Mission report of a Short Term Scientific Mission (STSM) of Filippos Tymvios from the Meteorological Service of Cyprus to University of Augsburg

Period: 03/03/2008 – 11/03/2008

"Estimation and rating of criteria for comparing circulation classification methods"

Filippos S. Tymvios, Meteorological Service of Cyprus

Christopher Beck, University of Augsburg

Andreas Philipp, University of Augsburg

Silas Michaelides, Meteorological Service of Cyprus

Summary

This mission was scheduled between the Cyprus team which is responsible for the neural network classification of weather partners for Cost 733 and the Augsburg team, responsible for the statistical interpretation of the classification methods. The mission was gravitated by the following topics:

- Presentation of the programming / scripting techniques used by the two teams
- Brief presentation of synoptic type weather classification applications concerning heavy rainfall events and dust concentration (Tymvios et al. 2007, 2008)
- A review of the objective comparison methods and the neural network classification methodology
- A tune-up of the neural classification method in order to investigate some topics discussed at the WG2 expert meeting in Augsburg using Domain 7 (central Europe)
- Evaluation of all the new classifications against the statistical methods proposed by the Augsburg team

We specifically focused on the varying number of classes' application. The motivations behind this were the recent discussions within WG-Meetings in Barcelona and Augsburg. The previous finding that performance in terms of evaluation criteria changes with varying number of epochs ignited the rerun of the models with different number of epochs.

The results of this mission will be circulated among the Cost 733 members and published on the website.

Database

For the synoptic classification of weather patterns, data from the ECMWF reanalysis project (ERA40) were used. More specifically, the 500hPa isobaric surfaces were utilized. The dataset consists of the 1200UTC isobaric heights, in geopotential meters, over an area defined by the geographical points 43^{o}N to 58N^{o} and 3^{o}E to 26^{o}E and for the years of 1957 to 2002 (01/10/1957 – 30/09/2002) at a resolution of 1^{o} x 1^{o}. Each day includes 16x24 points which are converted to a column vector of 384 points. This is Domain 7 or the Central Europe domain, as shown to the map below (marked in yellow colour)

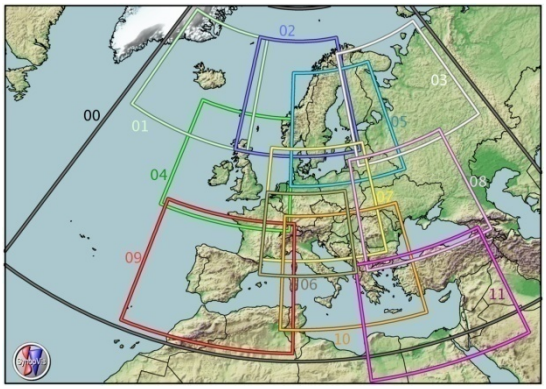

Figure 1. The Cost733 domains

Neural Network Classification

The Neural Network architecture proposed for the classification of weather patterns is the Kohonen's SOFM (Self-Organising Features Map), which is an unsupervised network (Kohonen, 1990). In unsupervised learning, there are no target values, as in the case of other methods of ANN. The Kohonen self-organizing feature map algorithm in its unsupervised mode, was chosen for building the neural network models, because neither the number of output classes (categories) nor the desired output are as *a priori*. This is a typical example where unsupervised learning is more appropriate, since the domain expert will be given the chance to see the results and then decide which model gives the best results and thus guided to decide the number of output classes that better represent the system (Pattichis et al., 1995).

The Kohonen self-organizing features map network is a type of unsupervised network, which has the ability to learn without being shown correct outputs in the sample patterns. These networks are able to separate data into a specified number of categories with only two layers: an input layer and an output layer, the latter consisting of one neuron for each possible output category.

The procedures followed for the design and the implementation of the neural networks are described in Michaelides et al. (2007) and presented briefly through the next paragraphs.

Given a training set, $X^{(k)}$, k=1, 2, … , p, the objective is to discover significant features or regularities in the training data (input data). In our case the input data was 16436 vectors, one for each day in the time span of 1957 to 2002. The neural network attempts to map the input feature vectors onto an array of neurons (usually one or two-dimensional). By doing so, the input feature vectors can be clustered into c-clusters, where c is less or equal to the number of neurons utilised, depending of the architecture of the learning algorithm used (Charalambous et al., 2001). Input vectors are presented sequentially in time without specifying the desired output (Michaelides et al., 2001). The two dimensional architecture of Kohonen's SOFM was adopted in the present research.

The input vector X is connected with each unit of the network through weights w_j, where j = 1, 2,…,M. M equals to 384, for the area in study (384 grid points). The neuron whose weight vector is closest to the input vector X (in terms of Euclidean distance) is the winner. This neuron is represented with I and is the winner neuron to input X if $||w_I - X|| = \min_i ||w_i - X||$, i = 1,…,M.

Note that $||w_I - X|| = [(w_{i1} - x_1)^2 + (w_{i2} - x_2)^2 + \ldots + (w_{iM} - x_M)2]^{½}$ is the Euclidean distance between weight vector w_i and input vector X (Charalambous et al., 2001). The number of the SOFM output nodes (possible classes) N in the Cost 733 classification effort varies according to the domain size, from 9 neurons for domain 6 (Alps) to 30 neurons for the large European domain.

The weight vectors of the winner neuron, as well as its neighbourhood neurons, are updated in such a way that they become closer to the input pattern. Learning follows the following rule:

$$w_i^{(new)} = \begin{cases} w_i^{(old)} + a\,(X - w_i^{(old)}) & , \; i \in N(I,R) \\ w_i^{(old)} & , \; I \notin N(I,R) \end{cases}$$

where the neighbourhood set N(I,R) of neuron I of radius R consists of the neurons 1,I±1,…,I±R, assuming these neurons exist, with maximum value being

around the winner I, in order for a larger number of neighbourhood units to share the experience of learning with the winner unit, and it becomes zero as the distance between the neighbourhood units and I increases. The coefficient a in the above relationship is called the learning factor and decreases to zero as the learning progresses. For simplicity, R is usually considered to have the shape of a geometric area, such as a rectangle or hexagon. In this work the hexagon setup was chosen.

The radius of the neighbourhood around the winner unit is relatively large to start with, to include all neurons. As the learning process continues, the neighbourhood is consecutively shrunk down to the point where only the winner unit is updated (Kohonen, 2001).

As more input vectors are represented to the network, the size of the neighbourhood decreases until it includes only the winning unit or the winning unit and some of its neighbours. Initially, the values of the weights are selected at random.

The learning method chosen for the SOFM architecture is described below:

1. The initial value of the weights was set to small random numbers while the learning rate and the neighbourhood were appointed with relatively large values. Steps 2 to 4 were repeated until the weights of the network were stabilized.

2. One vector X was chosen from the dataset as an input to the network.

3. The table for unit I with weight vector closest to X was determined by calculating

$||w_I - X|| = \min_i ||w_i - X||$.

4. The weight vector in (t + 1) iteration was updated according to:

$w_i(t + 1) = w_i(t) + a(t)(X - w_i(t))$, for units that belong in set N(I,R)

$w_i(t + 1) = w_i(t)$, for units that do not belong in set N(I,R).

5. The neighborhood and the learning rate of the parameters were decreased (Charalambous et al., 2001).

When all vectors in the training set were presented once at the input, the procedure was repeated many times with the vectors presented in order each time. This part of the algorithm at the end organizes the weights of the two-dimensional map, such that topologically close nodes become sensitive to input that is physically similar. Nodes were ordered in a natural manner, reflecting the different classes of the training set (Michaelides et al., 2001, 2007).

The 16436 vectors containing the data to be processed were merged into a single two-dimensional array, which was then fed as input to Matlab 7.5. The array resulting from the merging of the daily vectors has a dimension of 16436 columns and 384 rows. This merging was achieved by placing the 24 rows in each day sequentially, thus transforming each of the two-dimensional arrays corresponding to each day (24 rows and 16 columns) into a single column.

Using Matlab's own annotation, the SOFM implemented for Domain 7 is shown below :

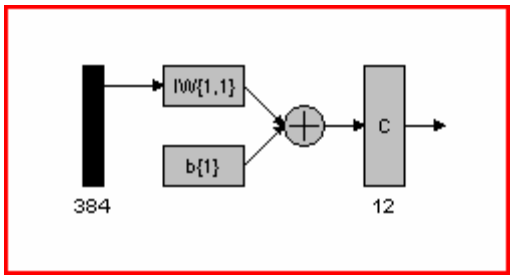

Figure 2. A schematic representation of the Neural Network Archtecture implemented (Matlab 7.5)

Experimental setup

Kohonen networks are trained for a fixed number of epochs. Small problems may train in as little as 50 epochs, but the number can be set to 10,000 ore even more for really large problems. These networks are extremely demanding as far as computing power and memory is concerned. The classification results included in cat. V.1.1 were obtained from models trained

with 100 epochs. The learning rate and neighbourhood size are automatically reduced as training progresses. The neighbourhood size should begin with a relatively high number, such as 90 percent of the number of neurons in the output layer. The network's initial weights, as well as the learning rate, were set as a random value of 0.5 and the size of the neighbourhood was depended each time from the number of outputs. The termination of the learning process (presentation of learning events) occurs after the completion of the preset number of epochs or when the neighbourhood becomes zero.

In order to investigate the sensitivity of the classification method to the number of classes, the domain 7 was classified using 4, 6, 9, 12, 16, 20, 25, 30, 36, 42 and 48 clusters. This setup is also convenient to compare this clustering method against other methods with different number of clusters to show if it's sensible to compare methods with different number of clusters or not.

We also run the classification algorithms for the domain 7 using more epochs to see if the number of epochs affects the results and if so, if there is a critical number after which further training will not cause improvement to the classification.

The neural classification of domain 7 with 12 and 20 clusters was materialized for 50, 100 and 500 epochs. The preliminary results showed a strong connection of the training period to the classification results and created the need to rerun the neural network classification for all domains altering the networks accordingly. Due to limited time for the STSM mission (more epochs increase the training time significantly), the new classification results will be send as soon as possible to Augsburg in order to be included in v.1.2 of the classification catalogue.

Classification results

The new classifications with variable number of clusters offer the possibility to compare the NNW method with all other classifications on the basis of (more or less) similiar numbers of classes.

Each of the COST733 classifications is compared to :

1. The NNW classification (100 epochs) that fits best with respect to the number of classes
2. The old NNW classification with 12 classes
3. The new NNW classification with 12 classes

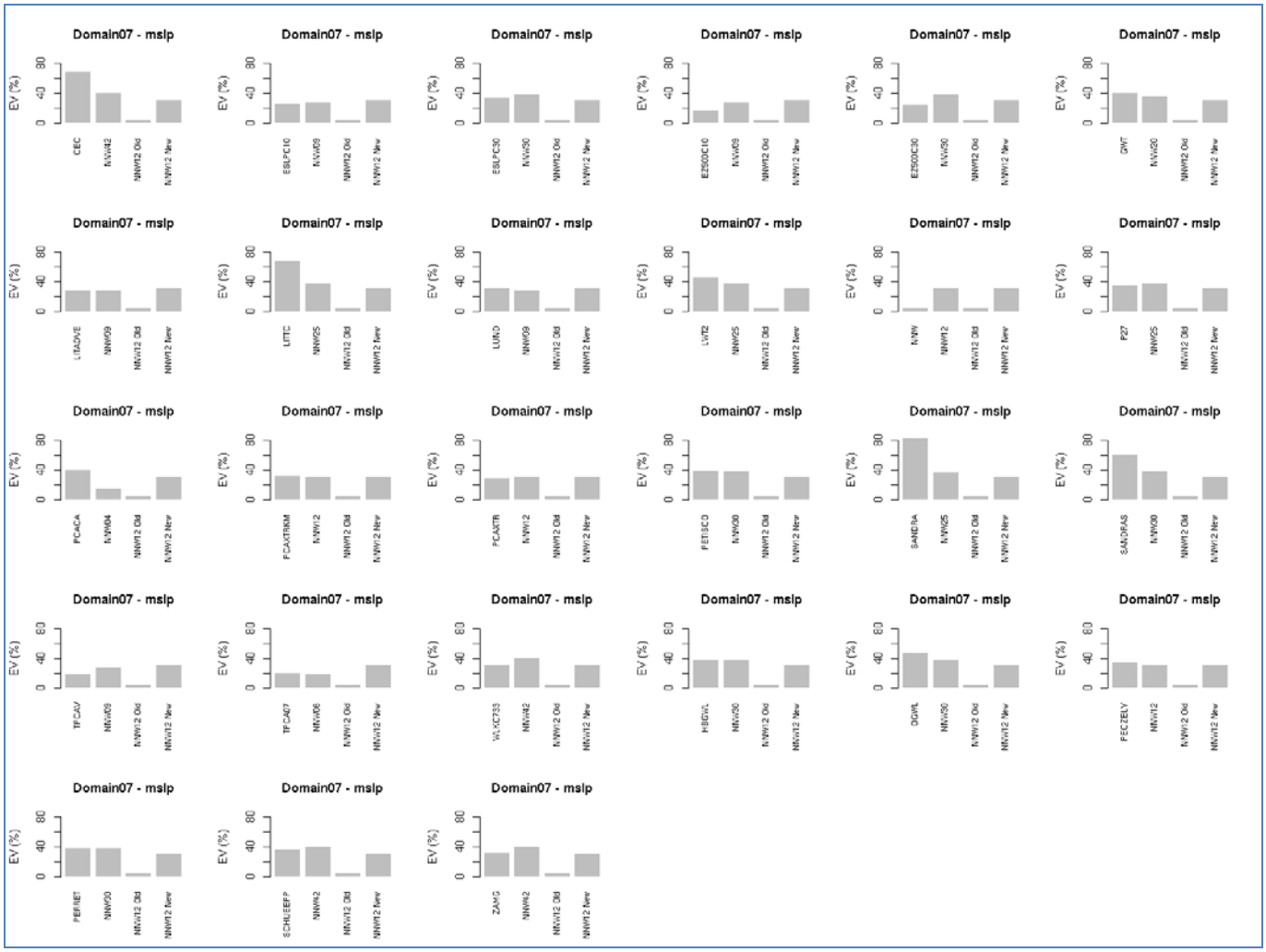

Figure 3. The explained variation (EV) estimated over the whole year

The same figure showing the Pseudo-F statistic instead of the Explained Variation follows.

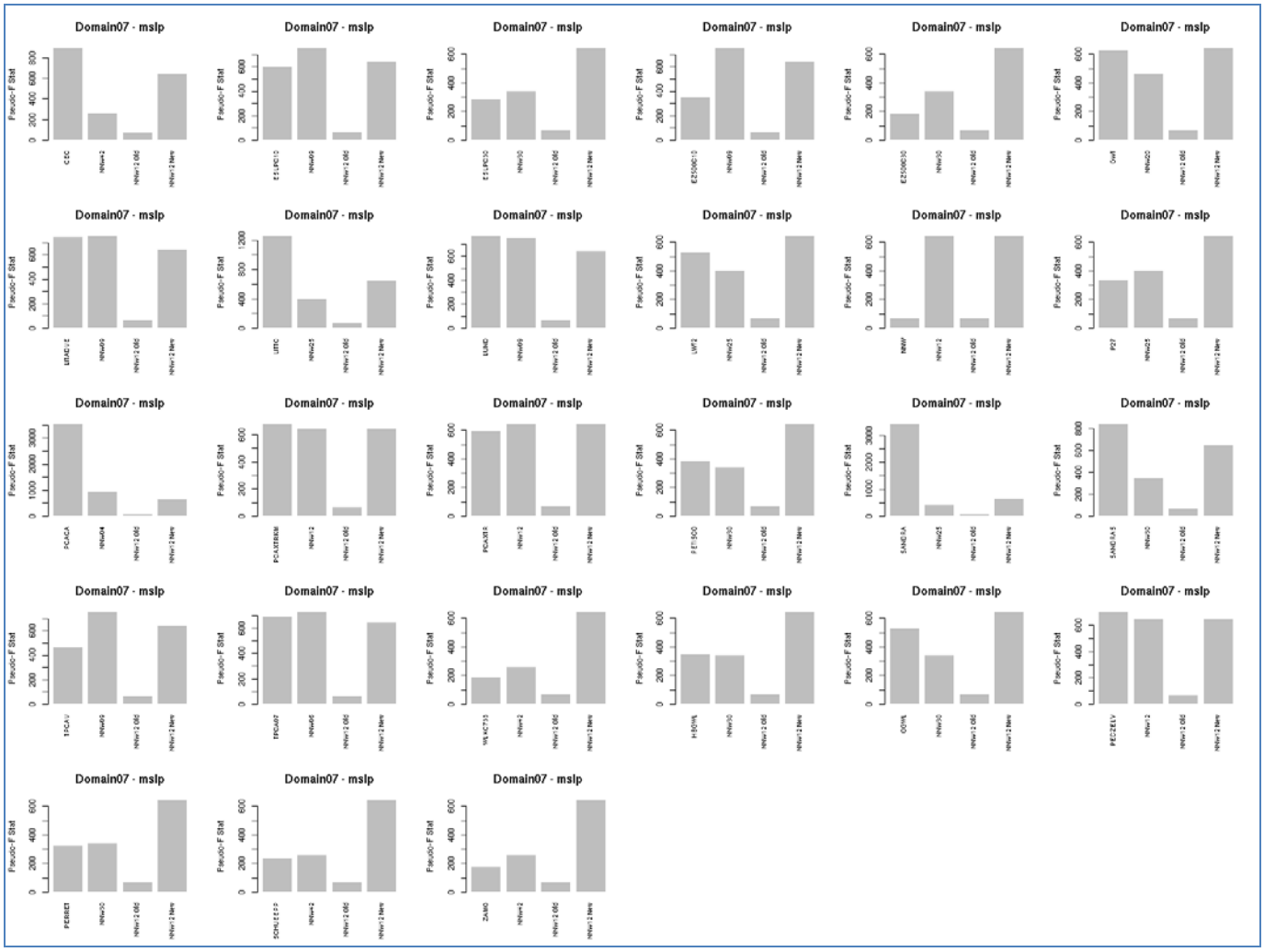

Figure 4. The Pseudo-F statistic estimated over the whole year

The "new" NNW12 classification performs much better than the "old" one and also, in most cases the differences between NNW and other classifications decrease if "similar" numbers of classes are used for comparison.

This provides us with a strong argument for the reruns of objective classifications with varying numbers of classes on which the WG2 decided in Barcelona and Augsburg.

Differences between the 12 and 20 classes' runs of NNW with varying numbers of epochs are in favour of the more epochs trained networks. The results also show that the more classes you have, the more epochs needed for the network to saturate. Since the number of classes is associated with the size of the domain, the training effort is multiplied accordingly. From more training

experiments contacted in Cyprus, the upper limit for the smallest domain (Domain 6) is 100 epochs, the limit for the medium sized domains (Domain 4,5,7,8,9,10 and 11) is 200 epochs, for the large domains (Domain 1, 2 and 3) is 400 epochs while for the large European domain (Domain) this limit is up to 600 epochs. Further training results no difference to the classification.

References

CHARALAMBOUS, C., CHARITOU, A., and KAOUROU, F. (2001), Comparative analysis of neural network models: Application in bankruptcy prediction, Annals of Operations Res. 99, 403–425

KOHONEN, T., (1990), The Self-Organizing Map. Proceedings of IEEE, 78(9), pp. 1464-1480

KOHONEN, T. (2001), Self-Organizing Maps, 3rd Ed. (Springer Series in Information Sciences)

MICHAELIDES, S.C., PATTICHIS, C.S., and KLEOVOULOU, G. (2001), Classification of rainfall variability by using artificial neural networks, Internat. J. Climatology 21, 1401–1414

MICHAELIDES, S.C., LIASSIDOU F. and SCHIZAS C.N. (2007), Synoptic classification and establishment of analogues with artificial neural networks. Pure and applied geophysics. 164 (2007) 1347-1364

PATTICHIS, C.S., SCHIZAS, C.N., and MIDDLETON, T.M. (1995), Neural Network Models in EMG Diagnosis, IEEE Transactions on Biomedical Engineering, vol. 42, no. 5, 486–495

SHIZAS C.N., MICHAELIDES S.C., PATTICHIS C.S, LIVERSAY R.R. (1991), Artificial neural networks in forecasting minimum temperature. Proceedings of Second

International Conference on Artificial Neural Networks, Bournemouth, UK. Institution of Electrical Engineers, Publ. No. 349, pp. 112-114.

TYMVIOS F.S., COSTANTINIDES P., RETALIS A., MICHAELIDES S.C., PARONIS S., EVRIPIDOU P. and KLEANTHOUS S. (2007), The AERAS project: data base implementation and Neural Network classification tests. Urban Air Quality Proceedings, 2007

TYMVIOS F.S., SAVVIDOU K., MICHAELIDES S.C., NIKOLAIDES K.A. (2008), Atmospheric circulation patterns associated with heavy precipitation over Cyprus. Geophysical Research Abstracts, Vol. 10, EGU2008-A-00000, 2008