Curtosis coefficient and Hopfield neural net application in classification of electricity customers.

López Vázquez, José Jesús; Aguado Sánchez, José Antonio; Martín Moreno, Francisco, Muñoz Gutiérrez; Francisco; Rodríguez Gómez, Alejandro, Ruiz González, José Ernesto

Departamento de Ingeniería Eléctrica
Universidad de Málaga
Campus El Ejido– Málaga, 29013 Málaga (España)
Tel.:+34 952 131306, fax:+34 952 131091, e-mail: jjlopez@uma.es;fimartin@uma.es

1. Introduction.

Liberalization of electricity markets will have many consequences in the future. Electricity utilities will offer to consumers attractive prices. For this purpose, utilities need to know how customers consume electricity.
The problem is very complicated due to the great amount of data, and can be simplified by classifying the customers in a reduced number of groups. In this way, utilities have to elaborate just a small number of offers.
Mathematical tools have to be developed for making classification without loss of relevant information.
The general objective is the internal variance in each cluster to be minimum, and the distance among different groups to be maximum. It is also very important to detect anomalous customers (due to data errors or to a very rare client). These anomalous clients can cause a low quality classification. [1,3,4].
In this work a system based on maximun and minimum curtosis with the Mahalanobis distance [2] is used for elaboration of groups (clustering) and detection of anomalous clients.
The recurrent Hopfield neural network [1] is used. Classifications are compared by relative validation indices, with and without maximum and minimum curtosis projections.

Keywords: Curtosis, Hopfield, Calinsky, Clustering

2. Anomalous client detection

The procedure for anomalous client detection is explained in fig. 1. One can see two data groups, DMTIP, that will be used in clustering, and DMATIP, atypical data, that will be studied separately.

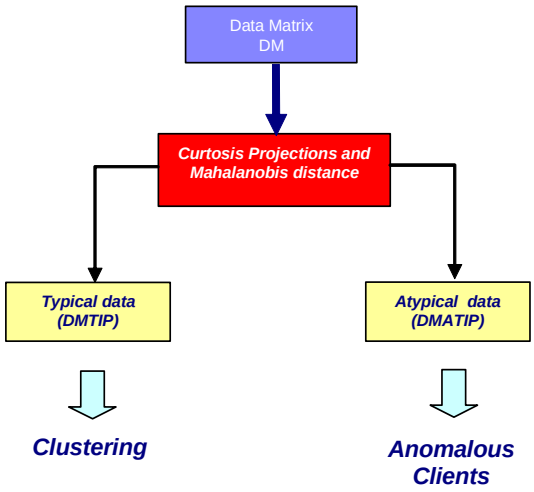

Figura 1: Aplicación de proyecciones de curtosis

3. Characterization and clustering .

Characterization and clustering of data matrix DMTYP procedure is explained in fig. 2.
The methods used ire in the time domain (Hour load profile (N) and form factors (FF)), in the frequency domain (harmonic analysis (F)), in the time-frequency domain (multi-resolution analysis, Wavelet transform (WT)) and main components (MC) [1,4].
In the clustering phase the recurrent Hopfield neural net and the Calinski, Davies Bouldin and WCBCR relative validation indices have been used. The number of clusters has been given from 5 to 30.

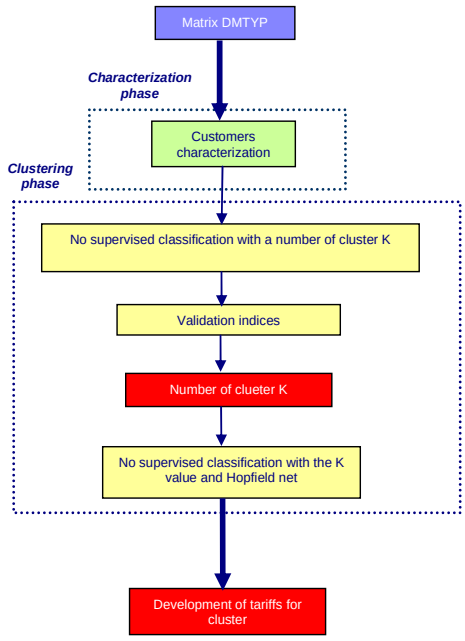

Fig. 2: Characterization and clustering schematics

4. Results.

A set of 230 medium voltage lines has been used in simulation, 19 feeding consumers of business type, 175 industrial type, and 36 residential. The results obtained of applying the procedure described above are listed in table I.

TABLE I
DATA MATRIX

	U	I	R	Total
MD	19	36	175	230
MDTYP	16	35	168	219
MDATYP	3	7	1	11

Data matrices MD and MDTIP have been clustered and results have been compared using the cluster validation indices. A summary of the obtained values for the CH (Calinski)) and the DB (Davies Boulding) indices is shown in table II.

TABLE II
RESULTS WITH CH Y DB AND HARMONIC ANALYSIS 3 (F3)

	MD		MDTYP	
	CH	DB	CH	DB
Valor	1068	0.6653	1215	0.65
N^{er} of clusters (K)	10	5	9	5

An example of the results obtained with the data matrices MD and MDTYP and the Calinski index is shown in fig. 3 and 4.

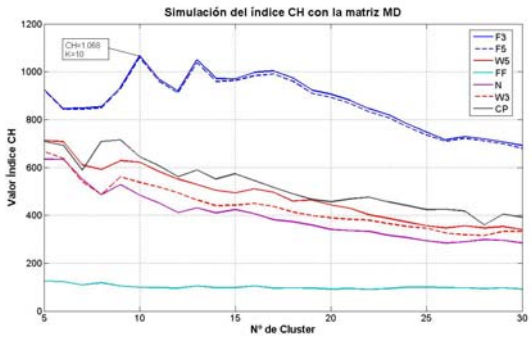

Fig. 3: Calinski index with matrix MD

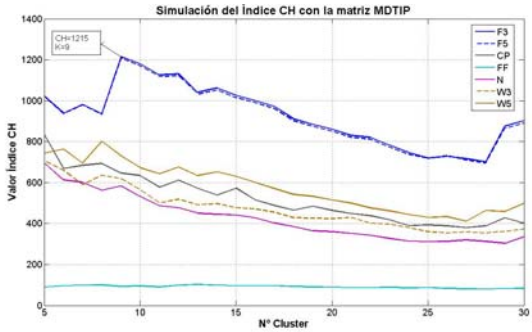

Fig. 4: Calinski index with matrix MDTYP

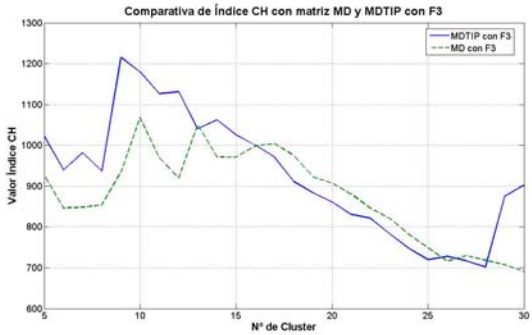

Fig. 5: CH index comparison with MD and MDTYP

4. Conclusions

Classification of customer has been improved with the application of maximum and minimum curtosis and Mahalabis distance.

The method has improved the identification of anomalous clients (MDATYP).

The quality of the clusters has been increased keeping the value K with respect to the initial clustering

References

[1] J.J. López, J.A. Aguado, F. Martín, F. Muñoz, A. Rodríguez and J.E. Ruiz, "Electric customer classification using Hopfield recurrent ANN", 5th International Conference on The European Electricity Market, Lisbon May 2008.

[2] Peña D. and Rodríguez, J. "Cluster identification using projections". Journal of American Statiscal Association, 96, 1433-1445.

[3] R., Xu, and D. Wunsch, "Survey of clustering algorithms", IEEE Transactions on Neural Networks, vol 16 nº 3, May 2005.

[4] G. Chicco, R. Napoli and F. Piglione, "Comparisons among clustering techniques for electricity customer classification", IEEE Trans. Power Systems, vol. 21, nº 2 may 2006.