HEIGHT CORRECTION OF ATMOSPHERIC MOTION VECTORS USING SPACE-BORNE LIDAR OBSERVATIONS FROM CALIPSO

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

vorgelegt von
Kathrin Folger
geboren in Göppingen

München, März 2016

1. Gutachter: DR. MARTIN WEISSMANN

2. Gutachter: PROF. DR. BERNHARD MAYER

Tag der mündlichen Prüfung: 9. Mai 2016

Parts of this thesis are contained in:

Kathrin Folger and Martin Weissmann, 2014: Height correction of atmospheric motion vectors using satellite lidar observations from CALIPSO. *Journal of Applied Meteorology and Climatology*, 53, 1809–1819.

Kathrin Folger and Martin Weissmann, 2016: Lidar-based height correction for the assimilation of atmospheric motion vectors. *Journal of Applied Meteorology and Climatology*, submitted.

CONTENTS

Abstract

Atmospheric Motion Vectors (AMVs) provide valuable wind information for the initial conditions of numerical weather prediction models. However, only a small fraction of the available observations is used in current data assimilation systems due to height assignment issues and horizontal error correlations.

The aim of this thesis is to investigate the feasibility of correcting the pressure heights of operational AMVs from the geostationary satellites Meteosat-9 and Meteosat-10 with cloud-top heights derived from independent lidar observations by the polar orbiting Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite. The intention is to reduce the height assignment error as well as the horizontal error correlation of AMVs for their use in data assimilation. Additionally, AMVs are treated as winds in a vertical layer as proposed by several recent studies. Corrected and uncorrected AMV winds are evaluated using radiosonde observations as well as short-term forecasts from the global forecasting system of the German Weather Service.

Firstly, a direct lidar-based height reassignment of AMVs with collocated CALIPSO observations is evaluated. Assigning AMV winds from Meteosat-10 to ~120 hPa deep layers below the lidar cloud top reduces the Vector Root Mean Square (VRMS) differences of AMVs from Meteosat-10 by 8-17% depending on the evaluation method, evaluation period and AMV altitude. In addition, the AMV error correlation is reduced by about 50 km through the correction.

Secondly, CALIPSO observations are used to derive statistical height bias correction functions for a general AMV height correction that can be applied to all operational AMVs from a geostationary satellite. Such a height bias correction achieves on average about 50% of the reduction of VRMS differences attained using the direct height reassignment, but has the clear advantage of avoiding the need for real-time lidar data and directly collocated lidar observations. Initial assimilation and forecast experiments with statistically corrected and layer-averaged Meteosat-10-AMVs in the framework of the current global forecasting system of the German Weather Service reveal encouraging results.

Overall, the results of this thesis demonstrate that height assignment errors of Meteosat-AMVs can be significantly reduced when information from lidar cloud-top observations is incorporated. Thus, lidar-based height correction methods exhibit a promising approach for an improved representation of AMVs in numerical weather prediction models in the future.

Zusammenfassung

Satelliten-Windvektoren (engl. Atmospheric Motion Vectors, kurz AMVs) liefern wertvolle Informationen zu atmosphärischen Windbedingungen, die für die Initialisierung von numerischen Wettervorhersage-Modellen benötigt werden. Allerdings wird nur ein Bruchteil aller verfügbaren AMVs wegen Problemen mit der Höhenzuordnung und horizontalen Fehlerkorrelationen in der Datenassimilation derzeit verwendet.

In dieser Arbeit soll untersucht werden, inwiefern die Druckhöhen von operationellen AMVs von den geostationären Satelliten Meteosat-9 und Meteosat-10 mit Hilfe von satelliten-gestützten Lidarmessungen des polar-umlaufenden Satelliten CALIPSO (engl. Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations) korrigiert werden können, um damit sowohl die fehlerhafte Höhenzuordnung als auch horizontale Fehlerkorrelationen zu verbessern. Zusätzlich werden AMVs, wie bereits von anderen aktuellen Studien vorgeschlagen, als vertikales Schichtmittel betrachtet. Korrigierte und unkorrigierte AMVs werden sowohl mit Radiosonden-Messungen als auch mit Modellfeldern von Kurzzeit-Vorhersagen des Globalmodells des deutschen Wetterdienstes ausgewertet.

Zuerst wird eine direkte Höhenkorrektur von Meteosat-10-AMVs mit Hilfe von nahen CALIPSO-Lidarmessungen der Wolkenoberkante analysiert. Dabei erzielen Schichtmittel einer vertikalen Ausdehnung von 120 hPa unterhalb der Lidar-Wolkenoberkante eine Verringerung der Vector Root Mean Square (VRMS) Differenzen von 8-15%, abhängig von Auswertungsmethode, Auswertungszeitraum und AMV-Höhe. Zusätzlich wird die horizontale Korrelation der AMV-Fehler um ca. 50 km verringert.

Als zweiter Ansatz werden CALIPSO-Lidarmessungen dazu verwendet, statistische Höhenkorrektur-Funktionen abzuleiten, die auf alle AMVs eines bestimmten Satelliten angewendet werden können. Diese statistische Höhenkorrektur erreicht ungefähr 50% der Verbesserung, die durch die direkte Höhenkorrektur erzielt wird, bietet aber den Vorteil, keine direkt benachbarten Lidarmessungen in Echt-Zeit zu benötigen. Erste Assimilations- und Vorhersage-Experimente mit statistisch korrigierten Meteosat-10-AMVs im Globalmodell des deutschen Wetterdienstes zeigen vielversprechende Ergebnisse.

Insgesamt zeigen die Ergebnisse dieser Arbeit, dass die Verwendung von Lidardaten einen signifikanten Beitrag zur Fehlerverringerung von AMVs leistet. Die im Zuge dieser Arbeit vorgestellten lidar-basierten Höhenkorrektur-Methoden bieten daher einen aussichtsreichen Ansatz, AMVs in Wettermodellen zukünftig besser repräsentieren zu können.

1. INTRODUCTION

Satellite observations are crucial ingredients for the initialization of Numerical Weather Prediction (NWP) models as they provide information on atmospheric processes with unprecedented coverage and resolution. The amount and variety of available satellite observations has increased drastically during recent decades and requires complex data processing techniques in order to provide suitable data products to NWP centres in real-time. Nowadays, sophisticated data assimilation systems make better use of observations by applying enhanced forward operators and quality control procedures, as well as better forecast models and assimilation algorithms. Overall, satellite observations continue to play an important role in the forecast performance of NWP models in both hemispheres (Bouttier and Kelly, 2001; Kelly and Thépaut, 2007; English et al., 2013).

One essential information source for the prevalent wind conditions is Atmospheric Motion Vectors (AMVs). AMVs are retrievals of the atmospheric wind derived by tracking cloud and water vapour structures in successive images from multispectral satellite imagery. The displacement of these structures generally characterizes tropospheric motions, and therefore the horizontal wind speed and wind direction can be determined. AMVs provide wind information with unique spatial and temporal coverage especially over the oceans, in polar regions and over the southern hemisphere where in-situ observations are typically rare. Given that the current global observing system is heavily skewed towards mass/temperature observations, reliable and area-covering wind observations in remote areas are an essential data source for global NWP models (Velden et al., 2005; Forsythe, 2007; Weissmann et al., 2012; Baker et al., 2014).

AMVs have been derived since the early 1970s and have been an important ingredient for NWP models since then. In the early stages, AMVs were derived mostly manually. Over the following years, the AMV retrieval greatly enhanced, and nowadays allows a fully automatic extraction of AMVs. Until recently, mainly geostationary satellites were used due to the regular image frequency recording the same scenery, which enables wind information from about 60°N to 60°S to be derived (Menzel, 2001). Since the launch of the first polar-orbiting

satellite carrying a Moderate Resolution Imaging Spectroradiometer (MODIS) instrument in 2002, new possibilities for extracting wind information have been exploited, allowing for AMV derivation also in polar regions (Key et al., 2003; Dworak et al., 2009; Key et al., 2014).

The accuracy as well as the amount of available AMVs steadily increased during previous decades. The sensor technology in the satellite instruments has been enhanced continuously, now offering more channels, shorter time intervals between consecutive images and a higher pixel resolution, which consequently leads to higher data volume and enhanced coverage. Improved computer capacities and efficient telecommunication systems are the basis for an efficient usage of AMVs in NWP models (Forsythe, 2007). Nowadays, AMVs from five geostationary satellites and several polar-orbiting satellites are assimilated routinely in all global NWP systems. Figure 1 illustrates the AMV distribution during one assimilation cycle at the German Weather Service (DWD for Deutscher Wetterdienst). By using imagery from geostationary and polar-orbiting satellites, AMVs are almost globally available. However, a data gap between geostationary and polar-orbiting satellites exists at about 60° latitude for both hemispheres, leading to difficulties in accounting for phenomena such as the polar jets. This problem will be tackled by exploiting the possibility of combining images from different satellites, either with two polar-orbiting satellites like the Meteorological Operational Satellites (METOP) A/B (Hautecoeur et al., 2014) or with one geostationary and one polar-orbiting satellite (Lazarra et al., 2014).

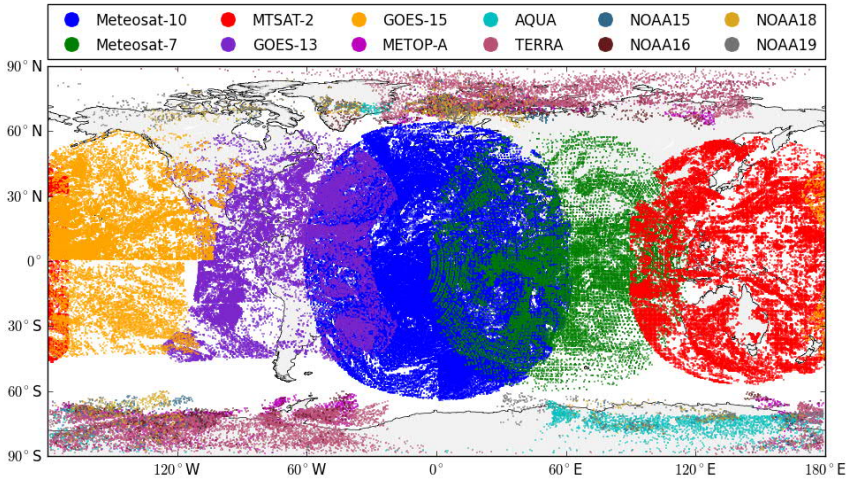

Figure 1: *Illustration of the AMV distribution during one assimilation cycle on 7 May 2013, 12 UTC for the global model of the German Weather Service. All available AMVs (before quality control procedures) from operational geostationary and polar-orbiting satellites with reference times between 10:30 UTC and 13:30 UTC (assimilation window) are shown.*

The positive impact of the assimilation of AMVs on the forecast skill of global NWP systems is frequently reviewed, and results generally emphasize the importance of AMV wind field observations for data assimilation. A recent inter-comparison project initialized by the Met Office and Météo France (Payan and Cotton, 2012) analysed the impact of AMVs at eight international forecasting centres and found that NWP systems continue to benefit from the assimilation of AMVs. Overall, the results of this collaborative impact study demonstrate the significant relative importance of AMVs in the global observing system for data assimilation. Several recent studies have shown that the assimilation of AMVs improves the NWP forecast skill, e.g. from DWD (Cress and Bitzer, 2012), the European Centre for Medium-Range Weather Forecasts (ECMWF) (Rohn et al., 2001; Bouttier and Kelly, 2001), the U.K. Met Office (Cotton et al., 2014), and the U.S. Navy (Baker et al., 2012). The benefit obtained by assimilating AMVs in global NWP models is not restricted to areas where the AMVs are derived but extends to regions further away from the actual AMV position (Santek, 2010). The positive impact of AMVs is particularly pronounced when assimilating AMVs for tropical cyclone track forecasts (Goerss et al., 1998; Soden et al., 2001; Wang et al., 2006). Furthermore, the beneficial contribution of AMVs derived from imagery of polar-orbiting satellites has been demonstrated in several studies (Key et al., 2003; Bormann and Thépaut, 2004; Riishojgaard and Zhu, 2004). Beyond their relevance for NWP models, AMVs additionally contribute to other fields of research, e.g. nowcasting systems (Mecikalski and Bedka, 2006; Merk and Zinner, 2013; Garcia-Pereda, 2014).

1.1 State of the art

Although AMVs have proven to be an important data source for the atmospheric wind field in NWP models, some issues remain. Several studies indicate that the assimilation of AMVs does not yet exploit its full potential. Quality problems may often be the reason why AMV assimilation does not always improve forecast skill more clearly (Bouttier and Kelly, 2001; Isaksen et al., 2010). Studies on AMV error characteristics are therefore an active field of research (e.g. Bresky et al., 2012; Lean et al., 2015; Salonen et al., 2015a).

In particular, the height assignment of AMVs continues to be a challenging task and introduces significant errors. Velden and Bedka (2009) estimated that height assignment errors contribute up to 70% to the total Vector Root Mean Square (VRMS) difference between AMV winds and radiosonde winds. In addition, these errors can be horizontally correlated up to several hundred kilometers (Bormann et al. 2003). A number of error sources

contribute to this: Temperature and humidity model profiles that are used to retrieve the AMV height may contain errors that are often correlated horizontally, and multi-layer clouds or semi-transparent clouds pose a further challenge for the height assignment process. In practice, these issues lead to the need for a massive thinning of the originally dense AMV dataset for data assimilation. Typical thinning strategies at NWP centres have scales of 200 km by 200 km (horizontal extent) by 100 hPa (in the vertical) for one single observation. As a consequence, only a small percentage of all available AMVs are used for data assimilation (Forsythe, 2007).

► Interpreting AMVs as layer-averages

AMVs are traditionally interpreted as single-level observations, and this continues to be the established strategy in operational data assimilation systems to date. However, this assumption might provide a suboptimal representation of AMVs. When clouds are observed by satellite imaging instruments, radiation is generally received from finite vertical layers and not only from the highest cloud level. In particular, upper-level clouds like cirrus are mostly semi-transparent and the detected signal may have contributions not only from the cloud top, but also from subjacent layers of often large vertical extent. Additionally, the motion of clouds that are used as tracking features to derive AMVs rather represents a vertically averaged wind over a cloud layer rather than the wind at the cloud top as is commonly understood (Schmetz et al., 1993; Hernandez-Carrascal and Bormann, 2014; Salonen et al., 2015a).

Several recent studies have revealed that AMVs represent vertical layers rather than discrete levels and that this should be accounted for in NWP models to reduce AMV errors. One of the first studies in this direction was conducted by Rao et al. (2002), suggesting benefits from spreading AMV wind information to more than one level. A subsequent study of Velden and Bedka (2009) compared AMV winds to layer-averaged radiosonde winds at three different locations in the U.S. for over three years. These layers extend from the operational AMV height downward with increasing layer depth from 10 hPa to 300 hPa. Generally, a consistently better agreement with radiosonde winds was found when a layer-averaging was applied, with varying optimal configurations for different satellite channels and altitude regions. Furthermore, Hernandez-Carrascal and Bormann (2014) showed in a simulated framework that assigning AMVs to layers that comprise the actual cloud from cloud top to cloud base shows advantages over the traditional single level assumption. Lean et al. (2015)

also quantified height assignment AMV error characteristics using a set of simulated AMVs and found the closest fit of AMVs to layer-averaged model winds that are most commonly located below the estimated cloud top.

Altogether, the potential of assigning AMVs to vertically extended layers instead of discrete levels has been clearly demonstrated. However, current assimilation systems do not use this information yet because the exact position and depth of this layer are relatively unknown and most likely depend on the AMV dataset and processing (Forsythe, 2007; Weissmann et al., 2013).

► AMV pressure height comparison with lidar cloud-top observations

Lidar observations provide reliable information on cloud-top heights that is independent from the AMV derivation procedures and from model fields used for the processing. For example, the space-borne Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) aboard the polar-orbiting Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite provides the framework for the comparison of AMV pressure heights and lidar cloud-top observations. Initial efforts in this direction at the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) were presented at the International Winds Workshop 2010 and showed a good agreement with CALIPSO lidar observations for low-level AMVs (Sèze et al., 2008). However, high- and mid-level AMVs exhibited large differences between AMV heights and collocated lidar observations. The authors ascribed these mismatches to very thin cirrus clouds at high levels. These can only be observed in a limited way by satellite imaging instruments whereas lidar observations allow for the detection of optically thin clouds. Di Michele et al. (2013) also compared AMV pressure heights from Meteosat-9 with CALIPSO lidar observations for a 10-day period and found similar characteristics. Both studies made first comparisons between AMV and lidar cloud-top observations, but did not investigate the effect of a direct AMV height correction by evaluating wind errors of the lidar-corrected AMVs with radiosondes or model fields. In addition, neither of these comparison studies had collocation constraints in the vertical, which can exclude situations where AMV and CALIPSO might not see the same cloud due to the temporal and spatial distance or diverging instrument capabilities.

A first study on a lidar-based AMV height correction with airborne lidar observations was conducted by Weissmann et al. (2013) and was intended to serve as a testbed for future space-based applications. A small, regional sample of airborne lidar observations was used to

correct AMVs from the Japanese geostationary Multi-Functional Transport Satellite 1R (MTSAT-1R) during The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC, for more information see e.g. Weissmann et al., 2011). The verification was done with aircraft dropsondes launched during the field campaign. Results indicated that assigning AMVs to 100–150 hPa deep layers below the lidar cloud top reduces the AMV wind error by 10–15% compared to layers of the same depth centered at the original AMV height.

1.2 Goals and outline

The aim of this thesis is to further elaborate the two approaches presented in the previous section that can potentially reduce the errors of AMVs. Firstly, satellite lidar cloud-top observations from the polar-orbiting satellite CALIPSO are used to correct AMV pressure heights. Secondly, AMVs are treated as vertically extended layer observations instead of single-level observations. As mentioned above, the optimal layer position and layer depth relative to the derived (operational) AMV pressure height may be highly situation-dependent. Lidar observations, in contrast, provide accurate information on the cloud-top height that is independent of different AMV datasets and processing algorithms. In order to find an appropriate layer that should be assigned to AMVs in data assimilation systems, a number of vertical layers relative to the lidar cloud top and relative to the original AMV height are investigated. Furthermore, different depths of these layers are tested and compared to the traditionally used single-level approach. This procedure is schematically illustrated in Figure 2.

In this work, two different AMV height correction methods based on lidar observations are evaluated. In the first part of this thesis, the effect of applying a direct height reassignment to individual AMVs with collocated CALIPSO observations is assessed. As proposed by Weissmann et al. (2013), this thesis conducts the transition from a limited-area AMV height correction with airborne lidar observations to larger scales using a sample of satellite lidar observations with significantly more AMV observations and a larger time period. Operational collocated radiosondes are used to evaluate AMV winds before and after the height correction. To overcome the limitations of spatially and temporally sparse radiosonde observations, model equivalents from the Global Model GME of DWD are used additionally for the wind evaluation.

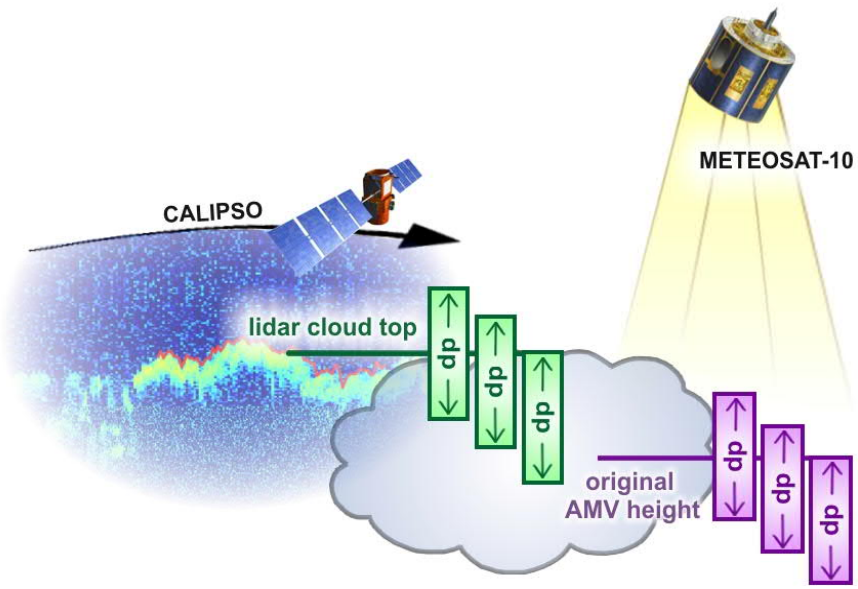

Figure 2: *Illustration of the AMV height correction using space-bo rne lidar observations: Pressure heights of Meteosat-10-AMVs are corrected with lidar observations from the polar-orbiting satellite CALIPSO. Different layer positions and layer depths relative to the lidar cloud-top height (green) and the original AMV height (purple) are evaluated.*

However, such a direct height reassignment can only be applied to collocated AMV/CALIPSO observations, which exhibit a comparatively small subset of all AMVs. As an alternative approach to the direct AMV height reassignment, a general height adjustment of operational AMVs is derived to correct systematic height biases. Applying this height bias correction allows proceeding from a direct height reassignment of individual AMVs to a larger scope of application, as this approach facilitates the usage of lidar information for the AMV height correction without the need for real-time lidar data and directly collocated lidar observations. To investigate the effect of statistically lidar-corrected and layer-averaged AMVs in an operational NWP system, assimilation and forecast experiments are conducted with the current Icosahedral Nonhydrostatic (ICON) global NWP model of DWD. The two methods to correct AMV heights presented in this work are referred to as DIRECT HEIGHT REASSIGNMENT and HEIGHT BIAS CORRECTION throughout the following chapters.

In summary, this thesis intends to address the following scientific questions:

1. Can **space-borne lidar observations** be used to correct pressure heights of AMVs, and which combination of **layer depth** and **layer position** relative to the lidar cloud-top observation provides an optimal representation of the AMV wind?

2. Can a statistical correction of systematic height biases be deduced from individual AMV height reassignment values as a general **height bias correction**, and does it improve **forecast skill** when assimilating statistically lidar-corrected and layer-averaged AMVs in an NWP system?

► Structure of this thesis

Section 2 summarizes the basic principles of the different observation types that are used in this thesis. This comprises the derivation of AMV observations and potential error sources during the derivation process and information on lidar observations. In addition, information about the operational global forecasting systems at DWD and the assimilation of AMVs is provided. **Section 3** gives an overview on the data sets and introduces the methodology of the two lidar-based height correction methods developed in this thesis. A brief overview of the wind evaluation methods completes this section. **Sections 4 and 5** present results on lidar-based height correction methods for the European geostationary Meteosat Second Generation (MSG) satellites. In **Section 4**, results from the direct height reassignment for individual AMVs with adjacent lidar observations are shown, followed by a brief analysis of error correlations. **Section 5** comprises results from the height bias correction, including assimilation and forecast experiments with statistically lidar-corrected AMVs. **Section 6** provides a brief outlook on potential future applications, such as an application of lidar-based height correction methods to AMVs from other geostationary satellites. A summary of the findings of this thesis and conclusions that can be drawn are provided in **Section 7.**

Parts of this thesis are contained in two publications that were submitted to the Journal of Applied Meteorology and Climatology (JAMC) during the PhD work. The verification of the direct height reassignment with radiosonde observations (i.e. results shown in Section 4.1) is contained in Folger and Weissmann (2014). Furthermore, the evaluation of the direct height correction with GME model equivalents (results shown in Section 4.2.1) and the height bias correction (comprising the Sections 5.1 and 5.2), including an outlook to the application of lidar-based height correction methods to other geostationary satellites (Section 6.1) are presented in Folger and Weissmann (2016). The corresponding parts about data and methodology used are also partly extracted from these papers, with additional and more detailed information presented in this thesis.

2. BASIC PRINCIPLES

This thesis uses two different types of observations, which will be introduced in this section. First, information on operational AMVs and their derivation process (including tracking and height assignment), as well as the associated error sources is provided. Second, space-borne lidar observations and their use to correct the operational AMV pressure heights are discussed. The basic principle of lidar measurements is introduced and details on space-borne CALIPSO lidar observations are presented. An overview of the global forecasting systems GME and ICON of DWD and their AMV assimilation procedures concludes this section.

2.1 AMV observations

2.1.1 Derivation of AMVs from cloud features

The basic strategy for the AMV derivation consists of four steps: (1) choose a feature that can be traced, (2) track that feature in consecutive satellite images and determine the magnitude and direction of the resulting displacement vector, (3) assign a representative pressure height to the vector, and (4) deduce a quality index for the derived AMV as an estimate on the product quality (Borde et al., 2014a). Although the algorithms used by the AMV producing centres vary slightly for both the tracking and the height assignment process, the basic strategy is similar. A general introduction to the principles of the AMV derivation and potential error sources is given in this section. As this thesis focuses on MSG-AMVs, an overview of MSG satellites as well as their specific AMV derivation scheme will be presented additionally.

2.1.1.1 Tracking of atmospheric motion

AMVs are derived by tracking the displacement of atmospheric structures in sequential satellite images. Typically, cloud structures are used for the tracking process, but water vapour gradients may also be traced (see e.g. Büche et al., 2006). As AMV pressure heights of tracked cloud structures are compared to lidar cloud-top observations in this thesis, AMVs derived from water vapour structures are not discussed further. Figure 3 shows the basic principle of the tracking process for a cloud feature. First, a suitable tracer is detected within a target box (orange), which is then tracked in several consecutive images within a predefined search area (purple) by applying cross-correlation techniques (see e.g. Schmetz et al., 1993). A horizontal wind speed and wind direction can be determined through the displacement of the respective feature. Generally, 3-4 images are used for the tracking process for geostationary satellites in order to allow for consistency checks between the individual wind vectors. The time interval of these images varies from 15-60 minutes for geostationary satellites depending on the respective instrument capabilities. In contrast, AMVs from polar-orbiting satellites generally use only two pictures for the tracking, as the time gap between the single images is about 100 minutes (one Earth orbit) (Menzel, 2001; Forsythe, 2007).

AMVs are derived by using images from different channels. Generally, visible channels (VIS) are used in the lower troposphere below pressure heights of 700 hPa during daylight periods. AMVs from infrared channels (IR) are derived throughout the troposphere, whereas AMVs from water vapour channels (WV) are mainly found in upper levels above 600 hPa (Velden et al., 2005; Velden and Bedka, 2009).

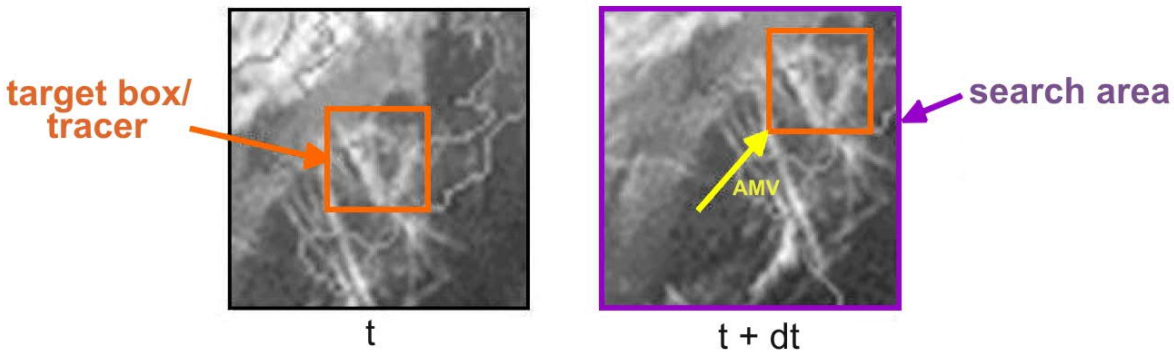

Figure 3*: Illustration of the tracking process of a cloud structure in two consecutive satellite images. The feature detected in the orange target box in the left image is traced in the subsequent right image within a pre-defined search area (purple).Through the displacement of the feature, a horizontal wind speed and direction can be determined (yellow arrow).*
Adapted from Forsythe, 2007.

The optimal configuration for the AMV derivation depends on parameters such as the tracer size and the time difference between the images, which may in turn depend on other parameters such as channel or image resolution (Borde and Garcia-Pereda, 2014b; Velden et al., 2000; Szantai et al., 2000). Garcia-Pereda and Borde (2014) found the lowest AMV errors for time differences of 10 - 15 minutes between the sequential satellite images in combination with tracer boxes of 16x16 or 24x24 pixels. Different processing centres pursue different strategies for the feature tracking and the subsequent height assignment. Inter-comparison studies between different processing centres have been conducted on a regular basis in recent years (Genkova et al., 2010; Santek et al., 2014).

2.1.1.2 AMV height assignment

Generally, the uppermost cloud pixels (e.g. the coldest peak or the coldest 25% of all pixels) within the target box are used to derive a representative height (Genkova et al., 2010). There are several AMV height assignment methods that are employed operationally for cloudy targets. A short introduction to the different concepts is provided in the following. A more detailed overview on height assignment methods can be found e.g. in Nieman et al. (1993, 1997).

► Brightness temperature conversion

The Equivalent Black Body Temperature (EBBT) method determines a representative pressure height by using temperature profiles of a model forecast and the brightness temperature of the respective cloud pixels of the satellite image. The brightness temperature is defined as the temperature an emitter (e.g. a 'grey' cloud structure detected by a satellite imager) would have if it was a black body (see e.g. Besançon, 1990). The AMV is then assigned to the model level that best fits the observed brightness temperature of the satellite image. This method is mainly used for low-level opaque clouds from channels in the infrared and visible range (Schmetz et al., 1993).

► Multi-channel techniques

Two multi-channel techniques are widely used at AMV processing centres to better account for the semi-transparency of clouds: the CO_2-slicing and the H_2O-intercept method. Both methods use the differences between two channels for the same satellite scene to derive an

AMV pressure height. The CO_2-slicing method uses the 13.9 micron band in combination with an IR image (Smith and Platt, 1979; Menzel et al., 1983). Water vapour sensitive channels provide an alterative approach to account for semi-transparency issues for satellites with no CO_2-channel available. The H_2O-intercept is conceptually similar to the CO_2-slicing, but uses a combination of one WV-image and one IR-image. (Szejwach, 1982; Schmetz et al., 1993). Multi-channel techniques are based on the following equation:

$$\underbrace{\frac{R_{CO_2/WV} - R^{cs}_{CO_2/WV}}{R_{IR} - R^{cs}_{IR}}}_{\text{observed}} = \underbrace{\frac{n\varepsilon_{CO_2/WV}[R^{bcd}_{CO_2/WV}(P) - R^{cs}_{CO_2/WV}]}{n\varepsilon_{IR}[R^{bcd}_{IR}(P) - R^{cs}_{IR}]}}_{\text{calculated}}$$

The left hand side represents the observed satellite radiances in the respective CO_2 or WV (numerator) and IR (denominator) channel. The difference between the radiances of the detected cloud structure and a clear-sky region (representing the surface radiation) is determined for both channels. The clear sky radiance R^{cs} is derived either from clear sky pixels adjacent to the cloud feature or from temperature profiles of a model short-term forecast. The right hand side includes the differences between a calculated "black" cloud (*bcd*) and a clear sky radiance and is computed for different pressure heights P. The one that best matches the observed radiance from the left hand side is selected (Nieman et al., 1993; Forsythe, 2007). The cloud fraction n is the same for both channels and can be cancelled in the fraction. To apply the equation to the CO_2-channel, it is assumed that the emissivity ε is roughly the same for ice clouds in the IR- and CO_2-channel and therefore can be eliminated in the fraction as well (Nieman et al., 1993). The H_2O-intercept method is based on the assumption of a constant ratio of the two cloud emissivities $\varepsilon(\mathrm{WV})$ and $\varepsilon(\mathrm{IR})$ (Schmetz et al., 1993; Nieman et al., 1997).

► Additional corrections for low-level AMVs

As an additional approach to improve the AMV height assignment, low-level AMVs are often assigned to the cloud base rather than the cloud top of the observed cloud feature. This strategy is based on a study of Hasler et al. (1979), who found a better agreement of AMV winds with wind observations at the cloud-base level for marine cumuli during a field

campaign. The assessment of the cloud base from satellite radiances is done by analysing cloudy and clear-sky pixels around the respective cloud structure. A brightness temperature histogram then enables the cloud-base height to be estimated (Le Marshall et al., 1994).

Additionally, some centres apply an inversion correction to low level AMVs. If a temperature inversion exists in the corresponding model profile of an AMV, this AMV is always relocated to the height of the minimum temperature, assuming that the cloud is not able to penetrate in stable inversion regions. Thereby, AMVs that are originally assigned too high in the atmosphere due to an under-representation of the inversion depth may be shifted to a more representative height (Forsythe, 2007).

2.1.1.3 AMV error sources

Different error sources may add to the total AMV error and can lead to wind vectors of poor quality. The main contribution arises from errors resulting from the height assignment process. Velden and Bedka (2009) estimated that about 70% of the total AMV error is due to height assignment issues. However, other error sources may contribute as well. The main error sources are discussed in the following.

► Height assignment errors

All height assignment techniques can potentially induce large errors. The EBBT-method can only provide reliable height estimates for opaque clouds. Uncertainties may arise for semi-transparent clouds when contributions from levels below the actual cloud height increase the brightness temperature, therefore leading to a height assignment that is too low in the atmosphere. (Schmetz et al., 1993). Although multi-channel techniques achieve superior results to the EBBT-method for semi-transparent clouds, the height assignment of upper-level AMVs constitutes a major issue. Both methods (CO_2-slicing and H_2O-intercept) have large limitations, especially for small differences between the two channels, which cannot be distinguished from atmospheric noise. In addition, the dependence on the (unknown) cloud microphysics and the amount of water vapour is an issue. Both methods perform poorly for thin and multi-layer clouds, often placing the AMV in between the two upper cloud layers (Borde and Dubuission, 2010). Generally, an optimal combination of techniques has not yet been found that tackles these problems satisfactorily (Borde, 2014c). Comparisons between

different multi-channel techniques have been conducted to assess their performance for various cloud situations, with differing results as to which technique outperforms the other (see e.g. Daniels et al., 2006; Borde and Dubuisson 2010; Schreiner et al., 2012).

In addition to the inherent error sources of the single height assignment methods, there are further issues that contribute to the height assignment error. In particular, the selection of the pixels that is used for the height assignment leads to uncertainties. Generally, tracked cloud features may contain pixels from different altitude levels. The more complex the detected cloud structure is, the more difficult is the assignment to a specific height (Forsythe, 2007). Consequently, the pixels that are actually tracked are not necessarily the same as those used for the height assignment afterwards. Additionally, features tracked in the visible channel cannot be converted directly to pressure levels, but corresponding IR images must be used. This can lead to large height assignment errors especially in multi-layer situations, when the brightest pixels of the target box in the respective IR image do not represent the tracked feature of the VIS image. (Borde et al., 2014a)

Generally, the height assignment is particularly error-prone when wind varies strongly with height (Salonen and Bormann, 2014). A pronounced “slow” speed bias is frequently found when comparing high-level and mid-level AMVs to in-situ wind observations. The main reason is an incorrect height assignment locating the AMV derived from a lower cloud feature with a slow wind speed too high in the atmosphere. However, a second contributing factor is the tracer size used in the tracking step, when the detected wind speed is averaged over the pixels of the target box. Small target box sizes imply a reduced shifting of the detected feature, and may introduce large errors due to sub-pixel displacements. However, larger target boxes lead to an enlarged averaging, thereby increasing the slow speed bias (Bresky et al., 2012).

► Representativity errors

The traditional approach of deriving AMVs is to assign them to a discrete altitude level. However, it was already suspected in the early stages of AMV derivation that representativity errors may arise when AMVs are interpreted as single-level observations instead of layer-averages over a vertically extended cloud layer (Schmetz et al., 1993). As already pointed out in Section 1, recent studies that compared AMV layer-winds to radiosonde/dropsonde observations indicate that AMVs represent winds in vertical layers rather than winds at

discrete levels (Velden and Bedka, 2009; Weissmann et al., 2013). Similar results are obtained by analysing AMV error characteristics within a simulated framework (Hernandez-Carrascal and Bormann, 2014; Lean et al., 2015), where AMVs are derived from simulated satellite images. This approach has the advantage of knowing the "true" atmospheric model state. However, care must be taken when interpreting the results, as model clouds can represent real clouds only to a limited extent and do not necessarily provide a realistic representation of cloud structures if the model resolution differs largely from scales resolved by the satellite imagery. Further challenges are often introduced by systematic errors in the model representation of clouds. Tracking processes of model-cloud tracers are therefore challenging and potentially error-prone (Lean et al., 2015).

Overall, it is commonly accepted that interpreting AMVs as layer-averages may yield benefits for data assimilation. However, AMVs are currently still assimilated as single level observations at all NWP centres.

► Additional error sources

In addition to height assignment and representativity errors, other error sources may introduce further uncertainties. As cloud structures are tracked over sequential images that are 15 - 100 minutes apart from each other, the detected cloud may evolve during that time period and the cloud-top height of the tracked feature may vary in subsequent images (Menzel, 2001). In addition, the detected wind speed of the cloud top constitutes a temporal and spatial average of the actual tropospheric motion and is not necessarily an optimal representation of the actual atmospheric wind field (Schmetz et al., 1993). Strictly speaking, AMVs can only represent an unbiased estimate of the atmospheric wind if clouds are randomly distributed and floating with the airflow (Schmetz and Nuret, 1989).

Furthermore, short-range forecast model profiles of temperature and humidity are used for the AMV height assignment, introducing additional errors due to an imperfect representation of the atmospheric conditions by the model (Schmetz et al., 1993). This is a particularly dangerous issue as the resulting AMVs then may contain errors that are correlated with the model short-term forecasts that are used commonly in data assimilation as first-guess (FG) fields.

As the sequential images used for the tracking process must coincide, errors may also arise for spinning radiometers such as the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) aboard the MSG satellites. For SEVIRI with a rotation rate of about 100 revolutions min^{-1}, the

Root Mean Square Error (RMS) for the image-to-image relative accuracy is 1.2 km (Schmetz et al., 2002), which can have a considerable effect especially on the accuracy of low-level AMVs with small wind speed values.

2.1.2 Meteosat Second Generation

The MSG satellites with a nominal life expectancy of 7 years each are the European series of geostationary satellites currently in operation. The MSG mission is a joint cooperation between the European Space Agency (ESA) and EUMETSAT and provides high-resolution imagery and improved sensor-technology compared to the preceding Meteosat First Generation (MFG) satellites. The first MSG satellite Meteosat-8 was brought into orbit in 2002, and was followed by three successful launches of further MSG satellites. The major instrument on board is SEVIRI, which provides full-disk imagery every 15 minutes in 12 channels ranging from the visible (0.6 μm) to the infrared spectrum (14 μm). SEVIRI provides images with a pixel size of 3 km at nadir for all channels, and 1 km for the high-resolution visible channel. As the angular stepping remains constant, the pixel resolution decreases for off-nadir views, i.e. the resolution decreases with increasing latitude and with increasing longitudinal distance from the satellite position (Schmetz et al., 2002). SEVIRI enhances the derivation of AMVs significantly compared to the precursor generation of MFG satellites both in quality and quantity as there are two water vapour channels and one CO_2-channel, therefore improving the error-prone height assignment for semi-transparent clouds (Borde, 2014c).

► Tracking of MSG-AMVs

For the derivation of MSG-AMVs, cloud and water vapour features are tracked from imagery of five MSG channels: The VIS channel at 0.8 μm, the IR channel at 10.8 μm, the two WV channels at 6.2 μm and 7.3 μm and the high-resolution VIS channel (broadband, about 0.4 - 1.1 μm) (Schmetz et al., 2002; Carranza et al., 2014). Other channels may be included in the future (EUMETSAT, 2011). The tracking algorithm for MSG-AMVs uses four consecutive images with 15 minutes in between each image pair. The reference time of the derived AMV is set to the second picture (Régis Borde, personal communication). The derivation process starts by targeting a suitable feature in a search box of 24x24 pixels. The cloud feature with the strongest contrast (with a minimum number of high-contrast pixels) is

then selected. Scene analysis algorithms classify each pixel as cloudy or clear sky, hence creating a cloud mask for each image. If a minimum amount of 50 cloudy pixels is not reached for a specific scene, a search for a clear-sky AMV by tracking water vapour gradients in the WV-channels is attempted. The selected feature is then tracked in the consecutive images in search boxes of 80x80 pixel size centered around the respective tracking element (Lutz, 1999; Schmetz et al., 2002; Borde et al., 2014a).

► Height assignment of MSG-AMVs

Cloud-top heights for MSG-AMVs from opaque clouds are derived by using the EBBT-method (prioritizing the 10.8 μm channel), including a low level inversion correction if necessary. Short-range forecast model profiles from ECMWF are used for the brightness temperature conversion. Several combinations of WV/IR channels and CO_2/IR channels are available for the height assignment for upper level semi-transparent clouds (prioritizing the CO_2-slicing). Clear sky radiances are extracted from clear sky pixels neighbouring the selected feature (EUMETSAT, 2011; Borde, 2014c).

The operational AMV height assignment strategy of EUMETSAT changed on 5 September 2012 to the Cross-Correlation Contribution (CCC) method. This method provides a more consistent height assignment as the pixels that contribute most to the tracking process are used to set the AMV height (Borde et al. 2014a). MSG-AMVs from different periods that comprise data from before and after the changeover are evaluated in this thesis. Before September 2012, the determination of a representative AMV height for MSG-AMVs was conducted for the coldest peak in the target box. In addition, a cloud-base height assignment for low-level MSG-AMVs was implemented. With the changeover to the CCC method in September 2012, the final pressure of the cloud feature is a weighted average of the pressure values assigned to the individual cloudy pixels that represent the dominant motion feature within the target box. In addition, AMVs are no longer assigned to the cloud base at low levels (EUMETSAT, 2011; Borde et al., 2014a). Unfortunately, the resulting operational Meteosat-AMV dataset contains no information about the height assignment method applied for deriving individual AMVs in the final product.

The performance of the new algorithm generally exceeds the old version for high-level and mid-level AMVs, exhibiting a higher average quality. An initial degradation of data quality in low levels (Salonen and Bormann, 2012) was improved by applying a patch introduced in

January 2013 (Carranza et al., 2014). The CCC-method is also tested by other AMV processing centres and achieves superior results compared to the traditional algorithms (e.g. Kim et al., 2015).

► AMV quality assessment

In order to rate the quality and representativeness of AMVs, a Quality Index (QI) is disseminated from EUMETSAT together with each AMV, enabling the forecasting centres to filter out wind vectors with questionable reliability. The automatic control scheme operational at EUMETSAT comprises several checks. First, a comparison between the three preliminary wind vectors derived from a consecutive image quartet ensures the temporal consistency between the single vectors. In addition, a spatial consistency check is applied by comparing the speed, direction and pressure of the final AMV to those of neighbouring AMVs in a similar pressure range. Furthermore, a forecast check against the FG field (short-range forecast) of the ECMWF-model, interpolated to the AMV location and pressure level, is applied. Other checks, such as an inter-channel consistency check, are not yet included, but considered as a future enhancement.

These quality checks are supposed to reject poor quality wind vectors with large deviations in wind speed and direction, but still allow for natural accelerations that may appear e.g. in the jet entrance-exit regions. Problems may arise if strict conformity rules are applied for the forecast check, when the AMV wind has to coincide closely to the FG field and therefore the model field is simply reproduced. Therefore, AMVs are not discarded for poor forecast consistency if temporal and spatial consistency checks achieve good results, but a lower quality is assigned. The information gained from all quality checks are combined to assign one QI for each AMV. The derived QI has values ranging from 0 to 100, with 100 indicating the best possible quality (Holmlund, 1998; EUMETSAT, 2011).

2.2 Lidar observations of cloud-top heights

Light detection and ranging (lidar) has been an important data source for profiling measurements for decades. In contrast to passive remote sensing techniques, a lidar actively emits photons of certain wavelengths and is not dependent on radiation from natural sources. This remote sensing technique enables a detailed analysis of cloud and aerosol layers, and the

determination of a variety of atmospheric parameters such as temperature, humidity, wind or trace gases (Weitkamp, 2005).

The basic principle of a backscatter lidar is described in the first subchapter of this section and is based on Weitkamp (2005). Subsequently, the space-borne lidar CALIOP on board the polar-orbiting satellite CALIPSO is introduced, including potential error sources of CALIPSO lidar observations. These chapters contain data content from the CALIOP Algorithm Theoretical Basis Documents about the CALIOP instrument and algorithm overview (Winker et al., 2006), scene classification algorithms (Liu et al., 2005) and feature detection and layer properties algorithms (Vaughan et al., 2005). In addition, information from Winker et al. (2009, 2010), Hunt et al. (2009) and Vaughan et al. (2009) is used.

2.2.1 Principles of a backscatter lidar

A lidar essentially consists of an emitting and a receiving device for radiation and can be applied from ground-based, air-borne or space-borne platforms. The basic principle of an air- or space-borne lidar is illustrated in Figure 4. Photons of a certain wavelength are emitted by short laser pulses in the nanosecond range and interact with particles in the atmosphere (air molecules, aerosols, cloud droplets) through scattering and absorption processes. A small fraction of the photons are scattered 180° backwards and can be detected by the lidar

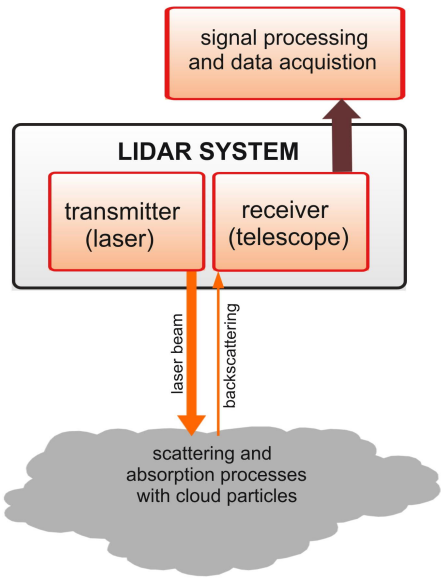

Figure 4: *Illustration of the basic components of a lidar system. Schematic adapted from Fischer, 2013*

telescope. When the wavelength remains unchanged during a scattering process, it is called elastic backscattering. By measuring the time difference between the emitted and received signals (travelling with the speed of light, ~$3 \cdot 10^8$ ms^{-1}), the position of the scattering process can be determined to a high accuracy. When the laser pulse hits a region of increased particle density or with larger particles, the backscattering signal is enhanced. Thus, lidar observations are highly sensitive to atmospheric layers of a higher optical depth such as cloud or aerosol layers. The backscattered lidar signal is described by the following 'lidar equation':

$$\underbrace{P(r,\lambda) = \underbrace{K \cdot \frac{1}{r^2}}_{(1)\ (2)} \cdot \underbrace{\beta(r,\lambda)}_{(3)} \cdot \underbrace{\exp\{-2 \cdot \int_0^r \alpha(r',\lambda) \cdot dr'\}}_{(4)}}$$

$P(r,\lambda)$ describes the signal strength for a certain wavelength λ as a function of distance *r*. The first term *K* is described by the following equation:

$$K = \frac{P_0 c \tau A \eta}{2}$$

K is a constant and contains the emitted pulse energy P_0, the speed of light *c*, the area of the receiver *A*, the overall system efficiency η, the temporal pulse length τ, and a factor ½ in order to account for the forward- and backward travelling of the laser pulse. As the signal strength decreases with larger distances, a correcting term is applied by dividing by r^2 (second term of the lidar equation). The backscatter coefficient $\beta(r,\lambda)$ represents the scatter probability at an angle of 180° to the receiver (third term). It consists of a fraction of air molecules and a fraction of atmospheric particles (aerosols, cloud droplets etc.).

$$\beta(r,\lambda) = \beta_{mol}(r,\lambda) + \beta_{aer}(r,\lambda)$$

The fourth term of the lidar equation represents the atmospheric transmission at the emitted wavelength λ and can take values from 0 to 1 by integrating over the extinction coefficient $\alpha(r,\lambda)$ from the emission source out to distance *r*. Analogously to $\beta(r,\lambda)$, $\alpha(r,\lambda)$ has contributions from air molecules and particulate matter. The backscatter and extinction coefficients for molecular backscattering can be assessed with relatively high confidence by using model profiles of the air molecules' density. However, the lidar equation is underdetermined, as both the particle backscatter and particle extinction coefficient are unknown. They are encapsulated in one variable called the lidar ratio S_{aer}:

$$S_{aer}(r) = \frac{\alpha_{aer}(r)}{\beta_{aer}(r)}$$

The lidar ratio has to be estimated to solve the lidar equation for the extinction coefficient $\alpha_{aer}(r)$. However, its assessment is often difficult, as it depends on several parameters such as wavelength, humidity and type and shape of the scattering particles.

2.2.2 Space-borne lidar observations from CALIPSO

The polar-orbiting satellite CALIPSO was launched in 2006 as a joint satellite mission of the National Aeronautics and Space Administration (NASA) and the French space agency Centre National d'Études Spatiales (CNES). CALIPSO flies at an inclination of 98.28° in a sun-synchronous orbit at 705 km altitude, encircling the earth in about 100 minutes. CALIPSO is part of the A-Train, which is a constellation of several international science satellites that fly in formation and therefore facilitate a wide variety of different observations of the same scenery from space (see Figure 5). The nominal life expectancy of CALIPSO was originally set to three years (until 2009), and is thus already exceeded by far, with CALIPSO still delivering reliable data from space.

► The lidar CALIOP

The lidar CALIOP is the major instrument on board and measures vertical profiles of the

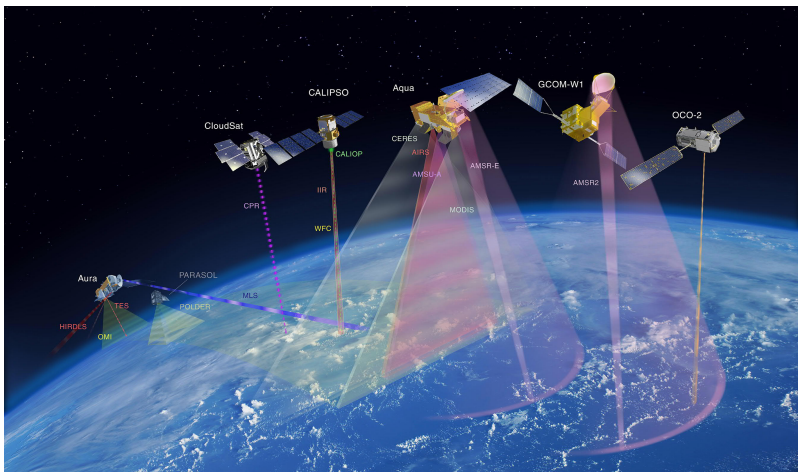

Figure 5: *The A-Train constellation: The leading satellite OCO-2 (Orbiting Carbon Observatory-2), followed by GCOM-W1 (Global Change Observation Mission W1), Aqua, CALIPSO, CloudSat and Aura, source: http://atrain.nasa.gov/)).*

atmospheric elastic backscatter at two wavelengths (532 nm and 1064 nm). CALIOP uses a Nd:YAG (Neodymium-doped Yttrium Aluminium Garnet) laser as transmitter, which has an original wavelength of 1064 nm that is then divided by a dichroic beam splitter. As particle diameters of aerosol or cloud droplets typically have similar dimensions to these wavelengths, this laser is well suited for remote sensing applications of the atmosphere. Additional measurements of the depolarization at 532 nm allow determining the cloud phase (ice/water).

CALIOP emits pulses of 20 nm of 110 mJ with a repetition rate of 20.16 Hz. At the Earth's surface, the beam diameter is 70 m. The continuous operation of CALIOP allows for lidar observations during the day and night. CALIOP points 3° from nadir in the forward along-track direction to avoid specular returns from still water surfaces or horizontally oriented cloud ice crystals. The full sampling resolution of the lidar is 30 m in the vertical and 335 m in the horizontal. As the atmosphere becomes spatially more uniform at higher altitudes and additionally, the thinner atmosphere in these regions requires a broader averaging for signal detection, the lidar profiles are averaged horizontally and vertically above an altitude level of 8.2 km. The horizontal and vertical resolution for the different altitude ranges are listed in Table 1.

► Data processing and feature detection

The uncalibrated, absolute backscatter intensity is processed using various algorithms. In

Altitude range [km]		lidar shots averaged	horizontal resolution [km]	vertical resolution [m]
top	base			
40.0	30.1	15	5.0	300
30.1	20.2	5	1.67	180
20.2	8.2	3	1.0	60
8.2	-0.5	1	0.33	30
-0.5	-2.0	1	0.33	300

Table 1*: Altitude-dependent horizontal and vertical resolution for CALIOP. Grey shading indicates the altitude ranges that are considered in this study. Adapted from Hunt et al., 2009.*

order to find aerosol or cloud features in the lidar profiles, the Selective Iterative Boundary Locator (SIBYL) is applied to the 532 nm attenuated backscatter profiles. A feature top is obtained by applying an altitude- and situation-dependent threshold in order to detect targets against the noisy background. To do this, either a minimum vertical feature distance with slightly enhanced backscatter or a very strong spike signal in a thin layer is required. If a feature cannot be detected at full resolution, an averaging of lidar profiles with increasing horizontal range is applied (5 km, 20 km and 80 km). This sophisticated multi-resolution averaging scheme with a range-varying detection threshold automatically adapting to the background conditions has proven to provide reliable information on layer-top heights from cloud features or aerosol layers. Afterwards, several scene classification algorithms (SCA) classify the retrieved layers as aerosol or cloud. In this context, the Cloud-Aerosol-Distinguisher (CAD) is defined as a quality index that indicates the reliability of the retrieved lidar information, ranging from +100 (cloud observation) to -100 (aerosol observation). The respective “clear air” profiles are derived from model data from the Goddard Earth Observing System Model, version 5 (GEOS-5) analysis product from NASA’s Global Modelling and Assimilation Office (GMAO).

2.2.3 Error sources of CALIPSO lidar observations

There are a number of error sources that can deteriorate the final lidar product. To analyse the pure lidar signal, undesired contributions from background light or detector noise have to be assessed carefully in order to be subtracted from the detected signal. Generally, space-borne lidars have a low signal-to-noise ratio due to the large distances between the location of the scattering process in the troposphere and the detector. Several sources of noise impair the actual lidar signal: Detector dark current and amplifier noise as well as statistical fluctuations of the background (especially during daylight when the sunlight deteriorates the signal) and of the lidar return signal itself may all contribute to the total noise rate. Weak signals as from subvisible high-level cirrus clouds can have the same order of magnitude as the noise and are sometimes difficult to detect. Averaging over a number of consecutive pulses helps to enhance the signal compared to background noise patterns, which is often necessary for tenuous aerosol layers. However, the spatial structure of these patterns may then be lost.

Some additional error sources also contribute to the total error of the lidar observation. For example, the GMAO-model used for the calculation of the molecular backscatter as well as

other external data (such as the accurate elevation above sea level) have limitations in representing the true atmosphere. Moreover, the geolocation has to be assessed carefully, as the lidar footprint moves at about 7 km/s across the ground. In addition, the background measurement of the solar background and the determination of the system efficiency may contain errors. Furthermore, the CALIOP laser beam is substantially attenuated for column optical depth values larger than ~3, so that multi-layer situations can be analysed only in a limited way.

Generally, the estimation of the particle lidar ratio S_{aer} represents a major error source for lidar retrievals of the extinction coefficient, as it is often complicated to assess for particular lidar scenes. Values vary strongly for different particles and incoming wavelengths and may therefore contribute to biased retrievals. However, the error-prone determination of S_{aer} is not necessary for the application of retrieving cloud-top heights. Thus, the major error source for the usage of lidar cloud-top information in this study is the vertical ranging error. Sophisticated detection algorithms and advanced sensor technologies ensure that the ranging error is estimated to be less than 30 m in the vertical. Therefore, CALIOP provides highly reliable information on cloud-top heights for the height correction of AMVs.

2.3 Assimilation of AMVs in the global models at DWD

2.3.1 DWD models (GME/ICON)

In this thesis, the two global models GME and ICON from DWD are used. GME was operational from 1 December 1999 until 20 January 2015 and was the first operational NWP model to be based on an icosahedral–hexagonal grid. The advantage of this grid construction is the relatively uniform grid size compared to latitude-longitude grids, where grid boxes become smaller towards the poles. The horizontal grid spacing is 20 km, with 60 vertical levels up to an altitude of 36 km. The system of equations is based on a hydrostatic assumption. Forecasts up to 174 hours are computed with 00 UTC and 12 UTC as initial times. Additionally, forecasts up to 48 hours are provided with 06 and 18 UTC as initial times. (DWD, 2015a; Majewski et al., 2002).

The global model ICON replaced GME as the operational global model in January 2015. ICON surpasses the capabilities of GME in several respects. The major difference is the

dynamical core using the non-hydrostatic equations on the global domain. In addition, new physical parameterization schemes allow for a better representation of various atmospheric processes such as convection, radiation or cloud cover. Furthermore, ICON offers more prognostic variables and provides e.g. better mass conservation properties. As ICON is comparatively new, occasional small error corrections and model adjustments have been needed since becoming operational (see e.g. DWD, 2015b). Generally, forecast quality has proven to improve significantly and the model bias (e.g. for the wind speed and wind direction) is reduced compared to GME (Harald Anlauf, personal communication). A comparison of some technical details for both models is provided in Table 2 (DWD, 2015a; Reinert et al., 2015; Zängl et al., 2015; Baldauf et al., 2015).

► Data assimilation system

Detailed information on data assimilation techniques can be found in Kalnay (2003), which serves also as basis for the following brief introduction. Generally, data assimilation techniques aim to combine a model background field (FG) with observational data in an optimal way. Typically, this background field is a short-term forecast initialized at the previous assimilation time step, which already contains information from earlier observations from the assimilation during preceding time steps ("cycling"). The observational errors and the errors of the FG field determine their respective weightings in the resulting analysis: a high observation error leads to a low weight for the respective observation, so that the

	ICON	**GME**
system of equations	non-hydrostatic	hydrostatic
mesh size [km]	13 km	20 km
mean grid area	173 km^2	346 km^2
vertical coordinates	height	pressure
vertical layers	90	60
model top margin	75 km	36 km

Table 2: *Comparison of ICON and GME for some technical aspects. Adapted from DWD, 2015.*

FG field dominates the analysis and vice versa. The analysis field is defined for every model grid point and provides then the initial conditions from which a forecast can be started. Numerous data assimilation strategies have been developed during recent decades. The decision of NWP centres over which data assimilation system to use is always a trade-off between computer capacities and the accuracy of the initial state and the subsequent forecast.

The data assimilation system at DWD (both for the former GME and the current ICON model) is based on the three-dimensional variational technique (3DVAR) with a 3-h cycling (00, 03, …, 18, 21 UTC). During one 3DVAR data assimilation cycle, all observations within a 3-h window (+/- 90 min from the corresponding time step) are combined with the model background field, assuming implicitly that these observations take place at exactly the same time. The assimilation window is usually equal to the time between consecutive analyses/forecast cycles. The 3DVAR algorithm aims to minimize a cost function containing the background field, the observations and their error covariance matrices to create an optimal analysis state for subsequent forecasts. From the resulting analysis, a 3-h short-term forecast is initialized, which serves again as the background field for the next assimilation cycle (Frank et al., 2014; Reinert et al., 2015).

2.3.2 Operational AMVs at DWD

AMVs from geostationary and polar-orbiting satellites are assimilated routinely in global NWP models. DWD operationally uses AMVs from five different geostationary satellites and seven polar-orbiting satellites.

The main European geostationary satellite is Meteosat-10, which is located at 0° longitude and covers Europe, Africa and large parts of the Atlantic Ocean. In addition, AMVs from the following geostationary satellites are assimilated routinely: Meteosat-7 at 57°E, the Multi-Functional Transport Satellite 2 (MTSAT-2) at 145°E and the two Geostationary Operational Environmental Satellites (GOES) at 135°W (GOES-West) and 75°W (GOES-East). Meteosat-10 belongs to the Meteosat Second Generation with 12 channels in total in the visible and infra-red range. Meteosat-7 (MFG) is less sophisticated with only three channels. Thus, fewer channels for the feature tracking and the subsequent height assignment are available and e.g. do not allow for CO_2-slicing due to the absence of the CO_2-channel. In addition, the interval between MFG images is 30 min compared to the 15 min interval for MSG imagery. Consequently, considerably fewer AMVs are available. Meteosat-AMVs are

derived operationally by EUMETSAT. GOES satellites have 6 channels each and GOES-AMVs are provided by the National Environmental Satellite Data and Information Service (NESDIS) of the National Oceanic and Atmospheric Administration (NOAA). MTSAT-2 was the operational Japanese geostationary satellite with 5 channels during the period evaluated in this study; MTSAT-2-AMVs were provided by the Japan Meteorological Agency (JMA).

In addition to geostationary satellites, AMVs from polar-orbiting satellites are used in data assimilation in order to extract information about the wind field in polar regions. AMVs from the satellites of the NOAA-series, from Aqua and Terra and from the METOP-satellites have been assimilated routinely for several years. Table 3 lists all satellites that are used for the derivation of AMVs that are assimilated operationally at DWD (Effective: October 2015).

2.3.3 AMV error correlation and thinning

AMVs exhibit significant spatially and temporally correlated errors for several reasons. Height assignment procedures use model temperature profiles that are already mutually correlated. In addition, AMVs that tend to be dissimilar to their neighbours are rejected during quality control processes, thereby enhancing existing error correlations. AMV errors can exhibit significant spatial correlations with horizontal lengths of up to 800 km (Bormann et al., 2003). However, observation errors are usually assumed to be spatially uncorrelated in

geostationary satellites		**polar-orbiting satellites**
ID	position	ID
Meteosat-10	0°	AQUA
Meteosat-7	57°E	TERRA
MTSAT-2	145°E	NOAA-15
GOES-13	75°W	NOAA-18
GOES-15	135°W	NOAA-19
		METOP-A
		METOP-B

Table 3: *Operational and monitored satellites used for the AMV derivation and operational assimilation at DWD. Source: Alexander Cress, DWD. Effective 25 Oct 2015.*

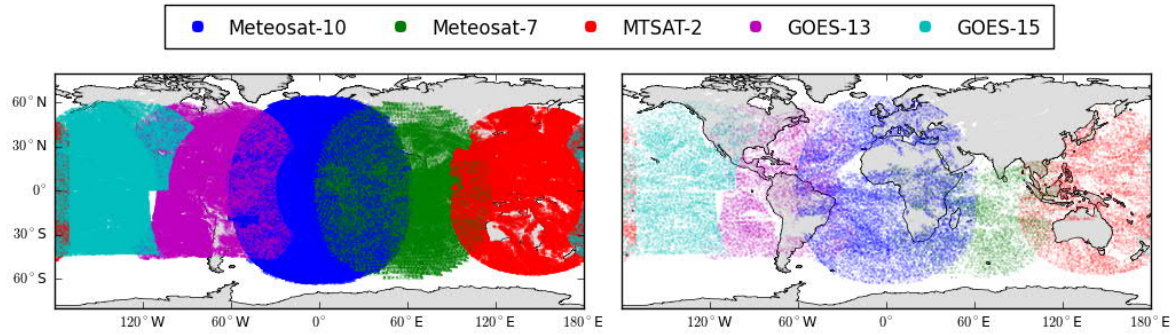

Figure 6: *Distribution of all available AMVs from geostationary satellites for the 10 th May 2013 before (left) and after (right) quality procedures are applied.*

data assimilation systems to save computational cost. Most data assimilation centres apply a spatial thinning of about 200 km in addition to an inflation of the corresponding observation errors as a trade-off between the density of AMV wind information and correlation magnitude (Forsythe, 2007). Moreover, the precise values for observation errors that can be assigned to AMVs are not known, resulting in a large fraction of available AMVs being used in a suboptimal way (Velden and Bedka, 2009).

The quality control and thinning strategy at DWD is made up of several aspects. Temporal and spatial blacklisting thins the data drastically: AMVs are assimilated in a 6-h interval (00, 06, 12, and 18 UTC) in the global model at DWD, thereby discarding half of the available AMV data set from 03, 09, 15 and 18 UTC. The minimum horizontal distance between adjacent AMVs is set to 240 km each. In each thinning box, the AMV with the highest QI is selected. This QI has to exceed a certain threshold that depends on AMV altitude, satellite, instrument channel and latitude region. AMVs from some geographic regions (e.g. the Himalayan Mountains) or channels (e. g. VIS-AMVs over land surfaces) are always denied, since such AMVs often show negative long-term monitoring statistics. An additional FG check rejects all AMV winds deviating strongly from the model background.

Overall, this thinning strategy leads to the rejection of a large part of the originally dense AMV data set. Figure 6 shows the distribution of all operationally derived AMVs on the left and the AMVs actually used (after thinning and quality control procedures were applied) on the right for 10 May 2013 in the global model GME for the five main geostationary satellites. Depending on the satellite, only 1% - 4% of the available AMV data set is used for data assimilation. For example, about 17000 AMVs of the original number of about 1086000 AMVs from Meteosat-10 (blue) are assimilated on the depicted day.

3. DATA AND METHODOLOGY

3.1 Data sets and evaluation methods

3.1.1 Observational data

This study mainly focuses on the height correction of **AMVs** derived from images from the European geostationary MSG satellites (Meteosat-9 and Meteosat-10). Details on the derivation of MSG-AMVs can be found in Section 2.1.2. The MSG satellites are located at 0° longitude and provide frequent satellite imagery for the operational weather service. Meteosat-9 was the operational European satellite until January 2013, and was then succeeded by Meteosat-10 (Salonen and Bormann, 2014). EUMETSAT is in charge of the derivation and quality control of MSG-AMVs and provides the final AMV data product for 1-h intervals to the end-users, disseminating over 10^6 MSG-AMVs per day. The EUMETSAT AMV product includes information about the geographical position, the pressure height, u- and v-components of the wind, channel and QI. In addition to a detailed evaluation of lidar-based height correction methods for MSG-AMVs (Sections 4 and 5), a brief assessment of results for AMVs from other geostationary satellites (GOES and MTSAT-2) is provided in Section 6.1 as an outlook for potential future applications.

For the AMV height correction, space-borne **lidar observations** from CALIPSO are used. As described in Section 2.2.2, the official CALIPSO Level-2 cloud layer product provides cloud-top heights with a horizontal sampling resolution of 335 m (for the lowest 8.2 km of the troposphere) and 1 km (from 8.2 km to 20.2 km altitude). In order to have a consistent data set for all altitude levels, the product with a horizontal resolution of 1 km is used in this thesis. Averaging schemes for a coarser horizontal resolution are also available in the official data product, with resolutions ranging from 5 km to 80 km. These broad averaging schemes are mainly used to detect thin cirrus or faint aerosol layers that cannot be distinguished from atmospheric noise patterns under higher resolution. Given that CALIPSO lidar observations

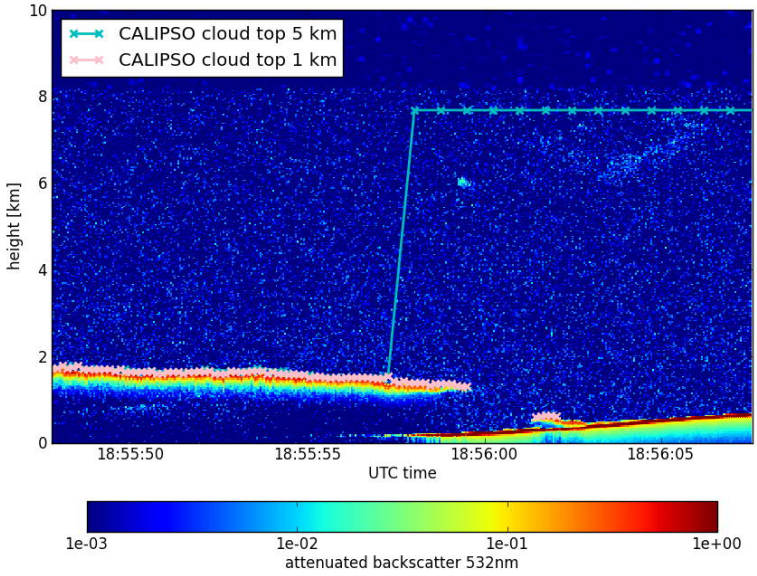

Figure 7: *Example case with two CALIPSO cloud-layer products with 1-km resolution (rose) and 5-km resolution (cyan) cloud-top height. A strong cloud signal (yellow-reddish colours) at about 1.5 km - 2 km altitude is characterized as cloud at both resolutions. A faint signal at about 7 km height is classified as cloud at 5-km resolution.*

are being compared to MSG imagery in this study, the detection of very thin cirrus by the lidar is not required, as satellite imaging instruments generally do not have the ability to detect these structures. Sensitivity studies have often shown a more reliable cloud-top detection for the 1-km product compared to the 5-km-averaged cloud top for the purpose of comparison studies to AMV heights. As an example, Figure 7 shows a strong backscatter signal from a low-level cloud at 1.5 km, which is detected for both 1-km and 5-km resolution. However, a faint signal at about 7 km altitude is observed by the 5-km averaging scheme. The 5-km cloud-top is thereby set to a constant height level at about 7.7 km due to the broader averaging that is able to detect also tenuous layers that cannot be directly observed in the high-resolution data that is depicted in Fig. 7. However, these layers generally do not represent the cloud-top observations of (optically thicker) clouds that are needed in this study for the comparison with AMVs derived from imaging instruments.

In summary, the 1-km cloud-layer product is chosen for the AMV height correction presented in the following chapters, as it provides reliable and high-resolution cloud-top heights, which will be averaged for the AMV height correction (see subsequent Section 3.1.2). The cloud layer product provides the 1-km horizontally averaged cloud-top height from the CALIOP lidar, and in addition the number of superimposed cloud layers, the cloud phase and a quality index for clouds.

3.1.2 Collocation of AMVs and CALIPSO lidar observations

In order to find suitable CALIPSO lidar cloud-top observations that are close to AMVs, different collocation criteria are applied. These generally follow Weissmann et al. (2013), but use stricter horizontal and temporal distance requirements as well as additional criteria for the selection of collocated CALIPSO lidar observations. In this study, AMVs are corrected with nearby CALIPSO lidar observations that are within 50 km horizontal distance and 30 min time difference from the location and time of each AMV. This principle is schematically illustrated in Figure 8. The median of all lidar cloud-top observations within this range (red dots) is taken as a representative cloud top, which is then compared to the operational AMV pressure height. By applying a threshold of at least 20 lidar observations (and at most 100 due to the maximum horizontal distance of 50 km), it is ensured that dimensions of the tracked AMV feature (which is situated within a 24x24 pixel box) and the CALIPSO cloud are similar. In addition, the RMS difference between single lidar cloud observations and their median value must not exceed 70 hPa in order to exclude scenes with strongly altering cloud-top heights within short distances. Multi-layer cloud scenes as well as

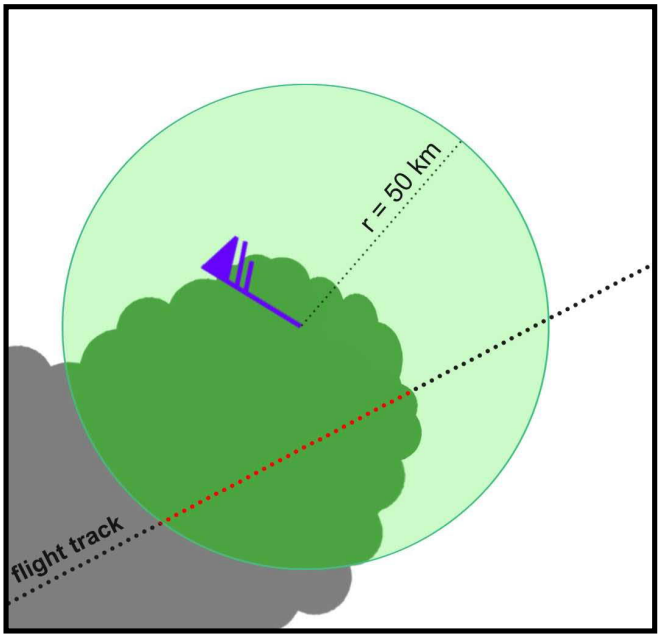

Figure 8: *Illustration of the collocation procedure: The lidar signal along the (dotted) flight track of CALIPSO may detect cloudy and cloud-free areas in the vicinity (50 km radius, green circle) of the corresponding (purple) AMV. All cloud signals within this range (red dots) are used for the calculation of the median cloud-top. Adapted from Folger (2012).*

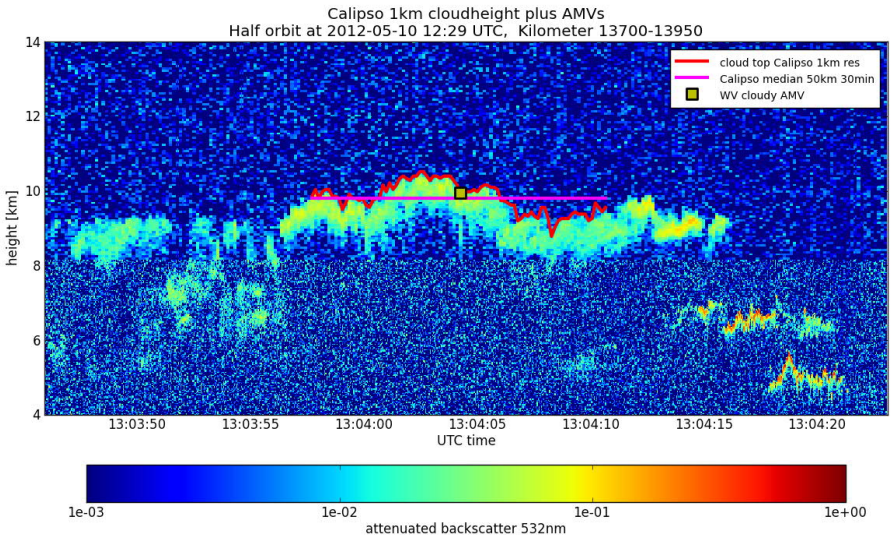

Figure 9: *Illustration of the calculation of the median cloud top (purple line) as representative cloud-top of all available CALIPSO lidar cloud-top observations within 50 km (corresponding to about 7 seconds on the x-axis) for one WV-AMV (yellow square) on 10 May 2012. The nominal AMV derivation time is 13:30 UTC; therefore all shown lidar cloud- top observations are within the maximum allowed temporal difference of 30 minutes of the AMV.*

parameter	**threshold**
horizontal distance	≤ 50 km
time difference	≤ 30 min
AMV QI	≥ 50
number of lidar cloud-top observations for median	≥ 20
RMS between individual lidar cloud-top observations and their median	≤ 70 hPa
CAD	≥ 90
multi-layer clouds	no
vertical distance	AMV ≤ 100 hPa above and ≤ 200 hPa below lidar cloud-top

Table 4: *Collocation criteria for AMV and CALIPSO observations*

cloud observations with a CAD < 90 are discarded. However, situations may still arise where CALIPSO observes different clouds than the ones used for deriving AMVs, either due to the temporal or horizontal distance of the lidar cloud and the corresponding AMV, or due to different instrument capabilities. To mitigate these issues, only AMVs that are at most 100 hPa above and 200 hPa below the respective median cloud-top heights are considered. This asymmetric interval is chosen based on the assumption that an AMV represents the atmospheric motion of a vertically extended cloud layer and is therefore located below the actual cloud top. All AMVs beyond this range are discarded. In addition, the AMV quality index (ranging from 0 to 100, with 100 indicating the best possible quality) must be greater than 50. An example case for the determination of the median is shown in Figure 9 for one WV-AMV on 10 May 2012.

These collocation parameters are chosen to account for the main issues that may arise when AMV pressure heights are compared to lidar cloud-top observations. A summary of the collocation criteria applied is provided in Table 4. Tighter thresholds may be desirable for some parameters (for example smaller horizontal and temporal distances), but the sample size then reduces drastically. Therefore, the presented threshold requirements provide a trade-off between the amount of available AMV/CALIPSO pairs and the matching accuracy. However,

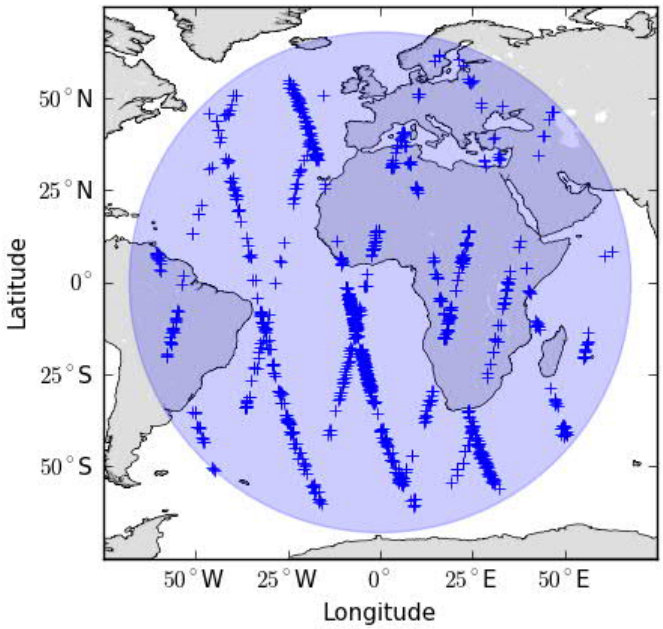

Figure 10: *Geographic position of the 1247 AMVs with collocated CALIPSO lidar observations on 1 Apr 2012 that fulfil the collocation requirements described in Section 3.1.2.*

sensitivity studies with modified collocation criteria revealed that tightening the parameters towards stricter values generally leads to similar findings. In this regard, the restriction of the vertical difference between the pressure heights seems to be the most important criterion to ensure capturing the same feature both from SEVIRI and CALIOP due to their different instrument characteristics.

Figure 10 shows the position of Meteosat-9 AMVs and CALIPSO lidar observations on 1 April 2012 matching the described collocation requirements. For this day, 1247 collocated observations were found within the Meteosat-9 domain (approximately 63° in each direction from 0° longitude and 0° latitude).

3.1.3 AMV wind evaluation methods

3.1.3.1 Radiosonde observations

The direct lidar height reassignment is evaluated using nearby operational radiosonde soundings in Section 4.1. As the wind field is usually horizontally more uniform than cloud-top heights, the collocation criterion for nearby radiosondes is extended to 150 km and 90 min from the corresponding AMV. Given that radiosonde wind observations are direct in-situ measurements, the inherent errors (originating e.g. from instrument errors) are small compared to errors of satellite-based retrievals, and observation errors are assumed to be uncorrelated between different radiosondes. Generally, an additional error source arises from position errors and temporal errors due to the radiosonde drift during the ascent, as the exact position and time of the respective wind observation is not stored in the DWD database. However, Seidel et al. (2011) assessed average radiosonde drift distances and ascending times to be typically in the range of 5 km (20km) and 0.07 h (0.89 h) for the lower troposphere (upper troposphere), which are smaller than the collocation criteria of 150 km and 90 min. Overall, radiosonde wind observations provide very accurate information on the wind conditions and will serve as a reference for the true state of the atmosphere in the following.

The calculation of the layer-average of the wind observations at the single radiosonde levels follows Folger (2012). All wind observations within the respective layer range are taken into account by allocating appropriate (boxcar) weightings to each radiosonde level according to its distance from the adjacent pressure levels: One radiosonde wind observation is weighted

with the pressure distance ranging from half-way to the overlying to half-way to the underlying pressure level (as percentage of the entire layer), resulting in high weights for isolated pressure levels and low weights for pressure levels with close neighboring values. If a layer exceeds the lowest or highest radiosonde level, the layer depth is reduced accordingly.

3.1.3.2 GME model equivalents

In-situ wind observations by radiosondes generally have small errors and therefore provide an ideal data source for the wind evaluation. However, the inhomogeneous spatial and temporal distribution of radiosonde profiles complicates their use for evaluation purposes. The additional need for verification radiosondes drastically limits the sample size of collocated AMVs and lidar observations to about one percent of the original number. To overcome this limitation, FG fields of the global model GME are additionally used for the wind evaluation in Section 4 (direct height reassignment) and 5 (height bias correction). As already mentioned in Section 2.3.1, FG fields are 3-h short-term forecasts from the previous time step, which are commonly used in data assimilation as background field. Current observations (such as AMVs) of the actual time step have not been assimilated yet in the FG fields. However, AMVs are assimilated operationally in all preceding time steps, thereby retaining this wind information for the actual time step through the cycling process. It should be kept in mind that this might influence the AMV evaluation for the current time step due to the temporal correlation of AMV errors.

Analogous to the verification with radiosonde data, the FG model equivalents serve as a reference for the real atmospheric state. This assumption should be treated with care, as model fields can deviate from the true state due to errors in the forecast model. However, model errors and AMV errors can be assumed to be uncorrelated. Based on this assumption, differences between observation and model winds provide a good data source for the evaluation of observations that is available area-wide for every observation and therefore commonly used in the context of data assimilation. Direct comparisons of results from radiosonde and model evaluation shown in Section 5.2 substantiate this assumption, as similar findings are yielded for the two evaluation methods.

Evaluating and assimilating layer-averaged observations requires a forward operator for treating AMVs as vertical layers. An observation operator for layer-averaged AMVs was recently implemented in the DWD system. This operator provides AMV model equivalents

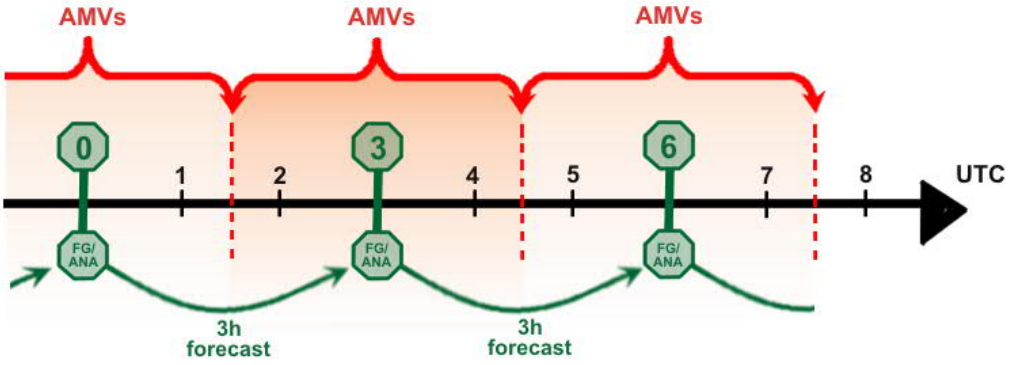

Figure 11: *Illustration of the data assimilation process and the AMV wind evaluation. All AMVs within a 3-h window around the respective time step (+/-90 min) are evaluated using the FG field, which corresponds to the 3-h forecast of the last time step. Combining the FG with current observations within the assimilation window results in the analysis (ANA), from which a new forecast can be started.*

derived from 3-h short-term (FG) forecast that are used for the wind evaluation in the present study. The calculation of the 3-h forecast uses the operational GME settings, which also includes the assimilation of single-level AMVs assigned to their original height in the preceding time steps. As the FG is used for the wind evaluation, new observations are not assimilated yet at the corresponding time step. All AMVs within the 3-h assimilation window (+/- 1.5 h) that match the collocation requirements (see Section 3.1.2) are compared to the FG field at the corresponding time step, leading to a maximum temporal difference between AMV and model equivalents of 90 minutes. This approach is illustrated in Figure 11. The geographical position of each AMV is horizontally interpolated between the model grid points, and then the vertical layer averaging is applied. Layer-averages are calculated according to Simpson's rule (see e.g. Süli and Mayers, 2003). Therefore, a weighted average of the interpolated wind values at the layer centre, layer top and layer base is computed. The weighting coefficients of both the layer-top and the layer-base wind are $^1/_6$, while the wind at the layer centre is weighted with $^4/_6$. If a layer exceeds the lowest model level, the layer depth is reduced accordingly.

The layer-averaging scheme used for the model evaluation is different from that of the verification with radiosonde data, as Simpson's approach for computing layer-averages constitutes a more feasible approach for the implementation in the observation operator and models generally exhibit smoother wind profiles than radiosondes. However, results that will

	evaluation method	time frame	total amount of collocated AMV/CALIPSO pairs	notation hereafter
Direct height correction	radiosonde	1 April – 6 October 2012 and 16 April – 13 June 2013 (220 days)	4478	"radiosonde evaluation period"
	model	31 May – 10 June 2013 (11 days)	13200	"first evaluation period"
Height bias correction	model	7 May – 12 May 2013 (6 days)	7410	"second evaluation period"

Table 5: *Overview of the study periods used for the different evaluation methods*

be shown in Section 4.1 and Section 4.2 (radiosonde and model evaluation, respectively) imply that the different calculations of the layer mean yield similar findings.

3.1.4 Evaluation periods

The present study uses eight months (1 April – 6 October 2012 and 16 April – 13 June 2013) of operational AMVs that were derived hourly from the geostationary MSG satellites Meteosat-9 (2012 period) and Meteosat-10 (2013 period) by EUMETSAT. CALIPSO observations are missing on 27 days of the 8-month study period so that 220 days of data are available. Altogether, 243097 matches of MSG-AMVs and CALIPSO lidar observations are found in this period.

AMVs are divided in different altitude regions according to Menzel (1996), where intervals are defined that are commonly used for AMV classification. AMVs are classified as high-level AMVs for pressure heights < 400 hPa. Mid-level AMVs are located between 400 hPa and 700 hPa and low-level AMVs have pressure heights > 700 hPa. One common feature of AMV height distributions is the relatively small amount of mid-level AMVs, which complicates statistical analyses in this region due to an often insufficient sample size. The

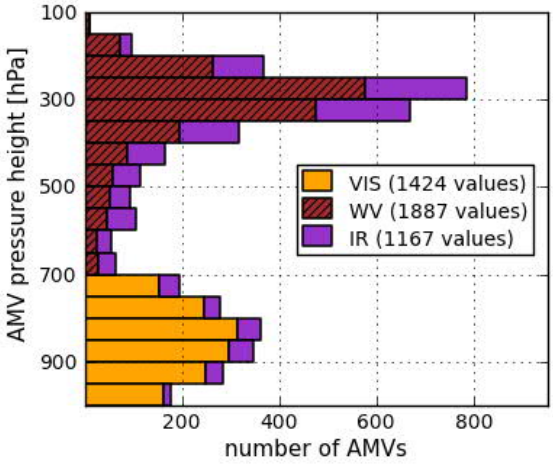

Figure 12: *Height distribution of all AMVs with collocated CALIPSO observations and radiosondes used during the radiosonde evaluation period*

small AMV amount in mid-levels is due to multi-layer situations when the uppermost cloud layer hides subjacent clouds. Thus, the distribution of AMVs is skewed towards upper levels (Régis Borde, personal communication). For the verification with radiosonde data in Section 4.1, where the overall sample size is comparatively small, this issue is circumvented by shifting the upper boundary of the "mid-level region" to a pressure height of 300 hPa. For the model evaluation in Sections 4.2 and 5.2, the original separation altitudes are used to classify high-level and low-level AMVs. Mid-level AMVs between 400 hPa and 700 hPa are not analysed for the model evaluation.

Generally, the verification with radiosonde data requires a long period to obtain an appropriate amount of AMV/CALIPSO pairs, as a collocated radiosonde must be additionally available for each pair. In contrast, the evaluation with model equivalents allows an analysis of all AMV/CALIPSO matches that fulfill the collocation requirements without the need of an additional radiosonde. Hence, already a small subset of all available days is sufficient for model evaluation purposes due to the larger amount of collocation pairs. The time periods used are summarized in Table 5 and described briefly in the following.

► Radiosonde evaluation period

For the verification with radiosonde data, all available 220 days are used. Given the comparatively low number of launched radiosondes per day, the sample size of collocated

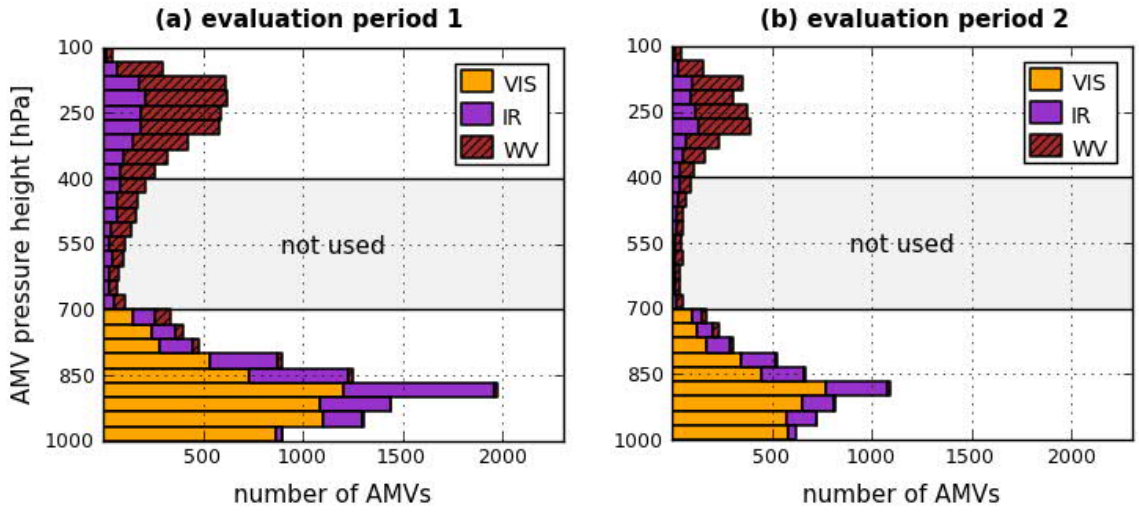

Figure 13*: Height distribution of Meteosat-10-AMVs with collocated CALIPSO lidar observations used for (a) the direct height reassignment (31 May 2013 – 10 June 2013) in Section 4.2 and (b) the height bias correction (5 May 2013 – 12 May 2013) in Section 5.2 for the model evaluation*

Meteosat-AMVs, CALIPSO lidar observations and operational radiosondes reduces to 4478 matches for the 8-month radiosonde evaluation period. Most of the height corrected AMVs are located over Europe and Africa, where radiosondes are available for the wind verification. Altogether, 1259 high-level AMVs derived from the IR- and WV-channels (337 and 922 matches, respectively) are available. The respective CALIPSO observations are all classified as ice clouds. The mid-level data set consists of 1576 AMVs (611 IR-AMVs and 965 WV-AMVs) and the corresponding CALIPSO cloud products comprise 67% ice clouds and 33% water clouds. The 1643 low-level AMVs from the IR- and VIS channels (219 and 1424 matches, respectively) are expected to correspond to water clouds only. Figure 12 shows the vertical distribution of all AMVs that are used during the radiosonde verification period.

► Model evaluation periods

For the model evaluation, two different periods are analysed. For the evaluation of the direct lidar height reassignment in Section 4.2, the evaluation period comprises 11 days (31 May – 10 June 2013) of operational AMVs. For the height bias correction in Section 5, a 6-day time interval is used that ranges from 7 May to 12 May 2013. These two periods are referred to as first and second evaluation period in the following. The slightly different timeframe is chosen because the height bias correction requires continuous (or nearly continuous) CALIPSO lidar

observations in the preceding 30-day time interval of the respective evaluation period. This criterion is not fulfilled for the first evaluation period due to significant gaps in the CALIPSO data set.

High-level AMVs above a pressure height of 400 hPa and low-level AMVs below a pressure height of 700 hPa are used for the model evaluation. Figure 13 shows the vertical distribution of AMVs for both evaluation periods for Meteosat-10. Altogether, 13200 AMVs in the first evaluation period and 7410 AMVs in the second evaluation period are analysed, with about 70% of them located in low-level and 30% in high-level regions.

3.2 AMV height correction methods

3.2.1 Direct height reassignment

Based on the collocation criteria defined in Section 3.1.2, the pressure heights of individual AMVs can be corrected with nearby CALIPSO lidar observations. This direct height reassignment is an adapted and extended version of the method presented in Weissmann et al. (2013), where airborne lidar observations during a field campaign are used for the height correction of MTSAT-AMVs.

3.2.1.1 Basic principle and error metrics

In order to find an optimal layer that represents the AMV wind, several combinations of different layer positions and layer depths are evaluated. An overview of the considered layers is illustrated in Figure 14. AMV winds are compared to radiosonde/model winds vertically averaged over layers of varying depth from 0 hPa to 200 hPa: firstly for layers relative to the originally assigned AMV height and secondly for layers relative to the CALIPSO lidar cloud top height. A layer depth of 0 hPa denotes a discrete level, which corresponds to the procedure applied operationally for the original AMV height. Three different layer positions are evaluated:

(i) layers centered at the corresponding AMV height or lidar cloud-top height,

(ii) layers with 25% of the layer range above and 75% of the layer range below the corresponding height and

(iii) layers from the corresponding height downward.

In order to assess the benefit of lidar-corrected and layer-averaged AMVs, two error metrics are applied for all considered layers: The VRMS difference and the wind speed bias. These are calculated as follows:

$$\mathrm{VRMS} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{du_i^2 + dv_i^2}$$

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(\mathrm{windspeed}_{\mathrm{oper,i}} - \mathrm{windspeed}_{\mathrm{model,i}})$$

with $du_i = u_i(\mathrm{AMV\_operational}) - u_i(\mathrm{AMV\_model/radiosonde})$, dv_i analogously.

N corresponds to the number of available AMVs with corresponding CALIPSO lidar cloud-

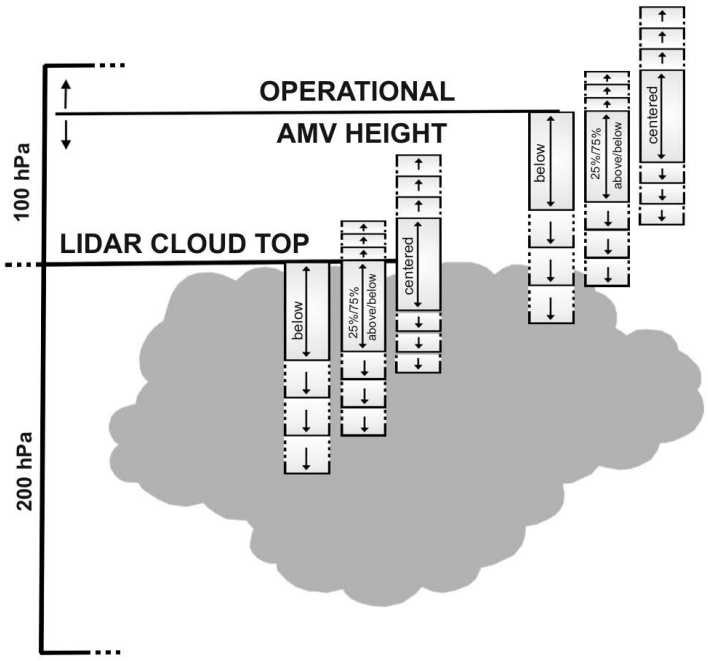

Figure 14: *Schematic illustration of the height correction method. Layers of varying depth relative to the lidar cloud-top height and relative to the operational AMV height are analysed. Layer depths vary from 0 hPa (discrete level) to 200 hPa. Only AMVs that are at most 100 hPa above and 200 hPa below the respective cloud-top height are evaluated. Not to scale.*

top observations, '*oper*' denotes operational wind observation values and '*model*' the GME model equivalents of the respective layers. The VRMS difference defined above is sometimes also referred to as Mean Vector Difference (MVD) (Menzel, 1996).

3.2.1.2 Assessment of AMV error correlations

AMV error correlations pose a major issue for data assimilation, as AMV errors can be correlated up to horizontal distances of several hundred kilometres (Bormann et al., 2003). Generally, the incorporation of model-independent data (such as lidar observations) in NWP systems is expected to reduce the error correlation of observations. This aspect is investigated for the horizontal AMV error correlation when independent lidar observations are included for the height correction.

Radiosonde observations are generally the optimal approach for evaluating such error correlations. Bormann et al. (2003) calculated the horizontal error correlation of differences between AMVs and radiosonde observations in order to characterize operational AMV error correlations. As observation errors between radiosondes are assumed to be uncorrelated, it was implied that the correlation of AMV-radiosonde differences can be directly ascribed to the correlation of AMV errors, which can therefore be evaluated directly.

Given that the limited number of both radiosondes and lidar observations massively restrict the sample size, the evaluation of error correlations with radiosonde observations in common with Bormann et al. (2013) unfortunately is not possible for the present study. Instead, the horizontal AMV error correlation of differences between AMV and model FG equivalents (FG departures) is calculated (i) for operational AMVs and (ii) for AMVs that are assigned to 120-hPa layer-averages below the lidar cloud-top height. Drawing conclusions from the correlation of AMV-FG differences (instead of AMV-radiosonde differences) to AMV error correlations is not unconditionally possible, as the inherent model errors are usually highly correlated. However, the *relative* reduction of the AMV error correlation can be analysed with FG model fields, as the model error is assumed to be the same for both settings (i) and (ii).

For the assessment of horizontal AMV error correlations, all Meteosat-10 AMVs with available collocated lidar observations during the first evaluation period (31 May – 10 June 2013) (called AMV sample "A" in the following) are evaluated. For these AMVs, circumjacent AMVs (without a direct lidar height correction available, AMV sample "B") are

divided into 50-km-bins around each member of AMV A for the same point in time. In addition, the vertical distance between the AMV pressure heights may not exceed 150 hPa for the AMV-A/AMV-B pairs obtained. The Pearson correlation coefficient $\rho^{\text{oper/CALIPSO}}$ for the u-component of the wind field is calculated for each 50-km-bin as follows (analogously for the v-component):

$$\rho_{A,B}^{\text{oper}}(u) = corr((u_{A,\text{oper}} - u_{A,\text{model\_oper}}),(u_{B,\text{oper}} - u_{B,\text{model\_oper}}))$$

$$\rho_{A,B}^{\text{CALIPSO}}(u) = corr((u_{A,\text{oper}} - u_{A,\text{model\_CALIPSO}}),(u_{B,\text{oper}} - u_{B,\text{model\_oper}}))$$

where the index *A* denotes the AMVs with a direct CALIPSO lidar height correction available and *B* the AMVs without collocated lidar observations that surround AMV A. The notation '*corr*' represents the calculation of Pearson's correlation coefficient of the following term in brackets. The index '*oper*' indicates the operationally assigned wind, '*model_oper*' the model winds at the operational AMV heights, and '*model_CALIPSO*' the lidar-corrected and layer-averaged model equivalents. The total error correlation of FG departures for u and v combined is then derived as

$$\rho_{A,B}^{\text{oper/CALIPSO}}(u,v) = \frac{\rho_{A,B}^{\text{oper/CALIPSO}}(u) + \rho_{A,B}^{\text{oper/CALIPSO}}(v)}{2}.$$

$\rho_{\text{A,B}}^{\text{CALIPSO}}(u,v)$ is the error correlation of FG departures between lidar-corrected AMVs (sample A) and operational (non-lidar-corrected) AMVs (sample B), whereas $\rho_{\text{A,B}}^{\text{oper}}(u,v)$ describes the error correlation of FG departures when AMVs from sample A and sample B are both not lidar-corrected, but retain their operational AMV height. Accordingly, the comparison between $\rho_{\text{A,B}}^{\text{CALIPSO}}(u,v)$ and $\rho_{\text{A,B}}^{\text{oper}}(u,v)$ provides information of altered error correlations when lidar-corrected AMVs are assimilated additionally to the operational AMV data set.

The additional calculation of the error correlation of FG departures when AMVs from both samples A and B are lidar-corrected is unfortunately not possible, as this is complicated to evaluate due to an insufficient sample size of closely collocated AMVs with direct lidar observations nearby.

3.2.2 Height bias correction

The previously described direct AMV height reassignment is based on actual cloud-top heights of lidar observations collocated to the respective AMV. However, this method is only applicable to a small number of operational AMVs, as the times and positions of the AMVs have to coincide with nearby CALIPSO lidar observations. As an alternative approach, a height bias correction for a general mean adjustment of all AMV heights from a respective satellite can be calculated.

3.2.2.1 Derivation of height bias correction functions

For the purpose of calculating height bias correction functions, the direct height reassignment is applied to all AMVs with available collocated lidar observations within a certain time period and then an average over the resulting height adjustment values is computed. This height bias correction is then applied to a subsequent independent evaluation period. Durations of 10 days and 30 days are used as averaging periods for computing the height bias correction functions. The resulting corrections are applied during the second evaluation period (7 May to 12 May 2013). The 30-day mean comprises the days from 1 April to 6 May 2013 (with missing CALIPSO data on six days within this interval). The 10-day mean is calculated from the ten days preceding the respective date. This means that for example the height bias correction derived from the period 2 May – 11 May 2013 is applied and evaluated on 12 May 2013. As a third approach, the 30-day period is subdivided for different latitude bands to determine separate correction functions for the northern hemisphere (latitude larger than 25°N), the southern hemisphere (latitude larger than 25°S) and a tropical region in between. Table 2 lists the numbers of AMVs with collocated CALIPSO lidar observations that are used for the different height correction periods for Meteosat-10.

Height bias correction functions are calculated separately for the different channels (VIS, IR and WV) for 50-hPa altitude-bins between 950 hPa and 200 hPa plus one additional bin each for AMVs below and above this range. Every bin must contain at least 30 individual adjustment values to determine a valid mean adjustment for the respective altitude bin and AMV channel.

Mean VRMS differences and wind speed bias values are calculated for all AMVs for (i) discrete operational levels, (ii) levels at 60 hPa below the actual lidar cloud-top

height correction period	counts	
	low level	high level
30 days	24027	12114
30 days northern hemisphere	3439	2901
30 days tropics	10889	6571
30 days southern hemisphere	9699	2606
10 days	7873	4114

Table 6: *Number of lidar-corrected AMVs used to calculate height bias correction functions over different periods for Meteosat 10. Counts of the 30-day height correction periods comprise AMVs in the period 1 Apr – 6 May 2013. Counts of the 10-day height correction period are averaged over the respective counts for each of the six days of the second evaluation period.*

observation and (iii) adjusted levels based on the height bias correction. In addition, 120-hPa-deep layer-averages centered at these levels are evaluated. The level 60 hPa below the lidar cloud top is chosen as it represents the mean pressure of the 120-hPa layer.

3.2.2.2 Assimilation and forecast experiments

► Experimental setup

Correcting the operational pressure height of AMVs with directly collocated lidar observations causes difficulties due to availability problems of real-time lidar data. Hence, a height bias correction has been developed that can be applied more easily in data assimilation systems. The benefit of statistically corrected Meteosat-10-AMVs is tested by conducting

assimilation and forecast experiments with ICON, which is the global forecasting system currently used at DWD (for details on ICON see Section 2.3.1). The experiments are conducted for a 16 day assimilation period (1 May 2013 – 16 May 2013). For that purpose, height bias correction functions are calculated separately for each day of the assimilation period from a 15-day training interval preceding the respective date. This height bias correction is then applied to all operational Meteosat-10-AMVs for the corresponding day. The training period duration of 15 days is chosen due to a CALIPSO data gap before 16 April 2013 that does not allow for longer training periods.

In the operational setting of ICON, AMVs are assimilated every 6 hours, discarding about half of the hourly derived AMV data set that remains after thinning and quality control procedures. This might lead to a suboptimal usage of wind information in data assimilation systems due to strong variations of wind observation numbers from one cycle to another and is considered to be changed in operational usage in the future. Thus, AMVs are assimilated in a 3-h cycling interval in this study. The horizontal resolution used in the experiments is 40 km, which is coarser than the current operational resolution (13 km) due to the limitation of data storage capacity. Apart from these changes, the experimental setup used is equivalent to the operational settings, including also other baseline observations from the global observing system.

Using this experimental setup, three different experiments for different settings for Meteosat-10-AMVs are conducted. The reference run (REF) assimilates Meteosat-10-AMVs at their originally derived pressure height. For two further experiments, the height bias correction is applied to the same set of Meteosat-10-AMVs as in the reference run, thereby shifting the AMVs to an adjusted pressure height based on a 15-day training period. These statistically corrected AMVs are then assimilated in two experiments: firstly at the discrete lidar-corrected level (LEV) and secondly as a layer-average extending over a 120-hPa deep layer centered at the lidar-corrected level (LYR). The three experiments are summarized in Table 7.

► Forecast evaluation methods

In contrast to the evaluation of AMVs (with individual radiosonde or model equivalents at the specific AMV coordinates), the evaluation of forecast model fields focuses on an area-covering impact assessment of the obtained forecast field. The evaluation of forecast

model fields can be conducted using different measures. For an optimal and accurate impact assessment, the model fields obtained should be compared with independent in-situ observations. In practice, this is only possible in a limited way due to a spatially and temporally inhomogeneous observation network that is incapable of providing global observations at every model grid point. However, the verification of the forecast wind field can be performed by comparing (irregularly distributed) radiosonde soundings to the corresponding model profile at the same location and time. This approach is used in this thesis for the evaluation of short-term forecasts (FG fields).

In order to evaluate the performance of long-term forecasts in a more homogeneous manner, a common evaluation approach is to compare the respective forecast field to the analysis field, i.e. the analysis field of the 3-h data assimilation cycle of the same experiment at the same point in time. As described in Section 2.3.1, the analysis field represents an estimate of the atmospheric conditions at the present time step by combining the FG field and current available observations in an optimal way. By cycling previous information during the data assimilation process, the analysis fields also contain information from all observations since the forecast started (and also from the current time step). Although these analysis fields are not equivalent to the true state of the atmosphere, they represent a best estimate of the truth that is available area-wide and is therefore commonly used for evaluation purposes in data assimilation.

As advanced global models like ICON assimilate a large number of observations and have, in general, a high forecast performance, effects on the forecast skill by changing the assimilation

short name	**experiment**	**description**
REF	reference run	AMVs assimilated at original pressure height
LEV	height bias correction: level assimilation	AMVs assimilated at adjusted pressure height based on a 15-day bias-correction
LYR	height bias correction: layer assimilation	AMVs assimilated as 120-hPa layer-averages centered at the adjusted pressure based on a 15-day bias-correction

Table 7: *Overview of the three assimilation experiments*

area	radiosonde ascents	assimilated AMVs (high | mid | low)
total	6702	513888 (38.6% | 12.9% | 48.5%)
Northern hemisphere	4808	123733 (46.5% | 14.4% | 39.1%)
Tropics	1416	228879 (39.7% | 5.7% | 54.6%)
Southern hemisphere	478	161276 (31.0% | 21.8% | 47.2%)

Table 8: *Number of radiosonde ascents and assimilated AMVs for the 16-day assimilation period (1 May – 16 May 2013) for a hemispheric subdivision of the Meteosat-10-domain. AMV counts are additionally split in the percentages of high-level, mid-level and low-level AMVs. The listed AMV counts are the respective average of the three experiments REF, LEV and LYR (exact numbers vary in the range of 0.05%-0.2%).*

parameters of one single observation type (such as AMVs) are expected to be relatively small. Evaluation of forecast runs is therefore commonly performed using the normalized difference of the RMS of the prognostic variables between the obtained forecast field and the reference field (which corresponds to the analysis of the 3-h cycle of the same experiment). The normalized RMS (or analogously VRMS) difference is calculated as follows:

$$\text{norm. RMS diff.}\,[h] = \frac{\text{mean}\left(\text{RMS}_{\text{LYR}}[h] - \text{RMS}_{\text{REF}}[h]\right)}{\text{mean}\left(0.5 \cdot \left(\text{RMS}_{\text{LYR}}[h] + \text{RMS}_{\text{REF}}[h]\right)\right)}$$

RMS_{LYR} denotes the RMS error of the experiment LYR, RMS_{REF} analogously for the experiment REF. The normalized RMS difference is computed for each forecast lead time h. The calculation of the mean both in the numerator and the denominator of the fraction comprises all available forecast runs for the respective lead time.

Altogether 127 3-h assimilation cycles are computed, leading to 513888 assimilated Meteosat-10-AMVs in total. The corresponding numbers of assimilated AMVs for a subdivision into the hemispheric and tropical parts of the Meteosat-10-domain are listed in Table 8 in the right column. In Section 5.3.1, all operational Meteosat-10-AMVs of the 16-day assimilation period are compared to the corresponding FG model equivalent of the respective experiment REF, LEV and LYR. In addition, the FG fields of the three experiments are compared with radiosonde soundings for all latitudes and longitudes between -65° and +65°, corresponding roughly to the Meteosat domain. The counts of the available radiosondes are also listed in Table 8.

In addition to the evaluation of short-term forecasts, preliminary results from 7-day forecast runs over the European region are shown in Section 5.3.2. The two experiments REF and LYR are evaluated by comparing forecast fields to the analysis field of the 3-h assimilation cycles. Forecasts are initialized at 00 UTC and 12 UTC of each day during the assimilation period. As the corresponding analysis field is needed for the evaluation of the forecast lead times, the forecasts of the last seven days of the assimilation period cannot be considered for evaluation. Altogether, 17 free forecast runs (the first one from 1 May 2013, 12 UTC and the last one from 9 May 2013, 12 UTC) are evaluated.

4. DIRECT HEIGHT REASSIGNMENT

This section aims to investigate whether space-based lidar observations can be used for the height correction of AMVs and evaluates the effect of reassigning AMVs to layers of varying depth and position relative to the lidar cloud-top height. Results for this individual height reassignment for AMVs with directly collocated satellite lidar observations are evaluated firstly with radiosonde wind observations (Section 4.1) and secondly with model equivalents from the global model GME (Section 4.2). A detailed description of the approach for the direct height reassignment can be found in Section 3.2.1.

4.1 Verification with radiosonde data

Accurate in-situ radiosonde wind observations are generally the optimal choice for the evaluation of AMV winds. However, the temporally and spatially rare distribution of operational radiosondes requires a long study period to achieve a sufficiently large data set of AMVs with collocated CALIPSO observations and radiosondes for an adequate evaluation. For this purpose, 220 days with altogether ~4500 matches are used for the verification with radiosonde data (see Section 3.1.4 for more details).

4.1.1 VRMS differences and wind speed bias

Figure 15 shows the mean VRMS differences of AMVs and radiosonde winds. VRMS values are calculated for assigning AMVs to vertical layers of increasing depth, which are computed by averaging radiosonde winds over the respective layer. The first set of layers uses the original AMV height as reference (red lines); the second set uses lidar cloud top observations as reference (green lines). The corresponding wind speed bias is shown in Figure 16.

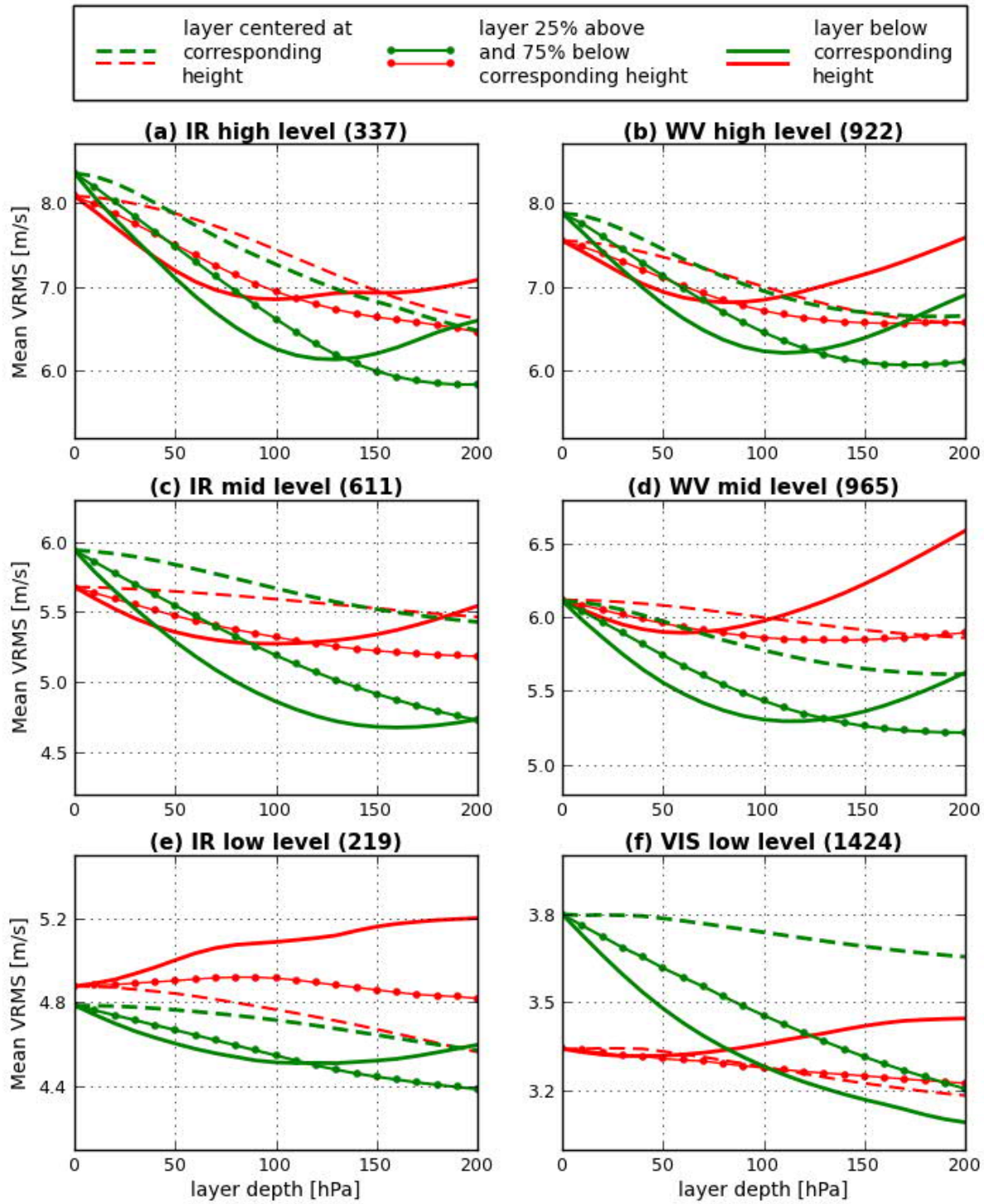

Figure 15: *Mean VRMS differences between AMV winds and layer-averaged radiosonde winds for (a) high-level IR-AMVs, (b) high-level WV-AMVs, (c) mid-level IR-AMVs, (d) mid-level WV-AMVs, (e) low-level IR-AMVs and (f) low-level VIS-AMVs. Numbers in brackets are AMV counts for the respective graph. Red lines represent layers relative to the original AMV pressure height, green lines relative to the lidar cloud-top height. The three different layer positions are indicated by different line styles (cf. legend).*

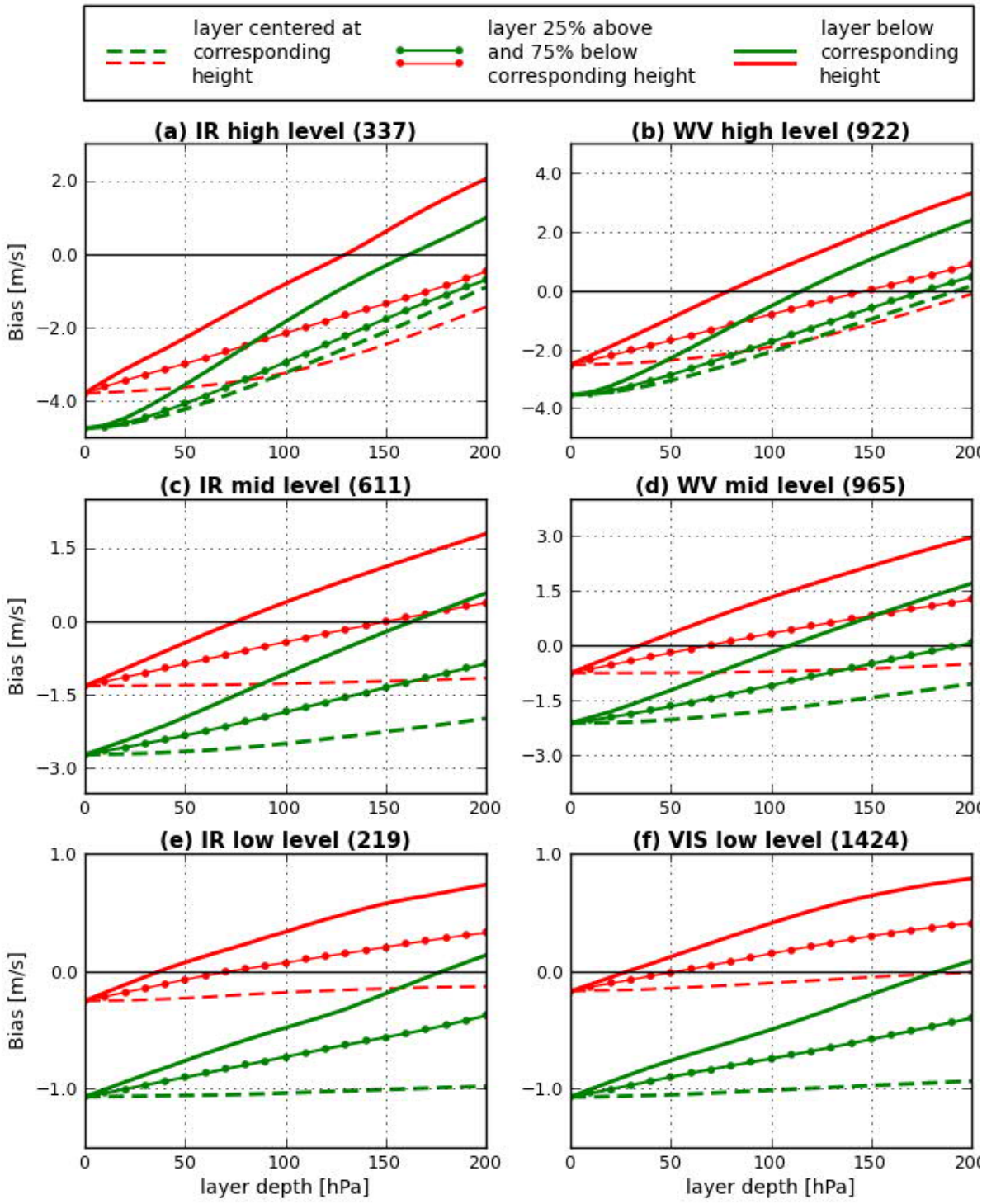

Figure 16: *As Figure 15 but for wind speed bias*

Results for high- and mid-level AMVs above 700 hPa (Figs. 15a – 15d) from WV- and IR-channels exhibit a distinct reduction of VRMS differences when AMVs are treated as vertically extended layers instead of as single level observations (which are the values for

0 hPa on the x-axis). Lowest VRMS differences are achieved either by layers below the lidar cloud top or by layers with 25% above and 75% below the lidar cloud top. The optimal depth of these layers varies from 120 hPa to 200 hPa. Layers below the lidar cloud top exhibit lowest VRMS differences for a depth of 100-150 hPa and layers with 25/75% above/below the lidar cloud top yield best results for a depth of 150-200 hPa.

Overall, the shape of the curves for these two lidar layers is fairly similar for the different subsets presented in Figs. 15a –15d and small differences in the position of the minimum may also be a result of the limited sample size of individual subsets instead of systematic differences in between them. For all these four subsets, the minimum of VRMS differences for layers relative to the lidar cloud top is in the range of 0.5 – 1.5 m s $^{-1}$ lower than the lowest values reached with layers relative to the original AMV height.

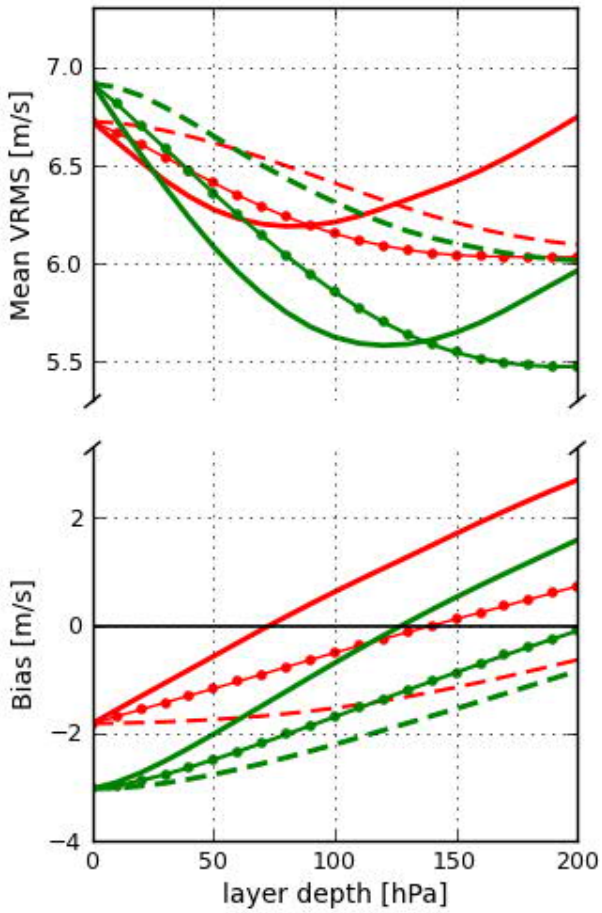

Figure 17: *Mean VRMS and wind speed bias of differences between AMV winds and layer-averaged radiosonde winds for upper level AMVs above 700 hPa (IR and WV combined). Altogether, 2835 AMVs are used (948 IR and 1887 WV). Red lines represent layers relative to the original AMV pressure height, green lines relative to the lidar cloud top height. Note that the scales for mean VRMS and wind speed bias values are different.*

Figs. 16a and 16b exhibit a significant “slow” bias of high-level AMVs assigned to their original discrete height (values for 0 hPa on the x-axis). Such a slow wind speed bias has also been found in other recent studies (e.g. Bresky et al., 2012). Generally, the wind speed bias is reduced when AMVs are assigned to deeper layers and results indicate that assigning them for example to layers of 100-150 hPa below the lidar cloud top largely removes the slow bias of current upper level AMVs. Overall, the results presented in Figure 16 show that layers leading to low VRMS differences tend to be similar to layers leading to a low wind speed bias.

In contrast to upper level AMVs, low-level AMVs (Figs. 15e and 15f) are typically assigned to an estimated cloud base height rather than a level near the cloud top. Averaging over layers that are centered at the original AMV height shows a slight advantage over the discrete value with increasing layer depth. Worse results are revealed for layers below the original AMV height, which might be due to the fact that these AMV heights are relative to the cloud base and consequently, a layer below the cloud base does not represent the wind conditions of the tracked cloud correctly. Slightly better results are obtained when lidar cloud-top information is incorporated, but the benefit is less distinct than for mid- and high-level AMVs. 200 hPa layers with 25/75% above/below the lidar cloud top and 200 hPa layers below the lidar cloud

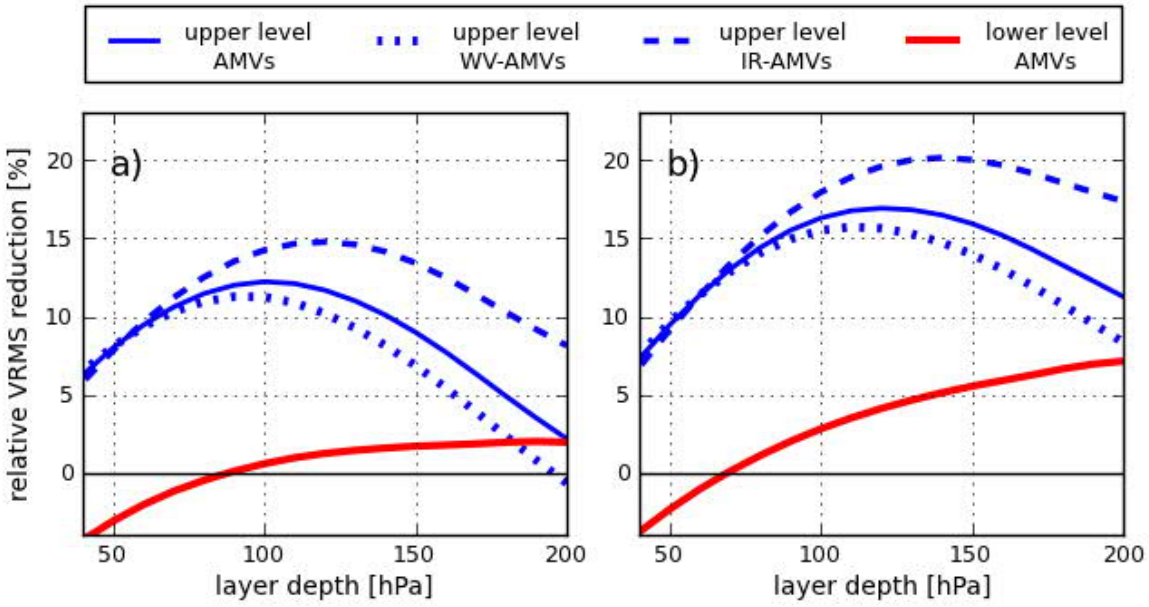

Figure 18: *Relative reduction of VRMS differences between AMV and radiosonde winds for assigning AMVs to layers below the lidar cloud top instead of (a) layers of the same depth centered at the original AMV height and (b) the discrete original AMV heights. Upper level AMVs above 700 hPa (blue solid line) are additionally divided into upper level WV-AMVs (blue dotted) and upper level IR-AMVs (blue dashed). The red solid line represents results for lower level AMVs (≥ 700 hPa).*

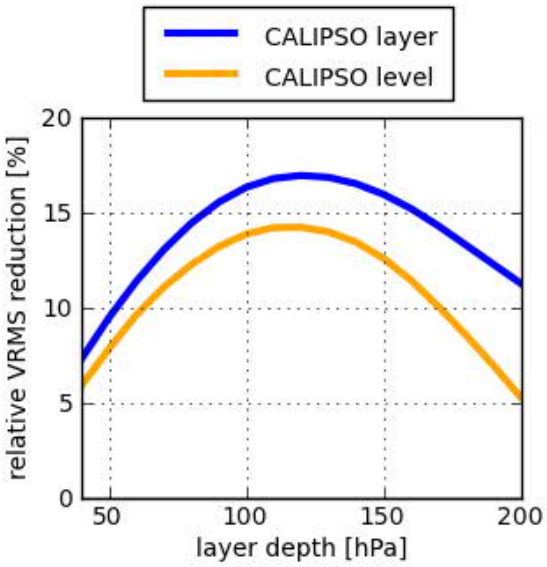

Figure 19: *Relative reduction of VRMS differences between AMV and radiosonde winds for assigning AMVs to layers below the lidar cloud top (blue line) and to the respective mean pressure levels of that layer below lidar cloud top (orange line) instead of the discrete original AMV heights.*

top (for IR and VIS, respectively) lead to the lowest VRMS differences, but results for layers of the same depth centered at the original AMV heights are only 0.1-0.2 m s $^{-1}$ higher. As low-level AMVs are located at pressure heights greater than 700 hPa, the 200 hPa layers below the lidar cloud top are mostly layers from the lidar cloud top to the lowest radiosonde level. The lower benefit of lidar cloud top heights for the reassignment of low-level AMVs may result from the relation of low-level AMVs to cloud-base winds and the inability of satellite lidars to observe these cloud bases.

High- and mid-level AMVs overall exhibit a similar behaviour and therefore all AMVs above 700 hPa are combined in Figure 17. The combination of high- and mid-level AMVs will be referred to as “upper level AMVs” in the following. Results indicate that lowest VRMS differences in combination with lowest wind speed bias values are achieved for either 120-130 hPa layers below the lidar cloud top or for 200 hPa layers with 25/75% above/below the lidar cloud top.

4.1.2 Relative VRMS reduction for lidar layers and lidar levels

Figure 18 shows the relative reduction of VRMS differences when results for layers below the lidar cloud top are compared to results of layers of the same depth centered at the original

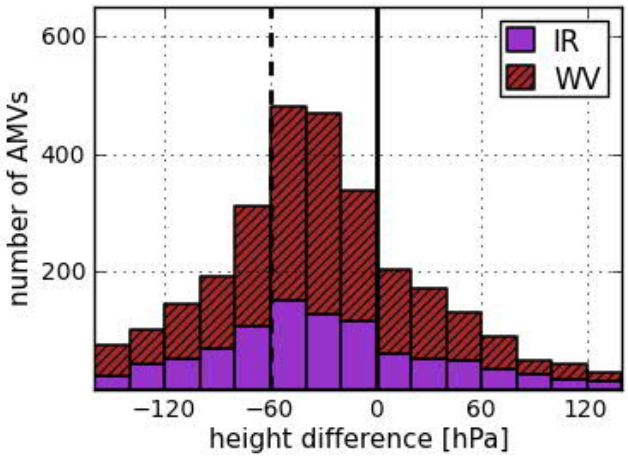

Figure 20: *Histogram of height differences (hPa) between the original AMV pressure height and the mean pressure of the corresponding 120 hPa layers below the lidar cloud top for upper level AMVs above 700 hPa (1887 WV-AMVs and 948 IR-AMVs). The dashed vertical line corresponds to the pressure height of the lidar cloud top.*

AMV height (Fig. 18a) and results using the discrete original AMV height (Fig. 18b). The shape of the curves in Fig. 18a and 18b is similar. For upper level AMVs (blue lines), best results are yielded for layer depths of 100-120 hPa. Highest reduction values are ~12% for lidar layers compared to layers centered at the original AMV height (Fig. 18a) and ~17% compared to the discrete original AMV height (Fig. 18b). The improvement is apparent in both upper level channels IR and WV (blue dotted and dashed lines). Dividing between upper level ice clouds and water clouds leads to a similar reduction and is therefore not shown. About 59.4% (64.6%) of the 2835 upper level AMVs show reduced VRMS differences for 120 hPa layers below the lidar cloud top in relation to 120 hPa layers centered at the original AMV height (to the discrete original AMV heights).

Correcting the height of low-level AMVs (red lines) with lidar information only leads to a small VRMS reduction, but the averaging over deep layers shows advantages over using discrete heights. The VRMS differences for 200 hPa layers below the lidar cloud top are predominantly superior to the VRMS differences of 200 hPa layers centered at the original AMV height and of the discrete original AMV heights (50.8% and 59.2%, respectively).

After demonstrating the benefit of assigning AMVs to vertical layers below the lidar cloud top, it is now investigated how much of that reduction could be achieved by assigning them to one representative discrete level relative to the lidar cloud top instead. The blue solid line in

Figure 19 represents the treatment of AMVs as a layer-average below the lidar cloud top (equivalent to the blue solid line in Figure 18b), whereas the orange line represents the assignment of AMVs to the discrete mean pressure height of that lidar layer, i.e. a discrete level located half of the layer depth below the lidar cloud top. Results indicate that assigning AMVs to the mean pressure of the lidar layers achieves most of the reduction of assigning AMVs to vertically extended lidar layers. However, interpreting AMVs as layer-averaged winds leads to a relative reduction that is ~3% higher. The maximum of the curves is for both approaches at ~120 hPa, which corresponds to using discrete levels 60 hPa below the lidar cloud top. The corresponding wind speed bias values at this maximum are for both approaches close to zero (not shown).

Figure 20 illustrates the distribution of differences between the original AMV pressure and the mean pressure level of 120 hPa deep layers below the lidar cloud top for upper level AMVs. About 75% of the AMVs are located above the mean pressure of the lidar layers and are thus shifted to lower altitudes (negative values) with the lidar height correction. As AMVs are derived by tracking the motion of the cloud, the lidar cloud top (dashed line) marks the natural upper edge where AMVs should be located. However, approximately 30% of the

		upper level AMVs		**low-level AMVs**	
	time period	**VRMS reduction**	**counts**	**VRMS reduction**	**counts**
	All (220 days)	16.9	2835	7.1	1645
(1)	**1 Apr. – 3 Sep. 2012** (142 days)	18.9	1725	5.1	999
(2)	**5 Sep. – 6 Oct. 2012** (32 days)	11.4	406	18.5	249
(3)	**16 Apr. – 12 June 2013** (46 days)	14.1	704	5.6	397

Table 9: *Relative VRMS reduction in percent and number of matches for different time periods for assigning AMVs to layers below the lidar cloud top instead of the discrete original AMV height. The depth of the assigned layers is 120 hPa (200 hPa) for upper (low) level AMVs with pressure heights above (below) 700 hPa.*

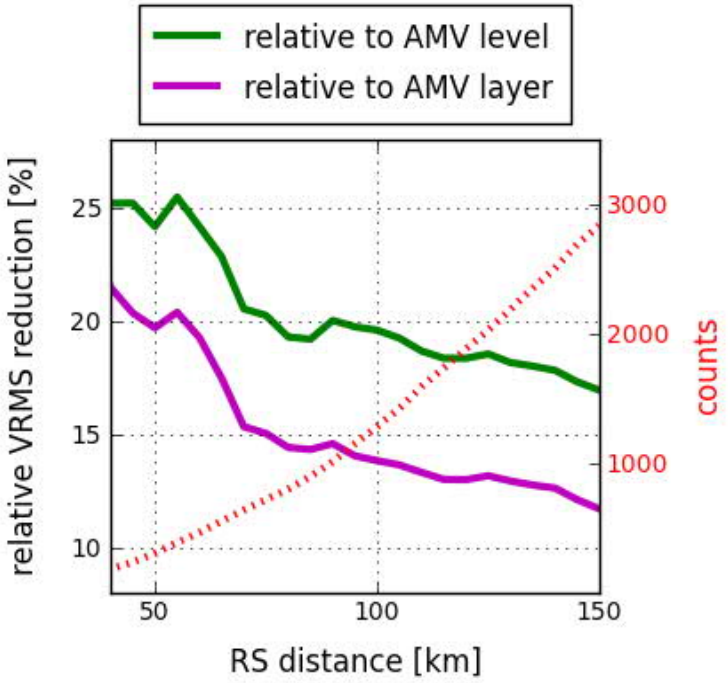

Figure 21: *Relative VRMS reduction of differences between AMV and radiosonde winds as a function of the horizontal distance between AMV and radiosonde (RS) for assigning AMVs to 120 hPa layers below the lidar cloud top instead of layers centered at the original AMV height (green line) and the original discrete AMV height (purple line). The dotted red line corresponds to the y-axis-label on the right and shows the sample size.*

AMVs are located above the cloud, which may be related to an erroneous height assignment as well as to the temporal and horizontal displacement of AMV and CALIPSO lidar observation. On average, upper-level AMVs are located 31 hPa above the lidar layer center (and correspondingly, 29 hPa below the lidar cloud top), with only small differences between the single channels WV and IR. In summary, this indicates that the operational processing of upper-level AMVs should consider that AMVs rather represent wind in a layer below the actual cloud top, but the systematic height differences are likely dependent on the applied AMV processing systems and its settings.

4.1.3 Effects of using different subsamples

To investigate the effect of changes in the height assignment algorithm of EUMETSAT, the analysed 220 days are divided into three different time periods in Table 9. The first one comprises 142 days before 5 September 2012, the day when the height assignment algorithm

was changed to the CCC-method (see Section 2.1.1.2). The second period consists of 32 days starting on 5 September 2012 and the last period consists of 46 days from 16 April until 12 June 2013. According to the preceding results (see Fig. 17), the lidar layer depth is set to 120 hPa for upper level AMVs and 200 hPa for low-level AMVs. For upper level AMVs, the VRMS reduction for assigning layers below the lidar cloud top instead of the discrete original AMV heights is apparent in all three periods ranging from 11.4% to 18.9%. As stated before, low-level AMVs do not show a clear reduction of VRMS differences through the direct height reassignment. However, one noticeable feature is the high VRMS reduction for low-level AMVs in the second period from 5 September to 6 October 2012. This is likely related to a temporary degradation of the quality of low-level AMVs in the time period after the height assignment algorithm changed to the CCC-method (Salonen and Bormann, 2012).

In order to utilize a reasonable large sample size, the collocation criterion for AMVs and radiosondes is set to 150 km and 90 minutes. However, the temporal and spatial displacement of AMVs and verification radiosondes introduces an additional error component that is expected to be independent of the AMV error itself and the height correction. Therefore, a weak collocation criterion leads to an underestimation of the actual relative VRMS reduction. Figure 21 shows how the relative reduction of VRMS differences for upper level AMVs increases as the horizontal collocation criterion is tightened. Naturally, the number of matches decreases for smaller distances. The error reduction for 120 hPa layers below the lidar cloud

	upper-level AMVs		**low-level AMVs**	
QI	**error reduction**	**counts**	**error reduction**	**counts**
≥ 50	16.9	2835	7.1	1643
≥ 60	16.8	2573	8.0	1439
≥ 70	16.6	2265	8.3	1254
≥ 80	14.5	1792	9.4	1003

Table 10: *Relative VRMS reduction in percent and number of matches for different quality indices QI for assigning AMVs to layers below the lidar cloud top instead of the discrete original AMV height. The layer depth is 120 hPa (200 hPa) for upper (low) level AMVs with pressure heights above (below) 700 hPa.*

top relative to layers centered at the originally assigned AMV heights shows a strong increase from ~12% at 150 km to ~21% at 40 km (purple line). When compared to the discrete original AMV height, the relative VRMS reduction increases from ~17% to ~25% (green line). Reducing the time difference does not lead to clearly larger improvements and is therefore not shown.

This study uses a threshold for the AMV quality index QI of 50. Restricting it to higher values (up to $\geq$ 80) reduces the sample size to up to ~60%. Table 10 lists the relative VRMS reduction for assigning 120 hPa layers (upper level AMVs) and 200 hPa layers (low-level AMVs) below the lidar cloud top instead of the discrete original AMV heights for different quality thresholds. Restricting the sample to upper level AMVs with QI $\geq$ 80 shows slightly less improvement than including lower quality AMVs, but the differences are smaller than 2.5%. For low-level AMVs, the VRMS reduction slightly increases when only AMVs with higher quality are regarded.

4.2 Comparison with GME model equivalents

In addition to the preceding verification with radiosonde data, AMV winds are now additionally evaluated with GME model equivalents to circumvent the constraint of temporally and spatially rare radiosondes. This allows analysing a considerably larger amount of AMVs, albeit this approach does not provide an entirely independent dataset for the validation. Based on the results of the previous chapter, layers of varying depth ranging from 0 hPa to 200 hPa at three positions are evaluated: (i) below the lidar cloud-top height, (ii) with 25% above and 75% below the lidar cloud-top height and (iii) centered at the operational AMV height as reference layers. Altogether, 13200 AMV/CALIPSO matches during the first evaluation period (31 May – 10 June 2013) are analysed.

Figure 22 illustrates the distribution of height differences between operational AMV heights and cloud-top heights derived from collocated CALIPSO lidar observations for all used Meteosat-10-AMVs. More than 80% of all operationally assigned AMVs are located below the actual lidar cloud top, corresponding to positive height differences on the x-axis. The highest number of AMVs occurs within the first 50 hPa below the lidar cloud top. A further subdivision into latitude bands reveals similar distributions for extra-tropical and tropical regions (not shown).

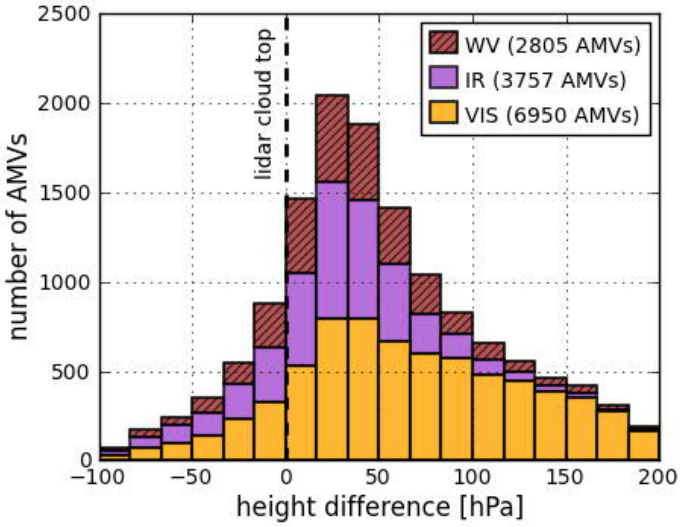

Figure 22: *Histogram of height differences (hPa) between original AMV pressure heights and lidar cloud-top heights for high-level and low-level AMVs combined. Positive values correspond to AMV heights that are below the respective lidar cloud top.*

4.2.1 VRMS differences and wind speed bias

Figure 23 shows the VRMS difference (upper panels) and wind speed bias (lower panels) between operational AMV winds and layer-averaged FG model winds. Red dashed lines represent layers that are centered at the operational AMV height, which serve as a reference for the reassigned layers relative to the lidar cloud-top height (green lines). High-level AMVs above a pressure height of 400 hPa comprise WV and IR AMVs, whereas low-level AMVs below a pressure height of 700 hPa consist mainly of VIS and IR AMVs. Dividing the AMVs data set for different channels used for their derivation shows similar results for both high-level and low-level AMVs and is therefore not shown.

For high-level AMVs (Fig. 23a), lowest VRMS differences are achieved for 120-hPa layers below the lidar cloud top, resulting in a relative VRMS reduction of about 10% when compared to reference layers of the same depth centered at the original AMV height (red dashed line) and of about 15% when compared to the discrete operational AMV heights (red dashed line at 0 hPa). The wind speed bias tends to be close to zero for 100-hPa layers below the lidar cloud top. Low-level AMVs (Fig. 23b) show lowest VRMS differences for 120-hPa

layers below the lidar cloud top and for 200-hPa layers with 25% above and 75% below the lidar cloud top. For 120-hPa layers below the lidar cloud top, the reassignment reduces the VRMS difference by 8 % and 15% compared to reference layers centered at the operational AMV height and to the discrete operational height, respectively. The wind speed bias is generally small for low-level AMVs, but layers below the lidar cloud top exhibit slightly smaller values than the 25%/75% layers.

In order to investigate the effect for different latitude bands, the AMV sample used in Figure 23 is subdivided in extra-tropical and tropical regions in Figure 24. Generally,

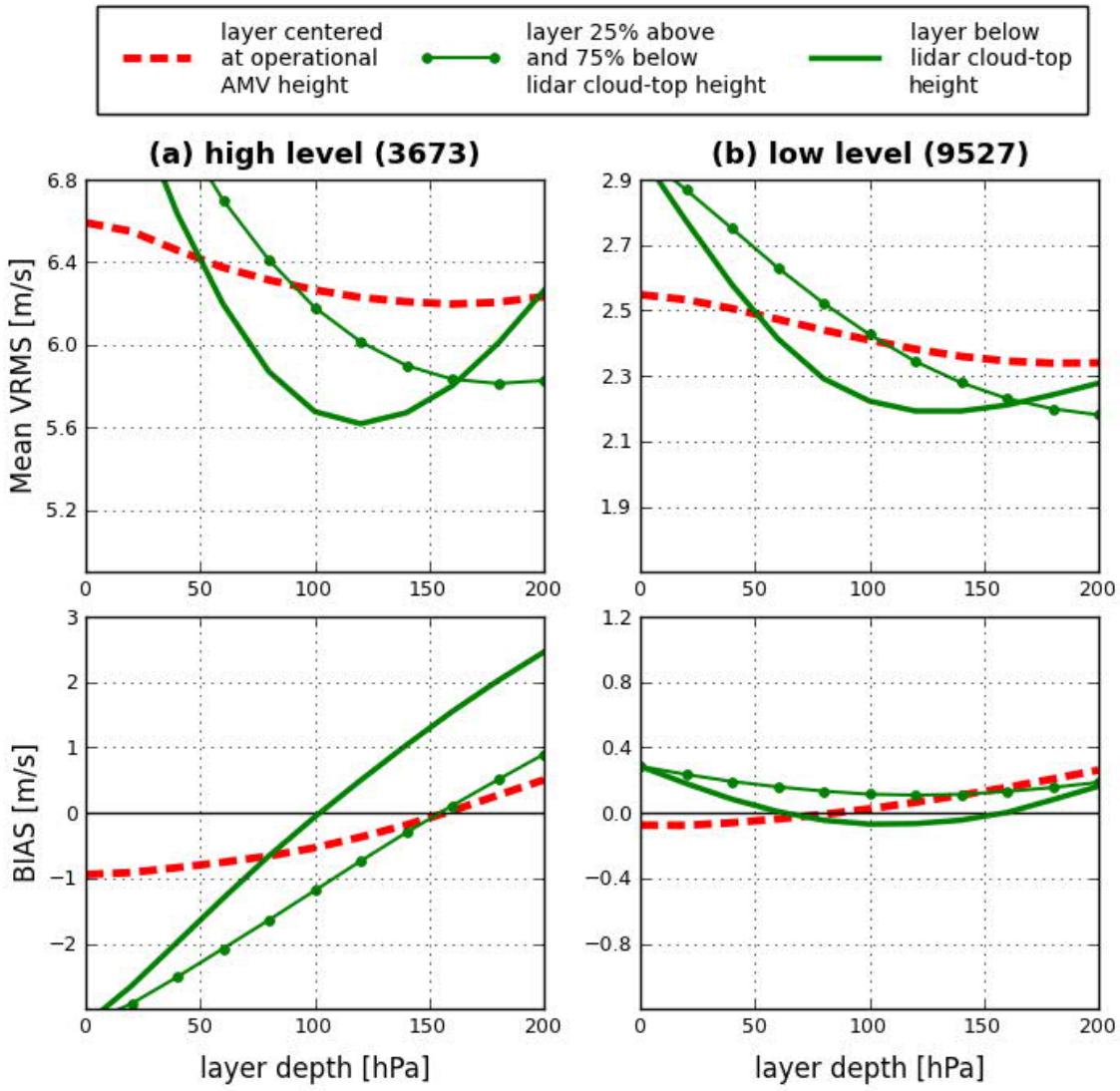

Figure 23: *Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV winds and layer-averaged model winds for (a) high-level and (b) low-level Meteosat-10-AMVs. Numbers in brackets are AMV counts. Red dashed lines represent layers centered at the original AMV pressure height; green solid lines represent layers below the lidar cloud-top height; green dotted lines represent layers with 25% above and 75% below the lidar cloud-top height (cf. legend).*

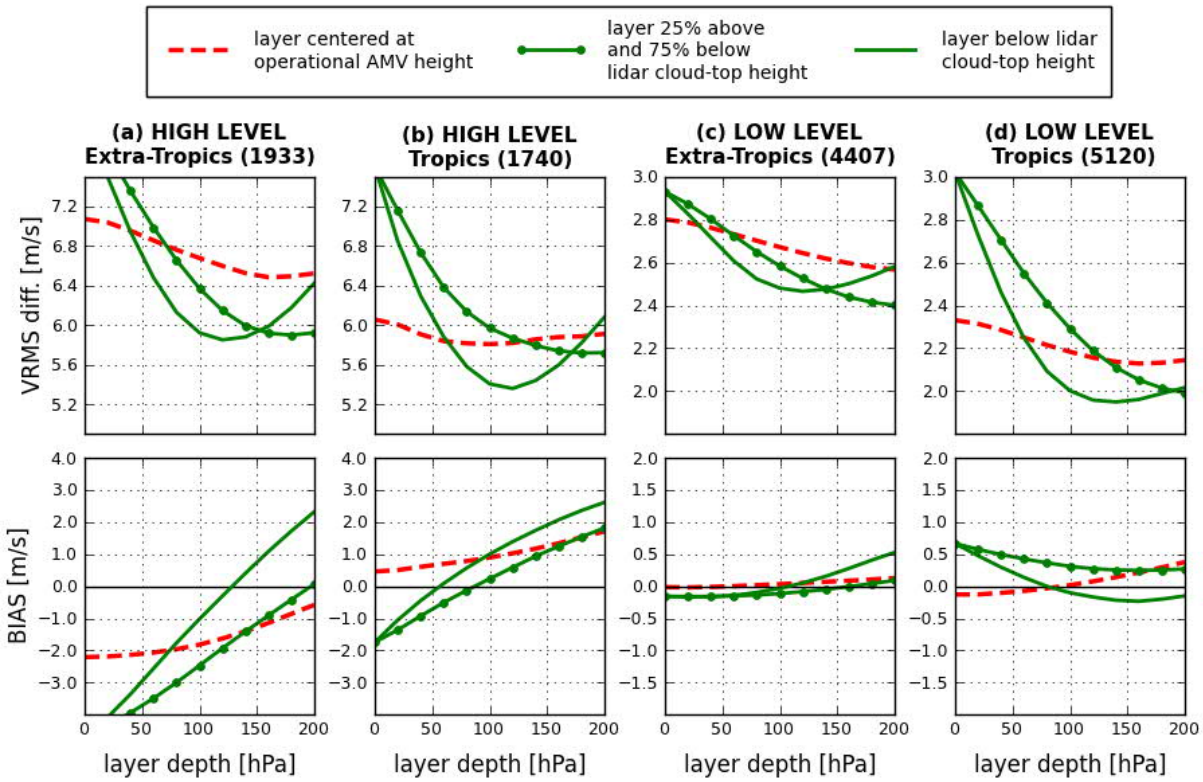

Figure 24: *Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV winds from Meteosat-10 and layer-averaged FG model winds for (a) high-level AMVs in extra-tropical regions, (b) high-level AMVs in the tropics, (c) low-level AMVs in extra-tropical regions and (d) low-level AMVs in the tropics. Numbers in brackets are AMV counts. Red dashed lines represent layers centered at the original AMV pressure height; green solid lines represent layers below the lidar cloud-top height; green dotted lines represent layers with 25% above and 75% below the lidar cloud-top height (cf. legend). Note that the scales for high-level and low-level AMVs are different for VRMS differences and the wind speed bias.*

120-hPa layers below the lidar cloud top achieve lowest VRMS differences for high-level and low-level AMVs in both regions (Fig. 24a – 24d, upper panels). For high-level AMVs in extra-tropical regions (Fig. 24a, lower panel), the wind speed bias is close to zero for 120-hPa layers below the lidar cloud top and thus coincides with lowest VRMS differences. The wind speed bias in the tropics shows larger values for the 120-hPa layer (Fig. 24b, lower panel), but still has about the same magnitude as the wind speed bias for layers of the same depth at the operational AMV height. As GME has some known shortcomings in high-level tropical regions due to a relatively poor convection scheme, the tropical wind speed bias may also result from the model error in that region. Wind speed bias values for low-level AMVs (Fig. 24c and 24d, lower panel) are generally small for extra-tropical and tropical regions.

4.2.2 AMV error correlations

To investigate the effect of the CALIPSO-based height reassignment on the horizontal correlation of AMV errors, correlation coefficients for the difference between AMVs and FG model equivalents (FG departures) are computed as a function of horizontal distance (Fig. 25) as described in Section 3.2.1.2. The red, dashed line in Figure 25 corresponds to the error correlation of FG departures for AMVs at the operational pressure height, showing a decrease of the correlation with increasing horizontal distance. The green line represents values for the error correlation of FG departures of lidar-corrected AMVs (using 120-hPa deep layers below the lidar cloud-top height) to operational AMVs. The lidar-corrected data set shows significantly lower error correlation values compared to the operational AMVs, which further emphasizes the potential benefit of lidar-corrected AMVs for data assimilation. On average, the lidar height reassignment reduces the correlation by about 50 km. As the evaluation is conducted with model equivalents, the absolute values of the correlation are

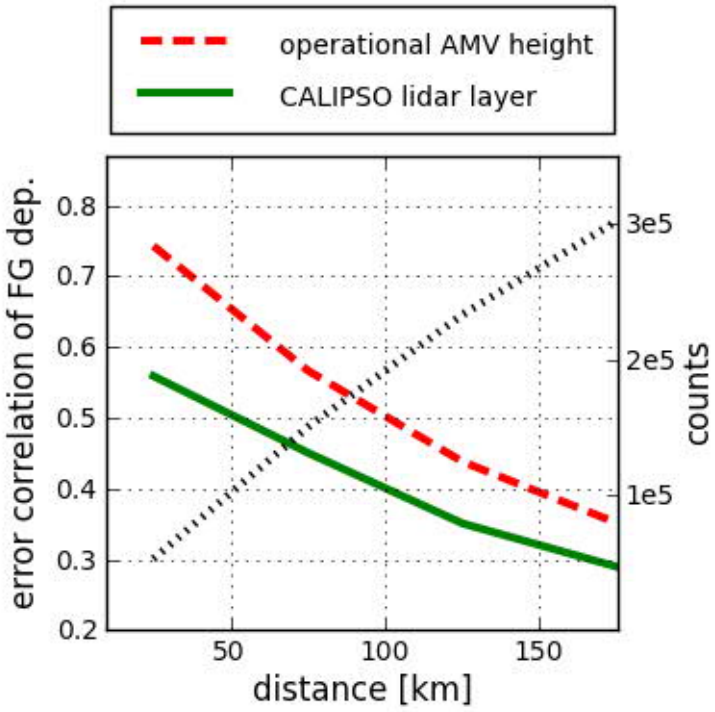

Figure 25: *Horizontal AMV error correlation of FG departures of operational AMVs (dashed red line) and lidar-corrected layer-averaged AMVs (green line) when compared to surrounding operational AMVs for Meteosat-10 from all levels as a function of horizontal distance between the AMVs. The black dotted line shows the number of collocations used and corresponds to the y-axis on the right.*

strongly affected by the horizontal correlation of model errors and should not be evaluated. However, the model correlation can be expected to be independent of the AMV errors and therefore results for the corrected and uncorrected AMV data set can be compared. Bormann et al. (2003) calculated the error correlation of AMVs using radiosondes for verification and found correlation values that were about 30% smaller than the values in Fig. 25. This difference is likely the contribution of the model correlation.

4.3 Summary and discussion

In this section, satellite lidar observations are used to directly correct the height of AMVs from MSG satellite imagery with lidar cloud-top observations from CALIPSO. Appropriate layer depths and layer positions relative to the lidar cloud top and relative to the original AMV height are investigated by comparing AMV winds to radiosonde or GME model winds averaged over layers of the respective depth and position.

For the verification with radiosonde data, 220 days of data with altogether about 4500 collocated AMVs, CALIPSO observations and radiosondes are analysed. Assigning upper-level AMVs to 120 hPa layers below the lidar cloud top leads to an improvement of ~12% compared to assigning layers of the same depth centered at the original AMV heights and of ~17% compared to using the discrete original AMV heights. Similar results are yielded for 200 hPa layers with 25% of the layer above and 75% below the lidar cloud top. The reduction of VRMS differences for AMVs below pressure levels of 700 hPa is less distinct when layers relative to the lidar cloud top are used instead of layers relative to the originally assigned AMV height. Although there is only a slight VRMS reduction for these AMVs using lidar information, there is indication that lidar observations can reduce VRMS differences in periods with lower AMV quality due to changes in the AMV processing. The reasons why the lidar height correction is showing much better results for upper level AMVs may be connected to the relation of low-level AMVs to cloud-base winds and the inability of satellite lidars to observe these cloud bases accurately.

A tighter threshold for the horizontal distance between AMVs and radiosondes used for verification even leads to a clearly larger relative effect of the direct height reassignment. The results imply that the direct height reassignment can actually reduce the AMV wind error by over 20% compared to assigning AMVs to layers relative to the original height and over 25%

compared to using the discrete original AMV height, but the sample size gets comparably small for a tight threshold.

In the second part of this section, the evaluation of the height reassignment of AMVs with directly collocated CALIPSO lidar observations is conducted with GME model equivalents instead of with radiosondes for an 11-day period with 13200 AMV/CALIPSO matches. For Meteosat-10, both high-level and low-level AMVs exhibit the lowest VRMS differences for assigning AMVs to 120-hPa deep layers below the lidar cloud top. This leads to a reduction of VRMS differences of 8-10% when compared to layers of the same depth centered at the operational AMV height, and about 15% when compared to the discrete operational AMV levels.

Overall, the results of the model evaluation of the direct height reassignment confirm the findings for the verification with radiosonde sounding data. For the radiosonde-evaluated upper-level AMVs that are mainly located over European continent, 120-hPa layers below the lidar cloud top yielded lowest VRMS differences and a wind speed bias close to zero when verified with radiosondes. These results closely coincide with the findings of the model evaluation for extra-tropical regions for both high-level and low-level AMVs. The different findings for low-level AMVs when verified with radiosonde observations are likely related to the selection of the evaluation period: The model evaluation uses data after the changeover to the CCC-method, which also switched off the application of the assignment of low-level AMVs to the estimated cloud-base. On the contrary, the radiosonde verification evaluated mostly AMVs before the CCC-method was introduced. Overall, the consistency of the findings for the model evaluation and the verification with radiosonde data implies that model error does not blur the results and emphasizes the validity of using short-term forecasts for the evaluation.

The positive impact of assigning AMVs to layers instead of discrete levels shown in this study coincides with findings of preceding studies (Velden and Bedka, 2009: Weissmann et al., 2013). Using a simulated model framework, Hernandez-Carrascal and Bormann (2014) illustrated that AMVs represent winds averaged over a cloud layer instead of the cloud-top or cloud-base level wind. Lean et al. (2015) also quantified height assignment AMV error characteristics using a set of simulated AMVs and found the closest fit of AMVs to layer-averaged model winds that are most commonly located below the estimated cloud-top. This corresponds well to the results presented in this study with lowest wind VRMS differences and wind speed bias values when assigning AMVs to layers below the lidar cloud-top height.

In addition, Lean et al. (2015) and Hernandez-Carrascal and Bormann (2014) stated that a part of the benefit that is gained by assigning AMVs to layer-averages over a cloud layer can be reached by using a discrete level positioned within the respective layer. This is also confirmed in the present study.

Overall, there are about 1000–1300 Meteosat AMVs with nearby CALIPSO observations that could be directly corrected with lidar information per day. About 3300-4500 operational Meteosat-10-AMVs are assimilated every 6 h in the current global forecasting model of DWD, leading to ~13200-18000 assimilated Meteosat-10-AMVs per day. Replacing the original AMV data set with lidar-corrected AMVs based on the direct height correction would lead to a large reduction of the already small number of assimilated AMVs. A more appropriate approach would be to include the lidar-corrected AMV data set in addition to the originally assimilated AMV data set, as both data sets generally do not overlap much. Hereby, the number of assimilated Meteosat-10-AMVs could be increased by 6% - 10%. However, the availability of real-time lidar data would be an essential prerequisite, which limits the operational application of the direct height correction with CALIPSO data for assimilation purposes.

5. HEIGHT BIAS CORRECTION AND DATA ASSIMILATION EXPERIMENTS

The previous section showed that a direct lidar-based height reassignment significantly reduces the VRMS difference and the wind speed bias of AMVs. However, this direct correction can only be applied to a small fraction of the AMV data set, where collocated CALIPSO observations are available for the individual reassignment. In addition, the need for real-time lidar data would pose a significant effort for data providers and NWP centres. To address this issue, a statistical lidar-based height bias correction of AMV pressure heights is evaluated in this section. After an introduction to the derivation of height bias correction functions, the potential of a height bias correction based on different training intervals is assessed using GME model equivalents. Subsequently, initial results of the forecast performance when lidar-corrected AMVs are assimilated in the global model ICON are presented.

5.1 Height bias correction functions

Results of a direct, CALIPSO-based height reassignment in the previous section identified 120-hPa deep layers below the lidar cloud top as an overall optimal configuration for the height reassignment. For this reason, such layers are used as the basis for deriving the height bias correction. As described in Section 3.2.2, three different options for the respective training interval are tested: 10 days, 30 days, and 30 days with a hemispheric/tropical subdivision. As an example, Figure 26 shows a 30-day height bias correction function (left panel) and a 10-day height bias correction function (right panel) for Meteosat-10 as a function of altitude and channel. The mean pressure of the optimal layer, meaning the discrete level at 60 hPa below the lidar cloud top, is used for the pressure adjustment values on the x-axis.

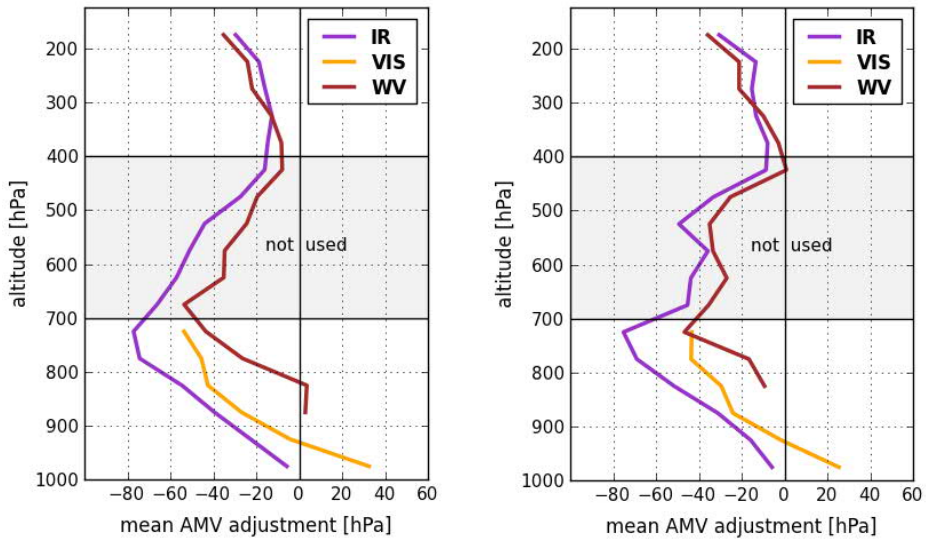

Figure 26: *Height bias correction functions for Meteosat-10 for a 30-day period (1 April 2013 – 6 May 2013, left panel) and a 10-day period (1 May 2013 – 10 May 2013, right panel) as a function of altitude. Different line styles indicate different satellite channels (cf. legend).*

Negative values indicate that the AMV is shifted downwards in the atmosphere. Mid-level AMVs between 400 hPa and 700 hPa are not used for the height bias correction because of the comparably small AMV sample size in this range. Typical AMV numbers for each vertical pressure bin are about 4000 (1300) for high levels and 2400 (840) for low level AMVs for the 30-day (10-day) correction. Generally, the shape of the curves for the 30-day height bias correction and the 10-day height bias correction is very similar (also for the other 10-day training periods), with a more jagged shape for the 10-day correction due to the smaller sample size. On average, the adjustment of high-level AMVs is of the order of - 20 hPa. Low-level AMVs are also shifted downwards at most altitude levels. The largest adjustment of 60-80 hPa occurs for AMVs with 700-800 hPa altitude. Generally, the curves of the height bias correction functions for the latitude subdivision (not shown) also tend to have a similar shape, with less pronounced height adjustment values for the tropics than for the extra-tropics.

5.2 Comparison of the direct height reassignment and the height bias correction

As described in Section 3.1.2, the direct height reassignment uses the actual collocated lidar observations for an individual height correction for each AMV. On the contrary, the height bias correction aims to correct systematic AMV height biases using statistical adjustments based on height bias correction functions derived from preceding training intervals. To investigate the potential benefit of the height bias correction, both lidar-based height correction methods are evaluated for the same sample of Meteosat-10-AMVs, i.e. all AMVs with a directly collocated CALIPSO lidar observation available during the second evaluation period (7 May – 12 May 2013).

Figure 27 shows the mean VRMS difference and wind speed bias between AMV and model winds for results for operational AMV heights (red bars), for applying the direct height reassignment (green bars), and for applying the three different height bias correction functions (blue bars). The left part of each panel shows the results when assigning AMVs to discrete levels, meaning the operational levels (red), levels at 60 hPa below the cloud top of a directly collocated lidar observation (green) or three “adjusted” levels based on the height bias correction (blue). Correspondingly, the right part represents layer-averaged values for 120-hPa deep layers centered at the respective heights. Results for the direct height reassignment are generally similar to the results presented in Section 4.2, with slight variations due to the different evaluation periods considered in the two sections.

For high-level AMVs (Fig. 27a), the lowest VRMS differences (upper panel) are achieved for 120-hPa layers based on the direct height reassignment. However, the height bias correction also yields a distinct reduction of VRMS differences compared to those from levels/layers relative to the operational AMV height. About 30-50% of the reduction of the direct height reassignment is achieved, with no clear preference for a particular correction function. In addition, the wind speed bias (lower panel) is clearly reduced for the height reassignment as well as for the height bias correction based on a 30-day mean and a 10-day mean when compared to the wind speed bias at the operational AMV height. Low-level AMVs (Fig. 27b) exhibit a similar pattern to high-level AMVs. Again, the direct height reassignment shows the best results when layer-averages are used. In addition, layers relative to the adjusted heights based on the height bias correction show a clear reduction of VRMS differences compared to

the operational values. In particular, the 10-day height bias correction exhibits VRMS differences that are almost as low as those of the direct height reassignment. The wind speed bias for low-level AMVs is strongly reduced for the direct height reassignment as well as for the height bias correction, especially when a layer-averaging is applied. Overall, VRMS differences are generally lower for layer-averages than for discrete levels, which further emphasizes the fact that AMVs represent the wind in a vertically extended layer.

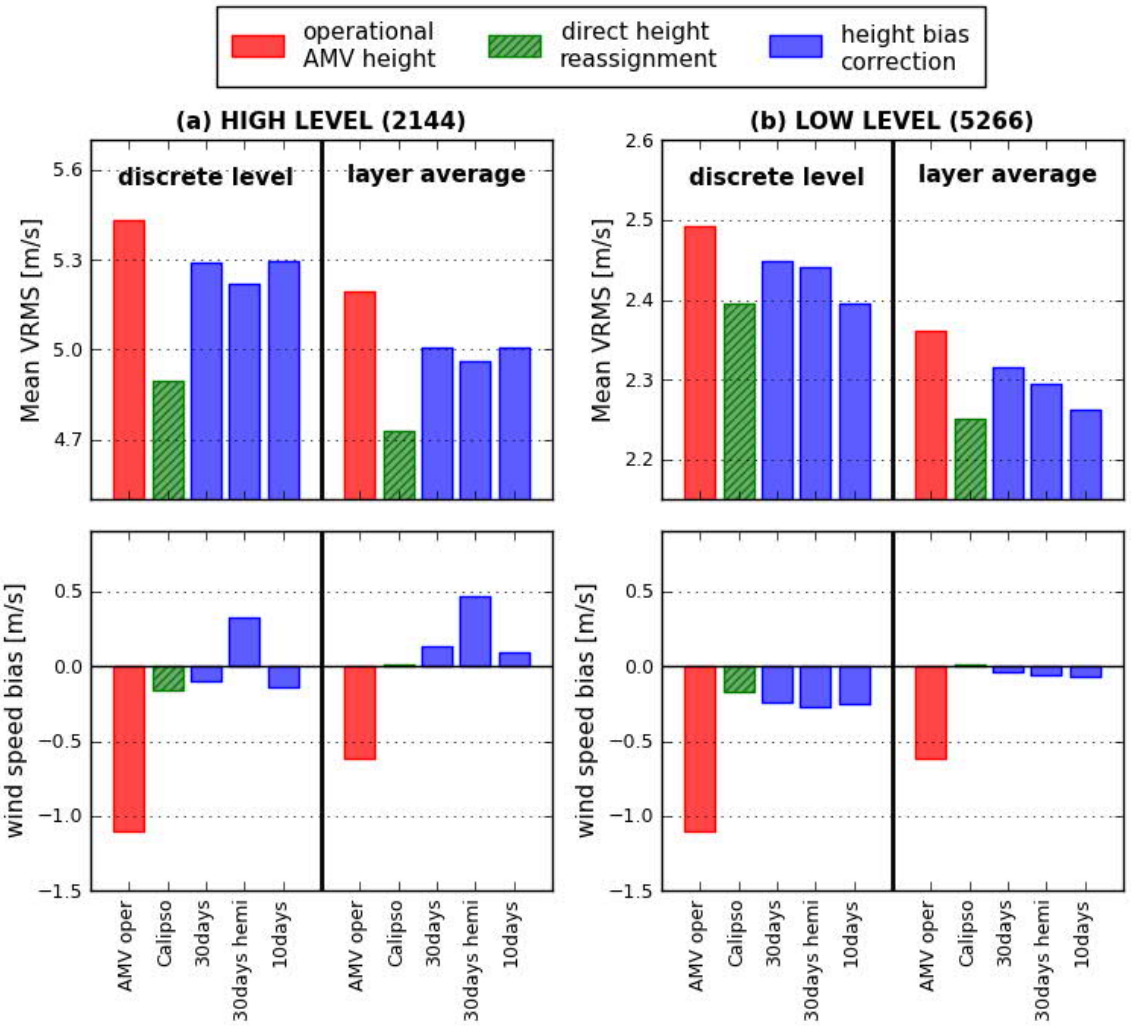

Figure 27*: Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV and model FG winds for (a) high-level and (b) low-level Meteosat-10-AMVs. Numbers in brackets are AMV counts. 'AMV oper' corresponds to the operational AMV height and 'CALIPSO' to the direct lidar height reassignment. Different height bias correction functions are designated as '30days' (30-day mean), '30days hemi' (30-day mean with hemispheric and tropical sub-divisions) and '10days' (10-day mean). Results are shown both for assigning AMVs to discrete levels (meaning the operational level, the level at 60 hPa below the actual lidar cloud top and the three height-bias- corrected levels) in the left part of each panel and to 120-hPa layer-averages centered at these levels in the right part of each panel.*

The subdivision into extra-tropical and tropical regions for the height bias correction is investigated in Figure 28 for both high-level and low-level AMVs. For clarity, each subfigure shows only results for the operational, discrete AMV level (red bar on the left of each panel) and for 120-hPa layer-averages centered at the operational level and at lidar-corrected levels (right part of each panel), omitting results for the lidar-based height correction for discrete levels. The general tendency seen in Figure 27 is reflected in extra-tropical as well as tropical regions: The direct height reassignment produces the lowest VRMS differences and a small wind speed bias in all subfigures (Fig. 28a – 28d, green bars), but the height bias correction on average achieves about 50% of this reduction. As pointed out earlier, high-level AMVs in the tropics tend to have a stronger wind speed bias than in the extra-tropics, which may be

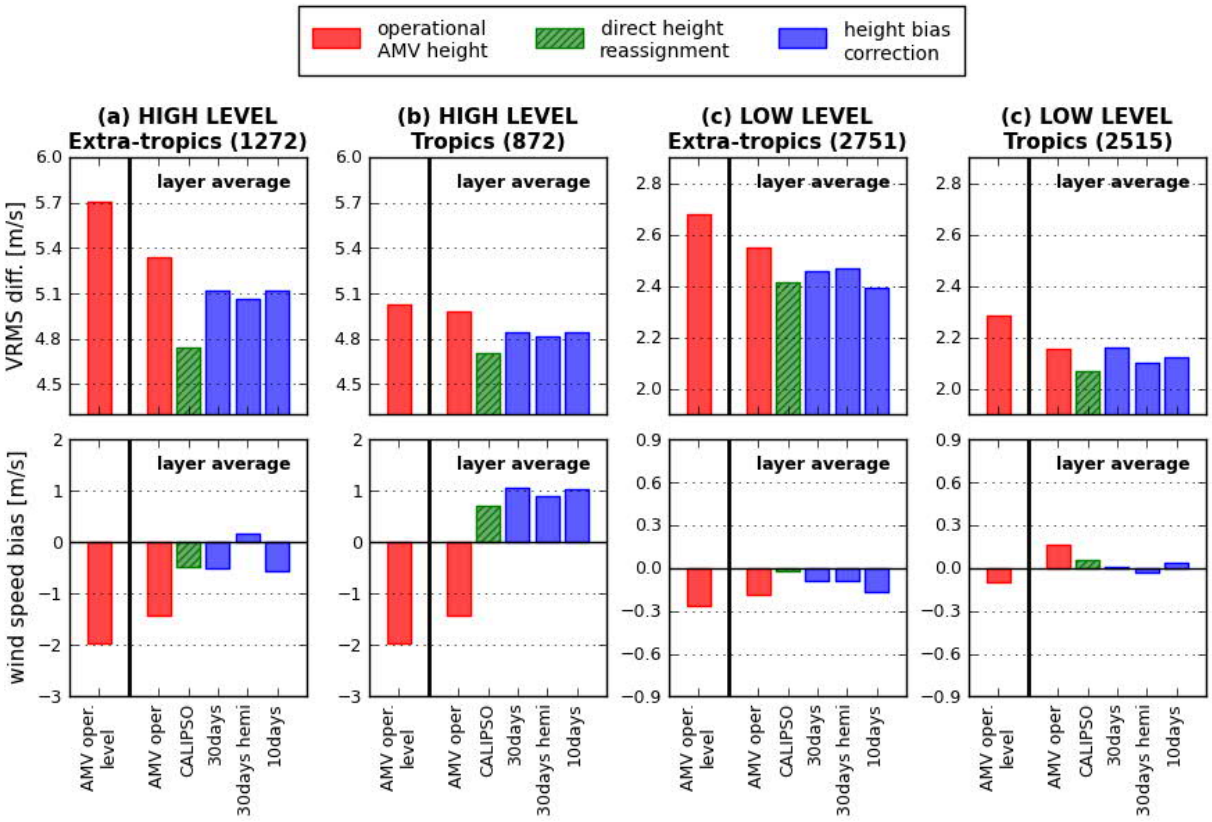

Figure 28: *Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV winds from Meteosat-10 and layer-averaged FG model winds for (a) high-level AMVs in extra-tropical regions, (b) high-level AMVs in the tropics, (c) low-level AMVs in extra-tropical regions and (d) low-level AMVs in the tropics. Numbers in brackets are AMV counts. 'AMV oper level' corresponds to the discrete operational AMV height. The right part of each panel shows results for 120-hPa layer-averages centred at the operational AMV height ('AMV oper'), below the actual lidar cloud top ('CALIPSO') and centered at the levels based on the three height bias correction functions ('30days', '30days hemi', '10days').*

related to an inadequate representation of the atmospheric state by GME. Nevertheless, the application of both height correction methods in the tropics overall leads to a smaller bias than the operational value.

To investigate the effect of different layer depths and level positions relative to the lidar cloud-top height, Figure 29 shows the relative reduction of VRMS differences of both lidar height correction methods (solid lines for the direct height reassignment and dashed lines for a 30-day height bias correction) for discrete levels and different layer depths when their results are compared directly to the results for the discrete operational AMV heights. The relative reduction of the VRMS difference is shown as a function of layer depth for all latitudes for low-level and high-level AMVs combined. Overall, the best results are achieved using the direct lidar height reassignment for 120-hPa layers below the lidar cloud top (solid green line) with a VRMS reduction of about 11% compared to the operational AMV heights. Again, this value deviates slightly from the reduction of VRMS differences of ~15% found in

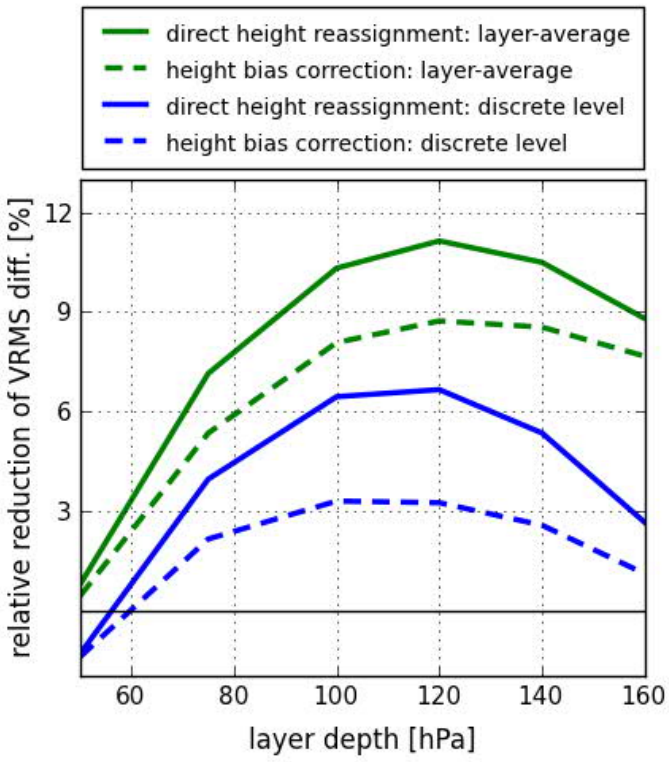

Figure 29: *Relative reduction of VRMS differences between AMV and model FG winds for assigning AMVs to layers/levels below the lidar cloud top (solid lines) and to layers/levels based on the 30-day height bias correction (dashed lines) instead of the discrete operational AMV heights. Green lines represent layer-averages and blue lines discrete levels relative to the respective height. Low-level and high-level AMVs are combined. The x-axis denotes the vertical depth of the layers. The reassigned levels are located at the mean pressure of the layers.*

Section 4.2, as two different evaluation periods are considered in the two sections. Using the height bias correction, the largest reduction of VRMS differences (~9%) is also achieved for 120-hPa layers (dashed green line), reaching about 80% of the VRMS reduction that is achieved with the direct height reassignment. The impact from using discrete levels below the lidar cloud top (solid blue line) is less distinct. The largest error reduction (about 7%) is achieved for levels at 50-60 hPa (drawn in the figure at 100-120 hPa) below the lidar cloud top. Results for the discrete pressure heights that are based on the 30-day height bias correction are least pronounced (dashed blue line) and only show a slightly positive effect (3.5% reduction).

5.3 Assimilation of lidar-corrected AMVs in ICON

As demonstrated in the preceding chapter, the height bias correction is capable of reducing the VRMS difference and the wind speed bias of Meteosat-10-AMVs, and may therefore provide an efficient way of using lidar information for the AMV height correction in NWP systems. To investigate this hypothesis, the impact of lidar-corrected and layer-averaged Meteosat-10-AMVs is assessed in assimilation and forecast experiments with the global NWP model ICON of DWD for a 16-day assimilation period (1 May 16 May 2013). Thereby, the forecast skill of 3-h short-term forecasts (FG field) as well as free forecasts up to 7 days is evaluated. As described in Section 0, three experiments are conducted: Firstly, all AMVs are assimilated on the discrete level of their original AMV height, serving as a reference (REF). Secondly, all Meteosat-10-AMVs are reassigned to lidar-corrected heights based on a 15-day training period (LEV). Thirdly, all Meteosat-10-AMVs are assimilated as 120-hPa layer-averages centered at the lidar-corrected height (LYR).

5.3.1 Evaluation of 3-h short-term forecasts (FG)

First, the forecast impact of observations is evaluated in the context of FG departures of the AMV wind. Figure 30 shows the VRMS difference and wind speed bias between AMV and FG model winds for the three experiments REF (red solid), LEV (green dashed) and LYR (green solid) for the 16-day assimilation period (127 assimilation cycles at 3-h intervals). The VRMS differences (upper panels) show a consistently better agreement

with the FG model equivalents throughout the troposphere when AMV are assigned to 120-hPa layers centered at the lidar-corrected height (LYR) instead of assigning them to the original discrete height (REF) both for extra-tropical (Fig. 30a) and tropical (Fig. 30b) regions. For the experiment LEV, the reduction of VRMS differences compared to the reference run is not as distinct as for the experiment LYR, but it does achieve either equal or lower VRMS differences than the experiment REF for extra-tropical regions. In the tropics, LEV results are generally also positive, but show a slight degradation of VRMS differences compared to REF in some upper-level regions. By integrating over the whole troposphere, the relative reduction of the VRMS differences from the experiment LYR compared to the reference run is 3.7% for the tropics and 4.0% for extra-tropical regions (1.5% and 0.1% for the corresponding values for the experiment LEV).

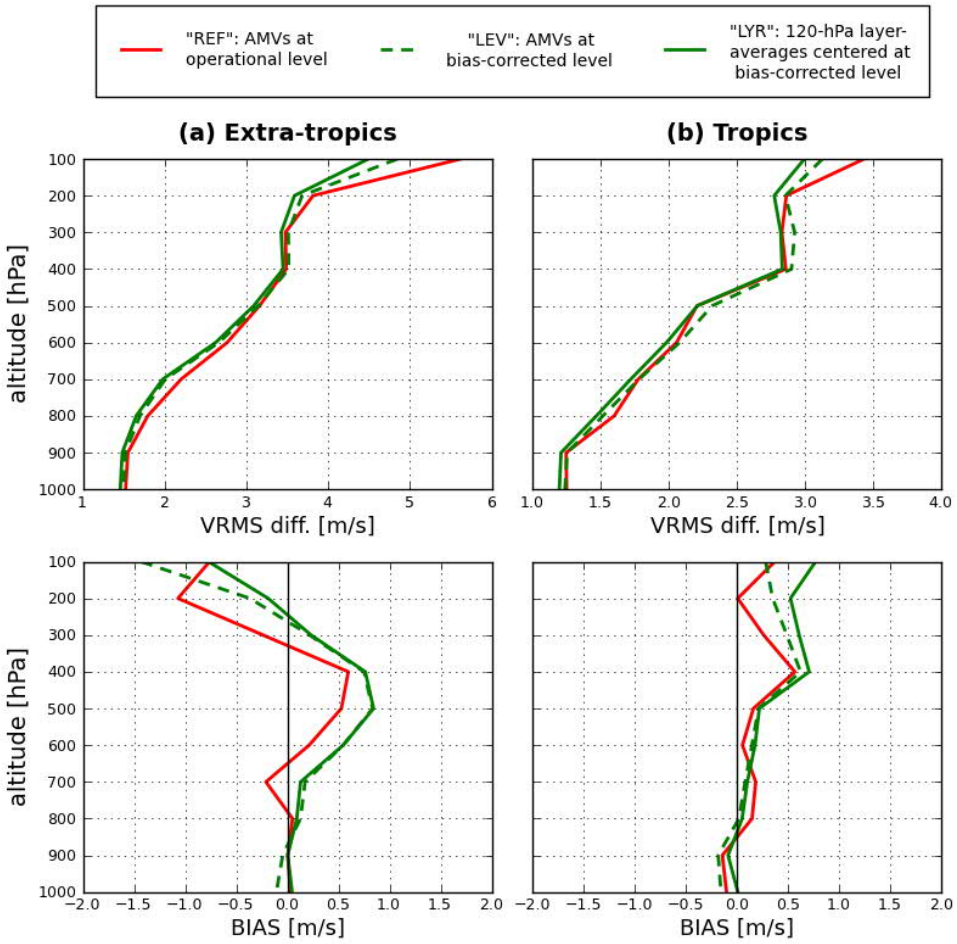

Figure 30: *VRMS differences (upper panels) and wind speed bias (lower panels) between AMV wind and corresponding FG model equivalent as a function of altitude for the three experiments REF, LEV and LYR (a) for extra-tropical and (b) for tropical regions for the 16-day assimilation period (1 May 2013 - 16 May 2013).*

The wind speed bias between AMV and model wind (Fig. 30, lower panels) shows absolute values for the experiments LEV and LYR that are of equal or lower magnitude than the operational values for low-level and high-level regions in the extra-tropics and low-level and mid-level regions in the tropics. In contrast, an increase of the wind speed bias is observed for extra-tropical regions for altitude regions from 300 hPa to 700 hPa and for tropical regions above 600 hPa. However, this increase might be related to a suboptimal model representation of the true atmospheric state, as the magnitude of the wind speed bias has the same order as long-term monitoring model bias values (not shown). Consequently, this increase may be related to a model bias instead of an observational bias.

To further investigate this hypothesis, the verification of the 3-h short-term forecasts (FG fields) against radiosonde observations is shown in Figure 31. For clarity, only results for

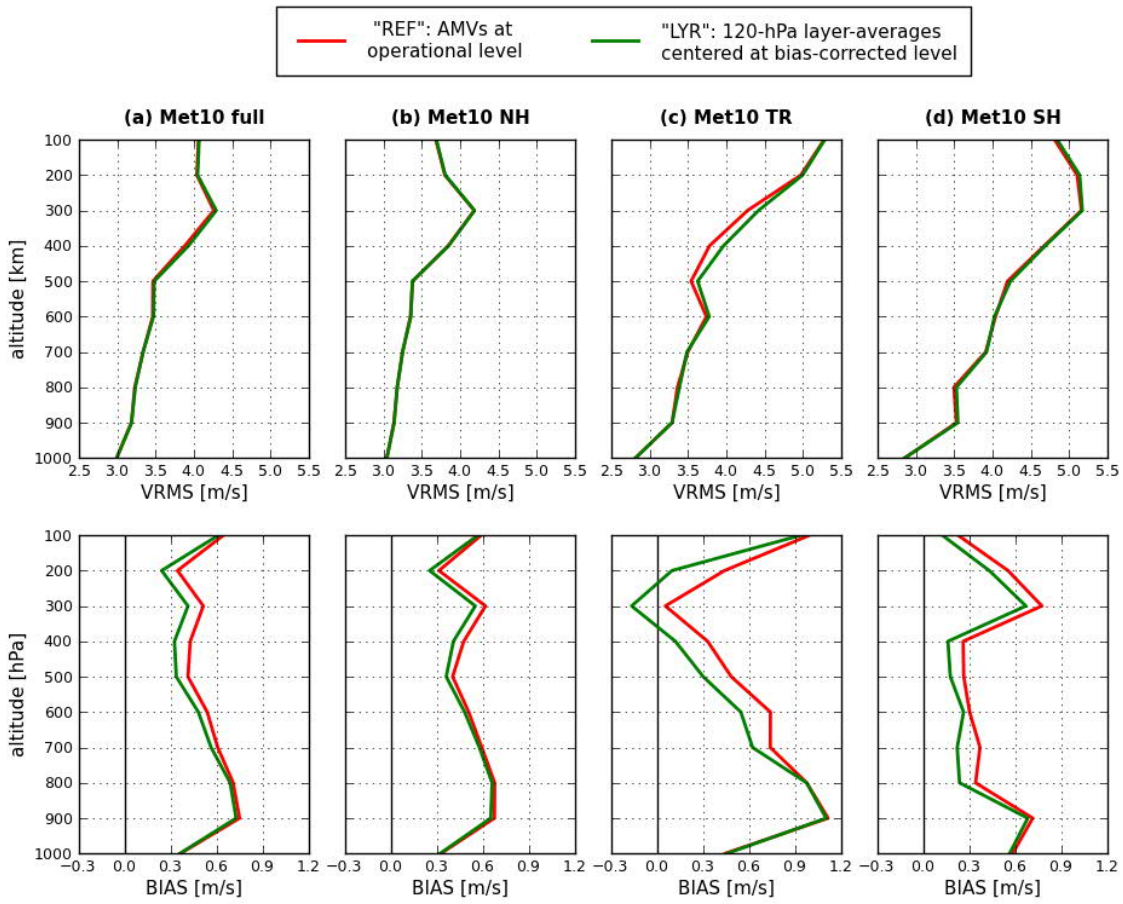

Figure 31: *VRMS differences (upper panels) and wind speed bias (lower panels) between radiosonde winds and FG model winds for the experiments REF(red lines) and LYR (green lines) for (a) the full Meteosat-10 domain (65° N/S/W/E), (b) its northern hemispheric part, (c) its tropical part and (d) its southern hemispheric part for the 16-day assimilation period (1 May – 16 May 2013).*

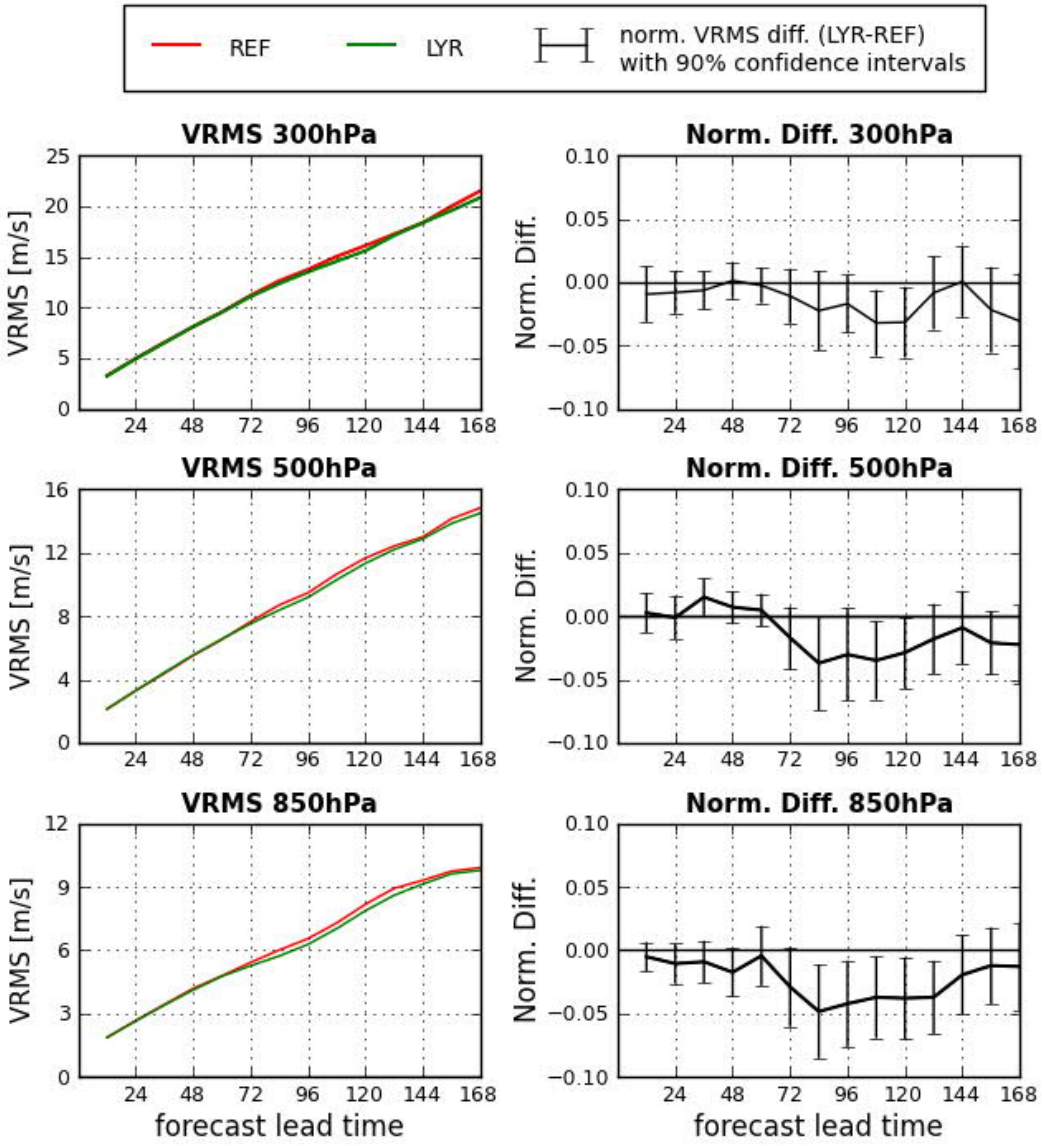

Figure 32: *VRMS (left panels) and normalized VRMS differences between the experiments LYR minus REF (right panels) as a function of forecast lead time at high levels (300 hPa, upper panels), mid levels (500 hPa, mid panels) and low levels (850 hPa, lower panels) for the evaluation period from 1 May 2013 – 9 May 2013*

the experiments REF and LYR are depicted. As the impact could be different due to a coarser or finer radiosonde observation network, a sub-division for the northern hemisphere, the tropics and the southern hemisphere is made. Upper panels show results for VRMS differences, and lower panels show wind speed bias values as a function of altitude for the full Meteosat-10-domain (Fig. 31a) and for the additional subdivision into its northern hemispheric part (Fig. 31b), its tropical part (Fig, 31c) and its southern hemispheric part (Fig. 31d). VRMS differences shown in the upper panels of Figure 31 do not exhibit clear

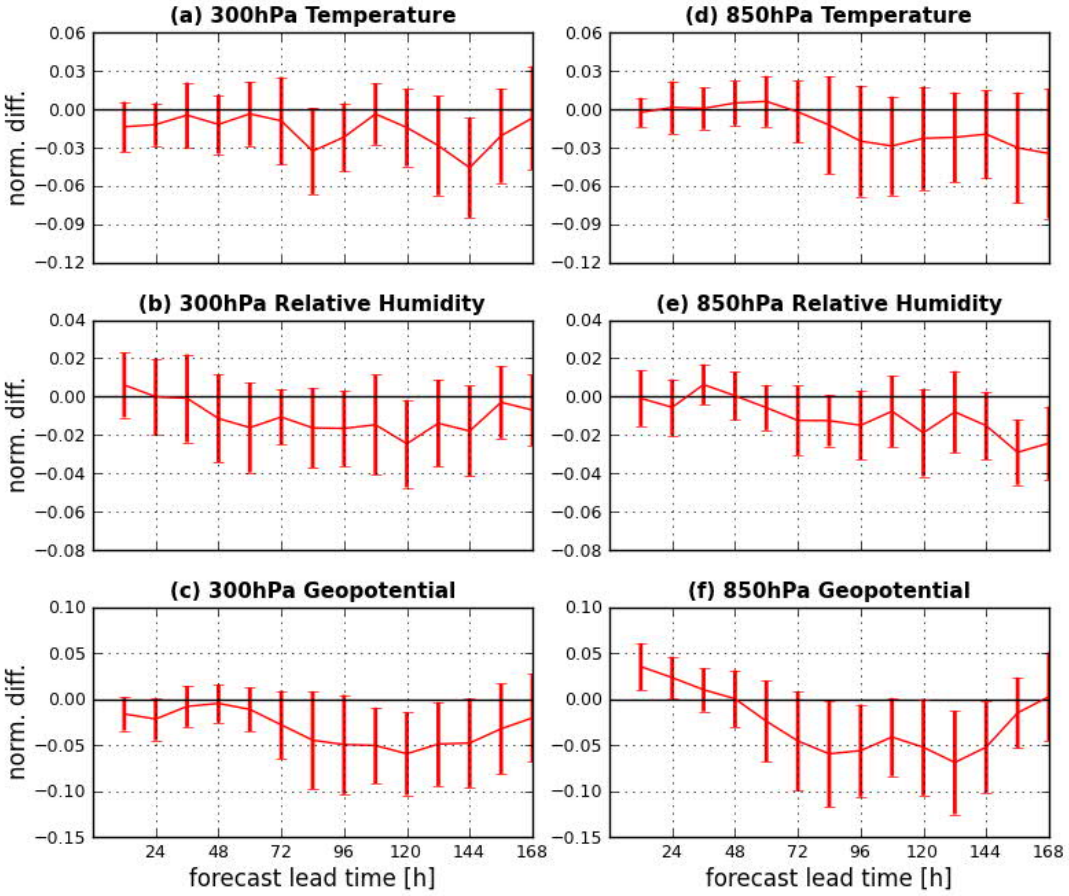

Figure 33: *Normalized RMS differences (LYR - REF) as a function of forecast lead time for the prognostic variables (a,d) temperature, (b,e) geopotential, (c,f) relative humidity for the evaluation period from 1 May 2013 – 9 May 2013. Left panels are for high levels (300 hPa), right panels for low levels (850 hPa)*

positive or negative effects when the lidar-corrected and layer-averaged AMVs are assimilated. However, a slight degradation compared to the reference run is seen, especially for tropical regions between 300 hPa and 600 hPa. Generally, the relative degradation when integrated over all altitude levels is rather small and reaches values of -0.1% for the northern hemisphere, -1.3% for the tropics and -0.4% for the southern hemisphere.

In contrast, a remarkable positive impact on the wind speed bias (lower panels of Fig. 31) can be seen throughout the atmosphere for the experiment LYR with respect to the reference run (REF). The effect is more apparent in the tropics and the southern hemisphere. Likely, this is related to the smaller amount of in-situ wind observations that are available for the

assimilation in global NWP models. On the contrary, a dense radiosonde network over Europe allows for the assimilation of frequent radiosonde observations, generally leading to comparatively well-defined wind conditions in that region. This usually results in a smaller impact from new observational data. Integrating over all altitude levels, the relative reduction of the wind speed bias compared to the reference run results in 6.7% for the northern hemisphere, 20.9% in the tropical region and 19.7% in the southern hemisphere of the Meteosat-10-domain.

5.3.2 Evaluation of 7-day forecast runs

The effect of assimilating lidar-corrected and layer-averaged AMVs on free forecasts is displayed in Figure 32. The left panels show the VRMS for the two experiments LYR (green lines) and REF (red lines) as a function of forecast lead-time up to seven days (168 hours) when evaluated with the analysis field of the 3-h cycling. The three different pressure levels at 300 hPa (upper panel), 500 hPa (middle panel) and 850 hPa (bottom panel) are representative of high levels, mid levels and low levels, respectively. Generally, the experiment LYR outperforms the reference run for almost all forecast lead times. As the differences between the two curves are small, the corresponding panels on the right hand side show the normalized VRMS differences of the two experiments (LYR minus REF) with 90% confidence intervals. Negative values indicate a positive effect from assimilating lidar-corrected and layer-averaged AMVs. Generally, the forecast impact is neutral for the first 2-3 forecast days for all three altitude levels shown. However, later forecast times reveal a positive (albeit mostly not significant) effect.

This tendency can also be seen when regarding other prognostic variables, as demonstrated in Figure 33. Normalized RMS differences for high levels (300 hPa, left panels) and low levels (850 hPa, right panels) for the temperature (upper panels), relative humidity (mid panels) and geopotential (lower panels) are shown as a function of forecast lead-time for the experiment LYR with respect to the reference run REF. Again, the first 2-3 forecast days exhibit a mostly neutral or only slightly positive effect, which is enhanced for longer forecast lead times.

5.4 Summary and discussion

In this section, height bias corrections functions are calculated based on the statistics of the differences between operational AMV height and a position relative to the lidar cloud-top height. Different lengths of the training period and settings for deriving the height bias correction are tested and the resulting corrections are then applied to a subsequent evaluation period. Overall, this adjustment of AMV pressure heights of Meteosat-10 leads to lower VRMS differences and a lower wind speed bias compared to using the operational AMV heights, with no clear preference for the duration of the training period. On average, the reduction is about 40-50% of the reduction by the direct reassignment, but has the clear advantage that all AMVs from a geostationary satellite can be corrected without the need for directly collocated lidar observations. In accordance with the results from the direct height reassignment in Section 4, it is again confirmed that AMVs are rather represented by a layer-averaged wind instead of the wind at a discrete level: Assigning AMVs to 120-hPa deep vertical layers based on a 30-day height bias correction leads to a reduction of VRMS differences that is three times larger than when assigning them to discrete levels at the mean pressure of the respective layers.

Initial assimilation and forecast experiments with lidar-corrected plus layer-averaged Meteosat-10-AMVs are performed with the global model ICON. Results indicate that the assimilation works well in terms of both the FG fit to radiosonde observations and forecast performance. Generally, the positive impact is largest in tropical and southern hemispheric regions of the Meteosat-10-domain, as the dense coverage of conventional observations in the northern hemisphere constrains the analysis more strongly. In particular, the wind speed of the FG fields decreases by up to 20% when verified against radiosonde soundings.

In addition, the positive impact of assimilating lidar-corrected and layer-averaged AMVs results in an improved accuracy of the 7-day forecasts for the wind field as well as for other prognostic variables when the forecast field is compared to the three-hourly cycled analysis fields from the same experiment. However, it should be noted that the duration of the evaluation period is comparatively short and a longer experiment is needed to achieve significant results.

Generally, an alternative strategy to improve the representation of AMVs in NWP systems would be either assigning AMVs to a lidar-corrected discrete height level, or assigning AMVs to a layer-average relative to the operational AMV height without incorporating lidar

information. Both approaches may achieve a part of the benefit of the proposed lidar-based height bias correction. However, the optimal position and depth of a layer relative to the operationally derived AMV height may be highly situation-dependent and therefore difficult to determine. Assigning AMVs to a lidar-corrected level based on a statistical height bias correction (without applying a layer-averaging) may be a compromise for current data assimilation systems as this is easier to implement in NWP systems than an observation operator for layer-averaged wind observations. However, its benefit tends be relatively small compared to what can be gained by using a layer-approach.

6. POTENTIAL FUTURE APPLICATIONS

6.1 Lidar-based AMV height correction for other geostationary satellites

After demonstrating the benefit of the direct height reassignment and the height bias correction for Meteosat-10-AMVs, these lidar-based height correction methods can also be applied to AMVs from other geostationary satellites. In the following, both methods are tested for AMVs from the GOES-satellites and MTSAT-2. The evaluation periods correspond to the ones used for Meteosat-10: Results are shown for the direct height reassignment applied during the first evaluation period and for the 30-day height bias correction applied during the second evaluation period.

AMVs from GOES or MTSAT-2 imagery are derived by different institutions than MSG-AMVs, using different processing algorithms and quality control procedures. Furthermore, the AMV height assignment is aided by different model fields and may therefore show different characteristics than Meteosat-10 for lidar-based height correction methods. Based on sensitivity studies using different settings for the direct height reassignment and height bias correction, the collocation criteria for GOES- and MTSAT-2-AMVs are tightened compared to the ones used for Meteosat-10. First, the AMV quality index threshold is raised from 50 to 80. In addition, AMVs are only used if the CALIPSO flight path approaches the AMV position to less than 10 km, which corresponds to an average distance of ca. 25 km between the AMV and the available lidar cloud-top observations within the 50-km radius that is used for the calculation of the median lidar cloud-top height. While these stricter collocation criteria improve the results for GOES-AMVs and MTSAT-2-AMVs, the results for the Meteosat-10 dataset are fairly independent of the applied criteria. This is illustrated in Figure 34, which shows the reduction of VRMS differences when results for layers below the lidar cloud top are compared to results of layers of the same depth centered at the original AMV height for Meteosat-10 during the

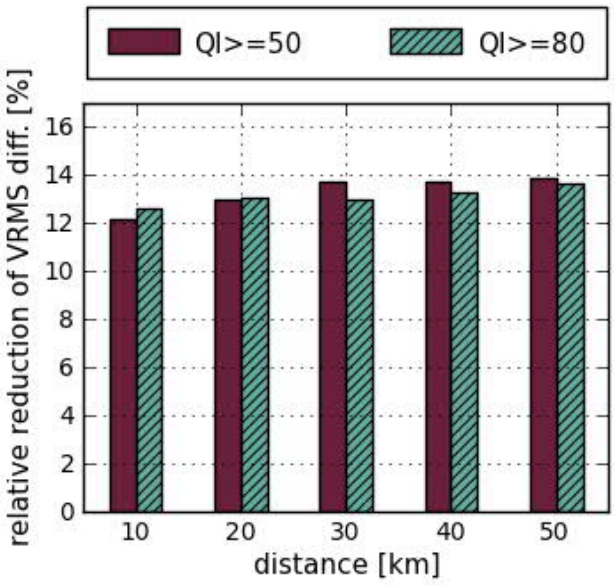

Figure 34: *Relative reduction of VRMS differences for Meteosat-10-AMVs when AMVs are assigned to 120-hPa layers below the lidar cloud top and compared to the operational AMV height. The x-axis denotes different thresholds for the maximum horizontal distance between corresponding AMV and CALIPSO observations. Dark purple bars represent results fo r a QI greater than 50, turquoise hatched bars for a QI greater than 80.*

first evaluation period. Hereby, a QI threshold of 50 (dark purple) and of 80 (turquoise) both lead to a relatively constant reduction of VRMS differences by about 12%-14% when constraining the horizontal distance. A subdivision for tropical and extra-tropical regions also exhibits similar results (not shown).

Given that the findings for the two GOES-satellites show very similar characteristics, results for GOES-East and GOES-West are combined in the following. In contrast to the evaluation with MSG-AMV, only a 30-day average is calculated for the height bias correction due to the smaller number of available AMVs. For GOES and MTSAT-2, there are considerably fewer available AMVs than for Meteosat-10. Table 11 lists the numbers of all used AMVs with collocated lidar observations for both periods. Only about 10% of the number of Meteosat-10-AMVs is found for MTSAT-2 and GOES-AMVs. This is due to the smaller number of operationally available AMVs from these satellites (GOES-AMVs are only available in 3-h intervals for the time of the study period compared to hourly MSG-AMVs or MTSAT-2-AMVs), as well as due to the stricter collocation criteria (higher QI threshold and smaller horizontal distance) applied for the height reassignment and height bias correction for AMVs from these satellites.

		Meteosat-10	MTSAT-2	GOES (East and West comb.)
first evaluation period (31 May – 11 June 2013)	high level	3673	621	226
	low level	9527	487	1192
second evaluation period (7 May – 12 May 2013)	high level	2144	295	143
	low level	5266	278	698

Table 11: *Number of AMVs with collocated lidar observations used for both evaluation periods*

6.1.1 Direct height reassignment

The mean VRMS difference and wind speed bias for the direct height reassignment for GOES-AMVs and MTSAT-2-AMVs is illustrated in Figure 35. For GOES high-level AMVs (Fig. 35a), 100-120 hPa layers below the lidar cloud top show a small benefit over the operational AMV heights, with about 3% relative reduction of VRMS differences compared to layers of the same depth centered at the operational AMV height and 9% compared to discrete operational AMV heights. The corresponding wind speed bias values are close to zero at layer depths of approximately 80 hPa.

Low-level GOES-AMVs (Fig. 35b) show a large reduction of VRMS differences for assigning AMVs to layers relative to the lidar cloud-top height for all layer depths, with minimum VRMS differences for 150-hPa layers. Here, the reduction reaches 22% (30%) when these layers are compared to layers (levels) at the operational AMV height. The corresponding wind speed bias is also clearly reduced compared to the operational values. High-level AMVs from MTSAT-2 (Fig. 35c) exhibit a similar pattern as high-level GOES-AMVs, achieving lowest VRMS differences for 100 hPa deep layers below the lidar cloud top. For low-level MTSAT-2-AMVs (Fig. 35d), no distinct improvement is found for assigning AMVs to layers relative to the lidar cloud-top height.

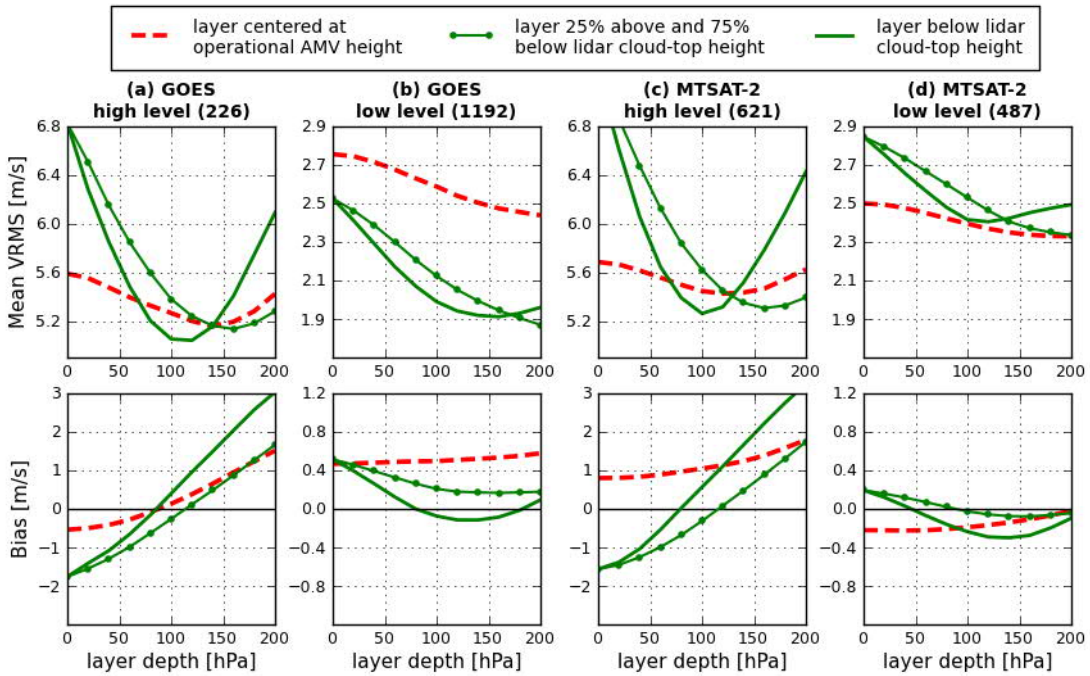

Figure 35: *Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV winds and layer-averaged model winds for (a) high-level GOES-AMVs, (b) low-level GOES-AMVs, (c) high-level MTSAT-2-AMVs and (d) low-level MTSAT-2-AMVs. Numbers in brackets are AMV counts. Red dashed lines represent layers centered at the original AMV pressure height; green lines represent layers relative to the lidar cloud-top height (cf. legend).*

6.1.2 Height bias correction

Figure 36 shows the VRMS difference and wind speed bias for the 30-day height bias correction for GOES- and MTSAT-2-AMVs. Results for levels/layers based on a 30-day bias correction (blue bars) are compared to levels/layers relative to the operational AMV height (red bars) and relative to the lidar cloud-top height (green bars). For high-level GOES-AMVs (Fig. 36a), the VRMS differences of the direct lidar height reassignment as well as the height bias correction do not show advantages over the operational height. In addition, the wind speed bias deteriorates for both height correction methods.

In contrast, low-level GOES-AMVs (Fig. 36b) exhibit clearly lower VRMS differences with similarly positive results for the direct height reassignment and the 30-day height bias correction. Again, layer-averaging exhibits additional benefits compared to using discrete levels. High-level AMVs from MTSAT-2 (Fig. 36c) show a similar pattern as Meteosat-10-AMVs in terms of VRMS differences. Lowest values are achieved for the direct height reassignment, but the height bias correction also leads to a small reduction of VRMS

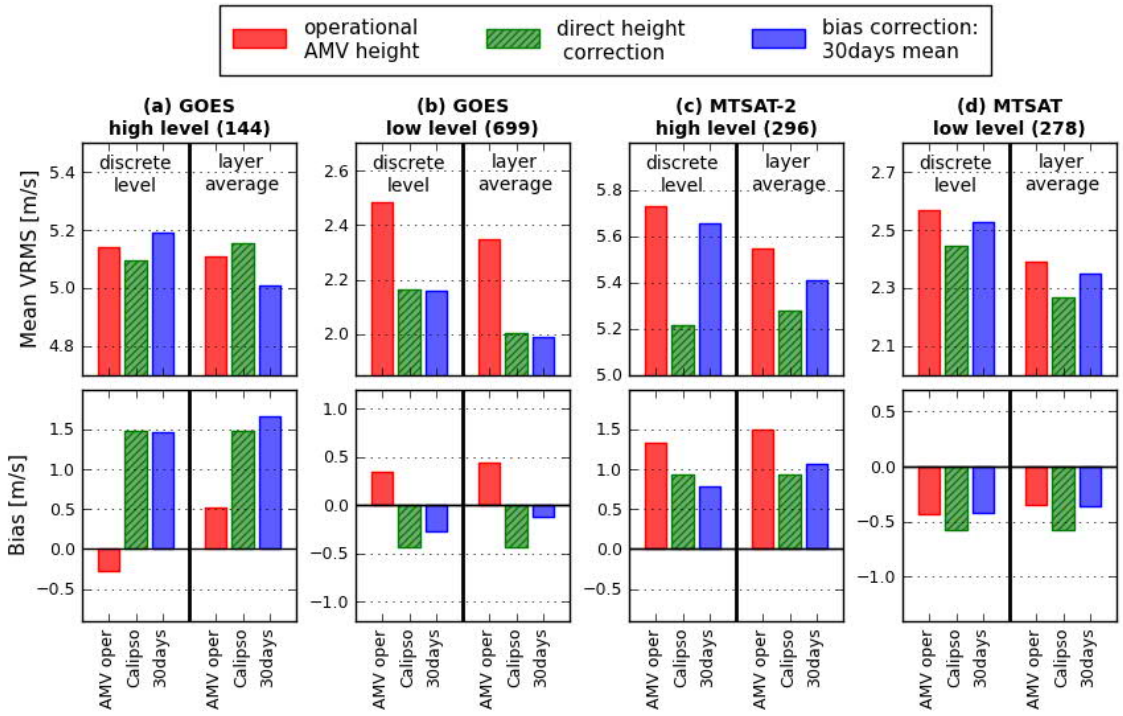

Figure 36: *Mean VRMS differences (upper panels) and wind speed bias (lower panels) between AMV and model winds for (a) GOES high-level, (b) GOES low-level, (c) MTSAT-2 high-level and (d) MTSAT-2 low-level AMVs. Numbers in brackets are AMV counts. 'AMV oper' corresponds to the operational AMV height and 'CALIPSO' to the direct lidar height reassignment . The applied height bias correction function is based on a 30-day mean ('30days'). Results are shown both for assigning AMVs to discrete levels (meaning the operational level, the level at 60 hPa below the actual lidar cloud top and the bias-corrected level) in the left part of each panel and to 120-hPa layer-averages centered at these levels in the right part of each panel.*

differences compared to the operational values. In addition, the wind speed bias is slightly reduced for both lidar height correction methods. For low-level AMVs of MTSAT-2 (Fig. 36d), the height bias correction shows slightly lower VRMS differences than the operational AMV height. Wind speed bias values for both height correction methods are of similar magnitude as the operational bias.

6.1.3 Summary and discussion

For other geostationary satellites, the positive effect of the direct height reassignment as well as of the height bias correction is less distinct than for Meteosat-AMVs. This is likely due to different AMV derivation algorithms that are aided by different model fields, and different

quality control procedures that have to be explored in more detail before final conclusions can be made. However, both height correction methods overall indicate benefits in terms of VRMS differences and wind speed bias when compared to the operational wind errors.

MTSAT-2-AMVs show both for high-level and low-level AMVs either neutral or slightly positive effects for both lidar height correction methods when compared to results at the operational AMV height. For low-level GOES-AMVs, results of the direct height reassignment as well as of the height bias correction indicate that lidar observations can reduce VRMS differences by up to 30%. This large reduction also reflects a well known feature of low-level GOES-AMVs in inversion regions, when AMVs are assigned too high in the atmosphere by NESDIS (Cotton, 2012). In contrast, high-level GOES-AMVs exhibit only small benefits for reducing VRMS differences, and even show degradation in terms of the wind speed bias. This may be related to a wind speed bias correction of +8% applied operationally by the data provider (NESDIS) to most GOES-AMVs above a pressure height of 300 hPa (Christopher Velden, personal communication). Consequently, there may be a need to either adapt the CALIPSO-based height bias correction for high-level GOES-AMVs or to apply the height bias correction to GOES-AMVs without the wind speed bias correction. However, it should also be kept in mind that the sample size for high-level GOES-AMVs is comparably small and further studies are required to draw robust conclusions regarding the benefits of a lidar-based height correction of high-level GOES-AMVs.

Generally, lidar-based height correction methods show promising results especially for low-level GOES-AMVs, but there is still need for research. Most reliable results for GOES-AMVs or MTSAT-2-AMVs could be obtained by verifying the AMV wind with radiosonde observations, as it is shown for MSG-AMVs in Section 4.1. However, the traditional launch times of operational radiosondes are around 00 UTC and 12 UTC, which coincide best with overpassing times of the sun-synchronous orbit of CALIPSO over the Meteosat domain. In other regions, the collocation of radiosondes and CALIPSO lidar observations is fairly complicated and would require an extensively long study period (several years) to receive enough collocated AMV/CALIPSO/radiosonde pairs.

6.2 Comparison of different statistical height correction strategies

The lidar-based height correction presented in this thesis constitutes a novel approach to categorize AMV errors with independent lidar observations. A different strategy of AMV

error characterization that is often used in AMV monitoring in NWP systems is the Level-of-Best-Fit (LBF) method. The AMV wind is compared to model profiles in order to find the model pressure level where the observed AMV wind fits best to the model wind. As the LBF method is highly dependent on the underlying model field, care should be taken when interpreting the results due to inherent model errors. In addition, it should be noted that LBF statistics are only based on a small (and, in contrast to lidar height correction statistics) not randomly distributed subsample of AMVs, as LBF can only be applied for certain wind conditions (e.g. without multiple minima in the model profile). However, a recent study that compared the LBF error statistics of the global models of ECMWF and the Met Office suggests that long-term height differences have overall similar characteristics (Salonen et al., 2015a).

Recently, an initial comparison between the lidar-based height bias correction and the LBF method used at ECMWF was initialized, and first joint results were presented by Kirsti Salonen (ECMWF) at the EUMETSAT Meteorological Satellite Conference in France in September 2015. Statistics from both methods were compared for a study period of about two

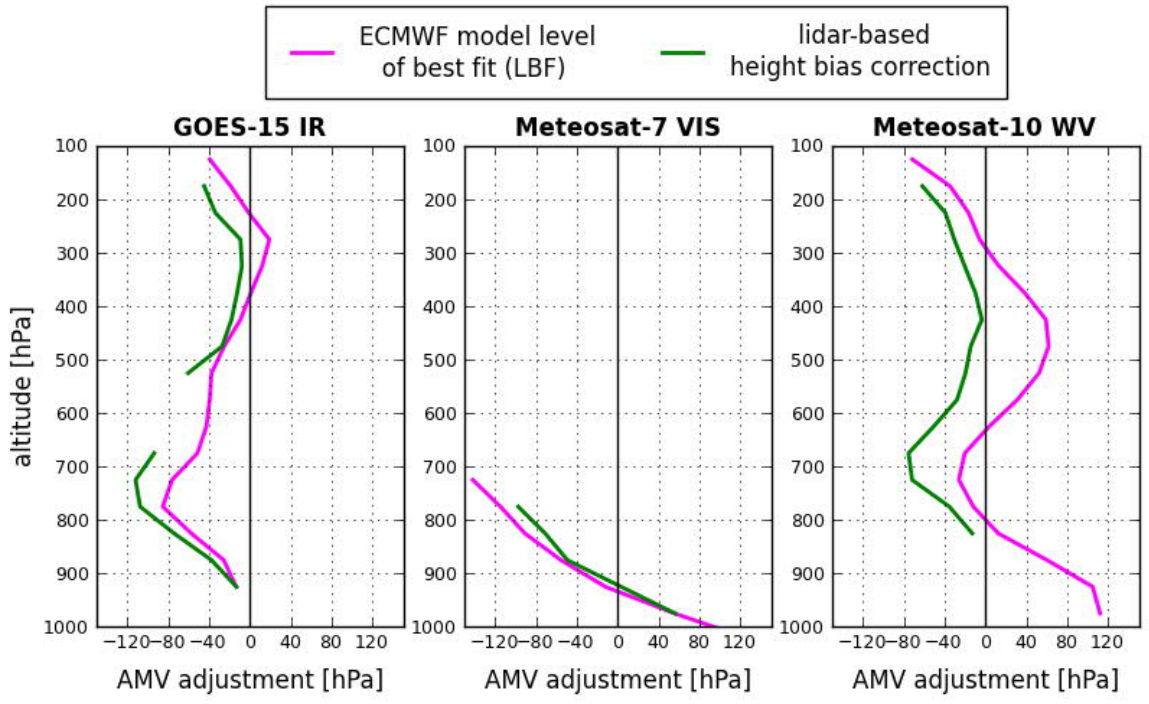

Figure 37: *Mean adjustment based on the ECMWF best fit pressure (magenta lines) and the lidar height bias correction (green lines) for GOES-15 IR-AMVs (left panel), Meteosat-7 VIS-AMVs (middle panel) and Meteosat-10 WV-AMVs (right panel). Adapted from Salonen et al., 2015b*

and a half months for the five main geostationary satellites Meteosat-10, Meteosat-7, MTSAT-2, GOES-13 and GOES-15 for the different channels VIS, IR and WV.

Figure 37 shows example cases for different satellite and channel combinations. Generally, the shape of the curves is similar for both methods, which holds as well for other channel and satellite combinations (not shown). However, the magnitude of the proposed height adjustment can vary considerably up to 50-60 hPa difference between the two methods, and there is a tendency for lidar-based corrections to locate AMVs lower than the LBF approach (see Salonen et al., 2015b for further details).

Generally, the lidar-based AMV height correction and the LBF method provide information on systematic height errors that are independent of each other. A comparison of these two methods can provide new insights about AMV error characteristics and their use in NWP models and should be exploited further in the future.

7. CONCLUSION AND OUTLOOK

Undoubtedly, AMVs are a crucial source of wind information for constructing the initial state for NWP models. However, major error sources such as height assignment issues, error correlations and representativity errors due to an interpretation as a single-level wind remain, so that the dense AMV data set must be heavily thinned in data assimilation systems. This thesis has, for the first time, investigated the potential of reducing the wind errors of AMVs through the incorporation of cloud-top observations from the lidar aboard the polar-orbiting satellite CALIPSO. High-resolution lidar cloud-top observations are expected to be independent of the AMV derivation and therefore exhibit a reliable information source that can be used for the AMV height correction.

The first part of this thesis addressed the research question, whether space-borne lidar observations can be used to improve the representation of AMVs by assigning them to a lidar-based layer position and layer depth. Individual AMVs were reassigned to different levels and layers relative to directly collocated lidar cloud-top observations. The wind evaluation was conducted using operational radiosondes and FG model fields from the global model GME. Both evaluation methods consistently showed the lowest VRMS differences and wind speed bias values for high-level AMVs for 120-hPa deep layers below the lidar cloud-top. The evaluation of low-level AMVs revealed slightly deviating results for the radiosonde and model evaluation due to a different operational height assignment approach (AMV assignment to cloud base and cloud top, respectively) during the corresponding evaluation period. Overall, the reduction of VRMS differences for the 120-hPa layers below the lidar cloud top was quantified as 15-17% when compared to discrete operational levels, and 8-12% when compared to reference layers of the same depth centered at the operational level. These results depended somewhat on the respective evaluation period, evaluation method and AMV altitude, but were generally consistent for all settings. Furthermore, the lidar-based height reassignment reduced the “slow” bias of current upper-level AMVs, and horizontal AMV error correlations were decreased by ~50 km. Generally, a subdivision into tropical and extra-

tropical regions led to similar findings, supporting the robustness of the results presented in this study.

As the direct height reassignment requires collocated lidar observations for each AMV, the applicability of this method is restricted to space-borne lidar observations available in real-time and is therefore rather complex to apply in operational data assimilation systems. Therefore, the second scientific question underlying this thesis was the issue of how NWP may benefit from incorporating lidar information for the AMV height correction in a statistical approach without the need for real-time lidar data. To address this, an alternative height correction to the direct AMV height reassignment was proposed by introducing a statistical correction of systematic AMV height biases that can be applied to all AMVs of the respective geostationary satellite. For that purpose, an average adjustment based on 30-day or 10-day training periods was calculated and then applied to a subsequent, independent evaluation period. Results indicated that this height bias correction achieves on average 50% of the VRMS reduction of the direct height reassignment. The height bias correction therefore provides a feasible way of correcting operational AMV heights that requires neither real-time lidar data nor directly collocated lidar observations and would be easy to implement in an NWP system. When anticipated for operational use, monthly or weekly updates of the height bias correction functions seem advisable in order to catch features due to seasonal variability or changes in the height assignment processing. However, the optimal update interval is still to be determined.

The applicability of using lidar-corrected and layer-averaged AMVs in NWP systems was tested by running assimilation and forecast experiments for a 16-day period with the global model ICON with overall encouraging results. Of special note was the reduction of the wind speed bias by up to 20% when the resulting FG model fields are compared with radiosonde soundings in the tropics and the southern hemisphere. However, a precise assessment of the impact on forecast skill will require further experimentation, such as applying longer study periods and taking into account seasonal variability.

In summary, the lidar-based height correction methods presented in this thesis have proven to be a valuable approach for the reduction of the VRMS difference, the wind speed bias and error correlations of operational AMVs, suggesting that NWP may benefit from assimilating lidar-corrected and layer-averaged AMVs. Options for future applications are to be explored

in more detail and may contain the expansion of lidar-based height correction techniques to other geostationary (or polar-orbiting) satellites. However, further investigations in this direction are needed before final conclusions can be made.

Moreover, the application of height correction methods based on space-borne lidar information is not restricted to CALIPSO. Other space-borne lidars are planned to be launched in the near future, e.g. the Earth Clouds, Aerosols and Radiation Explorer (EarthCARE, see e.g. Illingworth et al., 2015). Thus, the assimilation of AMVs as layer-averages in combination with lidar information for the AMV height correction is seen as a promising approach to increase the benefit of AMVs for NWP in future. Furthermore, a lidar-based height bias correction may be useful for deriving consistent data sets for climate research and to evaluate AMV height assignment methods.

The lidar-based height correction approach is planned to be performed as a part of the DWD contribution to the Satellite Application Facility for Numerical Weather Prediction (NWP-SAF) AMV monitoring in the future. This EUMETSAT-funded initiative collects monitoring statistics from different weather centres and provides statistical comparisons between observations and model background fields to produce a detailed analysis of errors for various observation types (Forsythe, 2007; Cotton, 2012).

List of abbreviations

AMV	Atmospheric Motion Vector
CAD	Cloud-Aerosol-Distinguisher
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CCC	Cross-Correlation Contribution method
CNES	Centre National d'Études Spatiales
DWD	Deutscher Wetterdienst
EBBT	Equivalent Black Body Temperature
ECMWF	European Centre for Medium-Range Weather Forecasts
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
ESA	European Space Agency
FG	First Guess
GCOM	Global Change Observation Mission
GME	Global Model
GEOS-5	Goddard Earth Observing System model, version 5
GMAO	Global Modelling and Assimilation Office
GOES	Geostationary Operational Environmental Satellite
ICON	Icosahedral Nonhydrostatic model
IR	Infra-Red
JMA	Japan Meteorological Agency
LBF	Level of Best Fit
LEV	Level assimilation with lidar-corrected AMVs
LIDAR	Light Detection And Ranging
LYR	Layer assimilation with lidar-corrected AMVs
METOP	Meteorological Operational Satellite
MFG	Meteosat First Generation
MODIS	Moderate resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
MTSAT	Multi-Functional Transport Satellite

MVD	Mean Vector Difference
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
Nd:YAG	Neodymium-doped Yttrium Aluminium Garnet
NESDIS	National Environmental Satellite Data and Information Service
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
NWP-SAF	NWP Satellite Application Facility
OCO	Orbiting Carbon Observatory
QI	Quality Index
REF	Reference run (assimilation experiments)
RMS	Root Mean Square error
SCA	Scene Classification Algorithms
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SIBYL	Selective Iterative Boundary Locator
THORPEX	The Observing System Research and Predictability Experiment
T-PARC	THORPEX Pacific Asian Regional Campaign
VIS	Visible
VRMS	Vector Root Mean Square error
WV	Water Vapour
3DVAR	3-Dimensional Variational data assimilation

Bibliography

Baker, N. L., R. Langland, P. M. Pauley1, L Xu, D. Merkova, R. Gelaro and C. Velden, 2012: The impact of satellite atmospheric motion vectors in the U.S. Navy global data assimilation system – NWP results. Proceedings of the 11th *International Winds Workshop* , Auckland, New Zealand, February 2012.

Baker, W. E., R. Atlas, C. Cardinali, A. Clement, G. D. Emmitt, B. M. Gentry, R. M. Hardesty, E. Källén, M. J. Kavaya, R. Langland, Z. Ma, M. Masutani, W. McCarty, R. B. Pierce, Z. Pu, L. P. Riishojgaard, J. Ryan, S. Tucker, M. Weissmann, and J. G. Yoe, 2014: Lidar-measured wind profiles – the missing link in the global observing system. *Bulletin of the American Meteorological Society*, 95, 543–564.

Baldauf, M., G. Zängl, M. Buchhold, R. Potthast, 2015: DWD report. 30th WGNE-meeting, National Centers for Environmental Prediction (NCEP), Washington, March 2015.

Besançon, R. M. (Editor), 1990: The encyclopedia of physics, third edition. Springer Science and Business Media New York.

Borde, R., and P. Dubuisson, 2010: Sensitivity of atmospheric motion vectors height assignment methods to semitransparent cloud properties using simulated Meteosat-8 radiances. *Journal of Applied Meteorology and Climatology* , 49, 1205–1218.

Borde, R., M. Doutriaux-Boucher, G. Dew, and M. Carranza, 2014a: A direct link between feature tracking and height assignment of operational EUMETSAT atmospheric motion vectors. *Journal of Atmospheric and Oceanic Technology* , 31, 33–46.

Borde, R. and J. García-Pereda, 2014b: Impact of wind guess on the tracking of atmospheric motion vectors. *Journal of Atmospheric and Oceanic Technology* , 31, 458–467.

Borde, R.: An overview of 10 years of research activities on AMVs at EUMETSAT, 2014c, *Proceedings of the 12th International Winds Workshop* , Copenhagen, Denmark, June 2014.

Bormann, N., S. Saarinen, G. Kelly, and J.-N. Thépaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Monthly Weather Review*, 131, 706–718.

Bormann, N., and J.-N. Thépaut, 2004: Impact of MODIS polar winds in ECMWFs 4DVAR data assimilation system. *Monthly Weather Review*, 132, 929–940.

Bouttier, F., and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quarterly Journal of the Royal Meteorological Society,* 127, 1469–1488.

Bresky, W. C., J. M. Daniels, A. A. Bailey, and S. T. Wanzong, 2012: New methods toward minimizing the slow speed bias associated with atmospheric motion vectors. *Journal of Applied Meteorology and Climatology* , 51, 2137–2151.

Büche, G., H. Karbstein, A. Kummer, and H. Fischer, 2006: Water vapor structure displacements from cloud-free Meteosat scenes and their interpretation for the wind field. *Journal of Applied Meteorology and Climatology* , 45, 556–575.

Carranza, M., R. Borde, M. Doutriaux-Boucher, 2014: Recent changes in the derivation of geostationary atmospheric motion vectors at EUMETSAT. *Proceedings of the 12th International Winds Workshop* , Copenhagen, Denmark, June 2014.

Cotton, J., 2012: Fifth analysis of the data displayed on the NWP SAF AMV monitoring website. Document NWPSAF-MO-TR-027. Available online at http://nwpsaf.eu/monitoring/amv/nwpsaf_mo_tr_027.pdf.

Cotton, J., M. Forsythe, F. Warrick, and R. Marriott, 2014: AMV impact studies at the Met Office, *Proceedings of the 12th International Winds Workshop*, Copenhagen, Denmark, June 2014.

Cress, A., and H. W. Bitzer, 2012: Recent progress in using satellite winds at the German Weather Service. *Proceedings of the 11th International Winds Workshop*, Auckland, New Zealand, February 2012.

Daniels, J., W. Bresky, C. Velden, and K. Bedka, 2006: Characterization of tracer height assignment errors: A recurring theme for atmospheric motion vectors. *Proceedings of the 8th International Winds Workshop*, Peking, China, April 2006.

Di Michele, S., T. McNally, P. Bauer, and I. Genkova, 2013: Quality assessment of cloud-top height estimates from satellite IR radiances using the CALIPSO lidar. *IEEE Transactions on Geoscience and Remote Sensing* , 51, 2454–2464.

DWD, 2015a: Operationelles NWV-System Änderungsmitteilung: Operationelle Einführung des neuen nichthydrostatischen Globalmodells ICON. Available online at: http://www.dwd.de/DE/fachnutzer/forschung_lehre/numerische_wettervorhersage/nwv_aenderungen/_functions/DownloadBox_modellaenderungen/2015_icon_Download_Box.html?nn=346850&lsbId=479218.

DWD, 2015b: Operationelles NWV-System Änderungsmitteilung: ICON, Version 2.0.04: Diverse Fehlerkorrekturen mit geringer Auswirkung auf die Vorhersagequalität. Available online at same link as DWD, 2015a (above).

Dworak, R., and J. R. Key, 2009: Twenty years of polar winds from AVHRR: Validation and comparison with ERA-40. *Journal of Applied Meteorology and Climatology* , 48, 24–40.

English, S., T. McNally, N. Bormann, K. Salonen, M. Matricardi, A. Horanyi, M. Rennie, M. Janisková, S. Di Michele, A. Geer, E. Di Tomaso, C. Cardinali, P. de Rosnay, J. Muñoz Sabater, M. Bonavita, C. Albergel, R. Engelen and J.-N. Thépaut, 2013: Impact of satellite data. *Research Department ECMWF*, Technical Memorandum No. 711, 46 pp.

EUMETSAT, 2011: MSG meteorological products extraction facility algorithm specification document. EUMETSAT Doc. EUM/MSG/SPE/022, 259 pp. Available online at: http://www.eumetsat.int/website/home/Data/TechnicalDocuments/index.html.

Fischer, L., 2013: Statistical characterisation of water vapour variability in the troposphere: a height-resolved analysis using airborne lidar observations and COSMO-DE model simulations. Ph.D. thesis, Ludwig-Maximilians-University Munich, Faculty of Physics.

Folger, K., 2012: Höhenkorrektur von Satelliten-Windvektoren (AMVs) mit flugzeug-getragenen Lidarmessungen. M.Sc. thesis, Meteorologisches Institut, Ludwig-Maximilians-University Munich, Germany, 81 pp.

Folger, K. and M. Weissmann, 2014: Height correction of atmospheric motion vectors using satellite lidar observations from CALIPSO. *Journal of Applied Meteorology and Climatology* , 53, 1809–1819.

Folger, K. and M. Weissmann, 2016: Lidar-based height correction for the assimilation of atmospheric motion vectors. *Journal of Applied Meteorology and Climatology* , submitted.

Forsythe, M., 2007: Atmospheric motion vectors: past, present and future. *Seminar on recent development in the use of satellite observations in NWP* , ECMWF, Reading, September 2007.

Frank, H., D. Liermann, D. Majewski, B. Ritter, 2014: Kurze Beschreibung des Globalmodells GME (20 km / L60) und seiner Datenbanken auf dem Datenserver des DWD. Technical Report, DWD, Geschäftsbereich Forschung und Entwicklung, 83 pp.

García-Pereda, J. and R. Borde, 2014: Impact of the tracer size and the temporal gap between images in the extraction of atmospheric motion vectors. *Journal of Atmospheric and Oceanic Technology*, 31, 1761–1770.

García-Pereda, J., 2014: NWC SAF/High resolution winds AMV software evolution between 2012 and 2014. *Proceedings of the 12th International Winds Workshop*, Copenhagen, Denmark, June 2014.

Genkova, I., R. Borde, J. Schmetz, C. Velden, K. Holmlund, N. Bormann, and P. Bauer, 2010: Global atmospheric motion vector intercomparison study. *Proceedings of the 10th International Winds Workshop* , Tokyo, Japan, February 2010.

Goerss, J. S., Velden, C. S., and Hawkins, J. D., 1998: The impact of multispectral GOES-8 wind information on Atlantic tropical cyclone track forecasts in 1995. Part II: NOGAPS forecasts. *Monthly Weather Review*, 126, 1219–1227.

Hasler, A.F., W.C. Skillman, W.E. Shenk, and J. Steranka, 1979: In situ aircraft verification of the quality of satellite cloud winds over oceanic regions. *Journal of Applied Meteorology* , 18, 1481–1489.

Hautecoeur, O., R. Borde, M. Doutriaux-Boucher, and M. Carranza, 2014: EUMETSAT operational dual-METOP winds products. *Proceedings of the 12th International Winds Workshop*, Copenhagen, Denmark, June 2014.

Hernandez-Carrascal, A., and N. Bormann, 2014: Atmospheric motion vectors from model simulations. Part II: Interpretation as spatial and vertical averages of wind and role of clouds. *Journal of Applied Meteorology and Climatology* , 53, 65–82.

Holmlund, K., 1998: The utilization of statistical properties of satellite-derived atmospheric motion vectors to derive quality indicators. *Weather Forecasting* , 13, 1093–1104.

Hunt, W. H., D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: CALIPSO lidar description and performance assessment . *Journal of Atmospheric and Oceanic Technology* , 26, 1214–1228.

Illingworth, A. J., H. W. Barker, A. Beljaars, M. Ceccaldi, H. Chepfer, J. Cole, J. Delanoë, C. Domenech, D. P. Donovan, S. Fukuda, M. Hirakata, R. J. Hogan, A. Huenerbein, P. Kollias, T. Kubota, T. Nakajima, T. Y. Nakajima, T. Nishizawa, Y. Ohno, H. Okamoto, R. Oki, K. Sato, M. Satoh, M. Shephard, U. Wandinger, T. Wehr, and G.-J. van Zadelhoff, 2015: The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation and radiation. *Bulletin of the American Meteorological Society* , 96, 1311–1332.

Isaksen, L., C. Cardinali and D. Li, 2010: Diagnosing the impact of atmospheric motion vectors in the ECMWF 4D-VAR assimilation system. *Proceedings of the 10th International Winds Workshop*, Tokyo, Japan, February 2010.

Kalnay, E., 2003: Atmospheric modeling, data assimilation and predictability. Cambridge University Press.

Kelly, G., and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the global observing system through observing system experiments. EUMETSAT/ECMWF Report Series. Available online at:
http://old.ecmwf.int/publications/library/ecpublications/_pdf/saf/Kelly_Thepaut.pdf.

Key, J. R., D. Santek, C. S. Velden, N. Bormann, J.-N. Thépaut, L. P. Riishojgaard, Y. Zhu, and W. P. Menzel, 2003: Cloud-drift and water vapor winds in the polar regions from MODIS. *IEEE Transactions on Geoscience and Remote Sensing* , 41, 482–492.

Key, J., J. Daniels, S. Wanzong, A. Bailey, H. Qi, W. Bresky, D. Santek, C. Velden, and W. Wolf, 2014: VIIRS polar winds status and use. *Proceedings of the 12th International Winds Workshop*, Copenhagen, Denmark, June 2014.

Kim, T.-M., B. Lee, S. Chung, H. Park, 2015: Study of the height assignment of AMV for the future GK-2A satellite, *Proceedings of the EUMETSAT Satellite Conference* , Toulouse, France, September 2015.

Lazzara, M., R Dworak, D. Santek, B. Hoover, C. Velden, and J. Key, 2014: High-latitude atmospheric motion vectors from composite satellite data. *Journal of Applied Meteorology and Climatol*ogy, *53*, 534–547.

Le Marshall, J., N. Pescod, B. Seaman, G. Mills, and P. Stewart, 1994: An operational system for generating cloud drift winds in the Australian region and their impact on numerical weather prediction. *Weather and Forecasting*, 9, 361–370.

Lean, P., S. Migliorini, and G. Kelly, 2015: Understanding atmospheric motion vector vertical representativity using a simulation study and first guess departure statistics. *Journal of Applied Meteorology and Climatology* , 54, 2479–2500.

Liu, Z., A. H. Omar, Y. Hu, M. A. Vaughan, and D. M. Winker, 2005: CALIOP algorithm theoretical basis document part 3: Scene classification algorithms. Document No: PC-SCI-202 Part 3. Available online at:
http://www-calipso.larc.nasa.gov/resources/project_documentation.php.

Lutz, H.-J., 1999: Cloud processing for Meteosat Second Generation. Technical Department EUMETSAT, Technical Memorandum No. 4, 26 pp.

Majewski, D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, 2002: The operational global icosahedral–hexagonal gridpoint model GME: Description and high-resolution tests. *Monthly Weather Review*, 130, 319–338.

Mecikalski J. R., and Kristopher M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Monthly Weather Review,* 134, 49–78.

Menzel, W. P., W. L. Smith, and T. R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment using VAS. *Journal of Climate and Applied Meteorology* , 22, 377-384.

Menzel, W. P., 1996: Report from the working group on verification statistics. *Proceedings of the 3rd International Winds Workshop* , Switzerland, June 1996. Available online at: cimss.ssec.wisc.edu/iwwg/iww3/index_3rdWindsWorkshop.htm.

Menzel, W. P., 2001: Cloud tracking with satellite imagery: From the pioneering work of Ted Fujita to the present. *Bulletin of the American Meteorological Society* , 82, 33–47.

Merk, D., and T. Zinner, 2013: Detection of convective initiation using Meteosat SEVIRI: Implementation in and verification with the tracking and nowcasting algorithm Cb-TRAM. *Atmospheric Measurement Techniques* , 6, 1903–1918.

Nieman, S. J., J. Schmetz, and W. P. Menzel, 1993: A comparison of several techniques to assign heights to cloud tracers. *Journal of Applied Meteorology* , 32, 1559–1568.

Nieman, S. J., W. P. Menzel, C. M. Hayden, D. Gray, S. T. Wanzong, C. S. Velden, and J. Daniels, 1997: Fully automated cloud-drift winds in NESDIS operations. *Bulletin of the American Meteorological Society* , 78, 1121–1133.

Payan, C., and J. Cotton, 2012: Collaborative satellite winds impact study. Report version 1.1. Online available at:
http://cimss.ssec.wisc.edu/iwwg/Docs/windsdenial-synthesisV1-1.pdf.

Rao, P. A., C. S. Velden, and S. A. Braun, 2002: The vertical error characteristics of GOES-derived winds: description and experiments with numerical weather prediction. *Journal of Applied Meteorology* , 41, 253–271.

Reinert, D., F. Prill, H. Frank, and G. Zängl, 2015: ICON database reference manual, version 1.1.1. Effective date July 21, 2015. Available online at:
http://www.dwd.de/SharedDocs/downloads/DE/modelldokumentationen/nwv/icon/icon_dbbeschr_aktuell.html.

Riishojgaard, L. P., and Y. Zhu, 2004: Impact experiments on GMAO data assimilation and forecast systems with MODIS winds during MOWSAP. *Proceedings of the 7th International Winds Workshop*, Helsinki, Finland, June 2004.

Rohn, M., G. Kelly, and R. W. Saunders, 2001: Impact of a new cloud motion wind product from Meteosat on NWP analyses and forecasts. *Monthly Weather Review*, 129, 2392–2403.

Salonen, K., and N. Bormann, 2012: Atmospheric motion vector observations in the ECMWF system: Second year report. EUMETSAT/ECMWF Fellowship Programme Research Report 28, 41 pp. Available online at
http://old.ecmwf.int/publications/library/ecpublications/_pdf/saf/rr28.pdf.

Salonen, K., and N. Bormann, 2014: AMVs in the operational ECMWF system. *Proceedings of the 12th International Winds Workshop* , Copenhagen, Denmark, June 2014.

Salonen, K., J. Cotton, N. Bormann, and M. Forsythe, 2015a: Characterizing AMV height-assignment error by comparing best-fit pressure statistics from the Met Office and ECMWF data assimilation systems. *Journal of Applied Meteorology and Climatology,* 54, 225–242.

Salonen, K., K. Folger, M. Weissmann and N. Bormann, 2015b: Investigations on accounting for the systematic height assignment errors for AMVs in NWP. *Proceedings of the EUMETSAT Meteorological Satellite Conference* , Toulouse, France, September 2015.

Santek, D., 2010: The impact of satellite-derived polar winds on lower-latitude forecasts. *Monthly Weather Review*, 138, 123–139.

Santek, D., J. García-Pereda, C. Velden, I. Genkova, S. Wanzong, D. Stettner, and M. Mindock, 2014: 2014 AMV intercomparison study report (comparison of NWC SAF/HRW AMVs with AMVs from other producers). Technical report available online at: www.nwcsaf.org/HD/files/vsadoc/CIMSS_AMV_Comparison_FinalReport_04July2014.pdf.

Schmetz, J., and M. Nuret, 1989: Cloud motion wind estimates in Europe. *Workshop on the Use of Satellite Data in Operational Numerical Weather Prediction:* 1989–1993, Vol. 2, ECMWF/EUMETSAT, 274–283.

Schmetz, J., K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, and L. Van De Berg, 1993: Operational cloud-motion winds from Meteosat infrared images. *Journal of Applied Meteorology* , 32, 1206–1225.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, 83, 977–992.

Schreiner, A., W. P. Menzel, W. Straka, and A. Heidinger, 2012: Comparing CO2/IRW and H2O/IRW CTPs. *Proceedings of the 11th International Winds Workshop* , Auckland, New Zealand, February 2012.

Seidel, D. J., B. Sun, M. Pettey, and A. Reale, 2011: Global radiosonde balloon drift statistics, *Journal of Geophysical Research* , 116, D07102.

Sèze, G., S. Marchand, J. Pelon and R. Borde, 2008: Comparison of AMV cloud top pressure derived from MSG with space based lidar observations (CALIPSO). *Proceedings of the 9th International Winds Workshop* , Annapolis, USA, April 2008.

Smith, W.L. and C. M. R. Platt, 1978: Comparison of satellite-deduced cloud heights with indications from radiosonde and ground-based laser measurements. *Journal of Applied Meteorology*, 17, 1796–1802.

Soden, B., C. Velden, and R. Tuleya, 2001: The impact of satellite winds on experimental GFDL hurricane model forecasts. *Monthly Weather Review*, 129, 835–852.

Süli, E., and D. F. Mayers, 2003: An introduction to numerical analysis. Cambridge University Press.

Szantai, A., F. Desalmand, M. Desbois, and P. Lecomte, 2000: Tracking low-level clouds over central Africa on Meteosat images. *Proceedings of the EUMETSAT Meteorological Satellite Data Users' Conference,* Bologna, Italy.

Szejwach, G., 1982: Determination of semi-transparent cirrus cloud temperature from infrared radiances: Application to Meteosat. *Journal of Applied Meteorology* , 21, 384–393.

Vaughan, M. A., D. M. Winker, K. A. Powell, 2005: CALIOP algorithm theoretical basis document part 2: Feature Detection and layer properties algorithms. Document No: PC-SCI-202 Part 2. Available online at:
http://www-calipso.larc.nasa.gov/resources/project_documentation.php.

Vaughan, M. A., K. A. Powell, D. M. Winker, C. A. Hostetler, R. E. Kuehn, W. H. Hunt, B. J. Getzewich, S. A. Young, Z. Liu, and M. J. McGill, 2009: Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements. *Journal of Atmospheric and Oceanic Technology* , 26, 2034–2050.

Velden, C., D. Stettner, and J. Daniels, 2000: Wind vector fields derived from GOES rapid-scan imagery. *Proceedings of the 10th Conference on Satellite Meteorology and Oceanography*, California, Amer. Meteor. Soc., 20–23.

Velden, C. S., J. Daniels, D. Stettner, D. Santek, J. Key, J. Dunion, K. Holmlund, G. Dengel, W. Bresky, and P. Menzel, 2005: Recent innovations in deriving tropospheric winds from meteorological satellites. *Bulletin of the American Meteorological Society* , 86, 205–223.

Velden, C. S., and K. M. Bedka, 2009: Identifying the uncertainty in determining satellite-derived atmospheric motion vector height attribution. *Journal of Applied Meteorology and Climatology*, 48, 450–463.

Wang, D., X. Liang, Y. Duan, and J. C. Chan, 2006: Impact of four-dimensional variational data assimilation of atmospheric motion vectors on tropical cyclone track forecasts. *Weather and Forecasting*, 21, 663–669.

Weissmann, M., F. Harnisch, C.-C. Wu, P.-H. Lin, Y. Ohta, K. Yamashita, Y.-H. Kim, E.-H. Jeon, T. Nakazawa, and S. Aberson, 2011: The influence of assimilating dropsonde data on typhoon track and midlatitude forecasts. *Monthly Weather Review*, 139, 908–920.

Weissmann, M., R. H. Langland, C. Cardinali, P. M. Pauley, and S. Rahm, 2012: Influence of airborne Doppler wind lidar profiles near typhoon Sinlaku on ECMWF and NOGAPS forecasts. *Quarterly Journal of the Royal Meteorological Society*, 138, 118–130.

Weissmann, M., K. Folger, and H. Lange, 2013: Height correction of atmospheric motion vectors using airborne lidar observations. *Journal of Applied Meteorology and Climatology* , 52, 1868–1877.

Weitkamp, C, 2005: Lidar: Range-resolved optical remote sensing of the atmosphere. Springer.

Winker, D. M., C. A. Hostetler, M. A. Vaughan, and A. H. Omar, 2006: CALIOP algorithm theoretical basis document part 1: CALIOP instrument, and algorithms overview. Document No: PC-SCI-202 Part 1. Available online at: http://www-calipso.larc.nasa.gov/resources/project_documentation.php.

Winker, D. M., M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, 2009: Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology* , 26, 2310–2323.

Winker, D. M., J. Pelon, J. A. Coakley Jr., S. A. Ackerman, R. J. Charlson, P. R. Colarco, P. Flamant, Q. Fu, R. M. Hoff, C. Kittaka, T. L. Kubar, H. Le Treut, M. P. McCormick, G. Mégie, L. Poole, K. Powell, C. Trepte, M. A. Vaughan, and B. A. Wielicki, 2010: The CALIPSO mission: A global 3D view of aerosols and clouds. *Bulletin of the American Meteorological Society*, 91, 1211–1229.

Zängl, G., D. Reinert, P. Rípodas and M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141, 563–579.

Acknowledgments

First and foremost, I want to express my sincere gratitude to Dr. Martin Weissmann for being my supervisor during the last three and a half years. He perfectly guided me through this thesis, always having time for constructive discussions, valuable advice or simply some motivating words. In my eyes, it could not have gone better – thank you!

I would also like to thank Prof. Dr. Bernhard Mayer for reviewing this thesis as co-examiner, for his interest in my work and helpful suggestions.

My thanks are further extended to Harald Anlauf and Alex Cress from DWD in Offenbach for their continuous support over the last years. Harald implemented an observation operator for layer-averaged AMVs that I used in this thesis, and was indeed an invaluable help with small and big problems with the operator. Alex was always a great information source for everything that is connected to AMVs at DWD.

My colleagues at the Meteorological Institute at the LMU Munich contributed a lot that I felt comfortable at work. Many thanks in particular to my long-term room-mates Katrin Scheufele, Fabian Brundke and Tobias Necker for making the daily routine in our office enjoyable and cheerful. Also a big thanks to Florian Harnisch and Richard Keane for proof-reading parts of this thesis and Lotte Bierdel for many sunny lunch (or afternoon) breaks on the balcony to clear one's mind again.

This work would not have been possible without the support of my great family, who backed me in all conceivable ways and always let me feel beloved. Last but not least: Thank you Oliver for all your support and motivating words, and for simply being always there.

This doctoral thesis was carried out at the Hans-Ertel-Centre for Weather Research. This German research network of universities, research institutes and DWD is funded by the BMVI (Federal Ministry of Transport and Digital Infrastructure). The CALIPSO data were obtained from the NASA Langley Research Center Atmospheric Science Data Center.