Applications of classification of synoptic patterns with artificial neural networks

F. Tymvios, S. Michaelides

Cyprus Meteorological Service

Introduction

- There is a strong relationship between large scale circulation patterns and regional surface variables such as rainfall, dust etc, as these are controlled by large scale circulation patterns

- Hereafter, reference to circulation patterns implies the use of synoptic upper air charts
  - Height patterns at 500hPa-level of no divergence

Artificial Neural Networks were used in the classification of synoptic patterns. The resulting classifications were subsequently used in three applications that will be presented here:

1. Synoptic classification and rainfall extremes in Cyprus (forecasting)
2. Synoptic classification and dust events in Cyprus (forecasting)
3. Synoptic classification: trends (climatology)

Artificial Neural Networks

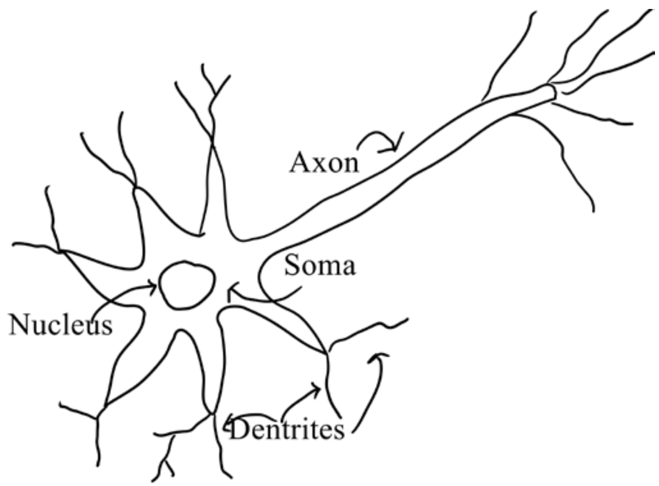
- What can they do for us?

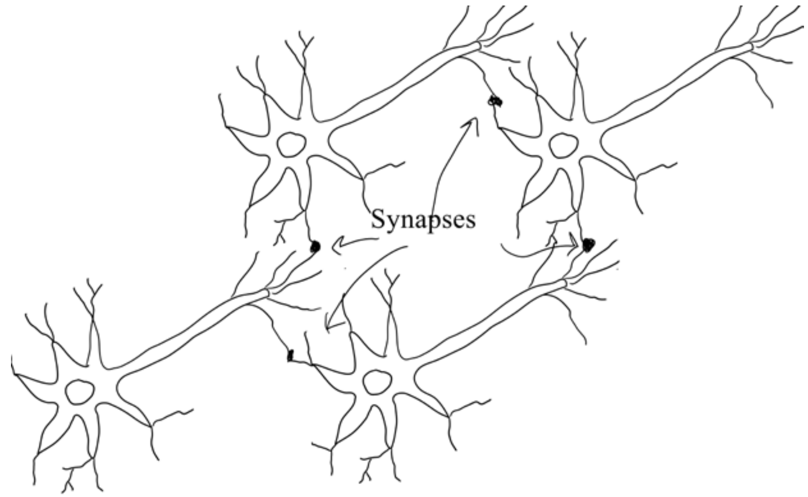
During the last two decades, ANN have proven to be excellent tools for research activities and real world applications :

- they are able to handle non-linear interrelations (non-linear function approximation)
- separate data (data classification)
- locate hidden relations in data groups (clustering)
- model natural systems (simulation)
- forecasting

Natural neurons versus Artificial Neural Networks

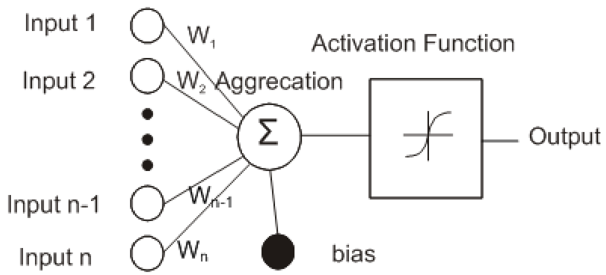
An Artificial Neural Network is an interconnected structure of simple processing units, whose functionality resembles of the biological processing elements, the neurons



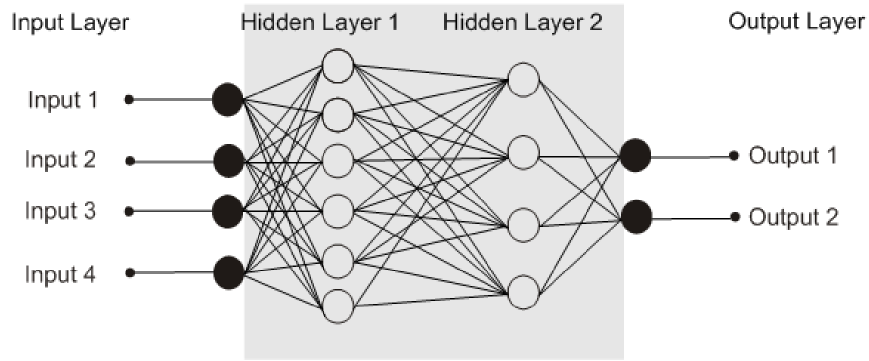


Artificial Neurons

Artificial neurons are algorithms that mimic nature: Two examples.



The perceptron



The multi-layer feed forward neural network

- Input is the synoptic analysis of 500hPa at a particular time
- Output is the number of class the analysis belongs to.

- The number of outputs (classes) is not
- *a priori* determined, but an “optimum” can be adopted by experimentation, in relation to the specific application.

- After a series of experiments and bearing in mind the applications that were implemented, it seems that a “reasonable” number of output (classes) for the European area is around 36.

- It can be said that this threshold exhibits the level of discretization required for the synoptic scale phenomena examined.

Kohonen Networks
Self-Organizing Feature Map (SOM)

- **Competitive**

Neurons compete to win. Only the winner (and to a far less extend its close neighbours) are "awarded"

- **Self-Organizing**

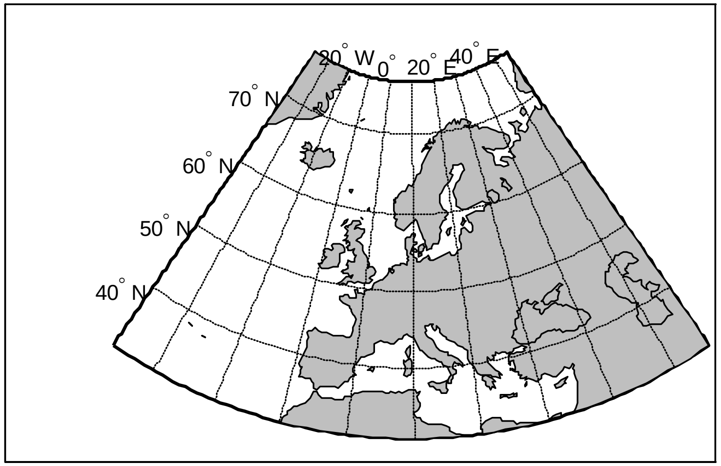
They have the ability to learn *without supervision* and *a priori* output patterns. They modify their connection weights based on the characteristics of the input values

Application 1 : Synoptic classification and extreme rainfall

- Events with the most extreme values of rainfall; the corresponding Cluster numbers refer to the 36-cluster classification

- The aim of this project was to identify synoptic patterns which are favorable for producing extreme precipitation events in Cyprus and thus provide guidance to weather forecasters on what to expect with such patterns.

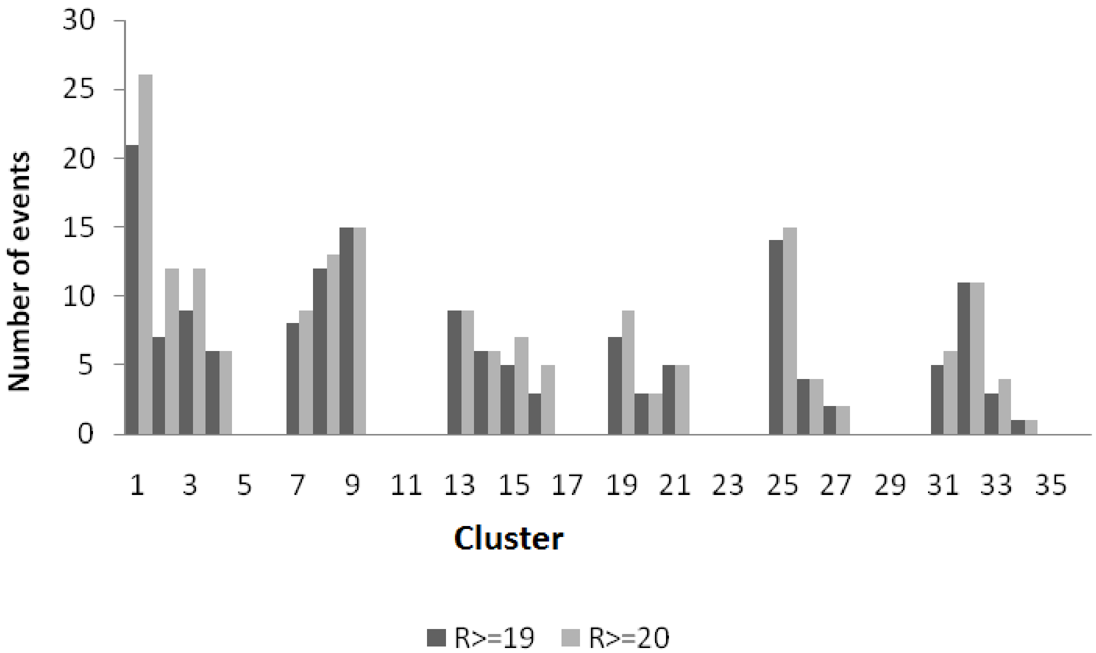
Database



- ERA-40 Reanalysis

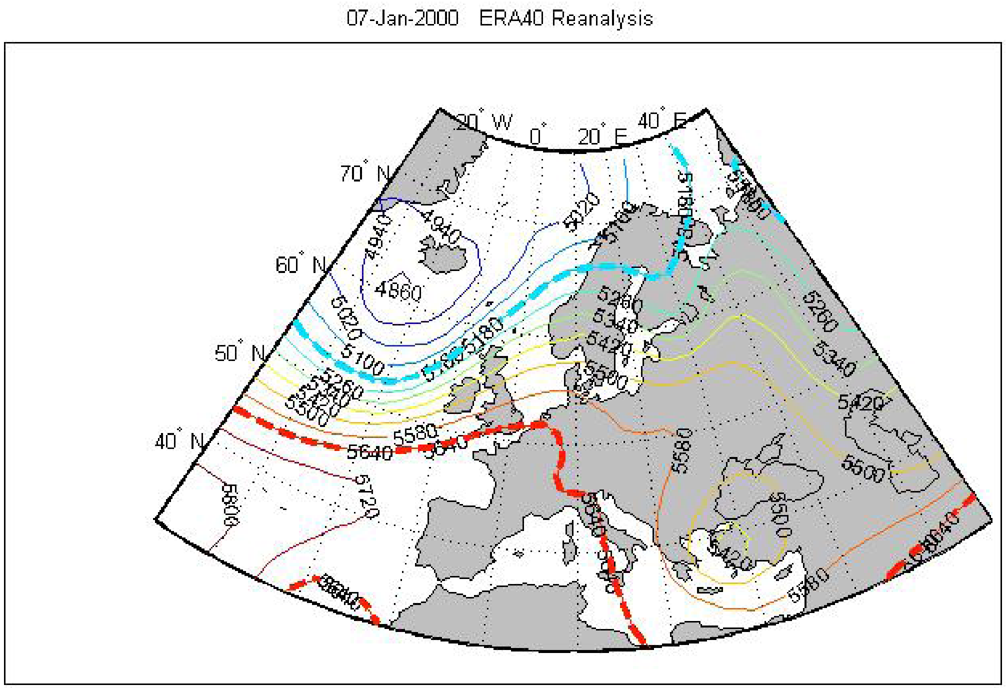
The dataset consists of the 1200UTC isobaric heights measured in geopotential meters, over an area defined by the geographical points 30ºN to 76Nº and 37ºW to 56ºE. The data refer to period from 1 October 1957 to 30 September 2002.

- This study used 23 rain-gauges in Cyprus (checked for homogeneity). The data refer to period from 1 October 1957 to 30 September 2002.
- From the distribution of the frequency of the average daily rainfall, **the upper 5% of the distribution and the upper 10% of the distribution** were calculated and are defined as the heavy rainfall thresholds for Cyprus; these thresholds are 20mm (denoted as $R \geq 20$) and 19mm (denoted as $R \geq 19$), respectively.
- The number of heavy rainfall days when the average rainfall over Cyprus exceeded the thresholds of 20mm and 19mm is 156 and 180, respectively.



Allocation of heavy rainfall events in the 36-cluster classification.

Cluster 1 is the one that is most frequently associated with extreme rainfall events (R≥20, 13%; R≥19, 14%), followed by Cluster 9 (R≥20, 10%; R≥19, 8%).



Typical geopotential height distribution pattern for Cluster 1, on 7 January 2000 (unit: gpm).

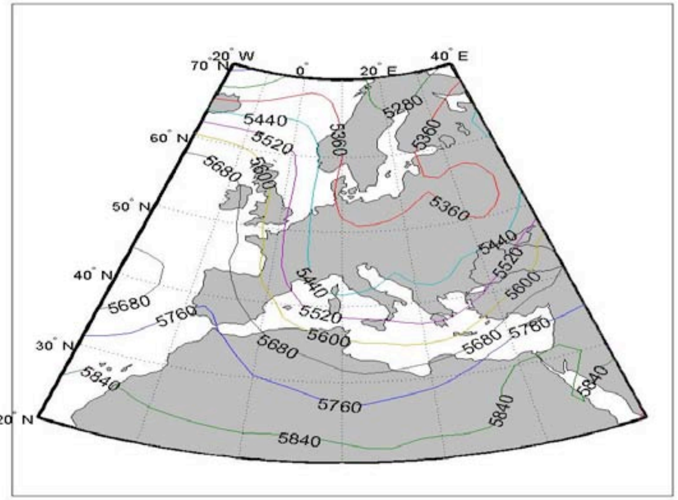
- From the selected cases studied, the association of heavy rainfall events with the specific patterns is in good agreement with the weather forecaster's perception of which synoptic situations can lead to excessive precipitation events.

Application 2 : Synoptic classification and dust events

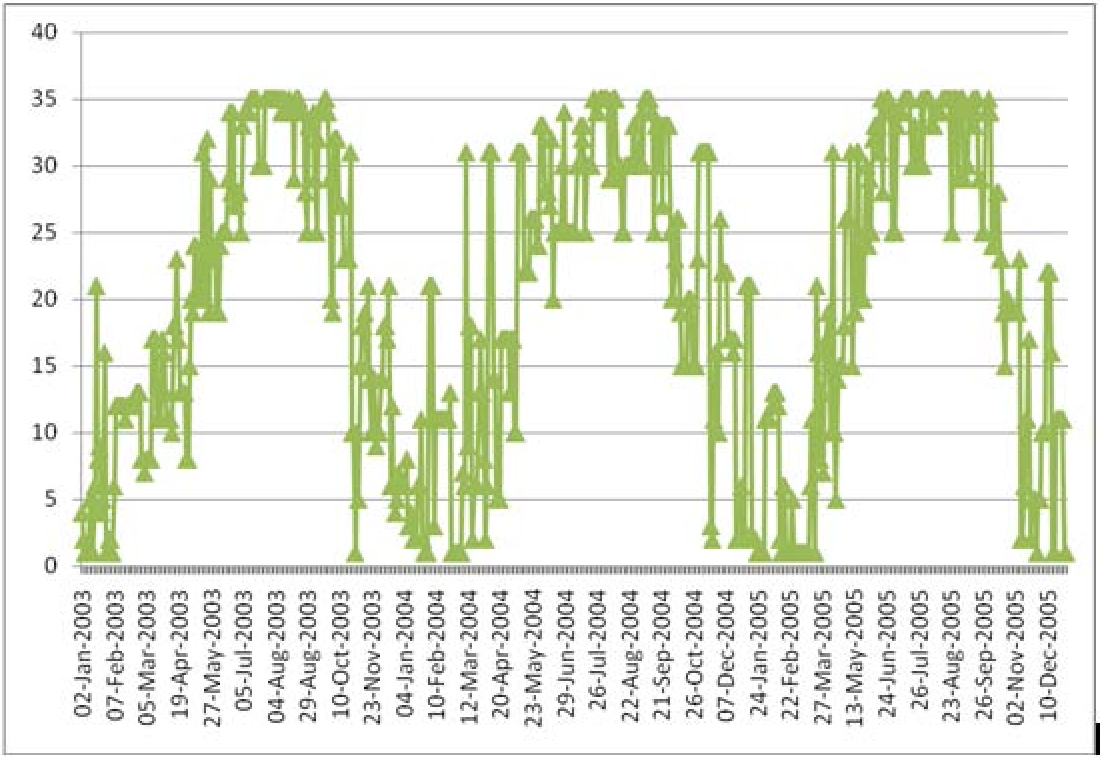
- This application refers to the occurrence of dust events over Cyprus and increased levels of atmospheric pollution through transfer of desert dust. In this research, synoptic patterns determined with ANN were subsequently related to measurements of atmospheric pollution determined on the ground.
- The aim of this project was to create a neural network and regression methodology that would be employed in forecasting dust loads in Cyprus, during dust events.

Database

- NCEP database



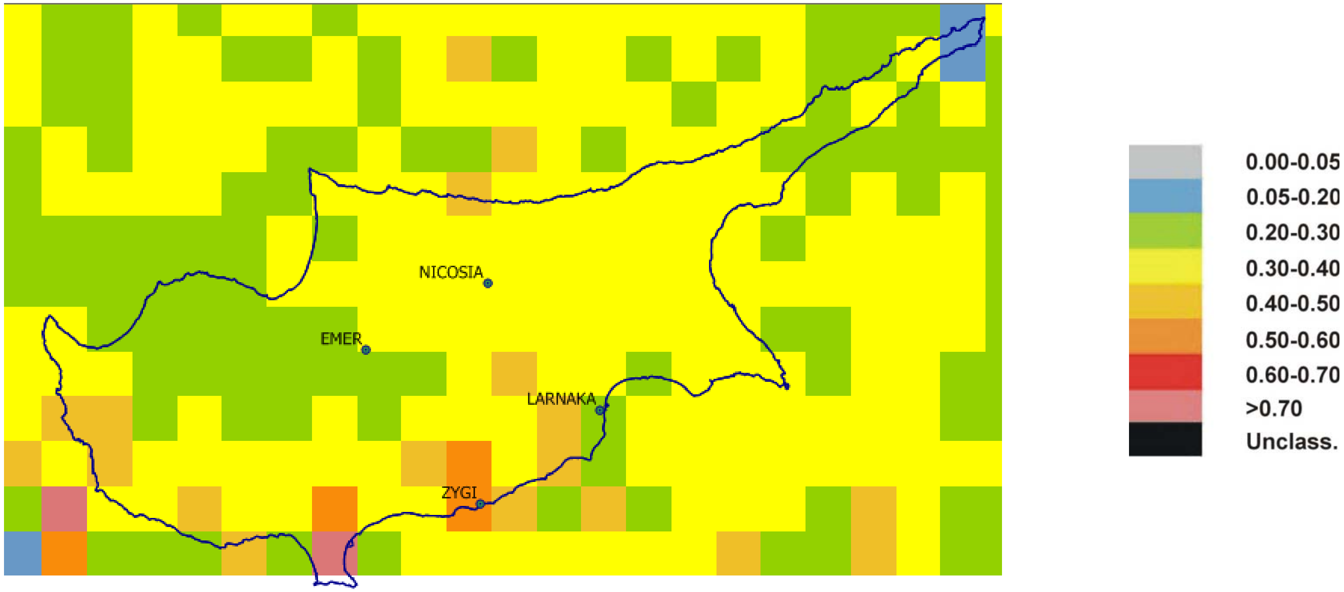
The data used in this application are the $2.5^o \times 2.5^o$ grid point values for the 500hPa isobaric surface at 1200UTC and for each day in the period from 1/1/1980 to 31/12/2005 retrieved from the National Centres for Environmental Prediction (NCEP, USA). The area used is defined by the points 60ºN, 20ºW (top left point) and 20ºN, 40ºE (bottom right).

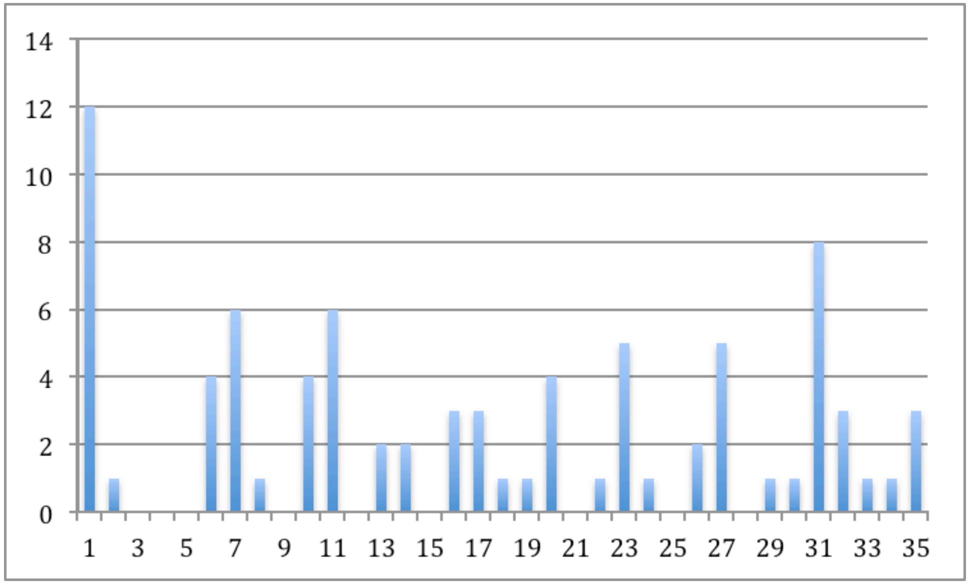


Time distribution of classes (synoptic types) in the three year period 2003-2005 [Strong seasonal dependence!]

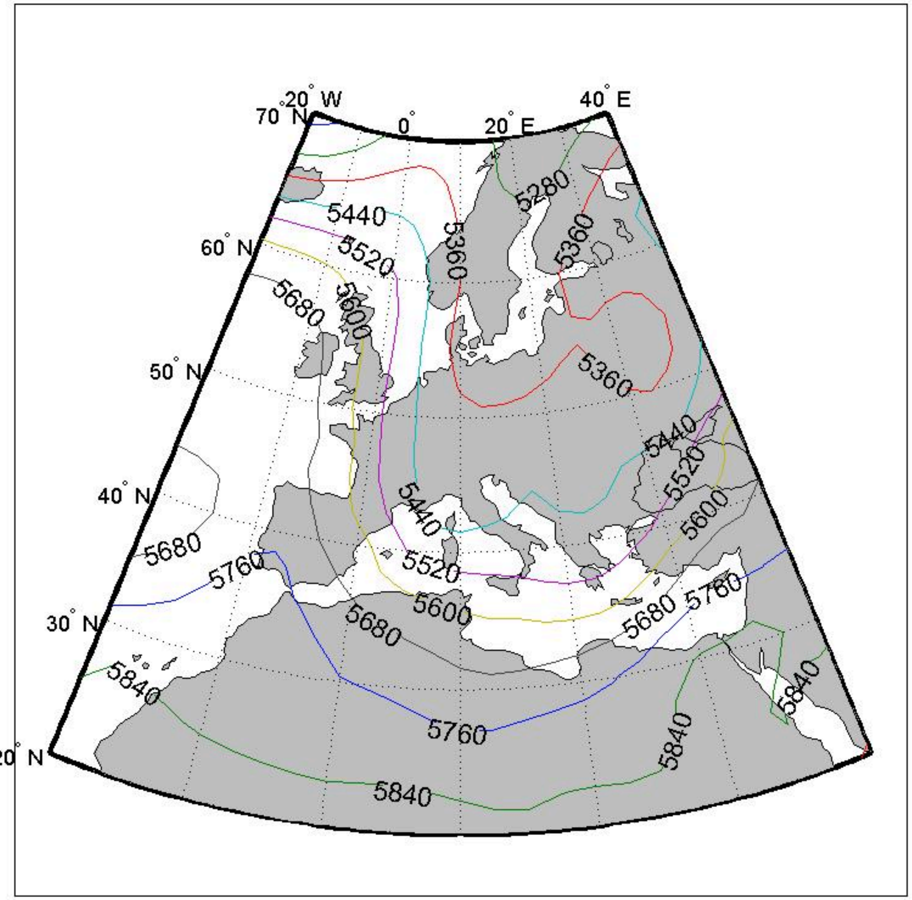
- The surface database consists of aerosol (PM10) concentrations from 2003, 2004 and 2005 at 4 stations representing four different types of area characteristics : coastal, urban, rural and industrial stations. Another station, completely isolated of any exogenous sources of aerosol pollution was used as a background station. The daily aerosol concentration at this station was used as the index that would associate high dust load events with circulation patterns. A high dust load event is defined as the day in which the daily total aerosol concentration exceeds 50mg/m^3.

Another source of input was the spot values of the Atmospheric Optical Depth (AOD) over the 5 stations, obtained from the Aqua and Terra satellites.

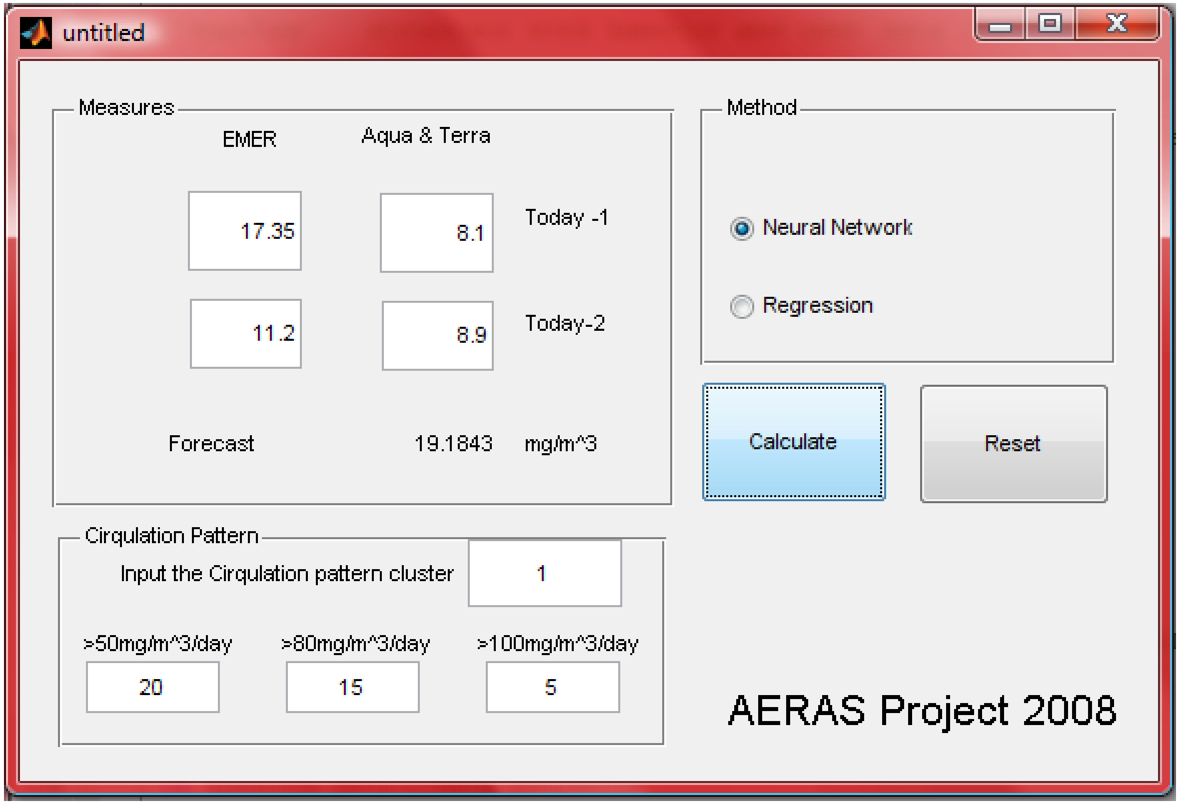




The distribution of the appearances of high dust load events (>50mg/m^3/day) in the three year period 2003-2005 for the 35 cluster classification – a preference to certain classes is evident

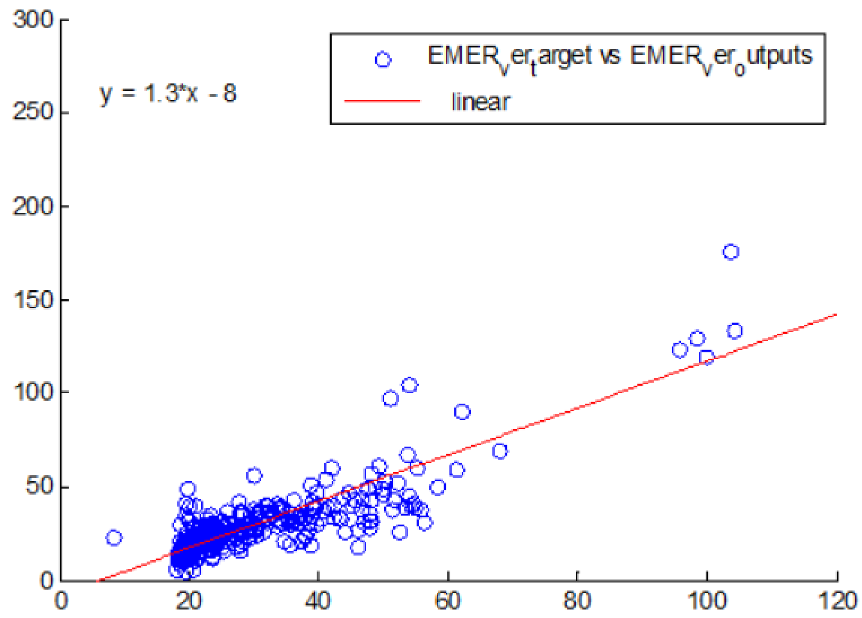


A favorable pattern for dust event over Cyprus



Forecasting Software: Interface for forecasting dust concentrations
[Forecast value and Probabilities for exceeding 50, 80 and 100 mg/m^3/day
[European Union allowable threshold]
Input: Day-1 and Day-2 measured concentrations and AOD
and Class (e.g. 1).
Methodology: Neural Network or Regression

Cyprus Meteorological Service



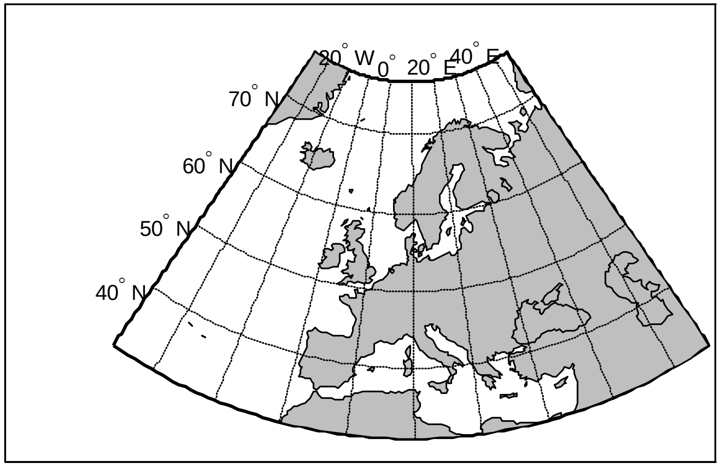
Relationship between forecast and measured concentrations for the background station (for the test set of data)

Application 3 : Trends in synoptic patterns

- The aim of this application was to identify possible trends in synoptic patterns that were established by using artificial neural networks’ classification.

Database

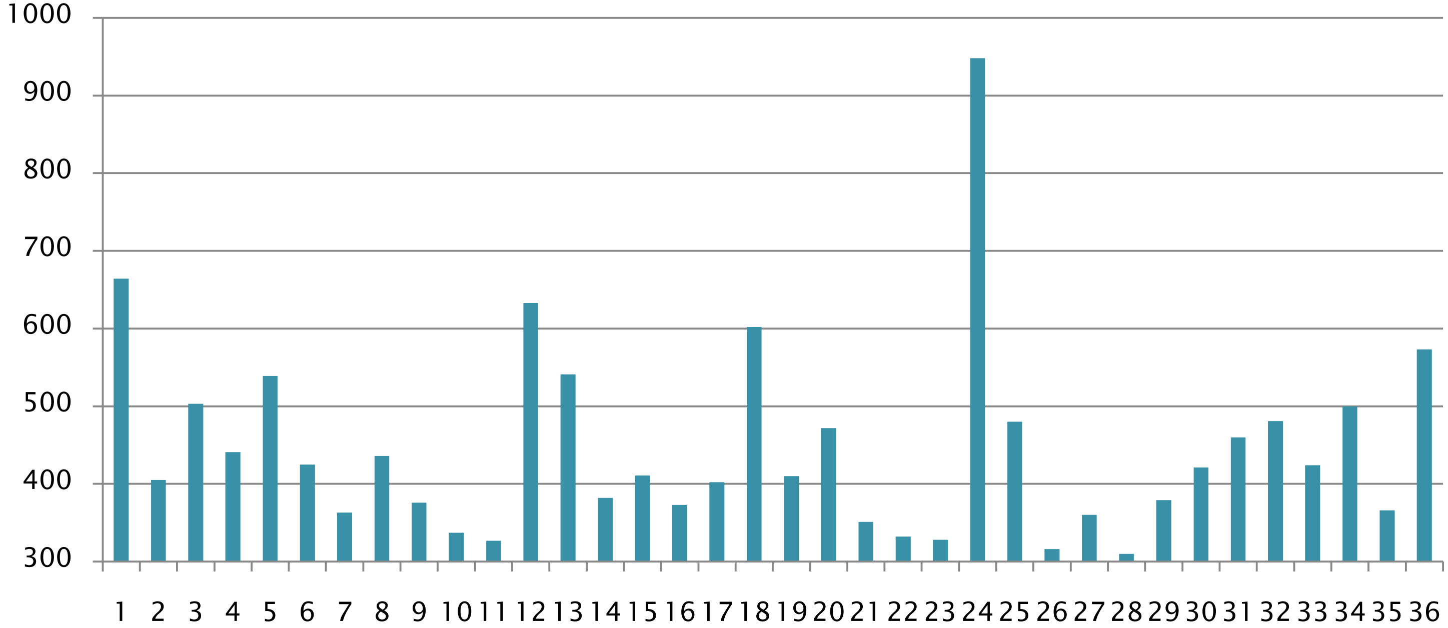
- ERA-40 Reanalysis



The dataset consists of the 1200UTC isobaric heights measured in geopotential meters, over an area defined by the geographical points 30°N to 76N° and 37°W to 56°E.

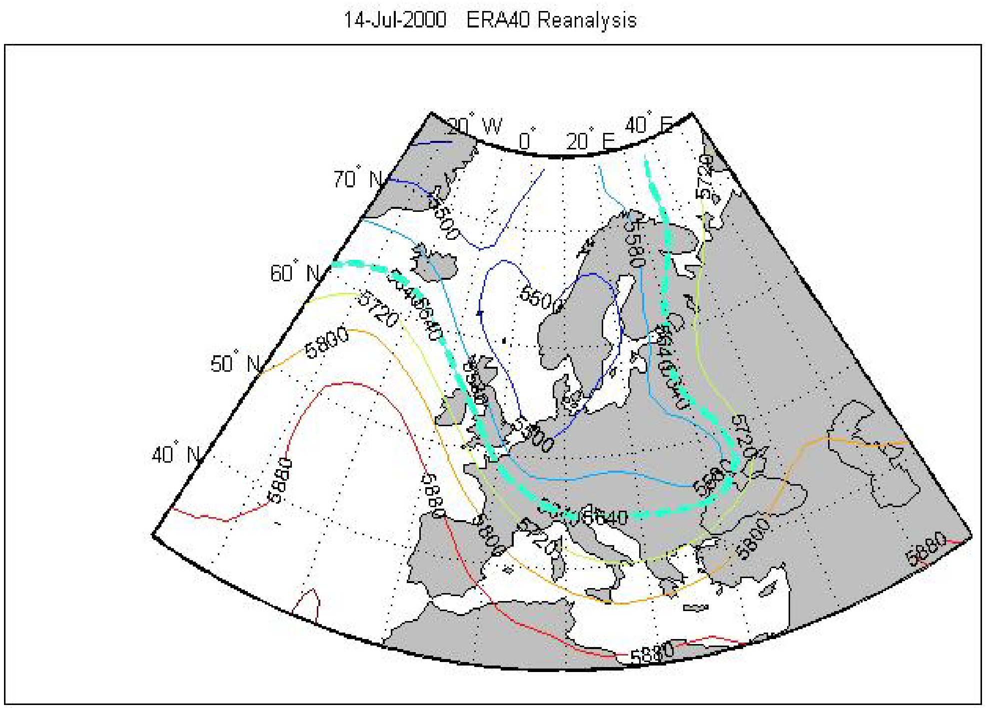
- A time period of 44 years (1958 to 2001) was considered.

Frequency of occurrence of clusters (36 clusters considered).



- Three important clusters will be briefly presented in the following:
- **Cluster 24** (the most frequent "warm and dry" cluster)
- **Cluster 1** (the most frequent "cold and wet" cluster)
- **Cluster 9** (a transitional period cluster, frequently associated with heavy rainfall events)

Cluster 24 : Typical example

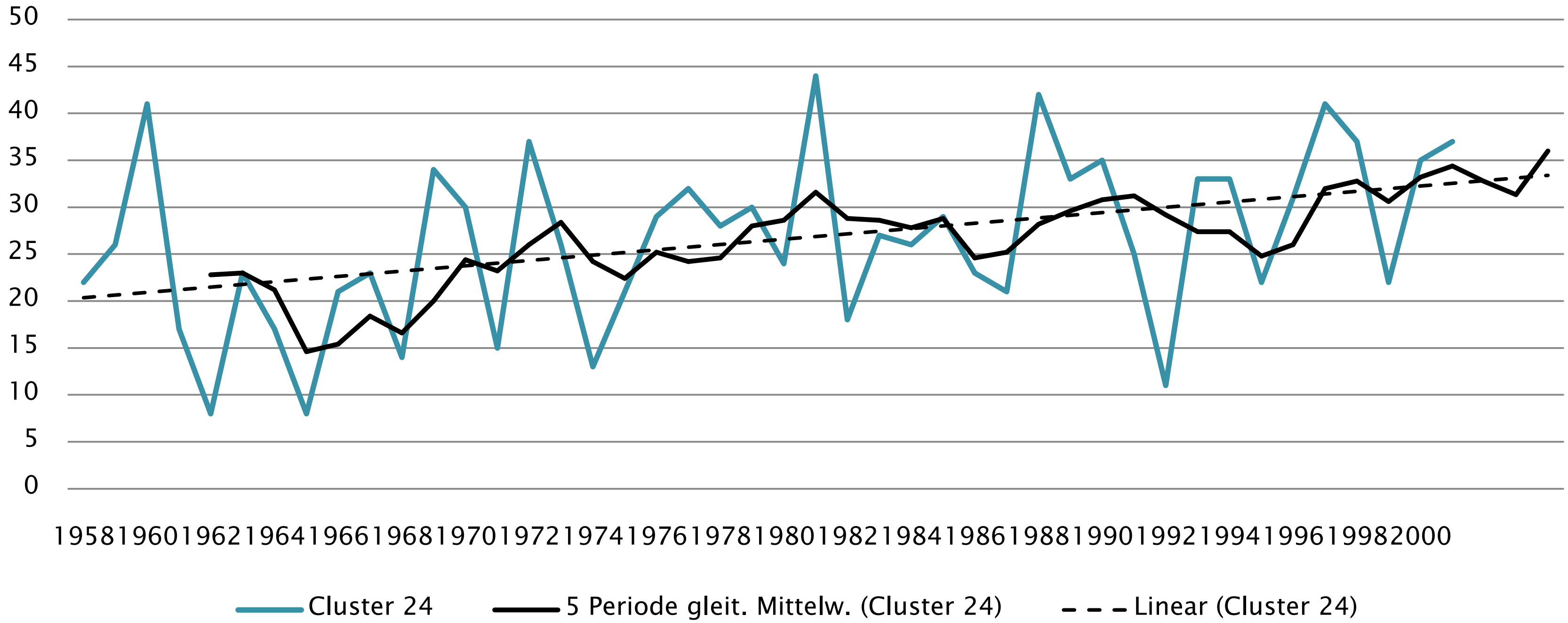


The most frequent is Cluster 24. This is a "warm and dry" cluster, dominant in July and August.
It corresponds to a persistent summer-time weather pattern. The frequency of the yearly occurrences of this cluster can be connected to the frequency of dry spells over eastern Mediterranean.

Example of cluster 24 height pattern over Europe at 14 July 2000

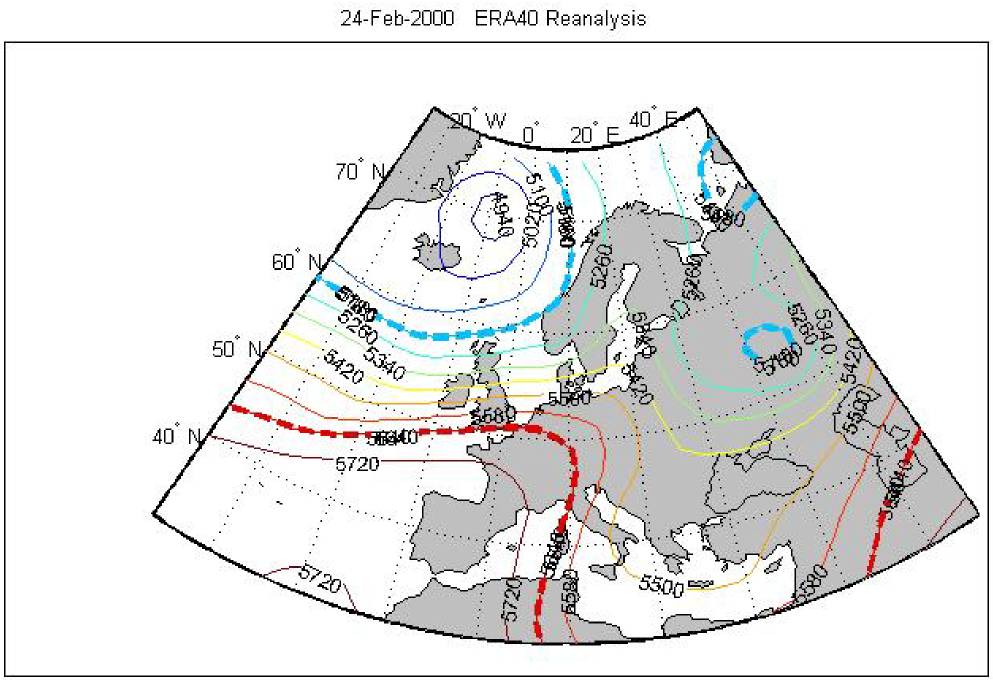
Cyprus Meteorological Service

Cluster 24 : Occurences wrt time



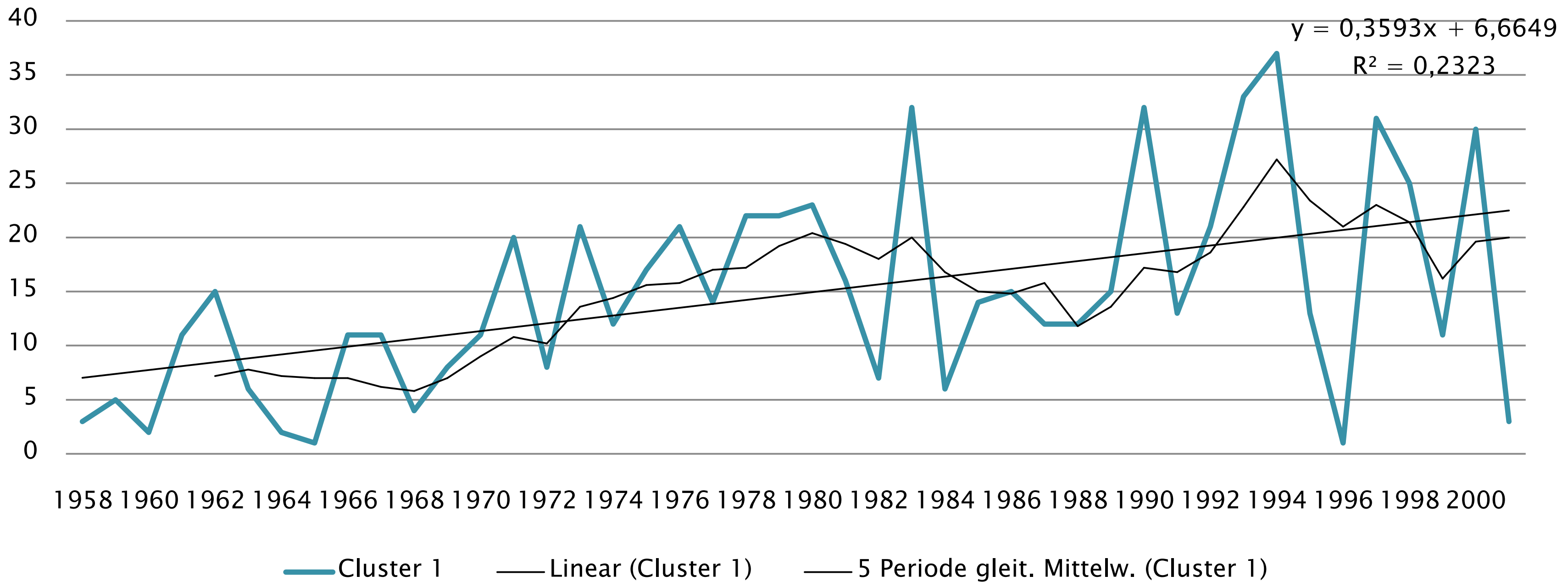
A noticeable increase in the number of Cluster 24 occurrences over time. The increase of the occurrences can be connected with dry periods over he island of Cyprus (e.g. 1957-1962, 1969-1974 and 1993-1997).

Cluster 1: Typical example



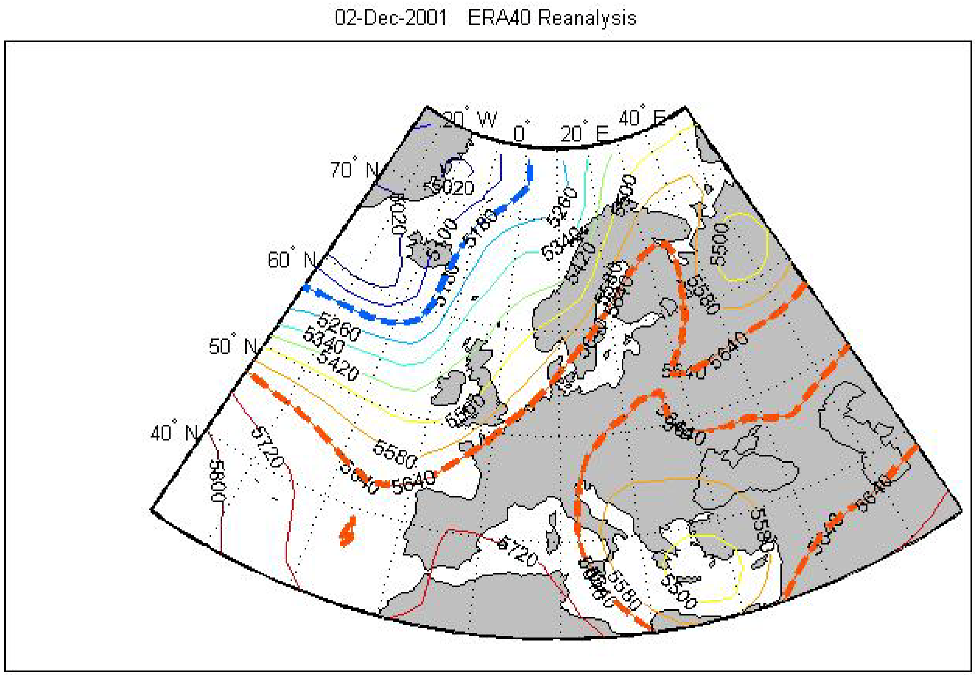
Cluster 1 is the most frequent "cold and wet" cluster.

Cluster 1 : Occurences wrt time



Although there is a notable variation in the frequency of this cluster in the last 15 years, overall, there is a tendency for increase.

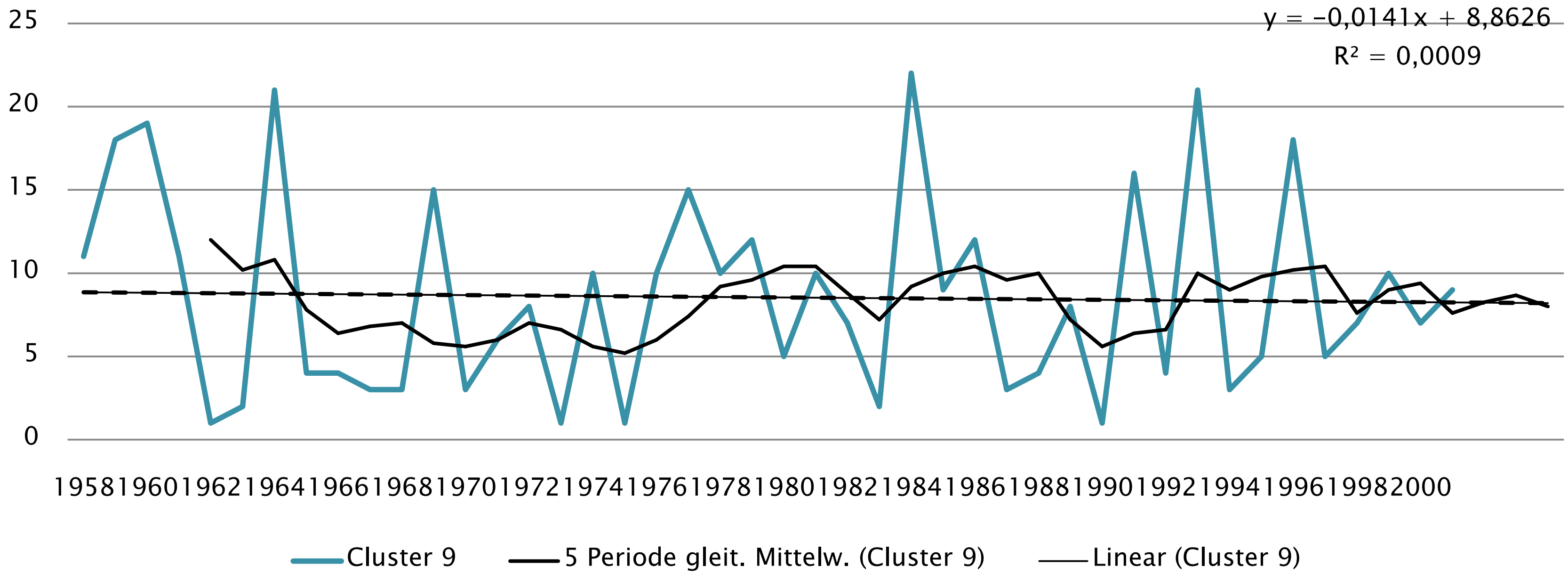
Cluster 9 : Typical example



Cluster 9 is a collection of significant synoptic situations, in terms of extreme rainfall events. It is a cluster that is almost equally shared between the “cool and wet” and “transitional” months.

Height pattern of 02/12/2001 with a daily average rainfall of 62.3mm over Cyprus.

Cluster 9 : Occurences wrt time



Overall, there is no significant change in the number occurrences of this cluster over time.

Comparing time periods

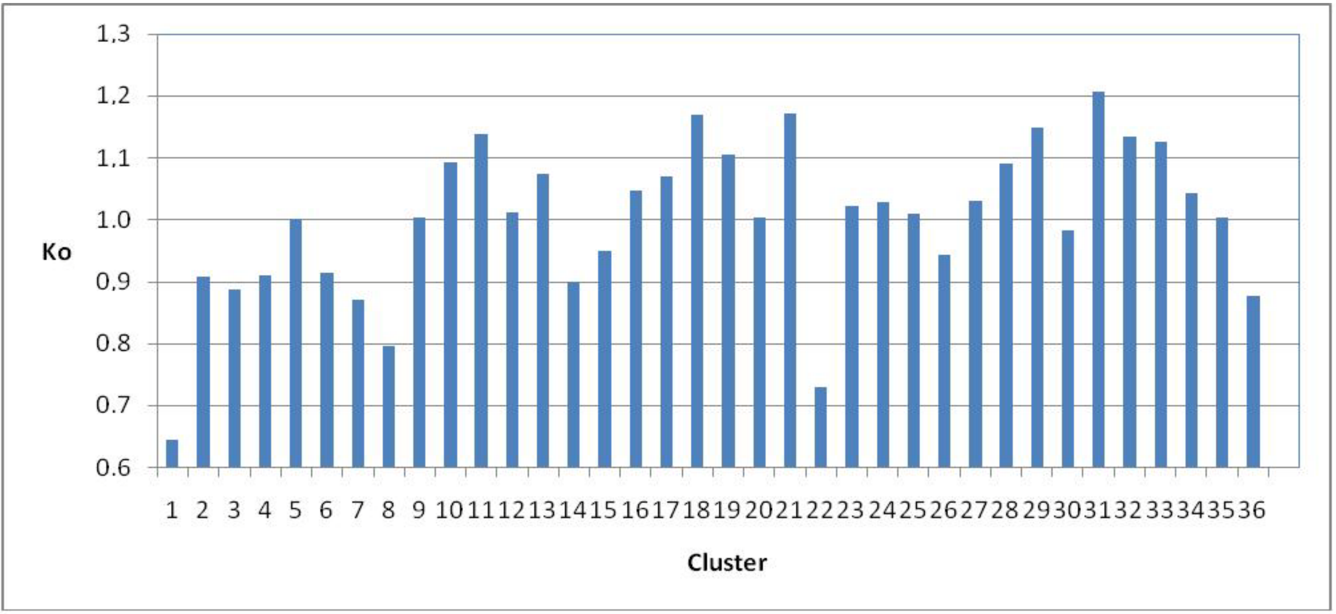
- The database is split to two periods:
  - Reference Period (1-1-1958 to 31-12-1987, 30 years)
  - Test Period(1-1-1988 to 31-12-2001, 14 years)

Using the distribution of the occurrences of each cluster for each period relative frequency) a Ko index is defined:

$$\text{Ko index} = \frac{\text{Relative frequency of clusters in the reference period}}{\text{Relative frequency of clusters in the test period}}$$

This index can be used to compare the number of occurrences of each cluster over the two periods. For **Ko equal to unity**, **there is no change in the frequency** of the respective cluster in the two periods; when **Ko is greater** (less) **than one**, **then the respective cluster is less** (more) **frequent in the "test" period than in the "reference" period.**

36 cluster classification – Ko index



- The most prominent change is noted for Cluster 1 (the most frequent "cold and wet" cluster)
- Regarding the most frequent "warm and dry" Cluster 24, the change between the two periods is very small.
- The same is true for the "transition" Cluster 9.

Concluding remarks

- 1. ANN classification can be used in assisting weather forecasters in forecasting extreme rainfall and dust events.

- 2. There seems to be a tendency for some patterns established with ANN to change over time.

Question:

- If there is a tendency for significant change in the frequency of synoptic patterns, will that affect the performance of any forecasting "tools" built on such a classification?

- Thank you!