SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

The STSM applicant submits this report for approval to the STSM coordinator

Action number: Cost Action 15211
STSM title: Short term scientific mission in Athena at the National Observatory of Athens
STSM start and end date: 05/03/2018 to 31/03/2018
Grantee name: Ion-Andrei Nita

PURPOSE OF THE STSM

The main purpose of this STSM was to analyse the synoptic patterns leading to lightning occurrence in Europe. The period we studied was 2005-2014, consisting from ZEUS long-range lightning network data which covers a domain from 30 W to 40 E longitude and 25 N to 60 N latitude. Since most of the lightning activity occurs around noon, we considered only the events from 9:00 to 15:00 UTC. We constructed a daily classification of synoptic patterns using mean sea level-pressure, the height of 500 mb geopotential, thickness between 850 hPa to 500 hPa and the relative vorticity at 500 hPa. The classification is based on an optimization algorithm that is converging by an iterative process all the objects into classes that are most similar. This was done by using a designed software written in FORTRAN and developed in a previous COST action, called cost733class. Using this method, we are calculated simply the frequency of lightning upon all the circulation types. The results consists in a several distribution patterns over the southern European continent, mostly of them showing a meridional circulation as being the major dynamic trigger for the atmospheric instability.

The second plan is to use a statistical method to predict the possibility of lightning activity using several predictors extracted from ECMWF reanalysis. Based on the location (lat/lon) and the day of a lightning event, we extracted the values at 12:00 UTC time for 12 parameters. We applied a logistic regression and in the last part, we tried to predict lightning activity using a statistical model.

The general purpose of this STSM was to find out more about the ZEUS data which is created and maintained by the National Observatory of Athens. In the same time, I wanted to start a possible collaboration with the researchers from this institute for future international researching projects or papers in common studying the lightning in relation with other atmospheric parameters

DESCRIPTION OF WORK CARRIED OUT DURING THE STSM

During my staying in Athens, the main tasks carried were:

1. Downloading from ECMWF all the reanalysis data (Interim version) for the period 1981-2016 for four parameters used to derive a classification of weather types in Europe. The parameters are: mean sea level

pressure, the height of 500 mb level, the thickness between 850-500 mb levels and the relative vorticity at 500 mb.

2. creating a synoptic classification with 27 types for the European continent and studying lightning activity upon atmospheric circulation types;

2. importing ZEUS data into R software and split it by years, seasons; calculating descriptive statistics (histograms, plots for hourly mean no. of flashes/each grid cell, plots for the mean number of days with lightning for each grid cell/months;

3. testing various models of logistic regression with different parameters to see which ones are having the best results for our final prediction model;

4. Applying a principal component analysis to remove multicollinearity among predictors. The new set of variables were orthogonal and therefore, could be included into the analysis.

5. Writing the introduction, data and methods and conclusions parts from a common scientific paper that is supposed to be sent to the review until the start of this summer.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

In this STSM, the main objectives were to assess the atmospheric circulation types (CT) leading to lightning occurrence in Europe. Since the study domain is a large one, we considered to focus on specific areas where a high density of flashes upon a specific CT can be observed.

For this study, we considered only the data recorded by ZEUS from 9:00 to 15:00 UTC time for the period April to September 2005-2014 for a spatial domain from -10 – 40 E lon and 25 – 50 N lat. Firstly, we run cost733class software to derive a classification of weather type using from ECMWF Era interim four inputs: mean sea level pressure, the height of 500 mb geopotential, thickness between 850-500 mb and the relative vorticity at 500 mb geopotential level on a spatial domain covering the European continent (D00) – fig.1 for 2005-2014.. We obtain 27 circulation types. The centroids for each type is presented in fig, 2 & 3.

We calculated a simply frequency value by dividing the number of occurrences for a CT (days) to the number of lightning discharges during that specific CT. This was done for every grid cell for the period 2005-2015 April to September. We considered only the CTs with a frequency higher than 10 days in order to eliminate possible misleading results (fig. 4). Prior to that, we decided to conduct a short analysis of 3 hour step lightning activity to prove that most of the convective activity occurs around noon hours; the results are expressed by plots (fig. 6). Also, the mean number of days with flashes for each month for each grid cell was calculated and the results are showing a higher activity especially in May, June and July (fig. 5).

We found out that most distinctive synoptic patterns leading to a high frequency of lightning flashes in the southern Europe are related to the meridional circulations (N-S advections) – fig.7 . These synoptic types are common when the zonal circulation on the continental scale is driven off by disturbances in the upper troposphere related to the Jet stream currents. (fig. 8) This pattern is associated with the advection of polar air masses towards the lower and warmer latitudes in the southern Europe. This is usually leading to powerful updrafts since the thermic contrast can be high and strong enough to maintain the air instability specific in the noon time. Other synoptic types conductive to high occurring of lightning are specific to southerly advections from the Mediterranean Sea towards the Central Europe. These types can occur mostly as a result of cyclogenesis in the Adriatic Sea when polar air meets with the warmer air in this region. As a result, this weather type leads to a higher frequency of flashes in the central part of the continent as a result of the advancing warmer air towards this area.

A second interest during this STSM was represented by the possibility of predicting lighting using a statistical model. We calculated a series of parameters which are known to be leading to air instability. Mostly were derived from ERA Interim version while others were extracted from CM-SAF products (cloud top height and cloud top temperature). In order to validate our presumption consisting in the high probability that all of these parameters are auto-correlated, we calculated a *variance inflation factor* score which showed as expected a high correlation among the predictors. Therefore, we decided to apply a principal components analysis in order to remove the multicollinearity. Prior to that, we applied a standardization on the data to avoid any erroneous results in the PC scores and have a comparable reference scale in the analysis. We kept only the PCs whose eigenvalues were higher than 0.3 (that were explaining about 95 % from the cumulative variance explained). The new obtained data was split into two datasets: the *training dataset* which had 75 % of the original size while the *testing dataset* was comprising 25 %. The first one was used to build the logistic model while the second dataset was used to assess its performance. In the preliminary phase, logistic regression was applied on monthly basis (e.g. all the months of June from 2005-2014)

Several performance scores were calculated to assess the performance of prediction. The *area under the curve* values was indicating generally a good model (values higher than 0.8). However, when comparing testing dataset (assessing 25 % from the original dataset) with the data that the model predicted, the results are yet to be improved. As a first assumption, we have reasons to believe that the model is working well on smaller timescale (i.e. a day or a group of days) rather than on multi-month basis.

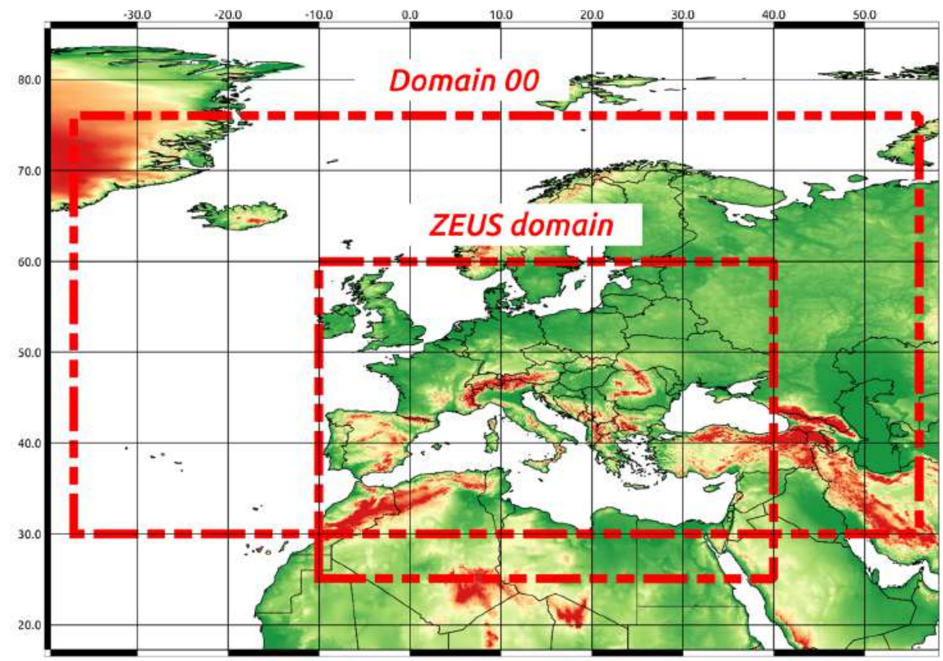

__Figure 1.__ Study domain for ZEUS data (ZEUS domain) and CTs classification (Domain 00)

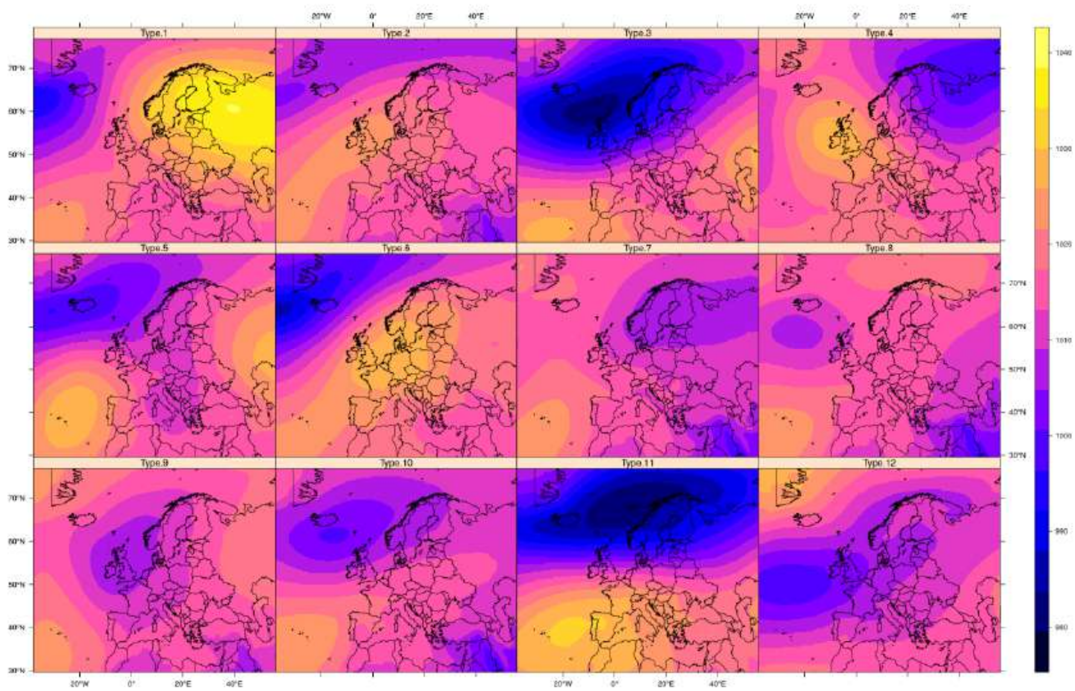

__Figure 2.__ Centroids at MSLP for each CT resulted (1)

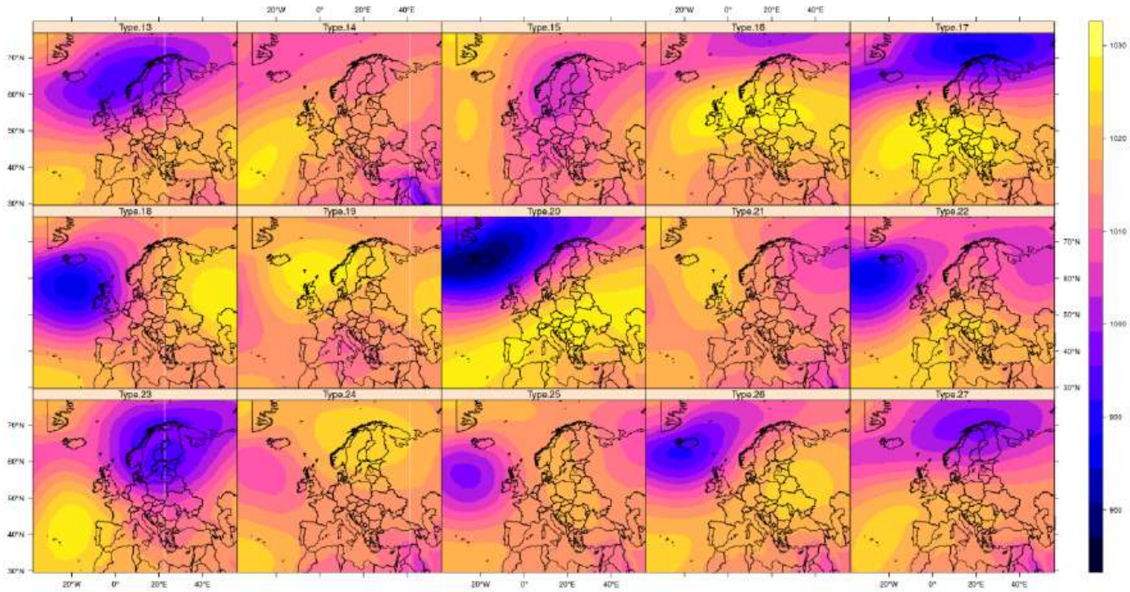

Figure 3. *Centroids at MSLP for each CT resulted (2)*

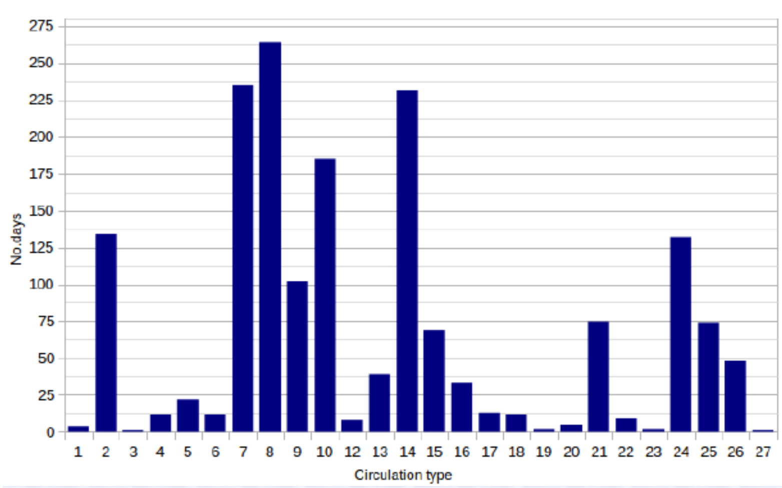

Figure 4. *Frequencies for each CT (April - September 2005-2014)*

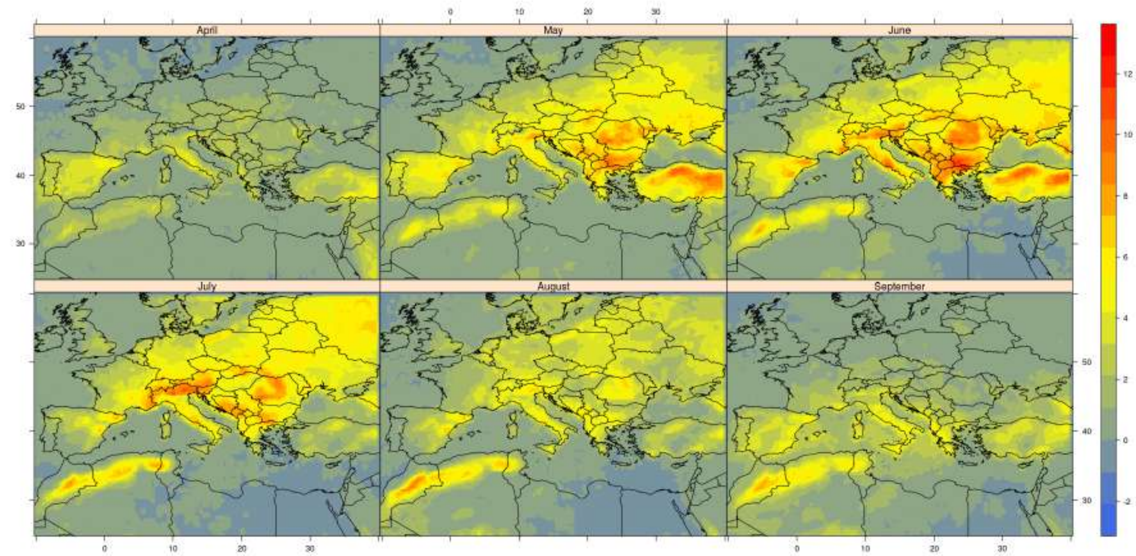

Figure 1. *Mean number of days with lightning activity detected by ZEUS*

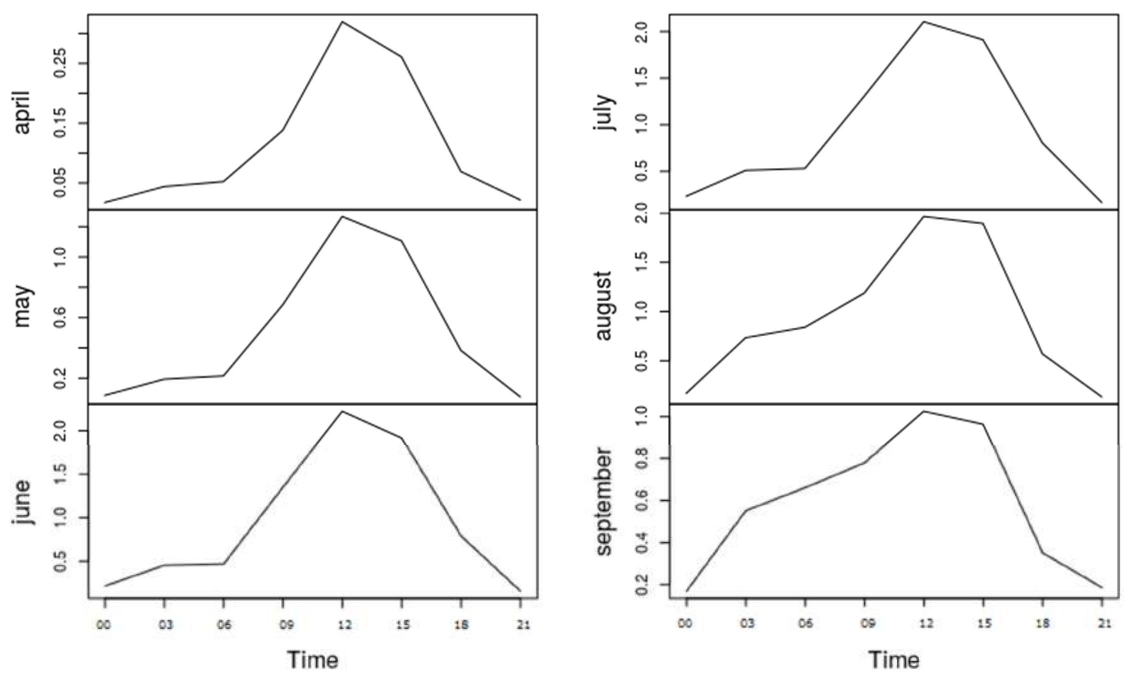

Figure 2. *Daily mean number of flashes recorded at every 3 hour time step*

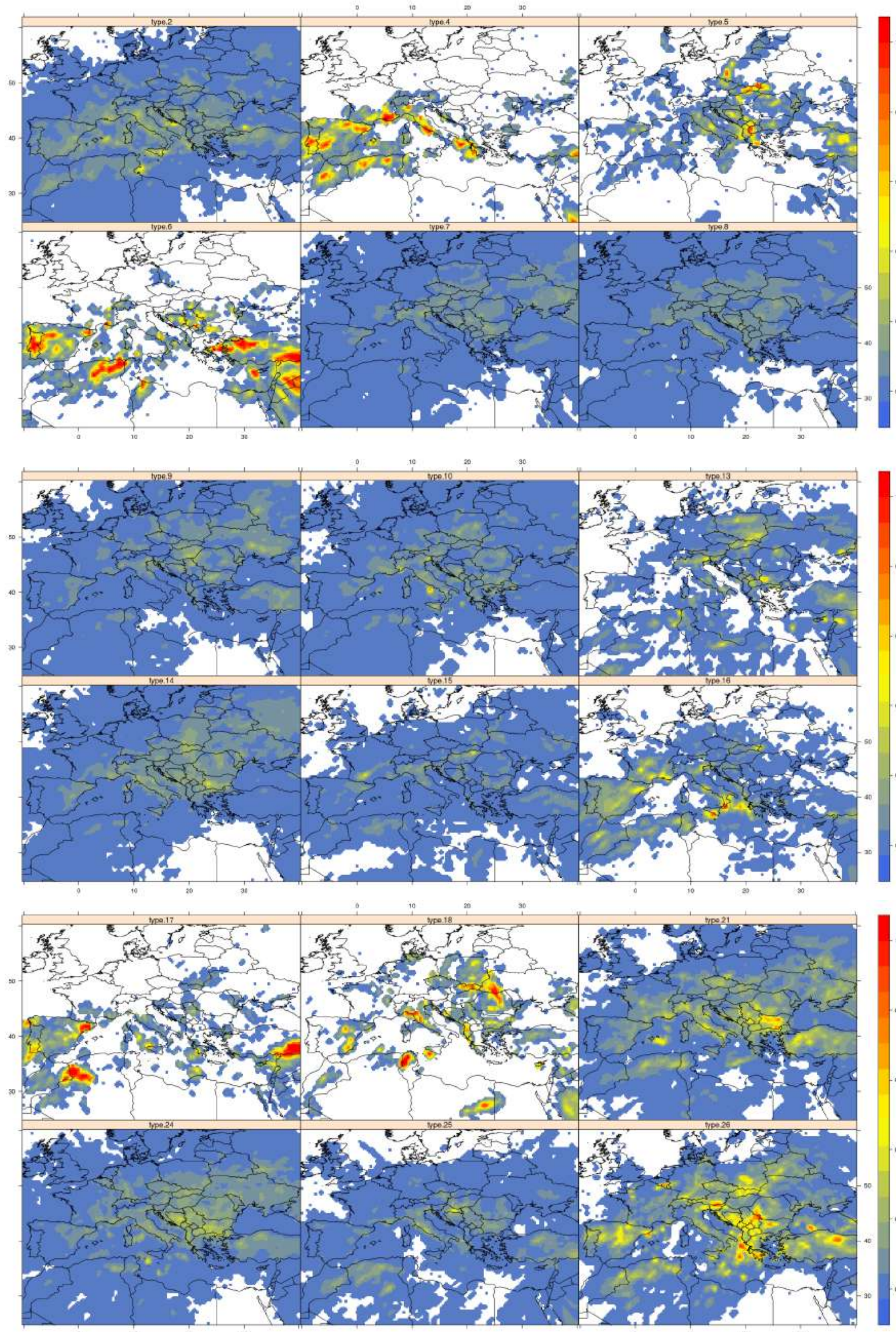

Figure 3. *Frequency of lightning occurrence upon CTs*

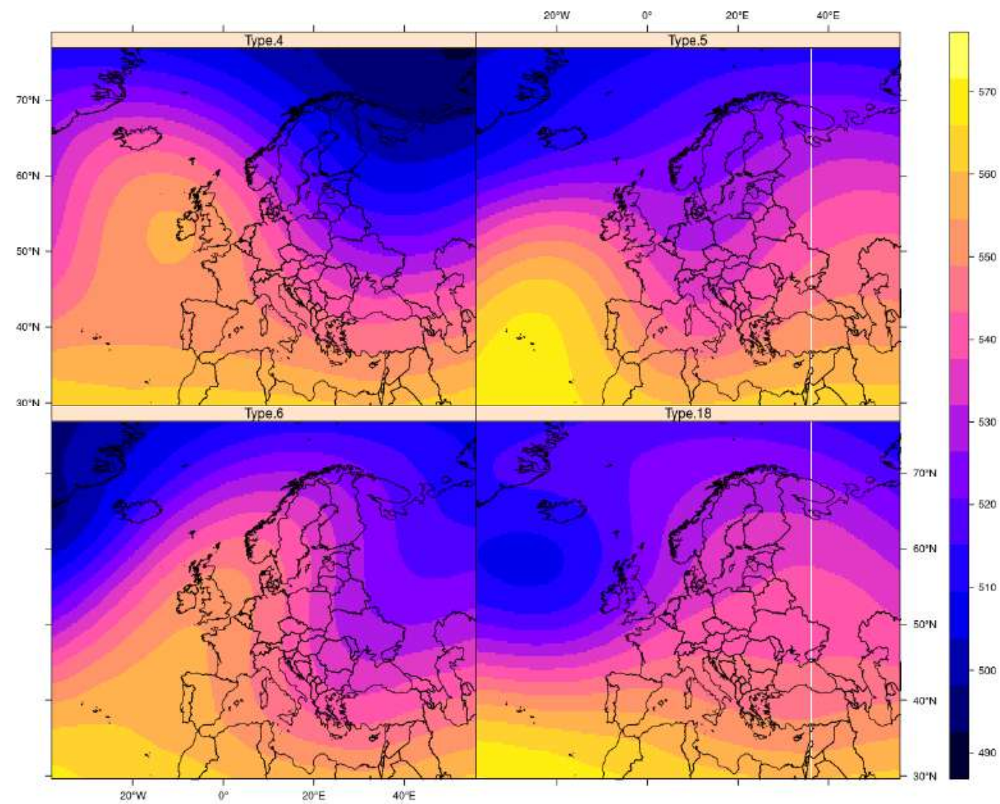

Figure 8*. Mean 500 hPa geopotential for synoptic types responsible for high frequency of lightning*

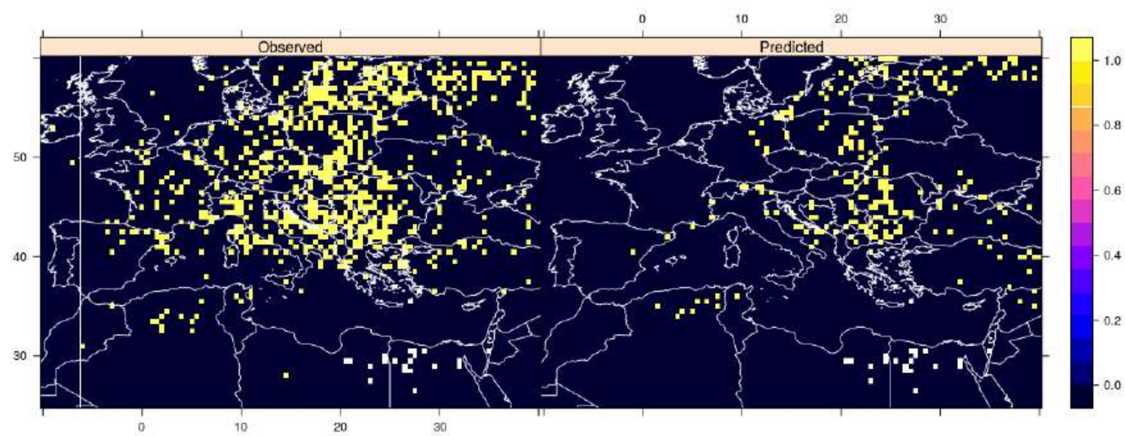

Figure 9. *Logistic regression model prediction output for the month of July (2005-2014). LEFT: Testing dataset (25 % of the total cells recording lightning flashes); RIGHT: the prediction map built on the testing dataset using the output of the regression model*

FUTURE COLLABORATIONS (if applicable)

As said before, the main objective of this STSM was to work along with researchers from National Observatory of Athens, especially Mr. Kostas Lagouvardos and Ms. Vassiliki Kotroni on a common scientific paper regarding lightning activity upon synoptic weather types. After the end of my STSM, I will plan to submit the draft paper to an according journal in the immediate period. ZEUS lightning data proved to be very valuable so far in investigating the atmospheric instability. After this paper will be done, we have in plan to continue our research especially in lightning activity and links with other atmospheric parameters.