

PRONÓSTICO EFICIENTE DE LA DEMANDA DIARIA DEL SISTEMA ELECTRICO INTERCONECTADO DEL PERU MEDIANTE ANALISIS ESTOCASTICO ARIMA CON SUCESOS EXTERNOS

EFFICIENT DEMAND DAILY FORECAST OF THE PERUVIAN ELECTIC SYSTEM THROUGH ARIMA STOCHASTIC ANALYSIS WITH ATYPICAL EVENTS

Salome Gonzales Chávez

RESUMEN

La demanda diaria del Sistema Eléctrico Interconectado Nacional-SEIN, posee características muy peculiares de tendencia, estacionalidad y aleatoriedad, situación que complica al proceso de estimación de su pronóstico. El objetivo del presente trabajo consiste en formular y calcular modelos ARIMA con Análisis de Sucesos Externos, a fin de lograr pronósticos eficientes de la demanda eléctrica de cada día siguiente, a nivel total y desagregado por áreas. Un buen pronóstico de la demanda diaria garantiza el despacho eficiente y económico de generación y transmisión, así como el aseguramiento y calidad de la demanda sectorial nacional. El enfoque metodológico lo constituye el tratamiento de cada serie temporal objetivo, mediante transformaciones estadístico-matemáticas apropiadas para alcanzar estabilidad tanto en varianzas como en medias regulares y estacionales; paralelamente filtrar los sucesos externos hasta alcanzar a un Modelo ARIMA predictivo de cada área del sistema eléctrico del Perú (Centro, Sur y Norte) y para cada día de la semana. Los resultados alcanzados en la presente investigación demuestran la eficiencia predictiva comparativa. Es decir, tomando como indicador de calidad de pronóstico al Error Absoluto Promedio Porcentual (MAPE), se han obtenido valores inferiores al 1% en las proyecciones de la demanda diaria total del SEIN, frente al 2% que se logra con actuales técnicas determinísticas.

Palabras clave.- Pronóstico de Demanda, Despacho Eléctrico, ARIMA, Sucesos Externos, Serie Temporal, Proceso Estocástico, MAPE, Sistema Interconectado Nacional.

ABSTRAC

The daily electric demand in Peruvian National Interconnected System-SEIN- has very particular trend, seasonality and characteristics external effects, a situation that complicates the process of estimating the short-term forecast. The aim of this paper is to formulate and calculate ARIMA models with External Events Analysis to achieve efficient forecasts of electricity demand each day, at total level and broken down by areas of the SEIN. The methodology is based on treating each time series using appropriate statistical-mathematical transformations to achieve stability in variance as regular seasonal averages, parallel external events to try to reach an optimal predictive model ARIMA each area of the electrical system of Peru (Central, South and North) and for each day of the week. The results demonstrate the predictive efficiency. Taking as a quality indicator forecast the Mean Absolute Percent Error (MAPE), have obtained values lower than 1% by the projections of the total daily demand SEIN versus 2%

¹Dr. Docente investigador de la Facultad de Ingeniería Mecánica de la Universidad Nacional de Ingeniería, Lima-Perú.

obtained with existing deterministic techniques.

Key words.- Demand Forecast, Electricity Demand, ARIMA, External Events, Time Series, Stochastic Process, MAPE, Electric Grid System.

INTRODUCCION

OBJETIVOS Y FUNDAMENTACION DE LA HIPOTESIS

Estructurar, calcular y aplicar una nueva metodología cuantitativa mediante el enfoque de procesos estocásticos estacionarios, Modelos ARIMA, para la determinación del Pronóstico de la Demanda Eléctrica de Corto Plazo del Sistema Eléctrico Interconectado Nacional PDCP-SEIN. Esto es, pronosticar la demanda diaria del SEIN, con un sólido sustento metodológico matemático-estadístico y alto índice de aproximación frente a otras técnicas actuales utilizadas.

ANTECEDENTES Y PROCEDIMIENTO ACTUAL

En el ámbito nacional, no se tiene experiencias en el uso de los procesos estocásticos dentro del análisis de proyecciones de variables temporales energéticas, en específico en las proyecciones diarias de la demanda eléctrica nacional.

El método actual del pronóstico de la demanda diaria del SEIN, posee las siguientes restricciones:

- Toma como data de trabajo solamente cinco días históricos para cada día proyectado de la semana, situación que al no considerar la suficiente cantidad de información histórica, prescinde el patrón tendencial representativo de la serie diaria, distorsiona la componente estacional y, no captura irregularidades de nivel ni la presencia de atípicos.
- El cálculo del pronóstico para cada punto medio horario del día objetivo, se obtiene del producto de la mediana de los cinco valores específicos de la demanda en dicho punto horario, multiplicado por la tendencia lineal de los valores extremos del dominio diario desagregado en tres subdominios (valor mínimo de 0:30-7:30, valor máximo de 08:00-17:30 y valor máximo de 18:00-23:59), correspondientes a cada uno de los cinco días. Los valores específicos de cada punto se

obtienen del cociente entre la demanda en tal punto y su valor extremo correspondiente a tal subdominio. Ello genera errores estructurales, como: la mediana es un estadístico de posición central, más no necesariamente representa a valores específicos dependientes de mínimos y máximos de subdominios; la tendencia al ajustarse de forma lineal no captura el efecto irregular-estacional-tendencial representativo de los cinco grupos muestrales históricos.

PLANTEAMIENTO METODOLOGICO Y CÁLCULO

Planteamiento metodológico

La Demanda Eléctrica Diaria es una serie temporal que posee patrones de tendencia, estacionalidad, ciclaje e irregularidades de nivel, por lo que de origen no puede ser explicada mediante un proceso estocástico estacionario (Ref. 1). Entonces, para que esta serie sea evaluada de forma aproximada como un proceso estocástico estacionario lineal, previamente se le tiene que eliminar dichos patrones, ello mediante transformadas matemáticas apropiadas.

Un modelo ARIMA aplicado a una serie temporal, parte de un proceso estocástico no estacionario lineal homogenizado (Ref. 2); esto es, que luego de estabilizar su varianza mediante transformaciones potenciales, se le practica diferenciaciones regulares y estacionales (integraciones), hasta aproximarse a un proceso estacionario del tipo ARMA (Autoregressive-Moving Average), el cual viene a ser la combinación de dos procesos estocásticos estacionarios lineales puros uno autorregresivo y otro medias móviles. En el análisis estocástico de series temporales, se define a un proceso estocástico como la familia de variables aleatorias de $\{X_t\}$, donde t es el tiempo, tal que para cada serie finita de elecciones de t (t_1, t_2, \dots, t_n), se define una distribución de probabilidad conjunta para las correspondientes variables aleatorias $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ (Ref. 3). Así, una serie temporal X_t se define como el conjunto de valores observados de

distintas variables aleatorias correspondientes a períodos de tiempo consecutivos; dichos períodos tienen la misma amplitud y la serie tiene un carácter discreto. Es decir, el valor observado de la serie en el instante t puede ser considerado como una muestra aleatoria de tamaño uno de la variable X_t del proceso estocástico definida en dicho instante. Entonces se puede decir que X_t y $X_{t'}$ están separadas por k retardos sí $|t - t'| = k$.

Para aplicaciones prácticas, un proceso estocástico se puede describir mediante sus momentos, entre los cuales se destacan los siguientes:

La *media* de un proceso estocástico definida por:

$$\mu_t = E(X_t) \tag{1}$$

La *función de autocovarianza* (covarianzas entre variables referidas a momentos distintos en el tiempo), dada por:

$$x_{t,s} \equiv Cov(X_t, X_{t+k}) \equiv E\{[X_t - E(X_t)][X_{t+k} - E(X_{t+k})]\} \tag{2}$$

$; k = 0, 1, 2, \dots$

La *varianza* del proceso ($k = 0$), dada por:

$$x_{t,t} \equiv var X_t \equiv E(X_t - \mu_t)^2 = \sigma_t^2 \tag{3}$$

La *función de autocorrelación*, definida por:

$$\dots_{t,t+k} = \frac{x_{t,t+k}}{\sqrt{x_{t,t}} \sqrt{x_{t+k,t+k}}} \tag{4}$$

Un proceso estocástico estacionario se puede describir bajo dos sentidos (Ref. 4): estricto y amplio, quedando el segundo para su aplicación.

La estacionariedad en el sentido amplio se caracteriza mediante las siguientes propiedades: Las esperanzas matemáticas de las variables aleatorias no dependen del tiempo; es decir son constantes:

$$E(X_t) = E(X_{t+m}), \forall m; \text{ ó } \mu_t = \mu \quad \forall t \tag{5}$$

Las varianzas tampoco dependen del tiempo y son finitas, es decir:

$$Var(X_t) = Var(X_{t+m}) < \infty, \forall m; \text{ ó } \sigma_t^2 = \sigma^2, \forall t \tag{6}$$

Las covarianzas entre dos períodos de tiempo distintos solamente dependen del lapso de tiempo transcurrido entre estos dos períodos; es decir:

$$Cov(X_t, X_{t'}) = Cov(X_{t+m}, X_{t'+m}) \quad \forall m \tag{7}$$

$\text{ó } x_{t,s} = x_{|t-s|} \quad \forall t,s$

Para estas condiciones de estacionariedad, la autocorrelación de orden k (\dots_k) es la correlación separada k períodos de la misma serie temporal; esto es:

$$\dots_k = \frac{x_k}{x_0} = \frac{cov(X_t, X_{t+k})}{var(X_t)} \tag{8}$$

Al conjunto de autocorrelaciones obtenidas para distintos valores de k se le denomina *función de autocorrelación* (FAC), estadístico de gran utilidad en la identificación de modelos.

Se precisa trabajar con procesos estocásticos estacionarios y ergódicos (Ref. 5) para poder efectuar el proceso de inferencia consistente en, dada una serie temporal, inferir cuál es el proceso estocástico que ha podido generar dicha serie temporal. Para ello se han de estimar los parámetros que configuran las funciones de autocovarianza y de autocorrelación. La ergodicidad posibilita obtener estimadores consistentes de dichos parámetros; por cuanto, si el valor de \dots_k tuviera valores elevados para órdenes k altos, significaría que al aumentar el tamaño de la muestra disponible se añadiría poca información nueva para caracterizar adecuadamente el proceso; ello llevaría a que los estimadores obtenidos no serían consistentes.

Un proceso estocástico estacionario es ergódico en la media, μ , si es posible estimar consistentemente este parámetro haciendo uso de la *media muestral temporal*, que se define como:

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t \tag{9}$$

De forma análoga, se puede decir de ergodicidad respecto a la autocovarianza. Así la función de autocorrelación \dots_k se estima mediante la *función de autocorrelación muestral* (FACM) expresada como:

$$r_k = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_t (X_t - \bar{X})^2}, \quad k = 1, 2, \dots \quad (10)$$

La representación gráfica de r_k , se le denomina *correlograma muestral* y constituye un gran instrumento de análisis en el proceso de identificación del modelo ARIMA.

En general, ninguna serie temporal energética es estacionaria de origen, dado que posee varianza no constante así como medias no constantes en las componentes regulares (tendenciales) y estacionales. Sin embargo, es posible transformar tales series reales no estacionarias en otras aproximadamente estacionarias, ello mediante adecuadas transformaciones algebraicas.

Efectuar un pronóstico bajo el enfoque estocástico ARIMA, es inferir la distribución de probabilidad de una observación futura X_{T+1} dada una serie X_1, X_2, \dots, X_T de valores pasados. Para determinar las características del proceso estocástico subyacente a la serie temporal, se debe considerar como un *proceso estocástico lineal discreto* (Ref. 6).

Los casos particulares del proceso estocástico lineal discreto son: modelo de medias móviles de orden p , MA(p); modelo autorregresivo de orden q , AR(q), y modelo mixto autorregresivo- medias móviles de orden p, q , ARMA (p, q).

Modelo de medias móviles MA(q).- Se define mediante la expresión:

$$X_t = \mu + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q} \quad (11)$$

ó:

$$X_t = \mu + W(L)u_t \quad (12)$$

μ es la esperanza matemática de X_t , L es el operador de retardos y $W(L)$ es el operador polinomial de retardos, dado como:

$$W(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q \quad (13)$$

Un modelo de medias móviles siempre es estacionario, y será invertible cuando pueda expresarse como un proceso autorregresivo de orden infinito.

Para ello deberá cumplirse que las raíces de $W(L) = 0$, caigan fuera del círculo unitario (Ref. 2).

Como caso particular se tiene el Modelo MA(1), definido por:

$$X_t = \mu + W(L)u_t; \quad W(L) = 1 - \theta_1 L \quad (14)$$

El modelo MA(1) será siempre estacionario; y para que sea invertible deberá cumplirse que la raíz de la ecuación siguiente caiga fuera del círculo unitario (es decir que se cumpla $|\theta_1| < 1$):

$$W(L) = 1 - \theta_1 L = 0 \quad (15)$$

La función de autocorrelación de MA(1) tendrá la forma:

$$\dots r_k = \begin{cases} -\frac{\theta_1}{1 + \theta_1^2} & \text{para } k = 1 \\ 0 & \text{para } k > 1 \end{cases} \quad (16)$$

Modelo autorregresivo AR(p).- Un modelo autorregresivo de orden p se define como:

$$X_t = w_1 X_{t-1} + w_2 X_{t-2} + \dots + w_p X_{t-p} + u + u_t \quad (17)$$

ó:

$$W(L)X_t = u + u_t \quad (18)$$

$W(L)$ es el operador polinomial de retardos.

$$W(L) = 1 - w_1 L - w_2 L^2 - \dots - w_p L^p \quad (19)$$

A diferencia de los modelos de medias móviles que siempre son estacionarios, los autorregresivos deben cumplir como condición de estacionariedad que las raíces del polinomio característico $W(L) = 0$ caigan fuera del círculo unidad.

Este modelo siempre está en forma invertida. Particularmente se tiene el Modelo AR(1), que se expresa como:

$$W(L)X_t = u + u_t; \quad W(L) = 1 - w_1 L; \quad (20)$$

Modelo mixto ARMA (p, q).- Este modelo mixto autorregresivo (AR)-medias móviles (MA) de orden p, q , se define mediante:

$$X_t = w_1 X_{t-1} + w_2 X_{t-2} + \dots + w_p X_{t-p} + u_t + u_{t-1} - \dots - u_{t-q} \quad (21)$$

ó:

$$w(L)X_t = u + \theta(L)u_t \quad (22)$$

El modelo ARMA se dice que es estacionario cuando lo es su parte autorregresiva AR; esto es, cuando las raíces de la ecuación $w(L) = 0$ caen fuera del círculo unidad, y se dice que es invertible cuando lo es su parte MA; esto es, cuando las

raíces de la ecuación $\theta(L) = 0$ caen fuera del círculo unidad. Adicionalmente a las condiciones de estacionalidad e invertibilidad también se supondrá que las raíces de $w(L) = 0$ y $\theta(L) = 0$ no son comunes.

Las funciones teóricas de autocorrelación (FAC) y de autocorrelación parcial (FACP), (ver Figura 1), sirven como patrones de referencia para identificar las funciones de autocorrelación regular y parcial muestrales de la serie temporal en estudio.

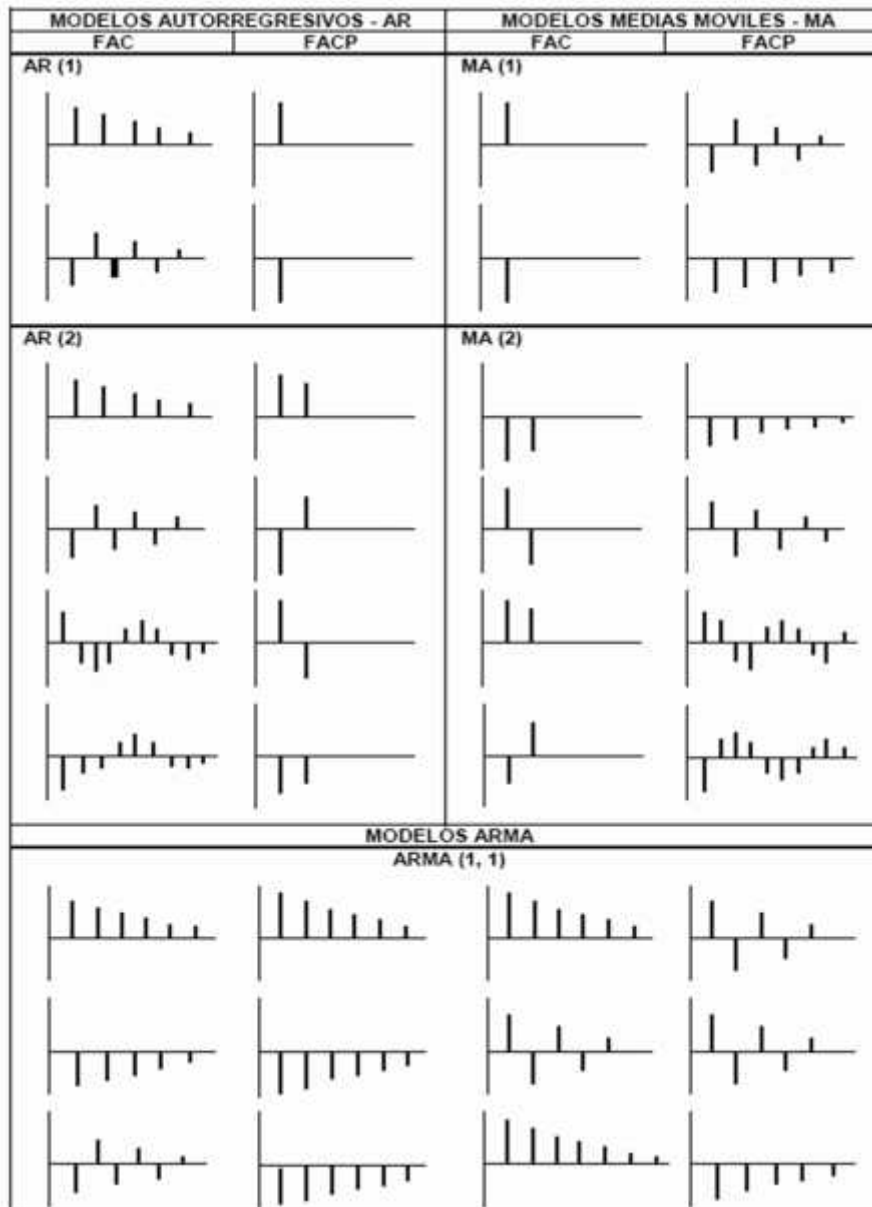


Fig. 1 Funciones teóricas de autocorrelación regular (FAC) y parcial (FACP).

FUNDAMENTACION MATEMATICA DE UN MODELO ARIMA

Modelos lineales no estacionarios homogéneos

Un proceso estocástico no estacionario es homogéneo cuando al diferenciar en el proceso original, el proceso transformado resultante es estacionario, y el número de veces que debe diferenciarse el proceso original para transformarse en estacionario constituye el grado u orden de homogeneidad o integración. Entonces una serie original X_t se puede convertir en aproximadamente estacionaria, al aplicarle diferencias en una o más órdenes; es decir:

$$\Delta^d X_t = (1-L)^d X_t ; t = 1, 2, \dots, T \quad (23)$$

A un proceso integrado X_t se le denomina proceso autorregresivo-medias móviles integrado, ARIMA(p, d, q), si tomando diferencias de orden d se obtiene un proceso estacionario Z_t del tipo ARMA(p, q).

El modelo ARIMA(p, d, q), se expresa de la siguiente forma:

$$Z_t = W_1 Z_{t-1} + W_2 Z_{t-2} + \dots + W_p Z_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad (24)$$

ó:

$$w(L)Z_t = \theta(L)u_t ; Z_t = (1-L)^d X_t \quad (25)$$

Finalmente quedando como ARIMA (p, d, q):

$$w(L)(1-L)^d X_t = \theta(L)u_t \quad (26)$$

No se incluye el término constante u dado que la media de la serie diferenciada Z_t es cero, como frecuentemente suele ocurrir. En caso contrario este parámetro se incluye en la expresión (26)

Al analizar la mayoría de las series temporales reales, se suele observar que éstas presentan una tendencia creciente o decreciente. La eliminación de esta tendencia (no estacionariedad en media) de la serie se suele conseguir mediante diferenciaciones implícitas en los modelos ARIMA. En ocasiones se observa también que existe una tendencia en la varianza, esto es, que la dispersión de las observaciones no es constante a

lo largo del tiempo (Ref. 7 y 5), la cual no se elimina mediante estas diferenciaciones. Cuando se presenta este hecho, se aplica la transformación Box-Cox (Ref. 6), quedando el modelo como:

$$w(L)(1-L)^d X_t^{(\lambda)} = \theta(L)u_t + u \quad (27)$$

Siendo:

$$X_t^{(\lambda)} = \begin{cases} \frac{X_t^{(\lambda)}}{\lambda} & \text{para } \lambda \neq 0 \\ \ln X_t & \text{para } \lambda = 0 \end{cases} \quad (28)$$

Modelos estacionales no estacionarios homogéneos. - Otra fuente de no estacionariedad en muchas de las series reales del ámbito energético lo constituye la estacionalidad (con periodicidad s). Para desestacionalizar las series, se procede a la *diferenciación estacional*.

Los modelos estacionales no estacionarios pero homogéneos, ARIMA(P,D,Q), se expresan como:

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_p Z_{t-ps} + u_t - \theta_1 u_{t-s} - \dots - \theta_q u_{t-qs} \quad (29)$$

$$Z_t = \Delta_s^D X_t = (1-L^s)^D X_t \quad (30)$$

La expresión resumida de ARIMA(P,D,Q) es:

$$\Phi_p(L^s)(1-L^s)^D X_t = \theta_q(L^s)u_t \quad (31)$$

Donde $\Phi_p(L^s)$ y $\theta_q(L^s)$, son los polinomios de retardos autoregresivo estacional y media móvil estacional, respectivamente.

$$\begin{aligned} \Phi_p(L^s) &= 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps} \\ \theta_q(L^s) &= 1 - \theta_1 L^s - \theta_2 L^{2s} - \dots - \theta_q L^{qs} \end{aligned} \quad (32)$$

Modelo multiplicativo general. - Los modelos que conjugan ambos tipos de interdependencias entre las observaciones regulares y las estacionales, son los modelos multiplicativos generales, los mismos que se denotan abreviadamente como ARIMA(p,d,q)xARIMA(P,D,Q). Se expresan como:

$$\begin{aligned} \Phi_p(L^s)w_p(L)(1-L^s)^D(1-L)^d X_t = \\ \theta_q(L^s)\theta_q(L)u_t + u \end{aligned} \quad (33)$$

Análisis de sucesos externos.- Una serie temporal real puede contener ciertos Sucesos Externos que afectan singularmente a la normal característica evolutiva de la serie; estos sucesos son denominados intervenciones y atípicos. En la serie demanda diaria, las intervenciones pueden ser picos generados días festivos, racionamientos, interrupciones, etc. Los atípicos, son suceso puntuales desconocidos que pueden ser originados por ejemplo por errores de medición y señal SCADA, errores de contabilidad, etc.

Una de las formas de detección de los posibles sucesos externos (intervenciones o atípicos), se realizan a través de la visualización de la serie objetivo en niveles o estacionarizada, y con mayor precisión observando el comportamiento de los residuos del correspondiente modelo ARIMA previamente calculado. Una vez detectados y ubicados los sucesos externos, dependiendo de su característica, se procede a su filtrado utilizando variables ficticias tales como pulso, escalón,

tendencias, etc. Luego se aplica al modelo ARIMA previamente seleccionado, obteniendo así el modelo ARIMA definitivo de proyección. La estructura matemática con la que se modela el filtrado de los Sucesos Externos, se presenta de forma sintetizada en el ANEXO

PROCESO DE CÁLCULO

Partiendo de una determinada serie temporal se trata de averiguar qué modelo ARIMA(p,d,q) x ARIMA(P,D,Q) es susceptible de haber generado dicha serie, es decir, qué modelo representa adecuadamente el comportamiento de la misma, con el fin de utilizarlo para obtener predicciones de valores futuros de la serie en estudio. Para ello se siguen cuatro etapas: identificación, estimación, validación y predicción. En la Figura 2, se ha confeccionado un diagrama de flujo que resume el proceso a seguir para el cálculo de pronósticos mediante modelos ARIMA.

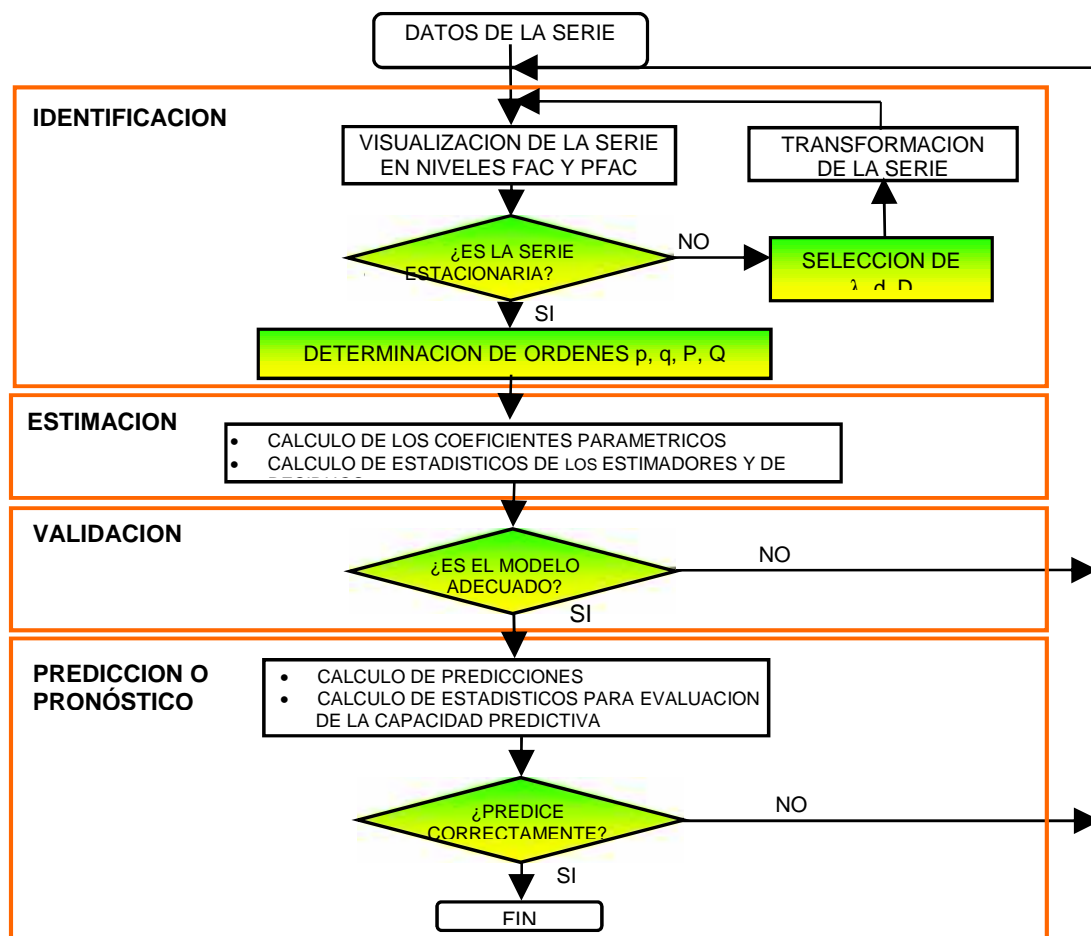


Fig. 2 Diagrama de flujo del procedimiento de cálculo y análisis ARIMA.

PROPUESTA DE SOLUCION. ANALISIS DE RESULTADOS

Cada día de la semana posee características particulares de demanda eléctrica en resolución media horaria, en magnitudes, tendencias, estacionalidades y aleatoriedades; por tanto se conforma una serie histórica para cada día de la semana.

La representatividad del comportamiento evolutivo de la data diaria lo marca el tamaño del horizonte

histórico; entonces se opta por un tamaño muestral móvil de 51 semanas.

El tratamiento de la información histórica de la demanda diaria se realiza a partir de los informes de evaluación diarios de operación del COES, agrupado por Áreas de influencia: Centro, Sur y Norte. En la Figura 3 se ilustra el comportamiento histórico de la demanda diaria por áreas en un determinado día, donde se observa la presencia de sucesos externos en cada caso.

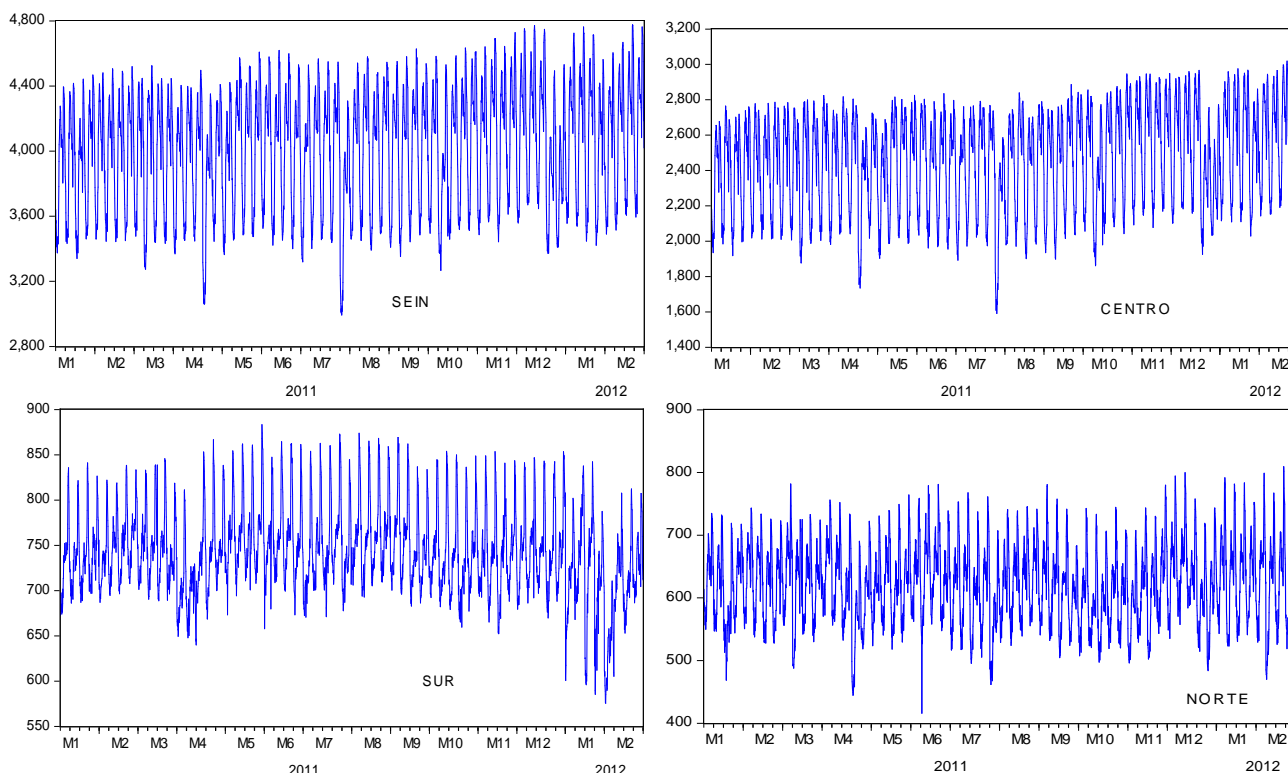


Fig. 3 Series muestrales del SEIN y Áreas Centro, Sur y Norte, de un día de la semana.

Identificación del modelo. Eligiendo la serie histórica de un día de la semana, se procede a identificar la estacionariedad de dicha serie.

Con las funciones de Autocorrelación Regular y Parcial, se comprueba la no estacionariedad, por lo que se efectúa su transformación en una nueva serie tal que cumpla la estabilidad, tanto en varianza como en medias regular y estacional. Una vez alcanzada la condición de estacionariedad, se identifica las órdenes autorregresivas regular (p) y estacional (P), así como las órdenes medias móviles regular (q) y estacional (Q).

Estimación de parámetros.- La estimación de los parámetros del modelo ARIMA calculado (W_i, n_i, Φ_i y Θ_i) se realiza por el método de mínimos cuadrados ordinarios. Asimismo se comprueba si los parámetros calculados son estadísticamente significantes, mediante la probabilidad asociada al estadístico *t*-Student (p-valor), para así contrastar la hipótesis nula de que el parámetro correspondiente es igual a cero.

Tomando un día cualquiera de la semana dentro del periodo de demostración de la propuesta, p.e. el pronóstico de un sábado 03 marzo 2012, luego

del proceso de identificación y estimación se obtiene que el modelo ARIMA, que describe el comportamiento predictivo de la demanda total SEIN del sábado siguiente, posee las siguientes características:

- ARIMA (0,1,2)x(0,1,1)
- Transformación: logaritmo natural
- Sucesos externos: sábado 23 de abril (efecto de Semana Santa 2011), sábado 30 de julio (efecto de fiestas patrias 2011), sábado 08 de

octubre 2011 (día feriado), sábado 24 de diciembre 2011 (efecto de fiestas de navidad), sábado 31 de diciembre 2011 (efecto de fiestas de fin de año) y sábado 28 de enero 2012 (efecto de averías del sistema por condiciones meteorológicas).

La ecuación de estimación del Modelo ARIMA, calculado en Programa Estructurado E-Views (Ref. 8), así como en Programa IBM SPSS (Ref. 9), tiene la siguiente forma (ver Tabla 1):

Tabla 1 Salida en E-Views de las estimaciones de los parámetros ARIMA.

Dependent Variable: D(LOG(SEINSA),1,48)				
Method: Least Squares				
Sample (adjusted): 1/15/2011 01:00 2/26/2012 00:00				
Included observations: 2831 after adjustments				
Convergence achieved after 7 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SA23ABR2011)	-0.138333	0.010192	-13.57226	0.0000
D(SA30JUL2011)	-0.161099	0.010198	-15.79711	0.0000
D(SA08OCT2011)	-0.018952	0.010192	-1.859490	0.0631
D(SA24DIC2011)	-0.073284	0.010309	-7.108733	0.0000
D(SA31DIC2011)	-0.023356	0.010346	-2.257455	0.0241
D(SA28ENE2012)	-0.053535	0.010281	-5.207002	0.0000
MA(1)	-0.320714	0.018670	-17.17800	0.0000
MA(2)	-0.152498	0.018664	-8.170745	0.0000
SMA(48)	-0.730552	0.012834	-56.92347	0.0000
R-squared	0.518586	Mean dependent var		-2.22E-05
Adjusted R-squared	0.517221	S.D. dependent var		0.017233
S.E. of regression	0.011974	Akaike info criterion		-6.009003
Sum squared resid	0.404602	Schwarz criterion		-5.990092
Log likelihood	8514.743	Hannan-Quinn criter.		-6.002180
Durbin-Watson stat	1.991141			

El modelo ARIMA calculado, posee la siguiente estructura en su versión amplia:

$$(1 - L)(1 - L^{48}) \log SEINSA_t = (1 - \alpha_1 L - \alpha_2 L^2)(1 - \Theta_1 L^{48}) u_t + (1 - L)(c_1 * SA_{23 ABR 2011} + c_2 * SA_{30 JUL 2011} + c_3 * SA_{08 OCT 2011} + c_4 * SA_{24 DIC 2011} + c_5 * SA_{31 DIC 2011} + c_6 * SA_{28 ENE 2012})$$

(34)

Donde: $\alpha_1 = -0.320$; $\alpha_2 = -0.152$; $\Theta_1 = -0.730$; $c_1 = -0.138$; $c_2 = -0.161$; $c_3 = -0.018$; $c_4 = -0.073$; $c_5 = -0.023$; $c_6 = -0.053$

El pronóstico de la demanda eléctrica SEIN total del día sábado (SEINSA), se expresa:

$$\log SEINSA_t = \log SEINSA_{t-1} + \log SEINSA_{t-48} - \log SEINSA_{t-49} + u_t - \alpha_1 u_{t-1} - \alpha_2 u_{t-2} - \Theta_1 u_{t-48} + \alpha_1 \Theta_1 u_{t-49} + \alpha_2 \Theta_1 u_{t-50} + c_1 * SA_{23 ABR 2011}_t + c_2 * SA_{30 JUL 2011}_t + c_3 * SA_{08 OCT 2011}_t + c_4 * SA_{24 DIC 2011}_t + c_5 * SA_{31 DIC 2011}_t + c_6 * SA_{28 ENE 2012}_t - c_1 * SA_{23 ABR 2011}_{t-1} - c_2 * SA_{30 JUL 2011}_{t-1} - c_3 * SA_{08 OCT 2011}_{t-1} - c_4 * SA_{24 DIC 2011}_{t-1} - c_5 * SA_{31 DIC 2011}_{t-1} - c_6 * SA_{28 ENE 2012}_{t-1}$$

(35)

Siendo la expresión del error:

$$u_t = SEINSA_t - \hat{SEINSA}_t$$

Si T es el instante correspondiente a la última observación histórica de la variable *SEINSA* (p.e. sábado 25 de febrero 2012 a las 23:59), entonces el pronóstico de cada media hora inicia en $t = T+1$, es decir sábado 03 de marzo 2012 a las 00:30 y finaliza en $t = T+48$, esto es el sábado 03 de marzo 2012 a las 23:59 (48 valores pronosticados a intervalo de media hora).

VALIDACION, PRONÓSTICO Y CAPACIDAD PREDICTIVA

Una de las formas de validación de la calidad del pronóstico, consiste en comprobar si se satisface la hipótesis relativa de que los residuos del modelo (u_t) obedecen al comportamiento de un ruido blanco; esto es, comprobar que la serie error u_t posee media nula, es estable en varianza y obedece a una distribución normal.

La capacidad predictiva, calidad o bondad de ajuste de un modelo ARIMA, se logra minimizando el estadístico Error Promedio Absoluto Porcentual (MAPE). Otra de las formas de evaluar la calidad predictiva del modelo ARIMA, es la comprobación ex-post del MAPE entre lo ejecutado y lo proyectado, y su comparación con el procedimiento tradicional.

A continuación se muestra los resultados proyectivos diarios para un día de la semana (caso sábado 03 marzo) de la demanda del SEIN total y por áreas Centro, Sur y Norte (Figuras 4 a 7).

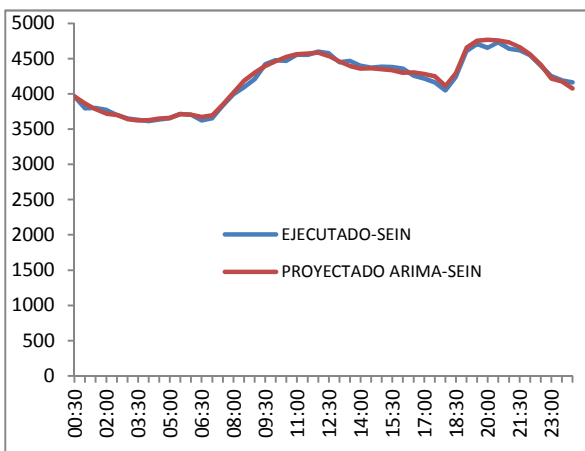


Fig. 4 Comparación pronóstico ARIMA vs Ejecutado, SEIN Total.

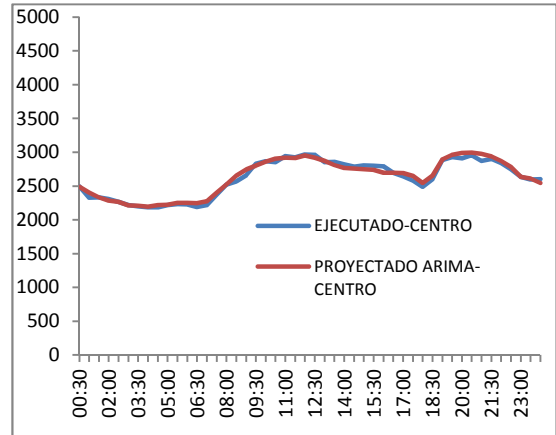


Fig. 5 Comparación pronóstico ARIMA vs Ejecutado, CENTRO.

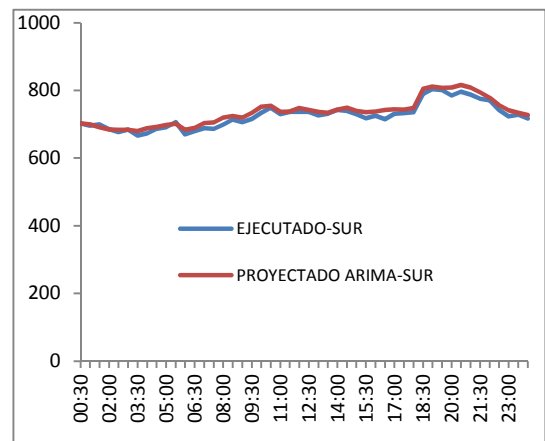


Fig. 6 Comparación pronóstico ARIMA vs Ejecutado, SUR.

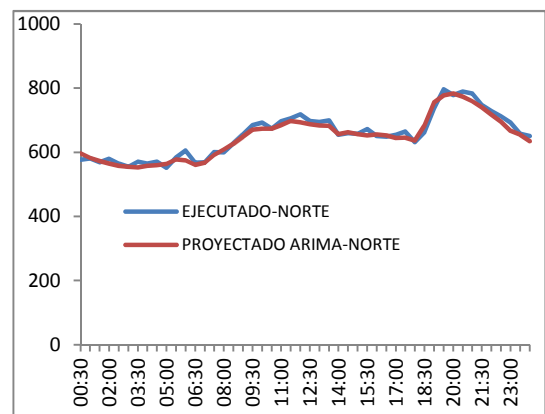


Fig. 7 Comparación pronóstico ARIMA vs Ejecutado, NORTE-

En la Tabla 2, se resume la comparación de la calidad predictiva entre la metodología propuesta y

el actual procedimiento, teniendo como estadístico de error al MAPE.

Tabla 2. Comparación de eficiencia predictiva.

SEIN Y AREAS	DÍA PROYECTADO	
	MAPE Propuesto ARIMA (%)	MAPE Procedimiento Tradicional (%)
SEIN Total	0.8	2.4
Área Centro	1.3	2.3
Área Sur	1.5	2.1
Área Norte	1.6	2.7

CONCLUSIONES Y RECOMENDACIONES

Se han elaborado estructuras de modelos ARIMA típicos para el pronóstico de la demanda de cada día de la semana, correspondientes a SEIN Total y cada Área: Centro, Sur y Norte. El tamaño del horizonte histórico elegido para cada día a proyectar, tiene característica móvil y obedece a patrones de tendencia, estacionalidad e irregularidades, propios de cada uno de los siete días de la semana.

Los resultados alcanzados demuestran la calidad del pronóstico, tanto por los bajos valores de estadísticos de error cuantificados, como por el cumplimiento de los condicionantes estadísticos. Para la demostración de la calidad predictiva de los modelos ARIMA formulados, se han realizado cálculos de pronósticos de la demanda eléctrica diaria del SEIN y AREAS a lo largo del 2012. Como indicador de la calidad de pronóstico se toma al MAPE, cuyos valores se han comparado tanto a nivel ex-post con los ejecutados diarios, como con los obtenidos con el procedimiento tradicional. De las evaluaciones, tomando un día cualquiera (p.e. sábado 03 marzo 2012), el MAPE para la proyección de la Demanda SEIN total mediante la nueva propuesta es 0.8%, mientras que con el procedimiento tradicional alcanza 2.3%. De forma similar, para las proyecciones diarias por Áreas, los valores MAPE de la propuesta, son menores que los del procedimiento actual.

- Dada las ventajas del nuevo modelo propuesto de pronóstico de la demanda eléctrica diaria, su aplicación se hace extensiva para proyectar demandas desagregadas diversas, tales como barras representativas, de distribución, cargas especiales, etc.
- Otra de las aplicaciones del modelo propuesto, son los pronósticos diarios de demandas de sistemas aislados del Perú, optimización de ventas eléctricas de distribuidores, etc.

REFERENCIAS

- 1) **Makridakis, S., Wheelwright, S. C.** "Forecasting Methods for Management". John Wiley, New York, 1989.
- 2) **Box, G. E. P., Jenkins, G. M.** "Time Series Analysis: Forecasting and Control". Holden Day, San Francisco, 1976.
- 3) **Parzen, E.,** "Procesos Estocásticos". Paraninfo, Madrid, 1972.
- 4) **Prawda, J.,** "Métodos y Modelos de Investigación de Operaciones, Modelos Estocásticos". Limusa, Madrid, 1980.
- 5) **Jenkins, G. M.** "Practical Experiences with Modelling and Forecasting time Series". Gwilym Jenkins and Partners, San Francisco, 1979.
- 6) **Pankratz, A.** "Forecasting with Univariate Box-Jenkins Model: Concepts and Cases". John Wiley, New York, 1983.
- 7) **Xiberta, J., Gonzáles, S.** "Energetic Plan of Asturias, Spain". Oviedo University, Spain, 1996.
- 8) **Quantitative Micro Software, E-VIEWS.** "Estimation, Forecasting, Advanced Statistical". CA USA, 2012
- 9) **IBM SPSS,** "Statistical Product and Service Solutions. Forecasting, Neural Networks". NY USA, 2012

E-mail: salome@uni.edu.pe

Aceptación de documentos: enero 2014

Aceptación de documentos: abril 2014

ANEXO

ESTRUCTURA MATEMÁTICA SINTETIZADA PARA FILTRADO DE LOS SUCEOS EXTERNOS

Una expresión que generaliza a los modelos de **intervención**, Y_t , tiene la forma siguiente:

$$Y_t = v(L) I_t^{t_0} ; v(L) = \frac{\check{S}(L)L^b}{u(L)}$$

$$\check{S}(L) = \check{S}_0 + \check{S}_1 L + \check{S}_2 L^2 + \dots \check{S}_s L^s$$

$$u(L) = 1 - u_1 L - u_2 L^2 - \dots - u_r L^r$$

$I_t^{t_0}$: Variable ficticia impulso (vale 1 en el periodo de ocurrencia del suceso, $t=t_0$, y 0 en los demás periodos) o escalón (vale 0 antes del suceso, $t < t_0$, y 1 después del suceso).

$v(L)$: *Función de respuesta* a la variable impulso $\check{S}_0, \check{S}_1, \dots, \check{S}_s; u_1, u_2, \dots, u_r$: parámetros desconocidos a estimar, también denominados factores de respuesta a la variable ficticia (en este caso impulso).

Los sucesos **atípicos** que pueden estar involucrados en una serie temporal, de acuerdo a las huellas que dejan en la serie, se pueden identificar y modelar bajo las siguientes formas:

Atípico aditivo: AO. Es un suceso externo que afecta a una serie temporal en un solo instante, de forma tal que se puede expresar como:

$$Y_t = Z_t + \check{S} I_t^{t_0}$$

$I_t^{t_0}$: variable impulso, que representa la presencia o ausencia del atípico en el tiempo t_0 y está definida por:

$$I_t^{t_0} = \begin{cases} 1 & \text{si } t = t_0 \\ 0 & \text{si } t \neq t_0 \end{cases}$$

\check{S} : es el efecto de dicho atípico, o factor de respuesta al impulso.

Z_t : obedece a un modelo ARIMA, libre de atípicos.

$$Z_t = \frac{\Theta_q(L) \omega_q(L)}{\Phi_p(L) \omega_p(L) (1-L^s)^D (1-L)^d} u_t$$

Atípico innovacional: IO. Es un suceso cuyo efecto se propaga en conformidad con el modelo ARIMA del proceso, afectando a todos los valores observados después de su ocurrencia. Se expresa como:

$$Y_t = Z_t (u_t + \check{S} I_t^{t_0})$$

Cambio de nivel: LS. Un cambio en nivel es un suceso que afecta a una serie en un tiempo t_0 , a partir del que su efecto es permanente. Se puede expresar bajo dos formas:

$$Y_t = Z_t + \frac{1}{1-L} \check{S} I_t^{t_0}$$

$$Y_t = Z_t + \check{S} S_t^{t_0}$$

Siendo, $I_t^{t_0}$ la variable impulso y $S_t^{t_0}$ la variable escalón definida por:

$$S_t^{t_0} = \begin{cases} 1 & \text{si } t \geq t_0 \\ 0 & \text{si } t < t_0 \end{cases}$$

Cambio temporal: TC. Este suceso que tiene un impacto inicial, cuyo efecto decae en forma exponencial de acuerdo con un factor de amortiguación u . Se expresa como:

$$Y_t = Z_t + \frac{1}{1-uL} \check{S} I_t^{t_0} ; 0 < u < 1$$