We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

8

Recurrent Self-Organizing Map for Severe Weather Patterns Recognition

José Alberto Sá, Brígida Rocha, Arthur Almeida and José Ricardo Souza
Federal University of Pará (UFPA)
Brazil

1. Introduction

Weather patterns recognition is very important to improve forecasting skills regarding severe storm conditions over a given area of the Earth. Severe weather can damage electric and telecommunication systems, besides generating material losses and even losses of life (Cooray et al., 2007; Lo Piparo, 2010; Santos et al., 2011). In especial for the electrical sector, is strategic to recognize weather patterns that may help predict weather caused damages. Severe weather forecast is crucial to reduce permanent damage to the systems equipments and outages in transmission or distribution lines.

This study aimed to evaluate the temporal extensions applicability of Self-Organizing Map (Kohonen, 1990, 2001) for severe weather patterns recognition over the eastern Amazon region, which may be used in improving weather forecasting and mitigation of the risks and damages associated.

A large part of this region is located at low latitudes, where severe weather is usually associated with the formation of peculiar meteorological systems that generate a large amount of local rainfall and a high number of lightning occurrences. These systems are noted for their intense convective activity (Jayaratne, 2008; Williams, 2008).

Convective indices pattern recognition has been studied by means of neural network to determine the best predictors among the sounding-based indices, for thunderstorm prediction and intensity classification (Manzato, 2007). The model was tested for the Northern Italy conditions (Manzato, 2008). Statistical regression methods have also been used for radiosonde and lightning observations data obtained over areas of Florida Peninsula in the U. S. A. (Shafer & Fuelberg, 2006).

These important contributions to this area of study have shown that the applications should be pursued to find out the best predicting statistical tools. Moreover, the achieved skills are largely dependent on the hour of the sounding and the local climatic conditions. So far few studies have been carried out for the extremely moist tropical conditions, prevailing over the Amazon region, where the data for the case studies analyzed in this chapter were obtained.

In this context, the convective patterns recognition for the Amazon region may be used in local weather forecast. These forecasts are subsidiary elements in decision-making regarding preventive actions to avoid further damage to the electrical system. These outages lead to

productivity and information losses in the industrial production processes, which contribute negatively to the composition of the electric power quality indices (Rakov & Uman, 2005).

This study sought to recognize severe weather indices patterns, starting from an atmospheric sounding database. It is known that the atmospheric instability may be inferred from available radiosondes atmospheric profiling data. The stability indices drawn from observed atmospheric conditions have been used to warn people of potential losses (Peppier, 1988). Thus, this work analyzed the capacity of the Self-Organizing Map (SOM) and two of its temporal extensions: Temporal Kohonen Map and Recurrent Self-Organizing Map (Chappell & Taylor, 1993; Koskela et al., 1998a, 1998b; Varsta et al., 2000; Varsta et al., 2001) for clustering and classification of atmospheric sounding patterns in order to contribute with the weather studies over the Brazilian Amazon. The option of using this type of neural network was due to the fact that it uses only the input parameters, making it ideal for problems where the patterns are unknown.

Although there are other temporal extensions of SOM, such as recursive SOM - RecSOM (Hammer et al., 2004; Voegtlin, 2002), Self-Organizing Map for Structured Data - SOMSD (Hagenbuchner et al., 2003) and Merge Self-Organizing Map – MSOM (Strickert & Hammer, 2005), all these of global context, the option in this work was to apply local context algorithms, leaving to future studies the application of global context algorithms in this knowledge area. It is also important to refer the existence of the recent studies on the TKM and RSOM networks (Cherif et al., 2011; Huang & Wu, 2010; Ni & Yin, 2009).

In summary, with the original SOM algorithm and its extensions TKM and RSOM; stability indices data (Peppier, 1988); multivariate statistical techniques (principal component analysis and k-means); confusion matrix (Han & Kamber, 2006) and receiver operating characteristics (ROC) analysis (Fawcett, 2006); it was possible to evaluate the usefulness of these recurrent neural networks for the severe weather patterns recognition.

2. SOM and temporal extensions (TKM and RSOM)

This section discusses the fundamental concepts of Self-Organizing Map (SOM) and the proposed changes to introduce temporal sequence processing by the SOM: Temporal Kohonen Map (TKM) and Recurrent Self-Organizing Map (RSOM).

2.1 Self-organizing map (SOM)

The SOM is a nonlinear algorithm used for data clustering and classification. This algorithm is characterized by the processing of static data, i.e., not considering the data timelines, with the output of this neural network dependent only on present input data (Kohonen, 1990, 2001). The SOM is a single-layer neural network, in which it is recorded the learning by algorithm. This layer usually has low dimension structure (1D or 2D).

The training of the SOM is based on unsupervised learning, by adjusting of prototypes, according to the distribution of the input data, performed as follows:

The weight vector of each unit in the map space V_O is compared to an input vector. A metric-based criterion is chosen to determine the unit that has the minimum distance (Best Matching Unit), i.e., the neuron with the most similar prototype is selected as winner or winning unit, according to equation 1:

$$b(t) = \arg\min_{i \in Vo}\left\{\left\|\mathbf{x}(t) - \mathbf{w}_i(t)\right\|\right\} \tag{1}$$

Where:

- $\mathbf{x}(t)$ is an input vector, at time t, from the input space V_I;
- $\mathbf{w}_i(t)$ is a prototype, at time t, from the map space V_O;
- $b(t)$ is the index (position) of the winner neuron, at time t.

The neurons of the SOM cooperate to receive future incoming stimuli in an organized manner around the winner neuron. The winner neuron will be the center of a topological neighbourhood where neurons help each other to receive input signals along the iterations of network training. Thus, after obtaining the winning neuron, its weights are adjusted to increase the similarity with the input vector, the same being done for the weights of its neighbours, by an update rule, according to equation 2:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \gamma(t)h_{ib}(t)(\mathbf{x}(t) - \mathbf{w}_i(t)) \tag{2}$$

Where:

- $\gamma(t)$ is a learning rate;
- $h_{ib}(t)$ is a neighbourhood function.

Usually, the learning rate $\gamma(t)$ is defined by equation 3:

$$\gamma(t) = \gamma_0\exp\left(-\frac{t}{\tau_1}\right) \tag{3}$$

Where:

- t is the number of iterations;
- γ_0 is the initial value of the learning rate (value between 0 and 1);
- τ_1 is the time constant.

The neural network decreases its ability to learn, gradually over time, in order to prevent the drastic change by new data, in the sedimented knowledge through several iterations. The time constant influences the network learning as follows: high τ_1 value generates long period of intensive learning.

The neighbourhood function in a SOM is a similar way to reproduce the interactions of biological neurons, which stimulate their neighbours, in decreasing order, by increasing the lateral distance between them. So, for the SOM, this feature is reproduced by the parameter $h_{ib}(t)$ that determines how each neuron will receive readjustment to gain the future input stimuli. The largest adjustments are applied to the winner neuron and its neighbours, and minors to the neurons further from the winner neuron, because this parameter decreases with increasing lateral distance. Normally it is used the Gaussian function to represent the rate of cooperation between the neurons, by equation 4:

$$h_{ib}(t) = \exp\left(-\frac{l_{ib}^2}{2\sigma(t)^2}\right) \tag{4}$$

Where:

- l_{ib} is the lateral distance between neurons *i* and *b*;
- $\sigma(t)$ is the effective width of the topological neighbourhood.

Considering that the effective width of the topological neighbourhood will diminish with time increasingly specialized network regions will be built for certain input patterns. Over the course of iterations the radius of a neighbourhood should be smaller, which implies lower $h_{ib}(t)$ values, over time, thereby resulting in a restricted and specialized neighbourhood. For this, the exponential function is usually used, according to equation 5:

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau_1}\right) \tag{5}$$

Where:

- σ_0 is the initial value of effective width;
- τ_1 is a time constant.

2.2 Temporal Kohonen Map (TKM)

The SOM was originally designed for the static data processing, but for the dynamic data patterns recognition, it becomes necessary to include the temporal dimension in this algorithm. A pioneer algorithm in this adaptation was the Temporal Kohonen Map - TKM (Chappell & Taylor, 1993). It introduces the temporal processing using the same update rule of the original SOM, just changing the way of choosing the winner neuron. It uses the neurons activation history, by equation 6:

$$V_i(t) = dV_i(t\text{-}1) - \frac{1}{2}\|\mathbf{x}(t) - \mathbf{w}_i(t)\|^2 \tag{6}$$

Where:

- $Vi(t)$ is the neuron activation, at time t;
- *d* is a time constant (value between 0 and 1);
- $\mathbf{x}(t)$ is a input vector, at time *t*;
- $\mathbf{w}_i(t)$ is a prototype, at time *t*.

A TKM algorithm flow diagram is displayed in Figure 1. The current activation of the neuron is dependent on previous activation.
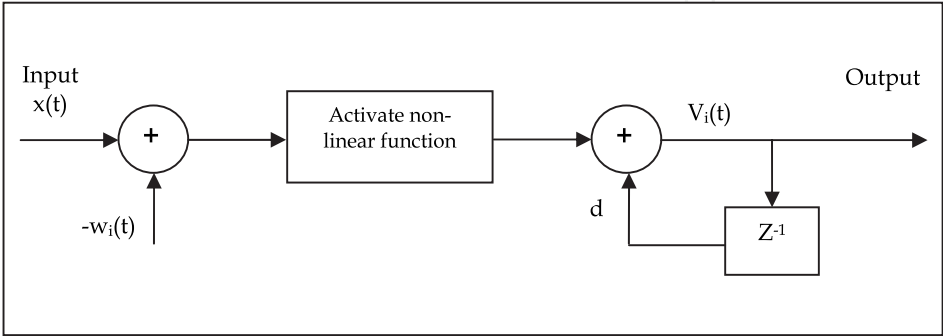

Fig. 1. TKM algorithm flow diagram

In the original SOM, for each new iteration, the value *1* is applied to the winner neuron and the value *0* to the other neurons. This creates an abrupt change in the neuron activation. In the TKM occurs a smooth change in the activation value (leaky integrator), because it uses the previous activation value, as shown in equation 6. In the TKM algorithm, the neuron that has the highest activation will be considered the winner neuron, according to the equation 7:

$$b(t) = \arg\max_{i \in Vo}\{V_i(t)\} \tag{7}$$

After choosing the winner neuron, the TKM network performs operations identical to the original SOM.

The basic differences between TKM and SOM networks are:

- For the determination of the winner neurons in TKM is necessary to calculate and record the activation $V_i(t)$, while in SOM is necessary to calculate the quantization error $\mathbf{x}(t)-\mathbf{w}_i(t)$;
- The winner neuron in TKM is one with greater activation $V_i(t)$, while in SOM is one with smallest quantization error $\mathbf{x}(t)-\mathbf{w}_i(t)$.

Interestingly, for *d*=0 the TKM network becomes equivalent to SOM network used for static data (Salhi et al., 2009).

2.3 Recurrent self-organizing map (RSOM)

Another algorithm that introduced the temporal processing to the SOM was the Recurrent Self-Organizing Map - RSOM using a new form of selection of the winner neuron and weights update rule (Koskela et al., 1998a, 1998b; Varsta et al., 2000; Varsta et al., 2001). This algorithm moved the leaky integrator from the unit outputs into the inputs. The RSOM allows storing information in the map units (difference vectors), considering the past input vectors, by equation 8:

$$\mathbf{y}_i(t) = (1-\alpha)\mathbf{y}_i(t-1) + \alpha(\mathbf{x}(t) - \mathbf{w}_i(t)) \tag{8}$$

Where:

- $\mathbf{y}_i(t)$ is called recursive difference of the neuron *i*, at time *t*;
- α is the leaking coefficient (value between 0 and 1).

Considering the term $\mathbf{x}(t)$ - $\mathbf{w}_i(t)$ with the quantization error, the winner neuron will be one that has the least recursive difference, i.e., the smallest sum of the present and past quantization errors, according to the equation 9:

$$b(t) = \arg\min_{i \in Vo}\{\|\mathbf{y}_i(t)\|\} \tag{9}$$

In the RSOM the weights update occur according to the equation 10:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \gamma(t)h_{ib}(t)\mathbf{y}_i(t) \tag{10}$$

Where:

- $\gamma(t)$ is a learning rate, at time t;
- $h_{ib}(t)$ is a neighbourhood function, at time t;
- $\mathbf{y}_i(t)$ is the recursive difference of the neuron i, at time t.

Thus, the RSOM takes into account the past inputs and also starts to remember explicitly the space-time patterns.

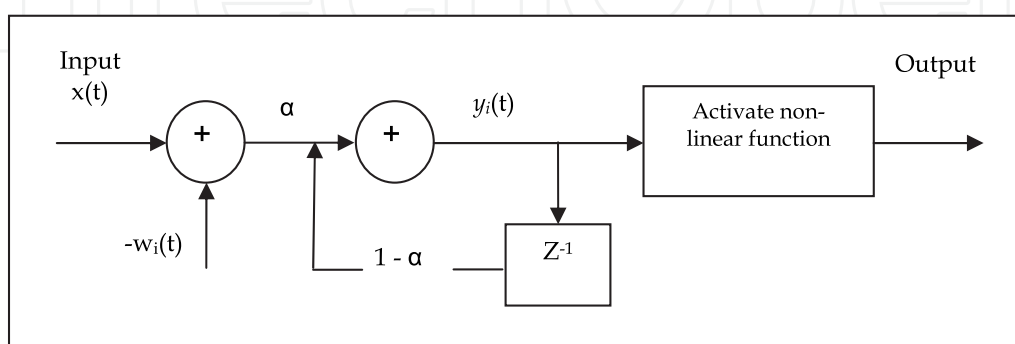A RSOM algorithm flow diagram is exhibited in Figure 2.

Fig. 2. RSOM algorithm flow diagram

The basic differences between RSOM and SOM networks are:

- For the determination of the winner neurons in RSOM is necessary to calculate and record the recursive difference $\mathbf{y}_i(t)$, while in SOM the choice criterion of the winner neurons is the quantization error;
- The winner neuron in RSOM is one with smallest recursive difference $\mathbf{y}_i(t)$, while in SOM is one with smallest quantization error.

To note that if $\alpha=1$ the RSOM network becomes identical to a SOM network (Salhi et al., 2009).

Angelovič (2005) discribes several advantages of the use RSOM for prediction systems. First, the small computing complexity, opposite to the global models. Then, the unsupervised learning. It allows building models from the data with only a little a priori knowledge.

3. Materials and methods

This section discusses the study area, data pre-processing and models training.

3.1 Study area and data pre-processing

The study used sounding data from weather station denominated SBBE, number 82193 (Belem airport), in the interval 2003 to 2010. The data were collected at the University of Wyoming website. The Figure 3 shows the station location, in the eastern Amazon Region, with their geographic coordinates (latitude: 1.38 S and longitude: 48.48 W).

3.1.1 Data characteristics

The sounding data are obtained by radiosondes transported by meteorological balloons. A radiosonde may determine various atmospheric parameters, such as atmospheric pressure,

temperature, dewpoint temperature, relative humidity, among others, in various atmospheric levels. These parameters are used to calculate sounding indices that seek to analyze the state of the atmosphere at a given time. Figure 4 shows an example of sounding indices collected from a radiosonde launched on January 1, 2010 at 12 h UTC.

Fig. 3. SBBE station localization (Belem airport).

For the evaluation of the atmospheric static stability, used for thunderstorms forecasting, several indicators have been developed (Peppier, 1988). Some indicators admit as instability factors the temperature difference and humidity difference between two pressure levels; while others, besides these factors, add the characteristics of the wind (speed and direction) at the same pressure levels. There are also indices based on the energy requirements for the occurrence of convective phenomena. Some indices and parameters used for the thunderstorms forecasting are: Showalter Index, K Index, Lifted Index, Cross Totals Index, Vertical Totals Index, Total Totals Index, SWEAT Index, Convective Inhibition, Convective Available Potential Energy, Level of Free Convection, Precipitable Water, among others.

```
                              Station identifier: SBBE
                                Station number: 82193
                              Observation time: 100101/1200
                               Station latitude: -1.38
                              Station longitude: -48.48
                              Station elevation: 16.0
                                Showalter index: 0.50
                                  Lifted index: -1.41
        LIFT computed using virtual temperature: -1.75
                                   SWEAT index: 208.81
                                       K index: 32.50
                             Cross totals index: 21.00
                          Vertical totals index: 22.10
                            Totals totals index: 43.10
              Convective Available Potential Energy: 198.35
                  CAPE using virtual temperature: 266.66
                          Convective Inhibition: -13.84
                  CINS using virtual temperature: -10.40
                                Equilibrum Level: 367.16
        Equilibrum Level using virtual temperature: 359.12
                        Level of Free Convection: 884.68
                  LFCT using virtual temperature: 893.44
                         Bulk Richardson Number: 20.60
               Bulk Richardson Number using CAPV: 27.70
        Temp [K] of the Lifted Condensation Level: 293.91
        Pres [hPa] of the Lifted Condensation Level: 955.56
              Mean mixed layer potential temperature: 297.76
                    Mean mixed layer mixing ratio: 16.44
                     1000 hPa to 500 hPa thickness: 5756.00
        Precipitable water [mm] for entire sounding: 56.66
```

Fig. 4. SBBE station information and sounding indices

3.1.2 Data selection

A selection algorithm was used to identify all 24 available atmospheric indices from radiosoundings performed at 12h UTC (9h Local Time) in the period analyzed. Subsequently, the indices calculated with virtual temperature were eliminated, leaving only 18 indices.

After normalization of these 18 indices by the standard deviation, the principal component analysis was used to reduce the number of variables. It was found that among the 18 principal components, the first three represented about 70% (seventy percent) of the total variance. Four variables related to severe weather conditions had considerable numerical values of the respective coefficients in the linear combinations of these principal components. Namely: SWEAT index (SWET), Convective Available Potential Energy (CAPE), Level of Free Convection (LFCT) and Precipitable Water (PWAT). Therefore, these elements were defined as the input vectors variables of the SOM, TKM and RSOM networks. Figure 5 shows a variance explained for principal components, until the ninth principal component.
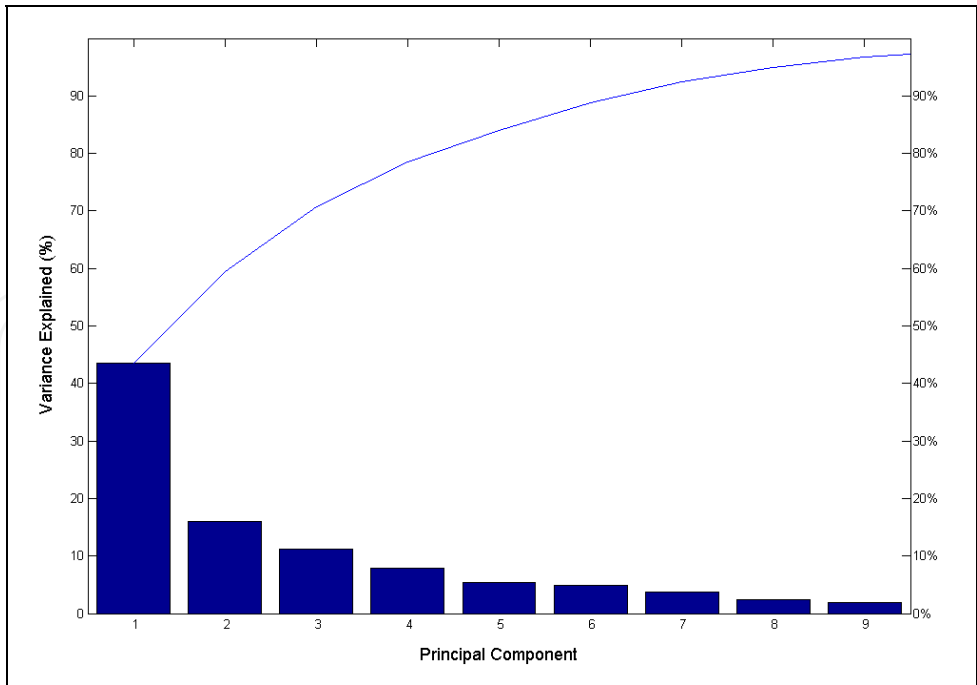
Fig. 5. Variance explained for principal components

The SWEAT index (or Severe Weather Threat Index) uses several variables (dewpoint, wind speed and direction, among others) to determine the likeliness of severe weather. The Convective Available Potential Energy (CAPE) is the integration of the positive area on a Skew-T sounding diagram. It exists when the difference between the equivalent potential temperature of the air parcel and the saturated equivalent potential temperature of the environment is positive. This means that the pseudo-adiabatic of the displaced air parcel is warmer than the environment (unstable condition). The Level of Free Convection (LFCT) is the CAPE region lower boundary. At this level a lifted air parcel will become equal in temperature to that of the environmental temperature. Once an air parcel is lifted to the LFCT it will rise all the way to the CAPE region top. The Precipitable Water (or Precipitable Water Vapor) is a parameter which gives the amount of moisture in the troposphere.

3.1.3 Data cleansing

The input vectors contained four variables: SWEAT index (SWET), Convective Available Potential Energy (CAPE), Level of Free Convection (LFCT) and Precipitable Water (PWAT). The vectors containing missing data from one or more variables were discarded. At the end of the cleansing process, a total of 1774 examples were obtained.

3.1.4 Data normalization

To reduce the discrepancies magnitude in the input vectors values, the min-max normalization was applied according to the equation 11. This transformed the original values of these input variables in normalized values within the range [0, 1].

$$value_{normalized} = \frac{value_{original} - \min A}{\max A - \min A} \tag{11}$$

Where:

- min A is the minimum value of the variable A;
- max A is the maximum value of the variable A.

3.1.5 Clusters formation for evaluation of the models

For the evaluation of the applicability of SOM and two of its temporal extensions: Temporal Kohonen Map (TKM) and Recurrent Self-Organizing Map (RSOM) for the weather patterns recognition related to atmospheric instability factors, clusters were built using the K-means technique, which generated three clusters, containing 697, 484 and 593 examples, for the cluster 1, 2 and 3, respectively. Figure 6 shows the characteristics of the three clusters according to the four variables analyzed.

In describing some of the differential characteristics between the clusters, it is noticed that in cluster 1 CAPE and PWAT have their concentrations at low values, while in cluster 2 the concentration of the CAPE is in low values, however for PWAT the values are high. In cluster 3 both CAPE and PWAT have their concentrations at high values. Another distinctive feature among clusters is the gradual rise of the LFCT median value for the clusters 1, 2 and 3, respectively. It is also noticed that the cluster 1 has a SWET median value lower when compared with clusters 2 and 3.
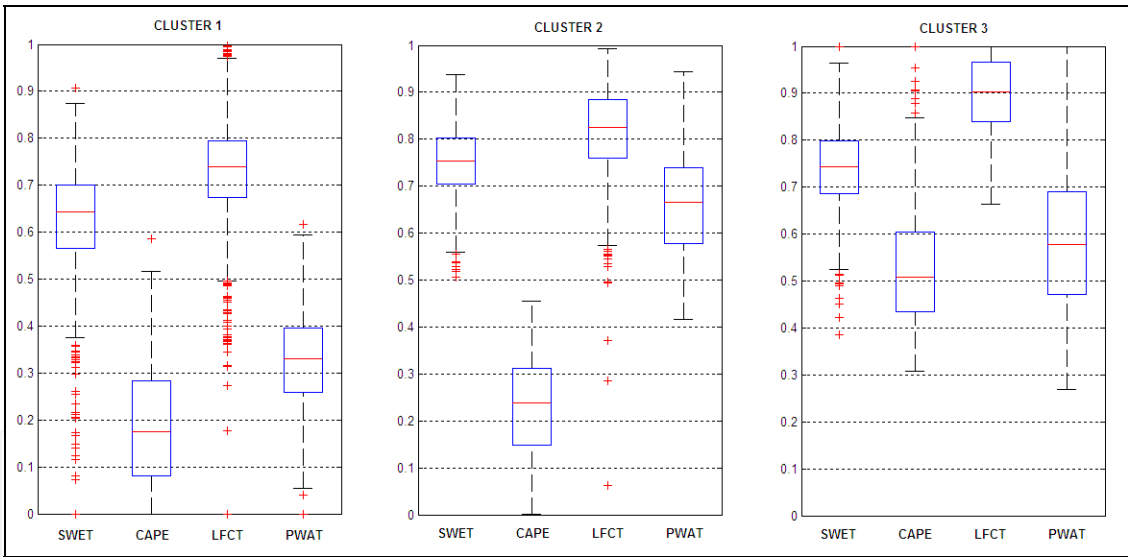
Fig. 6. Characteristics of the clusters

3.2 Training and evaluation of the models

For the performance analysis of the networks (SOM, TKM and RSOM) 3 maps were constructed, for each network type, in the grids: 5 x 5 units, 7 x 7 units and 9 x 9 units, therefore 9 maps in total.

Each network was trained with 600 examples extracted of the data set, with 200 examples of each cluster, randomly chosen. After training, the units of the maps were labeled according to their winning histories during training. This allowed that the networks were used as classifiers to evaluate their discrimination power.

The parameters used in the training of the SOM, TKM and RSOM networks were:

- Random weight initialization;
- Initial learning rate equal to 0.8;
- Final learning rate equal to 0.001;
- Number of epochs equal to 50.

Specifically for the TKM network was used the time constant d equal to 0.65 and for the RSOM network was used the leaking coefficient α equal to 0.35.

Subsequently it was evaluated the performance of the TKM and RSOM classifiers for different values of the constants α and d.

The results were presented in confusion matrices. In the confusion matrix each column represents the expected results, while each row corresponds to the actual results. During the simulation 1174 remaining examples of the data set were used.

After, a ROC analysis was done. The ROC graph is a technique for visualizing and evaluating classifiers based on their performance (Fawcett, 2006). A ROC graph allows identifying relative tradeoffs of a discrete classifier (one that your output is only a class label). In ROC graph the true positive rate (tp rate) of a classifier is plotted on the Y axis, while the false positive rate (fp rate) is plotted on the X axis. Fig. 7 shows a ROC graph with three classifiers labeled A through C.
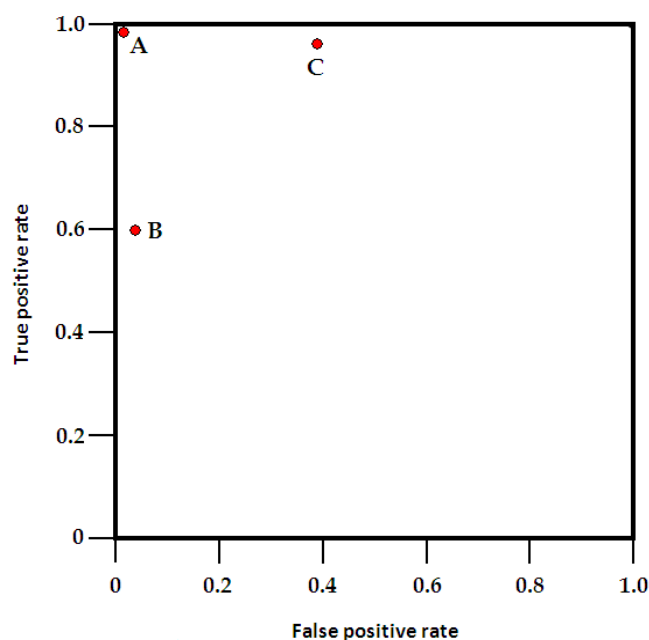
Fig. 7. ROC graph showing three discrete classifiers

Some points in ROC graph are very important. The lower left point (0, 0) represents a classifier that commits no false positive errors but also gains no true positives. The opposite situation is represented by the upper right point (1, 1). The upper left point (0, 1) represents a perfect classification (tp rate = 1 and fp rate = 0). In Fig 7, the point A is the ideal classifier; the point B represents a conservative classifier; and the point C represents a liberal classifier. Conservative classifiers make positive classifications only with strong evidence, i.e., their

false positive rates are low but they also have low true positive rates. Liberal classifiers have high true positive rates but they also have high false positive rates.

4. Results

This section presents the results of the assessment among the studied networks: Self-Organizing Map (SOM), Temporal Kohonen Map (TKM) and Recurrent Self-Organizing Map (RSOM) for the severe weather pattern classification.

4.1 Evaluation of the SOM, TKM and RSOM classifiers

Table 1 shows the confusion matrices and the global accuracy of the neural networks studied. It is noticed that with the TKM and RSOM classifiers were provided superior performances to the original SOM, and between the recurrent networks, the RSOM network showed the best results.

Fig. 8 exhibits the ROC graph for the SOM, TKM and RSOM when the grids of the neural networks are 5 x 5. One may notice that the RSOM presents a larger tp rate and a smaller fp rate for the labels 2 and 3. For the label 1, the SOM network presented itself as the most liberal, and the RSOM as the most conservative network. Therefore, for this grid, the results have indicated that the RSOM classifier has a better performance than the other networks analyzed in this work.

Grids	SOM classifier			TKM classifier			RSOM classifier		
5 x 5	447	39	11	415	67	15	441	53	3
	16	267	1	1	275	8	1	282	1
	28	51	314	8	24	361	2	22	369
	Global accuracy (87.6%)			Global accuracy (89.5%)			Global accuracy (93.0%)		
7 x 7	422	53	22	454	27	16	470	22	5
	4	272	8	21	245	18	3	280	1
	12	10	371	1	21	371	13	22	358
	Global accuracy (90.7%)			Global accuracy (91.1%)			Global accuracy (94.4%)		
9 x 9	434	42	21	464	21	12	473	19	5
	9	254	21	32	231	21	3	280	1
	11	15	367	12	27	354	7	17	369
	Global accuracy (89.9%)			Global accuracy (89.4%)			Global accuracy (95.6%)		

Legend:

confusion matrix		Predicted cluster		
		1	2	3
True cluster	1	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$
	2	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$
	3	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$

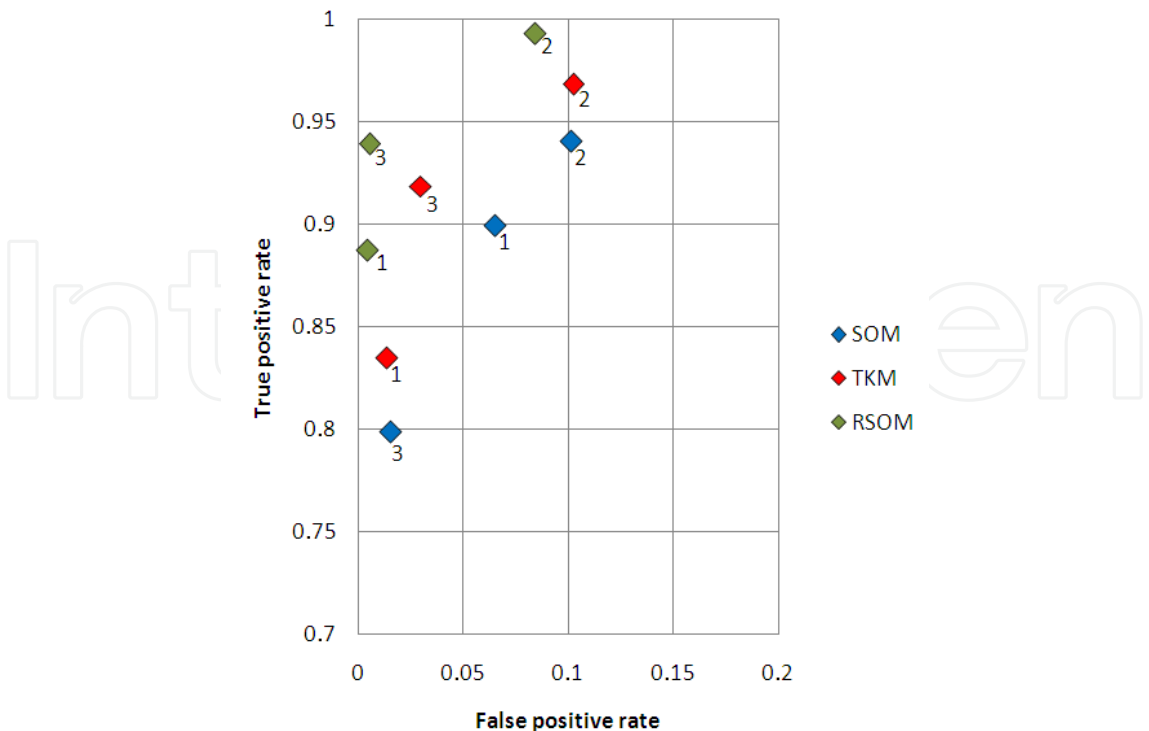Table 1. Evaluation of the SOM, TKM and RSOM classifiers in different grids

Fig. 8. ROC graph for the SOM, TKM and RSOM in 5x5 grid

Fig. 9 displays the ROC graph of the analyzed models for the 7x7 grid. It is evident that the RSOM network has a larger tp rate and a smaller fp rate for labels 1 and 2. For the label 3 the SOM and TKM networks presented similar liberal characteristics, while the RSOM network showed a more conservative behavior. For these dimensions the results also indicated a better performance of the RSOM classifier when compared to the SOM and TKM network options.
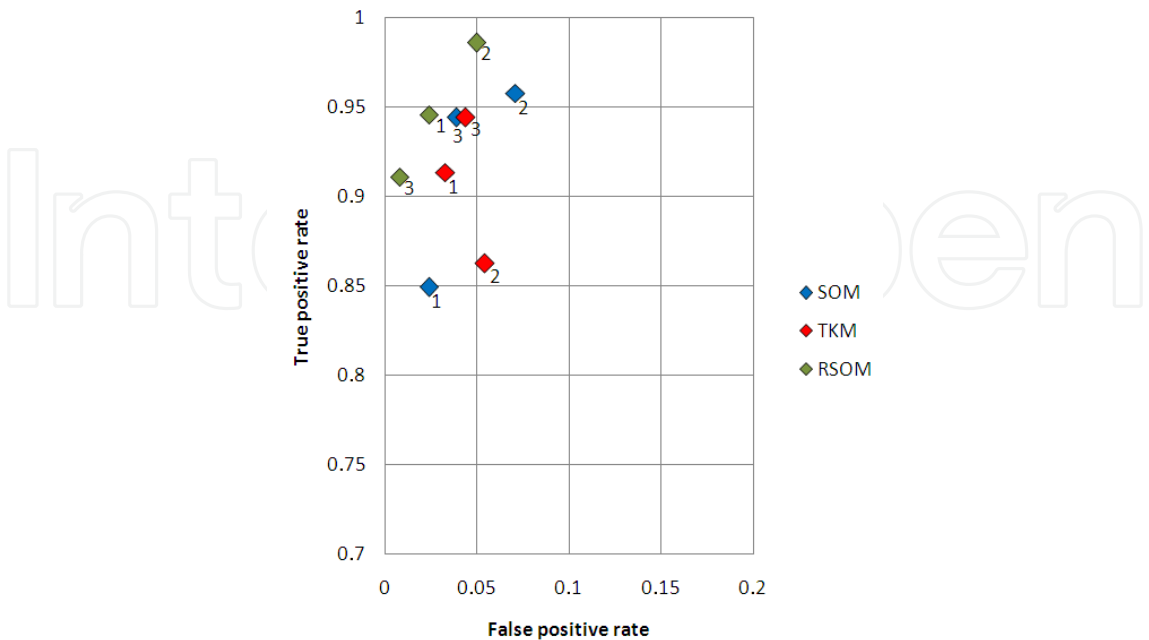
Fig. 9. ROC graph for the SOM, TKM and RSOM in 7x7 grid

The ROC graph for the SOM, TKM and RSOM in 9x9 grid is presented in Figure 10. One notices that for this grid, the RSOM network has a larger tp rate and smaller fp rate for all three labels considered. This fact confirms even more the best performance observed for the RSOM classifier, among all networks analyzed.
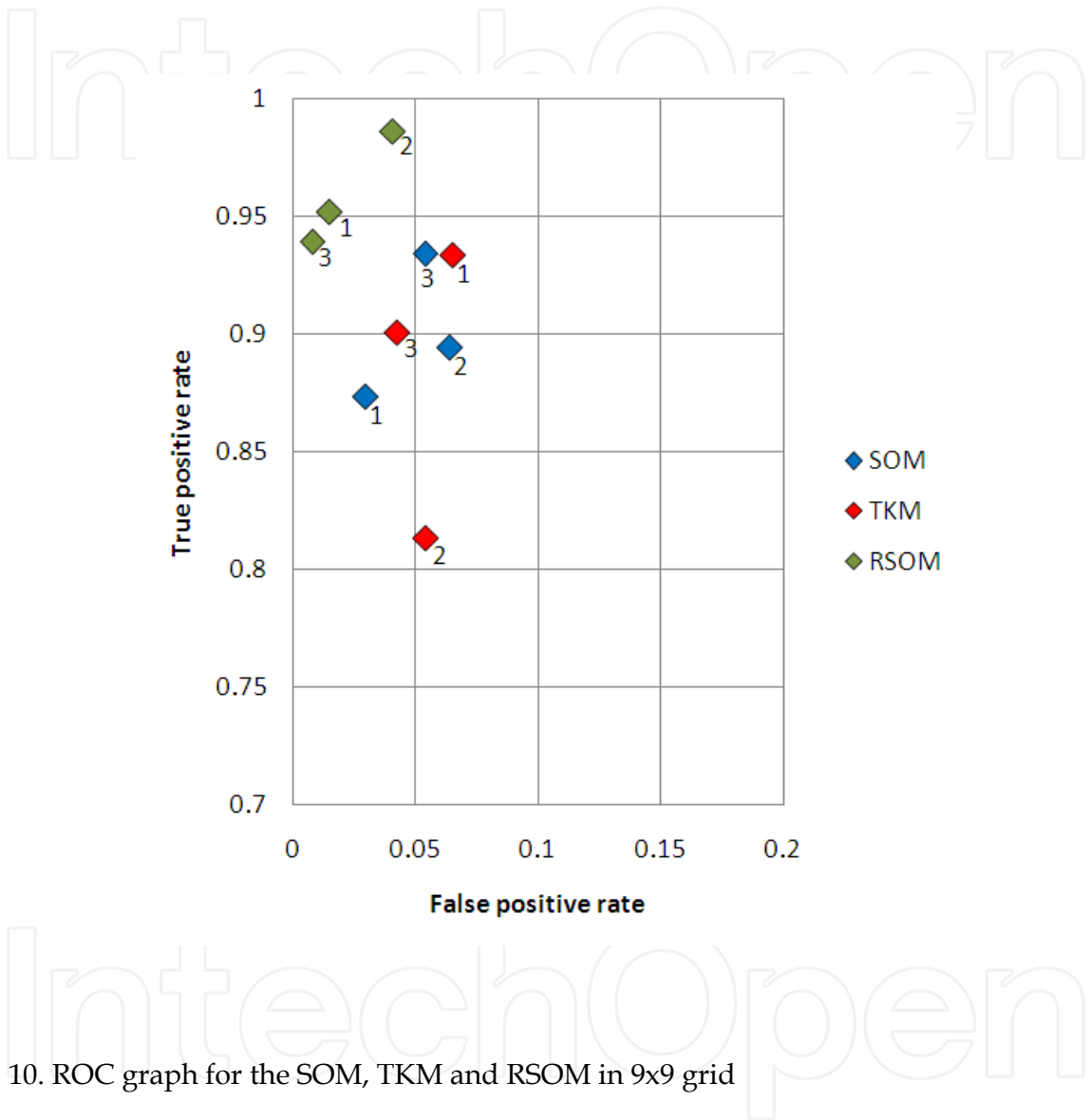
Fig. 10. ROC graph for the SOM, TKM and RSOM in 9x9 grid

4.2 U-matrix of the SOM, TKM and RSOM networks

Table 2 shows a comparison among the U-Matrices of the networks studied. The U-Matrices are representations of the self-organizing networks where the Euclidean distance between the codebook vector of the neighbouring neurons is represented in a two-dimensional color scale image.

It is observed that the RSOM network presented the best view among the networks studied, distinguishing clearly the existence of three clusters in the data set used for training this neural network.

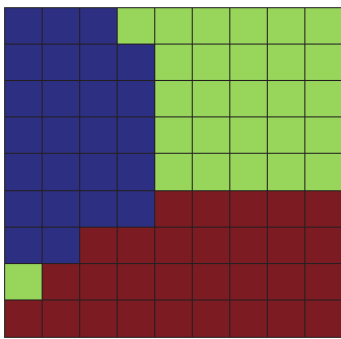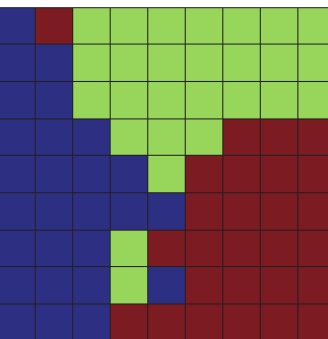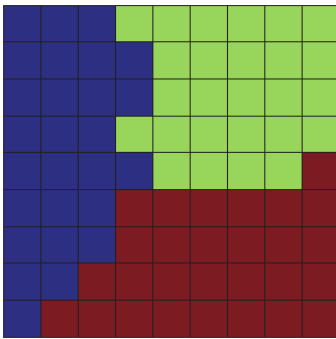| Grids	SOM network	TKM network	RSOM network
5 x 5			
7 x 7			
9 x 9			
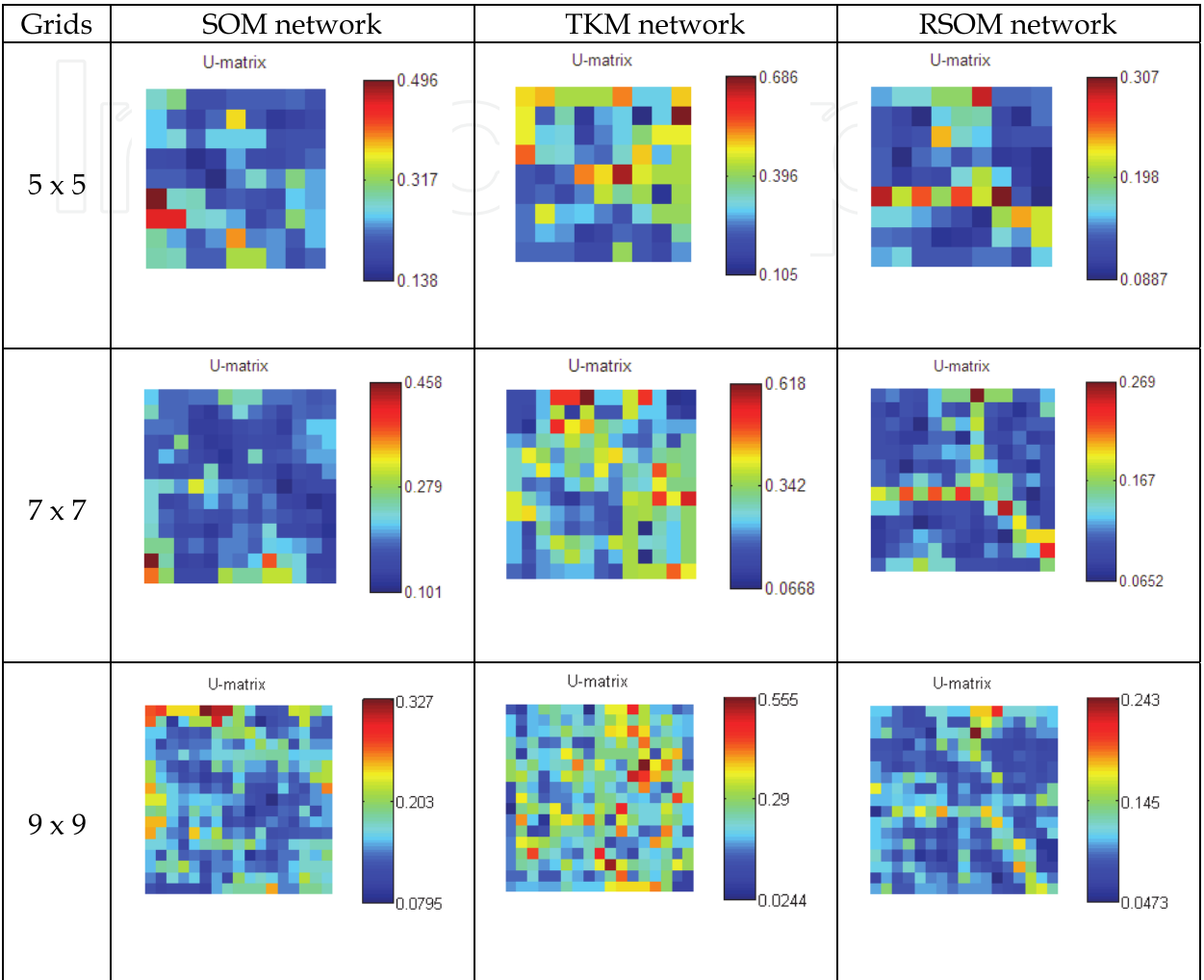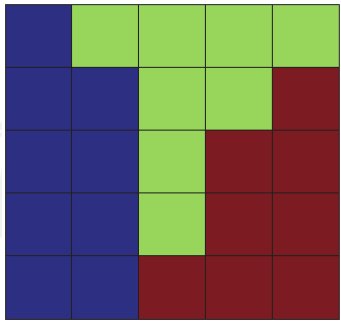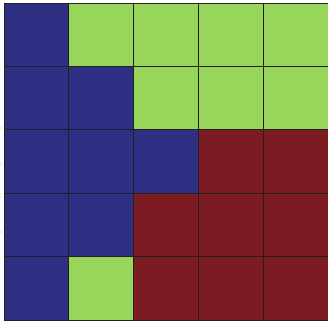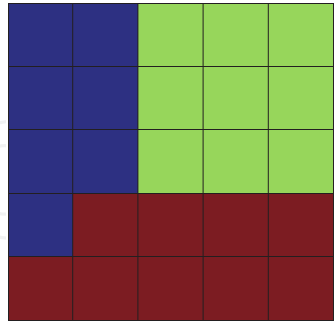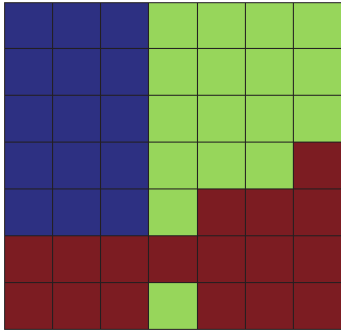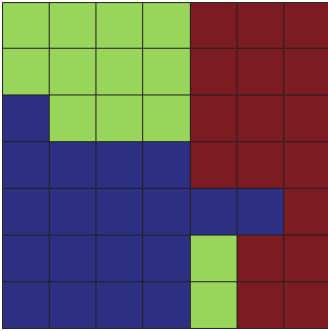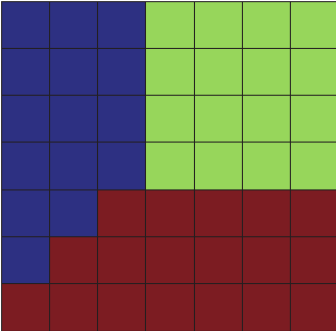
Legend: Color scale represent the Euclidean distance between the codebook vector of the neighbouring neurons

Table 2. U-matrix of the SOM, TKM and RSOM networks

4.3 Labeling of the neurons after the training of the SOM, TKM and RSOM networks

Table 3 shows a comparison between the labeling of the neurons after the training process, using as criteria the activation frequency. It is noticed that RSOM network has a higher organization when compared with the other networks. The labels: blue for the cluster 1, green for the cluster 2, and red for the cluster 3, were used.

Grids	SOM network	TKM network	RSOM network
5 x 5			
7 x 7			
9 x 9			

Legend: Blue for the cluster 1. Green for the cluster 2. Red for the cluster 3

Table 3. Labeling of the SOM, TKM and RSOM networks neurons in different grids

4.4 Time constants variation of the TKM and RSOM classifiers

One difference between the SOM network and its temporal extensions TKM and RSOM is the change in the performance when occur variation in the time constants. In the section 4.4.1 and 4.4.2 are shown the results when the time coefficients (d and α) vary in the range 0 to 1.

4.4.1 *d* variation

For different *d* values the TKM network presented different global accuracies, reducing their values in the range limits of 0 to 1. The table 4 shows the confusion matrices and Figure 11 shows the superposition of the global accuracies due to the *d* variation. For each TKM dimension studied the points were spaced at 0.25 intervals.

Grid: 5 x 5 units											
d = 0.10			d = 0.35			d = 0.60			d = 0.85		
433	47	17	446	35	16	440	48	9	387	39	71
0	276	8	15	266	3	55	218	11	50	233	1
4	57	332	2	43	348	20	22	351	0	81	312
Global accuracy (88.7%)			Global accuracy (90.3%)			Global accuracy (85.9%)			Global accuracy (79.4%)		
Grid: 7 x 7 units											
d = 0.10			d = 0.35			d = 0.60			d = 0.85		
448	38	11	455	33	9	412	70	15	372	23	102
5	251	28	10	262	12	0	280	4	62	215	7
16	15	362	4	13	376	1	51	341	18	83	292
Global accuracy (90.4%)			Global accuracy (93.1%)			Global accuracy (88.0%)			Global accuracy (74.9%)		
Grid: 9 x 9 units											
d = 0.10			d = 0.35			d = 0.60			d = 0.85		
455	29	13	448	37	12	456	37	4	406	31	60
11	258	15	21	251	12	7	276	1	23	249	12
17	34	342	12	15	366	10	24	359	16	54	323
Global accuracy (89.9%)			Global accuracy (90.7%)			Global accuracy (92.9%)			Global accuracy (83.3%)		

Legend:

confusion matrix	Predicted cluster		
	1	2	3
True cluster 1	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$
2	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$
3	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$

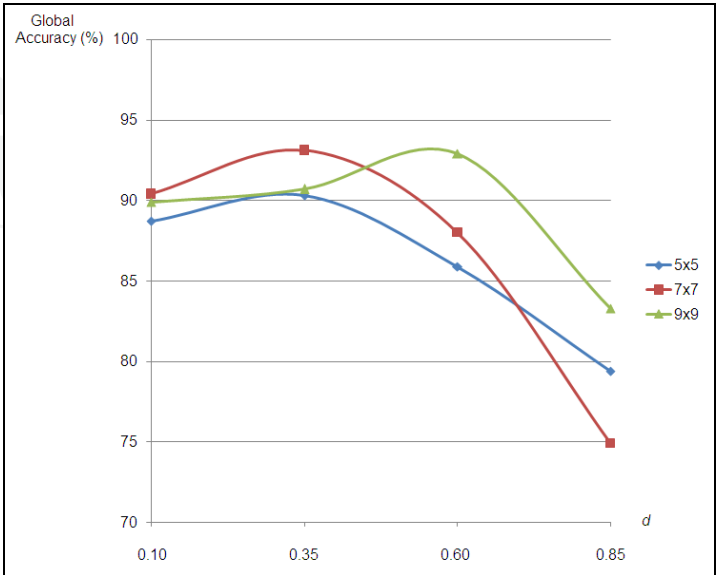Table 4. *d* variation and global accuracy of the TKM network in different grids

Fig. 11. Superposition of the TKM networks global accuracies due to the *d* variation

Figure 12 shows the ROC graph of the TKM model with 5x5 map units and *d* variations. It indicates that for lower values of *d* (*d*=0.10 and *d*=0.35) this classifier presented more conservative characteristics for labels 1 and 3, and the most liberal behaviour for the label 2. On the other hand, for higher values of *d* (*d*=0.60 and *d*=0.85), in general one notices a decrease of the tp rate values and an increment of the fp rate for all labels. One may conclude therefore, that the TKM better performances were observed for the lower values of *d*.
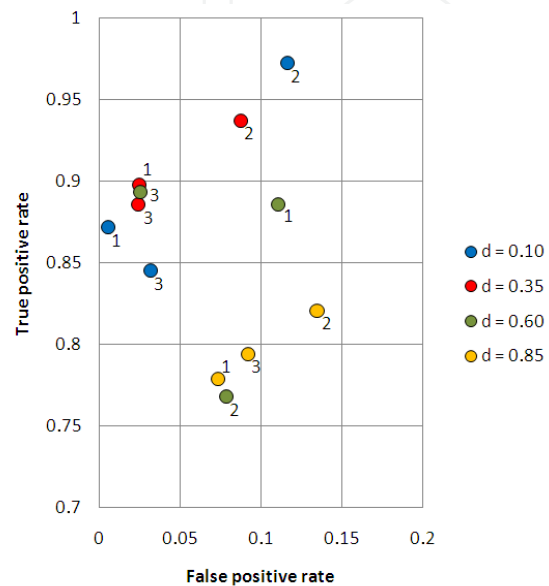
Fig. 12. ROC graph for the TKM model in 5x5 grid with *d* variation

Figure 13 displays a ROC graph for the TKM model with 7x7 grid and *d* variation. In this particular case, it is even more evident the superior performance of this classifier when one uses the lower values of d. Indeed, its best performance was found for *d*=0.35 and the worst corresponded to *d*=0.85.
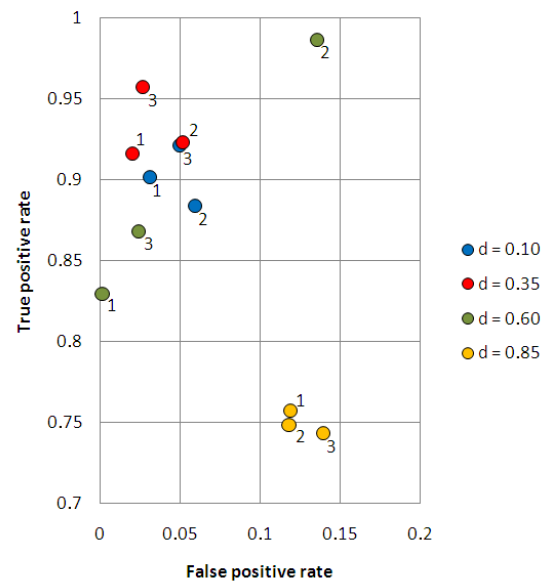
Fig. 13. ROC graph for the TKM model in 7x7 grid with *d* variation

The results relative to the TKM model with 9x9 grid and *d* variation are graphically displayed in Figure 14. For this case, one may notice an approximation among the performances of the model for *d*=0.60 and the results obtained for lower values of d, such as the cases for *d*=0.10 and *d*=0.35. This togetherness was also observed for the *d*=0.85 case, even though it remains as the worst performance case for the TKM model. Therefore, the conclusion was that the smaller values of d provided the best performances for the clusters classification by this network type.
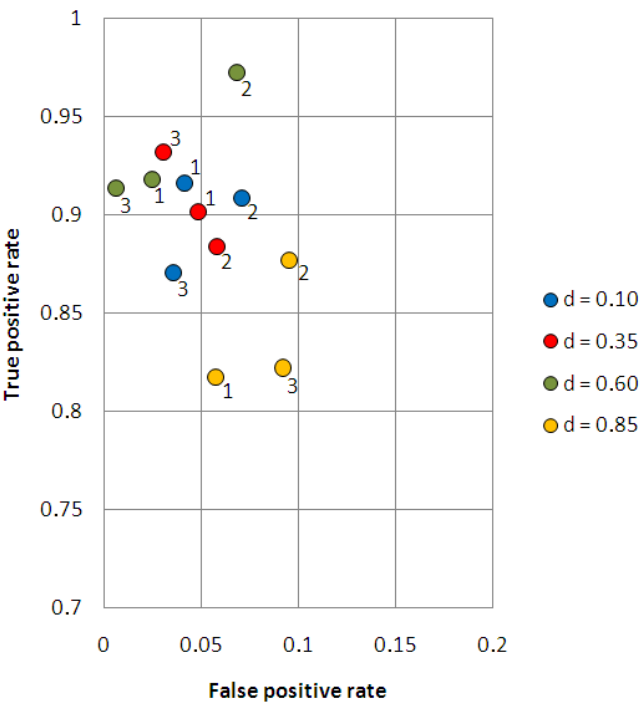
Fig. 14. ROC graph for the TKM model in 9x9 grid with *d* variation
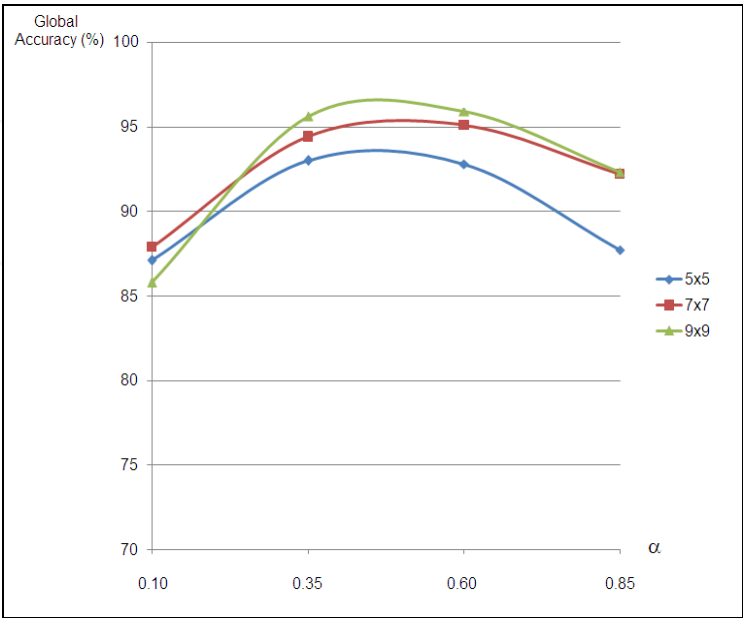
Fig. 15. Superposition of the RSOM networks global accuracies due to the α variation

4.4.2 α variation

For different α values the RSOM network presented a global accuracy significant variation. Table 5 shows the confusion matrices and figure 15 shows the superposition of the global accuracies due to the α variation, for each RSOM dimension studied.

Figure 16 shows a ROC graph for the RSOM model with 5x5 grid and α variation. One notices that with intermediate values of α (α=0.35 and α=0.60) the classifier presented the greatest performances (larger tp rate and smaller fp rate).

Figure 17 shows a ROC graph for the RSOM model with 7x7 grid and α variation. There is an approximation of the performances in this network, between the extreme values (α=0.10 and α=0.85) and the intermediate values of α (α=0.35 and α=0.60). Despite this fact, the intermediate α values still presented a general tendency to be the best performers, with larger tp rate and smaller fp rate.

Grid: 5 x 5 units											
α = 0.10			α = 0.35			α = 0.60			α = 0.85		
437	20	40	441	53	3	455	37	5	445	42	10
0	244	40	1	282	1	1	283	0	35	248	1
45	7	341	2	22	369	6	36	351	4	52	337
Global accuracy (87.1%)			Global accuracy (93.0%)			Global accuracy (92.8%)			Global accuracy (87.7%)		

Grid: 7 x 7 units											
α = 0.10			α = 0.35			α = 0.60			α = 0.85		
425	72	0	470	22	5	458	28	11	457	35	5
1	260	23	3	280	1	4	279	1	4	280	0
43	3	347	13	22	358	1	12	380	13	35	345
Global accuracy (87.9%)			Global accuracy (94.4%)			Global accuracy (95.1%)			Global accuracy (92.2%)		

Grid: 9 x 9 units											
α = 0.10			α = 0.35			α = 0.60			α = 0.85		
440	57	0	473	19	5	465	24	8	458	29	10
4	259	21	3	280	1	6	277	1	15	260	9
49	36	308	7	17	369	0	9	384	11	16	366
Global accuracy (85.8%)			Global accuracy (95.6%)			Global accuracy (95.9%)			Global accuracy (92.3%)		

Legend:

confusion matrix		Predicted cluster		
		1	2	3
True cluster	1	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$
	2	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$
	3	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$

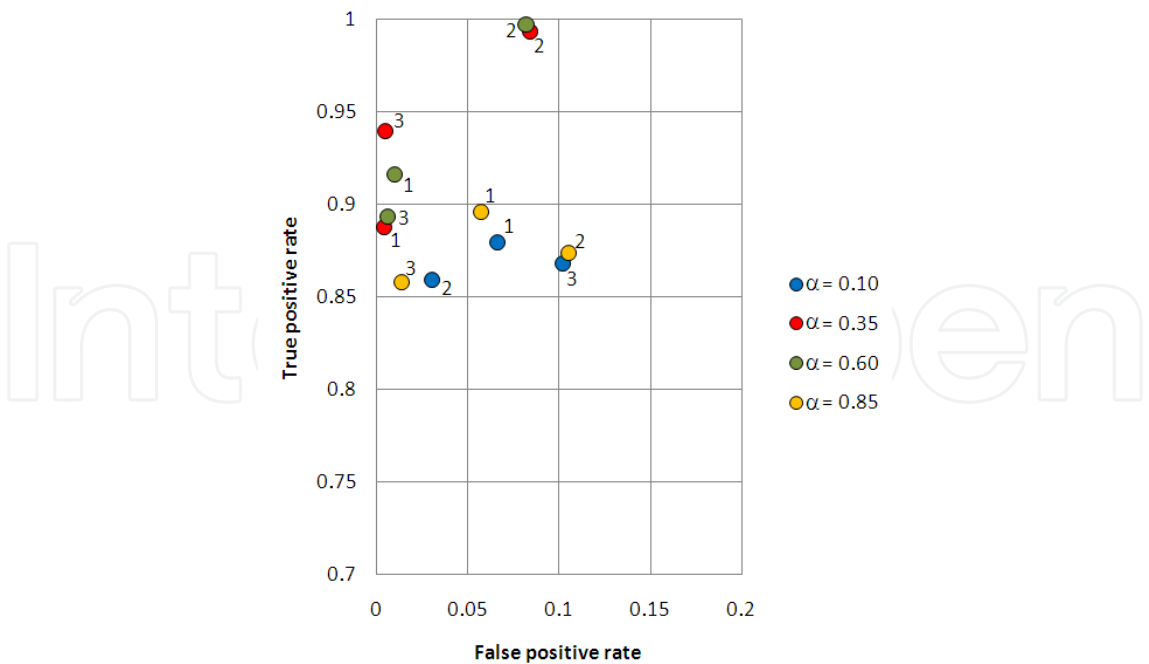Table 5. α variation and global accuracy of the RSOM network in different grids

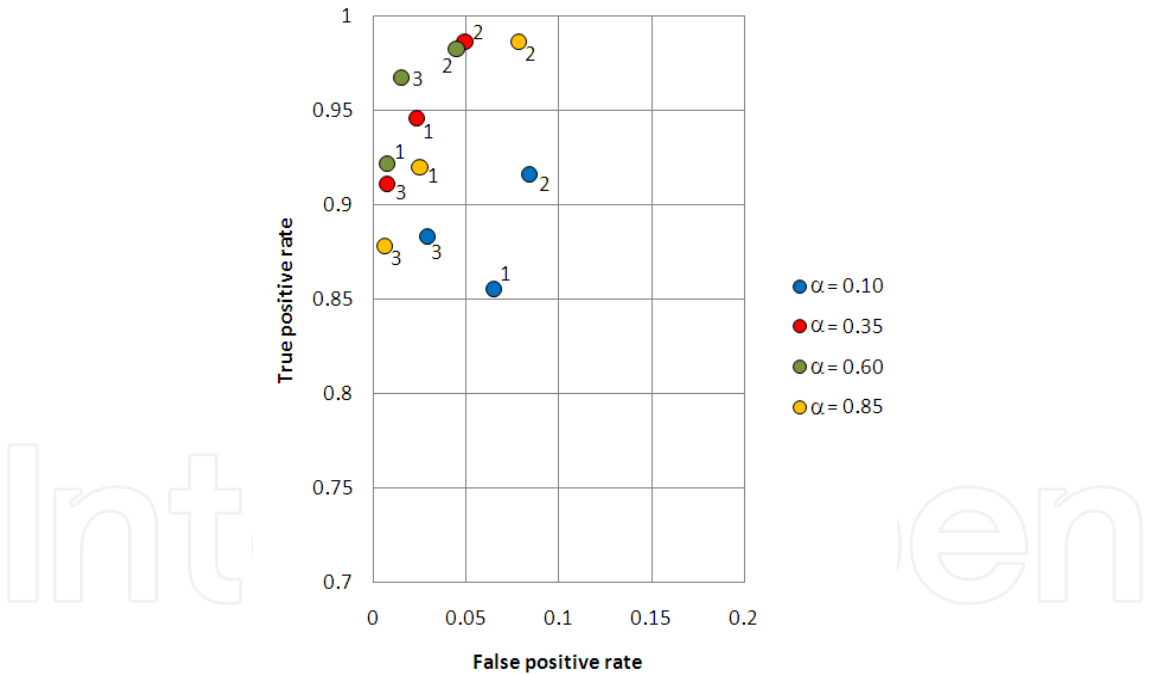Fig. 16. ROC graph for the RSOM model in 5x5 grid with α variation

Fig. 17. ROC graph for the RSOM model in 7x7 grid with α variation

Figure 18 shows a ROC graph for the RSOM model with 9x9 grid and α variation. This figure indicates a superior performance of this classifier, when one uses intermediate values of α (α=0.35 and α=0.60). In such cases the classifier becomes nearly ideal, with tp rate approaching 100% and fp rate near 0%. On the other hand, this classifier performance becomes very poor for the lowest extreme (α=0.10) when compared to the other α parameters.
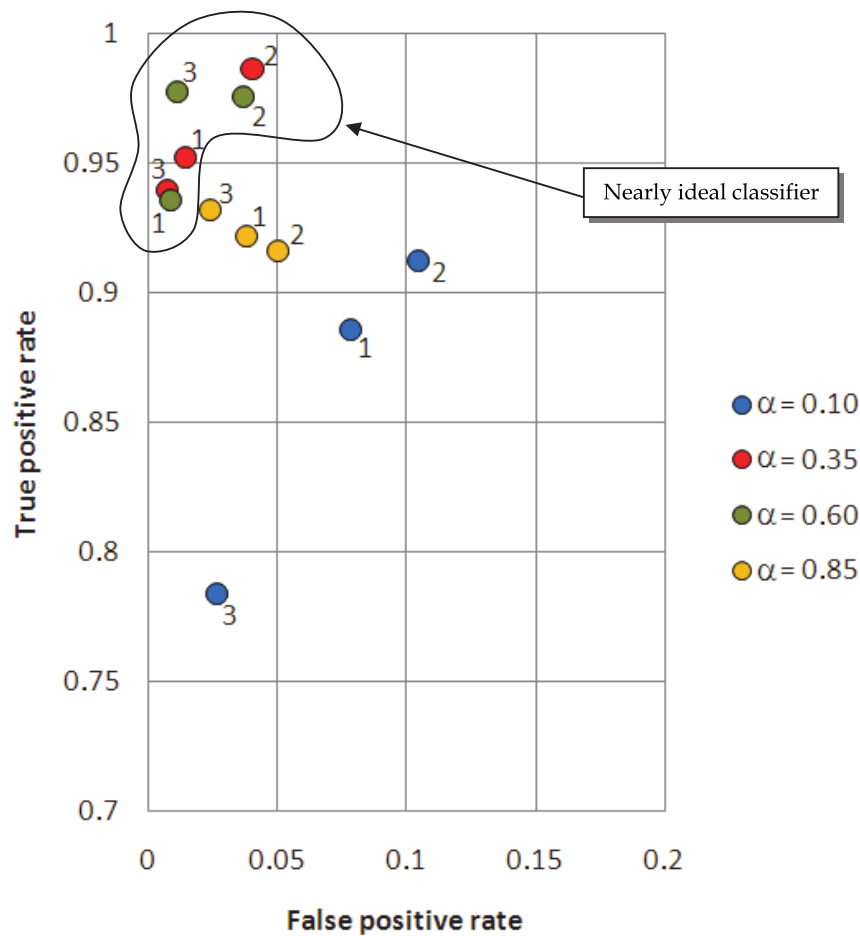
Fig. 18. ROC graph for the RSOM model in 9x9 grid with α variation

In summary, we can say that the RSOM network was the one which offered the best clusters classification performance, as long as, using intermediate α values.

5. Conclusion

This work aimed to evaluate the applicability of the self-organizing map local temporal extensions (Temporal Kohonen Map and Recurrent Self-Organizing Map) in the severe weather patterns recognition. The study area was the eastern Amazon region, one of the areas of the Earth with higher frequency and intensity of extreme weather events, especially in terms of the lightning occurrences.

Its purpose was to contribute in the identification of a useful tool for the weather studies, to reduce the damages associated with this natural phenomenon. The performance analysis of the recurrent neural networks (TKM and RSOM) with relation to the original SOM resulted in the following main conclusions:

- There is significant change in the global accuracy of the TKM and RSOM classifiers depending on the choice of the time coefficients (d and α), respectively;
- The recurrent neural networks (TKM and RSOM), used as classifiers, presented improved performance over the original SOM when adjusted its time coefficients with core values within the range of 0 to 1;

- The U-matrix of the RSOM network presented a better cluster visualization (in terms of separation) when compared with other networks studied, allowing a clear differentiation among the three clusters;
- The labeling of the neurons in the maps, after the training, was better defined for the RSOM network when compared with the other networks studied;
- Finally, after the ROC analysis, it was concluded that among the neural networks studied, the RSOM had the best performance as classifier, confirming its usefulness as a potential tool for studies related to the severe weather patterns recognition.

6. References

Angelovič, P. (2005). Time series prediction using RSOM and local models, *Proceedings of IIT. SRC 2005*, pp. 27–34, April 27, 2005

Chappell, G. & Taylor, J. (1993). The temporal Kohonen map. *Neural Networks*, Vol.6, No.3, pp. 441-445, ISSN: 0893-6080

Cherif, A.;Cardot, H. & Bone, R. (2011). SOM time series clustering and prediction with recurrent neural networks. *Neurocomputing*, Vol.74, Issue 11, (May 2011), pp. 1936-1944, ISSN 0925-2312, doi: 10.1016/j.neucom.2010.11.026

Cooray, V.; Cooray, C. & Andrews, C. (2007). Lightning caused injuries in humans. *Journal of Electrostatics*, Vol.65, Issues 5-6, (May 2007), pp. 386-394, ISSN 0304-3886, doi:10.1016/j.elstat.2006.09.016

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, Vol.27, Issue 8, (June 2006), pp. 861-874, ISSN 0167-8655, doi:10.1016/j.patrec.2005.10.010

Hagenbuchner, M.; Sperduti, A. & Tsoi, A. (2003). A self-organizing map for adaptive processing of structured data. *IEEE Transactions on Neural Networks*, Vol.14, No.3, pp.491–505, ISSN: 1045-9227

Hammer, B.; Micheli, A.; Sperduti, A. & Strickert, M. (2004). Recursive self-organizing network models. *Neural Networks*, Vol. 17, Issues 8-9, New Developments in Self-Organizing Systems (October-November 2004), pp. 1061-1085, ISSN 0893-6080, doi: 10.1016/j.neunet.2004.06.009

Han, J. & Kamber, M. (2006). *Data Mining: Concepts and Techniques*, 2nd ed., Morgan Kaufmann Publishers, San Francisco, USA, ISBN: 978-1-55860-901-3

Huang, S.-C. & Wu, T.-K. (2010). Integrating recurrent SOM with wavelet-based kernel partial least square regressions for financial forecasting. *Expert Systems with Applications*, Vol.37, Issue 8, (August 2010), pp. 5698-5705, ISSN 0957-4174, doi: 10.1016/j.eswa.2010.02.040

Jayaratne, R. (2008). Thunderstorm electrification mechanisms, In: *The Lightning Flash*, V. Cooray, (Ed.), pp. 17-44, Institution of Engineering and Technology,ISBN 978-0-85296-780-5, London, United Kingdom

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, Vol.78, No.9, (September 1990), pp. 1464-1480, ISSN 0018-9219

Kohonen, T. (2001). *Self-organizing maps*, Springer Series in Information Sciences, ISBN 3-540-67921-9, Berlin, Germany

Koskela, T.; Varsta, M.; Heikkonen, J. & Kaski, K. (1998a). Recurrent SOM with local linear models in time series prediction, *Proceedings of Sixth European Symposium on Artificial Neural Networks*, pp. 167–172

Koskela, T.; Varsta, M.; Heikkonen, J. & Kaski, K. (1998b). Time series prediction using recurrent SOM with local linear models, *International Journal of Knowledge-based and Intelligent Engineering Systems*, Vol.2, pp. 60–68, ISSN: 1327-2314

Lo Piparo, G. (2010). Lightning protection of telecommunication towers, In: *Lightning Protection,* V. Cooray, (Ed.), pp. 723-787, Institution of Engineering and Technology, ISBN 978-0-86341-744-3, London, United Kingdom

Manzato, A. (2007). Sounding-derived indices for neural network based short-term thunderstorm and rainfall forecasts. *Atmospheric Research*, Vol.83, Issues 2-4, (February 2007), pp. 349-365, ISSN 0169-8095, doi:10.1016/j.atmosres.2005.10.021

Manzato, A. (2008). A verification of numerical model forecasts for sounding-derived indices above Udine, Northeast Italy. *Weather and Forecasting - American Meteorological Society,* Vol.23, (June 2008), pp. 477-495, doi: 10.1175/2007WAF2007018.1

Ni, H. & Yin, H. (2009). Exchange rate prediction using hybrid neural networks and trading indicators. *Neurocomputing*, Vol.72, Issues 13-15, (August 2009), pp. 2815-2823, ISSN 0925-2312, doi: 10.1016/j.neucom.2008.09.023

Peppier, R. (1988). *A Review of Static Stability Indices and Related Thermodynamic Parameters.* Illinois State Water Survey Climate and Meteorology Section, USA

Rakov, V. & Uman, M. (2005). *Lightning: Physics and Effects*, Cambridge University Press, ISBN 978-0-521-58327-5, New York, USA

Salhi, M. S.; Arous, N. & Ellouze, N. (2009). Principal temporal extensions of SOM: Overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition* , Vol.2, (December 2009), pp. 61-84, ISSN 2005-4270

Santos, A. ; Souza, J. ; Souza, E. ; Carmo, A. ; Ribeiro, W. & Oliveira, R. (2011). The climatic context of lightning storms, associated to electric systems power outages in eastern Amazonia, *Proceedings of XIV International Conference on Atmospheric Electricity-ICAE*, Rio de Janeiro, Brazil, August 7-12, 2011

Shafer, P. E. & Fuelberg, H. E. (2006). A statistical procedure to forecast warm season lightning over portions of the Florida peninsula. *Weather and Forecasting - American Meteorological Society,* Vol.21, (October 2006), pp. 851-868, ISSN 0882-8156, doi: 10.1175/WAF954.1

Strickert, M. & Hammer, B. (2005). Merge SOM for temporal data. *Neurocomputing*, Vol.64, (March 2005), pp. 39-71, ISSN 0925-2312, doi: 10.1016/j.neucom.2004.11.014

Varsta, M.; Heikkonen, J. & Lampinen, J. (2000). Analytical comparison of the temporal Kohonen map and the recurrent self organizing map. In: M. Verleysen (Ed.),273–280, ESANN'2000

Varsta, M. ; Heikkonen, J. ; Lampinen, J. & Millan, J. (2001). Temporal Kohonen map and the recurrent self-organizing map: Analytical and experimental comparison. *Neural Processing Letters,* Vol.13, No.3, pp. 237 – 251, ISSN 1370-4621

Voegtlin, T. (2002). Recursive self-organizing maps. *Neural Networks,* Vol.15, Issues 8–9,(October 2002), pp. 979–991, doi :10.1016/S0893-6080(02)00072-2

Williams, E. (2008). Charge structure and geographical variation of thunderclouds, In: *The Lightning Flash,* V. Cooray, (Ed.), pp. 1-15, Institution of Engineering and Technology, ISBN 978-0-85296-780-5, London, United Kingdom
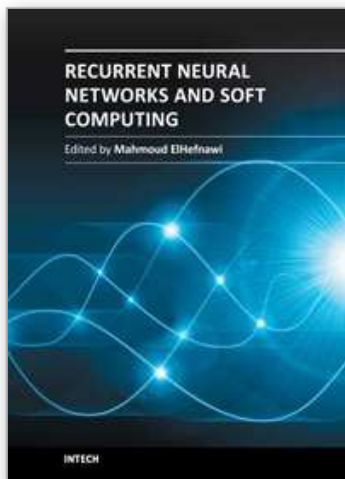
Recurrent Neural Networks and Soft Computing

Edited by Dr. Mahmoud ElHefnawi

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds