Atmos. Meas. Tech. Discuss., 6, 1269–1310, 2013
www.atmos-meas-tech-discuss.net/6/1269/2013/
doi:10.5194/amtd-6-1269-2013

Atmospheric Measurement Techniques Discussions

Validation of the METEOSAT storm detection and nowcasting system Cb-TRAM with lightning network data – Europe and South Africa

T. Zinner[1], C. Forster[2], E. de Coning[3], and H.-D. Betz[4,5]

[1]Meteorologisches Institut, Ludwig-Maximilians-Universität München, 80333 Munich, Germany
[2]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82230 Wessling, Germany
[3]South African Weather Service, South Africa
[4]Fakultät für Physik, Ludwig-Maximilians-Universität München, 80333 Munich, Germany
[5]Nowcast GmbH, 81377 Munich, Germany

Received: 11 January 2013 – Accepted: 23 January 2013 – Published: 4 February 2013

Correspondence to: T. Zinner (tobias.zinner@lmu.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In this manuscript, recent changes to the DLR METEOSAT thunderstorm TRacking And Monitoring algorithm (Cb-TRAM) are presented as well as a validation of Cb-TRAM against the European ground-based LIghtning NETwork data (LINET) of Nowcast GmbH and Lightning Detection Network (LDN) data of the South African Weather Service (SAWS). The validation is conducted along the well known skill scores probability of detection (POD) and false alarm ratio (FAR) on the basis of METEOSAT/SEVIRI pixels as well as on the basis of thunderstorm objects. The values obtained demonstrate the limits of Cb-TRAM in specific as well as the limits of satellite methods in general which are based on thermal emission and solar reflectivity information from thunderstorm tops.

Although the climatic conditions and the occurence of thunderstorms is quite different for Europe and South Africa, the quality score values are similar. Our conclusion is that Cb-TRAM provides robust results of well-defined quality for very different climatic regimes. The POD for a thunderstorm with intense lightning is about 80 % during the day. The FAR for a Cb-TRAM detected thunderstorm which is not at least close to intense lightning activity is about 50 %; if the proximity to any lightning activity is evaluated the FAR is even much lower at about 15 %. Pixel-based analysis shows that the detected thunderstorm object size is not indiscriminately large, but well within the physical limitations of the method. Nighttime POD and FAR are somewhat worse as the detection scheme can not use high resolution visible information. Nowcasting scores show useful values up to approximatelly 30 min.

1 Introduction

Today a wide range of possibilities for thunderstorm nowcasting based on satellite data is provided, due to the temporal and spatial coverage especially from a geostationary perspective, e.g. from METEOSAT SEVIRI (Spinning Enhanced Visible and InfraRed

AMTD
6, 1269–1310, 2013

Imager) data. One example is DLR's Cb-TRAM (Cumulonimbus Tracking and Monitoring) algorithm which detects, tracks, and nowcasts convection based on multi-channel METEOSAT SEVIRI data (Zinner et al., 2008). Geostationary satellites allow for a continous observation of thunderstorm development all over the observable part of the globe (between about −60 to +60° N and −60 to +60° E) independent from ground based networks like radar or lightning observation which are still only covering limited areas in the world with high sensitivity.

Moreover, satellite data generally allows for the observation of completely different stages of the storm development with the same sensor. For instance, instability indices are derived for cloud free areas before even first cloud development occurs, based on first-guess atmospheric profiles from numerical weather models and a successive adaptation to the vertical atmospheric temperature and moisture information derived from the water vapour and infra-red window channels of SEVIRI (Koenig and de Coning, 2009). The next stage of convective development is covered by detection schemes for the first appearance of clouds (convective initiation, e.g. Mecikalski and Bedka, 2006). Using a series of threshold tests (instantaneous as well as time trends) they indentify the cloudy areas which are most likely showing substantial convectively induced cloud growth about 45 min before they show considerable rain or even lightning. A similar detection scheme for convective initiation is also part of our Cb-TRAM algorithm, although this detection scheme is not the subject of analysis in the following (see Sect. 2, or Zinner et al., 2008). The third step, a detection of existing thunderstorms and monitoring of their life cycles is covered by techniques like the Rapid Development Thunderstorms tool (RDT) of MeteoFrance and Nowcasting SAF (Guillou, 2007) or again a detection scheme stage of Cb-TRAM (Zinner et al., 2008).

Thunderstorm detection and nowcasting using satellite observations is of increasing importance for aviation, as thunderstorms are related to hazardous phenomena like turbulence, icing, hail, and lightning that can lead to serious aircraft incidents. Information from thunderstorm detection and nowcasting algorithms like Cb-TRAM could help pilots in gaining a better overview of the weather situation as compared to what can

AMTD
6, 1269–1310, 2013

be provided by nowadays onboard observation systems (Senesi et al., 2009; Tafferner et al., 2009, 2010).

An important precondition before a pilot or other users can correctly use this information is the knowledge on its reliability. Thus, users and of course also developers of detection and nowcasting systems need quantitative characterization of the systems' capabilities. The value of pre-convective instability indices, of detections of convective initiation or developed storms, and eventually of nowcasting products derived have to be quantified. Only this way they can be used – and as far as the developer is concerned – compared and improved.

A variety of methods exist in order to quantify the capabilities of an algorithm or a numerical model. Most of them have been developed to validate model forecasts against observational data. The traditional validation approach is based on simple pixel-based grid overlays in which the forecast field is matched to an observation field or a set of observation points (Brown et al., 2004). Contingency tables are compiled which can then be used to compute verification measures and skill scores, such as the Probability of Detection (POD) and the False Alarm Ratio (FAR). For details on the scores see e.g. Wilks (1995) and Doswell et al. (1990). However, one problem with the traditional skill scores is the fact that they are insensitive to differences in location, timing and shape errors. For this reason, new approaches have been developed recently (see e.g. Casati et al., 2008 for a review of new verification approaches), one of them being the object- or feature-based approach (e.g. Ebert and McBride, 2000; Davis et al., 2006a,b; Marzban and Sandgathe, 2006) which identifies features in the forecast and observed fields and then assesses different attributes like position and size associated with each individual forecast-observation pair.

The main aim of this paper, besides a presentation of current improvements to the Cb-TRAM detection scheme, is the validation of Cb-TRAM against an independent observational data source. As lightning activity is an exclusive feature of thunderstorms (in contrast to, e.g. heavy precipitation), lightning data will be the independent data source of choice for the present analysis. For a validation over Europe the ground-based

AMTD
6, 1269–1310, 2013

LIghtning NETwork data (LINET) of Nowcast GmbH will be used, as they have a high accuracy over Europe and are continuously available over long time periods providing a good basis for a statistical analysis (Betz et al., 2008). Over South Africa, the Lightning Detection Network (LDN) data of the South African Weather Service (SAWS) is used (Gijben, 2012). A data set for a full 3-month period around the seasonal peak thunderstorm occurence is used for both regions. For other regions that are covered by the METEOSAT scan, no independent data source is available to date that constitutes an adequate source for a validation over long time periods. In order to provide a comprehensive assessment of the Cb-TRAM detections and nowcasts, both a traditional pixel-based and an object-based validation approach has been performed in this study. The following expands work started in Zinner and Betz (2009).

The paper is structured as follows: the two independent sources of thunderstorm detection, Cb-TRAM and lightning data, are introduced in the first sections. New developments and changes to the original Cb-TRAM METEOSAT algorithm by Zinner et al. (2008) are subject of Sect. 2. Object based detections and nowcasts of mature thunderstorms are provided. Section 3.1 presents the lightning networks LINET and LDN. In Sect. 4 these data are grouped to contiguous objects depending on different thresholds of measured lightning frequencies and their spread in time and space. The validation of Cb-TRAM objects against lightning data (pixel- and object-based) is presented in Sect. 5, and the results are discussed in Sect. 6.

2 The METEOSAT thunderstorm tracking and monitoring algorithm Cb-TRAM: recent improvements

Cb-TRAM is documented in Zinner et al. (2008) summing up work which has been going on at DLR for already more than 10 yr. It uses four different METEOSAT SEVIRI channels, namely the high resolution visible (HRV), the infra-red (IR) 10.8 µm, the IR 12.0 µm, and the water vapour (WV) 6.2 µm to detect three different development stages of thunderstorms: convection initiation, rapid growth, and mature stage.

AMTD
6, 1269–1310, 2013

Nowcasts are provided up to one hour (see Fig. 1). The core algorithms of Cb-TRAM, namely the image matching and motion vector derivation which enable the tracking and nowcasting of thunderstorm cells, were used for different purposes before: contrail detection (Mannstein et al., 1999), stereo imagery (Muller et al., 2007), but also first convective storm studies (Mannstein et al., 2002). Once the tool was established for day-to-day detection and tracking of convective cells on a project basis (EU projects RiskAware, 2004–2006; FLYSAFE, 2006–2009, Tafferner et al., 2008; ongoing DLR project Wetter & Fliegen; Forster and Tafferner, 2009, 2012) a rapid evolution of the detection schemes was initiated driven by weaknesses appearing during regular operation. Details on the detection schemes for earlier stages of the thunderstorm life cycle ("convective initiation" and "rapid growth") can be found in Zinner et al. (2008) and are only summarized here. The "mature" stage 3 detection scheme, however, experienced a major overhaul and is thus presented in detail.

The image matching technique analyses the motion field, or more precisely the transformation field, that describes the change from one image to the next. A continuous field of vectors is obtained from all features visible in the image regardless of its physical nature. The image is analysed stepwise from large scale to small scale features – the so called "pyramidal matching" procedure. This vector field can be utilized to generate intermediate or extrapolated synthetic images. The extrapolations are used throughout Cb-TRAM for several purposes.

First an extrapolation is used in the tracking scheme to facilitate the matching of cloud objects identified in the detection scheme at one time with its alter ego at the next time step. This feature improves the standard cloud object overlap matching technique by accounting for cloud motion. Especially the matching over long time periods or of small objects is improved. In a similar way the influence of cloud motion can be distinguished from the analysis of apparent growing tendencies and strong IR cooling trends as it can be estimated using an extrapolation beforehand. Finally, extrapolation in time is used to generate simple nowcasts of cloud object positions.

AMTD
6, 1269–1310, 2013

The complete detected area is sub-divided into objects, an object is a continuous group of pixels. Each object is labeled with the most severe development stage detected in any of its pixels ("convective initiation", "rapid growth", "mature"). To account for the oblique geostationary satellite viewing geometry, each object's position is parallax corrected using a cloud top height based on the mean 10.8 µm temperature observed within the object. For this parallax correction an uncertainty of a few kilometres in horizontal position has to be assumed (equivalent to one SEVIRI pixel). For all three storm stages a minimum size requirement of three connected pixels (8-connectivity) is implemented to avoid numerous spurious and fluctuating detections. A normal resolution METEOSAT pixel is about $4 \times 6\,\mathrm{km}^2$ (E–W by N–S) for Europe and $4.5 \times 4.5\,\mathrm{km}^2$ for South Africa, i.e. has an area close to $20\,\mathrm{km}^2$ for both areas.

Stage 1 "convective initiation" identifies cloud objects which show signs of convective growth (cumuli) without the display of clear thunderstorm activity yet. An object consists of all connected pixels which show an increase in HRV reflectivity which is accompanied by any IR 10.8 µm cooling.

Stage 2 "rapid development" identifies cloud objects which show a rapid cooling of more than $1\,\mathrm{K}\,15\,\mathrm{min}^{-1}$ in the water vapour (WV) 6.2 µm channel. Thereby, parts of cloud tops are detected which grow rapidly at heights at or close to the water vapor tropospheric background temperature. This is a usual sign of clouds growing close to strong inversions in the middle troposphere or at tropopause level.

Stage 3 "mature stage" detects clouds reaching or even overshooting tropopause levels. Originally ECMWF tropopause temperatures were derived for this detection scheme. Although this already constituted an improvement over the use of fixed temperature thresholds, detection failures occurred for low-capped thunderstorms and for the application in tropical environment (with a much less distinct cold point tropopause). These points lead to the following changes compared to the version presented in Zinner et al. (2008).

AMTD
6, 1269–1310, 2013

The stage 3 “mature thunderstorms” scheme is now composed of three main criteria:

1. as the new temperature criterion the difference $T_{6.2\,\mu m} - T_{10.8\,\mu m}$ is introduced which is complemented, as in the original Zinner et al. (2008) version,

2. by a HRV texture information during daylight hours (and a similar texture information from the WV 6.2 during night time),

3. and a removal of thin cirrus (which still can misleadingly match the other two crteria) by the use of a second temperature difference $T_{10.8} - T_{12.0}$.

The first two are combined in a way that a close miss of the storm threshold in one criterion can be compensated by a clear signal in the second. This fuzzy combination leads to much more consistent detections over a storm life cycle than the isolated use of arbitrary thresholds for the temperature criterion alone. In addition, and as demonstrated in Zinner et al. (2008), the use of HRV texture improves the separation of large areas of storm anvils and high cloud tops around frontal systems from the small cores of convective activity we are most interested in. A detection via the temperature difference alone was found to be too insensitive to make this separation and, at the same time, very sensitive to the exact value of the actually chosen temperature difference threshold. The third criterion masks out thin cirrus using a single threshold value.

First the $WV_{6.2\,\mu m} - IR_{10.8\,\mu m}$ difference is evaluated. Wherever it is positive, cloud tops are suspected to reach or overshoot the tropospheric background which is a clear sign of strong convective activity (Schmetz et al., 1997). This effect is attributed to tropospheric water vapor pushed into the stratosphere by towering convection. There, at increasing ambient temperatures above the tropopause, the additional water vapor emits radiation in the 6.2 µm channel while the measurement around 10.8 µm is not influenced by water vapor and shows the cloud tops around the cold point tropopause. As mentioned above, looking for a positive difference of these two channels alone leads to miss-detections of large cloud areas, especially in frontal systems. Raising this detection threshold to positive values, on the other hand, causes missed detections. Already

AMTD
6, 1269–1310, 2013

in the original setup similar problems with the insensitivity of the main temperature criterion led to the combination of different detectable signs of storm activity in a weighted non-binary sense. Namely the turbulent cloud top structure of active convective updraft cores is utilized in this context which is especially well detectable in the HRV channel during daytime.

During daylight hours (defined as local solar zenith angle SZA $< 75^{\circ}$) the "local standard deviation" is used as a texture measure for the HRV image. This standard deviation is obtained via application of a Gaussian weighting kernel centred on the pixel of interest to find a neighbourhood typical value and derive the weighted standard deviation from this value (Zinner et al., 2008). If the standard deviation is larger than the typical standard deviation found for 65 % of all thunderstorm (value obtained from Cb-TRAM test runs without texture criterion), the temperature difference is weighted with this standard deviation in a way that increases the likelihood for a detection. Technically the detection threshold for the $T_{6.2} - T_{10.8}$ difference could be lowered by up to 10 K this way: even a difference of −10 K could still be detected as mature storm, if the local standard deviation is large enough. During actual operation the most extreme values of local standard deviation observed lead to the detection of storms which show a negative difference of −3 K. Areas which do not show a clear texture signal of a turbulent thunderstorm cloud top, on the other hand, are less likely to be detected due to the combination of criteria. This excludes large cloud areas especially in situations of front passages. The dependence of the texture signal on solar zenith angle is accounted for, thereby becoming independent from geographic region, time of day, and season.

Emphasis in the following analysis is laid on the daytime version of the "mature thunderstorms" detection as the vast majority of all convective activity takes place during daylight hours. At the same time the spatially most detailed information, the high resolution visible channel, is only available during day light hours. During night time the HRV texture is replaced by an analogous WV6.2 texture signal. Although we are missing the high-resolution information provided by the HRV, there still is lower resolution information on the variability of the cloud top in the IR/WV channels, although it is less

AMTD
6, 1269–1310, 2013

specific in identification of most active cells. To provide comparable detection sensitivity between day and night, only the HRV texture contribution is swapped with the WV texture. The size of the texture contribution has to be tuned to match the daytime scheme: a local WV6.2 standard deviation exceeding the value which 75 % of all thunderstorms show is used (again value obtained from Cb-TRAM test runs without any texture criterion).

3 Lightning data

3.1 European lightning network data – LINET

Lightning detection can be performed by means of quite different techniques, but in many countries fully automated networks are most common, which utilize a number of antennae for the measurement of electric and/or magnetic fields emitted during lightning discharges. The sensor data are transmitted to a central processor, where lightning location is performed. LINET exploits the VLF/LF (Very Low Frequency/Low Frequency) regime and combines the measurement of cloud-to-ground (CG) and inter-cloud (IC) strokes within a single technology, employing baselines of 200–250 km for an adequate coverage in the central parts of the network (Betz et al., 2008). Presently, in many border areas with the inclusion of the Mediterranean Sea the baselines between stations are larger; consequently, the detection efficiency is reduced, i.e. weak IC and CG signals are not located. Figure 2 shows the sensor locations as of April 2008. The domain used in the following analysis covers the central LINET network with maximum sensitivity between a latitude of 40 and 54° North and a longitude of −5 to 16° East. For the analysis the lightning data is not divided into CG and IC strokes. The examined time period covers a Northern Hemisphere summer, namely June, July, and August 2008.

AMTD
6, 1269–1310, 2013

3.2 South African lightning data from LDN

During 2005, 19 VAISALA LS7000 sensors constituing the SAWS LDN have been installed across South Africa and are fully operational since the beginning of 2006 (Gijben, 2012). The SAWS LDN is only one of three ground-based lightning detection networks in the Southern Hemisphere; the others being in Brazil and Australia. Data from this network is supposed to provide primarly CG recordings. It constitutes a sufficient basis for a lightning climatology (Gijben, 2012).

In 2009, a major upgrade of the network was initiated. Four new sensors were added to the network between 2009 and 2010. The sensor network provides a detection efficiency of 90 % for all CG incidents and a location accuracy of 500 m within the boundaries of South Africa (Gijben, 2012). According to Zajac and Rutledge (2001) lightning detected at a distance of more than 100 km from the outer ring of lightning sensors is very often a false recording. The further data usage will thus be limited to continental South Africa; Cb-TRAM objects are only compared to LDN data over this region. The examined time period covers a Southern Hemisphere summer, namely December 2009, January, and February 2010.

4 Definition of lightning cells

The following validation is conducted on a SEVIRI pixel basis on the one hand, and on a storm object basis on the other hand. For the object-based comparison, the lightning reports are combined into contigous areas of certain minimum flash rate per area. This way the Cb-TRAM storm objects, based on the satellite data, can be validated against storm objects based on lightning data. In order to separate them from Cb-TRAM "objects" the latter are further called "cells". Figure 4 illustrates this process.

First, all reported lightning locations are allocated to the correct SEVIRI pixel. Multiple detections of a single event are filtered out by the requirement of a minimum time

AMTD
6, 1269–1310, 2013

and space separation (1 s, 5000 m). Several definitions of what a "good" storm detection has to identify are imaginable.

As mentioned in the introduction, the weather phenomena related to thunderstorms are hazardous for air traffic. However, if an aircraft intends to avoid a thunderstorm, the flight route has to be consolidated with other threats like other air traffic and ground collision. It is therefore helpful to have an indication on the severity of the hazard. A thunderstorm with only weak lightning activity is only a moderate hazard which will be avoided if possible, but which an aircraft could fly through if necessary. A thunderstorm with strong lightning activity, however, constitutes a severe hazard which should be avoided in any case. Following the literature (Steinacker et al., 2000; Oettinger et al., 2001; Betz et al., 2008) a series of possible lightning density thresholds are thinkable as a sign of convective activity. In the following the thresholds 0 and 5 flash reports within 3 km radius and 5 min are inspected closer, as they represent reasonable thresholds for "any" and for "severe" thunderstorm hazards with regard to aviation, respectively (e.g. Betz et al., 2008). These correspond to "any flash report per square kilometre and minute" and "more than 0.01 flash reports per square kilometre and minute". For the sake of clarity in the text the terms "any lightning" and "intense lightning activity" are used instead. The "intense lightning activity" level is approximately equivalent to "10 flash reports within a Meteosat pixel and a 15 min time period" (for Europe and Southern Africa).

For the object-based analysis all connected pixels (8-connectivity) wich show a lightning activity above the threshold are combined into lightning cells. In Fig. 4b (*any lightning*) and c (*intense lightning activity*) the contiguous orange or red colored areas in constitute the lightning cells (Cb-TRAM objects are colored in blue and red).

We require the Cb-TRAM detection to detect lighting activity even if confined to a single METEOSAT pixel, although Cb-TRAM is only able to detect storm objects of a mimimum size of "three connected pixels". Although this might deteriorate the detection quality, it seems a fair requirement as lightning activity in a single METEOSAT pixel usually is related to (satellite) detectable thunderstorm activity in clearly larger areas.

AMTD
6, 1269–1310, 2013

In addition to the variation in lightning activity (thresholds *any* and *intense*) we investigate three different levels of expected spatial accuracy for the detection: "overlap" with (no offset al.owed), "contact" with (one pixel offset), or "proximity" to lightning activity (two pixel offset). This is necessary as we cannot assume perfect matches of lightning activity and satellite detectable storm object for several reasons: (1) lightning activity does not necessarily happen directly beneath the most prominent cloud top characteristics detected by the satellite (e.g. through shear related tilt of the storm), (2) the localization of lightning activity is not perfect, a miss-location into an adjacent pixel is always possible, and (3) the parallax correction of Cb-TRAM detections carries an uncertainty of about one pixel as well, as it is done on an object basis only and not on a pixel-by-pixel basis. In the following validation the exact original object position as well as relaxed spatial accuracy requirements are evaluated to provide an exhaustive estimate of the skill.

4.1 Comparability of European and South African data

In general, Central Europe and South Africa obviously represent two very different thunderstorm regimes. The overall activity is to be expected clearly higher for the subtropical South Africa, which is identified as a hot spot of convection in global thunderstorm distributions (Zipser et al., 2006). While most of the thunderstorms in South Africa can be expected to be common multi-cell storms to mesoscale convective complexes not connected to frontal zones, European thunderstorm activity is often connected to fronts.

Given the high probability of different sensitivity for the two lighting detection networks and the fact that the SA network aims at providing CG events only while the European network provides both CG and IC detections, an adaptation of the activity thresholds used to allocate storm intensity seems inevitable. Unfortunately the problematic characterisation of lightning detections as CG or IC, which is done via an imprecise height detection for both networks, hardly allows to complete such an adjustment

AMTD
6, 1269–1310, 2013

in a fully correct way. In addition, a first analysis does not show any clear difference in detection efficiency for the European and the South African network.

In both analysis domains the land surface covers areas of comparable size (about 1.2 mio km^2 for South Africa, about 1.7 mio km^2 in Europe). During the analysed periods the overall lightning (detection) activity for South Africa is larger than in Europe (3.8 lightning detections per km^2 land surface in South Africa, 2.6 in Europe). This can be expected in a sub-tropical environment which is identified as favorable of convection in global thunderstorm distributions (Zipser et al., 2006). The difference in electrical activity is likely even larger, as the SAWS network only aims at detecting CG events and the most probable stroke current, a measure of sensitivity, in SAWS data is higher, i.e. the network less sensitive, compared to LINET data. Such sensitivity differences diminish if lightning activity is arranged into lightning cells. If done so, South Africa still displays a higher occurence of *intense* lightning cells by a factor of 1.5. The number of Cb-TRAM detected mature storm objects also points to very similarly increased activity in South Africa compared to Europe with a very factor of 1.6.

Summarizing, at first sight these total occurence numbers show very comparable relations with no unexpected dependance on region. Thus no adjustments of lightning activity numbers is introduced. The lightning records from both networks are used as they are.

5 Validation of Cb-TRAM against lightning data

We will provide skill characteristics in the form of the classical probability of detection (POD) and false alarm ratio (FAR) for Cb-TRAM detections and nowcasts in comparison to lightning cells on object- and pixel-basis. These two skill measures are the most widely used, and they have the advantages that they are, on the one hand, intuitively understandable and, on the other hand, the resulting values of direct instructive value for the user. POD and FAR are based on 2×2 contingency tables (Wilks, 2006). The

AMTD
6, 1269–1310, 2013

contingency Table 1 summarizes the combinations of criteria applied to generate the following skill results.

The object-based POD

$$\mathrm{POD} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{misses}} \quad (1)$$

for the Cb-TRAM detection scheme "mature thunderstorm" is obtained by checking for overlap with Cb-TRAM detections for each lightning cell at each METEOSAT time step (Fig. 5b). All spatial accuracy levels required are defined on the METEOSAT pixel grid: "overlap" is relaxed step-by-step by including lightning cells in "contact" (no gap between detection and lightning cell) and in "proximity" to Cb-TRAM detections (1-pixel gap) into the "hits". Following from the SEVIRI pixel size for Europe, *contact* means that the distance between lightning detection and Cb-TRAM detection is somewhere between 0 (!) and about 5 km, *proximity* means between about 5 and 10 km.

The pixel-based POD is provided in an analogous way (Fig. 5a). All pixels which are part of a Cb-TRAM mature thunderstorm detection as well as all METEOSAT pixels which show a lightning activtiy above the threshold (*any* or *intense*) are analysed.

The advantage of the pixel-based analysis is that a more objective skill information becomes available, as an analysis of objects disregards the size of the objects. When the size of the compared objects becomes very different the object-based POD and FAR become useless. E.g. when detected objects become large (or even contain all pixels) the POD becomes better and better (or even 100 %); at the same time the FAR would also improve (eventually become 0 %). The pixel-based analysis would represent the weakness of this approach by growing FAR.

The disadvantage of the pixel-based view is, in the first place, that lightning cells and satellite objects cannot not be expected to have the same size for physical reasons. While the satellite sensor can analyse the structure of the whole thunderstorm cloud body as visible from above, lightning activity cannot be expected over the whole area, but is mostly confined to certain dynamically active updraft regions. Thus only small parts of the Cb-TRAM object area can be expected to be covered by lightning activity.

AMTD
6, 1269–1310, 2013

In addition, horizontal offset of Cb-TRAM object and lightning cell can be caused by vertical tilt of thunderstorm development or/and by the Cb-TRAM parallax correction. These points together generally cause rather small PODs and very large FARs.

The object-based values are more informative from a user perspective, as they relate better to the expectations towards a storm forecast. For example, for air traffic applications the interesting information is whether a Cb-TRAM object in fact represents a mature intense thunderstorm. The exact position of lightning in the object is only of minor interest. The presence of downdrafts, heavy precipitation or hail, as well as clear air turbulence above the storm is at least of equal relevance for pilots and air traffic management (but cannot be accounted for with our validation data set).

The false alarm ratio

$$\mathrm{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \tag{2}$$

is generated from the numbers of Cb-TRAM detections separated in confirmed detections (hits) and false alarms. It is obtained in an analogous way as the POD in an object-based and pixel-based sense, separately for the two lighting activity levels and the three spatial accuracy levels.

Once a Cb-TRAM object has reached the mature stage, the nowcast of this object's position up to 60 min into the future is investigated also. For the purpose of an evaluation of the nowcast capabilities of Cb-TRAM alone, we could have completed the analysis on the basis of Cb-TRAM objects only: detected objects and extrapolated/nowcasted objects from an earlier time step could have been be compared. For clarity and the general validity of the results, we keep the above adopted approach to compare against the lightning activity of each nowcast object in the 15 min time frame around the time for which the nowcast was issued.

AMTD
6, 1269–1310, 2013

5.1 Detection and nowcasts – pixel based

Our validation data base for the Central European validation domain consists of 92 consecutive days of data (day and night) for a time period throughout the main thunderstorm season in summer 2008 (June, July, and August). The data for South Africa consists of 90 consecutive days of data (day and night) for a time period throughout the main thunderstorm season in the Southern Hemisphere in summer 2009/2010 (December, January, and February). Analyses are carried out for day and nighttime separately. "Daytime" is defined as a time step when the solar zenith angle is less than 75° for more than 75 % of the domain. The results over all lightning cell and Cb-TRAM detection and nowcast object pixels during daytime hours are presented for these regions and time span in the Table 2. POD and FAR for the analysis of pixels with the lightning activity level *intense* (> 10 flashes pixel^{-1} 15 min^{-1}) as well as the FAR for the level *any lightning* are shown.

Over Europe, about 69 % of all pixels showing *intense lightning activity* are detected by a Cb-TRAM mature stage detection for the same pixel. If the spatial tolerance of the analysis is increased, e.g. only *contact* or *proximity* to a Cb-TRAM detection is required, the POD even rises to 80 % and 84 % of all METEOSAT pixels with *intense* lightning activity.

At the same time, only 10 % of all Cb-TRAM detected pixels contain sl intense lightning activity (1-FAR, in Table 2), mainly due to the reasons discussed in the previous section. At least 29 % have *contact* to lightning activity (43 % are in *proximity*). Numbers improve when FAR for *any* activity is investigated. To pick just one value: at least about 41 % of all Cb-TRAM detected pixels have close contact to *any lighting* activity within the analysed 15 min time period.

As expectable, the skill scores for the nowcasts deteriorate the longer the forecast lead time. While the POD for an *intense* lightning pixel to be in direct *contact* with a Cb-TRAM detection is still 72 % for the 15 min nowcasts, a 60 min nowcast only provides a probability of a correct detection of 38 %. At the same time the likelihood

AMTD
6, 1269–1310, 2013

of a Cb-TRAM false alarm pixel with no lightning activity at least in *contact* increases from 69 % to 85 % for lead times between 15 and 60 min.

Over South Africa, the POD for *intense* lightning areas over all three different values of spatial accuracy are higher by about 9 percentage points (Table 2). Up to 93 % of the *intense* lightning pixels are at least in *proximity* of Cb-TRAM object pixels. At the same time all values of false alarm ratio are slightly higher for South Africa than for Central Europe by 1–3 points (apart from *any lightning* in *contact* or *proximity* where the difference amounts to 6 and 8 points). Higher POD and FAR for *intense* lightning suggests that the area with detected lightning is smaller or the area with Cb-TRAM detections is larger compared to the evaluation for Europe.

In summary, this all leads to the speculation that the lightning detection network of South Africa is slightly less sensitive or Cb-TRAM is more sensitive over South Africa. For instance, a less sensitive detection network would lead to smaller areas with a certain lightning activity compared to Europe. Only areas with a comparably stronger activity would be evaluated. Obviously these areas would be easier to be detected from space (higher POD), but at the same time less of the Cb-TRAM detections would contain such higher lightning activity (higher FAR).

The clear difference in detection between the two regions disappears within the first nowcast steps. Reason might be that the share of areas of lower lightning activity missed in the SAWS data favors correct nowcasts for the remaining more intense detections. Areas of strong activity tend to be better suited for extrapolation and thus the missing areas do not affect the numbers there.

In the light of the fact that lightning cells and Cb-TRAM objects cannot be expected to fully overlap or even have the same size for several physical reasons, the values in Table 2 are already very encouraging and reflect what thunderstorm dynamics and life cycle allow. Still this number is not specifically what a user or customer is interested in who uses a product flagging hazard areas due to mature thunderstorms. In addition, this pixel-based analysis is biased to large objects/cells which contribute many pixels and which, at the same time, are more likely to be detected. Small single cell

AMTD
6, 1269–1310, 2013

storms only covering a few METEOSAT pixels are not represented well in this average score values as they are much harder to detect, and even more pronounced, to be nowcasted.

5.2 Detection and nowcasts – object based

Opposed to the pixel-based analysis presented before, the object-based analysis treats each storm equally regardless of its size (compare Fig. 5b). This, of course, could in turn lead to an over-emphasis on the results for small cells, which might not be the most hazardous.

Without any additional detection tolerance 67 % of all cells showing *intense* lightning activity are detected (Table 3) over Europe. 71 % are in *contact* with a Cb-TRAM mature stage detection, and 73 % are within a one pixel vicinity of a detection. The FAR for cells of *intense* lightning for Cb-TRAM detections is 60 % for exact *overlap* (down to 52 % for a one pixel vicinity).

These values are not overwhelmingly good, but have to be put into perspective. One the one hand, removing small lightning cells from the analysis (e.g. with a mimimum size requirement of 3 pixels) stronlgy improves the POD to values up to around 80 % (not shown). an adjustment that could be considered fair as the Cb-TRAM detection is limited to this minimum object size as well. As Cb-TRAM's main application is related to air traffic safety, the large cells of heavy activity are of highest interest and a relatively high FAR (with regard to *intense* activity) could be accepted. On the other hand, as mentioned before, an object with at least *any lightning* activity cannot necessaryly be regarded a miss: checking Table 3 (lowest block of data) shows that under this assumption as few as 14 % of all Cb-TRAM objects do not have at least some lighting activity within *proximity*. I.e. of all detected Cb-TRAM objects, not all are of highest intensity, but the vast majority belongs to convective storms in a mature development stage when they produce lightning.

Figures 6 show the day-to-day variability for the three skill scores for "contacting objects" for Europe over the whole summer. One can see some variability although the

AMTD
6, 1269–1310, 2013

clearest outliers are often connected to days with only very few analyzed cases (LINET cells, bottom of 6). A line marks a moving average over all thunderstorms during an 11-day time frame. POD is roughly between 70 and 90 % and FAR between 50 and 75 % (both for "intense activity"). FAR (*any lightning*) is between 10 and 40 %.

The corresponding values for the four nowcast steps are shown in Table 3, too. They show interesting information and clear limits of our extrapolation technique. While the POD (*intense activity*, objects in *contact*) is still about 58 % for a 30 min nowcast, it drops down to around 44 % for the 60 min nowcast. At the same time even the tolerant FAR for *any lightning* activity in *contact* to Cb-TRAM objects reaches values of 35 (30 min) and 50 % (60 min). On the one hand, this is probably owed to the typical course of life of a convective cell which is ignored in our extrapolation algorithm. Even mature thunderstorms which are detected at one time can easily decay within 60 min. On the other hand, the results are a clear sign of technical characteristics of the extrapolation algorithm applied in Cb-TRAM. The motion or transformation vector fields derived are obtained from matching small scale brightness values in the context of a larger scale analysis step (pyramidal matcher). This can lead to sharp gradients in the vector field when small isolated features (clouds) move over large scale stationary background (surface). Vector fields extracted this way are well suited for the use in extrapolation for one or two time steps, as long as the motion still takes place in a similar embedding motion regime. That means, e.g. thunderstorms embedded in larger scale cloud systems or situations of broken cloud fields covering some area allow for better extrapolation results than small isolated convective cells. For general reliable nowcasts of more than 30 min improvements are necessary.

The object-based POD values for South Africa are even clearer positive than the pixel-based compared to Europe, about 11–14 percentage points (Table A4). Surprisingly, at the same time also the FAR values are better, i.e. smaller. FAR for cells with *intense* lightning activity are smaller than for Europe by 6–8 points while the values for *any* are smaller by 2–4 points. Generally differences are smaller for increasing nowcast horizon.

AMTD
6, 1269–1310, 2013

Certainly the most important reason for the presented differences for the two regions is the fact that South African convection is not obscured to the same extent by non-convective cloudiness, as it is the case for Europe with its frequent frontal passages. Especially for the satellite perspective, detection of active convective cores in wide spread frontal cloud layers present a challenge. Isolated thunderstorms are easier to detect and missing fronts do not provide an important source of false alarms.

The fact that the SAWS LDN is supposed to primarily provide CG detections could serve as a further explanatory approach. This is consistent with the expectation that with stronger thunderstorm intensity the ratio of IC vs. CG flashes grows. This would lead to an apparent lower flash rate for the strong SA cells compared to EU cells, while less intense cells could be less affected. This could lead to a loss of area containing lighting, especially on the edges of the *intense* lighting cells. Nonetheless, these *intense* cells would not disappear completely. Thus the false alarms on a pixel-basis would be more frequent, while the false alarms on an object-basis would not be affected at all. Further, the lower sensor density of the SAWS network leads to lower location accuracy and thus to slightly less compact cell derivations, which would be more similar to the Cb-TRAM objects. Also the tendency of CG flash rates dominating the later stages of a mature thunderstorms while IC flash rates peak earlier might play a role in this respect. Although these slightly speculative explanation approaches might point into the right direction, the specifics of the differences between the two lightning detection networks have to be left to further analysis.

5.3 Differences in day- and night-time detection

Up to now the focus of the analysis was on the daytime detection of mature thunderstorms using the high spatial resolution information of the SEVIRI HRV channel. The detection of mature storms during nighttime has to be based on IR channels alone (see Sect. 2). This affects the skill scores during the night. In addition, also the thunderstorm dynamics during night time differ from daylight hours. While many new and many short-lived storms develop during the day, mostly a few well organized thunderstorm

AMTD
6, 1269–1310, 2013

complexes or storms caused for synoptic reasons only (e.g. fronts) exist throughout the night.

Tables A1, A2, A3, and A4 show the skill scores for the night hours only and for all detected thunderstorms over all full 24 h days. Pixel-based POD is about 8–10 percentage points below the day values, while FAR increase by around 10 points as well (even more in the *proximity* cases). Object-based values of POD go down by 10 to 15 points, while FAR values go up mostly around 5 to 10 points and even 20 points for the *any lightning* activity threshold.

This all is an obvious consequence of the lower sensitivity of the nighttime detection due to the missing HRV information. The nighttime detection is obviously less specific and was adjusted in a compromise way to generate a smooth transition between day and night detection schemes for the most prominent thunderstorms. Clearly more nighttime detections are not related to thunderstorm activity at least in *proximity* (object-based FAR for *any lightning* around 30 % instead of around 13 %). At the same time more thunderstorms are missed due to the missing high resolution visible channel information (only about 60 % (POD) instead of about 77 % of all storm objects are detected on average over all spatial accuracies in the tables). A 30 min forecast projects the position of about 50 % of all storms correctly instead of around 60 % during daytime; after 60 min it is POD = 45 % during day and 41 % during night. This shrinking gap between night and day nowcast with forecast horizon could be a sign of the longer life-time, and that way, the better predictability of nighttime thunderstorms (although the forecast basis, the detection, is worse during the night).

Tables A3 and A4 show the numbers over all cases during day and nighttime. These total skill scores are closer to the daytime values as the majority of all storms and detections appear during the daylight hours.

AMTD
6, 1269–1310, 2013

6 Summary and conclusions

We presented a comparison of METEOSAT based thunderstorm identification and short-term forecasts with ground-based lightning data. This way a validation of the Cb-TRAM (Thunderstorm Tracking and Monitoring) algorithm for the detection of mature thunderstorms against lightning ground-truth is provided over 6 months in total in two different regions of the world (Europe and South Africa). The validation is conducted in the form of POD and FAR for different lightning intensity classes and different spatial accuracy requirements. Results are evaluated on a pixel- and on thunderstorm object-basis. The following summarizing values are averaged over all cases from South Africa and Europe.

The probability to detect a thunderstorm with *intense* lightning (approx. 10 flashes/pixel/15 min) with Cb-TRAM in Meteosat data reaches its highest object-based values during the day with 77 % for the medium spatial accuracy requirement (detected storm is in *contact* to lightning activity). False alarm ratios for a Cb-TRAM detection in fact not in *contact* to *intense* lightning are at 52 %; the false alarm ratio for a detection in *contact* with no lightning at all is much lower at 16 %.

Important additional information on top of these object-based results, are the results of the pixel-based analysis. It shows that the detected thunderstorm object size is not indiscriminately large, but well within the physical limits. As much as 85 % of all Meteosat pixels containing *intense* lightning are located at least next to a detected Cb-TRAM storm object. Vice versa, about 30 % of the area of detected Cb-TRAM objects really contain *intense* lightning activity (up to 40 % contain at least some lighting activity). This seems to be a reasonable value, as the satellite's detection of cloud top characteristics has to be less specific than the exact positioning of the active cores in lightning data.

Results degrade for the night-time detection scheme which can not use the high resolution visible information. Besides that, intensity of night-time thunderstorms might be lower and storms rather decay. This, of course, makes a correct detection more

AMTD
6, 1269–1310, 2013

difficult. At night POD for objects is lower by about 15 percentage points, while FAR for *intense* lightning objects increases by about 5 and for *any lighting* by about 20 points.

The scores for the nowcasts, which are generated through extrapolation of the current development state of the detected objects, degrade with forecast horizon. Still a 30 min daytime forecast of the position of a mature convective cell is in *proximity* (a maximum separation of one Meteosat pixel, about 5 km) of mature convective activity showing lightning activity in almost 75 % (= 1-FAR) of the cases and 67 % of all thunderstorms present at that time are forecasted with this accuracy (= POD). At 60 min this values become worse: POD = 55 % and FAR = 63 %.

All these quality scores, especially for the daytime detection scheme and the shorter range forcasts, are very encouraging given the fact that the vast majority of all strong potentially harmful cells is warned of while only very few do not contain any mature convective activity.

Main objective of this work was the objective characterisation of the detection and nowcast quality of Cb-TRAM. Of course, the values of all skill scores are decisively depending on the definition of success. We tried to show more than only one absolute criterion to allow the user to obtain a more complete view of Cb-TRAM's capabilities and limitations. Instead we provided different definitions of convective intensity and spatial accuracy of detections and nowcasts during the analysis. Although probably most direct and objective measure of mature convective activity was chosen, electrical activity, it became clear during the analysis that also our validation data is not the absolute truth. There exist differences in sensitivity in the used different lightning detection networks which affect the skill results (European LINET and South African SAWS LDN networks). Nonetheless, the effects are not clear enough to do more than speculate about the reasons. Together with the climatological fact that the typical South African locally triggered multi-cell thunderstorm is easier to be detected from a satellite perspective than part of the European storms which are triggered by and embedded into fronts, these lead to slightly better skill values for South Africa.

AMTD
6, 1269–1310, 2013

It has to be mentioned that scores become better with respect to POD for thunderstorm objects by about 10 percentage points on average during the day and more than 15 during the night, if not only the Cb-TRAM "mature stage" detection, but also its "rapid development" detection would be included. This means that 85 to 95 % of all thunderstorms are detected by one of the two Cb-TRAM schemes throughout the day (even the forecast POD reaching values above 70 % in this case). This development stage is supposed to preceed the mature stage, but of course includes the possibility of early electrical activity as the thunderstorm goes through a stage of strong updrafts quickly pushing the cloud top to higher levels. Of course, this hypothetic inclusion of this less specific stage drives false alarms to much larger values by as much as 20 percentage points, as the "rapid development" detection is not supposed to be a reliable sign of intense convection.

After the current state of Cb-TRAM nowcast skill is established with this study, an important next step will be an improvement of the nowcast scheme. So far it is a simple extrapolation of the currently observable trends. In the future it should include an additional life-cycle model to include the typical course of development into the nowcasts. Such a life-cycle model could, e.g. be included into a statistical forecast which could replace the existing deterministic after 15 or 30 min.

While this study was focused on the "mature stage" detection and the related forecasts, with the short excursion into the "rapid development" just mentioned, further validation work is in preparation which focuses on the "early stage" or "convective initiation" stage detected by Cb-TRAM.

Appendix A

Results for over nighttime only and whole day

Here the tables with nighttime and full day results for Europe and South Africa are provided for completeness. Pixel-based validation scores are given in Tables A1 and

AMTD
6, 1269–1310, 2013

A3. The related daytime results are given in Table 2. Object-based results for night and full day are given in Tables A2 and A4. Related daytime data can be found in Table 3.

Acknowledgements. This manuscript is dedicated to the memory of our colleague Hermann Mannstein who died unexpectedly and much too early on 25 January 2013. He was a scientific innovator behind many cloud remote sensing activities at DLR over nearly three decades. Among many other things, he developed the Cb-TRAM core algorithms. He will be deeply missed as expert, adviser and friend. The Meteosat SEVIRI data used within Cb-TRAM is copyrighted by EUMETSAT. Lightning data is provided by Nowcast GmbH and the South African Weather Service.

References

Betz, H. D., Schmidt, K., Oettinger, W. P., and Montag, B.: Cell-tracking with lightning data from LINET, Adv. Geosci., 17, 55–61, doi:10.5194/adgeo-17-55-2008, 2008. 1273, 1278, 1280

Brown, B., Bullock, R. R., Davis, C. A., Gotway, J. H., Chapman, M. B., Takacs, A., Gilleland, E., and Manning, K.: New verfication approaches for convective weather forecasts, in: 11th Conf. on Aviation, Range, and Aerospace, Hyannis, Massachusetts, USA, American Meteorological Society, No. 9.4, 13 pp., 3 October–7 October, 2004. 1272

Doswell, C., Davies-Jones, R., and Keller, D. L.: On summary measures of skill in rare event forecasting based on contingency tables, Weather Forecast., 5, 576–585, 1990. 1272

Forster, C. and Tafferner, A.: An integrated user-oriented weather forecast system for air traffic using real-time observations and model data, in: Proceedings of the European Air and Space Conference (CEAS), Manchester, UK, 11 pp., 26–29 October 2009, 2009. 1274

Forster, C. and Tafferner, A.: Nowcasting Thunderstorms for Munich Airport, DLR-Forschungsbericht ISSN 1434–8454, ISRN DLR-FB–2012-02, Deutsches Zentrum für Luft- und Raumfahrt e.V., Bibliotheks- und Informationswesen, Köln, 2012. 1274, 1308

Gijben, M.: The lightning climatology of South Africa, S. Afr. J. Sci., 108, 740, doi:10.4102/sajs.v108i3/4.740, 2012. 1273, 1279, 1306

Gill, T.: Initial steps in the development of a comprehensive lightning climatology of South Africa, Master's thesis, University of the Witwatersrand, Johannesburg, 2008. 1306

Guillou, Y.: Algorithm Theoretical Basis Document for "Rapid Development Thunderstorms", Tech. rep., Nowcasting Satellite Application Facility (NWC-SAF) Report,

AMTD
6, 1269–1310, 2013

SAF/NWC/CDOP/MFT/SCI/ATBD/11, Meteo France, Toulouse, France, Issue 1, Rev. 3, 2007. 1271

Koenig, M. and de Coning, E.: The MSG global instability indices product and its use as a nowcasting tool, Weather Forecast., 24, 272–285., 2009. 1271

Mannstein, H., Meyer, R., and Wendling, P.: Operational detection of contrails from NOAA-AVHRR data, Int. J. Remote Sens., 20, 1641–1660, 1999. 1274

Mannstein, H., Huntrieser, H., and Wimmer, S.: Determination of the mass flux in convective cells over Europe, in: Proceedings of the EUMETSAT Meteorological Satellite Conference 2002, Dublin, Ireland, 2–6 September 2002, EUMETSAT P.36, ISBN 92-9110-049-8, 264–269, 2002. 1274

Mecikalski, J. R. and Bedka, K. M.: SForecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery, Mon. Weather Rev., 134, 49–78, 2006. 1271

Muller, J.-P., Denis, M.-A., Dundas, R., Mitchell, K. L. M., Naud, C. M., and Mannstein, H.: Stereo cloud-top height and amount retrieval from ATSR2, Int. J. Remote Sens., 28, 1921–1938, 2007. 1274

Oettinger, W. P., Eisert, B., Betz, H.-D., and Gerl, A.: Automatische Erkennung extremer Wettersituationen im Nowcasting-Bereich, in: Proceedings DACH Tagung, Wien, 9 pp., 17–21 September, 2001. 1280

Schmetz, J., Tjemkes, S. A., Gube, M., and van de Berg, L.: Monitoring deep convection and convective overshooting with METEOSAT, Adv. Space Res., 19, 433–441, 1997. 1276

Senesi, S., Guillou, Y., Tafferner, A., and Forster, C.: Cb nowcasting in FLYSAFE: improving flight safety regarding thunderstorm hazards, in: WMO Symposium on Nowcasting, Whistler, B.C., Canada, 30 August–4 September 2009, No. 4.12, 14 pp., 2009. 1272

Steinacker, R., Dorninger, M., Wölfelmaier, F., and Krennert, T.: Automatic tracking of convective cells and cell complexes from lightning and radar data, Meteorol. Atmos. Phys., 72, 101–110, 2000. 1280

Tafferner, A., Forster, C., Hagen, M., Keil, C., Zinner, T., and Volkert, H.: Development and Propagation of Severe Thunderstorms in the Upper Danube Catchment Area: Towards an Integrated Nowcasting and Forecasting System using Real-time Data and High-resolution Simulations, Meteorol. Atmos. Phys., 101, 211–227, doi:10.1007/s00703-008-0322-7, 2008. 1274

AMTD
6, 1269–1310, 2013

Tafferner, A., Forster, C., Guillou, Y., and Sénési, S.: Nowcasting thunderstorm hazards for flight operations: the CB WIMS approach in FLYSAFE, in: Prodeedings of the European Air and Space Conference (CEAS), Manchester, UK, 26–29 October 2009, 16 pp., 2009. 1272

Tafferner, A., Forster, C., Hagen, M., Hauf, T., Lunnon, B., Mirza, A., Guillou, Y., and Zinner, T.: Improved thunderstorm weather information for pilots through ground and satellite based observing systems, in: 14th conference on Aviation, Range, and Aerospace Meteorology (ARAM), 90th AMS Annual Meeting, Atlanta, 17–21 January 2010, No. 161186, 2010. 1272

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, 467 pp., San Diego, USA, 1995. 1272

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, International Geophysics Series, 91, 2nd edn., Academic Press, Elsevier, Amsterdam, 627 pp., 2006. 1282

Zajac, B. A. and Rutledge, S. A.: Cloud to ground lightning in the contiguous United States from 1995 to 1999, Mon. Weather Rev., 129, 999–1019, 2001. 1279

Zinner, T. and Betz, H.-D.: Validation of Meteosat storm detection and nowcasting based on lightning network data, in: Proceedings of the 2009 EUMETSAT Meteorological Satellite Conference, Bath, UK, 21–25 September 2009, 9 pp., EUMETSAT P.55, iSBN 978-92-9110-086-6, ISSN 1011-3932, 2009. 1273

Zinner, T., Mannstein, H., and Tafferner, A.: Cb-TRAM: tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data, Meteorol. Atmos. Phys., 101, 191–210, doi:10.1007/s00703-008-0290-y, 2008. 1271, 1273, 1274, 1275, 1276, 1277

Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., and Yorty, D. P.: Where are the most intense thunderstorms on earth?, B. Am. Meteorol. Soc., 87, 1057–1071, 2006. 1281, 1282

AMTD
6, 1269–1310, 2013

Table 1. Contingency table for the comparison of LINET lightning and Cb-TRAM detections and nowcasts of mature thunderstorms on a pixel/object basis.

	lightning pixel/cell (*any* or *intense*)	no lightning pixel/cell (*any* or *intense*)
Cb-TRAM detect/nowc pixel/object	hit (pixel/cell within 0, 1, 2 px)	false alarm (pixel/cell within 0, 1, 2 px)
no Cb-TRAM detect/nowc pixel/object	miss (pixel/cell within 0, 1, 2 px)	correct negative

AMTD
6, 1269–1310, 2013

Table 2. Pixel based validation scores for the current Cb-TRAM detection scheme for mature storms during daytime, and the 15, 30, 45 and 60 min forecasts. top – POD for *intense lightning pixels* (> 10 flashes pixel^{-1} 15 min^{-1}), top; center – FAR with regard to *intense lightning pixels*; bottom – FAR$_{\text{any lightning}}$ with regard to *pixels* containing *any lightning* (> 0 flashes pixel^{-1} 15 min^{-1}).

accuracy	0 px *overlap*	1 px *contact*	2 px *proximity*	0 px *overlap*	1 px *contact*	2 px *proximity*
POD						
detection [%]	69.1	80.4	83.9	79.4	89.1	92.5
nowcast [%]						
15 min	59.0	71.9	78.4	63.4	72.7	79.7
30 min	46.4	59.5	68.0	50.7	59.9	67.6
45 min	35.5	47.8	57.1	40.1	48.6	56.3
60 min	27.4	38.3	47.0	31.9	39.2	46.2
FAR						
detection [%]	89.6	71.2	57.3	91.2	72.9	58.2
nowcast [%]						
15 min	92.7	77.6	64.5	93.7	79.1	65.8
30 min	94.9	83.1	71.8	95.5	84.3	73.2
45 min	96.2	86.9	77.3	96.6	87.7	78.5
60 min	97.0	89.4	81.2	97.2	89.8	81.8
FAR$_{\text{any lightning}}$						
detection [%]	84.9	59.1	40.7	87.7	65.0	48.8
nowcast [%]						
15 min	89.3	68.6	51.7	91.1	73.2	58.8
30 min	92.2	76.0	61.7	93.5	79.7	67.7
45 min	94.1	81.1	69.0	95.0	83.8	73.7
60 min	95.2	84.5	74.2	95.8	86.3	77.6

AMTD
6, 1269–1310, 2013

Table 3. Object based validation scores for the current Cb-TRAM detection scheme for mature storms during daytime, and the 15, 30, 45 and 60 min forecasted objects. top – POD for *intense lightning objects* (>10 flashes pixel^{-1} 15 min^{-1}); center – FAR with regard to *intense lightning objects*; bottom – $FAR_{any\ lightning}$ with regard to *objects* containing *any lightning* (>0 flashes pixel^{-1} 15 min^{-1}).

accuracy	Central Europe			South Africa		
	0 px *overlap*	1 px *contact*	2 px *proximity*	0 px *overlap*	1 px *contact*	2 px *proximity*
POD						
detection [%]	67.3	71.1	72.7	80.9	83.8	83.8
nowcast [%]						
15 min	58.9	65.2	68.6	66.1	74.5	78.3
30 min	49.2	57.6	62.4	55.3	65.4	71.4
45 min	40.6	50.5	56.0	45.8	57.2	63.8
60 min	33.0	43.5	49.9	38.3	49.6	57.0
FAR						
detection [%]	60.2	55.8	51.6	54.2	49.0	43.5
nowcast [%]						
15 min	64.7	57.5	52.4	62.3	52.6	45.6
30 min	70.8	62.1	56.0	69.4	58.7	50.5
45 min	76.1	66.9	60.5	74.9	64.4	56.2
60 min	80.1	71.5	64.9	78.8	69.6	61.8
$FAR_{any\ lightning}$						
detection [%]	24.0	16.9	14.0	20.4	14.8	12.0
nowcast [%]						
15 min	34.5	24.5	18.7	33.4	24.4	18.4
30 min	44.9	34.7	27.5	44.2	35.1	27.8
45 min	53.5	43.3	36.1	51.6	43.4	36.4
60 min	59.6	50.4	43.4	57.3	50.2	43.7

AMTD
6, 1269–1310, 2013

Table A1. Night: pixel based validation scores for the current Cb-TRAM nighttime detection scheme for mature storms, and the 15, 30, 45 and 60 min forecasts. Top – POD for *intense lightning pixels* (> flashes pixel^{-1} 15 min^{-1}), top; center – FAR with regard to *intense lightning pixels*; bottom – $FAR_{any\ lightning}$ with regard to *pixels* containing *any lightning* (> 0 flashes pixel^{-1} 15 min^{-1}).

accuracy	Central Europe 0 px *overlap*	Central Europe 1 px *contact*	Central Europe 2 px *proximity*	South Africa 0 px *overlap*	South Africa 1 px *contact*	South Africa 2 px *proximity*
POD						
detection [%]	66.1	72.4	75.3	71.5	79.8	83.7
nowcast [%]						
15 min	59.0	66.5	71.0	59.3	66.7	72.3
30 min	49.8	57.8	63.4	48.8	55.9	62.2
45 min	40.7	48.1	54.3	39.7	46.3	52.5
60 min	33.5	40.2	46.1	32.8	38.4	43.9
FAR						
detection [%]	92.8	78.3	67.2	95.1	84.1	74.0
nowcast [%]						
15 min	94.4	82.2	72.1	96.2	86.8	77.4
30 min	95.7	85.8	76.8	97.1	89.3	81.0
45 min	96.6	88.4	80.6	97.7	91.2	83.9
60 min	97.2	90.3	83.5	98.1	92.5	86.1
$FAR_{any\ lightning}$						
detection [%]	89.7	72.0	59.5	92.9	78.6	67.6
nowcast [%]						
15 min	91.9	77.2	65.9	94.4	82.1	71.9
30 min	93.7	81.6	71.8	95.5	85.2	76.3
45 min	94.9	84.9	76.4	96.3	87.6	79.7
60 min	95.8	87.2	79.8	96.9	89.2	82.2

AMTD
6, 1269–1310, 2013

Table A2. Night: object based validation scores for the current Cb-TRAM nighttime detection scheme for mature storms, and the 15, 30, 45 and 60 min forecasted objects. Top – POD for *intense lightning objects* (> 10 flashes pixel^{-1} 15 min^{-1}); center – FAR with regard to *intense lightning objects*; bottom – $FAR_{any\ lightning}$ with regard to *objects* containing *any lightning* (> 0 flashes pixel^{-1} 15 min^{-1}).

accuracy	Central Europe			South Africa		
	0 px *overlap*	1 px *contact*	2 px *proximity*	0 px *overlap*	1 px *contact*	2 px *proximity*
POD						
detection [%]	53.7	54.1	54.8	64.7	67.3	67.9
nowcast [%]						
15 min	49.5	51.9	53.5	53.9	61.1	65.0
30 min	44.5	48.4	50.6	46.7	55.4	59.9
45 min	39.3	44.4	47.1	39.7	49.2	54.9
60 min	34.4	40.2	43.4	33.7	43.3	49.7
FAR						
detection [%]	63.0	59.2	56.0	64.2	58.7	53.2
nowcast [%]						
15 min	66.9	61.0	57.1	69.0	61.1	54.9
30 min	72.1	64.9	60.1	73.4	64.7	58.0
45 min	76.2	69.4	64.1	77.4	68.6	61.9
60 min	79.3	73.3	68.2	80.5	72.7	66.0
$FAR_{any\ lightning}$						
detection [%]	39.2	35.1	32.4	39.3	34.5	30.9
nowcast [%]						
15 min	45.3	39.5	35.5	45.4	39.6	35.0
30 min	52.7	46.3	41.6	51.0	44.9	39.9
45 min	58.3	52.4	47.7	55.8	50.2	45.2
60 min	63.4	57.9	53.1	60.5	55.1	50.3

AMTD
6, 1269–1310, 2013

Table A3. Full day: pixel based validation scores for the current Cb-TRAM detection scheme (day and night) for mature storms, and the 15, 30, 45 and 60 min forecasts. Top – POD for *intense lightning pixels* (>10 flashes pixel^{-1} 15 min^{-1}), top; center – FAR with regard to *intense lightning pixels*; bottom – $FAR_{\text{any lightning}}$ with regard to *pixels* containing *any lightning* (>0 flashes pixel^{-1} 15 min^{-1}).

	Central Europe			South Africa		
accuracy	0 px *overlap*	1 px *contact*	2 px *proximity*	0 px *overlap*	1 px *contact*	2 px *proximity*
POD						
detection [%]	67.8	76.9	80.2	74.6	83.5	87.2
nowcast [%]						
15 min	59.0	69.7	75.4	61.0	69.2	75.5
30 min	47.7	58.9	66.2	49.7	57.7	64.6
45 min	37.3	47.9	56.1	39.8	47.4	54.2
60 min	29.5	39.0	46.7	32.4	38.8	45.0
FAR						
detection [%]	91.4	75.2	62.9	94.0	80.8	69.4
nowcast [%]						
15 min	93.7	80.2	68.8	95.5	84.5	73.9
30 min	95.4	84.6	74.6	96.6	87.8	78.6
45 min	96.4	87.7	79.1	97.4	90.2	82.3
60 min	97.1	89.9	82.5	97.8	91.7	84.8
$FAR_{\text{any lightning}}$						
detection [%]	87.6	66.3	51.2	91.3	74.6	62.1
nowcast [%]						
15 min	90.8	73.4	59.7	93.4	79.4	68.0
30 min	93.1	79.2	67.4	94.9	83.5	73.7
45 min	94.5	83.2	73.2	95.9	86.4	77.9
60 min	95.5	86.0	77.3	96.6	88.4	80.8

AMTD
6, 1269–1310, 2013

Table A4. Full day: object based validation scores for the current Cb-TRAM detection scheme (day and night) for mature storms, and the 15, 30, 45 and 60 min forecasted objects. Top – POD for *intense lightning objects* (> 10 flashes pixel^{-1} 15 min^{-1}); center – FAR with regard to *intense lightning objects*; bottom – FAR$_{\text{any lightning}}$ with regard to *objects* containing *any lightning* (> 0 flashes pixel^{-1} 15 min^{-1}).

	Central Europe			South Africa		
accuracy	0 px *overlap*	1 px *contact*	2 px *proximity*	0 px *overlap*	1 px *contact*	2 px *proximity*
POD						
detection [%] nowcast [%]	61.8	64.3	65.3	71.6	74.1	74.3
15 min	55.3	60.1	62.7	59.4	67.0	70.7
30 min	47.5	54.3	58.0	50.7	60.0	65.0
45 min	40.2	48.4	52.8	42.5	52.9	58.9
60 min	33.4	42.4	47.7	35.9	46.2	53.1
FAR						
detection [%] nowcast [%]	61.1	56.9	53.1	60.0	54.6	49.1
15 min	65.4	58.6	54.0	66.1	57.5	51.0
30 min	71.3	63.0	57.3	71.7	62.1	54.8
45 min	76.2	67.8	61.7	76.3	66.8	59.5
60 min	79.8	72.1	66.0	79.7	71.4	64.2
FAR$_{\text{any lightning}}$						
detection [%] nowcast [%]	29.0	22.9	20.0	31.3	26.1	22.9
15 min	38.0	29.4	24.2	40.3	33.1	27.9
30 min	47.5	38.5	32.1	48.1	40.7	34.8
45 min	55.0	46.3	39.9	54.0	47.3	41.4
60 min	60.8	52.8	46.6	59.1	53.0	47.5

AMTD
6, 1269–1310, 2013

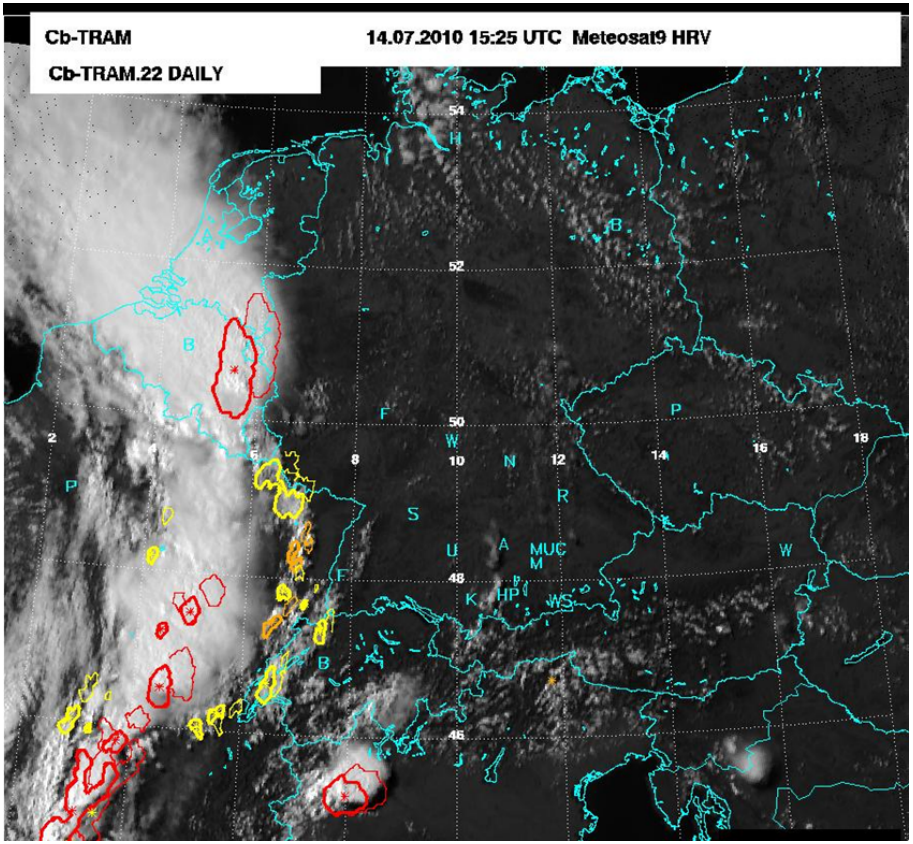

Fig. 1. METEOSAT-8 SEVIRI HRV image for the 14 July 2010, over central Europe, 14:25 UTC (time of data acquisition). Overlaid is the Cb-TRAM output (thick yellow, orange, red contours) and 30 min nowcasts (thin contours).

AMTD
6, 1269–1310, 2013

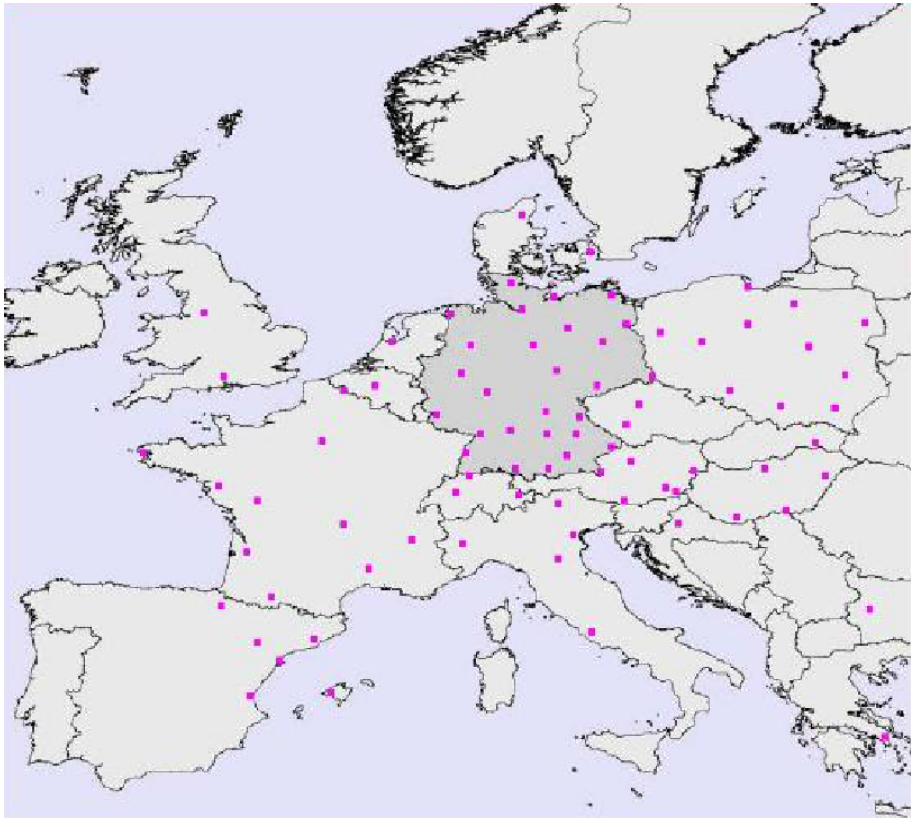

Fig. 2. Location of more than 100 sensor sites of the Nowcast lightning detection network LINET, as of April 2008 (provided by nowcast GmbH).

AMTD
6, 1269–1310, 2013

Fig. 3. 19 of 21 South African Weather Service LDN sensor sites operational during the time period December/January/February 2009/2010 (without the two newest sensors, at the time, at Springbok (Northern Cape) and Aliwal North (Eastern Cape); Gijben, 2012; image from Gill, 2008).

AMTD
6, 1269–1310, 2013

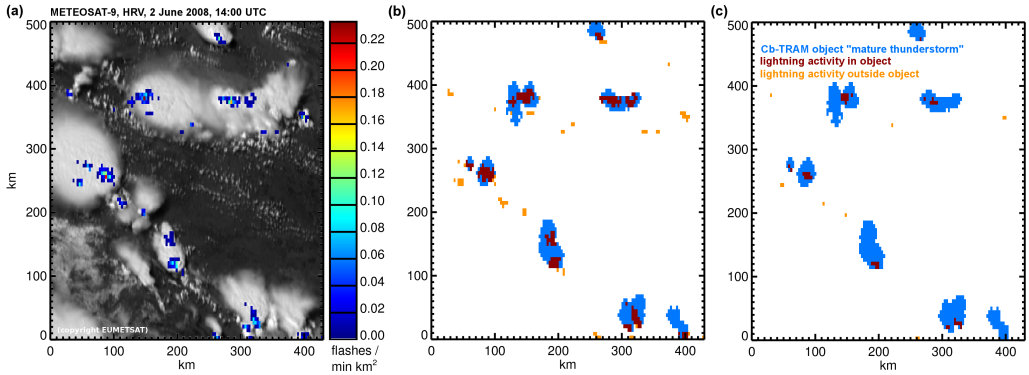

Fig. 4. (a) Shows METEOSAT SEVIRI HRV data (Germany, Alps in the south) on 2 June 2008 at 14:00 UTC (nominal measurement time, real data acquisition 14:07 UTC), overlaid is the flash rate from the LINET network mapped on METEOSAT normal resolution pixels (time of METEOSAT image ±7.5 min, parallax corrected). **(b)** Shows the comparison of Cb-TRAM mature thunderstorm detections in blue, "any" lightning activity within a detected Cb-TRAM object (red), and outside an Cb-TRAM object (orange). **(c)** As before for "intense" lightning activity.

AMTD
6, 1269–1310, 2013

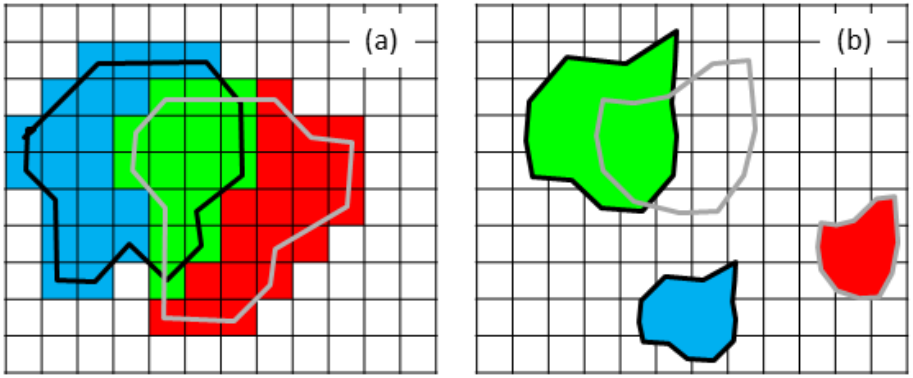

Fig. 5. Lightning (black) and Cb-TRAM (grey) objects. **(a)** Pixel-based analysis: the pixels covered by the lightning and the Cb-TRAM objects are counted. **(b)** Object-based analysis: the lightning and the Cb-TRAM objects are counted. Green are the hits, blue the misses, and red the false alarms (adapted from Forster and Tafferner, 2012).

AMTD
6, 1269–1310, 2013

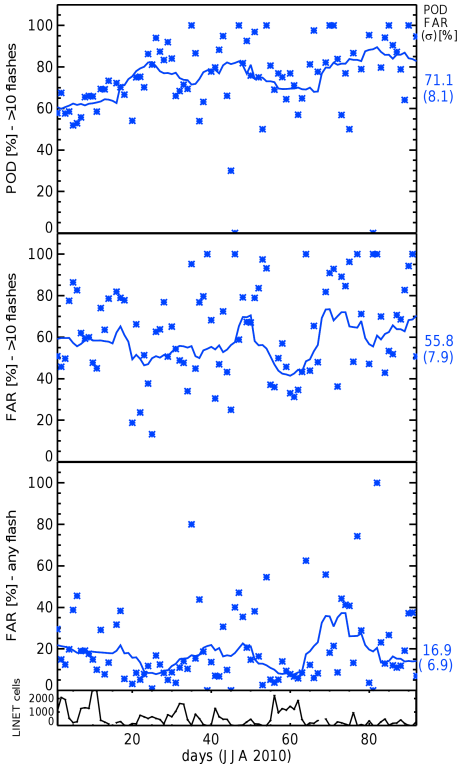

Fig. 6. Top: POD for a lightning cell of severe activity (>10 flashes pixel^{-1} 15 min^{-1}) which is detected by a Cb-TRAM object in *contact*. Shown are the day-to-day values as symbols and the moving average over all evaluated objects within 11 days as well as the mean and standard deviation of the 11 day-average to the right of the image (all values in percent). Center: FAR of Cb-TRAM objects which have no lightning cell showing *intense* activity in direct *contact*. Bottom: FAR of Cb-TRAM objects which do not have at least some lightning activity at all in direct *contact*. Very bottom: number of analyzed lightning cells on which the skill scores are based.

AMTD
6, 1269–1310, 2013

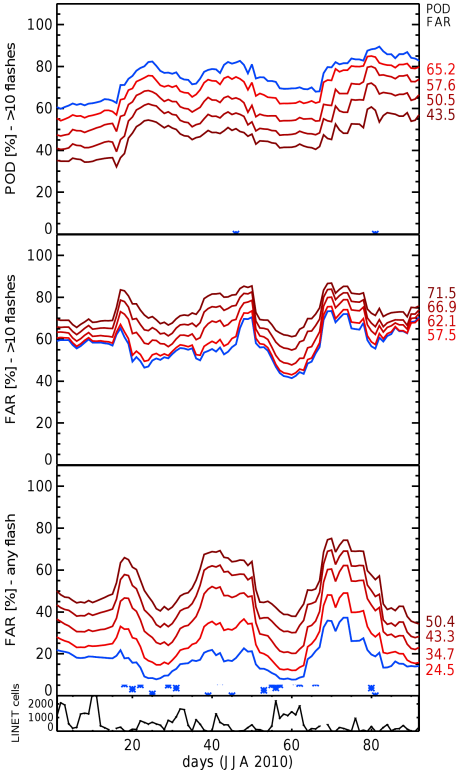

Fig. 7. 11-day moving averages for POD (top), FAR (center) for *intense* lightning cells, and FAR for cells of *any lightning* activity as in Fig. 6 in blue and in addition the 11-day averages for the 15, 30, 45 and 60 min Cb-TRAM nowcasts.

AMTD
6, 1269–1310, 2013