Applying AdaBoost algorithm on multiclass OvA-SVM for the delineation of rainy clouds using multispectral MSG-SEVIRI data

Amar Belghit[1*], *Mourad* Lazri[1], *Ali Hamroun*[1], *Karim* Labadi[2], and *Sadjia* Hamdad[1]

[1]Laboratoire LAMPA (laboratoire d'analyse et de modélisation des phénomènes aléatoires), Faculty of Electrical Engineering and Computer Science, University Mouloud MAMMERI of Tizi Ouzou (Algeria).

[2]Laboratoire LR2E, ECAM-EPMI / Quartz-Lab Cergy pontoise (France).

*Corresponding Author: amar.belghit@ummto.dz

Abstract. The use of artificial intelligence and machine learning methods has become a very useful and efficient choice in precipitation retrieval from meteorological satellite data. In this work, we implement the AdaBoost algorithm to optimize and enhance the performance of the classification and delineation of precipitating clouds in northern Algeria carried out by multiclass One-versus-All Support Vector Machine (OvA-SVM). The model developed which combines the AdaBoost algorithm with a multiclass OvA-SVM is applied to images from the MSG-SEVIRI (Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imaging) satellite, with Sétif meteorological Radar data for training and testing validation phases, in which we also did the tuning for setting the adequate number of iterations to stop the AdaBoost ensemble algorithm. In order to evaluate the elaborated model, two classification techniques used previously for rainy clouds delineation in our study region, namely the Convective/Stratiform Rain Area Delineation Technique (CS-RADT) and the Random Forest technique (RFT) are applied for comparison with our built model. The classification results obtained show that AdaBoost with OvA-SVM (AdaOvA-SVM) presents very interesting performances where the evaluation parameters POD, POFD, FAR, BIAS, CSI and PC indicate the values 95.2%, 12.4%, 14.7%, 0.9, 88.1% and 96.5% respectively. Indeed, the AdaOvA-SVM technique has outperformed the CS-RADT and RFT techniques showing better cloud classification performances. At the end of this study, it is shown that the AdaBoost can improve and optimize the classification accuracy of the multiclass OvA-SVM used as its weak classifier.

1. Introduction

Precipitation retrieval at high spatio-temporal resolution is of a very important value, and plays a significant role in the understanding of the hydrological cycle and climate parameter predictions, leading to a proper management of the water resources and the forecasting of climate variations, and also the delineation of different precipitation types and processes of precipitation which is in the front of our interests. However, quantifying precipitation for these applications remains a significant challenge, this is primarily due to the inadequate density of ground measurement instruments such as raingauges and meteorological radars, which remain critically sparse across several parts of the study area.

Indeed, these instruments traditionally used for the direct measurement of precipitation for their accurate ground truth information are limited in terms of spatial coverage in areas, such as desert, sea or mountainous regions, leaving us without much ground truth precipitation measurement.

For this purpose, in order to classify and measure precipitation, data acquired from sensors onboard meteorological geostationary weather satellites are called upon. Which offer large amounts of information on cloud properties, in addition to their high spatio-temporal and spectral resolution and availability, in such zones where ground truth measures are lacking.

However, the data provided by these satellites like the MSG (Meteosat Second Generation) satellite do not have a direct link to cloud types and measurements of precipitation rates. Therefore, several mathematical models are used to establish non-linear relationships between remote sensing meteorological satellite information and precipitation classes and rates [1,2]. Based on the conceptual model that precipitating stratiform clouds are characterized by a sufficiently large optical thickness and a large radius of cloud particles, in addition to convective precipitating clouds detected from cloud top temperatures given by InfraRed channels of these satellites , these methods have significantly improved the results of classifications and estimation of precipitation intensities in the mid-latitude regions where convective and stratiform precipitations coexist [2].

To further take advantage of the quantity and variability of these meteorological satellite data for precipitations estimation, classification algorithms based on artificial intelligence and machine learning are massively introduced and used in this field , such as artificial neural networks ANN and support vector machine SVM [3] , Random forest RF [4] and also some other more recent works combining multiple machine learning algorithms to build precipitation classification and estimation models. Which have proven to be very effective in the classification and estimation of precipitation intensities. However, despite the promising accuracy levels achieved by current Machine Learning (ML) classifiers, several challenges remain in satellite-based rainfall retrieval and cloud classification. Specifically, issues such as overfitting frequently limit the model's ability to generalize beyond the training phase. Furthermore, the rapid advancement in meteorological sensor technology is providing an unprecedented volume of high-resolution data, necessitating more robust and scalable classification methods. In an attempt to solve this issue some research works have tried to optimize and improve the classifiers accuracy by increasing the training databases. Other works made adjustments to classifier architectures or tried modifications to input parameters without obtaining any significant improvement.

The trend nowadays is moving towards the use of Ensemble learning techniques to increase classification accuracy and diminish the over-fitting issues. These methods combine multiple classifiers called weak learners to iteratively classify the data set and then give a reliable answer or model by combining the results of the iterations. One of the best-known ensemble methods are Boosting algorithms [5]. The Boosting algorithm is based on combining a succession of learned predictors sequentially, where each predictor attempts to improve on the performance of its predecessor by correcting its errors to then finally give us a final result combining these different predictors. Among these algorithms, AdaBoost acronym for ‘Adaptive Boosting’ [6]. The success and efficiency of AdaBoost can be attributed to the fact that it combines classification results of several diverse weaklearners composing this algorithm, which leads to improved classifier performance and decreasing the errors.

For the implementation of AdaBoost, the choice of it’s weak classifier is a crucial part and remains controversial. According to the fundamental theory of AdaBoost, a candidate model is considered a 'weak learner' provided its predictive performance marginally exceeds that of a random guess [6]. In the case of using SVMs as weak learners of the AdaBoost algorithm, research results diverge. Some works support the idea and have shown its good results [7]. Others consider the SVM to be a strong classifier not suitable to be used as a weaklearner for AdaBoost [8]. Indeed, in this study we aim at applying the AdaBoost algorithm on multiclass SVM model, in order to verify the applicability of our chosen multiclass SVM model namely the One-versus-All SVM (OvA-SVM) as the weaklearner for AdaBoost algorithm.

On the other hand, checking the ability of the AdaBoost algorithm in improving the accuracy of OvA-SVM and optimizing its performances in the classification of rainy clouds into (convective, stratiform and non-precipitating clouds) using the chosen multispectral input parameters from the MSG-SEVIRI satellite, by exploiting the optical and microphysical properties of clouds as input parameters of our algorithm. Data from the Sétif weather radar are used as references for learning and validating the technique.

The rest of this work is organized as follows:

A brief description on the study region and the data used is presented in section 2. Section 3 and 4 contain the scientific context of the study and the methodology where the mathematical concept of AdaBoost and multiclass OvA-SVM were exposed. In section 5 training and parameter tuning of our model are presented and in section 6 we give classification and evaluation results of our developed model. A conclusion and perspectives of the study are presented in section 7.

2. Study area and data

2.1. Study area

The study area is located in the North-East of Algeria (See Figure 1). This region is known for its Mediterranean climate where precipitation from convective systems and stratiform systems coexist.

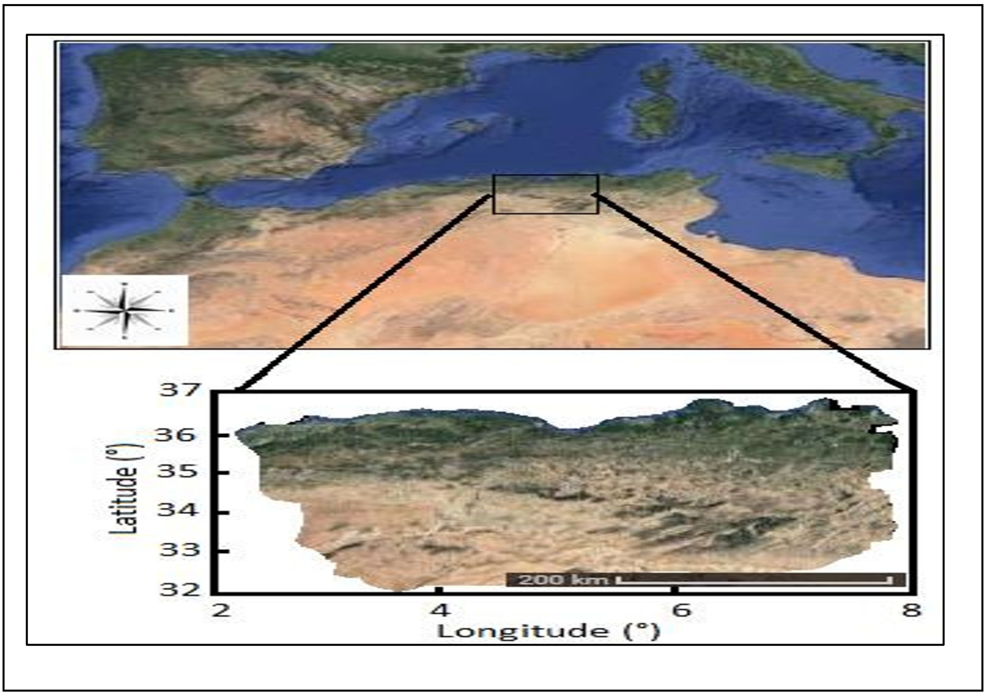

Figure.1. Study region (North-East of Algeria).

2.2. Data

This study uses data pairs (MSG-SEVIRI data/ Radar data). MSG satellite data is used as inputs for classification of precipitation types over the study region using our AdaBoost-SVM model.

As for Radar data, they are used as output references for learning and validation of elaborated models. Data from three rainy seasons is used to develop our model; for learning, we took the data collected during the 2008-2009 and 2009-2010 seasons, and for validation, we considered the data recorded during 2010-2011 season (Table 3).

2.2.1. MSG-SEVIRI data

MSG (Meteosat Second Generation) satellite is a capable meteorological satellite, providing useful earth and clouds observation information and data, thanks to the advanced resolution of it's SEVIRI (Spinning Enhanced Visible and Infrared Imaging) radiometer which provides 12 images every 15 minutes, captured in 12 different frequency bands [9].

MSG-SEVIRI data providing implicit information on the optical and microphysical properties of clouds are used as input parameters of our machine learning model, for the classification of precipitation into three classes (non-precipitating, convective and stratiform).

The optical and microphysical properties of clouds considered for the classification and delineation of precipitating clouds in this work are: The vertical extension of the clouds, existence of ice particles in the clouds, cloud top temperatures (CTT), the particle size in the clouds, the optical thickness of the clouds and the evolution of the clouds in time.

To obtain information on these optical and microphysical properties of clouds, we have selected the following different MSG-SEVIRI channels: VIS0.6, NIR1.6, IR3.9, WV6.2, WV7.3, IR8.7, IR10.8 and IR12 to use different combinations of their brightness temperature differences as inputs for our model. Depending on their (Daytime and Nighttime) availability, we have considered combinations for the day and others for the night (see Table 1).

Table 1. Spectral parameters used as inputs and corresponding possible values.

Input parameters	Possible values		Information on
	Daytime	Nighttime	
$T_{10.8}$(Kelvin)	207.2k to 283.9k	205.3k to 282.4k	Vertical cloud extent and cloud top temperature [2].
$\Delta T_{10.8-12.0}$(Kelvin)	-0.3k to 7.4k	-0.3k to 7.1k	Existence of ice particles in the clouds [2].
$\Delta T_{8.7-10.8}$(Kelvin)	-4.6k to 1.3k	-4.8k to 1.7k	Existence of ice particles in clouds [10].
$\Delta T_{6.2-10.8}$(Kelvin)	-50.1k to 6.4 k	-51.8k to 5.1 k	Vertical cloud extension, cloud top temperature [2].
$\Delta T_{7.3-12.1}$	-50.3k to 6.6 k	-52.0k to 5.7 k	Vertical cloud extension, cloud top temperature.
$R_{0.6}$ (µm)	0.02µm to 1 µm	No used	Cloud Particle Size and Cloud Optical Thickness [2].
$R_{1.6}$ (µm)	0.03µm to 1 µm	No used	Cloud Particle Size and Cloud Optical Thickness.
$\Delta T_{3.9-10.8}$ (Kelvin)	No used	-10.3k to 15.1k	Cloud Particle Size and Cloud Optical Thickness.
$\Delta T_{3.9-7.3}$ (Kelvin)	No used	-4.9 k to 25 k	Cloud Particle Size and Cloud Optical Thickness.
$\Delta T_{10.8(t)-10.8(t-1)}$	-75.8k to 75.4k	-76.4k to 76.2k	Cloud Time Evolution.

2.2.2. Radar Data

The radar data is considered as the reliable reference for training and evaluation of the developed model, since thresholds on the radar reflectivity values (Z) allow us to separate the cloud pixels into their true precipitating classes in order to train and evaluate the classifications achieved by our model.

The data in our study is collected with a spatial resolution of $1 \times 1 Km^2$ and a temporal resolution of 15 minutes by Sétif Rain Radar (SRR).

The radar is installed near the city of Sétif, Algeria at $36^0\ 11'N, 5^0 25'$E and 1700m of altitude. This SRR provides images in the Plan Position Indicator (PPI) in 512×512 pixel format, each pixel is coded on 4 bits giving 16 possible levels of reflectivity in dBZ [04, 12, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, 70] which represents the physical parameter of the radar denoted (Z) which can also be converted into precipitation intensity R(mm/h), using the adapted Z-R relation in this study region.

For the training and validation of the model's output, we took three precipitation classes which are: the convective class, stratiform and non-precipitating class, fixed from the radar data according to the corresponding interval values of the Reflectivity Z in dBZ (Table 2).

Table 2 shows the intervals and thresholds of the radar reflectivities (Z) in dBZ adopted for the delineation of the different precipitation classes.

Table 2. Classes of precipitation and radar reflectivity corresponding intervals in dBZ.

Radar Reflectivities (Z)	Classes	Intensity Level
pixel>38dBZ	Convective	High to very high intensities
04dBZ<pixel<=38dBZ	Stratiform	Low to moderate intensities
Pixel <=04dBZ	No rain	None at very low intensities

As shown in Table 2, the radar reflectivity of 38dBZ was adopted as the threshold to separate between stratiform and convective precipitation, while radar pixels with reflectivity below 04dBZ were considered non rainy [1].

2.2.3. Correspondence between satellite data and Radar data

To match the satellite data to the radar data, for spatial comparison between the two types of data, radar data with an original spatial resolution of (1 x 1 km^2) was reprojected and resampled to the spatial resolution of SEVIRI images in the study area which is approximately (4 x 5 km^2).

To reduce parallax shift and collocation errors, for the upper-level clouds, we averaged the pixels over 5x5 pixel windows and 3 x 3 pixel windows for the lower level clouds.

On the other hand, due to the low time shift between the two data types which is very minimal (3 min), it is not necessary to perform time synchronization.

3. Scientific background of the study

This ensemble machine learning concept is introduced to improve the classifications obtained from a single classifier and enhance the precision of the overall results. Multiple trainings of a classifier using weighted datasets, produce better and more reliable results when all these results are combined. The final classification is obtained either by taking an average or by majority vote of the individual classifications according to the strategy of the adopted classifier.

These ensemble methods aim to reduce the variance and/or the bias in classifiers, where the bias indicates how far the average classification obtained by the algorithm is from the real data. As for the variance, it indicates the differences of the results between several classifiers and iterations of the same algorithm.

AdaBoost which is one of the main ensemble algorithms allows also for the balancing of the accuracy/ diversity dilemma which is encountered by lots of machine learning classifiers in this level.

The AdaBoost (Adaptive Boost) algorithm was developed by Freund and Schapire [6], based on notions from the Boosting algorithm [5]. It is an algorithm that combines several sequentially trained models of an unreliable weak classifier to generate a robust final decision. The condition for AdaBoost to perform well in a classifier is that this base classifier is not robust but rather just slightly more accurate than a random classifier.

One of the objectives of this work is to answer this question concerning the practicality of using AdaBoost algorithm on a multiclass SVM variant. We therefore use the AdaBoost algorithm on the SVM classifier, to check if it is really capable of improving the accuracy of our multiclass SVM classifier and optimize its performance. To do this, we implement a certain classification strategy using a multiclass-SVM classifier on which we apply the AdaBoost algorithm.

4. Methodology

In order to reach the objectives of our study, namely the classification and delineation of precipitating cloud types, and the highlighting of the contribution of the AdaBoost algorithm on the enhancement and optimization of the classification accuracy given by the application of the chosen multiclass SVM, we have implemented an SVM multiclass model, which is the One-versus-all SVM (OvA-SVM).

Then the AdaBoost algorithm is applied downstream to study it's compatibility with our chosen multiclass SVM strategy, and its capacity to enhance and improve multiclass SVM classification accuracy. Before the description of our methods scheme and strategy, we give the mathematical concepts of the AdaBoost algorithm and the SVM classifier as well as our chosen multiclass OvA-SVM variant.

4.1. Mathematical concept and steps of AdaBoost

AdaBoost (Adaptive Boosting) is a sequential iterative algorithm, originally introduced by Freund and Schapire [6]. It is an ensemble learning method that seeks to build a "strong" learning algorithm based on a succession of "weak" classifiers.

The main intuition behind AdaBoost is to run a given weak learning algorithm repeatedly while modifying probability distributions *D*, on the training data. This distribution is initially uniform with the same weight attributed for all the samples of the data. As iterations go, it assigns higher weights to samples misclassified by a weak classifier, so that the new weak classifier of the following iteration can reduce the classification error by focusing on them. At the same time, lower weights will be assigned to correctly classified samples. Meanwhile, the *D* distribution representing the weights of the samples is updated after each iteration. In the end, the hypotheses h_t produced by the weak learner of each cycle 't' are combined into a single "strong" hypothesis.

Since the AdaBoost has been very effective in solving two-class classification problems. To obtain a multi-class classification, multiple binary classifications should be performed.

The AdaBoost is mathematically formulated as follows:

Consider the pairs (x_1, y_1), ….., (x_n, y_n) where $x_i \in X$ are the characteristics (features) of individual *I* and $y_i \in Y = \{-1, +1\}$ the variable to be predicted (labels).

AdaBoost initially assigns the weight associated with each sample 'i 'according to distribution D:

$$D_t(i) = {}^{1}/_{n}, i = 1, 2, \dots, n \tag{1}$$

For $t = 1, 2, \dots, T$:

- ➢ Determines the hypothesis h_t : X to $\{-1, +1\}$ which minimizes the classification error ε_t in iteration t as a function of the weights D_t according to the following equations:

$$h_t = arg \min_{h \in H} \sum_{i=1}^{n} D_t(i)[y_i \neq h(x_i)] \tag{2}$$

$$\varepsilon_t = \sum_{i=1}^{n} D_t(i)[y_i \neq h(x_i)] \tag{3}$$

ε_t is the model (weak learner) error in iteration t.

- ➢ If $\varepsilon_{min,t} < 0.5$ the function h_t is selected, otherwise the AdaBoost algorithm stops

- AdaBoost then calculates the coefficient of h_t : $\alpha_t \in R$, with:

$$\alpha_t = \frac{1}{2} ln \frac{1-\varepsilon_t}{\varepsilon_t} \tag{4}$$

- Then updates the weight of the samples for the next iteration

$$D_{t+1}(i) = \frac{D_t(i) e^{-\alpha_t y_i h_t(x_i)}}{2\sqrt{\varepsilon_t(1-\varepsilon_t)}} \tag{5}$$

When the AdaBoost algorithm stops at iteration K, the classifier resulting from the boosting process is:

$$H(x) = sign\ (\sum_{t=1}^{K} \alpha_t h_t(x)) \tag{6}$$

4.2. Mathematical concept of Support Vector Machine (SVM)

The SVM (Support Vector Machines) algorithm aims to find the optimal separating hyperplane (OSH) between two classes by maximizing the margin (m) between the two closest samples of these classes located on the boundaries [11].

In a binary classification problem, the decision function of SVM is:

$$f(x) = \langle w, \emptyset(x) \rangle + b \tag{7}$$

Where *∅(x)* is the mapping of element *x* from the input space into a higher dimensional feature space. The ⟨.,.⟩ denotes the scalar (dot) product in the new feature space. For nonlinearly separable point problems, SVM uses a function that handles the tradeoff between classification errors versus margin maximization rate. Consequently, a constraint relaxation variable or "slack": $\xi_i \geq 0$ is introduced for each training example, so that it sets a tolerated margin of error on the separation of the boundaries of the two classes, to classify all data samples.

The optimal values of *w* and *b* can be obtained by solving the following optimization problem:

$$\text{Minimize : } g(w, \xi) = \frac{1}{2}||w||^2 + C \sum_{i=1}^{N} \xi_i \tag{8}$$

$$\text{Subject to: } y_i(\langle w, \emptyset(x) \rangle + b) \geq 1 - \xi_i \ , \xi_i \ \geq 0 \tag{9}$$

Where ξ_i is the *ith* slack variable and 'C' is the parameter that controls the trade-off between the error tolerated in the classification, and the widening of the margin and $(C.\sum_i \xi_i)$ is a penalty term.

According to Wolfe's dual form, the minimization problem (equations 8 and 9) can be written as (equation 10 and 11):

$$\text{Minimize : } W(\alpha) = -\sum_{i=1}^{N} \alpha_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j k(x_i, x_j) \tag{10}$$

$$\text{Subject to: } \sum_{i=1}^{N} y_i \alpha_i = 0, \ \forall i: 0 \leq \alpha_i \leq C \tag{11}$$

Where α_i is a Lagrange multiplier that corresponds to the sample x_i, and *k(.,.)*: is a kernel function that implicitly maps the input vectors into an appropriate higher dimension feature space, by giving the scalar product between these vectors in the new higher dimension space.

$$k(x_i, x_j) = \langle \emptyset(x_i), \emptyset(x_j) \rangle \quad (12)$$

For classification with multiclass SVM, in this study we chose the use of the Gaussian kernel (RBF kernel), since studies like [12] showed that the application of SVM with this RBF kernel as the AdaBoost weaklearner has led to good accuracy and can be implemented efficiently.

The Gaussian RBF (Radial Basis Function) is given by the following expression:

$$k(x_i, x_j) = exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \quad (13)$$

The samples are transferred nonlinearly into a high-dimensional feature space. Where, an optimal separating hyperplane is constructed by the support vectors. The support vectors correspond to the centers of the RBF nuclei in the input space and they represent the samples situated on the edges of the margin separating the classes. The generalization performance of SVM depends mainly on the kernel parameters, in RBF kernel case it is parameter *σ*, and the adjustment parameter, *C*. They must be defined beforehand.

The SVM approach [11] is formally defined for two-class problems. For the case of multiclass SVM, it is necessary to extend it to the K-class problem with labels $Y \in \{Y_1, \ldots, Y_k\}$. There are several variants for the design of a multiclass SVM using a combination of binary classifiers [13,14].

In this work we opted for the use of the One-vs-All-SVM strategy to build our multiclass SVM model, which has the following functioning principle:

- **One versus All (OvA) variant:**

In this type of strategy, according to the number K of classes it is necessary to construct in turn K discriminating binary models, with the modality y_k for each class against all the others combined, $Y' \in \{Y_k=+1, \ldots, Y_{(k)}=-1\}$. We obtain K decision functions $f_k(x)$ which give us each the probability of element *x* belonging to class K. In the end of the classification, the class K with the highest score probability is selected as the class to which element *x* belongs using the following expression:

$$\hat{y} = arg\,\max_k f_k(x) \quad (14)$$

The One-versus-all strategy consists in training several binary models, which will specialize each one of them in the recognition of a class *K* amongst all the others. The implementation of this OvA-SVM variant requires *K* learnings to be performed on the data, each iteration will be recognizing and separating one of the *K* classes from all the others. However, for this variant, an imbalance of classes in the construction of individual binary models leads to poor learning, hence the optimizing role of AdaBoost algorithm which we apply on this OvA-SVM model is needed (see Figure 3).

With this OvA-SVM strategy, for our three-class study case, three binary classification nodes (binary SVM) in parallel are performed (see Figure 2); where classification1: distinguishes (no rain against convective and stratiform), classification2: distinguishes (stratiform against convective and no rain), classification3: distinguishes (convective against stratiform and no rain). Figure 2 shows the diagram of the OvA(OneVsAll) variant used for the construction of our multiclass SVM (OvA-SVM).

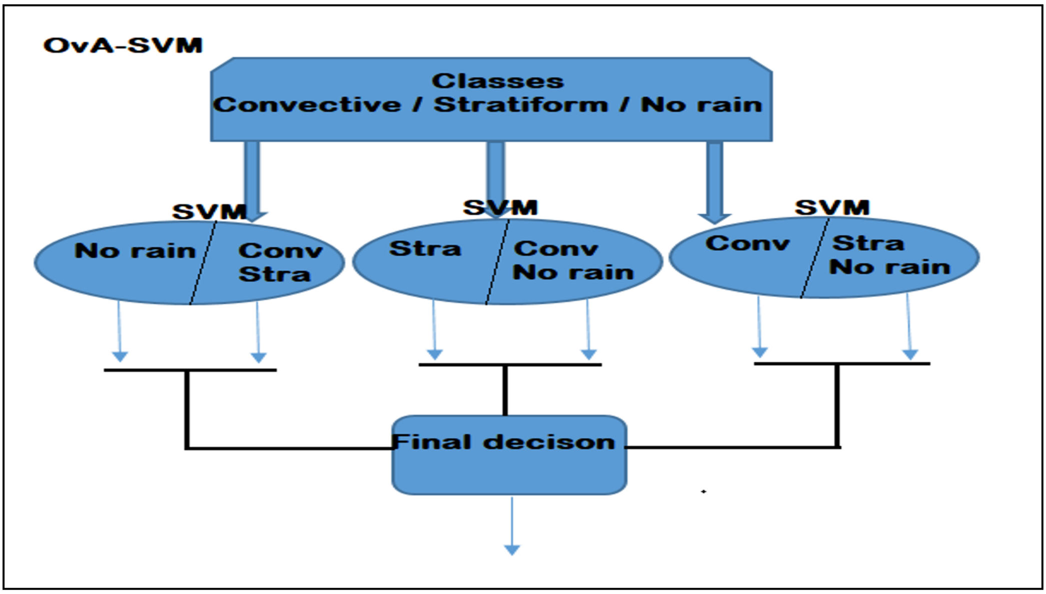

Figure.2. Diagram of the multiclass SVM model OvA-SVM.

4.3. Implementation of the classification model OvA-SVM with AdaBoost

In this part we describe the implementation of our classification method AdaBoost-OvA-SVM (Ada-OvA-SVM), by applying the AdaBoost algorithm on our chosen multiclass SVM model OvA-SVM.

The AdaBoost algorithm implementation is more efficient in the case of binary classifications problems. For which they were initially designed [6]. Thus, to better optimize the classification performance, for each binary separation node output by our multiclass OvA-SVM model, the AdaBoost algorithm is applied (see Figure 3).

Therefore, to solve multi-class problems in our study of precipitating cloud classification from MSG images, it suffices to apply a series of binary classifications by adopting the OvA (One vs All) strategy to build the multi-class SVM for the classification of precipitation types.

After that we applied for each binary output of the OvA-SVM classification a routine of the AdaBoost algorithm (Figure 3).

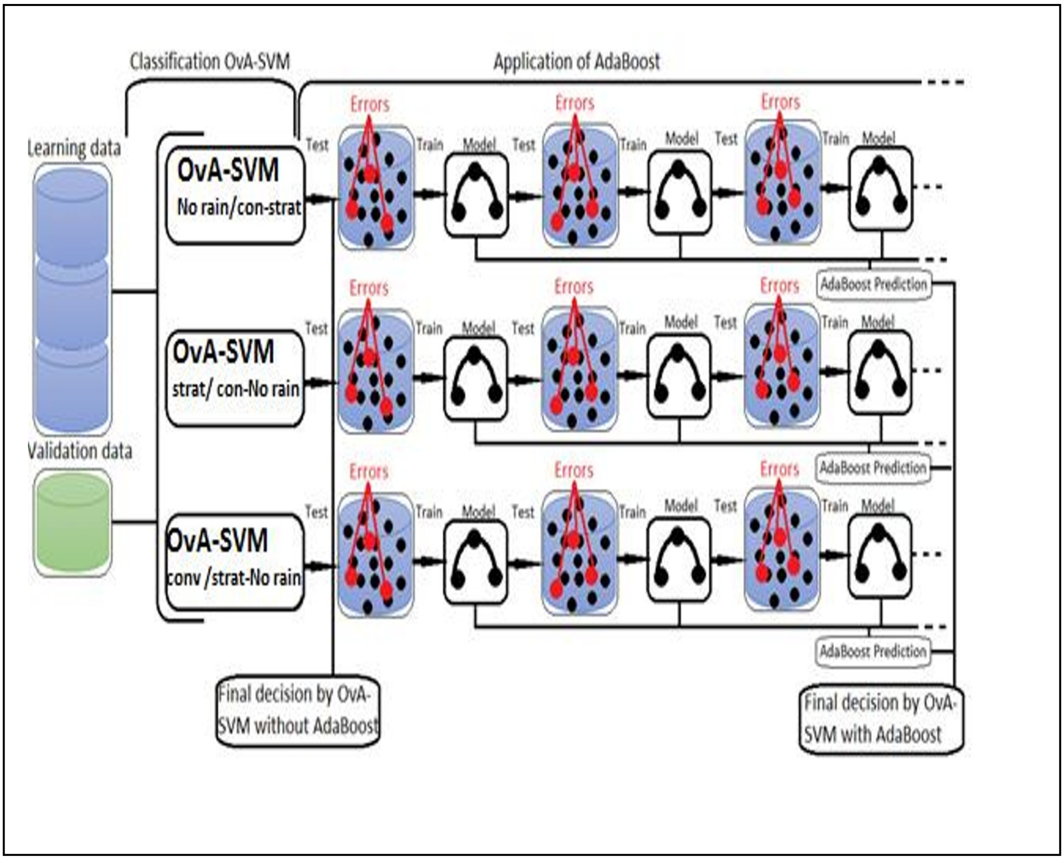

Figure. 3. Application diagram of the AdaBoost Algorithm on OvA-SVM model.

Figure.3 shows the combination of the AdaBoost algorithm with the One-versus-All-SVM multiclass classifier. Where three binary classifications (no rain/all; stratiform/all; convective/all) performed in parallel using binary SVMs to which the AdaBoost algorithm was applied for each binary output.

5. Training and parameter tuning of the model

5.1. Training and validation database

In order to implement the developed AdaBoost-OvA-SVM model in our study, we used a database consisting of three rainy seasons (2008/2009, 2009/2010 and 2010/2011). Table 3 gives the size and distribution of the dataset used during learning and validation steps of the developed model OvA-SVM with and without AdaBoost, and the number of pixels per class for the two periods.

Table 3. Periods of training and validation databases with the number of instant precipitation scenes (IPS) and of pixels per class for each period.

	Learning period						Validation period		
	Rainy season (2008-2009) 1799 IPS			Rainy season (2009-2010) 2080 IPS			Rainy season (2010-2011) 1287 IPS		
Class	Convective	Stratiform	No rain	Convective	Stratiform	No rain	Convective	Stratiform	No rain
Pixels observed by radar	241044	874452	3541165	296542	986855	3805477	195745	745210	3021456
OvA-SVM	Learning			Application			Classification		
AdaBoost	No operation			Learning			Optimization of classification		

Instant precipitation scene (IPS) is an MSG-SEVIRI scene that contains at least one pixel detected as precipitating. Situations corresponding to clear skies are not taken into account.

5.2. Tuning of the number iterations for AdaBoost

With the application of the built Ada-OvA-SVM model on classifying one instantaneous precipitation scene (IPS), we realized the tuning of the number of iterations for the AdaBoost algorithm, setting it to the adequate value which gives us the best classification accuracy of the Ada-OvA-SVM and allows us to avoid the over-learning problem.

The instantaneous precipitation scene on which we applied the developed classification model is taken during the passage of a cyclone above the study region. The cyclone is charged by convective and stratiform precipitation. The scene is collected by the MSG satellite on November 29, 2010 at 12:00 UTC pm.

The classification accuracy of the AdaBoost-OvA-SVM as a function of the number of AdaBoost iterations, is indicated in the Table 4. Where it shows the classification accuracy situations of our AdaBoost-OvA-SVM classifier at iteration 0 before AdaBoost was applied on the OvA-SVM, then after applying the AdaBoost at iterations: 100, 200, 300 and 400 to get an overall idea of the impact of AdaBoost on improving the classification done by the OvA-SVM, and the number of iterations at which the AdaBoost is most efficient giving us the best classification accuracy of AdaOvA-SVM model.

Table 4. Classification Accuracy after 0, 100, 200, 300 and 400 iterations for AdaBoost-OvA-SVM.

	Classification Accuracy (%) according to iterations number				
	0 iterations (OvA-SVM without AdaBoost)	100 iterations	200 iterations	300 iterations	400 iterations
AdaOvA-SVM	52.4	69.5	84.2	90.4	95.1

As shown in (Table 4) the application of the AdaBoost algorithm on the OvA-SVM classifier significantly reduced the number of misclassified pixels by reclassifying them correctly for all classes thus increasing the classification accuracy. The resulted classification accuracy for AdaOvA-SVM goes from 52.4% to 95.1% at iteration 400, passing by 84.2% at iteration 200.

In order to better illustrate the effect of the number of AdaBoost iterations on the accuracy of the resulting AdaOvA-SVM classifier, we plotted the curve representing the classification accuracy as a function of the AdaBoost iterations number (0 to 412 iterations) (see Figure 5).

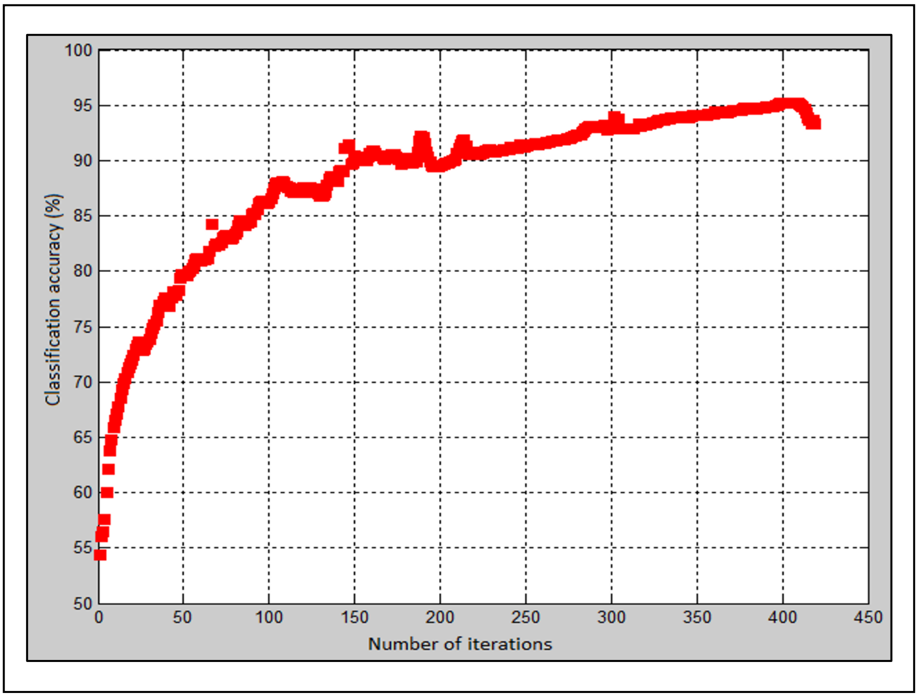

Figure. 5. AdaOvA-SVM Classification accuracy according to number of AdaBoost iterations.

The more we continue to apply the AdaBoost the classification accuracy increases, it stabilizes and reaches it's peak value of 95.1% at iteration number 400. Passed this value there is no improvement and the AdaBoost resulting classifier can overfit the training samples and therefore lead to poor generalization performance. It is therefore necessary to stop the learning cycles of AdaBoost at the appropriate time. Thus, we set the adequat number of iterations at 400, which we used in the rest of our study.

6. Application and evaluation of our model

6.1. Application on all scenes of Ada-OvA-SVM

In this section, we present the classification results by applying the built Ada-OvA-SVM model on all the precipitation scenes of the 2010/2011 rainy season (table 3). The goal is to show in an objective way the contribution made by the AdaBoost in improving and optimizing the classifying accuracy and performance of the multiclass OvA-SVM.

To do this, we calculated the evaluation parameters using the statistical coefficients: Probability Of Detection (POD), the Probability Of False Detection (POFD), the False Alarm Ratio (FAR), The Critical Success Index (CSI), and The Percentage of Correct (PC) and The Frequency BIAS index (Bias), obtained by comparisons with real classes of the pixels in meteorological radar data, to show the evaluation of the performances of both OvA-SVM without AdaBoost and after the application of AdaBoost for the AdaOvA-SVM model, we obtain the statistical results shown in Table 6.

Table 6. Evaluation parameters for performance of OvA-SVM with and without AdaBoost for convective and stratiform classes.

	Convective class						**Stratiform class**					
	POD (%)	POFD (%)	FAR (%)	BIAS	CSI (%)	PC (%)	POD (%)	POFD (%)	FAR (%)	BIAS	CSI (%)	PC (%)
OvA-SVM	52.2	33.2	39.3	0.6	51.2	53.1	50.4	33.5	40.2	1.6	50.2	51.7
AdaOvA-SVM	95.2	12.4	14.7	0.9	88.1	96.5	94.4	12.9	15.5	1.2	87.8	96.1
Optimal	100	0	0	1	100	100	100	0	0	1	100	100

According to results shown in Table 6, the performances of the OvA-SVM have been optimized by the implementation of the AdaBoost algorithm. All the values of the evaluation parameters indicate a clear improvement in the classifications for both the convective class and the stratiform class.

The OvA-SVM classifier has benefited a lot from the incorporation of the AdaBoost algorithm where its performance has increased. Indeed, for the convective class, the POD increased from 52.2% to 95.2% , while the POFD and FAR droped from 33.2% and 39.3% to 12.4% and 14.7% respectively. The Bias indicates an underestimation of 0.6 before it is reduced to 0.9. As for the CSI and PC, the values go from 51.2% and 53.1% to 88.1% and 96.5% respectively. The classification performance for the stratiform class mirrored the trends observed in the convective category. Specifically, the standalone SVM yielded a POD of 50.4%, a POFD of 31.5%, and a FAR of 40.2%. However, the integration of AdaBoost significantly enhanced these metrics, improving the POD to 94.4% while simultaneously reducing the POFD and FAR to 12.9% and 15.5%, respectively. There is a 1.6 overestimate shown through in the OvA-SVM without the AdaBoost; it was reduced to 1.2 with the AdaBoost. In terms of CSI and PC, the implementation of AdaBoost increased these parameters. The CSI goes from 50.2% to 87.8% and the PC goes from 51.7% to 96.1%.

All these results prove that the AdaBoost Algorithm has improved the accuracy of the classifications by implementing it on the OvA-SVM classifier, giving us an AdaOvA-SVM classifier with improved and optimized classification results.

The reason can be explained by the fact that for this type of MSG-SEVIRI data used in our study, there is a strong imbalance between the different classes especially between the non-precipitating class and the others. Where studies and experiments on unbalanced datasets showed that AdaBoost with SVM performed much better than SVM without AdaBoost [7, 12]. In this study, we leveraged the inherent sensitivity of the OvA-SVM to class imbalance. When implemented independently, the OvA-SVM exhibited suboptimal classification performance, effectively serving as a 'weak learner' within the AdaBoost framework. By integrating these approaches into a hybrid AdaOvA-SVM classifier, we achieved a more robust and accurate model tailored to the specific characteristics of our meteorological dataset.

So answering the problematic we wanted to solve in this work, we can conclude that AdaBoost algorithm can optimize and enhance the performances and accuracy of the multiclass OvA-SVM used as a weak classifier in our study case, which is applied on the type of MSG-SEVIRI data we had available to delineate precipitating clouds.

6.2. Inter-comparison of AdaOvA-SVM with other classifiers

To get an impression on the contribution of the model developed in this study AdaOvA-SVM compared to some other works already published in the literature, and used to delineate and classify rainy clouds in our study area, using the same datasets we used for our developed model AdaOvA-SVM, we implemented the following techniques for comparison:

- Convective/Stratiform Rain Area Delineation Technique (CS-RADT) [1] is a satellite rainfall retrieval technique based on various spectral parameters of SEVIRI expressing microphysical and optical cloud properties. It uses thresholding technique on multispectral parameters to distinguish between stratiform and convective clouds. The different thresholds were chosen by optimization from reference data of meteorological Radar in the study area.

- The Random Forest technique (RFT) [4] is used to classify rainfall into three classes (no-rain, convective and stratiform) using Random Forest classifier. The RFT in this method is optimized with 400 trees (ntree) and 6 features (mtry), this by minimizing the “out-of-bag” (OOB) error rate.

The overall statistical evaluation results given by applying theses different compared techniques and our developed Ada-OvA-SVM model on all the scenes in our testing datasets of MSG-SEVIRI images are shown in Table 7.

Table7. Statistical evaluation results for the different techniques compared to AdaOvA-SVM.

	POD (%)	POFD (%)	FAR (%)	Bias	CSI (%)	PC (%)
AdaOvA-SVM	94.6	9.02	24.47	1.11	85.53	95.31
CS-RADT	68.22	11.47	37.56	1.12	59.32	88.89
RFT	90.32	9.22	26.87	0.88	76.89	93.40
Optimal	100	0	0	1	100	100

The AdaOvA-SVM method shows very interesting performances. The statistical evaluation scores in Table.7. show that the AdaBoost-SVM exceeded by far the CS-RADT method and performs better than the RFT with it's very accurate POD score of 94.6% compared to 68.22% and 90.32% respectively of the other two compared methods, and also much lower error detecting scores like POFD and FAR of 9.02% and 24.47% compared to 11.47% and 37.56% for CS-RADT and to 9.22% and 26.87% for RFT. With also very much higher and more accurate scores of CSI and PC of 85.53% and 95.31% compared to 59.32% and 88.89% for the CS-RADT method and 76.89 and 93.40 for the RFT. While Ada-OvA-SVM and CS-RADT show a slight overestimation of respectively 1.11 and 1.12 and RFT show an underestimation of 0.88.

To give a visual impression on the contribution of the elaborated machine learning model AdaOvA-SVM in classifiying the precipitating clouds, we present in (Figure 6) the results of classification using our model and the different compared methods, applied on a scene of precipitations of an anticyclone taken on November 29, 2010 at 12:00 UTC pm by the MSG satellite.

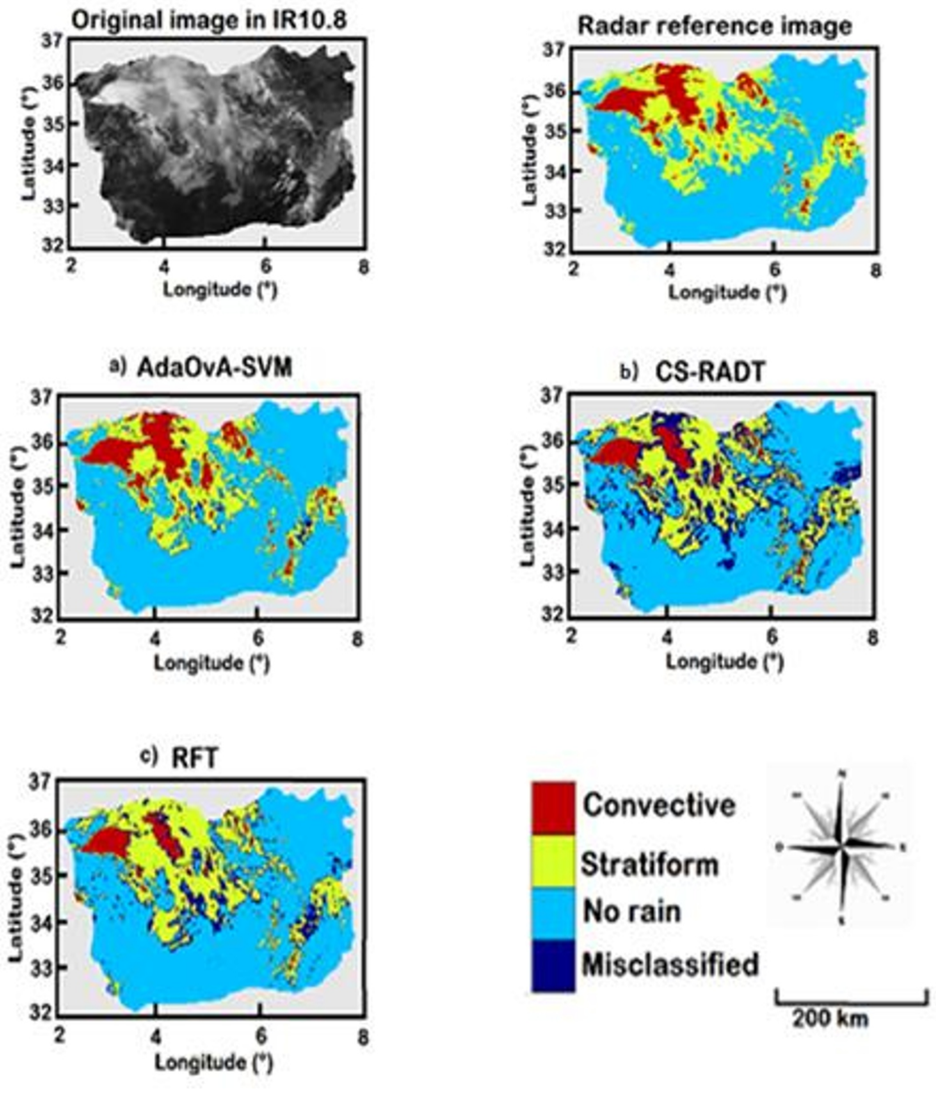

Figure 6. Classification results of the instantaneous precipitation scene (November 29, 2010 at 12:00 pm), a) classified by AdaOvA-SVM, b) classified by CS-RADT, c) classified by RFT.

Visually, as for the statistical results from the classification of all scenes, we notice that the misclassified pixels are less present in the case of AdaOvA-SVM compared to the CS-RADT and RFT methods, confirming the superiority of our developed model AdaOvA-SVM in classifying and delineating precipitating clouds compared to the other two models.

The misclassified pixels are mainly present in the edges between the convective class and the stratiform class. The no rain class is easier to classify.

7. Conclusion

The objective of this work is, to achieve a quality classification of precipitating clouds over our study region into different classes: convective, stratiform and non-precipitating, from MSG-SEVIRI satellite images, and on the other hand, to show the contribution and impact of the AdaBoost algorithm on the multiclass SVM classifier.

For this purpose AdaBoost algorithm was applied on the OvA-SVM multiclass model, with multispectral parameters from MSG-SEVIRI providing implicit information on optical and microphysical properties of the clouds that were selected and used as model inputs.

After the training of both the OvA-SVM multiclass model and AdaBoost algorithm using a training and testing database of MSG-SEVIRI parameters data, with corresponding meteorological radar images as validation data , The OvA-SVM model without and with AdaBoost have been applied to an instantaneous precipitation scene of an event where convective cells are embedded in a stratiform system, in order to do the tuning of the AdaOvA-SVM model and set the adequate value for the number of iterations of AdaBoost by which the AdaOvA-SVM gives the best accuracy of classification which we set at 400.

In the case of classification on all precipitation scenes, the evaluation of the performances of the OvA-SVM with and without AdaBoost showed that OvA-SVM classifier has benefited a lot from the incorporation of the AdaBoost algorithm, where its performance has increased and its classification accuracy has been significantly enhanced, giving us an AdaOvA-SVM classifier with improved and optimized classification results.

Thus, this study confirmed the power of the AdaBoost algorithm when combined with a weak classifier such as the OvA-SVM variant in this case study, which performed as a weaklearner due to its sensitivity to the imbalances in the study data classes.

We also carried out an intercomparison between the classifications and rainy clouds delineation results obtained by the AdaOvA-SVM method, with those of the CS-RADT and RFT methods. The AdaOvA-SVM method has shown very interesting performances. It exceeded the CS-RADT and RFT methods accuracy and classification quality results.

All these results show that the OvA-SVM can be used as a weak classifier for the AdaBoost algorithm in these types of precipitation classification studies using MSG-SEVIRI datasets, and that AdaBoost is capable of enhancing and optimizing OvA-SVM classification performance and accuracy.

Moreover, as a future perspective of this work, it is interesting to test other variants of SVM with AdaBoost, such as Newton Support Vector Machine (NSVM), Proximal Support Vector Machine (PSVM) to identify another variant that can be used as a weaklearner for AdaBoost algorithm and give better classifying results.

And it is also interesting to see the impact of using other types of data as inputs for these classifying models, such as textural information on the rainy clouds extracted from MSG-SEVIRI images, and increasing the size of the training and testing datasets.

References

[1] M. Lazri , Z. Ameur , S. Ameur , Y. Mohia , J.M. Brucker, J. Testud,. Rainfall estimation over a Mediterranean region using a method based on various spectral parameters of SEVIRI-MSG. Adv. Space Res. **52**, 1450–1466, (2013).

[2] B. Thies , T. Nauss , J. Bendix . Discriminating raining from non-raining clouds at mid-latitudes using meteosat second generation daytime data. Atmos. Chem. Phys. **8**, 2341–2349, (2008a).

[3] M.A.Tebbi., & B.Haddad., Artificial intelligence systems for rainy areas detection and convective cells' delineation for the south shore of Mediterranean Sea during day and nighttime using MSG satellite images. *Atmospheric research*, *178*, 380-392 (2016).

[4] F. Ouallouche , M. Lazri , S. Ameur . Improvement of rainfall estimation from MSG data using Random Forests classification and regression. Atmos. Res. **211**, 62–72, (2018).

[5] R.E. Schapire, . The strength of weak learnability. *Machine learning*, ***5***, 197-227 (1990).

[6] Y.Freund , R. E. Schapire, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of Computer and System Sciences, Volume **55**, Issue 1, Pages 119-139, (1997).

[7] Xuchun Li, Lei Wang and E. Sung, " A study of AdaBoost with SVM based weak learners," Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005., pp. 196-201 vol. **1**, (2005) doi: 10.1109/IJCNN.2005.1555829.

[8] J. Wickramaratna , S. Holden , B. Buxton : Performance degradation in boosting. In Kittler, J., Roli, F., eds.: Proceedings of the 2nd International Workshop on Multiple Classifier Systems MCS2001. Volume 2096 of LNCS. Springer 11–21 (2001).

[9] EUMETSAT, Applications of Meteosat Second Generation – Conversion from Counts to Radiances and from Radiances to Brightness Temperatures and Reflectance, (2004). http://oiswww.eumetsat. org/WEBOPS/msg_interpretation/index.html.

[10] T. Nauss and A. A. Kokhanovsky. Discriminating raining from non-raining clouds at mid latitudes using multispectral satellite data. *Atmos. Chem. Phys.,* ***6****, 5031–5036. (*2006).

[11] C. Cortes and V. Vapnik , Support-Vector Networks. Machine Learning, **20**, 273-297(1995).
http://dx.doi.org/10.1007/BF00994018

[12] X. Li, L. Wang., & E. Sung., AdaBoost with SVM-based component classifiers. *Engineering Applications of Artificial Intelligence*, ***21***(5), 785-795 (2008).

[13] A.C. Lorena , A.C. De Carvalho , & J.M. Gama., A review on the combination of binary classifiers in multiclass problems. *Artificial Intelligence Review*, ***30***, 19-37 (2008).

[14] P. Du, X. Bai, K. Tan. *et al.* Advances of Four Machine Learning Methods for Spatial Data Handling: a Review. *J geovis spat anal* **4,** 13 (2020). https://doi.org/10.1007/s41651-020-00048-5.