Patterns in Temporal Series of Meteorological Variables Using SOM & TDIDT

Marisa Cogliati, Paola Britos and Ramón García-Martínez

Geography Department. School of Human Sciences. National University of Comahue
cogliati@uncoma.edu.ar
Software & Knowledge Engineering Center. Graduate School. Buenos Aires Institute of Technology
Intelligent Systems Laboratory. School of Engineering. University of Buenos Aires
{pbritos,rgm}@itba.edu.ar

Abstract. The purpose of the present article is to investigate if there exist any such set of temporal stable patterns in temporal series of meteorological variables studying series of air temperature, wind speed and direction an atmospheric pressure in a period with meteorological conditions involving nocturnal inversion of air temperature in Allen, Río Negro, Argentina. Our conjecture is that there exist independent stable temporal activities, the mixture of which give rise to the weather variables; and these stable activities could be extracted by Self Organized Maps plus Top Down Induction Decision Trees analysis of the data arising from the weather patterns, viewing them as temporal signals.

1. Introduction

Classical laws of fluid motion govern the states of the atmosphere. Atmospheric states exhibit a great deal of correlations at various spatial and temporal scale. Diagnostic of such states attempt to capture the dynamics of various atmospheric variables (like temperature and pressure) and how physical processes influence the behaviour. Thus weather system can be thought as a complex system whose components interact in various spatial and temporal scales. It is also known that the atmospheric system is chaotic and there are limits to the predictability of its future state [Lorenz, 1963, 1965]. Nevertheless, even though daily weather may, under certain conditions, exhibit symptoms of chaos, long-term climatic trends are still meaningful and their study can provide significant information about climate changes. Statistical approaches to weather and climate prediction have a long and distinguished history that predates modelling based on physics and dynamics [Wilks, 1995; Santhanam and Patra, 2001].
Intelligent systems are appearing as useful alternatives to traditional statistical modelling techniques in many scientific disciplines [Hertz *et al*., 1991; Rich & Knight, 1991; Setiono & Liu, 1996; Yao & Liu, 1998; Dow & Sietsma, 1991; Gallant, 1993; Back *et al*., 1998; García Martínez & Borrajo, 2000; Grosser *et al.*, 2005]. In their overview of applications of neural networks (as example of intelligent system) in the atmospheric sciences, Gardner and Dorling [1998] concluded that neural networks

generally give as good or better results than linear methods. So far, little attention has been paid to combining linear methods with neural networks or other types of intelligent systems in order to enhance the power of the later. A general rule in this sort of applications says that the phenomenon to be learned by the intelligent system should be as simple as possible and all advance information should be utilized by pre-processing [Haykin, 1994]. This trend continues today with newer approaches based on machine learning algorithms [Hsieh and Tang, 1998; Monahan, 2000].
The term intelligent data mining [Evangelos & Han, 1996; Michalski *et al.*, 1998], is the application of automatic learning methods [Michalski *et al.*, 1983; Holsheimer & Siebes, 1991] to the non-trivial process of extract and present/display implicit knowledge, previously unknown, potentially useful and humanly comprehensible, from large data sets, with object to predict of automated form tendencies and behaviours; and to describe of automated form models previously unknown, [Chen *et al.*, 1996; Mannila, 1997; Piatetski-Shapiro *et al.*, 1991; 1996; Perichinsky & García-Martínez, 2000; Perichinsky *et al.*, 2003] involve the use of machine learning techniques and tools.

2. Problem

The central problem in weather and climate modelling is to predict the future states of the atmospheric system. Since the weather data are generally voluminous, they can be mined for occurrence of particular patterns that distinguish specific weather phenomena. It is therefore possible to view the weather variables as sources of spatio-temporal signals. The information from these spatio-temporal signals can be extracted using data mining techniques. The variation in the weather variables can be viewed as a mixture of several independently occurring spatio-temporal signals with different strengths. Independent component analysis (ICA) has been widely studied in the domain of signal and image processing where each signal is viewed as a mixture of several independently occurring source signals. Under the assumption of non-Gaussian mixtures, it is possible to extract the independently occurring signals from the mixtures under certain well known constraints. Therefore, if the assumption of independent stable activity in the weather variables holds true then it is also possible to extract them using the same technique of ICA. One basic assumption of this approach is viewing the weather phenomenon as a mixture of a certain number of signals with independent stable activity. By ‘stable activity’, meaning spatiotemporal stability, i.e., the activities that do not change over time and are spatially independent. The observed weather phenomenon is only a mixture of these stable activities. The weather changes due to the changes in the mixing patterns of these stable activities over time. For linear mixtures, the change in the mixing coefficients gives rise to the changing nature of the global weather [Stone, Porrill, Buchel, and Friston, 1999; Hyvarinen, 2001].
The purpose of the present article is to investigate if there exist any such set of temporal stable patterns related to the observed weather phenomena. Our conjecture is that there exist independent stable temporal activities, the mixture of which give rise to the weather variables; and these stable activities could be extracted by neural

networks analysis of the data arising from the weather and climate patterns, viewing them as temporal signals.

3. Proposed Solution

The variables as presented in the paper could not be considered random ones because of presence of temporal cycles. In addition, a linear behaviour as result of mixture of latent variables could not be assumed [Hyvarinen *et al.*, 2001]. In order to establish if there exist any such set of temporal stable patterns related to observed weather or climate phenomena we select weather station data described in [Flores *et al*, 1996]. The records of the observed weather temporal series [Ambroise *et al.*¸2000; Malmgren and Winter, 1999 ; Tian *et al.*, 1999] are clustered with SOM [Kohonen, 2001; Kasi, *et al.*, 2000; Tirri, 1991; Duller, 1998] and rules describing each obtained cluster were built applying TDIDT [Quinlan, 1993] to each cluster records. The described process is shown in figure 1.

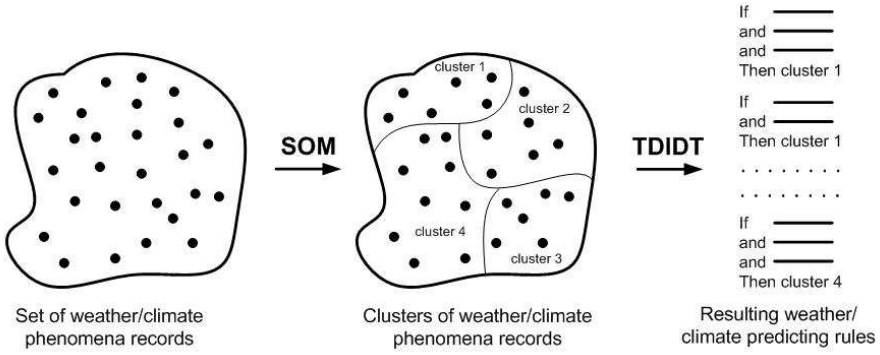

Fig. 1. Process for establishing temporal stable patterns related to observed weather/ climate phenomena

4. Data for experiments

The original data was a set of temperature, wind speed, wind direction and atmospheric pressure observations, taken every fifteen minutes from 13/10/94 to 17/10/94 in Allen, Río Negro province, Argentina. The weather station was located in the agricultural region called Upper Río Negro Valley (URNV) encompassing the lower valleys of the Limay and Neuquén rivers and the upper valley of the Negro river. The arable lands of best quality are located on the river terraces extending from the side pediments up to the floodplain. The terraces are limited by cliffs and the side pediments of the Patagonian plateau that surrounds the valleys. The valley is broad and shallow with steplike edges. The Negro river valley has a WNW-to-ESE orientation in the study area. The mean height differences with the North Patagonian Plateau is 120m for the Río Negro valley. The weather station data was obtained during MECIN (stands in spanish for: MEdiciones de la Capa de Inversión Nocturna:

Nocturnal Inversion Layer Measurements] field experience carried out in the URNV [Flores *et al*, 1996] from September through October of the years 1992 to 1997. The data was complete, without using any replacement technique. The so called Upper Río Negro Valley in Argentina is one of the most important fruit and vegetable production regions of the country. It comprises the lower valleys of the Limay and Neuquén rivers and the upper Negro river valley. Out of the 41,671 cultivated hectares, 84.6% are cultivated with fruit trees, especially apple, pear and stone fruit trees [Cogliati, 2001]. Late frosts occurring when trees are sensitive to low temperatures have a significant impact on the regional production. This study presents an analysis of meteorological variables in one weather station in the Upper Río Negro Valley. To such effect, observations made when synoptic-scale weather patterns were favourable for radiative frosts (light wind and clear sky) or nocturnal temperature inversion in the lower layer were used. Calm winds were more frequent within the valleys. In Allen, air flow behaviour might be associated with forced channelling with wind direction following valley direction. In the night time, some cases of very light NNE wind occurred, which may be associated with drainage winds from the *barda*.

5. Results of experiments

The first analysis implementing SOM analysis determined nine clusters, that could be associated to different wind directions, maximum and mean wind speed, atmospheric pressure and temperature. Air temperature includes periodic daily variation, that was included in the analysis to explore relationship with wind variations. Four of nine groups identified, included the 94 percent of cases and several statistically significant rules. The detected rules for each group (cluster) are described in tables 1 to 9.

Groups A and B describe strongest wind cases with maximum wind speed greater than 5.8 m/s and mean wind speed greater than 1.3 m/s. Group C describe cases considering greater temperatures and wind speed with wind direction from south. Group D describes cases of wind speed up to 5 m/s from north to south directions and wind speed up to 5 m/s. In groups F and G and H cases present non obvious characteristics. Group J discriminates calm wind and groups Z_1 and Z_2 describes undeterminated cases. The required frost analysis involve nocturnal and diurnal processes identification, so, the time of observation is a variable that might be included. The inclusion of date and time of observation produced a diminution of the quantity of groups involved, but an important increment in the number of rules (38 rules). This inclusion of new characteristics in the TDIDT analysis produced too much behaviour rules that produces confusion and detect obvious patterns as well as useful ones. This item would need further additional analysis. A confidence limit was pointed in order to study the rules.

Considering confidence level above 0.6 and rules involving more than 25 cases results in 11 rules. This rules pointed some groups characteristics. Group A includes 135 cases with relative higher air pressure mainly in the morning. The 324 cases in Group B present lower air pressure and wind speed. Prevailing wind direction was western sector. Group C discriminated weaker mean wind speed (less than 0.2 m/s) during the

morning and relative higher air pressure (371 cases) and cooler air temperature mainly from northern to southern direction. Group D presented westerly wind but cooler air temperature (96 cases) and F includes early afternoon and afternoon cases (154 cases).

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 992.84 AND C5294VVE < 0.20 AND HOUR < 10.33 THEN GROUP = A	26	0.85
IF C5294P >= 990.64 AND C5294P < 991.68 AND C5294VMX >= 0.65 AND HOUR >= 10.33 AND HOUR < 16.48 AND C5294TOU >= 16.75 THEN GROUP = A	12	0.92
IF C5294P >= 991.68 AND C5294VMX >= 0.65 AND HOUR >= 10:33 AND (1) (1) HOUR < 16:48 THEN GROUP = A	51	1
IF C5294P >= 989.61 AND C5294VVE >= 0.20 AND HOUR >= 6:57 AND HOUR < 10:33 THEN GROUP = A	26	1

Table 1. Rules from Group A (Cluster A)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P < 986.54 AND C5294VVE >= 0.20 AND HOUR < 10:33 THEN GROUP = B	44	0.91
IF C5294P < 989.24 AND C5294VMX >= 0.65 AND HOUR >= 10:33 AND C5294TOU >= 15.15 THEN GROUP = B	265	0.9
IF C5294P < 985.14 AND C5294VMX >= 0.65 AND C5294VMX < 1.55 AND HOUR >= 10:33 AND C5294TOU < 15.15 THEN GROUP = B	6	0.67
IF C5294P < 986.14 AND C5294VMX >= 1.55 AND HOUR >= 10:33 AND C5294TOU < 15.15 THEN GROUP = B	18	1
IF C5294P >= 986.14 AND C5294P < 989.24 AND C5294VMX >= 1.55 AND C5294VVE >= 1.10 AND HOUR >= 10:33 AND C5294TOU < 15.15 THEN GROUP = B	7	1
IF C5294P < 979.48 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR < 8:09 AND C5294TOU < 10.35 THEN GROUP = B	2	1

Table 2. Rules from Group B (Cluster B)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P < 992.84 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR < 8.09 THEN GROUP = C	225	0.99
IF C5294P < 992.84 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 8.09 AND HOUR < 10.33 AND C5294TOU < 5.40 THEN (1.00) (1) GROUP = C	5	1
IF C5294P >= 987.15 AND C5294P < 992.84 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 8.09 AND HOUR < 10.33 AND C5294TOU >= 5.40 THEN GROUP = C	6	1
IF C5294P >= 984.49 AND C5294P < 987.15 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 8.09 AND HOUR < 8.34 AND C5294TOU >= 5.40 THEN GROUP = C	4	0.75
IF C5294P >= 988.14 AND C5294P < 992.84 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND (1) (1) HOUR >= 8.09 AND (1) (1) HOUR < 10.33 THEN GROUP = C	29	0.83
IF C5294P < 992.84 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR <8.09 AND C5294TOU >= 10.35 THEN GROUP = C	50	0.74
IF C5294P >= 989.61 AND C5294VVE >= 0.20 AND HOUR < 6.57 THEN GROUP = C	35	0.60
IF C5294P >= 989.05 AND C5294P < 992.84 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR < 8.09 AND C5294TOU < 10.35 THEN GROUP = C	15	1
IF C5294P >= 989.24 AND C5294P < 990.64 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR >= 10.33 AND HOUR < 16.48 THEN GROUP = C	12	0.92

Table 3. Rules from Group C (Cluster C)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 979.48 AND C5294P < 989.05 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR < 8:09 AND C5294TOU < 10.35 THEN GROUP = D	25	0.68
IF C5294P >= 986.54 AND C5294P < 989.61 AND C5294VVE >= 0.20 AND HOUR < 10:33 THEN GROUP = D	8	0.63
IF C5294P >= 989.24 AND C5294VMX >= 2.00 AND HOUR >= 16:48 THEN GROUP = D	24	0.67
IF C5294P >= 989.24 AND C5294P < 990.64 AND C5294VMX >= 0.65 AND C5294VVE >= 0.20 AND HOUR >= 10:33 AND HOUR < 16:48 THEN GROUP = D	6	0.67
IF C5294P >= 990.64 AND C5294P < 991.68 AND C5294VMX >= 0.65 AND HOUR >= 10:33 AND HOUR < 16:48 AND C5294TOU < 16.75 THEN GROUP = D	8	0.63
IF C5294P >= 986.14 AND C5294P < 989.24 AND C5294VMX >= 1.55 AND C5294VVE < 1.10 AND HOUR >= 10:33 AND C5294TOU < 15.15 THEN GROUP = D	26	0.81
IF C5294P >= 984.92 AND C5294P < 988.14 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR >= 8:09 AND HOUR < 10:33 THEN GROUP = D	7	1

Table 4. Rules from Group D (Cluster D)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 981.17 AND C5294VMX < 0.65 AND HOUR >= 12:45 THEN GROUP = F	159	1
IF C5294P >= 985.14 AND C5294P < 989.24 AND C5294VMX >= 0.65 AND C5294VMX < 1.55 AND HOUR >= 20:24 AND C5294TOU >= 13.20 AND C5294TOU < 15.15 THEN GROUP = F	2	1

Table 5. Rules from Group F (Cluster F)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P < 981.17 AND C5294VMX < 0.65 AND HOUR >= 10:33 THEN GROUP = G	15	0.93
IF C5294P >= 985.14 AND C5294P < 989.24 AND C5294VMX >= 0.65 AND C5294VMX < 1.55 AND HOUR >= 20:24 AND C5294TOU < 13.20 THEN GROUP = G	12	0.92
IF C5294P >= 989.24 AND C5294VMX >= 0.65 AND C5294VMX < 2.00 AND HOUR >= 16:48 THEN GROUP = G	20	0.4

Table 6. Rules from Group G (Cluster G)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 981.17 AND C5294VMX < 0.20 AND HOUR >= 10:33 AND HOUR < 12:43 THEN GROUP = H	10	0.9
IF C5294P < 984.49 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 8.09 AND HOUR < 10:33 AND >= 5.40 THEN GROUP = H	11	1
IF C5294P >= 984.49 AND C5294P < 992.84 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 9:36 AND HOUR < 10:33 AND C5294TOU >= 5.40 THEN GROUP = H	7	1
IF C5294P >= 984.49 AND C5294P < 987.15 AND C5294VMX < 0.65 AND C5294VVE < 0.20 AND HOUR >= 8:38 AND HOUR < 9:36 AND C5294TOU >= 5.40 THEN GROUP = H	5	1

Table 7. Rules from Group H (Cluster H)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 985.14 AND C5294P < 989.24 AND C5294VMX >= 0.65 AND C5294VMX < 1.55 AND HOUR >= 10:33 AND HOUR < 20:24 AND C5294TOU < 15.15 THEN GROUP = J	4	0.5

Table 8. Rules from Group J (Cluster J)

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P >= 981.17 AND C5294VMX >= 0.20 AND C5294VMX < 0.65 AND HOUR >= 10:33 AND HOUR < 12:43 THEN GROUP = UNDETERMINATE	6	0.33

RULES	SUPORT DATA	CONFI-DENCE
IF C5294P < 984.92 AND C5294VMX >= 0.65 AND C5294VVE < 0.20 AND HOUR >= 8:09 AND HOUR < 10:33 THEN GROUP = UNDETERMINATE	10	0.4

Table 9. Rules from Group Z_1 and Z_2 (indeterminated cluster)

In "C5294…", "C52" is the meteorological station code and "94" is the year (1994). In "C5294vdd", "vdd" is the wind orientation. In "C5294vve", "vve" is average wind intensity. In "C5294tou", "tou" is air temperature (Cº). In "C5294P", "P" is pressure (hPa).

The figure 2 presents the maximum wind speed versus local time for the different groups selected. The discrimination of different meteorological situations could differentiate physical relationships in the analyzed cases, further analysis considering atmospheric temporal variations could improve the selection, discarding the obvious deterministic patterns.

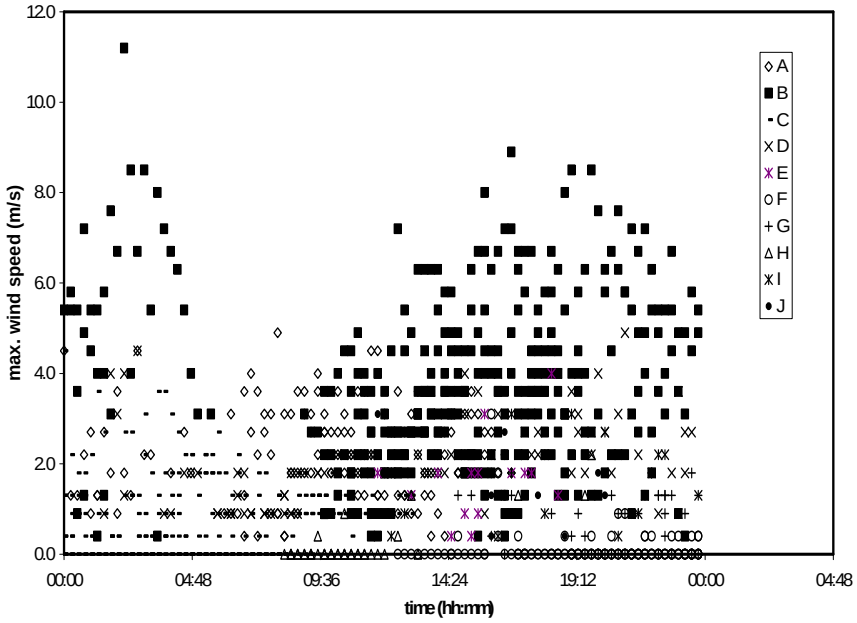

Fig. 2. Scatter plot of different groups data of maximum wind speed versus time in Allen (Río Negro Argentina) from 13/10/94 to 17/10/94.

6. Conclusions

The so called Upper Río Negro Valley in Argentina is one of the most important fruit and vegetable production regions of the country. It comprises the lower valleys of the

Limay and Neuquén rivers and the upper Negro river valley. Late frosts occurring when trees are sensitive to low temperatures have a significant impact on the regional production. Time series analysis of air temperature, atmospheric pressure, wind speed and direction involves a large amount of data and data mining could be an alternative to statistical traditional methods to find clusters with stable signals.
This study presents an analysis of meteorological variables in one weather station in the Upper Río Negro Valley by means of SOM analysis and applying TDIDT to build rules. To such effect, observations made when synoptic-scale weather patterns were favourable for radiative frosts (light wind and clear sky) or nocturnal temperature inversion in the lower layer were used. The obtained rules represent wind, temperature and pressure characteristics, the groups separate calm, and nocturnal and diurnal main characteristics according to prior traditional methods analysis (Cogliati, 2001), newer found relationships might be studied in advance.
The inclusion of a larger number of variables such time and date produces a large number of rules without defining precise intervals that produces confusion and detect obvious patterns as well as useful ones. This item would need further extensive study.
The variation in the weather variables can be viewed as a mixture of several independently occurring spatio-temporal signals with different strengths.

Acknowledgements: The authors would like to thank Jorge Lässig for providing the meteorological data obtained in MECIN field experiment.

7. References

Ambroise, C., Seze, G., Badran, F., and Thiria, S. 2000. *Hierarchical clustering of self-organizing maps for cloud classification*. Neurocomputing, 30(1):47–52

Back, B., Sere, K., & Vanharanta, H. 1998. *Managing complexity in large data bases using self-organizing maps*. Accounting Management & Information Technologies 8, 191-210.

Chen, M., Han, J., Yu, P. 1996. *Data mining: An overview from database perspective*. IEEE Transactions on Knowledge and Data Engineering 8(6): 866-883.

Cogliati, M.G. 2001. *Estudio térmico y del flujo del aire en septiembre y octubre en los valles de los ríos Limay, Neuquén y Negro*. Doctoral Dissertation. University of Buenos Aires.

Dow R. J. y Sietsma J. 1991. *Creating Artificial Neural Networks that Generalize*. Neural Networks . 4(1): 198-209.

Duller, A. W. G. 1998. *Self-organizing neural networks: their application to "real-world" problems*. Australian Journal of Intelligent Information Processing Systems, 5:175–80

Evangelos, S., Han, J. 1996. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. Portland, EE.UU.

Flores, A. ; Lässig, J. ; Cogliati, M. ; Palese, C., Bastanski, M. 1996. *Mediciones de la Capa de Inversión Nocturna en los valles de los ríos Limay, Neuquén y Negro*. Proceedings VII Argentine Congress on Meteorology. VII Latinamerican and Iberic Congress on Meteorology. Buenos Aires.

Gallant, S. 1993. *Neural Network Learning & Experts Systems*. MIT Press, Cambridge, MA.

García Martínez, R. y Borrajo, D. 2000. *An Integrated Approach of Learning, Planning & Executing*. Journal of Intelligent & Robotic Systems. 29(1): 47-78.

Gardner, M., Dorling, S. 1998. *Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences*. Atmospheric Environment 32: 2627-2636

Grosser, H., Britos, P. y García-Martínez, R. 2005. *Detecting Fraud in Mobile Telephony Using Neural Networks.* Lecture Notes in Artificial Intelligence 3533: 613-615.

Haykin, S., 1994. *Neural networks: A comprehensive foundation*. Prentice-Hall, Englewood Cliffs, NJ.

Hertz J., A. Krogh y R. Palmer 1991. *Introduction to the Theory of Neural Computation*. Reading, MA: Addison-Wesley.

Holsheimer, M., Siebes, A. 1991. *Data Mining: The Search for Knowledge in Databases*. Report CS-R9406, ISSN 0169-118X, Amersterdam, The Netherlands.

Hsieh, W. , and Tang, B. 1998. *Applying neural network models to prediction and data analysis in meteorology and oceanography*. Bulletin of American Meteorological Society 79: 1855-1870.

Hyvarinen, A. 2001. *Complexity pursuit: Separating interesting components from time-series*. Neural Computation 13: 883-898.

Hyvarinen, A., Karhunen, J. and Oja, E. 2001. *Independent Component Analysis*. John Wiley & Sons.

Kaski, S., Venna, J., and Kohonen, T. 2000. *Coloring that reveals cluster structures in multivariate data.* Australian Journal of Intelligent Information Processing Systems, 6:82–8.

Kohonen, T. 2001. *Self-Organizing Maps.* Springer Series in Information Sciences, Vol. 30, Springer, Berlin.

Lorenz, E. 1963. *Deterministic non-periodic flow*. Journal of Atmospheric Sciences 20: 130-141.

Malmgren, B. A. and Winter, A. 1999. *Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network*. Journal of Climate, 12:977–85

Mannila, H. 1997. *Methods and problems in data mining*. In Proc. of International Conference on Database Theory, Delphi, Greece.

Michalski, R., Carbonell, J., Mitchell, T. 1983. *Machine learning I: An AI Approach*. Morgan Kaufmann, Los Altos, CA.

Michalski, R.S., Bratko, I., Kubat, M. 1998. *Machine Learning and Data Mining, Methods and Applications*. John Wiley & Sons Ltd, West Sussex, England.

Monahan, A. 2000. *Nonlinear principal component analysis by neural networks: Theory and applications to the Lorenz system*. Journal of Climate 13: 821-835.

Perichinsky, G., García-Martínez, R. 2000. *A Data Mining Approach to Computational Taxonomy*. Proceedings Argentine Computer Science Researchers Worksop: 107-110.

Perichinsky, G., Servetto, A., García-Martínez, R., Orellana, R., Plastino, A. 2003. *Taxomic Evidence Applying Algorithms of Intelligent Data Minning Asteroid Families.* Proceedings de la International Conference on Computer Science, Software Engineering, Information Technology, e-Bussines & Applications 308-315.

Piatetski-Shapiro, G., Frawley, W., Matheus, C. 1991. *Knowledge discovery in databases: an overview*. AAAI-MIT Press, Menlo Park, California.

Piatetsky-Shapiro, G., Fayyad, U.M., Smyth, P. 1996. *From data mining to knowledge discovery*. AAAI Press/MIT Press, CA.

Quinlan, R. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers. San Mateo California.

Rich E. y Knight, K. 1991. *Introduction to Artificial Networks*. Mac Graw-Hill. Publications.

Santhanam M., and Patra, P. 2001. *Statistics of atmospheric correlations.* Physical Review E 64: 016102-1-1-7.

Setiono R. & Liu. H. 1996. *Symbolic representation of neural networks*. IEEE Computer Magazine 29(3): 71-77.

Stone, J., Porrill, J., Buchel, C., and Friston, K. 1999. *Spatial, temporal and spatiotemporal independent component analysis of fMRI data*. In 18th Leed Statistical Research Workshop on Spatiotemporal Mdeling and its Applications. University of Leeds.

Tian, B., Shaikh, M. A., Azimi Sadjadi, M. R., Vonder Haar, T. H., and Reinke, D. L. 1999. *Study of cloud classification with neural networks using spectral and textural features*. IEEE Transactions on Neural Networks, 10(1):138–151

Tirri, H. 1991. *Implementing Expert System Rule Conditions by Neural Networks*. New Generation Computing. 10(1): 55-71.

Wilks, D. 1995. *Statistical methods in Atmospheric Sciences*. Academic Press, London.

Yao X. y Liu Y. 1998. *Toward Designing Artificial Neural Networks by Evolution*. Applied Mathematics & Computation 91(1): 83-90.