



**HAL**  
open science

## A new index to estimate precipitation using cloud growing rate

Jean Claude Bergès, Isabelle Jobard, Rémy Roca

► **To cite this version:**

Jean Claude Bergès, Isabelle Jobard, Rémy Roca. A new index to estimate precipitation using cloud growing rate. *Geophysical Research Letters*, 2009, 36, pp.L08808. 10.1029/2008GL036665 . hal-00378664

**HAL Id: hal-00378664**

**<https://hal.science/hal-00378664v1>**

Submitted on 19 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# A new index to estimate precipitation using cloud growing rate

Jean Claude Bergès,<sup>1</sup> Isabelle Jobard,<sup>2</sup> and Rémy Roca<sup>2</sup>

Received 15 November 2008; revised 5 February 2009; accepted 18 March 2009; published 18 April 2009.

[1] A new index assessing the cloud growing rate is described in this paper. It has been designed to be integrated in rainfall estimation procedures. As the highest precipitation rates occur in growing convective cores, integrating this information should enhance the precipitation estimations. The index computation relies on image processing methods. It is composed of two steps: first a watershed segmentation is applied and then an original heritage process is performed. This second step, which is based on a simple matrix computation, is adapted to the watershed algorithm as it mitigates the over-segmentation artefact. This cloud growing rate index is compared with a cooling index which is usually computed to help in identifying raining cells and it is demonstrated to be more efficient. Moreover a simple integration of this index in a widespread rainfall estimation method leads to an improvement of the diurnal cycle retrieval. **Citation:** Bergès, J. C., I. Jobard, and R. Roca (2009), A new index to estimate precipitation using cloud growing rate, *Geophys. Res. Lett.*, 36, L08808, doi:10.1029/2008GL036665.

## 1. Introduction

[2] Estimating rainfall intensity from geostationary satellite IR data is a sensitive issue as the relationship between rainfall intensity and cloud top IR temperature is not straightforward and then unstable. As a matter of fact the thermal infrared channel provides a better estimation of precipitation probability than of rainfall rate because of the high intensity estimation issue. *Arkin and Meisner* [1987] based the GOES Precipitation Index (GPI) on this property. However the GPI resolution was defined for large time and space scales, mainly because the thermal infrared signal in a convective system is much more homogeneous than the precipitation distribution.

[3] Convection in the tropics is organized in the form of Mesoscale Convective System [*Houze*, 2004] composed of a convective core of about 10–100 km, made up of convective cells at different maturing stages of about 1–10 km and a stratiform region with no active convection but a large cold cloud shield of about 100–1000 km [*Roca et al.*, 2002]. These systems exhibit elaborated life cycle and relationship to the amount of rainfall they produce. For a convective cell the situation is less complex and the maximum of precipitation occurs during the growing phase of the convective cell [*Redelsperger*, 1997]. So integrating a growing rate index should improve the rainfall estimation accuracy. But while many algorithms have been developed for the full cloud tracking, few works have been dedicated to a proper estima-

tion of this parameter. For instance, the EUMETSAT SAF Rapid Detection of Thunderstorms algorithm [*Morel and Senesi*, 2002] provides similar index which is used to characterize the initial phase of convective system but are not used for rainfall estimation. *Machado and Laurent* [2004] focused on the cloud area expansion that reveals skill for rainfall nowcasting (L. A. T. Machado et al., Convective system area expansion, high-level wind divergence and vertical velocity: A tool for nowcasting, paper presented at the World Weather Research Program Symposium on Nowcasting and Very Short Range Forecasting, World Weather Research Programme, Toulouse, France, 5–9 September 2005). *Vicente and Scofield* [1998] use a simple cooling index defined as the difference from two consecutive images, but the efficiency of the cooling indicator appears even slightly lower than the image spatial variance which has a tenuous link with the precipitation process.

[4] The main reason for this lack of efficiency is that a cloud growing rate cannot be assessed at the pixel level but at a cluster level which should represent the convective cell scale. In this paper a new growing rate index, based on watershed segmentation, is described. The main issue has been to keep it stable regarding cluster over-segmentation. This growing rate index has been computed on a six months series of geostationary satellite data and its validity has been checked against passive microwave data.

## 2. Data

[5] Geostationary dataset, supplied by SATMOS, is composed of the whole series of GOES12 thermal infrared 10.8  $\mu\text{m}$  channel images from February to July 2006 over South America. In order to avoid missing sectors in temporal series, extractions have been performed on a geographic window 19S:11N/90W:30W. Excluding the incomplete slots, this dataset is composed of 7640 images.

[6] The reference rainfall dataset is based on the 3G68L product delivered in the TRMM/NASA database and the rainfall occurrence is derived from the passive microwave precipitation intensity [*Kummerow et al.*, 2001]. TMI rainfall estimation is based on a Bayesian procedure relating passive microwave to TRMM precipitation radar reflectivities. Over land this estimator is mainly related to 85 GHz radiation scattering by ice particles; thus in a deep convection context, estimated rainfall is based on cloud top properties and can differ from ground rainfall measurements. Hereafter it is used only for rain detection.

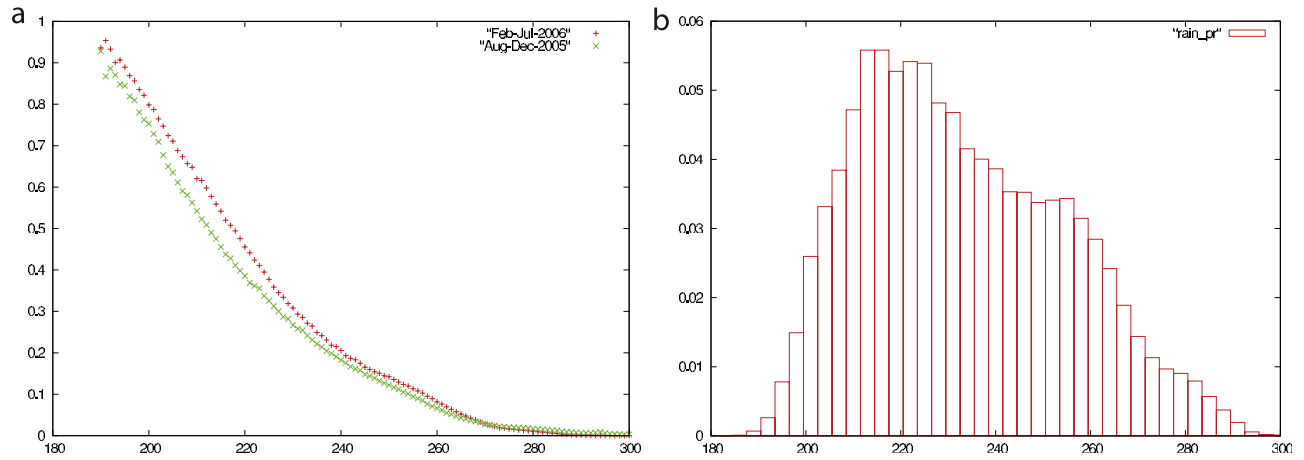
## 3. Method

### 3.1. Rainfall Probability Function and Kernel Delineation

[7] As quoted before, there is a statistical relation between probability of rainfall occurrence and IR brightness temper-

<sup>1</sup>PRODIG, Université Paris 1, CNRS, Paris, France.

<sup>2</sup>LMD, Ecole Polytechnique, IPSL, CNRS, Palaiseau, France.



**Figure 1.** Rainfall probability function estimated by histogram matching. (a) Empirical functions are computed for two different time periods. (b) Rainfall probability distribution (unitless) versus  $10.8 \mu\text{m}$  brightness temperature.

ature. This relation is statistically valid in inter-tropical area where deep convection is predominant. In this paper a rainfall probability function (RPF) is used as an empirical estimator. A RPF, computed by histogram matching, has been plotted in Figure 1a. For each bin of 1K the percentage of rainy pixels according to TMI has been assimilated to a rainfall probability. A second RPF, estimated on a different period, has been displayed on the same graph and shows the low stability of this relation. Whereas the two curves look smooth and coherent, using them directly in a quantitative estimation would produce significant errors. Figure 1b shows the distribution of detected rainfall versus IR; it appears that the two RPF differ mainly in the temperature range corresponding to the distribution peak. Nevertheless at this huge aggregation level the empirical functions appear as monotonic with quasi-null value above 280 K. Of course a RPF could be computed a more sophisticated way including multi-spectral or aerologic properties but it should remain correlated with the main statistical factor which is temperature.

[8] Given a RPF, the computation of a growing rate index relies on a rainy cluster delineation. A simple identification process is to apply a fixed temperature threshold and to look at aggregates as rainy area clusters. But the main drawback of this method is the choice of a threshold value: a too cold value would ignore small systems in expansion phase whereas a too warm value would produce large area with unpredictable connections. To avoid this effect, the kernel delineation is based on image segmentation. The watershed algorithm [Vincent and Soille, 1991] has been selected as it is widespread and fast. In an intuitive presentation of this method the RPF is interpreted as a digital elevation model. The first step is to identify local maxima, each of them will be associated with a cluster. Then by a recursive process these kernels are enlarged by all neighbouring pixels with lower or equal altitude. On the termination, when all pixels are classified, each cluster represents a runoff basin. Watershed segmentation relies on difference of altitude and not in magnitude of this effect. Therefore it is sensitive to oversegmentation mainly in quasi-flat area.

[9] Applying a median filter could suppress significant information as the size of an updraft area size significant part is of the same order as the geostationary satellite pixel size. The GATE experiment has shown a log normal

distribution of convective core diameter with more than 50% below  $10 \text{ km}^2$  [Houze and Betts, 1981]. However a slight image smoothing reduces the number of clusters and speeds up computation, while keeping the most intense precipitation area.

### 3.2. Growing Rate Index Definition

[10] Due to RPF heterogeneity, the most significant extension parameter is not the surface but the volume. Hereafter volume will not mean a physical vertical extension but the product of the RPF by the cluster characteristic function. For simple events the growing rate would be the volume ratio of the same cluster in two successive images. However the clusters produced by the watershed algorithm are contiguous and a more sophisticated heritage process has to be defined rather than a one to one correspondence. On this topic growing rate estimation is deeply different from cloud tracking where separated kernels are generated and heritage process is defined by a significant intersection surface discarding merging and splitting clusters. Properly managing these situations is a key issue for growing rate assessment as it mitigates oversegmentation effect.

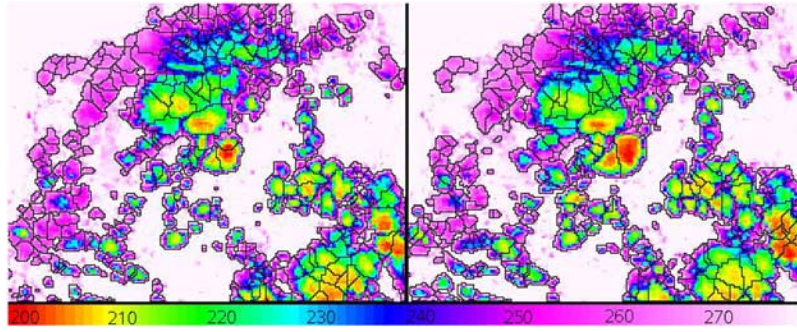
[11] The heritage process will be defined in a probabilistic way. Let  $\mathbf{S}$  be a geographic domain,  $\mathbf{A}$  a partition on this domain (i.e., a set of clusters  $\mathbf{a}$ ), a fuzzy partition  $\mathbf{fA}$  is a set of functions  $\mathbf{fa}$  on  $\mathbf{S}$  satisfying the two conditions:

$$\mathbf{fa}(\mathbf{x}) \in [0, 1] \text{ for } \mathbf{x} \in \mathbf{S} \quad (1)$$

$$\sum_{\mathbf{a} \in \mathbf{A}} \mathbf{fa}(\mathbf{x}) \in [0, 1] \text{ for } \mathbf{x} \in \mathbf{S} \quad (2)$$

Fuzzy partitions are an extension of partition characteristic functions. Obviously the product of a RPF by the characteristic functions of a partition is a fuzzy partition. The function  $\mathbf{fa}$  can be described as a partition cluster weighted by a RPF.

[12] Instead of associating to each cluster of the first image a cluster in the second one, the heritage process computes a new fuzzy partition on the first image. This new fuzzy partition will be indexed by the same cluster set as the original fuzzy partition. Let  $\mathbf{A}$  and  $\mathbf{B}$  the partitions generated



**Figure 2.** 10.8  $\mu\text{m}$  brightness temperature on the window 72.5W-62.5W/3.5N-4S for July 17th 2006, (left) at 17:30 and (right) at 18:00. Clusters boundaries, obtained by watershed segmentation, are overlaid. Color scale is coded in Kelvin.

on two successive images and  $\mathbf{fA}$  and  $\mathbf{fB}$  the fuzzy partitions computed from their respective RPF as described above, then  $\mathbf{hA}$  the heir of  $\mathbf{fA}$  is a set of functions  $\mathbf{ha}$  defined by:

$$\mathbf{ha} = \sum_{\mathbf{b} \in \mathbf{B}} \lambda_{\mathbf{ab}} \mathbf{fb} \quad (3)$$

$$\lambda_{\mathbf{ab}} = \mathbf{sab} / \sum_{\mathbf{a}' \in \mathbf{A}} \mathbf{sa}'\mathbf{b} \quad (4)$$

$$\mathbf{sab} = \int_{\mathbf{x} \in \mathbf{S}} \mathbf{fa}(\mathbf{x}) \cdot \mathbf{fb}(\mathbf{x}) \, \mathbf{d}\mathbf{x} \quad (5)$$

Should the denominator in equation (4) be zero the coefficient  $\lambda_{\mathbf{ab}}$  would be zeroed.  $\mathbf{hA}$  satisfies the conditions (1) and (2) and is therefore a fuzzy partition.  $\mathbf{sab}$  can be interpreted as the intersecting volume of clusters  $\mathbf{a}$  and  $\mathbf{b}$  and  $\lambda_{\mathbf{ab}}$  as the relative importance of cluster  $\mathbf{a}$  compared with other clusters of  $\mathbf{A}$  intersecting  $\mathbf{b}$ . The fuzzy cluster function  $\mathbf{ha}$  is a mapping of  $\mathbf{fa}$  (a fuzzy cluster function on the first image) among the fuzzy clusters of the second image. This mapping is dependent of the RPF and not only of the intersecting area. In some simple case,  $\mathbf{ha}$  computation is rather straightforward. When a cluster  $\mathbf{a}$  splits in two clusters  $\mathbf{b}$  and  $\mathbf{b}'$  without any interference with other clusters, then  $\mathbf{ha}$  is the sum of  $\mathbf{fb}$  and  $\mathbf{fb}'$ . In the opposite situation where two clusters  $\mathbf{a}$  and  $\mathbf{a}'$  merge in one  $\mathbf{b}$ ,  $\mathbf{ha}$  and  $\mathbf{ha}'$  constitute a linear decomposition of  $\mathbf{fb}$ . The cluster  $\mathbf{fb}$  is split between  $\mathbf{ha}$  and  $\mathbf{ha}'$  proportionally to the intersecting volumes  $\mathbf{a} \cap \mathbf{b}$  and  $\mathbf{a}' \cap \mathbf{b}$ .

[13] The growing rate index for cluster  $\mathbf{a}$  will be simply defined as the volume ratio between  $\mathbf{fa}$  and  $\mathbf{ha}$ :

$$\int_{\mathbf{x} \in \mathbf{S}} \mathbf{ha}(\mathbf{x}) \, \mathbf{d}\mathbf{x} / \int_{\mathbf{x} \in \mathbf{S}} \mathbf{fa}(\mathbf{x}) \, \mathbf{d}\mathbf{x} \quad (6)$$

This index is based on simple matrix operations. First a similarity matrix is computed where the coefficient on line  $\mathbf{a}$  and column  $\mathbf{b}$  is the intersecting volume of clusters  $\mathbf{a}$  (first partition) and  $\mathbf{b}$  (second partition) and this matrix is normalized so that the sum of each line is 1 or 0. Then, the vector of inherited cluster volumes is computed by a matrix product between the normalized similarity matrix and the vector of cluster volume on the second fuzzy partition,  $\mathbf{fB}$ .

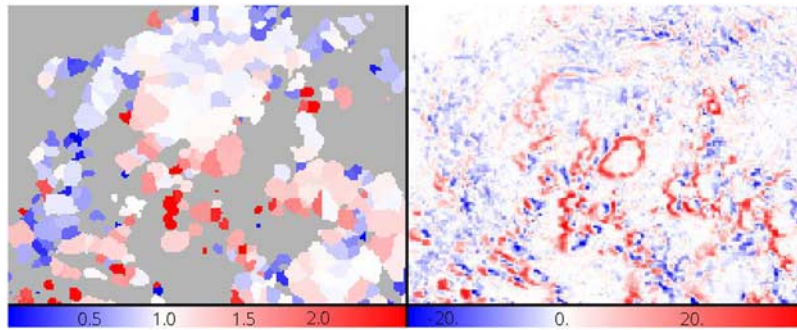
[14] An important property is that the growing rate index is not sensitive to any oversegmentation on  $\mathbf{B}$ . Splitting in two, by a proportional factor, a column of the similarity matrix would not modify the volume vector of inherited fuzzy partition.

#### 4. Applications

[15] The actual computations rely on a temperature linear RPF ranging from zero at 273 K to one at 183 K. A median filter on a  $3 \times 3$  window has been applied on temperature images to lower memory requirements. In this implementation both watershed segmentation and matrix operations run quickly and the resulting index is well suited for integration in an operational procedure.

[16] Figures 2 (left) and 2 (right) present two successive RPF images overlaid with watershed segmentation lines. The local maxima of the RPF are clearly separated but the segmentation algorithm sensitivity to slight temperature variations can be noticed. The resulting growing rate index is displayed in Figure 3 (left) and the cooling rate defined as the pixel temperature difference is shown for comparison in Figure 3 (right). The two indices look clearly different. For example in the center of the image, whereas the cooling index produces the highest values on the periphery of the convective cells, the growing rate index identifies properly the whole growing area. Moreover on the growing rate index figure, two merging clusters get close values and, even in such a complex convective system organization, expansion and decay areas appear clearly. It can be noticed also that the cooling index maxima would correspond to the local variance maxima area on cloud edge and this explains the results obtained by Vicente who found similar performances for these two indices.

[17] Assessing impact of the growing rate index for rainfall estimation is difficult because it depends both on the estimation procedure design and on the retrieval scale. But it appears as a valuable second order factor for the cloud top temperature index. Extensive computation has been carried out on the whole dataset for brightness temperature below 273 K and results are displayed in Figure 4. In Figure 4a the mean value of the logarithm of the growing rate index has been plotted versus the corresponding 10.8  $\mu\text{m}$  temperature in red (resp. in green) for all pixels classified by TMI as rainy (resp. non rainy). The error bars have been plotted on the same graph. Due to the sample size the errors are small except for the non-rainy cold pixels. Comparison



**Figure 3.** The cluster evolution computed from the figure 2 images by two different methods: (left) the growing rate index (dimensionless) and (right) the cooling index (Kelvin difference).

of these two curves shows a discrimination efficiency quickly increasing with temperature above 223 K. At the analysis scale, there is no significant differences for low temperature clouds and, at a first glance, the index does not discriminate the high altitude non precipitating clouds. The index looks as more efficient to detect young convective cores producing rainfall despite a warm top temperature. This results have been obtained without a prior filtering by cloud classification and all the precipitation types encountered in the study area contribute. Should we focus on an identified convective system the results could be better as suggested by Figures 2 and 3 (left) where all the frontal expansion areas are associated with high index values.

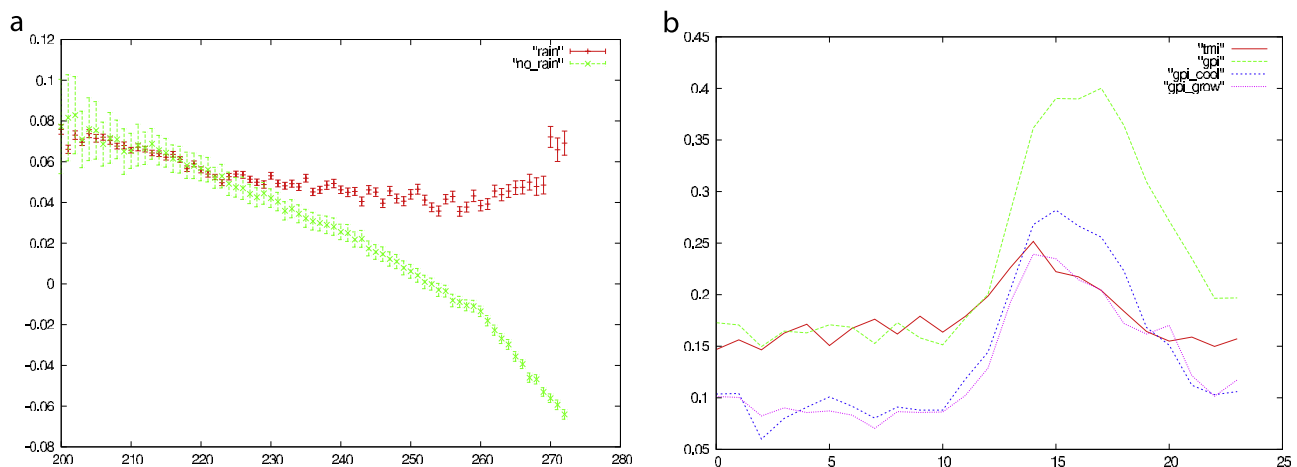
[18] Assessment of diurnal cycle retrieval by various indicators is presented in Figure 4b. GPI is a global estimator with some efficiency at the roughest scales, but even at these scales it fails to properly render the diurnal precipitation cycle. The mean TMI rainfall probability has been plotted (red curve) versus local time. On the same graph the green curve represents a GPI index (proportion of pixels with temperature colder than 235 K). A lag appears between these two curves. Whereas the peak of TMI probability is, as expected, around 14 h the maximum of GPI index is around 17 h. Two GPI-enhanced indices have been computed. For the “GPI-grow” index (magenta curve) a pixel is rainy if its temperature is colder than GPI

threshold and if its growing index is greater than 1. The “GPI-cool” index (blue curve) selects all the pixels colder than the threshold and with a positive cooling rate. Comparing the peak times, the “GPI-cool” is closer to the TMI than the GPI but the “GPI-grow” is still better. This result is not highly dependent on the GPI threshold and other values would show very similar patterns.

## 5. Conclusions

[19] The growing rate index appears as consistent and fairly related to microwave rainfall estimator. The primary motivation of this work was to enhance EPSAT-SG, the rainfall estimation procedure designed for the AMMA (African Monsoon Multiscale Analysis) experiment; now it is planned to integrate the new index in the next version of the rainfall algorithm. But the scope of this index goes far beyond, it is not specific to this algorithm. Only a thermal infrared channel is needed to compute this index and it can be used in backprocessing series of geostationary satellite data.

[20] Its capacity to enhance the diurnal cycle retrieval makes this index a valuable parameter for rainfall estimation methods using passive microwave data since most of microwave sensors are onboard sun synchronous satellites and any information at the diurnal cycle scale can bring a sensible improvement.



**Figure 4.** Growing rate index application to rainfall estimation. (a) Mean values of index logarithm with error bars versus  $10.8 \mu\text{m}$  brightness temperature. (b) Diurnal cycle retrieval by various indicators (mean rainfall probability versus local time).

[21] **Acknowledgments.** SATMOS is a satellite data service funded by IPSL/CNRS and hosted by Météo-France in Lannion (France).

## References

- Arkin, P. A., and B. N. Meisner (1987), The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982–84, *Mon. Weather Rev.*, *115*, 51–74.
- Houze, R. A., Jr. (2004), Mesoscale convective systems, *Rev. Geophys.*, *42*, RG4003, doi:10.1029/2004RG000150.
- Houze, R. A., Jr., and A. K. Betts (1981), Convection in GATE, *Rev. Geophys. Space Phys.*, *19*, 541–576.
- Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, D. B. Shin, and T. T. Wilheit (2001), The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors, *J. Appl. Meteorol.*, *40*, 1801–1840.
- Machado, L. A. T., and H. Laurent (2004), The convective system area expansion over Amazonia and its relationships with convective system life duration and high-level wind divergence, *Mon. Weather Rev.*, *132*, 714–725.
- Morel, C., and S. Senesi (2002), A climatology of mesoscale convective systems over Europe using satellite infrared imagery. Part I: Methodology, *Q. J. R. Meteorol. Soc.*, *128*, 1953–1971.
- Redelsperger, J. L. (1997), The mesoscale organization of deep convection, in *The Physics and Parameterization of Moist Atmospheric Convection*, edited by R. K. Smith, pp. 159–160, Springer, London.
- Roca, R., M. Viollier, L. Picon, and M. Desbois (2002), A multisatellite analysis of deep convection and its moist environment over the Indian Ocean during the winter monsoon, *J. Geophys. Res.*, *107*(D19), 8012, doi:10.1029/2000JD000040.
- Vicente, G., and R. A. Scofield (1998), The operational GOES infrared rainfall estimation technique, *Bull. Am. Meteorol. Soc.*, *79*, 1883–1898.
- Vincent, L., and P. Soille (1991), Watersheds in digital spaces: An efficient algorithm based on immersion simulations, *IEEE Trans. Pattern Anal. Mach. Intel.*, *13*, 583–598.

---

J. C. Bergès, PRODIG, Université Paris 1, CNRS, 5 rue valette, F-75005 Paris CEDEX, France. (zebulon@univ-paris1.fr)

I. Jobard and R. Roca, LMD, Ecole Polytechnique, IPSL, CNRS, F-91128 Palaiseau CEDEX, France.