



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

T. Zinner<sup>1</sup>, C. Forster<sup>2</sup>, E. de Coning<sup>3</sup>, and H.-D. Betz<sup>4,5</sup>

<sup>1</sup>Meteorologisches Institut, Ludwig-Maximilians-Universität München, Munich, Germany

<sup>2</sup>Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

<sup>3</sup>South African Weather Service, Pretoria, South Africa

<sup>4</sup>Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany

<sup>5</sup>Nowcast GmbH, Munich, Germany

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

Received: 11 January 2013 – Published in Atmos. Meas. Tech. Discuss.: 4 February 2013

Revised: 7 May 2013 – Accepted: 17 May 2013 – Published: 19 June 2013

**Abstract.** In this paper, recent changes to the Meteosat thunderstorm TRacking And Monitoring algorithm (Cb-TRAM) are presented as well as a validation of Cb-TRAM against data from the European ground-based LIghtning NETwork (LINET) of Nowcast GmbH and the South African Weather Service Lightning Detection Network (SAWS LDN). Validation is conducted along the well-known skill measures 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 specific limitations of Cb-TRAM, as well as limitations of satellite methods in general which are based on thermal emission and solar reflectivity information from thunderstorm cloud tops.

Although the climatic conditions and the occurrence of thunderstorms are quite different for Europe and South Africa, 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 detection which is not even close to intense lightning is about 50 %. If only proximity to any lightning activity is required, FAR is much lower at about 15 %. Pixel-based analysis shows that detected thunderstorm object size is not indiscriminately large, but well within physical limitations of the satellite method. Night-time POD and FAR are

somewhat worse as the detection scheme does not use the high-resolution visible information during night-time hours. Nowcasting scores show useful values up to approximately 30 min in advance.

## 1 Introduction

In recent years a wide range of possibilities for thunderstorm nowcasting based on satellite data have become available, due to the temporal and spatial coverage especially from a geostationary satellite perspective. An example is the data from Meteosat SEVIRI (Spinning Enhanced Visible and InfraRed Imager). The Cb-TRAM (Cumulonimbus Tracking and Monitoring) algorithm of the German Aerospace Center (DLR) detects, tracks, and nowcasts convection based on multi-channel Meteosat SEVIRI data (Zinner et al., 2008). Geostationary satellites allow continuous observations of thunderstorm development all over the observable part of the globe (between about  $-60$  to  $+60^\circ$  N and  $-60$  to  $+60^\circ$  E) independent of ground-based networks such as radar or lightning observation networks. These still only cover limited areas in the world with high sensitivity.

Moreover, satellite data generally allow for observation of completely different stages of storm development with the same sensor. For instance, instability indices are derived for

cloud-free areas even before first cloud development even occurs. To this end, first-guess atmospheric profiles from numerical weather prediction models are adapted to vertical atmospheric temperature and moisture information which can be observed in the water vapour and infrared 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. Meczalski and Bedka, 2006). Using a series of threshold tests (instantaneous as well as time trends), they identify cloudy areas which most likely show substantial convectively induced cloud growth with about 45 min lead time or show considerable rain or even lightning. A similar detection scheme for convective initiation is part of the Cb-TRAM algorithm too (see Sect. 2 or Zinner et al., 2008), but will not be subject of the following analysis. The third step, a detection of existing thunderstorms and monitoring of their life cycles, is covered by techniques like the Rapid Developing Thunderstorm (RDT) product of MeteoFrance and Nowcasting Satellite Application Facility (Autones, 2012) or, again, the detection scheme stage of Cb-TRAM (Zinner et al., 2008).

Thunderstorm detection and nowcasting using satellite observations is of increasing importance for aviation. Thunderstorms are related to hazardous phenomena like turbulence, icing, hail, and lightning that can lead to serious air traffic incidents. Information from thunderstorm detection and nowcasting algorithms such as Cb-TRAM could help pilots in gaining a better overview of the weather situation compared to possibilities that can be provided by on-board observation systems (Senesi et al., 2009; Tafferner et al., 2009, 2010). Similarly, warnings of heavy precipitation, hailstorms, flash floods or severe wind gusts are important for public authorities, for example, fire departments, organisers of open-air activities as well as the general public.

An important requirement before a pilot or other user can correctly use this information is the knowledge of its reliability. Users and developers of detection and nowcasting systems need quantitative characterisation of the systems' capabilities. The quality of pre-convective instability indices, of convective initiation or developed storm detections, and eventually of the nowcasting products derived has to be quantified.

A variety of methods exist for the quantification of the capabilities of an algorithm or a numerical weather prediction model. Most of them have been developed to validate model forecasts against reference forecasts or 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, such as the probability of detection (POD) and the false alarm ratio (FAR). For details on these quantities, see for example Wilks (2006) and Doswell et al. (1990). However, one problem with

traditional skill measures 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). One of them is the object- or feature-based approach (e.g. Ebert and McBride, 2000; Marzban and Sandgathe, 2006) which identifies features in forecasted and observed fields. Then it assesses different attributes (e.g. position and size) associated with each individual forecast–observation pair.

The aims of this paper are twofold: firstly a presentation of recent improvements to the Cb-TRAM detection scheme will be shown, and, secondly, the validation of Cb-TRAM against an independent observational data source is presented. As lightning activity is an exclusive feature of thunderstorms (in contrast to, for example, heavy precipitation), lightning data will be the independent data source of choice for the presented analysis. For a validation over Europe ground-based Lightning Network data (LINET) of Nowcast GmbH will be used. This data set has a high accuracy over Europe and is continuously available over long time periods (Betz et al., 2008). This way a good basis for a statistical analysis is provided. Over South Africa, Lightning Detection Network (LDN) data of the South African Weather Service (SAWS) will be used (Gijben, 2012). A data set for a full 3-month period around the seasonal peak thunderstorm occurrence is used for both regions. For other regions that are covered by the Meteosat scan, no independent data source of comparable precision is available to date for a long time period. In order to provide a comprehensive assessment of Cb-TRAM detection and nowcast features, both a traditional pixel-based and an object-based validation approach has been performed in this study. The work presented here extends a previous study of 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 algorithm by Zinner et al. (2008) are the subjects of Sect. 2. Object-based detections and nowcasts of mature thunderstorms are provided. Sect. 3.1 presents the lightning networks LINET and LDN. In Sect. 3.3 these data are grouped to contiguous objects according to different thresholds of measured lightning frequencies and their variation in time and space. The validation of Cb-TRAM objects against lightning data (pixel- and object-based) is presented in Sect. 4, and results are discussed in Sect. 5.

## 2 The Meteosat thunderstorm tracking and monitoring algorithm Cb-TRAM

Cb-TRAM is documented in Zinner et al. (2008) summarising work which has been done at DLR for more than 10 yr. It uses four different Meteosat SEVIRI channels, namely the high-resolution visible (HRV), the infrared (IR) 10.8  $\mu\text{m}$ , the

IR 12.0  $\mu\text{m}$ , and the water vapour (WV) 6.2  $\mu\text{m}$ . These are used to detect three different stages of thunderstorms: convection initiation, rapid growth, and mature stage. Not necessarily all stages appear in nature in this order or are detectable from a satellite perspective. Nonetheless they mark an increasing risk of severe thunderstorm impact. The HRV is used whenever the local solar zenith angle is smaller than 75° (defined as daytime in Cb-TRAM).

Nowcasts are provided for up to one hour (see Fig. 1). In this study forecasts for 15, 30, 45 and 60 min are generated. The image matching and motion vector derivation which enable tracking and nowcasting of thunderstorm cells are core algorithms of Cb-TRAM. They were used for different purposes before: contrail detection (Mannstein et al., 1999), stereo imagery (Muller et al., 2007), and 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 less severe stages of thunderstorms (“convective initiation” and “rapid growth”) can be found in Zinner et al. (2008) and are only summarised here. The “mature” stage 3 detection scheme, however, experienced a major overhaul and is, thus, presented in detail in the following section.

Stage 1 “early development/convective initiation” identifies cloud objects which show signs of convective growth (cumuli) without clear thunderstorm activity yet. An object consists of all connected SEVIRI pixels which show an increase in HRV reflectivity which is accompanied by IR 10.8  $\mu\text{m}$  cooling.

Stage 2 “rapid development” identifies cloud objects which show a rapid cooling of more than 1 K 15 min<sup>-1</sup> in the water vapour (WV) 6.2  $\mu\text{m}$  channel. This way, parts of cloud tops are detected which grow rapidly at heights at or close to the water vapour tropospheric background temperature. This is a common 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. In the original version (Zinner et al., 2008) tropopause temperature from ECMWF forecast model runs was used 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 application in tropical environments (with a much less distinct cold point tropopause).

Apart from its use in detection schemes, the image matching technique is a central part of Cb-TRAM. It is used for the analyses of the motion field or, more precisely, of the transformation field that describes changes from one image to the next (Zinner et al., 2008). A continuous field of vectors is obtained from all features visible in an image pair regardless of

its physical nature. The image is analysed step by step from large-scale to small-scale features – the so-called “pyramidal matching” procedure. The obtained vector field can be utilised to generate intermediate images or extrapolated images. The extrapolations are used throughout Cb-TRAM for several purposes.

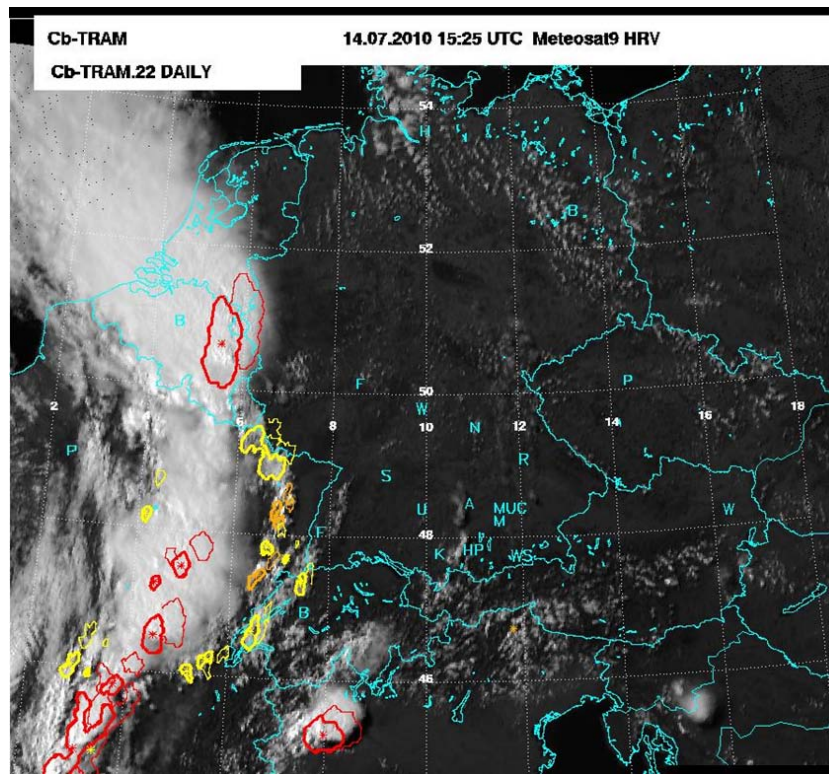
First of all, an extrapolation is used in the tracking scheme to facilitate the association of cloud objects identified at one time with its future manifestation at the next time step. In other words, detected object positions at a time  $t$  are extrapolated to forecasted positions at time  $t + 1$ . These forecasted positions are compared to new detections at time  $t + 1$  (via overlap analysis) in order to construct the track of specific objects. Specifically tracking over long periods of time and/or tracking of small objects is improved by this method. In a similar procedure the influence of cloud motion can be separated from cloud horizontal growth or strong IR cooling trends which constitute detection criteria, e.g. in the stage 1 and 2 schemes. To this end, expected changes due to cloud motion are removed from image differences before remaining differences are regarded as time trends. Finally, extrapolation in time is used to generate simple nowcasts of cloud object positions (for details see Zinner et al., 2008).

The pixel area detected is sub-divided into objects. An object is a continuous group of pixels. Each object is labelled with the most severe development stage detected in any of its pixels. 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  $\mu\text{m}$  temperature observed within the object and the related height from a temperature profile provided by numerical weather prediction model (here forecasts of the ECMWF, European Centre for Medium-Range Weather Forecast). 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 × 6 km<sup>2</sup> (E–W by N–S) for Europe and 4.5 × 4.5 km<sup>2</sup> for South Africa, i.e. has an area close to 20 km<sup>2</sup> for both areas.

### Recent improvements

The new version of the stage 3 “mature thunderstorms” scheme is composed of three main criteria:

1. a new temperature criterion is introduced: the difference  $T_{6.2\mu\text{m}} - T_{10.8\mu\text{m}}$ . As in the original version (Zinner et al., 2008), it is complemented by
2. a HRV texture information during daylight hours (and a similar texture information from the WV 6.2 during night-time).



**Fig. 1.** Meteosat-8 SEVIRI HRV image for 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).

- Thin cirrus, which still can misleadingly match the other two criteria, is removed using a second temperature difference  $T_{10.8} - T_{12.0}$ .

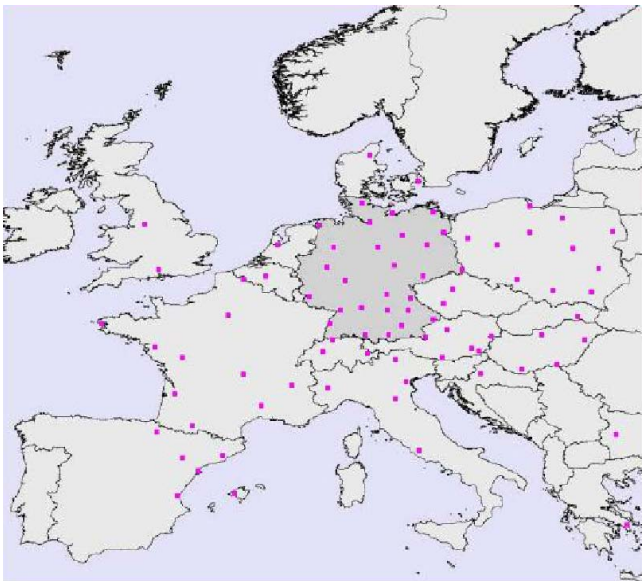
The first two are combined in a way that a close miss of the storm threshold in one criterion can be compensated by a clear signal in the second. This fuzzy combination leads to more consistent detections over a storm life cycle compared to isolated use of single thresholds for the temperature criterion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

First the  $WV_{6.2\mu\text{m}} - IR_{10.8\mu\text{m}}$  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 vapour pushed into stratosphere by towering convection. There, at increasing ambient temperatures above the tropopause, the additional water vapour emits radiation in the  $6.2\mu\text{m}$  channel. At the same time the measurement around  $10.8\mu\text{m}$  is not influ-

enced by water vapour and shows cloud tops around the cold point tropopause. As mentioned above, a positive difference of these two channels is not sufficient for a clear identification of deep convection. It can lead to miss detections of large cloud areas especially in frontal systems. Changing this detection threshold to positive values, on the other hand, causes missed detections.

In the original set-up the insensitivity of the main temperature remained an issue. This led to a weighted combination where additional detectable signs of storm activity were included. The turbulent cloud top structure of active convective updraft cores is utilised in this context. It is particularly well detectable in the HRV channel during daytime (shadows).

The “local standard deviation” is used as texture measure for the HRV image during daytime hours. 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 thunderstorms (value obtained from Cb-TRAM test runs without texture criterion), this is considered to be an additional warning sign. In this case the temperature difference is weighted with the standard deviation in a way that increases the likelihood for detection. Technically, the detection threshold for the  $T_{6.2} - T_{10.8}$  difference could be lowered



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

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 negative differences 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 frontal passages.

The dependence of brightness in the HRV channel on solar zenith angle is accounted for by the use of reflectivity (brightness normalised with available solar irradiance). The variation of the texture signal of cloud top structures with solar zenith angle is considered by normalisation with the typical dependence obtained from Meteosat data. This way the method becomes independent of geographic region, time of day, and season.

The emphasis in the following analysis is on the daytime version of the “mature thunderstorms” detection. 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 daylight hours. During night-time the HRV texture is replaced by an analogous WV6.2 texture signal. There is still low resolution information on the variability of the cloud top in the IR/WV channels available, although it is less characteristic for identification of the most active cells. The size of the texture contribution has to be tuned to match the daytime scheme: a local WV6.2 standard deviation is implemented. If this standard deviation exceeds the value that 75 % of all thunderstorms show, it increases the like-



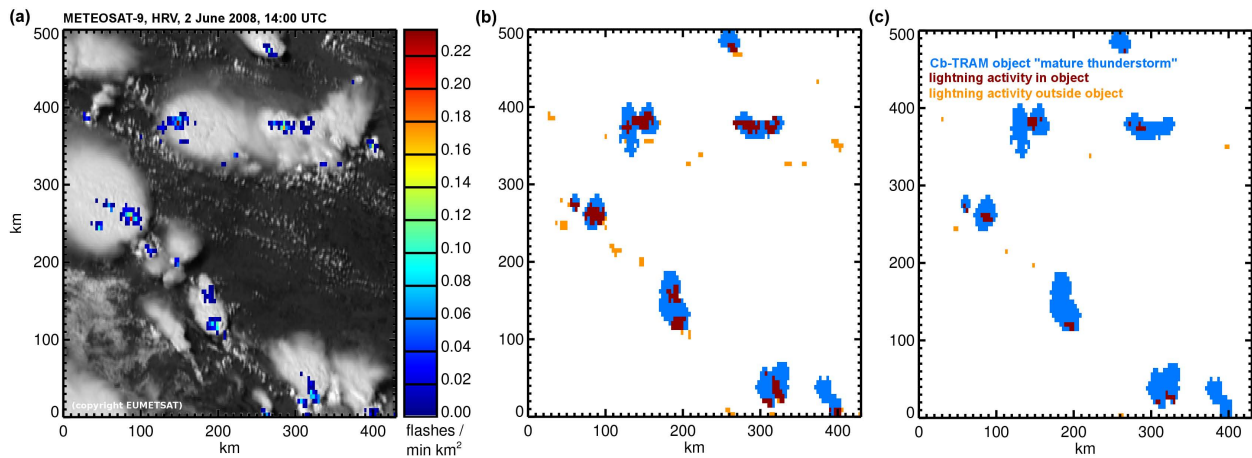
**Fig. 3.** 19 of 21 South African Weather Service LDN sensor sites operational during the time period December/January/February 2009/2010 (without the two newest sensors, at the time, at Springbok (Northern Cape) and Aliwal North (Eastern Cape); Gijben, 2012; image from Gill, 2008).

lihood for detection. This threshold value is obtained from Cb-TRAM test runs.

### 3 Lightning data

#### 3.1 European lightning network data – LINET

Lightning detection can be performed by means of a few varying techniques, but in some countries fully automated ground-based networks are installed. These networks utilise 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- and intra-cloud strokes (IC, not separated here). Baselines of 200–250 km are employed for an adequate coverage in the central parts of the network (Betz et al., 2008). In the relevant area LINET detects strokes with range-normalised currents down to 4 kA with a detection efficiency of more than 95%. The statistical average location accuracy is 150 m as verified by strikes into towers of known position and by evaluation of lightning-induced damages from insurance cases. In the border areas of the network (e.g. the Mediterranean), baselines between stations are larger. Consequently, detection efficiency is reduced in these areas: weak IC and CG signals are not located. Figure 2 shows the sensor locations in April 2008. The domain used in the following analysis covers the central LINET network with maximum sensitivity between a latitude of 40 and 54° N and a longitude of  $-5$  to 16° E. For



**Fig. 4.** (a) 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  $\pm 7.5$  min, parallax-corrected). (b) 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.

the analysis, lightning data are not separated into CG and IC strokes. The examined time period covers a Northern Hemisphere summer, namely June, July and August 2008.

### 3.2 South African lightning data from LDN

During 2005, 19 VAISALA LS7000 sensors constituting the SAWS LDN were installed across South Africa and have been fully operational since the beginning of 2006 (Fig. 3). 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 are supposed to provide primarily 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. Further data usage is thus 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, as well as January and February 2010.

### 3.3 Definition of lightning cells

The following validation is conducted on a SEVIRI pixel basis, on one hand, and on a storm object basis, on the other hand. For the object-based comparison, lightning reports are combined into contiguous areas of a certain minimum flash rate per area. This way Cb-TRAM storm objects, based on

the satellite data, can be validated against storm objects based on lightning data. In order to distinguish them from Cb-TRAM “objects”, the latter are called “cells” from now on. 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 separation in time and space (1 s, 5000 m).

Several definitions of a “good” storm detection are possible. As mentioned in the introduction, 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 such as other air traffic and ground collision. It is therefore helpful to have an indication of 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 at all costs. Following the literature (Steinacker et al., 2000; Oettinger et al., 2001; Betz et al., 2008), a series of possible lightning density thresholds are possible as a sign of convective activity. Herein the thresholds 0 and 5 flash reports within 3 km radius and 5 min are inspected closer. They represent reasonable thresholds for “any” and for “severe” thunderstorm hazards with regard to aviation (e.g. Betz et al., 2008). These correspond to “any flash report per square kilometre and minute” and “more than 0.035 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 South Africa).

**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 ( <i>any or intense</i> )	no lightning pixel/cell ( <i>any or intense</i> )
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

For the object-based analysis, all connected pixels (8-connectivity) with lightning activity above the threshold form a lightning cell. In Fig. 4b (*any lightning*) and c (*intense lightning activity*), orange and red coloured areas represent the lightning cells (Cb-TRAM objects are coloured in blue and red).

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

In addition to variations in lightning activity (thresholds *any* and *intense*), we investigate three different levels of expected spatial accuracy for the detection: “overlap” with (no offset allowed), “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 as detected by satellite (e.g. through shear-related tilt of the storm); (2) the localisation of lightning activity is not perfect, with a miss location into an adjacent pixel always possible; and (3) the parallax correction of Cb-TRAM detections carries an uncertainty of about one pixel also, as it is done on an object basis only and not on a pixel-by-pixel basis. In the following validation, exact object positions as well as relaxed spatial accuracy requirements are evaluated to provide an exhaustive estimate of the skill.

In principle there are inaccuracies in time which should be considered too. If the timing of lightning detections is assumed to be perfect, the uncertainty results from the timing precision of the Meteosat measurements. A full Meteosat SEVIRI scan takes about 12.5 min to cover the visible earth between about 75° southern latitude and 75° northern latitude. Although the resulting differences in the time

of data collection for South African and European SEVIRI data are taken into account, an inaccuracy of about 1–2 min can be expected. Thunderstorms usually move several metres per second. In other words, an assumed storm motion around 10 ms<sup>-1</sup> would translate into a spatial inaccuracy up to around 1000 m within 2 min. Consequently the effect of temporal mismatches is well below the different spatial accuracy levels, which will be tested in the following (up to two SEVIRI pixels or about 10 km).

### 3.4 Comparability of European and South African data

In general, Central Europe and South Africa represent two very different thunderstorm regimes. The overall activity is expected to be higher for sub-tropical South Africa, which is identified as a hotspot of convection in global thunderstorm distributions (Christian et al., 2003; Brooks et al., 2003). Most of the thunderstorms in South Africa can be expected to be multi-cell storms or mesoscale convective complexes which usually are not connected to frontal zones. European thunderstorm activity is often connected to fronts due to its location in the mid-latitude westerlies (see e.g. Doswell, 2001).

Given likely different sensitivities of the two lightning detection networks (e.g. the South African network detects primarily CG events), an adaptation of activity thresholds used to allocate storm intensity seems necessary. Unfortunately the problematic characterisation of lightning detections as CG or IC makes it difficult to make such an adjustment accurately. The characterisation for the European network is done by means of an imprecise height detection, and this information is not provided for the South African network at all. In addition, a first analysis does not show any clear difference in overall 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 × 10<sup>6</sup> km<sup>2</sup> for South Africa, about 1.7 × 10<sup>6</sup> km<sup>2</sup> in Europe). During the analysed periods, detected lightning activity for South Africa is larger than in Europe (3.8 lightning detections per km<sup>2</sup> land surface in South Africa, 2.6 in Europe). This is consistent with expectations. It is likely that the difference in real electrical activity is even larger. The SAWS network only aims to detect CG events, and the most probable stroke current – a measure of sensitivity – in SAWS data are higher (i.e. the network is less sensitive compared to LINET data). Such sensitivity differences diminish if lightning activity is arranged into lightning cells. If this is done, South Africa still displays a more frequent occurrence of *intense* lightning cells by a factor of 1.5. The number of Cb-TRAM-detected mature storm objects also points to a very similar difference in convective activity in South Africa compared to Europe with a factor of 1.6.

Summarising, at first sight these total occurrence numbers show very comparable relations with no unexpected

**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  $\text{pixel}^{-1} 15 \text{ min}^{-1}$ ); centre – FAR with regard to *intense lightning pixels*; bottom – FAR<sub>any lightning</sub> with regard to *pixels containing any lightning* ( $> 0$  flashes  $\text{pixel}^{-1} 15 \text{ min}^{-1}$ ).

accuracy	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
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 <sub>any lightning</sub>						
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

dependence on region. Thus no adjustments of lightning activity numbers are introduced. The lightning records from both networks are used as they are.

#### 4 Validation of Cb-TRAM against lightning data

Skill characteristics will be provided in the form of the probability of detection (POD) and false alarm ratio (FAR). These will be provided for Cb-TRAM detections and Cb-TRAM nowcasts in comparison to lightning cells on object and pixel basis. These two skill measures are long-established, and they have advantages: on one hand, they are intuitively understandable and, on the other hand, the resulting values are of direct instructive value for users. POD and FAR are based on  $2 \times 2$  contingency tables (Wilks, 2006). Contingency Table 1 summarises the combinations of criteria applied to generate the following skill results.

In order to provide the object-based POD,

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

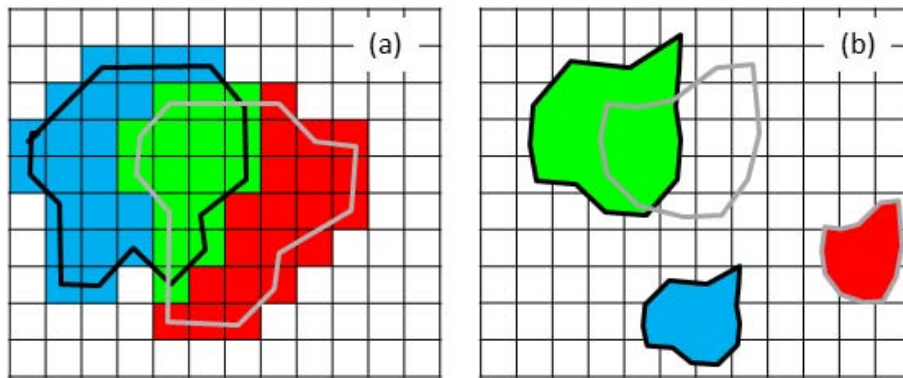
for the Cb-TRAM detection scheme “mature thunderstorm”, for each lightning cell, overlap with Cb-TRAM detections is analysed for each Meteosat time step (Fig. 5b). After the

analysis for “overlap”, the spatial accuracy is reduced step-wise, and lightning cells in “contact” (no gap between detection and lightning cell) and in “proximity” to Cb-TRAM detections (1-pixel gap) are included into the “hits”. In other words, the “contact” analysis includes all objects as hits that show overlap *or* contact. “Proximity” includes objects which show proximity, contact or overlap. For the SEVIRI pixel size for Europe, *contact* means that the distance between lightning detection and Cb-TRAM detection is 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 are analysed, as well as all Meteosat pixels which show lightning activity above the threshold (*any* or *intense*).

The advantage of the pixel-based analysis is that more objective skill information becomes available, because the analysis of objects disregards the size of objects. Whenever the sizes of the compared objects are very different, the object-based POD and FAR become useless. For example, when detected objects become large (in the extreme case it can contain all pixels), POD becomes better and better (in the extreme case even 100%). At the same time the FAR would improve (and eventually become 0%) too. A pixel-based





**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).

analysis conducted in parallel would show the weakness of this approach by a growing FAR.

The disadvantage of the pixel-based analysis is, firstly, that lightning cells and satellite objects cannot be expected to have the same size for physical reasons. While the satellite sensor can analyse the structure of the whole thunderstorm cloud as visible from above, lightning activity cannot be expected over the whole area. The latter is mostly confined to certain dynamically active updraft regions. Thus only small parts of the Cb-TRAM object area can be expected to contain lightning activity. In addition, a horizontal offset of Cb-TRAM object and lightning cell can be caused by a vertical tilt of the thunderstorm development or by the Cb-TRAM parallax correction. The consequence of all these points could be a rather small POD and very large FAR.

The object-based values are more informative from a user perspective, as they are more closely related to the expectations on a storm forecast. For example, for air traffic applications information whether a Cb-TRAM object in fact represents a mature intense thunderstorm is important. The exact position of lightning in the object is only of secondary 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 considered with our validation data set).

The false alarm ratio

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad (2)$$

is obtained from the number of all Cb-TRAM detections separated in confirmed detections (hits) and false alarms. It is provided analogous to the POD in an object-based and pixel-based sense, for the two lightning activity levels, and the three spatial accuracy levels.

Once a Cb-TRAM object has reached the mature stage, the nowcasts of this object's position up to 60 min into the future are investigated too. For simplicity, we apply the approach presented above to compare nowcasted positions

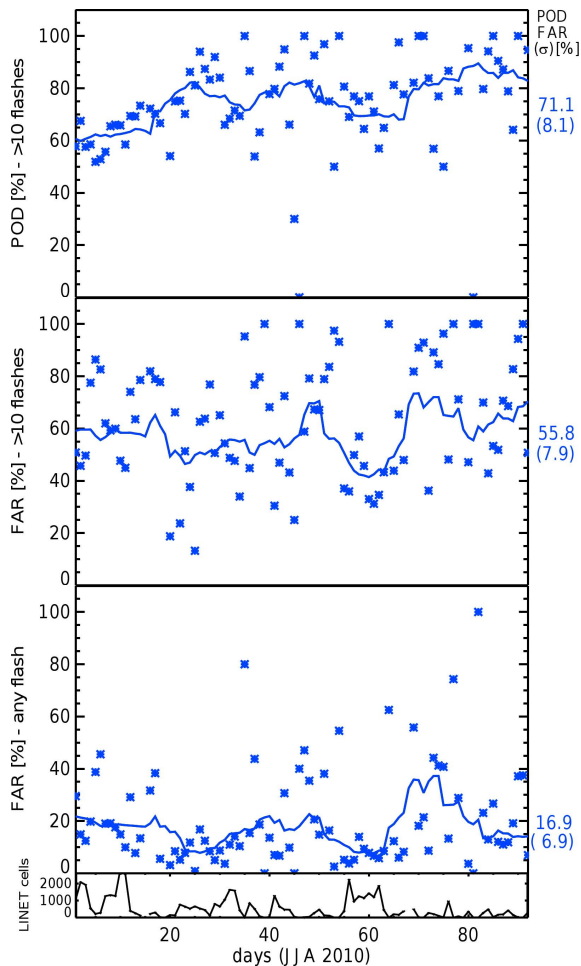
against measured lightning activity (in 15 min time frames around the forecast times 15, 30, 45 and 60 min). Although this way errors due to the advection and/or forecast scheme are mixed with errors due to the original detection, we want to avoid the introduction of another system of quality assessment at this stage (e.g. the comparison of nowcast to future Cb-TRAM detections). In addition, this approach seems to be the most user-oriented. A user is not interested in the source of error, but only in the fact whether a forecast is correct or not, i.e. whether lightning occurs at the forecasted location.

#### 4.1 Detection and nowcasts – pixel-based

For the Central European domain, 92 consecutive days of data (day and night) are evaluated throughout the main thunderstorm season in summer 2008 (June, July, and August). For South Africa 90 consecutive days are used from the Southern Hemisphere summer thunderstorm season 2009/2010 (December, January, and February). Analyses are carried out for daytime and night-time separately. “Daytime” is defined as a time step for which the solar zenith angle is less than  $75^\circ$  for more than 75 % of the domain. The results over all pixels of lightning cells and Cb-TRAM detection or nowcast objects are presented in Table 2. POD and FAR for the analysis of pixels with the lightning activity level *intense* ( $> 10$  flashes  $\text{pixel}^{-1} 15 \text{ min}^{-1}$ ) 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 required spatial accuracy in the analysis is reduced (i.e. only *contact* or *proximity* to a Cb-TRAM detection is required), POD improves. A total of 80 % or 84 % of all Meteosat pixels with *intense* lightning activity are detected.

At the same time, only 10 % of all Cb-TRAM-detected pixels contain *intense lightning activity* (1-FAR, in Table 2). This is mainly due to the reasons discussed in the previous section. At least 29 % have *contact* with lightning activity,



**Fig. 6.** Top: POD for a lightning cell of severe activity ( $> 10$  flashes  $\text{pixel}^{-1}$   $15 \text{ min}^{-1}$ ) 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). Centre: 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 analysed lightning cells on which the skill scores are based.

whereas 43 % are at least in *proximity*. Numbers improve further if FAR for *any* activity is investigated. To pick just one value – at least about 41 % of all Cb-TRAM-detected pixels have close contact with *any lightning* activity within the analysed 15 min time period.

The skill values for the nowcasts deteriorate with longer lead time. While 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 of a Cb-TRAM false alarm pixel with lightning

activity not even in *contact* increases from 69 to 85 % for lead times between 15 and 60 min.

Over South Africa, POD values for *intense* lightning are higher over all three evaluated accuracies 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 almost all values of the false alarm ratio are slightly higher compared to Central Europe by 1–3 points. The higher values of POD and FAR for *intense* lightning suggest that the overall lightning area (the total number of pixels) is smaller compared to the European domain (or the area with Cb-TRAM detections is larger).

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

The difference in detection between the two regions disappears within the first nowcast steps. The reason might be that the mentioned preference for areas of stronger activity facilitates extrapolation. The undetected pixels with weaker activity have less effect on the skill of nowcasts.

Given that lightning cells and Cb-TRAM objects cannot be expected to fully overlap or even to have the same size for several physical reasons, the values in Table 2 are already very encouraging. They may already reflect limitations caused by thunderstorm dynamics and life cycle. However, this pixel-based analysis is biased to large objects/cells which contribute many pixels to the analysis and which, at the same time, are more likely to be detected. Small single cell storms that only cover a few Meteosat pixels are not represented well. They are much harder to detect and even harder to forecast. Nonetheless, a user might be just as interested in these smaller scale events.

#### 4.2 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 obviously could lead to an over-emphasis of the results for small cells which might not be the most hazardous.

If *overlap* is required, 67 % of all cells that show *intense* lightning activity are detected over Europe (Table 3). A total of 71 % are in *contact* with a Cb-TRAM mature stage detection, and 73 % are within *proximity* 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 *proximity*).

These values are not overwhelmingly good, but have to be put into perspective. On one hand, removing small lightning

**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  $\text{pixel}^{-1} 15 \text{ min}^{-1}$ ); centre – FAR with regard to *intense lightning objects*; bottom – FAR<sub>any lightning</sub> with regard to *objects containing any lightning* ( $>0$  flashes  $\text{pixel}^{-1} 15 \text{ min}^{-1}$ ).

accuracy	Central Europe			South Africa		
	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
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 <sub>any lightning</sub>						
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

cells from the analysis strongly improves the POD to values up to around 80 % (e.g. by application of a minimum size requirement of 3 pixels, not shown). Such a size requirement could be considered a fair adjustment as the Cb-TRAM detection is limited to this minimum object size as well. On the other hand, as mentioned before, an object with at least *any lightning* activity cannot necessarily be regarded a miss. A check of Table 3 (lowest block of data) reveals that under this assumption as few as 14 % of all Cb-TRAM objects do not have at least some lighting activity within *proximity*. In other words, of all detected Cb-TRAM objects not all are of highest intensity, but the vast majority belong to convective storms in a mature development stage when they produce lightning.

Figure 6 shows the day-to-day variability for the three skill measures for “contacting objects” for Europe over the whole summer. One can see some variability, although the clearest outliers are often connected to days with only very few analysed cases (LINET cells, bottom of Fig. 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 also shown in Table 3. They show interesting information and clear limits of our extrapolation technique. While 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 FAR for *any lightning* activity in *contact* with Cb-TRAM objects reaches values of 35 (30 min) and 50 % (60 min). On one hand, this is probably due to the typical life cycle 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 if 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. In this case the motion still takes place in a similar motion regime. As a consequence, 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.

Compared to Europe, object-based POD values for South Africa show an even greater difference than the pixel-based.

They are better by about 11–14 percentage points (Table A4). Surprisingly, at the same time FAR values are better too. FAR for cells with *intense* lightning activity is smaller by 6–8 points while values for *any* are smaller by 2–4 points. Generally differences are smaller for increasing lead time.

An important reason for the presented differences for the two regions is the fact that South African convection is usually not as obscured from the satellite perspective as European convection is. In Europe non-convective cloudiness related to frequent frontal passages often embeds and covers the convection. For satellite detection, active convective cores in widespread frontal cloud layers present a challenge. Isolated thunderstorms are easier to detect, and fewer frontal systems do not provide an important source of false alarms for South Africa.

The fact that the SAWS LDN is supposed to primarily provide CG detections could serve as a further explanatory approach. Consistent with the expectation that with increasing thunderstorm intensity the ratio of IC vs. CG flashes grows, this would lead to an apparent lower flash rate for strong South African cells compared to European cells. Less intense cells could be less affected. This could lead to a loss of area containing lightning, especially on the edges of *intense* lightning cells. Nonetheless, these *intense* cells would not disappear completely. Thus false alarms on a pixel basis would be more frequent. The false alarms on an object basis would not be affected at all. Furthermore, the lower sensor density of the SAWS network leads to lower location accuracy. Thus derived cells are slightly less compact. This would make them more similar in size to Cb-TRAM objects. The tendency of CG flashes dominating in later stages of mature thunderstorms while IC flash rates peak earlier might play a role in this respect too. Although these slightly speculative explanations might point in the right direction, specifics of the differences between the two lightning detection networks have to be left to further analysis.

### 4.3 Differences in daytime and night-time detection

Up to now the analysis was focused on the daytime detection of mature thunderstorms. The high spatial resolution information of the SEVIRI HRV channel was exploited. The detection of mature storms during night-time has to be based on IR channels alone (see Sect. 2). This affects the skill measures during the night. In addition, thunderstorm dynamics during night-time also differs from daylight hours. While many new and many short-lived storms develop during the day, throughout the night mostly only a few well-organised thunderstorm complexes or storms exist (caused by synoptic reasons, e.g. frontal systems).

Tables A1, A2, A3, and A4 show skill measures for the night hours only as well as for all detected thunderstorms over the full 24 h day. Pixel-based POD is about 8–10 percentage points below the day values while FAR increases by around 10 points (even more in the *proximity* cases). Object-

based values of POD go down by 10 to 15 points while FAR values go up around 5 to 10 points and even 20 points for the *any lightning* activity threshold.

These changes are an obvious consequence of the lower sensitivity of the night-time detection due to missing HRV information. The night-time detection is obviously less specific. Clearly more night-time detections are not related to thunderstorm activity at least in *proximity*. The object-based FAR for *any lightning* is around 30 % instead of around 13 %. At the same time more thunderstorms are missed. 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 correctly predicts the position of about 50 % of all storms instead of around 60 % during daytime. After 60 min POD is 45 % during day and 41 % during night. This shrinking gap with nowcast lead time between night and day could be a sign of longer lifetimes and a related better predictability of night-time thunderstorms.

Tables A3 and A4 show results over all cases during daytime and night-time. These total skill measures are closer to the daytime values as the majority of all storms and detections appear during the daylight hours.

## 5 Summary and conclusions

We presented a comparison of Meteosat-based thunderstorm detection 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 two different regions of the world (Europe and South Africa). The validation is conducted by evaluation of POD and FAR for different lightning intensity classes and different spatial accuracy requirements. Results are evaluated on a pixel basis and on a thunderstorm object basis. The following values are averaged over all cases from South Africa and Europe.

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

The results of the pixel-based analysis adds important information to these object-based results. It shows that the detected thunderstorm object size is not indiscriminately large but well within the physical limits. As much as 85 % of all areas containing *intense* lightning are at least in *contact* with a detected Cb-TRAM storm object. On the other hand, about 30 % of the area of detected Cb-TRAM objects really contains *intense* lightning activity (up to 40 % contains at least some lightning activity). This seems to be a reasonable value,

as the satellite's detection of cloud top characteristics cannot be expected to be as specific as the exact positioning of the active cores in lightning data.

Results are worse for the night-time detection scheme when high-resolution visible information cannot be used. In addition, the intensity of night-time thunderstorms might be lower and storms might be more likely to decay. This makes a correct detection more 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 lightning* by about 20 points.

Nowcasts are generated through extrapolation of the current development state of detected objects. The scores for nowcasts degrade with forecast lead time. Still a 30 min daytime forecast of the position of a mature convective cell is at least close to thunderstorms in most cases: mature convection with lightning activity is in *proximity* (a maximum separation of one Meteosat pixel, about 5 km) in almost 75 % ( $= 1 - \text{FAR}$ ) of the cases. A total of 67 % of all thunderstorms present after 30 min have been forecasted with this accuracy (POD). At 60 min these values become worse:  $\text{POD} = 55\%$  and  $\text{FAR} = 63\%$ .

All these quality scores, especially for the daytime detection scheme and the shorter range forecasts, are very encouraging. Warnings can be issued for the majority of strong and potentially harmful cells, while only few false warnings would not be connected to mature convective activity.

The main goal of this work was the objective characterisation of the detection and nowcast quality of Cb-TRAM. Of course, values of all skill measures crucially depend on the definition of success. We tried to show more than only one absolute criterion to present a more complete view of Cb-TRAM's capabilities and limitations. We provided different definitions of convective intensity and spatial accuracy of detections and nowcasts during the analysis. Although probably the most direct and objective measure of mature convective activity was chosen, electrical activity, it became clear during the analysis that our validation data are not the absolute truth either. There exist differences in sensitivity in the 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 here. Reasons for slightly better skill values for South Africa might be a combination of these differences and the climatological characteristics of thunderstorms there. The typical South African locally triggered multi-cell thunderstorm is more likely to be detected from a satellite perspective than many European storms which are triggered by and embedded into fronts.

It has to be mentioned that the POD for thunderstorm objects becomes better if the "rapid development" detection of Cb-TRAM is included into the analysis. POD improves by about 10 to 15 percentage points on average if either the "rapid development" or the "mature stage" detection is required. 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 reaches values above 70 % in this case. This rapid development stage is supposed to precede the mature stage. Nonetheless, it includes the possibility of early electrical activity, because the thunderstorm goes through a stage of strong updrafts which quickly push the cloud top to higher levels. The inclusion of this detection stage, which is not supposed to be a reliable sign of intense convection, of course, also drives false alarms to much larger values (by as much as 20 percentage points).

In this study the current state of Cb-TRAM nowcast skill is established. The important next step will be an improvement of the nowcast scheme. So far it is a simple extrapolation of currently observable trends. The promising results mentioned previously for the 30 min daytime forecast with a relaxed spatial accuracy requirement seem to corroborate the implementation of a more probabilistic element in the nowcast. The spatial accuracy could be relaxed by a dilatation of object size from forecast step to forecast step. This way small forecast objects with precise position but low forecast skill could be replaced with less specific larger areas of high thunderstorm probability with better skill. Consequently, areas with many small storm objects could be merged into wider areas of general thunderstorm risk. In addition a life-cycle model could be included into a probabilistic nowcast to account for the typical course of development.

Generally, approaches to fuse the nowcast based on satellite imagery with numerical weather prediction to extend the forecast range are under evaluation as well. One possibility could be the best member selection as part of an ensemble forecasting system. In addition, precise global lightning detection, which is planned for the next generation of geostationary satellites (EUMETSAT MTG LI, 2011; Goodman et al., 2013), will strongly improve the detection of mature thunderstorms from satellite.

While this study was focused on the "mature stage" detection and related forecasts, validation work which focuses on the "early development" or "convective initiation" stage detected by Cb-TRAM can be found in Merk and Zinner (2013).

## Appendix A

### Results for over night-time and whole day

The tables with night-time and full day results for Europe and South Africa are provided for completeness here. Pixel-based validation scores are given in Tables A1 and 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.

**Table A1.** Night: pixel-based validation scores for the current Cb-TRAM night-time detection scheme for mature storms, and the 15, 30, 45 and 60 min forecasts. Top – POD for *intense lightning pixels* ( $> 10$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>); centre – FAR with regard to *intense lightning pixels*; bottom – FAR<sub>any lightning</sub> with regard to *pixels containing any lightning* ( $> 0$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>).

accuracy	Central Europe			South Africa		
	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
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 <sub>any lightning</sub>						
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

**Table A2.** Night: object-based validation scores for the current Cb-TRAM night-time detection scheme for mature storms, and the 15, 30, 45 and 60 min forecasted objects. Top – POD for *intense lightning objects* ( $> 10$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>); centre – FAR with regard to *intense lightning objects*; bottom – FAR<sub>any lightning</sub> with regard to *objects containing any lightning* ( $> 0$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>).

accuracy	Central Europe			South Africa		
	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
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 <sub>any lightning</sub>						
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

**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<sup>-1</sup> 15 min<sup>-1</sup>); centre – FAR with regard to *intense lightning pixels*; bottom – FAR<sub>any lightning</sub> with regard to *pixels containing any lightning* ( $> 0$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>).

accuracy	Central Europe			South Africa		
	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
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 <sub>any lightning</sub>						
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

**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<sup>-1</sup> 15 min<sup>-1</sup>); centre – FAR with regard to *intense lightning objects*; bottom – FAR<sub>any lightning</sub> with regard to *objects containing any lightning* ( $> 0$  flashes pixel<sup>-1</sup> 15 min<sup>-1</sup>).

accuracy	Central Europe			South Africa		
	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>	0 px <i>overlap</i>	1 px <i>contact</i>	2 px <i>proximity</i>
POD						
detection [%]	61.8	64.3	65.3	71.6	74.1	74.3
nowcast [%]						
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 [%]	61.1	56.9	53.1	60.0	54.6	49.1
nowcast [%]						
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 <sub>any lightning</sub>						
detection [%]	29.0	22.9	20.0	31.3	26.1	22.9
nowcast [%]						
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

*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 an expert, adviser and friend. The Meteosat SEVIRI data used within Cb-TRAM are copyrighted by EUMETSAT. Numerical weather prediction data from the European Centre for Medium-Range Weather Forecasts (ECMWF) were used within Cb-TRAM. We thank Gerard Kilroy for his English proofreading as well as two anonymous referees and Jean-Marc Moisselin for their helpful reviews.

Edited by: P. Stammes

## References

- Autones, F.: Algorithm Theoretical Basis Document for “Rapid Development Thunderstorms” (RDT-PGE11 v2.3), Tech. rep., Nowcasting Satellite Application Facility (NWC-SAF) Report, SAF/NWC/CDOP/MFT/SCI/ATBD/11, Météo France, Toulouse, France, Issue 2, Rev. 3, 2012.
- 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.
- Brooks, H. E., Lee, W. L., and Craven, J. P.: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data, *Atmos. Res.*, 67–68, 73–94, 2003.
- Brown, B., Bullock, R. R., Davis, C. A., Gotway, J. H., Chapman, M. B., Takacs, A., Gilleland, E., and Manning, K.: New verification 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.
- Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocerlich, M., Damrath, U., Ebert, E. E., Brown, B. G., and Mason, S.: Forecast verification: current status and future directions, *Meteorol. Appl.*, 15, 3–18, 2008.
- Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Boeck, W. L., Buechler, D. E., Driscoll, K. T., Goodman, S. J., Hall, J. M., Koshak, W. J., Mach, D. M., and Stewart, M. F.: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector, *J. Geophys. Res.*, 108, 4005, doi:10.1029/2002JD002347, 2003.
- Doswell, C. A. (Ed.): Severe Convective Storms, Meteorological Monographs, American Meteorological Society, 45 Beacon Street, Boston, Massachusetts, 2001.
- 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.
- Ebert, E. E. and McBride, J. L.: Verification of precipitation in weather systems: Determination of systematic errors, *J. Hydrol.*, 239, 179–202, 2000.
- EUMETSAT MTG LI, Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data, Doc. No.: EUM/MTG/DOC/11/0155, EUMETSAT, Darmstadt, October, 2011.
- 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.
- 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.
- Gijben, M.: The lightning climatology of South Africa, *S. Afr. J. Sci.*, 108, 740, doi:10.4102/sajs.v108i3/4.740, 2012.
- Gill, T.: Initial steps in the development of a comprehensive lightning climatology of South Africa, Master’s thesis, University of the Witwatersrand, Johannesburg, 2008.
- Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., Carey, L., Schultz, C., Bateman, M., McCaul Jr., E., and Stano, G.: The GOES-R Geostationary Lightning Mapper (GLM), *Atmos. Res.*, 125–126, 34–49, doi:10.1016/j.atmosres.2013.01.006, 2013.
- 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.
- Mannstein, H., Meyer, R., and Wendling, P.: Operational detection of contrails from NOAA-AVHRR data, *Int. J. Remote Sens.*, 20, 1641–1660, 1999.
- 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.
- Marzban, C. and Sandgathe, S.: Cluster analysis for verification of precipitation fields, *Weather Forecast.*, 11, 824–838, 2006.
- Mecikalski, J. R. and Bedka, K. M.: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery, *Mon. Weather Rev.*, 134, 49–78, 2006.
- Merk, D. and Zinner, T.: Detection of convective initiation using Meteosat SEVIRI: implementation in and verification with the tracking and nowcasting algorithm Cb-TRAM, *Atmos. Meas. Tech. Discuss.*, 6, 1771–1813, doi:10.5194/amtd-6-1771-2013, 2013.
- 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.
- 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.
- 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.
- 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.
- 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.



- 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.
- Tafferner, A., Forster, C., Guillou, Y., and S n si, S.: Nowcasting thunderstorm hazards for flight operations: the CB WIMS approach in FLYSAFE, in: Proceedings of the European Air and Space Conference (CEAS), Manchester, UK, 26–29 October 2009, 16 pp., 2009.
- 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.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, International Geophysics Series, 91, 2nd Edn., Academic Press, Elsevier, Amsterdam, 627 pp., 2006.
- 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.
- 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.
- 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.