ISSN: 3067-8358
Research Article
Journal of Global Perspectives on Society, Culture, and Development

Combining Classifiers for Delineation of Raining Cloud over North Algeria and Mediterranean Sea

Bensafi Noureddine[1]*, Lazri Mourad[2], Absi Rafik[2], Labadi Karim[3] & Ameur Soltane[3]

[1]Laboratoire LAMPA, Université UMMTO, Tizi Ouzou, Algeria MCB, Université UMBB, Boumerdès, Algeria
[2]Laboratoire LAMPA, Université UMMTO, Tizi Ouzou, Algeria
[3]Quartz-Lab, ECAM- EPMI, Cergy-Pontoise, France

***Corresponding author:** Bensafi Noureddine, Laboratoire LAMPA, Université UMMTO, Tizi Ouzou, Algeria MCB, Université UMBB, Boumerdès, Algeria.

Submitted: 02 February 2026 **Accepted:** 11 February 2026 **Published:** 17 February 2026

***Citation:** Bensafi, N., Lazri, M., Absi, R., Labadi, K., Ameur, S. (2026). Combining classifiers for delineation of raining cloud over North Algeria and the Mediterranean Sea. J of Glob Perspect Soc Cult Dev, 2(1), 01-07.*

Abstract
In this paper, we presented the application of the classifier combination technique to classify cloud systems, namely stratiform and convective systems. The data used come from the SEVIRI (Spinning Enhanced Visible and Infra-Red Imager) radiometric imager on board the MSG (Meteosat Second Generation) satellite. The methods were trained and evaluated with satellite and radar data of the rain event of October 10, 2018 in the south of France provided by Météo-France. This allowed us to evaluate and compare the performances for cloud identification of two methods frequently used in combination of classifiers

Keywords: Machine Learning, Combining Classifiers, Bagging, Boosting, Stacking, Rainfall Estimation, Delineation Raining Cloud, SEVIRI, MSG, Radar.

Introduction

The techniques based on the supervised learning methods proposed, for solving the problem of identifying cloud systems and estimating precipitation, exploit the information from the multi-spectral channels of the SEVIRI (Spinning Enhanced Visible and Infra-Red Imager) radiometer imager [2, 3, 11, 12]. Some of these methods have shown a performance that surpasses the others to predict one or more classes. The idea of making these different methods cooperate in order to improve the precision of the classification appeared with the development of new called hybrid techniques. Two approaches are used for multi-classifier cooperation. The first approach is to merge information from different classifiers for decision making, considering each classifier as an independent source of information and exploiting the characteristic performance of these sources to favor the most reliable classifier to predict the class membership.

To do this, classifier fusion relies on three main strategies that are the voting principle [10, 16], possibility theory and belief theory [7, 17,1, 4]. Each of these strategies is adapted for the type of information processed or the relevant decision criterion for classification. The second approach is to use strategies or mechanisms for classifiers using different subset of data for each classifier or using different classifiers for a set of training data. These strategies are at the origin of several popular multi-classifier algorithms such as Boosting and Bagging [9,5]. It is the second approach that we will focus on this study. To do this, this paper we will detail the principle of these different mechanisms of combining classifiers used in Section 2, and by subjecting our database to tests that will be analyzed in Section 3. Finally, we draw some conclusions in Section 4.

Combining Classifiers

Combining classifiers consists of combining the prediction of a set of independent classifiers. What has been at the origin of a very active field of research in supervised learning, by exploiting the local precision of the different learning methods to form a set of classifiers, which improves the global precision for the same application [6]. Several methods have been implemented by combining several classifiers, in this section we will expose some methods, which have shown their effectiveness for the data processed in our research study.

Bagging

The principle of Bagging is "Bootstrap aggregation", it is a re-sampling of the training dataset, by randomly generating other subsets of data from the original training set, while keeping the number of samples it contains, which means the appearance of the same sample several times and the disappearance of several others in the same subset. Indeed, a larger classifier prediction error for the absent or less represented classes in the subsets. However, the combination of classifiers used for the subsets gives higher prediction accuracy than a single classifier and a training dataset. The advantage of bagging is the reduction of the variance without influencing the accuracy of the prediction. Also, the study conducted by the effectiveness of Bagging is more significant for learning methods whose prediction decision is significantly different when there is an insignificant change in the learning set [5].

Boosting

Boosting consists of a series of classifiers of different learning methods whose prediction is weak. Its principle is that the samples misclassified by the first classifier are chosen more often unlike the correctly predicted samples, with the aim of constituting the training data set submitted to the next classifier, this step is repeated for all of the classifier until the end of the series. Methods based on the Boosting principle have been developed, the Arcing method and the Ada-Boost method [9]. The Arcing method [5], the frequency of samples misclassified by a classifier is the selection criterion in the training dataset submitted to the next classifier; this step is repeated until the end of the series of classifiers. The Ada-Boost method uses a series of classifiers, which exploits the probabilities of the set of samples at the output of a classifier, and then weight the prediction error of each sample, the new set retains the completeness of the samples in the training set[8].

Stacking

The mechanism at the base of Stacking consists of two levels, a Meta level where a classifier is found that allows the combination of the classification of a set of lower level classifiers, which aims to improve the prediction of all the classifiers compared to those obtained by these same classifiers individually. The selection algorithm is based on the principle of cross-validation, a base level classifier having collected the best vote for the prediction of a class among the set of classifiers submitted to the different subsets of the training set, will be chosen by the meta level classifier to predict this same class [16].

Data and Application

Three types of climate for which their precipitation rates are important. Namely, the oceanic climate found in Western Europe bordered by the Atlantic Ocean to the west, with annual precipitation varying between 900 and 1800 mm in the form of drizzle and fog. In addition, the Mediterranean climate, with annual precipitation varying between 250 and 1150 mm, Also, the transitional climate of the mountain ranges of the Tell Atlas and the Saharan Atlas, which separates the region bordered by the Mediterranean Sea from the desert from the Sahara to the south, with annual precipitation decreasing to less than 100 mm south of the Saharan Atlas. The exclusion of the arid desert climate from our study is due to very rare precipitation in the form of downpours and of little interest for our study, with annual quantities which often do not exceed 50 mm. The rainy season lasts from October to March. A band is defined with a latitude of 30° to 50° North and a longitude of 6° West to 12° East as shown in following figure.

Data

The data used in this study are satellite and radar images, which represent a rainy and stormy event in the Mediterranean and going back to the south of France from October 10, 2018 from 08h GMT to 18h GMT. In addition, precipitation intensities exceeding 30 mm/h and cumulative rainfall more than 250 mm in certain regions, which is an indication of the presence of an intense convective system, caused by an anti-cyclone in the center of Europe, a depression in the North Atlantic Ocean and upwelling of a warm air masses from Africa. The satellite data are images acquired by the SEVIRI radiometer imager (Spinning Enhanced Visible and Infra-Red Imager) on board the MSG4 (Meteosat Second Generation 4) satellite. These images are captured every 15 minutes in 12 spectral channels and spatial resolution of 3 km above the equator. Also, the images were processed in real time in order to correct the radiometric and geometric effects induced during the capture and geo-located, then the information is localized on the Earth, calibrated and the radiances linearized so that they are exploitable [8].

These data exploited in order to provide parameters for the identification of the characteristics of clouds and the types of precipitation of each pixel of the images captured, the parameters we will discuss in this section. Two types of clouds are the source of significant precipitation in the region of our study, stratiform and convective clouds, which can be detected by exploiting the parameters of the cloud as detailed in the study [3]. Relationship which connects them to the data provided by SEVIRI channels and summarized by the following table:

Table 1: Relationship between cloud parameters and SEVIRI data

Parameters	Channel	Relationship
Cloud Top Tempera-ture	$IR_{10.8}$	$AT = \frac{\sum_{i=1}^{N} T_{IR10.8}(i)}{N}$
Vertical Extension of the Cloud	$\Delta T_{IR10.8-IR12.0}$	$H = \frac{T_{max} - T_{min}}{6.5} - dH$
Altitude of the Cloud	$\Delta T_{WV6.2-IR10.8}$ Or $\Delta T_{WV7.3-IR12.0}$	Very negative value => Low altitude Very low value => High altitude

Cloud Water Path	During the day ΔT_(VIS0.6-NIR1.6) During the night ΔT_(IR3.9-IR10.8)	$CWP = \frac{2}{3} * \rho * \tau * r_e$
Cloud Phase	$\Delta T_{IR8.7-IR10.8}$	Low and negative value => Low altitude

With:

- N is the number of pixels that form the cloud.
- dH is the minimum height of the cloud or the rain column which is 650 m.
- ρ is the density of the water contained in the cloud.
- (τ) is the optical thickness and the effective particle radius (r_e) of the cloud.
- (r_e) is the effective particle radius of the cloud.

Radar data are images captured by the Météo-France radar network located in the south of France. These images are captured every 5 minutes with a spatial resolution of 1 km, processing has been applied to eliminate noise caused by the terrain and neighboring radars [13]. The information is the reflectivity, which is the measurement of the energy reflected by the precipitation in (dBZ) and it gives us explicit information about the water content of the cloud.

For a spatial and temporal correspondence of the data manipulated, we carried out the co-localization and the synchronization of the satellite and radar data.

Table 1: illustrates the relationship for calculating the precipitation rate R (mm/h) from the reflectivity measured by the radar Z (dBZ) [12].

Table 2: Reflectivity and precipitation rate relation

Type	Relationship	Corr. Coeff.
Strat. Precipitation Z≤ 38 dBZ	$Z=248.28\times R^{1.58}$.	0.79
Conv. Precipitation Z≥ 42 dBZ	$Z=412.35\times R^{1.49}$.	0.84

With

$$Z(dBZ) = 10\log(Z) \quad (1)$$

Application

The input variables of our models are calculated from the data of multispectral channels of the SEVIRI radiometer of the Meteosat Second Generation satellite. Which contains characteristic parameters of the clouds deduced from the SEVIRI data, which correspond to the ground truth data, which are the radar data. The model allows assigning a class to a novel sample using the samples from the training set. For this, we present as input the variables that characterize this novel samples; the class thus assigned by the model represents one of the three classes non-precipitant, stratiform precipitant and convective precipitant.

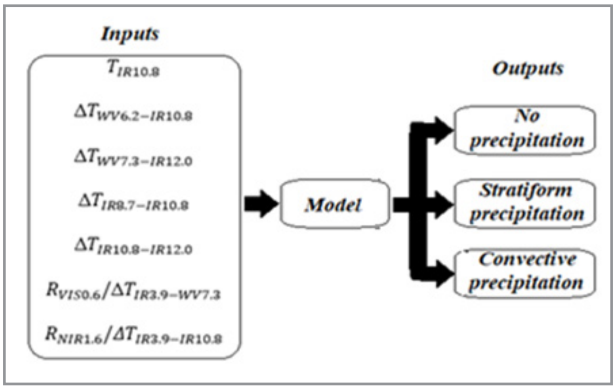

Figure 1: Structure of the Bagging and Ada-Boost model that combines spectral parameters from SEVIRI images

Results and Discussion

To evaluate the performance of the Bagging and Ada-Boost methods for cloud identification were tested to the SEVIRI and radar data samples, these datasets being taken simultaneously and respecting the spatial correspondence between them. We chose a database belonging to the stormy event of October 10, 2018. This database contains 3,187,904 samples, which correspond to eight scenes captured between 12 h UTC and 14 h UTC, which were divided in two, 80% for training and 20% for testing of the two ensemble classifiers, Bagging and Ada-Boost that we chose for our application because they showed good performance in other application areas. In order to compare the Bagging and Ada-Boost methods with other methods, the two stratiform and convective classes are in a single class called precipitant [14, 15].

The appropriate statistical parameters used to quantify the estimation errors of the methods with radar measurements calculated from the contingency table summarized in Table 4.

Table 3: Contingency table of estimates and measures

Identified by WkNN	Observed by the radar		
	Precipitating	**Non-precipitating**	**Total**
Precipitating	a	b	a+b
Non-precipitating	c	d	c+d
Total	a+c	b+d	T

Where a,b,c and d are contingency table values and T=a+b+c+d

- The rate of correctly identified events is calculated with the Probability Of Detection (POD).

$$POD = \frac{a}{a+b} \tag{2}$$

- The rate of number of pixels misidentified by the WKNN method is calculated with The Probability Of False Detection (POFD).

$$POFD = \frac{b}{b+d} \tag{3}$$

- The rate of events estimated when they were not events is calculated with the False Alarm Ratio (FAR).

$$FAR = \frac{b}{a+b} \tag{4}$$

- The difference between the estimate and the measurements is calculated with the Frequency BIAS index (Bias):

$$Bias = \frac{a+b}{a+c} \tag{5}$$

- The estimated correctly diagnosed event rate is calculated the Critical Success Index (CSI).

$$CSI = \frac{a}{a+b+c} \tag{6}$$

- The rate of correct estimates is calculated with the Percentage of Corrects (PC).

$$PC = \frac{a+d}{T} \tag{7}$$

The comparison between the Bagging and Ada-Boost methods and WkNN method the ECST technique is summarized in Table

Table 4: The statistical results of the verification for Ada-Boost and Bagging

	POD	POFD	FAR	Bias	CSI	PC
ECST	0.64	0.06	0.29	0.85	0.58	0.93
WkNN	0.68	0.05	0.26	0.89	0.61	0.95
Bagging	0.62	0.08	0.33	1.18	0.57	0.89
Ada-Boost	0.56	0.1	0.35	1.21	0.55	0.86
Optimal values	1	0	0	1	1	1

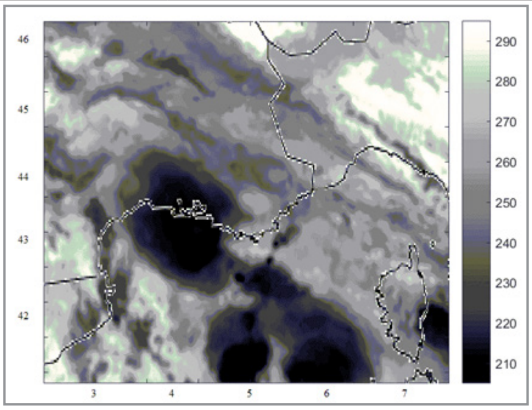

Figure 2: $T_{IR10.8}$ image scene from October 10, 2018 at 12:40 UTC

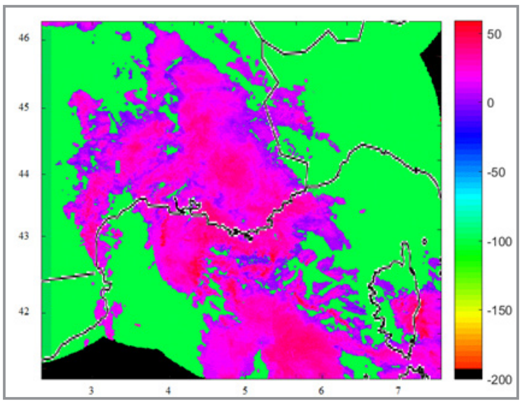

Figure 3: The rain area delineated by the Météo-France radar network in dBZ

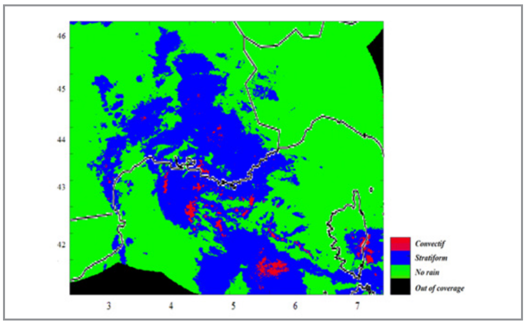

Figure 4: The rain area delineated by the Météo-France radar network in 4 levels

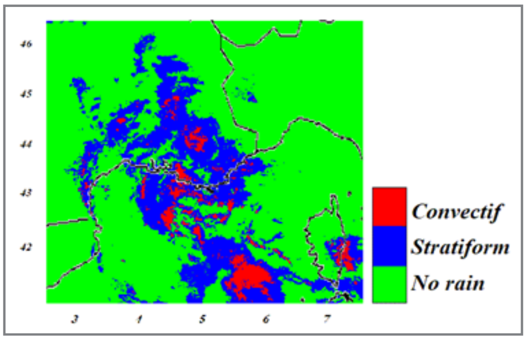

Figure 5: The rain area delineated by the Bagging method in South East of France

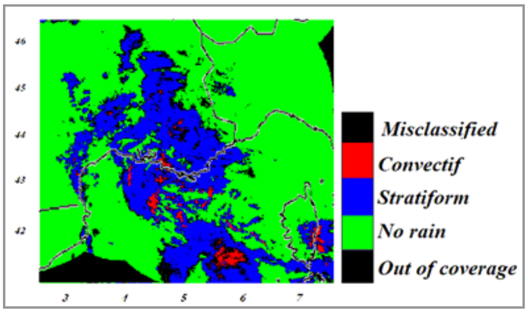

Figure 6: Comparison between Bagging method and the Météo-France radar network delineation

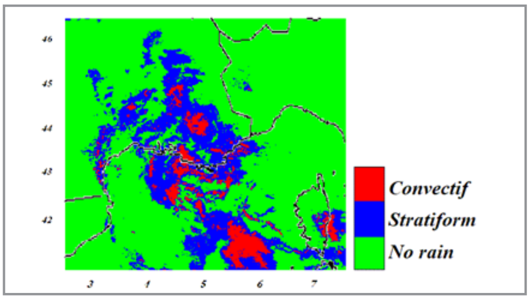

Figure 7: The rain area delineated by the Ada Boost method in South East of France

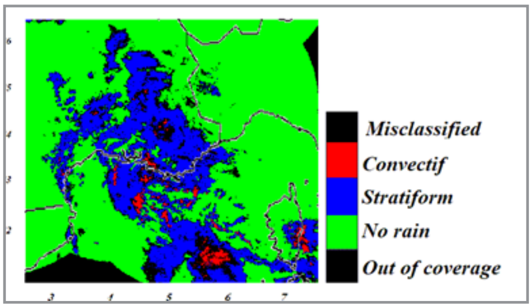

Figure 8: Comparison between Ada Boost method and the Météo-France radar network delineation

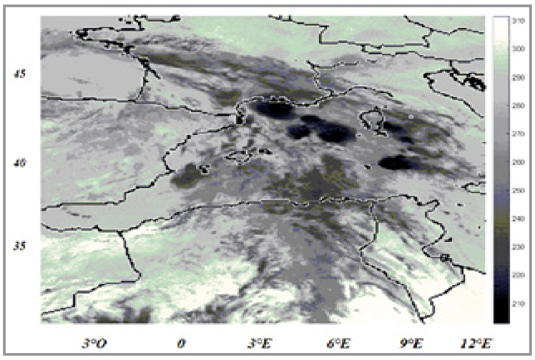

Figure 9: T_IR10.8 image scene from October 10, 2018 at 12:40 UTC over North of Algeria and Mediterranean sea

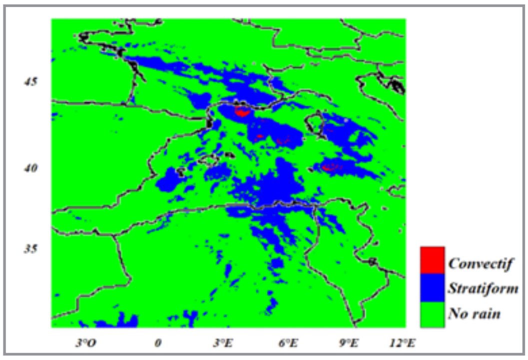

Figure 10: The rain area delineated by the Bagging method over North of Algeria and Mediterranean sea

Conclusion

In light of the results obtained, we deduce that the Bagging classifiers have better accuracy than the Ada-Boost classifier; this is valid for the four predefined classes. When the number of samples of the same class weakly represented in the database, the prediction error is greater for the Ada-Boost method. This is not the case for the Bagging method. The Bagging method is more suitable for a heterogeneous database, unlike the Ada-Boost method.

As a perspective to this study, we recommend to use a homogeneous database for an application to Ada-Boost classifiers [18].

References

1. Appriou, A. (2002). Discrimination multi-signal par la théorie de l'évidence. In Décision et reconnaissance des formes en signal (pp. 219-258). Hermès Science Publications.
2. Bensafi, N., Lazri, M., Ameur, S. (2019). Novel WkNN-based technique to improve instantaneous rainfall estimation over the north of Algeria using the multispectral MSG SEVIRI imagery. Journal of Atmospheric and Solar-Terrestrial Physics, 183, 110-119.
3. Bensafi, N., Attaf, Y., Lazri, M., Ameur, S. (2023). Delineation of raining cloud using a WkNN from multispectral data of SEVIRI radiometer. In M. R. Laouar, V. E. Balas, B. Lejdel, S. Eom, & M. A. Boudia , 12th International Conference on Information Systems and Advanced Technologies "ICISAT 2022" (Lecture Notes in Networks and Systems, Vol. 624). Springer. https://doi.org/10.1007/978-3-031-25344-7-21
4. Bloch, I. (2003). Fusion d'informations en traitement du signal et des images. Hermès Science Publications.
5. Breiman, L. (1996). Bagging predictors. Machine Learning, 24(3), 123-140.
6. Dietterich, T. G. (1997). Machine-learning research: Four current directions. AI Magazine, 18(4), 97-136.
7. Dubois, D., Prade, H. (1988). Possibility theory. Plenum Press.
8. European Organisation for the Exploitation of Meteorological Satellites. (n.d.). EUMETSAT: Monitoring the weather and climate from space. Retrieved February 10, 2026, from https://www.eumetsat.int/
9. Freund, Y., Schapire, R. E. (1996). Experiments with a new boosting algorithm. In Proceedings of the Thirteenth International Conference on Machine Learning (ICML '96) (pp. 148-156).
10. Lam, L., Suen, C. Y. (1997). Application of majority voting to pattern recognition: An analysis of its behavior and performance. IEEE Transactions on Systems, Man, and Cybernetics, 27(5), 553-568.
11. Lazri, M., Ameur, S. (2018). Combination of support vector machine, artificial neural network and random forest for improving the classification of convective and stratiform rain using spectral features of SEVIRI data. Atmospheric Research, 203, 118-129.
12. Lazri, M., Ameur, Z., Ameur, S., Mohia, Y., Brucker, J. M., Testud, J. (2013). Rainfall estimation over a Mediterranean region using a method based on various spectral parameters of SEVIRI-MSG. Advances in Space Research, 52(8), 1450-1466.
13. Ouallouche, F., Lazri, M., Ameur, S. (2018). Improvement of rainfall estimation from MSG data using random forests classification and regression. Atmospheric Research, 211, 62–72.
14. Tabary, P., Augros, C., Champeaux, J.-L., Chèze, J.-L., Faure, D., Idziorek, D., Lorandel, R., Urban, B., & Vogt, V. (2013). Le réseau et les produits radars de Météo-France. La

Météorologie, 8, 15.
15. Ting, K. M., Witten, I. H. (1997). Stacking bagged and dagged models. In Proceedings of the Fourteenth International Conference on Machine Learning (ICML 1997) (pp. 367-375).
16. Wolpert, D. H. (1992). Stacked generalization. Neural Networks, 5(2), 241-260.
17. Xu, B. V., Krzyzak, A., Suen, C. Y. (1992). Methods of combining multiple classifiers and their application to handwriting recognition. IEEE Transactions on Systems, Man, and Cybernetics, 22(3), 418-435.
18. Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1(1), 3-28.