Basis for a Rainfall Estimation Technique Using IR–VIS Cloud Classification and Parameters over the Life Cycle of Mesoscale Convective Systems

G. DELGADO
Astronomy and Meteorology Department, University of Barcelona, Barcelona, Spain

LUÍZ A. T. MACHADO, CARLOS F. ANGELIS, AND MARCUS J. BOTTINO
Instituto Nacional de Pesquisas Espaciais, Centro de Previsão de Tempo e Estudos Climáticos, Cachoeira Paulista-São Paulo, Brazil

Á. REDAÑO AND J. LORENTE
Astronomy and Meteorology Department, University of Barcelona, Barcelona, Spain

L. GIMENO AND R. NIETO
Universidad de Vigo, Facultad de Ciencias de Ourense, Ourense, Spain, and University of Lisbon, CGUL, IDL, Lisbon, Portugal

(Manuscript received 5 January 2007, in final form 2 July 2007)

ABSTRACT

This paper discusses the basis for a new rainfall estimation method using geostationary infrared and visible data. The precipitation radar on board the Tropical Rainfall Measuring Mission satellite is used to train the algorithm presented (which is the basis of the estimation method) and the further intercomparison. The algorithm uses daily Geostationary Operational Environmental Satellite infrared–visible (IR–VIS) cloud classifications together with radiative and evolution properties of clouds over the life cycle of mesoscale convective systems (MCSs) in different brightness temperature ($T_b$) ranges. Despite recognition of the importance of the relationship between the life cycle of MCSs and the rainfall rate they produce, this relationship has not previously been quantified precisely. An empirical relationship is found between the characteristics that describe the MCSs' life cycle and the magnitude of rainfall rate they produce. Numerous earlier studies focus on this subject using cloud-patch or pixel-based techniques; this work combines the two techniques. The algorithm performs reasonably well in the case of convective systems and also for stratiform clouds, although it tends to overestimate rainfall rates. Despite only using satellite information to initialize the algorithm, satisfactory results were obtained relative to the hydroestimator technique, which in addition to the IR information uses extra satellite data such as moisture and orographic corrections. This shows that the use of IR–VIS cloud classification and MCS properties provides a robust basis for creating a future estimation method incorporating humidity Eta field outputs for a moisture correction, digital elevation models combined with low-level moisture advection for an orographic correction, and a nighttime cloud classification.

1. Introduction

Several studies have focused on the rainfall estimation using infrared–visible (IR–VIS) imagery highlighting its importance due to the low sampling of the ground-based radars and the sparse distribution of the rain gauges. Estimations by satellite provide high spatial and temporal sampling frequencies, but the information is inferred through indirect methods, which result in significant errors in the final rainfall estimation. Previous works discuss the use of satellite rainfall-rate estimation in the hydrological, meteorological, and climatological sciences: flood forecasting, verification of rainfall forecast from numerical weather prediction (NWP) models, moisture budget studies, and the evalu-

Corresponding author address: Germán Delgado, Dept. D’Astronomia i Meteorologia, Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Spain.
E-mail: isohipsa@yahoo.es

DOI: 10.1175/2007JAMC1684.1
© 2008 American Meteorological Society
Estimating rainfall rate by satellite using IR imagery has been a topic of study since the late 1970s. Griffith et al. (1978) present the Griffith–Woodley technique based on the study of the life cycle of clouds. The hydroestimator (HE) technique (Vicente et al. 1998, 2002) is based on the methodology developed by Scofield (1987). To compute rainfall rates, it uses a non-linear power-law relationship between cloud-top temperature and radar-derived rainfall-rate estimates, a gradient and growth-rate mask, and a humidity mask. The HE was developed to estimate rainfall rates in intense convective systems. Arkin and Meisner (1987) developed the Geostationary Operational Environmental Satellite (GOES) precipitation index (GPI) method, which explores statistical relationships between cloud-top temperatures and rainfall. The technique assigns a mean rainfall rate to all the pixels below a threshold temperature. The GPI is better suited to estimating monthly cumulative rainfall than shorter time ranges. The GOES Multispectral Rainfall Algorithm (GMSRA), described by Ba and Gruber (2001), uses five channels of the GOES satellite and incorporates cloud-top particle information by taking advantage of the spectral resolution. Like the HE, GMSRA uses additional moisture correction to account for evaporation of rainfall that falls from the clouds but does not reach the surface (Kidd et al., 2003). Both HE and GMSRA estimates improve the GPI, with better correlations, bias, and RMS.

Joyce et al. (2004) present the Climate Prediction Center Morphing Method (CMORPH); the method uses motion vectors derived from geostationary satellite IR imagery to propagate the precipitation derived from passive microwave data. The shape and intensity of the precipitation’s features are modified (morphed) during the time between microwave sensor scans by time-weighted linear interpolation. Laing et al. (1999) present a relationship between SSM/I-Derived precipitation characteristics and IR data to estimate the precipitation produced by African mesoscale convective complexes (MCCs). Precipitation is diagnosed using the IR-observed cold cloud area as a function of time over the MCC’s life cycle. Xu et al. (1999) present a method for identifying and removing no-rain cold clouds from IR imagery. Seven cloud-patch features are used to describe cloud-top properties and produce classification rules. Last, some rainfall estimation methods using neural networks are summarized. Hong et al. (2004) describe the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIAN) Cloud Classification System (CCS). Local and regional cloud features are extracted to calibrate the cloud-top temperature and rainfall relationships for the classified cloud groups. The results show good correlation coefficients for a 0.25° grid. Tapiador et al. (2004) present an evaluation of operational procedures using passive microwave and infrared satellite measurements. Several neural networks are tested, from multilayer perceptron to adaptive resonance theory architectures. Bellerby et al. (2000) describe the development of an algorithm using the precipitation radar (PR) on board the Tropical Rainfall Measuring Mission (TRMM) satellite and multispectral GOES imagery. Coincident PR measurements were matched with four-band GOES image data to form the training dataset for a neural network.

This work presents a daily rainfall estimation algorithm using satellite imagery (RESAT) for the area of South America. RESAT uses only satellite data, without including moisture or orographic corrections. The satisfactory results obtained when compared with HE—which does use moisture and orographic corrections—during the intercomparison period demonstrate that RESAT is a valid basis for a new estimation method that will include moisture and orographic corrections. The South American region, which includes the Amazonian forest, covers a tropical and subtropical area where the formation of mesoscale convective systems (MCSs) with the potential to produce heavy rainfall is common. This makes RESAT a valuable tool for monitoring rainfall originating in both convective and stratiform cloud types in the zone.

All the works mentioned above can be classified as pixel-based or cloud-patch-based methods and all of them have a bearing on this paper. RESAT is based on both pixel and cloud-patch properties. Griffith et al. (1978), Scofield (1987), Feidas and Cartalis (2001), Maton and Laurent (2001), and Schumacher and Johnson (2005) point out the importance of the relationship between the life cycle of the MCS and the rainfall it produces. In spite of this, in some cases only qualitative relationships are given and in others cases some problems arise due to the limitations of cloud-patch-based methods. These limitations include the difficulty in discriminating between cirrus anvils and convective clouds in the Griffith–Woodley technique, and the requirement for the direct interaction of a meteorologist in the case of Scofield (1987). In this work, a cluster rainfall estimation is first computed using radiative and evolu-
tion parameters over the life cycle of the MCS and then a pixel-based estimation is derived by adding a correction using cloud-top pixel radiative properties. Uddstrom and Gray (1996) obtain accuracy upward of 60% in delineating rain and no-rain samples of a cloud classification that shows the dependence of the final rainfall estimation on the cloud type. Other authors emphasize the importance of the relation between the microstructure of clouds and precipitation potential using mid-IR thermal channels (Rosenfeld and Gutman 1994; Lensky and Rosenfeld 2003). Finally, Levizzani et al. (2001) analyze the potential of VIS–IR algorithms. These results motivated us to use an IR–VIS cloud classification in order to assign different coefficients to each cloud type in the final rainfall-rate estimation and to discriminate between rain and no-rain states.

2. Data and satellite products

All the data used in this work was provided by the satellite division (Divisão de Satélites e Sistemas Ambientais; DSA) of the Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), which belongs to the Instituto Nacional de Pesquisas Espaciais (INPE) located in the state of São Paulo, Brazil. The satellite division receives data from many satellites, such as GOES, Meteosat, Terra, Aqua, and the National Oceanic and Atmospheric Administration (NOAA). The following satellite products used in this paper are produced in the DSA in real time: an IR–VIS cloud classification; an improved version of the National Environmental Satellite, Data, and Information Service (NESDIS) HE (Vicente et al. 1998), the DSA HE; and the operational product Forecast and Tracking of Active Convective Cells (FORTRACC) that tracks MCSs and displays information on the expansion, speed, and other parameters of the systems (Machado et al. 1998; Machado and Laurent 2004). Both FORTRACC and the IR–VIS cloud classification algorithm are run in parallel with RESAT to produce the arrays containing the inputs to initialize the algorithm.

Two months of TRMM PR data collected over South America during November and December of 2004 were also provided by CPTEC’s satellite division. Kummerow et al. (1998) and Bellerby et al. (2000) emphasize the quality of the PR as a source of training data, so TRMM PR rainfall values were used for both the training and the intercomparison periods. The PR scans a 215-km swath with vertical and horizontal resolutions of 250 m and $4.3 \times 4.3$ km$^2$, respectively, at nadir, while the minimum detectable signal is approximately 17 dBZ. We also extracted the cloud-top temperature from the IR GOES channel. The upper side of the flowchart in Fig. 1 summarizes the dataset used to initialize the algorithm. All the input data including the IR GOES images; the TRMM PR product, the IR–VIS cloud classification, the DSA HE maps, and the arrays containing the values of the parameters of each MCS, were reprojected to a Mercator projection with a spatial resolution of 0.038° of longitude and 0.036° of latitude.

The South America region is the domain chosen for this study (Fig. 2). All the cases are located inside a window centered over Brazil (35°S–5°N, 35°–75°W).

To select the cases of study, several restrictions were imposed within two months of data. Because of the use of the VIS GOES channel to obtain the cloud classification, only daily data were chosen. There is a ±15 min time difference and a maximum navigation error of 10 km between the TRMM pass and the IR GOES scanning. The IR data closest to the time of the TRMM observation were chosen to obtain a final time difference of less than ±5 min and a navigation error of less than 1 pixel. Finally, the availability of GOES images covering the Southern Hemisphere is not guaranteed by NOAA, so when images are missed, no products are generated. This imposed a limitation on the data that we were able to use when selecting the cases. In total, 27 cases were selected for the study; 18 during the training period and 9 for the later intercomparison. During the intercomparison period we used different rainy systems from those used in the training period, in order to guarantee the independence of the data during the evaluation. RESAT and the DSA HE were compared using TRMM data. The HE has been rigorously validated using rain gauges and ground-based radar (Vicente et al. 1998). Comparison with HE is therefore a valid technique for assessing the performance of RESAT, although further validations using ground-based radar information will be needed in the future to test the estimation method (incorporating the necessary corrections) based on RESAT. Table 1 summarizes the date and hour of all the cases selected. In sections 2a–c we briefly explain the different satellite products used to initialize RESAT.

a. IR–VIS cloud classification

Clouds were classified using a cluster analysis with IR and VIS GOES images; hence, the classification is only operative during the daytime. The basis of the algorithm has been described in several works (Delgado et al. 2007; Porcú and Levizzani 1992; Rossow and Gardner 1993), so only a brief description is provided here. Four parameters are used to create a multispectral histogram: the IR brightness temperature ($T_b$), the visible reflectivity, and two texture indices computed as the
FIG. 1. Flow diagram showing the three modules of the rainfall estimation algorithm ($T_b$-range-tracking algorithm, rain/no-rain discrimination, and rainfall-rate estimation) and the dataset initializing the model.
standard deviation between the first neighbors of each pixel in both channels. Each point in the histogram corresponds to the number of pixels in the image with the same four parameter values. After analyzing the cluster distribution in the multispectral histogram over several months, 30 seed points are chosen. Every seed point has four components, corresponding to each of the four parameters described below and each seed point will flag a class in the histogram. For the classification algorithm to be calibrated properly (i.e., so that a physical correspondence between the values of the parameters defining each seed point and a cloud or surface type may be established) these parameters require a different value depending on the time of the day. Three daily segments were used: morning, midday, and afternoon. The algorithm associates each point in the multispectral histogram with the closest seed point, minimizing the Euclidean distance. The main classes were surface (3 types), cumulus (10 types), cirrus (6 types), stratus (5 types), and multilayer (6 types).

b. FORTRACC

The high temporal resolution of GOES allows the development of operational methods in order to monitor meteorological phenomena, such as the expansion and shifting velocity of MCSs. The expansion, mean temperature difference, and other parameters of MCSs used in this paper and defined below were obtained using FORTRACC. The technique is based on an algorithm that allows tracking of the MCS, computing of the radiative and morphological properties, and forecasting of the evolution of these physical properties (based on cloud-top brightness temperature) up to 120 min, using IR satellite imagery. The algorithm consists of four steps:

1) Cloud clusters are detected on the basis of a size and temperature threshold.
2) Morphological and radiative parameters of each MCS are determined using a statistical module.
3) MCSs are tracked, on the basis of MCS-overlapping areas between two successive images.
4) Forecasting is based on the evolution of MCS in previous time steps.

The fourth step is not used in this work. The algorithm is based on the work of Machado et al. (1998).
The author analyses the life cycle of convective systems and suggests that the surface expansion of a convective system could be associated with the high level of wind divergence and the length of its life cycle.

c. DSA HE

The results of the rainfall-rate estimation algorithm presented in this work were compared during the inter-comparison period with the estimations of the improved version of the NESDIS HE (Vicente et al. 1998), hereafter known as DSA HE, which assimilates GOES (10.5 μm) and Meteosat (11.5 μm) data, wind, and humidity data produced by CPTEC’s numerical weather forecast model (Eta model), and topography information (including an orographic correction). The NESDIS HE mathematical function, which assesses information (including an orographic correction). The NESDIS HE mathematical function, which assesses information assimilated by the DSA HE. These improvements to the NESDIS HE were made in order to estimate rainfall over South America more accurately, mainly in its tropical part.

3. Description of the algorithm

Most of the works mentioned in the introduction develop the estimation technique through general assumptions. We make similar assumptions, taking into account the importance of the development stage of the convective system and the cloud classification. Scofield (1987) summarizes some of these assumptions as having particular importance:

1) clouds with cold tops in the IR imagery and high reflectivity in the VIS imagery produce more rainfall than those with warmer tops and
2) low reflectivity and decaying clouds produce little or no rainfall, whereas clouds in the forming stage tend to produce heavy rainfall.

Since these are general assumptions, we do not have to expect colder pixels to produce more rain than warmer pixels in all the cases. A warmer cloud system can produce more rainfall than a cooler one if the warmer system is in a growing phase.

We do not work directly with the VIS channel, although the information of this channel is contained in the cloud classification. Thus, those clouds with a high reflectivity top (convective in the cloud classification) are treated differently in the algorithm. In fact, every cloud type has different parameterization in the algorithm. We also pay attention to variability in the MCSs internal brightness temperature, as well as in the value of \( T_b \). Concerning the second general approach mentioned above, different parameters of the evolution of each convective system were computed using FORTRACC. These parameters (such as the expansion or the mean temperature difference between two consecutive images of the same system) describe the stage of each convective system along its life cycle and its dynamic conditions. As we explain in section 3c, those convective systems in a growing phase (some of them in an exploding phase during the initial stage) are associated with a high rainfall rate.

Other works (Wylie 1979; Hong et al. 2004) focus on the importance of working with different \( T_b \) thresholds in order to detect correctly the convective cells embedded in the MCSs. Convective cells are the coldest pixels in a MCSs, and although they are smaller, they produce heavier rainfall. As we explain in the next section, five different \( T_b \) thresholds are considered in order to allocate the proper dynamic parameters to the convective cells embedded in the detected convective systems. This hypothesis helps to reduce the errors in the final rainfall-rate estimation.

Many authors (Mapes and Houze 1992; Machado and Rosow 1993; Machado et al. 1998) consider that \( T_b = 245 \text{ K} \) is a satisfactory threshold for identifying MCSs. The near-linear relationship between the convective system area and its threshold indicates the insensitivity of the choice of a specific value within a range of 10–20 K (Carvalho and Jones 2001). In this work, we consider a \( T_b \) threshold of 250 K to select MCSs.

The estimation algorithm is organized in three basic modules: the \( T_b \)-range-tracking algorithm, the rain/no-rain discrimination, and the rainfall-rate estimation. The first one divides every MCS into the appropriate \( T_b \) ranges and creates the arrays containing the internal variability (\( T_{\text{var}} \)) and the MCS parameters. The flowchart in Fig. 1 shows the organization of the algorithm and how all the input data is set into the different modules.

a. Threshold divisions

RESAT is based on the correlation between the radiative and evolution parameters of the convective systems and its mean rainfall rate. Therefore, only pixels colder than 250 K are considered for the estimation. A deep convective cloud system is identified by adjacent satellite image pixels with an infrared brightness temperature colder than a given threshold. To separate effectively the convective cells contained in the MCS, the following structure and evolution parameters are computed for IR cloud clusters with \( T_b \) colder than five different thresholds (250, 240, 230, 220, and 210 K): mean and minimum temperature (\( T_m \) and \( T_{\text{min}} \)), mean temperature difference between two consecutive images (\( \Delta T_m \)), minimum temperature difference between
two consecutive images ($\Delta T_{\text{min}}$), and the expansion computed as the normalized difference between areas divided by time (30 min) in two consecutive images, defined as $\Delta E = 1/A \times (dA/dt)$ (see Machado et al. 1998; Machado and Laurent 2004), where $A$ is the area of a given MCS at a given time. The computation of MCS parameters is independent of the IR–VIS cloud classification; that is to say, there is no database with the values of the parameters for every cloud type. Each $T_b$ threshold will define a $T_b$ range in each MCS (see Fig. 3), the most external being the one defined as all the pixels colder than 250 K (first $T_b$ range) and the most internal being the one defined as all the pixels colder than 210 K (fifth $T_b$ range). Both rain/no-rain discrimination and rainfall-rate estimation were computed for each of the cloud $T_b$ ranges obtained by applying the $T_b$-range-tracking algorithm, summarized in the flowchart of Fig. 3.

Fig. 3. Flowchart of the $T_b$ range-tracking algorithm.
The $T_b$-range-tracking algorithm first assigns the value of the estimated rainfall computed for the first $T_b$ range of each MCS. If the MCS is divided into more $T_b$ ranges, the algorithm superposes the estimated values of rainfall rate to the pixels in the second $T_b$ range. The process continues for the rest of the $T_b$ ranges down to the last one.

The difference between $T_b$ of each pixel and $T_m$ of the cloud cluster where the pixel is located is also used. The difference is denoted by $T_v$. This variable indicates the internal variability of the MCS with respect to $T_b$. Pixels with negatives values of $T_v$ are located in the coldest region of the $T_b$ range, with a higher probability of having rainfall. The definition of $T_v$ helps us to discriminate better between two pixels with the same $T_b$ but different values of the rainfall rate.

b. Rain/no-rain discrimination

Nearly 97% of the rain pixels of the selected cases during the analyzed period fall into six cloud types: four convective, one cold stratiform, and one cumulus (a single cumuliform vertically developed cloud in its first stage of development). One of the four convective types was flagged as a deep convective cloud (DCC) and the other convective ones can be interpreted as the coldest part of the anvil of the DCC in different temperature ranges. We refer to these three as convective (C1, C2, and C3) to differentiate them from DCC. Pixels classified as C1 are the colder ones, being located close to the cold core (DCC) of the MCS; their reflectivity in the VIS channel is also higher. Pixels classified as C2 are also colder than C3, but with similar values of VIS reflectivity and considerable differences in the texture indexes, C3 being the roughest and C2 the smoothest.

The percentage of pixels of each class for every $T_b$ range is computed in each MCS. Each $T_b$ range is associated with the class of which it contains most of the pixels. Only MCS $T_b$ ranges associated with the rainy cloud types mentioned above are classified as rain clusters. Only pixels belonging to the six cloud types mentioned above, in each $T_b$ range that is classified as rainy, are flagged as possible rain pixels. The final rain/no-rain discrimination depends on the cloud type:

1) Cumulus pixels embedded in $T_b$ ranges associated with convective, DCC, and cold stratiform cloud types are classified as rain pixels. Those embedded in $T_b$ ranges associated with the cirrus cloud type are not classified as rain pixels.

2) Cold stratiform pixels embedded in $T_b$ ranges associated with DCC and cold or warm stratiform cloud types are classified as rain pixels.

3) Convective pixels embedded in $T_b$ ranges associated with convective, DCC, and cold stratiform cloud types and with $T_v < 0$ are classified as rain pixels.

4) DCC pixels embedded in $T_b$ ranges associated with DCC and convective cloud types and with $T_v < 0$ are classified as rain pixels.

c. Rainfall-rate estimation

The rainfall-rate estimation of the pixels classified as rainy pixels consists of three steps.

1) In the first step, the mean rainfall rates from the TRMM PR data are computed for each MCS for each $T_b$ threshold. This value is compared with the parameters that describe the stage of evolution of the MCS for the same $T_b$ threshold. A multiple linear regression is computed between these variables and the mean rainfall rate. A rain value (the cluster rainfall estimation, $R_c$) is associated with each $T_b$ range of each MCS by superposing the value of $R_c$ for each $T_b$ threshold.

2) In the second step, $T_v$ is used to assign a more accurate value of rainfall rate to each pixel, the pixel rainfall estimation ($R_p$). In this step, a pixel correction (pixel rainfall correction, $r_p$) is added to $R_c$ using a nonlinear relation between $T_v$ and the PR data for each pixel of the MCS in a given $T_b$ range.

3) In the third step, the frequency distribution computed over all the cases of the training period of the PR data and $R_p$ is approximated to an exponential distribution. The $R_p$ cumulated distribution is fitted to the PR data distribution in order to increment its range and obtain the final value of the estimated rainfall rate $R$.

1) Cluster rainfall estimation

This first step assesses the development stage of every MCS in every $T_b$ range. Systems in a growing phase will be assigned a higher value of $R_c$. As commented in section 2, a warmer MCS in a growing phase can have more rain associated with it than a colder more mature MCS.

A multiple linear regression is computed between $\Delta E$, $T_m$, $\Delta T_m$, $T_{min}$, $\Delta T_{min}$, and the mean TRMM PR rainfall rate for each $T_b$ threshold in each MCS. Thus, each pixel from the same MCS in the same $T_b$ range is associated with the same cluster rainfall estimation ($R_c$), given by the following expression:

$$R_c = a_i\Delta E + b_iT_m + c_iT_{min} + d_i\Delta T_{min} + e_iT_{min} + f_i,$$

(1)

where $i$ is the $T_b$ range; $a_i$, $b_i$, $c_i$, $d_i$, and $e_i$ are the linear coefficients of each variable; and $f_i$ is the constant term.
Figure 4 shows the mean TRMM PR value versus the different cluster parameters for a 230-K threshold. Similar plots are obtained for the rest of the $T_b$ thresholds. Table 2 shows the coefficients and correlation index of the multiple linear regressions for each $T_b$ threshold. Using (1), a value of $R_c$ is assigned to all the pixels of one MCS in each $T_b$ range. First, all the pixels of the MCS are assigned the value of $R_c$ computed for the first $T_b$ threshold (pixels colder than 250 K); then the value of $R_c$ computed for the second $T_b$ threshold (pixels colder than 240 K) is superposed on the first one. The same occurs with the other values of $R_c$ for the

![Figure 4](image)

**FIG. 4.** Scatterplots between the parameters (see text) describing the evolution of the MCSs and the mean rain for a $T_b$ threshold of 230 K. A linear correlation is observed. All the parameters present a negative slope, except for the expansion. This shows a relation between the growth of the MCSs and their potential to produce heavy rainfall.
dependence between
of
rain values. This dependence is weaker in colder ranges
tive clouds and it is clear that colder pixels produce higher
mean value of the difference between the PR value and
Tb
shape) this is only true for the colder ranges of
MCS) are associated with higher values of
Figure 5 shows the mean TRMM PR value versus
stratiform clouds, a general tendency is observed: low
found for the cumulus, convective 2, convective 3, and

tionship for the six cloud types. The points were fitted
MCS parameters are associated to the whole cloud sys-
Tb
is the result of the comparison between the
and the cloud type. A similar methodology is
– VIS cloud classification
is a func-
and
PR and
is a function of
Rc
is the contribution of pixels to the
total rainfall rate and is a function of
Tc
which gives an
idea of the value of the spatial temperature gradient
within the MCS. However, both contributions are de-
vised as functions of the mean value of the TRMM PR
data. As a consequence, the range of
Rp in the cumu-
lated frequency distribution is smaller than the range of
the PR data (Fig. 7), but both PR and
Rp distributions
fit into an exponential distribution. The average value
of the PR data in different grid sizes was also computed
and its spatial distribution was compared to the
Rp distribution: the greater the grid size is, the more similar
are the ranges of the two distributions.

2) PIXEL RAINFALL ESTIMATION

In the second step, the following rainfall rate is as-
signed to each pixel that belongs to one of the six cloud
types producing rainfall:

\[ R_p = R_c + r_c, \]

where \( r_c \) is the result of the comparison between the
mean value of the difference between the PR value and
Rc for each 1-K interval of
Tc
. Therefore, \( R_p \) is a
function of
Tc
and the cloud type. A similar methodology is
used in Vicente et al. (1998). Figure 6 shows this
relationship for the six cloud types. The points were fitted
to a third degree polynomial curve. The curves in Fig. 6
have different shapes depending on the cloud type. Al-
though only a weak dependence between
r_c
and
Tc
is
found for the cumulus, convective 2, convective 3, and
stratiform clouds, a general tendency is observed: low
values of
Tc (corresponding to the coldest pixels of the
MCS) are associated with higher values of
r_c. In the
case of convective 2 pixels (with a concave downward
shape) this is only true for the colder ranges of
Tc. The
dependence between
Tc
and
r_c
is strong for deep convective clouds and it is clear that
colder pixels produce higher rain values. This dependence is weaker in colder ranges
of
Tc in the convective 1 cloud pixels, which are not in
the core of the MCS. Table 3 summarizes the coefficients of the polynomial curve for all the cloud types.

Both
r_c
and
r_p
contribute to the value of
R_p. The
r_c
is the cluster contribution and its value depends on the
radiative and evolution parameters that characterize
each MCS. The
r_p
is the contribution of pixels to the
total rainfall rate and is a function of
Tc
which gives an
idea of the value of the spatial temperature gradient
within the MCS. However, both contributions are de-
vised as functions of the mean value of the TRMM PR
data. As a consequence, the range of
Rp in the cumu-
lated frequency distribution is smaller than the range of
the PR data (Fig. 7), but both PR and
Rp distributions
fit into an exponential distribution. The average value
of the PR data in different grid sizes was also computed
and its spatial distribution was compared to the
Rp distribution: the greater the grid size is, the more similar
are the ranges of the two distributions.

3) RANGE INCREMENT

In the third step of the rainfall-rate estimation, both
PR and
Rp frequency distributions are approximated to
an exponential probability density function (pdf) for
each cloud type. The final value of the estimated rain-
fall rate
R is given by fitting the
Rp exponential pdf to
the PR exponential pdf. Figure 7 shows both distribu-
tions in the case of DCC. To fit one pdf into the other,
we assume the following approximation: the PR data
pdf and the final estimated rainfall-rate pdf have the
same rate of decay. The fitting method associates a
certain value of
Rp with the value of the PR data that
has the same rate of decay in both exponential pdfs.
This value assigned to
Rp will be the final rainfall-rate-
estimated value, \( \bar{R} \), given in the following expression:

Table 2. Coefficient values and correlation coefficients of the multiple linear regression to obtain the cluster rainfall estimation for
all the
T Thresholds.

<table>
<thead>
<tr>
<th>( T )</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( d )</th>
<th>( e )</th>
<th>( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>250 K</td>
<td>Coef value</td>
<td>0.000 81</td>
<td>-0.048 26</td>
<td>-0.083 93</td>
<td>-0.021 99</td>
<td>-0.020 15</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.32</td>
<td>-0.45</td>
<td>-0.41</td>
<td>-0.46</td>
<td>-0.37</td>
</tr>
<tr>
<td>240 K</td>
<td>Coef value</td>
<td>0.002 36</td>
<td>-0.019 61</td>
<td>-0.063 05</td>
<td>-0.050 48</td>
<td>0.007 24</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.35</td>
<td>-0.44</td>
<td>-0.28</td>
<td>-0.44</td>
<td>-0.24</td>
</tr>
<tr>
<td>230 K</td>
<td>Coef value</td>
<td>0.001 94</td>
<td>-0.070 76</td>
<td>-0.174 29</td>
<td>-0.011 76</td>
<td>-0.013 25</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.41</td>
<td>-0.44</td>
<td>-0.46</td>
<td>-0.43</td>
<td>-0.40</td>
</tr>
<tr>
<td>220 K</td>
<td>Coef value</td>
<td>0.002 54</td>
<td>-0.110 85</td>
<td>-0.123 12</td>
<td>-0.108 22</td>
<td>-0.020 18</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.39</td>
<td>-0.32</td>
<td>-0.35</td>
<td>-0.37</td>
<td>-0.39</td>
</tr>
<tr>
<td>210 K</td>
<td>Coef value</td>
<td>0.001 37</td>
<td>0.007 20</td>
<td>-0.119 89</td>
<td>-0.127 44</td>
<td>-0.073 76</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>0.45</td>
<td>-0.41</td>
<td>-0.41</td>
<td>-0.47</td>
<td>-0.48</td>
</tr>
</tbody>
</table>
\[ R(R_p, \lambda_{R_p}, \lambda_{r}) = (\lambda_{R_p}/\lambda_r)R_p, \]  \hspace{1cm} (3)

where \( \lambda \) is the rate parameter of the distribution.

Table 4 summarizes the values of \( \lambda_{R_p} \) and \( \lambda_r \) for all the accumulated histograms of the different cloud types.

4. Results

Several statistical measures were used to compare the rainfall-rate estimation with the PR data. In the case of the rain/no-rain discrimination, we used the false-alarm ratio (FAR), the probability of detection...
(POD), the error (ERR, and the frequency bias index (FBI). The range of POD, FAR, and ERR is 0 to 1. The perfect score for FAR and ERR is 0 and for POD is 1. Values of FBI greater (less) than 1 indicate overestimation (underestimation).

In the case of rainfall estimation, we used the correlation coefficient, the RMSE, the bias, and the standard deviation. Values of bias greater (less) than 0 indicate overestimation (underestimation).

RESAT was compared with the HE method (Vicente

![Diagram](image.png)

**Fig. 6.** Mean value of $r_c$ computed for each 1-K interval of $T_v$ for the six cloud types considered. The curve is fitted with a third degree polynomial function. A different curve is obtained for every cloud type, which allows discrimination between cloud types in the final rainfall estimation.

<table>
<thead>
<tr>
<th>Cloud type</th>
<th>Third degree</th>
<th>Second degree</th>
<th>First degree</th>
<th>Constant term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulus</td>
<td>3.09 × 10^{-4}</td>
<td>-64.21 × 10^{-4}</td>
<td>-0.049 499</td>
<td>-0.584 657</td>
</tr>
<tr>
<td>Convective 3</td>
<td>-2.47 × 10^{-4}</td>
<td>78.36 × 10^{-4}</td>
<td>-0.118 129</td>
<td>-1.784 454</td>
</tr>
<tr>
<td>Convective 2</td>
<td>-2.30 × 10^{-4}</td>
<td>0.014 565</td>
<td>-0.215 432</td>
<td>-1.047 433</td>
</tr>
<tr>
<td>Convective 1</td>
<td>4.68 × 10^{-4}</td>
<td>-0.019 028</td>
<td>0.103 186</td>
<td>-3.014 308</td>
</tr>
<tr>
<td>Stratiform</td>
<td>-1.77 × 10^{-4}</td>
<td>-11.12 × 10^{-4}</td>
<td>-0.015 940</td>
<td>-1.693 500</td>
</tr>
<tr>
<td>Deep convective</td>
<td>-23.40 × 10^{-4}</td>
<td>0.037 950</td>
<td>-0.074 900</td>
<td>-2.930 100</td>
</tr>
</tbody>
</table>
et al. 1998). The HE method uses a nonlinear power-law relationship between the cloud-top temperature and radar-derived rainfall estimates to compute rainfall rates. The HE uses a gradient and a growth rate mask to make the rain/no-rain discrimination. The main differences between RESAT and HE are that the HE does not take into account the cloud types and uses moisture and orographic corrections. The HE was originally developed to estimate rainfall rates in intense convective systems.

Table 4. Rate parameters of the pixel rainfall estimation ($\lambda_{RP}$) and radar frequency distributions ($\lambda_r$) for all cloud types.

<table>
<thead>
<tr>
<th>Cloud type</th>
<th>$\lambda_{RP}$</th>
<th>$\lambda_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulus</td>
<td>0.98</td>
<td>0.19</td>
</tr>
<tr>
<td>Convective 3</td>
<td>1.12</td>
<td>0.18</td>
</tr>
<tr>
<td>Convective 2</td>
<td>0.77</td>
<td>0.14</td>
</tr>
<tr>
<td>Convective 1</td>
<td>2.13</td>
<td>0.22</td>
</tr>
<tr>
<td>Stratiform</td>
<td>2.63</td>
<td>0.24</td>
</tr>
<tr>
<td>Deep convective</td>
<td>0.90</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Table 5. Statistical index describing the accuracy of the rain/no-rain discrimination of the RESAT method (training and intercomparison periods) and the HE method (intercomparison period).

<table>
<thead>
<tr>
<th></th>
<th>Training period</th>
<th>Intercomparison period</th>
<th>Intercomparison HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>POD</td>
<td>0.61</td>
<td>0.55</td>
<td>0.43</td>
</tr>
<tr>
<td>FAR</td>
<td>0.43</td>
<td>0.32</td>
<td>0.42</td>
</tr>
<tr>
<td>ERR</td>
<td>0.26</td>
<td>0.27</td>
<td>0.33</td>
</tr>
<tr>
<td>FBI</td>
<td>1.06</td>
<td>0.81</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Fig. 7. Frequency distribution of rainfall from radar and the pixel rainfall estimation ($R_p$) for DCC pixels. Both distributions can be approximated by an exponential distribution. The algorithm fits the $R_p$ distribution into the radar distribution to obtain the final rainfall estimation.

a. Rain/no rain

Table 5 shows the values of the different statistical index for the training and intercomparison periods. There are few differences between both periods. POD and FBI perform better during the training period, but FAR improves during the intercomparison period. This behavior seems logical, because FBI shows an underestimation of the rainy pixels in the intercomparison period.
period and an overestimation during the training period. The overestimation leads to better POD scores and the underestimation leads to better FAR scores. The difference in the values of FBI between the training and intercomparison periods may be due to the high percentage of DCC pixels (associated with high probability of rainfall) in the training period.

The entire index performs better in the RESAT than in the HE (Table 5). As already stated, the HE was developed to estimate intense convective rain. In this study, it was used for different kinds of meteorological situations; some involving convective storms characterized by the presence of deep convective clouds, others involving cloud systems that contain cumulus, stratus, and warm convective clouds. The HE clearly underestimates the rain pixels as the value $FBI_{\text{RESAT}} = 0.74$ shows, this underestimation may be due to the fact that HE tends to classify mainly DCC as rainy pixels. Table 6 shows the sensitivity of the cloud type in the rain/no-rain discrimination for both RESAT and HE methods. Results show how RESAT much performs better in the case of cold stratiform and growing cumulus clouds, but clearly RESAT scores better than HE. The underestimation of rainy pixels in the HE with a value of $FBI_{\text{HE}} = 0.11$, as opposed to the RESAT value of 0.78, leads to a better score of the FAR for the HE for this type of cloud. The differences between the two techniques are clear in the case of cumulus clouds, where RESAT scores a POD of 0.83, which is clearly higher than the HE score (0.16). In short, cloud classification as an input in the algorithm gives better results for rain/no-rain discrimination, especially for cold stratiform and growing cumulus clouds, but also in the convective ones.

### Rainfall estimation

The statistical variables used to validate RESAT were computed for different grid sizes: $5 \times 5$ pixels $(20 \times 20 \text{ km})$, $9 \times 9$ pixels $(36 \times 36 \text{ km})$, $15 \times 15$ pixels $(60 \times 60 \text{ km})$, and $25 \times 25$ pixels $(100 \times 100 \text{ km})$. The spatial mean rainfall rate was computed with the estimated rainfall rate; 1 h of cumulated rainfall would lead to better results (Vicente et al. 1998), but the TRMM satellite does not pass frequently enough to obtain the required data. The rain/no-rain statistical parameters (POD, FAR, ERR, and FBI) were also computed for the different grid sizes, which generated results similar to those already discussed. Table 7 shows the results for both the training and intercomparison periods.

The parameters used for the intercomparison score slightly better during the training period, without significant difference. The statistical parameters tend to converge to a similar value in both the training and intercomparison period when the grid size is increased, except for the bias. In both periods, RESAT tends to overestimate (bias > 0) the value of rainfall, with greater values of the bias during the intercomparison. This pattern is also observed in each of the studied cases. RESAT seems to perform reasonably well for a $60 \times 60 \text{ km}$ grid size, with a correlation factor of 0.54 in

### Table 6. Statistical index showing the accuracy of the rain/no-rain discrimination for different cloud types (deep convective, cumulus and cold stratiform) for both the RESAT and HE methods. Results show how RESAT much performs better in the case of cold stratiform and growing cumulus clouds.

<table>
<thead>
<tr>
<th></th>
<th>DCC</th>
<th>Cold stratiform</th>
<th>Cumulus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RESAT</td>
<td>HE</td>
<td>RESAT</td>
</tr>
<tr>
<td>POD</td>
<td>0.65</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>FAR</td>
<td>0.42</td>
<td>0.38</td>
<td>0.42</td>
</tr>
<tr>
<td>ERR</td>
<td>0.43</td>
<td>0.41</td>
<td>0.36</td>
</tr>
<tr>
<td>FBI</td>
<td>1.12</td>
<td>0.91</td>
<td>0.78</td>
</tr>
</tbody>
</table>

### Table 7. Statistical index showing the accuracy of the RESAT rainfall estimation for different grid sizes during the training and intercomparison periods.

<table>
<thead>
<tr>
<th>Grid size (km)</th>
<th>Training</th>
<th>Intercomparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>4460</td>
<td>1781 704 244</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.32</td>
<td>0.45 0.54 0.64</td>
</tr>
<tr>
<td>Satellite std dev (mm h$^{-1}$)</td>
<td>2.61 2.26 1.87 1.37</td>
<td>3.42 3.01 2.44 1.90</td>
</tr>
<tr>
<td>Radar std dev (mm h$^{-1}$)</td>
<td>2.35 1.86 1.38 0.92</td>
<td>2.76 2.22 1.68 1.24</td>
</tr>
<tr>
<td>RMSE (mm h$^{-1}$)</td>
<td>4.45 4.15 3.95 3.99</td>
<td>5.92 5.63 5.23 5</td>
</tr>
<tr>
<td>Bias (mm h$^{-1}$)</td>
<td>1.1 1.01 1 1.11</td>
<td>1.63 1.55 1.38 1.42</td>
</tr>
<tr>
<td>POD</td>
<td>0.81</td>
<td>0.88 0.94 0.98</td>
</tr>
<tr>
<td>FAR</td>
<td>0.16</td>
<td>0.11 0.07 0.03</td>
</tr>
<tr>
<td>ERR</td>
<td>0.23</td>
<td>0.17 0.11 0.03</td>
</tr>
<tr>
<td>FBI</td>
<td>0.97</td>
<td>1 1.02 1.01</td>
</tr>
</tbody>
</table>
the training period and 0.53 in the intercomparison period. The RMSE also decreases when the grid size increases. The bias seems to be independent of the grid size. Figure 8 show the scatterplots of the RESAT estimates for four different grid sizes.

COMPARISON WITH THE HE METHOD

Table 8 summarizes the values of the statistical parameters computed for both methods during the intercomparison period and Fig. 9 shows the scatterplots of HE estimates for two grid sizes. In general, RESAT performs better than the HE with clearly greater correlation coefficients and RMSE in all grid sizes. Even though both RESAT and HE present similar values of bias, bias seems to converge in the case of the HE. As was noted in the last section, in the case of RESAT, the bias does not decrease when the grid size decreases. This behavior can be explained by the fact that pixels within the same $T_b$ range that belong to the same cloud type are associated with a similar value of rainfall (the small differences would be given during the second step of the estimation); thus, the calculation of a spatial mean of the rainfall rate does not present high variability when the grid size is decreased.

Figure 10 shows an event during the intercomparison period with the RESAT and the HE estimations. It is a convective system in a growing phase, located near the Brazilian coast. Both the RESAT and TRMM PR images show how the cold core of the system, with a greater value of expansion, produce a heavier rainfall than the anvil, with a lower brightness temperature and lower value of expansion.

5. Summary and future work

The satellite rainfall-rate estimation algorithm we present, RESAT, mixes cloud-patch and pixel-based techniques. It is compared with the HE technique (Vicente et al. 1998), obtaining satisfactory results despite the fact that RESAT does not incorporate either moisture or orographic corrections. The TRMM precipitation radar product was used in the training and intercomparison periods in order to design and test the accuracy of the algorithm. The algorithm first discriminates the rainy pixels associated with rainfall only with respect to the pixels classified as cumulus, convective clouds, or cold stratiform clouds. To estimate rain in those pixels classified as rainy, the algorithm uses some radiation and evolution properties of MCSs at different brightness temperature thresholds. The MCS is divided into brightness temperature ranges using the $T_b$-range-tracking algorithm (Fig. 3), in order to detect the convective cells embedded in the MCS. MCS properties, such as the expansions of the MCS or its mean temperature difference, correlate linearly with the mean rainfall rate computed using the TRMM radar data for each MCS in each $T_b$ range. This permits us to assign an estimated value of rainfall rate to each MCS in each $T_b$ range. The use of IR–VIS cloud classification is accurate for both rain/no-rain discrimination and rainfall-rate estimation. The TRMM PR brightness temperature relationship and the cumulative TRMM data his-
Table 8. Statistical index comparing the rainfall-rate estimation of RESAT and HE for different grid sizes.

<table>
<thead>
<tr>
<th>Grid size (km)</th>
<th>RESAT</th>
<th>HE</th>
<th>RESAT</th>
<th>HE</th>
<th>RESAT</th>
<th>HE</th>
<th>RESAT</th>
<th>HE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 × 20</td>
<td>36 × 36</td>
<td>60 × 60</td>
<td>100 × 100</td>
<td>20 × 20</td>
<td>36 × 36</td>
<td>60 × 60</td>
<td>100 × 100</td>
</tr>
<tr>
<td>Sample size</td>
<td>1488</td>
<td>524</td>
<td>205</td>
<td>75</td>
<td>1488</td>
<td>524</td>
<td>205</td>
<td>75</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.26</td>
<td>0.39</td>
<td>0.53</td>
<td>0.64</td>
<td>0.22</td>
<td>0.39</td>
<td>0.41</td>
<td>0.49</td>
</tr>
<tr>
<td>Satellite std dev (mm h⁻¹)</td>
<td>3.42</td>
<td>3.01</td>
<td>2.44</td>
<td>1.90</td>
<td>5.14</td>
<td>3.76</td>
<td>2.67</td>
<td>1.72</td>
</tr>
<tr>
<td>Radar std dev (mm h⁻¹)</td>
<td>2.76</td>
<td>2.22</td>
<td>1.68</td>
<td>1.24</td>
<td>2.35</td>
<td>1.86</td>
<td>1.38</td>
<td>0.92</td>
</tr>
<tr>
<td>RMSE (mm h⁻¹)</td>
<td>5.92</td>
<td>5.63</td>
<td>5.23</td>
<td>5</td>
<td>6.01</td>
<td>5.66</td>
<td>5.54</td>
<td>5.56</td>
</tr>
<tr>
<td>Bias (mm h⁻¹)</td>
<td>1.63</td>
<td>1.55</td>
<td>1.38</td>
<td>1.42</td>
<td>1.66</td>
<td>1.46</td>
<td>1.37</td>
<td>1.36</td>
</tr>
<tr>
<td>POD</td>
<td>0.73</td>
<td>0.81</td>
<td>0.87</td>
<td>0.93</td>
<td>0.64</td>
<td>0.73</td>
<td>0.84</td>
<td>0.93</td>
</tr>
<tr>
<td>FAR</td>
<td>0.17</td>
<td>0.13</td>
<td>0.08</td>
<td>0.04</td>
<td>0.18</td>
<td>0.13</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td>ERR</td>
<td>0.25</td>
<td>0.2</td>
<td>0.13</td>
<td>0.07</td>
<td>0.39</td>
<td>0.33</td>
<td>0.21</td>
<td>0.09</td>
</tr>
<tr>
<td>FBI</td>
<td>0.89</td>
<td>0.94</td>
<td>0.95</td>
<td>0.97</td>
<td>0.81</td>
<td>0.87</td>
<td>0.92</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Fig. 9. Scatterplot of HE for (left) 60 × 60 km and (right) 100 × 100 km grid sizes for the validation period.

Fig. 10. Rainfall rate at 1045 UTC 17 Nov 2004: (top left) RESAT, (top right) HE, and (bottom left) radar. The square box on the map shows the area of the case study.
tograms strongly depend on the cloud type. Different parameterizations are set in the algorithm, according to the cloud type. As a result, the rainfall estimation yields good results over large temperature ranges.

POD and FBI obtain better scores during the training period (0.61 and 1.06, respectively) than during the intercomparison period (0.55 and 0.81), while the FAR scores are the worst (0.43 in the training period and 0.32 in the intercomparison period). This behavior is logical, as RESAT overestimates rainy pixels during the training period and underestimates them during the intercomparison period. The comparison was performed for different grid sizes. All the statistical variables seem to converge at a similar value for the two periods when the grid size is increased, except for the bias, which is always positive, indicating overestimation. The statistical variables show that RESAT gives satisfactory results for a grid size of 60 $\times$ 60 km, with a correlation coefficient of 0.53 and almost perfect POD (0.87) and FAR (0.08) scores. When compared with HE, RESAT always has a better correlation. The use of the cloud classification acquires special importance in the rain/no-rain discrimination. Few differences are observed between the two methods for cold pixels belonging to convective clouds. These differences increase significantly in RESAT’s favor when the methods are compared for cumulus and stratiform clouds. RESAT also presents a better correlation and lower values of RMSE than HE for all grid sizes (Table 8).

The comparison with HE shows that RESAT is ready to be put into operation and used as a complement to HE. A final future version of RESAT (able to replace HE) should have moisture and orographic corrections incorporated as well as a night cloud classification using the IR band channels of the GOES satellite.

Acknowledgments. This work has been supported by the Spanish Ministry of Science and Education under Grants REN2002-04558-C04-04 and CGL2005-07288-C05-04.

REFERENCES


