Neural Networks Applied to Spatial Load Forecasting in GIS

J. N. FIDALGO
INESC Porto and Dept. of Electrical Engineering and Computers
Engineering Faculty of University of Porto
Rua Dr. Roberto Frias, nº 378, 4200-465, Porto
PORTUGAL

Abstract: - Quality spatial load forecasting is a major prerequisite for energy distribution systems planning. The load evolution outline depends on the urban expansion and its land usage. This paper presents a methodology for knowledge extraction of the data provided by a GIS (Geographical Information Systems) platform. The main goal consists of developing studies that lead to the understanding of the influence of geographical factors on the load growth patterns and energetic potential development. Kohonen maps and Artificial Neural Networks are used for data interpretation and spatial load forecasting purposes.

Key-Words: - Neural networks, Kohonen maps, geographical information systems, spatial load forecasting

1 Introduction

Transmission and distribution planning of power systems networks require the prediction of future geographical distribution of loads. This is the role of Spatial Load Forecasting (SLF) [13]: to forecast the future electric demand and its geographical distribution, generating expansion-planning scenarios and allowing the interpretation of load growth patterns. Pattern recognition tools are usually applied to detect and shape the sources of the load evolution. The influences of geographical factors on the land-use pattern development should be modeled in order to increase the forecasting and allow a suitable interpretation of the system behavior. For instance, a basic study of GIS data tells us clearly that load growth is generally higher when the distance to the road is smaller. Although important, this information is not enough; it should be quantified and integrated with the other potential factors. A considerable number of techniques have been applied to deal with this problem: fuzzy inference models [3-5, 10,11], cellular automata [4, 14,15], self-organization maps [12, 13, 14], neuro-fuzzy [12] and many others.

In this paper, the data analysis follows the following sequence. First, it portraits basic draw of the development (potential load growth) as a function of the geographical variables. This procedure allows to reduction the total amount of data to be deal with. Then, a feature selection procedure [1] is used to reduce the number of variables (potential factors for load growth), stressing the influence of the most important ones. The study follows with a clustering application, based on Kohonen Self-Organizing Maps [2]. Here, some conclusions are drawn about the characteristics of classes (prototypes) that lead to higher load growths. Finally, Artificial Neural Networks (ANN) are used as a regression tool to forecast the load growth as a function of the geographical factors.

2 Nomenclature

dev(i) – development (potential low growth) in the reference year i (prediction year);
dev(i-1) – development in the previous year;
dev(i-2) – development two years ago;
dev(i-3) – development three years ago;
altitude – altitude;
dist_mc – distance to main urban center;
dist_ac – distance to secondary urban center;
dist_r – distance to road;
slope – terrain slope;
sat(i) – saturation factor in the reference year;
sat(i-1) – saturation factor in the previous year;
sat(i-2) – saturation factor two years ago.

3 Main objectives

The main objectives of the current work are:

a) To develop a regression tool able to estimate dev(i) as a function of the other factors;
b) To handle the data in order to make possible the system behavior interpretation.

4 Data cleaning

A common problem when dealing with data provided by GIS is the amount of data to be dealt with. Generally, a large amount of the GIS data is irrelevant because it refers to non-urban areas like forests, rivers or ecologically protected territory, where building is not allowed. In these cases, the potential load growth is null or residual. Besides, there are also other cases where expected development is also negligible like the point where the slope, altitude or the distance to road are too high. In order to filter these cases, a picture of dev(i) as a function of the geographical factors was drawn. Fig. 1 illustrates changes in the development dev(i) as a function of the considered geographical variables. It is clearly shown that for some of these variables domains, there are "regions" where dev(i) approximately null. For instance, when the variable slope is higher that 15, dev(i) is null for the great majority of points. In order to discard this kind of meaningless data and decrease the quantity of data to handle the following rules where considered: Eliminate pattern if (either or):

a) altitude > 750 m;
b) dist_ac > 18000 m;
c) dist_rc > 1300 m;
d) slope > 16 m.

5 Identification of a minimal set of relevant features

This step consists of finding a methodology able to select a subset of features with low cardinal, but still able to characterize the system state, in this case represented by dev(i), as a function of the selected independent variables. In general, one cannot assure that the number of features achieved is really the smallest possible. To authors' knowledge, that can only be done by extensive search where all possible combinations of inputs are experienced. However, that's not the aim of this research; one is more interested in a practical (fast and efficient) method for simplifying ANN training and system behavior interpretation.

The proposed feature selection approach is fully described in [1], but it may be easily understood by analyzing the general scheme on Fig. 2. Here, the numbers near the arrows represent the actual number of variables: N1 is the initial number of variables, N2 (<N1) is the number of variables after correlated features are dismissed and N3 (<N2) is the final number of variables (after the rejection of the less-sensitive features).

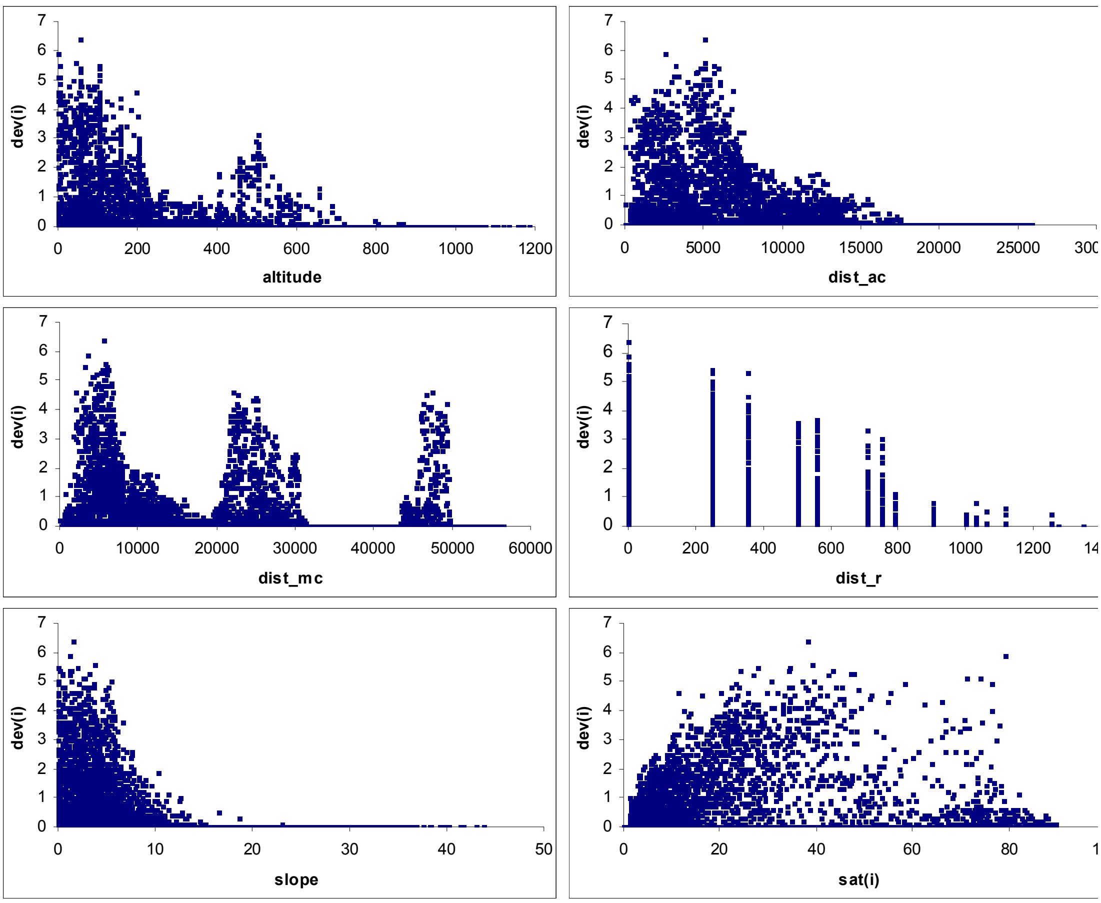

Fig. 1- Development as a function of the considered geographical variables

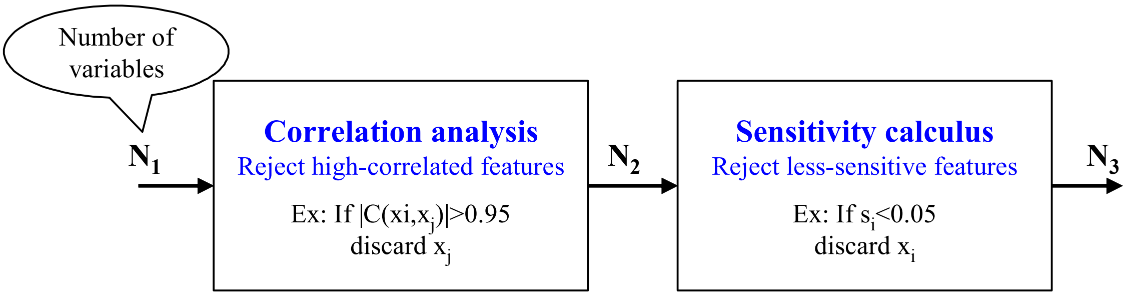

Fig. 2 – General FSS method shceme

Table 1 – Correlation analysis

	altitude	*dist_ac*	*dist_mc*	*dist_r(i)*	*slope*	*dev(i-3)*	*dev(i-2)*	*dev(i-1)*	*sat(i-2)*	*sat(i-1)*	*sat(i)*	*dev(i)*
altitude	1.00											
dist_ac	-0.03	1.00										
dist_mc	0.19	-0.33	1.00									
dist_r(i)	-0.03	-0.10	-0.05	1.00								
slope	0.33	-0.13	0.20	0.00	1.00							
dev(i-3)	-0.20	-0.34	-0.18	-0.16	-0.16	1.00						
dev(i-2)	-0.22	-0.35	-0.18	-0.14	-0.15	0.78	1.00					
dev(i-1)	-0.24	-0.35	-0.14	-0.16	-0.20	0.56	0.78	1.00				
sat(i-2)	-0.26	-0.47	-0.30	-0.01	-0.11	0.55	0.43	0.28	1.00			
sat(i-1)	-0.29	-0.50	-0.31	0.02	-0.11	0.60	0.51	0.36	**0.99**	1.00		
sat(i)	-0.32	-0.53	-0.34	0.05	-0.11	0.61	0.57	0.46	0.94	**0.98**	1.00	
dev(i)	-0.24	-0.30	-0.08	-0.18	-0.20	0.26	0.46	0.70	0.08	0.15	0.24	1.00

From the results shown in Table 1, one may conclude that there is a strong correlations among sat(i), sat(i-1) and sat(i-2). According to proposed feature selection methodology, two of these three variables are redundant – only one of these features is needed for ANN multiregression purposes, without lost of performance. Then, an ANN with these inputs is trained and its performance is evaluated in order to check if the remaining variables really contain enough discriminating power.

The inputs of the ANN shown in Fig. 3 contain no correlated variables. Once trained, this ANN is used as support for sensitivity indexes (si) calculus. According to the results presented in Table 2 and considering a threshold of 0.05 for the sensitivity index, the two last features (grey shadow) must be discarded.

The establishment of thresholds is a common problem in a good number of engineering techniques. In fact, adopted threshold values in the correlation analysis and sensitivity evaluation phases may affect global results.

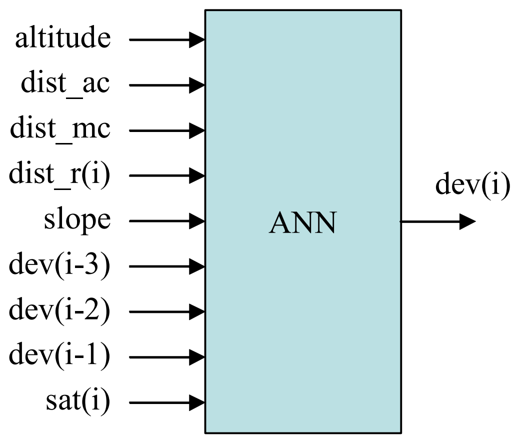

Fig. 3 – Intermediate ANN (after the elimination of the correlated features)

Table 2 – Sensitivity indexes ranking

Rank	Variable	s_i
1	dist_r(i)	0.235
2	saturation(i)	0.202
3	altitude	0.155
4	dist_mc	0.118
5	slope	0.103
6	development(i-1)	0.084
7	dist_ac	0.058
8	development(i-3)	0.023
9	development(i-2)	0.021

For instance, if sensitivity threshold value is settled to high one may discard some variables that may be important, at least under some operation condition and degrade final performance. If that happens, that threshold should be decreased and the final step (ANN training and performance evaluation) should be repeated. In general, the importance of variables may vary from the uninformative or redundant to the most significant ones.

After the elimination of the uninformative variables, one arrives at a point where a trade-off between number of variables and performance is patent. At this point, one may decide to keep a larger variable set and a given performance or decrease that set and allow some degradation at final performance.

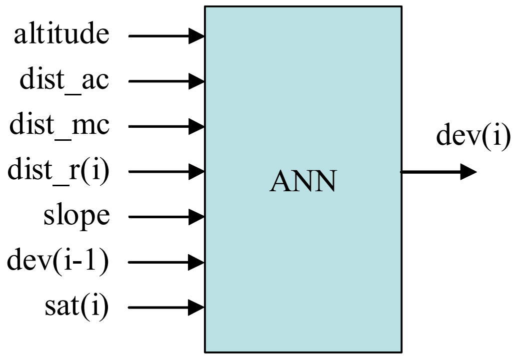

Fig. 4 – ANN input/output architecture after the elimination of the most irrelevant features

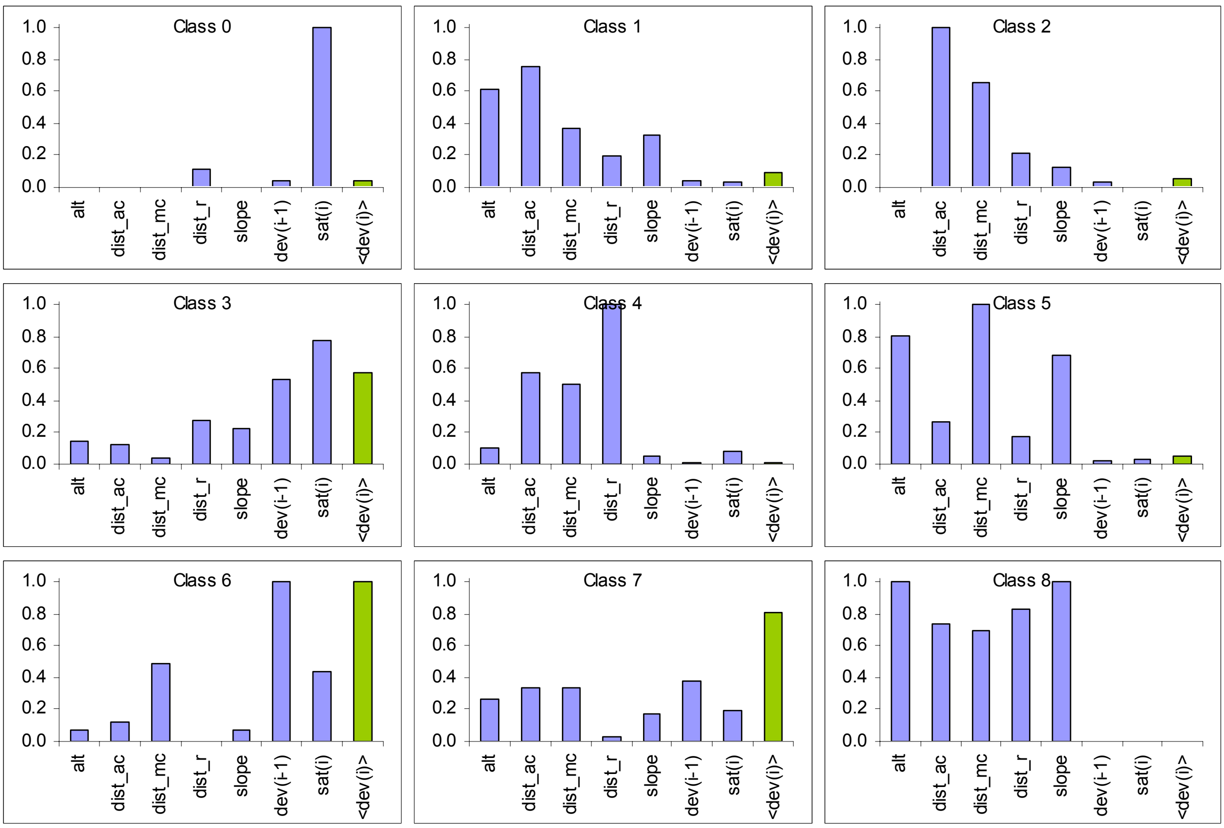

Fig. 5 – Classes attained by the Kohonen clustering algorithm. Variable <dev(i)> - mean development in each class - was not used as input of the clustering process.

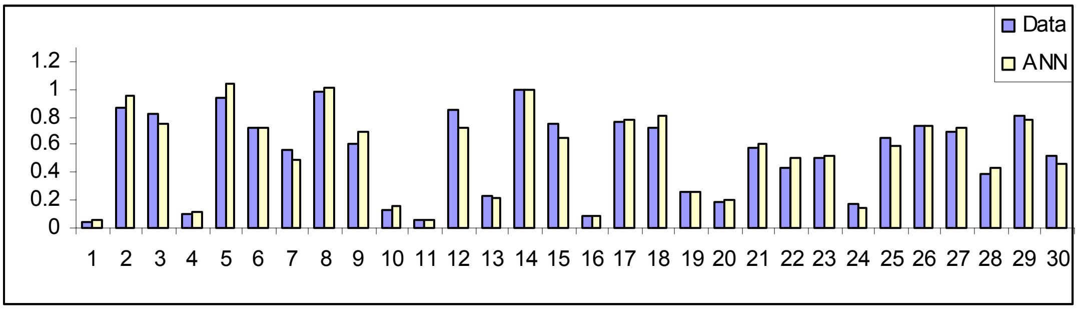

Fig. 6 – Examples of ANN dev(i) prediction compared to the real data

5 Development prediction

Fig. 4 show the final ANN input /output scheme after discarding the less significant features, according to the sensitivities index ranking illustrated in Table 2. After training this ANN presents a mean absolute deviation error (mean absolute error / mean absolute value) of 0.37. At the first sight this performance error may be seem somewhat poor. However, there are two factors that contribute for this event:

1. There is a considerable part of GIS data that correspond to null or quite insignificant development (rural areas, sloppy areas and so on). These data points where dev(i)=0 bring the mean absolute value to a very small level;
2. GIS data deals with social (human) behavior (where do I want my house?) and this behavior is typically hard to forecast.

Fig. 6 shows some prediction examples. As one can see, ANN output generally approximates the target values.

4 Clustering

The remaining features were then subjected to a clustering process based on Kohonen algorithm [2]. Several experiences were conducted considering different hypothesis in the output Kohonen grid (different numbers of classes). Fig. 5 shows the most successful outcome. In this figure, the variable dev(i) was also included for interpretation purposes, although it was not used for clusters determination. In order to make the interpretation easier, all variables are normalized in the range [0;1] – minimum value of each variable is settled to 0 and the maximum to 1

The analysis of Fig. 5 allow us to conclude that higher values of dev(i) occur in the left-bottom corner of the Kohonen grid (classes 3, 6 and 7). A straight reading of these results lead to the following conclusion: higher development dev(i) arise when the development in the previous year dev(i-1) is high if the region is not over-saturated (sat(i)<0.8).

The other classes (0, 1, 2, 4, 5 and 8) exhibit a poor development dev(i) potential. Being so, they are not considered as relevant in the previous analysis.

6 Conclusions

Kohonen maps, correlation analysis and Artificial Neural Networks were used for understanding of the geographical data relations. One was particularly interested on the interpretation of the potential influence of the available geographical factors in the development potential. The analysis of the results show that proposed methodology is able to discover the most relevant features, providing at the same time a (clustered) view of the geographical data classes.

References:

[1] J. N. Fidalgo, "Feature Selection Based on ANN Sensitivity Analysis - a Practical Study", in *Proc. 2001 International Conference on Neural Networks and Applications, WSES2001*, Tenerife, Feb. 2001, pp. 206-211.

[2] T. Kohonen, *Self Organization and Associative Memory*, Springer-Verlag, Berlin, 1984

[3] Vladimiro Miranda, Cláudio Monteiro, "Probabilities and Fuzzy sets in the Evaluation of Renewable Energies - The SOLARGIS Experience", *Proceedings of PMAPS'97 - 5th International Probabilistic Methods Applied to Power Systems Conference*, vol.1, September, 1997

[4] Vladimiro Miranda, Cláudio Monteiro, "Fuzzy Inference in Spatial Load Forecasting",

Proceedings of PES Winter Meeting 2000, January, 2000

[5] Cláudio Monteiro, Vladimiro Miranda, Maria Teresa Ponce de Leão, "Scenario Identification Process in Fuzzy Spatial Load Forecasting", *Proceedings of PMAPS'2000 - 6th International Conference on Probabilistic Methods Applied to Power Systems*, September, 2000

[6] D. E. Rumelheart, G. E. Hinton, R. J. Williams, "Learning Internal Representations By Error Propagation", *In Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*, D. E. Rumelhart, J. L. McClelland, (Eds.), MIT Press, Cambridge, Ma, USA, 1986.

[7] Simon Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed., Prentice-Hall, 1999.

[8] F. M. Silva, L. B. Almeida, "Acceleration Techniques For The Backpropagation Algorithm", *In Neural Networks*, L. B. Almeida and C. J. Wellekens (Eds.), Springer-Verlag, 1990.

[10] Cláudio Monteiro, Vladimiro Miranda, Maria Teresa Ponce de Leão, "Scenario Identification Process in Fuzzy Spatial Load Forecasting", Proceedings of PMAPS'2000 - 6th International Conference on Probabilistic Methods Applied to Power Systems, September, 2000.

[11] Cláudio Monteiro, Vladimiro Miranda, Maria Teresa Ponce de Leão, "Validation Process for Fuzzy Spatial Load Forecasting", Proceedings of PMAPS'2002 - International Conference on Probabilistic Methods Applied to Power Systems, September, 2002.

[12] Vladimiro Miranda, Cláudio Monteiro, Tatjana Konjic, "Spatial Offer and demand forecasting with Neuro-Fuzzy Inference Systems in GIS", Proceedings of WM2001 - IEEE PES Winter Meeting, (IEEE CD-ROM 01CH37194C), January, 2001.

[13] H. L. Willis, Spatial Load Forecasting, Marcel Dekker Inc.,New York, 1996. ISBN: 0-8247-9425-7

[14] R. White, G. Engelen, "Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land use-use patterns", Environment and Planning A, No25,1993, pp. 1175-1189.

[15] M. Batty, "Cellular automata and urban form: a primer", Journal of American Planning Association, No63, 1997, pp. 266-274