
Convection Initiation
-
Detection and Nowcasting with
multiple data sources

Dennis Stich

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Dennis Stich
aus München

München, den 20. Dezember 2012

Erstgutachter: Prof. Dr. G. Craig

Zweitgutachter: Prof. Dr. U. Schumann

Tag der mündlichen Prüfung: 15.05.2013

Abstract

Newly developing convective clouds can be detected and monitored by satellite, but which of these clouds will grow to mature thunderstorms is difficult to predict from one data source alone. Within this thesis it is shown how the quality of satellite-based convection initiation (CI) detections can be raised substantially by the use of additional data sources which quantify available moisture, airmass instability, and lift for the analyzed clouds, the necessary ingredients for thunderstorms to develop.

Regions of interest for possible CI are detected by the Cb-TRAM¹ algorithm using 5 minute rapid scan satellite data. Cb-TRAM combines satellite channel data making it possible to distinguish newly developing (initiating), fast growing, and mature convective storms. Furthermore, these detections are extrapolated into the future, producing nowcasts for up to 60 minutes.

For evaluating the quality of the satellite-based CI detection and to quantify the achievable improvement by the use of additional data, a suitable verification method for these CI detections is very important. An object-based verification approach for these Cb-TRAM CI objects is introduced, which has been newly developed within this study. In order to derive sound statistics, the verification is performed over a whole summer period (May 15 - August 31, 2009) and for the whole Central European area. The CI detections can be categorized as developing (hits) and non-developing (false alarms). The verification results show a large amount of false alarms which has to be reduced in order to get more meaningful CI detections.

The possibility of individual Cb-TRAM CI detections to grow further is analyzed using additional data from surface observations and numerical weather prediction (NWP) model output, in order to gain the information on available moisture, instability, and lift for each CI detection object. This information is combined using fuzzy logic to obtain a so-called "CI forcing" value per object. Finally the CI forcing value is translated into a probability of further development to a thunderstorm for each cell.

Within this thesis the benefit of using multiple data sources to improve CI nowcasting is demonstrated. The additional information provided by the newly incorporated data raises the CI detection and nowcast quality by allowing, depending on the user-selectable amount of omitted hits (in the range of 0 - 25 %), a substantial reduction of the amount of false alarms (5 - 65 %). The methodology can also easily be adapted or extended for further additional data sources. An early identification of regions where mature storms will evolve allows for more adequate, user-oriented warnings.

¹Cumulonimbus Tracking And Monitoring

Kurzfassung

Die Entstehung neuer, konvektiver Wolken kann mit Hilfe von Satellitendaten beobachtet werden, aber eine Vorhersage, welche dieser Wolken sich zu ausgewachsenen Gewittern entwickeln, ist mit einer Datenquelle allein nur schwer möglich. In dieser Arbeit wird gezeigt, wie die satellitenbasierte Erkennung der Initiierung von Konvektion (convection initiation oder kurz CI) substantiell verbessert werden kann durch die Nutzung zusätzlicher Datenquellen, die die zur Verfügung stehende Feuchte, Instabilität und Hebung - also die notwendigen Zutaten für die Entstehung eines Gewitters - für die detektierten Wolken quantifizieren.

In fünf minütigen Satellitendaten werden durch den Cb-TRAM Algorithmus (Cumulonimbus Tracking And Monitoring) Gebiete mit wahrscheinlicher CI bestimmt. Cb-TRAM kombiniert Daten unterschiedlicher Satellitenkanäle um neu entstehende (CI), schnell wachsende und ausgewachsene konvektive Ereignisse zu identifizieren. Diese Erkennungen werden in die Zukunft extrapoliert um Nowcasts für bis zu 60 Minuten zu erzeugen.

Zur Beurteilung der Qualität der satellitenbasierten CI Erkennung und um die erreichbare Verbesserung durch die Verwendung zusätzlicher Datenquellen zu quantifizieren, ist eine sinnvolle Verifikationsmethode für diese CI Erkennungen von Nöten. Ein speziell für diese Cb-TRAM CI Erkennungen innerhalb dieser Arbeit entwickelter, neuer, objekt-basierter Ansatz wird hierzu eingeführt. Um verlässliche Statistiken zu erhalten wurde die Verifikation für eine ganze Sommerperiode (15. Mai 2009 - 31. August 2009) über ganz Mitteleuropa durchgeführt. Die CI Erkennungen können als sich weiterentwickelnde (Treffer bzw. hits) oder sich nicht weiterentwickelnde (Fehlalarme bzw. false alarms) klassifiziert werden. Die Verifikationsergebnisse zeigen eine hohe Anzahl an false alarms, welche reduziert werden muss um aussagekräftigere CI Erkennungen zu erhalten.

Mit Hilfe von Beobachtungsdaten und Daten eines numerischen Wettervorhersagemodells werden Informationen über verfügbare Feuchte, Instabilität und Hebung für jedes detektierte Objekt gewonnen um damit die Möglichkeit einer Weiterentwicklung einzelner Cb-TRAM CI Erkennungen zu analysieren. Diese Informationen werden mit Fuzzy Logik kombiniert um eine sogenannte "CI Forcierung" für jedes der Objekte zu erhalten. Abschließend wird die "CI Forcierung" in eine Weiterentwicklungswahrscheinlichkeit (in ein Gewitter) für das jeweilige Objekt übersetzt.

Im Rahmen dieser Doktorarbeit wird der Nutzen der Verwendung zusätzlicher Datenquellen zur Verbesserung des CI Nowcasting demonstriert. Die zusätzlichen Informationen durch die neu verwendeten Daten erhöhen die CI Erkennungs- und Nowcastqualität indem die Anzahl der Fehlalarme, abhängig vom Anteil

der "aussortierten" Treffer (vom Nutzer wählbar im Bereich zwischen 0 - 25 %), substantiell reduziert werden kann (5 - 65 %). Die Methode kann leicht angepasst oder auch erweitert werden um weitere zusätzliche Daten zu testen beziehungsweise zu verwenden. Eine frühzeitige Erkennung von Gebieten in denen sich Gewitter bilden werden, ermöglicht verbesserte, nutzerorientierte Warnungen.

Contents

1. Introduction	3
1.1. Motivation	3
1.2. Basic concept, aims & outline	5
2. Atmospheric Convection and Convection Initiation	11
2.1. Convection and Convection Initiation	11
2.2. Types of organization	15
2.3. Nowcasting of Convection	19
3. Methods and data sources	23
3.1. Cb-TRAM CI detection	23
3.1.1. Cb-TRAM overview	23
3.1.2. Detection stages	26
3.2. Verification	29
3.2.1. Verification methods	29
3.2.2. How to verify CI?	36
3.3. Additional data sources	37
3.3.1. Lightning data	39
3.3.2. Surface observation analysis	40
3.3.3. NWP model data	42
3.4. Fuzzy logic to combine different data sources	45
4. Analysis of the existing CI detection and nowcasting in Cb-TRAM	51
4.1. CI verification methods for Cb-TRAM	51
4.2. Performance of Cb-TRAM CI stage	56
5. Combination with additional data sources	59
5.1. Lightning data	60
5.2. Surface observations & model data	62
5.2.1. Indicating low-level moisture	62
5.2.2. Indicating instability	66
5.2.3. Indicating lift	69
5.3. Combination of indicators	74
6. Discussion	81
7. Conclusions and outlook	91
A. Data basis	95

B. VERA moisture flux	97
C. Fuzzy logic rulebase	99
D. Abbreviations and acronyms	101
E. Links	105
F. Additional figures	107
List of Figures	114
List of Tables	117
Bibliography	131

1. Introduction

1.1. Motivation

If you sit outside and watch the sky on a warm and sunny summer day you can often watch the typical cauliflower-like cumulus clouds starting to develop. At some point these cumulus humilis ("harmless") might start towering and develop to a more tremendous cumulus congestus ("massive"). Most of these cumulus clouds just collapse again and no storms develop. However, sometimes these clouds grow even further and develop to a thunderstorm cloud with an anvil clearly visible from afar, a so-called cumulonimbus cloud ("raining cumulus"). Here the curiosity of many atmospheric physicists is raised. What is the difference in the atmospheric conditions and why did the cloud further develop at one point or why not? Therefore the atmospheric conditions in and around the storm are quite interesting research topics. In general, thunderstorms are among the most impressive and exciting phenomena to observe in atmospheric physics. The beauty of the tremendous cloud formations or the nocturnal view of lightning fascinates many people. They also have for example important impact on the hydrological cycle (Wulfmeyer et al., 2011), their vertical transport of trace-gases (Fischer et al., 2003) and their production of lightning-induced NO_x (Schumann & Huntrieser, 2007) influences the atmospheric chemistry in the UTLS (Upper Troposphere Lower Stratosphere) region, and the forecast of a thunderstorm with all its related attributes is still quite difficult (Wulfmeyer et al., 2011). Besides these characteristics thunderstorms use to frighten many people and they are the cause for high weather-related risks. Common attributes of a thunderstorm like hail, heavy precipitation, wind gusts, tornadoes, turbulence, and lightning cause a lot of damage. This leads to a total damage estimate for Europe of 5 to 8 billion Euro per year (Dotzek & Forster, 2011).

Thus the reduction of these risks and the damage caused by thunderstorms is one of the main reasons for further research on thunderstorms. Weather forecasts for thunderstorms are sometimes based on very short range forecasting (forecast for up to 12 hours), but mostly thunderstorm forecasts are based on so called nowcasting, with a time horizon of zero to two hours. The time horizons for these forecast ranges fluctuate for different studies. The values above are selected as defined by the World Meteorological Organization (WMO-No. 485, 2012). Nowcasting often focusses on observational data (e.g. from radar or satellites) and its extrapolation for the upcoming nowcasting time steps. A second possibility to forecast thunderstorms is to use numerical weather prediction (NWP) models. For a thunderstorm forecast they have to be capable of resolving convective events which means, they need high horizontal and vertical resolutions, and

a very good understanding and numerical realization of the ongoing physics. The main problem with convective forecasting in NWP models is to get the position, timing, and amount of points where convection really starts, thus the convective initiation, represented right. "Convective initiation" (CI) names the phase at the beginning of a thunderstorm life-cycle where a convective cloud starts to grow.

The aim of nowcasting is to warn as precisely in space and time as possible. Thunderstorm warnings can be produced for regions, events, or companies, informing if a risk has to be expected. Industrial sectors highly interested on reliable storm warnings are for example energy suppliers or the whole air transportation sector as well as organizers of any kind of open-air concerts, sporting events or festivities. An important fact about nowcasting is the lead time with reasonable nowcast quality that can be achieved. Lead time describes the time between the prediction of a storm hitting a special area at a point in time in the future, and the real occurrence of the storm there at, or at least close to, the predicted point in time.

There are many nowcasting tools that generate thunderstorm nowcasts as will be explained in Chapter 2 but most of them focus on fully developed storms, so called mature storms and some, later on in a storms life-cycle, on their decay. One idea to raise the achievable lead time is to try to detect a storm earlier within its life-cycle in the observation data used for the nowcasts. Detecting a storm in the CI stage of its life-cycle might on the one hand help to gain more lead time for some warnings and it might on the other hand help to warn on phenomena which already exist within this early stage and may be "overlooked" by later detections. As an example, showing the need for improved CI detection and nowcasting, I want to focus on nowcasting applications for aviation here.

Thunderstorms heavily disturb all kinds of in-flight and ground aviation operations. A Eurocontrol Performance Review Commission (2011) study shows that thunderstorms are responsible for 30.9 % of the weather related delays in European air traffic flow management, which makes thunderstorms the weather phenomena with the largest impact on the air traffic over Europe. Leighton (2006) states that thunderstorm activity is the reason for up to 90 % of all delays in the airspace over the USA during the summer months. An early detection of a developing storm with high lead time for the first warning might for example help an airport and the air-traffic control to react on a storm arriving at an airport. If the lead time before an estimated time of arrival is sufficient some flights from the airport later affected by the storm might be rescheduled to get the airplanes out of the airport before it is affected. Or flights to the affected airport might be rescheduled to keep planes on the ground which would not be able to land at their destination until the airport is not affected anymore. The nowcasts also help to decide if it makes sense that a plane waits, circling an aerodrome, or if it has to be redirected. Besides the advantages for the planning, a detection of initiating storm cells can also be useful to warn in areas where turbulence due to newly growing convective clouds may appear. Convectively Induced Turbulence (CIT) is responsible for over 60% of turbulence-related aircraft accidents and nearly 23% of these CIT-related accidents resulted in fatal injuries to the occupants of the aircraft (Bedka et al., 2010).

The use of nowcasting in the aviation sector is quite multifaceted. There is a multitude of tools tested for a use in aviation, especially in the US. Usually flight safety agencies, airports, airliners, and pilots get information and forecasts from national weather services. However the main problem is that more or less everybody has his own tool or method or data provider he looks at and sometimes trusts more. The key problem is the multitude of different information used. It would be quite useful to get the same easy understandable information about the current situation and its nowcast, without any room for eventually wrong interpretations, to all persons who are involved in a decision making processes in aviation. This could lead to a use of nowcasting which might help to realize better strategic planning on how to operate as well as possible despite adverse weather conditions, and not just reacting rapidly on them.

The advantages of earlier, reliable detections and warnings are one big motivation for CI research. A particular challenge of the CI detection and nowcasting is, that the detection of a new developing cloud, for example with the help of satellite data, might be correct due to some serious cloud development, but if the cloud collapses again within the next time steps the CI nowcast will be wrong. If the detection is based only on satellite data, for example, the cloud is observed only from above. The horizontal growth of the cloud and the cooling of the cloud top can for example be evaluated but it is difficult to evaluate the atmosphere below the cloud. The alternative use of NWP model data might show realistic storm structures also in lower atmospheric levels than observed by the satellite data, but often with errors in positioning or timing of the CI in comparison to reality. Surface observations or instability information by atmospheric soundings, usually have problems with the resolution of the measurements in space and time and contain none or only little upper air information. Thus they have deficiencies in resembling the atmospheric reality as well. Thus many data sources have advantages and drawbacks for their use in conduction with CI detection and nowcasting. Within this thesis some different data sources which are expected to be useful for operational (near-)realtime CI detection are tested and the most promising are combined to overcome the possible weaknesses of using only one data source with the aim of a better CI nowcasting.

1.2. Basic concept, aims & outline

The key scientific question that should be answered as an aim of this thesis is:

- How much can CI nowcasting be improved by the use of multiple data sources?

Within the approach shown here three crucial questions arise on the journey to this key point:

- How can CI nowcast quality be evaluated?
- What kind of additional data is most promising?

and

- How should the data be combined?

The conceptional idea of this thesis can be described in three steps. The first step is to find an existing tool expected to have some skill for CI detection and nowcasting and to develop a verification method suitable for its evaluation. The second step is to combine the output of the existing tool with information of additional data sources. The third and final step is to repeat the verification for the combined information to find the most useful of the tested data combinations (Fig. 1.1).

The search for an existing tool with skill for CI detection and nowcasting leads to the question which data is most useful for a first detection of CI cells. Radar is certainly useful to monitor areas of existing convective development, "but most operational radar networks are currently not set up to adequately monitor vertical cloud growth prior to the development of significantly sizable hydrometeors" (Walker et al., 2012). Thus CI detection with the help of satellite data is achievable earlier in the development of the storm than the detection of a signal in the radar data that can be identified as a proof for CI. The higher lead time is the reason for preferring satellite and not radar information when choosing the "basic tool".

The basic tool used here is the Cb-TRAM algorithm (Zinner et al., 2008; Zinner & Betz, 2009), a state-of-the-art satellite data based thunderstorm detection and nowcasting tool which was developed at the Institute of Atmospheric Physics of the German Aerospace Center (DLR). The acronym Cb-TRAM stands for Cumulonimbus TRacking And Monitoring. The tool will be explained in Chapter 3.1.

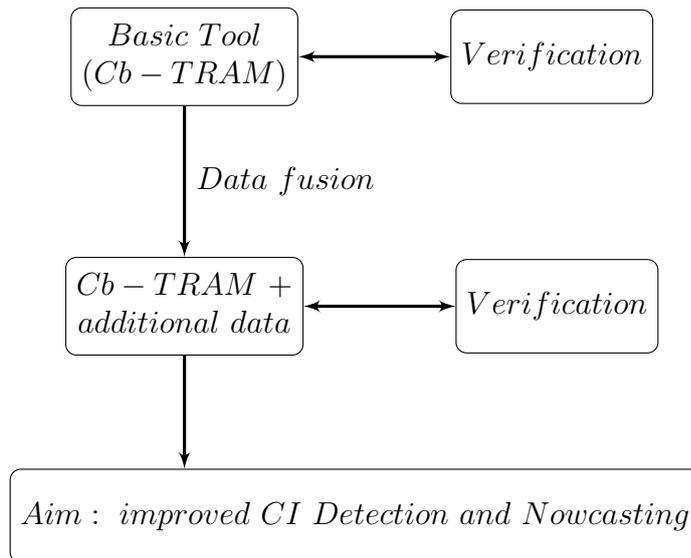


Figure 1.1.: Outline of the basic concept of this thesis (Stich et al., 2011). For further description see text.

One innovative key aspect of this thesis is the verification of the CI nowcasts which is performed not only for few test cases but for a whole summer period from May

15 2009 until August 31 2009 including many different synoptic situations over Central Europe. The crucial part here is that, more or less no well-established and reasonable way for the verification of CI events was developed until now. The key problem is to find something that can be used as proxy for CI. The data used as proof for CI occurrence usually describes convective parameters like the first reflectivity signal greater than 35 dBZ in radar data (Roberts & Rutledge, 2003; Mecikalski & Bedka, 2006) or something like that. The approaches usually can be described by a statement like, "if I observe this result, CI must have happened before". The proof is later in time than the CI process itself, thus you have to expect that you can not seriously use the classical verification ideas where you compare your forecast and the observed data pixel by pixel. In the same way all the non-pixelbased, highly sophisticated, verification ideas for precipitation where objects or amplitudes are compared are developed for comparison of two situations (forecast and observation) at one point in time. All of these verification approaches do not try to relate some forecast of an event (CI) with a later on existing observation being a consequence of the event (e.g. lightning, 35 dBZ radar signal, etc.). Usually verification approaches try to compare one situation, the event or the consequence, in both, forecast and observation data. This is the crucial difference between many verification tasks and the CI verification problem. Thus the choice of an appropriate verification method is worth to spent some time and thoughts on.

The aims of the combination with additional data sources are to get better verification results and more reliable nowcasts than without the additional data. In the satellite image the first development of cumulus clouds can be detected and the additional data should add the information if the environment is favorable for a further development of a thunderstorm or if the forcing is not sufficient. Thus the additional information should help to reduce false alarms for "non-CI-favorable" environments. Furthermore it should help to retrieve a kind of CI probability for cells in, at least to some degree, "CI-favorable" environments. The combination is performed by a fuzzy logic data fusion approach (for explanation see Chapter 3.4) to incorporate meteorological expert knowledge in the combination and to avoid to rely on crisp thresholds for the combination of different data sources. The use of fuzzy logic for meteorological applications in combining datasets and applying conceptual models is summarized for example in Mueller et al. (2003). A fuzzy logic combination is capable to distinguish between "non-CI-favorable" and "CI-favorable" environments as will be shown later. Fuzzy logic deals with "fuzzy" or imprecise reasoning. Instead of incorporating fixed thresholds for binary decisions if something is "true" or "false", fuzzy logic can handle the concept of partial truth. Thus it is capable to translate human reasoning into mathematic decisions in a more appropriate way than binary logic. A fuzzy logic system is a way to use expert knowledge about a process to build a kind of decision support tool.

Concerning the additional data the focus of this study is on two issues: First of all the data should be regularly available and not any kind of special measurements only available for test cases, a measurement campaign, or whatever. This is due to the fact that the CI detection and nowcasting method should get

operational. The second point is that the testing of the additional material does not focus on test cases where in few examples the combination works out well, but on a statistical approach where the whole aforementioned summer period in 2009 is used to evaluate the abilities of a combination. In Chapter 5 not only the results of the best performing combination are shown but also some information on tested data which did not help to significantly increase the abilities of the CI detection.

It is a particular concern within this study not to come up with a new kind of "magic number", like many thunderstorm studies before, which performs under certain conditions eventually better for forecasting new storm development than the showalter index, lifted index, total totals or whatever of the almost unlimited amount of different convective indices you want to compare. Summaries on many different of these indices and their functionality as predictors over parts of Europe can be found for example in Huntrieser et al. (1996), Haklander & Van Delden (2003), Kunz (2007), or Kaltenböck et al. (2009). In addition, the combination of different input fields into an index and using only the index number out of a black box is something many forecasters understandably regard with a lot of distrust. In Doswell & Schultz (2006) the authors "assert that no single number can replace the value of a forecaster simply looking at the soundings, as well as looking at diverse diagnostic variable computations based on those soundings". The purpose of the data combination for the CI detection is, in line with the argumentation in Doswell & Schultz (2006), to "add value beyond what a hypothetical forecaster can see simply by viewing the data" and to raise situational awareness of a forecaster. Moreover the basic idea of the CI detection and now-casting within Cb-TRAM is not to use it especially as tool for forecasters but to inform decision makers with an easy understandable visualization about the situation. In relation to air traffic the aim would be that all involved parties, the air traffic control, the airports, airlines, pilots, etc. should get the same information of the current weather conditions as an understandable base for their Collaborative Decision Making (CDM). Thus the output of the CI detection with additional data sources is aimed to be a useful stand-alone product for CDM as well.

Another important point of the approach within this thesis is that the additional data is not evaluated or interpreted for the whole analyzed, Central European, area at once. The CI detection of Cb-TRAM marks polygons around clouds that are expected to develop to a storm and all the additional fields tested are evaluated for each of the polygons. Furthermore it is quite easy to extend the testing with the methodology developed here and add an additional field which is not yet incorporated. This object-based framework has significant advantages. The full information of satellite-derived properties of the observed cloud, along with NWP fields, surface observations, lightning data, radar data, etc. can be carried along with the object for its full lifetime. This also establishes the future possibility to evaluate trends within the additional fields. I expect that the possibilities of such an object-based framework are worth to stick to these objects detected by the satellite CI detection. An alternative would have been to combine all these fields for the whole investigated area and to use the satellite detection

just as one field of many. This would have been a possible way if the focus is more on a probabilistic forecast of CI occurrence. However this would quite sure lead to a loss of small scale objects and lead to larger areas of more or less high potential for CI.

In summary, the unique features of the thesis are:

- a newly developed verification method for CI nowcasts.
- the analysis of the verification method for Cb-TRAM CI nowcasts for a full convective season over Central Europe (not just a few test cases).
- an object-based framework for the analysis of additional data within polygons/objects detected by Cb-TRAM.
- the fuzzy logic based data combination as Cb-TRAM post-processing.
- an improved CI detection and thus nowcasting with Cb-TRAM.

The outline of the thesis is as follows: In Chapter 2 some basic information about CI, thunderstorms and the nowcasting of both are given. Chapter 3 includes a short overview on the different methods and data sources used within this study. Where Section 3.1 describes Cb-TRAM, Section 3.2 is on some aspects on state-of-the-art verification methods and their "difficulties" with CI, Section 3.3 gives information on the used additional data sources (lightning measurements, surface observations, and NWP model data) and Section 3.4 is a short introduction into the basics of fuzzy logic as it will be used later to combine the different data sources. In Chapter 4 the verification method used for the Cb-TRAM CI detection verification will be described in more detail and the results of the analysis of the existing CI detection and nowcasting in Cb-TRAM will be presented. The following Chapter 5 introduces the potential improvement of the CI detection and nowcasting by adding the different additional data sources and points out the results of the final combination. Followed by the discussion of the results in Chapter 6 and finally conclusions and an outlook are presented out in Chapter 7.

2. Atmospheric Convection and Convection Initiation

2.1. Convection and Convection Initiation

In general, the term convection applies to the transport of some property (e.g. heat or mass) by fluid movement. Thus it is, for example, besides radiation and conduction the third main process by which heat is transported. In meteorology convection is usually associated with the vertical and buoyant component of the flow. Nonbuoyant atmospheric flows which transport a property are called advection. Advection can describe either horizontal or vertical transport. Atmospheric convection is a vast field which is treated for example in the monograph of Emanuel (1994).

In the following the focus will be on different forms of Deep Moist Convection (DMC). DMC can be used synonymous with the word thunderstorm. A thunderstorm is defined as such if lightning and therewith thunder occurs within a storm produced by a cumulonimbus cloud (Cb). Thunderstorms develop whenever and wherever the ingredients for their formation come together: instability, moisture, and lift (Zinner & Groenemeijer, 2012). The instability of an air parcel is determined by the temperature and moisture of both, the environmental air and the parcel itself. The lift serves as a trigger mechanism. A simplified description of the behavior is given by the lifted parcel theory following Bjerknes (1938), which is discussed in more detail in Manzato & Morgan (2003). If a parcel in an unstable environment experiences some lift by a flow over a topographic barrier or because it is warmer than the surrounding air it starts to rise until it reaches its Equilibrium Level (EL). The level from which a parcel of air can rise due to its temperature difference to the environment without any further lift or triggering needed is called Level of Free Convection (LFC). During the rise the parcel might get saturated and at this point the (deep moist) convection initiates: condensation and thus the growth of a cumulus cloud starts. This lower boundary of the cloud is, due to this mechanism, called Lifted Condensation Level (LCL). The condensation releases latent heat to the parcel which helps to keep up the temperature difference between the rising air and its environment. The latent heat release decreases the cooling of the parcel, the so-called lapse rate, from the dry adiabatic lapse rate of -9.8 K/km to a moist adiabatic lapse rate. The moist adiabatic lapse rate varies with the temperature of the parcel and ranges between -4 K/km and the dry adiabatic value. Due to this lowering of the lapse rate it is not uncommon for the atmosphere to be stable when dry, and unstable when saturated (Durran & Klemp, 1982). This behavior is described as conditional

instability.

Theoretical concepts on instability name conditional, latent, and potential instability. A layer is conditionally unstable when the environmental lapse rate curve is between the dry and the moist adiabatic lapse rate curves on a thermodynamic diagram. Thus the moisture within the layer conditions the instability. Latent instability describes the state when the actual lapse rate above the LFC is lower than the moist-adiabatic lapse rate. There, conditional instability can be caused by a moist air parcel rising from below the LFC in the relatively dry environment (Kunz, 2007). Last but not least, potential instability describes an unsaturated column of air where the equivalent potential temperature θ_e decreases with height (Emanuel, 1994). If such a column is lifted until it is fully saturated, it will become unstable insensitive to its initial stratification. After Kunz (2007) this kind of instability is described by the KO-Index which will be used in Chapter 5.

In many studies on thunderstorms moisture and instability are evaluated by looking at the Convective Available Potential Energy (CAPE). The value of CAPE shall give the information if the moisture content of the air is high enough that conditional instability actually contains the potential for parcels to become buoyant (Doswell, 2001). Predicting convection initiation (CI) gets more difficult due to the fact that the presence of CAPE is not a sufficient condition for CI. Air typically requires triggering, thus some forced ascent is needed to reach the LFC. It is important to quantify this lift, the third of the earlier mentioned ingredients, and introduce a measure of the energy a parcel starting from the surface needs to reach the LFC. This quantity is called "convective inhibition" (CIN). Following Markowski & Richardson (2010) CIN and CAPE are calculated as in Eq. 2.1 and Eq. 2.2.

$$CIN = - \int_0^{LFC} B dz \approx -g \int_{SFC}^{LFC} \frac{T'_v}{\bar{T}_v} dz \quad (2.1)$$

$$CAPE = \int_{LFC}^{EL} B dz \approx g \int_{LFC}^{EL} \frac{T'_v}{\bar{T}_v} dz \quad (2.2)$$

where

$z = 0$ at the surface (SFC),

g = acceleration of gravity,

and the buoyancy B is expressed in terms of the virtual temperature perturbation of the lifted parcel (T'_v) relative to the virtual temperature of the environment (\bar{T}_v), with a virtual temperature of the lifted parcel written as $T_v = \bar{T}_v + T'_v$.

T_v is the temperature that dry air would need to attain in order to have the same density as the moist air at the same pressure (Wallace & Hobbs, 2006). The size of the correction climbs for warm and/or moist conditions to the order of 10 % (for example for 30 degrees Celsius and 50 % relative humidity in 900 hPa the correction is almost exactly 3 degrees.)

As an example for typical severe storm environments in the U.S. Markowski & Richardson (2010) state that CAPE values $\leq 1000 J kg^{-1}$ are considered small and large values are $\geq 2500 J kg^{-1}$, whereas large CIN values are $\geq 50 J kg^{-1}$ and values $\leq 10 J kg^{-1}$ are considered small. In Europe "high values of CAPE

are much more unlikely than in North America" (Brooks et al., 2003). European storm environments often show similarities to conditions observed in the southeastern U.S. during the cool season (Brooks, 2009) with low LCL heights and moderate CAPE. Frequently synoptic forcing and local (e.g. orographic) influences are crucial for thunderstorm initiation (Kaltenböck et al., 2009). CAPE values above 2000 J kg^{-1} are rare and occur primarily at the Mediterranean and Black Seas (Romero et al., 2007). CAPE and CIN frequently are computed using temperature instead of virtual temperature which leads to significant differences in the resulting values. Especially for the in comparison to CAPE relatively small values of CIN this difference due to the calculation might be the difference between a more or less insurmountable high amount of CIN (which, as the name says, inhibits the initiation of convection) or practically no CIN at all (making convection initiation very likely). By including the effects of moisture for the profiles CAPE values are usually increased and CIN values are typically decreased Markowski & Richardson (2010).

For a better understanding of CAPE, CIN, LCL, LFC, and EL, a visualization with the help of a skew-T log p diagram of a radio sounding is shown in Fig. 2.1. Within the Figure the dew-point temperature and the temperature of a sounding are used.

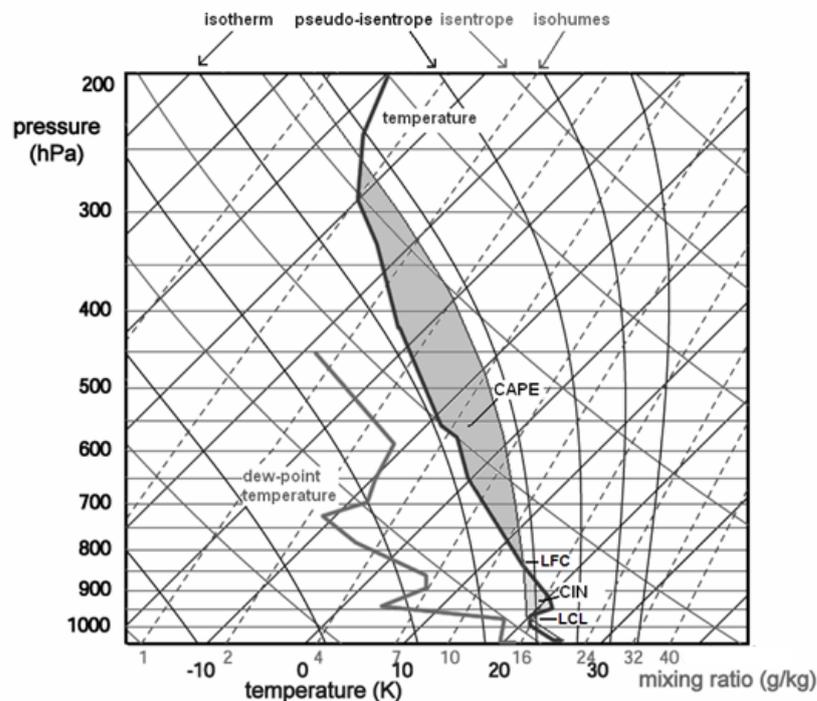


Figure 2.1.: Example of a skew-T log p diagram. The temperature and dew point temperature data from a radiosonde ascent are plotted. The diagram shows isobars, isotherms, isentropes (lines of equal potential temperature or dry adiabats), pseudo-isentropes (lines of equal pseudo-equivalent potential temperature or moist adiabats), and isohumes (lines of equal mixing ratio). Further descriptions see text. Diagram taken from <http://www.estofex.org/guide/>

Use of the virtual temperature would cause a slight right shift of the temperature curve. Following the line of constant mixing ratio from the dew-point temperature at the surface and the dry adiabat from the temperature at the surface we find the LCL (Lifted Condensation Level) at their intersection. A parcel from the surface would follow the dry adiabat until it reaches the LCL. Afterwards a rising parcel would follow the moist adiabat. If the parcel is colder than the environmental temperature (at the left of the temperature curve) for some height interval it needs energy to rise further. Thus this area between moist adiabat and environmental temperature curve, above the LCL and below the LFC (Level of Free Convection), defines the CIN. The LFC is defined as the height where the lifted parcel becomes warmer than the environmental temperature. Afterwards the parcel is warmer than its environment and can rise without further triggering until it reaches the EL (Equilibrium Level). This level is also often called Level of Neutral Buoyancy (LNB) and is usually close to or at the tropopause (Zinner & Groenemeijer, 2012). The lines of the adiabat and the environmental temperature between the heights of the LFC and EL limit the area which is defined as CAPE.

As mentioned already earlier, the presence of CAPE is not a sufficient condition for CI. Neither is the absence of CIN a sufficient condition for CI. It seems to be more promising to look at quantities which describe the three ingredients (instability, moisture, and lift) more directly. CAPE and CIN values which are used as kind of "forecast thresholds" for CI also heavily depend on the region you look at. As already mentioned, the values of CAPE in the central U.S. tend to be a lot higher than for most parts of Europe.

Zimmer et al. (2011) state that "the initiation and lifecycle of a convective cloud is directly a result of processes local to the cloud itself: the conditional instability of the column, the absence of a capping inversion or other inhibiting factors, and the boundary layer variability that can trigger an updraft". Thus background flow which causes a lift heavily increases the CI probability. Following Zinner & Groenemeijer (2012), causes for growing CAPE, or decreasing CIN, and thus for CI can be, for example, increasing surface moisture, or surface warming during the course of the day. Besides these surface driven processes, mid-tropospheric cold air advection or infrared radiation has similar destabilizing effects. In mid-latitudes destabilization and thus a growing thunderstorm potential is commonly observed ahead of an approaching front where colder air arrives at elevated levels ahead of the surface front due to surface friction. CI usually takes place along air mass boundaries (like synoptic fronts, drylines, or sea breezes) or in conjunction with orographic circulations driven by differential heating of sloped or elevated terrain, and by forced lifting of atmosphere layers. CI might as well be caused by synoptic scale forcing, which results in upper tropospheric divergence and related large-scale lifting, as can be observed in the area of the jet stream maxima in mid-latitudes (Zinner & Groenemeijer, 2012; Georgiev & Santurette, 2010). The lift to overcome potentially existing CIN and to get an air mass to or above its LFC is crucial for CI, independent of the process which causes the lift.

If CI takes place for example somewhere along an air mass boundary the probability for further CI close to this event decreases heavily. The one successful CI case uses the potential instability of the air mass and causes a downward flow

around the developing column of air which both suppresses other developments nearby.

In summary CI needs sufficient initial lift of an sufficiently unstable and sufficiently moist airmass. Further details on the thermodynamics of a thunderstorm can be found for example in the afore mentioned monographs of Emanuel (1994), Markowski & Richardson (2010), Doswell (2001), or in Houze (1993).

2.2. Types of organization

As background knowledge on the development, lifetime, and especially lifecycle of thunderstorms this Section gives a short basic overview on different types of thunderstorms. Thunderstorms can be classified into three simplified, basic types (Hagen et al., 1999). These types are:

- Single cell
- Multicell
- Supercell

If these thunderstorm types start to build up larger groups these are called Mesoscale Convective Systems (MCSs). A MCS might be a squall line, a Mesoscale Convective Complex (MCC) or even a tropical cyclone. In the following the three simplified types are described briefly.

All thunderstorm types go through the same stages within their life cycle. After CI takes place three main stages are distinguished: the developing stage, the mature stage, and the dissipation stage.

These stages can be shown quite nicely with the example of an isolated *single cell* storm (see Fig. 2.2). The developing stage describes a rapid vertical growth of the initialized cumulus cloud and the updraft strength increases. The vertical growth of the cloud is stopped at its EL which is usually close to the tropopause. Then the lifted air is forced to spread laterally which leads to the development of the characteristic anvil shape of a cumulonimbus cloud (Cb). If a storm has a quite vigorous updraft it may even penetrate to higher levels, then we see a short lived so-called overshooting top above the anvil (Bedka et al., 2010). During the growth of the cloud the water droplets within get larger and many freeze and become ice particles. The strength of the updraft inside the storm cloud determines how big the particles might get before they start to fall out of the cloud and the storm begins to precipitate. If the updraft is strong enough the frozen particles can get large and might not melt until they reach the surface, then they fall as hail. For details on cloud microphysics and the formation of different precipitation types like rain, graupel, or hail, see, for example, Tessendorf et al. (2005) or the monograph of Houze (1993). In the part of the cloud where the precipitation is situated the evaporational cooling due to the falling rain is responsible for the creation of a downdraft. This coexistence of an updraft and a downdraft area in and below the stormcloud defines the mature stage. If there is no or only little wind shear, as it is typical for single cell storms, the thunderstorm

reaches its dissipating stage quite fast. The dissipating stage will be reached after approximately 25 – 30 minutes under low shear conditions. Then the downdraft due to the precipitation spreads over the whole cloud. The storm kind of "rains out" and thus the storm dissipates. The cold downdraft below the dissipating storm cloud spreads radially when it reaches the ground and cuts off the potential for moist and warm inflow in the storm. If this phenomenon occurs in a quite vigorous way, it is called downburst (Byers & Braham, 1949; Fujita & Wakimoto, 1981) and is for example a potential threat for aircraft due to rapid changes of wind direction and wind speed. The aircraft has to face a strong headwind, then it is within the downdraft and finally it is exposed to a strong tailwind. The overall lifetime of such an isolated single cell is ~ 1 h.

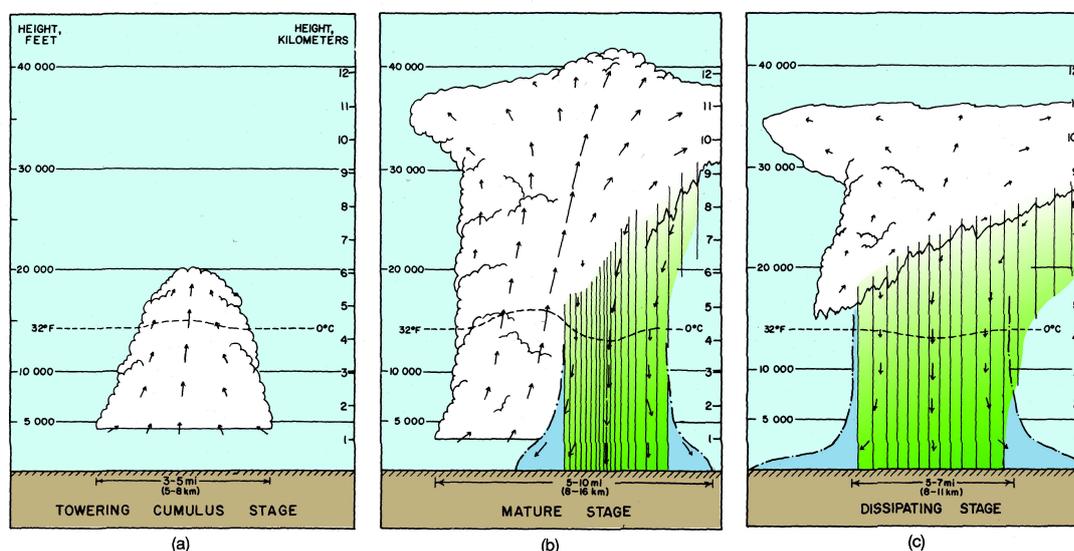


Figure 2.2.: Different stages within the life cycle of a thunderstorm. Taken from Markowski & Richardson (2010). Adapted from Byers & Braham (1949) and Doswell (1985).

If the outflow of the storm is able to trigger a new development, we talk of a *multicell* storm. The air in the outflow from convective downdrafts is cold and builds a density current along the earth's surface. The density current forces the warmer surrounding air to ascend. Thus, these so-called cold pools are responsible for the generation of a storm cell in multicell storms as a new source of lift. While the cold pool spreads out over a large area and becomes shallow, the lift may no longer be sufficient to cause air parcels to reach their LFC, then the growth of new cells will stop (Doswell, 2001, Chapter 3). Those multicell storms can last for a few hours. The different stages of development observed in the life cycle of a single cell coexist with each one forming a part of the same storm system. A multicell storm can be handled as an aggregate of cells. The pattern of cells within the multicell thunderstorm is continually changing (Houze, 1993). One important additional ingredient for the development of multicells is an environment with moderate vertical wind shear. The cold pool and the environmental vertical wind shear induce horizontal vorticity which helps at one flank of the cold pool to

generate a more vigorous lift which then leads to a better chance of an updraft reaching the LFC. The structure of a multicell is displayed in Fig. 2.3a).

Storms having strong updraft rotation are known as *supercells*. Large vertical wind shear is conducive to supercell development. The wind shear separates up- and downdraft locations within the storm. The tilting of horizontal vorticity leads to storm rotation. In addition the dynamic vertical pressure gradients that accompany large-shear environments can also enhance updrafts at altitudes high above the gust front (Markowski & Richardson, 2010). The presence of a mesocyclone within the updraft is the major difference between supercells and nonsupercells. Supercells generally have larger updrafts than ordinary storms and are responsible for an immoderate proportion of severe weather reports due to large hail or tornadoes. Their lifetime exceeds the lifetime of multicells and can be up to 10 h. An schematic is shown in Fig. 2.3b).

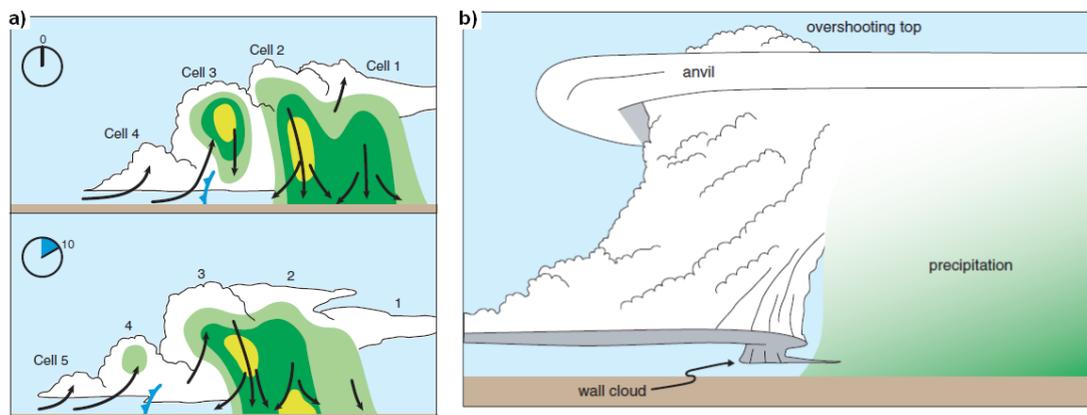


Figure 2.3.: The development of a multicell storm with time is displayed in a). b) shows the typical structure of a supercell with its rotating updraft and an overshooting top which reaches the lower stratosphere. Taken from Markowski & Richardson (2010).

Storms that produce a lot of weather related hazards are usually called "severe" convective storms. The likelihood of severe weather tends to increase with the degree of organization of the convection (Doswell, 2001). In the U.S. "severe" usually means a minimum hail size, damaging convective winds, and the development of tornadoes. The minimum hail sizes and wind speeds differ in the literature. In the U.S. the definitions of the Storm Prediction Center (SPC) of NOAA's National Weather Service¹ are commonly used:

- Hail at least 1 inch (= 2.54 cm) in diameter or larger, and/or
- wind gusts to 58 mph (= 93.34 $km h^{-1}$) or greater, and/or
- tornado occurrence

Nowadays heavy precipitation with the risk of flash floods is quite often added to this list of definitions. In Europe slightly different values (due to the use of

¹<http://www.spc.noaa.gov/>

SI units) are introduced for example by the European Severe Storms Laboratory (ESSL)² who operates the European Severe Weather Database (ESWD)³. After Dotzek et al. (2009) the diameter of large hail in the ESWD database is > 2 cm and the term severe convective wind is used for wind speeds above 25 m s^{-1} ($= 90 \text{ km h}^{-1}$). Albeit lightning is a big hazard accompanying thunderstorms, lightning can't be added as proof of a severe storm because than every thunderstorm would be severe by definition. In some cases flash rates or lightning density are used as a measure for severity (Williams et al., 1999; Lang et al., 2004; Betz et al., 2008). Besides large hail, wind gusts, and tornadoes, the ESWD database collects information about damaging lightning, heavy rain, funnel clouds, gustnados, dust devils and some hazards connected with wintery weather.

In a nutshell, the main difference between these three types of storms is the available vertical wind shear within the environmental air (as in Fig. 2.4). The originally satellite based CI detection examined within my study does not focus on CI of multi- or supercells but on CI of all types of DMC. While instability, moisture, and lift are important ingredients for every deep moist convective storm type the wind shear is only important for multicell or supercell development. There is no focus on storms which are expected to be extremely severe or anything like this. That is the reason, why wind shear is not further examined later within this study.

For a more detailed description of the different types or the above mentioned organization of storms as MCS I strongly recommend again to have a look at Markowski & Richardson (2010), Doswell (2001), or Houze (1993).

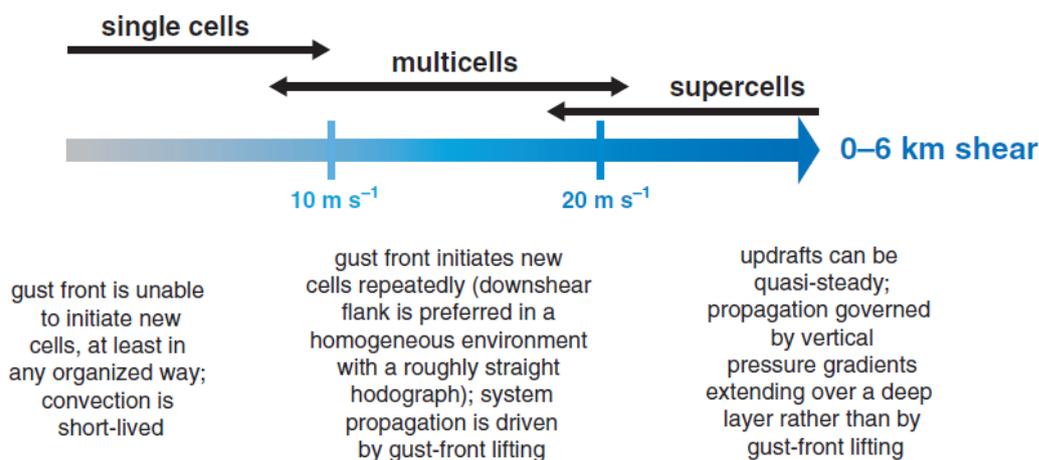


Figure 2.4.: Different amounts of wind shear are conducive for different storm types. For more information see text. Taken from Markowski & Richardson (2010).

²<http://www.essl.org/>

³<http://www.eswd.eu/>

2.3. Nowcasting of Convection

The term "nowcasting" describes weather forecasting in the very short range of the next 0 - 2 hours. Many nowcasting tools even focus only on the first hour. Historically, nowcasting evolved already almost 50 years ago. The first short-term predictions of the movement of convection by temporal extrapolation of meteorological radar imagery is introduced in Wilson (1966) and around ten years later satellite imagery is used for this purpose (Purdom, 1976) too. The majority of the first generation satellite-based nowcasting methods focused on nowcasting attempts for CI and storm development (Mass, 2012). The radar-based Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN) is not only capable of tracking individual cells, but also allows temporal changes of the cell structure, in terms of size or intensity (Dixon & Wiener, 1993). In the late 1990s the UK Met Office introduced powerful tools for nowcasts of heavy precipitation. These tools are based on radar tracking and temporal extrapolation of convection: NIMROD (Nowcasting and Initialization for Modeling using Regional Observation Data) combines radar data, satellite data and NWP output for precipitation nowcasts of up to 6 hours (Golding, 1998) furthermore GANDOLF (Generating Advanced Nowcasts for Deployment in Operational Land-based flood Forecasts) is a nowcasting and warning system using the same data sources with the aim of reducing hazards due to convection-induced flash flooding (Pierce et al., 2000). At the same time the German weather service DWD (Deutscher Wetterdienst) introduces the warning and cell tracking tool CONRAD (Lang, 2001). CONRAD (CONvection in RADar) identifies, evaluates, and tracks convective cells, and produces local warnings for severe rainfall, wind, and hail (Hafner et al., 2004) based on the data from the DWD radar network. Within the DWD project RADVOR-OP different radar and model based methods for extrapolation were added to further improve the nowcast quality and to extend the nowcasts lead time from 1 hour (CONRAD) to 2 hours (Bartels et al., 2005). RADVOR-OP is a German acronym about radar-based (near) real-time very short-range forecasting for operational purposes. Within the project the radar data was used for data assimilation in the non-hydrostatic limited area model COSMO-DE to improve the model quality (Klink et al., 2004) and thus to make the model winds more realistic and more useful for extrapolation purposes (Winterrath et al., 2007).

Little more than 30 years after his radar nowcasting paper Wilson et al. (1998) gave "a status report" on the nowcasting of thunderstorms. Since then a multitude of other nowcasting tools was developed as can be seen for example from the short summaries in Dance et al. (2010), Mass (2012), or Ruzanski et al. (2011). During the 2008 Summer Olympic Games, the World Meteorological Organization (WMO) World Weather Research Program (WWRP), organized an intercomparison campaign for nowcasting tools conducted in Beijing, China. This demonstration is called the Beijing 2008 Forecast Demonstration Project (Wilson et al., 2010). A similar Forecast Demonstration Project was already conducted in Sydney, Australia, during the 2000 Sydney Summer Olympics Games (Keenan et al., 2003; Wilson et al., 2004) combined with the intercomparison of multiple verification methods (Ebert et al., 2004; Casati et al., 2008). In re-

cent developments a multitude of data sources is used. NCAR uses a fuzzy logic combination of radar, satellite, upper air, surface, and NWP data within its Auto-Nowcaster (ANC) as described by Mueller et al. (2003) and Saxen et al. (2008). ANC also has a forecaster-interactive capability which allows a forecaster to enter locations of surface convergence boundaries manually (Roberts et al., 2012). Sometimes orography is included in the nowcasting system to improve spatial interpolation as in the INCA (Integrated Nowcasting through Comprehensive Analysis) system of the Austrian ZAMG (Steinheimer & Haiden, 2007; Haiden et al., 2011). Model Output Statistics (MOS) after Glahn & Lowry (1972) are also included for nowcasting like in the DWD routine CellMOS (Hoffmann, 2008; Wapler et al., 2012), the Czech Statistical Advection Model SAM (Sokol & Pesice, 2012) or the Localized Aviation MOS Product short LAMP (Ghirardelli & Glahn, 2010; Rudack & Ghirardelli, 2010).

Besides the increased amount of incorporated data sources also new nowcasting methodologies were and are developed. In Rivolta et al. (2006) three techniques are introduced as "conventional" methods for the example of "satellite nowcasting of infrared radiance fields": persistence method, steady-state method and linear method. Persistence assumes that each pixel in a forecasted satellite or radar image shows the same value than the pixel in the previous image. Steady-state describes a shift of the image by a suitable motion vector without changes in the size or intensity of structures within the image. The motion vector for producing a forecast is usually calculated by cross-correlation between the two latest images. The linear method gains the future values of each pixel by linear extrapolation of the trend within the two latest images. In a more sophisticated approach the steady-state or the linear extrapolation can also be used for whole objects and not just pixel-wise. Other methods also rely on conceptual life-cycle models, image processing to gain the motion vector fields, or on neural-networks.

Typically nowcasts are superior to NWP models up to a lead time of 3–6 hours (Bowler et al., 2006; Pinto et al., 2010; Kober et al., 2012). For the time interval from 2 or 3 to 6 or even 8 hour forecasts more and more groups work on blending approaches, gradually combining radar based nowcast for the first hours with later on superior probabilistic NWP output for gaining better results in the transition hours (Wolfson et al., 2008; Dupree et al., 2009; Kober et al., 2012). In Kober et al. (2012) probabilistic nowcasts are developed by extending the deterministic radar tracker Rad-TRAM (RADar TRacking And Monitoring) which is described in Kober & Tafferner (2009) and combining these nowcasts with probabilistic forecasts based on the output of the COSMO-DE Ensemble Prediction System (EPS). COSMO-DE EPS is the operational ensemble prediction system of the DWD using the COSMO-DE model (see Chapter 3.3.3).

For detecting CI before a signal in most of the radar based tools is available, geostationary satellite based CI detection algorithms are developed. Algorithms like the SATellite Convection Analysis and Tracking algorithm SATCAST (Mecikalski & Bedka, 2006; Mecikalski et al., 2008; Siewert et al., 2010) with its version 2 (SATCASTv2) as described in Walker et al. (2012), or the UW-CIMSS CI nowcast algorithm UWCI (Sieglaff et al., 2011). In SATCAST cloud-top cooling rates, and instantaneous and time trends of channel differences are evaluated pixel by

pixel in different numbers for the American GOES (Geostationary Operational Environmental Satellite) and the European MSG data. In the recently developed SATCASTv2 an object-tracking method inspired by Zinner et al. (2008) is incorporated. Mesoscale atmospheric motion vectors are used for the extrapolation of the detected object to retrieve a first guess (Walker et al., 2012) then the overlap is evaluated. The tracking is used to improve the calculation of time trends which is heavily dependent on realistic matching of the right pixels. UWCI uses GOES infrared window channel box-averaged Cloud-Top Cooling (CTC) rates. CI nowcasts in UWCI are "based on a combination of CTC rates and satellite-derived cloud-top type-phase trends" (Sieglaff et al., 2011). One advantage of UWCI is its day-night independence. Both, SATCASTv2 and UWCI, produce only CI detections and they do not show an expected future position of the cloudpixels marked as CI. The output is more or less comparable to the CI detections of DLR's aforementioned Cb-TRAM, but they do not produce nowcasts, like within Cb-TRAM.

Algorithms focusing not solely on CI detection which are using geostationary satellite data as well are for example the Rapid Developing Thunderstorms (RDT) tool (Morel et al., 2000) or Cb-TRAM (Zinner et al., 2008; Zinner & Betz, 2009). Cb-TRAM is capable of tracking and monitoring of severe convection from onset (CI) over rapid development to mature phase using multi-channel Meteosat-8 or 9 SEVIRI data. It will be applied in this study and will be explained in the following chapter in more detail. The aim within this thesis now is to add information about lift, instability and moisture from other data sources to the satellite based CI detection to improve the quality of the CI nowcasting as will be described in Chapter 5.

3. Methods and data sources

3.1. Cb-TRAM CI detection

The acronym Cb-TRAM stands for **C**umulonimbus **T**Racking **A**nd **M**onitoring. Cb-TRAM is a state-of-the-art nowcasting tool for satellite based CI detection which had at the point where this study was started the unique feature of an object based methodology. Nowadays other tools like SATCASTv2 (Walker et al., 2012) have adapted the object tracking approach of Cb-TRAM. Due to the fact that the object tracking and the satellite channels used for the detection are very close to each other in Cb-TRAM and SATCASTv2, it seems reasonable that the results of this study on the use of additional data sources could be generalized for these type of CI detection algorithms.

The following section is divided into two parts. The first one introduces Cb-TRAM and the MSG data used by that tool. The other describes the different detection stages of Cb-TRAM with main focus on its CI detection.

3.1.1. Cb-TRAM overview

Cb-TRAM is a tool based on data of the SEVIRI (Spinning Enhanced Visible and Infra-Red Imager) instrument on the geostationary Meteosat Second Generation (MSG) satellites Meteosat-8 or Meteosat-9. The aim of Cb-TRAM is to detect convection before the onset of precipitation, and to distinguish between different development stages in the life cycle of thunderstorms (Zinner et al., 2008). This is especially useful in areas without radar coverage. Cb-TRAM is well documented in Zinner et al. (2008) and Zinner & Betz (2009). In the following a short overview on the MSG data and on the basic steps in Cb-TRAM is given.

MSG data is provided by EUMETSAT, the European organisation for the exploitation of METeorological SATellites. The Meteosat-9 satellite has its geostationary position at 0 degrees longitude in about 36000 km above the equator and has an imaging repeat cycle of 15 minutes (normal scan). The MSG SEVIRI instrument on board of the satellite observes the atmosphere in 12 spectral channels ranging from wavelengths in the visible (VIS) part of the spectra to wavelengths in the infrared (IR) incorporating water vapor (WV) absorption bands as shown in Table 3.1. Eleven of the channels have a spatial sampling distance of 3 km at nadir, what means at the sub-satellite point, and cover the full disk (Fig. 3.1a). The twelfth channel, the so called high-resolution visible (HRV) channel has a 1 km spatial sampling at nadir and covers half of the full disc (Schmetz et al., 2002). The HRV is split in two windows for optimized coverage of the landmasses during their daytime hours, despite the smaller area which is covered

by this channel.

Channel no.		Characteristics of spectral band (μm)			Main gaseous absorber or window
		λ_{cen}	λ_{min}	λ_{max}	
1	VIS 0.6	0.635	0.56	0.71	Window
2	VIS 0.8	0.81	0.74	0.88	Window
3	NIR 1.6	1.64	1.50	1.78	Window
4	IR 3.9	3.90	3.48	4.36	Window
5	WV 6.2	6.25	5.35	7.15	Water vapor
6	WV 7.3	7.35	6.85	7.85	Water vapor
7	IR 8.7	8.70	8.30	9.10	Window
8	IR 9.7	9.66	9.38	9.94	Ozone
9	IR 10.8	10.80	9.80	11.80	Window
10	IR 12.0	12.00	11.00	13.00	Window
11	IR 13.4	13.40	12.40	14.40	Carbon dioxide
12	HRV	Broadband (about 0.4-1.1)			Window/water vapor

Table 3.1.: SEVIRI channels with bandwidth after Schmetz et al. (2002). Channels used by Cb-TRAM are marked blue.

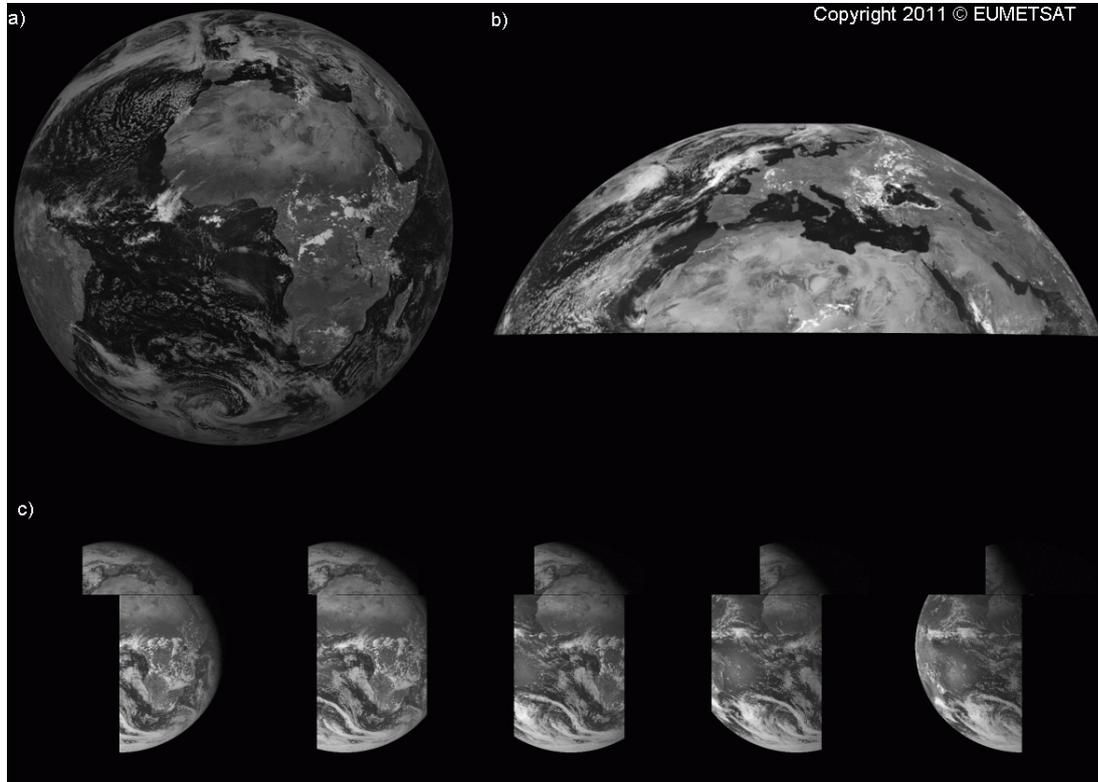


Figure 3.1.: Example images showing a) the area covered by a full disk scan, b) the area covered by rapid scan, and c) the shift of the southern window of the HRV scan pattern

As described on the EUMETSAT website¹, the upper window is fixed over Europe and northern Africa for optimized coverage of the European region. The lower window shifts to follow the daily illumination over the remaining African continent (Fig. 3.1c).

Meteosat-8 is located at 9.5 degrees east with a repeat cycle of 5 minutes (rapid scan). The SEVIRI instrument on Meteosat-8 scans a reduced area that corresponds approximately to the top third of a full disk scan (Fig. 3.1b) to realize the shorter repeat cycle. The rapid scan data has a limitation if desired for operational use because of its service interruptions. After 26 days of operational rapid scan service the service is interrupted for two days of full disc scanning used for maintenance activities and intercomparison of Meteosat-8 and -9. Nevertheless the rapid scan update circle of 5 minutes is very useful for observations of an often quite short living process like CI. A convective cell might easily develop and grow beyond its CI stage within 10 to 20 minutes and then the normal scan data would capture this cell at most once but often not even once.

Cb-TRAM uses the four SEVIRI channels marked blue in Table 3.1: WV 6.2, IR 10.8, IR 12.0, and HRV. The brightness temperature of the WV 6.2 channel gives information about the water vapor tropospheric background temperature. IR 10.8 temperature resembles the cloud top temperature. IR 12.0 is used to filter thin cirrus shields. The HRV is a visible channel and therefore only available during daylight hours. Usage of the HRV is defined in dependency of the local solar zenith angle ($SZA < 75^\circ$) as in Zinner & Betz (2009). Its higher spatial resolution increases the detection abilities of Cb-TRAM during the daylight hours. For different purposes not only the reflectivity but also the roughness of the HRV signal is evaluated (see Subsection 3.1.2).

The key algorithms of Cb-TRAM extract a general transformation vector field from consecutive images, describing the change within the two scenes. The pyramidal image matching analyzes the images stepwise, in different resolutions, to start with large scale features and refine the vector field with the help of the smaller scales. The resulting vector field can be used to generate an extrapolated synthetic image. These synthetic images are utilized several times throughout Cb-TRAM: for the detection, tracking, and nowcasting.

- Detection:

The criteria for the detection (see Subsection 3.1.2) make use of time-trends/changes in different fields. If the difference between the values of the fields for the different time steps would be calculated, the result would account not only for newly developing or decaying signals in the field but also for advection of signals. The deformation vector field gained by the analysis of a field should lead from one image to the other and therefore always includes the trend beside the advection. This has the biggest impact in the analysis of the visible channel. By choosing not the deformation vector field gained with the help of the channel and time step we want to analyze but with the deformation of one of the other channels or a slightly earlier time step we get a first guess that is more appropriate to filter the

¹www.eumetsat.int/Home/Main/Satellites/MeteosatSecondGeneration/Services/index.htm

advection. That means by producing the extrapolated first guess field with the most adequate deformation vector field the extrapolation field shows only advection and almost no trend. Therefore, the difference between the synthetic extrapolation and the observation for the corresponding time step shows mainly the trend of the field and the advection is excluded almost completely.

- **Tracking:**
A combination of masks, showing pixels with fulfilled criteria for detection, leads to the creation of cloud objects. This is described in the Subsection on detection stages (3.1.2) and in more detail in (Zinner et al., 2008). After determining these cloud objects or cells at a time step the extrapolated synthetic images, with the expected position of the cloud objects, strongly evolve the quality of the cloud object matching within different time steps in the tracking scheme. A successfully tracked cell gets the same cell ID as in the earlier time step. Cell merging and splitting can be handled as well. For details on this topic see again Zinner et al. (2008) or Zinner & Betz (2009).
- **Nowcasting:**
In addition, these extrapolations are used to generate nowcasts of the objects for up to one hour. The available time interval of the nowcasts depends on the used input data. For normal scan data a 15, 30, 45, and 60 minute nowcast are produced. For the rapid scan data the time interval is reduced to 5 minutes. A more detailed description can be found in Zinner et al. (2008).

For all cells detected by Cb-TRAM we get information about its history – in terms of merging and splitting of cells along their life cycles and in terms of a track showing the trajectory of the center of gravity since its first detection – and its future in terms of nowcasts for up to one hour.

The pyramidal image matching method to extract the vector field is used for multiple applications with and besides Cb-TRAM, e.g. for contrail detection (Mannstein et al., 1999), verification of Quantitative Precipitation Forecasts (QPFs) with the “Displacement and Amplitude Score” (Keil & Craig, 2007, 2009), radar tracking with Rad-TRAM (Kober & Tafferner, 2009), and within other studies on convective storms (Mannstein et al., 2002; Tafferner et al., 2008; Dotzek & Forster, 2011).

3.1.2. Detection stages

Thunderstorm detection in Cb-TRAM is split up in three different detection stages describing different stages in the life cycle of a storm. Stage 1 is the convection initiation (CI) stage, stage 2 marks rapid development, and stage 3 detects mature thunderstorm cells. A perfectly detected and tracked storm would therefore start in stage 1, then develop to a stage 2 detection and finally change

to a stage 3 detection. This extension to development stages later in the life cycle than the CI stage is one of the main differences between Cb-TRAM and other satellite based CI detection algorithms like SATCAST (Mecikalski & Bedka, 2006; Mecikalski et al., 2008; Siewert et al., 2010) or UWCI (Sieglaff et al., 2011) which both focus only on CI. This difference will be particularly beneficial for the evaluation. The main focus of the following description of the Cb-TRAM detection lies on the CI stage, but a short introduction to the detection of the second and third stage is given as well as they are used in the evaluation. The explanations can again be found in more detail in Zinner et al. (2008) and Zinner & Betz (2009).

Boundary layer convergence is expected to be one key trigger mechanism for CI (Wilson & Mueller, 1993; Wilson et al., 1998). Low-level convergence is often marked by the development of shallow cumulus humilis fields. However, not each of these clouds can be expected to grow to a mature thunderstorm, thus only the newly and rapid developing cumulus clouds shall be detected. Therefore the detection of CI stage objects uses the development of signals in channels that can represent vertical and horizontal growth of a cell. IR 10.8 cooling is a sign for vertical growth and the HRV is used to detect changes in the horizontal extent of a cell. Due to the usage of the HRV channel the CI stage detection is a daytime-only product. To estimate the growth of the cloud, time trends of these channels are evaluated utilizing extrapolated, synthetic first guess fields as mentioned before to determine the development without the influence of advection (see Subsection 3.1.1). Additional preconditions like the minimal size of a cell, or the appropriate choice of the deformation vector field used for the extrapolation are discussed in Zinner et al. (2008). Areas with HRV reflectivity less than 0.5 are excluded to filter miss-detections due to moving fields of thin (dark) cirrus clouds and to limit the detections to real low-level development. The resulting Δ IR 10.8 and Δ HRV are not combined by two independent thresholds for each field but in a fuzzy logic approach to allow a weighted consideration of the signals in the two fields. That means that for example also cumulus clouds which show only little horizontal growth within the last time step can be detected as CI object, if their vertical growth is strong enough and vice versa. The objects detected by the stage 1 criteria are marked as yellow polygons on Cb-TRAM plots as can be seen in the example (Fig. 3.2).

The rapid development stage or stage 2 identifies cells with rapid cooling of more than $1\text{ K}/15\text{ min}$ in the WV 6.2 channel. The calculation of the trend is realized analogous to stage 1 trend determination. These clouds grow rapidly to middle or higher tropospheric levels but their cloud top temperature in IR 10.8 is not yet cold enough to be detected as mature. The cells detected by the stage 2 criteria are marked as orange polygons on Cb-TRAM plots (Fig. 3.2).

Stage 3 denotes mature thunderstorms. Two criteria are combined for the detection. First the WV 6.2 – IR 10.8 difference is evaluated. If the difference approaches zero or gets positive, the cloud top of the cumulus is expected to reach the tropopause (Schmetz et al., 1997) or even overshoot that level and form a so-called "overshooting top" (Bedka et al., 2010; Aumann et al., 2011). The second criterion is the HRV texture during daytime and WV 6.2 texture during

nighttime to focus on the turbulent areas of the cloud top. The combination of the criteria is again made by a fuzzy logic approach. The texture criteria helps to select only the most active parts of the cell patterns detected by the temperature criteria and therefore filters large parts of the anvil cloud of a storm. The aim is to focus on those parts of a storm cloud where the main lightning activity and the most vigorous updrafts are expected. In addition to the two criteria a filter for thin cirrus clouds is utilized that evaluates the IR 10.8 – IR 12.0 difference which shows positive values for thin cirrus clouds (Krebs et al., 2007). The storm objects detected by stage 3 are marked as red polygons in Fig. 3.2.

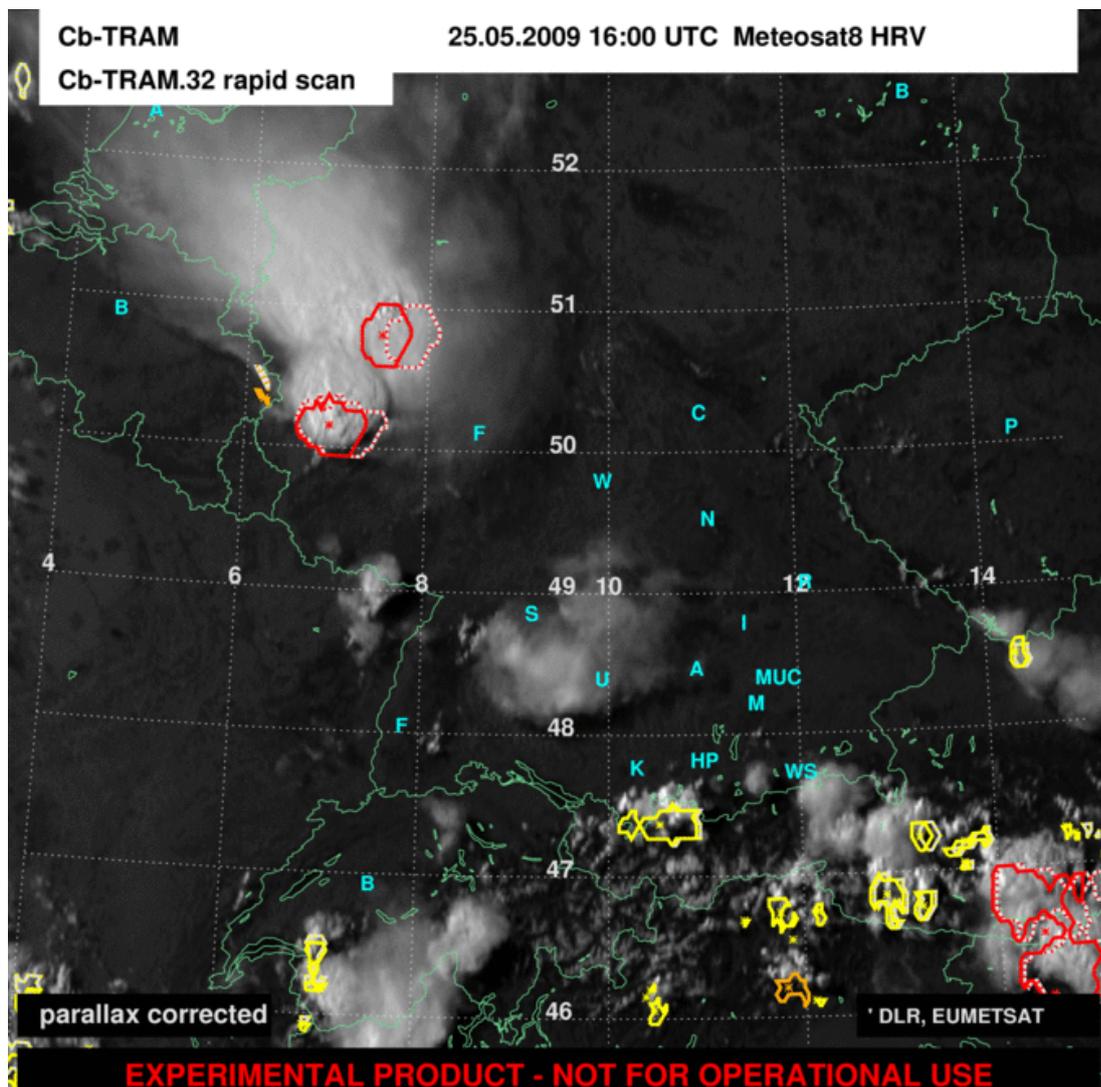


Figure 3.2.: Example of an Cb-TRAM plot with yellow polygons representing stage 1 detections (CI), orange polygons are stage 2 detections (rapid development), and red polygons are stage 3 detections (mature thunderstorms). The dotted lines for each object show the 30 minute nowcast. The star inside the objects marks the center of gravity.

These polygons or objects, as final product of the detection, are constructed by

different pixel based masks, build up by the fulfilled detection criteria in the detection stages. Areas of pixels with values above the earlier mentioned fuzzy logic criteria are selected and stored as cloud mask for each detection stage. After all three stages are checked a combined cloud mask is build. In areas where pixels fulfill the criteria for more than one detection stage, the most developed stage is chosen as characteristic for the whole pattern. Areas which are at least built up by three neighboring pixels are combined to an object. If two of these objects are separated by not more than two pixels (less than 10 km) they are merged due to the fact that larger patterns simplify the tracking. All parts of the cloud mask which are too small to build up an object will be ignored as noise. The process is again described in more detail in Zinner et al. (2008) and Zinner & Betz (2009).

3.2. Verification

The aim of this thesis is to evaluate the quality of the short-time forecast for CI events in Cb-TRAM and, furthermore to improve the quality by combining different data sources to result in a "*better*" forecast with the combination than with each individual source. Therefore it is important to understand what is meant by a good forecast and how its types of goodness (Murphy, 1993) – *quality*, *value* and *consistency* – can be evaluated. Murphy (1993) defines *quality* as a term for the correspondence between a forecast and the matching observation and the *value* of a forecast is its benefit for a decision-making process. The *consistency* describes the correspondence of a forecasters' judgment and their given forecast. This is heavily related to the question if, or how the uncertainty of a forecast is communicated. To be able to fill the analysis of the CI forecasts with an objective information how good the forecast is – with main focus on its quality – a deeper understanding of the vast field of forecast verification is needed. In the following a brief, general overview on different verification methods and verification measures is given, and the second part of this Section focuses on methods that can be used specifically for CI forecast verification.

3.2.1. Verification methods

A lot of different verification methods exist. This Subsection cannot be meant as a full overview on verification methods. The whole research area of forecast verification can be split up in two main categories: probabilistic and deterministic forecasts. The field of verification methods for probabilistic and ensemble forecasts lies beyond the scope of this text. For further reading on forecast verification in general a lot of references will be given in this Subsection. Possible starting points to get further information on the field of forecast verification research are the CAWCR (Centre for Australian Weather and Climate Research) website of the Joint Working Group on Forecast Verification Research² or the textbooks of Jolliffe & Stephenson (2003) and Wilks (2006). The focus of the following text lies on the second of the main categories mentioned before: deterministic forecasts.

²<http://www.cawcr.gov.au/projects/verification/>

There is a multitude of different verification methods for deterministic forecasts that can be categorized as visual, dichotomous, multi-category, continuous, and spatial methods.

The visual method is the so-called and often used "eyeball" verification. The problem here is that it is a highly subjective and not quantitative way of data comparison.

The dichotomous forecasts are also called yes/no forecasts or binary (1/0) forecasts. The most prominent example for a binary forecast are Finley's tornado forecasts which will be introduced later in this Subsection. To verify binary forecasts 2×2 contingency tables are used (Table 3.2). They show the four different members of the so-called *joint distribution*:

- *hit* – if an event is forecast and observed
- *false alarm* – if an event is forecast but not observed
- *miss* – if an event is not forecast but observed
- *correct negative* – if an event is neither forecast nor observed

It is a very intuitive way to categorize the results and is thus often used. It helps to understand the errors made with the forecast. A perfect forecast would produce only hits and correct negatives.

		Observation	
		Yes	No
Forecast	Yes	<i>hits</i>	<i>false alarms</i>
	No	<i>misses</i>	<i>correct negatives</i>

Table 3.2.: A 2×2 contingency table: The four possible outcomes for a deterministic forecast of a binary event (Wilks, 2006).

Many verification statistics can be calculated based on the entries of a contingency table. Depending on the chosen score it is possible to describe particular aspects of the forecast quality. Which score or which combination of scores to choose depends heavily on the questions that should be answered by the verification, and on the type of forecast. For instance, for rare events like Cbs it does not make sense to use scores that utilize correct negatives, because for rare events the non-events are usually not forecast and if they would be, the number of correct negatives would be orders of magnitude larger than the other three entries of the contingency table (Jolliffe & Stephenson, 2003). Some of the most common scores without correct negatives are:

- POD = Probability of detection (also called hit rate)

$$POD = \frac{hits}{hits + misses} \quad (3.1)$$

POD is the fraction of the observed events correctly forecast. The perfect score is 1.

- FAR = False alarm ratio

$$FAR = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad (3.2)$$

FAR is the fraction of the forecast events not observed. The perfect score is 0.

- SR = Success ratio

$$SR = \frac{\text{hits}}{\text{hits} + \text{false alarms}} \quad (3.3)$$

SR is the fraction of the forecast events that are observed and therefore equal to $1 - FAR$. The perfect score is 1.

- CSI = Critical Success Index (also called Threat score)

$$CSI = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \quad (3.4)$$

CSI is the ratio of correctly predicted events to all events (observed and predicted). The perfect score is 1. If the event is statistically rare the number of hits by chance decreases as well. Thus the CSI depends on the climatological frequency of the event and shows poorer scores for rare events (Mason, 1989).

- Bias

$$Bias = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \quad (3.5)$$

Bias is the fraction of all forecast events to all observed events. Perfect score is Bias = 1. It is a measure for overforecasting (Bias > 1) and underforecasting (Bias < 1).

The most prominent example for a discussion about forecast verification and scores is also an evidence for the long history of forecast verification and furthermore it underlines the importance to use scores which are independent of "correct negatives" for rare events: Finley's 1884 published tornado forecasts raised a heavy discussion about the calculation and interpretation of scores. This discussion is often referred to as "The Finley Affair" and documented in detail by Murphy (1996). Sergeant Finley of the U.S. Army Signal Corps produced an experimental tornado forecast and published the results (Finley, 1884) shown in Table 3.3.

		Observation		
		Tornado	No Tornado	Total
Forecast	Tornado	28	72	100
	No Tornado	23	2680	2703
	Total	51	2752	2803

Table 3.3.: Contingency table showing the results for Finley's Tornado forecasts (Jolliffe & Stephenson, 2003).

He calculated a score nowadays called *fraction correct* or *accuracy* of his forecasts ($= (\text{hits} + \text{correct negatives}/\text{total}) * 100$) of more than 96%. As a response Gilbert (1884) pointed out that this score does not make sense for Finley's tornado predictions. He mentioned that the assumption that the forecast of "tornado" and "no tornado" are equally difficult is a "serious fallacy" and gave the example of only "no tornado" forecasts. If all forecasts say "no tornado", the contingency table would show no hits, but 2752 correct negatives, what results in a fraction correct of more than 98%. Thus relying on this score Gilbert's forecast that does not require "any study of the meteorological record" seems to be "better" than Finley's forecasts. Gilbert pointed out the important difference between statistically rare and normally frequent events. This was the historical starting signal for the discussion about forecast verification.

Coming back to the different methods of verification for deterministic forecasts, the next type mentioned before are multi-category forecasts. An example might be a precipitation forecast with not just a yes or no decision, but different categories like rain, freezing rain, graupel, and snow. An $i \times j$ contingency table also called multi-category contingency table can be used to show the frequency of the different forecast and observation pairs per category. Here, histograms and the accuracy describing what fraction of the forecast was in the right category are common methods for verification.

Continuous forecasts of variables like temperature are usually evaluated with the help of scatter plots, mean squared error, root mean squared error or correlation between forecast and observation.

The drawback of classical verification methods is that they are usually evaluated pixel by pixel. Thus traditional verification scores are not capable to give reasonable credit to high resolution forecasts (Mass et al., 2002). If a high resolution forecast is capable to resolve some features more realistic than a coarse resolution forecast but with a slight positioning error, the pixel based verification might even show a worse result for the high resolution version than for the coarse resolution forecast as illustrated in Fig. 3.3.

Fig. 3.3 shows the observation of a trough in pressure measurements and three model forecasts with different model resolution (36 km, 12 km, and 4 km). The forecast troughs have a positioning error of 40 to 50 km. The model with the highest resolution represents the observed structure very well, and is obviously the best of the three forecasts. However, if, as an example, the absolute error

at the observation location marked by the red rectangle is evaluated the deviation between forecast and observation is worst for the 4 km resolution. Pixel based methods are very harsh and require an exact match between forecasts and observations at every grid point to score a hit. A slight displacement between forecast and observation results in the so-called double penalty problem. The pixels where forecast and observation overlap are counted as hits but the pixels outside of the overlapping area are penalized in two ways: The non-overlapping forecast pixels are counted as false alarms and the non-overlapping observation pixels are counted as misses (see Fig. 3.4 b).

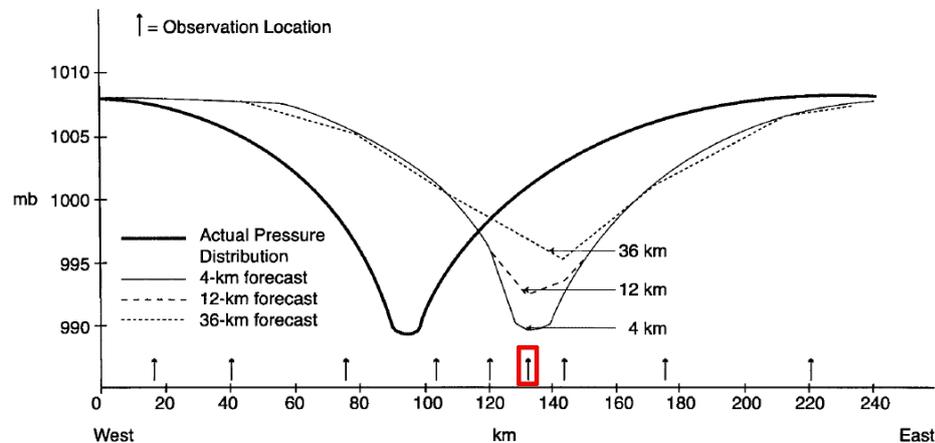


Figure 3.3.: Example for problems of point by point evaluation or pixel based verification methods to evaluate structures with a slight positioning error. A description of the data shown here can be found in the text. Taken from Mass et al. (2002).

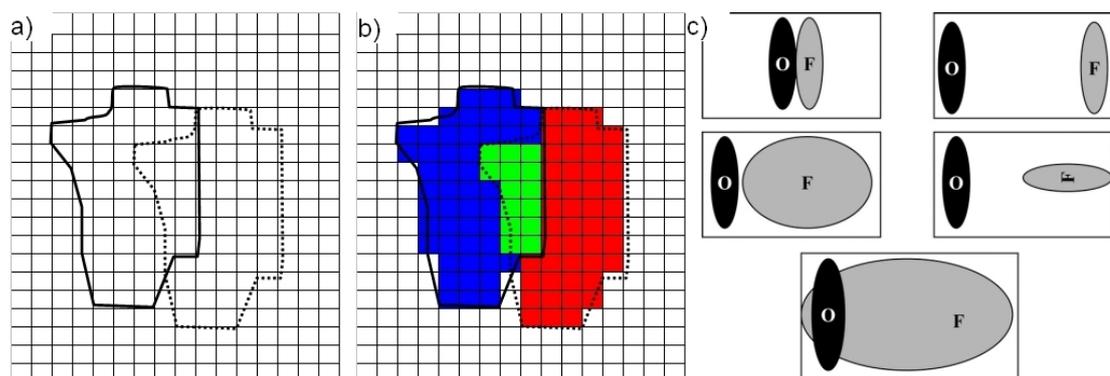


Figure 3.4.: Image a) shows a grid with an observation (solid contour) and a forecast structure (dotted contour) which is slightly displaced. Image b) shows which pixels are counted as hits (green), misses (blue), and false alarms (red) to illustrate the double penalty problem and c) shows O(bservation) and F(orecast) pairs with different types of forecast errors classical scores cannot evaluate. Fig. 3.4 c) is taken from Davis et al. (2006a).

The structure of the forecast and observation fits perfectly here, it is just shifted by a few pixels (Fig. 3.4 a).

The pixels that are counted as hits are only few compared to the amount of pixels that are counted as misses or false alarms. In addition these methods are not able to give credit for being close and they might reward a forecast for wrong reasons as can be seen from Fig. 3.4 c). The upper four examples of observation and forecast pairs in Fig. 3.4 c) show examples for different forecasts which all have the same scores for pixel based measures (e.g. $CSI = 0$, $FAR = 1$, $POD = 0$). Thus the measures give no information if the structure of the forecast is right and just slightly wrong in position (but non-overlapping) or if the structure is oriented wrong and far away and therefore the forecast was quite poor. The example on the bottom of Fig. 3.4 c) shows that forecasts are generally rewarded for being smooth by these measures. The smoother the structure the smaller the absolute errors between forecast and observation (as in Fig. 3.3) and the more often an overlap between forecast structures and observation can be realized.

Due to these restrictions of classical methods, a lot of new verification measures with focus on spatial structures and presence of desired features in the forecast were developed (Casati et al., 2008). The main difference between these measures and the classical verification statistics is that they are not evaluated on a point-to-point basis with no regard to spatial information (Baldwin & Kain, 2006). A review on many of the different approaches is given by Casati et al. (2008) and Gilleland et al. (2009). These spatial verification methods split up in four different types (see Fig. 3.5): Scale-separation (or scale decomposition) methods, neighborhood (or fuzzy) methods, feature-based (or object based) methods, and field deformation methods.

They can be described as two filtering approaches (neighborhood and scale decomposition) and two displacement approaches (object based and field deformation) as in Gilleland et al. (2009).

Many of these methods focus on the verification of Quantitative Precipitation Forecasts (QPFs) and are compared in the so-called Spatial Forecast Verification Methods Intercomparison Project (ICP³) described in detail by Ahijevych et al. (2009) and Gilleland et al. (2010a).

Examples for scale separation methods are the Intensity-Scale Technique (Casati et al., 2004; Casati, 2010) or the Fractions Skill Score (Roberts & Lean, 2008; Mittermaier & Roberts, 2010) which give information on the skill of a forecast system on different spatial scales. Field deformation methods use optical flow or image warping techniques to quantify the difference between forecast and observation fields. Examples here are the Image Warp Statistic (Gilleland et al., 2010b) and the Displacement and Amplitude Score or short DAS (Keil & Craig, 2007, 2009), which describes the displacement error and the amplitude error of QPF objects. Fuzzy or neighboring methods have the aim to reward close forecasts. Examples for these methods are given in Ebert (2008) and Ebert (2009) and are often closely related to methods from other categories of spatial verification. Object or feature based methods identify features in forecast and observation fields and then

³<http://www.ral.ucar.edu/projects/icp/>

compare attributes of these features (Casati et al., 2008). These attributes might be the size and position of an object and in case of QPFs also their amplitude in the radar echo, for example. Examples for object based scores are the SAL technique (Wernli et al., 2008, 2009), which assesses the Structure–Area–Location error for objects, CRA utilizing Contiguous Rain Areas to define the objects (Ebert & McBride, 2000; McBride & Ebert, 2000; Ebert & Gallus, 2009) and the Method for Object-based Diagnostic Evaluation (MODE) which is also a method to compare QPFs with observations (Davis et al., 2006a,b, 2009). Many methods belong to more than one category like a multiscale object based method by Lack et al. (2010) or the "cluster analysis for object-oriented verification of fields" by Marzban & Sandgathe (2008), and other approaches studied by Marzban et al. (2009).

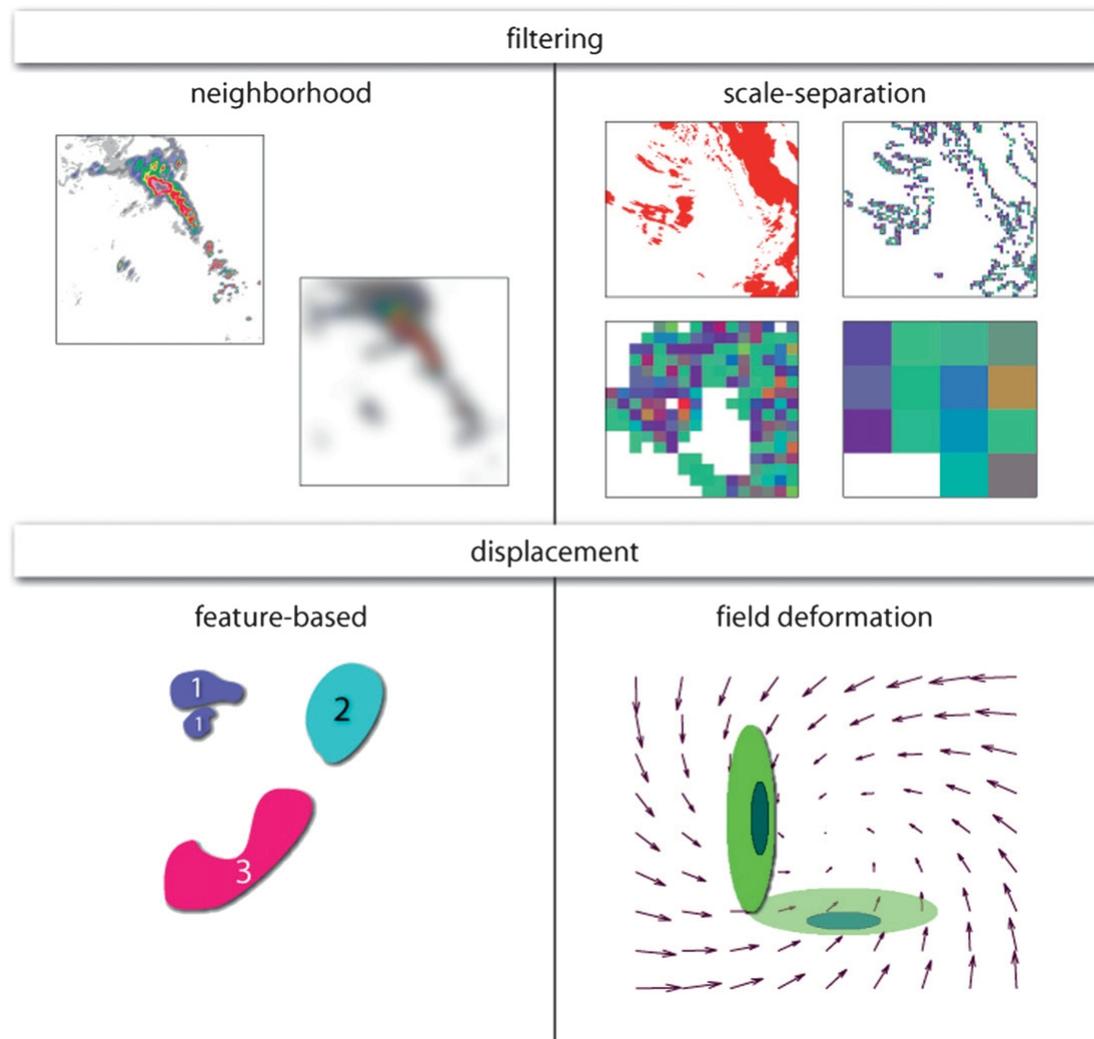


Figure 3.5.: The four different types of spatial verification methods split up in filtering (top) and displacement (bottom) approaches. Taken from Gilleland et al. (2009).

3.2.2. How to verify CI?

The verification of CI has many aspects that make it different from the verification of QPFs. One crucial point is the definition of CI and the choice of the data source used as "truth" for CI verification. Usually a CI event is verified as true if later in the lifecycle of the cell, detected as CI, a thunderstorm developed. Thus CI is defined by a kind of thunderstorm definition for the data source used for the verification. The statement used for that kind of definition is: If we see a thunderstorm right now there must have been CI earlier.

Different kinds of definitions for verification of a thunderstorm, based on radar or – preferably – on lightning data, exist in the literature. Roberts & Rutledge (2003) and Mecikalski & Bedka (2006) utilize the first radar echo ≥ 35 dBZ as a proof that CI has happened. The radar echo intensity can be used to surrogate lightning information. The definition of a thunderstorm with the help of radar echo intensities ≥ 35 dBZ is related to studies by Dye et al. (1989) and Gremillion & Orville (1999). They "have shown that the onset of storm electrification generally occurs > 5 minutes after storm echoes of 30 dBZ or greater have reached subfreezing levels" (Roberts et al., 2012). This 35 dBZ threshold discriminates quite good between weakly precipitating storms and vigorous convective storms with heavy rain and is therefore also used in radar tracking algorithms. Rad-TRAM (Kober & Tafferner, 2009) uses 37 dBZ to define precipitation cells. TITAN⁴ the Thunderstorm Identification, Tracking, Analysis and Nowcasting tool (Dixon & Wiener, 1993) utilizes two thresholds at 35 dBZ and 45 dBZ.

The already mentioned second data source commonly used for the proof of a thunderstorm is lightning data. Donovan et al. (2008) and Zinner & Betz (2009) use "flash rates" or "lightning density" as criteria to verify mature storms. Thus they use a cluster of flashes with certain criteria for the density in time and space as evidence for a thunderstorm. Zinner & Betz (2009) verified the mature storm detection stage of Cb-TRAM against lightning, showing that Cb-TRAM is capable to detect up to 90% of the storms that are marked by the lightning data with a false alarm ratio below 25%. Thus the Cb-TRAM detection of mature storm cells seems to be sufficient, too, to be used as truth for CI verification.

The main problem for CI verification is an objective and automated matching of CI cells to the later detected thunderstorms. For storm cells some studies go back to a kind of "eyeball" verification of objects to match storm detections with verification data (Donovan et al., 2008) and fill their contingency table. QPF objects can be compared more or less "directly" to the precipitation observations with many of the spatial methods mentioned before. A CI object, in contrast, is not supposed to match the later existing thunderstorm perfectly. It should just identify the cloud that later on develops to a storm, but almost all attributes of the cell like its position, size, shape etc. are even expected to change during the development from CI to a mature storm. Thus most of the spatial methods for QPFs that evaluate differences in the shape, size, and/or amplitude of an object cannot be used here. Mecikalski et al. (2008) tries to verify the SATCAST results for some case studies. SATCAST produces pixel based detections and no objects,

⁴<http://www.ral.ucar.edu/projects/titan/home/>

or cells. They use a kind of fuzzy pixel based measure to evaluate if the pixels, they expect to represent CI, develop to a storm in the radar signal of neighboring positions within the next 30 to 60 minutes and calculate classical statistics with these "fuzzy" hits, misses, and false alarms. For SATCASTv2 the analysis is performed with a pixel based comparison for few test cases around some radar sites using classical scores, where some are incorporating also correct negative in the calculation.

The extrapolation or nowcast of a CI detection like in Cb-TRAM cannot include the correct development of the object size or shape, but it should be in the right position where the changed cell is expected. Therefore a strict classical verification with a pixel-per-pixel comparison does not make sense here. The aim of the Cb-TRAM CI detection verification is to get objective information on the amount of CI detections that develop to the next detection stages within the nowcast timespan and the amount of developing cells that are detected as CI cells. Thus a criterion evaluating the object overlap between a CI stage nowcast position and the according analysis of the cell seems to be desirable for the purpose of Cb-TRAM CI detection verification. Additionally the verification method has to be extendable to a longer period and not just a few case study days to get reliable statistics on Cb-TRAM performance. For evaluation of the uncertainty in the nowcast position a fuzzified verification method is tested which allows a small positioning error due to nowcasting errors. Further answers to the question on the realization of such a spatial verification method and the results for the Cb-TRAM CI detection are given in Chapter 4.

3.3. Additional data sources

This section includes descriptions of the additional data sources that are analyzed for this study and are tested in combination with Cb-TRAM. The area over which the analysis is performed is shown by the Cb-TRAM output in Fig. 3.6. It covers large parts of Central Europe.

The intention is to keep the following data descriptions rather short. References and websites are added for further reading about the different data sources. A table summing up the collected data can be found as Appendix A. As already mentioned the different data sources should add information about lift, instability and moisture to the satellite based CI detection. Thus one aim is to filter CI false alarms for "non-CI-favorable" circumstances. A second aim is to retrieve a kind of CI probability for cells with "CI-favorable" circumstances.

Lightning data is used as a proof that a thunderstorm already exists in the area detected as CI. This means, if a cell is detected as CI and it incorporates already lightning, it cannot be a false alarm, in the sense of a storm non-occurrence, and thus will not be filtered, independent from the information of the other additional data sources. The classification error that a cell which shows already lightning should be assigned to one of the later detection stages is not treated within this study. One common reason for this kind of misclassification might be that the vertical extent of the detected storm is simply not large enough and the top of

the storm does not reach the upper troposphere. Thus it is not classified as a mature storm. In situations with small, short-living, and showery storms these misclassifications are most common.

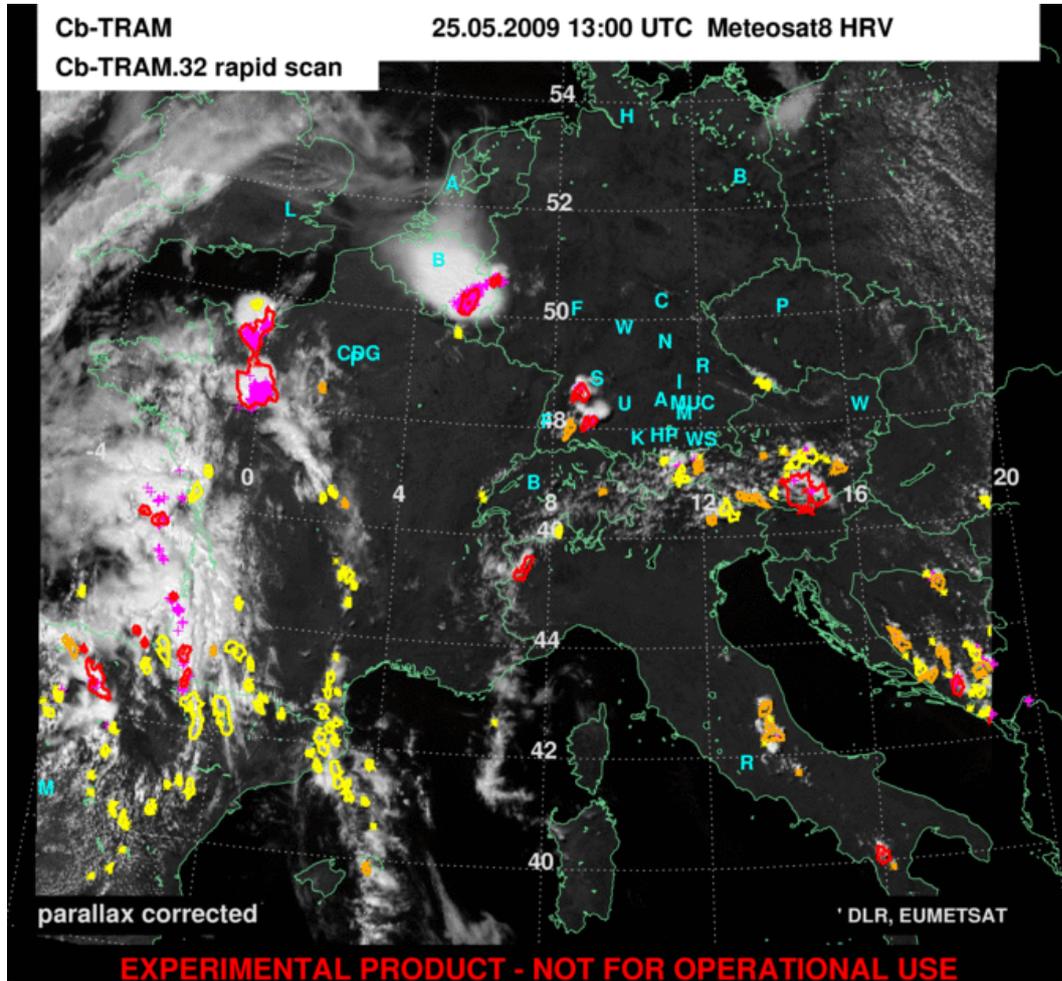


Figure 3.6.: Cb-TRAM plot as in Fig. 3.2 but without nowcast contours and center of gravity. Additionally magenta plus signs mark LINET lightning detections. For more information on LINET see Section 3.3.1.

All the other additional data is associated with one of the above mentioned ingredients. Moisture is important in low levels for the development of a storm. The moisture within this study is investigated with the help of surface observations. The concrete data tested here are moisture flux convergence and equivalent potential temperature. For evaluation of the instability, upper air data (mid- or upper-tropospheric) is needed as well. All available regular upper air observation data is not sufficient in spacial and/or temporal resolution, thus NWP model data is incorporated here, too. A combination of surface observations and mid-tropospheric NWP model data is used to calculate an instability index. The used KO-Index is based on the equivalent potential temperature in four different atmospheric levels. To get information on the available lift both are tested - surface observations and upper air model data. One idea is to evaluate the

surface observations on mass convergence. If we observe a convergent flow at the surface the air has to rise somehow and generates lift there. In addition the upper air updrafts in the omega (ω) data (vertical motion) of a NWP model are investigated. Therefore different model levels between 850 hPa and 400 hPa are incorporated in the evaluation. In comparison to the surface convergence, which describes one of the processes causing lift (see Ch. 2.1), the updraft in the model data is expected to be more useful to describe the lift. The NWP model should be able to generate a signal in the ω data, for any of the processes which might be responsible for triggering the lift.

3.3.1. Lightning data

The lightning data used within this study are ground-based LINET measurements provided by *nowcast*⁵. LINET has a very precise detection of the lightning locations with errors less than 150 m provided that the sensor baseline does not exceed ~ 250 km (Betz et al., 2009), which is given at least for the whole mainland covered by this study as shown in Fig. 3.7. More information about the abilities of this network can be found e.g. in Betz et al. (2008) and Betz et al. (2009). The data is available in near-realtime and is updated every five minutes.

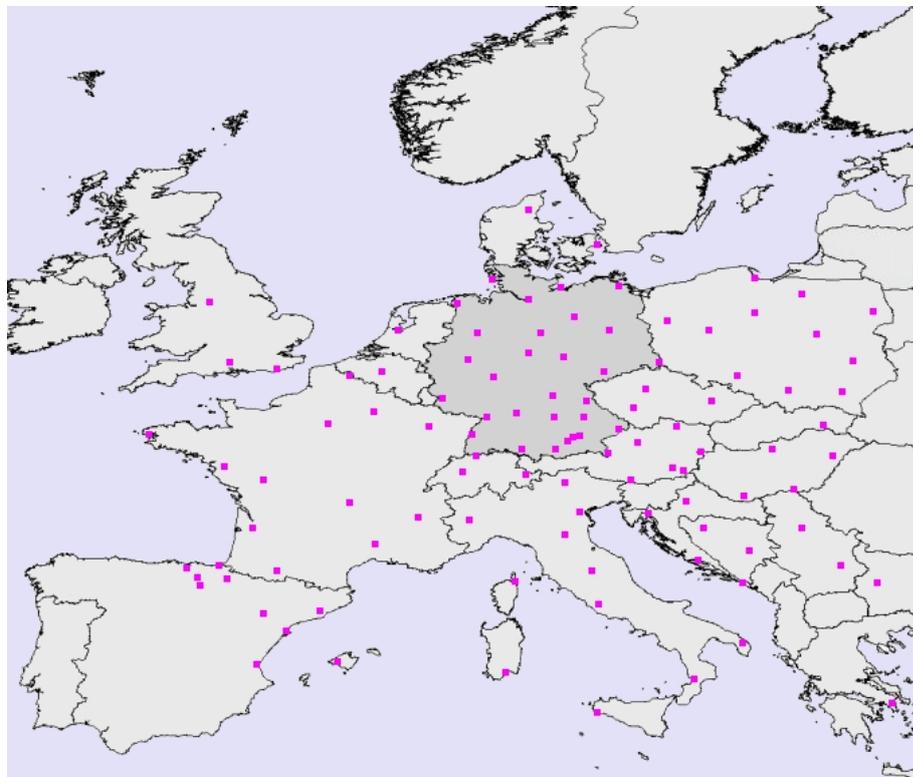


Figure 3.7.: LINET stations around central Europe. Coverage of 2009. Figure taken from the *nowcast* webpage

⁵<http://www.nowcast.de/>

3.3.2. Surface observation analysis

The Vienna Enhanced Resolution Analysis⁶ (VERA) scheme is an objective high resolution analysis of meteorological fields based on surface observations. It is developed by the University of Vienna. The surface synop station observations are brought on a 2 D grid with the help of a thin plate smoothing spline interpolation (Steinacker et al., 2000b). No model first guess fields are needed. The interpolation utilizes additional information via so-called “fingerprints” (Steinacker et al., 2006; Bica et al., 2007) – a priori knowledge for areas with complex terrain or sparse data – and the data is quality controlled (Häberli et al., 2004). Additional studies on the interpolation of precipitation fields in complex topography (Dorninger et al., 2008; Schneider & Steinacker, 2009) and on the tracking (Steinacker et al., 2000a) and nowcasting (Schneider et al., 2008) of convective storms with VERA give further insight to the methods. VERA can be run with different grid resolutions and on different domains depending on the available input synop station data.

Output fields are e.g. u- & v-component of wind, potential temperature, θ_e , pressure reduced to MSL, 6hourly accumulated precipitation, or horizontal moisture flux convergence/divergence (MFC/MFD).

Moisture flux convergence, convergence itself and θ_e from VERA will be further examined within this thesis. The used grid resolution for the interpolation of the data is 8 km as can be seen from the German header in Figs. 3.8 and 3.9. In visual tests this 8 km resolution performed quite good, while the 4 km resolution output showed artificial structures, especially around the coastlines. These structures seem to be mainly due to the available density of synop stations used within this study with its loss of information over the seas.

MFC is calculated based on Banacos & Schultz (2005) as:

$$MFC = -\nabla \cdot (q\vec{v}_h) \quad (3.6)$$

or for the moisture flux divergence as VERA output as:

$$MFD = \nabla \cdot (q\vec{v}_h) \quad (3.7)$$

where

q = specific humidity in $g\ kg^{-1}$, and

\vec{v}_h = the horizontal wind vector containing the components u and v in $m\ s^{-1}$.

For further information on the calculation of MFD see Appendix B or Kaufmann (2006). Fig. 3.8 shows an example plot of the VERA standard output for MFD together with the 10 m wind field. Positive values resemble moisture flux divergence, negative values moisture flux convergence of the order of $10^{-4}\ g\ kg^{-1}\ s^{-1}$.

The equivalent potential temperature θ_e is the potential temperature that a parcel of air would have, if all its moisture content were condensed and the resultant latent heat of condensation were used to warm the parcel. Therefore, the temperature of a parcel can be brought to its θ_e value by raising (expanding) the

⁶<http://www.univie.ac.at/amk/vera/>

parcel from its original level until all the water vapor in the parcel has condensed and fallen out, and then compressing the parcel dry adiabatically to a pressure of 1000 hPa.

Montag, 25. Mai 2009, 15:00 UTC, Mitteleuropa (8 km Gitter)

Feuchtefflussdivergenz (Farbflächen), Einheit: $10^{-4} \text{ g}/(\text{kg s}) [2]$, Beobachtungen: 969, Symbol: *, Min: -30.58, Max: 21.67, μ : -0.05, σ^2 : 19.13
10 m Wind (Pfeile), Einheit: m/s, Beobachtungen: 1067, Symbol: o, Min: 0.06, Max: 11.71, μ : 2.97, σ^2 : 2.67 \rightarrow (10 m/s)

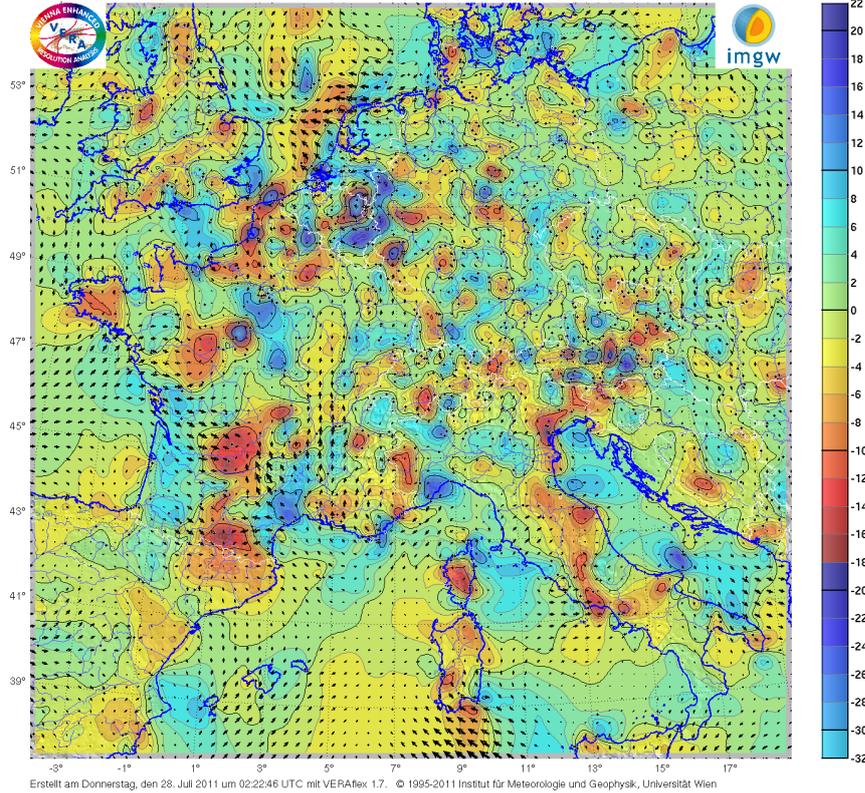


Figure 3.8.: Example of an original VERA MFD plot. Color shows MFD, arrows show 10 m wind.

θ_e is a measure for the static energy of a parcel of air and can after Holton (2004) be written as:

$$\theta_e \approx \theta \exp \left[\frac{Lm}{c_p T} \right] \quad (3.8)$$

where

θ = potential temperature in Kelvin,

T = current temperature in Kelvin,

m = mixing ratio in $10^{-3} \text{ g kg}^{-1}$,

c_p = specific heat at constant pressure ($1004 \text{ J kg}^{-1} \text{ K}^{-1}$, Holton, 2004), and

L = latent heat of condensation ($2.5 \times 10^6 \text{ J kg}^{-1}$, Holton, 2004).

The θ_e in VERA which is used later within this thesis is given in degree Centigrade and not in Kelvin thus $\theta_e (\text{VERA}) = \theta_e (\text{Holton}) - 273.15 \text{ K}$. Fig. 3.9 shows an example plot of the VERA standard output for θ_e together with mean sea-level pressure isobars and the 10 m wind field.

Montag, 25. Mai 2009, 15:00 UTC, Mitteleuropa (8 km Gitter)

Aquivalentpotentielle Temperatur (Farbflächen), Einheit: °C [2], Beobachtungen: 1017, Symbol: Diamant, Min: 25.51, Max: 80.27, μ : 50.12, σ^2 : 124.44
 Reduzierter Bodendruck (Isolinien), Einheit: hPa [1], Beobachtungen: 947, Symbol: +, Min: 1008.37, Max: 1021.14, μ : 1015.3, σ^2 : 4.78
 10 m Wind (Pfeile), Einheit: m/s, Beobachtungen: 1067, Symbol: o, Min: 0.06, Max: 11.71, μ : 2.97, σ^2 : 2.67 → (10 m/s)

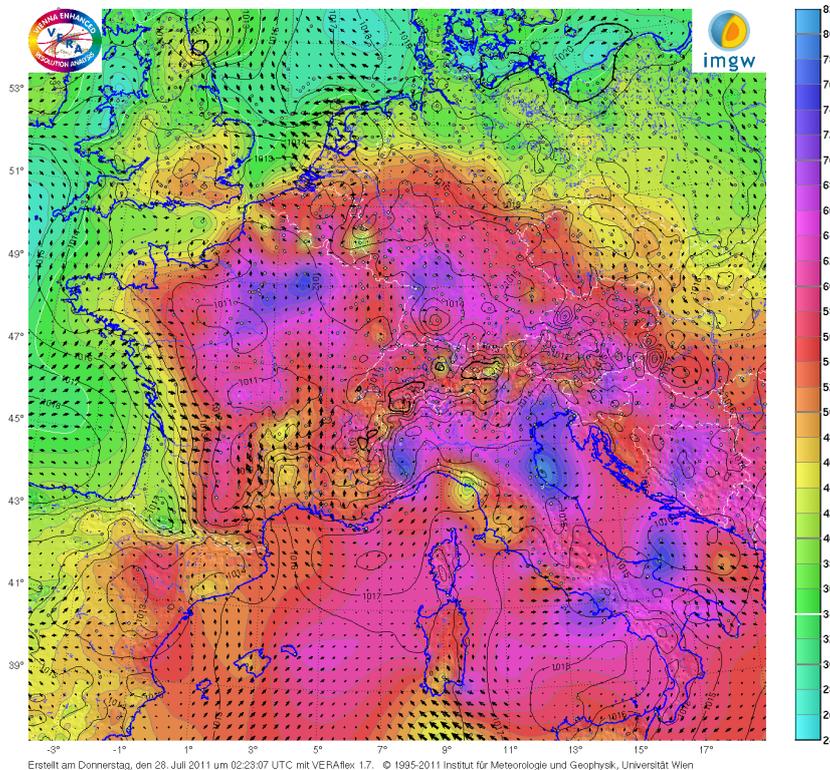


Figure 3.9.: Example of an original VERA θ_e plot. Color shows θ_e , arrows show 10 m wind, isolines show pressure reduced to MSL.

The synop information as input for the VERA algorithm is available hourly. In most measurement routines of the European national weather services the information needed as input is measured more often within one hour, but the availability of this data is quite restricted. The hourly data used here is available via a database of the DWD less than five minutes after the full hour. Only few minutes later, usually around six to seven minutes after the full hour the VERA analysis can be finished.

3.3.3. NWP model data

The DWD runs two different operational versions of the non-hydrostatic limited-area atmospheric prediction model COSMO⁷. They are named COSMO-EU and COSMO-DE. The COSMO-EU model focusses on the meso- β scale with a grid spacing of 7 km. The main aim here is the accurate prediction of near-surface weather conditions concentrating on clouds, frontal precipitation, fog, and orographically and thermally forced local wind systems (Schättler et al., 2011). In contrast the COSMO-DE model is a meso- γ scale version with 2.8 km grid spacing. The aim here is a direct simulation of severe weather events triggered by

⁷<http://www.COSMO-model.org>

deep moist convection. Examples for this kind of events are supercell thunderstorms, intense mesoscale convective complexes, and prefrontal squall-line storms (Schättler et al., 2011). The COSMO-EU domain covers more or less Europe while COSMO-DE concentrates on Germany with little parts from the neighboring countries as can be seen in Fig. 3.10. A summary of the different values for resolution, grid points, etc. characterizing the two different COSMO versions is given in Table 3.4 at the end of this chapter.

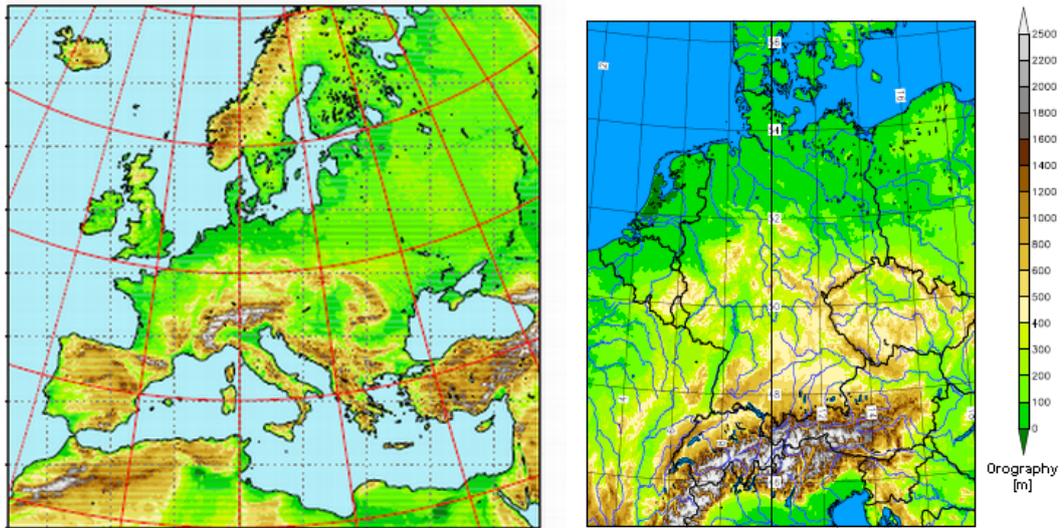


Figure 3.10.: The two maps show the domains of the DWD COSMO versions with COSMO-EU on the left side and COSMO-DE on the right side. Images taken from Schulz & Schättler (2011) and Baldauf et al. (2011a).

The different phenomena that should be predicted by the different models also influence their forecast time: The smaller the scale of the atmospheric phenomena simulated by the model, the shorter their life time and therefore the lower their predictability. Thus the COSMO-EU model produces forecasts for up to 72 hours while the COSMO-DE produces only forecast of up to 21 hours.

In contrast to the COSMO-EU where convection is fully parameterized the COSMO-DE is able to resolve large convective elements. Analyzing the forecasts of COSMO-DE in comparison to observations these elements sometimes show errors in the timing and/or location of the convective cells (Dahl et al., 2011) which might to a big part be related to deficiencies in the quality of the boundary layer parameterization (Baldauf et al., 2011b). Despite COSMO-DE is able to forecast deep convection explicitly, individual convective cells are hardly predictable. To overcome this issue the COSMO-DE model is extended to a convective-scale ensemble prediction system (Baldauf et al., 2011b) at DWD right now.

Due to the fact that the direct thunderstorm measures in the COSMO-DE model like the output field "thunderstorm probability" are influenced by the location and timing errors mentioned before I preferred measures to check for the availability of two of the ingredients for convection with the help of the model data: instability and lift. The model fields tested for my study are θ_e in 850, 700,

and 500 hPa, used for the calculation of an instability index (the KO-Index), as well as the model updraft ω in five levels between the 850 and 400 hPa levels. These fields are available hourly by the use of COSMO-DE as well as COSMO-EU, but by using the COSMO-EU fields the data covers the whole domain I wanted to use for this study. The COSMO-EU model is as mentioned before quite good in resembling the local wind systems. The orography is represented quite well with the 7 km grid spacing with not to big differences in comparison to the 2.8 km spacing of the COSMO-DE as can be seen in Fig. 3.11. The ω fields are quite noisy and are interfered with wave structures which are smoothed by a Gaussian kernel in the later described analysis.

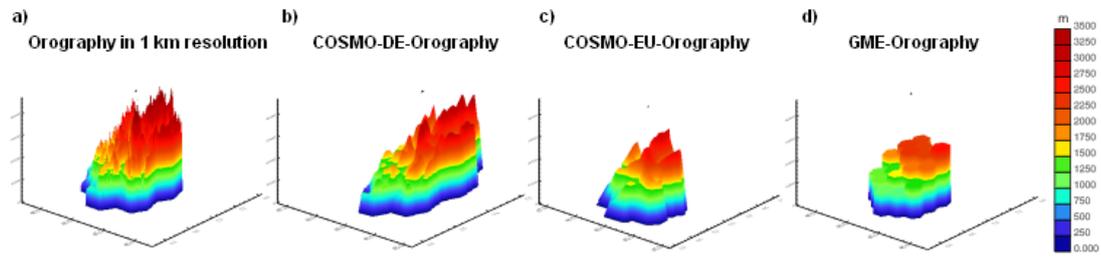


Figure 3.11.: Orography in an area around Mont Blanc. a) raw data with 1 km resolution (maximum height: more than 4000 m a.s.l.); b) COSMO-DE orography derived from the raw data as mean values each over a 8 km² grid element (maximum height: nearly 3700 m a.s.l.); c) COSMO-EU orography derived from the raw data as mean values each over a 49 km² grid element (maximum height: more than 3000 m a.s.l.); d) GME orography derived as mean values each over a 346 km² grid element (maximum height: 2920 m a.s.l.). Images and text after the description on the DWD webpage⁸

The KO-Index (German: "Konvektiv-Index") is developed and still used by the DWD. It is calculated as difference of the equivalent potential temperature between mid levels (700 to 500 hPa) and low levels (1000 to 850 hPa) to evaluate the potential instability between these levels (Haklander & Van Delden, 2003).

$$KO = 0.5(\theta_{e_{500hPa}} + \theta_{e_{700hPa}}) - 0.5(\theta_{e_{850hPa}} + \theta_{e_{1000hPa}}). \quad (3.9)$$

In this study it is calculated as combination of COSMO-EU model data (M) and surface observations (VERA) as:

$$KO = 0.5(\theta_{e_{500hPa_M}} + \theta_{e_{700hPa_M}}) - 0.5(\theta_{e_{850hPa_M}} + \theta_{e_{sfc_VERA}}). \quad (3.10)$$

⁸http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_spezielle_nutzer_forschung_analyse&T12203837091139841917821gsbDocumentPath=Navigation%2FForschung%2FAnalyse__Modellierung%2FFU__NM__LMK__node.html%3F__nnn%3Dtrue&_pageLabel=_dwdwww_spezielle_nutzer_forschung_analyse&switchLang=en [30.11.2012]

	COSMO-EU	COSMO-DE
Resolution	7 km	2.8 km
Grid-points	665 x 657	421 x 461
Layers	40	50
Forecast up to	78 h	21 h
Update cycle	6 h	3 h
Coordinates of the domain corners (λ_g, φ_g):		
lower left corner	9.19°W, 27.70°N	2.98°E, 44.77°N
upper left corner	34.24°W, 65.58°N	1.04°E, 56.20°N
upper right corner	63.47°E, 62.40°N	19.84°E, 56.14°N
lower right corner	34.67°E, 26.12°N	17.72°E, 44.72°N

Table 3.4.: Comparison of the two operational COSMO versions at DWD. Information taken from Schulz & Schättler (2011) and Baldauf et al. (2011a).

3.4. Fuzzy logic to combine different data sources

Fuzzy logic, in contrast to conventional (Boolean) logic, is capable of handling the idea of partial truth. Conventional logic uses the principle of bivalence, meaning that propositions are either fully true (1 in the sense of a binary decision) or fully false (0). Whereas fuzzy logic introduces some vagueness, stating that a proposition might be partially true and false to some degree. A fundamental concept of fuzzy logic is that of fuzzy sets which was first published by Zadeh (1965). Fuzzy sets are an extension of the classical notion of (crisp) sets: in a classical set there are two distinct categories of objects/values, those in the set and those outside of the set (complement), a fuzzy set permits intermediate degrees of membership. A depiction of classical and fuzzy sets with illustration of the operations complement, intersection, and union is shown in Fig. 3.12. The reasoning in fuzzy logic is closer to human reasoning. This can for example be shown with the help of an example from the monograph of Klir & Yuan (1995) shown in Fig. 3.13. In human reasoning there are no distinct boundaries between "cold", "cool", "moderate", "warm" or "hot" which might resemble the (very) low to (very) high values in the Figure. And often a temperature is somewhere "between" two of these categories in reality as can be visualized like in Fig. 3.13a). With the fuzzy logic resembling of these sets a temperature might be to 40 % within one category and to 60 % within the neighboring category instead of being by 100 % in one category due to some strict threshold like it would be in the case of Fig. 3.13b). Here all values within an interval will be changed to one crisp value representing the whole interval. For the fuzzy variable like in Fig. 3.13a), depending on the symmetry or non-symmetry of the transitions between two categories, the membership grades might add up to 100% if the consequent fuzzy sets are symmetric but they don't have to be symmetric anyway. It is possible to use membership functions where for example the membership grade

of a temperature is 1 for category "very low" and at the same time already non-zero for "low" etc. The definition of the fuzzy sets allows incorporating expert knowledge to get reasonable fuzzy sets for the task they are constructed for. Due to the fact that only symmetric cases will be shown within this study I will not go further into details here and suggest that readers, whose interest is aroused now, refer for example to the books of Haupt et al. (2009), Klir & Yuan (1995), or Ross (2010).

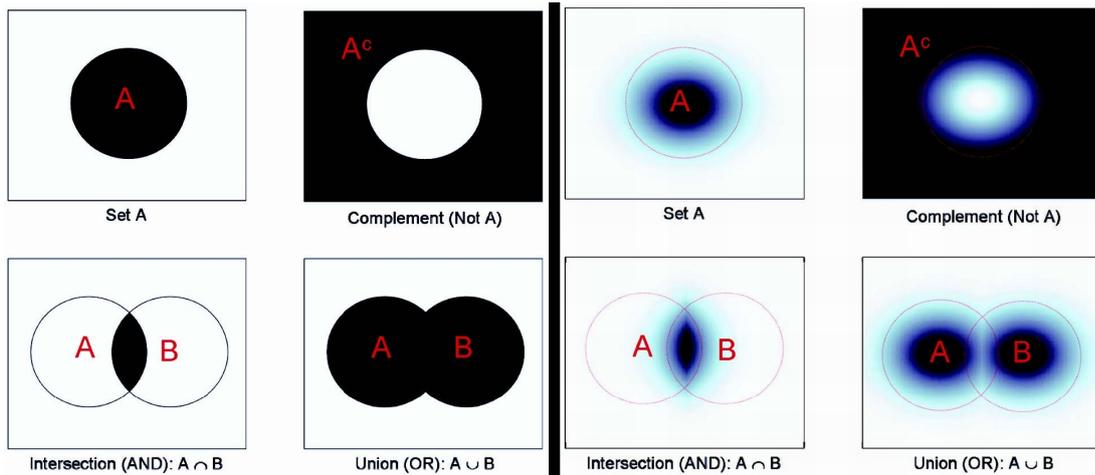


Figure 3.12.: Left: Classical set operations with black shading representing elements in the resultant set. Right: Fuzzy set operations where the degree of membership in the resultant set is represented by shades of gray, with white points having membership 0 and black points having membership 1. Dotted red circles represent the 0.1 membership contours of the original fuzzy sets. Taken from Haupt et al. (2009).

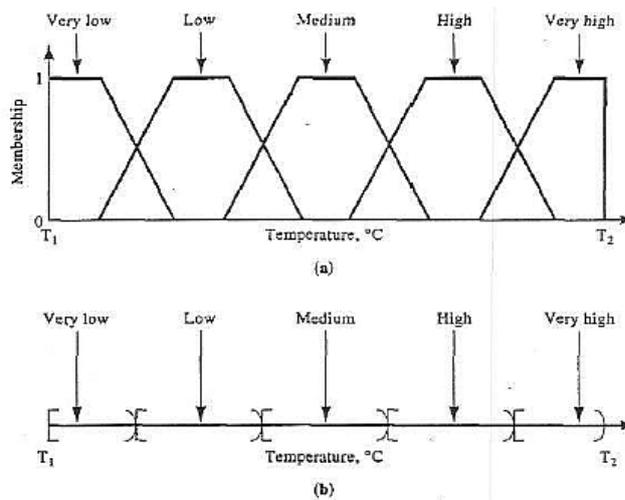


Figure 3.13.: Temperature ranging from T_1 to T_2 shown as a) fuzzy variable and b) "crisp" variable. Taken from Klir & Yuan (1995).

It is important to mention the difference between fuzzy logic and probabilistic

logic. 40 % in probabilistic sense would explain that the likelihood (or probability) that a proposition is true is 40 %. In fuzzy logic it means that a proposition is to 40 % true (and to 60 % not true).

The combination of data with fuzzy logic can be described by three main steps:

- Definition of the fuzzy sets for each data
- Declaration of a rulebase for the combination
- Defuzzification

In the following these three steps and thus the data fusion used to build a fuzzy logic based expert system for CI will be explained in more detail.

Definition of the fuzzy sets for each data:

As a first point you need to define fuzzy sets for each data type you want to use for the combination. The aim here was to use data which gives information about the basic ingredients mentioned in Chapter 2: moisture, instability, and lift. By the analysis that will be shown in Chapter 4 one data type out of the different available data sources is chosen for each of the ingredients which describes best if there is a CI forcing expected or not. Thus there are fuzzy sets containing the information if we expect stable, neutral, or unstable conditions; dry, medium, or wet conditions and downward motion, neutral conditions, or lift. This means for each of the ingredients we analyze a field to find out if the conditions are pro CI, neutral, or contra CI (see Fig. 3.14).

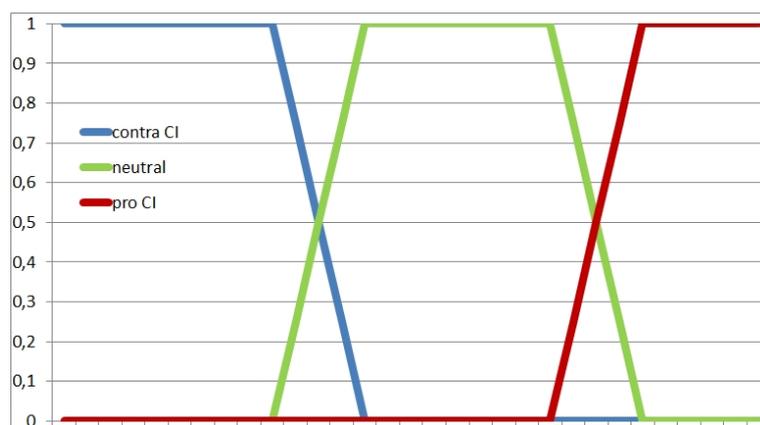


Figure 3.14.: Example shows fuzzy sets for an arbitrary data source used for the CI detection. The set marked blue is an area where no CI development is expected, green is neutral and red is a set where CI forcing is expected. For further description see text.

Besides these fuzzy sets as input, some output fuzzy sets have to be defined as well. For CI these sets describe five stages of CI forcing (instead of three stages for the input sets): very low, low, medium, high, and very high CI forcing are the output sets.

Declaration of a rulebase for the combination:

For the combination of the data a so called rulebase is needed. This is a multitude of rules that defines memberships within the different output fuzzy sets for the possible combinations of "states" of ingredients. Here these rules are build based on expert knowledge and statistical information gained from the verification results. These rules are IF-THEN relationships like:

IF $data1(proCI) > 0$ AND $data2(proCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "very high"* > 0

or:

IF $data1(contraCI) > 0$ AND $data2(contraCI) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "very low"* > 0

The standard use of AND in fuzzy logic means, if two values x and y are true (membership grade > 0) then the combination with both being true has the lower membership grade:

$$\text{truth}(x \text{ and } y) = \text{minimum}(\text{truth}(x), \text{truth}(y))$$

In our case we combine three different data sources with three different categories in the input fuzzy sets. Thus we have $3 \times 3 \times 3 = 27$ rules for different combinations and they are clustered, as mentioned above, in five resulting "CI forcing" sets. Each rule assigns a membership degree for the combination to one of the output fuzzy sets. The assignment of input to output sets by the rules is visualized in Tab. 3.5. The arrays shown in the table indicate if the input set components of the three different data sources are pro CI (+), neutral (0), or contra CI (-).

					CI forcing				
very low		low	medium	high	very high				
-	-	-	0	+	+	+			
-	-	+	0	-	+	+			
		($\times 3$)		($\times 3$)					
		-	0	+	+	0	($\times 3$)	($\times 6$)	($\times 3$)
		0	+	0	+	0	($\times 3$)	($\times 3$)	($\times 3$)
		0	+	0	+	0	($\times 3$)	($\times 3$)	($\times 3$)
		($\times 3$)		($\times 3$)					
					no. of rules per output fuzzy set				
very low		low	medium	high	very high				
1		9	7	9	1				

Table 3.5.: The scheme shows how the input fuzzy sets build up the output fuzzy sets. Each input might have components pro CI (+), neutral (0), or contra CI (-). The value in parentheses shows the amount of permutations available of the shown array. For more explanation see text.

The CI forcing is classified as "very low" if all additional data sources are "contra CI", "low" for more contra than pro, "medium" for an equal amount of contra and pro, "high" for more pro than contra, and "very high" for pro CI conditions in all the additional information. If you look at possible realizations of, for example, two pro CI and one neutral input, there are three possible permutations of this array within the rulebase. The neutral forcing might be in each one of the three additional fields. This amount of permutations is given in parentheses in the table.

Usually you have more than one rule per set, getting an array of membership grades. The combined membership grade for the set is the maximum value of the array of membership grades. The output fuzzy sets are illustrated for one example in Fig. 3.15 and the whole rulebase is added in Appendix C. A detailed example with the input fuzzy sets and the resulting output fuzzy sets will be presented in Chapter (5.3)

Defuzzification:

Defuzzification takes the fuzzified results and converts them to a crisp value as final result of the combination. The defuzzification technique used is the center of gravity method. There, the triangles or trapezoids of the output fuzzy sets are chopped off by a horizontal line marking the membership grade for each set as shown by the hachures in Fig. 3.15. Then the resulting trapezoids which are marked with hachures are taken as one single geometric shape and the x coordinate of the calculated centroid is the defuzzified value (6.7 in Fig. 3.15). The interpretation of the resulting value of the combination will be described later in Chapter 5.

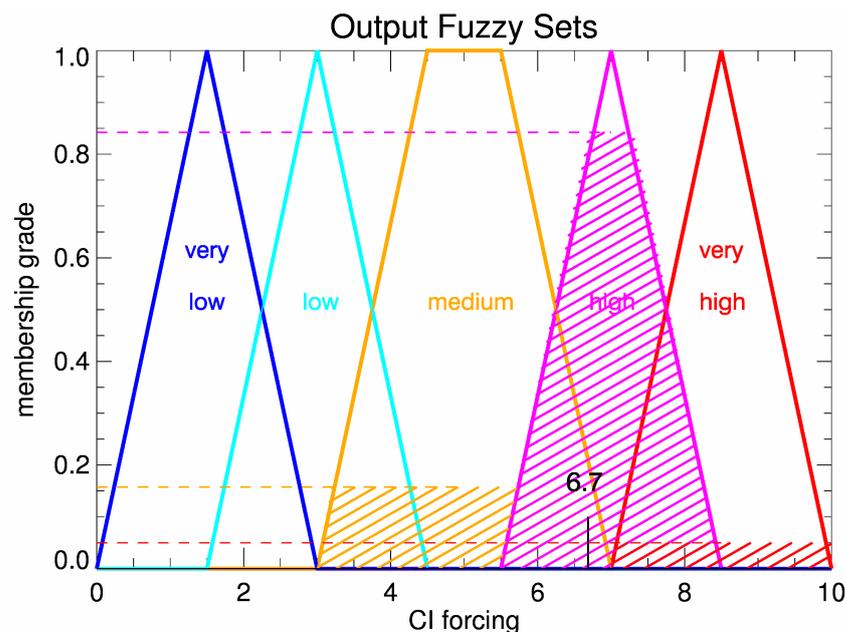


Figure 3.15.: Image shows the output fuzzy sets for an arbitrary example. For further description see text.

4. Analysis of the existing CI detection and nowcasting in Cb-TRAM

For being in a position to declare the use of additional data for CI nowcasting as beneficial, and to underline this statement with some hard facts, a reasonable methodology for the analysis of the existing Cb-TRAM CI detection and nowcasting and the development of both by the use of additional data sources is essential. Within this Chapter the evolution of the CI verification methodology newly developed for Cb-TRAM within this thesis will be described. Some slightly different verification concepts which are suggested to be appropriate for CI verification will be presented and discussed in parallel in Section 4.1. This leads to the results of the methods presented in Section 4.2. Based on these results one of the shown verification methods will be used for the comparison between the performance of Cb-TRAM alone against the combined information of Cb-TRAM and the additional data in Chapter 5.

4.1. CI verification methods for Cb-TRAM

This Section aims at the description of the verification approach used to gain verification results shown and discussed later within this thesis. For background information on forecast verification in general, CI verification in detail, and descriptions of the verification specific nomenclature the reader is referred to Section 3.2.

A key point of the developed verification method is that it is not aimed at the detection with respect to the full lifecycle of each individual stormcell, giving one verification result for the whole lifetime of the cell. The focus is on the usefulness of the CI nowcasts. Thus the verification is performed for all CI detections per time step where the nowcasts are launched. At each Cb-TRAM analysis time where CI detections exist, a nowcast for the next 15, 30, 45 and 60 minutes for the expected position of the detected developing cloud is available. For each of this nowcast time steps the nowcast will be compared to the corresponding analysis as illustrated in Fig. 4.1. This results in four contingency tables with hits, misses, and false alarms for each of the above mentioned nowcast lead times for every time step. Now the question is: What is a hit, miss, or false alarm? How are they defined? Correct negatives will not be evaluated for reasons described in Section 3.2.

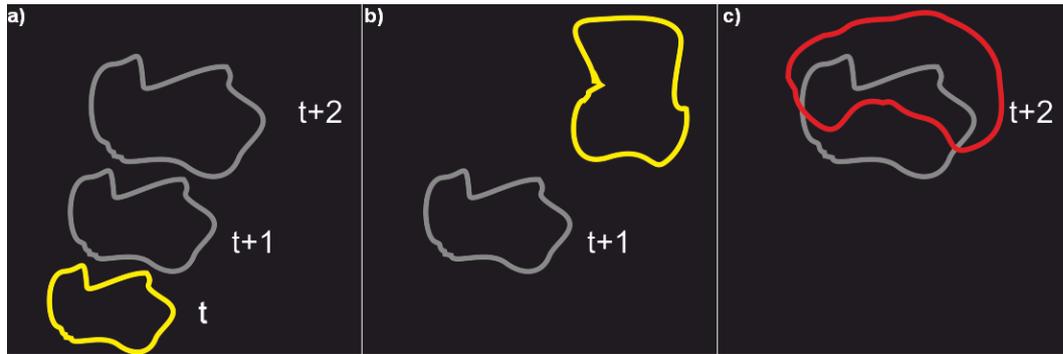


Figure 4.1.: Image a) shows a Cb-TRAM CI detection at time step t (yellow contour) with its nowcasts for the next two time steps $t+1$ and $t+2$ (gray contours). Image b) shows time step $t+1$ with the corresponding nowcast of the CI cell (gray) and a new CI detection and c) shows time step $t+2$ again with the corresponding nowcast of the first CI cell (gray) overlapping a mature cell (red).

Independent of pixel or object evaluation two possible ways of evaluation exist here. One possibility is to allow only developments to stage 2 or stage 3 detections as hits, the second possibility is to allow also long lasting CI cells as hits. Thus, explained for pixels at time $t+x$, if a pixel in the nowcast is part of a CI object and in the analysis it is part of a detection too, meeting the criteria (with or without long lasting CI stage), a hit is counted. If there is a CI pixel in the nowcast but not in the analysis it is counted as false alarm and if there is a pixel from a stage 2 or 3 detection in the analysis but no (CI) detection in the nowcast, the pixel will be counted as miss. For a visualization see also Fig. 3.4a) & b) or Fig. 4.1. The objects of stage 2 and 3 in the time step t where the nowcast is launched, which can be tracked into the analysis time step $t+x$ and can be assigned to the same cell ID like in t , will not be evaluated for the CI analysis here.

Based on the background on verification given earlier in Section 3.2 different approaches are tested to find the appropriate method. First of all a classical **pixel based** approach for filling a contingency table was performed to use it as a reference for comparison to the other approaches. This means every nowcast and the according analyzes are compared pixel by pixel. As already explained earlier, pixel based methods are very harsh and require an exact match between forecasts and observations, what can definitely not be expected for the CI nowcasts. Thus the pixel based approach is definitely inappropriate for the task of CI verification. It is performed and mentioned for the sake of completeness but the results will not be shown in detail.

A much more suitable approach is to use an **object based** verification method for the CI detection polygons and their CI nowcasts. A first object based approach would be to use the internal matching in Cb-TRAM itself. Every detection has a **cell ID** which stays unchanged for the next time steps if a cell is recognized as a follow-up detection, by object overlap, in the later time steps. Thus if a cell ID of a yellow CI detection cell exists in a later time step for a stage 2 or stage 3 cell, it could be counted as hit. This approach would be quite easy to realize, but it has

a key problem: the merging and splitting of cells. First of all, the CI detections can quite often be observed as a small cluster of more than one developing cloud where several independent cells are detected. During the growth of the clouds to the next development stages these "clusters" often merge and only one of the IDs will be kept for the newly developed cell. This could lead to several false alarms for the cell IDs which will not exist anymore. Also the splitting of cells in detection stage 2 or 3 would cause trouble for this ID based verification. The smaller part of the cell would get a new cell ID after the split. If only the IDs are evaluated these ones would be counted as miss. Additionally, for the case where one yellow cell overlaps with two later on developing cells, the ID would fit to only one (or even none) of the developing cells and the other one would be counted as miss.

To overcome these issues an **object overlap** based approach is introduced, where for any CI detection the nowcasts are compared to later on existing detections. And a simple object overlap is the criteria for defining a hit. CI nowcast objects which do not overlap with any detection of the following time steps are counted as false alarms and stage 2 or 3 detections which are not preceded by a CI detection are counted as miss. As already mentioned above, this is not done with focus on the overall lifetime of a cell but with focus on usability of the nowcasts. This means for the nowcast time steps +15, +30, +45, and +60 minutes from the CI detection the overlap criteria is evaluated and for each lead time verification scores will be calculated (see below). As additional information also a so-called **accumulated evaluation** (acc) will be calculated. The accumulated version is less focused on the verification results for a precise leadtime (of 15, 30, 45, or 60 minutes), instead it shall help to gain information if the nowcasts produced for the next hour prove to be true at least at one point of time within this hour, independent of the exact leadtime. This means, if a CI cell overlaps with a later on existing detection at any time step between 15 and 60 minutes lead time it will be counted as hit, evaluating all hits within the one hour nowcast lead time of a current detection. CI cells which do not overlap with other detections anytime within this lead time time-frame will be counted as false alarms and cells which are developing to orange or red cells within this time frame without a premonition by a CI detection are counted as misses.

The overlap criterion allows for small imprecision in location and shape of the extrapolated nowcasts. If a CI cell nowcast position does not overlap with a later detected object between stage 1 and 3 this might have different reasons. First of all the detection might have been wrong. The second possibility is that the nowcast was wrong and the cell moved faster, too slow, or too far to the left or right from the expected track to overlap with the previous expected position. In addition a third possibility has to be considered too. Especially for longer lead times like 45 or 60 minutes it might happen that the detection was right, but the cell might have diminished or might have decayed already.

To overcome errors in the extrapolated nowcast position for correctly detected cells two variations of the verification algorithm are presented. The aim is to reduce the double penalty problem where nowcasts objects close to but not overlapping with an analysis object are counted as false alarms and the analysis object

is counted as miss. The first idea is to allow a kind of **search radius** around the objects (see Fig. 4.2) which grows with time. For the 15 minute nowcast 20 km around the nowcast object are added as area where an overlap is still be evaluated as hit. For the 60 minute nowcast this radius is up to 50 km. This is an approach of a fuzzy verification method.

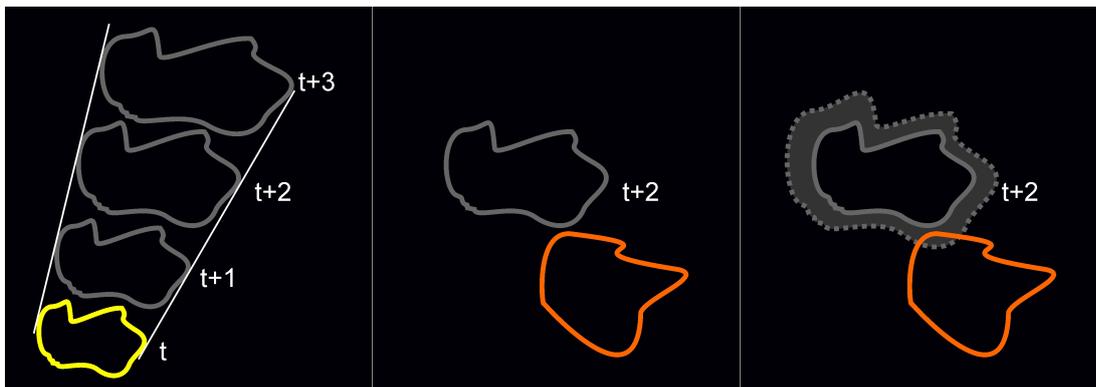


Figure 4.2.: Left: Cb-TRAM CI detection at time step t (yellow contour) with its nowcasts for the next three time steps $t+1$, $t+2$ and $t+3$ (gray contours). Middle: Time step $t+2$ with the corresponding nowcast of the CI cell (gray) and a new stage 2 detection nearby (orange). Right: Time step $t+2$ with a search radius around the contour of the CI cell (shaded area between contour and dotted line) now overlapping with the nearby detection.

The second variation should help to overcome the problem of "too fast moving" nowcasts which can sometimes be observed quite clearly for quasi stationary development. Quite often, especially for storm development triggered in mountainous areas, the yellow cells start bubbling and the overall movement of the surrounding air creates a normal nowcast with different positions for the time steps away from the original cloud object. These positions are due to extrapolation with the cloud motion vectors gained by digital image processing within Cb-TRAM. Thus, these "too fast moving" nowcasts are due to inaccuracies in the Cb-TRAM image matching process and they are dependent on the overall cloud coverage and movement within the observed scene. However, in these quasi stationary cases the clouds keep bubbling at a more or less unchanged position and after a development they move very slowly. This leads to storms moving much slower as originally expected. Therefore the old nowcast positions are used as well as so-called **nowcast track** (see Fig. 4.3). This means, for example, for the 45 minute lead time nowcast of time t the polygons of the 45, 30, and 15 minute nowcast as well as the position of the original detection are compared with the new stage 2 or 3 detections at $t + 45$.

Both of these variations, the search radius and the nowcast track, will only be evaluated for developing cells and will not be applied for overlap with (long living) yellow cells. A detection of a yellow cell in $t + 45$, overlapping with one of the nowcast polygons inside this track of polygons which form more or less a whole sector from the original detection at t , is a very weak criteria for a hit and would

make not much sense here. The same argument votes against the evaluation of long living cells with the search radius around the nowcasts.



Figure 4.3.: Left: Cb-TRAM CI detection at time step t (yellow contour) with its nowcasts for the next three time steps $t+1$, $t+2$ and $t+3$ (gray contours). Middle: Time step $t+3$ with the corresponding nowcast of the CI cell (gray) and a new stage 2 detection nearby (orange). Right: Time step $t+3$ with its "nowcast track" (all older positions of this detection in gray) now overlapping with the nearby detection.

The overlap criterion does not use any minimum overlap of a specified amount of pixels within the object or a percentage of its area, due to the fact that the CI objects usually have a small size compared to the areas covered by mature storms. One drawback of the overlap criteria used is that it is not a "strictly proper" forecast evaluation procedure (Wilks, 2006). It might encourage to so-called hedging. Hedging means forecasting some future weather events in order to achieve a better score but not necessarily to get a better or more realistic forecast. For this score it means you might raise your possibility to overlap with later on existing cells by simply enlarging the CI objects, which can be done up to unrealistic sizes, just to get a better verification result. The bigger the CI objects the likelier they will overlap with future "rapid development" or "mature storm" objects. In contrast, for small objects the probability of a random overlap between nowcast and later existing stage 2 or stage 3 objects is quite small.

Which scores are now used for the ongoing analysis? With these different types of overlap criteria contingency tables can be filled for each nowcast lead time, and the accumulated evaluation, for each time step. To keep it quite easy understandable common measures like the probability of detection and the false alarm ratio are used for the evaluation of these objects. Adapting these measures, used very often, helps to keep the results comparable to other studies. The key difference is the definition of hits, misses, and false alarms for filling the contingency table and not a fancy new calculation with the table entries. To clarify the difference to the usually pixel based versions of these scores they will be abbreviated as oPOD (object-based probability of detection) and oFAR (object-based false alarm ratio). Besides that, they are calculated, as described

already earlier (see Chapter 3.2), as

$$oPOD = \frac{hits}{hits + misses} \quad (4.1)$$

and

$$oFAR = \frac{false\ alarms}{hits + false\ alarms}. \quad (4.2)$$

More information on the concrete calculation of the scores will be given along with the results for the Cb-TRAM CI stage verification within the next Section.

4.2. Performance of Cb-TRAM CI stage

The performance of the Cb-TRAM CI stage is evaluated for 86 days in summer 2009. These are all days within the summer period from May 15 2009 until August 31 2009 where no bigger data gaps during daytime in the rapid scan satellite data or one of the additional data sources obscures the evaluation. The restriction that only time steps during daytime are evaluated is due to the need of availability of the HRV channel for the Cb-TRAM CI detection. The evaluation starts at least half an hour after the daytime detection with the HRV channel is available over the whole analyzed area, from that time Cb-TRAM is able to determine the needed timetrends in the HRV properly. Furthermore it ends for a time step where one more hour of daytime detection data is available, to be able to check for long living yellow cells, which will not be detected anymore after sunset. Also days without or with only few noteworthy convective developments above Europe are within the analyzed days which is an important difference to many other verification approaches for CI tools. Days where over large areas first cumulus development can be observed from the satellite view but the overall forcing is too weak to finally initiate usually show a larger amount of false alarms.

The values presented within this Section are mean values over all time steps and days. The verification routine computes oPOD and oFAR for each of the used daytime time steps and calculates a mean value for the whole day. Afterwards the 86 daily mean values are used to build a mean value for the performance. Table 4.1 shows the results for the 15, 30, 45, and 60 minute nowcasts as well as for the accumulated evaluation (acc). "Accumulated" means overlap of nowcast and analysis in at least one of the analyzed time steps (15 to 60 minute nowcasts) as criterion for a hit with no focus on any special nowcast time step, as described earlier. The results for the cell ID based evaluation are shown as well. The first two rows show the results allowing (long living) yellow cells along with stage 2 or 3 detections as hits too, the latter rows show the results if further development to stages 2 or 3 (dev) is required for a hit (oPOD dev and oFAR dev).

As shown in Table 4.1 the object based probability of detection oPOD for the 15 minute nowcasts is close to 60 % (0.592) which means that in a statistical mean 60 % of the cells detected in the analysis, which were not already in stage 2 or 3 15 minutes earlier, are detected as CI cells at least 15 minutes earlier. The object based false alarm ratio oFAR is slightly above 60 % (0.611). The values for oPOD

decrease for the following nowcast lead times from 0.421 (30 minute nowcasts) to 0.238 for the 60 minute nowcast. Thus the 60 minute nowcasts overlap with slightly less than 25 % of the cells detected one hour later. The oFAR rises up to 85 % (0.852) for the 60 minute nowcasts. This means that 85 % of the 60 minute CI nowcasts will not overlap with a cell existing in the analysis one hour later. The accumulated evaluation shows that 40.9 % of the cells detected within 15 to 60 minutes from the time where the CI nowcast is launched are successfully hit by a nowcast. Again this includes only cells which are not already within stage 2 or 3 when the nowcast is launched. The score based on the cell ID, which is calculated for the same period as the accumulated evaluation, is a little lower (around 2 %). This is for example due to cells where one bigger stage 2 or 3 cell develops from more than one small CI cell where the cell ID of only one of the merged cells is kept but they all overlap with the later existing cell. For the oFAR the accumulated value, which describes CI nowcasts which never overlap with another cell in the evaluated time span of 15 to 60 minute lead time, is 54.5 %, where the cell ID based value is worse by 5 % (0.593). The available five and ten minute nowcast are not included in the calculation of the accumulated score. Due to the fact that here also overlap with cells in the CI stage is counted as hit and no development to stage 2 or 3 is needed as a hit criteria the high probability that these nowcasts overlap with 5 or 10 minute later detected ones would result in a not quite meaningful calculation of the scores.

	15 min	30 min	45 min	60 min	acc	ID
oPOD	0.592	0.421	0.310	0.238	0.409	0.387
oFAR	0.611	0.755	0.821	0.852	0.545	0.593
oPOD dev	0.228	0.199	0.174	0.154	0.170	0.091
oFAR dev	0.885	0.884	0.886	0.887	0.818	0.873

Table 4.1.: Statistics on the Cb-TRAM CI stage verification. For description of the nowcast lead time, accumulated, and cell ID based scores, and for the different types of oPOD and oFAR (without or with further development required) see text.

The lower half of Table 4.1 shows the same calculations for the case that a development from a stage 1 nowcast to a stage 2 or 3 detection is required for a hit. The oPOD dev is thus lower than the corresponding oPOD and sinks from 22.8 % for the 15 minute nowcasts to 15.4 % for the 60 minute nowcasts. The accumulated oPOD dev is 17.0 % and the difference to the cell ID based score is quite striking when development is required, its oPOD dev value is only 9.1 %. The oFAR dev lies above 88 % for all single nowcast lead times evaluated. The accumulated oFAR dev is 81.8 % which is again better than the cell ID based score of 87.3 %.

The accumulated version (acc) of the scores with further development to stages 2 or 3 (dev), marked blue in Table 4.1, will be further investigated in the following. As described in Section 4.1 two methods to incorporate a correction for eventually existing errors in the nowcasting, and not in the detection, where developed.

Table 4.2 shows the results for the accumulated evaluation (acc) if on of these methods, a search radius around the objects (acc_{radius}) or the use of the nowcast track (acc_{track}), is applied.

While the search radius, growing with lead time, is capable of rising the accumulated oPOD dev to 20.3 % the oFAR dev sinks to 79.2 %. The incorporation of the nowcast track performs slightly worse for the oFAR dev with a result of 81.0 % which is quite close to the original value without the track. Regarding the oPOD dev the track is rising the value to almost 25 % (0.249). This means the nowcasts produced for the next hour successfully hit one out of four cells developing to stage 2 or 3 anytime within the next 15 to 60 minutes. Additional discussion about this value will be presented in Chapter 6, but an oFAR dev of around 80 % is far too much despite of any possible discussions to relativize this value.

	acc	acc_{radius}	acc_{track}
oPOD dev	0.170	0.203	0.249
oFAR dev	0.818	0.792	0.810

Table 4.2.: Intercomparison of the Cb-TRAM CI stage verification for the accumulated evaluation (with development of the nowcasts to stage 2 or 3 detections) with both corrections for nowcasting errors – search radius and nowcast track.

This large value for the oFAR dev is the key motivation why the additional data should be used to filter as much as possible of the false alarms while losing as few as possible hits to not reduce the oPOD dev. Furthermore, in cases where a strict filtering is not possible, the incorporation of the additional data should at least give a kind of confidence level for the produced CI nowcasts. This confidence level should evaluate how likely, with regard to the used data, a further development of the detected CI cell to the next detection stages is. Thus this confidence level, gained by postprocessing of the Cb-TRAM detections with the additional data, shall give additional information on the reliability of the original Cb-TRAM CI detection.

For comparison of the performance with and without the additional data the accumulated measures incorporating the nowcast track will be used in the following Chapter (acc_{track} – marked blue in Table 4.2). It is preferred due to the best oPOD dev value of the accumulated scores. The small differences in the oFAR dev, which shall be reduced by the postprocessing anyway, are less important for the decision. Further description why this measure is used will be given at the beginning of the next Chapter.

5. Combination with additional data sources

Within the following Sections the combination and the data analysis for each of the ingredients - moisture, instability, and lift - will be described. As explained earlier, the object-based, accumulated verification approach for developing cells with use of the nowcast track (acc_{track} after Ch. 4) is applied to classify each CI cell as hit or false alarm. Due to the fact, that the analysis for the ingredients is performed exclusively for CI cells, misses are not evaluated. They would concern stage 2 or stage 3 objects. In the following, the CI cells, classified as hit or false alarm, will be checked for lightning occurrence within the cell in the last five minutes before the detection time step (Section 5.1), and for each cell mean values per object of each one of the possible additional data is calculated (Section 5.2). The gained statistics will be presented for the data used later on as part of the fuzzy logic data combination. Furthermore, they will be used to describe why some of the additional fields are more useful than others. The development of the fuzzy sets for each of the additional data used for the combination will be described as well. Finally, the results of the combination will be given (Section 5.3).

The analysis of the additional data is not performed for each Cb-TRAM daytime detection time step within the analyzed 86 days but only hourly, due to the one hour lead time of the nowcasts. The use of all the 5 minute rapid scan CI detections and not the hourly data would cause some problems in the analysis of the additional data. The additional data shall be examined for each cell, building up statistics for the additional data clustered in hit cells and in false alarm cells. In case of the 5 minute rapid scan time steps as analysis times, long-living (and eventually non-developing) CI cells would be weighted much more than fast developing or fast decaying CI cells. If every available rapid scan time step would be analyzed the cells which live longer than one time step would get a mean value for each of the additional data at each time step. A cell which is detected as a CI cell in consecutive time steps and does not develop to stage two or three, shall not be incorporated in the statistics more often than for example a cell which needs only one time step to develop from the CI stage to stage two or three. These multiple counts for one cell are acceptable in a verification approach which should judge the usefulness of the produced nowcasts for each time step, but not for an statistical approach to get information about the possibility to use additional data to raise the detection quality. Analyzing only hourly data will not fully delete these multiple counts but it is sufficient to reduce the amount of multiple counts to a minimum for only few long-living cells. The exact leadtime of the nowcast is not crucial for this analysis, thus the accumulated evaluation (acc_{track}) is chosen. Cells living more than one hour in the CI stage shall not be

treated as one continuous CI event as will be discussed in Chapter 6.

Some Figures which do not describe the later on used additional data, but are only examples to clarify statements on possible data that showed to be less useful to evaluate the ingredients for the CI forcing, are not included within this Chapter. These images are added in Appendix F and will only be mentioned briefly within this chapter.

In summary, more than 34000 CI cells are evaluated for the following statistics. The sample includes 6169 hits and 28137 false alarms. Due to the change in data set size and no time averaging this would lead to an oFAR dev¹ of 0.820 if we use the above given values for the sample to calculate an oFAR dev (slight difference between this value and the one in Tab. 4.1 is due to the different amount of time steps - 5 minute data to hourly data - in the sample used for the calculation of the mean value). Each of these cells is checked for lightning occurrence within the cell. Furthermore, the mean values per CI object of each of the additional data are calculated to evaluate the three above mentioned ingredients.

As short clarification of the wording used within the text, data source always describes the "origin" of the data, thus data sources are satellite, lightning, surface, and model data. Ingredients, are the mentioned ingredients for the development of a thunderstorm, moisture, instability, and lift. Data, additional data, data type, etc. are used to describe the different physical values from the data sources, for example θ_e , MFC, or ω . Indicator is used for the data types which indicate the availability of one of the ingredients.

5.1. Lightning data

As already discussed in Section 3.3 CI cells which include lightning signals are classified wrong. The cell development is underestimated for these cases and the filtering shall only be done for the case of CI overestimation. Only cells which will not further develop and will not start to produce lightning shall be rejected. An example for the case of a CI cell including lightning is given in Fig. 5.1, which shows southern Germany.

The displacement between the yellow polygon and the visible cloud contour in Fig. 5.1 is an example for the so-called parallax shift due to the satellite angle of view and the cloud height as illustrated in Fig. 5.2. In line with the example in Fig. 5.1, I will briefly explain what is done by the parallax correction which is included in Cb-TRAM and which has to be incorporated for the whole analysis of the additional data. Without the correction the CI contour would fit quite nicely to the cumulus cloud which is shifted to the north and very slightly to the east of the polygon. The cloud top has a certain height above ground. Due to the angle of view from the satellite position (see Chapter 3.1) at 9.5 degrees east in about 36000 km above the equator, clouds over the analyzed area of this study appear to be further north than they are in reality. The shift depends on the cloud top height and the latitude of the cloud pixels. The lateral shift within

¹oFAR dev = accumulated object-based FAR using the nowcast track, with cell development as a hit criterion (as introduced in Ch. 4)

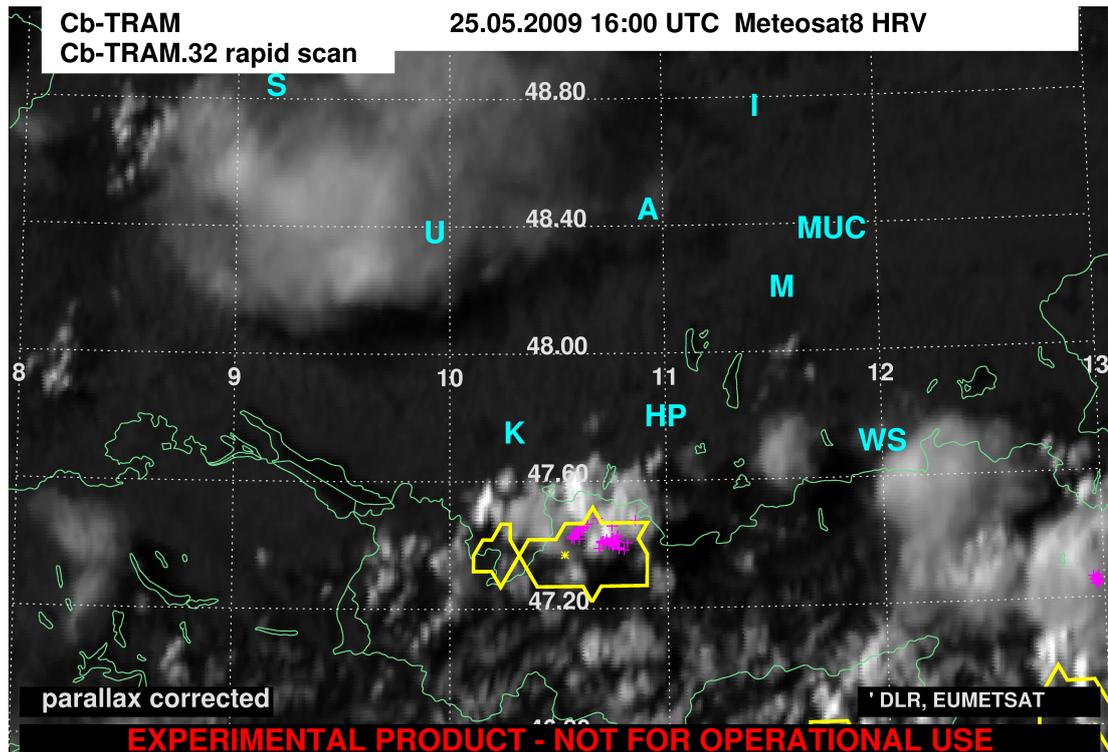


Figure 5.1.: This plot shows an example plot of Cb-TRAM where Cb-TRAM CI detections (yellow cells) include LINET lightning measurements (magenta plus signs).

this study is a lot smaller, especially for the example in Fig. 5.1, due to the very small difference in longitude between the cloud pixels and the satellite position. The parallax correction compensates this latitudinal and longitudinal shift. The parallax corrected coordinates of the polygon have to be used as the real position at the surface for the CI detection within the analysis of the additional data. Following the calculation for a thunderstorm over Austria in Radová & Seidl (2008), the parallax shift for a cloud top height between 10 to 15 km in Central Europe is 15 to 25 km. The parallax positioning error is quite small in comparison to the scales of change of some of the additional data tested and thus would often cause only small errors, but especially for the very precisely located lightning information it has to be taken into account.

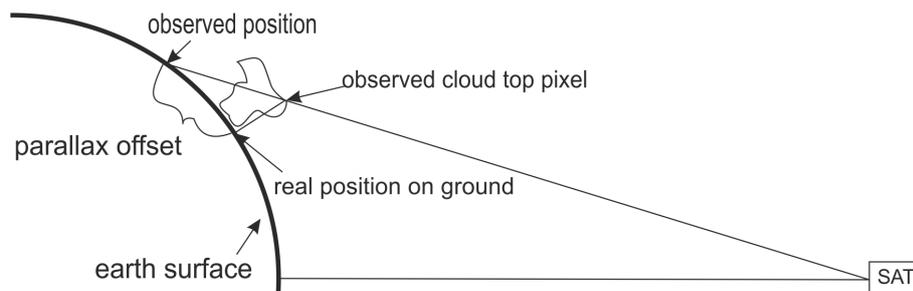


Figure 5.2.: Schematic showing the parallax shift due to the satellite angle of view and the cloud height.

In the whole analyzed 86 day sample almost 1500 CI cells include lightning signals. Half of them (735) do not develop to stage 2 or 3 detections in Cb-TRAM. Thus these would be evaluated as false alarm without the lightning information. In the following, after incorporation of the lightning data, a cell is classified as hit if either

- it develops to a Cb-TRAM stage two or three detection,
- it includes lightning already in the CI stage, or
- both.

Thus after incorporation of the lightning data the amount of hits is raised to 6904 and the amount of false alarms in the sample sinks to 27402 (after acc_{track} method – Ch. 4). These values result in a reduced oFAR dev of 0.799 instead of 0.820 without the lightning information. Furthermore, oPOD dev increases from 0.249 to 0.271.

5.2. Surface observations & model data

In the following Subsections surface observations processed by VERA and model data from COSMO-EU are used to find data which can be used as indicator for the different ingredients (moisture, instability, and lift). Background information on VERA and COSMO-EU is given in Ch. 3.3

With the aim to enable reasonable filtering of satellite based CI false alarms, additional data which helps to discriminate between hits and false alarms is presented for each of the ingredients. This is done in line with an explanation of the developed fuzzy set for the data which will be used for the combination. In cases where more than one data type is tested as proxy for this ingredient, further descriptions, why the one chosen is the better one for the combination, are given.

Main discussion points on the fuzzy sets per each data, like the choice of the transition zones and the sensitivity of this choice, or the final combination and its use for possible customers, will be considered in more detail in the discussion in Chapter 6.

5.2.1. Indicating low-level moisture

As already described earlier (see Chapter 3.3) two different VERA output products indicating moisture at the surface are analyzed here. These two are the Moisture Flux Convergence (MFC) and the equivalent potential temperature (θ_e). First of all, for each of the CI cells within this analysis a respective mean value of MFC and θ_e is calculated. In case of a small CI cell which does not incorporate a grid point of the additional data, the size of the object is gradually raised, symmetric around its center, until it incorporates a grid point of the additional data. This method is applied for the two additional data sources with lower resolution than the used satellite data, VERA data (8 km resolution) and the COSMO-EU model data (7 km resolution).

Furthermore, to get an optical impression of the data, plots displaying the VERA information along with lightning data, Cb-TRAM output, and the verification information are produced. Two examples for θ_e are shown in Fig. 5.3 (next page). The gray shading in the background represent the θ_e values while the LINET information is marked by magenta plus signs. The Cb-TRAM output is shown slightly changed to its earlier plots. The stage 2 (rapid development/orange) and stage 3 (mature/red) detections are represented as earlier. The CI or stage 1 detections are plotted in yellow (as usual) for the cells which are verified as hits and they are plotted in green for cells that represent false alarms.

Fig. 5.3 shows a common scene from May 25 2009, 15 UTC (left side) where it is at first glance quite difficult to detect any special features but the second scene from June 12 2009, 15 UTC (right side) indicates that low θ_e values (here mostly below 38°C in areas where CI cells occur) seem to correlate with false alarms.

In Fig. 5.4 all analyzed cells for the hits and false alarms which do not include lightning are categorized in two degree bins from lower than 24°C to lower than 98°C . The size of the bins is chosen freely to get a reasonable detailed distribution of the data. The hits per bin are normalized by the total amount of hits, and the false alarms per bin are normalized by their overall amount, too.

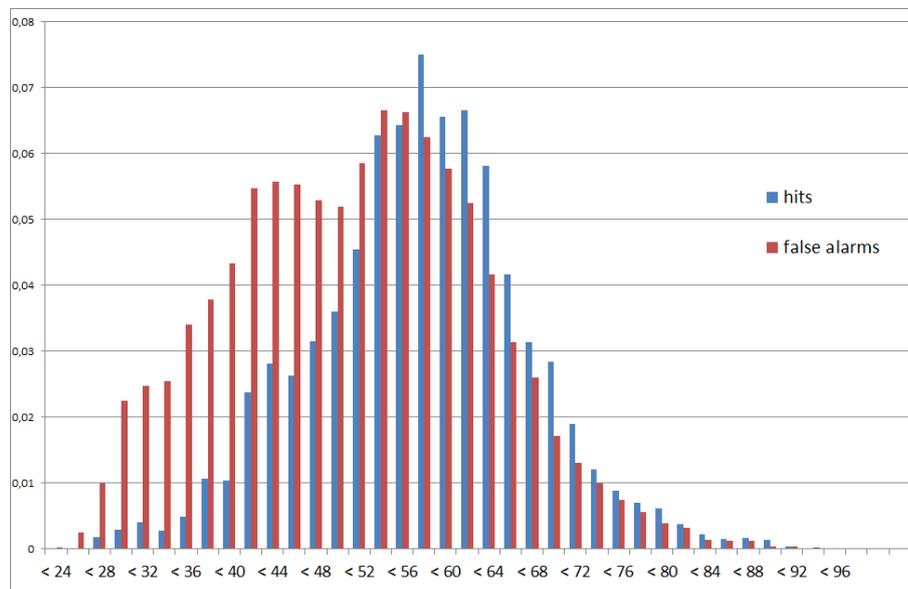


Figure 5.4.: The amount of hits(blue)/false alarms(red) per two degree bins for θ_e values between lower than 24°C and lower than 98°C normalized by the total amount of hits/false alarms. The statistics shown here and in the following plots of this type are calculated for the whole 86 day sample period.

From the analysis shown in Fig. 5.4 we can see that for cold θ_e values below 36°C the amount of hits per bin is always far below 1 % of the overall hits. In contrast the percentage of false alarms per bin is already above 2 or even 3 % for some of these temperature ranges. If we look at the sampled data, θ_e values

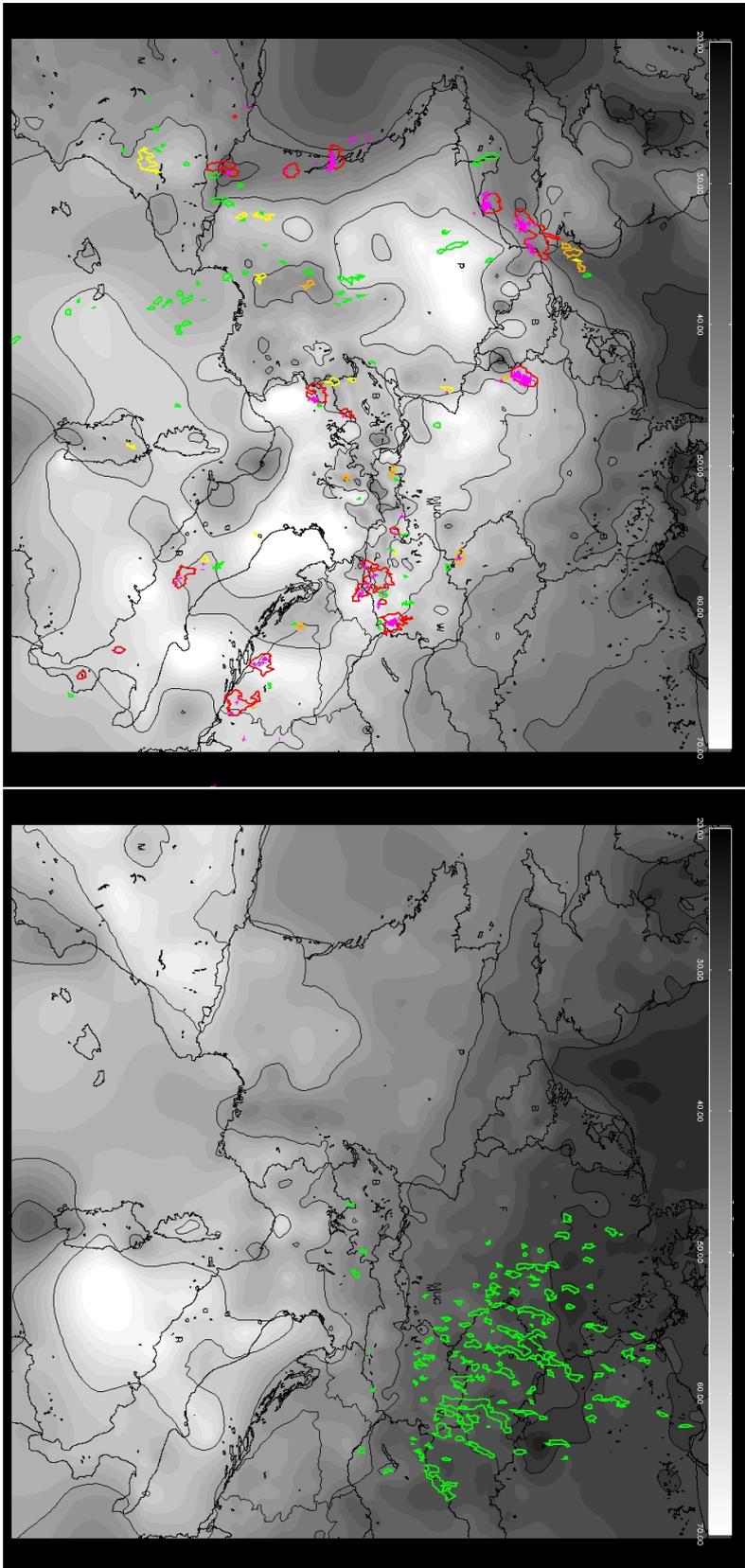


Figure 5.3.: Example plots of Cb-TRAM output and LINET data over VERA θ_e plots from May 25 2009, 15 UTC (left) and June 12 2009, 15 UTC. The gray shading represents θ_e values ranging from 20 to 70 °C. Dark gray for low and light gray for high (warm and moist air) θ_e . The yellow, orange, and red contours have the common colors as in Cb-TRAM, the green polygons show CI cells which do not develop to a stage 2 or 3 detection. Thus it shows CI cells which would be counted as false alarms.

below 36 °C can be found for 1.7 % of all hits (118 cells) but for 12.0 % of all false alarms (3311 cells). If we look at θ_e values below 41 °C, for example, we see that 4.8 % of all hits (340 cells) and 22.7 % of all false alarms (6284 cells) are colder than 41 °C. This cumulative percentage is shown in Fig. 5.5 calculated for all θ_e value frequencies below the given temperatures. The cumulative values for the false alarms indicate how many false alarm cells we could filter by introducing a simple lower boundary for θ_e . The corresponding value for the hits shows how many hits would be lost if such a strict threshold would be introduced. The hits do not add up to 100 percent because the cells including lightning are not integrated here. The cells with lightning account for 21,3 % of all hits and they would not be filtered anyway. The difference of the cumulative values peaks at 30.7 % for cells colder than 56 °C. This is the kind of border where in Fig. 5.4 the red bars are higher than the blue ones on the left side and the blue bars are higher than the red ones on the right side. Two thirds of all false alarms are colder than a θ_e of 56 °C, but also 35.6 % of all hits are below that 56 °C. Filtering here with a strict threshold would lead to an oFAR dev of 0.675 but would cost way to much hits and thus reduce the oPOD dev considerably.

The construction of a fuzzy set that resembles the expert knowledge gained by the analysis of the θ_e data now aims at building three classes, where, for one of them, further development of the cells detected as CI seems to be not very likely

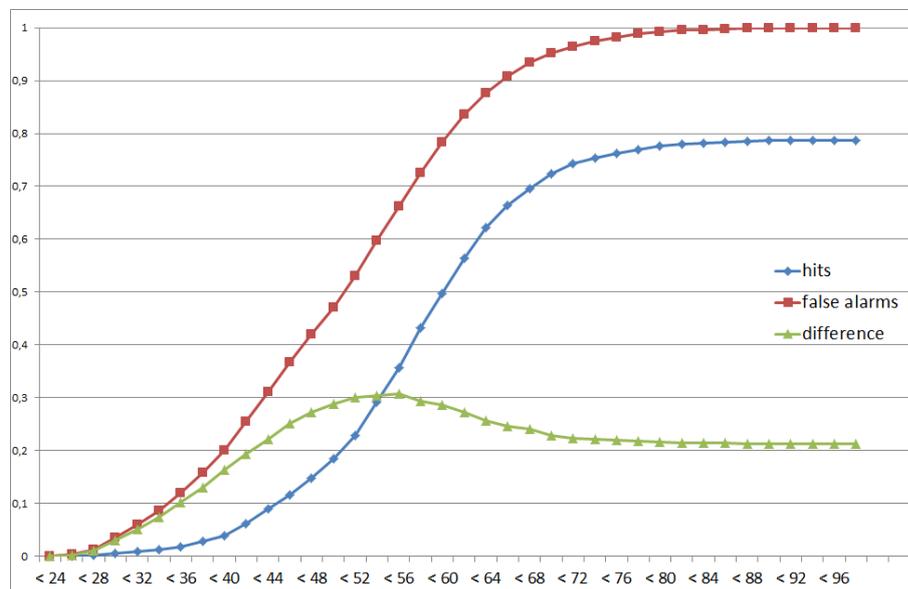


Figure 5.5.: Cumulative percentage of VERA θ_e value frequencies below the given temperatures for hits (blue), false alarms (red), and the difference of the two values (green). For further description see text.

(low θ_e), for one it seems to be a bad idea to filter the CI cells (high θ_e) and one lies in between these two (middle θ_e). The values where the difference in Fig. 5.5 is peaking seems good to find a transition from the range defined as middle to the range for high θ_e values. This transition is thus defined between 50 and 56 °C. The transition from low θ_e to the range in the middle is chosen by filtering almost no

hits at this boundary. Thus the transition is from 36 to 41 °C where only few hits but already quite a lot of false alarm cells can be found as mentioned above. This results in the fuzzy sets for θ_e shown in Fig. 5.6. A discussion on the influence of these choices on the final results will be added in Chapter 6.

θ_e is chosen as proxy for the moisture simply because the MFC is, after the same analysis, less useful for the filtering. An example of a moisture flux plot is given in the Appendix in Fig. F.1. The equivalent statistical plots to Fig. 5.4 and Fig. 5.5 can be seen there as well (Fig. F.2 and Fig. F.3). The statistics show less

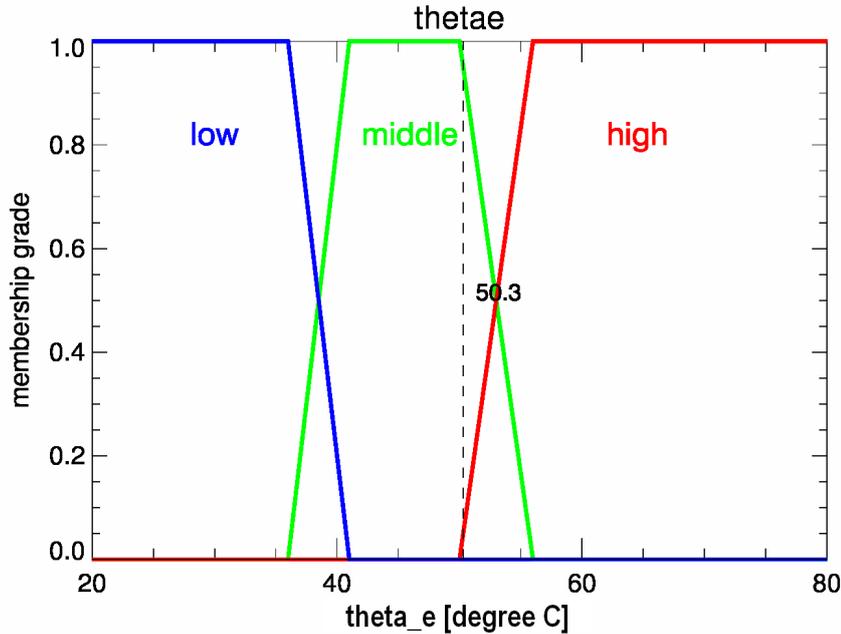


Figure 5.6.: VERA θ_e fuzzy set with an arbitrary example value. A θ_e value of 50.3 °C is some 95 % in the range of middle θ_e values and some 5 % in the range of high θ_e values.

difference between the cumulative values for false alarms and hits in both interesting ranges. The difference for the areas with negative MFC, equivalent to MFD, where no CI triggering is expected, thus where the normalized amount of false alarms is higher than the normalized amount of hits, is quite low. In addition the peak value is only slightly higher than the final difference value due to the incorporated lightning data, and is reached for an MFC value where already two thirds of the hits would be filtered.

5.2.2. Indicating instability

As indicator for instability one of the numerous instability indices is used. As already explained the chosen KO-Index describes the potential instability between mid levels (700 to 500 hPa) and low levels (1000 to 850 hPa) by calculating a difference of θ_e values for the four levels (see Equation 3.10). The low level value is again the θ_e gained by the VERA output, for the three more elevated levels

COSMO-EU data is used. Following the literature, KO values above six describe quite stable circumstances with low thunderstorm potential, values below two correspond to a high potential for thunderstorms. Results between two and six can be interpreted with a moderate potential for thunderstorms. This is summed up in in Tab. 5.1.

KO Index	potential for thunderstorm occurrence (airmass stratification)
$KO > 6$	weak (stable)
$2 \leq KO \leq 6$	moderate (indifferent)
$KO < 2$	high (unstable)

Table 5.1.: Common classification of KO-Index values (Kalthoff et al., 2010).

Within the following analysis the usual classification given in Tab. 5.1 is shifted to lower values as can be seen from Fig. 5.7. With the common thresholds almost all evaluated cells would be considered to be in the space of unstable atmospheric conditions, only a low percentage would be classified having a moderate potential for storm development and even less would be assigned a weak potential. In case the surface information of the COSMO-EU model is used instead of the surface VERA information, for the calculation of the KO-Index, almost all cells would be considered to be in the space of stable atmospheric conditions (not shown).

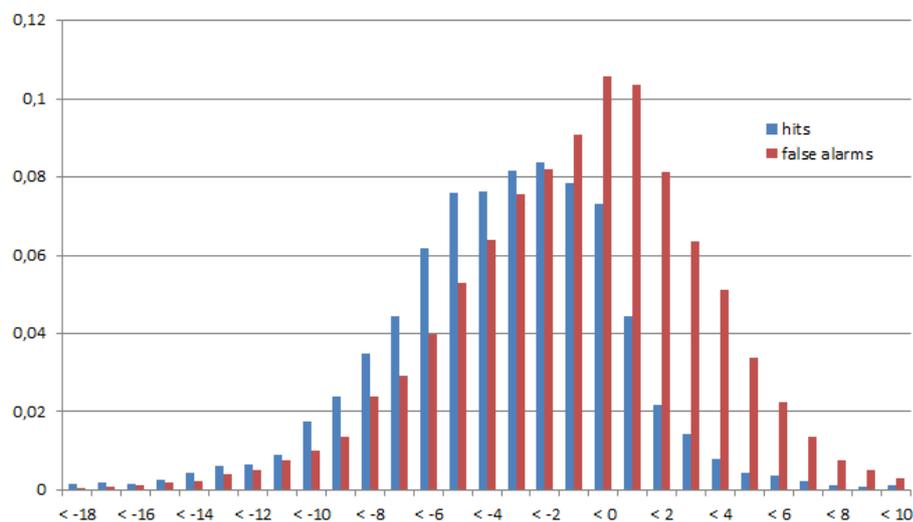


Figure 5.7.: The amount of hits(blue)/false alarms(red) for each KO Index value between lower than -18 and lower than 10, normalized by the total amount of hits/false alarms.

Differences between literature values and observed values for meteorological instability indices are often due to the region they were originally designed for. Most of these indices were originally developed for the continental U.S. There, more extreme values for these indices are common, due to the higher degrees of

instability often observed over the great plains in comparison to typical European conditions. This is not the case for the KO-Index which was developed at the DWD. The "KO" in KO-Index abbreviates the German translation of the word "convective". The literature values for the KO-Index are calculated for radiosonde measurements and are applicable over Europe. Thus, the key difference lies in a too stable thermal stratification within the model data. It is a known problem, that - at least for the summer months - the COSMO-EU model on average is too stable (Baldauf et al., 2007; Pflüger, 2004). This is the reason for KO values indicating stable conditions if only COSMO-EU data is used. The inconsistency between the COSMO-EU data and the VERA surface analysis leads to the results shown here with KO values indicating mainly too unstable airmass compared to the KO values in the literature. This leads to a shift of the classification to lower KO values here, based on the statistics.

There are two main reasons why the version with the VERA surface θ_e (as in Eq. 3.10) and not the COSMO-EU surface θ_e (as in Eq. 3.9) is used for further analysis. First of all, the VERA information is expected to be a lot more realistic than surface data from the model forecast. Furthermore, the statistical information of the KO-Index values for hits and false alarms is more useful for the version using VERA than for the version using only COSMO-EU. Using VERA improves the ability to distinguish between the distributions for hits and false alarms.

In Fig. 5.7 the two distributions are separated quite nicely for higher KO values which again leads to the possibility to filter false alarms with the KO data. Only 2.2 % of all hits (158 cells) show KO values above 3, in contrast to 14.0 % of all false alarms (3881 cells). KO values greater than 1 are found for 5.8 % of all hits (410 cells) and 28.5 % of all false alarms (7887 cells). An overview on these cumulative percentages for the KO Index is shown in Fig. 5.8.

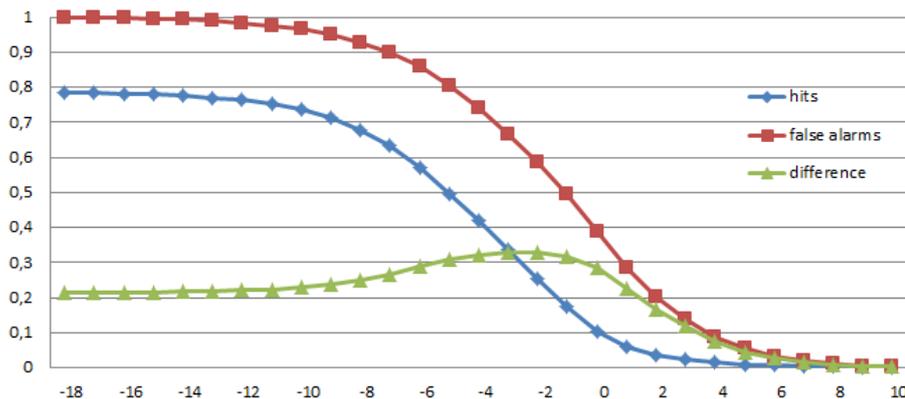


Figure 5.8.: Cumulative percentage of KO value frequencies above the given KO-Index for hits (blue), false alarms (red), and the difference of the two values (green). For further description see text.

In contrast to θ_e the cumulative percentages now are calculated for all KO value frequencies above the given KO value. The difference peaks at a KO of -2 with a value of 33.1 % (25.4 % of all hits and 58.5 % of all false alarms have higher values). The argumentation for constructing the fuzzy set is the same as described

earlier for the fuzzy set of the θ_e data. Thus the values around this peak in the difference will again be used to pinpoint a transition in the fuzzy set, here between neutral and unstable, while the low, positive values mentioned earlier, where only few hits lie above, will be used for the transition area between stable and neutral conditions. This leads to the fuzzy set shown in Fig. 5.9 with a transition from unstable to neutral between KO values of -4 to -1 and a transition from neutral to stable between 1 and 3.

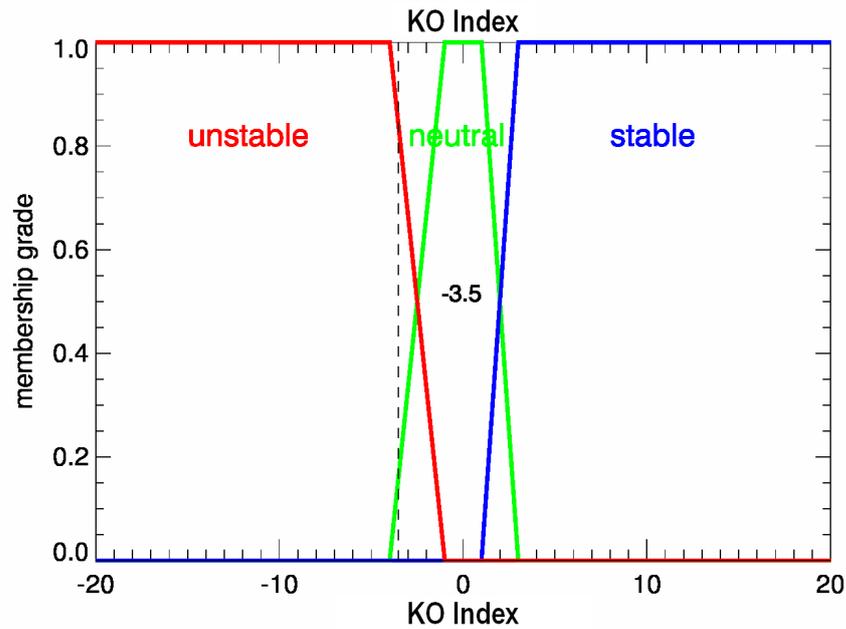


Figure 5.9.: KO-Index fuzzy set with an arbitrary example value. A KO value of -3.5 is some 18 % in the range of neutral KO values and some 82 % in the range of unstable KO values.

Due to the convincing results in the analysis of the KO-Index data no other type of instability measure has been tested within this study.

5.2.3. Indicating lift

For information on the available lift in the middle of the troposphere an analysis of the COSMO-EU vertical movement ω in 500 hPA is presented. The original ω data of the COSMO-EU is a quite noisy field with wavelike structures that might change the algebraic sign of ω quite often. This leads to the corrugated structures in ω fields that can be observed quite nicely in Fig. 5.10. Its quite easily seen on the shown example map, especially in the southwestern part over Spain and Southern France and in the eastern part over the Czech Republic, Austria, and south to the former Yugoslavian countries. These small scale wave features will be smoothed with a Gaussian kernel for the upcoming analysis to get an image of the large scale features, in which sufficient lift for CI might occur.

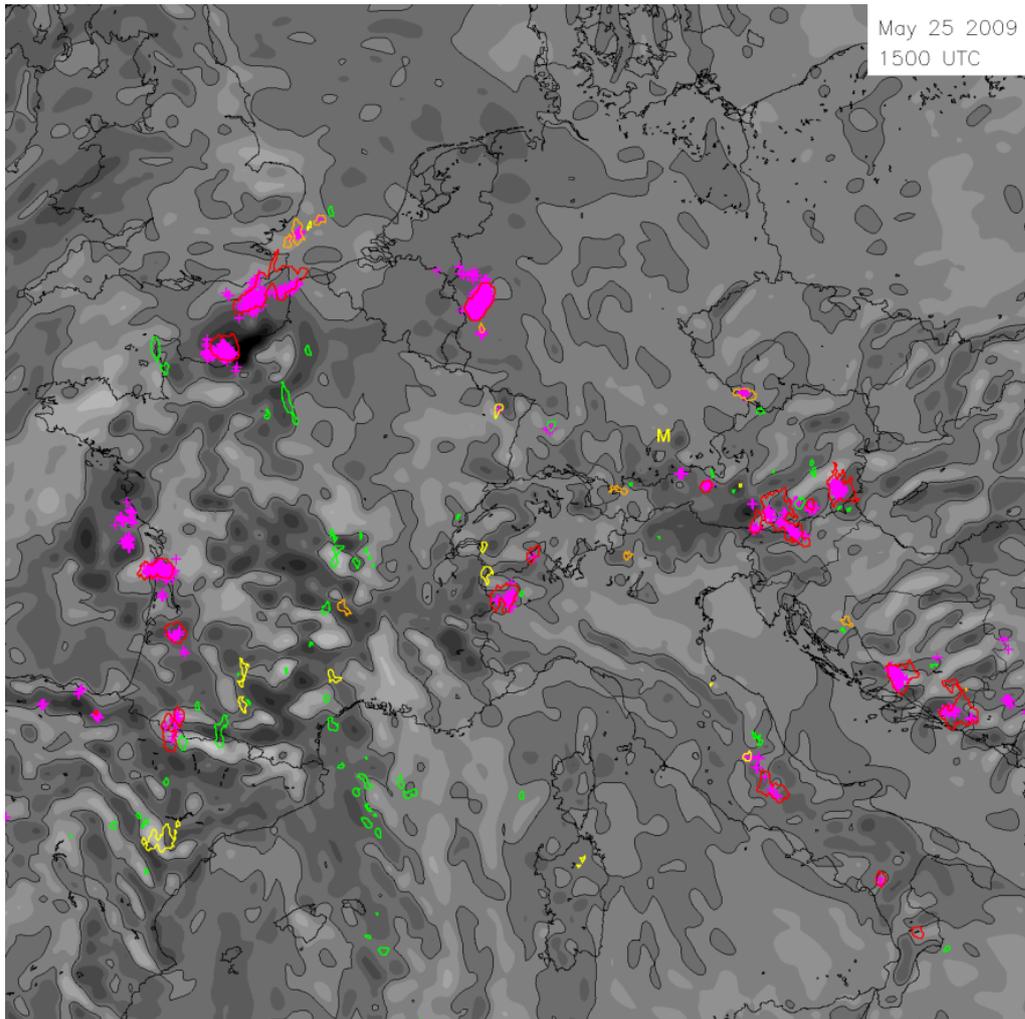


Figure 5.10.: Example plot of original ω data in 500 hPA. The thin black line is the 0 hPA/h contour. Darker gray resembles negative values (upward motion) and light gray is used for positive values (downward). The range of resolved values is from -300 hPA/h to 300 hPA/h.

Variations of the edge length for the kernel of 5 pixels (35 km), 9 pixels (63 km), and 17 pixels (119 km) were tested to find the minimum smoothing needed. The version with the 17 pixel edge length is the first which does not show the corrugated structures and is shown in Fig. 5.11.

Plots of the same example time step for the two tested versions that were not used for the analysis can be found in the Appendix. In line with the original and the final smoothed versions from Fig. 5.10 and Fig. 5.11, they are shown in Fig. F.4 and Fig. F.5. In Fig. F.4 the overall structure does not change very much for the small kernel in comparison to the original image only the amplitude of maxima and minima is smoothed obviously. The wave structures in the medium kernel part of Fig. F.5 are still slightly visible. The smoothing by this kernel is still too weak to eliminate all of these waves. To be sure to eliminate all of these unwanted structures and thus to be able to focus on large scale movement the largest of the presented kernels is used for the analysis.

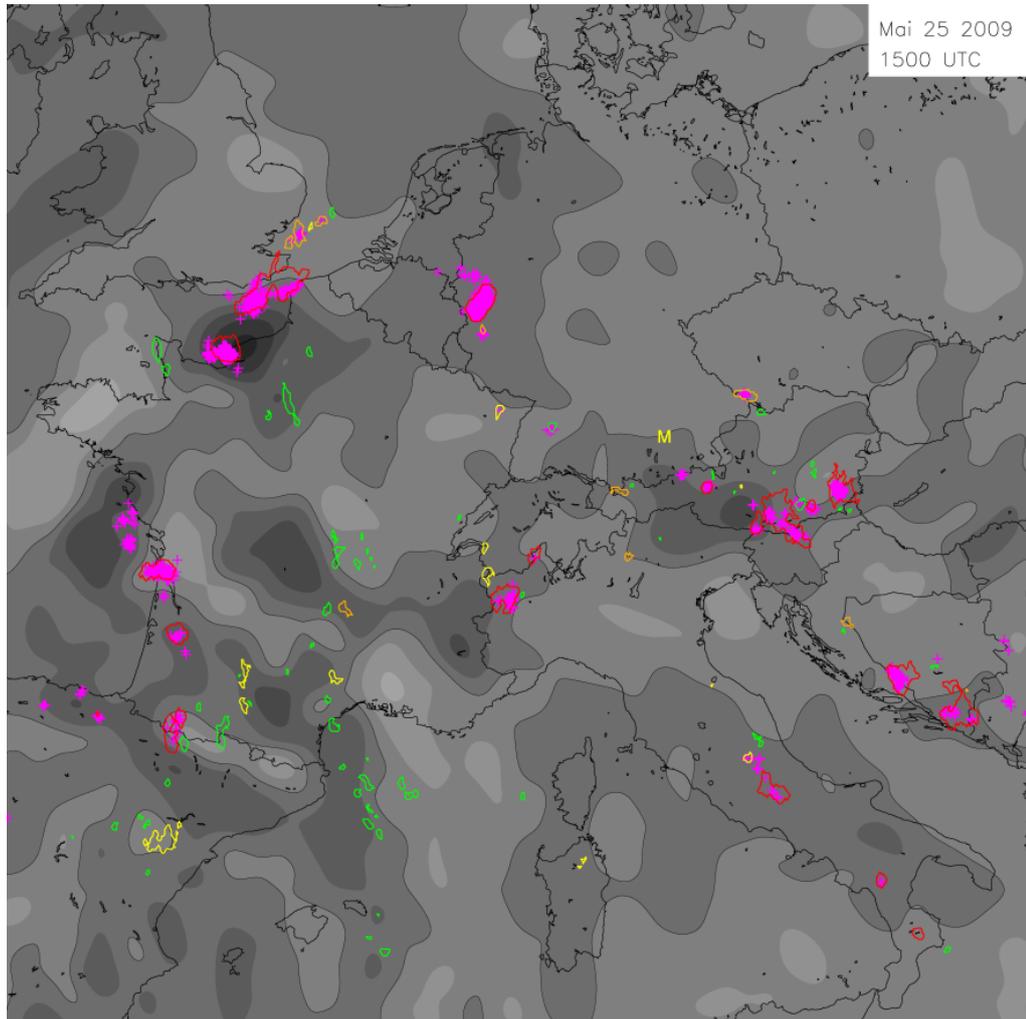


Figure 5.11.: Example from Fig. 5.10 after appliance of a Gaussian kernel for smoothing the data.

The range of data analyzed for the smoothed ω data begins with values lower than -100 hPa/h and ends at values higher than 80 hPa/h. However from Fig. F.6 in the Appendix, we can see that most of the cells are within a range much closer around 0 hPa/h. To be able to visualize the interesting area in more detail the range for the statistical distributions shown within the next plots is zoomed in and is shown from -15 hPa/h to 15 hPa/h. The normalized distributions in Fig. 5.12 show that their peaks are not separated as nicely as in the earlier cases.

For both earlier cases the normalized hits had a part at the edge of their distribution where they had low values during a still quite high distribution for the false alarms. Here the hits do not drop that nicely at the edge of their distribution. This leads to the fact, that for low percentages of possibly filtered hits, the spread to the possibly filtered false alarms is less than for the data shown earlier. If we would filter for example all cells where ω is greater than 10 hPa/h this would have an impact on 3.4 % of all hits (238 cells) and with 5.8 % on a slightly higher percentage of all false alarms (2542 cells). Using a threshold of ω greater than 5 hPa/h would eliminate 9.2 % of all hits (650 cells) and, at least,

already 22.4 % of all false alarms (6224 cells).

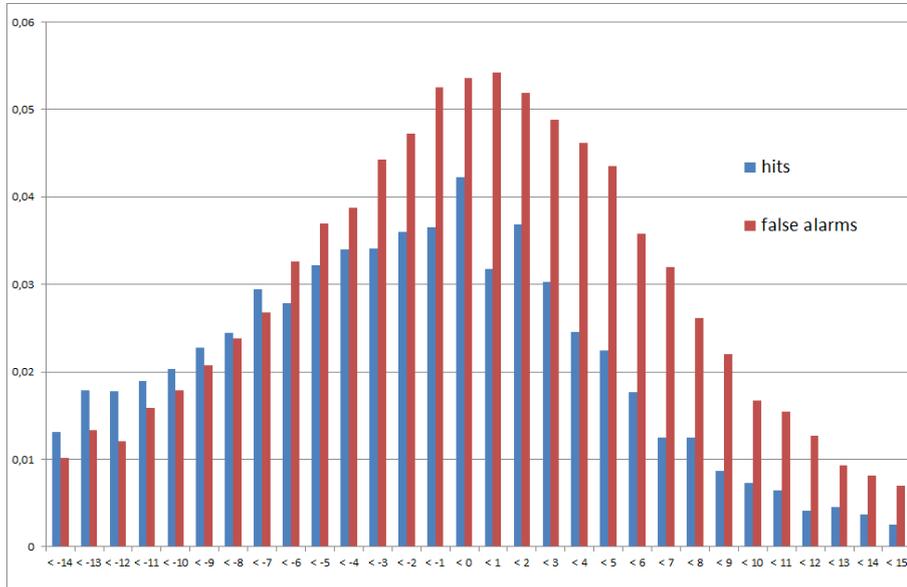


Figure 5.12.: The amount of hits(blue)/false alarms(red) for each ω value between lower than -14 hPa/h and lower than 15 hPa/h, normalized by the total amount of hits/false alarms. Zoom in smaller range of the data from Fig. F.6 (top).

At the other end of the distribution, where the normalized percentage of the hits per 1 hPa/h is higher than the normalized percentage of the false alarms, the spread is quite small as well. This leads to an effect visible in Fig. 5.13.

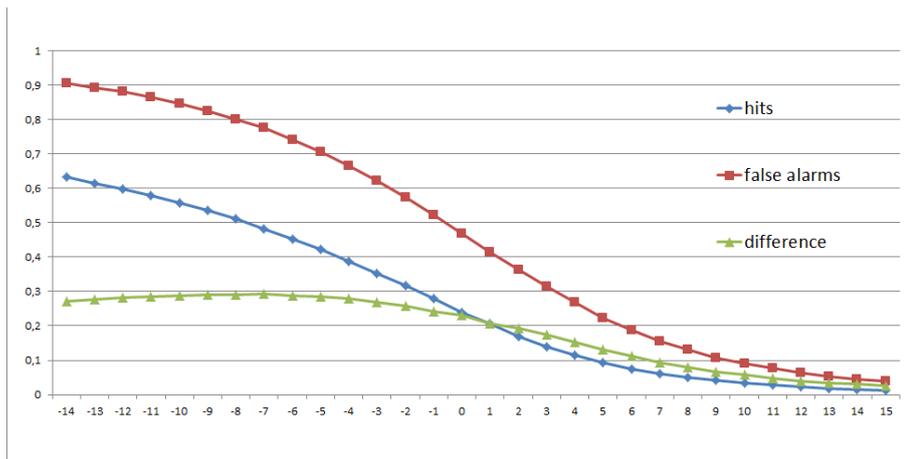


Figure 5.13.: Cumulative percentage of ω value frequencies above the given ω in hPa/h for hits (blue), false alarms (red), and the difference of the two values (green). For further description see text. Zoom in smaller range of the data from Fig. F.6 (bottom).

Within the shown range, the peak for the calculated difference between the cumulative percentages of hits and false alarms is not as distinct as in the earlier

shown data analysis cases. The difference peaks at 29.3 % for -7 hPa/h, where already 48.1 % of all hits and 77.4 % of all false alarms include higher values.

Cells with ω values shown by a reasonable number of false alarms might be filtered when almost no hits are lost, and cells with ω values between that range and the peak of the difference are difficult to judge clearly. Thus, relying on the same concept as earlier within this chapter to translate the statistical information into a fuzzy set of the COSMO-EU ω data in 500 hPa results in the set shown in Fig. 5.14. The shift from upward motion (CI forcing) to neutral conditions lies between -7 hPa/h and 0 hPa/h and in the range from 5 hPa/h to 10 hPa/h the transition from neutral to downward motion (non-CI-favorable) takes place.

After the analysis of the ω data the question might arise, why this ω data is used anyway due to the results which showed to be less convincing than the data taken for the other two ingredients. To answer this a short overview on some other data tested as proxy for lift is presented here. The other ideas which data might be used for indications of sufficient lift focussed on a horizontal flux analysis from VERA and more ω data by COSMO-EU using different levels or ω profiles incorporating the information of multiple levels.

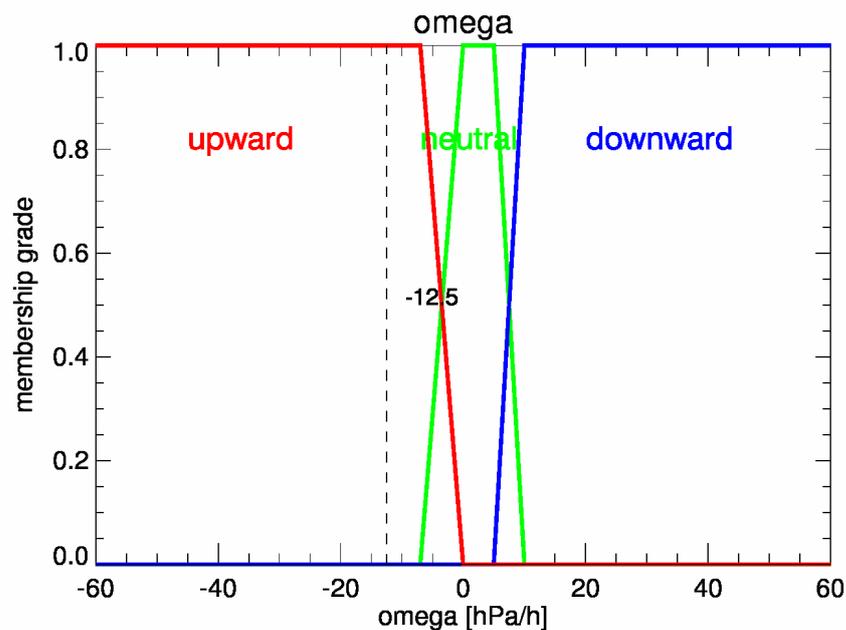


Figure 5.14.: Fuzzy set for ω in 500 hPa with an arbitrary example value. An ω value of -12.5 hPa/h is 100 % in the range of upward ω values.

The horizontal flux analysis at surface level is a postproduction from VERA output. This information on divergence and convergence, possibly triggering lift, showed up to be not distinctive enough to get better information than from the ω data. The statistics (not shown) indicate that the VERA convergence is even a slightly less distinctive feature, for hits and false alarms, than the MFC shown in the Appendix in Figs. F.1 to F.3.

For evaluating the impact of model updraft strength five ω levels (850, 700, 600, 500, and 400 hPa) were examined in more detail. Furthermore profiles in-

corporating some or all of these levels were tested, too. The information using more than one level would be classified as CI-favorable if the value is CI-favorable in each single level and vice versa for non-CI-favorable conditions. A situation where not all levels had the same algebraic sign would be neutral. This showed to be not distinctive enough again. As extension a second approach to define fuzzy sets, like the one shown in Fig. 5.14 for the 500 hPa level, for each of these levels was performed. Aiming at the construction of a second fuzzy logic combination to get one ω -forcing value for each cell, this was examined briefly and rejected after not very encouraging results (not shown). The combination approaches did not result in better abilities for filtering than the easier, straight forward way of simply using only one level.

In addition to put the use of the 500 hPa data into question an analysis for other single levels was performed and is summed up briefly. The normalized percentages per bin are broader distributed for the other levels which means the values in the central part of the distributions sink for the other levels and the bins to the sides get higher values which is true for both, hits and false alarms. The spread between hits and false alarms per bin is almost unchanged. An example for these points is shown for the three levels of 500, 600, and 700 hPa/h in Fig. F.7 in the corresponding Appendix. Adding more levels would not change this result but obscure the plot even more, that is why only three of the levels are plotted here. The difference of the cumulative percentages, as key value used for the translation from the analysis of possible filtering to the fuzzy set, is changed only very slightly for the different levels. The peak value is the highest for the 500 hPa level and the chosen level does not change the ω value at which the peak is to find (as can be seen from Fig. F.8). These results lead to the conclusion the the ω data in 500 hPa is still the most useful, of the tested data, for the evaluation of the available mid-tropospheric lift.

5.3. Combination of indicators

The conceptual approach for the combination is to build a routine which runs as kind of Cb-TRAM postprocessing and then checks each CI cell detected at a Cb-TRAM daytime detection time step in two steps. The first step is to check the CI cell for occurrence of lightning flashes within the cell. If there is a lightning detection inside the cell, it is kept as a cell which cannot be a false alarm. In the second step all cells which do not incorporate lightning are tested for the "CI forcing" that can be found at the location of the cell. The location is defined by the polygon of the parallax corrected CI detection. This CI forcing is based on the availability of the three thunderstorm ingredients discussed earlier - moisture, instability, and lift. They are represented by the surface θ_e analysis from VERA for indicating low-level moisture, the KO-Index calculated from VERA data and COSMO-EU model data for an instability analysis, and mid-tropospheric vertical motion in the COSMO-EU ω data (in 500 hPa) for the indication of lift. The data chosen here as useful for the combination and the other data tested and described in Sections 5.1 and 5.2, which are not used for the combination, are

summed up in Tab. 5.2.

data type	benefit	finally used
lightning data (LINET)	high	yes
θ_e (VERA)	high	yes
MFC (VERA)	medium	no
KO Index (COSMO-EU & VERA)	high	yes
KO Index (only COSMO-EU)	medium	no
Convergence (VERA)	low	no
negative ω (updraft) in 500 hPa (COSMO-EU)	high - medium	yes
negative ω (updraft) in other levels (COSMO-EU)	medium - low	no
negative ω (updraft) with combination of levels (COSMO-EU)	medium - low	no

Table 5.2.: Classification of the benefit gained by each of the additional data tested for the task of CI detection and nowcasting. Rows with data used for the combination are additionally marked blue.

As an additional motivation for the fuzzy logic combination concrete results for a type of combination with fixed thresholds shall be discussed. Therefore I would like to illustrate with an example, the additional information about the CI cells that can be gained, if the data analysis showed earlier is used to find fixed thresholds for filtering false alarms.

The example is visualized with the help of Fig. 5.15. It shows an illustration of a fuzzy set (upper part), like the ones shown earlier, and its reduction to a number line (lower part). The color represents the type of forcing for the sets on the left and right side of the range with blue as non-CI-favorable and red as CI-favorable part.

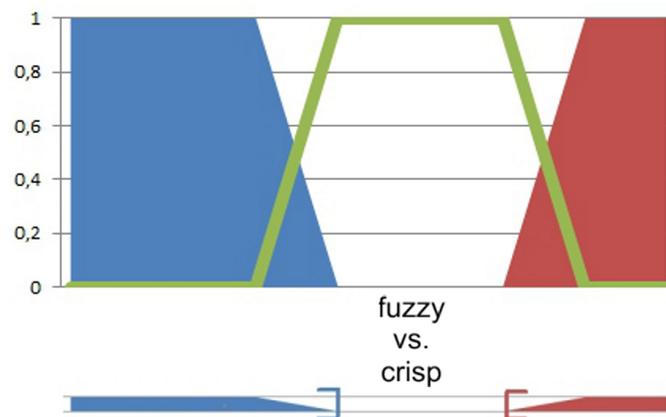


Figure 5.15.: The upper part shows an arbitrary input fuzzy sets with blue as non-CI-favorable and red as CI-favorable part of the fuzzy set. In the lower part the fuzzy set is reduced to a number line. The blue bracket marks the fixed threshold between non-CI-favorable and neutral circumstances, the red bracket marks the crisp threshold between neutral and CI-favorable values

If the values representing the beginning transition from the neutral to the blue or red area would be taken as fixed thresholds (brackets on the number line) for each of the sets boundaries we can calculate an amount of false alarms that could be filtered with this approach. The fixed thresholds at the transition to the blue area (the blue bracket) bounds, for each data source, all cells which are at least partially within the non-CI-favorable fuzzy set. Combining the data with fixed thresholds in the way, that a CI cell is judged as non-CI-favorable if all ingredients are non-CI-favorable, results in the following numbers for the available sample.

As shown in Tab. 5.3 the amount of false alarms where all ingredients are judged as at least partially non-CI-favorable is 1195 cells (4.4 % of all false alarms). Within this criteria hits are only represented by 12 cells (< 0.2 % of all hits). Thus if this combination of fixed thresholds would be used these false alarms could be filtered without losing a meaningful amount of hits. The other thresholds at the transition to the red area (red bracket) bounds all cells which are at least partially within the CI-favorable fuzzy set. Following the same combination this CI-favorable range above the thresholds (see Tab. 5.3) incorporates 6619 false alarms (24.2 % of all false alarms) and 4041 hits (58.5 % of all hits). Filtering all cells which are not within this set would thus lead to filtering 20783 false alarms (75,8 %), but also to losing the tremendous amount of 2863 hits (41.5 %).

fixed threshold with all input at least partially...	hits	false alarms
...non forcing (blue in Fig. 5.15)	12 (< 0.2 %)	1195 (4.4 %)
...forcing (red in Fig. 5.15)	4041 (58.5 %)	6619 (24.2 %)

Table 5.3.: The table shows results for the example of a combination with fixed thresholds. For each amount of hits/false alarms the percentage of all hits/false alarms is added in parentheses. For further descriptions see text.

With the aim of losing as few hits as possible only a small amount of false alarms (4.4 %) could be filtered directly. In addition, this would happen without gaining a lot information on the other cells. The second threshold would filter too much of the hits which would result in a low oPOD dev of 0.185 in contrast to an at least fairly reduced oFAR dev of 0.621. For comparison see the blue column for acc_{track} in Tab. 4.2 (oPOD dev and oFAR dev values for the larger verification dataset in Ch. 4).

The idea now in using a fuzzy logic combination is first of all to allow a degree of uncertainty in the choice of the threshold values, by using a smooth transition and, in addition, it helps to gain further statistical information on the cells after incorporating the additional data. The cells are not just classified as below or above a combined threshold but they gain a new attribute with additional information by using the fuzzy logic. For each cell (without lightning occurrence) we get an index value for the CI forcing, indicated by the fuzzy logic combination of the fuzzy sets for each additional data, as described earlier.

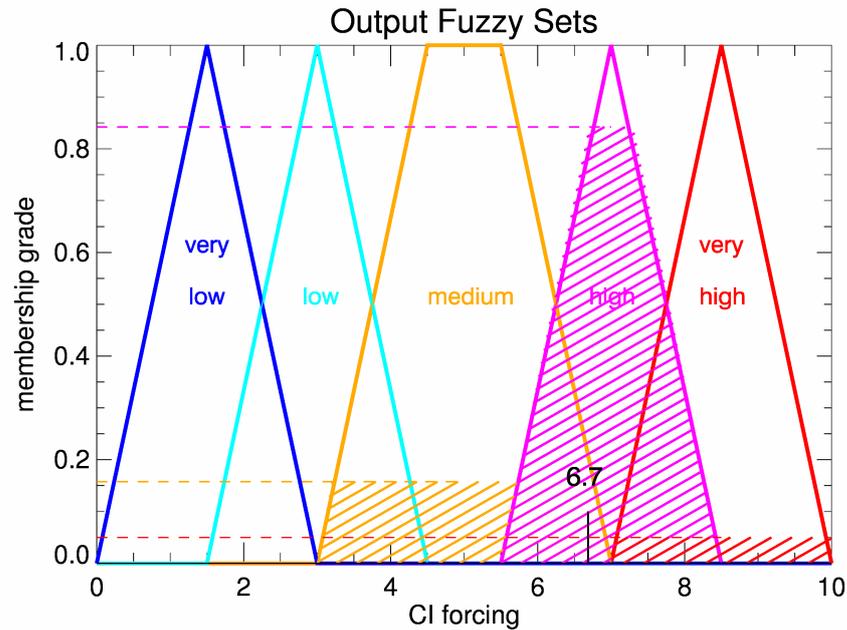


Figure 5.16.: Visualization of the output fuzzy sets for the arbitrary examples of input fuzzy sets shown earlier within this Chapter. For further description see text.

In Fig. 5.16 (same as Fig. 3.15) we see the output fuzzy sets with the resulting CI forcing for the arbitrary example values in the input fuzzy sets shown in Figs. 5.6, 5.9, and 5.14. The output fuzzy sets are constructed as described in Section 3.4, following the description in Tab. 3.5, which explains the assignment of the different rules in the rulebase (see Appendix C) to the different output fuzzy sets. Thus, this resulting CI forcing value of 6.7 is an arbitrary index interpretation of the input fuzzy sets in an arbitrary range from 1.5 to 8.5. As covered in the passage below, the CI forcing value can not be interpreted directly as a kind of CI probability ranging from 0 % to 100 %. Neither can it be translated, corresponding to the minimum and maximum value reached in the CI forcing value range, as 15 % to 85 %, as might be an intuitive procedure, too. However, by the construction of a meaningful rulebase to get this CI forcing index value from the output sets, it can be expected, that the probability of a CI cell verifying as hit is higher for higher CI forcing values.

This leads to the same type of statistical analysis as already performed for the input data, to get further information on a kind of CI probability for the different CI forcing index values. Clustering the CI forcing index in 0.5 index value bins and afterwards counting the amount of hits and false alarms per CI forcing index bin leads to the possibility to calculate the percentage of hits per bin. This results in a CI probability for each bin. The normalized amount of CI hit and false alarm cells per bin is shown in Fig. 5.17. The size of the bins is chosen freely again to get a reasonable detailed distribution of the data.

The design of the rulebase (see Tab. 3.5) is somehow reflected in the distribution in Fig. 5.17. In contrast, to a combination with fixed thresholds, the fuzzy logic

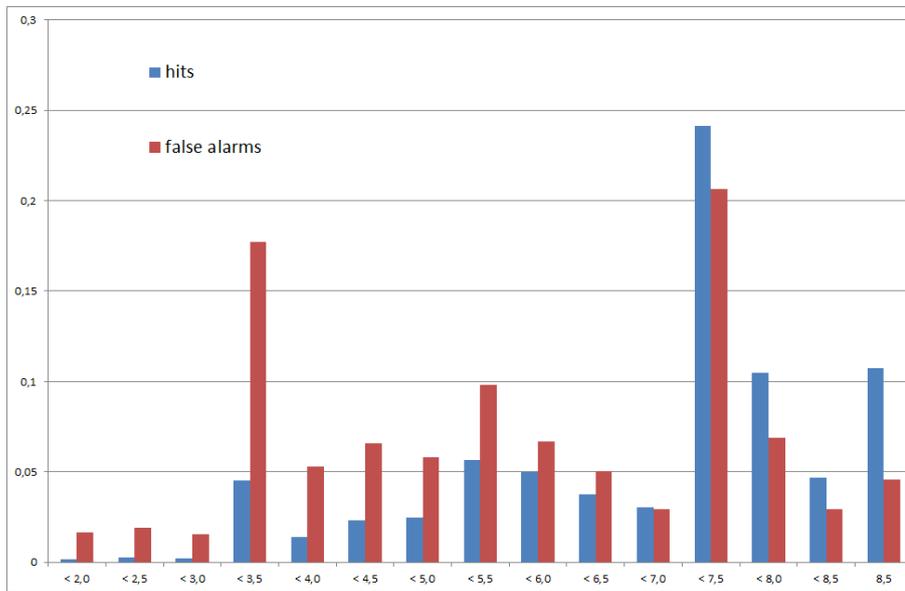


Figure 5.17.: The amount of hits(blue)/false alarms(red) for each 0.5 CI forcing index value bin between lower than 2.0 and lower than 8.5 and the ones equal to 8.5, normalized by the total amount of hits/false alarms.

combination results in a whole spectrum of CI forcing values instead of only few possible discrete values. The amount of hits per output set is at least to some degree related to the amount of rules leading to this set. The centers of the five output fuzzy sets for the CI forcing shown in Fig. 5.16 ("very low", "low", "medium", "high", and "very high") are at 1.5, 3.0, 5.0, 7.0, and 8.5. There are no values below 1.5 and above 8.5 due to the lack of a neighboring set leading to a shift to these lower or higher values. The set with very low CI forcing has only one decision rule leading to this output and contains only few hits and false alarms. For the medium forcing six rules point to this result and the distribution shows a slight peak for the lower than 5.5 bin, where the CI forcing value of 5.0 belongs to. The sets for low and high forcing are met most often by the rules (nine times each) and are quite populated (≤ 3.5 and ≤ 7.5). The set with very high forcing is very populated despite only one rule resulting in this set.

For all normalized percentages per bin with values lower than 6.5 the false alarm value is higher than the value for the hits, for the one with lower than 7.0 it is about equal and for the following ones the normalized amount of hits is higher than the one of false alarms. The low amount of hits in the very low forcing category, is an important feature for possible filtering approaches. In case all cells with CI forcing values lower than 3.0 would be filtered only 0.6 % (42 cells) of all hits would be lost, but 5.1 % (1401 cells) of the false alarms might be dismissed. This can be seen from the cumulative distributions shown in Fig. 5.18.

As already mentioned the likelihood of a detected CI cell to show further development can be related to the CI forcing value. CI forcing values greater or equal 7.0 are found for 9604 false alarms (35 % of all false alarms) and for 3452 hits (50 % of all hits). Furthermore for the hits the cells including lightning have

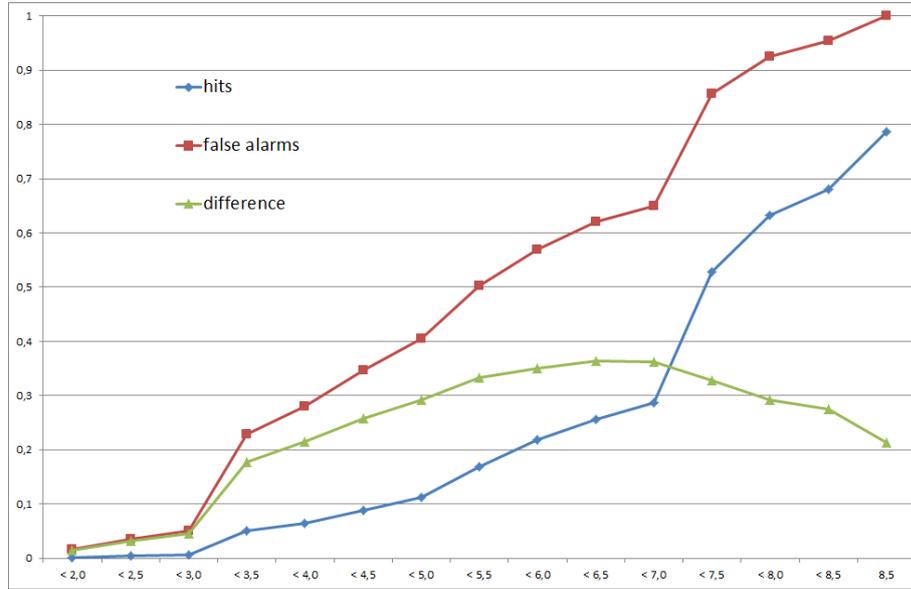


Figure 5.18.: Cumulative percentage of CI forcing index value frequencies below (or at 8.5) the given CI forcing index for hits (blue), false alarms (red), and the difference of the two values (green). For further description see text.

to be incorporated in the analysis of CI-favorable conditions too. Thus 4922 cells show forcing values ≥ 7.0 or lightning, which represents 71.3 % of all hits. These numbers would result in an oFAR dev of 0.661 for a CI forcing value greater or equal 7.0. If the amount of hits ≥ 7.0 (+ cells with lightning) is divided by the total amount of cells with values ≥ 7.0 the result can be interpreted as a CI probability for cells with a CI forcing value ≥ 7.0 . The probability values are based on the performed statistical analysis. The resulting probability for CI forcing ≥ 7.0 is 0.339 or 33.9 %. The calculation of the CI probability is defined as hits divided by overall CI cells and corresponds to $1 - \text{oFAR dev}$. This is the same equation as the Success Ratio SR in Eq. 3.3.

$$CIprobability = \frac{hits}{all\ CI\ cells} = \frac{hits}{hits + false\ alarms} \quad (= SR) \quad (5.1)$$

These calculations of oFAR dev and the CI probability can be performed for each CI forcing value and are shown also as cumulated values greater or equal the CI forcing values in Fig. 5.19. The bins are interpreted differently here as in Fig. 5.17 or Fig. 5.18. The separation here is not analyzed for values lower than the given value but for values greater or equal the CI forcing values. This shows that CI detections with CI forcing value of greater or equal 8.0 or lightning detection within the cell in comparison to the available false alarms within this category result in an oFAR dev of 44.8 % or a CI probability of 55.2 %. For the lower CI forcing values Fig. 5.19 shows always the values of oFAR dev and CI probability for the cells in each bin and cumulative for all cells within this bin or the ones with higher forcing values.

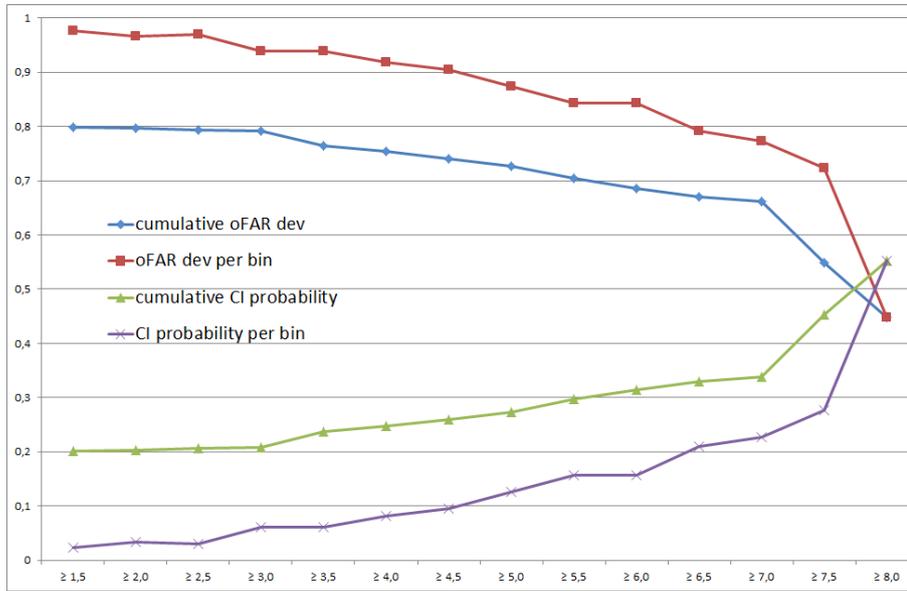


Figure 5.19.: Cumulative calculated oFAR dev (blue) and CI probability (green) above the given CI forcing index. Furthermore oFAR dev (red) and CI probability (purple) calculated for each individual bin.

A discussion of the information gained by the combination and how it could finally be used will follow in Chapter 6. However the cumulated oFAR dev or CI probability shows nicely that the level of confidence in the CI detection rises with higher CI forcing values. In case the CI forcing value is not taken into account one out of five cells develops to a stage two or three detection. For cells with a CI forcing value ≥ 4.0 one out of four cells develops, for cells with a CI forcing value ≥ 6.5 already one out of three cells develops, and finally for cells with a CI forcing of around ≥ 7.8 even every second detection develops to a higher detection stage.

The values, calculated per bin, illustrate how small the amount of hits compared to the amount of false alarms is for low CI forcing values. In the three bins with the lowest CI forcing values the CI probability for a detection is between 2.4 % (11 hits and 455 false alarms) and 3.3 % (18 hits and 526 false alarms). In summary these three bins incorporate the above mentioned 42 hits and 1401 false alarms which signifies that approximately 1 out of 33 cells develops to one of the next development stages (2.9 % CI probability).

The choice which value to use as CI forcing threshold for filtering of false alarms depends on the needs and expectations of potential users. If situational awareness in areas with first developments shall be raised, only these aforementioned low CI forcing values might be filtered. In cases where a high probability of further development is desired, minimizing false alarms despite losing hits, the threshold for filtering could be raised to higher CI forcing values. More discussion and ideas on usability and further development of the method will follow in the next two chapters.

6. Discussion

Within the previous Chapters results for the verification of the original Cb-TRAM CI detection performance (Ch. 4) and the improvement of the CI detection and nowcasting by the use of additional data sources (Ch. 5) are presented. Within this chapter some questions related to these results shall be discussed.

The discussion is divided in two parts, one related to the results of the Cb-TRAM CI detection verification and the second one with focus on the results of the combination with additional data. First of all the choice of the verification method is advocated and the verification results are put into perspective compared to other studies, both relating to the results in Ch. 4. Besides the use of Cb-TRAM output of other detection stages as verification data set, the question, how Cb-TRAM itself might be further developed for the CI nowcasting, is discussed as well. In a second part of the discussion the focus is more on the results presented in Ch. 5. Here the key points are the usefulness of the additional data, the choice of the data sources, the fuzzy logic based data combination, and the applicability of the resulting Cb-TRAM CI detection and nowcast. In addition a view on the question if the methodology for the postprocessing might be applicable for other CI nowcasting methods, or even more general for other nowcasting methods, is given as well.

Verification of Cb-TRAM CI detection

To begin with, an oPOD dev (acc_{track} in Tab. 4.2) of approximately 0.25 is not as bad as one might think at first. This result means that the detections and their nowcast for up to one hour cover 25 % of the storms that develop within the next 60 minutes. The nowcast is based on a detection of a CI suspect and there are no nowcasts of future developments in areas where nothing is detected right now, thus CI cells which do not exist already in a "detectable" shape but start to develop in 30 or 40 minutes are impossible to hit by this type of nowcast which does not predict future CI events. This leads to a high count of misses for cells whose development from CI to a new cell has a short lead time. If the lead time of a CI detection until cell development to stage 2 or 3 is only 15 minutes it will be counted as miss in the 20 to 60 minutes earlier performed analysis.

The mean lead time was not evaluated within this analysis. To evaluate the leadtime the analysis would need to be focussed on the lifecycle of the individual cells and not on the analysis of the detected contours and their nowcasts per time step. A "perfect" storm, which develops independent from any other storm systems, is expected to grow in approximately 10 to 15 minutes from first cloud development until it starts to precipitate. Which means 10 to 15 minutes from the first CI detection (stage 1) until cell development to stage 3. In cases the cells

grow very fast, the Cb-TRAM criteria for CI and for rapid development (stage 2) might both be fulfilled within the same pixels. In these cases the cell is defined to be in the further developed stage in the sense of the lifecycle. Furthermore this means if CI and the rapid development of stage 2 are detected close to each other for the first time in the same time step, this cell will start with a stage 2 detection instead of a CI detection. This problem occurs as well for storms developing in the direct environment of an existing storm. If the difference in location is very small (lower or equal two MSG SEVIRI pixels as described in Ch. 3.1) the objects of lower detection stages will be incorporated into the mature object within Cb-TRAM.

Another problematic case is the existence of cirrus clouds. Cb-TRAM cannot detect CI below preexisting cirrus clouds despite they are very thin. This is typical to satellite based CI detections as can be seen from comparable other CI detection tools like SATCASTv2 (Walker et al., 2012).

The type of evaluation here should give an information on the usefulness of the nowcasts produced with the Cb-TRAM CI stage detections, as they are right now. That is the reason why the cell nowcasts are evaluated in this manner. Another approach would be to check for each orange stage 2 cell, and each red stage 3 cell, if a yellow CI detection can be found at the beginning of their cell histories. This is more appropriate if the lifecycle of the cells shall be in the focus of the study. In cases where one cell develops from more than one CI detection this method would lead to a wrong count for false alarms. Only the one CI cell which is tracked within the history would be counted correctly as hit. For further description of the tracking see Ch. 3.1 or Zinner et al. (2008).

Despite the high amount of false alarms, the bias (see Eq. 3.5), with a value of approximately 1.3, is still close to 1 and thus the Cb-TRAM CI stage does not too heavily overforecast the amount of new developments.

For the high numbers of false alarms and misses one way of interpretation is obvious. This would be that the CI detection does not work well and the high number of false alarm cells and the high number of miss cells simply are not related at all. However, as already mentioned, this is the verification result for the usability of up to 1 hour thunderstorm nowcasting for new developing cells based on CI detections and not a CI prediction for future CI events. In addition a kind of eyeball verification on many test cases brought up the two following example scenarios which can be observed quite regularly.

On the occasion that CI cells are detected for the first time in an area where no other clouds exist, no meaningful transition vector field for the area can be calculated. In these cases the nowcasts cannot be extrapolated and the future cell positions are expected to be at the same place where the cell is right now. If then a cell development to a higher detection stage takes place already in the next time step with the cell being slightly shifted from the earlier CI detection and not overlapping, it might happen that these two cells are misleadingly not tracked as one cell. In these cases the first cell without the nowcast is counted as false alarm and the further developed cell nearby is counted as miss. This resembles a classical double penalty problem for close forecasts described earlier. In the verification types tested, this one would only be resolved as a hit if the

version with the search radius would be used, and then only if the chosen radius is large enough. This can be classified as a tracking error which can be observed at developments within large cloud free areas.

The second scenario can often be observed with cells in mountainous area in weak wind situations. Cells often start to bubble in valleys and are detected as CI suspect but do not develop to the next stage within the next time steps. Sometimes these cells eventually develop to one of the next stages with a lead time of more than 1 hour. Due to the construction of the verification, cells with lead times above 60 minutes are all counted as false alarms, for their first detection timesteps until they are within one hour to their further development. Despite the fact, that these clouds, being in the initiation stage for such a long time, seem to be quite unrealistic results, the detections still can be useful to raise situational awareness of forecasters to observe this area more closely. Even though showing not a CI event in the strict understanding of convection initiation, which means a first development of a then ongoing convective growth, these repeating CI events can prove beneficial. An area with initiation lasting for more than one hour definitely does not describe the initiation of the later on developing cell solely, but it marks first, weaker developments which collapsed again and can be seen as precursors of the later on successful CI event.

As an additional point to mention, the performance of the Cb-TRAM CI detection surely can get better by further studies on the detection itself. This was to some degree done within a master thesis by Merk (2012) who successfully incorporated a cloud mask after Berendes et al. (2008) and some of the MSG adapted SATCAST interest fields after Siewert et al. (2010) into the CI detection of Cb-TRAM. With a verification approach focussing on the lifecycle of the cells and only few test cases presented, this work already shows encouraging results. The key problem here as on any detailed work on Cb-TRAM before is that the focus was on the normal scan data. This thesis presented here is the first one which solely focused on rapid scan data. Comparison of the detection for both update rates leads to the impression that the rapid scan CI detection is more noisy than the normal scan detection. This might possibly be solved simply by small changes in the detection criteria or the use of different time steps to calculate the trends for the detection. However, further work needs to be done here clearly.

The use of stage two or three Cb-TRAM detections to verify stage one nowcasts might be surprising in the first moment. The next few sentences shall clarify the choice of a verification data produced by the same tool. Radar data is usually the first choice as data source to verify any kind of thunderstorm nowcasting. As described already in Chapter 3 a radar reflectivity threshold around 35 dBz (Roberts et al., 2012) to 37 dBz (Kober & Tafferner, 2009) is commonly used to define thunderstorm signals in radar data. However radar data is not the appropriate data source within this study. The quality of all the available European radar composite data is simply too unreliable for verification purposes over some parts of the area examined here. The main reason for the partially low quality are technical problems with the construction of a European radar composite due to the fact that there is no common scan strategy used by the different national radar

networks across Europe (Kober & Tafferner, 2009). In Zinner & Betz (2009) and Zinner et al. (2012) good verification results for a comparison between LINET lightning data and Cb-TRAM mature cells are presented for normal scan data. They suggest that Cb-TRAM mature detections resemble the "truth" shown by the lightning data quite good. Furthermore the differences between rapid scan and normal scan detection objects are a lot less pronounced for stage two and three detections. Due to that argument the results of Zinner & Betz (2009) for the Cb-TRAM normal scan mature cells, seem to be realistic for rapid scan mature storm detections as well. The Cb-TRAM detection of stage two or three is independent of the detection for stage one, which is important, too. This led to the approach using directly the Cb-TRAM polygons for further developed stages as proxy for the "truth", as well facilitating the implementation of the cell ID comparison, the object matching, and thus the object-based comparison explained earlier.

In comparison to other studies the completely different type of evaluation, as described below, and the strict usage of stage 1 only has to be taken into account. Comparability of studies would eventually be higher if stage 1 (CI) and 2 (rapid development) of Cb-TRAM would be evaluated together. Often fast developing cells are detected as orange right from the beginning even with 5 minute rapid scan data. Other verification approaches for CI nowcasts can be found, for example in Mecikalski & Bedka (2006), Mecikalski et al. (2008), Siewert et al. (2010), or Walker et al. (2012), here with the aim to verify SATCAST or SATCASTv2 CI nowcasts. A big difference between the study presented here and these others is the very large size of the test dataset used here, and the evaluation of weak, or almost non-convective days within this study. The choice of the verification method and the interpretation of the chosen methods is difficult to compare too. The other studies show higher POD values and slightly lower FAR values, but for a non-comparable, pixel-based fuzzy verification method for matching the SATCAST CI detections with radar data. Furthermore an important reason for the lower scores within this study presented here, besides the mentioned differences in evaluation, seems to be the choice of the test dataset. In contrast to other studies, using only days with strong convective activity, the test dataset that is used within this study here also includes days where only few or even no mature convection occurs. This raises the possibility of producing false alarms in the detections. The focus within the other studies was more on the achieved lead time per detection. Thus for each cell detected the lead time, until it can be matched with radar signals anyhow, was measured. Within this thesis the focus is more on the overall useability of the nowcast for up to 1 hour.

Given the limited predictability of convection, one might wonder how good a CI nowcast could ever be. CI occurs on a small scale and is only a short period at the very beginning of the convective lifecycle. After the detection of CI the further development is heavily dependent on the current convective environment. If one cell successfully develops to a mature storm it might influence the convective environment for other cells which are nearby. It might hamper the development of other CI suspects as well as it might trigger new thunderstorms, as described in Ch. 2. A meaningful estimate of an achievable minimal oFAR dev value is

quite difficult. The oFAR dev value could easily be reduced below the current values, but not without losing a substantial amount of hits. oFAR dev might even be artificially reduced to 0.0 if only CI suspects including lightning are kept as CI detections. For oPOD dev it seems definitely unrealistic to expect a value of 0.5 or even more for the nowcasts with up to 60 minutes leadtime. The short timescale of CI and its development speaks in favor of probabilistic information for such extended leadtimes. As already explained, the verification result gives the information that approximately one quarter of the storms of the next hour are correctly detected within the CI detections and their nowcasts. Moreover this work suggests how these nowcasts can be used and interpreted with the help of additional data to gain a probabilistic information on the future development for each CI detection object which leads to the next part of the discussion.

Combination with additional data

The first point to be discussed here is the choice of the additional data sources. The aim was to get additional information about the three basic ingredients for the development of thunderstorms – low-level moisture, instability, and lift. The data should have sufficient resolution in space and time to generate a relevant analysis, reflecting the current atmospheric circumstances as detailed as possible. Due to the desire to get the most realistic information possible, observational data is preferred if available with sufficient resolution. Therefore the VERA tool with its highly-sophisticated interpolation (Steinacker et al., 2000b) was used to produce continuous surface analysis data. For upper air information no sufficient observational data is available, thus a numerical weather prediction model is needed to get additional information for the analysis. The used COSMO-EU model focusses on the meso- β scale with a grid spacing of 7 km. The focus of the model is the accurate prediction of near-surface weather conditions concentrating on clouds, frontal precipitation, fog, and orographically and thermally forced local wind systems (Schättler et al., 2011), whereas convection is fully parameterized. The wind systems, near-surface conditions, and a good forecast of the frontal position are important for the desired information on the above mentioned moisture, instability and lift. The second DWD version of the COSMO model, the COSMO-DE model, is a meso- γ scale version with 2.8 km grid spacing. The aim here is a direct simulation of severe weather events triggered by deep moist convection. At the first moment this raises the expectation that the COSMO-DE would be the better alternative for CI nowcasting. Thus, why the COSMO-EU is used within this study? There are two key factors militating for the use of COSMO-EU. First of all, a quite mundane reason for the use of COSMO-EU instead of -DE is the area covered by the models. The COSMO-DE domain is smaller than the area analyzed within this study. As second reason, the COSMO-DE sometimes shows errors in the timing and/or location of the convective cells (Dahl et al., 2011; Köhler, 2011) which might to a big part be related to deficiencies in the quality of the boundary layer parameterization (Baldauf et al., 2011b). COSMO-DE is able to forecast deep convection explicitly, but individual convective cells are hardly predictable. Potentially, the COSMO-DE

convective-scale ensemble prediction system (EPS) (Baldauf et al., 2011b) is capable to overcome this issue to some degree. However, the COSMO-DE EPS is operational since summer 2012 and thus its quality in producing realistic convection might only be tested in future studies as will be briefly discussed later in the outlook of this thesis. Anyway, if the aim is to represent the (large scale) convective environment around CI suspects, high resolution is not needed. Neither is the trend of the environmental information with one hour temporal resolution important for CI nowcasting. The convective timescale is quite short in comparison to changes of the large scale environment and CI itself is only a short process at the beginning of storm lifecycles.

The data tested and finally used from these data sources was discussed already in Chapter 5 and is summed up in Table 5.2. With θ_e , KO-Index, and ω (updraft) in 500 hPa, one additional information per thunderstorm ingredient was chosen which shall help to distinct between developing and non-developing CI cells. In addition lightning data was used aside of the data combination to avoid to filter CI cells erroneously, as described in Section 3.3.

In the following some aspects on the fuzzy logic combination shall be discussed. The first question to address, after the additional information described above, is why only three additional data types (one per ingredient) and not more potentially useful data is combined within the fuzzy logic. More data per ingredient might lead to even more additional information. It seems reasonable to keep the weighting between the ingredients equal, despite it is proofed that the data for one ingredient is less reliable. For one information per ingredient we have three input fuzzy sets with three classes (non-forcing, neutral, forcing). This leads to *input classes*^(no.ofindicators) different rules in the fuzzy logic rulebase ($3^3 = 27$ for the three ingredients right now). Raising the additional information per ingredient just to two fields describing moisture, two for instability, and two for lift would already result in a rulebase with, for example, $3^6 = 729$ rules. In this case it might as well make sense to include a weighting within the additional information describing the same ingredient. This would lead to a fuzzy logic system which is quite challenging to understand and to tune. Thus in case of an extension of the fuzzy logic system the results have to be interpreted carefully and the performance needs to be tested properly. Beside the additional work needed the question is if there is enough additional information which would be useful if the fuzzy logic combination shall not be too heavily dependent on the quality of the used model data. Some ideas to this point will be presented in the outlook as well.

However, instead of discussing potential extensions in more detail, the current fuzzy logic system shall be central here. A crucial part of the combination is the definition of the fuzzy input fuzzy sets and the sensitivity of the results to the chosen values for the transition ranges. The input fuzzy sets where, as mentioned above, chosen to differentiate between values of the additional data which proofed to be non-forcing, neutral, or forcing for CI development with the help of the performed statistical analysis. After the statistical analysis the aim was to lose as few hits as possible while filtering as much false alarms as possible. The transition from the non-forcing to the neutral range is chosen by filtering almost

no hits at this boundary. The transition between neutral and forcing is located at values of the additional data where the difference between the percentages of false alarms and hits, that might be filtered successfully, peaks. Thus the definition of the values where the transition ranges are located can be advocated with the statistical analysis shown earlier, but how important are the exact values of the boundaries of the transition ranges? To answer this question, two aspects have to be scrutinized: the influence of the exact values and of the size of the transition range. The exact boundaries seem to be less important. Small shifts here shall cause only small shifts in the distribution of the results. More important is the choice of the transition range. In case of too small transition ranges the fuzzy logic combination almost loses its fuzzyness. Too steep transitions, thus small ranges, result in distinct peaks around the center values of the different output fuzzy sets. In contrast, too shallow transition ranges might smooth the distribution of the resulting values too much and might lead to less differentiation by weakening the influence and validity of the different input fuzzy sets. In addition, if the transition gets too smooth the considerations where to place the transition ranges might become effectively meaningless.

A potential point of criticism of the data combination presented within this thesis might be, that the VERA surface θ_e is used twice. As indicator for moisture and, together with model θ_e on other levels, to quantify instability with the KO-Index. This leads to the assumption that both ingredients would indicate a possible filtering for the same cells due to the low θ_e value which makes the additional information of the other levels used within the KO-Index less important. The intersection of the cells that might be filtered due to low θ_e (low moisture) and the ones with high KO-Index value (stable) was tested during the analysis of the additional data (not shown). The cells with low θ_e have KO values covering more or less the whole spectrum of the KO values from fairly unstable to stable conditions. If only cells with KO values indicating stable conditions are tested, the corresponding θ_e values cover again a wide range covering not only dry conditions. Thus low moisture might coexist with unstable conditions and vice versa despite using the same surface θ_e information for the calculation.

In the following the gain for the Cb-TRAM CI nowcast by using the additional data in line with the applicability of the resulting nowcast will be considered in more detail. As shown in Fig. 5.18, the achievable amount of filtered false alarms without losing hits is around 5 % of all false alarms. If the loss of hits is accepted the same Figure shows, that the maximum difference in the percentage of filtered false alarms and hits is achieved by a sacrifice of around 25 % of all hits. This would lead to the possibility of filtering almost 65 % of the false alarms. The most useful filtering depends on the aims of the user of the nowcast. If false alarms are problematic for the user the amount of false alarms should be filtered despite losing many hits as well. In case the situational awareness for new developments is important for the user and false alarms are less problematic in comparison to the awareness, the filtering shall be performed carefully and only few or no hits shall be filtered. Of course, there are not just these two possibilities to chose for a user, any value within the transition zone between filtering 5 % and 65 % of the false alarms (losing 0 - 25 % hits) can be chosen and translated to a CI forcing

value. An aspect of the CI warnings which is not analyzed within this thesis are warnings concerning CIT (Convectively Induced Turbulence). Also detections of cells which did not develop to a mature storms might be useful to send off a CIT warning. This aspect is difficult to evaluate due to a lack of sufficient verification data.

An important point about the applicability of the data fusion is the sufficiently frequent and near-realtime availability of the additional data sources. In the case of the data used within this study the five minute rapid scan data is used for the Cb-TRAM analysis. The time difference between the satellite scanning time and the final Cb-TRAM detection results is around six minutes for the current domain. The lightning data of the last five minutes before the analyzed time step is easily available then. The additional data used for the data combination is updated only hourly with the data used right now. Thus all Cb-TRAM analysis time steps between full hours need to use the VERA and COSMO-EU data for the last full hour within the postprocessing. However, the data to use is available in time for every Cb-TRAM analyzed time step. The time the postprocessing needs is dependent on the amount of objects and the size of the domain, but does not take more than one additional minute. If the methodology is going to be used operationally the input data might be updated more often depending on the domain. The synop data which is needed as input for VERA is measured in 10 or 15 minute intervals by most of the European national weather services. Important for the quality of the VERA output is the availability of enough data points within and also around the domain that should be analyzed. Thus if the analysis is performed for example for the Terminal Maneuvering Area (TMA) of the airports Frankfurt or Munich the DWD synop data has to be complemented with synop data of the respective neighboring countries. For the model data there are two possible ways of raising the update ratio of the data. The model output which is available hourly might be interpolated for timesteps in between or the model used might be changed to one with a higher frequency of output time steps. The change of the model would lead to a repetition of the statistical analysis to define the input fuzzy sets for the combination properly. However, the change of the domain might lead to a shift in the statistical distributions anyway. Thus a change of the analyzed domain will require adequate testing and, in most cases, some appropriate fine tuning to optimize the results for the area. In addition it would make sense to build up a permanently growing (for an extended period) statistical dataset for each domain where the postprocessing is used. The newly gained information by enlargement of the statistics can not only be used for tuning of the existing combination, it might as well enable to include information on diurnal or seasonal changes of some indicators. The same θ_e value, for example, might imply a different future behavior for a CI detection on a March morning than for a CI cell on an early afternoon in August.

In summary, the gain of the presented object-based concept to build a combination of multiple data sources for postprocessing of satellite-based CI detections and their nowcasts was shown successfully. Adapting the methodology for different CI nowcasting methods can be realized quite easily. The most important points for generalization of the results can be summed up in a few requirements.

Beside a nowcasting tool detecting CI suspect objects the keys to maximize the achievable gain for the detection quality are

- a reasonable verification method to distinct honestly between hits and false alarms.
- a sufficient sample size for the statistical analysis to design reasonable input fuzzy sets.
- the quality of the additional data sources in reflecting the available ingredients and thus in differentiating between hits and false alarms.

In line with the results of the study presented here similar studies were launched by other groups as well. Walker et al. (2012) announce for example statistical approaches that include nonsatellite datasets for the reduction of false alarms which are currently developed for SATCASTv2. The UW-CIMSS (University of Wisconsin – Cooperative Institute for Meteorological Satellite Studies) is currently working on the analysis of additional data and its time trends within satellite detected (CI) objects as well (Wayne Feltz, personal communication, March 2012). Parts of their study, that will be presented in Hartung et al. (2012) and Sieglaff et al. (2012), introduce a "satellite-based object tracking framework" which "functions as an independent vehicle from which a fused array of meteorological data including satellite-derived [...] Cloud-Top Cooling (CTC) rates, radar, and lightning information for each cloud-object can be examined simultaneously". For analyzing the additional data, they are remapping the multiple data fields into the objects and do not apply any further data fusion. The research on SATCASTv2 and the work around UWCI at UW-CIMSS are the only other groups which are referenced here within the discussion despite of the huge amount of other convective nowcasting studies because of the common and quite unique focus on (originally satellite-based) CI detection within these tools and this thesis.

All in all, the results shown within this thesis and these other studies mentioned here show opportunities how to successfully improve CI nowcasting. An important point to mention is that the approach presented here can be a value-adding postprocessing on top of the existing CI detection abilities for either of the mentioned satellite-based CI detection approaches, Cb-TRAM stage 1, UWCI, or SATCASTv2, or comparable tools. Furthermore, this object based analysis of multiple data sources seems to be an excellent methodology for nowcasting purposes anyway, not just for CI, as will be further explained in the outlook at the end of the next Chapter.

7. Conclusions and outlook

The key scientific question of this thesis is, how much can CI nowcasting be improved by the use of multiple data sources? The results produced with the newly developed concept demonstrate the gain of using multiple data sources. The concept of the newly developed methodology and the results produced therewith will be concluded here briefly, leading to an outlook of possible future extensions, developments, and adaptations of the ideas presented within this thesis.

Summing up the approach, Cb-TRAM rapid scan CI detections are evaluated with a new object-based CI verification methodology. For reduction of false alarms a postprocessing with multiple additional data sources is introduced. CI detections, which do not already include lightning signals, are further analyzed with regard to their convective environment with the help of surface observation analysis data and NWP model data. Information on available moisture, instability, and lift is combined with a fuzzy logic data fusion approach, resulting in a CI forcing value per each Cb-TRAM CI detection. These CI forcing values are translated into a probability of further development for each CI object, which is a valuable additional information.

The satellite-based Cb-TRAM algorithm is used as a state-of-the-art tool for the early detection of initiating convection. An object-based verification method, developed especially for these CI nowcasts is used to evaluate the detection abilities of the algorithm. With the aim to get proper statistics, the verification was performed for all CI suspects within a full convective season over Central Europe. Then a postprocessing methodology is applied to get additional information on these Cb-TRAM CI detections. Within the detected objects multiple data sources are analyzed to get additional information on moisture, instability, and lift for each cloud object. Low-level moisture is represented by surface θ_e from VERA, instability by a KO-Index calculation using VERA and COSMO-EU θ_e data, and lift by COSMO-EU ω data in 500 hPA. The data is combined with fuzzy logic, incorporating expert knowledge refined by the statistical information from the verification results. This fuzzy logic combination leads to a CI forcing value resembling the available ingredients. The CI forcing value can be translated into a CI probability for each single object, which describes, based on the forcing in the additional data, the probability that the CI detection develops to a mature storm. The statistical data gained within this study helps to decrease the amount of false alarms in the satellite based detections tremendously without losing too many hits.

Between 5 and up to tremendous 65 % of the false alarms can be filtered this way, depending on the amount of omitted hits ($\sim 0 - 25$ %). This results in an oFAR dev ranging between 0.790 and 0.649 and an oPOD dev between 0.270 and

0.204. These values in comparison to the values without the postprocessing are summed up in Tab. 7.1.

	Cb-TRAM	+ lightning	CI forcing ≥ 3.0 - 5 % false alarms	CI forcing ≥ 7.0 - 25 % hits - 65 % false alarms
oPOD dev	0.249	0.271	0.270	0.204
oFAR dev	0.810	0.799	0.790	0.649

Table 7.1.: Summary of the verification results for Cb-TRAM alone, with the lightning data, and for the lower and upper boundary of CI forcing value filtering thresholds discussed in the text.

One might wonder why removing 65 % of all false alarms reduces oFAR dev only to 0.649. The original oFAR dev close to 0.8 means that approximately four out of five cells are false alarms, or in other words, false alarms are four times as much as hits. After the maximum filtering only two out of three cells remain false alarms, thus false alarms are reduced to only twice as much as the hits.

The highest CI forcing values of the data combination show a CI probability slightly above 50 % while the cells with the lowest CI forcing values have a CI probability as little as 3 % or even less. This new additional CI probability information is available for each CI suspect cloud object after the postprocessing and is a valuable additional information gained by the additional data. Depending on the needs of potential users of the detections and nowcasts the cloud objects with low CI forcing values might be marked as less probable or they might even be filtered completely, which raises the quality of the detections and nowcasts.

This work is a proof of concept that the object-based analysis of additional data sources within these objects marked as CI suspects in advance, is of value and helps, as expected, to increase the usefulness of the detections and the nowcasts produced from them. Beside the improvement with the postprocessing the Cb-TRAM rapid scan CI detection itself seems to be noisier than the normal scan CI detection and might be improved by some tuning.

This leads directly to an outlook on possible future work. It is quite easy to test and extend the analysis within the polygons for additional fields which are not yet incorporated with the methodology developed here. If an input fuzzy set for any other data is defined it might be simply used to exchange one of the currently used additional data. In that case the fuzzy logic rulebase needs only minor adjustment. In cases where additional data should join the existing ones, additional work on a meaningful combination within the rulebase is needed.

Possible additional data sources for the future of the CI detection are for example:

- satellite-based upper tropospheric divergence fields:
EUMETSAT Meteosat Product Extraction Facilities (MPEF) generate an Upper-troposphere Divergence (MPEF DIV) product (EUMETSAT, 2005) with Atmospheric Motion Vectors (AMVs) from SEVIRI WV 6.2 channel data (for SEVIRI channels see Tab. 3.1). The MPEF DIV product is useful

for analyzing the preconvective environment and nowcasting of areas with higher CI potential as shown in Georgiev & Santurette (2010). Within this study, 76 % of all convective cells initiated at areas of divergence seen by MPEF DIV product.

- different NWP model data, e.g. the COSMO-DE EPS:
There are several possible gains for CI nowcasting with the COSMO-DE EPS. First of all the rapid update cycle of the COSMO-DE EPS in comparison to the COSMO-EU might be helpful. Moreover the direct CI forecast by the model mean might be better in the EPS than within the single run and besides this the background fields describing the moisture, instability, and lift might be more realistic in the model mean of the EPS and thus more useful for CI nowcasting. A keyword in using the COSMO-DE EPS for the type of CI nowcasting presented in this study, which might be very useful, is the "best member selection". If Cb-TRAM is not only applied on the observational data but also on the synthetic satellite images from the model, the EPS member which resembles the observational analysis best might be used preferably for the nowcasting of the detections.
- moisture from GPS (Global Positioning System) tomography:
GPS tomography, to investigate water vapor variability, might get interesting for CI nowcasting purposes in the future as well if sufficient resolution is achieved.

Another important development in the future will be the change from the current MSG satellites to the next generation of geostationary EUMETSAT satellites, Meteosat Third Generation (MTG)¹, in a few years which might increase the abilities of the CI detection already before the postprocessing with additional data. The so-called Flexible Combined Imager (FCI) on MTG will continue the operation of the MSG SEVIRI instrument with higher spatial resolution and faster repeat cycles. The spatial sampling distance will be 1 km for all visible channels (5) and all near infrared channels (3) and 2 km for all infrared channels (8). The normal scan will be repeated every 10 minutes and the European rapid scan, which covers approximately one quarter of the full disk, will have a repeat cycle of 2.5 minutes. In the rapid scan mode there will be two additional channels in the solar domain (visible and near infrared), with a spatial resolution of 0.5 km, and two in the thermal domain (infrared), with a spatial resolution of 1 km. Besides these improvements two additional important improvements will be available with MTG. A lightning imager which delivers information on total lightning in near real-time and an infrared sounder which will provide profiles of water vapor (2 km vertical resolution) and temperature (1 km vertical resolution) hourly. More information can be found within the instrument description on the EUMETSAT MTG website¹.

In addition to analyzing the current conditions for an evaluation of the CI forcing, the concept of the object-based analysis of multiple data sources also

¹<http://www.eumetsat.int/Home/Main/Satellites/MeteosatThirdGeneration/index.htm>

establishes the possibility to evaluate trends within the additional fields, like in Hartung et al. (2012).

Besides all the discussed aspects of using this object-based analysis of multiple additional data sources for the nowcasting of CI, this concept shall be adapted for studies on the whole lifecycle of a storm cell in the near future. The ingredients used for evaluating CI forcing have to be exchanged if other development stages shall be examined. For example, data like trends of a flash density, trends in the cloud top temperature or trends in the strength of a radar echo, etc., have to be evaluated if the current state (intensifying, steady, or decaying) of a mature storm shall be analyzed. However the methodology itself can be used more or less unaltered. The key is and will be a sufficient size of a data set to get reasonable statistics for the construction of the fuzzy sets. The same approach might be possible for other phenomena too, as long as they can be described with objects, thus areas of interest for a phenomenon, and additional data influencing the future development of the phenomenon in the object. Examples for other phenomena which might be analyzed with object-based fusion of multiple data sources are aircraft icing, freezing rain, or winter weather objects.

A. Data basis

data source	data type	timestep Δt	resolution Δx
MSG ¹ (rapid scan)	HRV	5 min	1 km (nadir)
	WV 6.2	5 min	3 km (nadir)
	IR 10.8	5 min	3 km (nadir)
	IR 12.0	5 min	3 km (nadir)
LINET	lightning data	5 min files ²	≤ 150 m error ³
COSMO-EU ⁴	ω	1 h	7 km
	T	1 h	7 km
	geopotential	1 h	7 km
	relative humidity	1 h	7 km
Surface observations	synop data	1 h	irregular spacing of stations
already further processed (input data):			
Cb-TRAM (MSG rapid scan)	detection polygons	5 min	3 km (nadir)
VERA (synop data)	u, v, θ, θ_e ⁵	1 h	8 km
	MFD ⁶	1 h	8 km
	T, p, vorticity, etc. ⁷	1 h	8 km

- 1) overview over all MSG channels in Tab. 3.1
- 2) files listing all current lightning detections with update ratio of 5 minutes
- 3) positioning of the lightning detection might have an error ≤ 150 m
- 4) on 10 standard pressure levels between 1000 - 200 hPa
- 5) utilized VERA output fields
- 6) calculated analogous to VERA (see Appendix B) with different smoothing
- 7) non used VERA output, for further description see VERA webpage⁸

Table A.1.: Overview of the different data collected for the analysis presented within this thesis. The data covers the period from May 15 2009 until August 31 2009. The Central European domain over which the analysis is performed is shown for example by the Cb-TRAM output in Fig. 3.6 or the maps in Appendix F. The different data sources are described in more detail in Ch. 3.

⁸<http://www.univie.ac.at/amk/vera/>

B. VERA moisture flux

The equation for the calculation of moisture flux convergence (MFC) can be derived from the conservation of water vapor in pressure coordinates as shown in Banacos & Schultz (2005):

$$\frac{dq}{dt} = S \quad (\text{B.1})$$

where

$$\frac{dq}{dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \omega \frac{\partial}{\partial p} \quad (\text{B.2})$$

u , v , and ω represent the wind components in pressure coordinates, q is the specific humidity, S is a storage term for water vapor which is replaced by evaporation and precipitation in the following ($S = E - P$). Using the mass continuity equation ($\partial u / \partial x + \partial v / \partial y + \partial \omega / \partial p = 0$) the Equation (B.1) can be rewritten in flux form:

$$\frac{\partial q}{\partial t} + u \frac{\partial q}{\partial x} + v \frac{\partial q}{\partial y} + \omega \frac{\partial q}{\partial p} + q \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} \right) = E - P \quad (\text{B.3})$$

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} (qu) + \frac{\partial}{\partial y} (qv) + \frac{\partial}{\partial p} (q\omega) = E - P \quad (\text{B.4})$$

$$\underbrace{\frac{\partial q}{\partial t}}_{\text{local rate of change of } q} + \underbrace{\nabla \cdot (q\vec{v}_h)}_{\text{-horizontal MFC}} + \underbrace{\frac{\partial}{\partial p} (q\omega)}_{\text{-vertical MFC}} = \underbrace{E - P}_{\text{sources and sinks}} \quad (\text{B.5})$$

Therefore the horizontal moisture flux convergence can be written as:

$$MFC = -\nabla \cdot (q\vec{v}_h) = -\vec{v}_h \cdot \nabla q - q \nabla \cdot \vec{v}_h \quad (\text{B.6})$$

$$MFC = \underbrace{-u \frac{\partial q}{\partial x} - v \frac{\partial q}{\partial y}}_{\text{advection term}} - \underbrace{q \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)}_{\text{convergence term}} \quad (\text{B.7})$$

The specific humidity q is defined as ratio of the density of water vapor to the density of wet air, whereas the mixing ratio m is the ratio of the density of water vapor to the density of dry air:

$$q = \frac{\rho_W}{\rho} = \frac{\rho_W}{\rho_L + \rho_W} \quad (\text{B.8})$$

$$m = \frac{\rho_W}{\rho_L} \quad (\text{B.9})$$

The values for q and m are nearly equivalent (Wallace & Hobbs, 2006). For the calculation of MFC in VERA the specific humidity q is replaced by the mixing ratio m . m can then be calculated from Equation (3.8) by usage of the θ and θ_e fields. Looking at the parameters in the exponent, $L \sim 10^6$, $m \sim 10^{-3}$, $c_p \sim 10^3$, and $T \sim 10^2$, the exponent is very small and can be approximated by $e^x \cong 1 + x$. For pressure values close to $p_0 = 1000 \text{ hPa}$ the ratio of θ/T can be approximated as 1.

$$\theta_e \cong \theta \left[1 + \frac{Lm}{c_p T} \right] = \theta + \frac{Lm}{c_p} \underbrace{\frac{\theta}{T}}_{\sim 1} \quad (\text{B.10})$$

Resulting in the following equation for the mixing ratio m :

$$m \cong (\theta_e - \theta) \frac{c_p}{L} \quad (\text{B.11})$$

Using the mixing ratio m instead of q in Equation (B.7) and replacing the spatial derivations by finite differences leads to:

$$MFC = -u \frac{\Delta m}{\Delta x} - m \frac{\Delta u}{\Delta x} - v \frac{\Delta m}{\Delta y} - m \frac{\Delta v}{\Delta y} \quad (\text{B.12})$$

where $\Delta x = \Delta y = \Delta_{xy}$.

In contrast to the MFC calculation in Banacos & Schultz (2005) VERA produces moisture flux divergence (MFD) fields (Kaufmann, 2006): $MFD = -1 \times MFC$. For the calculation of Δm , Δu , and Δv for one gridpoint the values of the four neighboring gridpoints are used as illustrated in Fig. B.1. Therefore, the resulting MFD field is one pixel per edge smaller than the original fields for m , u , and v . Using index i for east-west direction and index j for north-south direction we get:

$$MFD_{i,j} = (m_{i+1,j}u_{i+1,j} - m_{i-1,j}u_{i-1,j} + m_{i,j+1}v_{i,j+1} - m_{i,j-1}v_{i,j-1}) \frac{1}{2\Delta_{xy}} \quad (\text{B.13})$$

Positive values resemble moisture flux divergence, negative values moisture flux convergence of the order of $10^{-4} \text{ g kg}^{-1} \text{ s}^{-1}$.

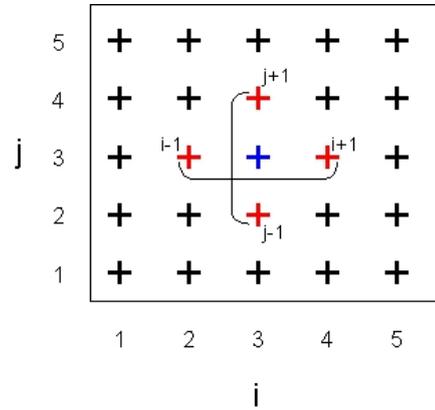


Figure B.1.: MFD calculation in VERA. For the MFD value at the central gridpoint (blue) the wind speed and mixing ratio values at the four neighboring gridpoints (red) are utilized.

C. Fuzzy logic rulebase

IF $data1(proCI) > 0$ AND $data2(proCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "very high"* > 0

IF $data1(proCI) > 0$ AND $data2(proCI) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(proCI) > 0$ AND $data2(proCI) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(proCI) > 0$ AND $data2(neutral) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(proCI) > 0$ AND $data2(neutral) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(proCI) > 0$ AND $data2(neutral) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "medium"* > 0

IF $data1(proCI) > 0$ AND $data2(contraCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(proCI) > 0$ AND $data2(contraCI) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "medium"* > 0

IF $data1(proCI) > 0$ AND $data2(contraCI) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "low"* > 0

IF $data1(neutral) > 0$ AND $data2(proCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(neutral) > 0$ AND $data2(proCI) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(neutral) > 0$ AND $data2(proCI) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "medium"* > 0

IF $data1(neutral) > 0$ AND $data2(neutral) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(neutral) > 0$ AND $data2(neutral) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "medium"* > 0

IF $data1(neutral) > 0$ AND $data2(neutral) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "low"* > 0

IF $data1(neutral) > 0$ AND $data2(contraCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "medium"* > 0

IF $data1(neutral) > 0$ AND $data2(contraCI) > 0$ AND $data3(neutral) > 0$;
THEN *CI forcing "low"* > 0

IF $data1(neutral) > 0$ AND $data2(contraCI) > 0$ AND $data3(contraCI) > 0$;
THEN *CI forcing "low"* > 0

IF $data1(contraCI) > 0$ AND $data2(proCI) > 0$ AND $data3(proCI) > 0$;
THEN *CI forcing "high"* > 0

IF $data1(contraCI) > 0$ AND $data2(proCI) > 0$ AND $data3(neutral) > 0$;

```
    THEN CI forcing "medium" > 0
IF data1(contraCI) > 0 AND data2(proCI) > 0 AND data3(contraCI) > 0;
    THEN CI forcing "low" > 0
IF data1(contraCI) > 0 AND data2(neutral) > 0 AND data3(proCI) > 0;
    THEN CI forcing "medium" > 0
IF data1(contraCI) > 0 AND data2(neutral) > 0 AND data3(neutral) > 0;
    THEN CI forcing "low" > 0
IF data1(contraCI) > 0 AND data2(neutral) > 0 AND data3(contraCI) > 0;
    THEN CI forcing "low" > 0
IF data1(contraCI) > 0 AND data2(contraCI) > 0 AND data3(proCI) > 0;
    THEN CI forcing "low" > 0
IF data1(contraCI) > 0 AND data2(contraCI) > 0 AND data3(neutral) > 0;
    THEN CI forcing "low" > 0
IF data1(contraCI) > 0 AND data2(contraCI) > 0 AND data3(contraCI) > 0;
    THEN CI forcing "very low" > 0
```

D. Abbreviations and acronyms

acc	accumulated evaluation
AMV	Atmospheric Motion Vector
ANC	Auto-Nowcaster
CAPE	Convective Available Potential Energy
CAWCR	Centre for Australian Weather and Climate Research
Cb	Cumulonimbus
Cb-TRAM	Cumulonimbus Tracking And Monitoring
CDM	Collaborative Decision Making
CI	convection initiation
CIN	Convective Inhibition
CIT	Convectively Induced Turbulence
CONRAD	CONvection in RADar
COSMO	CONsortium for Small-scale MOdeling
COSMO-DE	DWD COSMO model version with a grid spacing of 2.8 km covering Germany
COSMO-EU	DWD COSMO model version with a grid spacing of 7 km covering Europe
CRA	Contiguous Rain Areas
CSI	Critical Success Index
CTC	Cloud-Top Cooling
DAS	Displacement and Amplitude Score
dev	further development to stages 2 or 3
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
DMC	Deep Moist Convection
DWD	Deutscher Wetterdienst
EL	Equilibrium Level
EPS	Ensemble Prediction System
ESSL	European Severe Storms Laboratory

ESWD	European Severe Weather Database
EUMETSAT	EUropean organisation for the exploitation of METeorological SATellites
FAR	False Alarm Ratio
GANDOLF	Generating Advanced Nowcasts for Deployment in Operational Land-based flood Forecasts
GOES	Geostationary Operational Environmental Satellite
GPS	Global Positioning System
HRV	high-resolution visible
ICP	spatial forecast verification methods InterComparison Project
INCA	Integrated Nowcasting through Comprehensive Analysis
IR	infrared
KO-Index	German: KOnvektiv-Index
LAMP	Localized Aviation MOS Product
LCL	Lifted Condensation Level
LFC	Level of Free Convection
LINET	Llghtning NETwork
LNB	Level of Neutral Buoyancy
MCC	Mesoscale Convective Complexes
MCS	Mesoscale Convective System
MFC	Moisture Flux Convergence
MFD	Moisture Flux Divergence
MODE	Method for Object-based Diagnostic Evaluation
MOS	Model Output Statistics
MPEF	Meteosat Product Extraction Facilities
MPEF DIV	MPEF Upper-troposphere Divergence
MSG	Meteosat Second Generation
MSL	Mean Sea Level
MTG	Meteosat Third Generation
NCAR	National Center for Atmospheric Research
NIMROD	Nowcasting and Initialization for Modeling using Regional Observation Data
NIR	near infrared
NOAA	National Oceanic and Atmospheric Administration

NO _x	nitrogen oxide (NO _x = NO+NO ₂)
NWP	numerical weather prediction
oFAR	object-based False Alarm Ratio
ω	vertical motion in the COSMO model data
oPOD	object-based Probability Of Detection
POD	Probability Of Detection
QPF	Quantitative Precipitation Forecast
Rad-TRAM	RADAR TRACKING AND MONITORING
RADVOR-OP	German acronym of a DWD nowcasting tool
RDT	Rapid Developing Thunderstorms
SAL	Structure–Area–Location error
SAM	Statistical Advection Model
SATCAST	SATELLITE CONVECTION ANALYSIS AND TRACKING ALGORITHM
SEVIRI	SPINNING ENHANCED VISIBLE AND INFRA-RED IMAGER
SI	French: Système International d'unités
SPC	Storm Prediction Center
SR	Success Ratio
SZA	Solar Zenith Angle
θ_e	Equivalent potential temperature
TITAN	Thunderstorm Identification, Tracking, Analysis and Nowcasting
TMA	Terminal Maneuvering Area
UTC	Universal Time Coordinated
UTLS	Upper Troposphere Lower Stratosphere
UW-CIMSS	University of Wisconsin - Cooperative Institute for Meteorological Satellite Studies
UWCI	UW-CIMSS CI nowcast algorithm
VERA	Vienna Enhanced Resolution Analysis
VIS	visible
WMO	World Meteorological Organization
WV	water vapor
WWRP	World Weather Research Program
ZAMG	Zentralanstalt für Meteorologie und Geodynamik (Central Institute for Meteorology and Geodynamics, Vienna, Austria)

E. Links

Links in order of appearance in the text:

Guide "Forecasting Severe Convective Storms" on the ESTOFEX website:

<http://www.estofex.org/guide/>

Storm Prediction Center (SPC) of NOAA's National Weather Service:

<http://www.spc.noaa.gov/>

Website of the European Severe Storms Laboratory (ESSL):

<http://www.essl.org/>

The European Severe Weather Database (ESWD) website:

<http://www.eswd.eu>

EUMETSAT MSG website:

<http://www.eumetsat.int/Home/Main/Satellites/MeteosatSecondGeneration/Services/index.htm>

CAWCR website of the Joint Working Group on Forecast Verification Research:

<http://www.cawcr.gov.au/projects/verification/>

NOWCAST website for information on LINET:

<http://www.nowcast.de/>

Spatial Forecast Verification Methods Intercomparison Project website:

<http://www.ral.ucar.edu/projects/icp/>

TITAN website:

<http://www.ral.ucar.edu/projects/titan/home/>

VERA website:

<http://www.univie.ac.at/amk/vera/>

COSMO website:

<http://www.COSMO-model.org>

COSMO-EU website at DWD:

http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&state=maximized&_windowLabel=T12203837091139841917821&T12203837091139841917821gsbDocumentPath=Navigation%2FForschung%2FAnalyse__Modellierung%2FFU__NM__LME__node.html%3F__nnn%3Dtrue&_pageLabel=_dwdwww_spezielle_nutzer_forschung_analyse&switchLang=en

COSMO-DE website at DWD:

http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&state=maximized&_windowLabel=T12203837091139841917821&T12203837091139841917821gsbDocumentPath=Navigation%2FForschung%2FAnalyse__Modellierung%2FFU__NM__LME__node.html%3F__nnn%3Dtrue&_pageLabel=_dwdwww_spezielle_nutzer_forschung_analyse&switchLang=en

41917821gsbDocumentPath=Navigation%2FForschung%2FAnalyse__Modellieru
ng%2FFU__NM__LMK__node.html%3F__nnn%3Dtrue&_pageLabel=_dwdwww_spezi
elle_nutzer_forschung_analyse&switchLang=en

EUMETSAT MTG website:

[http://www.eumetsat.int/Home/Main/Satellites/MeteosatThirdGeneration/
index.htm](http://www.eumetsat.int/Home/Main/Satellites/MeteosatThirdGeneration/index.htm)

F. Additional figures

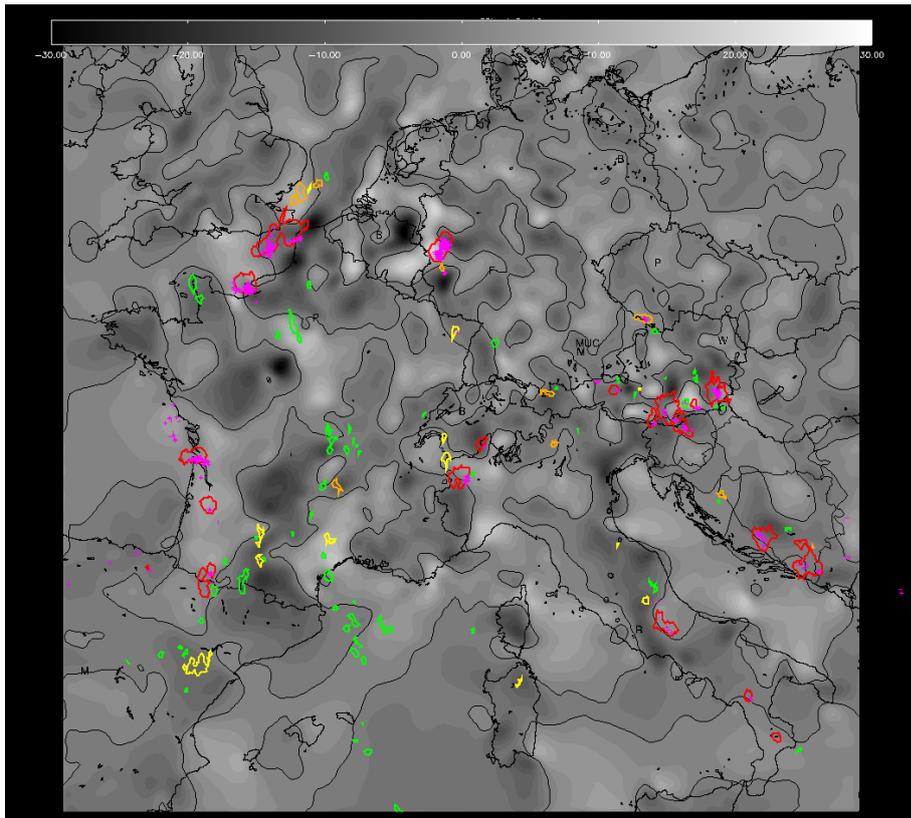


Figure F.1.: Example plot of Cb-TRAM output and LINET data over VERA moisture flux divergence plots from May 25 2009, 15 UTC. The gray shading represents the moisture flux in $\text{g}/(\text{kg s})$ ranging from -30 to $30 \text{ g}/(\text{kg s})$. Light gray represents divergence (positive values) and dark gray represents convergence (negative). The yellow, orange, and red contours have the common colors as in Cb-TRAM, the green polygons show CI cells which do not develop to a stage 2 or 3 detection. Thus it shows cells which would be counted as false alarms.

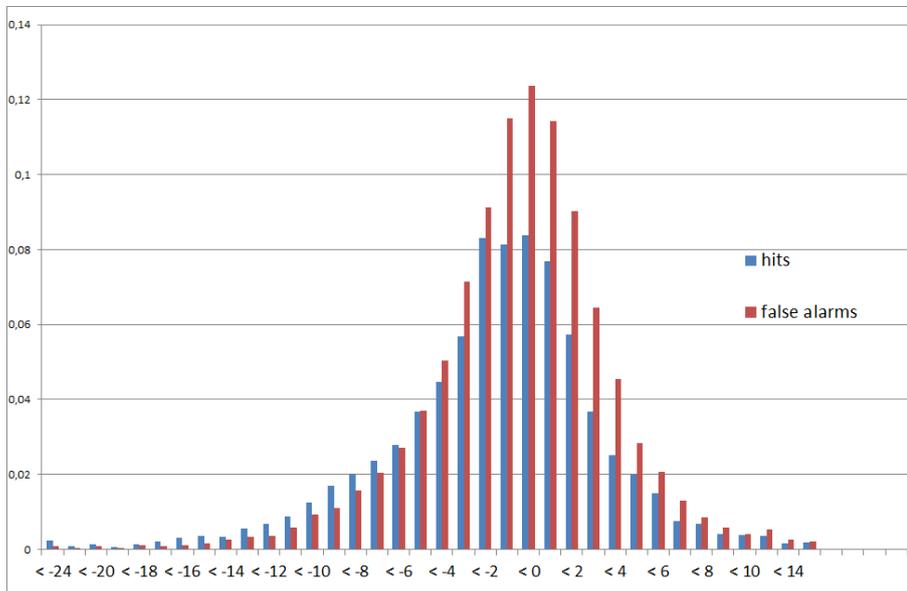


Figure F.2.: The amount of hits(blue)/false alarms(red) per 1 g/(kg s) bins for moisture flux values between lower than -20 g/(kg s) (negative means MFC) and below 10 g/(kg s) (positive means MFD), with larger bins outside of that range, normalized by the total amount of hits/false alarms.

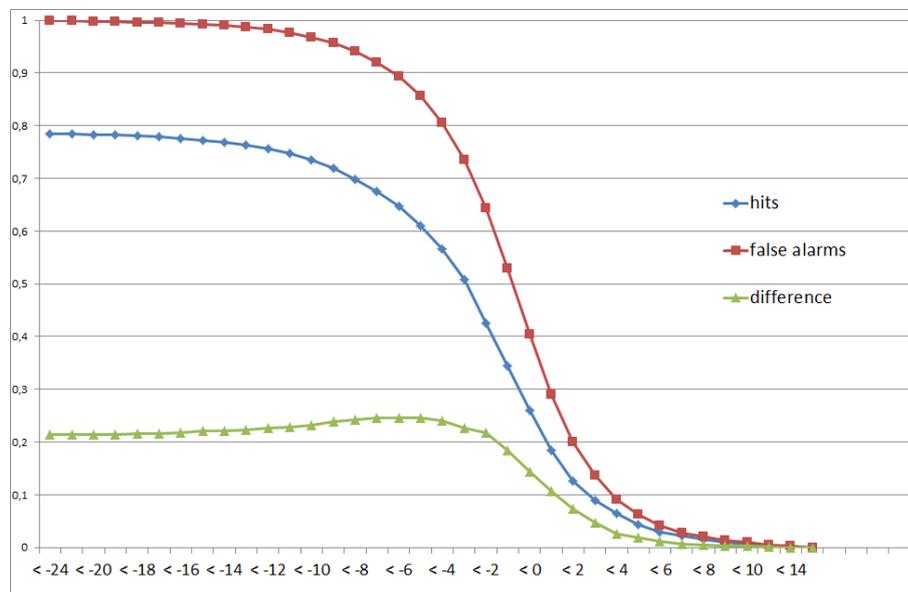


Figure F.3.: Cumulative percentage of MFD value frequencies of the given moisture flux values for hits (blue), false alarms (red), and the difference of the two values (green).

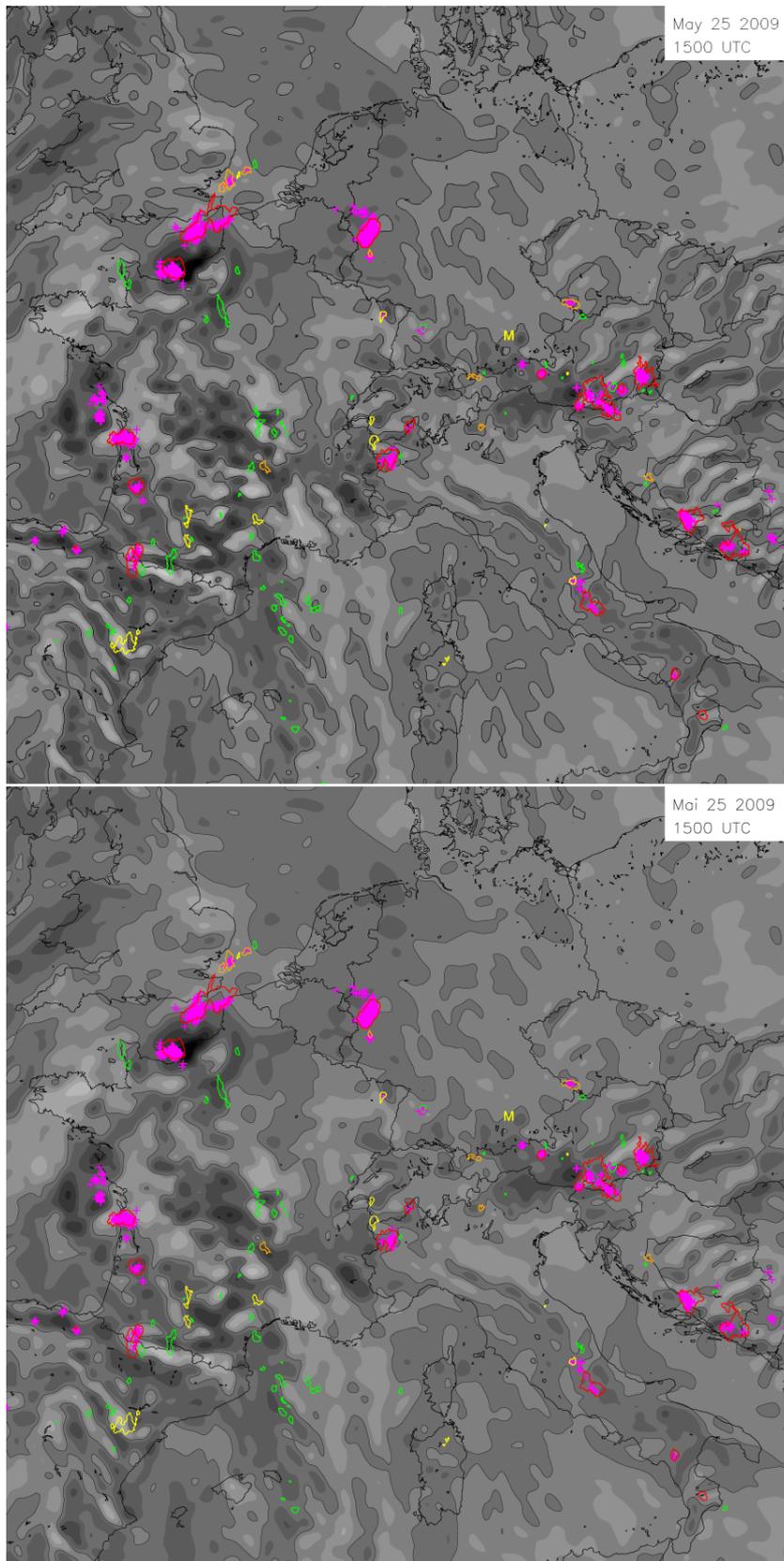


Figure F.4.: Example plot for COSMO-EU ω data in 500 hPa showing the data without any smoothing by a Gaussian kernel (top) and smoothed by a small kernel with an edge length of 5 pixels (bottom). For further descriptions see Chapter 5.

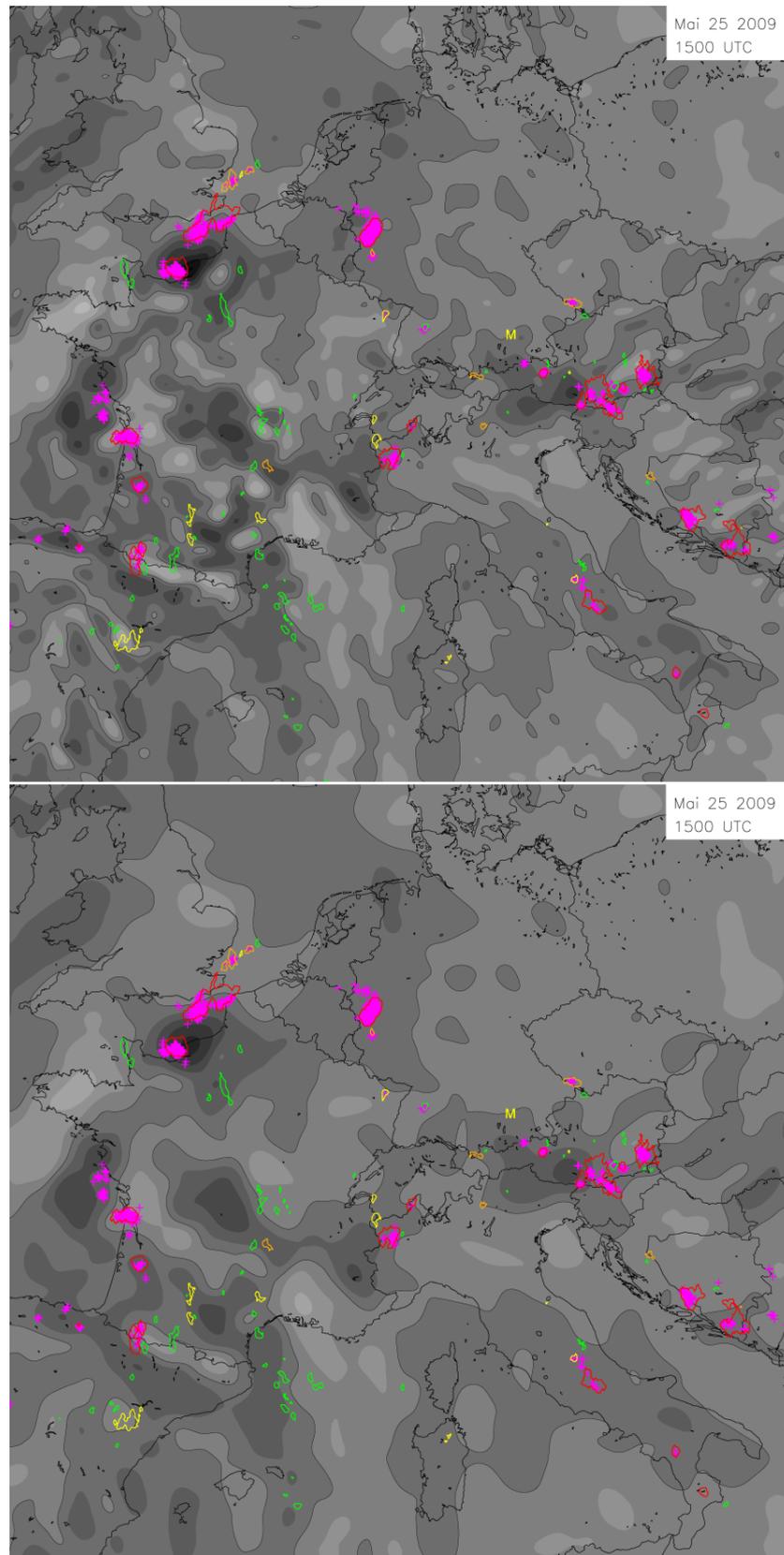


Figure F.5.: As Fig. F.4 but for a kernel edge length of 9 pixels (top) and the finally used edge length of 17 pixels (bottom).

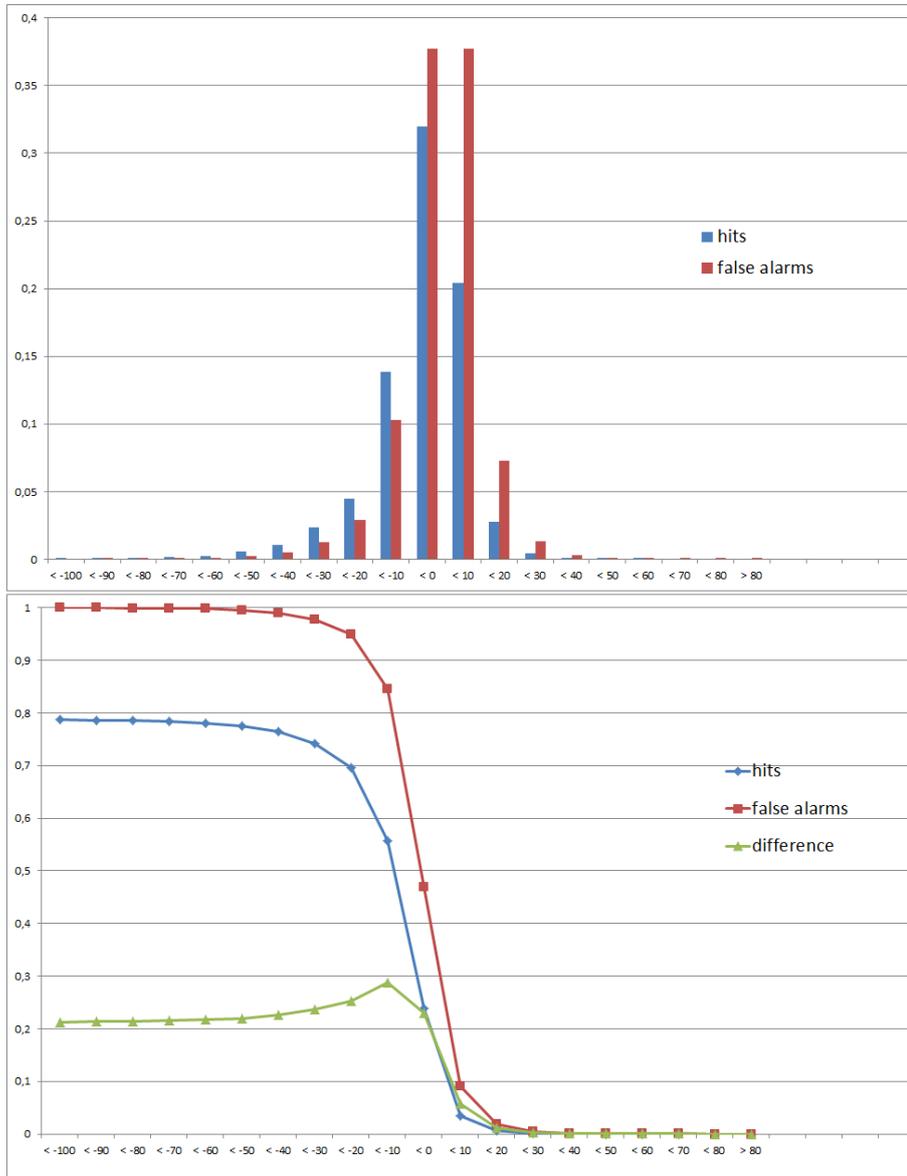


Figure F.6.: The amount of hits(blue)/false alarms(red) for each ω value between lower than -100 hPa/h and higher than 80 hPa/h in 10 hPa/h bins, normalized by the total amount of hits/false alarms (top). Cumulative percentage (bottom) of ω value frequencies above the given ω in hPa/h for hits (blue), false alarms (red), and the difference of the two values (green).

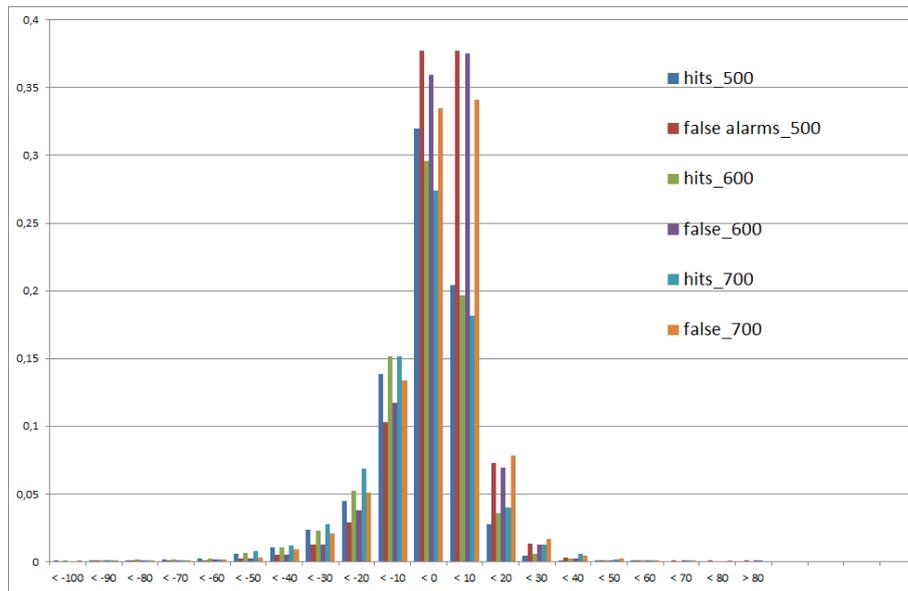


Figure F.7.: As Fig. F.6 (top) but for 500, 600, and 700 hPa/h.

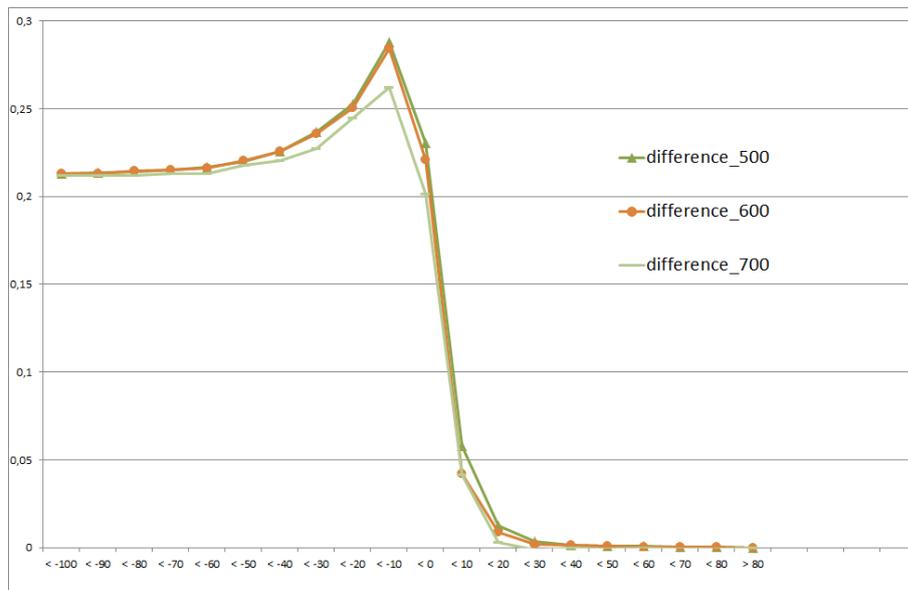


Figure F.8.: As Fig. F.6 (bottom), but showing the difference values of cumulative percentages, for 500, 600, and 700 hPa/h.

List of Figures

1.1. Concept	6
2.1. skew T log p diagram	13
2.2. Life cycle of a storm	16
2.3. Schematic of multi- and supercells	17
2.4. Wind shear and storm type	18
3.1. MSG areas	24
3.2. Cb-TRAM example plot	28
3.3. Pixel based verification and positioning errors	33
3.4. Double penalty and closeness	33
3.5. Spatial verification methods	35
3.6. Cb-TRAM analysis area	38
3.7. LINET coverage	39
3.8. VERA MFD example	41
3.9. VERA θ_e example	42
3.10. DWD COSMO model domains	43
3.11. Orography within DWD COSMO models	44
3.12. Conventional and fuzzy sets	46
3.13. Fuzzy logic classification for temperature	46
3.14. Fuzzy sets example for CI	47
3.15. Example output fuzzy sets	49
4.1. Matching of Nowcast and Analysis	52
4.2. Verification search radius	54
4.3. Verification with nowcast track	55
5.1. Lightning measurement in CI cell	61
5.2. Schematic of parallax offset	61
5.4. VERA θ_e statistics 1	63
5.3. Cb-TRAM detections on VERA θ_e	64
5.5. VERA θ_e statistics 2	65
5.6. VERA θ_e fuzzy set	66
5.7. KO Index statistics 1	67
5.8. KO Index statistics 2	68
5.9. KO Index fuzzy set	69
5.10. ω no kernel	70
5.11. ω with kernel	71
5.12. ω statistics 1	72

5.13. ω statistics 2	72
5.14. ω fuzzy set	73
5.15. Example fixed thresholds	75
5.16. Example output fuzzy sets	77
5.17. Combination statistics 1	78
5.18. Combination statistics 2	79
5.19. Combination statistics 3	80
B.1. VERA grid for MFD calculation	98
F.1. Cb-TRAM detections on VERA MFD	107
F.2. VERA MFD statistics 1	108
F.3. VERA MFD statistics 2	109
F.4. ω kernel intercomparison part 1	110
F.5. ω kernel intercomparison part 2	111
F.6. ω statistics 3	112
F.7. ω level comparison statistics 1	113
F.8. ω level comparison statistics 2	113

List of Tables

3.1. SEVIRI channels	24
3.2. Contingency table	30
3.3. Finley's Tornado forecasts	32
3.4. COSMO model comparison	45
3.5. Rulebase scheme	48
4.1. Cb-TRAM CI stage verification statistics 1	57
4.2. Cb-TRAM CI stage verification statistics 2	58
5.1. KO values literature	67
5.2. Benefit by tested additional data	75
5.3. Combination with fixed thresholds	76
7.1. Summary of the verification results	92
A.1. Data basis	95

Bibliography

- Ahijevych, D., Gilleland, E., Brown, B., & Ebert, E. (2009). Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Wea. Forecasting*, 24, 1485–1497.
- Aumann, H. H., DeSouza-Machado, S. G., & Behrangi, A. (2011). Deep convective clouds at the tropopause. *Atmos. Chem. Phys.*, 11, 1167–1176.
- Baldauf, M., Förstner, J., Klink, S., Reinhardt, T., Schraff, C., Seifert, A., & Stephan, K. (2011a). Kurze Beschreibung des Lokal-Modells Kürzestfrist COSMO-DE (LMK) und seiner Datenbanken auf dem Datenserver des DWD. *Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Offenbach, Germany*.
- Baldauf, M., Helmert, K., Haßler, B., Förstner, J., Reinhardt, T., Seifert, A., & Lenz, C.-J. (2007). The new Very Short Range Forecast Model COSMO-LMK for the Convection-Resolving scale. *COSMO User Seminar, Langen, Germany, 05 - 07 March 2007*.
- Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011b). Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Mon. Wea. Rev.*, 139, 3887–3905.
- Baldwin, M. & Kain, J. S. (2006). Sensitivity of several performance measures to displacement error, bias, and event frequency. *Wea. Forecasting*, 21, 636–648.
- Banacos, P. C. & Schultz, D. M. (2005). The Use of Moisture Flux Convergence in Forecasting Convective Initiation: Historical and Operational Perspectives. *Wea. Forecasting*, 20, 351–366.
- Bartels, H., Weigl, E., Klink, S., Kohler, O., Reich, T., Rosenow, W., Lang, P., Podlasly, C., Winterrath, T., Adrian, G., Majewski, D., & Lang, J. (2005). Projekt RADVOR-OP: Radargestützte, zeitnahe Niederschlagsvorhersage für den operationellen Einsatz (Niederschlag-Nowcasting-System). *Final Report, Deutscher Wetterdienst available via: <http://www.dwd.de/radvor-op>*.
- Bedka, K., Brunner, J., Dworak, R., Feltz, W., Otkin, J., & Greenwald, T. (2010). Objective Satellite-Based Detection of Overshooting Tops Using Infrared Window Channel Brightness Temperature Gradients. *J. Appl. Meteorol. Clim.*, 49, 181–202.

- Berendes, T., Mecikalski, J., MacKenzie Jr., W., Bedka, K., & Nair, U. (2008). Convective cloud identification and classification in daytime satellite imagery using standard deviation limited adaptive clustering. *J. Geophys. Res.*, 113.
- Betz, H. D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W. P., Defer, E., Dziewit, Z., & Konarski, J. (2009). LINET—An international lightning detection network in Europe. *Atmos. Res.*, 91, 564–573.
- Betz, H. D., Schmidt, K., Oettinger, W. P., & Montag, B. (2008). Cell-tracking with lightning data from LINET. *Adv. Geosci.*, 17, 55–61.
- Bica, B., Knabl, T., Steinacker, R., Ratheiser, M., Dorninger, M., Lotteraner, C., Schneider, S., Chimani, B., Gepp, W., & Tschannett, S. (2007). Thermally and Dynamically Induced Pressure Features over Complex Terrain from High-Resolution Analyses. *J. Appl. Meteorol. Clim.*, 46, 50–65.
- Bjerknes, J. (1938). Saturated-adiabatic ascent of air through dry-adiabatically descending environment. *Q. J. Roy. Meteor. Soc.*, 64, 325–330.
- Bowler, N., Pierce, C., & Seed, A. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Q. J. Roy. Meteor. Soc.*, 132, 2127–2155.
- Brooks, H. E. (2009). Proximity soundings for severe convection for Europe and the United States from reanalysis data. *Atmos. Res.*, 93(1–3), 546 – 553.
- Brooks, H. E., Lee, J. W., & Craven, J. P. (2003). The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, 67–68, 73 – 94.
- Byers, H. R. & Braham, R. R. (1949). The Thunderstorm. *Report of the Thunderstorm Project, U.S. Weather Bureau, Dept. of Commerce.*
- Casati, B. (2010). New Developments of the Intensity-Scale Technique within the Spatial Verification Methods Intercomparison Project. *Wea. Forecasting*, 25, 113–143.
- Casati, B., Ross, G., & Stephenson, D. (2004). A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.*, 11, 141–154.
- Casati, B., Wilson, L., Stephenson, D., Nurmi, P., Ghelli, A., Pocerlich, M., Damrath, U., Ebert, E., Brown, B., & Mason, S. (2008). Forecast verification: Current status and future directions. *Meteorol. Appl.*, 15, 3–18.
- Dahl, J. M. L., Höller, H., & Schumann, U. (2011). Modeling the Flash Rate of Thunderstorms. Part II: Implementation. *Mon. Wea. Rev.*, 139, 3112–3124.
- Dance, S., Ebert, E., & Scurrah, D. (2010). Thunderstorm Strike Probability Nowcasting. *J. Atmos. Ocean. Tech.*, 27, 79–93.

- Davis, C., Brown, B., & Bullock, R. (2006a). Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, 134, 1772–1784.
- Davis, C., Brown, B., & Bullock, R. (2006b). Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, 134, 1785–1795.
- Davis, C., Brown, B., Bullock, R., & Halley-Gotway, J. (2009). The method for object-based diagnostic evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Wea. Forecasting*, 24, 1252–1267.
- Dixon, M. & Wiener, G. (1993). TITAN: Thunderstorm Identification, Tracking, Analysis and Nowcasting – A radar based methodology. *J. Atmos. Ocean. Tech.*, 10, 785–797.
- Donovan, M. F., Williams, E. R., Kessinger, C., Blackburn, G., Herzegh, P. H., Bankert, R. L., Miller, S., & Mosher, F. R. (2008). The Identification and Verification of Hazardous Convective Cells over Oceans Using Visible and Infrared Satellite Observations. *J. Appl. Meteorol. Clim.*, 47, 164–184.
- Dorninger, M., Schneider, S., & Steinacker, R. (2008). On the interpolation of precipitation data over complex terrain. *Meteorol. Atmos. Phys.*, 101.
- Doswell, C. (1985). The operational meteorology of convective weather. Vol. 2: Storm scale analysis. *NOAA Tech. Memo. ERL ESG-15*.
- Doswell, C., Ed. (2001). *Severe Convective Storms*. Meteorological Monographs, Vol. 28, No. 50, pp. 561, American Meteorological Society, Boston.
- Doswell, C. & Schultz, D. (2006). On the Use of Indices and Parameters in Forecasting Severe Storms. *E-Journal of Severe Storms Meteorology*, 1.
- Dotzek, N. & Forster, C. (2011). Quantitative comparison of METEOSAT thunderstorm detection and nowcasting with in situ reports in the European Severe Weather Database (ESWD). *Atmos. Res.*, 100, 511–522.
- Dotzek, N., Groenemeijer, P., Feuerstein, B., & Holzer, A. M. (2009). Overview of ESSL’s severe convective storms research using the European Severe Weather Database ESWD. *Atmos. Res.*, 93(1–3), 575 – 586.
- Dupree, W., Morse, D., Chan, M., Tao, X., Reiche, C., Iskenderian, H., Wolfson, M., Pinto, J., Williams, J., Albo, D., et al. (2009). The 2008 CoSPA Forecast Demonstration. *ARAM Special Symposium on Weather-Air Traffic Management Integration*.
- Durrán, D. & Klemp, J. (1982). On the effects of moisture on the Brunt–Väisälä frequency. *J. Atmos. Sci.*, 39(10), 2152–2158.

- Dye, J., Winn, W., Jones, J., & Breed, D. (1989). The electrification of New Mexico thunderstorms 1. Relationship between precipitation development and the onset of electrification. *J. Geophys. Res.*, 94, 8643–8656.
- Ebert, E. & McBride, J. (2000). Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrology*, 239, 179–202.
- Ebert, E., Wilson, L. J., Brown, B. G., Nurmi, P., Brooks, H. E., Bally, J., & Jaeneke, M. (2004). Verification of nowcasts from the WWRP Sydney 2000 Forecast Demonstration Project. *Wea. Forecasting*, 19, 73–96.
- Ebert, E. E. (2008). Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteorol. Appl.*, 15, 51–64.
- Ebert, E. E. (2009). Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, 24, 1498–1510.
- Ebert, E. E. & Gallus, W. A. (2009). Toward Better Understanding of the Contiguous Rain Area (CRA) Method for Spatial Forecast Verification. *Wea. Forecasting*, 24, 1401–1415.
- Emanuel, K. (1994). *Atmospheric Convection*. Oxford University Press, New York.
- EUMETSAT (2005). Upper Level Divergence Product Algorithm description. *EUM/MET/REP/05/0163*.
- Eurocontrol Performance Review Commission (2011). Performance Review Report 2010. Available at <http://www.eurocontrol.int/articles/european-ans-performance-review>.
- Finley, J. P. (1884). Tornado predictions. *Amer. Meteor. J.*, 1, 85–88.
- Fischer, H., De Reus, M., Traub, M., Williams, J., Lelieveld, J., De Gouw, J., Warneke, C., Schlager, H., Minikin, A., Scheele, R., et al. (2003). Deep convective injection of boundary layer air into the lowermost stratosphere at mid-latitudes. *Atmos. Chem. Phys.*, 3(3), 739–745.
- Fujita, T. & Wakimoto, R. (1981). Five Scales of Airflow Associated with a Series of Downbursts on 16 July 1980. *Mon. Wea. Rev.*, 109, 1438–1456.
- Georgiev, C. & Santurette, P. (2010). Quality of mpef divergence product as a tool for very short range forecasting of convection. *Proceedings of 2010 EUMETSAT Meteorological Satellite Conference, Córdoba, Spain, 20–24 September 2010*.
- Ghirardelli, J. E. & Glahn, B. (2010). The Meteorological Development Laboratory's Aviation Weather Prediction System. *Wea. Forecasting*, 25, 1027–1051.
- Gilbert, G. K. (1884). Finley's tornado predictions. *Amer. Meteor. J.*, 1, 166–172.

- Gilleland, E., Ahijevych, D., Brown, B., Casati, B., & Ebert, E. (2009). Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, 24, 1416–1430.
- Gilleland, E., Ahijevych, D., Brown, B., & Ebert, E. (2010a). Verifying forecasts spatially. *B. Am. Meteorol. Soc.*, 91, 1365–1373.
- Gilleland, E., Lindström, J., & Lindgren, F. (2010b). Analyzing the Image Warp Forecast Verification Method on Precipitation Fields from the ICP. *Wea. Forecasting*, 25, 1249–1262.
- Glahn, H. & Lowry, D. (1972). The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteorol.*, 11(8), 1203–1211.
- Golding, B. (1998). Nimrod: a system for generating automated very short range forecasts. *Meteorol. Appl.*, 5, 1–16.
- Gremillion, M. & Orville, R. (1999). Thunderstorm characteristics of cloud-to-ground lightning at the Kennedy Space Center, Florida: A study of lightning initiation signatures as indicated by the WSR-88D. *Wea. forecasting*, 14(5), 640–649.
- Hafner, S., Reitter, R., & Seltmann, J. (2004). Developments in the DWD radar-network. *Proceedings of ERAD 2004*, (pp. 425–427).
- Hagen, M., Bartenschlager, B., & Finke, U. (1999). Motion characteristics of thunderstorms in southern Germany. *Meteorol. Appl.*, 6, 227–239.
- Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., & Gruber, C. (2011). The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. *Wea. Forecasting*, 26, 166–183.
- Haklander, A. & Van Delden, A. (2003). Thunderstorm predictors and their forecast skill for the Netherlands. *Atmos. Res.*, 67, 273–299.
- Hartung, D. C., Sieglaff, J. M., Cronce, L. M., & Feltz, W. F. (2012). An Inter-Comparison of UWCI-CTC Algorithm Cloud-Top Cooling Rates with WSR-88D Radar Data. *Submitted to Mon. Wea. Rev.*
- Haupt, S., Pasini, A., & Marzban, C. (2009). *Artificial intelligence methods in the environmental sciences*. Springer Verlag.
- Häberli, C., Groehn, I., Steinacker, R., Pöttschacher, W., & Dorninger, M. (2004). Performance of the surface observation network during MAP. *Meteorol. Z.*, 13, 109–121.
- Hoffmann, J. M. (2008). Entwicklung und Anwendung von statistischen Vorhersage-Interpretationsverfahren fuer Gewitternowcasting und Unwetterwarnungen unter Einbeziehung von Fernerkundungsdaten. *Dissertation, FU-Berlin*, 205 pp.

- Holton, J. R. (2004). *An introduction to dynamic meteorology*. International geophysics series, San Diego, New York: Academic Press, Elsevier, 4th ed.
- Houze, R. (1993). *Cloud dynamics*. International Geophysics Series, Vol. 53, Academic Press, Inc.
- Huntrieser, H., Schiesser, H. H., Schmid, W., & Waldvogel, A. (1996). Comparison of Traditional and Newly Developed Thunderstorm Indices for Switzerland. *Wea. Forecasting*, 12, 108–125.
- Jolliffe, I. T. & Stephenson, D. B. (2003). *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley.
- Kaltenböck, R., Diendorfer, G., & Dotzek, N. (2009). Evaluation of thunderstorm indices from ECMWF analyses, lightning data and severe storm reports. *Atmos. Res.*, 93, 381–396.
- Kalthoff, N., Kohler, M., Barthlott, C., Adler, B., Mobbs, S., Corsmeier, U., Träumner, K., Foken, T., Eigenmann, R., Krauss, L., Khodayar, S., & Di Girolamo, P. (2010). The dependence of convection-related parameters on surface and boundary-layer conditions over complex terrain. *Q. J. Roy. Meteor. Soc.*
- Kaufmann, H. (2006). Die mesoskalige Analyse der Feuchteflussdivergenz im Alpenraum. *Diplomarbeit zur Erlangung des akademischen Grades Magistra der Naturwissenschaften an der Fakultät für Geowissenschaften, Geographie und Astronomie der Universität Wien*.
- Keenan, T., Joe, P. and Wilson, J., Collier, C., Golding, B., Burgess, D., May, P., Pierce, C., Bally, J., Crook, A., Sills, D., Berry, L. and Bell, I., Fox, N., Pielke Jr., R., Ebert, E., Eilts, M., O'Loughlin, K., Webb, R., Carbone, R., Browning, K., Roberts, R., & Mueller, C. (2003). The Sydney 2000 World Weather Research Programme Forecast Demonstration Project: Overview and current status. *B. Am. Meteorol. Soc.*, 84, 1041–1054.
- Keil, C. & Craig, G. (2007). A displacement-based error measure applied in a regional ensemble forecasting system. *Mon. Wea. Rev.*, 135, 3248–3259.
- Keil, C. & Craig, G. (2009). A Displacement and Amplitude Score Employing an Optical Flow Technique. *Wea. Forecasting*, 24, 1297–1308.
- Köhler, M. (2011). Untersuchung der Auslösung von Gewittern während der DLR „Wetter & Fliegen“ Sommerkampagne 2010. *Masterarbeit, Meteorologisches Institut München, LMU München*.
- Klink, S., Stephan, K., & Schraff, C. (2004). Assimilation of radar data in the mesoscale NWP-model of DWD. *ERAD Publication Series*, 2, 155–161.
- Klir, G. & Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall New Jersey.

- Kober, K., Craig, G. C., Keil, C., & Dörnbrack, A. (2012). Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts. *Q. J. Roy. Meteor. Soc.*, 138, 755–768.
- Kober, K. & Tafferter, A. (2009). Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. *Meteorol. Z.*, 18, 75–84.
- Krebs, W., Mannstein, H., Bugliaro, L., & Mayer, B. (2007). Technical note: A new day- and night-time Meteosat Second Generation Cirrus Detection Algorithm MeCiDA. *Atmos. Chem. Phys.*, 7, 6145–6159.
- Kunz, M. (2007). The skill of convective parameters and indices to predict isolated and severe thunderstorms. *Nat. Hazard. Earth Sys.*, 7, 327–342.
- Lack, S., Limpert, G., & Fox, N. (2010). An Object-Oriented Multiscale Verification Scheme. *Wea. Forecasting*, 25, 79–92.
- Lang, P. (2001). Cell Tracking and Warning Indicators derived from operational Radar Products. *AMS Radar Conference, Munich, 2001*.
- Lang, T., Miller, L., Weisman, M., Rutledge, S., Barker III, L., Bringi, V., Chandrasekar, V., Detwiler, A., Doesken, N., Helsdon, J., et al. (2004). The Severe Thunderstorm Electrification and Precipitation Study. *B. Am. Meteorol. Soc.*, 85(8), 1107–1125.
- Leighton, Q. (2006). Modeling and simulation needs for next generation air transportation system research. *AIAA modeling and simulation technologies conference and exhibit, Keystone, Colorado, 21–24 August 2006. AIAA 2006-6109, pp 1–8*.
- Mannstein, H., Huntrieser, H., & Wimmer, S. (2002). Determination of the mass flux in convective cells over Europe. *Proceedings EUMETSAT Meteorological Satellite Conference 2002, Dublin, Ireland, EUMETSAT*, (pp. 264–269).
- Mannstein, H., Meyer, R., & Wendling, P. (1999). Operational detection of contrails from NOAA-AVHRR data. *Int. J. Remote Sens.*, 20, 1641–1660.
- Manzato, A. & Morgan, G. (2003). Evaluating the sounding instability with the Lifted Parcel Theory. *Atmos. Res.*, 67, 455–473.
- Markowski, P. & Richardson, Y. (2010). *Mesoscale Meteorology in Midlatitudes*. John Wiley & Sons, Ltd.
- Marzban, C. & Sandgathe, S. (2008). Cluster analysis for object-oriented verification of fields: A variation. *Mon. Wea. Rev.*, 136, 1013–1025.
- Marzban, C., Sandgathe, S., Lyons, H., & Lederer, N. (2009). Three spatial verification techniques: Cluster analysis, variogram, and optical flow. *Wea. Forecasting*, 24, 1457–1471.

- Mason, I. (1989). Dependence of the critical success index on sample climate and threshold probability. *Aust. Meteorol. Mag.*, 37, 75–81.
- Mass, C. (2012). Nowcasting: The Promise of New Technologies of Communication, Modeling, and Observation. *B. Am. Meteorol. Soc.*, 93, 797–809.
- Mass, C. F., Ovens, D., Westrick, K., & Colle, B. A. (2002). Does Increasing Horizontal Resolution Produce More Skillful Forecasts? *B. Am. Meteorol. Soc.*, 83, 407–430.
- McBride, J. & Ebert, E. (2000). Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. *Wea. Forecasting*, 15, 103–121.
- Mecikalski, J. & Bedka, K. (2006). Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, 134, 49–78.
- Mecikalski, J., Bedka, K., Paech, S., & Litten, L. (2008). A Statistical Evaluation of GOES Cloud-Top Properties for Nowcasting Convective Initiation. *Mon. Wea. Rev.*, 136, 4899–4914.
- Merk, D. (2012). Erkennung und Nowcasting von Convective Initiation basierend auf Meteosat-Daten. *Masterarbeit, Meteorologisches Institut München, LMU München*.
- Mittermaier, M. & Roberts, N. (2010). Intercomparison of Spatial Forecast Verification Methods: Identifying Skillful Spatial Scales Using the Fractions Skill Score. *Wea. Forecasting*, 25, 343–354.
- Morel, C., S en esi, S., Autones, F., & Labatut, L. (2000). The Rapid Developing Thunderstorms (RDT) product of the nowcasting SAF. Prototyping activities and quality assessment using GOES images. *Proceedings, The 2000 Meteorological Satellite Data Users' Conference, Eumetsat and CNR, Bologna, Italy*, (pp. 698–705).
- Mueller, C., Saxen, T., Roberts, R., Wilson, J., Betancourt, T., Dettling, S., Oien, N., & Yee, J. (2003). NCAR Auto-Nowcast System. *Wea. Forecasting*, 18, 545–561.
- Murphy, A. (1993). What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, 8, 281–293.
- Murphy, A. H. (1996). The Finley Affair: A Signal Event in the History of Forecast Verification. *Wea. Forecasting*, 11, 3–20.
- Pfl uger, U. (2004). Operational verification of vertical profiles at DWD. *COSMO Newsletter*, 4, 95–103.

- Pierce, C. E., Hardaker, P. J., Collier, C. G., & Haggett, C. M. (2000). GANDOLF: a system for generating automated nowcasts of convective precipitation. *Meteorol. Appl.*, 7, 341–360.
- Pinto, J., Dupree, W., Weygandt, S., Wolfson, M., Benjamin, S., & Steiner, M. (2010). Advances in the Consolidated Storm Prediction for Aviation (CoSPA). *Preprints, 14th Conf. on Aviation, Range, and Aerospace Meteorology, Atlanta, GA, Amer. Meteor. Soc., J11.2. Available online at <http://ams.confex.com/ams/pdfpapers/163811.pdf>.*
- Purdom, J. (1976). Some uses of high-resolution GOES imagery in the mesoscale forecasting of convection and its behavior. *Mon. Wea. Rev.*, 104, 1474–1483.
- Radová, M. & Seidl, J. (2008). Parallax applications when comparing radar and satellite data. *2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, Germany 8 - 12 September 2008.*
- Rivolta, G., Marzano, F., Coppola, E., Verdecchia, M., et al. (2006). Artificial neural-network technique for precipitation nowcasting from satellite imagery. *Adv. Geosci.*, 7, 97–103.
- Roberts, N. M. & Lean, H. W. (2008). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, 136, 78—97.
- Roberts, R. D., Anderson, A. R. S., Nelson, E., Brown, B. G., Wilson, J. W., Pocerlich, M., & Saxen, T. (2012). Impacts of Forecaster Involvement on Convective Storm Initiation and Evolution Nowcasting. *Wea. Forecasting*, 27, 1061—1089.
- Roberts, R. D. & Rutledge, S. (2003). Nowcasting storm initiation and growth using GOES-8 and WSR-88D data. *Wea. Forecasting*, 18, 562–584.
- Romero, R., Gayà, M., & Doswell, C. A. (2007). European climatology of severe convective storm environmental parameters: A test for significant tornado events. *Atmos. Res.*, 83(2–4), 389 – 404.
- Ross, T. (2010). *Fuzzy Logic with Engineering Applications, Third Edition*. John Wiley & Sons, Ltd.
- Rudack, D. E. & Ghirardelli, J. E. (2010). A Comparative Verification of Localized Aviation Model Output Statistics Program (LAMP) and Numerical Weather Prediction (NWP) Model Forecasts of Ceiling Height and Visibility. *Wea. Forecasting*, 25, 1161–1178.
- Ruzanski, E., Chandrasekar, V., & Wang, Y. (2011). The CASA Nowcasting System. *J. Atmos. Ocean. Tech.*, 28, 640–655.

- Saxen, T. R., Mueller, C. K., Warner, T. T., Steiner, M., Ellison, E. E., Hatfield, E. W., Betancourt, T. L., Dettling, S. M., & Oien, N. A. (2008). The Operational Mesogamma-Scale Analysis and Forecast System of the U.S. Army Test and Evaluation Command. Part IV: The White Sands Missile Range Auto-Nowcast System. *J. Appl. Meteorol. Clim.*, 47, 1123–1139.
- Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). An Introduction to Meteosat Second Generation (MSG). *B. Am. Meteorol. Soc.*, 83, 977–992.
- Schmetz, J., Tjemkes, S. A., Gube, M., & van de Berg, L. (1997). Monitoring deep convection and convective overshooting with METEOSAT. *Adv. Space Res.*, 19, 433–441.
- Schneider, S., Chimani, B., Kaufmann, H., Bica, B., Lotteraner, C., Tschanett, S., & Steinacker, R. (2008). Nowcasting of a supercell storm with VERA. *Meteorol. Atmos. Phys.*, 102, 23–36.
- Schneider, S. & Steinacker, R. (2009). Utilisation of radar information to refine precipitation fields by a variational approach. *Meteorol. Atmos. Phys.*, 103, 137–144.
- Schättler, U., Doms, G., & Schraff, C. (2011). A description of the nonhydrostatic regional COSMO-Model, Part VII: User's Guide - COSMO-Model 4.21. *Deutscher Wetterdienst, www.COSMO-model.org*.
- Schulz, J. & Schättler, U. (2011). Kurze Beschreibung des Lokal-Modells Europa COSMO-EU (LME) und seiner Datenbanken auf dem Datenserver des DWD. *Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Offenbach, Germany*.
- Schumann, U. & Huntrieser, H. (2007). The global lightning - induced nitrogen oxides source. *Atmos. Chem. Phys.*, 7, 3823–3907.
- Sieglaff, J. M., Counce, L. M., Feltz, W. F., Bedka, K. M., Pavolonis, M. J., & Heidinger, A. K. (2011). Nowcasting Convective Storm Initiation Using Satellite Based Box-Averaged Cloud Top Cooling and Cloud Type Trends. *J. Appl. Meteorol. Clim.*, 50, 110–126.
- Sieglaff, J. M., Hartung, D. C., Feltz, W. F., Counce, L. M., & Lakshmanan, V. (2012). Development and application of a satellite-based convective cloud object-tracking methodology: A multipurpose data fusion tool. *Submitted to J. Appl. Meteorol. Clim.*
- Siewert, C., Koenig, M., & Mecikalski, J. (2010). Application of Meteostat Second Generation data towards improving the nowcasting of convective initiation. *Meteorol. Appl.*
- Sokol, Z. & Pesice, P. (2012). Nowcasting of precipitation – Advective statistical forecast model (SAM) for the Czech Republic. *Atmos. Res.*, 103, 70–79.

- Steinacker, R., Dorninger, M., Wölfelmaier, F., & Krennert, T. (2000a). Automatic tracking of convective cells and cell complexes from lightning and radar data. *Meteorol. Atmos. Phys.*, 72, 101–110.
- Steinacker, R., Häberli, C., & Pöttschacher, W. (2000b). A transparent method for the analysis and quality evaluation of irregularly distributed and noisy observational data. *Mon. Wea. Rev.*, 128, 2303–2316.
- Steinacker, R., Ratheiser, M., Bica, B., Chimani, B., Dorninger, M., Gepp, W., Lotteraner, C., Schneider, S., & Tschannett, S. (2006). A mesoscale data analysis and downscaling method over complex terrain. *Mon. Wea. Rev.*, 134, 2758–2771.
- Steinheimer, M. & Haiden, T. (2007). Improved nowcasting of precipitation based on convective analysis fields. *Adv. Geosci.*, 10, 125–131.
- Stich, D., Forster, C., Zinner, T., & Tafferner, A. (2011). Convection Initiation - Nowcasting by data fusion and its verification. *6th European Conference on Severe Storms (ECSS 2011), 3 - 7 October 2011, Palma de Mallorca, Spain.*
- Tafferner, A., Forster, C., Hagen, M., Keil, C., Zinner, T., & Volkert, H. (2008). 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.
- Tessendorf, S., Miller, L., Wiens, K., & Rutledge, S. (2005). The 29 June 2000 supercell observed during STEPS. Part I: Kinematics and microphysics. *J. Atmos. Sci.*, 62, 4127–4150.
- Walker, J., MacKenzie Jr., W., Mecikalski, J., & Jewett, C. (2012). An Enhanced Geostationary Satellite-based Convective Initiation Algorithm for 0-2 Hour Nowcasting with Object Tracking. *J. Appl. Meteorol. Clim.*, 51, 1931–1949.
- Wallace, J. M. & Hobbs, P. V. (2006). *Atmospheric science : an introductory survey*. International geophysics series: Academic Press, Elsevier, 2nd ed.
- Wapler, K., Goeber, M., & Trepte, S. (2012). Comparative verification of different nowcasting systems to support optimisation of thunderstorm warnings. *Adv. Sci. Res.*, 8, 121–127.
- Wernli, H., Hofmann, C., & Zimmer, M. (2009). Spatial forecast verification methods intercomparison project: Application of the SAL technique. *Wea. Forecasting*, 24, 1472–1484.
- Wernli, H., Paulat, M., Hagen, M., & Frei, C. (2008). SAL - A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, 136, 4470–4487.

- Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences*. International geophysics series, San Diego, New York: Academic Press, Elsevier, 2nd ed.
- Williams, E., Boldi, B., Matlin, A., Weber, M., Hodanish, S., Sharp, D., Goodman, S., Raghavan, R., & Buechler, D. (1999). The behavior of total lightning activity in severe Florida thunderstorms. *Atmos. Res.*, 51, 245–265.
- Wilson, J., Crook, N., Mueller, C., Sun, J., & Dixon, M. (1998). Nowcasting Thunderstorms: A Status Report. *B. Am. Meteorol. Soc.*, 79, 2079–2099.
- Wilson, J., Ebert, E., Saxen, T., Roberts, R., Mueller, C., Sleigh, M., Pierce, C., & Seed, A. (2004). Sydney 2000 Forecast Demonstration Project: Convective Storm Nowcasting. *Wea. Forecasting*, 19, 131–150.
- Wilson, J. & Mueller, C. (1993). Nowcasts of Thunderstorm Initiation and Evolution. *Wea. Forecasting*, 8, 113–131.
- Wilson, J. W. (1966). *Movement and predictability of radar echoes*. National Severe Storms Laboratory Tech. Memo ERTM-NSSL-28.
- Wilson, J. W., Feng, Y., Chen, M., & Roberts, R. D. (2010). Nowcasting Challenges during the Beijing Olympics: Successes, Failures, and Implications for Future Nowcasting Systems. *Wea. Forecasting*, 25, 1691–1714.
- Winterrath, T., Rosenow, W., et al. (2007). A new module for the tracking of radar-derived precipitation with model-derived winds. *Adv. Geosci.*, (pp. 77–83).
- WMO-No. 485 (2012). *Manual on the Global Data-Processing and Forecasting System*. Vol. I, Annex IV to WMO Technical Regulations, Geneva, 2012.
- Wolfson, M., Dupree, W., Rasmussen, R., Steiner, M., Benjamin, S., & Weygandt, S. (2008). Consolidated storm prediction for aviation (CoSPA). *Integrated Communications, Navigation and Surveillance Conference, 2008. ICNS 2008*, (pp. 1–19).
- Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G. C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C. N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G. and Richard, E., Rotach, M. W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D. D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A., & Wirth, M. (2011). The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights. *Q. J. Roy. Meteor. Soc.*

- Zadeh, L. (1965). Fuzzy sets. *Inform. Control*, 8, 338–353.
- Zimmer, M., Craig, G. C., Keil, C., & Wernli, H. (2011). Classification of precipitation events with a convective response timescale and their forecasting characteristics. *Geophys. Res. Lett.*, 38.
- Zinner, T. & Betz, H.-D. (2009). Validation of Meteosat storm detection and nowcasting based on lightning network data. *2009 EUMETSAT Meteorological Satellite Conference, Bath, UK, 21 - 25 September 2009*.
- Zinner, T., Forster, C., de Coning, E., & Betz, H.-D. (2012). Validation of the METEOSAT storm detection and nowcasting system Cb-TRAM with lightning network data - Europe and South Africa. *Submitted to Atmos. Meas. Tech.*
- Zinner, T. & Groenemeijer, P. (2012). Thunderstorms: Thermodynamics and Organization. In U. Schumann (Ed.), *Atmospheric Physics, Research Topics in Aerospace* chapter 7, (pp. 101–114). Springer-Verlag Berlin Heidelberg.
- Zinner, T., Mannstein, H., & Tafferner, A. (2008). 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.

Data acknowledgments

Summary of acknowledgments for the support with data and tools. Some of the data and products used in the text have been provided or produced by the following organizations:

- The European Organisation for the Exploitation of Meteorological Satellites, EUMETSAT, provides the MSG SEVIRI rapid scan data used here within Cb-TRAM.
- The University of Vienna supported this work by making available a version of the Vienna Enhanced Resolution Analysis tool VERA
- The German Weather Service, DWD, provides data from the COSMO-EU model and surface observations which are obtained via the DWD Databases.
- The LINET lightning data used within this study is supplied by *nowcast*.

Acknowledgments

This work would not have been possible without the help of several people. Here, I would like to take the opportunity and acknowledge (hopefully all) the people that supported me to complete this thesis.

First of all I would like to thank Caroline Forster, my advisor, for her constant enthusiasm and help. Our discussions improved my way of working enormously and were a crucial assistance in not losing track in one of the many side topics. When I needed support on any topic, I always knew that I could bother you at all times. Thank you!

I want to express my gratitude to my Ph.D. supervisor George Craig. His conceptual ideas and his useful feedback on any of our discussions supported and improved this thesis considerably.

Additionally I am indebted to my co-supervisor Ulrich Schumann for discussions about my work and that he agreed to examine the thesis.

Moreover I am thankful to Thomas Gerz for providing the possibility to do my Ph.D. thesis in his department and for the interest he showed for my studies as well as for his conviction in my prospective engagement.

Furthermore I'd like to thank Arnold Tafferner for many constructive discussions on fuzzy logic, convection initiation, or IDL, as well as Tobias Zinner for fruitful discussions on the verification of Cb-TRAM.

Thanks to everyone in "Abteilung" 4 for the nice atmosphere and a great time with a lot of fun and productive discussions.

In addition I would like to appreciate Dieter Mayer for his help to get in touch with the input data, settings, and post-processing of my VERA data, and Uli Schättler who was and is my contact person for any question about the DWD databases.

At this point I want to thank my parents, Monica and Jochen, as well as my brother Daniel for their interest in my research and their constant encouragements.

Last but definitely not least I want to thank (Dr.) Theresa, my one and only!