y experimento numérico.

- En el Diseño de Estudio interviene la recopilación de los datos y su depuración o verificación de su calidad, para luego pasar al análisis estadístico preliminar (estadística descriptiva).

En este paso ordenamos cronológicamente los datos, verificamos

su calidad con un Test estadístico, calculamos algunos estadísticos (medias, moda, mediana, desviación estándar, otros), analizamos el histograma y la ojiva, y también postulamos alguna hipótesis para la distribución estadística de los 2.557 datos:

2.012,89
2.083,15
2.159,67
.....
2.448,70
2.476.18
2.498,88

Donde el primer dato de la columna corresponde al 1 de enero 2007 y el último al 31 de diciembre 2013, siendo un total de 2.557 datos encolumnados cronológicamente. Todos los datos corresponden a los caudales de agua por día, ya corregidos y medidos en litros / segundo.

Para la corrección de los datos con la consecuente eliminación de los datos anómalos utilizamos el estadístico τ , donde t es el valor de cada dato, \bar{x} es la media de todos los datos y s es su desviación estándar

Criterio de la distribución $t_{(p,n-2)}$:

Valores límites:

$$t_{(5\%,2557-2)} = 1.6479$$

$$t_{(0,1\%,2557-2)} = 3.1066$$

$$\tau_{(p,n)} = \frac{t_{(p,n-2)}\sqrt{n-1}}{\sqrt{n-2 + (t_{(p,n-2)})^2}}$$

$$\tau_{(5\%,2557)} = \frac{t_{(5\%,2555)}\sqrt{2556}}{\sqrt{2555 + (t_{(5\%,2555)})^2}} = 1.65$$

$$\tau_{(0,1\%,2557)} = \frac{t_{(0,1\%,2555)}\sqrt{2556}}{\sqrt{2555 + (t_{(0,1\%,2555)})^2}} = 3.10$$

Criterios:

1) $\tau \leq \tau_{(5\%,n)}$ no eliminables

2) $\tau_{(5\%,n)} < \tau < \tau_{(0,1\%,n)}$ eliminables por conveniencia

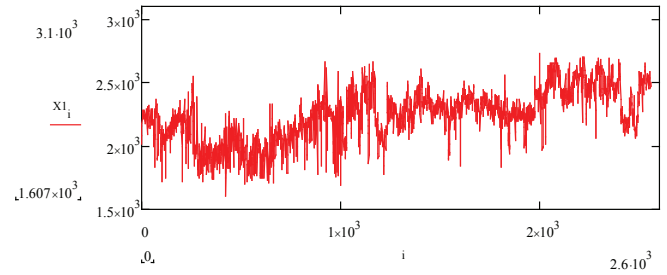
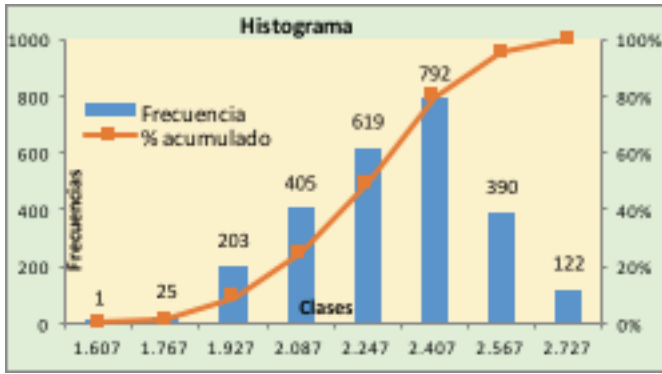
3) $\tau > \tau_{(0,1\%,n)}$ eliminables

Medidas de Centralización y de Dispersión

Media	2233.84
Error típico	4.09
Mediana	2252.00
Moda	2170.96
Desviación estándar	206.79
Varianza de la muestra	42760.31
Curtosis	-0.44
Coficiente de asimetría	-0.22
Rango	1119.89
Mínimo	1607.40
Máximo	2727.29
Suma	5711922.84
Cuenta	2557.00
Nivel de confianza (95.0%)	8.02

#	Clases:	
1	1607.4	1767.4
2	1767.4	1927.4
3	1927.4	2087.4
4	2087.4	2247.4
5	2247.4	2407.4
6	2407.4	2567.4
7	2567.4	2727.4
8	2727.4	2887.4
Clase	Frecuencia	% acumulado
1607.4	1	0.04%
1767.4	25	1.02%
1927.4	203	8.96%
2087.4	405	24.79%
2247.4	619	49.00%
2407.4	792	79.98%
2567.4	390	95.23%
2727.4	122	100.00%

Test χ^2 para comprobar la normalidad de los datos observados (los valores por filas corresponden a cada clase):



$m=2$ es el número de parámetros a estimarse (media y desviación típica).

$\nu = k - 1 - m = 4$ son los grados de libertad = (n° de clases $- 1 - n^\circ$ de parámetros = 4).

Al cumplirse la relación $(\chi^2_{(2,10)} = 4,605) < (\hat{\chi}^2 = 7166,5)$ para los datos observados del vector original X_1 , se rechaza la hipótesis de que los datos sigan una distribución normal.

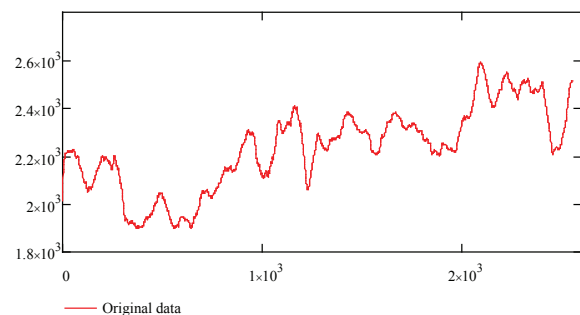
Con el Criterio de Kolmogorov-Smirnov tenemos:

Frec. obs. F_o	Frec. esper. F_e	F_o acumulada	F_e acumulada	$\frac{F_o - F_e}{n}$
1	0.00224336	1	0.002	0.000
13	0.04882143	14	0.051	0.005
32	0.69062335	46	0.742	0.018
2	6.35025903	48	7.092	0.016
14	37.9543078	62	45.046	0.007
29	147.451828	91	192.498	0.040
37	372.356294	128	564.854	0.171
267	611.204704	395	1176.059	0.305
515	652.129922	910	1828.189	0.359
981	452.273372	1891	2280.462	0.152
537	203.88598	2428	2484.348	0.022
122	59.7438432	2550	2544.092	0.002
6	11.3793748	2556	2555.472	0.000
n =			max =	0.359
2557				

b) El suavizamiento de los datos comprende el proceso de eliminación de los picos o de las oscilaciones existentes de un día a otro. Este proceso numérico elimina las grandes variaciones existentes entre los distintos valores diarios vecinos. Un proceso de suavizamiento puede realizarse mediante el uso de splines, otros procesos, o mediante un proceso de media móvil de longitud w , por ejemplo $w = 50$, es decir, generamos la nueva serie temporal o vector X , cuya primera componente es el promedio de los primeros 50 valores ($x_{1,1}, \dots, x_{1,50}$) del vector original X_1 , la segunda componente de X es el promedio de los 50 valores ($x_{1,2}, \dots, x_{1,51}$) del vector original X_1 , la tercera componente de X es el promedio de los 50 valores ($x_{1,3}, \dots, x_{1,52}$) del vector original X_1 , y así sucesivamente hasta llegar al dato número 2.557 del vector original X_1 . El nuevo vector X es:

2.013
2.048
2.085
.....

La gráfica del vector suavizado X es:



Como se cumple la relación

$$(\hat{D} = \frac{\max|F_o - F_e|}{n} = 0,359) \geq (D_{(2557;0,10)} =$$

0,12), se rechaza la hipótesis de que los datos del vector original observado X_1 sigan una distribución normal.

Gráfica del vector de datos originales:

c) El modelo de ajuste del vector X y de sus datos se basa en el modelo aditivo:

$$x(t) = T^*(t) + P^*(t) + E^*(t),$$

donde: $T^*(t)$ es la tendencia general de los datos X , $P^*(t)$ es la componente periódica y $E^*(t)$ es la componente aleatoria o ruido.

Utilizando el método de mínimos cuadrados, y suponiendo

(acorde a la gráfica o comportamiento general de X) que los datos del vector X poseen una tendencia general lineal del tipo $H(t)=at+b$, obtenemos $H(t)=0.179t+2002$.

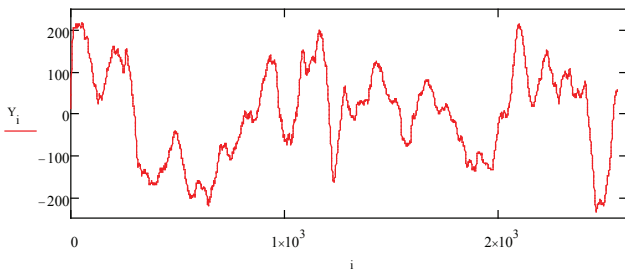
De la ecuación lineal de la tendencia H vemos que la pendiente de la recta tiene un valor $a=0.179$ lo que nos dice que la demanda del caudal de agua potable es creciente con el tiempo.

El polinomio trigonométrico o componente t tiene la forma $Pat=k=1mbk\cos 2\pi f_k t + ck\sen 2\pi f_k t$, donde los b_k y c_k son incógnitas a ser buscadas, $n=2557$, f_k son las frecuencias buscadas, α es el número de polinomios trigonométricos necesarios que permitirán en forma iterativa obtener una aproximación óptima o final. El proceso de búsqueda de los polinomios $P(t)$ y de sus coeficientes está estrechamente enlazado con el método de los mínimos cuadrados y con la transformada de Fourier (búsqueda de las frecuencias y periodos).

La aproximación mediante los polinomios trigonométricos se la hace inicialmente partiendo de la serie estacionalizada (para nuestro caso eliminando la tendencia H del vector X), es decir, de nueva serie estacionalizada toma la forma:

$$Y(t)=X(t)-H(t)$$

En este caso, la gráfica $Y(t)$ de tendrá la forma:



la misma que muestra claramente estar estacionalizada (sin tendencia).

Aplicamos la transformada rápida y discreta de Fourier para los números complejos x_0, \dots, x_{N-1} , siendo la transformada discreta de Fourier (DFT, por sus siglas en inglés) la expresión definida como:

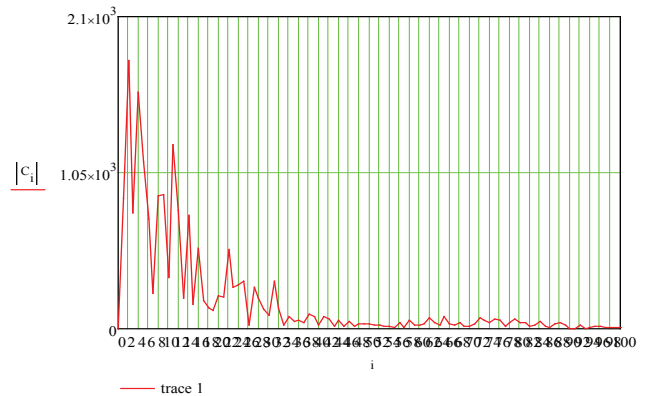
$$X_k = \sum_{n=0}^{N-1} x_n e^{i2\pi k \frac{n}{N}}, \quad k = 0, 1, \dots, N - 1$$

La evaluación directa de esa fórmula requiere $O(n^2)$ operaciones aritméticas. Mediante un algoritmo FFT se puede obtener el mismo resultado con sólo $O(n \log n)$ operaciones. Dado que la transformada discreta de Fourier inversa es análoga a la transformada discreta de Fourier, con distinto signo en el exponente y un factor $1/n$, cualquier algoritmo FFT puede ser fácilmente adaptado para el cálculo de la

transformada inversa. Por lo general tenemos que:

$$x_n = IDFT\{X_k\} = \frac{1}{N} (DFT\{X_k^*\})^*$$

Calculamos la Transformada de Fourier y el módulo del vector $Y(t)=X(t)-H(t)$ lo que representamos mediante $C=cfft(T)$ y $|cfft(Y)|$, respectivamente, dándonos el así denominado esquema de frecuencias (o periodograma):



De la gráfica anterior se obtienen las primeras frecuencias fundamentales: $f_k = 2, 4, 8, 9, 11, 14, 16, 20, 22, 25, 27, 31$, las mismas que permiten armar el vector $P(t)$ de 24 componentes o 12 armónicas $b_k \cos(\frac{2\pi f_k t}{n})$ y $c_k \sen(\frac{2\pi f_k t}{n})$, $k = 1, 2, \dots, 12$, donde los coeficientes son las incógnitas a ser buscadas con el método de mínimos cuadrados, siendo $n=2557$, t toma los valores $0, 1, 2, \dots, 2556$. Al aplicar el método de los mínimos cuadrados obtenemos los $12 + 12 = 24$ coeficientes b_k y c_k :

66.935
-44.837
-34.474
31.339
45.765
13.075
18.325
2.133
15.294
-8.835
-5.708
-4.54
-25.234
44.213
7.4
17.064



17.482
27.093
11.304
8.387
14.541
-9.105
9.557
11.724

$$P(t) = \sum_{k=1}^{12} \left(b_k \cos\left(\frac{2\pi f_k t}{n}\right) + c_k \sin\left(\frac{2\pi f_k t}{n}\right) \right)$$

Reemplazando todos los datos anteriores, obtenemos la forma explícita del polinomio trigonométrico anterior $P(t)$.

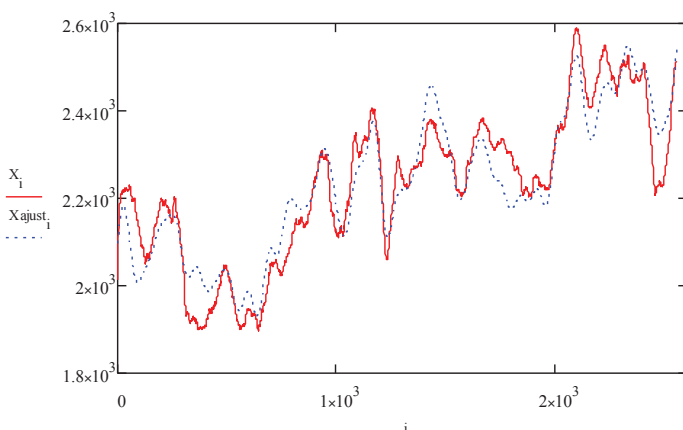
$$X_{ajust_i} = g(t_i) + P(t_i), \quad \text{donde } g(t_i) = H(t_i)$$

El primer coeficiente de determinación viene dado por:

$$R_{21} := 1 - \frac{\sum_{i=0}^{m1-1} |X_i - X_{ajust_i}|^2}{\sum_{i=0}^{m1-1} (X_i - \text{mean}(X))^2}$$

$$R_{21} = 0.877$$

La gráfica comparativa entre el vector original X y el del primer ajuste X_{ajust} , tiene la forma:



Procediendo de la misma manera y en forma iterativa con 4 Transformadas de Fourier, hasta obtener coeficientes b_K y c_K lo suficientemente pequeños (*del orden $\varepsilon=10^{-4}$*), es decir, coeficientes despreciables en el modelo de regresión multilínea, y paralelamente un coeficiente de determinación lo suficientemente cercano a 1, por ejemplo, si es posible, superior al 90%, obtenemos la función de ajuste predictiva

definitiva:

$$X_{pred_j} = g(t_j) + P(t_j) + P1(t_j) + P2(t_j) + P3(t_j),$$

Donde: $j=0, 1, \dots, 2556$; $g(t_j) = H(t_j) = 0.179t_j + 2002$; $P1(t_j)$ ya fue señalado anteriormente.

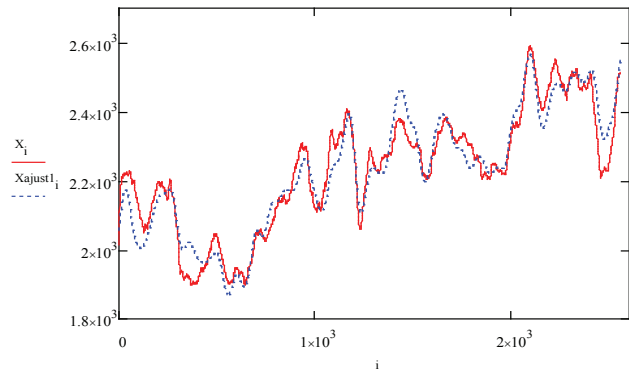
$$P1(t) = \sum_{k=1}^{15} \left(b_{1k} \cos\left(\frac{2\pi f_k t}{n}\right) + c_{1k} \sin\left(\frac{2\pi f_k t}{n}\right) \right)$$

Donde la Transformada de Fourier es $C1 := \text{cfft}(X - X_{ajust})$, las frecuencias fundamentales f_k son 1, 3, 5, 10, 12, 15, 17, 21, 23, 24, 28, 32, 34, 38, 41, para los $15 + 15 = 30$ coeficientes siendo **los valores de los coeficientes**

10.206
-6.242
-10.558
-0.278
-6.934
-8.605
-3.908
-9.297
0.068
-1.209
-0.748
-0.073
1.211
-2.363
-0.057
-40.678
10.327
-7.111
-0.192
8.393
0.371
2.715
-4.88
-0.7
-0.03
2.038
-1.451
1.546
0.367
0.393

En este caso pondremos: $X_{ajust1_i} := X_{ajust_i} + P1(t_i)$

El respectivo coeficiente de determinación es $R_{22} = 0.916$ y la gráfica comparativa del ajuste es:



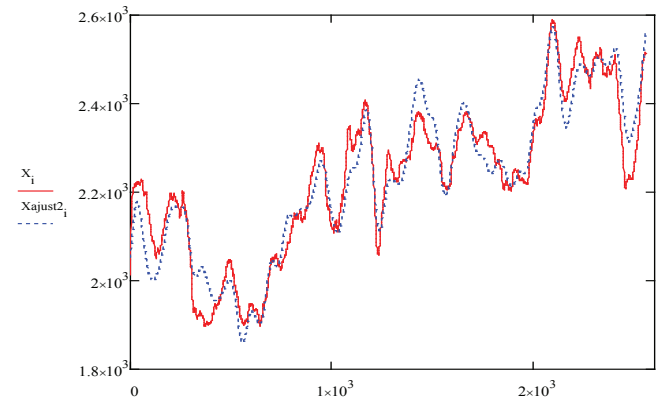
Para $P1(t_j)$ tenemos:

$$P2(t) = \sum_{k=1}^{14} \left(b_{2k} \cos\left(\frac{2\pi f_k t}{n}\right) + c_{2k} \text{sen}\left(\frac{2\pi f_k t}{n}\right) \right)$$

donde la Transformada de Fourier es $C2 := \text{cfft}(X - Xajust1)$, las frecuencias fundamentales f_k son 3, 5, 10, 12, 15, 16, 18, 21, 24, 28, 32, 34, 38, 41, para los $14 + 14 = 28$ coeficientes b_{2K} Y c_{2K} , $k = 1, 2, \dots, 14$; $n = 2557$ siendo los valores de los coeficientes:

-0.074
0.012
-0.381
-0.334
0.218
-5.337
0.018
0.365
-0.014
0.064
0.091
0.0046
0.019
0.08
-0.133
-0.031
-0.095
-0.432
-0.642
7.702
-0.453
0.296
0.216
-0.025
0.117
0.112
-0.011
0.04

El respectivo coeficiente de determinación es $R_{22} = 0.918$
La gráfica comparativa del ajuste tiene la forma:



Para $P3(t_j)$ tenemos:

$$P3(t) = \sum_{k=1}^{16} \left(b_{3k} \cos\left(\frac{2\pi f_k t}{n}\right) + c_{3k} \text{sen}\left(\frac{2\pi f_k t}{n}\right) \right)$$

donde la Transformada de Fourier es $C3 := \text{cfft}(X - Xajust2)$, las frecuencias fundamentales f_k son 3, 5, 10, 12, 15, 16, 21, 24, 28, 32, 34, 36, 38, 41, 44, 46, para los $16 + 16 = 32$ coeficientes b_{3K} Y c_{3K} , $k = 1, 2, \dots, 16$; $n = 2557$ siendo los valores de los coeficientes:

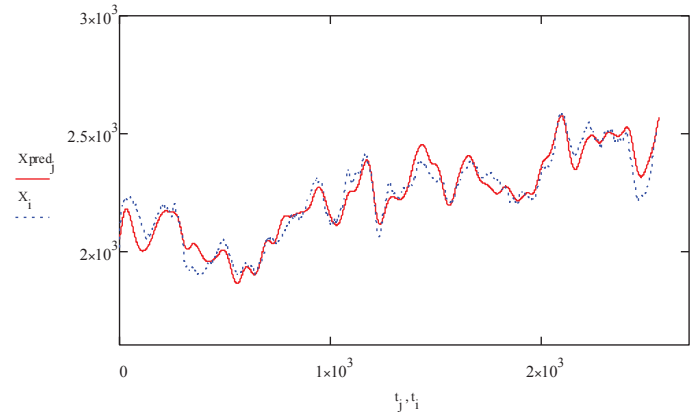
-9.55E-03
-0.023
5.12E-03
6.64E-03
-0.019
-0.026
6.62E-03
1.23E-03
-0.051
0.013
0.057
-1.346
0.035
0.012
-0.88
-0.562
0.014
2.54E-03
-0.015
1.56E-03
6.13E-03

En este caso pondremos: $Xajust2_i := Xajust1_i + P2|t_i$



-2.47E-03
-0.028
3.98E-03
-0.032
-0.096
-0.072
0.502
-0.01
-0.052
0.769
0.475

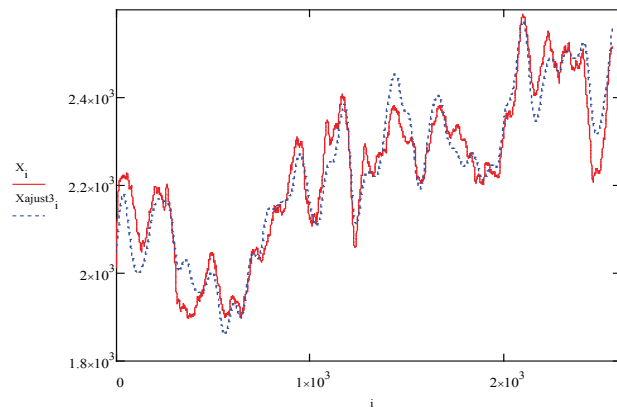
$$X_{pred_j} = g(t_j) + P(t_j) + P1(t_j) + P2(t_j) + P3(t_j)$$



En este caso se tiene: $X_{ajust3_i} := X_{ajust2_i} + P3(t_i)$

El respectivo coeficiente de determinación prácticamente permanece en $R_{23} = 0.918$, y la gráfica comparativa del ajuste tiene la forma:

Valores de pronóstico del caudal de agua potable para el año 2014:



El modelo aditivo definitivo para la función de ajuste o predictiva tiene la forma:

$$X_{pred_j} = g(t_j) + P(t_j) + P1(t_j) + P2(t_j) + P3(t_j),$$

donde cada componente o sumando ya se especificó anteriormente. El coeficiente de determinación entre los vectores X y X_{pred} es de $R^2 = 0,918$ lo que nos da cierta medida sobre la exactitud del ajuste de los datos X .

6. Resultados

Modelo aditivo de análisis de una serie temporal X basado en la aplicación de la Transformada de Fourier para determinar periodicidades o frecuencias escondidas en el fenómeno y en la aplicación del método de mínimos cuadrados MMC.

Metodología constructivista aplicable a problemas de determinación de caudales o de oferta - demanda de agua a nivel urbano o de otro tipo.

Función explícita de ajuste o predictiva y la gráfica comparativa del vector original X versus el vector de ajuste o de predicción X_{pred}

Pronóstico de la demanda de agua potable para el año 2014

Día	Valor	Día	Valor	Día	Valor
1	2571	122	2482	243	2573
2	2576	123	2484	244	2574
3	2581	124	2485	245	2574
4	2585	125	2486	246	2575
5	2589	126	2488	247	2575
6	2594	127	2489	248	2576
7	2598	128	2490	249	2576
8	2602	129	2491	250	2576
9	2605	130	2492	251	2577
10	2609	131	2493	252	2577
11	2612	132	2494	253	2577
12	2615	133	2495	254	2577
13	2618	134	2495	255	2577
14	2621	135	2496	256	2577
15	2624	136	2497	257	2577
16	2626	137	2498	258	2577
17	2628	138	2498	259	2576
18	2630	139	2499	260	2576
19	2632	140	2500	261	2575
20	2634	141	2500	262	2575
21	2635	142	2501	263	2574
22	2636	143	2501	264	2573
23	2637	144	2502	265	2572

24	2638	145	2503	266	2571	56	2578	177	2542	298	2474
25	2639	146	2503	267	2569	57	2574	178	2543	299	2471
26	2639	147	2504	268	2568	58	2570	179	2545	300	2467
27	2639	148	2505	269	2566	59	2566	180	2547	301	2463
28	2639	149	2505	270	2565	60	2562	181	2548	302	2459
29	2639	150	2506	271	2563	61	2558	182	2549	303	2456
30	2639	151	2507	272	2561	62	2554	183	2551	304	2452
31	2638	152	2508	273	2559	63	2549	184	2552	305	2449
32	2637	153	2509	274	2556	64	2545	185	2554	306	2446
33	2636	154	2510	275	2554	65	2541	186	2555	307	2442
34	2635	155	2511	276	2552	66	2537	187	2556	308	2439
35	2634	156	2512	277	2549	67	2533	188	2557	309	2436
36	2633	157	2513	278	2546	68	2529	189	2558	310	2433
37	2631	158	2514	279	2543	69	2524	190	2559	311	2431
38	2629	159	2515	280	2540	70	2520	191	2560	312	2428
39	2628	160	2516	281	2537	71	2516	192	2561	313	2425
40	2626	161	2517	282	2534	72	2512	193	2562	314	2423
41	2623	162	2519	283	2531	73	2508	194	2563	315	2421
42	2621	163	2520	284	2527	74	2504	195	2564	316	2419
43	2619	164	2521	285	2524	75	2500	196	2564	317	2417
44	2616	165	2523	286	2520	76	2497	197	2565	318	2415
45	2614	166	2524	287	2516	77	2493	198	2566	319	2413
46	2611	167	2526	288	2513	78	2489	199	2566	320	2411
47	2608	168	2527	289	2509	79	2486	200	2567	321	2410
48	2605	169	2529	290	2505	80	2483	201	2567	322	2408
49	2602	170	2531	291	2501	81	2480	202	2567	323	2407
50	2599	171	2532	292	2497	82	2477	203	2568	324	2406
51	2595	172	2534	293	2494	83	2474	204	2568	325	2405
52	2592	173	2535	294	2490	84	2471	205	2568	326	2405
53	2589	174	2537	295	2486	85	2469	206	2568	327	2404
54	2585	175	2539	296	2482	86	2466	207	2568	328	2403
55	2581	176	2540	297	2478	87	2464	208	2568	329	2403



88	2462	209	2568	330	2403
89	2461	210	2568	331	2402
90	2459	211	2568	332	2402
91	2458	212	2568	333	2402
92	2457	213	2568	334	2402
93	2456	214	2568	335	2403
94	2455	215	2568	336	2403
95	2454	216	2568	337	2403
96	2454	217	2568	338	2404
97	2454	218	2567	339	2404
98	2454	219	2567	340	2405
99	2454	220	2567	341	2405
100	2454	221	2567	342	2406
101	2454	222	2567	343	2407
102	2455	223	2567	344	2408
103	2456	224	2567	345	2408
104	2457	225	2567	346	2409
105	2458	226	2567	347	2410
106	2459	227	2567	348	2411
107	2460	228	2568	349	2412
108	2461	229	2568	350	2412
109	2463	230	2568	351	2413
110	2464	231	2568	352	2414
111	2465	232	2568	353	2415
112	2467	233	2569	354	2416
113	2468	234	2569	355	2416
114	2470	235	2570	356	2417
115	2472	236	2570	357	2418
116	2473	237	2570	358	2418
117	2475	238	2571	359	2419
118	2476	239	2571	360	2419
119	2478	240	2572	361	2420

120	2479	241	2572	362	2420
121	2481	242	2573	363	2421
				364	2421
				365	2421

7. Conclusiones

Se generará ahorro económico en las inversiones y costos de operación del servicio de agua potable de Quito y otras ciudades del País donde se aplique el sistema de pronóstico de demanda que entregará la presente investigación.

La función de pronóstico o predictiva del caudal de agua potable para Quito viene dada en forma explícita lo que permite hacer cálculos de caudales a futuro inmediato o cercano.

La función predictiva X_{pred} obtenida para el caudal de Quito puede ser precisada y mejorada en forma dinámica con la incorporación de los datos de los años 2014-2016. Los pronósticos irán desde el año 2017 en adelante y el Modelo obtenido será confrontado o comparado con modelos estocásticos u otros para tener la posibilidad de hacer un análisis comparativo de modelos y mirar que tan bueno es el presente modelo (Segunda etapa de la investigación).

Obtención de un modelo y metodología de pronóstico del caudal de agua potable aplicable a otras ciudades.

El coeficiente de determinación final, obtenido entre el vector de datos originales X y el vector correspondiente a la función de pronóstico o predictiva X_{pred} nos indica que esta función de ajuste justifica la información original en el orden del 91.8%.

Con ayuda de la función de ajuste o predictiva X_{pred} obtenida para el caudal se puede determinar los valores promedio o cargas por horas específicas, por días de la semana, por semanas o por meses, durante cualquier intervalo de tiempo entre los años 2007-2016 (Segunda etapa de la investigación).

Los errores relativos entre los datos originales X y los ajustados X_{pred} se encuentran en un rango que va desde un mínimo de 120 l/s a un máximo de 150 l/s.

Eficiencia en la gestión del agua urbana y la mitigación de riesgos y conflictos vinculados con el agua.

Bienestar ciudadano.

Reducción de conflictos por el uso del agua.

Con ayuda de los coeficientes obtenidos en el modelo de regresión multilínea de la función de pronóstico, es factible realizar capacitaciones para el personal de la EPMAPS en el manejo de la función predictiva del caudal de agua potable, por ejemplo, utilizando Excel.

8. Agradecimiento

Expresamos nuestro agradecimiento a las autoridades de la EPMAPS por la valiosa ayuda de sus funcionarios, colaboración con los datos y por el interés prestado a la presente investigación.

Agradecimiento al Dr. Luis Horna, Ph.D., catedrático de la EPN, por sus valiosas observaciones y recomendaciones al artículo.

9. Referencias

Adamowski J., Karapataki Ch. 2010. Comparison of Multivariate Regression and artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. *Journal of Hydrologic Engineering*, Vol. 15, No. 10.

Adamowski J. 2008. Peak Daily Water Demand Forecast Modeling Using Artificial Neural Networks. *Journal of Water Resources Planning and Management*.

Ayabaca Edgar. et al. 2006. Proyecto de Agua Potable Ríos Orientales. EMAAP-Q. Quito. Ecuador.

Bakker M, Vreeburg H.G, Van Schagen K.M. Rietveld L.C. (2013). A fully adaptive forecasting model for short-term drinking water Demand. *Environmental Modelling & Software* Vol.8, p. 141-151.

Box G., Jenkins G.M., Reinsel G. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco, California.

Campisi-Pinto S. Adamowski J. Oron G. 2012. Forecasting Urban Water Demand Via Wavelet-Denoising and Neural Network Models. Case Study: City of Syracuse, Italy. *Water Resources Management* (2012). 26.

CEDEX. 2012. Estudio de los Impactos del Cambio Climático en los Recursos Hídricos y las Masas de Agua. Efectos Potenciales del Cambio en las Demandas de Agua y Estrategias de Adaptación. Informe Técnico. Madrid. España.

Chen J. Boccelli D.L. (2014). Demand forecasting for water distribution systems. *Procedia Engineering* 70 (2014) 339 –

342. ELSIEVER.

Daza Sánchez F. (2008) *Demanda de Agua en Zonas Urbanas de Andalucía (Tesis Doctoral)*. Universidad de Córdoba. España.

Department of Environment and Conservation (ENVC). 2010. Newfoundland Labrador. Study on Water Quality and Demand on Public Water Supplies with Variable Flow Regimes and Water Demand. Canada. 2010.

Diao, K., Barjenbruch, M. and Bracklow, U. (2010) Study on the impacts of peaking factors on a water distribution system in Germany, *Water Science and Technology: Water Supply*, 10(2), pp. 165-172

EPMAPS – HAZEN and SAWYER. 2011. Estudios de Actualización del Plan Maestro Integrado de agua Potable y Alcantarillado para el DMQ. Quito. Ecuador.

EPMAPS (2008). Normas de Diseño de Sistemas de Agua Potable para la EMAAP-Q. Quito. Ecuador.

Herrera M. Torgo L. Izquierdo J. Pérez-García R. (2010). Predictive models for forecasting hourly urban water demand. *Journal of Hydrology* 387 (2010).

House-Peters, L. A., & Chang, H. (2011). Urban water demand modeling: Review of concepts, methods, and organizing principles. [Article]. *Water Resources Research*, 47, 15.

INEN (1992) Código Ecuatoriano de la construcción. Normas para estudio y diseño de sistemas de agua potable y disposición de aguas residuales para poblaciones mayores a 1 000 habitantes. Primera Edición.

Ministerio de Desarrollo Económico. (2000) Reglamento Técnico del sector de agua potable y saneamiento básico. RAS – 2000. Sección ii. Título b. Sistemas de Acueducto. Bogotá.

Nasseri M. Moeini A. Tabesh M. (2011). Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming. *Journal Expert Systems with Applications* 38 (2011) 7387–7395.

Yun Bai, Pu Wang, Chuan Li, Jingjing Xie, Yin Wang (2014). A multi-scale relevance vector regression approach for daily urban water demand forecasting. *Journal of Hydrology* 517 (2014) 236–245.