

CARACTERIZACIÓN DEL CONSUMO DE ENERGÍA ELÉCTRICA EN FUNCIÓN DEL TIEMPO: UN ENFOQUE MULTIVARIADO*

Lydia Lera, ICIMAF, CITMA. INTA, Universidad de Chile
Minerva Montero y Pedro González, ICIMAF, CITMA

RESUMEN

Durante los meses en los que las condiciones ambientales presentan mayor diferencia (enero y junio, 1995) se tomaron datos del consumo de energía eléctrica del hotel Sevilla con el fin de determinar acciones para su ahorro. El consumo de energía eléctrica se analizó en relación con la temperatura ambiental. Con el objetivo de buscar grupos que caracterizaran el consumo eléctrico según la hora del día, se aplicaron técnicas multivariadas, entre las que se encuentran la Taxonomía Numérica o Análisis de Conglomerados ("Cluster Analysis") (Crivisqui, 1999) y el Método STATIS (Structuration des Tableaux A Trois Indices de la Statistique) (Lavit, 1988; Grupo GERI, 1996).

ABSTRACT

During the two months that the environmental conditions differ the most (January and June, 1995) data on the consumption of electrical power in Hotel Sevilla were obtained. The consumption of electrical power was analyzed in relation with the environmental temperature to determine the appropriate actions towards its saving. With the objective of looking for groups characterizing the electrical consumption according to the hour of the day, multivariate techniques were applied, such as Cluster Analysis (Crivisqui, 1999) and the STATIS (Structuration des Tableaux A Trois Indices de la Statistique) Method (Lavit, 1988; Grupo GERI, 1996).

Key words: common structure, cluster analysis, STATIS.

MSC: 62P30.

1. INTRODUCCIÓN

En el hotel Sevilla se llevó a cabo un proyecto cuyo objetivo central era optimizar el consumo de energía sin afectar las prestaciones, entre ellas, el confort que requiere una instalación de ese tipo. Para ello era necesario caracterizar, entre otros aspectos a estudiar, el consumo energético diario en función del tiempo.

Para el estudio se seleccionaron los meses de enero y junio, como representantes de las temporadas de invierno y verano. Junio tiene además, la característica interesante de ser muy lluvioso y esto hace que la temperatura baje repentinamente, lo cual pone al predictor en condiciones difíciles. Se midió el consumo de energía cada 15 minutos (96 registros), comenzando a las 00:00 horas (González, 1996, 1997, 1998). Los datos de temperatura ambiente se tomaron de los registros del Instituto de Meteorología.

Para el análisis de los datos se emplearon dos métodos multivariados: el método de "Cluster Analysis" o método de Taxonomía Numérica, que se utiliza para agrupar observaciones o variables en clases, sobre la base de sus relaciones y/o atributos que tienen en común; y el método STATIS, que es un método de tipo factorial cuyo objetivo fundamental es la búsqueda de una estructura común entre diferentes conjuntos de datos. Estos datos se organizan en matrices cúbicas o triples (tablas de 3 entradas donde una o más de las entradas puede ser el tiempo). Las tres entradas de las dos matrices cúbicas empleadas en este trabajo están determinadas por los meses (enero y junio), los días del mes y los registros de consumo de energía o temperatura.

2. MATERIALES Y MÉTODOS

Para caracterizar el consumo de energía diario se utilizó el método de Taxonomía Numérica. Con vistas a analizar las características comunes de los dos meses se aplicó la técnica factorial STATIS. A continuación se hace una breve descripción de los métodos utilizados:

*Este trabajo fue presentado en el III Taller Internacional sobre Calidad y Desarrollo, QUALITAS'99, La Habana, Cuba, 1999.

Cluster Analysis

El propósito de este método es agrupar las observaciones de forma que los individuos sean homogéneos dentro de los grupos y que estos grupos sean lo más heterogéneos posible entre ellos. Las distancias entre los individuos se definen a través de una medida de distancia. En este trabajo se utilizó como medida de disimilitud la distancia euclidiana. En el primer paso, cada individuo representa su propio grupo. Una vez que se han determinado los pares de individuos más cercanos, las distancias entre los nuevos grupos se determinan a través de diferentes reglas de agrupación. Las utilizadas en este trabajo son:

Ligamiento completo: la distancia entre los grupos se determina por la mayor distancia entre dos individuos cualquiera en los diferentes grupos.

Ligamiento simple: la distancia entre dos grupos se determina por la distancia de los dos individuos más cercanos en los diferentes grupos.

Ligamiento promedio entre grupos: la distancia entre dos grupos se calcula como la distancia promedio entre todos los pares de individuos en los dos grupos diferentes.

Método Ward: este método usa un enfoque de análisis de varianza para evaluar la distancia entre los grupos y tiene como estrategia minimizar la suma de cuadrados dentro de los grupos que se pueden formar en cada paso.

Método Statist

El método STATIS es un método exploratorio del Análisis de Datos cuyo principal objetivo es la búsqueda de una estructura común entre k grupos que pueden estar formados por los mismos individuos a los que se les han hecho mediciones a través del tiempo o por individuos diferentes a los que se les miden las mismas variables.

Sea X_k la tabla (matriz de datos) formada por n filas y p columnas en la k -ésima ocasión, entonces $W_k = X_k X_k^T$ describe la estructura de los individuos en el interior de las tablas. Esta matriz de dimensión $n \times n$ contiene los productos escalares entre los n individuos de la tabla k -ésima, y viene a ser lo que se conoce como infraestructura. Contrariamente a este término, se llama interestructura a la relación entre las tablas está definida por las distancias entre las matrices W_k . El método STATIS utiliza el producto escalar de Hilbert-Schmidt (H-S), para definir una distancia entre dos matrices W_{k_1} y W_{k_2} . El producto escalar de H-S se define como:

$$\langle W_{k_1}, W_{k_2} \rangle = \text{tr}(D W_{k_1} D W_{k_2})$$

Este producto equivale a la suma de los cuadrados de las covarianzas entre todas las parejas de variables, una en la ocasión k_1 y otra, en la ocasión k_2 . Luego, los elementos del producto escalar de H-S son los coeficientes de la covarianza vectorial entre las matrices W_k , la matriz resultante, de dimensión $t \times t$, se representa por:

$$S = \begin{bmatrix} \ddots & & \ddots \\ \cdots & \langle W_{k_1}, W_{k_2} \rangle & \cdots \\ \ddots & & \ddots \end{bmatrix}$$

Si los pesos de los individuos son iguales

$$D = (1/n)I_n \text{ y } S_{k_1 k_2} = \text{tr}(D W_{k_1} D W_{k_2}) = (1/n^2) \text{tr}(X_{k_1} X_{k_1}^T X_{k_2} X_{k_2}^T)$$

Entonces en función de la norma de W_k , $\|W_k\| = S_{kk}^{1/2}$, se pueden calcular los coeficientes RV (correlación vectorial) como:

$$RV(k_1, k_2) = \frac{S_{k_1 k_2}}{S_{k_1 k_1}^{1/2} S_{k_2 k_2}^{1/2}}$$

Este coeficiente es importante para la interpretación de la interestructura y se interpreta como la evolución promedio de los individuos en las ocasiones k_1 y k_2 . Un coeficiente RV cercano a 1 significa que existe la misma estructura de los individuos en el interior de las tablas X_{k1} y X_{k2} , y que las posiciones de los individuos son estables. Si RV es 0, entonces son ortogonales. Para el caso de las variables el coeficiente de similitud es el coeficiente de correlación de Pearson. El método es como una generalización del Análisis de Componentes Principales y brinda la posibilidad de representaciones gráficas. Este método sigue un enfoque descriptivo, centrándose en la estructura algebraica de los datos sin tener en cuenta supuestos distribucionales.

Para la aplicación del método STATIS en este estudio se utilizaron programas realizados en MATLAB (Lacourly, 1997). El método STATIS se encuentra programado en TURBO PASCAL en Lavit (1988) y existe el programa ADE-4 (Thioulouse, J. et al., 1996), que es posible bajarlo en la dirección de internet <http://pbil.univ-lyon1.fr/ADE-4/>.

Para los datos de temperatura se utilizó la siguiente transformación:

$$t = (\text{temperatura} - \text{temperatura (mínima)}) / (\text{temperatura (máxima)} - \text{temperatura (mínima)})$$

considerando la temperatura mínima como 10 grados centígrados y la máxima 35. El consumo se midió en kw/h .

3. RESULTADOS Y DISCUSIÓN

Para analizar la tendencia del consumo en el transcurso del día se hicieron gráficos de perfiles para los meses analizados (Figuras 1 y 2). Se puede observar que, en los dos meses, los mayores consumos de energía se encuentran en el horario del día. En el mes de enero se destaca una mayor variación del consumo diario.

Mediante la aplicación del método de Taxonomía Numérica se obtuvieron grupos que caracterizan el consumo eléctrico según la hora del día. Se utilizaron las 4 técnicas mencionadas anteriormente y se decidió tomar las agrupaciones que más se repetían, obteniéndose 5 grupos. A modo de ilustración, las Figuras 3 y 4 muestran los dendogramas obtenidos para el método ligamiento completo.

El análisis de las observaciones de cada uno de los grupos obtenidos permitió la siguiente caracterización del consumo en función del tiempo:

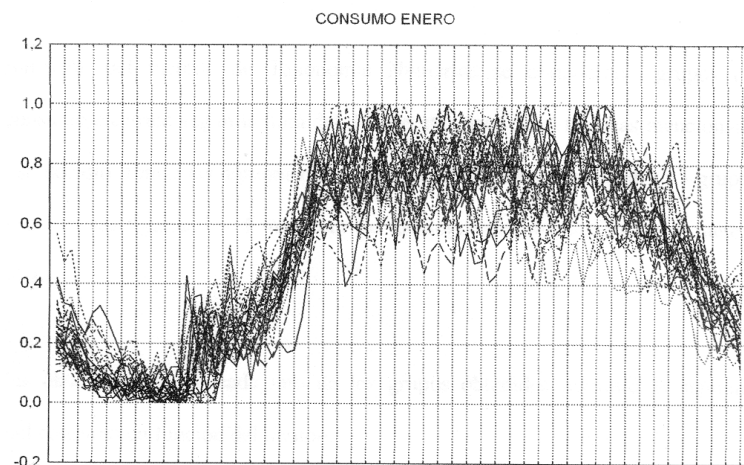


Figura 1. Perfiles diarios de consumo del mes de enero.

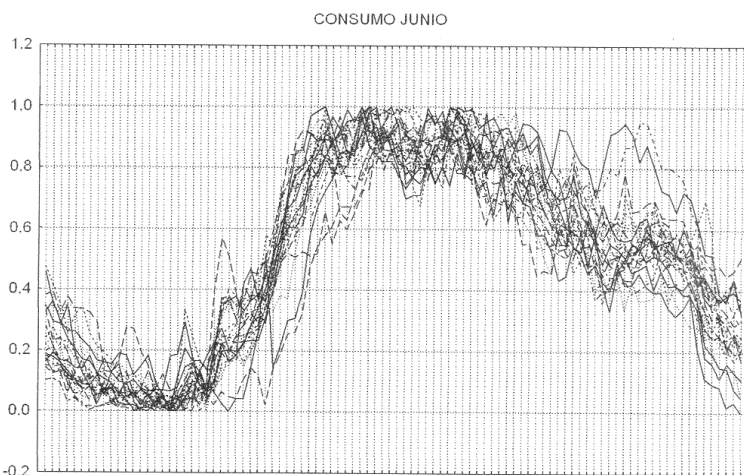


Figura 2. Perfiles diarios de consumo del mes de junio.

GRUPO I	GRUPO II	GRUPO III	GRUPO IV	GRUPO V
11:00 pm – 1:30 am	1:45 – 5:45 am	8:00 – 9:15 am	9:30 am – 2:45 pm	3:00 – 5:15 pm
6:00 – 7:45 am		5:30 – 10:45 pm		

Los grupos V y IV son los de mayores consumos, lo que es de esperar ya que se corresponde con los períodos donde la temperatura es más elevada. Esto coincide con lo que se observa en los perfiles (Figuras 1 y 2). Le sigue el grupo III. Los grupos II y I son los que presentan menor consumo, lo que también es de esperar, ya que el grupo II abarca la madrugada y el I parte de la medianoche y el amanecer.

A continuación se presentan las medias y desviaciones estándar del consumo total por grupo:

		GRUPO I	GRUPO II	GRUPO III	GRUPO IV	GRUPO V
ENERO	Media	0.2536	0.097	0.6191	0.7705	0.7569
	Desviación Estándar	0.0493	0.0211	0.0759	0.0764	0.0764
JUNIO	Media	0.2431	0.0991	0.5488	0.8612	0.7929
	Desviación Estándar	0.0838	0.1452	0.0931	0.0535	0.1200

El mayor consumo de energía es de hecho al mediodía, ya que a esa hora comienzan a dar servicio los restaurantes y tiendas del hotel. En ese horario la temperatura es más alta, por lo que los equipos de climatización necesitan trabajar más tiempo para alcanzar la temperatura deseada. Por la tarde y por la noche, además de los equipos de climatización, el alumbrado es el consumidor principal, pero como la temperatura ambiental es menor y muchos grandes locales están fuera de servicio, el consumo total es menor que a otras horas del día.

Figura 3. Dendograma del consumo (kw/h) de enero de 1995.

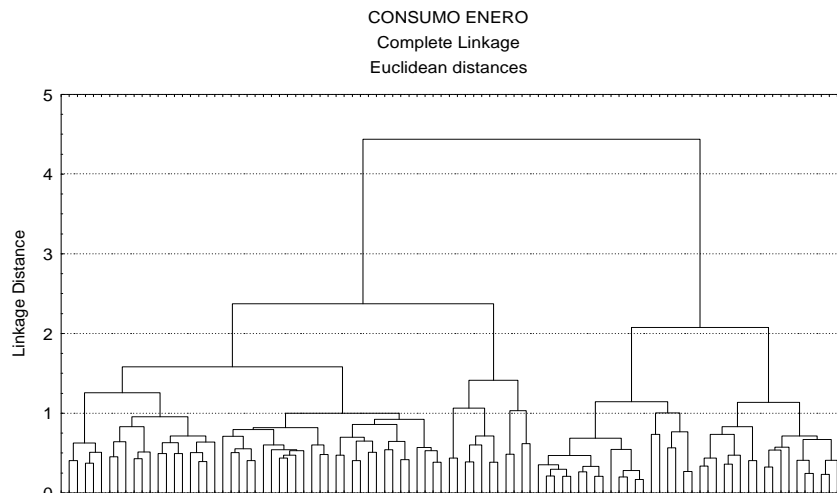
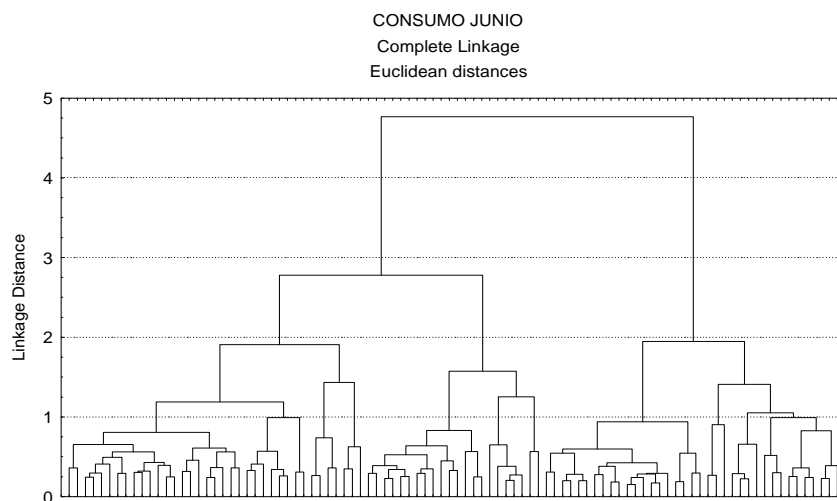


Figura 4. Dendograma del consumo (kw/h) de junio de 1995.



Como resultado de la aplicación del método STATIS se obtuvo que el coeficiente de la matriz RV relacionado con el consumo de los meses de enero y junio es cercano a 1. Éste se interpreta como un índice de similitud, por lo que se puede decir que los dos meses presentan una estructura similar en cuanto al comportamiento de los momentos del día de mayor consumo, así como también presentan relaciones de estabilidad entre los momentos de mayor y menor consumo.

Tabla 1. Matriz de coeficientes RV.

1.0000	0.9388
0.9388	1.0000

Para interpretar las posiciones de compromiso de las observaciones en el plano (Figura 5), se calculan las correlaciones entre las variables y los ejes de compromiso, habiendo correlación con todo el eje 1. La imagen euclidiana del compromiso del consumo muestra el comportamiento diferente de éste según la hora del día, poniéndose en evidencia la diferencia de consumo durante el periodo estudiado. Sólo interesa el análisis del primer eje ya que éste explica el 87.96 % de la varianza total. Al analizar la representación del compromiso se observa la formación de grupos con similares características a las obtenidas con el método de Taxonomía Numérica, los grupos del I al III son los mismos y los grupos IV y V se unen para formar un solo grupo.

Figura 5. Representación euclidiana del compromiso del consumo.

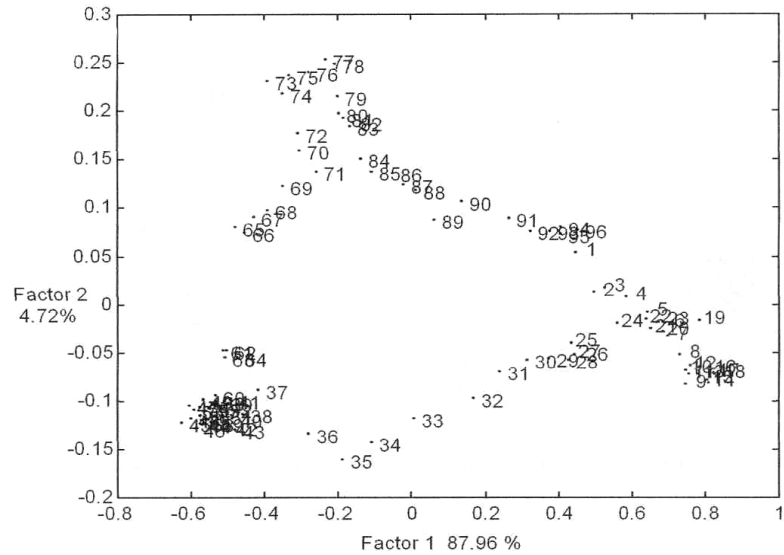
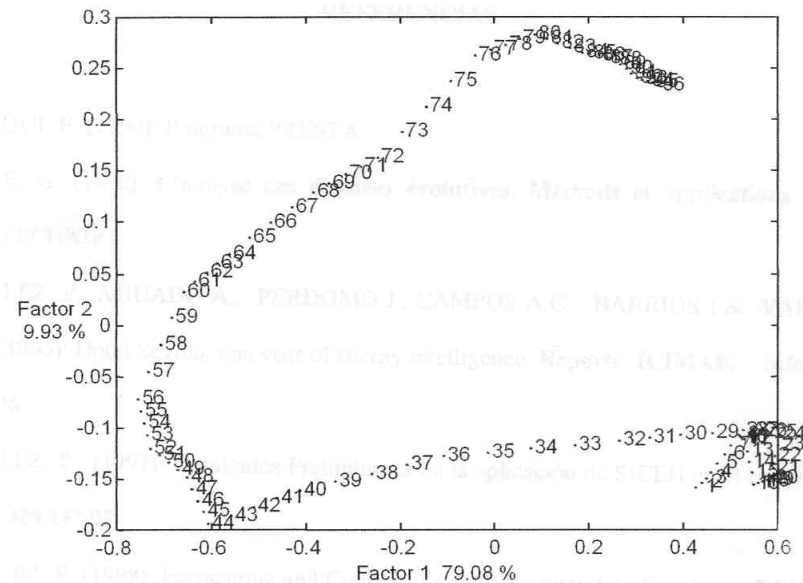
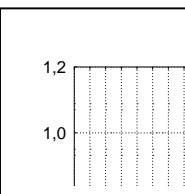


Figura 6. Imagen euclidiana del compromiso de la temperatura.



También se hizo el análisis para la temperatura y los resultados fueron similares. La Figura 6 nos refleja el comportamiento de la temperatura a lo largo del día.



CONCLUSIONES

La aplicación combinada de las técnicas multivariadas expuestas, hizo posible la determinación de grupos homogéneos de consumo energético con respecto a las características medidas. Esto permitió caracterizar el consumo energético del hotel en función del tiempo, lo que sirvió para conformar acciones para el ahorro de energía durante las horas de mayor consumo. A partir de estos resultados se pudieron evaluar posibles acciones para distribuir racionalmente el consumo de energía y evitar posibles sobrecargas, sin afectar las prestaciones del hotel.

REFERENCIAS

- CRIVISQUI, E. (1999): Programa PRESTA.
- GROUPE, G. (1996): **L'anayse des données évolutives. Méthods et applications.** Editions TECHNIP.
- GONZÁLEZ, P. **et al.** (1996): "Hotel Sevilla, one year of energy intelligence", **Reporte ICIMAF. Informática 96.**
- GONZÁLEZ, P. (1997): "Resultados Preliminares de la aplicación de SICEH en el Hotel Sevilla", **CIMAF' 97.**
- GONZÁLEZ, P. (1998): "Forecasting and Control of energy consumption based on a RBF Neural Network", **Neural Computation 98. Viena.**
- LACOURLY, N. (1997): Comunicación personal. Universidad de Chile.
- LAVIT, CH. (1988): **Analyse Conjointe de Tableaux Quantitatifs.** Masson, Paris.
- THIOULOUSE J.; D. CHESSEL; S. DOLÉDEC & J.M. OLIVIER (1997): "ADE-4: a multivariate analysis and graphical display software", **Statistics and Computing, 7, 1, 75-83.**