Convective-Scale Data Assimilation of Thermodynamic Lidar Data into the Weather Research and Forecasting Model

**Dissertation to obtain the doctoral degree of
Natural Sciences (Dr. rer. nat.)**

**Faculty of Natural Sciences
University of Hohenheim**

Institute of Physics and Meteorology

submitted by
Rohith Muraleedharan THUNDATHIL

from Kodungallur, India

2022

ii

Dean: Prof. Dr. rer. nat. Uwe Beifuß

Chairman of the Doctoral Committee: Prof. Dr. Reinhard Kohlus

1. Mentor and Supervisor: Prof. Dr. rer. nat. Volker Wulfmeyer

2. Mentor: Dr. Olivier Caumont

3. Mentor: Prof. Dr. Thilo Streck

Thesis submitted on: 09 March 2022
Doctoral defense examination on: 23 September 2022

The thesis titled "Convective-Scale Data Assimilation of Thermodynamic Lidar Data into the Weather Research and Forecasting Model" was accepted on 23 September, 2022 by the Faculty of Natural Sciences at the University of Hohenheim as a "Dissertation for obtaining the doctoral degree of Natural Sciences (Dr. rer. nat.)".

Abstract

This thesis studies the impact of assimilating temperature and humidity profiles from ground-based lidar systems and demonstrates its value for future short-range forecast. Thermodynamic profile obtained from the temperature Raman lidar and the water-vapour differential absorption lidar of the University of Hohenheim during the High Definition of Clouds and Precipitation for advancing Climate Prediction (HD(CP)^2) project Observation Prototype Experiment (HOPE) are assimilated into the Weather Research and Forecasting model Data Assimilation (WRFDA) system by means of a new forward operator. The impact study assimilating the high-resolution thermodynamic lidar data was conducted using variational and ensemble-based data assimilation methods. The first part of the thesis describes the development of the thermodynamic lidar operator and its implementation through a deterministic DA impact study. The operator facilitates the direct assimilation of water vapour mixing ratio (WVMR), a prognostic variable in the WRF model, without conversion to relative humidity. Undesirable cross sensitivities to temperature are avoided here so that the complete information content of the observation with respect to the water vapour is provided. The assimilation experiments were performed with the three-dimensional variational (3DVAR) DA system with a rapid update cycle (RUC) with hourly frequency over ten hours. The DA experiments with the new operator outperformed the previously used relative humidity operator, and the overall humidity and temperature analyses improved. The simultaneous assimilation of temperature and WVMR resulted in a degradation of the temperature analysis compared to the improvement observed in the sole temperature assimilation experiment. The static background error covariance matrix (B) in the 3DVAR was identified as the reason behind this behaviour. The correlation between the temperature and WVMR variables in the background error covariance matrix of the 3DVAR, which is static and not flow-dependent, limited the improvement in temperature. The second part of the thesis provides a solution for overcoming the static B matrix issue. A hybrid, ensemble-based approach was applied using the Ensemble Transform Kalman Filter (ETKF) and the 3DVAR to add flow dependency to the B matrix. The hybrid experiment resulted in a 50% lower temperature and water vapour root mean square error (RMSE) than the 3DVAR experiment. Comparisons against independent radiosonde observations showed a reduction of RMSE by 26% for water vapour and 38% for temperature. The planetary boundary layer (PBL) height of the analyses also showed an improvement compared to the available ceilometer. The impact of assimilating a single lidar vertical profile spreads over a 100 km radius, which is promising for future assimilation of water vapour and temperature data from operational lidar networks for short-range weather forecasting. A forecast improvement was observed for 7 hours lead time compared with the ceilometer derived planetary boundary layer height observations and 4 hours with

iv

Global Navigation Satellite System (GNSS) derived integrated water vapour observations. With the help of sophisticated DA systems and a robust network of lidar systems, the thesis throws light on the future of short-range operational forecasting.

Zusammenfassung

Die Einfluss der Integration von Temperatur- und Feuchtigkeitsprofilen aus bodengestützten aktiven Lidar-Systemen wird untersucht und ihr Nutzen für künftige Kurzstreckenvorhersagen wird demonstriert. Thermodynamische Profile, die mit dem Temperatur-Raman-Lidar und dem Wasserdampf-DIAL der Universität Hohenheim während des „Observation Prototype Experiment" (HOPE), das Teil des Projekts „High Definition of Clouds and Precipitation for advancing Climate Prediction" (HD(CP)^2) war, gewonnen wurden, werden mit Hilfe eines neuen Vorwärtsoperators in das Datenasimilations-System (DA) des Wetterforschungs- und -vorhersagemodells (WRF) Modells assimiliert. Die Untersuchungen zum Einfluss der Assimilation der hochauflösenden thermodynamischen Lidar-Daten wurden dabei mit Variations- und Ensemble-basierten Datenassimilationsmethoden durchgeführt.

Der erste Teil des Arbeit beschreibt die Entwicklung des Lidar-Operators und seine Implementierung mit Hilfe einer deterministischen Datenassimilationsstudie. Der Operator ermöglicht die direkte Assimilation des Wasserdampf-Mischungsverhältnisses (WVMR), einer prognostischen Variable im WRF-Modell, ohne Umrechnung in relative Feuchte. Unerwünschte Querempfindlichkeiten zur Temperatur werden hierbei vermieden, und der vollständige Informationsgehalt der Beobachtung in Bezug auf den Wasserdampf wird genutzt. Das Assimilations-Experiment wurde mit einer 3-dimensionalen Variationsdatenassimilation (3DVAR) durchgeführt, wobei über einen Zeitraum von zehn Stunden jede Stunde eine 3DVAR durchgeführt wurde. Die DA-Experimente mit dem neuen Operator verbesserten die Ergebnisse gegenüber dem zuvor verwendeten Operator für die relative Luftfeuchtigkeit, und die Wasserdampf- und Temperaturanalysen wurden insgesamt optimiert. Die gleichzeitige Assimilation von Temperatur und WVMR führte dabei zu einer geringfügigen Verschlechterung des Temperaturfeldes in der Analyse, während eine Verbesserung des Temperaturfeldes beobachtet wurde, wenn die Temperatur allein assimiliert wurde. Die statische Hintergrundfehler-Kovarianzmatrix (B) in der 3DVAR wurde als Grund für dieses Verhalten identifiziert. Die Korrelation zwischen den Temperatur- und den WVMR-Variablen in der statischen und nicht-strömungsbedingten Hintergrundfehler-Kovarianzmatrix der 3DVAR begrenzte die Verbesserung im Hinblick auf die Temperatur.

Der zweite Teil der Arbeit zeigt eine Lösung zur Überwindung des Problems der statischen B-Matrix auf. Es wurde ein hybrider Ansatz angewandt, der den Ensemble Transform Kalman Filter (ETKF) zusammen mit der 3DVAR verwendet, um der Hintergrundfehler-Kovarianzmatrix eine Strömungsabhängigkeit hinzuzufügen. Das Hybridexperiment führte, im Vergleich zum 3DVAR, zu einem 50% niedrigeren mittleren quadratischen Fehler (RMSE) für die Temperatur und Wasserdampf. Vergleiche mit unabhängigen Radiosondenbeobachtungen zeigten eine Verringerung des RMSE um 26% für Wasserdampf und 38% für die

vi

Temperatur. Vergleiche mit Ceilometern, die während HOPE zur Verfügung standen, zeigten, dass die prognostizierte Höhe der planetarischen Grenzschicht (PBL) deutlich näher an den Beobachtungen war. Der Einflussbereich der Assimilation eines einzelnen Lidar-Vertikalprofils erstreckte sich über einen Radius von 100 km, was für die Assimilation von Wasserdampf- und Temperaturdaten aus operationellen Lidar-Netzwerken für die kurzfristige Vorhersage vielversprechend ist. Eine Verbesserung der Vorhersage bezüglich der Entwicklung der planetarischen Grenzschicht konnte in den ersten 7 Stunden nach der letzten 3DVAR erreicht werden. Ein Vergleich mit vom Globalen Navigationssatellitensystem (GNSS) abgeleiteten Beobachtungen des integrierten Wasserdampfs ergab eine Verbesserung der Vorhersage während der ersten 4 Stunden nach dem letzten 3DVAR. Mit Hilfe von hochentwickelten DA-Systemen und einem robusten Netzwerk von Lidar-Systemen wirft die Arbeit ein Licht auf die Verbesserung der Zukunft der operativen Vorhersage im Nahbereich.

Dedicated to Achan, Amma, Maheshetan and Avinash.

Contents

Chapter 1

Introduction

1.1 Motivation

The charm of weather prediction is the ability to deduce the future state of the atmosphere from the current weather conditions. A brilliant idea was proposed at the turn of the nineteenth century by two of the greatest meteorologists in history, Cleveland Abbe (Abbe, 1901) and Vilhelm Bjerknes (Bjerknes et al., 2009). The revolutionary idea was to predict the state of the atmosphere with current weather conditions based on the laws of physics, assuming weather as an initial value problem. With a lack of computational resources as well as sparse and irregular observations during their times, this idea could not gain much acceptance among the scientific community, but now it is a reality. Currently, weather prediction is performed daily at major operational forecasting centres worldwide.

Although weather forecasts have been predominantly used in the risk and emergency management sector, they are also used in other areas. Many commercial sectors rely on timely weather reports for efficient performance and planning, such as agriculture, energy, transportation, logistics, and utilities. Mitigation and prevention of economic loss are benefits of a good weather forecast. Although the effort to increase the capabilities in numerical weather prediction (NWP) research is higher, it is often substantially lower than the resultant effects due to not having a profound knowledge of the same.

The prerequisite of an advanced weather forecast is observations. The quality of the forecast depends on the quality of observations and how well we treat them. Since we already understood that the weather prediction problem is an initial value problem, we need to get the current observations of the atmosphere, land, and oceans. The estimation of the initial conditions from a set of observations is done by a process called Data Assimilation (DA). DA can be explained as a mathematical process used to determine the most optimal initial state of an NWP model with the help of observations and short-range forecasts from the model. DA is usually done in a sequential mode where a previous model forecast is compared with a new set of observations and then updating the model. A new forecast is initiated from the updated model state, which we call the analysis. This process continues until the forecast skill improves. Hence DA provides an analysis of the initial state of

the land-atmosphere (L-A) system by merging the output of a previous forecast with all relevant observations.

Since the 1950s, with the advancements in scientific and technical infrastructure, numerical weather systems have been developed with increasing complexities (Lynch, 2008). The quality of the forecasts has constantly been improving, even for a forecast beyond two weeks (Bauer et al., 2015; Buizza and Leutbecher, 2015; Straaten et al., 2020). This excellence can be attributed to higher accuracy in the initial conditions of the model with the help of sophisticated DA techniques and a better representation of physical and dynamical processes by the model. If any of the above factors are impaired, the weather forecast skill becomes detrimental. The weather predictability limit of the most advanced NWP model at the European Centre for Medium-Range Weather Forecasts (ECMWF) is up to around ten days (Zhang et al., 2019). The predictability limit was seven days 30 years ago (Simmons and Hollingsworth, 2002; Bauer et al., 2015).

Initial conditions can result in the rapid growth of forecast error, leading to a level beyond which the forecast skill is not valid (Vannitsem, 2017). Degradation of the forecast skill can also result from the errors in boundary conditions (Collins and Allen, 2002; Nicolis, 2007). Model errors include errors arising from a poor representation of the subgrid physical and dynamical processes (Lorenz, 1982; Nicolis et al., 2009). The numerical scheme inaccuracies associated with the model also belong to model errors. All these factors induce errors that amplify rapidly in time due to the chaotic nature of the model.

An approach to improve the initial fields for NWP is by assimilating high-resolution observations related to the model's prognostic variables. Among these atmospheric observation fields, water vapour and temperature are some of the most critical variables determining the evolution of weather systems. The vertical and horizontal distribution of these variables is crucial for the evolution of weather on all spatial and temporal scales. In mesoscale models, the representation of L-A interaction and convection initiation (CI) is deficient, resulting from substantial observational gaps. These observational gaps limit the forecast skill of the numerical weather systems considerably (Wulfmeyer et al., 2015). CI is an important research topic in modelling since NWP models suffer from an inaccurate representation of clouds and precipitation processes. The lower troposphere is the critical region where strong gradients of moisture profiles are visible. Hence, the lower troposphere should be well captured to understand CI. High-resolution, three-dimensional moisture measurements are unavailable for many atmospheric science applications (Weckwerth et al., 1999). High-resolution distribution of water vapour has been identified as the limiting factor for predicting convective precipitation (Council et al., 1998; Emanuel et al., 1995; Dabberdt et al., 1996). With advanced remote-sensing instruments, we have the liberty to acquire temperature and water vapour observations at higher resolution, both temporally and spatially, as required by the NWP models.

The models can produce quality forecasts with advanced DA methodologies and suitable operators. The operators, either forward or inverse, should be designed in such a way that the observations are incorporated into the model optimally. Passive remote-sensing systems, like microwave radiometers (MWRs) and infrared (IR) spectrometers, claim to provide vertical profiles. However, their ability to capture the fine-scale processes in the atmosphere is limited (Wulfmeyer et al., 2015). The vertical resolution of MWRs spans from 300 to 1000 m in the lowest 2000 m above ground (Cadeddu et al., 2002; Blumberg et al., 2015; Wulfmeyer et al., 2015) and higher than 2000m above 2000 m AGL. Wulfmeyer et al. (2015) has shown the distribution of information content as a function of height illustrated by the cumulative degrees of freedom profiles, which show the MWR range is higher than 2000 m above 2000 m above ground level (AGL). IR spectrometers also have limited resolution of several tens of meters close to the land surface degrading to hundreds of meters towards the atmospheric boundary layer (ABL) top (Turner and Löhnert, 2014). On the other hand, active remote-sensing lidar instruments possess substantially higher temporal and vertical resolutions. Temperature rotational Raman lidar (TRL, e.g. Behrendt and Reichardt, 2000; Di Girolamo et al., 2004; Arshinov et al., 2005; Radlach et al., 2008; Hammann et al., 2015; Behrendt et al., 2015) provide temperature profiles, water vapour Raman lidar (WVRL, e.g. Veselovskii et al., 2009; Leblanc et al., 2012; Sakai et al., 2019) or water vapour differential absorption lidar (WVDIAL, e.g. Ismail and Browell, 1989; Wulfmeyer and Bösenberg, 1998; Behrendt et al., 2009; Repasky et al., 2013; Spuler et al., 2015; Weckwerth et al., 2016; Späth et al., 2016) provide moisture profiles. Several Raman lidar systems independently provide water vapour mixing ratio and temperature measurements simultaneously (e.g. Behrendt et al., 2002; Bhawar et al., 2011; Di Girolamo et al., 2017; Lange et al., 2019). A few such combined water vapour and temperature Raman lidar systems operate autonomously in an uninterrupted manner (e.g. Goldsmith et al., 1998; Turner and Goldsmith, 1999; Reichardt et al., 2012; Dinoev et al., 2013; Newsom et al., 2013).

Only limited DA research on the assimilation of thermodynamic lidar profiles has been conducted so far since the thermodynamic systems are few. Most of these were based on lidar data from field campaigns or observing systems simulation experiments such as the International H2O Project (IHOP) 2002 (Weckwerth and Parsons, 2006) in the US, the Convection Storm Initiation Project (CSIP) (Browning et al., 2007) in the UK 2004, the Convective and Orographically-induced Precipitation Study (COPS, Wulfmeyer et al., 2011) in Germany and D-PHASE 2007 (Rotach et al., 2009) in the Alpine region, the Plains Elevated Convection At Night (PECAN) 2015 in the US (Geerts et al., 2017), the Hydrological Cycle in the Mediterranean Experiment (HYMEX) (Drobinski et al., 2014), and soon the Water vapour Lidar Network Assimilation (WaLiNeAs) 2022 (Flamant et al., 2021). There are several promising studies on lidar DA, of which one exciting work is the assimilation of WVDIAL data into a global atmospheric model using **3DVar** DA (Kamineni et al.,

2003; Kamineni et al., 2006). The research showed improvements in hurricane track fore-casts. During NASA's airborne Lidar Atmospheric Sensing Experiment (LASE), within the IHOP campaign, WVDIAL data was assimilated into the 5th generation Pennsylvania State University-National Centre for Atmospheric Research Mesoscale Model (MM5). The assimilation was conducted using a **4DVar** (Wulfmeyer et al., 2006) which demonstrated significant improvements in the analyses and prediction of water vapour and wind fields and CI. Another important study was the assimilation of lower-tropospheric water vapour fields measured by two airborne lidars into the **3DVar** of the Application of Research to Operations at MEsoscale (AROME) NWP mesoscale model during the COPS measure-ment campaign (Bielli et al., 2012). The assimilation impact was positive; however not significant for the 24 h precipitation. It was found positive and significant for the six-hour precipitation, with 24 hours lasting improvement. The critical finding was that selected case studies showed improvement by better depicting CI and more accurate positioning of the precipitation systems. Similarly, Grzeschik et al. (2008) assimilated water-vapour data from three ground-based WVRLs, aligned in a triangle, into the MM5 model. The data was obtained from the Lindenberg Campaign for Assessment of Humidity and Cloud Profiling Systems and its Impact on High-Resolution Modelling (LAUNCH, Engelbart and Haas, 2006). The DA corrected the initial water vapour field by about 1 g kg^{-1}. It was detectable in the forecast up to 12 h. The first assimilation of TRL data into the WRF model using the **3DVar** DA system (Adam et al., 2016) showed improvements in the temperature analyses combined with corrected PBL heights. Yoshida et al. (2020) conducted OSSEs assimilating water vapour profiles. The Raman lidar data was assimilated into the Japan Meteorolog-ical Agency Non-Hydrostatic Model (JMA-NHM, Saito et al., 2006) through the LETKF DA system. A precipitation forecast was improved on the windward side of developing precipitation events.

With the promising impacts of lidar DA research, an efficient DA system for the optimal use of high-resolution data is necessary. For this purpose, a robust operator for the direct assimilation of lidar data products needs to be developed. In the WRFDA system, the water vapour mixing ratio (WVMR) or other moisture variables are assimilated with the help of a single radiosonde relative humidity (RH) operator. It is sub-optimal because RH is strongly sensitive to temperature. Therefore, a new forward operator for the direct assimilation of absolute humidity, mixing ratio or specific humidity needs to be developed. The value of these observations is independent of any cross-sensitivity to the temperature, which is the sole criteria for optimum use of information. The operator enables the models to acquire the complete information content of the observations. The assimilation impact of high-resolution lidar systems is a crucial topic that needs more research and exploration. Future operational forecasting can be revolutionary with the implementation of not only thermodynamic lidar systems but also the retrieval of other atmospheric variables through the active remote sensing instruments on an operational level around the world. With this

motivation, the thesis will address specific questions about using thermodynamic lidar profile data from a single ground-based lidar system. This thesis will study the impact of assimilating thermodynamic profiles with advanced deterministic and probabilistic DA methods. The objectives of the thesis are listed in the following subsection of this chapter.

1.1.1 Objectives of the PhD thesis

The thesis revolves around a central aim to study the impact of high-resolution, thermodynamic lidar systems on a convection-permitting scale weather prediction model and lay a foundation for the potential use of lidar systems for operational weather prediction. The promising results from this thesis will be an excellent contribution to the research community focused on the development of lidar DA techniques. The thesis is structured as a cumulative thesis with two pivotal publications addressing the objectives and an application-oriented publication.

The objectives of the thesis are addressed by two publications here. The first publication details the development and implementation of the forward operator and the assimilation of lidar thermodynamic profile observations into a deterministic **3DVar** DA system. The objectives of the first publication are described in chapter two and are as follows:

1. To develop a forward operator for the direct assimilation of WVMR and temperature data and to find whether the operator provides a reasonable impact on the analyses of the WV field.

2. To study the impact of moisture and temperature assimilation and the simultaneous assimilation of both.

3. To study the inter-dependency of moisture and temperature in the DA system.

The second publication introduces the ensemble approach to the assimilation of lidar thermodynamic profiles using a hybrid **3DVar-ETKF** DA system and short-range forecasts. The objectives are described in chapter three and are:

1. Does the addition of flow-dependency in the B matrix improve the analyses?

2. What is the spatial extent of the thermodynamic lidar profiles assimilation impact from a single geographical location using an ensemble approach versus a purely deterministic approach?

3. Up to what time range is a forecast from a single ground-based lidar system treated valid with respect to observations?

4. Will these results pave the way for future networks of lidar systems in operational forecasting?

Chapter four describes the potential application of lidar DA for a network of lidar systems in the future. Since the chapters are in the publication format, the reader is free to read the chapters independently of the others. The following sections of this chapter give the reader a brief, conceptual understanding of the DA methods, lidar systems and the forward operator.

1.2 Data Assimilation

The beauty of DA lies in its interdisciplinary nature of combining different areas of science such as numerical methods, algebra, probability and statistics, and dynamical systems. DA is a mathematical tool to estimate the state of the system at a point in the past, present or future. Although DA has applications in other areas of science and economics, the most crucial application of DA is weather prediction. In the weather prediction problem, we rely on two sources of information: the observations and the model. DA extracts synergies between model and observations by exploiting their respective informational content. From a statistical perspective, the definition of DA is "an approach for fusing data (observations) with prior knowledge (e.g., mathematical representations of physical laws; model output) to obtain an estimate of the distribution of the true state of a process (Wikle and Berliner, 2007)."

DA techniques consistently developed from the past discovering new methodologies since the 1950s. One of the earliest DA techniques is Newtonian relaxation or nudging (Hoke and Anthes, 1976; Lakshmivarahan and Lewis, 2013). In this technique, a relaxation term is added to the prognostic equation that nudges the solution towards the observations interpolated to the model grid (Kalnay, 2003). With this advancement, Bayesian strategies took control of the solution of the DA problem, which gave a start to the variational and ensemble-based DA methods. Variational (Var) DA works on implementing a constraint-based cost function minimisation approach. Var DA includes three-dimensional (3-D) Var and four-dimensional (4-D) Var. The Var provides a single, quasi-optimal analysis state as the output based on an a priori state or the background state (or forecast from the previous analysis state), a set of observations, and the prescribed Gaussian uncertainty statistics for the background and observations (Bannister, 2017). The ensemble-based DA methods are based on the ensemble Kalman filter (EnKF) that offers a priori statistics instead of a priori state in the Var approach. The error statistics are obtained from an ensemble instead of a prescribed source, which changes as the system evolves. A combination of the Var and ensemble-based approaches is now being used, the hybrid DA method, exploiting the advantages of both systems.

The three main DA strategies currently employed in the significant operational forecast centres are: (1) variational approaches, namely Three-Dimensional VARiational (**3DVar**, Courtier et al., 1998; Barker et al., 2004) DA and Four-Dimensional VARiational (**4DVar**, Huang et al., 2009) DA; (2) ensemble-based approaches using several variants of the ensemble Kalman filter (Evensen, 2003), and (3) hybrid combinations of the two approaches mentioned above (Clayton et al., 2013; Ingleby et al., 2013). Another option that facilitates the representation of the likely non-linearities of the weather forecasts, the particle filter, is still too demanding and in the research state (Milan et al., 2014; Poterjoy et al., 2017; Poterjoy et al., 2019; Potthast et al., 2019; Kawabata and Ueno, 2020).

Many operational forecasting centres are working towards the research, development, and service providers of weather forecasts at the regional level. The most important operational centres with their DA methodologies are:

1. **ECMWF:** European Center for Medium Weather Range Forecasting.
   ECMWF uses a **4DVar** DA system. Currently, work to implement the Object-Oriented Prediction System (OOPS) as a new **4DVar** data assimilation framework for ECMWF's Integrated Forecasting System (IFS) is in progress. OOPS is an international effort involving significant input from ECMWF, Météo-France and the HIRLAM-ALADIN community.

2. **DWD:** Deutscher Wetterdienst or the German Meteorological Service.
   The DWD started to use a regional numerical weather prediction (NWP) model called Icosahedral Nonhydrostatic Limited Area Model (ICON-LAM) in 2021, replacing the COSMO model. A novel Ensemble VARiational (EnVAR) setup is currently evaluated with stochastic physics perturbations implemented by DWD and compared to the deterministic operational run of the Kilometer Scale Ensemble Data Assimilation-Local Ensemble Transform Kalman Filter (KENDA-LETKF).

3. **Météo-France** or the national weather service of France.
   The Météo-France Applications of Research to Operations at Mesoscale–Ensemble Prediction System (AROME-EPS, Bouttier et al., 2012; Vié et al., 2012) applies Stochastic Perturbed Parameterization Tendencies (SPPT) and physics parameter perturbations in turbulence which provided promising impacts. Currently, Météo-France uses the **3DVar** for their operational forecast (Brousseau et al., 2016). Météo-France also incorporates a 3DEnVar DA sytem and its operational implementation is summarised in (Montmerle et al., 2018). With the WaLiNeAs initiative, an operational **4DEnVar** DA system will be tested.

4. **ACCORD:** A Consortium for COnvection-scale modelling Research and Development' or the ACCORD consortium is a partnership of three regional consortiums. They are: (1) the High Resolution Limited Area Model (HIRLAM) Aire Limitée Adaptation dynamique Développement InterNational (ALADIN) Research on Mesoscale Operational NWP In Europe or in short the HARMONIE model; (2) the ALADIN, and; (3) the Regional Cooperation for Limited-Area modelling in Central Europe (RC LACE) consortia. HARMONIE, ALADIN, and RC LACE merged in the year 2020 to form the ACCORD consortium. ACCORD is used for the operational short-range weather forecasts in Denmark, Estonia, Finland, Iceland, Ireland, Lithuania, the Netherlands, Norway, Spain, Sweden, Algeria, Austria, Belgium, Bulgaria, Czech Republic, Croatia, France, Hungary, Morocco, Poland, Portugal, Romania, Slovakia,

Slovenia, Turkey and Tunisia. An Incremental **3DVar** method is implemented in the operational model.

5. **JMA:** Japan Meteorological Agency.
   The operational mesoscale data assimilation system incorporated a **4DVar** method based on JMA's ASUCA non-hydrostatic model. JMA has applied its own hybrid method composed of a LETKF (Hunt et al., 2007) and 4D-Var within its operational system since December 2019.

6. **NOAA:** National Oceanic and Atmospheric Administration of the USA - Incremental hybrid **3DEnVar** and non-variational cloud analysis.
   The NOAA NCEP provides high-resolution CP forecasts for the USA in the US with the 3 km CON-terminous United States-North American (CONUS-NAM, Wu et al., 2017) and the High-Resolution Rapid Refresh (HRRR, Benjamin et al., 2016).

7. **UK Met Office:** National weather service of the UK.
   The Met Office Global and Regional Ensemble Prediction System (MOGREPS-UK, Tennant, 2015; Hagelin et al., 2017) applies perturbations to the analysis, which are consistent with the initial conditions DA (Bowler et al., 2008) in a hybrid **4DVar** DA system.

The **3DVar** approach assimilates data at specific analysis time-steps. The **4DVar** assimilates observations during a pre-defined DA time window with an adjoint model to minimise a cost function. Ensemble DA systems incorporate flow dependency in the background error covariance (B) matrix estimated from the ensemble (Houtekamer and Mitchell, 1998; Bishop et al., 2001; Evensen, 2003), which is superior to a purely deterministic DA system (Hamill and Snyder, 2000; Buehner, 2005; Wang et al., 2008a; Wang et al., 2008b; Kutty and Wang, 2015; Kutty et al., 2018) because it represents more realistically the "error of the day" in our chaotic weather system. The hybrid technique combines the flow-dependent error covariance matrix with a static error covariance matrix. The main advantage of a hybrid DA system over **3DVar** is that it uses a flow-dependent ensemble covariance rather than a static background error covariance. Also, since the analysis is based on Var, the filter divergence is minimal. Compared to EnKF (Evensen, 2003), a smaller number of ensemble members is required in hybrid DA (Wang et al., 2007; Zhang et al., 2013; Pan et al., 2014) because static error covariance is also incorporated in addition to the ensemble covariance. Through the extended control variable approach (Lorenc, 2003), the pure **3DVar** cost function is adapted to include the ensemble covariance. We will describe the DA methods in detail in the next section.

1.2.1   Variational Data Assimilation: 3DVar and 4DVar

In order to produce the best estimate of the current state of the system and accurately predict future states variational DA approach is widely used in numerical weather prediction combining observations with a model forecast. An iterative optimisation technique is used to minimise a weighted nonlinear least-squares measure of the error between the model forecast and the available observations to determine the estimate. The solution to the variational DA problem, the analysis, is the maximum a posteriori Bayesian estimate of the state of the system. An incremental version of the variational scheme is currently implemented in many operational centres. An approximate Gauss-Newton method derived analysis is determined by solving a nonlinear least-squares problem in the incremental method.

Bayes theorem describes the joint probability P for the occurrence of two events. If we consider $\vec{x}_b$ the background state vector and $\vec{y}$ the observation state vector,

$$P\left(\vec{y}\right) \cdot P\left(\vec{x}_b | \vec{y}\right) = P\left(\vec{y} | \vec{x}_b\right) \cdot P\left(\vec{x}_b\right) \tag{1.1}$$

where the $P\left(\vec{y}\right)$ is the a priori probability density function (pdf) of the observation $\vec{y}$, $P\left(\vec{x}_b | \vec{y}\right)$ denotes the conditional probability to determine $\vec{x}_b$ for a given observation $\vec{y}$, $P\left(\vec{y} | \vec{x}_b\right)$ shows the probability of the observation $\vec{y}$ given the background $\vec{x}_b$, and finally $P\left(\vec{x}_b\right)$ is the a priori pdf of the background $\vec{x}_b$. However, for a measured observation, we have $P\left(\vec{y}\right)$ equal to one. Hence we can rewrite the equation as:

$$P\left(\vec{x}_b | \vec{y}\right) = P\left(\vec{y} | \vec{x}_b\right) \cdot P\left(\vec{x}_b\right) \tag{1.2}$$

For DA systems, we assume the pdfs to have Gaussian error characteristics. Under this valuable assumption, we can write the prior pdf of the model state as,

$$P\left(\vec{x}\right) \propto \exp\{-\frac{1}{2}\left(\vec{x} - \vec{x}_b\right)^T \mathbf{B}^{-1}\left(\vec{x} - \vec{x}_b\right)\} \tag{1.3}$$

where we encounter a significant term **B** which is the covariance of the background error. To explain the term background error covariance we have to look into a nonlinear dynamical system. Consider a nonlinear model operator M that evolves model state at $t = n$ to $t = n + 1$. This can be mathematically represented as,

$$\vec{x}|_{t=n+1} = M\vec{x}|_{t=n} + \epsilon^m|_{t=n} \tag{1.4}$$

where ϵ^m is the model error. Although DA systems have model error, while explaining DA methods we assume the model to be perfect thereby treating $\epsilon^m = 0$. Now assuming we have $\vec{y}$ observations at $t = n$, the true state of the system $\vec{x}_t$ can be written in terms of the following equation.

$$\vec{y}|_{t=n} = H\left(\vec{x}_t\right)|_{t=n} + \epsilon^o|_{t=n} \tag{1.5}$$

A new term in Eq. 1.5, H, is the nonlinear observation operator that maps the true state into observation space. The observation error is denoted by the term ϵ^o. The mean of the observation error is assumed to be zero. The observation error covariance matrix, denoted by $\mathbf{R}$, if $E(.)$ is the expected value is given by:

$$\mathbf{R} = E(\epsilon^o \cdot \epsilon^{oT}) \tag{1.6}$$

Now coming to the background error, denoted by ϵ^b, is related to the true state and the background as:

$$\vec{x_b}|_{t=n} = \vec{x_t}|_{t=n} + \epsilon^b|_{t=n} \tag{1.7}$$

Assuming the expected value of the background error to be zero, the background error covariance matrix is given by:

$$\mathbf{B} = E(\epsilon^b \cdot \epsilon^{bT}) \tag{1.8}$$

The $\mathbf{B}$ matrix is static and reflects a climatological error covariance. $\mathbf{B}$ can be calculated mainly by three methods, namely, the NMC-method (Parrish and Derber, 1992), the analysis ensemble-method (Fisher, 2003), and by using innovation statistics (Hollingsworth and Lönnberg, 1986). All these methods have their own merits. The NMC-method is the most widely used method for the generation in which climatological background error covariances are estimated. This thesis uses the NMC method since it provides physically reasonable results in regional model domains and is computationally less expensive than the ensemble method. In the NMC-method, forecast difference statistics are computed, from which the forecast error covariance is derived. The forecast error covariance is derived from incorporating the errors applicable to that domain. However, the NMC method has certain drawbacks. It overestimates the covariances at large scale simulations and poorly observed regions (Berre, 2000; Fisher, 2003; Pereira and Berre, 2006).

Now having the knowledge of the covariance matrices, we come back to Eq. 1.3 to proceed finding the probability of the observation occurring given the model state. This can be written as,

$$P(\vec{y}|\vec{x}) \propto \exp\{-\frac{1}{2}(\vec{y} - H\vec{x})^T \mathbf{R^{-1}}(\vec{y} - H\vec{x})\} \tag{1.9}$$

Using the equations 1.3 and 1.9, and applying the Bayes' Theorem with the assumption that the errors in observations and background are independent we derive an expression for the posterior pdf of the state given the observations. The relation reads,

$$P(\vec{x}|\vec{y}) = \exp\{-\frac{1}{2}(\vec{x} - \vec{x_b})^T \mathbf{B^{-1}}(\vec{x} - \vec{x_b}) - \frac{1}{2}(\vec{y} - H\vec{x})^T \mathbf{R^{-1}}(\vec{y} - H\vec{x})\} \tag{1.10}$$

Maximising the value of $P(\vec{x}|\vec{y})$, known as the maximum a posteriori (MAP) estimate, will give us the estimate of the analysis. From Eq. 1.10, the maximum value of $P(\vec{x}|\vec{y})$ is attained when the exponential argument is close to zero. This is equivalent to determining

the minimum variance by minimising the cost function,

$$J(\vec{x}) = \frac{1}{2}(\vec{x} - \vec{x}_b)^T \mathbf{B^{-1}}(\vec{x} - \vec{x}_b) + \frac{1}{2}(\vec{y} - H\vec{x})^T \mathbf{R^{-1}}(\vec{y} - H\vec{x}) \tag{1.11}$$

The equation estimates the analysis based on the distance between the solution and the background that is weighted by the error in the background and the distance between the observations and the solutions weighted by the error in the observations. This is the basic formulation of the **3DVar**. If we try solving the **3DVar** equation explicitly, the analysis $\vec{x}_a$ is given by,

$$\vec{x}_a = \vec{x}_b + \mathbf{K}(\vec{y} - H(\vec{x}_b)) \tag{1.12}$$

where $\mathbf{K}$ is the Kalman gain matrix given by $\mathbf{K} = \mathbf{BH^T}(\mathbf{HBH^T} + \mathbf{R})^{-1}$. $\mathbf{H}$ is the linearised observation operator. A sequential application of the analysis and forecast steps called the best linear unbiased estimator (BLUE, Lewis et al., 2006) is explained below for an efficient cycling of the **3DVar** DA system.

1. Analysis calculation using Eq. 1.12:

2. The analysis is forecasted to obtain the background of the next time step:

$$\vec{x}_b|_{t=n+1} = M(\vec{x}_a|_{t=n}) \tag{1.13}$$

The cost function Eq. 1.11 can be minimised numerically using gradient descent algorithm over several iterations. The gradient of the cost function is derived with respect to the model state as,

$$\nabla J(\vec{x}) = \mathbf{B}^{-1}(\vec{x} - \vec{x}_b) - \mathbf{H}^T \mathbf{R}^{-1}(\vec{y} - H(\vec{x})) \tag{1.14}$$

The **3DVar** cost function in Eq. 1.11 can be rewritten in terms of the linearised observation operator $\mathbf{H}$ as,

$$\mathbf{J}(\vec{x}) = \frac{1}{2}(\vec{x} - \vec{x}_b)^T \mathbf{B^{-1}}(\vec{x} - \vec{x}_b) + \frac{1}{2}(\vec{y} - \mathbf{H}\vec{x})^T \mathbf{R^{-1}}(\vec{y} - \mathbf{H}\vec{x}) \tag{1.15}$$

3DVar is a more effective method for large systems since it is computationally less expensive than other methods like **4DVar** and ensemble-based methods. Another assumption in **3DVar** is that all observations are valid at one time rather than over a time window around the assimilation time which does not take advantage of a time series measurement.

The equations of **3DVar** must be extended to take into account the time dependence of the observations. The extended method is **4DVar** (Huang et al., 2009) that uses the dynamical model to allow observations to be assimilated at the correct time. **4DVar** ingests observations at multiple times across a specified time window. The **4DVar** technique aims to minimise a four-dimensional cost function with respect to the initial model state. The

cost function of the **3DVar** is modified to incorporate the observations along with the time information as,

$$J(\vec{x}) = \frac{1}{2} (\vec{x} - \vec{x_b})^T \mathbf{B}^{-1} (\vec{x} - \vec{x_b}) + \frac{1}{2} \sum_{i=0}^{N} (\vec{y_i} - H\vec{x_i})^T \mathbf{R}^{-1} (\vec{y_i} - H\vec{x_i}) \qquad (1.16)$$

4DVar is a computationally expensive operation if the domain is large and observations are ample. Also, **4DVar** uses tangent linear and adjoint models (Errico, 1997) to produce the propagation of analysis increment over the assimilation window, which requires additional operations which are computationally more expensive compared to **3DVar**.

1.2.2 Ensemble-based Data Assimilation

The atmosphere is a chaotic system whose actual state cannot be accurately determined. Therefore it is likely that initial conditions used to initiate a forecast will contain errors. Large deviations in the forecast can arise from small perturbations in the initial conditions. Hence it is imperative to represent the uncertainty in the initial conditions. One approach is to represent uncertainty in the initial conditions by considering the prior distribution of the initial state via an "ensemble" of initial conditions. Forecasting each ensemble member results in an ensemble of forecasts combined with observations. Ensemble DA schemes are anchored upon this principle.

The main advantage of using flavours of ensemble Kalman filters is that it provides estimates of the pdfs associated with the analysis. The ensembles give information on the uncertainty of the forecast since they provide a Monte Carlo estimate of the evolution of the pdf using the forecast model. Most beneficial background error covariance information at each time can be determined by reducing the costly computations of the error covariance forecast and the error covariance update. However, a reasonable estimate for this flow-dependent background matrix is obtained when there are many ensemble members. Otherwise, the flow-dependent error covariance matrix is contaminated with sampling errors. The larger the size of the model state, the higher is the ensemble size. If fewer ensemble members are used, the ensemble will not be statistically representative of the background error resulting in undersampling causing filter divergence. Hybrid DA methods combine the best of the variational and ensemble techniques. A detailed formulation of the hybrid DA methodology is introduced in the next section, which is the principal DA methodology used in this thesis.

1.2.3 Hybrid Data Assimilation

The Hybrid **3DVar-ETKF** DA technique (Wang et al., 2008a; Wang et al., 2008b) was used for assimilating thermodynamic lidar data in the second part of the thesis explained in the

third chapter. The results were compared with those of the **3DVar** DA system. The main advantage of a hybrid DA system over a **3DVar** (e.g. Barker et al., 2004) DA system is using a flow-dependent ensemble covariance matrix rather than a static background error covariance.

A brief but conceptual description of how the hybrid **3DVar** cost function is applied in the WRFDA framework and the transformation of ensemble members are given in this section. According to Wang et al. (2008), the hybrid **3DVar** cost function is defined as follows:

$$\mathbf{J}\left(\vec{x_1},\vec{a}\right) = \beta_1 \mathbf{J_1} + \beta_2 \mathbf{J_e} + \mathbf{J_o} \tag{1.17}$$

where $J\left(\vec{x_1},\vec{a}\right)$ is the hybrid cost function expressed as the sum of the observational component J_o and the background components J_1 (static component) and J_e (ensemble component) extended by the control variables and applied with appropriate weighting factors β_1 and β_2. The hybrid cost function $J\left(\vec{x_1},\vec{a}\right)$ is a function of $\vec{x_1}$ and $\vec{a}$. $\vec{x_1}$ is the analysis increment due to the static background error covariance and $\vec{a}$ is the concatenated vector of extended control variable for each ensemble member. The weighting factors β_1 and β_2 have to match the following rule:

$$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1 \tag{1.18}$$

The extended control variables in the WRF hybrid system are the three-dimensional wind components **U**, **V**, **W**, the perturbation geopotential variable **PH**, the perturbation dry mass in column **MU**, temperature **T**, and water vapour mixing ratio **QVAPOUR**. The cost function in Eq. 1.17 then expands to

$$\mathbf{J}\left(\vec{x_1},\vec{a}\right) = \beta_1 \frac{1}{2} \left(\vec{x_1}\right)^T \mathbf{B^{-1}} \left(\vec{x_1}\right) + \beta_2 \frac{1}{2} \sum_{k=1}^{N} \left(\vec{a_k}\right)^T \mathbf{C^{-1}} \left(\vec{a_k}\right) + \frac{1}{2} \left(\vec{y^o} - H\vec{x}\right)^T \mathbf{R^{-1}} \left(\vec{y^o} - H\vec{x}\right) \tag{1.19}$$

where $\vec{x}$ is the total analysis increment vector. $\vec{y^o}$ is the innovation vector where $\vec{y^o} = \vec{y} - H\left(\vec{x_b}\right)$. $\vec{y}$ denotes the observations assimilated and $\vec{x_b}$ denotes the background state vector. H is the nonlinear observation forward operator which maps the model grid points to the observations. **B** and **R** are static error covariance matrices of the background and observations, respectively. **B** is an almost diagonal matrix in the control variable space whereas **R** is assumed to be a fully diagonal matrix since the observations are assumed mutually independent. However, this is not the case when the observations are correlated. Vertical and horizontal correlations can appear as the off-diagonal terms due to correlations of observed variables. The correlation is inevitable and weaker for lidar-observed variables but predominantly stronger in passive remote-sensing systems due to the coarser resolution compared to active remote-sensing systems. **C** is the covariance matrix containing the flow dependent error covariance.

The analysis increment ($\vec{x}$) in the hybrid cost function is the sum of increments associated with static background error covariance ($\vec{x_1}$) and one with flow-dependent ensemble covariance.

$$\vec{x} = \vec{x_1} + \sum_{k=1}^{N} \left(\vec{a_k} \circ \vec{x_k^e} \right) \tag{1.20}$$

The second term in Eq. 1.20 contains the information of the ensemble members at each assimilation time-steps. $\vec{a_k}$ is the extended control variable for each ensemble member, where, k is denoted for the kth ensemble member. $\vec{x_k^e}$ is the kth ensemble perturbation normalized by $\sqrt{N-1}$ where N is the ensemble size, which expands to:

$$\vec{x_k^e} = \frac{\vec{x_k} - \bar{x}}{\sqrt{N-1}} \tag{1.21}$$

where $\bar{x}$ is the mean of N ensemble forecasts. The symbol $\circ$ denotes the Schur product which operates as element by element product. The term $\sum_{k=1}^{N} \left(\vec{a_k} \circ \vec{x_k^e} \right)$ is computed at every assimilation time-step so that the flow-dependent errors are retained in the background error covariance matrix.

The ensemble transformation for the cycling of the DA system starts with the computation of perturbations of each ensemble member. Therefore, the ensemble mean and the corresponding variance of each member are calculated. The difference between each ensemble member and the ensemble mean is the so-called member perturbation. The ensemble perturbations are updated by the ETKF (Bishop et al., 2001). ETKF is a type of an ensemble square root filter (EnSRF, Tippett et al., 2003). To derive the analysis perturbation matrix $\mathbf{X}^a$, we multiply the ensemble perturbation matrix $\mathbf{X}^e$ (which is also the background perturbation matrix for the DA cycle) with a matrix of weights or a transformation matrix denoted by $\mathbf{T}$ which then reads:

$$\mathbf{X}^a = \mathbf{X}^e \mathbf{T} \tag{1.22}$$

The weight matrix is given by:

$$\mathbf{T} = r\mathbf{E} \left(\rho\lambda + \mathbf{I} \right)^{-\frac{1}{2}} \mathbf{E}^T \tag{1.23}$$

where $\mathbf{E}$ and λ are the Eigen vectors and Eigen values of the $N \times N$ matrix $\frac{(\mathbf{X}^e)^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{X}^e}{N-1}$. $\mathbf{X}^e$ is the ensemble perturbation matrix formed by the horizontal concatenation of perturbation vectors, $\vec{x_k^e}$, where k extends from 1 to N. Here $\mathbf{H}$ is the linearized observation forward operator. The factors r and ρ are scalar tunable factors which are the inflation factor r and the fraction of the forecast-error covariance projected onto the ensemble subspace ρ. In an adaptive inflation configuration, both the tunable factors are set to one. The inflation factors stabilize to a nearly constant value after a series of DA cycles.
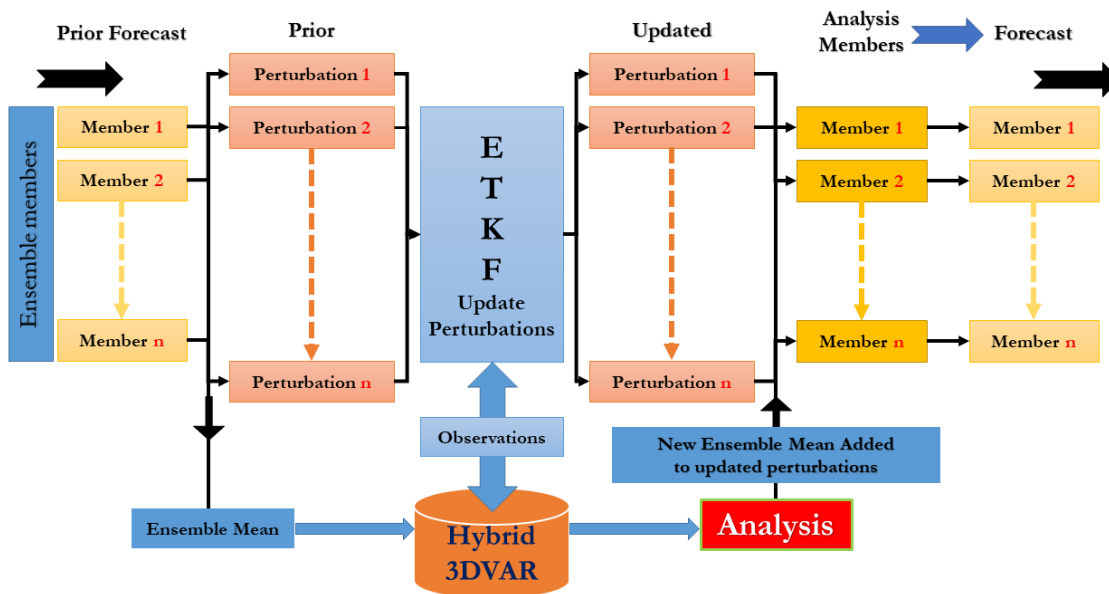
FIGURE 1.1: Schematic diagram of the hybrid **3DVar-ETKF** DA cycle. The
n forecast ensemble members from the first guess or the previous analysis
forecast cycle are the input for the subsequent cycle. The ETKF updates
prior perturbations to analysis perturbations. The analysis perturbations are
added to the updated mean to get a new set of analysis members for the next
DA cycle.

The columns of the analysis perturbation matrix are the new set of updated ensemble perturbations. The ensemble mean is applied as the background state $\vec{x}_b$ in the hybrid **3DVar** cost function Eq. 1.19 to get an updated ensemble mean. The updated perturbations from the ETKF are added to the updated mean to get the set of new analysis ensemble members. A free forecast is performed with the new set of ensemble members until the next assimilation time step. A flow chart for the ensemble treatment at each assimilation cycle is given in Fig. 1.1. First of all, in the initial input of the ETKF, perturbations of the ensemble members are produced. Then the arithmetic mean of the ensemble members is calculated. The difference between each ensemble member and the ensemble mean gives the corresponding member perturbation. The ETKF then updates the perturbations. The ensemble mean is the first guess for assimilation with the Hybrid **3DVar** cost function to get the analysis, which is the new ensemble mean. The updated perturbations are added to the analysis to produce the set of new analysis ensemble members. The analysis ensemble members then drive free forecasts until the next assimilation step.

Like the hybrid **3DVar-ETKF**, **ETKF-4DVar** also uses the same approach but with the **4DVar** framework. **ETKF-4DVar** uses a cost function written as,

$$J(\vec{x}) = \frac{1}{2}(\vec{x} - \vec{x}_b)^T \tilde{\mathbf{B}}^{-1}(\vec{x} - \vec{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\vec{y}_i - H\vec{x}_i)^T \mathbf{R}^{-1}(\vec{y}_i - H\vec{x}_i) \tag{1.24}$$

where all the variables are similar to that of the **4DVar** cost function in Eq. 1.16 except the background error covariance matrix $\tilde{\mathbf{B}}$. Here $\tilde{\mathbf{B}}$ is represented as a function of the climatological background error covariance **B** and the background error covariance matrix generated by the ETKF $\mathbf{P^b}$ given by the equation:

$$\tilde{\mathbf{B}} = \frac{1}{2}\mathbf{B} + \frac{1}{2}\mathbf{P^e} \tag{1.25}$$

A more popular version is the ensemble variational method, the **4DEnVar** illustrated in Fairbairn et al. (2014) and Liu et al. (2008). In **4DEnVar**, the temporal evolution of the background error covariance is determined using the covariances of ensemble perturbations at discrete times. The tangent linear and adjoint model of the forecast model is avoided in the **4DEnVar**. The **4DEnVar** approach has been demonstrated in several studied like Buehner et al. (2010a), Buehner et al. (2010b), Wang and Lei (2014), and Lorenc et al. (2015). The DA system of the Canadian operational global numerical weather prediction model (Buehner et al., 2010b), and the U.S. Navy's operational ensemble forecasting system (Bishop and Hodyss, 2011) uses this technique. The cost function of a **4DEnVar** is written as,

$$J(\vec{x}) = \frac{1}{2}(\vec{x} - \vec{x}_b)^T (\mathbf{P^e})^{-1}(\vec{x} - \vec{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\vec{y}_i - H\vec{x}_i)^T \mathbf{R}^{-1}(\vec{y}_i - H\vec{x}_i) \tag{1.26}$$

where again all the variables are similar to the **4DVar** equation Eq. 1.16 except the **B** matrix is replaced by the ensemble error covariance matrix $\mathbf{P^e}$. A detailed explanation of the different DA techniques is given in Goodliff et al. (2015).

1.3 Lidar systems: Differential absorption lidar and Temperature Raman lidar

This research study uses profile observations of temperature and moisture from two lidar systems: temperature Raman lidar and differential absorption lidar, respectively. Both the lidar systems were designed and developed, as well as are owned and operated by the UHOH. Raman lidar can measure both temperature and WVMR, and DIAL can measure moisture in the form of absolute humidity. The principles of operation of the two lidar systems are entirely different.

The principle of operation of the DIAL depends on two laser signals: online P_{ON} and offline P_{OFF} (Bösenberg, 1998; Wulfmeyer and Bösenberg, 1998). The P_{ON} signal's wavelength is tuned to a strong water vapour absorption line, resulting in a smaller backscatter signal. The P_{OFF} signal is tuned for weak absorption. The equation below derives the number density of the water vapour molecules using the online and offline laser pulses,

$$N_{WV}(r) = \frac{1}{2\left(\sigma_{ON}(r) - \sigma_{OFF}(r)\right)} \frac{d}{dr} \ln\left(\frac{P_{OFF}(r) - P_{B,OFF}}{P_{ON}(r) - P_{B,ON}}\right) \tag{1.27}$$

where N_{WV} is the number density of water vapour in the atmosphere, σ is the absorption cross section, P_B is the background signal, and r is the distance from lidar transmitter to the scattering volume along the line of sight. More details of the DIAL system of the UHOH can be found in Wagner et al. (2013) and Späth et al. (2016).

The UHOH TRL uses the rotational Raman technique (Cooney, 1972; Behrendt and Reichardt, 2000; Behrendt et al., 2004) to measure atmospheric temperature profiles. The method relies on temperature-dependent inelastic scattering in the atmosphere that occurs due to the emitted UV laser pulses. The rotational Raman spectrum of air consists of the Stokes and the anti-Stokes branches. The Stokes branch is found at wavelengths greater than the incident radiation, while the anti-Stokes branch is at shorter wavelengths. The UHOH TRL extracts only signals of the latter in order to avoid errors due to atmospheric fluorescence. The temperature is determined with the ratio Q of the two background-corrected Raman signals, $RR2$ and $RR1$. P_{RR2} and P_{RR1} are the signals for low and high quantum number transition settings of the filter, respectively, so that

$$Q(r) = \frac{P_{RR2}(r)}{P_{RR1}(r)} \tag{1.28}$$

Then the temperature profile can be derived as,

$$T(r) = \frac{b}{a - \ln Q(r)} \tag{1.29}$$

where a and b are calibration constants. The statistical error of the temperature measurements is calculated by Poisson statistics derived from the signal intensities of the photon-counting data. The statistical error of a single standard deviation σ for a signal count number **s** is given by the square root of **s**. These error characteristics are detailed in Behrendt and Reichardt (2000), Behrendt et al. (2015), and Wulfmeyer et al. (2016).

1.4 The Forward Operator

The direct assimilation of WVMR was performed via a new forward operator developed in the WRFDA system. This new forward operator enables using WVMR data directly without converting to RH, for which temperature data is also needed. So far, the WRFDA system ingests humidity data in the form of RH through the conventional radiosonde operator. All the vertical profile data products, namely radiosondes, ground-based microwave radiometers, and other humidity profiling instruments, have been using the radiosonde operator to assimilate humidity in the form of RH (Bielli et al., 2012). The sole RH operator for all the moisture observations was a severe limitation of the WRFDA system. Expressing moisture in the form of WVMR has benefits. The quantity is not affected by the atmospheric temperature and pressure changes enabling maximum moisture information content to be passed into the model in the area of interest and ruling out unnecessary cross sensitivities.

The RH operator used to assimilate moisture in the WRFDA system depends on the relationship,

$$\mathbf{RH} = \frac{m}{1+m} \frac{R_W}{R_L} \frac{p}{E(T) \left[1 + \frac{m}{1+m} \frac{R_W - R_L}{R_L}\right]} \approx 1.607 m \frac{p}{E(T)} \qquad (1.30)$$

where WVMR is denoted by m in units $kgkg^{-1}$, T is the temperature in units of K, p is the atmospheric pressure in units of Pa, and R_W and R_L are the specific gas constants of water vapour and dry air respectively in units of $Jkg^{-1}K^{-1}$. From the final approximated form of the Eq. 1.30, we know that RH has a strong dependency on the factors m, p, and the water vapour saturation pressure $E(T)$, of which $E(T)$ depends upon the temperature parameter. The relationship of $E(T)$ (Bolton, 1980) is given by,

$$E(T) = 611.2 \exp\left[17.67 \left(\frac{T - 273.15}{T - 29.65}\right)\right] \ [Pa] \qquad (1.31)$$

where T is in unit K. Hence the sensitivity of temperature in Eq. 1.30 through Eq. 1.31 is comparable to the sensitivity of WVMR m. The quantification of the sensitivity is calculated by finding a total derivative of the **RH**. The total derivative is,

$$\frac{\delta \mathbf{RH}}{\mathbf{RH}} = \frac{\delta m}{m} + \frac{\delta p}{p} - \left(\frac{17.67 \cdot 243.5}{(T - 29.65)^2}\right) (\delta T)^2 \qquad (1.32)$$

Suppose we substitute absolute values to the Eq. 1.32. In that case, the first and third terms will be comparable and the second term to be negligible compared to the other two terms (Thundathil et al., 2020). Hence the value of RH depends on temperature equally like the WVMR, which strengthens the requirement of a new operator with little influence

of temperature on the total moisture information. Therefore, within this study, a new operator was implemented in the WRFDA, which focuses on the analysis increment of the WVMR field. Since the WVDIAL measures absolute humidity, the operator was designed to convert absolute humidity to WVMR, which is not critically sensitive to temperature compared to that of the RH operator.

The conversion of absolute humidity ρ_{WV} to WVMR is given by the following relation,

$$m = \frac{\rho_{WV}}{\rho - \rho_{WV}} \approx \frac{\rho_{WV}}{\frac{p}{R_L T}\left[1 + 0.607m\right] - \rho_{WV}} \approx \frac{\rho_{WV}}{\frac{p}{R_L T} - \rho_{WV}} \approx \frac{\rho_{WV} R_L T}{p} \tag{1.33}$$

The approximated version of the relationship depends on the temperature, pressure and absolute humidity. Similar to the sensitivity analysis of the RH, we do this again for the equation as the total derivative with respect to the variables temperature, pressure and absolute humidity. The total derivative is as follows:

$$\frac{\delta m}{m} = \frac{R_L T}{p}\delta\rho_{WV} + \frac{\rho_{WV} R_L}{p}\delta T - \frac{\rho_{WV} R_L T}{p^2}\delta p \approx \frac{R_L T}{p}\delta\rho_{WV} \tag{1.34}$$

in which the term $\frac{\rho_{WV} R_L}{p}\delta T$ is much less than $\frac{R_L T}{p}\delta\rho_{WV}$ and the term $\frac{\rho_{WV} R_L T}{p^2}\delta p$ is still very much less than the other two terms (Thundathil et al., 2020).

Since the WRFDA system assimilates observations from various meteorological instruments, the new operator contains essential data processing steps. Initially, raw observations are converted from the instruments to the LITTLE R format. LITTLE R is an ASCII-based file format that is an intermediate format for the WRFDA to assimilate any number of observation types universally. The observation preprocessor (OBSPROC) of the WRFDA package reads only observations in the LITTLE R format. The OBSPROC removes the observations that do not fit the specified temporal and spatial domain. Also, it applies several other tasks like reordering and merging or deleting duplicate data. As a starting point in developing an exclusive forward operator for the atmospheric products derived by lidars, an already existing atmospheric infrared sounding retrieval (AIRSRET) or the FM-133 observation operator was used. We tested the AIRSRET operator because the operator has temperature and WVMR fields in the model, which are the end products of lidars. The AIRSRET operator takes in RH and temperature data and then converts to WVMR given by the relation:

$$m = \frac{\mathbf{RH}E(T)}{1.607p} \tag{1.35}$$

which is a rearrangement of the Eq. 1.30. In the new operator, the WRFDA code was modified so that the WVMR data field replaces the RH field by using Eq. 1.33. The new operator is called the thermodynamic lidar (TDLIDAR) operator since it can simultaneously ingest both temperature and moisture.

In near-real-time, the noise error is determined by the autocovariance function of the high-resolution absolute humidity time series at each height (Lenschow et al. 2000 and Wulfmeyer et al. 2016). It is routinely implemented in the IPM data processing algorithms. The representative error is calculated complying with Taylor's hypothesis of an air parcel. Suppose we apply the Taylor hypothesis for the water vapour time-height cross-section measured in a grid box of the model system. In that case, the water vapour variability will represent this box for a time period $\Delta \mathbf{t} \approx \frac{\Delta \mathbf{x}}{\mathbf{v}}$ where Δx is the horizontal model grid increment and $\mathbf{v}$ is the horizontal wind speed. Using the autocovariance function analysis, it is possible to separate atmospheric variance and noise variance delivering information about the accuracy of the measurement and the atmospheric variability. Suppose the autocovariance is taken at lag 0, which is equivalent to calculating the total variance of the time series. In that case, we can take this as an estimate for the total error consisting of the noise error variance and the representativeness error so that we can write,

$$\delta \rho_{WV}(z) \approx \sqrt{var(\rho_{WV}(t,z))} \approx \sqrt{(\delta \rho_{WV,noise}(z))^2 + (\delta \rho_{WV,represent}(z))^2} \; [kgm^{-3}] \quad (1.36)$$

Around the time step of assimilation, a window of $\pm$ 10 minutes is averaged to create the error profiles. The total observation error of moisture measurements from lidar is much lower than that of conventional datasets. Hence a new error factor *max_error_q_DIAL* was introduced in the WRFDA registry. This error factor enables the user to adjust the error window through which the observations are ingested into the model. The filtering process works with the equation:

$$m_{err} = \delta m \times max\_error\_q\_DIAL \quad (1.37)$$

where δm is the observation error. Suppose if the difference between the observation and the first guess falls within the quantity m_{err}, the observations are ingested. Thereby the model filters off low-quality WVDIAL observations that have a considerable bias difference between the first guess of the model. The error factor can be changed within the namelist of WRFDA in the wrfvar 5 optional arguments. The modified version of the WRFDA can be added to the public WRF repository for public use by the scientific community for thermodynamic profile observations from lidar systems or other profiling instruments.

Chapter 2

Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model

Paper by:

2.1   Overview

The first and foremost step in incorporating the high-resolution thermodynamic lidar profile observations into the NWP model is to ensure no loss of information. The task of blending the observations into the model is accomplished via a tool that can map the prognostic variables of the model to the lidar data products. The mapping function is called a forward operator. Through this forward operator, the lidar profile observations are assimilated into the model, and after that, the profile impact can be studied through various DA techniques.

In the WRF model, water vapour variables had been assimilated through the radiosonde operator, which assimilates water vapour contents in the form of RH. All the vertical profile observations from different instruments like the radiosondes, ground-based microwave radiometers, and other humidity profiling instruments have been using the radiosonde operator to assimilate humidity. Assimilation of water vapour contents in terms of RH is not the optimal approach because it is a moisture variable strongly sensitive to temperature.

WVMR is a variable that is not sensitive to temperature and only depends on the amount or quantity of water present in the air parcel.

Direct assimilation of WVMR called for the development of a new forward operator that had to be incorporated into the WRFDA system. This new forward operator enables using WVMR data directly without converting to RH, for which temperature data is also needed. The direct assimilation, in turn, avoids undesirable cross sensitivities to temperature enabling maximum information of the observation with respect to the water vapour contents in the area of interest to be propagated into the model. This chapter introduces the theoretical development of the new forward operator called the thermodynamic lidar (TDLIDAR) used for the study and, through the operator, the impact of thermodynamic profile observations with the **3DVar** DA system.

The performance of the TDLIDAR operator was compared with the RH operator to quantify the improvement. Hence, the impact of assimilating the temperature and WVMR profiles from lidar systems was analysed, comparing the impact through the TDLIDAR operator and the RH operator. The impact experiments were conducted on a convection-permitting horizontal resolution of 2.5 km over central Europe. A sequential DA was adopted for the study, which comprised an hourly rapid update cycle mode for ten hours. The case study selected for the experiment was an intensive observation period (IOP) from the High Definition of Clouds and Precipitation for advancing Climate Prediction (HD(CP)^2) project Observation Prototype Experiment (HOPE). The IOP took place from 0900 UTC to 1800 UTC on the 24 April 2013, in Western Germany on a clear sky day with hardly any optically thick clouds. Apart from the lidar measurements, four radiosonde launches were performed at 9, 11, 13 and 15 UTC.

For the impact study, six experiments were performed in total.

1. **NO_DA**   No data assimilated.

2. **CONV_DA (Control run)**   Only conventional data[1] assimilated.

3. **T_DA**   Control run + temperature data.

4. **Q_DA**   Control run + WVMR data.

5. **RH_DA**   Control run + RH data.

6. **QT_DA**   Control run + WVMR and temperature data.

Quantification of the analyses using RMSEs with respect to the assimilated lidar observations and independent radiosonde observations were performed to see the overall impact. From the results of the five DA runs, assimilation of both temperature and WVMR

---

[1]Observations from the European Centre for Medium-Range Weather Forecasts (ECMWF)

lidar observations improved the thermodynamic profiles in the analyses. In the thermo-dynamic lidar profile impact or the QT_DA analysis, the WVMR RMSE compared to WV-DIAL observations reduced by 40% compared to the control run, whereas the RH_DA did not show any improvement. The TDLIDAR operator showed a better impact than the RH operator.

However, in the analysis from the simultaneous assimilation of temperature and WVMR, the temperature did not improve compared to the improvement observed in the sole tem-perature assimilated case. A significant dependency of WVMR and temperature variables was observed. The reason behind this issue was identified as the influence of the static climatological background error covariance or the so-called "B matrix" in the **3DVar** DA system. Background meteorological fields provide information in non-observed areas and provide realistic reference states in some nonlinear observation operators in a DA system. As observations, the background state is prone to error, which is taken into account using the B matrix. The B matrix is derived from computing statistics on differences between forecasts over a particular domain of interest for a month or more to mimic the forecast errors (Bannister, 2008). In real-time operational forecasting with data assimilated from ac-tive remote-sensing instruments like lidar systems, which provide very low observational bias, correlation statistics derived from a set of forecast error differences might not pro-vide the best analysis. Chapter 3 explains how this issue can be alleviated by introducing a flow-dependent background error covariance matrix with the help of ensemble-based DA systems. Until then, this chapter gives an insight into the TDLIDAR operator and the thermodynamic lidar profile impact with the help of a **3DVar** DA system.

2.2   Publication

The publication titled "Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model" was is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model

**Rohith THUNDATHIL, Thomas SCHWITALLA, Andreas BEHRENDT,
Shravan Kumar MUPPA[1], Stephan ADAM, and Volker WULFMEYER**

Institute of Physics and Meteorology, University of Hohenheim, Germany

Abstract

The impact of assimilating thermodynamic profiles measured with lidars into the Weather Research and Forecasting (WRF)-Noah-Multiparameterization model system on a 2.5-km convection-permitting scale was investigated. We implemented a new forward operator for direct assimilation of the water vapor mixing ratio (WVMR). Data from two lidar systems of the University of Hohenheim were used: the water vapor differential absorption lidar (UHOH WVDIAL) and the temperature rotational Raman lidar (UHOH TRL). Six experiments were conducted with 1-hour assimilation cycles over a 10-hour period by applying a 3DVAR rapid update cycle (RUC): 1) no data assimilation 2) assimilation of conventional observations (control run), 3) lidar–temperature added, 4) lidar–moisture added with relative humidity (RH) operator, 5) same as 4) but with the WVMR operator, 6) both lidar–temperature and moisture profiles assimilated (impact run). The root-mean-square-error (RMSE) of the temperature with respect to the lidar observations was reduced from 1.1 K in the control run to 0.4 K in the lidar–temperature assimilation run. The RMSE of the WVMR with respect to the lidar observations was reduced from 0.87 g kg^{-1} in the control run to 0.53 g kg^{-1} in the lidar–moisture assimilation run with the WVMR operator, while no improvement was found with the RH operator; it was reduced further to 0.51 g kg^{-1} in the impact run. However, the RMSE of the temperature in the impact run did not show further improvement. Compared to independent radiosonde measurements, the temperature assimilation showed a slight improvement of 0.71 K in the RMSE to 0.63 K, while there was no conclusive improvement in the moisture impact. The correlation between the temperature and WVMR variables in the static-background error-covariance matrix affected the improvement in the analysis of both fields simultaneously. In the future, we expect better results with a flow-dependent error covariance matrix. In any case, the initial attempt to develop an exclusive thermodynamic lidar operator gave promising results for assimilating humidity observations directly into the WRF data assimilation system.

1.   Introduction

The vertical and horizontal distribution of water vapor and temperature in the atmosphere is crucial for the evolution of weather on all spatial and temporal scales. Detailed observations are important for improving the initial fields for numerical weather prediction (NWP) from nowcasting to the very short-range, the short-range, and the medium range. However, our

Corresponding author: Rohith Thundathil, Institute of Physics and Meteorology, University of Hohenheim, Garbenstrasse 30, Stuttgart 70599, Germany
E-mail: rohith.thundathil@uni-hohenheim.de
[1] Present affiliation: Micrometeorology Group, University of Bayreuth, Germany
J-stage Advance Published Date: 14 July 2020

present representation of land–atmosphere (L–A) interaction and convection initiation (CI) suffers in mesoscale models largely from huge observational gaps, consequently also limiting the predictive skill of NWP. Therefore, it is essential to enhance these observations and to investigate the impact of new remote sensing systems which are capable of measuring water vapor and temperature profiles into NWP models by means of data assimilation (DA).

Small-scale variations in moisture due to collision of boundaries (Kingsmill 1995), horizontal convective rolls and mesocyclones (Weckwerth et al. 1996; Murphey et al. 2006), and intersections between boundaries and horizontal convective rolls (Dailey and Fovell 1999) influences the location and timing of CI. The amount of moisture and variations in the vertical gradients of moisture and temperature at lower levels of the atmosphere can change the strength of CI significantly (Lee et al. 1991; Crook 1996). Several field campaigns have been conducted to understand the relationship between the three-dimensional thermodynamic fields and CI as well as the impact of assimilation of thermodynamic profiles. These have included the Mesoscale Alpine Program 1990 (Richard et al. 2007); the International H_2O Project (IHOP) 2002 (Weckwerth and Parsons 2006); the Convection Storm Initiation Project conducted in the summer period of 2004 and 2005 (Browning et al. 2007) and which provided sufficient data for impact studies using the Met Office unified model (Dixon et al. 2009); the Lindenberg Campaign for Assessment of Humidity and Cloud Profiling Systems and its Impact on High-Resolution Modeling (LAUNCH, Engelbart and Haas (2006) in the late summer of 2005; the Convective and Orographically-induced Precipitation Study (COPS) 2007 (Wulfmeyer et al. 2011); and the Plains Elevated Convection At Night (Geerts et al. 2017) campaign in summer 2015.

Recently, studies of land–atmosphere (L–A) feedback have also become the focus of improving the quality of weather forecast models as it was realized that a realistic representation of L–A interaction in mesoscale models is crucial for an accurate prediction of the pre-convective, dynamic, and thermodynamic environments. The first extensive study was the Land Atmosphere Feedback Experiment (Wulfmeyer et al. 2018) conducted in August 2017, which also provided a large data set for the assimilation of thermodynamic profiles measured with lidar in mesoscale models. The importance and sensitivity of L–A feedback for the simulation and prediction of the formation and organization of clouds and precipitation was exemplified in

Santanello et al. (2018).

At the major forecast centers, there are mainly three DA approaches which are currently used: (1) variational techniques like 3DVAR and 4DVAR (Courtier 1998; Barker et al. 2004; Huang et al. 2009); (2) ensemble-based approaches which include flavors of the ensemble Kalman filter (Evensen 2003), and (3) hybrid combinations of these (Ingleby et al. 2013). In 3DVAR, the data is assimilated at specific analysis time-steps, whereas in 4DVAR there is an adjoint model so that the cost function is minimized over a time period and not at a particular time-step. The drawback of the 3DVAR is the static nature of the background error covariance (B) matrix in the cost function. This prevents the model from incorporating the present dynamics of the atmosphere. Although 4DVAR implicitly incorporates a time-evolving background error covariance model (Lorenc 2003), the same static matrix, B, is propagated implicitly to a later time-step. However, the 4DVAR is superior to the 3DVAR scheme due to the evolution of the background error covariance matrix and the reduction of the model imbalance at the analysis time. Meteo-France uses the incremental 3DVAR in the Aire Limitée Adaptation dynamique Développement Inter National (ALADIN) model (Brousseau et al. 2011; Berre 2000); the German Weather Service (DWD) and MeteoSwiss uses the Local Ensemble Transform Kalman Filter (LETKF) DA in the Consortium for Small-scale Modelling (COSMO) model (Schraff et al. 2016); the UK Met Office has implemented incremental 3DVAR and 4DVAR (Ingleby et al. 2013); NOAA's National Centers for Environmental Prediction uses incremental hybrid 3DEnVar and non-variational cloud analysis (Wu et al. 2017; Hu et al. 2006, 2017; Benjamin et al. 2004, 2016); and the Japan Meteorological Agency (JMA) applies incremental 4DVAR and 3DVAR (Honda et al. 2006; Aranami et al. 2015). A recent discussion of the DA methods used in various forecast centers is given by Gustafsson et al. (2018). All of these DA techniques are capable of assimilating profiles of the thermodynamics and dynamics of the atmosphere.

Radiosonde and aircraft measurements are the only conventional data observation sources currently providing water vapor and temperature data within the planetary boundary layer (PBL) and lower troposphere. Radiosondes provide a vertical thermodynamic profile of the atmosphere from the surface layer through the lower troposphere whereas weather stations provide only surface measurements with limited impact on the vertical thermodynamic structure. Radiosondes

provide instantaneous data only at the time of ascent, giving more or less a snapshot of the atmosphere along their vertical track. Therefore, the soundings suffer from significant sampling errors, especially in the boundary layer with its highly turbulent fluctuations (Weckwerth et al. 1999). The coverage of the radiosonde network is quite coarse, and the number of radiosonde stations is decreasing rather than increasing in most countries due to their high cost of operation.

Another option is the application of passive and active remote sensing data. Wulfmeyer et al. (2015) gave a comprehensive overview of the current observational capabilities of remote sensing techniques with respect to thermodynamic fields in the lower troposphere. It was demonstrated that using space-borne passive remote sensing systems for thermodynamic observations does not provide the necessary vertical resolution in the lower troposphere to recover its vertical structure. Ground-based passive remote sensing instruments like microwave radiometers produce reliable data but have a coarse resolution of around 300 m to 1000 m in the lower 2000 m above the ground (Blumberg et al. 2015; Cadeddu et al. 2002; Wulfmeyer et al. 2015). IR spectrometers have higher vertical resolutions due to having more spectroscopic lines which can be evaluated; however, their vertical resolution is still limited to 100 m to 800 m up to 2000 m above ground level (Turner and Löhnert 2014). Convection-permitting models have vertical resolutions in the range of 100 m or less within the boundary layer, where fine-scale processes are crucial, in order to recover the thermodynamic structure of the atmosphere. Therefore the observation systems must fulfill the data requirements of convective-scale DA models to ensure higher representativeness (Wulfmeyer et al. 2015). Therefore, microwave radiometers and IR spectrometers are not capable of resolving the vertical structure of the lower troposphere, including the top of the PBL, the inversion strength at the PBL top, or the elevated inversion layers and the moisture structure in the free lower troposphere. However, this capability is expected to be crucial to achieving an improved prediction of L–A feedback and CI. Typical temporal resolutions of passive remote sensing instruments are 5–10 minutes, but further processing time is required either for the inversion of the spectra to vertical water vapor and temperature information or for the assimilation of the spectra through a forward operator in a DA system.

Active remote sensing techniques offer high temporal and spatial resolution data simultaneously to accurately capture the atmospheric fields without much loss of temperature and moisture gradient information. Two main techniques for humidity profiling are available: water vapor differential absorption lidar (WVDIAL) and water vapor Raman lidar (WVRL). Both systems achieve a high vertical and temporal resolution during both day- and night-time (Lange et al. 2018; Späth et al. 2018). Whereas WVDIAL does not require calibration (Ismail and Browell 1989; Bösenberg 1998), it has been demonstrated that, for WVRL, the calibration has long-term stability, and a high accuracy can be maintained for the measurements. Ground-based WVDIAL has been implemented for tropospheric measurements at various centers. Depending on the efficiency of the receiver and the average power of the laser transmitter, the combination of temporal and spatial resolution ranges from 1 s, 15 m (Metzendorf 2019) to 5 min, 300 m (Spuler et al. 2015). The NCAR and Montana State University have developed a compact, field-deployable micro-pulse DIAL (Spuler et al. 2015; Weckwerth et al. 2016) with a range resolution of 300 m and a temporal resolution of 1–5 min. The vertically pointing WVDIAL of the Institute of Physics and Meteorology (IPM, Wagner et al. 2011, 2013; Metzendorf 2019) has a range resolution of 15–300 m and temporal resolution of 1–10 s. The first WVDIAL with a 3-D scanner was also developed at the IPM of the University of Hohenheim (UHOH, Behrendt et al. 2009; Späth et al. 2014). Typical accuracies of the absolute humidity for the IPM's WVDIAL are in the range of 5–10 % within the PBL during the daytime. WVRLs have been making continuous measurements at various centers, such as the operational WVRL (Goldsmith et al. 1998; Turner and Goldsmith 1999) at the Atmospheric Radiation Measurements Southern Great Plains site in the U.S; the Raman Lidar for Meteorological Observations (RALMO, Dinoev et al. 2013; Brocard et al. 2013) in Payerne, Switzerland used by MeteoSwiss; the Raman Lidar for Atmospheric Moisture Sensing (RAMSES, Reichardt et al. 2012) in Lindenberg, Germany, used by the German Meteorological Service (DWD); and the WVRL at the Cabauw Experimental Site for Atmospheric Research (CESAR, Apituley et al. 2009) in the Netherlands. Typical resolutions of WVDIALs are around 150 m for the spatial resolution and 10 s for the temporal resolution, with an accuracy of < 5 %.

For temperature profiling in the lower troposphere, the temperature rotational Raman lidar (TRL) technique demonstrated the best performance (Behrendt et al. 2004; Di Girolamo et al. 2004; Arshinov et al. 2005; Radlach et al. 2008). It is now possible to

measure temperature profiles from close to the surface to the lower troposphere with a temporal resolution of a few minutes and a vertical resolution of approximately 100 m. This performance permits the detection of inversion layers and the characterization of the temperature gradient with a high degree of accuracy (Hammann et al. 2015). Continuous time–height cross-sections of the atmospheric thermodynamic profile are a unique feature of these lidar systems which enables promising research and applications in the direction of mesoscale DA. Therefore, WVDIAL, WVRL, and TRL are suitable and ready for application in DA impact studies.

The subject of this work is the analysis of the impact of two relatively new lidar systems used for water vapor and temperature profiling in mesoscale DA. The two active remote sensing system are the high-power, high-efficiency, 3D scanning WVDIAL which has an extraordinary resolution, accuracy, and range (Wagner et al. 2013; Späth et al. 2014, 2016) and the TRL for daytime and night-time temperature profiling (Radlach et al. 2008; Hammann et al. 2015; Behrendt et al. 2015; Lange et al. 2018), both developed and operated at the IPM in Stuttgart, Germany.

The experimental setup was based on the Weather Research and Forecasting-Noah-Multiparameterization (WRF-Noah-MP) model system and the WRF DA (WRFDA) system using a 3DVAR rapid update cycle (RUC). This RUC was developed and optimized for Europe (Schwitalla and Wulfmeyer 2014) and is operated on the convection-permitting scale. Previously, the water vapor mixing ratio (WVMR) or other water vapor variables were assimilated by applying the radiosonde relative humidity (RH) operator. It is obvious that this is not the optimal approach because the RH is strongly sensitive to temperature. Therefore, we developed a new forward operator for the assimilation of absolute humidity, mixing ratio or specific humidity independent of any cross-sensitivity to temperature. This forward operator was based on an already-existing atmospheric infrared sounding retrieval (AIRSRET) observation operator in the WRFDA system. We expected that this new operator would provide a strong and direct impact. The first key objective of this work was to quantify this impact.

So far, there have been only a few impact studies using thermodynamic lidar data. During IHOP 2002, Wulfmeyer et al. (2006) assimilated airborne water vapor DIAL data from the NASA LASE system into the 5th generation Pennsylvania State University-NCAR Mesoscale Model (MM5), which was based on a 4DVAR DA system. The results from the assimila-

tion resulted in a considerably improved prediction of CI due to strong and positive analysis increments, not only with respect to water vapor but also to dynamics. During LAUNCH, Grzeschik et al. (2008) assimilated water vapor data from a triangle of three WVRLs, again into the MM5. The initial water vapor field was corrected by about 1 g kg^{-1} and the WVRL impact on the water vapor field continued for up to 12 h in the forecast model. Airborne water vapor data from the Water Vapour Lidar Experiment in Space demonstrator was assimilated into the ECMWF 4DVAR global model by Harnisch et al. (2011). The analysis error was reduced after the assimilation of WVDIAL observations. COPS (Wulfmeyer et al. 2011) had two airborne lidars which measured lower tropospheric water vapor fields: these were assimilated into the 3DVAR assimilation system of the Application of Research to Operations at MEsoscale (AROME) numerical weather prediction model (Bielli et al. 2012). Temperature data from TRL were assimilated into the WRF model by Adam et al. (2016), which produced positive results. Also recently, as described in Yoshida et al. (2020), water vapor profiles from Raman lidar were assimilated using the LETKF system to investigate the effects on precipitation forecasts. All of these results confirm the positive impact of thermodynamic lidar DA on NWP models. The first study where WV and T profiles from active remote sensing measurements were assimilated simultaneously into a forecast system will be presented here.

For this purpose, we investigated the impact of assimilating high-resolution temperature profiles from the UHOH TRL and water vapor profiles from the UHOH WVDIAL into our version of the WRFDA model using a 3DVAR RUC.

This work describes how well the new forward operator can assimilate WVMR and temperature data from the lidar instruments and focuses on the following questions:
 - Does the new operator work and have a reasonable impact on the analysis of the WV field?
 - What is the impact of WV DA alone, the impact of T DA alone, and the combined impact?
 - How large is the inter-dependency of the WVMR and temperature variables in the DA system?

The manuscript is arranged as follows. Section 2 gives a brief overview about the HOPE campaign. Section 3 describes the WRFDA system, the configuration of the RUC applied in our study, as well as the new water vapor operator. The lidar observations are shown at the end of Section 3 together with a brief description of their principles. Section 4 describes the

results of the impact study with respect to temperature and moisture. The manuscript finishes with a summary of our results and an outlook.

2. Observations

2.1 The HOPE measurement campaign

The High Definition Clouds and Precipitation for advancing Climate Prediction HD(CP)[2] project aimed at improving the representation of clouds and precipitation in atmospheric models. By resolving clouds and precipitation processes, the uncertainty in climate change predictions can be significantly reduced (Stevens and Bony (2013); see http://www.hdcp2.eu for more information). The project was initiated by the German Federal Ministry of Education and Research in coordination with the German Meteorological Service (DWD) in October 2012. In order to evaluate the performance of models, the HD(CP)[2] Observation Prototype Experiment (HOPE) campaign (Macke et al. 2017) was conducted to provide high-resolution observations. The HOPE campaign focused on multi-sensor synergy within a micro- to mesoscale domain. The campaign took place in north-western Germany around the Jülich Research Centre during April and May 2013. The HOPE field campaign was conducted mainly at three supersites, which covered an approximately 10-km radius around the Jülich Research Centre. The supersites were designed in such a way to derive data concerning moisture, temperature, and wind at a resolution of 100 m for a volume of around 10 km × 10 km × 10 km. The three supersites used, where the main remote sensing facilities were deployed, were Jülich (JUE), Krauthausen (KRA), and Hambach (HAM). The IPM lidar systems were deployed at the Hambach site, where radiosondes were also launched during intensive observation periods (IOPs). The radiosonde type used during the IOPs was the DFM-09 model from GRAW (https://www.graw.de/products/radiosondes/dfm-09/). The WVDIAL (Späth et al. 2016) and the TRL (Hammann et al. 2015) from UHOH were positioned at 50°53′50.55″N, 6°27′50.27″E and 110 m above sea level (Fig. 1). The IPM lidar systems were designed to observe the three-dimensional thermodynamic temperature and moisture fields along with their turbulent fluctuations (Muppa et al. 2016; Behrendt et al. 2015; Wulfmeyer et al. 2016).

2.2 UHOH WVDIAL

In the DIAL technique, two laser signals are used, namely P_{on} and P_{off}, the online and offline signals, respectively. The wavelength of the P_{on} signal is
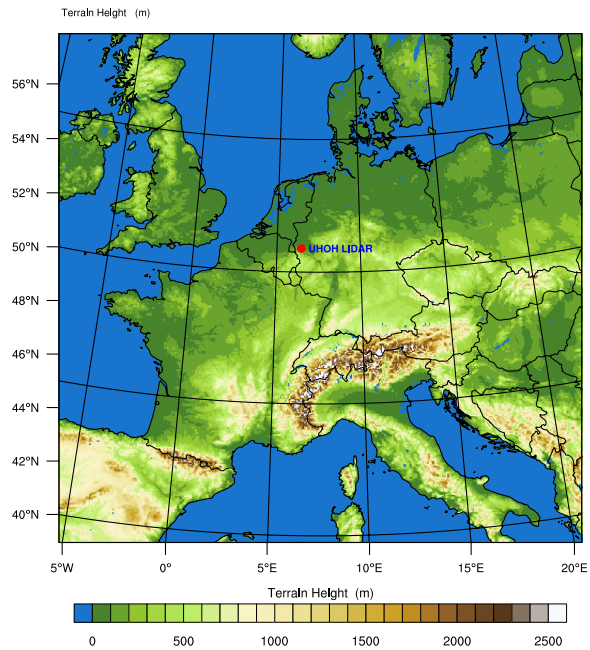
Fig. 1. The WRF model domain at a horizontal resolution of 2.5 km with orography and the location of the TRL and WVDIAL of UHOH during the HOPE campaign.

tuned in such a way that there is a strong absorption of water vapor in the atmospheric signal resulting in a reduction in the backscatter, whereas the P_{off} signal wavelength is tuned for weak absorption. The number density of the water vapor molecules is derived from the differential absorption of the online and offline signals (Schotland 1966):

$$N_{WV}(r) = \frac{1}{2(\sigma_{on}(r) - \sigma_{off}(r))} \frac{d}{dr} \ln\left(\frac{P_{off}(r) - P_{B,off}}{P_{on}(r) - P_{B,on}}\right),$$
(1)

where N_{WV} is the water vapor number density, σ denotes the absorption cross section, P_B is the background signal, and the argument r is the distance measured from the lidar system to the scattering volume along the line of sight of the laser beam. Further details of the UHOH WVDIAL can be found in Wagner et al. (2013) and Späth et al. (2016).

The data acquisition system had a sampling rate of 10 MHz, which allowed the atmospheric backscatter signals to be recorded with a fine vertical resolution of 15 m. The data were recorded for each laser shot and averaged over a period of 1 s to 10 s. The raw data
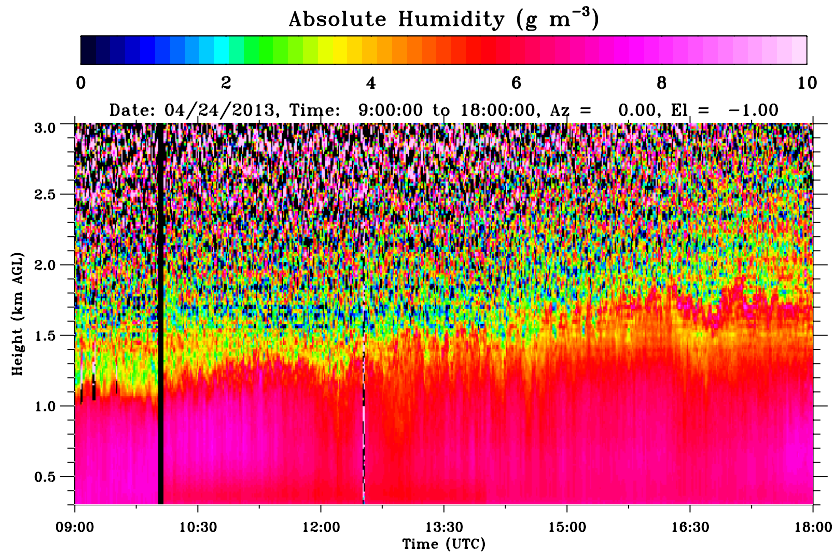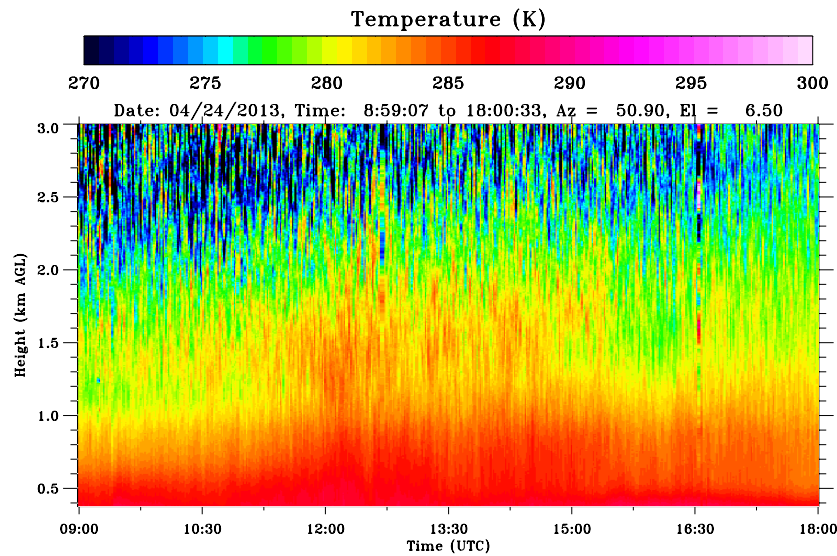
Fig. 2.   Absolute humidity time series from the WVDIAL with a vertical resolution of 15 m and a temporal resolution of 60 s between 09 UTC and 18 UTC on 24 April 2013.

used for the present study had a temporal resolution of 10 s. In Eq. (1), the derivative with respect to the range is derived by the Savitsky–Golay (SaGo) algorithm (Savitzky and Golay 1964). The window length in the SaGo algorithm was set to 135 m up to a height of 1500 m above ground level based on a consideration of the average height of the PBL. Between 1500 m and 3000 m, a window length of 285 m was applied since the signal-to-noise-ratio (SNR) of the signals decreased due to a reduction in the signal strength and differential optical thickness.

Time windows of ±10 minutes around the assimilation time step were chosen. A total of 120 lidar profiles from the high-resolution absolute humidity data (Fig. 2) which fall into these 20-minute windows were averaged for input at each assimilation time-step. The absolute humidity data and the corresponding error derived from the number density were then converted to WVMR data with associated errors. We ensured that the input data for the assimilation had a resolution roughly similar to that of the model. Hence the WVMR data, which was in 15-m steps, was fed into the assimilation data in 30-m steps. The WVDIAL error for the resolutions that were used ranged from 0.01 g kg^{-1} at a height of 400 m to a maximum of 1 g kg^{-1} at heights above 2 km (Späth et al. 2014).

2.3   UHOH TRL

The UHOH TRL measures atmospheric tempera-

ture profiles through the rotational Raman technique (Cooney 1972; Behrendt and Reichardt 2000; Behrendt et al. 2004). This method relies on the temperature-dependent inelastic scattering of UV laser pulses when collided with, Nitrogen and Oxygen molecules, the major gaseous constituents of the atmosphere.

The rotational Raman spectrum of air consists of two parts, the Stokes and the anti-Stokes branches. The Stokes branch is found at wavelengths greater than that of the incident radiation while the anti-Stokes branch is found at shorter wavelengths. The UHOH TRL extracts only signals of the latter. The temperature is determined using the ratio Q of the two background-corrected Raman signals $RR2$ and $RR1$. P_{RR2} and P_{RR1} are the signals for low and high quantum-number transition settings of the filter, respectively, so that

$$Q(r) = \frac{P_{RR2}(r)}{P_{RR1}(r)}. \qquad (2)$$

The temperature profile of the atmosphere $T(r)$ is obtained from

$$T(r) = \frac{b}{a - \ln Q(r)}, \qquad (3)$$

where a and b are calibration constants. The statistical error in the temperature measurements are derived

Fig. 3. Temperature time series from the TRL with a vertical resolution of 3.75 m and a temporal resolution of 60 s between 09 UTC and 18 UTC on 24 April 2013.

from Poisson statistics applied to the signal intensities of the photon-counting data. For a signal count number s, σ denoting the standard deviation, 1σ statistical error is given by the square root of s. The error characteristics are detailed in Behrendt et al. (2015), Behrendt and Reichardt, (2000), and Wulf-meyer et al. (2015, 2016). The temperature profiles are also averaged over a time window of 20 minutes at each assimilation time-step before assimilation into the model. The vertical profile from the TRL was smoothed with a running-average window of 108.75 m and then thinned to one value of 3.75 m. The error range of the profiles was from 0.1 K at 500 m to 1.1 K at 3000 m (Hammann et al. 2015). Figure 3 shows the time series data prior to further temporal averaging over 20 minutes. The averaged data were then used for assimilation.

2.4 Conventional observations

The DA system was augmented by a dense network of surface reports, SYNOP and METAR, over Europe. A set of radiosonde (RS) measurements, TEMP, provided a snapshot of the thermodynamic structure of the atmosphere from the point of launch. A set of wind profilers (PROFL) provided wind measurements along with the wind data provided along with the radiosonde products. Aircraft measurements (AMDAR) were also assimilated as part of the conventional observations. All of these observations were obtained from the

Global Telecommunication System data archive of the WMO, which are stored at the ECMWF. Satellite Atmospheric Motion Vectors (AMVs) above 700 hPa from the Meteosat Second Generation satellite were also included in the assimilation dataset. The AMVs data below 700 hPa were discarded since the data retrieval algorithm is not reliable (Horváth et al. 2017). Apart from these observations, global navigation satellite systems-zenith total delay (GNSS-ZTD) data were used in the DA system for improving the accuracy for humidity distributions over the domain. These data were obtained from the E-GVAP network (http://egvap.dmi.dk/). Table 1 shows a summary of the already large number of observations assimilated into the DA system within the conventional DA run, which meant that it was quite a challenge for the lidar data to achieve any additional impact. Figure 4 depicts the conventional observations assimilated into the model for 09 UTC: this was roughly the same for all the subsequent assimilation cycles.

3. Model setup

3.1 WRF model and configuration

The WRF model (Skamarock et al. 2008), version 3.8.1, was used for the impact study presented here. The WRF model has been applied for research at various characteristic spatial scales like the synoptic-scale, mesoscale, and large eddy simulation (LES) scale (Talbot et al. 2012; Wei et al. 2017; Muppa et al.

Table 1. Details of assimilated observations with their corresponding observation operators.

Type	Conventional Observations				
Observation Operator	AMDAR	AMV	GNSS-ZTD	METAR	PROFL*
Assimilated observations	1374–1875	2045–3095	1050–1076	261–310	52–57
Average number	1624	2570	1063	286	54

Type	Conventional Observations				TRL	WVDIAL
Observation Operator	SYNOP	SHIP	BUOY	TEMP*	TEMP*	AIRSRET*
Assimilated observations	968–1128	77–104	7–9	0–26	1	1
Average number	1048	90	8	13	1	1

*Profile observations
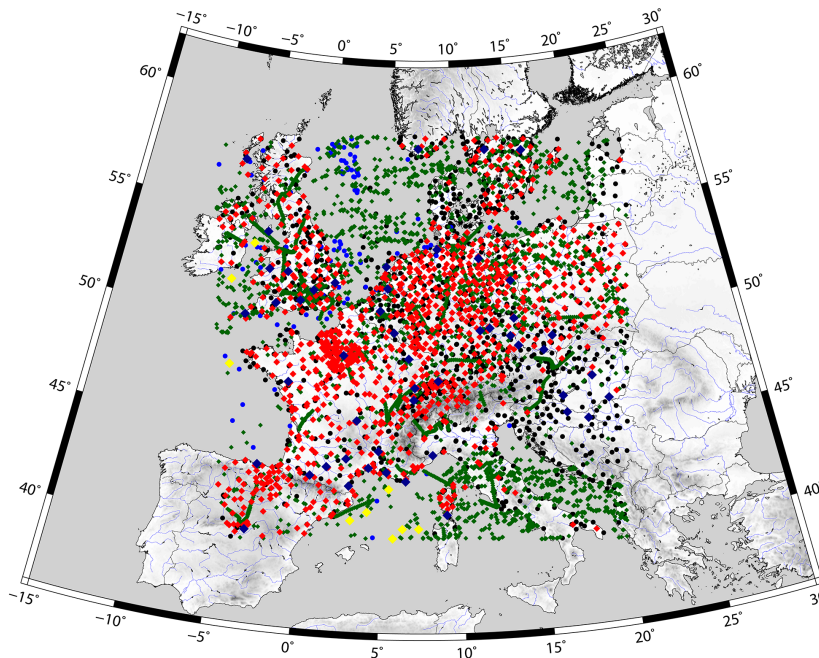
Fig. 4. Observation types and their locations for the assimilation time-step 09 UTC 24 April 2013. Black: surface stations (SYNOP + METAR), blue: ship observations (SHIP), green: aircraft observations and atmospheric motion vectors from satellite (AMDAR + SATOB), red: GPS zenith total delay, yellow: radiosondes (TEMP), and brown: wind profiler (PROFL).

2018; Schwitalla et al. 2017). Furthermore, the WRF model is extensively used for operational forecasting in various weather forecasting centers across the world (Powers et al. 2017). The WRF has two dynamical solvers – the Advanced Research WRF (ARW) core (Skamarock et al. 2008) and the Nonhydrostatic Mesoscale Model core (Janjic 2003). The former was applied in our study.

Compressible and nonhydrostatic Euler equations are integrated in the ARW dynamic solver. The prognostic variables in the model are the velocity components u and v in Cartesian coordinates and w in the vertical coordinate, the perturbation potential temperature θ, the perturbation geopotential φ, and the perturbation surface pressure p_s. The WVMR q_v is also a prognostic variable in the ARW solver.

The model was configured with a spatial resolution of 2.5 km and 856×832 grid cells (Fig. 1). The vertical resolution of the model was set to 100 levels up to 50 hPa with 27 levels within the PBL. Compared to the study of Adam et al. (2016), the number of vertical levels in the model was increased from 57 to 100 in

Table 2. WRF physics configuration used for the experiments.

Physics	Options implemented
Long wave radiation	RRTMG (Iacono et al. 2008)
Short wave radiation	RRTMG (Iacono et al. 2008)
Cloud microphysics	Thompson scheme (Thompson et al. 2008)
Planetary boundary layer	MYNN (Nakanishi and Niino 2006)
Land surface scheme	NOAH-MP LSM (Niu et al. 2011)

order to even better resolve gradients. The model time step for the simulation was set to 15 s. All simulations were initialized using European Centre for Medium Range Weather Forecasts (ECMWF) analysis with a spatial resolution of 0.125° (approximately 13.5 km). Also the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA; Donlon et al. 2012) data provided by the Met Office were applied to accurately initialize the sea surface temperatures.

The WRF model physics configuration used for the simulations is summarized in Table 2. The physics configuration used for the study was based on previous research and DA efforts (Adam et al. 2016; Schwitalla and Wulfmeyer 2014; Bauer et al. 2015; Schwitalla et al. 2011). The WRF was coupled with the Noah–MP Land Surface Model (Niu et al. 2011; Yang et al. 2011) which includes a canopy layer, three layers of snow, and four layers of soil. The skin temperature of the canopy and snow or soil surface are predicted by an interactive energy balance method. Shortwave and longwave radiation are parameterized with the RRTMG scheme (Iacono et al. 2008). Microphysical properties are represented by the Thompson double-moment scheme (Thompson et al. 2008), which explicitly predicts mixing ratios of cloud water, rain, cloud ice, snow, and graupel. The Mellor–Yamada Nakanishi Niino (MYNN; Nakanishi and Niino 2006) Level-2.5 scheme (Nakanishi and Niino 2009) was used as the PBL scheme. A new formulation of the turbulent length scales and parameterization of the pressure covariance as well as parameterization of the stability functions of third-order turbulent fluxes were incorporated in this MYNN scheme.

Deep-convection parameterization was not used in the study since we were running the model at the convection-permitting scale (Weisman et al. 2008). For shallow cumulus parameterization, the Global/Regional Integrated Model System Scheme (Hong et al. 2013) was used.

3.2 Data assimilation system

The WRFDA system incorporates a number of DA techniques which can be broadly classified as being based on the deterministic approach or the probabilistic approach. Deterministic approaches include the variational DA systems like the 3DVAR and 4DVAR. In this study, we applied the 3DVAR DA system in a RUC mode with an hourly update cycle. The code of the RUC is completely automated from the pre-processing stage to post-processing of the analysis and is designed for variable assimilation time windows. The WRFDA 3DVAR system is based upon the principle of iteratively minimizing the cost function $J(x)$, whose independent variable or the control variable is the analysis state vector x. The equation of the cost function for 3DVAR reads

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b)$$
$$+ \frac{1}{2}(y - H(x))^T R^{-1}(y - H(x)). \qquad (4)$$

The cost function $J(x)$ consists of two terms, a background and an observation term. The vector fields x, x_b and y are the analysis state, the background or the first guess, and the observation state vectors, respectively. H is the forward operator, which maps the analysis state vector space to the observation vector space. For instance, a corresponding operator is required for the DA of WVMR profiles, but this did not exist at the time this project started.

Apart from the general column vectors, there are two square matrices which play a major role in the cost function minimization: the background error covariance matrix B and the observation error covariance matrix R. In the DA system, R is a diagonal matrix since we assume that there is no correlation among the observation errors between different instruments or height levels. B is a square, positive, semi-definite and symmetric matrix whose eigenvalues are positive. B consists of the variances of the background forecast errors as the diagonal elements, and the covariance between them as the symmetric upper and lower triangular elements. The variances and covariances of B strongly contribute to the response of an analysis after an observation has been assimilated. The ratio of these values to the RMS errors of R determine the impact on the analysis. Hence an appropriate determination of B is crucial in a variational DA system.

B can be calculated mainly by three methods, namely, the NMC method (Parrish and Derber 1992), the analysis ensemble method (Fisher 2003), and by

using innovation statistics (Hollingsworth and Lönnberg 1986). All these methods have their own merits (and drawbacks). The NMC method, in which climatological background error covariances are estimated, is the most widely used method for the generation of *B*. We used the NMC method in our study since it provides physically reasonable results in regional model domains and is computationally less expensive than the ensemble method. In the NMC method, forecast difference statistics are computed, from which the forecast error covariance is then derived. The forecast error covariance is specifically derived for the domain in order to incorporate the errors applicable to that domain. However, the NMC method has certain drawbacks: it overestimates the covariances in large-scale simulations and poorly observed regions (Berre 2000; Fischer 2013; Berre et al. 2006). The statistics were derived for a period of a month from forecast differences of 24 hours and 12 hours since we were performing a regional simulation. The month of April 2013 was selected to derive the statistics. We used the CV6 option for implementing multivariate background error statistics in the *B* matrix. In the CV6 option, the moisture analysis is multivariate, which means that moisture increments are derived from temperature and wind increments and vice-versa.

3.3 WVMR forward operator

To assimilate the WVMR directly, a new forward operator had to be developed and incorporated in the WRFDA. This new forward operator allowed WVMR data to be used directly without converting it to RH, for which temperature data is also needed. Until now, the WRFDA system has ingested humidity data in the form of RH through the conventional radiosonde operator. Previously, all the vertical profile data products from radiosondes, ground-based microwave radiometers, and other humidity profiling instruments have used the radiosonde operator for the assimilation of humidity in the form of RH (Bielli et al. 2012).

The advantage of expressing moisture in the form of the WVMR is that the variable is a tracer and remains insensitive if there are changes in the atmospheric temperature or pressure fields. Consequently, the maximum information content of the observation is used with respect to the WV budget in the area of interest and unnecessary cross-sensitivities are avoided.

When the RH operator in the WRFDA is used for assimilating mixing ratio measurements *m*, the following relationship is used:

$$RH = \frac{m}{1+m}\frac{R_w}{R_L}\frac{p}{E(T)\left(1+\frac{m}{1+m}\frac{R_w-R_L}{R_L}\right)}$$

$$\cong m\frac{R_w}{R_L}\frac{p}{E(T)\left(1+m\frac{R_w-R_L}{R_L}\right)}$$

$$= 1.607m\frac{p}{E(T)(1+0.607m)}$$

$$\cong 1.607m\frac{p}{E(T)}. \tag{5}$$

Here, *T* is the ambient temperature in units of *K* and *p* is the total atmospheric pressure exerted by moist and dry air in units of *Pa*. R_w and R_L are the specific gas constants of water vapor and dry air, respectively, in units of $J\ kg^{-1}\ K^{-1}$. This relationship confirms that it is not the best idea to assimilate the WVMR using an RH operator because the sensitivity to temperature in the equation for the water vapor saturation pressure $E(T)$ (Bolton 1980),

$$E(T) = 611.2\exp\left[17.67\left(\frac{r-273.15}{r-273.15+243.5}\right)\right], \tag{6}$$

is comparable with the sensitivity to *m* and thus not negligible. This can be proved by deriving the total derivative of *RH* with respect to the variables *m*, *p*, and *T*. Starting from the total derivative of Eq. 5 with reference to Eq. 6, we finally get the expression for δRH as

$$\delta RH = RH\left[\frac{\delta m}{m} + \frac{\delta p}{p} - \left(\frac{17.67\times243.5}{(T-29.65)^2}\right)\delta T\right]. \tag{7}$$

Considering the absolute values of the terms within the square brackets in Eq. 7, the third term $\left(\frac{17.67\times243.5}{(T-29.65)^2}\right)\delta T$ is comparable with the first term $\frac{\delta m}{m}$. The second term $\frac{\delta p}{p}$ is very small compared to the other two terms. Please refer to the appendix section for a quantified analysis. From Eq. 7, we infer that the value of *RH* is dependent on *T* and *m*.

Therefore, a new operator that focuses on increased analysis of the WVMR field was implemented in the WRFDA in this study. In the case of the measurement of WVMR, the conversion is trivial because this is the prognostic variable used in the WRFDA. It should be noted that, in contrast to the WVRL, the WVDIAL measures absolute humidity and not the WVMR as the primary product. However, the conversion of absolute

humidity to WVMR is not as critically sensitive to temperature as the conversion to *RH* is.

When the absolute humidity ρ_{wv} is measured, the conversion is very simple and reads

$$m = \frac{\rho_{wv}}{\rho - \rho_{wv}} \cong \frac{\rho_{wv}}{\frac{p}{R_L T}(1 + 0.607m) - \rho_{wv}}$$

$$\cong \frac{\rho_{wv}}{\frac{p}{R_L T} - \rho_{wv}} \cong \frac{\rho_{wv} R_L T}{p}. \tag{8}$$

For the conversion, simply the model temperature and pressure variables are used. The WVMR error becomes mainly dependent on the error in the absolute humidity and reads

$$\delta m \cong \frac{R_L T}{p} \delta \rho_w + \frac{\partial m}{\partial T} \delta T + \frac{\partial m}{\partial p} \delta p$$

$$= \frac{R_L T}{p} \delta \rho_{WV} + \frac{\rho_{WV} R_L}{p} \delta T - \frac{\rho_{WV} R_L T}{p^2} \delta p$$

$$\cong \frac{R_L T}{p} \delta \rho_{WV}, \tag{9}$$

since $\frac{\rho_{WV} R_L}{p} \delta T \ll \frac{R_L T}{p} \delta \rho_{WV}$ and since $\frac{\rho_{WV} R_L T}{p^2} \delta p$ is less than the other two terms. Please find a numerical example in the appendix.

The error in *m* was determined with the total error in the absolute humidity data, which is the sum of a time-independent systematic error, the noise error, and the representativeness error. The systematic error was obtained from previous comparisons with other sensors (Bhawar et al. 2011) and the WVDIAL equation error propagation (Wulfmeyer and Bösenberg 1998). Due to the self-calibration property of the WVDIAL, the results revealed a very low systematic error of approximately 3 %, and so this error could be neglected in the DA process. It is one of the big advantages of the WVDIAL methodology that the corresponding measurements can be considered as bias-free or very small and unknown, and thus used as a reference. Hence we can only consider the statistical uncertainty for DA studies. Regarding the bias of the model, we constrained ourselves to the quality control of the data input to the model at the time of assimilation by introducing a new variable *max_error_q_DIAL*, into the WRF model registry that is described later in this section. The model bias greatly depends upon the model physics, which was not modified in this research.

The noise error can be determined in near-real-time by the determination of the autocovariance function of the high-resolution absolute humidity time series at each height. This method is explained in detail in Lenschow et al. (2000) and Wulfmeyer et al. (2016) and is routinely implemented in the IPM data-processing algorithms. Another advantage of the temporal resolution of time series data is that it allows an estimate of the representativeness error to be obtained. If we apply the Taylor hypothesis to the water vapor time–height cross section measured in a grid box of the model system, the water vapor variability will be representative for this box for a time period $\Delta T \approx \frac{\Delta x}{V}$, where Δx is the horizontal grid increment of the model and *V* is the horizontal wind speed. Using autocovariance function analysis, it is possible to separate atmospheric variance and noise variance to produce information about the accuracy of the measurement and the atmospheric variability. If the autocovariance is taken at lag 0, which is equivalent to calculating the total variance of the time series, we can take this as an estimate of the total error consisting of the noise error variance and the variance of the representativeness error so that we can write

$$\delta \rho_{WV}(z) \cong \sqrt{\text{var}(\rho_{WV}(t, z))}$$

$$\cong \sqrt{\delta \rho_{WV, \text{noise}}(z)^2 + \delta \rho_{WV, \text{represent}}(z)^2}. \tag{10}$$

These error profiles were calculated by averaging temporally over a 20-minute window of ±10 minutes around the time-step of the assimilation.

The new operator contains a couple of further essential data-processing steps. The WRFDA system assimilates observations obtained from various instruments. The initial step is the conversion of raw observations from these instruments to the LITTLE R format. LITTLE R is an ASCII-based file format and is an intermediate format used by the WRFDA to assimilate any number of observation types in a universal manner. The observation preprocessor (OBSPROC) of the WRFDA package reads only observations in the LITTLE R format. The OBSPROC removes the observations which do not fit in the specified temporal and spatial domain. Also it applies a number of other tasks like reordering and merging or deleting duplicate data.

As a starting point in our efforts toward developing an exclusive forward operator for the atmospheric products derived by lidar, an already-existing atmospheric infrared sounding retrieval (AIRSRET) or the FM-133 observation operator was used. We tested the AIRSRET operator because this operator has temperature and WVMR fields, which are basically the lidar end-products, in the model. The AIRSRET operator
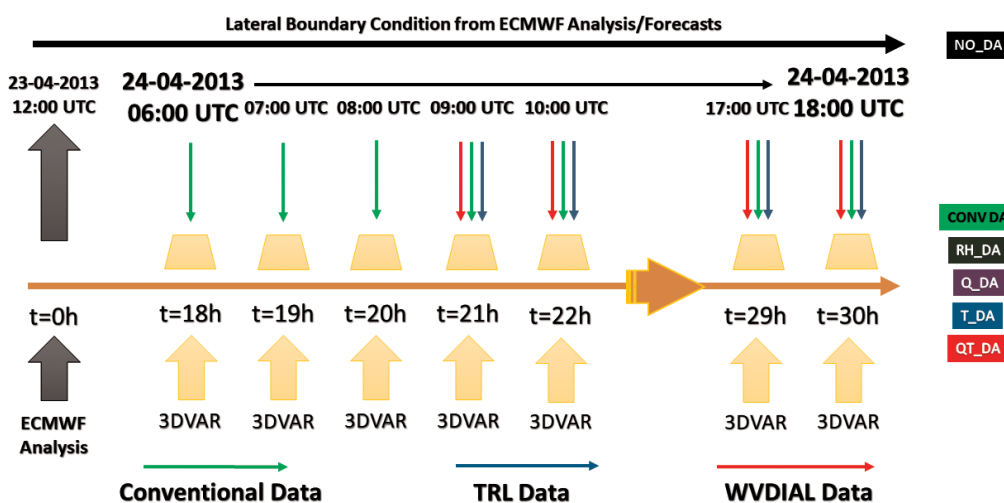
Fig. 5.   Schematic of the 3DVAR rapid update cycle initialized from the ECMWF analysis. A spin-up of 18 hours was performed until 06 UTC on 24 April 24 2013. Five experiments with different setups were performed. NO_DA (black) is the run with no data assimilation, CONV_DA (green) is the control run assimilating conventional data from 06 UTC to 18 UTC, RH_DA (olive green) is the assimilation with WVDIAL and conventional data from 09 UTC to 18 UTC using the RH operator, T_DA (blue) is the assimilation with TRL and conventional data from 09 UTC to 18 UTC, Q_DA (dark purple) is the assimilation with WVDIAL and conventional data from 09 UTC to 18 UTC using the TDLIDAR operator, and QT_DA (red) is the assimilation with WVDIAL, TRL, and conventional data from 09 UTC to 18 UTC using the TDLIDAR operator.

takes RH and temperature data and then converts them to WVMR:

$$m = \frac{RH\, E(T)}{1.607\, p} \quad [\text{kg kg}^{-1}], \tag{11}$$

which is basically Eq. (5). In the new operator, the WRFDA code was modified in such a way that the RH field was replaced by the WVMR data field by using Eq. (8). We call this new operator the thermodynamic lidar (TDLIDAR) operator.

The vertical profiles of the WVMR and temperature fields are linearly interpolated from the model levels to the observation data levels according to

$$(\rho_w, T(l_{in})) = [(\rho_w, T(l+1)) - (\rho_w, T(l))]\delta z$$
$$+ (\rho_w, T(l)) \qquad [\text{kg m}^{-3}, \text{K}]. \tag{12}$$

Here l is the model vertical level and l_{in} is the observation point within the model levels $l + 1$ and l. z is the height difference between two model levels.

As the total observation error for moisture measurements obtained from lidar is much lower than that for conventional datasets, a new error factor *max_error_q_DIAL* was incorporated in the WRFDA registry.

This new error factor enables the user to adjust the size of the error window through which the observations are ingested by the model. The observations are ingested only if the innovation or the difference between the observation and the first guess fall within m_{err} (Eq. 13). The model filters out low-quality WVDIAL observations that have a significant difference with the first guess of the model. The filtering is done with the help of this variable. The error factor is a scalar quantity which is multiplied by δm, the observation error, to get

$$m_{err} = \delta m \times max\_error\_q\_DIAL. \tag{13}$$

The error factor can be included in the WRFDA name list under section wrfvar 5 as *max_error_q_DIAL*. We did not yet introduce a separate registry variable for the temperature. However, we will incorporate the error factor for temperature in the next version of the operator.

3.4   Experimental setup

The assimilation was designed with 10 assimilation time-steps with hourly intervals between them. As shown in Fig. 5, the RUC was started after a spin-up period of 18 hours from 12 UTC 23rd April to 06

UTC 24th April, 2013. This spin-up was necessary for the model to stabilize itself with the initial and boundary conditions so that the model could then be forced in any desired manner. Only after a minimum spin-up time period are the model forecasts reliable for further analysis through DA.

We conducted 6 experiments: 1) a run (NO_DA) with no assimilation, 2) a conventional run (CONV_DA) with all the conventional data assimilated—the control run, 3) a TRL DA (T_DA) with TRL data assimilated along with conventional data using the standard TEMP forward operator, 4) a WVDIAL DA (Q_DA) with WVMR data assimilated along with conventional data using the TDLIDAR operator, 5) a WVDIAL DA (RH_DA) with RH data assimilated along with conventional data using the RH operator, and 6) finally the combined WVDIAL and TRL DA run (QT_DA) with WVMR and temperature lidar data along with conventional data assimilated using the TDLIDAR operator. In the Q_DA run, since the new operator also required the input of a temperature profile, we used for this the background temperature. After the initial spin-up of 18 hours, the CONV_DA run was initiated for three cycles starting from 0600 UTC each hour. At 0900 UTC, the other DA runs commenced with the forecast based on the 0800 UTC analysis that was valid for 0900 UTC as the background for that assimilation time-step. From 0900 UTC, all DA runs including the CONV_DA initiated from 0600 UTC were cycled till 1800 UTC (Fig. 5). In addition, a preconditioning DA run that included only hourly conventional data between 0600 and 0800 UTC was carried out to prepare the lidar DA and then to analyze the exclusive impact of the lidar data.

4.   Results

We analyzed the impact of assimilating the temperature and WVMR by applying TDLIDAR and also the RH forward operator for comparison. This section is divided into 4 subsections: first, the single observation tests for WVMR and temperature are described followed by an analysis of the sensitivity to the WVMR error factor, the impact of the temperature, and finally the impact of WVMR. The results of the assimilations are compared with available, independent radiosondes, which were launched every two hours during the IOP. It is important to note that the radiosonde measurements performed during the IOP were not assimilated in any of the experiments conducted.

4.1   Single observation tests

The spatial impact of assimilating an observation into the 3DVAR DA system is dependent on the structure of B. In order to understand the behavior of B, single observation tests (SOTs) were conducted. As we assimilated the WVMR and temperature profiles that also included experiments with background temperature profiles into the WRFDA system, the correlation of WVMR and temperature needs to be understood to interpret the combined impact with B. Since we were interested in the impact of WVMR data in the WRFDA system, an increment of 1 g kg^{-1} with a unit error of 1 g kg^{-1} was assigned at model level 10, which was approximately 255 m above ground level. This height was chosen to investigate the impact of assimilating near-surface observations. The impact on the vertical profile of the SOT is shown in Fig. 6a. The assimilation of a pseudo-WVMR observation of 1 g kg^{-1} results in an analysis increment of 0.3 g kg^{-1} at model level 10. As there is an increment in the WVMR analysis, there is a corresponding decrement in the temperature analysis at the same sigma level describing the correlation of temperature and WVMR in the DA system. The temperature at sigma level 10 has undergone an analysis decrement of 0.15 K. The impact of the assimilated WVMR pseudo-variable has a Gaussian-like distribution response across the vertical levels. While the WVMR assimilation created an increment in the WVMR variable, not only in the model level where assimilation was done but also in the model levels above, the temperature showed an opposite response. Figures 6b and 6c show the spatial impact of the SOT conducted at model level 10 for an assimilation carried out over the whole model domain. The impact of the assimilation has the highest WVMR increment at the point of assimilation and decreases radially with a Gaussian-like shape. The results for the temperature are similar but with the opposite sign. The WVMR increment was 0.1 g kg^{-1} to 0.3 g kg^{-1} over a region 250 km in diameter (Fig. 6b), while the temperature decrement was 0.1 K to 0.15 K over a region with a 300-km diameter (Fig. 6c). A similar SOT with a 1-K temperature increment and error of 1 K was also carried out at model level 10. Figure 6d shows the vertical profile of the SOT used for the temperature increment. An analysis increment of 0.28 K resulted from the SOT with a corresponding decremented response of 0.17 g kg^{-1} for the WVMR. The temperature increment was 0.1 K to 0.28 K over a region 300 km in diameter (Fig. 6e), while the WVMR decrement was 0.1 g kg^{-1} to 0.17 g kg^{-1} over a region 150 km in diameter (Fig. 6f).
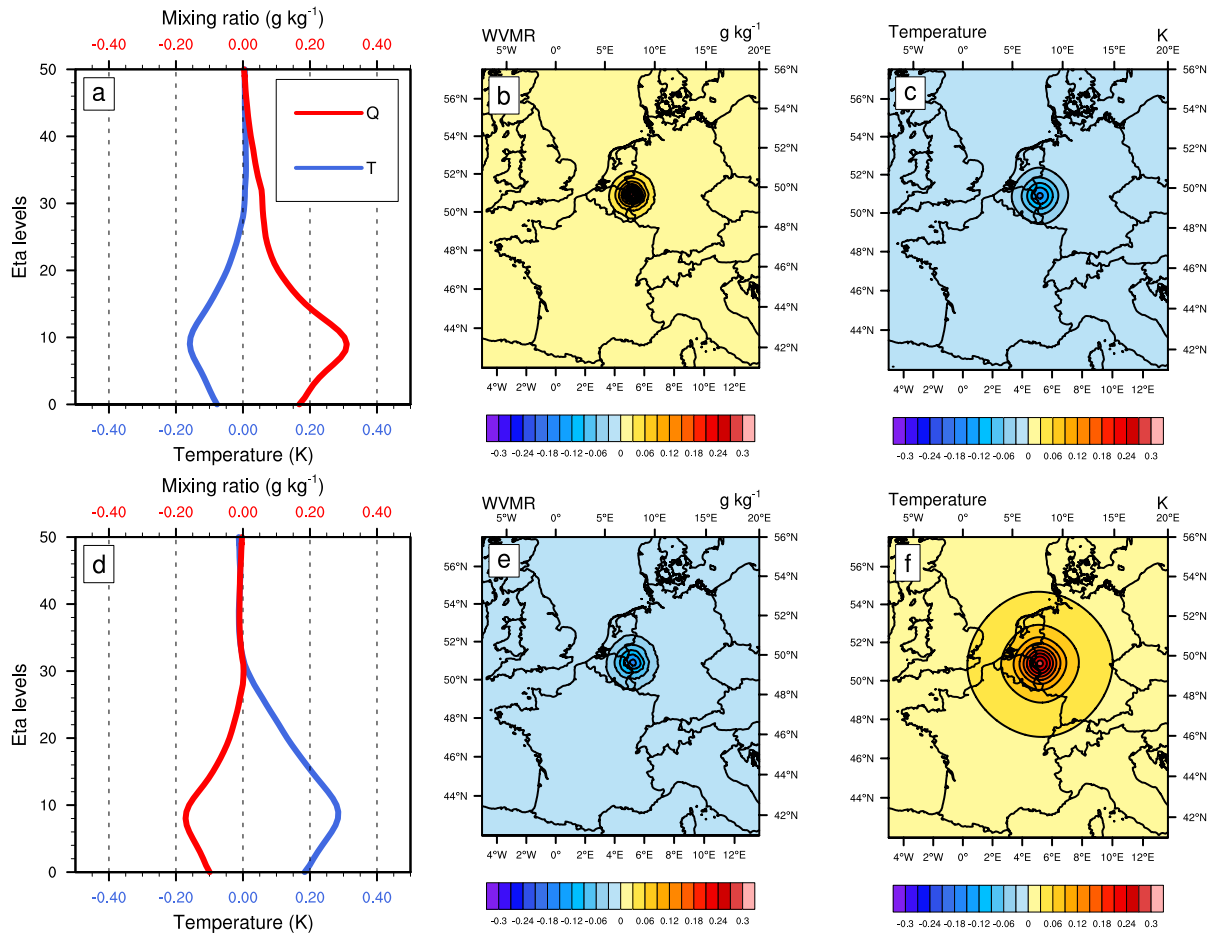
Fig. 6.   Vertical profiles and spatial distribution of analysis increments from single observation tests (SOTs) performed for a WVMR increment of 1 g kg^{-1} and 1-K temperature increment at model level 10 (255 m AGL). (a) Vertical profile, (b) and (c) spatial distribution of the analysis increments resulting from the WVMR SOT. (c), (e) and (f) results of the temperature SOT.

4.2   Sensitivity to WVMR error factor

In order to test the sensitivity to the error factor, the QT_DA experiment was conducted in two modes: one with the factor *max_error_q_DIAL* = 1 (QE1) and the other with *max_error_q_DIAL* = 4 (QE4). There were considerable differences in the model outputs of the two experiments since the number of observations assimilated was different in QE1 and QE4. Although the number of observations assimilated in QE1 and QE4 at 09 UTC were similar at 46 and 51, respectively, the later time-steps differed in terms of the number of observations assimilated, which was greater for QE4 than for QE1. The total number of observations during each assimilation cycle was 70.

The model rejected most of the observations in the interfacial layer, where the gradient of WVMR was high, since the observations were too far away from the first guess. The difference between the observation and the first guess value of any variable (innovation) decides whether the observation should be assimilated or not. The vertical profile of the analysis, profile of the background, and the WVDIAL WVMR observation profile along with its error bars are depicted in Fig. 7. From Fig. 7 we can see that at 09 UTC, QE4, which used 51 observations, shows a clearer impact on the vertical profile at 09 UTC than QE1, which used 46 observations. The QE1 profile has a higher deviation from the WVDIAL observations in the PBL than QE4. The WVMR profile from the WVDIAL has a low observation error until a height of 1300 m but
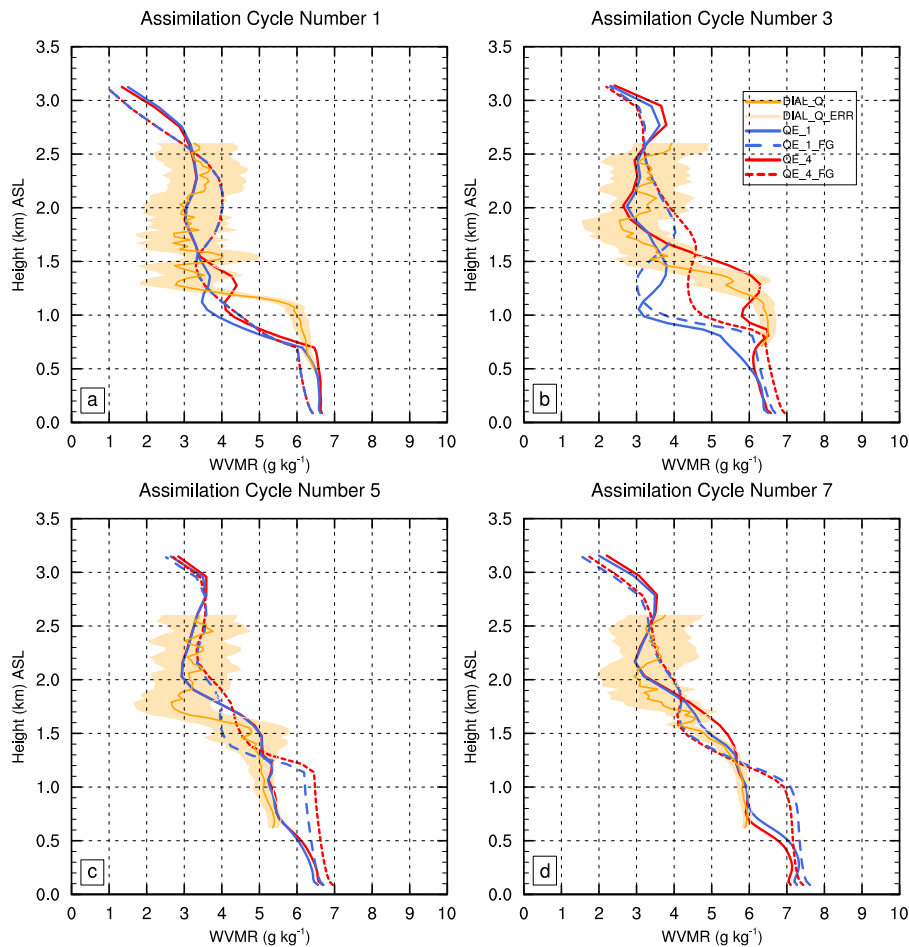
Fig. 7. Vertical profiles of the WVMR from 09 UTC, 11 UTC, 13 UTC, and 15 UTC on 24 April 2013 for the QE1 and QE4 experiments along with WVDIAL observations and associated error bars. The solid line represents the analysis profile and the dashed line the background profile.

grows significantly above this height. Hence the error window in the PBL is too small for the observations to be ingested into the DA.

The choice of the error factor is crucial for the quality of the model output. If it is too low, the model rejects most of the observations, not letting the model adapt toward the observations, which in turn does not improve the analysis. Otherwise, the model ingests all the observations including observations with considerable errors compared to the real-time observations, and this can cause the quality of the analysis to decrease. In this study, the error factor was fixed as four times the DIAL WVMR observation error, which was considered enough for the experiments to pass the quality check.

4.3 Temperature

Figure 8 depicts the temperature profiles at the assimilation time steps 09, 11, 13, and 15 UTC of all five experiments together with TRL and radiosonde observations. The radiosonde observations provided by the KIT cube (Kalthoff et al. 2013) were quality controlled before validation of the temperature profiles since GRAW DFM-09 radiosondes have a significant bias (Ingleby 2017). At these time-steps in the PBL, the NO_DA experiment showed a maximum deviation of around 2 K, which was less than the difference between the other DA experiments and the radiosonde observations. In the other five profiles where DA was performed, the temperature profiles significantly improved in the PBL. The CONV_DA and T_DA runs show a significant improvement in the temperature
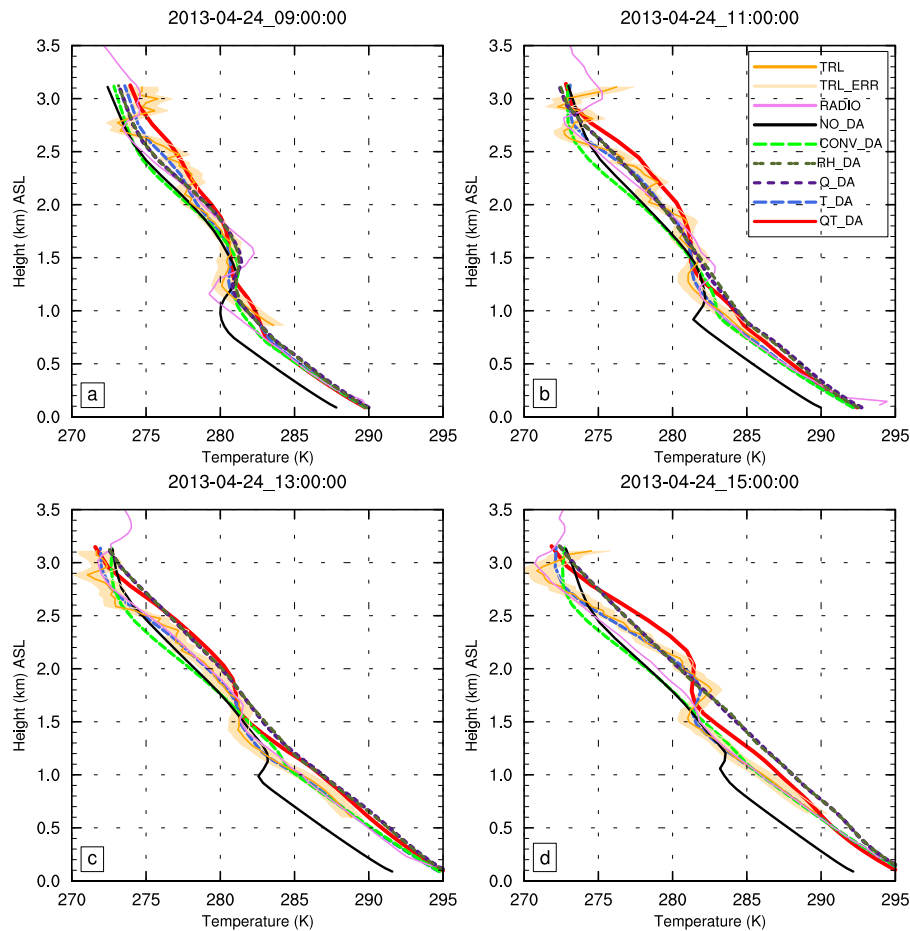
Fig. 8.   Temperature profiles of TRL, radiosondes, and analyses at (a) 09 UTC, (b) 11 UTC, (c) 13 UTC, and (d) 15 UTC. The TRL observations (orange) along with their total errors shown by error bars are plotted up to 3000 m AGL. Radiosonde observations (violet) which were not assimilated are plotted for reference. Black: NO_DA, green: CONV_DA, olive green: RH_DA, dark purple: Q_DA, blue: T_DA, and red: QT_DA.

profile in the PBL compared to the NO_DA run for all four time-steps. Q_DA, RH_DA, and QT_DA agree well with the radiosonde at 09 UTC in the PBL but start to deviate slowly to a higher temperature value after the first time-step. The Q_DA and RH_DA deviate by more in the PBL compared to the other three DA runs since no external temperature profile was assimilated. As the height increases, the CONV_DA profile becomes similar to the NO_DA profile. This is due to a lack of data points above the PBL in the conventional observations. However, after assimilating the TRL data along with the conventional data into the model, the deviation is reduced. In the interfacial layer and the lower free troposphere above this, the T_DA temperature profile, now having assimilated ample data points, is in good agreement with the TRL

profile at all four assimilation time-steps. The radiosonde profile is almost the same as the TRL profile for 09, 11, and 13 UTC but deviates above the PBL at 15 UTC. There is a difference of almost 1 K above the PBL; this gradually decreases with increasing height. This difference occurs due to the decrease of the SNR in the TRL profiles with height and the increase in distance between the sensors. The Q_DA, RH_DA, and QT_DA profiles in the lower free troposphere, deviate by less than 1 K and 2 K in the morning and afternoon, respectively, compared to the radiosonde observations. However, in the interfacial layer, the QT_DA is able to capture the inversion at all four time-steps, which Q_DA and RH_DA cannot. Figure 8 shows that Q_DA, RH_DA, and QT_DA deviate by more at higher ambient temperatures. In short, Q_
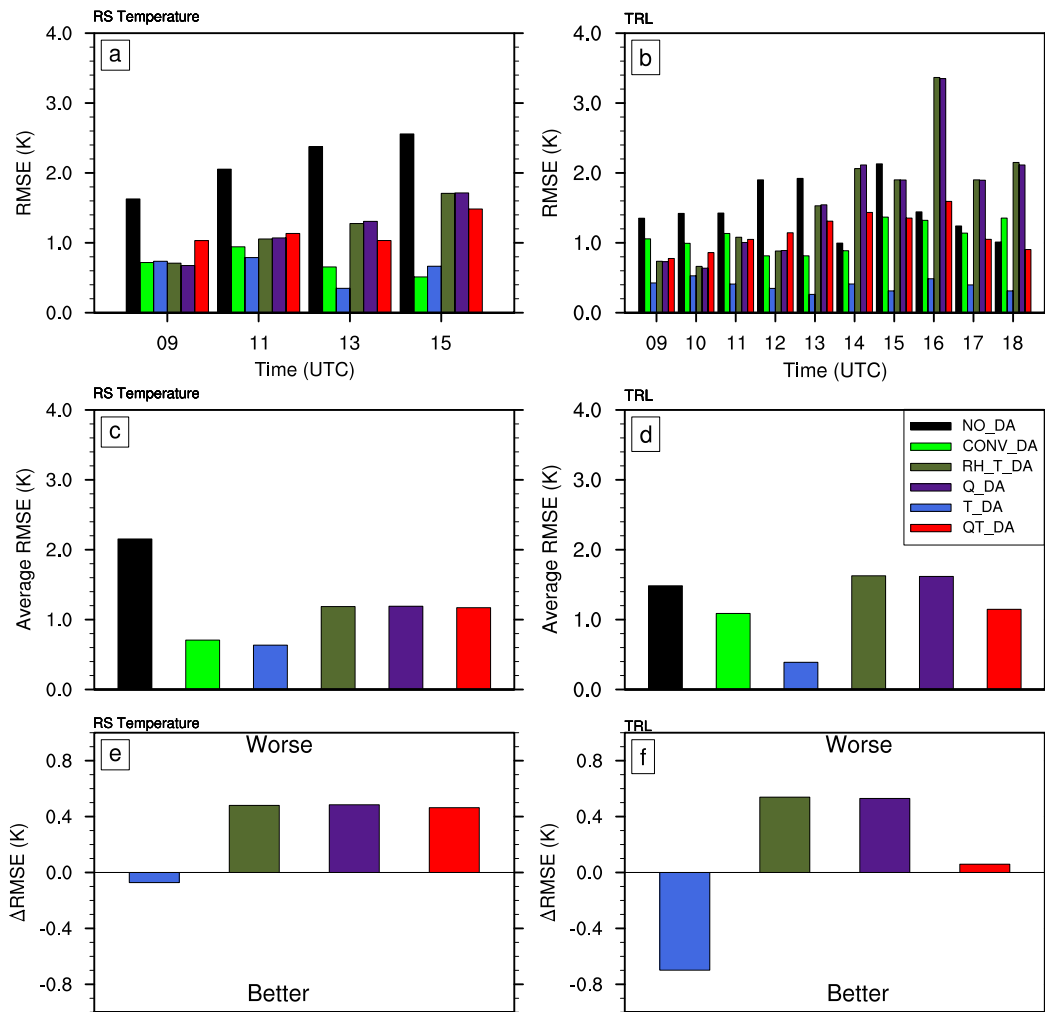
Fig. 9. Temperature RMSE of the analyses compared to local radiosonde data not assimilated into the model togeth-
er with assimilated TRL observations. (a) Comparison of the RMSE at the four assimilation time-steps (09, 11, 13,
15 UTC) with respect to the radiosonde data and (b) comparison of the RMSE with respect to the TRL observa-
tions at the 10 assimilation time-steps from 09 UTC to 18 UTC 24 April 2013. (c) and (d) comparison of the over-
all temperature RMSE for the corresponding time-steps for (a) and (b), respectively. (e) and (f) depict the relative
change in the average RMSE of (c) and (d), respectively, compared to the RMSE of CONV_DA.

DA, RH_DA, and QT_DA do not further improve
the temperature profiles of the model compared to the
improvement made by T_DA.

Figures 9a and 9b shows the RMSE with respect
to the radiosonde data for all four assimilation times
shown in Fig. 8 and the RMSE with respect to lidar
data for all ten assimilation times, respectively. The
overall average RMSE for each experiment (Figs. 9c,
d), and the relative change in the RMSE for other DA
experiments with respect to CONV_DA (Figs. 9e, f).
At 09 UTC, in Fig. 9c, CONV_DA and T_DA have

almost the same RMSE, though the radiosonde tem-
perature profile deviates from the TRL observations
slightly in the upper part of the PBL region. At 11 and
13 UTC, the RMSE of T_DA has the lowest value. At
15 UTC, the RMSE is higher due to the difference be-
tween the TRL and radiosonde profiles above the PBL
which has been discussed earlier. Q_DA, RH_DA, and
QT_DA have a slightly higher RMSE than the other
two DA runs but show an improvement compared to
the NO_DA experiment. Compared to CONV_DA,
the relative change in the RMSE (ΔRMSE) in Fig. 9e

for T_DA shows a decrease of 0.1 K, but Q_DA, RH_DA, and QT_DA show an increase of 0.5 K, 0.5K, and 0.45 K, respectively.

The RMSE of the analysis compared to the lidar observations is shown in Fig. 9b for all 10 assimilation time-steps. Q_DA, RH_DA, and QT_DA overestimated the temperature during daytime and, hence, the temperature RMSE with respect to the TRL observations increases from the first assimilation to the later cycles and decreases again for the final cycle. Q_DA and RH_DA have a higher RMSE than NO_DA for later cycles. The interesting feature to note here is that when the amount of moisture in the boundary layer is higher—that is, from 0900 UTC to 1100 UTC and from 1400 UTC to 1800 UTC—than between 1200 UTC and 1300 UTC, the assimilation has a higher impact. The RMSE for QT_DA is less than for CONV_DA during this time period. Between 1200 UTC and 1300 UTC, when the moisture in this region is lower, the temperature is overestimated, leading to a higher RMSE during this time. This is again a clear impact of the static nature of the background error covariance. Due to these counteracting impacts of the assimilation at different time-periods, the RMSEs for CONV_DA and QT_DA are similar in magnitude. The T_DA temperature RMSE is mostly constant over the assimilation cycles although there is a decrease of 0.2 K at around 1300 UTC from 0.4 K at 0900 UTC. From Fig. 9f, QT_DA has an increase of less than 0.05 K in ΔRMSE, which means that QT_DA did not worsen CONV_DA much, whereas Q_DA and RH_DA showed an increase of 0.5 K in ΔRMSE. T_DA shows a decrease of 0.7 K in ΔRMSE. In summary, T_DA outperformed all the other experiments in terms of the temperature impact.

4.4   Water vapor mixing ratio

Figure 10 depicts the profiles of the analyzed WVMR at the assimilation time steps 09, 11, 13, and 15 UTC for all the different DA experiments including the observations. The DIAL WVMR observations were limited to a height of 2.5 km since the observation error was higher than the observed value.

All the assimilation runs do not show much difference from NO_DA in the PBL at 09 UTC. The surface observations were well captured by all the experiments at 11 UTC except for Q_DA which shows insignificant values of WVMR. But in the PBL above the surface layer, Q_DA and QT_DA are in good agreement with the radiosonde and DIAL observations at later time-steps. The Q_DA and QT_DA profiles agree with the radiosonde and DIAL observations at

13 UTC and 15 UTC, whereas NO_DA, CONV_DA, and T_DA have higher values of WVMR in the PBL. The Q_DA and QT_DA profiles are similar to those of the other two assimilation experiments in the surface layer since there were no lidar observations available at those levels. NO_DA shows an overestimation in the WVMR of around 1 g kg^{-1} in the PBL. RH_DA did not outperform Q_DA and QT_DA as expected although it was close to QT_DA at 09 UTC.

The interfacial layer was best captured by Q_DA and QT_DA at all time-steps apart from the first assimilation time-step at 09 UTC. NO_DA and CONV_DA underestimated the WVMR at all time-steps, whereas T_DA shows a positive deviation at 13 UTC and 15 UTC in the interfacial layer. RH_DA shows a negative deviation at 11 UTC and a positive deviation at 15 UTC. The lower free troposphere impact for Q_DA, RH_DA, and QT_DA is in better agreement with the observations than compared to the other runs, which have mixed results. NO_DA and CONV_DA always have a positive deviation. T_DA has positive and negative deviations at 09 UTC and 15 UTC, respectively, but matches with Q_DA, RH_DA, and QT_DA at 11 UTC and 13 UTC. In short, Q_DA and QT_DA had a more major impact on the WVMR than the other experiments.

Figures 11a and 11b depict the WVMR RMSE compared to the radiosonde observations at 09, 11, 13, and 15 UTC, and the WVMR RMSE compared to the lidar observations at all ten assimilation time steps from 09 UTC to 18 UTC, respectively. The overall average of RMSE for each experiment are shown in Figs. 11c and 11d. The relative change in RMSE for the other DA experiments compared to CONV_DA are shown in Figs. 11e and 11f. Keeping in mind the radiosonde error due to drifting, Q_DA and QT_DA performed better than the other experiments although the difference with T_DA was less. The RMSE for RH_DA is the same as for CONV_DA although slightly better than T_DA and QT_DA. From Fig. 11a, the decline in WVMR RMSE as the assimilation cycle progresses is visible. Although the QT_DA RMSE decline rate is small, the decrease is consistent. Although the overall RMSE for Q_DA and QT_DA is closer to that for T_DA and CONV_DA, it is lower (Fig. 11c). The RMSE differences compared to CONV_DA are considerably less with magnitudes of +0.01 g kg^{-1} for RH_DA, +0.03 g kg^{-1} for QT_DA, and +0.05 g kg^{-1} for T_DA. Q_DA has a difference of −0.05 g kg^{-1} compared to CONV_DA.

Compared to the WVDIAL observations, the RMSEs in the WVMR (Figs. 11b, d) also have a
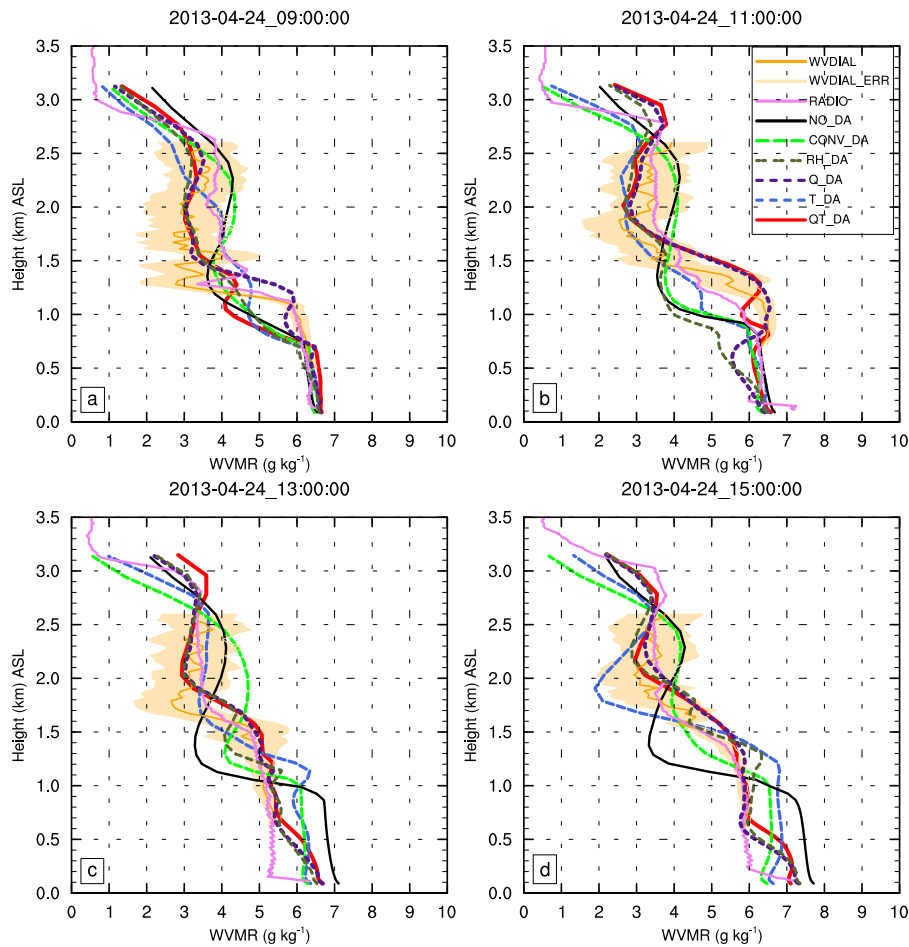
Fig. 10.   WVMR profiles of WVDIAL, radiosondes, and analyses at (a) 09 UTC, (b) 11 UTC, (c) 13 UTC, and (d) 15 UTC. The WVDIAL observations (orange) along with their total errors shown by error bars are plotted up to 2500 m AGL. Radiosonde observations (violet) which were not assimilated are plotted for reference. Black: NO_DA, green: CONV_DA, olive green: RH_DA, dark purple: Q_DA, blue: T_DA, and red: QT_DA are shown.

similar declining trend to those seen in the radiosonde comparisons in consecutive assimilation cycles, but the decline is higher. An important difference between the WVDIAL and radiosonde observations which needs to be considered is the error due to the temporal coverage of the two datasets. The WVDIAL dataset gives a complete profile of the atmosphere every 10 s, while the radiosonde provides data only from the point of ascent. The mean rate of ascent of the radiosondes launched during IOP 6 of the HOPE campaign was around 5 m s^{-1}. This means that the time taken for a radiosonde to cross the PBL (taking its height to be 1500 m) would be 5 minutes, which is still 30 times higher than the time required for obtaining a single lidar profile. This temporal resolution is not optimal

if the atmosphere is rapidly changing. Hence, the DIAL dataset is a continuous measurement whereas the radiosonde data are instantaneous ones. This also explains the reason why the DIAL dataset does not have such a smooth profile as the radiosonde data because the DIAL data capture all the fluctuations in the atmosphere. Q_DA and QT_DA (Fig. 11b) have the lowest RMSE in all the assimilation cycles; also, the declining trend for the RMSE in the successive assimilations proves that the model successfully corrects the WVMR. T_DA does not show a visible impact for successive assimilations. Hence, the WVMR RMSE for T_DA in Fig. 11b is always higher than for Q_DA and QT_DA. However, the WVMR RMSE for T_DA has a value similar to the CONV_DA. Although RH_
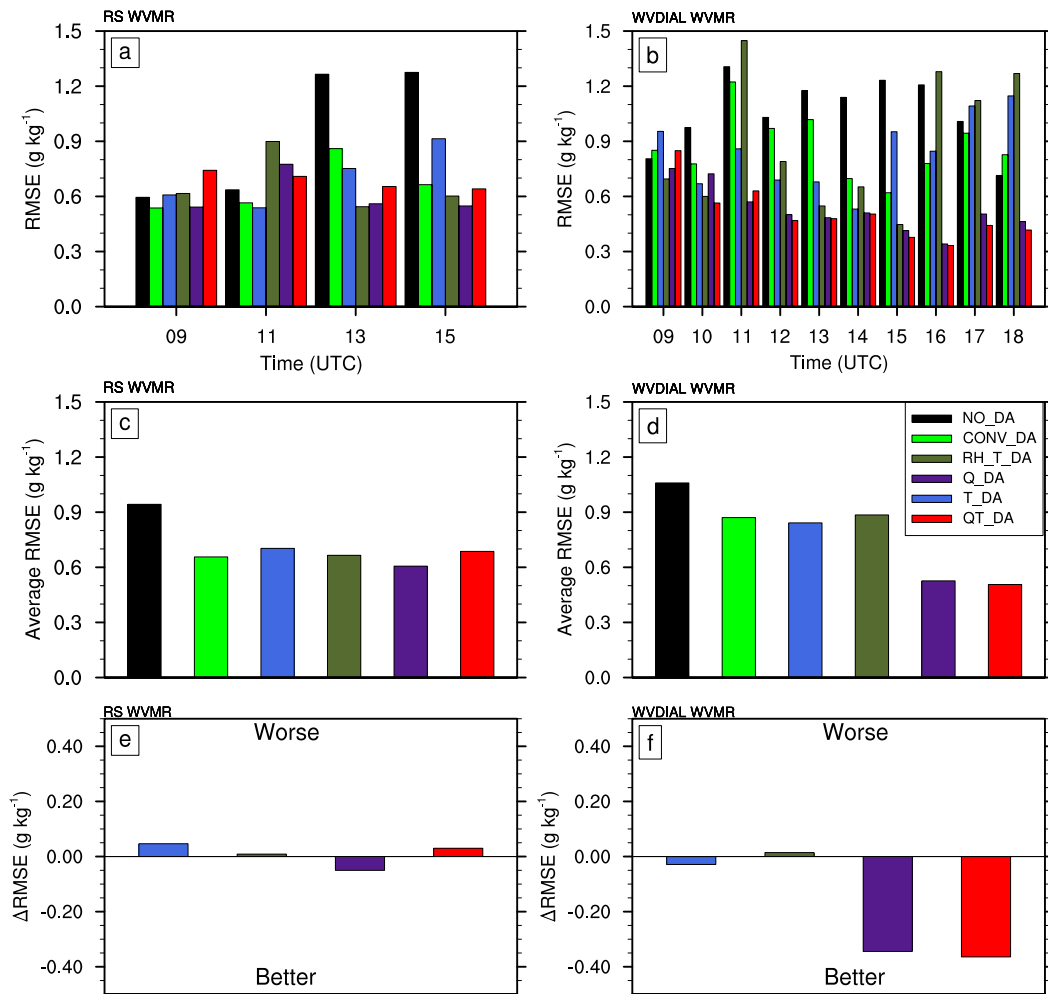
Fig. 11.   WVMR RMSE of the analyses compared to local radiosonde data not assimilated into the model together with assimilated WVDIAL observations. (a) Comparison of the RMSE at the four assimilation time-steps (09, 11, 13, 15 UTC) with respect to the radiosonde data and (b) comparison of the RMSE with respect to the WVDIAL observations at the 10 assimilation time-steps from 09 UTC to 18 UTC 24 April 2013. (c) and (d) compare the overall WVMR RMSE for the corresponding time-steps for (a) and (b), respectively. (e) and (f) depict the relative change in the average RMSE of (c) and (d) respectively, compared to the RMSE of CONV_DA.

DA has lower RMSE values than CONV_DA at 09 UTC and 10 UTC, later cycles have a higher RMSE. In Fig. 11d, the overall RMSE for QT_DA is the lowest for all the experiments. The ΔRMSE in Fig. 11f indicates that there is a decrease of 0.36 g kg^{-1} for QT_DA and 0.3 g kg^{-1} for Q_DA but only 0.03 g kg^{-1} for T_DA when compared to CONV_DA. RH_DA shows an increase of 0.02 g kg^{-1} compared to CONV_DA. Figure 12 shows an analysis of the difference between QT_DA and CONV_DA. The spatial analysis difference at 09 UTC and 18 UTC on 24 April 2013

are shown in Figs. 12a and 12b, respectively. A vertical cross-section of the analysis difference at 09 UTC is shown in Fig. 12c. In order to analyze the impact of the assimilated lidar data, a 6-hr forecast difference between QT_DA and CONV_DA initiated from 18 UTC is shown in Fig. 12d. However the assimilation impact cannot be due completely to the lidar observations and, presumably, the number of observations in the conventional data should be considered. The spatial analyses shown in Figs. 12a, 12b, and 12d are for a height of 2000 m, which is assumed to be the
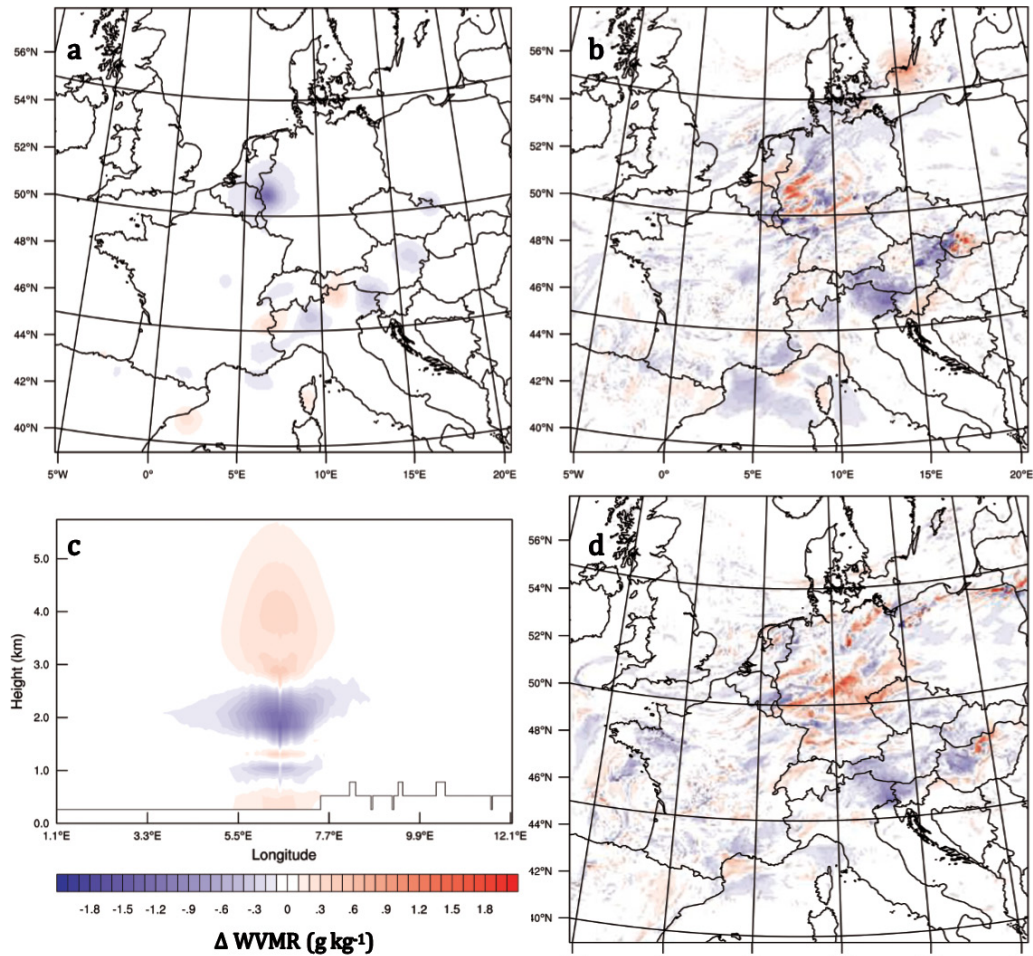
Fig. 12.   Analysis of the difference between QT_DA and CONV_DA at (a) 09 UTC and (b) 18 UTC 24 April 2013; (c) shows the vertical cross section of the difference, valid at 09 UTC; (d) is the six-hour forecast difference between QT_DA and CONV_DA initiated from 18 UTC. The spatial distributions are valid at a height of 2000 m ASL.

PBL top, where the impact is significant. The impact of a single lidar profile spreading over an area with a diameter of 300 km shows the potential of a network of lidars. The forecast difference after six hours initiated from 18 UTC (Fig. 12d) clearly shows that the impact of the assimilation is both enduring and stable since the impact of the assimilated lidar data lasts for short-range forecasts and does not lead to significant errors during this forecast range. The six-hour forecast difference does not exceed an absolute value of 1.2 g kg^{-1} in the areas near the lidar instrument location, which accounts for the stability of the atmosphere after assimilating the thermodynamic lidar data.

5.   Summary and outlook

In this study, we investigated the impact of assimilating WVMR and temperature data from lidar systems on the vertical structure of temperature and moisture inside the PBL. For this purpose, we applied WRF version 3.8.1 together with its 3DVAR DA system at a convection-permitting horizontal resolution of 2.5 km over central Europe. The DA system was operated in the RUC mode, meaning that the assimilations were hourly. For the present study, lidar data from the HOPE campaign were used for the assimilation. The IOP took place from 0900 UTC to 1800 UTC on the 24 April 2013 in western Germany on a clear-sky day with hardly any optically thick

clouds. Temperature data from heights of 500 m to 3000 m above the ground were taken for the experiment. WVMR data were taken from 400 m to 2500 m above the ground level. Data from lower levels had to be discarded due to the overlap error. Apart from the lidar measurements, there were four radiosonde launches at 09, 11, 13 and 15 UTC. The mean of these radiosonde measurements was used for calibrating the TRL and as an independent measurement for comparison with the model output since these radiosonde measurements were not assimilated in the DA system.

Six model runs were conducted for the whole impact study. A run (NO_DA) with no data assimilated, conventional data assimilation (CONV_DA) or the control run with only conventional observations from the ECMWF, TRL data assimilation (T_DA) along with the conventional dataset, WVMR data assimilation (Q_DA) along with the conventional dataset, RH data assimilation (RH_DA) along with the conventional dataset, and finally the WVMR and TRL data assimilation (QT_DA) along with the conventional data.

In this study, we introduced a new forward operator called TDLIDAR for direct WVMR DA, which was developed through the modification of an already-existing operator in the WRFDA system, the AIRSRET operator. Also, separate sensitivity tests were conducted with the QT_DA to study the sensitivity of the newly introduced error factor (*max_error_q_DIAL*) in the WRFDA registry. SOTs were conducted to analyze the response of the input WVMR and temperature data separately in the DA system. An increase in the WVMR resulted in a subsequent cooling at the point of assimilation in the model. On the other hand, an increase in the temperature resulted in a subsequent drying.

The impact of the assimilation of WVMR and temperature lidar data through the new forward operator was, overall, positive. The input observations were assimilated with a very low number of rejected observations: the model only rejected a few observations during the first assimilation cycle. The WVMR and temperature profiles of the model output indicated that the input lidar observations could correct the first guess during the assimilation process to a reasonable extent. From the results of the five DA runs, we conclude that, the assimilation of both temperature and WVMR lidar observations improved the thermodynamic profiles in the analyses. T_DA and Q_DA improved the temperature and moisture profiles, respectively, whereas QT_DA improved both compared to CONV_DA. RH_DA did not outperform either Q_DA or QT_DA in the study, showing that the TDLIDAR operator leads to a better impact than the RH operator. We quantified the analyses by their RMSE with respect to the assimilated lidar observations as well as independent radiosonde observations. However, the lidar observations were more suitable for model verification than radiosonde data because they point exactly to the zenith rather than along an irregular vertical track. The WVMR RMSE computed with respect to the WVDIAL observations for QT_DA reduced by 40 % compared to those computed for CONV_DA run whereas RH_DA did not show an overall improvement. This highlights that using the forward operator for the data input had a positive impact on the modeled WVMR variable. However, at the same time, the impact on the temperature was reduced due to the significant dependency between the WVMR and temperature variables in the analysis.

In real-time operational forecasting with data assimilated from active remote-sensing instruments like lidars, which provide data with a very low observational bias, a deterministic DA system whose correlation statistics are derived from a set of forecast error differences might not provide the best analysis. With the introduction of a flow-dependent background error-covariance matrix with the help of ensemble-based DA systems, the cross-correlation between the temperature and humidity variables is expected to be a better representative of the real-time scenario. The matrix **B** in ensemble-based DA systems reflects the dynamic nature of the atmosphere. Thus, we plan to assimilate thermodynamic lidar data with ensemble DA techniques in the future. Furthermore, modules for the conversion of absolute humidity and specific humidity to the WVMR will be incorporated. Currently, with a limited number of lidars, we limited our studies to convective-scale DA. However, in the future, with a larger number of lidars which operate as a network, we can enhance our studies to synoptic-scale DA. We foresee synoptic-scale DA of lidar networks as very beneficial for operational numerical weather forecasting centers.

Appendix

The total derivative of *RH* expands as per the equation.

$$\delta RH = 1.607 \frac{p}{E(T)} \delta m + 1.607 \frac{m}{E(T)} \delta p$$

$$- 1.607 \frac{mp}{E(T)^2} \delta E(T)$$

$$= \frac{1.607}{E(T)} \left[p\delta m + m\delta p - mp \left(\frac{17.67 \times 243.5}{(T - 29.65)^2} \right) \delta T \right].$$

$$(A1)$$

After dividing Eq. A1 by *RH* we get the relative error equation

$$\frac{\delta RH}{RH} = \frac{\delta m}{m} + \frac{\delta p}{p} - \left(\frac{17.67 \times 243.5}{(T - 29.65)^2} \right) \delta T. \qquad (A2)$$

For normal atmospheric conditions

$T = 300 \text{ K}, \quad p = 100{,}000 \text{ Pa}, \quad m = 0.01 \text{ kg kg}^{-1},$
$\delta m = 0.001 \text{ kg kg}^{-1}, \quad \delta T = 1.1 \text{ K}, \quad \delta p = 50 \text{ Pa}.$

We took a normal value of 10 g kg^{-1} and an error of 1 g kg^{-1} for the mixing ratio in the numerical example, which are similar to the values for the absolute humidity measurements from the WVDIAL. Similarly, a temperature error of 1.1 K was taken for the TRL measurements. Substituting the above values in Eq. A2, we get these values for the individual terms:

$$\frac{\delta m}{m} = 10 \%,$$

$$\frac{\delta p}{p} = 0.05 \%,$$

$$\left(\frac{17.67 \times 243.5}{(T - 29.65)^2} \right) \delta T = 6.5 \%.$$

The WVMR error δm expands to

$$\delta m = \frac{R_L T}{p} \delta \rho_{WV} + \frac{\rho_{WV} R_L}{p} \delta T - \frac{\rho_{WV} R_L T}{p^2} \delta p. \qquad (A3)$$

For normal atmospheric conditions

$T = 300 \text{ K}, \quad p = 100{,}000 \text{ Pa},$
$R_L = 287 \text{ J K}^{-1} \text{ kg}^{-1}, \quad \rho_{WV} = 0.01 \text{ kg kg}^{-1},$
$\delta \rho_{WV} = 0.001 \text{ kg kg}^{-1}, \quad \delta T = 1.1 \text{ K}, \quad \delta p = 50 \text{ Pa}.$

Substituting the above values we get

$$\frac{R_L T}{p} \delta \rho_{WV} = \underline{861} \times 10^{-6} \text{ kg kg}^{-1},$$

$$\frac{\rho_{WV} R_L}{p} \delta T = \underline{31.57} \times 10^{-6} \text{ kg kg}^{-1},$$

$$\frac{\rho_{WV} R_L T}{p^2} \delta p = \underline{4.305} \times 10^{-6} \text{ kg kg}^{-1}.$$

References

Adam, S., A. Behrendt, T. Schwitalla, E. Hammann, and V. Wulfmeyer, 2016: First assimilation of temperature lidar data into an NWP model: Impact on the simulation of the temperature field, inversion strength and PBL depth. *Quart. J. Roy. Meteor. Soc.*, **142**, 2882–2896.

Apituley, A., K. M. Wilson, C. Potma, H. Volten, and M. de Graaf, 2009: Performance assessment and application of Caeli — A high-performance Raman lidar for diurnal profiling of water vapour, aerosols and clouds. *Proceedings of the 8th International Symposium on Tropospheric Profiling*, Delft, Netherlands, S06–O10, 4 pp. [Available at http://projects.knmi.nl/cesar/istp8/data/1753005.pdf.]

Aranami, K., T. A. Hara, Y. A. Ikuta, K. O. Kawano, K. E. Matsubayashi, H. I. Kusabiraki, T. A. Ito, T. A. Egawa, K. O. Yamashita, Y. U. Ota, and Y. O. Ishikawa, 2015: A new operational regional model for convection-permitting numerical weather prediction at JMA. *Research activities in atmospheric and oceanic modelling*. WCRP Report No. 12/2015, Astakhova, E. (ed.), CAS/JSC Working Group on Numerical Experimentation, 2 pp. [Available at https://www.wcrp-climate.org/WGNE/BlueBook/2015/individual-articles/05_Aranami_Kohei_asuca.pdf.]

Arshinov, Y., S. Bobrovnikov, I. Serikov, A. Ansmann, U. Wandinger, D. Althausen, I. Mattis, and D. Müller, 2005: Daytime operation of a pure rotational Raman lidar by use of a Fabry–Perot interferometer. *Appl. Opt.*, **44**, 3593–3603.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A Three-Dimensional Variational Data Assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914.

Bauer, H.-S., T. Schwitalla, V. Wulfmeyer, A. Bakhshaii, U. Ehret, M. Neuper, and O. Caumont, 2015: Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF – A performance test. *Tellus*, **67**, 25047, doi:10.3402/tellusa.v67.25047.

Behrendt, A., and J. Reichardt, 2000: Atmospheric temperature profiling in the presence of clouds with a pure rotational Raman lidar by use of an interference-filter-based polychromator. *Appl. Opt.*, **39**, 1372–1378.

Behrendt, A., T. Nakamura, and T. Tsuda, 2004: Combined

temperature lidar for measurements in the troposphere, stratosphere, and mesosphere. *Appl. Opt.*, **43**, 2930, doi:10.1364/ao.43.002930.

Behrendt, A., V. Wulfmeyer, A. Riede, G. Wagner, S. Pal, H. Bauer, M. Radlach, and F. Späth, 2009: Three-dimensional observations of atmospheric humidity with a scanning differential absorption lidar. *Remote Sens. Clouds Atmos. XIV*, **7475**, 74750L, doi:10.1117/12.835143.

Behrendt, A., V. Wulfmeyer, E. Hammann, S. K. Muppa, and S. Pal, 2015: Profiles of second- to fourth-order moments of turbulent temperature fluctuations in the convective boundary layer: First measurements with rotational Raman lidar. Atmos. *Chem. Phys.*, **15**, 5485–5500.

Benjamin, S. G., S. S. Weygandt, J. M. Brown, T. L. Smith, T. Smirnova, W. R. Moninger, B. Schwartz, E. J. Szoke, and K. Brundage, 2004: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.

Benjamin, S. G., S. S. Weygandt, J. M. Brown, M. Hu, C. R. Alexander, T. G. Smirnova, J. B. Olson, E. P. James, D. C. Dowell, G. A. Grell, and H. Lin, 2016: A North American hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev.*, **144**, 1669–1694.

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Wea. Rev.*, **128**, 644–667.

Berre, L., S. E. Ştefănescu, and M. B. Pereira, 2006: The representation of the analysis effect in three error simulation techniques. *Tellus*, **58**, 196–209.

Bhawar, R., P. Di Girolamo, D. Summa, C. Flamant, D. Althausen, A. Behrendt, C. Kiemle, P. Bosser, M. Cacciani, C. Champollion, T. Di Iorio, R. Engelmann, C. Herold, D. Müller, S. Pal, M. Wirth, and V. Wulfmeyer, 2011: The water vapour intercomparison effort in the framework of the convective and orographically-induced precipitation study: Airborne-to-ground-based and airborne-to-airborne lidar systems. *Quart. J. Roy. Meteor. Soc.*, **137**, 325–348.

Bielli, S., M. Grzeschik, E. Richard, C. Flamant, C. Champollion, C. Kiemle, M. Dorninger, and P. Brousseau, 2012: Assimilation of water-vapour airborne lidar observations: Impact study on the COPS precipitation forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 1652–1667.

Blumberg, W. G., D. D. Turner, U. Löhnert, and S. Castleberry, 2015: Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part II: Actual retrieval performance in clear-sky and cloudy conditions. *J. Appl. Meteor. Climatol.*, **54**, 2305–2319.

Bolton, D., 1980: The computation of equivalent potential temperature. *Mon. Wea. Rev.*, **108**, 1046–1053.

Bösenberg, J., 1998: Ground-based differential absorption lidar for water-vapor and temperature profiling: Meth-

odology. *Appl. Opt.*, **37**, 3845–3860.

Brocard, E., R. Philipona, A. Haefele, G. Romanens, A. Mueller, D. Ruffieux, V. Simeonov, and B. Calpini, 2013: Raman lidar for meteorological observations, RALMO. Part 2: Validation of water vapor measurements. *Atmos. Meas. Tech.*, **6**, 1347–1358.

Brousseau, P., L. Berre, F. Bouttier, and G. Desroziers, 2011: Background-error covariances for a convective-scale data-assimilation system: AROME-France 3D-Var. *Quart. J. Roy. Meteor. Soc.*, **137**, 409–422.

Browning, K. A., A. M. Blyth, P. A. Clark, U. Corsmeier, C. J. Morcrette, J. L. Agnew, S. P. Ballard, D. Bamber, C. Barthlott, L. J. Bennett, and K. M. Beswick, 2007: The convective storm initiation project. *Bull. Amer. Meteor. Soc.*, **88**, 1939–1955.

Cadeddu, M. P., G. E. Peckham, and C. Gaffard, 2002: The vertical resolution of ground-based microwave radiometers analyzed through a multiresolution wavelet technique. *IEEE Trans. Geosci. Remote Sens.*, **40**, 531–540.

Cooney, J., 1972: Measurement of atmospheric temperature profiles by Raman backscatter. *J. Appl. Meteor.*, **11**, 108–112.

Courtier, P., E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.

Dailey, P. S., and R. G. Fovell, 1999: Numerical simulation of the interaction between the sea-breeze front and horizontal convective rolls. Part I: Offshore ambient flow. *Mon. Wea. Rev.*, **127**, 858–878.

Di Girolamo, P., R. Marchese, D. N. Whiteman, and B. B. Demoz, 2004: Rotational Raman Lidar measurements of atmospheric temperature in the UV. *Geophys. Res. Lett.*, **31**, L01106, doi:10.1029/2003GL018342.

Dinoev, T., V. B. Simeonov, Y. Arshinov, S. Bobrovnikov, P. Ristori, B. Calpini, M. Parlange, and H. van den Bergh, 2013: Raman Lidar for meteorological observations, RALMO. Part 1: Instrument description. *Atmos. Meas. Tech.*, **6**, 1329–1346.

Dixon, M., Z. Li, H. Lean, N. Roberts, and S. Balland, 2009: Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the met office unified model. *Mon. Wea. Rev.*, **137**, 1562–1584.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158.

Engelbart, D. A. M., and E. Haas, 2006: LAUNCH-2005-International Lindenberg campaign for assessment of

humidity and cloud profiling systems and its impact on high-resolution modelling. *Proceeding of 7th International Symposium on Tropospheric profiling: Needs and technologies*, 11–17.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.

Fischer, L., 2013: *Statistical characterisation of water vapour variability in the troposphere*. Doctoral dissertation, Ludwig Maximilian University of Munich, Germany, 124 pp. [Available at https://edoc.ub.uni-muenchen.de/16208/1/Fischer_Lucas.pdf.]

Fisher, M., 2003: Background error covariance modelling. *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean*, 45–63.

Geerts, B., D. Parsons, C. L. Ziegler, T. M. Weckwerth, M. I. Biggerstaff, R. D. Clark, M. C. Coniglio, B. B. Demoz, R. A. Ferrare, W. A. Gallus, Jr., K. Haghi, J. M. Hanesiak, P. M. Klein, K. R. Knupp, K. Kosiba, G. M. McFarquhar, J. A. Moore, A. R. Nehrir, M. D. Parker, J. O. Pinto, R. M. Rauber, R. S. Schumacher, D. D. Turner, Q. Wang, X. Wang, Z. Wang, and J. Wurman, 2017: The 2015 plains elevated convection at night field project. *Bull. Amer. Meteor. Soc.*, **98**, 767–786.

Goldsmith, J. E. M., F. H. Blair, S. E. Bisson, and D. D. Turner, 1998: Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols. *Appl. Opt.*, **37**, 4979–4990.

Grzeschik, M., H. S. Bauer, V. Wulfmeyer, D. Engelbart, U. Wandinger, I. Mattis, D. Althausen, R. Engelmann, M. Tesche, and A. Riede, 2008: Four-dimensional variational data analysis of water vapor Raman lidar data and their impact on mesoscale forecasts. *J. Atmos. Oceanic Technol.*, **25**, 1437–1453.

Gustafsson, N., T. Janjić, C. Schraff, D. Leuenberger, M. Weissmann, H. Reich, P. Brousseau, T. Montmerle, E. Wattrelot, A. Bučánek, and M. Mile, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1218–1256.

Hammann, E., A. Behrendt, F. Le Mounier, and V. Wulfmeyer, 2015: Temperature profiling of the atmospheric boundary layer with rotational Raman lidar during the HD(CP)^2 Observational Prototype Experiment. *Atmos. Chem. Phys.*, **15**, 2867–2881.

Harnisch, F., M. Weissmann, C. Cardinali, and M. Wirth, 2011: Experimental assimilation of DIAL water vapour observations in the ECMWF global model. *Quart. J. Roy. Meteor. Soc.*, **137**, 1532–1546.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus A*, **38**, 111–136.

Honda, Y., M. Nishijima, K. Koizumi, Y. Ohta, K. Tamiya, T. Kawabata, and T. Tsuyuki, 2006: A pre-operational variational data assimilation system for a non-hydrostatic model at the Japan Meteorological Agency: Formulation and preliminary results. *Quart. J. Roy. Meteor. Soc.*, **131**, 3465–3475.

Hong, S.-Y., H. Park, H. B. Cheong, J. E. E. Kim, M. S. Koo, J. Jang, S. Ham, S. O. Hwang, B. K. Park, E. C. Chang, and H. Li, 2013: The global/regional integrated model system (GRIMS). *Asia-Pacific J. Atmos. Sci.*, **49**, 219–243.

Horváth, Á., O. Hautecoeur, R. Borde, H. Deneke, and S. A. Buehler, 2017: Evaluation of the EUMETSAT global AVHRR wind product. *J. Appl. Meteor. Climatol.*, **56**, 2353–2376.

Hu, M., M. Xue, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698.

Hu, M., S. G. Benjamin, T. T. Ladwig, D. C. Dowell, S. S. Weygandt, C. R. Alexander, and J. S. Whitaker, 2017: GSI three-dimensional ensemble-variational hybrid data assimilation using a global ensemble for the regional rapid refresh model. *Mon. Wea. Rev.*, **145**, 4205–4225.

Huang, X. Y., Q. Xiao, D. M. Barker, X. Zhang, J. Michalakes, W. Huang, T. Henderson, J. Bray, Y. Chen, Z. M. J. Dudhia, Y. Guo, X. Zhang, D.-J. Won, H.-C. Lin, and Y.-H. Kuo, 2009: Four-dimensional variational data assimilation for WRF: Formulation and preliminary results. *Mon. Wea. Rev.*, **137**, 299–314.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, doi:10.1029/2008JD009944.

Ingleby, B., 2017: An assessment of different radiosonde types 2015/2016. *ECMWF Tech. Memo.*, No. 87, 71 pp, doi:10.21957/0nje0wpsa.

Ingleby, N. B., A. C. Lorenc, K. Ngan, F. Rawlins, and D. R. Jackson, 2013: Improved variational analyses using a nonlinear humidity control variable. *Quart. J. Roy. Meteor. Soc.*, **139**, 1875–1887.

Ismail, S., and E. V Browell, 1989: Airborne and spaceborne lidar measurements of water vapor profiles: a sensitivity analysis; erratum. *Appl. Opt.*, **28**, 4981, doi: 10.1364/ao.28.004981.

Janjic, Z. I., 2003: A nonhydrostatic model based on a new approach. *Meteor. Atmos. Phys.*, **82**, 271–285.

Kalthoff, N., B. Adler, A. Wieser, M. Kohler, K. Träumner, J. Handwerker, U. Corsmeier, S. Khodayar, D. Lambert, A. Kopmann, N. Kunka, G. Dick, M. Ramatschi, J. Wickert, and C. Kottmeier, 2013: KITcube - A mobile observation platform for convection studies deployed during HyMeX. *Meteor. Z.*, **22**, 633–647.

Kingsmill, D. E., 1995: Convection initiation associated with a sea-breeze front, a gust front, and their collision. *Mon. Wea. Rev.*, **123**, 2913–2933.

Lange, D., A. Behrendt, and V. Wulfmeyer, 2018: Compact automatic rotational Raman lidar system for continuous day- and nighttime temperature and humidity mapping. *Proceeding of 20th EGU General Assembly, EGU2018*, 4–13 April, 2018, Vienna, Austria, 9114–9114.

Lee, B. D., R. D. Farley, and M. R. Hjelmfelt, 1991: A numerical case study of convection initiation along colliding convergence boundaries in northeast Colorado. *J. Atmos. Sci.*, **48**, 2350–2366.

Lenschow, D. H., V. Wulfmeyer, and C. Senff, 2000: Measuring second- through fourth-order moments in noisy data. *J. Atmos. Oceanic Technol.*, **17**, 1330–1347.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.

Macke, A., Seifert, P. Baars, H. Barthlott, C. Beekmans, C. Behrendt, A. Behrendt, B. Bohn, M. Brueck, J. Bühl, S. Crewell, T. Damian, H. Deneke, S. Düsing, A. Foth, P. DiGirolamo, E. Hammann, R. Heinze, A. Hirsikko, J. Kalisch, N. Kalthoff, S. Kinne, M. Kohler, U. Löhnert, B. L. Madhavan, V. Maurer, S. K. Muppa, J. Schween, I. Serikov, H. Siebert, C. Simmer, F. Späth, S. Steinke, K. Träumner, S. Trömel, B. Wehner, A. Wieser, V. Wulfmeyer, and X. Xie, 2017: The HD (CP)^2 observational prototype experiment (HOPE)–An overview. *Atmos. Chem. Phys.*, **17**, 4887–4914.

Metzendorf, S., 2019: *10 W-Average-Power Single-Frequency Ti:sapphire laser with tuning agility–A breakthrough in high-resolution 3D water-vapor measurement.* Doctoral dissertation, University of Hohenheim of Stuttgart, Germany, 242 pp. [Available at http://opus.uni-hohenheim.de/volltexte/2019/1558/pdf/Metzendorf_2018_Dissertation_Hohenheim.pdf.]

Muppa, S. K., A. Behrendt, F. Späth, V. Wulfmeyer, S. Metzendorf, and A. Riede, 2016: Turbulent humidity fluctuations in the convective boundary layer: Case studies using water vapour differential absorption lidar measurements. *Bound.-Layer Meteor.*, **158**, 43–66.

Muppa, S. K., A. Behrendt, H.-S. Bauer, F. Späth, R. Heinze, C. Moseley, and V. Wulfmeyer, 2018: Characterizing the turbulent higher-order moments in the convective boundary layer: Evaluation of large eddy simulations with high-resolution lidar observations. *Proceeding of 20th EGU General Assembly, EGU2018*, 4–13 April, 2018, Vienna, Austria, 14632–14632.

Murphey, H. V., R. M. Wakimoto, C. Flamant, and D. E. Kingsmill, 2006: Dryline on 19 June 2002 during IHOP. Part I: Airborne Doppler and LEANDRE II analyses of the thin line structure and convection initiation. *Mon. Wea. Rev.*, **134**, 406–430.

Nakanishi, M., and H. Niino, 2006: An improved Mellor-Yamada Level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407.

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912.

Niu, G. Y., Z. L. Yang, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari, and Y. Xia, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.: Atmos.*, **116**, D12109, doi:10.1029/2010JD015139.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Powers, J. G., J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen, D. J. Gochis, R. Ahmadov, S. E. Peckham, G. A. Grell, J. Michalakes, S. Trahan, S. G. Benjamin, C. R. Alexander, G. J. Dimego, W. Wang, C. S. Schwartz, G. S. Romine, Z. Liu, C. Snyder, F. Chen, M. J. Barlage, W. Yu, and M. G. Duda, 2017: The weather research and forecasting model: Overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc.*, **98**, 1717–1737.

Radlach, M., A. Behrendt, and V. Wulfmeyer, 2008: Scanning rotational Raman lidar at 355 nm for the measurement of tropospheric temperature fields. *Atmos. Chem. Phys.*, **8**, 159–169.

Reichardt, J., U. Wandinger, V. Klein, I. Mattis, B. Hilber, and R. Begbie, 2012: RAMSES: German Meteorological Service autonomous Raman lidar for water vapor, temperature, aerosol, and cloud measurements. *Appl. Opt.*, **51**, 8111–8131.

Richard, E., A. Buzzi, and G. Zängl, 2007: Quantitative precipitation forecasting in the Alps: The advances achieved by the Mesoscale Alpine Programme. *Quart. J. Roy. Meteor. Soc.*, **133**, 831–846.

Santanello, J. A., P. A. Dirmeyer, C. R. Ferguson, K. L. Findell, A. B. Tawfik, A. Berg, M. Ek, P. Gentine, B. P. Guillod, C. van Heerwaarden, J. Roundy, and V. Wulfmeyer, 2018: Land-atmosphere interactions: The LoCo perspective. *Bull. Amer. Meteor. Soc.*, **99**, 1253–1272.

Savitzky, A., and M. J. E. Golay, 1964: Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.*, **36**, 1627–1639.

Schotland, R. M., 1966: Some observations of the vertical profile of water vapor by means of a laser optical radar (measurement and theory, including signal to noise and transfer function calculations, of atmospheric water vapor using ruby laser optical radar). *Proceeding of the 4th Symposium on Remote Sensing of Environment*, Rev. Edition, 4864-11-X, Ann Arbor, MI, Willow Run Lab., Inst. Sci. & Technol., Univ. of Michigan, 273–283.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periáñez, and R. Potthast, 2016: Kilometre-scale ensemble data assimilation for the COSMO model

(KENDA). *Quart. J. Roy. Meteor. Soc.*, **142**, 1453–1472.

Schwitalla, T., and V. Wulfmeyer, 2014: Radar data assimilation experiments using the IPM WRF Rapid Update Cycle. *Meteor. Z.*, **23**, 79–102.

Schwitalla, T., H. S. Bauer, V. Wulfmeyer, and F. Aoshima, 2011: High-resolution simulation over Central Europe: Assimilation experiments during COPS IOP 9c. *Quart. J. Roy. Meteor. Soc.*, **137**, 156–175.

Schwitalla, T., H. S. Bauer, V. Wulfmeyer, and K. Warrach-Sagi, 2017: Continuous high-resolution midlatitude-belt simulations for July-August 2013 with WRF. *Geosci. Model Dev.*, **10**, 2031–2055.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2008: *A description of the advanced research WRF Version 3*. NCAR Technical Note, NCAR/TN-475+STR. University Corporation for Atmospheric Research, 113 pp, doi: 10.5065/D68S4MVH.

Späth, F., A. Behrendt, S. K. Muppa, S. Metzendorf, A. Riede, and V. Wulfmeyer, 2014: High-resolution atmospheric water vapor measurements with a scanning differential absorption lidar. *Atmos. Chem. Phys. Discuss.*, **14**, 29057–29099.

Späth, F., A. Behrendt, S. Kumar Muppa, S. Metzendorf, A. Riede, and V. Wulfmeyer, 2016: 3-D water vapor field in the atmospheric boundary layer observed with scanning differential absorption lidar. *Atmos. Meas. Tech.*, **9**, 1701–1720.

Späth, F., S. Metzendorf, S. K. Muppa, A. Behrendt, and V. Wulfmeyer, 2018: Investigation of Humidity Structures from the Land-Surface to the Lower Troposphere Using a Scanning DIAL. *Proceedings of 20th EGU General Assembly, EGU2018*, 4–13 April, 2018, Vienna, Austria, 15198–15198.

Spuler, S. M., K. S. Repasky, B. Morley, D. Moen, M. Hayman, and A. R. Nehrir, 2015: Field-deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor. *Atmos. Meas. Tech.*, **8**, 1073–1087.

Stevens, B., and S. Bony, 2013: What are climate models missing? *Science*, **340**, 1053–1054.

Talbot, C., E. Bou-Zeid, and J. Smith, 2012: Nested mesoscale large-eddy simulations with WRF: Performance in real test cases. *J. Hydrometeor.*, **13**, 1421–1441.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115.

Turner, D. D., and J. E. M. Goldsmith, 1999: 24-Hour Raman lidar measurements during the Atmospheric Radiation Measurement program's 1996 and 1997 water vapor intensive observation periods. *J. Atmos. Oceanic Technol.*, **16**, 1062–1076.

Turner, D. D., and U. Löhnert, 2014: Information content

and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor. Climatol.*, **53**, 752–771.

Wagner, G., V. Wulfmeyer, and A. Behrendt, 2011: Detailed performance modeling of a pulsed high-power single-frequency Ti:sapphire laser. *Appl. Opt.*, **50**, 5921–5937.

Wagner, G., A. Behrendt, V. Wulfmeyer, F. Späth, and M. Schiller, 2013: High-power Ti:sapphire laser at 820 nm for scanning ground-based water–vapor differential absorption lidar. *Appl. Opt.*, **52**, 2454–2469.

Weckwerth, T. M., and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev.*, **134**, 5–22.

Weckwerth, T. M., J. W. Wilson, and R. M. Wakimoto, 1996: Thermodynamic variability within the convective boundary layer due to horizontal convective rolls. *Mon. Wea. Rev.*, **124**, 769–784.

Weckwerth, T. M., V. Wulfmeyer, R. M. Wakimoto, R. M. Hardesty, J. W. Wilson, and R. M. Banta, 1999: NCAR–NOAA lower-tropospheric water vapor workshop. *Bull. Amer. Meteor. Soc.*, **80**, 2339–2357.

Weckwerth, T. M., K. J. Weber, D. D. Turner, and S. M. Spuler, 2016: Validation of a water vapor micropulse differential absorption lidar (DIAL). *J. Atmos. Oceanic Technol.*, **33**, 2353–2372.

Wei, J., O. Branch, H.-S. Bauer, H.-D. Wizemann, and V. Wulfmeyer, 2017: Impact of enhanced hydrological processes on land-atmosphere feedbacks for the Attert catchment in Luxembourg – LES runs with WRF-Hydro-NOAH-MP vs. WRF-NOAH-MP. *Proceedings of 20th EGU General Assembly, EGU2018*, 4–13 April, 2018, Vienna, Austria, 16392–16392.

Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437.

Wu, W. S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional ensemble-variational data assimilation using global ensemble forecasts. *Wea. Forecasting*, **32**, 83–96.

Wulfmeyer, V., and J. Bösenberg, 1998: Ground-based differential absorption lidar for water-vapor profiling: Assessment of accuracy, resolution, and meteorological applications. *Appl. Opt.*, **37**, 3825–3844.

Wulfmeyer, V., H. S. Bauer, M. Grzeschik, A. Behrendt, F. Vandenberghe, E. V Browell, S. Ismail, and R. A. Ferrare, 2006: Four-dimensional variational assimilation of water vapor differential absorption lidar data: The first case study within IHOP_2002. *Mon. Wea. Rev.*, **134**, 209–230.

Wulfmeyer, V., A. Behrendt, C. Kottmeier, U. Corsmeier, C. Barthlott, G. C. Craig, M. Hagen, D. Althausen, F. Aoshima, M. Arpagaus, H. S. Bauer, L. Bennett, A. Blyth, C. Brandau, C. Champollion, S. Crewell, G.

Dick, P. D. Girolamo, M. Dorninger, Y. Dufournet, R. Eigenmann, R. Engelmann, C. Flamant, T. Foken, T. Gorgas, M. Grzeschik, J. Handwerker, C. Hauck, H. Höller, W. Junkermann, N. Kalthoff, C. Kiemle, S. Klink, M. König, L. Krauss, C. N. Long, F. Madonna, S. Mobbs, B. Neininger, S. Pal, G. Peters, G. Pigeon, E. Richard, M. W. Rotach, H. Russchenberg, T. Schwitalla, V. Smith, R. Steinacker, J. Trentmann, D. D. Turner, J. V. Baelen, S. Vogt, H. Volkert, T. Weckwerth, H. Wernli, A. Wieser, and M. Wirth, 2011: The Convective and Orographically-induced Precipitation Study (COPS): The scientific strategy, the field phase, and research highlights. *Quart. J. Roy. Meteor. Soc.*, **137**, 3–30.

Wulfmeyer, V., R. M. Hardesty, D. D. Turner, A. Behrendt, M. P. Cadeddu, P. Di Girolamo, P. Schlüssel, J. Van Baelen, and F. Zus, 2015: A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles. *Rev. Geophys.*, **53**, 819–895.

Wulfmeyer, V., S. K. Muppa, A. Behrendt, E. Hammann, F. Späth, Z. Sorbjan, D. D. Turner, and R. M. Hardesty, 2016: Determination of convective boundary layer en-trainment fluxes, dissipation rates, and the molecular destruction of variances: Theoretical description and a strategy for its confirmation with a novel lidar system synergy. *J. Atmos. Sci.*, **73**, 667–692.

Wulfmeyer, V., D. D. Turner, B. Baker, R. Banta, A. Behrendt, T. Bonin, W. A. Brewer, M. Buban, A. Choukulkar, E. Dumas R., M. Hardesty, T. Heus, J. Ingwersen, D. Lange, T. R. Lee, S. Metzendorf, S. K. Muppa, T. Meyers, R. Newsom, M. Osman, S. Raasch, J. Santanello, C. Senff, F. Späth, T. Wagner, and T. Weckwerth, 2018: A new research approach for observing and characterizing land–atmosphere feedback. *Bull. Amer. Meteor. Soc.*, **99**, 1639–1667.

Yang, Z. L., G.-Y. Niu, K. E. Mitchell, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, L. Longuevergne, K. Manning, D. Niyogi, M. Tewari, and Y. Xia, 2011: The community Noah land surface model with mul-tiparameterization options (Noah-MP): 2. Evaluation over global river basins. *J. Geophys. Res.: Atmos.*, **116**, D12110, doi:10.1029/2010JD015140.

Yoshida, S., S. Yokota, H. Seko, T. Sakai, and T. Nagai, 2020: Observation system simulation experiments of water vapor profiles observed by Raman lidar using LETKF system. *SOLA*, **16**, 43–50.

Chapter 3

Impact of Assimilating Lidar Water Vapour and Temperature Profiles with a Hybrid Ensemble Transform Kalman Filter – Three-Dimensional Variational Analysis on the Convection Permitting Scale

3.1   Overview

The atmosphere is not static but chaotic and dynamic. This dynamic nature calls for a DA technique that confirms the same. Incorporating a small perturbation or change in the initial condition of a deterministic forecast creates a considerable divergence in its future predicted state, defining a specific limit to deterministic predictability. Here comes the role of an ensemble forecast where a set of initial conditions is perturbed differently from each other. A spread is obtained through the forecasts from different ensemble members. A forecast can be determined, assigning a measure of confidence. Hence, instead of a single deterministic forecast, we are left with an ensemble of forecasts from which we can deduce

the uncertainty of the model prognostic variables. The uncertainty thereby can be used to determine the estimate of the current atmospheric state.

The previous chapter explained the assimilation of lidar temperature and water vapour mixing ratio profiles with the help of a deterministic **3DVar** DA system. It also introduced a new thermodynamic lidar forward operator to accomplish the direct assimilation of WVMR. Although an overall improvement in the analyses was observed, the study encountered an issue. The simultaneous assimilation of moisture and temperature showed less marked improvement than the independent assimilation of either of the parameters. As discussed in chapter 2, the problem was identified as the influence of the B matrix. This chapter introduces an ensemble approach to the thermodynamic lidar DA using a Hybrid **3DVar-ETKF** DA system to add flow-dependent error covariance on top of the static B matrix. The added flow-dependency in the background error covariance aims to capture better the thermodynamic flow in the atmosphere than the static background error covariance used in the **3DVar**.

Following the study in chapter 2, this study aims to use the high spatio-temporal resolution of lidar observations and investigate the extent to which the assimilation of these high-resolution data through advanced ensemble-based DA systems improves the analyses and corresponding forecasts. The following research questions are addressed in this chapter:

1. Does the addition of flow-dependency in the B matrix improve the analyses?

2. What is the spatial extent of the thermodynamic lidar profiles assimilation impact from a single geographical location using an ensemble approach versus a purely deterministic approach?

3. Up to what time range is a forecast from a single ground-based lidar system treated valid with respect to observations?

4. Will these results pave the way for future networks of lidar systems in operational forecasting?

The experiments presented in this study were conducted with the same convection-permitting resolution, likewise in (Thundathil et al., 2020). The assimilation experiment was also conducted with the same hourly rapid-update cycle approach. Ten ensemble members were generated by random perturbations of the initial condition for the whole study. In order to answer the research questions above, four experiments were conducted:

1. **NOLIDAR_3DVAR**   **3DVar** Conventional data assimilation.

2. **LIDAR_3DVAR**   Exp. 1 + WVMR and temperature.

3. **NOLIDAR_HYBRID**   HYBRID Conventional data assimilation.

4. **LIDAR_HYBRID**   Exp. 3 + WVMR and temperature.

Quantification of the results confirms the improvement of the LIDAR_HYBRID experiment over the LIDAR_3DVAR experiment. The hybrid experiment resulted in a 50% lower temperature and water vapour RMSE than the **3DVar** experiment when taking the assimilated lidar data as a reference. Comparing against independent radiosonde observations collocated with the lidar site, RMSE reduced 26% for water vapour and reduced 38% for temperature (Thundathil et al., 2021). Compared to independent ceilometer data, the planetary boundary layer height of the analyses provided additional evidence of improvement. **3DVar** analysis RMSE showed 140 m, whereas hybrid showed 60 m. Although limited to a single case study, we attribute these improvements to the flow-dependent background error covariance matrix in the hybrid **3DVar-ETKF** approach. The analysis increments of a single lidar vertical profile established a spatial impact spread over a radius of 100 km. The forecast impact of Planetary Boundary Layer Height (PBLH) persisted for about 7 hours, and the Precipitable Water Vapor (PWV) impact persisted for about a roughly 4-hour forecast lead time. The forecast validation is conducted to compare the ceilometer data for PBLH and Global Positioning System observations for PWV. In the future, the assimilation of water vapour and temperature data from an operational lidar network will benefit the weather prediction community. The next chapter will give insight into an application to test the improvement of forecast skills using several lidar systems operated simultaneously.

3.2   Publication

The publication titled "Impact of Assimilating Lidar Water Vapour and Temperature Profiles with a Hybrid Ensemble Transform Kalman Filter – Three-Dimensional Variational Analysis on the Convection Permitting Scale" is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

RESEARCH ARTICLE

Impact of assimilating lidar water vapour and temperature profiles with a hybrid ensemble transform Kalman filter: Three-dimensional variational analysis on the convection-permitting scale

Rohith Thundathil | **Thomas Schwitalla** | **Andreas Behrendt** | **Volker Wulfmeyer**

Institute of Physics and Meteorology, University of Hohenheim, Stuttgart, Germany

Correspondence
R. Thundathil, Institute of Physics and Meteorology, University of Hohenheim, Garbenstrasse 30, Stuttgart 70599, Germany.
Email:
rohith.thundathil@uni-hohenheim.de

Abstract

We discuss the analysis impact of the ensemble-based assimilation of differential absorption lidar observed water vapour and Raman lidar observed temperature profiles into the Weather Research and Forecasting model at convection-permitting scale. The impact of flow-dependent background error covariance in the data assimilation (DA) system that uses the hybrid three-dimensional variational (3DVAR) ensemble transform Kalman filter (ETKF) was compared to 3DVAR DA. The 3DVAR-ETKF experiment resulted in a 50% lower temperature and water vapour RMSE than the 3DVAR experiment when taking the assimilated lidar data as reference and 26% (38%) lower water vapour (temperature) RMSE when comparing against independent radiosonde observations collocated with the lidar site. The planetary boundary-layer height of the analyses compared to independent ceilometer data provided additional evidence of improvement. The 3DVAR analysis RMSE showed 140 m, whereas 3DVAR-ETKF showed 60 m. Although limited to a single case study, we attribute these improvements to the flow-dependent background error covariance matrix in the 3DVAR-ETKF approach. The vertical profile measured from a single stationary lidar system established a spatial impact with a 100 km radius. This seems to indicate future assimilation of water vapour and temperature data from an operational lidar network. The assimilation impact persisted 7 hr into the forecast time compared with the ceilometer data and 4 hr with GPS observations.

KEYWORDS

data assimilation, ensemble, lidar, numerical weather prediction, temperature, water vapour

1 | INTRODUCTION

The lack of quality observations limits the accuracy of the initial conditions used for numerical weather prediction (NWP) models resulting in suboptimal forecasts. Enhanced data assimilation (DA) techniques, with suitable observational forward operators, enable the model to absorb quality observations efficiently. The assimilation of temperature and moisture profiles obtained from active remote-sensing lidar systems (e.g. Wulfmeyer *et al.*, 2015) offers great potential for improving the predictive skills of NWP models (Wulfmeyer *et al.*, 2006; 2015; Adam *et al.*, 2016; Leuenberger *et al.*, 2020; Thundathil *et al.*, 2020).

Although passive remote-sensing systems, like microwave radiometers and infrared (IR) spectrometers, provide some vertical information on the atmosphere, their ability to resolve vertical structures is quite limited (Wulfmeyer *et al.*, 2015). Microwave radiometers have range resolutions of 300 to 1,000 m in the lowest 2,000 m above ground (Cadeddu *et al.*, 2002; Blumberg *et al.*, 2015; Wulfmeyer *et al.*, 2015). IR spectrometers have a slightly higher vertical resolution than microwave radiometers but are still limited to several hundreds of metres close to the land surface (Turner and Löhnert, 2014).

In contrast to passive remote-sensing techniques, the data of active remote-sensing lidar instruments possess substantially higher temporal and vertical resolutions. The latter is due to the range-resolved measurements by the run-time of the signals and the laser pulse duration combined with a high signal-to-noise ratio (SNR) of the backscatter signals. Temperature profiles are provided by temperature rotational Raman lidar (TRL: e.g. Behrendt and Reichardt, 2000; Di Girolamo *et al.*, 2004; Arshinov *et al.*, 2005; Radlach *et al.*, 2008), moisture profiles by water vapour Raman lidar (WVRL: e.g. Veselovskii *et al.*, 2009; Leblanc *et al.*, 2012; Sakai *et al.*, 2019) or water vapour differential absorption lidar (WVDIAL: e.g. Ismail and Browell, 1989; Bösenberg, 1998; Behrendt *et al.*, 2009; Wagner *et al.*, 2011; Repasky *et al.*, 2013; Wagner *et al.*, 2013; Spuler *et al.*, 2015; Späth *et al.*, 2016; Weckwerth *et al.*, 2016; Metzendorf, 2018). Several Raman lidar systems (e.g. Behrendt *et al.*, 2002; Di Girolamo *et al.*, 2004; 2006; 2009; 2016; 2017; Bhawar *et al.*, 2011; Lange *et al.*, 2019) independently provide measurements of both water vapour mixing ratio and temperature. A few such combined water vapour and temperature Raman lidar systems even operate autonomously in an uninterrupted manner (e.g. Goldsmith *et al.*, 1998; Turner and Goldsmith, 1999; Reichardt *et al.*, 2012, Dinoev *et al.*, 2013; Newsom *et al.*, 2013; Lange *et al.*, 2019). High range-resolution eliminates data gaps in the planetary boundary layer (PBL) and the lower troposphere. Thus, it is valuable for resolving the gradients at the top of the PBL and lids in the free troposphere, which influences atmospheric stability, cloud formation, and transport processes.

Advanced lidar systems provide observations with high accuracy and temporal resolution, which is highly beneficial for DA. The high accuracy avoids the need for a challenging bias correction of the data. It also simplifies operational use and minimizes the latency of the lidar data available for DA. The temporal resolution is up to a few seconds, which even permits investigation of statistics of turbulent fluctuations (Wulfmeyer, 1999a; 1999b; Kiemle *et al.*, 2007; Wulfmeyer *et al.*, 2010; Turner *et al.*, 2014a; 2014b; Behrendt *et al.*, 2015; Muppa *et al.*, 2016), covariances and fluxes (Behrendt *et al.*, 2020), and the subgrid-scale variability of water vapour mixing ratio and temperature. This makes it possible to estimate representativeness errors in the region of interest.

The 3DVAR (Courtier, 1998; Barker *et al.*, 2004) and 4DVAR (Huang *et al.*, 2009) DA systems use a static background error covariance matrix, usually derived according to the National Meteorological Centre (NMC) method (Parrish and Derber, 1992). Although 4DVAR relies on the evolution of the initial errors under tangent linear dynamics (Lorenc, 2003), each 4DVAR cycle initializes with the same static error covariance. Hence, the deterministic variational approaches have been enhanced using ensemble approaches in various forecast centres. Incorporating flow dependency into the DA system by an ensemble approach is superior to purely deterministic DA methods (Hamill and Snyder, 2000; Buehner, 2005; Wang *et al.*, 2008a; Wang, 2011; Kutty and Wang, 2015; Kutty *et al.*, 2018). With the use of ensemble DA techniques (Houtekamer and Mitchell, 1998; Bishop *et al.*, 2001; Evensen, 2003), a flow-dependent forecast error covariance matrix is estimated from the ensemble. In the hybrid technique, the flow-dependent error covariance is combined with a static error covariance matrix.

The first hybrid ensemble 3DVAR DA system was proposed by Hamill and Snyder (2000). Later, Lorenc (2003) devised an efficient method for introducing flow-dependent error covariances from ensemble perturbations, making it easier to adapt to the variational cost function framework. In order to update the ensemble perturbations, an ensemble-based DA system such as ensemble square-root filter (EnSRF: Whitaker and Hamill, 2002) or ensemble transform Kalman filter (ETKF: Bishop *et al.*, 2001) can be used. Several studies (Wang *et al.*, 2008b; Wang, 2011; Gao *et al.*, 2013; Schwartz *et al.*, 2013; Wang *et al.*, 2013; Pan *et al.*, 2014; Schwartz and Liu, 2014; Schwartz *et al.*, 2015; Shen *et al.*, 2016; 2017) have shown that using a hybrid 3DVAR-ETKF approach improves performance compared to a pure 3DVAR DA approach.

Recently, ensemble DA studies of thermodynamic and wind profiler data have been conducted (Chipilski *et al.*, 2020; Cui *et al.*, 2020). Chipilski *et al.* (2020) assimilated Atmospheric Emitted Radiance Interferometer (AERI) thermodynamic retrievals and wind profiler data at a high temporal frequency to improve bore-driven convection resulting in considerable forecast improvements. Cui *et al.* (2020) used high spatio-temporal resolution airborne Doppler aerosol wind lidar profiles to create the initial conditions for a short-range forecast that showed improvements in the wind and precipitation forecasts. DA impact studies using either airborne or ground-based lidar systems produced remarkable results in several case studies. WVDIAL data were assimilated into a global atmospheric model in Kamineni *et al.* (2003; 2006) using a 3DVAR. The results showed forecast improvements in hurricane tracks. Wulfmeyer *et al.* (2006) assimilated airborne water vapour DIAL data of the NASA Lidar Atmospheric Sensing Experiment system (LASE) into the fifth-generation Pennsylvania State University–National Centre for Atmospheric Research Mesoscale Model (MM5) by applying a 4DVAR DA. The study showed an improved prediction of convection initiation (CI) concerning water vapour and wind. During the Convective and Orographically-induced Precipitation Study (COPS: Wulfmeyer *et al.*, 2011), lower-tropospheric water vapour fields measured by two airborne lidars were assimilated into the 3DVAR assimilation system of the Application of Research to Operations at MEsoscale (AROME) NWP mesoscale model (Bielli *et al.*, 2012). Grzeschik *et al.* (2008) assimilated water vapour data from three ground-based WVRLs, aligned in a triangle, into the MM5 model from the Lindenberg Campaign for Assessment of Humidity and Cloud Profiling Systems and its Impact on High-Resolution Modelling (LAUNCH: Engelbart and Haas, 2006). The assimilation succeeded in correcting the initial water vapour field by about 1 g·kg^{-1} and extended the impact on the analysis water vapour field up to 12 hr in the forecast. In addition, TRL data were assimilated into the WRF model using the 3DVAR DA system in Adam *et al.* (2016). It showed temperature analysis improvements with corrected PBL heights.

Observation System Simulation Experiments (OSSEs) conducted by Otkin *et al.* (2011) assimilated data from various ground-based remote-sensing boundary-layer profiling instruments, namely, Doppler wind lidar, Raman lidar, microwave radiometer, and AERI. Otkin *et al.* (2011) showed an overall improvement in analysis illustrating that both mass and momentum observations are necessary to improve analysis accuracy. Later in Hartung *et al.* (2012), short-range ensemble forecasts were initiated from the analyses of Otkin *et al.* (2011). They contributed accurate forecasts of moisture flux convergence and the intensity and location of accumulated precipitation. Recently, some OSSEs were conducted assimilating water vapour profiles as observed from Raman lidar into the Japan Meteorological Agency Non-Hydrostatic Model (JMA-NHM: Saito *et al.*, 2006) through the LETKF DA system by Yoshida *et al.* (2020). The study showed an improvement in precipitation forecast by assimilating water vapour profiles from Raman lidar on the windward side of heavy precipitation.

In a previous study (Thundathil *et al.* (2020), hereafter referred to as T20), temperature and water vapour mixing ratio profiles from lidar systems were simultaneously assimilated. A new thermodynamic lidar (TDLIDAR) forward operator for the direct assimilation of the water vapour mixing ratio profile rather than relative humidity was introduced and implemented. Several data-denial DA experiments with temperature, WVMR and relative humidity were conducted with the 3DVAR DA system. However, the simultaneous assimilation of moisture and temperature showed less marked improvement than the independent assimilation of either of the parameters.

This study aims to make use of the high spatio-temporal resolution of lidar observations, and to investigate the extent to which the assimilation of these high-resolution data through advanced DA systems improves the analyses and corresponding forecasts. Following up on T20, we present the impact of high-resolution temperature and water vapour mixing ratio profiles from TRL and WVDIAL by adopting the WRF hybrid 3DVAR-ETKF approach (Wang *et al.*, 2008a; 2008b) to address the following questions:

1. Does the addition of a flow-dependent error in the background error covariance matrix improve the analyses?
2. What is the spatial extent of the impact of assimilating thermodynamic lidar profiles from a single geographical location using an ensemble approach versus a purely deterministic approach, and will it pave the way for future networks of lidar systems in operational forecasting?

The article is structured as follows. Section 2 describes the hybrid DA methodology. Section 3 gives an overview of the methodology for generating ensembles, the experimental set-up and the assimilated observations. Section 4 presents our results. The article concludes with a summary of our findings and gives an outlook on future research.

2 | HYBRID 3DVAR-ETKF DA SYSTEM

The Hybrid 3DVAR-ETKF (referred to as hybrid hereafter) DA technique (Wang *et al.*, 2008a; 2008b) was used to

assimilate thermodynamic lidar data in this study. The results were compared with those of the 3DVAR DA system. The main advantage of a hybrid DA system over 3DVAR is that it uses flow-dependent ensemble covariance rather than static background error covariance. Compared to EnKF (Evensen, 2003), a smaller number of ensemble members is required (Wang *et al.*, 2007; Zhang *et al.*, 2013; Pan *et al.*, 2014) in hybrid DA because, in addition to ensemble covariance, static error covariance is also incorporated. Through the extended control variable approach (Lorenc, 2003), the pure 3DVAR cost function is adapted to include the ensemble covariance. In the following, we briefly describe the hybrid 3DVAR cost function as it is applied in the WRFDA framework.

According to Wang *et al.* (2008b), the hybrid 3DVAR cost function is defined as follows:

$$J\left(\vec{x}_1, \vec{a}\right) = \beta_1 J_1 + \beta_2 J_e + J_o, \tag{1}$$

where $J\left(\vec{x}_1, \vec{a}\right)$ is the hybrid cost function expressed as the sum of the observational component (J_o) and the background components J_1 (static component) and J_e (ensemble component) extended by the control variables and applied with appropriate weighting factors β_1 and β_2 with:

$$\frac{1}{\beta_1} + \frac{1}{\beta_2} = 1. \tag{2}$$

The extended control variables in the WRF hybrid system are the three-dimensional wind components U, V, W, the perturbation geopotential variable PH, the perturbation dry mass in column MU, temperature T, and water vapour mixing ratio $QVAPOUR$.

The cost function in Equation (1) then expands to

$$J\left(\vec{x}_1, \vec{a}\right) = \frac{1}{2}\beta_1\left(\vec{x}_1\right)^T B^{-1}\left(\vec{x}_1\right) + \frac{1}{2}\beta_2 \sum_{k=1}^{N}\left(\vec{a}_k\right)^T A^{-1}\left(\vec{a}_k\right)$$
$$+ \frac{1}{2}\left(\vec{y^o} - H\vec{x}\right)^T R^{-1}\left(\vec{y^o} - H\vec{x}\right), \tag{3}$$

where $\vec{x}$ is the total analysis increment vector, $\vec{x}_1$ is the analysis increment due to the static background error covariance and $\vec{a}$ is the concatenated vector of extended control variable for each ensemble member. $\vec{y^o}$ is the innovation vector where $\vec{y^o} = \vec{y} - H\left(\vec{x}_b\right)$. $\vec{y}$ denotes the observations assimilated and $\vec{x}_b$ denotes the background state vector. H is the nonlinear observation forward operator which maps the model grid points to the observations. B and R are static error covariance matrices of the background and observations, respectively. B is assumed to start with a diagonal matrix in the control variable space whereas the off-diagonal terms, the covariances, exist in the model space which are generated with series of transformations. R is a fully diagonal matrix since the observations are

deemed to be mutually independent. A is the localization matrix that contributes to flow-dependent error covariance.

The ensemble transformation for the cycling of the DA system starts with the computation of each ensemble member's perturbation from the ensemble mean. Therefore, the ensemble mean and the corresponding variance of each member are calculated. The ensemble perturbations are updated by the ETKF (Bishop *et al.*, 2001) with a transformation matrix T:

$$T = rE(\rho\lambda + I)^{-1/2}E^T. \tag{4}$$

E and λ are the eigenvectors and eigenvalues of the $N \times N$ matrix $\frac{(X^e)^T H^T R^{-1} H X^e}{N-1}$.

X^e is the ensemble perturbation matrix, and r (inflation factor) and ρ are scalar tunable factors. ρ accounts for the fraction of the forecast-error covariance projected onto the ensemble subspace. To derive the analysis perturbation matrix X^a, we multiply the ensemble perturbation matrix X^e with T which then reads:

$$X^a = X^e T. \tag{5}$$

The columns of the analysis perturbation matrix constitute the new set of updated ensemble perturbations. The ensemble mean is applied as the background state $\vec{x}_b$ in the hybrid 3DVAR cost function Equation (3) to get an updated ensemble mean. The updated perturbations from the ETKF are added to the updated mean to obtain the set of new analysis ensemble members. With the new set of ensemble members, a free forecast is performed till the next assimilation time step (Figure 1).

3 | METHODS AND DATA

3.1 | Model and DA configuration

The WRF model (Skamarock *et al.*, 2008), version 3.8.1, was used in this study. The experiments presented in this study were conducted with the same spatial and vertical resolutions of 2.5 km and 100 levels, as in T20. The vertical grid was configured to include 27 levels within the lowest 1,500 m above ground. The model domain (Figure 2) covers central Europe with 856×832 grid cells. The European Centre for Medium-range Weather Forecasting (ECMWF) analysis with a spatial resolution of 0.125° (approx. 13.5 km) was used to initialize the model and generate lateral boundary conditions. The sea-surface temperatures (SST) were initialized using the Operational Sea surface Temperature and sea Ice Analysis (OSTIA: Donlon *et al.*, 2012) data provided by the UK Met Office.
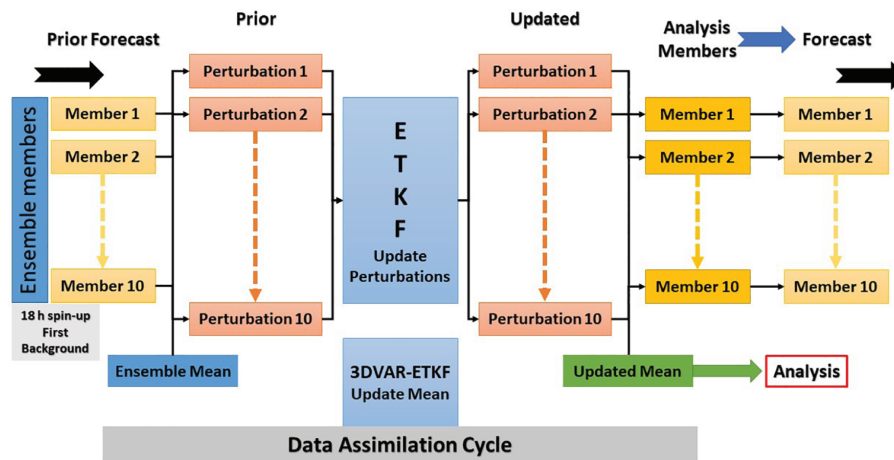
FIGURE 1 Schematic diagram of the hybrid 3DVAR-ETKF data assimilation cycle. The 10 forecast ensemble members from the first guess or from the previous analysis forecast cycle are the input for the subsequent cycle. Prior perturbations are updated by the ETKF to analysis perturbations. The analysis perturbations are added to the updated mean to get a new set of analysis members for the next DA cycle. The first background is from 18 hr spin-up ensemble members [Colour figure can be viewed at wileyonlinelibrary.com]
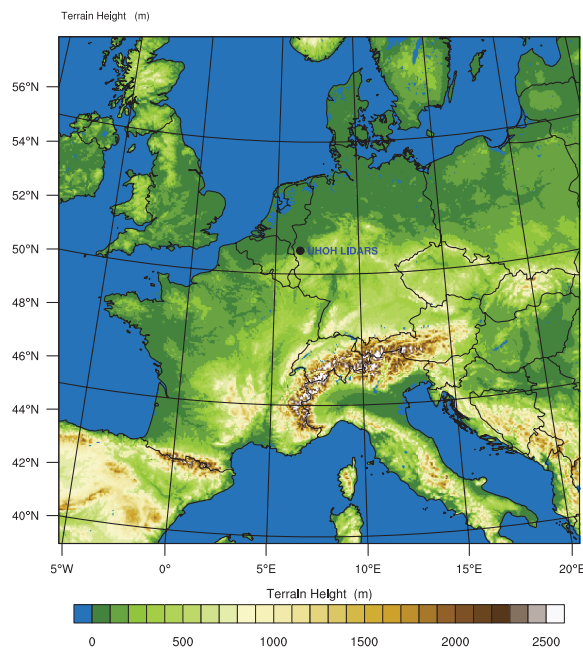

FIGURE 2 The WRF model domain at a horizontal resolution of 2.5 km with orography. The location of the TRL and WVDIAL of UHOH during the HOPE campaign is marked with a black dot [Colour figure can be viewed at wileyonlinelibrary.com]

Since we are interested in the impact of lidar profiles using different DA approaches, the WRF model physics configuration was the same as in T20. It was also evaluated in several previous studies (Schwitalla *et al.*, 2011; Schwitalla and Wulfmeyer, 2014; Bauer

et al., 2015; Adam *et al.*, 2016). The selected physics options comprised the WRF-NOAH multiparametrization (NOAH-MP) Land Surface Model (LSM: Niu *et al.*, 2011; Yang *et al.*, 2011), the Rapid Radiative Transfer Model General Circulation Model (RRTMG) scheme (Iacono *et al.*, 2008) for the parametrization of short-wave and long-wave radiation, the Thompson double moment cloud microphysics scheme (Thompson *et al.*, 2008) and the Mellor–Yamada–Nakanishi–Niino (MYNN: Nakanishi and Niino, 2006) Level 2.5 PBL scheme. Since the model was run on the convection-permitting scale, a deep cumulus parametrization scheme was not applied (Weisman *et al.*, 2008). The development of shallow cumulus was represented by the Global/Regional Integrated Modelling System (GRIMS) scheme (Hong *et al.*, 2013).

The background covariance matrix was created with the CV6 option of WRFDA. The NMC method's statistics (Parrish and Derber, 1992) were derived over April 2013 from forecast differences of 24 and 12 hr. The CV6 option is a multivariate background error covariance where moisture analysis is multivariate. The moisture control variable in CV6 is the unbalanced portion of the pseudo-relative humidity. A change in moisture in the CV6 can lead to a temperature and wind change and vice versa. All five of the correlation length-scales of the CV6 **B** matrix, namely, stream function, unbalanced velocity potential, unbalanced temperature, pseudo-relative humidity and unbalanced surface pressure, were assigned a value of 1. The same **B** matrix was used in the hybrid DA as well.

The ensemble members were generated by random perturbations of the initial condition. This ensemble generation method was already built in the WRFDA
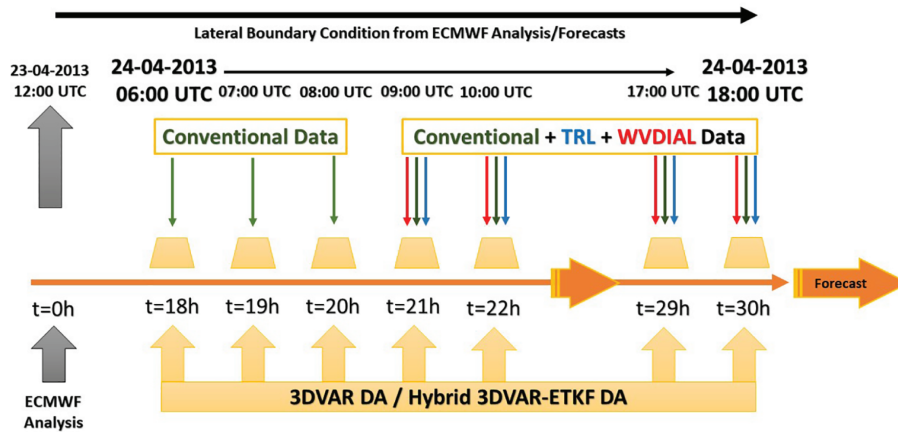
FIGURE 3   Schematic diagram of the 3DVAR/HYBRID 3DVAR-ETKF rapid update cycle initialized by the ECMWF analysis. The spin-up period lasted 18 hr until 0600 UTC on 24 April 2013. Four experiments with different set-ups were performed: NOLIDAR_3DVAR, assimilating conventional data from 0600 to 1800 UTC with 3DVAR DA, and LIDAR_3DVAR, the assimilation of WVDIAL, TRL and conventional data from 0900 to 1800 UTC using the TDLIDAR operator with 3DVAR DA. NOLIDAR_HYBRID and LIDAR_HYBRID are performed in the same way but with HYBRID 3DVAR-ETKF DA. A forecast for 6 hr is performed from the final analyses of all experiments [Colour figure can be viewed at wileyonlinelibrary.com]

framework, where a random perturbation in the model space can be performed (put_rand_seed = true). We tried to maintain the reproducibility of ensemble members by adding fixed random perturbations through the RANDOM CV method predefined in the WRFDA. Ten ensemble members were generated, assigning different integers from 1 to 10 to the "seed_array1" variable. The background covariance matrix was created with the CV6 option of WRFDA.

3.2 | Experimental set-up

The assimilation experiment was carried out with a rapid-update cycle (RUC) approach, as shown in Figure 3. The assimilation of the water vapour mixing ratio (WVMR) and temperature commenced from the fourth assimilation time step at 0900 UTC. It continued until 1800 UTC as lidar observations were only available during this period. We made a preconditioning hourly DA cycle assimilating only conventional data between 0600 to 0800 UTC to analyse the exclusive impact of the lidar data. We focused on four experiments: (a) NOLIDAR_3DVAR, with only the conventional data assimilated using 3DVAR, (b) LIDAR_3DVAR with WVMR and temperature lidar data assimilated through the new thermodynamic lidar operator (TDLIDAR: T20) along with the conventional data, (c) NOLIDAR_HYBRID similar to (a) but with hybrid 3DVAR-ETKF, and (d) LIDAR_HYBRID similar to (b) but with a hybrid 3DVAR-ETKF approach.

Ten ensemble members were generated with 18 hr of spin-up time before the first assimilation time step at 0600 UTC 24 April 2013. Due to computational constraints, we restricted the number of ensemble members to 10 in this study. The hybrid DA system is well-known for its performance at smaller ensemble sizes compared to a pure EnKF DA system (Wang *et al.*, 2007; Zhang *et al.*, 2013; Pan *et al.*, 2014). The weightage factors, β_1 and β_2, were set to 0.5 which means that the static component from the background error covariance matrix is attributed equal importance. The experiments were performed with a hybrid covariance localization scale of 60 km. During the HOPE intensive observation period (IOP), the horizontal wind speeds were 5 m·s^{-1} near the surface (Adam *et al.*, 2016). Since the assimilated lidar data were averaged over a 20 min time window, the distance travelled by an air parcel was approximately 6 km. A spatial scale of 10 times the surface distance was used since we had 10 assimilation cycles. A Gaussian correlation function was used in the present study. We used adaptive inflation scheme WG07 in the WRFDA (Wang *et al.*, 2007), with both tunable factors r and ρ set to 1, to incorporate sufficient spread in the successive assimilation cycles of the hourly RUC.

3.3 | Observations

The observations used for assimilation in this study are the same as those used in T20: SYNOP, METAR, radiosondes (TEMP), wind profilers, aircraft measurements and

atmospheric motion vectors (AMVs). In order to improve the moisture distribution, Global Positioning System (GPS) zenith total delay (ZTD) observations were used in addition. Lidar profiles from the WVDIAL (Späth *et al.*, 2016) and the TRL (Hammann *et al.*, 2015) of the University of Hohenheim (UHOH) obtained during the High Definition Clouds and Precipitation for advancing Climate Prediction (HD[CP]^2) Observation Prototype Experiment (HOPE) campaign (Macke *et al.*, 2017), were used for the impact studies. The IOP on 24 April 2013 was selected for the case-study. The lidar systems were fixed at the Hambach site in western Germany (Figure 2). The meteorological conditions on that day were clear sky with a few cirrus clouds and a warm moist advection from the southwest towards the lidar location with average wind speeds around 5 m·s^{-1} during the IOP (Adam *et al.*, 2016). Four radiosondes were launched at 0900, 1100, 1300 and 1500 UTC, close to the location where the lidar system was located. They were used for evaluation only.

Time windows of ±10 min around each assimilation time step in the lidar time-series data were chosen in order to assimilate water vapour and temperature observations. The WVDIAL absolute humidity data and the corresponding error, derived from the number density, were converted to WVMR data with associated errors. The original WVMR profile data with 15 m vertical resolution was assimilated as 30 m gliding averages to comply roughly with the model vertical resolution. The total error of the lidar profiles within the 20 min time window was computed as the square root of the sum of squares of the representativity error and the observation error. The WVDIAL total errors ranging from 0.01 g·kg^{-1} at heights of 400 m to a maximum of 1 g·kg^{-1} at heights above 2 km (Späth *et al.*, 2016) were used in the DA system. Similarly, the TRL profiles were also averaged over a time window of ±10 min around each assimilation time step. The TRL data were fed to the assimilation system with a vertical resolution of 37.5 m along with the errors. The TRL total errors ranging from 0.1 K at 500 m to 1.1 K at 3,000 m (Hammann *et al.*, 2015) were input into the DA system. For further information on the lidar measurements and errors, please refer to T20.

4 | RESULTS

In order to understand how the flow-dependent error covariances influence the background error covariance matrix, we first present results from single observation tests (SOTs) conducted with the hybrid approach. Then we compare the lidar profiles, which are assimilated into the model, with independent radiosonde measurements to understand their strength and limitations to be accounted for when analysing the results. Most of this

section will deal with the quantitative impact on WVMR and temperature analyses of the hybrid, LIDAR_HYBRID and NOLIDAR_HYBRID, compared with the 3DVAR LIDAR_3DVAR and NOLIDAR_3DVAR experiments. The analyses are compared to the independent radiosondes launched four times throughout the IOP (see the previous section). The RMSEs of the analyses profiles against the radiosonde measurements were calculated for each time step. We also present a comparison with independent planetary boundary-layer height (PBLH) observations from the ceilometers present during this HOPE IOP. Lidar DA analyses were compared to the assimilated lidar profiles to understand how well the model captured the observations. Finally, the predicted PBLH and precipitable water vapour (PWV) fields will be compared with ceilometer and GPS-PWV observations, respectively. The GPS ZTD data were already assimilated into the model during all the assimilation time steps. Hence the improvement in the PWV forecast was minimal compared to the independent ceilometers.

4.1 | Single observation tests

When assimilated with the lidar temperature and moisture data using a deterministic 3DVAR, the temperature analyses profiles had certain disadvantages (T20). Although the assimilation of temperature and WVMR with the TDLIDAR operator resulted in considerable improvement in the WVMR analysis, the temperature improvement was less than expected. The single observation tests conducted separately for temperature and WVMR enabled us to understand the background error covariance statistics of the **B** matrix. A pseudo-WVMR observation with an innovation of 1 g·kg^{-1} and an error of the same magnitude were assimilated in 3DVAR at model level 10 (255 m AGL). The assimilation resulted in an analysis increment of 0.3 g·kg^{-1}. Simultaneously, the temperature response of the analysis at the same height led to an analysis increment of -0.15 K. A pseudo-temperature observation of 1 K innovation value, and an error of 1 K were also assimilated at the same height. The assimilation resulted in a temperature increment of 0.28 K and a corresponding WVMR increment of -0.17 g·kg^{-1}. In 3DVAR DA, the assimilation of WVMR produced an analysis increment in the WVMR analysis, which in contrast, created a directly opposite response to temperature. This is purely a statistical correction based on climatological covariance and not a real-time scenario. The exact negative correlation happened with the assimilation of temperature as well. The spatial extent of the area influenced by the analysis increment is shown in Figure 4. Gao and Gao (2020) have presented impacts of multivariate **B** matrix using the CV6
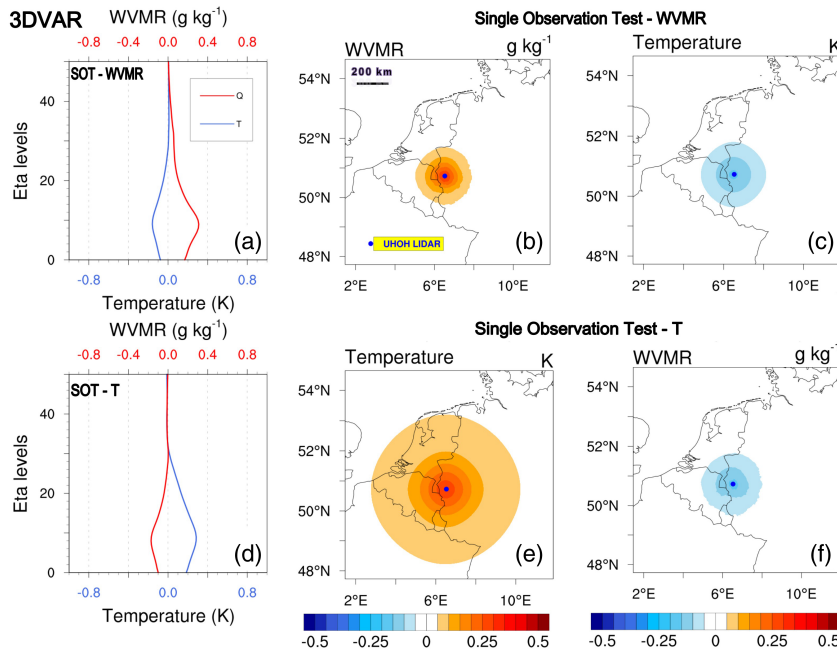
FIGURE 4 Vertical profiles and spatial distribution of analysis increments from the single observation tests (SOTs) performed for a WVMR increment of 1 g·kg⁻¹ and temperature increment of 1 K at model level 10 (255 m AGL) using the 3DVAR approach. (a) The vertical profile, (b,c) the spatial distribution of the analysis increment as a result of the WVMR SOT. (d,e,f) Results of the temperature SOT [Colour figure can be viewed at wileyonlinelibrary.com]

on the 3DVAR DA system of WRF. A dominant-negative correlation is visible in the temperature and WVMR variables. The CV6 is a climatological background error covariance that varies monthly and under the influence of transient weather systems resulting in rapid variations in the lower troposphere. By enhancing the **B** matrix with additional flow-dependent characteristics derived from ensemble members and with a large number of ensemble members, we can increase the sample space and hence bring the response of the error covariance matrix close to reality. We found the response due to the assimilation of WVMR and temperature arising from the **B** matrix statistics to be detrimental to the forward operator. The innate ability of the forward operator to correct the WVMR and temperature variables in the model is inhibited due to feedback response from the **B** matrix, which is not accurate. Real-time lidar observations incorporated in the model are tampered with due to the response from the **B** matrix. The forward operator was designed for the direct assimilation of the prognostic variables, WVMR and temperature, thereby minimizing the error propagation resulting from the relative humidity variable conversions. When assimilating the WVMR into the model, it was so far internally converted to relative humidity. Since relative humidity is sensitive to temperature (T20), undesirable cross sensitivities between moisture and temperature are triggered, leading to more significant errors. The T20 operator enables complete information concerning the water vapour contents in the area of interest to be propagated into the model, reducing error.

We performed the same SOTs with the hybrid DA system in order to understand how the temperature analysis also improved simultaneously with the WVMR (Figure 5). A pseudo-observation of 1 g·kg⁻¹ and 1 K, with an error of the same magnitude, was assimilated at the same level (10, 255 m AGL) for WVMR and temperature, respectively. The WVMR analysis increment was 0.53 g·kg⁻¹, and the temperature increment was 0.82 K, higher than the 3DVAR. The WVMR SOT had a positive temperature response of 0.28 K in the analysis, whereas temperature SOT had a response of 0.09 g·kg⁻¹ in the analysis. The hybrid DA showed positive analysis increments as responses to the SOTs compared to the negative analysis increment responses in 3DVAR. The forward operator is designed to assimilate the lidar WVMR and temperature data directly into the model. Since the SOT response due to assimilation of either WVMR or temperature, in the hybrid approach, is not negatively correlated to each other like in the 3DVAR, temperature analyses are not affected by WVMR assimilation or vice versa. In the hybrid, the temperature analysis response due to WVMR assimilation or vice versa is significantly smaller in magnitude; thereby, one variable does not strongly influence the other. This was a result of the hybrid approach introducing flow-dependent information in the background error covariance. We calculated the correlation coefficient of the temperature and WVMR at the location of the SOTs and covering all the model levels to understand the overall response. The correlation coefficients of the hybrid SOTs were 0.79 and 0.56 for WVMR SOT and temperature SOT, respectively. In

FIGURE 5 Same as Figure 4 but using hybrid [Colour figure can be viewed at wileyonlinelibrary.com]
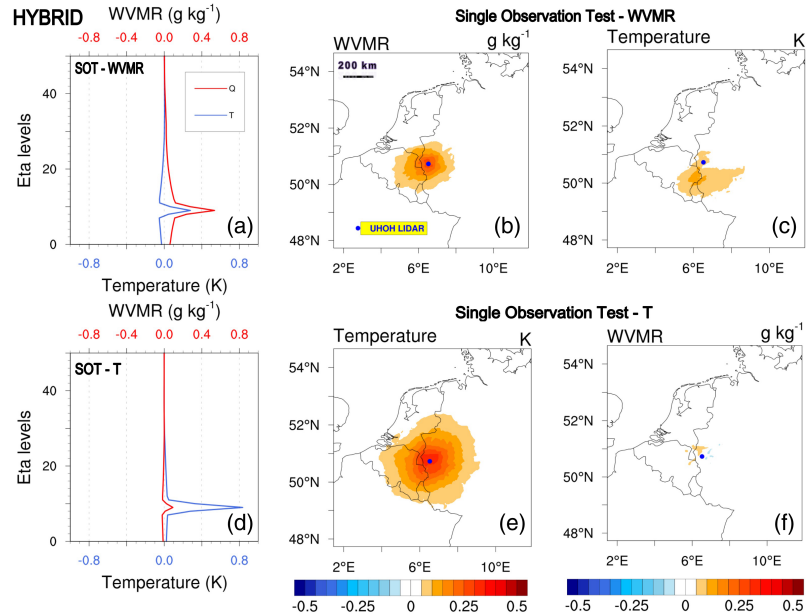
FIGURE 6 Vertical profiles of analysis increments from the SOTs performed for a WVMR increment of 1 g kg^{-1} at model levels 10 (255 m AGL), 20 (760 m AGL), and 40 (3030 m AGL), and the corresponding responses to temperature [Colour figure can be viewed at wileyonlinelibrary.com]
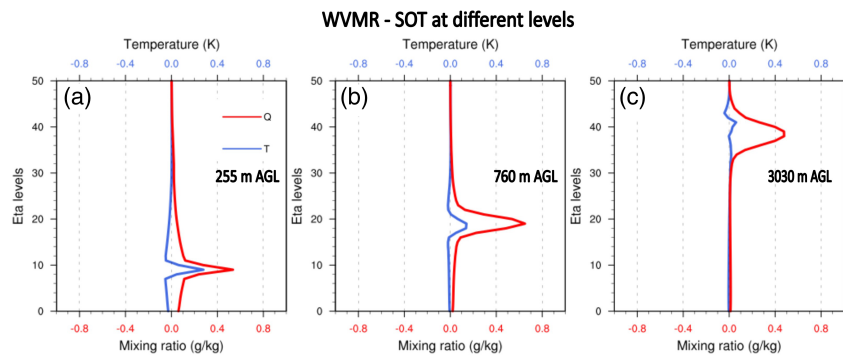
contrast, the 3DVAR SOTs had the same coefficients with -0.97 and -0.98 for WVMR and temperature, respectively.

The spatial extent of the impact of WVMR increment has a diameter of 250 km, similar to that in 3DVAR (refer to T20), with a magnitude ranging from 0.1 g·kg^{-1} at the outer region to 0.53 g·kg^{-1} at the centre. The temperature increment spatially extended to a diameter of 300 km with temperatures ranging from 0.1 to 0.82 K (centre). The spatial extent is the same in the hybrid and 3DVAR but the magnitude of impact at the central region of impact is higher for the hybrid. The hybrid also used the same **B** matrix with the CV6 option used in the 3DVAR with a correlation length-scale for all five parameters set to 1. In order to understand the impact at different levels, we also made two other SOTs for WVMR at levels 20 and 40 (Figure 6), which are at altitudes of 760 and 3,030 m, respectively. The SOTs had a similar positive temperature response trend at higher model levels, too, with a lower

magnitude than the surface levels. The SOTs show how the behaviour of the temperature response with the assimilation of WVMR weakens as we move to higher model levels. The change in response linked to height may help pre-processing profile measurements before assimilation.

4.2 | Comparison of lidar profiles with radiosonde profiles

The radiosondes were launched at the Hambach site in Germany close to the UHOH's Institute of Physics and Meteorology (IPM) lidar systems. The type of radiosonde used during the IOP was the DFM-09 model from GRAW.[1] There were four launches during our case-study, at 0900, 1100, 1300 and 1500 UTC. According to the assessment of

---

[1]https://www.graw.de/products/radiosondes/dfm-09/
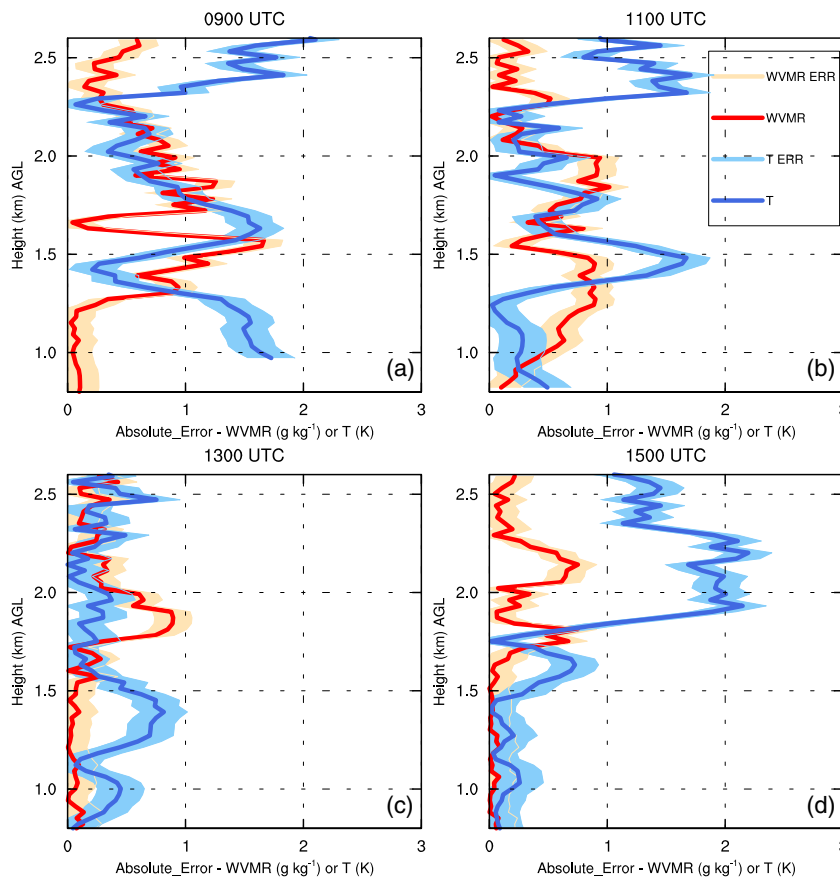
radiosondes in Ingleby (2017), GRAW radiosondes have a significant temperature bias that had to be accounted for. The manufacturer of the DFM-09 model confirms a statistical error of around 0.2 K for temperature and 4% for relative humidity. Another factor that contributes to the uncertainty of radiosondes is "drift" during their ascent. During the IOP, a surface wind speed of around 5 m·s^{-1} was recorded (Adam *et al.*, 2016). The radiosonde took about 5 min to cover the PBL and almost 11 min to reach 3,000 m. This provides some evidence about the track the radiosonde would have followed during its ascent. The radiosonde may have moved slightly away from the launch location assuming that the wind speed above the surface was not less than the surface wind speed. On the other hand, lidar observations are vertical profiles pointed straight upward for a continuous measurement. Since lidar systems provide continuous measurements, higher sampling is achieved, leading to reduced errors. Sampling does not apply to radiosonde measurements because it is a snapshot.

Lidar observations provide better representativity than radiosondes in time and space. Lidar observations are averaged over time according to the model grid configuration

and the meteorological conditions before assimilation. In this study, the temporal averaging of the profiles was for a period of 20 min (T20). Figure 7 shows the absolute error profile of the radiosonde profiles with the lidar profiles both for temperature and WVMR. Presumably, for the above reasons, the temperature error has a magnitude of 1.6 K at 0900 UTC at a height of 1,000 m. Similarly, at a height of 2,500 m, the errors for 0900, 1100 and 1500 UTC average to 1.5 K. The temperature error rises to a peak of 2 K between a height of 1,900 and 2,300 m at 1500 UTC. The high error values in the temperature reduce confidence in the quantitative analysis of the temperature analyses.

4.3 | Assimilation impact on WVMR

For our analysis, we evaluated the RMSEs of the NOLIDAR_3DVAR, LIDAR_3DVAR, NOLIDAR_HYBRID and LIDAR_HYBRID with respect to lidar observations. Figure 8 gives the analyses at 0900, 1100, 1300 and 1500 UTC when radiosonde data were available in addition. The LIDAR_HYBRID analyses were closer to

FIGURE 8   WVMR profiles of WVDIAL, radiosondes, and analyses at (a) 0900, (b) 1100, (c) 1300 and (d) 1500 UTC. The WVDIAL observations (orange) along with its total error are plotted. Radiosonde observations (magenta), which were not assimilated, are plotted for reference. The experiments NOLIDAR_3DVAR (green), LIDAR_3DVAR (red) and LIDAR_HYBRID (purple) are shown in the panel. The background profile (first guess: LIDAR_HYB_FG) of the LIDAR_HYBRID (black) is also shown [Colour figure can be viewed at wileyonlinelibrary.com]
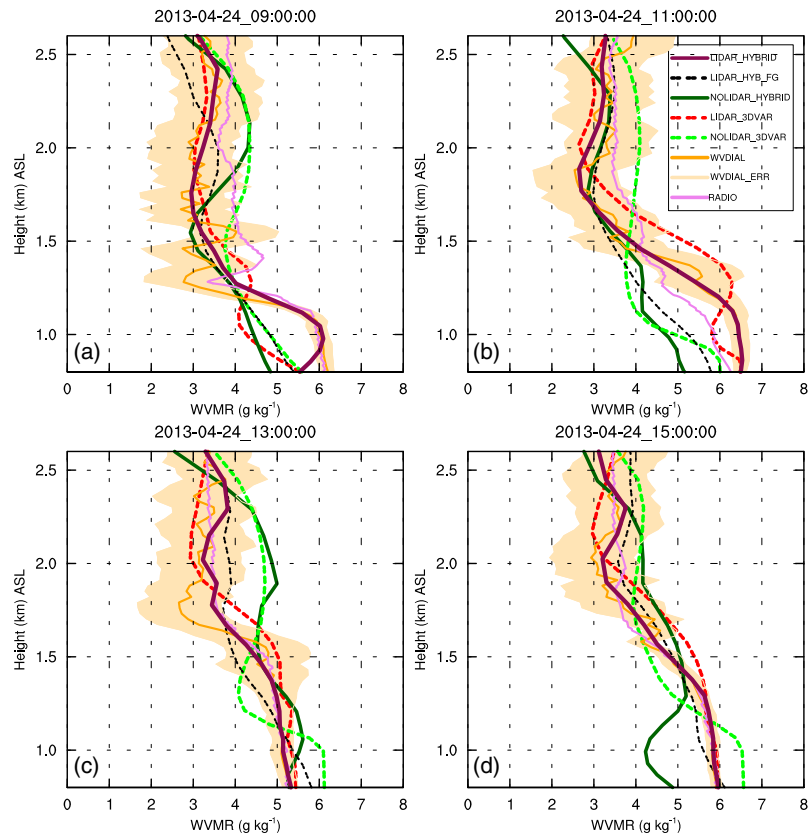
the lidar observations than that of the LIDAR_3DVAR at all four assimilation time-steps. Inside the PBL, until ~1,200 m AGL, the LIDAR_HYBRID results closely correlated with the WVDIAL and radiosonde profiles, whereas LIDAR_3DVAR, NOLIDAR_3DVAR and NOLIDAR_HYBRID showed larger differences to the observations. Especially at 0900 UTC, LIDAR_HYBRID captured the WVMR gradient inside the PBL while the other experiments did not. The LIDAR_HYBRID showed a sudden adjustment to match the lidar observations at an altitude of 900 m when the data became available. Overlap error in lidars occurs due to the incomplete overlapping of the transmitter and receiver telescope field of view. Overlap error is a drawback that does not enable lidar systems to measure parameters at near ranges. The interfacial layer at 1,200 m (Figure 8a) was better captured by the LIDAR_HYBRID than by the LIDAR_3DVAR. The lower troposphere was also very well captured by LIDAR_HYBRID at all four time steps compared to LIDAR_3DVAR.

The individual RMSE of WVMR compared to the WVDIAL observations at the 10 assimilation time steps, the average RMSEs of those time steps, and the change in RMSEs are presented in Figure 9a–c, respectively. The

RMSEs were calculated against the lidar observations between 800 and 2,500 m, where the lidar observations were available. The RMSEs of LIDAR_HYBRID (Figure 9a) were consistently lowest at all time steps. The average RMSE (Figure 9b) fell from $0.52\,\mathrm{g\cdot kg^{-1}}$ in LIDAR_3DVAR to $0.25\,\mathrm{g\cdot kg^{-1}}$ in LIDAR_HYBRID. The RMSE (Figure 9c) of LIDAR_HYBRID fell by $0.64\,\mathrm{g\cdot kg^{-1}}$ compared to NOLIDAR_HYBRID whereas, the LIDAR_3DVAR fell $0.36\,\mathrm{g\cdot kg^{-1}}$ compared to the NOLIDAR_3DVAR experiment. This reduction in RMSEs once again demonstrated the superior performance of the hybrid approach over the deterministic 3DVAR. The relative difference between the RMSEs of NOLIDAR_HYBRID and NOLIDAR_3DVAR was less than $0.01\,\mathrm{g\cdot kg^{-1}}$. The above results clearly demonstrated that the hybrid approach is promising for integrating high-resolution WVMR and temperature profiles from a potential future network of lidar systems operated over the region of interest.

The RMSEs of WVMR with respect to radiosondes were also calculated (Figure 10). The RMSE (Figure 10a) fell from $0.64\,\mathrm{g\cdot kg^{-1}}$ in LIDAR_3DVAR to $0.47\,\mathrm{g\cdot kg^{-1}}$ in the LIDAR_HYBRID experiment. The RMSE (Figure 10b) compared to NOLIDAR_3DVAR fell by $0.1\ \mathrm{g\cdot kg^{-1}}$ in
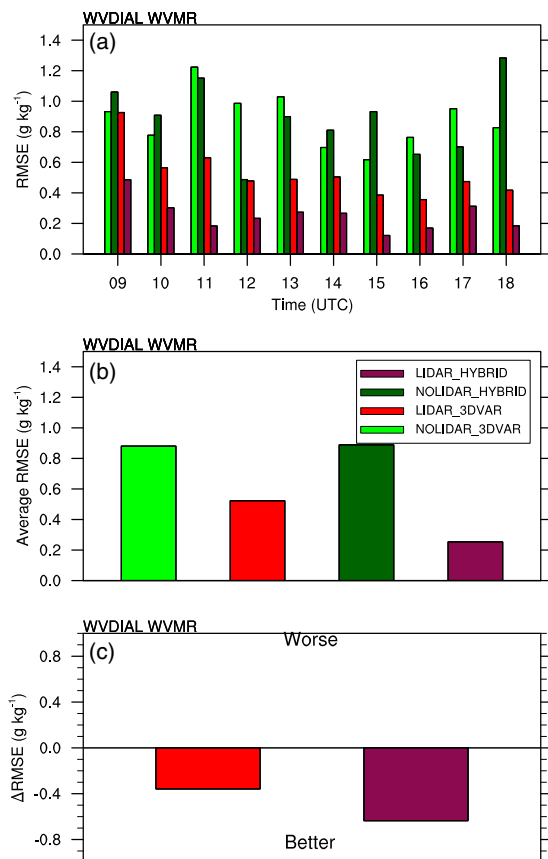
FIGURE 9   WVMR RMSE of the analyses compared to assimilated WVDIAL observations. (a) The RMSE with respect to the WVDIAL observations at the 10 assimilation time steps from 0900 to 1800 UTC 24 April 2013. (b) The mean WVMR RMSE of (a). (c) The relative change in the average RMSD of (b) compared to the RMSE of the respective NOLIDAR analysis [Colour figure can be viewed at wileyonlinelibrary.com]

LIDAR_3DVAR whereas, LIDAR_HYBRID RMSE fell 0.44 g·kg^{-1} compared to the NOLIDAR_HYBRID experiment. The average RMSE of NOLIDAR_HYBRID was about 0.17 g·kg^{-1} higher than in the NOLIDAR_3DVAR experiment. Nevertheless, the improvement in the lidar data-assimilated hybrid DA analysis was more prominent than in the 3DVAR analysis. Considering the overall WVMR RMSE compared to radiosonde observations, LIDAR_HYBRID showed considerable improvement over the LIDAR_3DVAR experiment.

4.4 | **Assimilation impact on temperature**

Figure 11 gives a comparison of temperature profiles with radiosondes and TRL observations for the analyses
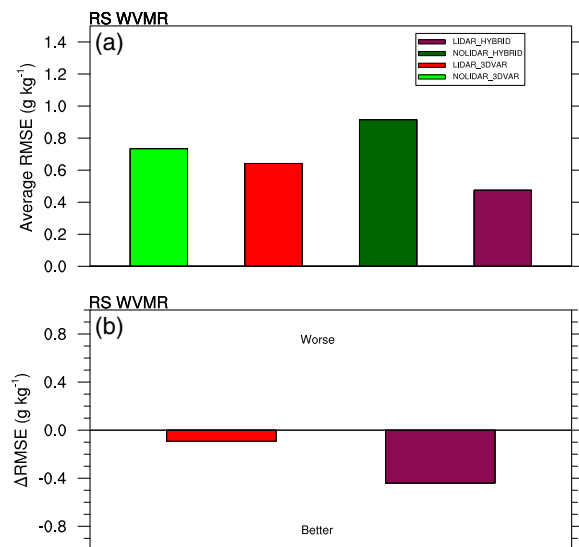
FIGURE 10   WVMR RMSE of the analyses compared to independent radiosonde observations. (a) The mean RMSE with respect to the WVDIAL observations averaged over four time steps 0900, 1100, 1300 and 1500 UTC on 24 April 2013. (b) The relative change in the average RMSE of (a) compared to the respective RMSE of the NOLIDAR analysis [Colour figure can be viewed at wileyonlinelibrary.com]

at 0900, 1100, 1300 and 1500 UTC. Inside the PBL, until 1,200 m AGL, LIDAR_HYBRID correlated best with the observed temperature profiles. The observed interfacial layer temperature at 1200 m AGL was captured by the LIDAR_HYBRID but not by the LIDAR_3DVAR. LIDAR_HYBRID also correlated best with the TRL observations from the interfacial layer to 3,000 m, whereas in the lower troposphere between 1,600 and 3,000 m, LIDAR_HYBRID agreed best with the radiosonde observations as well. Here we have to bear in mind the spatial representativeness errors associated with the drift of radiosondes, as discussed in the previous section. The radiosondes measure a snapshot of the atmosphere rather than a continuous profile like the lidar systems, which may not always be comparable to a temporally averaged set of lidar profiles. We average the lidar profiles in 20 min windows, consisting of 1,200 profiles for TRL and 120 profiles for DIAL since the temporal resolutions are 1 and 10 s, respectively. The averaged lidar profile might not show the temperature spikes visible in the radiosonde profiles at 0900 UTC between 1,000 and 1,600 m and at 1100 UTC between 1,300 and 1,500 m AGL.

With respect to the TRL profiles the LIDAR_HYBRID temperature RMSEs improved considerably at all assimilation time steps (Figure 12a). The average RMSE in the LIDAR_HYBRID experiment showed the lowest value

FIGURE 11 Same as Figure 8 but for temperature [Colour figure can be viewed at wileyonlinelibrary.com]
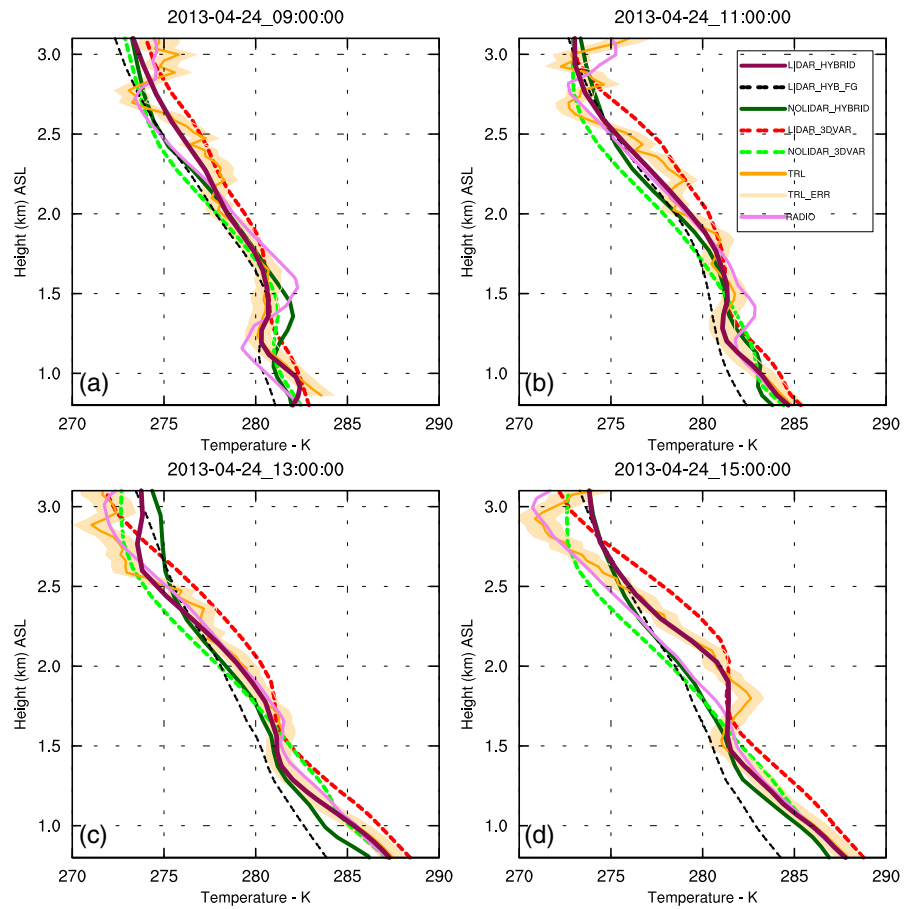
of 0.55 K while it stayed at 1.15 K in the LIDAR_3DVAR experiment. The relative reduction in the RMSE of LIDAR_HYBRID compared to NOLIDAR_HYBRID was 0.54 K. This constitutes a significant improvement over LIDAR_3DVAR, which, when compared to NOLIDAR_3DVAR, had an increment of 0.06 K. The assimilation of both WVMR and temperature together with the 3DVAR DA system did not succeed in simultaneously improving WVMR and temperature.

The RMSEs of temperature with respect to the radiosonde observations were also calculated (Figure 13). The average temperature RMSE in LIDAR_HYBRID fell to 0.87 K from 1.4 K in LIDAR_3DVAR. The change in temperature RMSE compared to radiosondes with respect to the NOLIDAR_HYBRID experiment fell by 0.17 K in the LIDAR_HYBRID while LIDAR_3DVAR, compared to NOLIDAR_3DVAR, increased by 0.59 K. From section 4.2, the temperature absolute error profile of the radiosonde with lidar temperature profiles at 0900, 1100 and 1500 UTC was higher (1.5 K average), and was seen as detrimental in the overall RMSEs. However, we cannot conclude from the error profiles that either radiosonde or lidar profiles

are close to the truth. In order to evaluate the improvement, we need to consider comparing analyses with other independent observations.

4.5 | Impact on planetary boundary-layer height

The impact of assimilating WVMR and temperature on PBLH using the different approaches was validated with ceilometer data in addition to the independent radiosonde measurements. Figure 14 shows the comparison of the analyses with the ceilometer data. The PBLH height was defined by the height of the largest positive potential temperature gradient in the lowest 2,000 m. These heights obtained from all four experiments were compared with ceilometer data. The overall RMSE compared to the ceilometer for the LIDAR_HYBRID experiment was 60 m. The LIDAR_3DVAR experiment performed better than the assimilation experiments without lidar data, still showing an RMSE of 140 m (Figure 14b). The improvement in the hybrid DA is visible from the evaluation of analyses
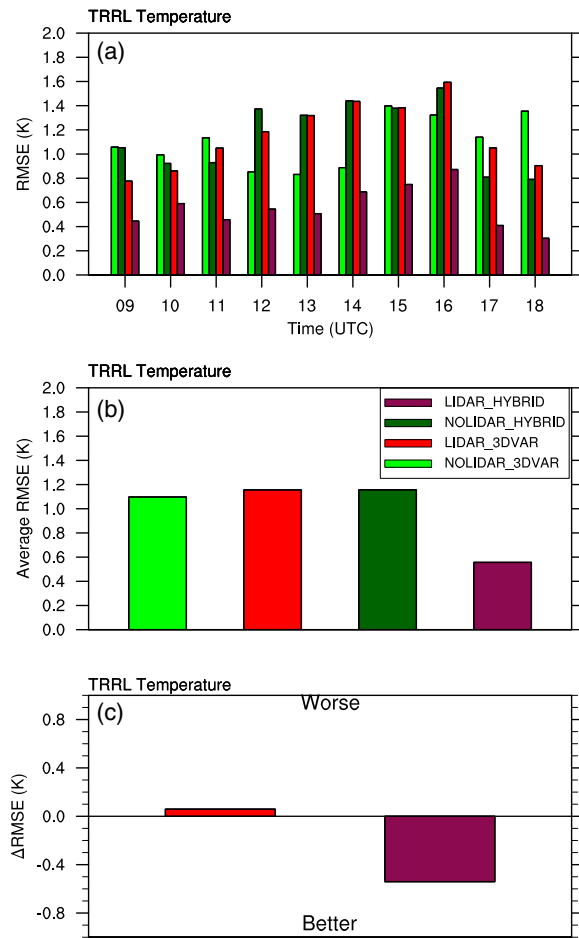
FIGURE 12   Same as Figure 9 but for temperature [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 13   Same as Figure 10 but for temperature [Colour figure can be viewed at wileyonlinelibrary.com]

with the independent ceilometer observations, although radiosonde comparisons were not completely convincing.

4.6 | Spatial extent of the assimilation impact

Figure 15 shows the spatial impact of assimilating lidar profiles by displaying the analysis difference between hybrid DA and 3DVAR experiments. The analysis difference provides insight into the spatial extent of the impact produced by the assimilation of the single lidar profiles. The tuning parameters of the background error correlations were kept identical for both 3DVAR and hybrid DA systems. The spatial plot was averaged over all assimilation times and then vertically averaged until 2,000 m AGL to obtain the spatial impact at different levels. A lower limit of ± 0.2 g·kg^{-1} and ± 0.2 K, kept as thresholds in the plots, for

WVMR and temperature, respectively, was used to eliminate negligible impacts in both horizontal and vertical cross-sections. The radius of the impact of the WVMR was spread over a diameter of 200 km in the hybrid DA with a minimum impact of 0.3 g·kg^{-1}. The maximum impact was 0.8 g·kg^{-1}. The temperature impact spreads over a diameter of 250 km ranging between 0.3 and 0.6 K. The areas of influence on WVMR and temperature were well beyond the hybrid covariance localization scale of 60 km. In the hybrid 3DVAR-ETKF DA system, the background or first-guess of each assimilation time-step is a mean of the ensemble member forecasts from the previous cycle. The ensemble members contain the propagated error produced due to the random perturbation of the initial state of the model state variable while they are generated. Therefore the mean of the ensemble member forecasts (first-guess) should possibly be corrected by the DA system to a greater extent with a larger innovation in DA. The correlation function used for the localization is Gaussian, and the localization impact declined over 60 km gradually but not abruptly.

In the 3DVAR (Figure 15a), the WVMR impact was limited to a 60 km radius, whereas the temperature impacted 150 km. The domain-maximum values were 0.8 g·kg^{-1} and 0.9 K for WVMR and temperature, respectively. The hybrid DA showed a uniform circular impact region with a radius of 100 km rather than 3DVAR. It had a non-uniform spatial impact for WVMR and temperature. The minimum area of impact was found to be a 100 km radius for a localization length-scale of 60 km.
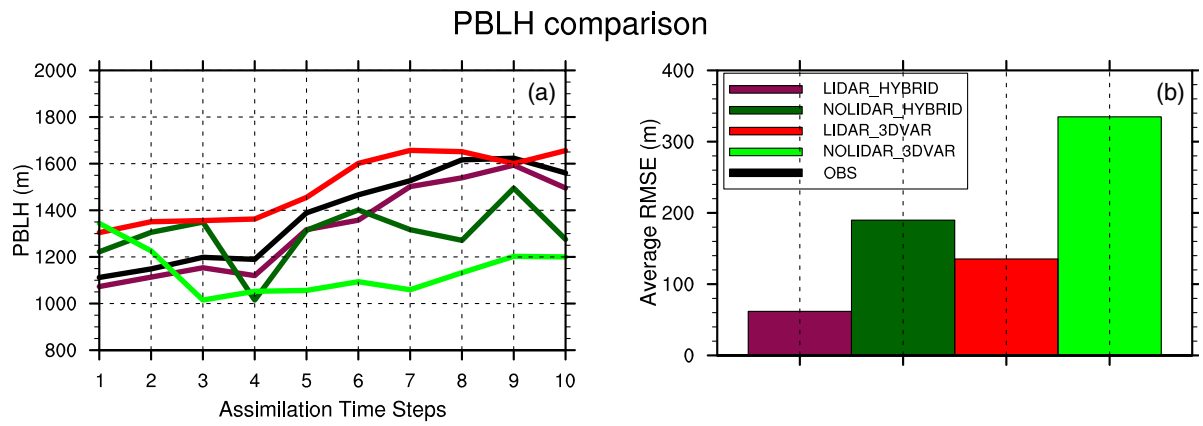
FIGURE 14 Planetary boundary-layer height (PBLH) at the 10 analyses time steps compared to ceilometer-derived PBLH. (a) The PBLH time series comparison with the ceilometer, and (b) the average RMSE of all 10 assimilation time steps for the different experiments [Colour figure can be viewed at wileyonlinelibrary.com]
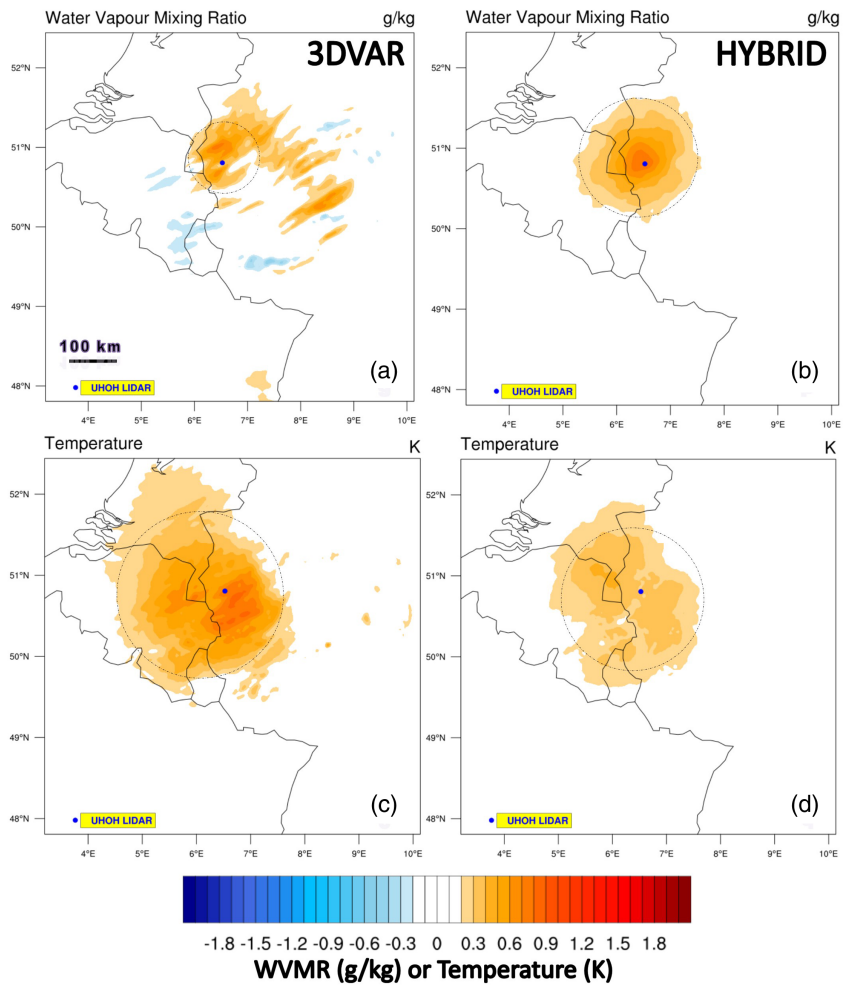
FIGURE 15 Spatial extent of the 3DVAR and hybrid DA impact: spatial averaged analysis difference between hybrid DA with-lidar and without-lidar experiments for 10 assimilation cycles from 0900 to 1800 UTC 24 April 2013 and vertically averaged field till 2000 m AGL. (a,b) The WVMR difference; (c,d) the temperature difference, respectively. The lidar site is marked with a blue dot. A dashed circle has been marked to give some understanding of the circular area of impact [Colour figure can be viewed at wileyonlinelibrary.com]
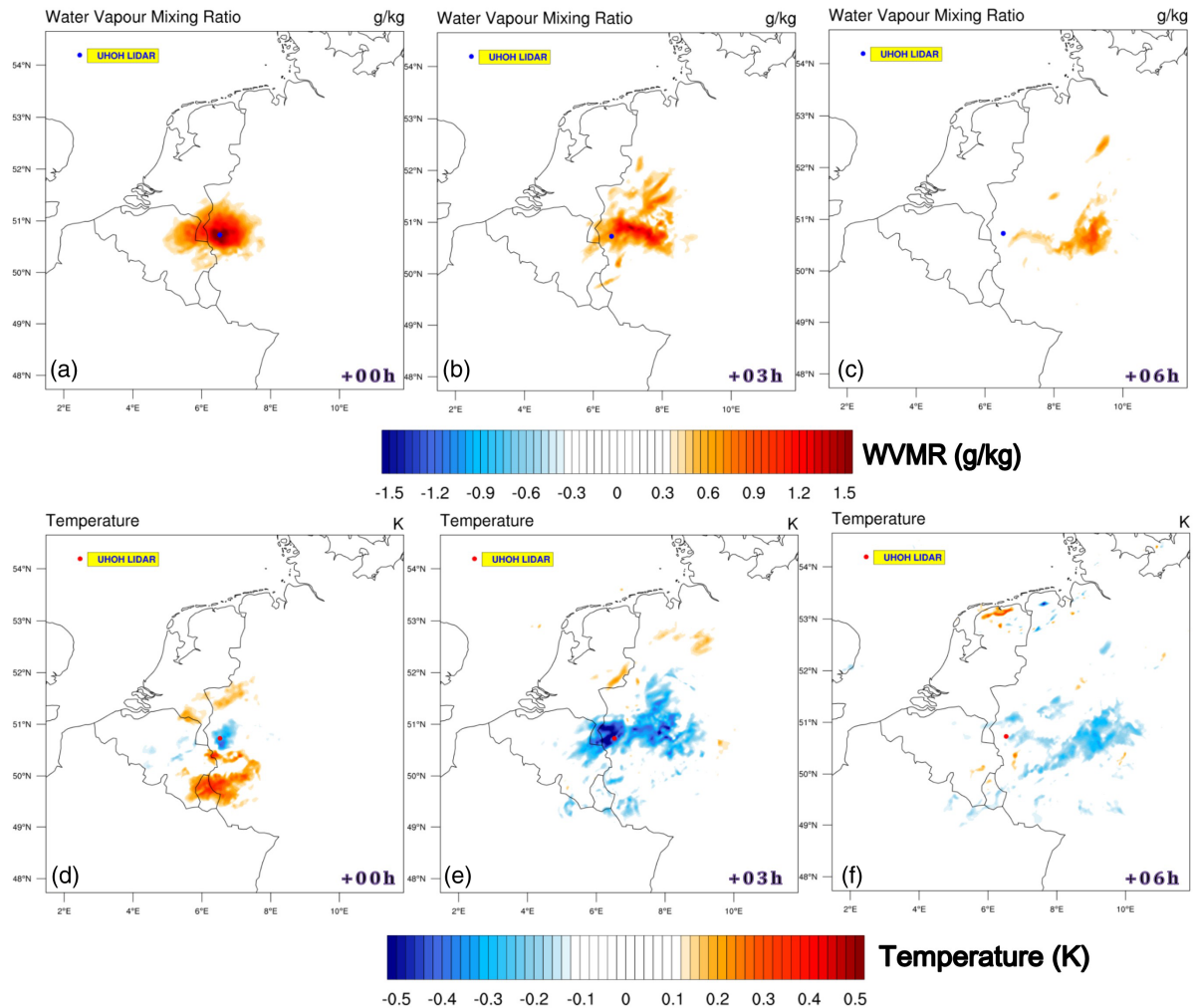
FIGURE 16   Impact of the hybrid assimilation of the lidar data for the forecast: hybrid DA forecast difference between hybrid DA with-lidar and without-lidar experiments, starting from the final assimilation time step at 1800 UTC with 0, 3 and 6 hr lead-time. Differences of (a–c) the WVMR fields and (d–f) the temperature fields. The blue dot in the upper plots and the red dot in the lower plots marks the lidar site [Colour figure can be viewed at wileyonlinelibrary.com]

4.7 | Forecast impact of the assimilation

The spatio-temporal forecast differences between the hybrid DA experiment with-lidar and without-lidar for WVMR and temperature are presented in Figure 16. The spatial forecast plot shows how long the impact of the lidar profiles remains within the model domain. The lower limit threshold, in the plot, of ± 0.3 g·kg^{-1} and ± 0.1 K, for WVMR and temperature, respectively, are applied here. The temperature difference showed a cooling effect during high increments of WVMR. The domain-maximum values of the forecast differences of water vapour fell in the next forecast hours from 1.4 to 0.6 g·kg^{-1} and from 0.5 to 0.25 K for temperature. Also, the forecast difference did not

exhibit unphysical correlations between temperature and WVMR, which could have been introduced by the small ensemble size. The hybrid analysis impact was free of any long-range correlations further away from the lidar location and downstream, apart from the minor differences possibly created by sound or gravity waves.

In order to evaluate the analysis improvement due to assimilation, a forecast evaluation was conducted with respect to the PBLH observations and PWV observations for a 9 hr forecast lead time from the final analysis time step of 1800 UTC. The PBLH forecast results (Figure 17) compared to PBLH observations showed the most positive impact of the lidar data. Comparison of the PWV forecast with the GPS-PWV observations revealed a minor

PBLH Forecast comparison
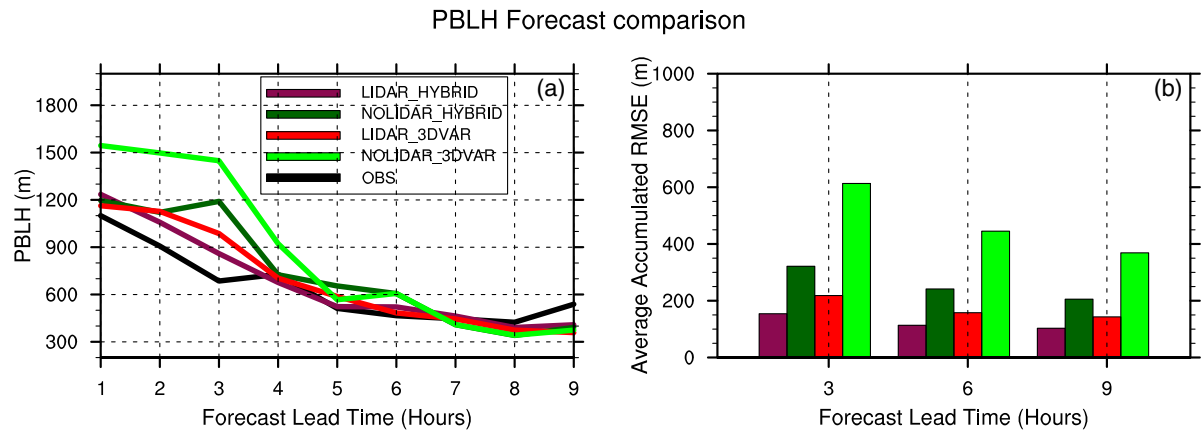
FIGURE 17 Forecast of PBLH from different experiments compared to ceilometer-derived PBLH for 9 hr forecast lead time starting at 1800 UTC. (a) The PBLH time series comparison of the forecasts with the observation, and (b) the average accumulated RMSE every 3 hr [Colour figure can be viewed at wileyonlinelibrary.com]
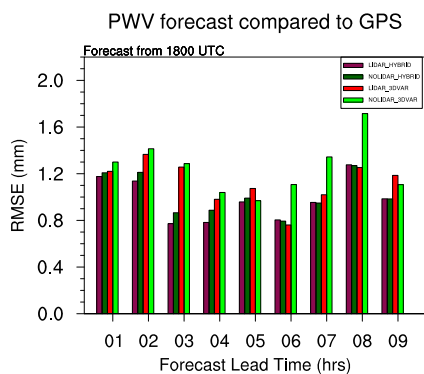
FIGURE 18 RMSE of precipitable water vapour (PWV) forecast compared to GPS-PWV observations for the different experiments for different forecast lead times (hr) starting at 1800 UTC 24 April 2013 [Colour figure can be viewed at wileyonlinelibrary.com]

improvement in the short-range forecast. As shown in Figure 17a, the PBLH forecast of LIDAR_HYBRID performed best compared to other DA runs. The impact of assimilation could be observed in the system for about 7 hr, later converging to the conventional DA run. Although the hybrid simulation performed better than the 3DVAR, LIDAR_3DVAR was better than NOLI-DAR_HYBRID. Figure 17b shows the accumulated average RMSE with respect to the observations every 3 hr for 3, 6 and 9 hr forecast lead time, where LIDAR_HYBRID showed the lowest values of all the experiments. PWV RMSE (Figure 18) is also calculated for the same forecast lead time as made for the PBLH forecast. When comparing the model-derived PWV values with the GPS-PWV observations, the LIDAR_HYBRID simulation again performed

better than other DA runs for the first 4 hr, although the improvement was relatively small. This low magnitude of improvement is expected since we had already assimilated GPS-ZTD observations in all DA runs.

5 | SUMMARY AND OUTLOOK

The study presented promising evidence of the potential of assimilating high-resolution active remote-sensing thermodynamic lidar to improve NWP analyses and forecasts through a case-study from the HOPE campaign in 2013. We measured the scale of improvement by assimilating WVMR and temperature profiles from WVDIAL and TRL with a hybrid ensemble DA system using the TDLIDAR operator from T20. For this purpose, we used the hybrid 3DVAR-ETKF DA system that comes with WRF version 3.8.1. We compared the results of the hybrid DA with a deterministic 3DVAR DA run. We also performed several SOTs to understand how the flow-dependent error covariance matrix influenced the analyses.

The WRF model was set up on the convection-permitting horizontal resolution of 2.5 km over central Europe with 100 vertical levels. A rapid-update cycle approach was adopted. Data of the UHOH-WVDIAL and TRL from the HOPE campaign in western Germany from 0900 to 1800 UTC on 24 April 2013 were assimilated hourly. We analysed four experiments: a simulation with conventional data assimilation using 3DVAR (NOLI-DAR_3DVAR); a similar experiment using the hybrid 3DVAR-ETKF (NOLIDAR_HYBRID); an experiment with lidar data assimilation in addition to the conventional data through 3DVAR (LIDAR_3DVAR), and the same with hybrid DA (LIDAR_HYBRID).

The LIDAR_HYBRID experiment showed an overall improvement over the LIDAR_3DVAR experiment. In comparison to the WVDIAL observations, the WVMR RMSE fell by $0.62\,\mathrm{g\cdot kg^{-1}}$ compared to NOLIDAR_3DVAR whereas the LIDAR_3DVAR experiment showed a lower reduction of $0.36\,\mathrm{g\cdot kg^{-1}}$. A previous study (T20) showed that the simultaneous assimilation of WVDIAL and TRL data utilizing a 3DVAR with a static background error covariance matrix did not improve the temperature analyses. Simultaneous improvement of temperature and WVMR was observed when adopting the hybrid DA approach. The reduction in the temperature RMSE from the LIDAR_HYBRID to NOLIDAR_3DVAR experiments compared to the TRL observations was 0.54 K, whereas the LIDAR_3DVAR experiment even showed a slight RMSE increase of 0.06 K. The RMSEs were calculated in comparison with independent radiosonde observations. The radiosonde observations could not be deemed to be true due to the significant difference between the radiosonde measurements and the lidar profiles. However, the overall RMSE for all times where the radiosondes were launched showed an improvement in hybrid over the 3DVAR.

Apart from radiosondes, PBLH estimates from ceilometers were compared, which showed positive results for the hybrid DA. The LIDAR_HYBRID DA run had the lowest RMSE of 60 m compared to 140 m in LIDAR_3DVAR. Forecast validations were conducted with respect to PBLH estimates and also GPS PWV. A 9 hr forecast starting at 1800 UTC was initiated from each of the different experiments. For both PBLH and IWV forecast, the LIDAR_HYBRID performed better than LIDAR_3DVAR DA. The forecast impact of PBLH persisted for about 7 hr, whereas the PWV impact persisted for about a roughly 4 hr forecast lead time.

In summary, the hybrid DA analyses were considerably improved compared to the 3DVAR DA analyses. We can state that the hybrid assimilation of the high-resolution thermodynamic lidar data has great potential to simultaneously improve both the moisture and temperature analyses on the convection-permitting scale. However, our investigation was currently limited to a single case-study.

We can also conclude that the added flow-dependency in the background error covariance supports the view that the model is able to better capture the thermodynamic flow in the atmosphere than the static background error covariance used in the deterministic 3DVAR.

The spatial extent of the impact when assimilating lidar profiles with an ensemble approach from a single geographical location was significantly higher than with the 3DVAR approach. However, the spatial scale can vary depending on the tuning of background correlations. The tuning parameters of the background correlations were the same for both the DA systems used in this study to understand impact. The analysis increments already of a single lidar vertical profile were spread over a radius of 100 km. The impact area provides insight into the minimum distance between two lidar systems to be set up adjacently in the potential future lidar network. A slightly less than 200 km grid-to-grid spacing of the lidar network could be a new design for a future lidar network from this case-study. However, a single case-study is not enough to derive any conclusions about future lidar network design strategies. We would want to do further research at different locations with different environmental conditions. Several DA research studies have been conducted using the water vapour lidar data which had created a spatial impact over a large area (Grzeschik *et al.*, 2008; Leuenberger *et al.*, 2020; Yoshida *et al.*, 2020). A new initiative, Water vapour Lidar Network Assimilation (WaLiNeAs), to assimilate high-resolution vertical water vapour profile data from a network of six water vapour lidar systems is planned to start in early September 2022 in the western Mediterranean (Flamant *et al.*, 2021). Lidar network DA promises excellent potential for future operational forecasting.

In the future, it will be very interesting to test the sensitivity of hybrid DA with more ensemble members and also with even more advanced DA techniques such as 4DVAR-hybrid or 4DEnVar. Furthermore, it will be interesting to explore the additional incorporation of wind profiles of Doppler lidars into the model. This is expected to simultaneously improve the model dynamics, too. Studies on the assimilation of Doppler wind lidar data alone have already produced very promising results (e.g. Sawada *et al.*, 2015).Our results suggest that hybrid DA approaches with their flow-dependent background error covariance are very well suited for this.

AUTHOR CONTRIBUTIONS
Rohith Thundathil: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; validation; visualization; writing – original draft; writing – review and editing. **Thomas Schwitalla:** Formal analysis; project

administration; resources; supervision; writing – review and editing. **Andreas Behrendt:** Formal analysis; project administration; resources; supervision; writing – review and editing. **Volker Wulfmeyer:** Formal analysis; funding acquisition; project administration; supervision; writing – review and editing.

DATA AVAILABILITY STATEMENT
The numerical model simulations dataset upon which this article is based is too large to be publicly archived with resources currently available. However, necessary data, documentation and methods used to support this study can be made available upon request from the corresponding author.

ORCID
Rohith Thundathil https://orcid.org/0000-0003-4492-1971
Thomas Schwitalla https://orcid.org/0000-0002-7898-8499

REFERENCES
Adam, S., Behrendt, A., Schwitalla, T., Hammann, E. and Wulfmeyer, V. (2016) First assimilation of temperature lidar data into an NWP model: impact on the simulation of the temperature field, inversion strength and PBL depth. *Quarterly Journal of the Royal Meteorological Society*, 142(700), 2882–2896. https://doi.org/10.1002/qj.2875.

Arshinov, Y., Bobrovnikov, S., Serikov, I., Ansmann, A., Wandinger, U., Althausen, D., Mattis, I. and Müller, D. (2005) Daytime operation of a pure rotational Raman lidar by use of a Fabry–Perot interferometer. *Applied Optics*, 44, 3593–3603. https://doi.org/10.1364/ao.44.003593.

Barker, D.M., Huang, W., Guo, Y.-R., Bourgeois, A.J. and Xiao, Q.N. (2004) A three-dimensional variational data assimilation system for MM5: implementation and initial results. *Monthly Weather Review*, 132, 897–914. https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

Bauer, H.-S., Schwitalla, T., Wulfmeyer, V., Bakhshaii, A., Ehret, U., Neuper, M. and Caumont, O. (2015) Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF – a performance test. *Tellus A*, 67, 25047. https://doi.org/10.3402/tellusa.v67.25047.

Behrendt, A., Nakamura, T., Onishi, M., Baumgart, R. and Tsuda, T. (2002) Combined Raman lidar for the measurement of atmospheric temperature, water vapor, particle extinction coefficient, and particle backscatter coefficient. *Applied Optics*, 41, 7657–7666. https://doi.org/10.1364/AO.41.007657.

Behrendt, A. and Reichardt, J. (2000) Atmospheric temperature profiling in the presence of clouds with a pure rotational Raman lidar by use of an interference-filter-based polychromator. *Applied Optics*, 39, 1372–1378. https://doi.org/10.1364/ao.39.001372.

Behrendt, A., Wulfmeyer, V., Hammann, E., Muppa, S.K. and Pal, S. (2015) Profiles of second- to fourth-order moments of turbulent temperature fluctuations in the convective boundary layer: first measurements with rotational Raman lidar. *Atmospheric*

Chemistry and Physics, 15, 5485–5500. https://doi.org/10.5194/acp-15-5485-2015.

Behrendt, A., Wulfmeyer, V., Riede, A., Wagner, G., Pal, S., Bauer, H., Radlach, M. and Späth, F. (2009) Three-dimensional observations of atmospheric humidity with a scanning differential absorption lidar. *Remote Sensing of Clouds and the Atmosphere XIV*, 7475, 74750L. https://doi.org/10.1117/12.835143.

Behrendt, A., Wulfmeyer, V., Senff, C., Muppa, S.K., Späth, F., Lange, D., Kalthoff, N. and Wieser, A. (2020) Observation of sensible and latent heat flux profiles with lidar. *Atmospheric Measurement Techniques*, 13, 3221–3233. https://doi.org/10.5194/amt-13-3221-2020.

Bhawar, R., Di Girolamo, P., Summa, D., Flamant, C., Althausen, D., Behrendt, A., Kiemle, C., Bosser, P., Cacciani, M., Champollion, C., Di Iorio, T., Engelmann, R., Herold, C., Müller, D., Pal, S., Wirth, M. and Wulfmeyer, V. (2011) The water vapour intercomparison effort in the framework of the Convective and Orographically-induced Precipitation Study: airborne-to-ground-based and airborne-to-airborne lidar systems. *Quarterly Journal of the Royal Meteorological Society*, 137(S1), 325–348. https://doi.org/10.1002/qj.697.

Bielli, S., Grzeschik, M., Richard, E., Flamant, C., Champollion, C., Kiemle, C., Dorninger, M. and Brousseau, P. (2012) Assimilation of water-vapour airborne lidar observations: impact study on the COPS precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, 138(667), 1652–1667. https://doi.org/10.1002/qj.1864.

Bishop, C.H., Etherton, B.J. and Majumdar, S.J. (2001) Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, 129, 420–436.

Blumberg, W.G., Turner, D.D., Löhnert, U. and Castleberry, S. (2015) Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part II: Actual retrieval performance in clear-sky and cloudy conditions. *Journal of Applied Meteorology and Climatology*, 54, 2305–2319. https://doi.org/10.1175/JAMC-D-15-0005.1.

Bösenberg, J. (1998) Ground-based differential absorption lidar for water-vapor and temperature profiling: methodology. *Applied Optics*, 37, 3845–3860. https://doi.org/10.1364/AO.37.003845.

Buehner, M. (2005) Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, 131(607), 1013–1043.

Cadeddu, M.P., Peckham, G.E. and Gaffard, C. (2002) The vertical resolution of ground-based microwave radiometers analyzed through a multiresolution wavelet technique. *IEEE Transactions on Geoscience and Remote Sensing*, 40, 531–540. https://doi.org/10.1109/TGRS.2002.1000313.

Chipilski, H.G., Wang, X. and Parsons, D.B. (2020) Impact of assimilating PECAN profilers on the prediction of bore-driven nocturnal convection: a multiscale forecast evaluation for the 6 July 2015 case study. *Monthly Weather Review*, 148, 1147–1175. https://doi.org/10.1175/MWR-D-19-0171.1.

Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M. and Pailleux, J. (1998) The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550), 1783–1807. https://doi.org/10.1256/smsqj.55001.

Cui, Z., Pu, Z., Emmitt, G.D. and Greco, S. (2020) The impact of airborne Doppler aerosol wind (DAWN) lidar wind profiles on numerical simulations of tropical convective systems during the NASA Convective Processes Experiment (CPEX). *Journal of Atmospheric and Oceanic Technology*, 37, 705–722. https://doi.org/10.1175/JTECH-D-19-0123.1.

Di Girolamo, P., Behrendt, A. and Wulfmeyer, V. (2006) Spaceborne profiling of atmospheric temperature and particle extinction with pure rotational Raman lidar and of relative humidity in combination with differential absorption lidar: performance simulations. *Applied Optics*, 45, 2474–2494.

Di Girolamo, P., Cacciani, M., Summa, D., Scoccione, A., De Rosa, B., Behrendt, A. and Wulfmeyer, V. (2017) Characterisation of boundary layer turbulent processes by the Raman lidar BASIL in the frame of HD(CP)^2 observational prototype experiment. *Atmospheric Chemistry and Physics*, 17, 745–767. https://doi.org/10.5194/acp-17-745-2017.

Di Girolamo, P., Flamant, C., Cacciani, M., Richard, E., Ducrocq, V., Summa, D., Stelitano, D., Fourrié, N. and Saïd, F. (2016) Observation of low-level wind reversals in the Gulf of Lion area and their impact on the water vapour variability. *Quarterly Journal of the Royal Meteorological Society*, 142(S1), 153–172.

Di Girolamo, P., Marchese, R., Whiteman, D.N. and Demoz, B.B. (2004) Rotational Raman lidar measurements of atmospheric temperature in the UV. *Geophysical Research Letters*, 31, 1–5. https://doi.org/10.1029/2003GL018342.

Di Girolamo, P., Summa, D. and Ferretti, R. (2009) Rotational Raman lidar measurements for the characterization of stratosphere–troposphere exchange mechanisms. *Journal of Atmospheric and Oceanic Technology*, 26, 1742–1762.

Dinoev, T., Simeonov, V., Arshinov, Y., Bobrovnikov, S., Ristori, P., Calpini, B., Parlange, M. and Bergh, H. (2013) Raman lidar for meteorological observations, RALMO–Part 1: Instrument description. *Atmospheric Measurement Techniques*, 6(5), 1329–1346.

Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012) The Operational Sea surface Temperature and sea Ice Analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158. https://doi.org/10.1016/j.rse.2010.10.017.

Engelbart, D. A. M., and E. Haas, (2006) LAUNCH-2005 – international Lindenberg campaign for assessment of humidity and cloud profiling systems and its impact on high-resolution modelling. *7th International Symposium on Tropospheric profiling: Needs and technologies*, 11–17.

Evensen, G. (2003) The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, 343–367. https://doi.org/10.1007/s10236-003-0036-9.

Flamant, C., Chazette, P., Caumont, O., Di Girolamo, P., Behrendt, A., Sicard M., Totems, J., Lange, D., Fourrié, N., Brousseau, P., Augros, C., Baron, A., Cacciani, M., Comerón, A., De Rosa, B., Ducrocq, V., Genau, P., Labatut, L., Muñoz-Porcar, C., Rodríguez-Gómez, A., Summa, D., Thundathil, R., Wulfmeyer, V. (2021) A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative. *Bulletin of Atmospheric Science and Technology*, 2(1–4). http://dx.doi.org/10.1007/s42865-021-00037-6

Gao, J., M. Xue, and D. J. Stensrud, (2013) The development of a hybrid EnKF-3DVAR algorithm for storm-scale data assimilation. *Advances in Meteorology*, 2013, 512656.

Gao, X. and Gao, S. (2020) Impact of multivariate background error covariance on the WRF-3DVAR assimilation for the Yellow Sea fog modeling. *Advances in Meteorology*, 2020, 8816185.

Goldsmith, J.E.M., Blair, F.H., Bisson, S.E. and Turner, D.D. (1998) Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols. *Applied Optics*, 37(21), 4979–4990.

Grzeschik, M., Bauer, H.-S. and Wulfmeyer, V. (2008) Four-dimensional variational data analysis of water vapor Raman lidar data and their impact on mesoscale forecasts. *Journal of Atmospheric and Oceanic Technology*, 25, 1437–1453.

Hamill, T.M. and Snyder, C. (2000) A hybrid ensemble Kalman filter–3D variational analysis scheme. *Monthly Weather Review*, 128, 2905–2919. https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Hammann, E., Behrendt, A., Le Mounier, F. and Wulfmeyer, V. (2015) Temperature profiling of the atmospheric boundary layer with rotational Raman lidar during the HD(CP)2 observational prototype experiment. *Atmospheric Chemistry and Physics*, 15, 2867–2881. https://doi.org/10.5194/acp-15-2867-2015.

Hartung, D.C., Otkin, J.A., Petersen, R.A., Turner, D.D. and Feltz, W.F. (2012) Assimilation of surface-based boundary layer profiler observations during a cool-season weather event using an observing system simulation experiment. Part II: Forecast assessment. *Monthly Weather Review*, 139, 2327–2346. https://doi.org/10.1175/2011MWR3623.1.

Hong, S.-Y., Park, H., Cheong, H.-B., Kim, J.-E.E., Koo, M.-S., Jang, J., Ham, S., Hwang, S.-O., Park, B.-K., Chan, E.-C. and Li, H. (2013) The Global/Regional Integrated Model system (GRIMs). *Asia-Pacific Journal of Atmospheric Sciences*, 49, 219–243.

Houtekamer, P.L. and Mitchell, H.L. (1998) Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126, 796–811. https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Huang, X.Y., Xiao, Q., Barker, D.M., Zhang, X., Michalakes, J., Huang, W., Henderson, T., Bray, J., Chen, Y., Ma, Z., Dudhia, J., Guo, Y., Zhang, X., Won, D.-J., Lin, H.-C. and Kuo, Y.-H. (2009) Four-dimensional variational data assimilation for WRF: formulation and preliminary results. *Monthly Weather Review*, 137, 299–314. https://doi.org/10.1175/2008MWR2577.1.

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A. and Collins, W.D. (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113(D13), 2–9. https://doi.org/10.1029/2008JD009944.

Ingleby, B. (2017) *An assessment of different radiosonde types 2015/2016*. ECMWF Technical Memorandum 807, 69 pp.

Ismail, S. and Browell, E.V. (1989) Airborne and spaceborne lidar measurements of water vapor profiles: a sensitivity analysis; erratum. *Applied Optics*, 28, 4981. https://doi.org/10.1364/ao.28.004981.

Kamineni, R., Krishnamurti, T.N., Pattnaik, S., Browell, E.V., Ismail, S. and Ferrare, R.A. (2006) Impact of CAMEX-4 datasets for hurricane forecasts using a global model. *Journal of the Atmospheric Sciences*, 63, 151–174.

Kamineni, R., Krishnamurti, T.N., Ferrare, R.A., Ismail, S. and Browell, E.V. (2003) Impact of high resolution water vapor cross-sectional data on hurricane forecasting. *Geophysical Research Letters*, 30(5), 1–4.

Kiemle, C., Ehret, G., Fix, A., Wirth, M., Poberaj, G., Brewer, W.A., Hardesty, R.M., Senff, C. and LeMone, M.A. (2007) Latent heat

flux profiles from collocated airborne water vapor and wind lidars during IHOP_2002. *Journal of Atmospheric and Oceanic Technology*, 24, 627–639.

Kutty, G., Muraleedharan, R. and Kesarkar, A.P. (2018) Impact of representing model error in a hybrid ensemble-variational data assimilation system for track forecast of tropical cyclones over the Bay of Bengal. *Pure and Applied Geophysics*, 175, 1155–1167. https://doi.org/10.1007/s00024-017-1747-z.

Kutty, G. and Wang, X. (2015) A comparison of the impacts of radiosonde and AMSU radiance observations in GSI based 3DEnsVar and 3DVar data assimilation systems for NCEP GFS. *Advances in Meteorology*, 2015, 280546.

Lange, D., Behrendt, A. and Wulfmeyer, V. (2019) Compact operational tropospheric water vapor and temperature Raman lidar with turbulence resolution. *Geophysical Research Letters*, 46, 14844–14853. https://doi.org/10.1029/2019GL085774.

Leblanc, T., McDermid, I.S. and Walsh, T.D. (2012) Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring. *Atmospheric Measurement Techniques*, 5, 17–36. https://doi.org/10.5194/amt-5-17-2012.

Leuenberger, D., Haefele, A., Omanovic, N., Fengler, M., Martucci, G., Calpini, B., Fuhrer, O. and Rossa, A. (2020) Improving high-impact numerical weather prediction with lidar and drone observations. *Bulletin of the American Meteorological Society*, 101, E1036–E1051. https://doi.org/10.1175/BAMS-D-19-0119.1.

Lorenc, A.C. (2003) The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 129(595), 3183–3203. https://doi.org/10.1256/qj.02.132.

Macke, A., Seifert, P., Baars, H., Barthlott, C., Beekmans, C., Behrendt, A., Bohn, B., Brueck, M., Bühl, J., Crewell, S., Damian, T., Deneke, H., Düsing, S., Foth, A., Di Girolamo, P., Hammann, E., Heinze, R., Hirsikko, A., Kalisch, J., Kalthoff, N., Kinne, S., Kohler, M., Löhnert, U., Madhavan, B.L., Maurer, V., Muppa, S.K., Schween, J., Serikov, I., Siebert, H., Simmer, C., Späth, F., Steinke, S., Träumner, K., Trömel, S., Wehner, B., Wieser, A., Wulfmeyer, V. and Xie, X. (2017) The HD(CP)² Observational Prototype Experiment (HOPE) – an overview. *Atmospheric Chemistry and Physics*, 17, 4887–4914.

Metzendorf, S. (2018) *10 W-average-power single-frequency Ti: sapphire laser with tuning agility – a breakthrough in high-resolution 3D water-vapor measurement*. Dr. Dissertation UHOH.

Muppa, S.K., Behrendt, A., Späth, F., Wulfmeyer, V., Metzendorf, S. and Riede, A. (2016) Turbulent humidity fluctuations in the convective boundary layer: case studies using water vapour differential absorption lidar measurements. *Boundary-Layer Meteorology*, 158, 43–66. https://doi.org/10.1007/s10546-015-0078-9.

Nakanishi, M. and Niino, H. (2006) An improved Mellor–Yamada Level-3 model: its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, 119, 397–407. https://doi.org/10.1007/s10546-005-9030-8.

Newsom, R.K., Turner, D.D. and Goldsmith, J.E. (2013) Long-term evaluation of temperature profiles measured by an operational Raman lidar. *Journal of Atmospheric and Oceanic Technology*, 30(8), 1616–1634.

Niu, G.Y., Yang, Z.-L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M. and Xia, Y. (2011) The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, 116(D12), 1–19. https://doi.org/10.1029/2010JD015139.

Otkin, J.A., Hartung, D.C., Turner, D.D., Petersen, R.A., Feltz, W.F. and Janzon, E. (2011) Assimilation of surface-based boundary layer profiler observations during a cool-season weather event using an observing system simulation experiment. Part I: Analysis impact. *Monthly Weather Review*, 139, 2309–2326. https://doi.org/10.1175/2011MWR3622.1.

Pan, Y., Zhu, K., Xue, M., Wang, X., Hu, M., Benjamin, S.G., Weygandt, S.S. and Whitaker, J.S. (2014) A GSI-based coupled EnSRF-En3DVar hybrid data assimilation system for the operational rapid refresh model: tests at a reduced resolution. *Monthly Weather Review*, 142, 3756–3780. https://doi.org/10.1175/MWR-D-13-00242.1.

Parrish, D.F. and Derber, J.C. (1992) The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120, 1747–1763. https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Radlach, M., Behrendt, A. and Wulfmeyer, V. (2008) Scanning rotational Raman lidar at 355 nm for the measurement of tropospheric temperature fields. *Atmospheric Chemistry and Physics*, 8, 159–169. https://doi.org/10.5194/acp-8-159-2008.

Reichardt, J., Wandinger, U., Klein, V., Mattis, I., Hilber, B. and Begbie, R. (2012) RAMSES: German Meteorological Service autonomous Raman lidar for water vapor, temperature, aerosol, and cloud measurements. *Applied Optics*, 51(34), 8111–8131.

Repasky, K.S., Moen, E., Spuler, S., Nehrir, A.R. and Carlsten, J.L. (2013) Progress towards an autonomous field deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor in the lower troposphere. *Remote Sensing*, 5, 6241–6259. https://doi.org/10.3390/rs5126241.

Saito, K., Fujita, T., Yamada, Y., Ishida, J.-i., Kumagai, Y., Aranami, K., Ohmori, S., Nagasawa, R., Kumagai, S., Muroi, C., Kato, T., Eito, H. and Yamazaki, Y. (2006) The operational JMA nonhydrostatic mesoscale model. *Monthly Weather Review*, 134, 1266–1298.

Sakai, T., Nagai, T., Izumi, T., Yoshida, S. and Shoji, Y. (2019) Automated compact mobile Raman lidar for water vapor measurement: instrument description and validation by comparison with radiosonde, GNSS, and high-resolution objective analysis. *Atmospheric Measurement Techniques*, 12, 313–326. https://doi.org/10.5194/amt-12-313-2019.

Sawada, M., Sakai, T., Iwasaki, T., Seko, H., Saito, K. and Miyoshi, T. (2015) Assimilating high-resolution winds from a Doppler lidar using an ensemble Kalman filter with lateral boundary adjustment. *Tellus A*, 67, 23473. https://doi.org/10.3402/tellusa.v67.23473.

Schwartz, C.S. and Liu, Z. (2014) Convection-permitting forecasts initialized with continuously cycling limited-area 3DVAR, ensemble Kalman filter, and "hybrid" variational-ensemble data assimilation systems. *Monthly Weather Review*, 142, 716–738. https://doi.org/10.1175/MWR-D-13-00100.1.

Schwartz, C.S., Liu, Z. and Huang, X.-Y. (2015) Sensitivity of limited-area hybrid variational-ensemble analyses and forecasts to ensemble perturbation resolution. *Monthly Weather Review*, 143, 3454–3477. https://doi.org/10.1175/MWR-D-14-00259.1.

Schwartz, C.S., Liu, Z., Huang, X.-Y., Kuo, Y.-H. and Fong, C.-T. (2013) Comparing limited-area 3DVAR and hybrid variational-ensemble data assimilation methods for typhoon track forecasts: sensitivity to outer loops and vortex relocation.

THUNDATHIL ET AL.

Monthly Weather Review, 141, 4350–4372. https://doi.org/10.1175/MWR-D-13-00028.1.

Schwitalla, T., Bauer, H.S., Wulfmeyer, V. and Aoshima, F. (2011) High-resolution simulation over central Europe: assimilation experiments during COPS IOP 9c. *Quarterly Journal of the Royal Meteorological Society*, 137(S1), 156–175. https://doi.org/10.1002/qj.721.

Schwitalla, T. and Wulfmeyer, V. (2014) Radar data assimilation experiments using the IPM WRF rapid update cycle. *Meteorologische Zeitschrift*, 23, 79–102. https://doi.org/10.1127/0941-2948/2014/0513.

Shen, F., Min, J. and Xu, D. (2016) Assimilation of radar radial velocity data with the WRF hybrid ETKF-3DVAR system for the prediction of hurricane *Ike* (2008). *Atmospheric Research*, 169, 127–138. https://doi.org/10.1016/j.atmosres.2015.09.019.

Shen, F., Xue, M. and Min, J. (2017) A comparison of limited-area 3DVAR and ETKF-En3DVAR data assimilation using radar observations at convective scale for the prediction of typhoon *Saomai* (2006). *Meteorological Applications*, 24, 628–641. https://doi.org/10.1002/met.1663.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W. and Powers, J. G. (2008) *A description of the Advanced Research WRF version 3*. NCAR Technical Note 475+STR. https://doi.org/10.5065/D6DZ069T

Späth, F., Behrendt, A., Kumar Muppa, S., Metzendorf, S., Riede, A. and Wulfmeyer, V. (2016) 3-D water vapor field in the atmospheric boundary layer observed with scanning differential absorption lidar. *Atmospheric Measurement Techniques*, 9, 1701–1720. https://doi.org/10.5194/amt-9-1701-2016.

Spuler, S.M., Repasky, K.S., Morley, B., Moen, D., Hayman, M. and Nehrir, A.R. (2015) Field-deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor. *Atmospheric Measurement Techniques*, 8, 1073–1087. https://doi.org/10.5194/amt-8-1073-2015.

Thompson, G., Field, P.R., Rasmussen, R.M. and Hall, W.D. (2008) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, 136, 5095–5115. https://doi.org/10.1175/2008MWR2387.1.

Thundathil, R., Schwitalla, T., Behrendt, A., Muppa, S.K., Adam, S. and Wulfmeyer, V. (2020) Assimilation of lidar water vapour mixing ratio and temperature profiles into a convection-permitting model. *Journal of the Meteorological Society of Japan*, 98, 959–986. https://doi.org/10.2151/jmsj.2020-049.

Turner, D.D. and Goldsmith, J.E.M. (1999) Twenty-four-hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement Program's 1996 and 1997 water vapor intensive observation periods. *Journal of Atmospheric and Oceanic Technology*, 16(8), 1062–1076.

Turner, D.D., Ferrare, R.A., Wulfmeyer, V. and Scarino, A.J. (2014a) Aircraft evaluation of ground-based Raman lidar water vapor turbulence profiles in convective mixed layers. *Journal of Atmospheric and Oceanic Technology*, 31, 1078–1088.

Turner, D.D. and Löhnert, U. (2014) Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *Journal of Applied Meteorology and Climatology*, 53, 752–771. https://doi.org/10.1175/JAMC-D-13-0126.1.

Turner, D.D., Wulfmeyer, V., Berg, L.K. and Schween, J.H. (2014b) Water vapor turbulence profiles in stationary continental convective mixed layers. *Journal of Geophysical Research – Atmospheres*, 119, 11151–11165.

Veselovskii, I., Whiteman, D.N., Kolgotin, A., Andrews, E. and Korenskii, M. (2009) Demonstration of aerosol property profiling by multiwavelength lidar under varying relative humidity conditions. *Journal of Atmospheric and Oceanic Technology*, 26, 1543–1557. https://doi.org/10.1175/2009JTECHA1254.1.

Wagner, G., Behrendt, A., Wulfmeyer, V., Späth, F. and Schiller, M. (2013) High-power Ti: sapphire laser at 820 nm for scanning ground-based water–vapor differential absorption lidar. *Applied Optics*, 52, 2454–2469. https://doi.org/10.1364/AO.52.002454.

Wagner, G., Wulfmeyer, V. and Behrendt, A. (2011) Detailed performance modeling of a pulsed high-power single-frequency Ti: sapphire laser. *Applied Optics*, 50, 5921–5937. https://doi.org/10.1364/AO.50.005921.

Wang, X. (2011) Application of the WRF hybrid ETKF–3DVAR data assimilation system for hurricane track forecasts. *Weather and Forecasting*, 26, 868–884. https://doi.org/10.1175/WAF-D-10-05058.1.

Wang, X., Barker, D.M., Snyder, C. and Hamill, T.M. (2008a) A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Monthly Weather Review*, 136, 5132–5147. https://doi.org/10.1175/2008MWR2445.1.

Wang, X., Barker, D.M., Snyder, C. and Hamill, T.M. (2008b) A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Monthly Weather Review*, 136, 5116–5131. https://doi.org/10.1175/2008MWR2444.1.

Wang, X., Hamill, T.M., Whitaker, J.S. and Bishop, C.H. (2007) A comparison of hybrid ensemble transform Kalman filter–optimum interpolation and ensemble square root filter analysis schemes. *Monthly Weather Review*, 135, 1055–1076. https://doi.org/10.1175/MWR3307.1.

Wang, X., Parrish, D., Kleist, D. and Whitaker, J. (2013) GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: single-resolution experiments. *Monthly Weather Review*, 141, 4098–4117. https://doi.org/10.1175/MWR-D-12-00141.1.

Weckwerth, T.M., Weber, K.J., Turner, D.D. and Spuler, S.M. (2016) Validation of a water vapor micropulse differential absorption lidar (DIAL). *Journal of Atmospheric and Oceanic Technology*, 33, 2353–2372. https://doi.org/10.1175/JTECH-D-16-0119.1.

Weisman, M.L., Davis, C., Wang, W., Manning, K.W. and Klemp, J.B. (2008) Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Weather and Forecasting*, 23, 407–437. https://doi.org/10.1175/2007WAF2007005.1.

Whitaker, J.S. and Hamill, T.M. (2002) Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130, 1913–1924.

Wulfmeyer, V. (1999a) Investigation of turbulent processes in the lower troposphere with water vapor DIAL and radar–RASS. *Journal of the Atmospheric Sciences*, 56, 1055–1076.

Wulfmeyer, V. (1999b) Investigations of humidity skewness and variance profiles in the convective boundary layer and comparison of the latter with large eddy simulation results. *Journal of the Atmospheric Sciences*, 56, 1077–1087.

Wulfmeyer, V., Bauer, H-S., Grzeschik, M., Behrendt, A., Vandenberghe, F., Browell, E.V., Ismail, S. and Ferrare, R.A. (2006)

Four-dimensional variational assimilation of water vapor differential absorption lidar data: the first case study within IHOP_2002. *Monthly Weather Review*, 134, 209–230. https://doi.org/10.1175/MWR3070.1.

Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G.C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C.N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G., Richard, E., Rotach, M.W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D.D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A. and Wirth, M. (2011) The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights. *Quarterly Journal of the Royal Meteorological Society*, 137(S1), 3–30. https://doi.org/10.1002/qj.752.

Wulfmeyer, V., Hardesty, R.M., Turner, D.D., Behrendt, A., Cadeddu, M.P., Di Girolamo, P., Schlüssel, P., Van Baelen, J. and Zus, F. (2015) A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles. *Reviews of Geophysics*, 53, 819–895. https://doi.org/10.1002/2014RG000476.

Wulfmeyer, V., Pal, S., Turner, D.D. and Wagner, E. (2010) Can water vapour Raman lidar resolve profiles of turbulent variables in the convective boundary layer? *Boundary-Layer Meteorology*, 136, 253–284. https://doi.org/10.1007/s10546-010-9494-z.

Yang, Z.L., Niu, G.-Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M. and Xia, Y. (2011) The community Noah land surface model with multi-parameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research*, 116(D12), 1–16. https://doi.org/10.1029/2010JD015140.

Yoshida, S., Yokota, S., Seko, H., Sakai, T. and Nagai, T. (2020) Observation system simulation experiments of water vapor profiles observed by Raman lidar using LETKF system. *SOLA*, 16, 43–50. https://doi.org/10.2151/sola.2020-008.

Zhang, F., Zhang, M. and Poterjoy, J. (2013) E3DVar: coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Monthly Weather Review*, 141, 900–917. https://doi.org/10.1175/MWR-D-12-00075.1.

Chapter 4

A Network of Water Vapor Raman Lidars for improving Heavy Precipitation Forecasting in Southern France: Introducing the WaLiNeAs initiative

4.1 Overview

Implementing lidar systems as a part of our daily weather prediction instrumentation is a vision into advanced forecasts. As a starting point, it is essential to gain more insight into the impact of the DA of thermodynamic profiles and other prognostic variables from lidar systems before implementing lidar systems on an operational level with other observation instruments. One of the key objectives of this thesis is to make a foundation for more lidar

impact studies and to understand how lidar systems can solve our problems like observational gaps and the inability to characterise the fine-scale weather phenomena which other instrumentation techniques lack.

This chapter provides an overview of an excellent initiative in the direction of paving the way for a new approach of operational forecasting using a network of water vapour Raman lidar systems. A unique and innovative French initiative project called the Water Vapor Lidar Network assimilation (WaLiNeAs) will start a measurement campaign in early September 2022. The measurement campaign will deploy a network of autonomous water vapour lidars from France, Germany, and Italy research groups across the Western Mediterranean. At the University of Hohenheim (UHOH), the transportable and automatic Atmospheric Raman Temperature and Humidity Sounder (ARTHUS, Lange et al., 2019) has been developed. Between November 2018 and October 2021, ARTHUS collected more than one year of data, including temperature and water-vapour mixing ratio with unprecedented temporal and range resolutions and accuracies. It was operated at the UHOH Land-Atmosphere Feedback Observatory (LAFO[1]) as well as in various field campaigns such as on a shipborne platform during the Elucidating the Role of Cloud-Circulation Coupling in Climate (EUREC[4]A, Stevens et al., 2021) campaign and recently at the Meteorologisches Observatorium Lindenberg - Richard Aßmann-Observatorium (MOL-RAO) DWD observatory in Lindenberg between July and September 2021. Probabilistic quantitative precipitation forecasting (PrQPF) is a challenging field of meteorology, which is fundamental for predicting and quantifying extreme precipitation events. With advanced remote-sensing instruments such as lidar systems, acquiring the high-resolution temporal and spatial dynamical and thermodynamic data for input to the numerical weather prediction (NWP) models through data assimilation (DA) techniques is possible. The project aims to implement an integrated prediction tool to enhance the forecast of HPEs in southern France, primarily demonstrating the benefit of assimilating vertically resolved water vapour data in the new version of the French operational AROME NWP system. Flamant et al., 2021 has detailed the goals and objectives of the research initiative in the next section of this chapter. As a result of this research initiative, future assimilation strategies can be better designed, and the importance and density of lidar networks to help predict short-range forecasts can be better understood.

---

[1] **https://lafo.uni-hohenheim.de/en**.

4.2 Publication

RESEARCH ARTICLE

A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative

Cyrille Flamant[1] · **Patrick Chazette[2]** · **Olivier Caumont[3]** · **Paolo Di Girolamo[4]** ·
Andreas Behrendt[5] · **Michaël Sicard[6,7]** · **Julien Totems[2]** · **Diego Lange[5]** ·
Nadia Fourrié[3] · **Pierre Brousseau[3]** · **Clotilde Augros[3]** · **Alexandre Baron[2]** ·
Marco Cacciani[4] · **Adolfo Comerón[6]** · **Benedetto De Rosa[4]** · **Veronique Ducrocq[3]** ·
Pascal Genau[1] · **Laurent Labatut[3]** · **Constantino Muñoz-Porcar[6]** ·
Alejandro Rodríguez-Gómez[6] · **Donato Summa[4,8]** · **Rohith Thundathil[5]** ·
Volker Wulfmeyer[5]

Abstract

Extreme heavy precipitation events (HPEs) pose a threat to human life but remain difficult to predict because of the lack of adequate high frequency and high-resolution water vapor (WV) observations in the low troposphere (below 3 km). To fill this observational gap, we aim at implementing an integrated prediction tool, coupling network measurements of WV profiles, and a numerical weather prediction model to precisely estimate the amount, timing, and location of rainfall associated with HPEs in southern France (struck by ~7 HPEs per year on average during the fall). The Water vapor Lidar Network Assimilation (WaLiNeAs) project will deploy a network of 6 autonomous Raman WV lidars around the Western Mediterranean to provide measurements with high vertical resolution and accuracy to be assimilated in the French Application of Research to Operations at Mesoscale (AROME-France) model, using a four-dimensional ensemble-variational approach with 15-min updates. This integrated prediction tool is expected to enhance the model capability for kilometer-scale prediction of HPEs over southern France up to 48 h in advance. The field campaign is scheduled to start early September 2022, to cover the period most propitious to heavy precipitation events in southern France. The Raman WV lidar network will be operated by a consortium of French, German, Italian, and Spanish research groups. This project will lead to recommendations on the lidar data processing for future operational exploitation in numerical weather prediction (NWP) systems.

Keywords Remote sensing · Numerical weather prediction model · AROME ·
Assimilation · Western Mediterranean

✉ Cyrille Flamant
    cyrille.flamant@latmos.ipsl.fr

Extended author information available on the last page of the article

                                                                    Springer

1 Introduction

Heavy precipitation events (HPEs) pose a threat to human lives (e.g., Llasat et al. 2013) as well as the economy and the environment of impacted regions. HPEs occurring in small and steep watersheds are responsible for the triggering of flash floods with a sudden and often violent onset and rapid rise of rivers, typically from 1 to 6 h following the causative rainfall (Gaume et al. 2009). Flash floods and landslides lead to fatalities, loss of crops and livestocks, damage to infrastructures, as well as disruption of transport and communication. HPEs remain difficult to predict. Considerable efforts to improve the forecast skill for such severe events have been made in recent years, and significant progress has been realized through the development of kilometer-scale numerical weather NWP systems (Ducrocq et al. 2014) and data assimilation techniques (e.g., Kwon et al. 2018, Gustafsson et al. 2018). However, our ability to predict such high-impact events remains limited because of the lack of adequate high frequency, high-resolution vertically resolved water vapor (WV) observations in the low troposphere to be assimilated in NWP systems, and especially in the boundary layer (Weckwerth et al. 2004; Wulfmeyer et al. 2015; Leuenberger et al. 2020).

The implementation of an integrated prediction tool, coupling network measurements of WV profiles, and a NWP model, to precisely estimate the amount, timing, and locations of rainfall associated with HPEs up to 48 h in advance, is a strong societal demand, especially in regions of France most exposed to heavy rainfall (defined as maximum accumulation in excess of 150 mm per day, Ricard et al. 2012) as those located along the Mediterranean coast. Figure 1 shows the geographical distribution of HPEs obtained from rain gauges for the period 1970–2019) in southern France. Over this period, HPEs are most numerous in Languedoc-Roussillon, along the southern edge of the Cevennes range, between the Mediterranean coastline and the southern Alps, and along the eastern side of Corsica. The two most important HPEs seen in Fig. 1 are located in the Aude department and are related to the 12–13 November 1999 event (Nuissier et al. 2008; Ducrocq et al. 2008) and the 14–15 October 2018 event (Caumont et al. 2021). Large amounts of rainfall associated with these
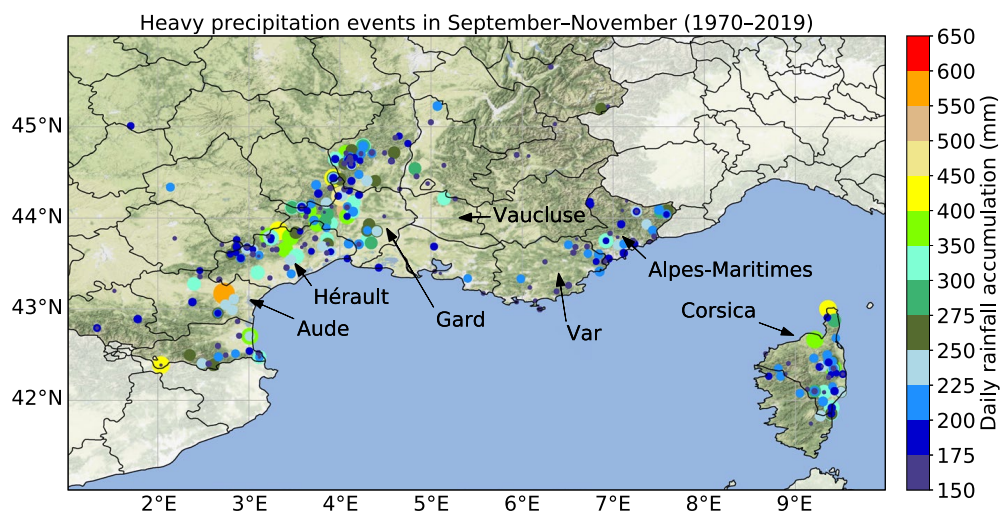
Fig. 1 Geographical distribution of heavy precipitation events defined as maximum accumulation > 150 mm/day and separated from other events by a distance of more 100 km. The size of the circle is a function of accumulated precipitation for a given event. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap

cases were attributed to strong synoptic forcing and associated with quasi-stationary mesoscale convective systems (MCSs) which, for instance, led to accumulated surface precipitation reaching about 620 mm for the former event (Ducrocq et al. 2008). Other remarkable events in the same area (Languedoc-Roussillon and Cevennes) have been observed in the Hérault and Gard which are also linked to catastrophic events such as the 13–14 October 1995 and 8–9 September 2002 cases, respectively (Ducrocq et al. 2008) also related to torrential rainfall cause by stationary MCSs. In the southern Alps area, remarkable HPE cases include the Vaison-La-Romaine event in the Vaucluse department (22 September 1992) and the Côte d'Azur event in the Alpes Maritimes department (3 October 2015). In addition, two other cases are included that have been identified during the hydrological Cycle in the Mediterranean Experiment First Special Observing Period (HyMeX SOP1) on 14 and 26 October and 2012 (Duffourg et al. 2016, 2018). Finally, a famous HPE in Corsica was the 31 October–1 November 1993 case which affected the eastern side of Island with up to 450 mm of rain in a day in several locations.

Southern France is a region stricken by an average of 7 HPEs per year during the fall (September to November) (Ricard et al. 2012, Ducrocq et al., 2014) as illustrated by Fig. 2 for the period 1970–2019. It is worth noting that HPEs can also occur outside of the September–November period, as for instance in December when more than an event per year can be observed. All other months exhibit less than one HPE per year, with the minimum in the monthly climatology being observed in June and July, which are the months less favorable for sustainability of HPEs (lower sea surface temperatures than in the fall and less eastward moving low pressure disturbances across the Mediterranean).

Furthermore, and in close connection with climate change issues, the analysis of extreme Mediterranean rainfall events for the French regions over the last few decades shows an intensification of heavy rainfall between 1961 and 2015 ($+22\%$ on the annual maximum daily totals) and an increase in the frequency of the strongest Mediterranean episodes, particularly those exceeding the 200 mm threshold in 24 h (Ribes et al. 2019).

Accurate characterization of WV in the lower atmosphere is essential for quantitative precipitation forecasting associated with HPEs (e.g., Behrendt et al. 2011, and references therein). However, the spatial and temporal variability of the WV field is very high, notably due to the fact that water coexists in three phases in the atmosphere. Furthermore, WV is an essential atmospheric meteorological and climatological variable but one that

Fig. 2 Frequency of monthly heavy precipitation events per year averaged over the period 1970–2019, highlighting in red the three most likely months (September–October–November)
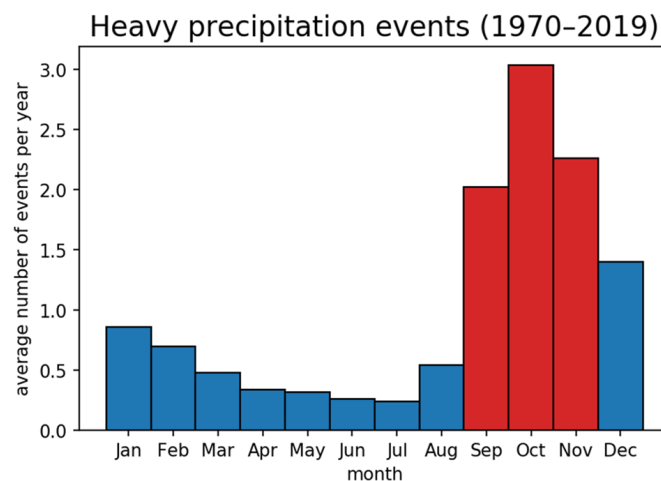
is still difficult to measure, as WV concentrations can vary by three orders of magnitude in the troposphere (typically from 10 g kg^{-1} near the surface to 0.01 g kg^{-1} in the upper troposphere).

Despite the undeniable contribution of microwave and infrared sounders onboard satellites, the assimilation of atmospheric WV-related observations from space still suffers from many limitations: (i) insufficient vertical resolution and accuracy below 3 km to describe precisely the very strong gradients of the moisture profiles observed in the lower troposphere (Chazette et al. 2014; Wulfmeyer et al. 2015), i.e., in a region key to understand convective initiation and the life cycle of heavy precipitating systems, and (ii) poor temporal sampling due to the fact that infrared and microwave sounders are embarked on meteorological satellites and cannot fully grasp the strong temporal variability of the WV field. Global Navigation Satellite System Radio Occultation (GNSS RO) observation from the satellite provide information on humidity at high vertical resolution (~ 100 m) down to 1 km above the Earth surface through the limb sounding. However, they suffer from 3 major drawbacks that currently limit their interest for data assimilation in the French limited area operational Application of Research to Operations at MEsoscale (AROME-France) model. First, the horizontal resolution of GNSS RO products is too coarse in the troposphere (a few hundred km, as they consist of path-integrated measurements) which is not appropriate for a fine scale NWP system (e.g., AROME-France has a 1.3-km grid size, and 90 vertical levels between the surface and 10 hPa, 33 of which are below 2000 m). Second, GNSS RO observations do not allow to observe the WV in the first kilometer of the atmosphere which is of great importance for monitoring moisture upstream of HPE hotspots. Third, there are very few GNSS RO observations available in each AROME-France 1-h assimilation window. For instance, the WV products currently assimilated from the plethora of microwave, and infrared sensors on polar orbiting spacecraft only represent 15% of the data ingested in the AROME-France operational model. In conclusion, currently available GNSS RO products are not suitable for improving HPE forecast through data assimilation.

Currently, 85% of the WV products assimilated in AROME-France come from surface stations, radar reflectivity near surface, aircraft, surface-based GNSS, and radiosoundings. However, only the latter can provide vertically resolved WV profiles above the surface, and this twice a day at best. Ground-based GNSS networks only provide integrated WV contents, whereas radars just provide indirect information on WV in precipitating systems. Surface stations only provide information on moisture very close to the surface, and aircrafts are currently very poorly equipped with moisture sensors.

Much of the AROME-France domain covers the Western Mediterranean which is usually upstream of convective systems, while observation systems used for assimilation (radar, GNSS, SYNOP stations) are primarily terrestrial. This is a major caveat because this is where the fast-evolving evaporation and air mass moistening processes take place that are crucial to understand and anticipate the development of HPEs downstream over southern France. This explains why HPEs in southern France are quite challenging to forecast with sufficient lead-time compatible with hazard warnings. Hence, the assimilation of data in the lower layers across the Western Mediterranean and available in near real-time will benefit prediction of HPEs at the mesoscale.

Unlike the instruments stated above, Raman water vapor lidars have the ability to measure water vapor profiles with high temporal and vertical resolution in the troposphere, making them ideal instruments for studying the evolution of water vapor in the troposphere in a fast evolving pre-HPE moist environment, provided that they are properly calibrated. Because they are rugged, easy-to-use instruments, Raman lidars are nowadays intensively used in the framework of measurement campaigns (e.g., Chazette

et al. 2016a, b; Lange et al. 2019; Di Girolamo et al. 2020, among others). Furthermore, an increasing number of Raman lidars around the world have been automated and are operated in a continuous manner for the purpose of numerical weather prediction and climate monitoring (Goldsmith et al. 1998; Reichardt et al. 2012; Dinoev et al. 2013). The major drawbacks associated with the use of Raman lidars is that their emitted laser cannot penetrate clouds or fog more than few tens of meters.

The project WaLiNeAs (Water vapor Lidar Network Assimilation) aims at bringing together French, Italian, German, and Spanish scientists concerned with improving HPEs forecasts around the Mediterranean. The members of the international consortium have joined forces to tackle the issues and challenges highlighted above. This paper describes the rationale of the WaLiNeAs initiative as well as the measurement and assimilation strategies central to project. It also highlights the key expectations from the program funded by several agencies in France, Italy, Germany, and Spain.

2 Challenges, objectives, and strategy

In the framework of the WaLiNeAs initiative, we aim at implementing an integrated prediction tool to enhance the forecast of HPEs in southern France, based on the sub-hourly assimilation of vertically resolved water vapor observations in the lower troposphere. A consortium of French, German, Italian, and Spanish research groups will deploy a network of 6 autonomous WV lidars for providing measurements with high vertical resolution and accuracy across the Western Mediterranean in fall 2022, closing critical gaps in lower troposphere WV observations by current operational networks and satellites. The proposed WV lidar network has been designed to account for all relevant WV sources and transport patterns known to contribute to the generation of HPEs in southern France. This network will aim at demonstrating the benefit of the assimilation of vertically resolved WV data in the forthcoming version of the operational 1.3-km grid size AROME-France NWP system (Brousseau et al. 2016; Montmerle et al. 2018) which enables ensemble-variational data assimilation for kilometer-scale prediction of heavy precipitation over southeastern France (Desroziers et al 2014). The ensemble-variational data assimilation system that will be available in 2022 will produce hourly analyses and will be operated with a rapid update cycle of assimilation of new observations at least each 15 min.

The WaLiNeAs project is a unique, innovative initiative that will allow for assimilating lidar-derived WV profiles in the lower troposphere with hectometer-scale vertical resolution in near real-time conditions for a continuous period of at least 3 consecutive months. The benefit of WaLiNeAs to the academic and operational communities is dual: advance knowledge of the processes at play in the life cycle of HPEs and enhance the predictability of HPEs in southern France at scales relevant for meteorological studies. Both aspects are dealt with in the framework of WaLiNeAs.

Finally, the breakthrough science that will be carried out in the framework of WaLiN-eAs concerns:

- Characterizing the predictability of HPEs and uncertainties in the prediction of the position, evolution, and the rainfall amount of the precipitating systems

- Assessing the role of water vapor distribution over the Western Mediterranean on the characteristics of the moist inflow (origin, evolution, pathways) feeding deep convection leading to HPEs
- Investigating the role of Mediterranean cyclones in the water cycle and HPEs in the Western Mediterranean
- Understanding the impact of elevated tropical moist plumes on the life cycle of deep convection and related HPEs
- Advancing knowledge on the role of dry intrusions on HPE in cases of frontal precipitation

The Raman lidar-derived WV data acquired in the framework of the WaLiNeAs project will be made available to Météo-France shortly after being acquired and will be assimilated up to 96 times per day. The focus of the project will be on providing high quality lidar-derived WV data in the first 3 km of the atmosphere, where no other observational technique can provide adequate data. To that respect, the two most cutting edge aspects of the project are as follows: (i) the near real-time processing and qualification of the lidar data (with WV profiles in the lower troposphere available every 15 min) and (ii) the proposed system (four-dimensional ensemble variational—4DEnVar (Desroziers et al. 2014)—with kilometric resolution) for the assimilation of qualified lidar data acquired in the lower troposphere, day and night.

Besides demonstrating the potential of WV lidar data assimilation in the AROME-France system, an ancillary objective of the project is also to show that Raman lidars can be left to operate continuously almost unattended for a period of at least 3 months. It is a prerequisite in the perspective of future/further deployment of operational Raman lidar systems meant to fill the observational gaps in water vapor in the lower troposphere of the current operational observation networks and satellites. This project can be considered a test bed for the concept of operational use of Raman lidars to be assimilated in a kilometer-scale NWP system. In all cases, in order for Raman lidars to be used more broadly in an operational context by meteorological services, reducing initial cost for installation, as well as maintenance frequency and cost, will be necessary.

The WaLiNeAs project builds on previous experience of the consortium partners, namely:

- Wulfmeyer et al. (2006) have assimilated airborne WV lidar observations acquired during a case study of the International H_2O project (IHOP) with the MM5 mesoscale NWP system and its four-dimensional variational (4D-Var) assimilation system.
- Grzeschik et al. (2008) extended the previous work to the assimilation of the WV observations provided by a ground-based network of Raman lidar systems.
- Bielli et al. (2012) used a pre-operational version of the Météo-France AROME-France NWP system (Seity et al. 2011) and its associated 3D-Var high-resolution assimilation system to evaluate the impact on the quantitative precipitation forecasts of the Convective and Orographically-driven Precipitation Study (COPS) airborne lidar observations collected during July 2007. A continuous assimilation cycle updated every 3 h was run over a month to provide the initial conditions of a sequence of 30 h forecasts carried out over 19 consecutive days.
- An ambitious effort has been conducted to assimilate water vapor mixing ratio observations from the airborne lidar LEANDRE 2 (Lidar Embarqué pour l'étude des Aérosols, des Nuages, de la Dynamique, du Rayonnement et des Espèces minoritaires, developed at LATMOS, Guyancourt, France) and the two ground-based Raman lidars located in

Menorca (the Weather Atmospheric LIdar (WALI) developed by LSCE, Gif-sur-Yvette, France) and Candillargues (the BASILicata Lidar (BASIL) developed by the University de la Basilicata, Potenza, Italy) as part of the reanalysis project aiming at assimilating the research observations collected during the HyMeX SOP1 (Richard et al., 2014; Fourrié et al. 2015, 2019, 2021). One month of ground-based Raman Lidar data and airborne lidar observations from approximately 20 flights were used in the reanalyses conducted by Fourrié et al. (2015, 2019, 2021).

- Thundathil et al. (2020) assimilated even both water vapor and temperature lidar data into the weather and research Forecast model on the convection-permitting scale applying a 3D-Var rapid update cycle and found that especially the assimilation of moisture results in a significant improvement of the model fields.

All these studies have shown encouraging results in terms of the impact of lidar-derived WV data assimilation on quantitative precipitation forecasts. However, they are based on a limited number of cases, date from a limited number of lidar systems (2 at most) and too short assimilation periods to allow for general conclusions to be drawn in a statistical sense regarding the impact of WV lidar observation assimilation on the quantitative precipitation forecast.

3 Implementation of WaLiNeAs

As part of the WaLiNeAs initiative, the WV Raman lidars will be operated continuously during 3 months starting early September 2022, to cover the period most propitious to HPEs in southern France (see Figs. 1 and 2). The duration of the operation is imposed by the necessity to have a long enough record in order for the WV profiles assimilated to have a significant impact on the moisture fields in the model forecasts. This long data set is crucial to assess statistical errors and to genuinely evaluate the benefit of Raman lidar water vapor data assimilation for operational NWPS.

This proposal accounts for the fact that Raman lidar technology has reached the level of maturity needed for unattended, continuous operation. Since more than 10 years, automated Raman lidar systems are operated in automatic mode at several observatories and research institutions (Goldsmith et al. 1998; Balin et al. 2004; Reichardt et al. 2012; Dinoev et al. 2013; Brocard et al. 2013; Leuenberger et al. 2020). Recently, also mobile systems became available which can be moved for field experiments: This is attested by the large data sets acquired by WALI and BASIL in the field during HyMeX SOP1 (Chazette et al. 2016a, b; Di Girolamo et al. 2020) or by the automated Raman lidar ARTHUS (Atmospheric Raman Temperature and Humidity Sounder, Lange et al. 2019) of the University of Hohenheim that operated from a ship for over a month during the EUREC[4]A campaign (Stevens et al., 2021). Figure 3 shows examples of time-height cross-sections of WV mixing ratio measured with WALI from 17 September to 28 October 2012 over Menorca, Spain (Fig. 3a), with BASIL over the 12-day time period from 0000 UTC on 17 October 2012 to 0000 UTC on 27 October 2012 over Candillargues, France (Fig. 3b), and with ARTHUS between 11 and 19 February 2020 on-board the research vessel Maria S. Merian over the Tropical Atlantic Ocean during the EUREC[4]A campaign (Fig. 3c). During HyMeX, WALI acquired 1000 h of WV mixing ratio profiles while being operated continuously during SOP1 and BASIL acquired over 600 h of data during the same period, but was operated continuously for 3 days at a time, with a sampling strategy focused on

Fig. 3 **a** Time-height cross-section of water vapor mixing ratio acquired during the HyMeX SOP1 with the WALI ▶ system of LSCE (~1000 h of data) from 17 September to 28 October 2012 over Menorca. The temporal resolution is 1 h, and a gliding average of 15 m was applied in the vertical. **b** Same as (**a**) but measured by BASIL over the 12-day time period from 0000 UTC on 17 October 2012 to 0000 UTC on 27 October 2012 during HyMeX. The temporal resolution is 5 min, and a gliding average of 150 m was applied in the vertical. **c** Same as (**a**) measured by ARTHUS collected between 11 and 19 February 2020 onboard the research vessel Maria S. Merian within the EUREC4A deployment. The temporal resolution is 10 s, and a gliding average of 50 m was applied in the vertical. The black/white areas correspond to missing data due to the limitation of the detection system during the daytime. The black areas correspond to missing data due to the limitation of the detection system during the daytime or lidar system operation stoppage

intensive observation periods only. During EUREC⁴A, ARTHUS collected useful data between 24 January and 19 February 2020, for approximately 620 h.

We aim at operating a dedicated network of 6 autonomous Raman WV lidars over the Western Mediterranean in locations shown in Fig. 4 to monitor the low-level and elevated moisture towards regions most likely to be hampered by HPEs in southern France in the fall as illustrated by Fig. 5, and also discussed by Ricard et al. (2012, see their Fig. 1), namely, Languedoc-Roussillon (LR), Cévennes-Vivarais (CV), southern Alps (SA), and Corsica (CO). Five of the six Raman lidar systems will be deployed specifically for WaLiNeAs. The 6th system is operating as a fixed, long-term monitoring station. We did not conduct a data targeting study (as in Majumdar 2016) to select the location of the 5 mobile lidar stations because of the lack of tools to perform such studies at the convective scale with the AROME model. The low-level moisture pathways are inspired by the composite analysis for 40 HPEs made by Ricard et al. (2012), for each of the target areas based on 700–1000-hPa integrated moisture flux and 925-hPa wind speed (see their Fig. 11). We are considering sites such as Barcelona and the islands of Menorca in Spain (to monitor the southerly flow upwind of LR, CV and SA, as well as elevated plumes from tropical Africa), Ajaccio in Corsica (to monitor the southerly flow upwind of CO and the easterly flow upwind of LR), Narbonne in the Aude Valley (to monitor the westerly flow from the Atlantic Ocean as well as event over LR), Cannes (to monitor the easterly flow upwind LR and SA), and Montpellier (to monitor the southerly flow and boundary layer upstream of CV). The main moisture patterns leading to HPE in southern France shown in Fig. 5 are also based on prior knowledge (Duffourg and Ducrocq 2013) and on the most recent work conducted on HyMeX SOP1, e.g., Chazette et al. (2016a), Di Girolamo et al. (2016), Duffourg et al. (2018), and Khodayar et al. (2018), among others. In an Observing System Simulation Experiment context, Yoshida et al. (2020) have shown that assimilating Raman lidar water vapor data on the windward side of heavy precipitation was likely to improve precipitation forecasts.

The worth of installing a Raman lidar system to monitor WV upstream of HPE-prone target area is further illustrated in Fig. 6. In Fig. 6a, we show the density distribution of all WV vertical profiles acquired with the WALI system installed in the city of La Ciutadella, in Menorca, during the entire HyMeX SOP1 (see Fig. 4 for the location of La Ciutadella). WALI was operating upstream of target areas in southern France, namely, LR, CV, and SA. The distribution highlights the spread in WV mixing ratio observed upstream of the Western Mediterranean coastline (color) around the mean profiles (black solid line). Figure 6a stresses out that in the course of the HyMeX SOP1, the highest occurrence of WV was found below 2 km amsl (where WV mixing ratio in on average comprised between 5 and 15 g kg^{-1}) and above 5 km amsl (where values are very low, less than 1.5 g kg^{-1}). The spread below 2 km amsl reaches 6 g kg^{-1} for WV mixing ratio values occurring more than 30%, and more than doubles for WV mixing ratio values occurring more than 10% of the time. Figure 6b and

(a) WALI

(b) BASIL

(c) ARTHUS
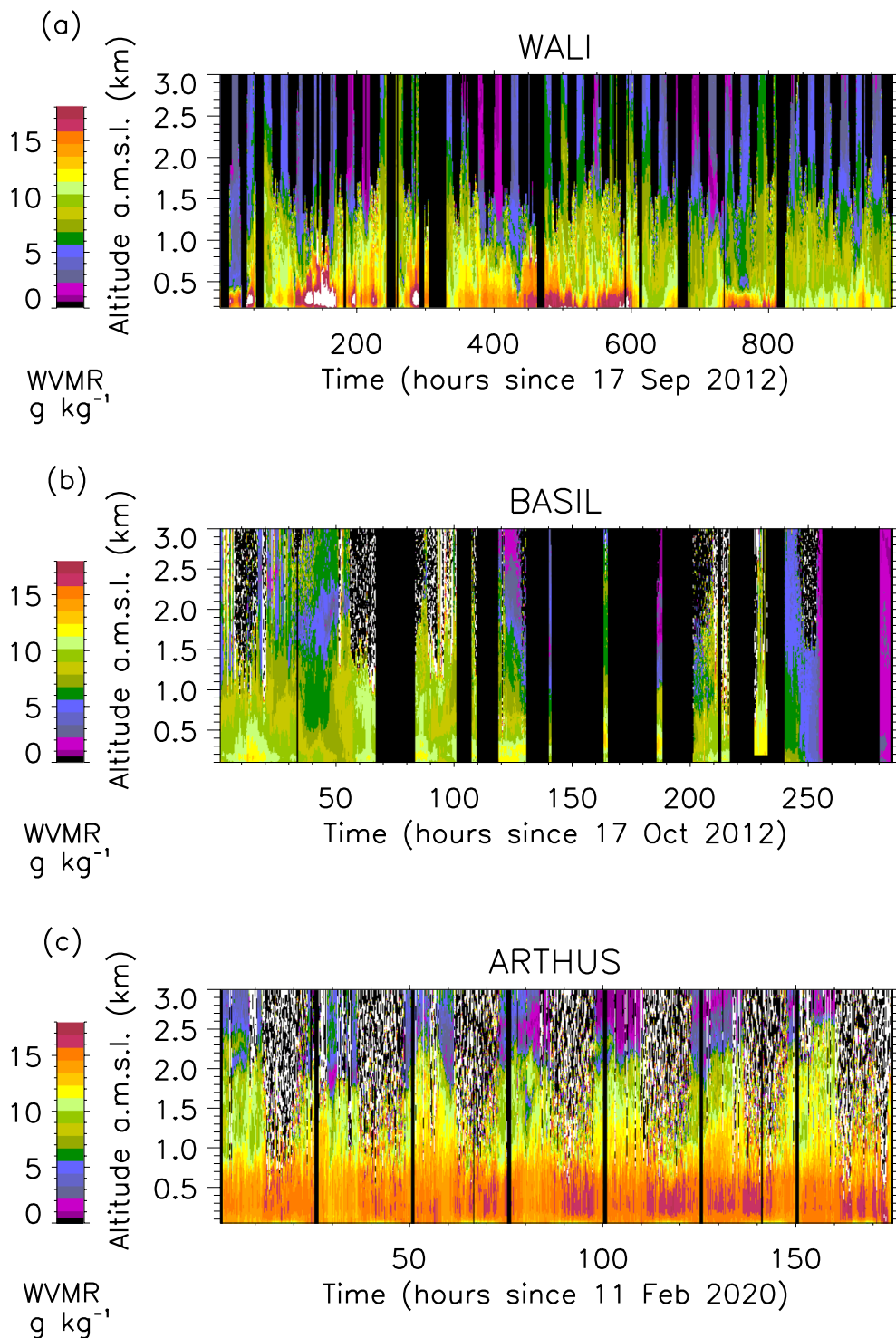
c show the density distribution of WV mixing ratio in HPE-free conditions and in HPE conditions, respectively, for two 5-day periods. The HPE-free period selected is from 2 to 6 October 2012, during which no HPEs were observed in the Western Mediterranean basin.
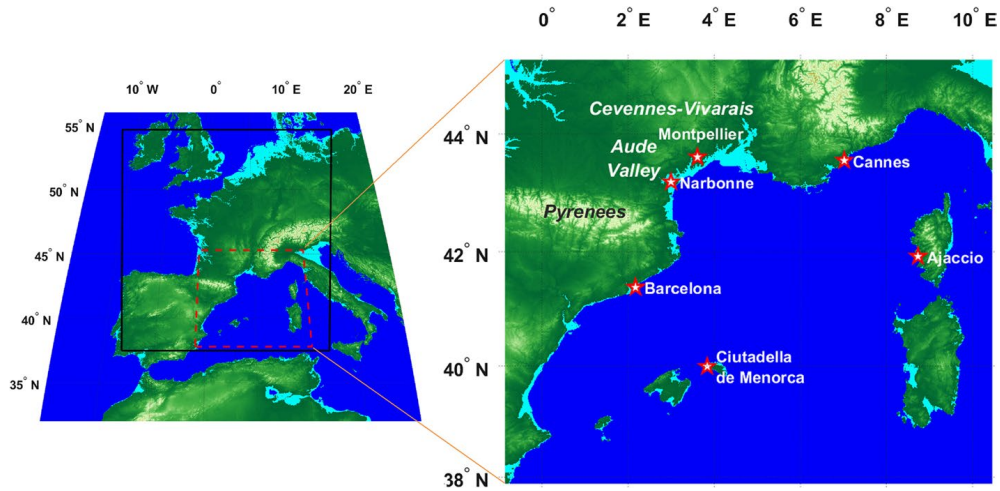
Fig. 4 Left: orography of the AROME-France domain (area delimited by the black contour). Right: zoom on the region of operation of the Raman Lidar Network with the location (city) of each lidar highlighted with red and white stars

For the HPE conditions, we selected the period from 7 to 11 October 2012, up to 4 days ahead of HPE events observed in the CV, namely, IOP12a on 11–12 October (Khodayar
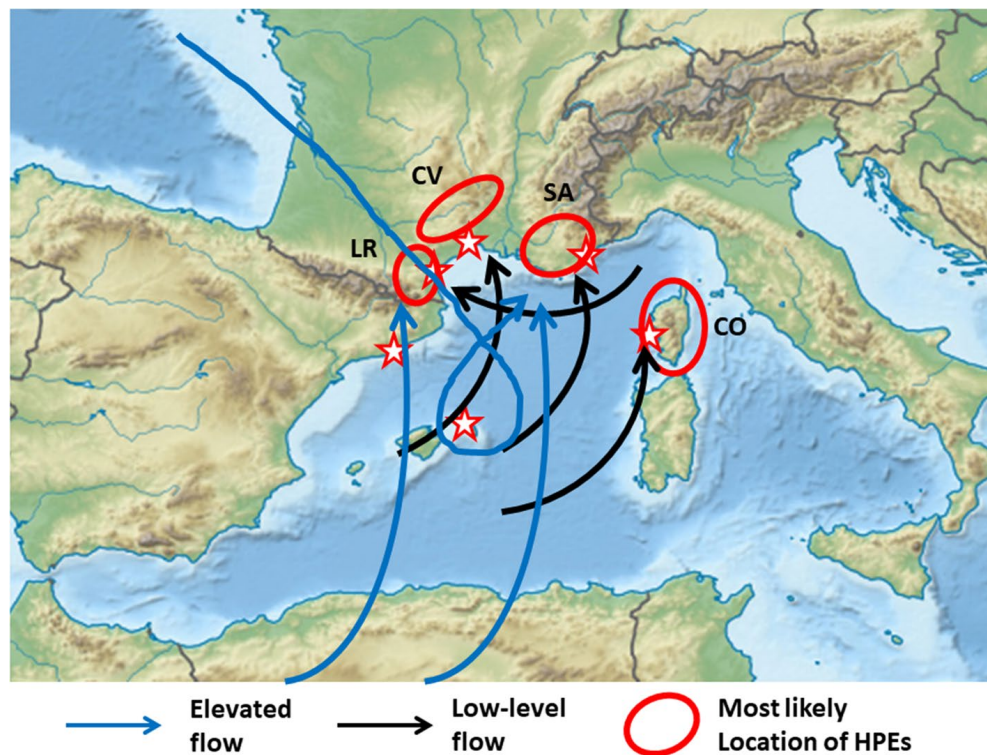
Fig. 5 Regions most likely to be impacted by HPEs (red circles) together with main flow patterns in the low levels (black arrows) and in altitude (2–4 km, blue arrows). LR is Languedoc-Roussillon, CV is, Cévennes-Vivarais, SA is southern Alps, and CO is Corsica. The red and white stars indicate the location of Raman Lidar systems in the WaLiNeAs Network
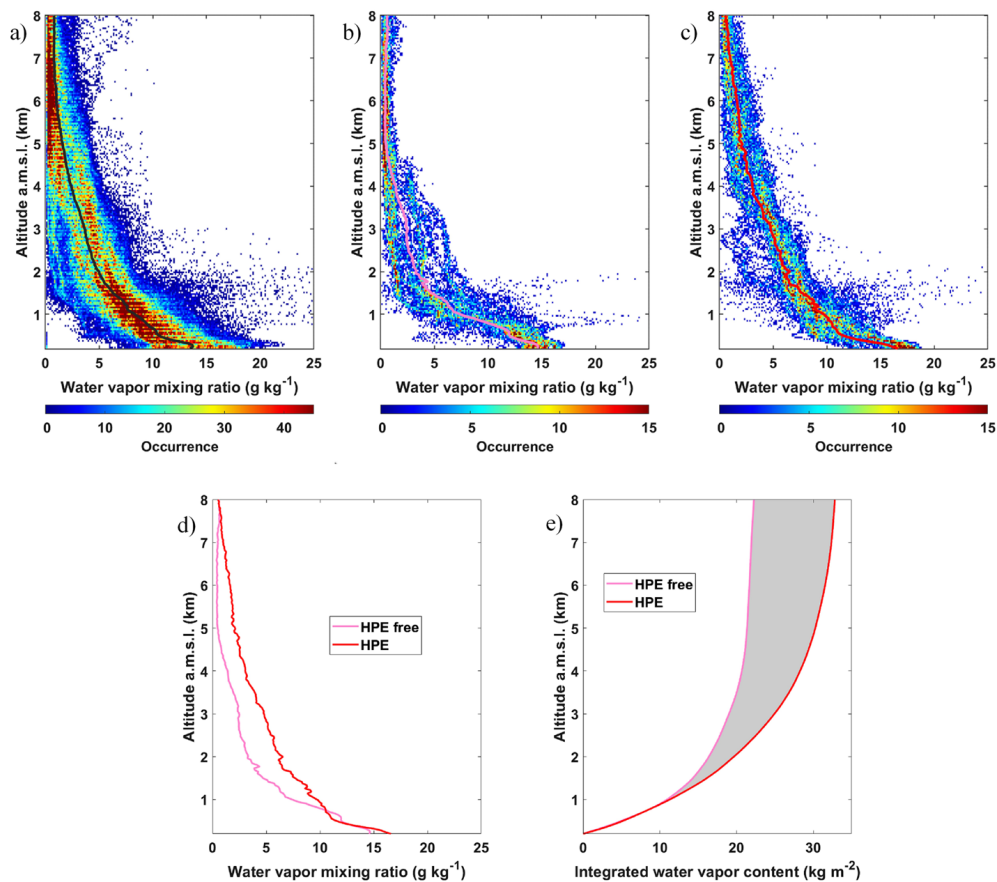
Fig. 6 a Density distribution of all WV vertical profiles acquired with WALI during the HyMeX SOP1 (see Fig. 5a). The black solid line indicates the barycenter of the distribution at each altitude bin. The occurrence of water vapor mixing ratio values is color code. **b** Same as (a), but for the profiles acquired in HPE-free conditions from 2 to 6 October 2012. The pink solid line indicates the barycenter of the distribution at each altitude bin. **c** Same as (b), profiles acquired in HPE conditions from 7 to 11 October 2012. The red solid line indicates the barycenter of the distribution at each altitude bin. **d** Average vertical distribution of WV mixing ratio as a function of altitude in the HPE-free case (pink solid line) and the HPE case (red solid line). **e** Integrated water vapor content as a function of altitude for the HPE-free case (pink solid line) and the HPE case (red solid line). The integrated content is computed at the integral of the lidar-derived specific humidity profiles from the ground upward. The shaded area highlights the difference between the HPE-free and HPE distribution

et al., 2018) and IOP 13 on 14–16 October (Duffourg et al. 2018). Above 2 km amsl, the WV mixing ratio distribution for the no HPE period (Fig. 6b) is skewed towards small values than in the HPE period (Fig. 6c) up to 8 km amsl. The average WV mixing ratio profiles, computed as the barycenters of the distributions at each altitude bin (pink and red solid lines, for the HPE-free and HPE periods, respectively), overlain in Fig. 6b and c, are also shown in Fig. 6d. They highlight that, between 0.75 and 8 km amsl, the troposphere is significantly moister in HPE conditions than in HPE-free conditions. The marine boundary layer is also slightly more moist in HPE conditions below 300 m amsl. The integrated water vapor content profile, computed at the integral of the lidar-derived specific humidity profiles from the ground upward, in HPE-free conditions (pink solid line) and in HPE conditions (red solid line), is shown in Fig. 6e. Over the depth of the lower troposphere, i.e., below 8 km amsl, the difference between HPE and HPE-free conditions reaches 10 kg m^{-2}. The integrated water

vapor content in HPE conditions (~ 33 kg m^{-2}) is comparable to those derived from global positioning system by Khodayar et al. (2018) during IOP 12, even though slightly smaller due to the fact that we are only considering partial columns. The above analysis clearly suggests that Raman lidar-derived WV mixing ratio profiles collected in Menorca in a period up to 4 days ahead of HPEs in the CV area are significantly moister than for a 4-day period when no HPE events are observed in southern France, thereby providing data worthy of assimilation in a HPE forecasting context.

The 5 mobile Raman lidar systems that will be operating in the field are rugged and transportable systems that have operated in many locations in recent years:

- The Weather Atmospheric LIdar (WALI, Chazette et al. 2014) developed at LSCE, which was involved in the SOP1 of HyMeX (Chazette et al. 2016a,b; Di Girolamo et al. 2020) or during the Pollution in the ARCtic System (PARCS) project (Totems et al. 2019) and recently during the Lacustrine-Water vApor Isotope inVentory Experiment (L-WAIVE) project (Chazette et al. 2021)
- The Airborne Lidar for Atmospheric Studies (ALiAS, Chazette et al. 2012, 2017, 2019, 2020) developed at LSCE
- The Lidar for Automatic Atmospheric Surveys using Raman Scattering (LAASURS; Baron et al. 2020) developed at LSCE
- The University of BASILicata ground-based Raman Lidar system (BASIL), which was involved in HyMeX (Di Girolamo et al., 2009, 2016, 2017, 2020; Stelitano et al. 2019)
- The Atmospheric Raman Temperature and Humidity Sounder (ARTHUS, Lange et al. 2019) of the University of Hohenheim

The fixed Raman lidar system that will operate in Barcelona is the lidar system of Universitat Politècnica de Catalunya (UPC, Muñoz-Porcar et al. 2018, 2021) which is in operation since 1993 and is part of the European Aerosol Research Lidar Network (EARLINET) since 2000 and of the European Research Infrastructure for the observation of Aerosol, Clouds, and Trace Gases (ACTRIS) since 2011.

For details on the instruments (emission unit, reception unit, spatio-temporal sampling strategy, etc.), the reader is referred to the existing body of literature that are listed above and are summarized in Table 1. LAASURS and ALIAS will be upgraded to WV Raman lidar prior to the start of the fall 2022 field campaign.

The lidar data will be collected, disseminated, and monitored in real time, as would be done in an operational context. Since it is unrealistic to set up a full real-time operational NWP system dedicated to a single observing system (the cost of real-time high-performance computing with human supervision far exceeds funds allotted to WaLiNeAs), the lidar data assimilation will be evaluated in a quasi-operational environment, i.e., with the same tools and data, but without the associated resources necessary for real-time operations. This framework will allow deriving results that will hold true for an operational context, at an economic cost.

4 Project organization and structure

The project is organized around 4 scientific tasks aimed at developing an innovative integrated forecasting tool in order to pave the way towards an operational, breakthrough HPE-related hazard warning capability for southern France and Corsica as well as work on the cutting-edge science linked with the development of such an innovative tool.

Table 1 Main technical characteristics of the lidar instruments

	WALI	ARTHUS	BASIL	UPC/EARLINET
Laser	Nd:YAG	Nd:YAG	Nd:YAG	Nd:YAG
Pulse energy and initial wavelength	60 mJ at 355 nm	100 mJ at 355 nm	500 mJ at 355 nm 300 mJ at 532 nm 400 mJ at 1064 nm	60 mJ at 355 nm 130 mJ at 532 nm 130 mJ at 1064 nm
Frequency	20 Hz	200 Hz	20 Hz	20 Hz
Reception channels	Elastic total 354.67 nm Elastic −354.67 nm Elastic ⊥ 354.67 nm Roto-vib. Raman N_2 386.63 nm Roto-vib. Raman H_2O 407.5 nm Rotot. Raman (N_2&O_2)—Low J 354.3 nm Rotot. Raman (N_2&O_2)—High J 352.9 nm	Elastic, 355 nm Rot. Ram. 1, CWL selectable Rot. Ram. 2, CWL selectable Raman H_2O, 408 nm	Elastic total 354.67 nm Elastic ∥354.67 nm Elastic ⊥ 354.67 nm Elastic total 532.05 nm Elastic ∥532.05 nm Elastic total 1064.1 nm Roto-vib. Raman N_2 386.63 nm Roto-vib. Raman H_2O 407.5 nm Roto-vib. Raman N_2 607.4 nm Rotot. Raman (N_2&O_2)—Low J 354.3 nm Rotot. Raman (N_2&O_2)—High J 352.9 nm	Elastic total 354.7 nm Elastic ⊥ 354.7 nm Elastic total 532.1 nm Elastic ⊥ 532.1 nm Elastic total 1064.1 nm Rot. Raman (N_2&O_2) 353.9 nm Rot.-vib. Raman H_2O 407.5 nm Rot. Raman (N_2&O_2) 530.2 nm
Receiver diameter	~15 cm	~40 cm	45 cm	40 cm
Field of view	~2.3 mrad	<1 mrad	0.45 mrad (FWHM)	1.2 mrad
Lowest useful range bin	~180 m	~50 m	~500 m (~100 for 532.05 nm total &∥ and 1064.1 nm total)	~400 m after overlap correction
Detector	Photomultiplier tubes	Photomultiplier tubes	Photomultiplier tubes	Photomultiplier tubes
Filter bandwidths	0.2–0.3 nm	0.3–0.5 nm	0.2–1.0 nm	0.8–1.0 nm
Vertical sampling	0.75 m (analog) 15 m (photon counting)	7.5 m (analog) 7.5 m (photon counting)	7.5 m (analog and photon counting)	3.75 m (analog and photon counting)
Vertical resolution	~30 m	~30 m	7.5–150 m	≥3.75 m (adjustable)
Acquisition system	PXI technology at 200 MHz	LICEL GmbH	Licel at 200 MHz, FastComTec at 250 MHz	LICEL GmbH

4.1 WP1 Field campaign

WP1 (field campaign) will drive the necessary experimental/instrumental deployments to achieve the acquisition of WV profiles and the near real-time transmission of the data to the French Weather Service. This will include the upgrade of the Raman lidar systems to be deployed, as well as the definition of the exact location of the implementation sites and complementary instruments, the data transmission protocols and data quality assurance, as well as the characterization of system calibration. WP1 is organized around 3 main activities.

4.1.1 Upgrade and preparation of lidar systems

The WV-Raman upgrades consist in adding a H_2O-Raman channel on each lidar and the development of an acquisition chain. The lasers will be upgraded in energy to improve the signal to noise ratio. This will ensure that the output energy of the systems as well as their performances (precision, systematic error on WV profiles, etc.) will be same as the operational WALI system. Note that these two mini-lidar systems will be autonomous and connected via Internet. Prior to the field campaign, a thorough intercomparison of WV profiles between the operational WALI Raman lidar and the upgraded Raman systems (LAASURS and ALIAS) will be undertaken.

4.1.2 Real-time data processing and data transmission

For real-time purpose, the six Raman lidars will deliver 2–4 profiles per hour which are averaged over 15 min with vertical resolution of 100 m and a targeted root-mean-square-error of 0.4 g kg^{-1} in the first 3 km, day and night. Performances are expected to exceed these target values during the night for all systems (WALI, ALIAS, LAASURS, BASIL, ARTHUS, and UPC/EARLINET). In addition to the water vapor mixing ratio profiles, the statistical uncertainties of these profiles as well as the atmospheric variance determined with the auto-covariance technique (Lenschow et al. 2000) will be provided. It is a significant advantage of the lidar technique that also these error profiles can be determined and consequently be used for the data assimilation.

The WV profile acquisition, processing, and transmission sequence is anticipated to span over 25 min (Fig. 7). For the analysis at XXXX UTC, the assimilation system will
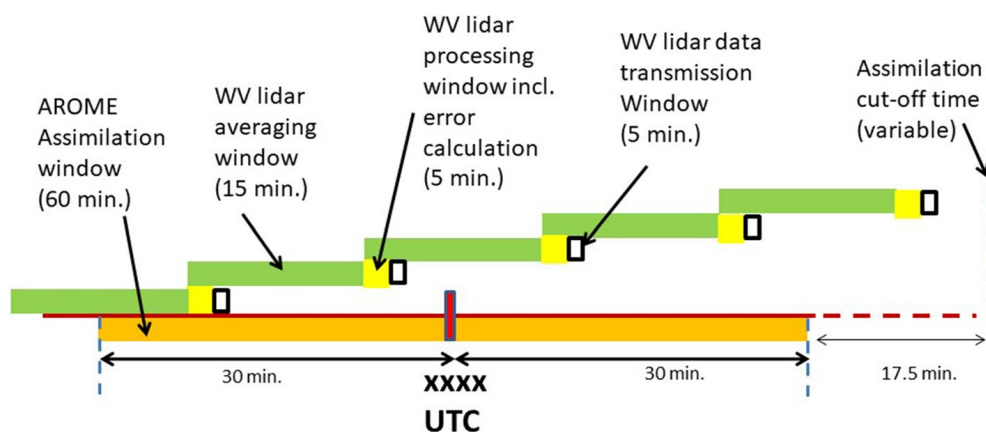
Fig. 7 Timeline of the WV lidar data transmission sequence from the operating sites to Météo-France for assimilation of near real-time WV profiles in the AROME-France NWP system

ingest observations made between XXXX UTC—30 min and XXXX UTC + 30 min. Every 15 min between XXXX UTC—30 min and XXXX UTC + 30 min, the WV lidar data (resulting from a 15-min average) will be processed, and errors calculated within a 5-min window and then transmitted in an additional 5-min window (Fig. 7). There will be some overlap in the processing of each profile: As soon as the 15-min measurement period (in green in Fig. 7) is over, there is both the start of the next 15-min measurement period and the start of the processing (5 min, in yellow) + transmission (5 min, in white framed in black) of the one that has just ended. The measuring periods for an assimilation window would be [XXXX–30 min; XXXX + 30 min] , [XXXX–37.5 min; XXXX–22.5 min], [XXXX–22.5 min; XXXX–7.5 min], [XXXX–7.5 min; XXXX + 7.5 min], [XXXX + 7.5 min; XXXX + 22.5 min], [XXXX + 22.5 min; XXXX + 37.5 min]. So, for a cut-off time greater than or equal to XXXX + 47.5 min, there are 4 profiles for the corresponding analysis. The end of the last 5-min transmission window should occur before the so-called cut-off time. The cut-off time is the time after the hour of analysis until which one waits for the observations to arrive before starting the calculations. In the current AROME-France assimilation system, this cut-off is not constant and varies according to the time of day (it varies from T + 20 min. to T + 3h15). Depending on the cut-off time, up to 6 WV lidar profiles will be assimilated in each hourly 4DEnVar analysis. This sequence is subject to adaptation depending on the evolution of the high computational performances at Météo-France and performances of the WV lidars in the field.

4.1.3 WV lidar intercomparison

After the conduct of the field campaign, the consortium will also deliver a consistent, self-coherent, and validated WV data set of lidar profiles, including uncertainties at high spatio-temporal resolution for data assimilation experiments. This effort will include intercomparison of WV lidar-derived profiles between the 3 operational systems WALI, BASIL, and ARTHUS. Comparison between WALI, LAASURS, and ALIAS will be conducted at LSCE before and after the field campaign. The long-term stability and calibration of the WV lidar systems will also be monitored throughout the field campaign using a mobile radiosounding unit that will be operated alongside each of the 6 Raman lidar systems at the beginning, mid-way through, and at the end of the campaign.

4.2 WP2 Data monitoring

WP2 (data monitoring) aims at Raman lidar data assimilation in AROME-France. This includes evaluating and optimizing how the observations can be used in the data assimilation system. For this, the ability of the model to simulate the physical quantity that is observed must be ensured. In practice, this is done through the "monitoring" of observations, i.e., the computation of observation-minus-background statistics, where "background" refers to short-term forecasts that will be blended with observations during the data assimilation process. Such a monitoring requires some preliminary work. To compare observations and the background, a common physical space has to be chosen, which needs to be close to that of the raw measurements to avoid the introduction of retrieval errors, while lending itself to the simulation from the model prognostic variables. The choice of this physical space in terms of geometry, physical quantity, and observation processing (time–space averaging, filtering) will be carefully assessed based on the horizontal and vertical resolutions of the current AROME-France version (e.g., 1.3 km) and the

15-min period of observations use. The current version of AROME-France uses a three-dimensional variational (3D-Var) algorithm to assimilate weather data. A four-dimensional ensemble-variational (4DEnVar) data assimilation system will be available in 2022 to replace the current 3D-Var data assimilation system. It will be used in place of the 3D-Var data assimilation system for the WaLiNeAs project.

A near real-time monitoring will be set up during the field campaign scheduled in the fall of 2022 that will enable the evaluation of the statistical consistency between the observations and the model background. Such a real-time monitoring is usually performed at operational weather forecasting centers to detect gross errors such as hardware failure, calibration drift, or transmission losses. The monitoring performed in the framework of WaLiNeAs will ensure that the data are collected as expected and, if needed, allow for corrective action to be taken immediately so as to minimize any data loss.

After the field campaign, observation-minus-background statistics will be performed on the consistent, self-coherent, and validated data set of lidar profiles, once it is available (WP1). The objective is to compare to which extent the real-time and post-processed lidar data differ with respect to the model. Depending on the results, observation-minus-background biases will be removed so as to comply with the data assimilation technique assumptions. The resulting lidar data sets and observation operator will be used in WP3 for the data assimilation experiments.

4.3 WP3 Lidar data assimilation

WP3 (lidar data assimilation) is focused on the post-campaign work on the assessment of the lidar data impact in the assimilation scheme. Since the objective of the project is to prove the feasibility and benefit of assimilating lidar data in an operational context, the assessment of the impact of the lidar data assimilation will be carried out by performing data assimilation experiments with the AROME-France system. The proven methodology of observing system experiments will be used. It consists in running two different experiments: the reference experiment, while the data assimilation experiment will additionally assimilate lidar data. So, the reference experiment will already assimilate all routinely available observations, and thus the impact of lidar observations will translate the ability of this new observing system to complement existing observing systems. With this methodology, the impact of the lidar data assimilation is simply obtained by contrasting the weather forecasts obtained by each of the two experiments with respect to an independent observational data set (e.g., precipitation amounts measured by rain gauges).

Data assimilation experiments will be performed for the two data sets prepared in WP1 and WP2: a set of lidar data collected in real time and a set of consistent, self-coherent, and validated lidar data. The real-time data set will provide baseline results that will show which benefit can readily be obtained with the current real-time data processing. The post-processed data set will show to which extent additional processing may improve the quality of the weather forecasts. This will likely lead to recommendations on the lidar data processing for future operational exploitation in NWP systems.

4.4 WP4 HPE-related science

WP4 (HPE-related science) is expected to provide an improved representation of the highly variable spatial–temporal distribution of WV in the AROME-France analyses from the advanced data assimilation implemented in this project that will in turn lead to an overall

improvement of the complex thermodynamical and dynamical processes controlling the life cycle of HPEs. We will investigate the impact of the WV profiles processing (real-time vs validated lidar data) on (i) the WV distribution over the Western Mediterranean in the AROME-France model and (ii) the prediction of the position, evolution, and the rainfall amount of the precipitating systems and HPEs encountered during the 3-month field campaign. The results will also be compared to AROME-France reference simulations in which lidar-derived WV profiles are not assimilated to further emphasize the worth (or lack thereof) of assimilating such data in the French NWP system. In addition, with a similar approach, we will examine what is gained in terms of advancing our knowledge of complex processes pertaining to the characteristics of the moist inflow (origin, evolution, pathways) feeding deep convection leading to HPEs. We will also study the impact of dry intrusions from the upper troposphere and moist tropical plumes on HPEs encountered during the campaign.

5 Outlook

This project aims at the development of all-weather, unattended, rugged, and operational Raman lidar systems for smart monitoring of the environment, and WV in particular. The WaLiNeAs project aims at developing the test bed of an integrated prediction tool, coupling network measurements of WV profiles, and a weather forecast model to precisely estimate precipitable water upstream of an event up to 48 h in advance in southern France. This project is highly innovative and will lay the foundation for a future integrated warning tool aiming to prevent natural hazards associated with HPEs as often experienced along the Mediterranean coastline. Once the proof of concept is validated in the framework of the WaLiNeAs project, similar integrated tools may be applied in other parts of the world to avoid similar natural hazards.

The highest risk for the project lies with the meteorology and the possible lack of heavy precipitation events during the fall of 2022. However, the length of the field campaign (3 months) is the best insurance that extreme events will happen somewhere in northwestern Mediterranean. Nevertheless, on average, $\sim$ 7 HPEs (daily rainfall > 150 mm) occur every year between September and November (Ricard et al. 2012, Fig. 2). Furthermore, even in the case of lower than average HPE activity in southern France, the network data will be beneficial to the AROME-France forecasts, and a positive impact is expected on average skill scores and in the case of southern maritime inflow situation. Furthermore, experience learned from the lidar data processing in near real time and assimilation in NWP systems will still be extremely valuable to make recommendation on the use of WV lidars for future operational NWP systems. In all cases, the uttermost important objective of the project is to contribute to increase the accuracy of forecasts of quantitative precipitation in order to satisfy the societal demands in terms of amount, timing, and basin-specific locations of rainfall and flash flooding.

Declarations

Conflict of interest The authors declare no competing interests.

References

Balin I, Serikov I, Bobrovnikov S, Simeonov V, Calpini B, Arshinoc Y, van den Bergh H (2004) Simultaneous measurement of atmospheric temperature, humidity, and aerosol extinction and backscatter coefficients by a combined vibrational–pure-rotational Raman lidar. Appl Phys B 79:775–782

Baron A, Chazette P, Totems J (2020) Remote sensing of two exceptional winter aerosol pollution events and representativeness of ground-based measurements. Atmos Chem Phys 20:6749–6768. https://doi.org/10.5194/acp-20-6749-2020

Behrendt A, Pal S, Aoshima F, Bender M, Blyth A, Corsmeier U, Cuesta J, Dick G, Dorninger M, Flamant C, Di Girolamo P, Gorgas T, Huang Y, Kalthoff N, Khodayar S, Mannstein H, Träumner K, Wieser A, Wulfmeyer V (2011) Observation of convection initiation processes with a suite of state-of-the-art research instruments during COPS IOP8b. Q J Roy Meteorol Soc 137(S1):81–100

Bielli S, Grzeschik M, Richard E, Flamant C, Champollion C, Kiemle C, Dorninger M, Brousseau P (2012) Assimilation of water vapour airborne lidar observations: impact study on the COPS precipitation forecasts. Q J Roy Meteorol Soc 138:1652–1667

Brocard E, Philipona R, Haefele A, Romanens G, Mueller A, Ruffieux D, Simeonov V, Calpini B (2013) 2013: Raman Lidar for Meteorological Observations, RALMO – Part 2: validation of water vapor measurements. Atmos Meas Tech 6:1347–1358. https://doi.org/10.5194/amt-6-1347-2013

Brousseau P, Seity Y, Ricard D, Léger J (2016) Improvement of the forecast of convective activity from the AROME-France system. Q J R Meteorol Soc 142(699):2231–2243

Caumont O, Mandement M, Bouttier F, Eeckman J, Lebeaupin Brossier C, Lovat A, Nuissier O, Laurantin O (2021) The heavy precipitation event of 14–15 October 2018 in the Aude catchment: a meteorological study based on operational numerical weather prediction systems and standard and personal observations. Nat Hazards Earth Syst Sci 21:1135–1157

Chazette P, Dabas A, Sanak J, Lardier M, Royer P (2012) French airborne lidar measurements for Eyjafjallajökull ash plume survey. Atmos Chem Phys 12:7059–7072. https://doi.org/10.5194/acp-12-7059-2012

Chazette P, Marnas F, Totems J, Shang X (2014) Comparison of IASI water vapor retrieval with H_2O-Raman lidar in the framework of the Mediterranean HyMeX and ChArMEx programs. Atmos Chem Phys 14:9583–9596. https://doi.org/10.5194/acp-14-9583-2014

Chazette P, Flamant C, Raut J-C, Totems J, Shang X (2016a) Tropical moisture enriched storm tracks over the Mediterranean and their link with intense rainfall in the Cevennes-Vivarais area during HyMeX. Q J Roy Meteorol Soc 142(S1):320–334. https://doi.org/10.1002/qj.2674

Chazette P, Flamant C, Shang X, Totems J, Raut J-C, Doerenbecher A, Ducrocq V, Fourrié N, Bock O, Dorenbecher A, Cloché S (2016b) multi-instrument and multi-model assessment of atmospheric moisture variability over the Western Mediterranean during HyMeX. Q J R Meteorol Soc 142(S1):7–22. https://doi.org/10.1002/qj.2671

Chazette P, Totems J, Baron A, Flamant C, Bony S (2020) Trade-wind clouds and aerosols characterized by airborne horizontal lidar measurements during the EUREC⁴A field campaign, *Earth Syst. Sci. Data*. Earth Syst Sci Data 12:2919–2936

Chazette P, Totems J, Shang X (2017) Atmospheric aerosol variability above the Paris Area during the 2015 heat wave - Comparison with the 2003 and 2006 heat waves. Atmos Environ 170. https://doi.org/10.1016/j.atmosenv.2017.09.055

Chazette P, Totems J, Shang X (2019) Transport of aerosols over the French Riviera – link between ground-based lidar and spaceborne observations. Atmos Chem Phys 19:3885–3904. https://doi.org/10.5194/acp-19-3885-2019

Chazette P, Flamant C, Sodemann H, Totems J, Monod A, Dieudonné E, Baron A, Seidl A, Steen-Larsen H-C, Doira P, Durand A, Ravier S (2021) The lacustrine-water vapor isotope inventory experiment L-WAIVE. Atmos Chem Phys 21:10911–10937

Desroziers G, Camino J-T, Berre L (2014) 4DEnVar: link with 4D state formulation of variational assimilation and different possible implementations. Q.J.R. Meteorol Soc 140:2097–2110. https://doi.org/10.1002/qj.2325

Di Girolamo P, De Rosa B, Flamant C, Summa D, Bousquet O, Chazette P, Totems J, Cacciani M (2020) Water vapour mixing ratio and temperature intercomparison results in the framework of the

hydrological cycle in the Mediterranean experiment – special observation period 1. Bull Atmos Sci Tech 1:133–153

Di Girolamo P, Cacciani M, Summa D, Scoccione A, De Rosa B, Behrendt A, Wulfmeyer V (2017) Characterisation of boundary layer turbulent processes by the Raman Lidar BASIL in the frame of HD(CP)2 observational prototype experiment. Atmos Chem Phys 17:745–767. https://doi.org/10.5194/acp-17-745-2017

Di Girolamo P, Flamant C, Cacciani M, Richard E, Ducrocq V, Summa D, Stelitano D, Fourrié N, Said F (2016) Observation of low-level wind reversals over the Gulf of Lion and their impact on the water vapour variability. Q J R Meteorol Soc 142(S1):153–172. https://doi.org/10.1002/qj.2767

Di Girolamo P, Summa D, Sabatino D, Ferretti R, Faccani C (2009) Multiparameter Raman Lidar measurements for the characterization of a dry stratospheric intrusion event. J Atmos Oceanic Technol 26:1742–1762. https://doi.org/10.1175/2009JTECHA1253.1 (ISSN: 0739-0572)

Dinoev T, Simeonov VB, Arshinov Y, Bobrovnikov S, Ristori P, Calpini B, Parlange M, van den Bergh H (2013) Raman Lidar for meteorological observations, RALMO. Part 1: Instrument description. Atmos Meas Tech 6:1329–1346

Ducrocq V, Nuissier O, Ricard D, Lebeaupin C, Thouvenin T (2008) A numerical study of three catastrophic precipitating events over southern France. II: Mesoscale triggering and stationarity factors. Q J R Meteorol Soc 134:131–145. https://doi.org/10.1002/qj.199

Ducrocq V et al (2014) HyMeX-SOP1, the field campaign dedicated to heavy precipitation and flash flooding in the northwestern Mediterranean. Bull Am Meteo Soc 95:1083–1100

Duffourg F, Nuissier O, Ducrocq V, Flamant C, Chazette P, Delanoë J, Doerenbecher A, Fourrié N, Di Girolamo P, Lac C, Legain D, Martinet M, Saïd F, Bock O (2016) Offshore deep convection initiation and maintenance during HyMeX IOP 16a heavy precipitation event. Q J R Meteorol Soc 142(S1):259–274. https://doi.org/10.1002/qj.2725

Duffourg F, Lee K-O, Ducrocq V, Flamant C, Chazette P, Di Girolamo P (2018) Role of moisture patterns in the backbuilding formation of HyMeX IOP13 Heavy Precipitating Systems. Q J R Meteorol Soc 144:291–303

Duffourg F, Ducrocq V (2013) Assessment of the water supply to Mediterranean heavy precipitation : a method based on finely designed water budgets. Atmos Sci Lett 14(3):133–138

Fourrié N, Nuret M, Brousseau P, Caumont O (2021) Data assimilation impact studies with the AROME-WMED reanalysis of the first special observation period of the Hydrological cycle in the Mediterranean Experiment. Nat Hazards Earth Syst Sci 21:463–480. https://doi.org/10.5194/nhess-21-463-2021

Fourrié N, Nuret M, Brousseau P, Caumont O, Doerenbecher A, Wattrelot E, Moll P, Bénichou H, Puech D, Bock O, Bosser O, Chazette P, Flamant C, Di Girolamo P, Richard E, Saïd F (2019) The AROME-WMED reanalyses of the first special observation period of the hydrological cycle in the Mediterranean experiment. Geophys Model Dev 12:2657–2678

Fourrié N, Bresson É, Nuret M, Jany C, Brousseau P, Doerenbecher A, Kreitz M, Nuissier O, Sevault E, Bénichou H, Amodei M, Pouponneau F (2015) AROME-WMED, a real-time mesoscale model designed for the HyMeX special observation periods. Geosci Model Dev 8:1919–1941

Gaume E et al (2009) A compilation of data on European flash floods. J Hydrol 367:70–78

Grzeschik M, Bauer H-S, Wulfmeyer V, Engelbart D, Wandinger U, Mattis I, Althausen D, Engelmann R, Tesche M, Riede A (2008) Four dimensional variational data analysis of water vapor Raman lidar data and their impact on mesocale forecasts. J Atmos Ocean Technol 25:1437–1453

Goldsmith JE, Blair FH, Bisson SE, Turner DD (1998) Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols. Appl Opt 37:4979–4990

Gustafsson N, Janjić T, Schraff C, Leuenberger D, Weissmann M, Reich H, Brousseau P, Montmerle T, Wattrelot E, Bučánek A, Mile M, Hamdi R, Lindskog M, Barkmeijer J, Dahlbom M, Macpherson B, Ballard S, Inverarity G, Carley J, Alexander C, Dowell D, Liu S, Ikuta Y, Fujita T (2018) Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. Q J R Meteorol Soc. https://doi.org/10.1002/qj.3179

Khodayar S, Czajka B, Caldas-Alvarez A, Helgert S, Flamant C, Di Girolamo P, Bock O, Chazette P (2018) Multi-scale observations of moisture feeding heavy precipitating systems in the Northwestern Mediterranean during HyMeX IOP12. Q J R Meteorol Soc 144:2761–2780

Kwon I-H, English S, Bell W, Potthast R, Collard A, Ruston B (2018) Assessment of progress and status of data assimilation in numerical weather prediction. Bull Amer Meteorol Soc 98:ES75–ES79

Lange D, Behrendt A, Wulfmeyer V (2019) Compact operational tropospheric water vapor and temperature Raman lidar with turbulence resolution. Geophys Res Lett 46:14844–14853. https://doi.org/10.1029/2019GL085774

Lenschow DH, Wulfmeyer V, Senff C (2000) Measuring second-through fourth-order moments in noisy data. J Atmos Oceanic Technol 17(10):1330–1347

Leuenberger D, Haefele A, Omanovic N, Fengler M, Martucci G, Calpini B, Fuhrer O, Rossa A (2020) Improving high-impact numerical weather prediction with Lidar and drone observations. Bull Amer Meteor Soc 101:1036–1051. https://doi.org/10.1175/BAMS-D-19-0119.1

Llasat C, Llasat-Botija M, Petrucci O, Pasqua AA, Rosselló J, Vinet F, Boissier L (2013) Towards a database on societal impact of Mediterranean floods within the framework of the HYMEX project. Nat Hazards Earth Syst Sci 13:1337–1350

Majumdar SJ (2016) A review of targeted observations. Bull Am Meteor Soc 97(12):2287–2303. https://doi.org/10.1175/bams-d-14-00259.1

Montmerle T, Michel Y, Arbogast E, Ménétrier B, Brousseau P (2018) A 3D ensemble variational data assimilation scheme for the limited-area AROME model: formulation and preliminary results. Q J R Meteorol Soc 144:2196–2215. https://doi.org/10.1002/qj.3334

Muñoz-Porcar C, Sicard M, Granados-Muñoz MJ, Barragán R, Comerón A, Rocadenbosh F, Rodríguez-Gómez A, Garcia-Vizcaino D (2021) Synergy of Raman Lidar and modeled temperature for relative humidity profiling: assessment and uncertainty analysis. IEEE Trans Geosci Remote Sens 59:1–12

Muñoz-Porcar C, Comerón A, Sicard M, Barragán R, Garcia-Vizcaino D, Rodríguez-Gómez A, Rocadenbosh F, Granados-Muñoz MJ (2018) Calibration of Raman lidar water vapor mixing ratio measurements using zenithal measurements of diffuse sunlight and a radiative transfer model. IEEE Trans Geosci Remote Sens 56:7405–7414

Nuissier O, Ducrocq V, Ricard D, Lebeaupin C, Anquetin S (2008) A numerical study of three catastrophic precipitating events over southern France. I: Numerical framework and synoptic ingredients. Q J R Meteorol Soc 134:111–130. https://doi.org/10.1002/qj.200

Reichardt J, Wandinger U, Klein V, Mattis I, Hilber B, Begbie R (2012) RAMSES: German Meteorological Service autonomous Raman lidar for water vapor, temperature, aerosol, and cloud measurements. Appl Opt 51:8111–8131

Ribes A, Thao S, Vautard R, Dubuisson B, Somot S, Colin J, Planton S, Soubeyroux J-M (2019) Observed increase in extreme daily rainfall in the French Mediterranean. Clim Dyn 52:1095–1114. https://doi.org/10.1007/s00382-018-4179-2

Ricard DV, Ducrocq, and L. Auger, (2012) A climatology of the mesoscale environment associated with heavily precipitating events over a northwestern Mediterranean area. J Appl Meteor Climatol 51:468–488

Richard E, Fourrié N, Faivre M, Bielli S, Flamant C (2014) Assimilation of LEANDRE II water observations: impact study on the HyMeX SOP1 precipitation forecasts, 8th HyMeX workshop, Sep 2014. La Valletta, Malta, pp 15–18

Seity Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F, Lac C, Masson V (2011) The AROME-France convective scale operational model. Mon Weather Rev 139:976–991

Stelitano D, Di Girolamo P, Scoccione A, Summa D, Cacciani M (2019) Characterization of atmospheric aerosol optical properties based on the combined use of a ground-based Raman lidar and an airborne optical particle counter in the framework of the Hydrological Cycle in the Mediterranean Experiment – Special Observation Period 1. Atmos Meas Tech 12:2183–2199. https://doi.org/10.5194/amt-12-2183-2019

Stevens B, Bony S, Farrell D, Ament F, Blyth A, Fairall C, Karstensen J, Quinn P, Speich S, Aemisegger F, Albright AL, Bodenschatz E, Caesar K-A, Chewitt-Lucas R, Delanoë J, De Boer G, Ewald F, Forde M, George G, Hausold A, Hagen M, Hirsch L, Jansen F, Kinne S, Klocke D, Kölling T, Konow H, Mohr W, Naumann AK, Nuijens L, Pincus R, Reverdin G, Roberts G, Schnitt S, Sullivan P, Touzé-Peiffer L, Vial J, Vogel R, Acquistapace C, Alexander N, Alves L, Arixi S, Asmath H, Bagheri G, Bailey A, Baranowski D, Baron A, Barrau S, Barrett PA, Behrendt A, Bellenger H, Bendinger A, Beucher F, Bigorre S, Blossey P, Bock O, Bosser P, Bourras D, Bouruet-Aubertot P, Bower K, Branger H, Brennek M, Brewer A, Brilouet P-E, Brügmann B, Buehler SA, Burke E, Burton R, Böing S, Calmer R, Canonici J-C, Carton X, Cato GJ, Charles JA, Chazette P, Chen Y, Choularton T, Chuang P, Clarke S, Coe H, Cornet C, Coutris P, Couvreux F, Crewell S, Cronin T, Cui Z, Cuypers Y, Daley A, Damerell GM, Dauhut T, De Graaf D, De Groot G, De Szoeke SP, Denby L, Deneke H, Desbios J-P, Douet V, Drushka K, Dütsch M, Ehrlich A, Emanuel K, Emmanouilidis A, Etienne J-C, Etienne-Leblanc S, Faure G, Feingold G, Ferrero L, Fildier B, Fix A, Flamant C, Flatau PJ, Foltz GR, Forster L, Gadian A, Galewsky J, Gallagher M, Gallimore P, Gaston C, Gentemann C, Geyskens N, Giez A, Gourbeyre C, Gross S, Grosz R, Gutleben M, Güttler J, Hall K, Harris G, Helfer KC, Henze D, Herbert C, Heywood KJ, Holanda B, Ibanez-Landeta A, Intrieri J, Iyer S, Jacob M, Julien F, Kalesse H, Kazil J, Kellman A, Kirchner U, Klingebiel M, Kremper LA, Kretzschmar J, Krüger O, Kurz A, Körner M, L'Hégaret P, Lachlan-Cope T, Laing A, Landschützer P, Lang T, Lange D, Lange I, Laplace C, Laxenaire R, Le Bihan C, Leandro M, Lefevre N, Lenschow D, Li Q, Lloyd G, Los S, Losi N, Lothon M, Lovell O, Luneau C, Makuch P, Malinowski S, Manta G, Marinou E, Marsden N, Matthieu L, Maury N, Mayer B, Mayers-Als M, Mazel C, McGeary W, McWilliams JC, Mech M, Mehlmann M, Meroni AN, Mieslinger T, Minikin A, Morfa Avalos Y,

Muller C, Musat I, Möller G, Napoli A, Neuberger A, Noone D, Nordsiek F, Nowak JL, Olivier L, Oswald L, Parker DJ, Peck C, Person R, Plueddemann A, Pologne L, Posyniak M, Prange M, Pöhlker M, Pöhlker C, Pörtge V, Pöschl U, Quiñones Meléndez E, Radtke J, Ramage K, Reimann J, Renault L, Reus K, Reyes A, Ribbe J, Ringel M, Ritschel M, Rocha CB, Rochetin N, Rollo C, Röttenbacher J, Saffin L, Sandiford S, Sandu I, Schemann V, Schirmacher I, Schlenczek O, Schmidt J, Schröder M, Schulz H, Schwarzenboeck A, Schäfer M, Sealy A, Serikov I, Shohan S, Siddle E, Siebesma AP, Späth F, Stephan CC, Stolla MK, Szkółka W, Tarot S, Tetoni E, Thompson E, Thomson J, Tomassini L, Totems J, Villiger L, Walther A, Webber B, Wendisch M, Whitehall S, Wiltshire A, Wing AA, Wirth M, Wiskand J, Wolf K, Worbes L, Wright E, Wulfmeyer V, Young S, Zhang D, Zhang C, Ziemen F, Zinner T, Zuidema P, Zöger M (2021) EUREC4A. Earth Syst Sci Data. https://doi.org/10.5194/essd-2021-18

Thundathil R, Schwitalla T, Behrendt A, Muppa SK, Adam S, Wulfmeyer V (2020) Assimilation of lidar water vapour mixing ratio and temperature profiles into a convection-permitting model. J Meteorol Soc Jpn 98:959–986

Totems J, Chazette P, Raut J-C, J. (2019) Accuracy of current Arctic springtime water vapour estimates, assessed by Raman lidar. Q J R Meteorol Soc 145(720):1234–1249. https://doi.org/10.1002/qj.3492

Weckwerth T, Parsons D, Koch S, Moore J, Lemone P, Demoz B, Flamant C, Geerts B, Wang J, Feltz W (2004) An overview of the International H2O Project (IHOP_2002) and some preliminary highlights. Bull Amer Meteorol Soc 85:253–277

Wulfmeyer V, Hardesty RM, Turner DD, Behrendt A, Cadeddu MP, Di Girolamo P, Schlüssel P, Van Baelen J, Zus F (2015) A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles. Rev Geophys 53:819–895

Wulfmeyer V, Bauer H-S, Grzeschik M, Behrendt A, Vandenberghe F, Browell EV, Ismail S, Ferrare RA (2006) Four-dimensional variational assimilation of water vapor differential absorption lidar data: the first case study within IHOP-2002. Mon Weather Rev 134:209–230

Yoshida S, Yokota S, Seko H, Sakai T, Nagai T (2020) Observation system simulation experiments of water vapor profiles observed by Raman lidar using LETKF system. SOLA 16:43–50. https://doi.org/10.2151/sola.2020-008

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Cyrille Flamant[1] [ORCID] **· Patrick Chazette[2] · Olivier Caumont[3] · Paolo Di Girolamo[4] · Andreas Behrendt[5] · Michaël Sicard[6,7] · Julien Totems[2] · Diego Lange[5] · Nadia Fourrié[3] · Pierre Brousseau[3] · Clotilde Augros[3] · Alexandre Baron[2] · Marco Cacciani[4] · Adolfo Comerón[6] · Benedetto De Rosa[4] · Veronique Ducrocq[3] · Pascal Genau[1] · Laurent Labatut[3] · Constantino Muñoz-Porcar[6] · Alejandro Rodríguez-Gómez[6] · Donato Summa[4,8] · Rohith Thundathil[5] · Volker Wulfmeyer[5]**

1   LATMOS/IPSL, CNRS-SU-UVSQ, Sorbonne Université, Paris, France

2   LSCE/IPSL, CNRS-CEA-UVSQ, University Paris-Saclay, CEA Saclay, Gif sur Yvette, France

3   CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

4   Scuola Di Ingegneria, Università Della Basilicata, Potenza, Italy

5   Institut Für Physik Und Meteorologie, Universitat Hohenheim, Stuttgart, Germany

6   CommSensLab, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain

7   Institut D'Estudis Espacials de Catalunya, Ciències I Tecnologies de L'Espai-Centre de Recerca de L'Aeronàutica I de L'Espai, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain

8   CNR-IMAA, Tito Scalo center, Potenza, Italy

Chapter 5

Summary and Outlook

The ultimate goal of the thesis is to study the impact of assimilating thermodynamic profiles from lidar systems on a convection-permitting scale and explore the potential of using lidar systems for short-range forecasting. The whole impact study was conducted with the Weather Research and Forecasting model data assimilation (WRFDA) system on a convection-permitting horizontal resolution of 2.5 km over central Europe. A rapid update cycle (RUC) with hourly assimilations was performed for the DA experiments.

The thermodynamic profile observations were derived from two lidar systems of the University of Hohenheim (UHOH). Temperature and moisture profile observations were derived from the temperature Raman lidar (TRL) and differential absorption lidar (DIAL). Lidar profile observations were obtained from the High Definition of Clouds and Precipitation for advancing Climate Prediction (HD(CP)^2) project Observation Prototype Experiment (HOPE) campaign. The intensive observation period (IOP) from 0900 UTC to 1800 UTC was selected on 24 April 2013. Apart from the lidar measurements, four radiosondes were launched at 9, 11, 13 and 15 UTC.

In the WRF model, moisture information derived from the various instruments, whether in-situ or remote-sensing, was assimilated through a relative humidity (RH) operator. The prognostic variable in the WRF model, which is the water vapour mixing ratio (WVMR), was determined through the conversion of the RH parameter. RH has a strong dependency on temperature; however, WVMR is a conservative variable for processes involving a change in temperature. The assimilation of moisture and temperature profile observations obtained from the lidar systems had to be incorporated into the WRF model via a forward operator exclusively for WVMR. Hence for the direct assimilation of WVMR and temperature, a new forward operator called the thermodynamic lidar (TDLIDAR) operator was introduced. The operator was developed by modifying an already existing operator in the WRFDA system, the atmospheric infrared sounding retrieval (AIRSRET) operator.

In order to analyse the performance of the operator and to understand the impact of assimilating the lidar derived thermodynamic profiles, we conducted six model simulations. A **3DVar** DA system was configured for the whole impact study in a RUC mode. The model runs are: 1) NO_DA or the model run with no data assimilated; 2) CONV_DA or the

control run with only conventional observations from the ECMWF assimilated; 3) T_DA, where temperature data along with the conventional dataset are assimilated; 4) Q_DA, where WVMR data along with the conventional dataset are assimilated; 5) RH_DA, where RH along with the conventional dataset are assimilated, and finally 6) QT_DA, where both WVMR and temperature data along with the conventional data are assimilated.

From the results, an overall positive impact of WVMR and temperature lidar data assimilation was observed with the help of the new forward operator. With a deficient number of rejected observations, the WRF model assimilated the profile observations. The WVMR and temperature analyses indicated that the lidar observations could correct the first guess or the background of the model state during the assimilation process to a reasonable extent.

The study showed that the TDLIDAR operator has a better impact than the traditional RH operator, which was one of the research highlights. The WVMR and temperature data input through the forward operator positively impacted the WVMR analysis. The root-mean-square error (RMSE) of WVMR compared to WVDIAL observations for QT_DA, where both WVMR and temperature data were assimilated, reduced by 40% compared to the control run. In contrast, RH_DA, where RH data was assimilated, did not improve overall. Therefore the data input through the forward operator positively impacted the WVMR analysis. However, the impact on the temperature analysis did not provide a satisfactory improvement. Temperature analysis did not improve due to the significant dependency of WVMR with temperature in the background error covariance matrix **B**. With several single observation tests with different temperature and WVMR assimilations, we concluded that the static nature of the background error covariance matrix was the sole reason for the detrimental impact on the temperature analysis. A positive increment in the WVMR created a strong negative increment in the temperature anaylsis which did not agree with the real-time temperature conditions of the case study.

A flow-dependent **B** matrix was expected to alleviate the cross-correlation of the temperature and humidity variables to represent the real-time scenario better. The **B** matrix in ensemble-based DA systems reflects the dynamic nature of the atmosphere. Thus, thermodynamic lidar data assimilation with ensemble DA techniques was performed as the next step. As expected, the ensemble-based DA study produced promising improved results using flow-dependent error covariance compared to the deterministic 3DVar experiment.

A hybrid **3DVar**-ensemble transform Kalman filter (ETKF) DA system was used for the DA study with the same DA cycling architecture as the previous deterministic **3DVar** DA experiment, and also the same observations were assimilated. The results of the hybrid DA with a deterministic **3DVar** DA run were compared. Several single observation tests (SOTs) were also conducted to understand how the flow-dependent error covariance matrix influences the analyses.

Similar to the previous impact study, four experiments were conducted, namely: conventional data assimilation using **3DVar** (NOLIDAR_3DVAR); similar experiment using hybrid **3DVar-ETKF** (NOLIDAR_HYBRID); lidar data assimilation experiment in addition to the conventional data with **3DVar** (LIDAR_3DVAR), and same with hybrid DA (LIDAR_HYBRID).

Simultaneous temperature and WVMR improvement were observed adopting the hybrid DA approach rather than the **3DVar** experiment. Compared to the WVDIAL observations, the WVMR RMSE of the LIDAR_HYBRID fell by 0.62 g kg^{-1} compared to NOLIDAR_3DVAR. In contrast, the LIDAR_3DVAR experiment showed a lower reduction of 0.36 g kg^{-1}. The reduction in the temperature RMSE from the LIDAR_HYBRID to NOLIDAR_3DVAR experiments compared to the TRL observations was 0.54 K, whereas the LIDAR_3DVAR experiment even showed a slight RMSE increase of 0.06 K. The RMSEs were calculated in comparison with independent radiosonde observations. The overall RMSE for all times where the radiosondes were launched showed an improvement of the hybrid approach over the **3DVar**. Added flow-dependency in the **B** matrix enables the model to capture better the thermodynamic flow of the atmosphere than the static **B** matrix used in the deterministic 3DVar.

Apart from radiosondes, planetary boundary layer height (PBLH) estimates from ceilometers were compared, which showed positive results for the hybrid DA. Forecast validations were conducted with respect to PBLH estimates and also Global Navigation Satellite System-precipitable water vapour (GNSS-PWV). A 9-hour forecast starting at 18 UTC was initiated from each experiment. For both PBLH and integrated water vapour (IWV) forecasts, the hybrid performed better than **3DVar**. The forecast impact of PBLH persisted for about 7 hours. Considering the overall PBLH RMSEs, LIDAR_HYBRID DA run had the lowest RMSE of 60 m compared to 140 m in LIDAR_3DVAR. The PWV impact persisted for about a roughly 4-hour forecast lead time. The results from the study indicate the future potential of assimilating water vapour and temperature data from an operational lidar network which is explained as an application in Chapter four. Another important finding was the spatial extent of the impact when assimilating lidar profiles with an ensemble approach from a single geographical location. The area influenced was significantly higher than with the **3DVar** approach. The analysis increments of a single lidar vertical profile were spread over a radius of 100 km using the hybrid approach.

As an outlook from this thesis, the aim should be to develop a robust network of operational lidar systems in the future, capturing all the prognostic variables essential for a good weather forecast. This aim can be accomplished with sophisticated DA methods and a network of lidar systems soon. As seen from the results, the analysis increments of a single lidar vertical profile were spread over a radius of 100 km. The impact area provides insight into the minimum distance between two lidar systems to be set up adjacently in the potential future lidar network. From this study, a slightly less than 200 km spacing of the

lidar network is sufficient. However, a single case study is insufficient to derive conclusions about future lidar network design strategies. To obtain more robust conclusions, we want to do further research at different locations with different environmental conditions.

The thesis also gives an insight into a new initiative where the application of the thesis can be extended. A French initiative called the Water Vapor Lidar Network assimilation (WaLiNeAs) will deploy several autonomous water vapour Raman lidars in the Southern region of France facing the Western Mediterranean Sea. The WaLiNeAs initiative will strengthen the foundation for a new approach to operational forecasting with the help of lidar systems operated simultaneously. Research groups from France, Germany and Italy will operate their Raman lidars together. The transportable and automatic Atmospheric Raman Temperature and Humidity Sounder (ARTHUS, Lange et al., 2019) from the University of Hohenheim (UHOH) will also be a part of the research campaign. With the research initiative, the French operational AROME numerical weather prediction (NWP) system will test the **4DEnVar** DA technique operationally. With the new operational system, probabilistic quantitative precipitation forecasting (PrQPF), a challenging field of meteorology, will be extensively studied, analysing a number of heavy precipitation events. The WaLiNeAs initiative will be a starting point to test and design future assimilation strategies. Also, from a scientific perspective, the role of a higher density of lidar systems operating as a network to help predict short-range forecasts can be better understood. Lidar network DA promises excellent potential to improve future operational forecasting.

Also, it would be beneficial to incorporate wind data in the model. This thesis focused on thermodynamic variables, whereas wind is also vital for weather forecasting. Wind DA is expected to improve the model dynamics. Wind observations are a critical component in the mesoscale data assimilation that determines the weather evolution. In the absence of a strong forcing such as orography, the model may lead to noise and error growth. Also, high-resolution wind observations are essential to initialise turbulent scale phenomena. The absence of turbulence characterisation leads to high variances in high-resolution NWP resulting in faster changes and smaller scales in the background error covariances. With such poor initialisation, analysis of large scales is potentially compromised (Stoffelen et al., 2020). Future DA studies will focus on advanced DA methodologies simultaneously incorporating Doppler wind data and thermodynamic profiles.

Another area that remains a challenge and needs more research is the modelling of the observation error covariance matrix **R**. As we proceed to high-resolution operational forecasting, demand for the assimilation of high density and high-resolution observations is also high. Restricting the observational error covariance to the idea that observations are uncorrelated will lead to problems in the accuracy of short-range forecasts like now casting for thunderstorms or flash floods. Uncorrelated error covariances lead to analysis improvement only in the region of influence where high-density observations are available. Desroziers et al., 2005 had defined a new method to separate correlated observation and

background errors. This method is based on a statistical approximation of the covariance of observation errors using post-analysis diagnostics from linear estimation theory.

Computational resources involved in designing a DA system depends on the resolution of the NWP model implemented. The higher the resolution, the higher is the computational expense. Hybrid DA using variational and probabilistic DA approaches has great potential in future operational forecasting using high-resolution thermodynamic lidar data for assimilation. Advancement in the quality and speed of computational resources has taken NWP to a different level. As the computational capabilities increase, it is not as difficult as before to incorporate the high-resolution time-series measurement from lidar systems anymore. The advantage of lidar data is its temporal and horizontal resolution which can detect even a minor disturbance in the atmosphere. More DA techniques that can use the continuous measurement of the lidar systems like the **4DVar**, **ETKF-4DVar**, and **4DEn-Var** should be explored in the future, which considers the temporal information of the observations. In the future, it will be exciting to test the sensitivity of ensemble-based DA techniques with more ensemble members with advanced computing power. Increasing the number of ensemble members is also essential to alleviate errors arising due to sampling issues.

Bibliography

Abbe, Cleveland (1901). "The physical basis of long-range weather forecasts". In: *Monthly Weather Review* 29.12, pp. 551–561. DOI: 10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2.

Adam, Stephan, Andreas Behrendt, Thomas Schwitalla, Eva Hammann, and Volker Wulfmeyer (2016). "First assimilation of temperature lidar data into an NWP model: Impact on the simulation of the temperature field, inversion strength and PBL depth". In: *Quarterly Journal of the Royal Meteorological Society* 142.700, pp. 2882–2896. DOI: 10.1002/qj.2875.

Arshinov, Yuri, Sergey Bobrovnikov, Ilya Serikov, Albert Ansmann, Ulla Wandinger, Dietrich Althausen, Ina Mattis, and Detlef Müller (2005). "Daytime operation of a pure rotational Raman lidar by use of a Fabry–Perot interferometer". In: *Applied optics* 44.17, pp. 3593–3603. DOI: 10.1364/AO.44.003593.

Bannister, RN (2017). "A review of operational methods of variational and ensemble-variational data assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 143.703, pp. 607–633. DOI: 10.1002/qj.2982.

Bannister, Ross N (2008). "A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 134.637, pp. 1951–1970. DOI: 10.1002/qj.339.

Barker, Dale M, Wei Huang, Yong-Run Guo, AJ Bourgeois, and QN Xiao (2004). "A three-dimensional variational data assimilation system for MM5: Implementation and initial results". In: *Monthly Weather Review* 132.4, pp. 897–914. DOI: 10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2.

Bauer, Peter, Alan Thorpe, and Gilbert Brunet (2015). "The quiet revolution of numerical weather prediction". In: *Nature* 525.7567, pp. 47–55. DOI: 10.1038/nature14956.

Behrendt, A, V Wulfmeyer, E Hammann, SK Muppa, and S Pal (2015). "Profiles of second- to fourth-order moments of turbulent temperature fluctuations in the convective boundary layer: first measurements with rotational Raman lidar". In: *Atmospheric Chemistry and Physics* 15.10, pp. 5485–5500. DOI: 10.5194/acp-15-5485-2015.

Behrendt, Andreas, Takuji Nakamura, Michitaka Onishi, Rudolf Baumgart, and Toshitaka Tsuda (2002). "Combined Raman lidar for the measurement of atmospheric temperature, water vapor, particle extinction coefficient, and particle backscatter coefficient". In: *Applied optics* 41.36, pp. 7657–7666. DOI: `10.1364/AO.41.007657`.

Behrendt, Andreas, Takuji Nakamura, and Toshitaka Tsuda (2004). "Combined temperature lidar for measurements in the troposphere, stratosphere, and mesosphere". In: *Applied optics* 43.14, pp. 2930–2939. DOI: `10.1364/AO.43.002930`.

Behrendt, Andreas and Jens Reichardt (2000). "Atmospheric temperature profiling in the presence of clouds with a pure rotational Raman lidar by use of an interference-filter-based polychromator". In: *Applied Optics* 39.9, pp. 1372–1378. DOI: `10.1364/AO.39.001372`.

Behrendt, Andreas, Volker Wulfmeyer, Andrea Riede, Gerd Wagner, Sandip Pal, Heinz Bauer, Marcus Radlach, and Florian Späth (2009). "Three-dimensional observations of atmospheric humidity with a scanning differential absorption lidar". In: *Remote sensing of clouds and the atmosphere XIV*. Vol. 7475. SPIE, pp. 157–166. DOI: `10.1117/12.835143`.

Benjamin, Stanley G, Stephen S Weygandt, John M Brown, Ming Hu, Curtis R Alexander, Tatiana G Smirnova, Joseph B Olson, Eric P James, David C Dowell, Georg A Grell, et al. (2016). "A North American hourly assimilation and model forecast cycle: The Rapid Refresh". In: *Monthly Weather Review* 144.4, pp. 1669–1694. DOI: `10.1175/MWR-D-15-0242.1`.

Berre, Loïk (2000). "Estimation of synoptic and mesoscale forecast error covariances in a limited-area model". In: *Monthly weather review* 128.3, pp. 644–667. DOI: `10.1175/1520-0493(2000)128<0644:EOSAMF>2.0.CO;2`.

Bhawar, Rohini, Paolo Di Girolamo, Donato Summa, Cyrille Flamant, Dietrich Althausen, Andreas Behrendt, Christoph Kiemle, P Bosser, M Cacciani, Cedric Champollion, et al. (2011). "The water vapour intercomparison effort in the framework of the Convective and Orographically-induced Precipitation Study: airborne-to-ground-based and airborne-to-airborne lidar systems". In: *Quarterly Journal of the Royal Meteorological Society* 137.S1, pp. 325–348. DOI: `10.1002/qj.697`.

Bielli, Soline, Matthias Grzeschik, Evelyne Richard, Cyrille Flamant, Cedric Champollion, Christoph Kiemle, M Dorninger, and Pierre Brousseau (2012). "Assimilation of water-vapour airborne lidar observations: impact study on the COPS precipitation forecasts". In: *Quarterly Journal of the Royal Meteorological Society* 138.667, pp. 1652–1667. DOI: `10.1002/qj.1864`.

Bishop, Craig H, Brian J Etherton, and Sharanya J Majumdar (2001). "Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects". In: *Monthly weather review* 129.3, pp. 420–436. DOI: `10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2`.

Bishop, Craig H and Daniel Hodyss (2011). "Adaptive ensemble covariance localization in ensemble 4D-VAR state estimation". In: *Monthly Weather Review* 139.4, pp. 1241–1255. DOI: 10.1175/2010MWR3403.1.

Bjerknes, Vilhelm, Esther Volken, and S Bronnimann (2009). "The problem of weather prediction, considered from the viewpoints of mechanics and physics". In: *Meteorologische Zeitschrift* 18.6, p. 663. DOI: 10.1127/0941-2948/2009/416.

Blumberg, WG, DD Turner, U Löhnert, and S Castleberry (2015). "Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part II: Actual retrieval performance in clear-sky and cloudy conditions". In: *Journal of Applied Meteorology and Climatology* 54.11, pp. 2305–2319. DOI: 10.1175/JAMC-D-15-0005.1.

Bolton, David (1980). "The computation of equivalent potential temperature". In: *Monthly weather review* 108.7, pp. 1046–1053. DOI: 10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Bösenberg, Jens (1998). "Ground-based differential absorption lidar for water-vapor and temperature profiling: methodology". In: *Applied Optics* 37.18, pp. 3845–3860. DOI: 10.1364/AO.37.003845.

Bouttier, François, Benoît Vié, Olivier Nuissier, and Laure Raynaud (2012). "Impact of stochastic physics in a convection-permitting ensemble". In: *Monthly Weather Review* 140.11, pp. 3706–3721. DOI: 10.1175/MWR-D-12-00031.1.

Bowler, Neill E, Alberto Arribas, Kenneth R Mylne, Kelvyn B Robertson, and Sarah E Beare (2008). "The MOGREPS short-range ensemble prediction system". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 134.632, pp. 703–722. DOI: 10.1002/qj.234.

Brousseau, Pierre, Yann Seity, Didier Ricard, and Julien Léger (2016). "Improvement of the forecast of convective activity from the AROME-France system". In: *Quarterly Journal of the Royal Meteorological Society* 142.699, pp. 2231–2243. DOI: 10.1002/qj.2822.

Browning, Keith A, Alan M Blyth, Peter A Clark, Ulrich Corsmeier, Cyril J Morcrette, Judith L Agnew, Sue P Ballard, Dave Bamber, Christian Barthlott, Lindsay J Bennett, et al. (2007). "The convective storm initiation project". In: *Bulletin of the American Meteorological Society* 88.12, pp. 1939–1956. DOI: 10.1175/BAMS-88-12-1939.

Buehner, Mark (2005). "Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.607, pp. 1013–1043. DOI: 10.1256/qj.04.15.

Buehner, Mark, PL Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He (2010a). "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments". In: *Monthly Weather Review* 138.5, pp. 1550–1566. DOI: 10.1175/2009MWR3157.1.

Buehner, Mark, PL Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He (2010b). "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations". In: *Monthly Weather Review* 138.5, pp. 1567–1586. DOI: 10.1175/2009MWR3158.1.

Buizza, Roberto and Martin Leutbecher (2015). "The forecast skill horizon". In: *Quarterly Journal of the Royal Meteorological Society* 141.693, pp. 3366–3382. DOI: 10.1002/qj.2619.

Cadeddu, Maria P, GE Peckham, and C Gaffard (2002). "The vertical resolution of ground-based microwave radiometers analyzed through a multiresolution wavelet technique". In: *IEEE transactions on geoscience and remote sensing* 40.3, pp. 531–540. DOI: 10.1109/TGRS.2002.1000313.

Clayton, Adam M, Andrew C Lorenc, and Dale M Barker (2013). "Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office". In: *Quarterly Journal of the Royal Meteorological Society* 139.675, pp. 1445–1461. DOI: 10.1002/qj.2054.

Collins, Matthew and Myles R Allen (2002). "Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability". In: *Journal of Climate* 15.21, pp. 3104–3109. DOI: 10.1175/1520-0442(2002)015<3104:ATRROI>2.0.CO;2.

Cooney, John (1972). "Measurement of atmospheric temperature profiles by Raman backscatter". In: *Journal of Applied Meteorology and Climatology* 11.1, pp. 108–112. DOI: 10.1175/1520-0450(1972)011<0108:MOATPB>2.0.CO;2.

Council, National Research et al. (1998). *The Atmospheric Sciences: Entering the Twenty-First Century*. National Academies Press.

Courtier, Philippe, E Andersson, W Heckley, D Vasiljevic, M Hamrud, A Hollingsworth, F Rabier, M Fisher, and J Pailleux (1998). "The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation". In: *Quarterly Journal of the Royal Meteorological Society* 124.550, pp. 1783–1807. DOI: 10.1002/qj.49712455002.

Dabberdt, Walter F, Thomas W Schlatter, and with contributions from the rest of the PDT-2 (1996). "Research opportunities from emerging atmospheric observing and modeling capabilities". In: *Bulletin of the American Meteorological Society* 77.2, pp. 305–324. DOI: 10.1175/1520-0477(1996)077<0305:ROFEAO>2.0.CO;2.

Desroziers, Gérald, Loic Berre, Bernard Chapnik, and Paul Poli (2005). "Diagnosis of observation, background and analysis-error statistics in observation space". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.613, pp. 3385–3396. DOI: 10.1256/qj.05.108.

Di Girolamo, Paolo, Marco Cacciani, Donato Summa, Andrea Scoccione, Benedetto De Rosa, Andreas Behrendt, and Volker Wulfmeyer (2017). "Characterisation of boundary layer turbulent processes by the Raman lidar BASIL in the frame of HD (CP) 2 Observational Prototype Experiment". In: *Atmospheric Chemistry and Physics* 17.1, pp. 745–767. DOI: 10.5194/acp-17-745-2017.

Di Girolamo, Paolo, R Marchese, DN Whiteman, and BB Demoz (2004). "Rotational Raman Lidar measurements of atmospheric temperature in the UV". In: *Geophysical research letters* 31.1. DOI: 10.1029/2003GL018342.

Dinoev, T, V Simeonov, Y Arshinov, S Bobrovnikov, Pablo Ristori, B Calpini, M Parlange, and Hubert van den Bergh (2013). "Raman lidar for meteorological observations, RALMO– Part 1: Instrument description". In: *Atmospheric Measurement Techniques* 6.5, pp. 1329– 1346. DOI: 10.5194/amt-6-1329-2013.

Drobinski, Philippe, Véronique Ducrocq, P Alpert, E Anagnostou, K Béranger, M Borga, Isabelle Braud, Andre Chanzy, S Davolio, G Delrieu, et al. (2014). "HyMeX: A 10-year multidisciplinary program on the Mediterranean water cycle". In: *Bulletin of the American Meteorological Society* 95.7, pp. 1063–1082. DOI: BAMS-D-12-00242.1.

Emanuel, Kerry, David Raymond, Alan Betts, Lance Bosart, Christopher Bretherton, Kelvin Droegemeier, Brian Farrell, J Michael Fritsch, Robert Houze, Margaret Le Mone, et al. (1995). "Report of the first prospectus development team of the US Weather Research Program to NOAA and the NSF". In: *Bulletin of the American Meteorological Society*, pp. 1194–1208. URL: https://www.jstor.org/stable/26232528.

Engelbart, DAM and E Haas (2006). "LAUNCH-2005: International Lindenberg campaign for assessment of humidity and cloud profiling systems and its impact on high-resolution modelling". In: *7th International Symposium on Tropospheric profiling: Needs and technologies*, pp. 11–17. URL: https://library.wmo.int/pmb_ged/wmo-td_1354_en/P2(07)_Engelbart_Germany.pdf.

Errico, Ronald M (1997). "What is an adjoint model?" In: *Bulletin of the American Meteorological Society* 78.11, pp. 2577–2592. DOI: 10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2.

Evensen, Geir (2003). "The ensemble Kalman filter: Theoretical formulation and practical implementation". In: *Ocean dynamics* 53.4, pp. 343–367. DOI: 10.1007/s10236-003-0036-9.

Fairbairn, David, Stephen R Pring, Andrew C Lorenc, and Ian Roulstone (2014). "A comparison of 4DVar with ensemble data assimilation methods". In: *Quarterly Journal of the Royal Meteorological Society* 140.678, pp. 281–294. DOI: 10.1002/qj.2135.

Fisher, Mike (2003). "Background error covariance modelling". In: *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean*. Shinfield Park, Reading, pp. 45– 63.

Flamant, Cyrille, Patrick Chazette, Olivier Caumont, Paolo Di Girolamo, Andreas Behrendt, Michaël Sicard, Julien Totems, Diego Lange, Nadia Fourrié, Pierre Brousseau, et al. (2021). "A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative". In: *Bulletin of Atmospheric Science and Technology* 2.1, pp. 1–21. DOI: 10.1007/s42865-021-00037-6.

Geerts, Bart, David Parsons, Conrad L Ziegler, Tammy M Weckwerth, Michael I Biggerstaff, Richard D Clark, Michael C Coniglio, Belay B Demoz, Richard A Ferrare, William A Gallus Jr, et al. (2017). "The 2015 plains elevated convection at night field project". In: *Bulletin of the American Meteorological Society* 98.4, pp. 767–786. DOI: `10.1175/BAMS-D-15-00257.1`.

Goldsmith, JEM, Forest H Blair, Scott E Bisson, and David D Turner (1998). "Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols". In: *Applied Optics* 37.21, pp. 4979–4990. DOI: `10.1364/AO.37.004979`.

Goodliff, Michael, Javier Amezcua, and Peter Jan Van Leeuwen (2015). "Comparing hybrid data assimilation methods on the Lorenz 1963 model with increasing non-linearity". In: *Tellus A: Dynamic Meteorology and Oceanography* 67.1, p. 26928. DOI: `10.3402/tellusa.v67.26928`.

Grzeschik, Matthias, Hans-Stefan Bauer, Volker Wulfmeyer, Dirk Engelbart, Ulla Wandinger, Ina Mattis, Dietrich Althausen, Ronny Engelmann, Matthias Tesche, and Andrea Riede (2008). "Four-dimensional variational data analysis of water vapor Raman lidar data and their impact on mesoscale forecasts". In: *Journal of Atmospheric and Oceanic Technology* 25.8, pp. 1437–1453. DOI: `10.1175/2007JTECHA974.1`.

Hagelin, Susanna, Joohyung Son, Richard Swinbank, Anne McCabe, Nigel Roberts, and Warren Tennant (2017). "The Met Office convective-scale ensemble, MOGREPS-UK". In: *Quarterly Journal of the Royal Meteorological Society* 143.708, pp. 2846–2861. DOI: `10.1002/qj.3135`.

Hamill, Thomas M and Chris Snyder (2000). "A hybrid ensemble Kalman filter–3D variational analysis scheme". In: *Monthly Weather Review* 128.8, pp. 2905–2919. DOI: `10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2`.

Hammann, E, A Behrendt, F Le Mounier, and V Wulfmeyer (2015). "Temperature profiling of the atmospheric boundary layer with rotational Raman lidar during the HD (CP) 2 Observational Prototype Experiment". In: *Atmospheric Chemistry and Physics* 15.5, pp. 2867–2881. DOI: `10.5194/acp-15-2867-2015`.

Hoke, James E and Richard A Anthes (1976). "The initialization of numerical models by a dynamic-initialization technique". In: *Monthly Weather Review* 104.12, pp. 1551–1556. DOI: `10.1175/1520-0493(1976)104<1551:TIONMB>2.0.CO;2`.

Hollingsworth, A and P Lönnberg (1986). "The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field". In: *Tellus A* 38.2, pp. 111–136. DOI: `10.1111/j.1600-0870.1986.tb00460.x`.

Houtekamer, Peter L and Herschel L Mitchell (1998). "Data assimilation using an ensemble Kalman filter technique". In: *Monthly Weather Review* 126.3, pp. 796–811. DOI: `10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2`.

Huang, Xiang-Yu, Qingnong Xiao, Dale M Barker, Xin Zhang, John Michalakes, Wei Huang, Tom Henderson, John Bray, Yongsheng Chen, Zaizhong Ma, et al. (2009). "Four-dimensional

variational data assimilation for WRF: Formulation and preliminary results". In: *Monthly Weather Review* 137.1, pp. 299–314. DOI: 10.1175/2008MWR2577.1.

Hunt, Brian R, Eric J Kostelich, and Istvan Szunyogh (2007). "Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter". In: *Physica D: Nonlinear Phenomena* 230.1-2, pp. 112–126. DOI: https://doi.org/10.1016/j.physd.2006.11.008.

Ingleby, NB, Andrew C Lorenc, K Ngan, F Rawlins, and DR Jackson (2013). "Improved variational analyses using a nonlinear humidity control variable". In: *Quarterly Journal of the Royal Meteorological Society* 139.676, pp. 1875–1887. DOI: 10.1002/qj.2073.

Ismail, Syed and Edward V Browell (1989). "Airborne and spaceborne lidar measurements of water vapor profiles: a sensitivity analysis". In: *Applied Optics* 28.17, pp. 3603–3615. DOI: 10.1364/AO.28.003603.

Kalnay, Eugenia (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

Kamineni, Rupa, TN Krishnamurti, Richard A Ferrare, Syed Ismail, and Edward V Browell (2003). "Impact of high resolution water vapor cross-sectional data on hurricane forecasting". In: *Geophysical research letters* 30.5. DOI: 10.1029/2002GL016741.

Kamineni, Rupa, TN Krishnamurti, S Pattnaik, Edward V Browell, Syed Ismail, and Richard A Ferrare (2006). "Impact of CAMEX-4 datasets for hurricane forecasts using a global model". In: *Journal of the atmospheric sciences* 63.1, pp. 151–174. DOI: 10.1175/JAS3588.1.

Kawabata, Takuya and Genta Ueno (2020). "Non-Gaussian probability densities of convection initiation and development investigated using a particle filter with a storm-scale numerical weather prediction model". In: *Monthly Weather Review* 148.1, pp. 3–20. DOI: 10.1175/MWR-D-18-0367.1.

Kutty, Govindan, Rohit Muraleedharan, and Amit P Kesarkar (2018). "Impact of representing model error in a hybrid ensemble-variational data assimilation system for track forecast of tropical cyclones over the Bay of Bengal". In: *Pure and Applied Geophysics* 175.3, pp. 1155–1167. DOI: 10.1007/s00024-017-1747-z.

Kutty, Govindan and Xuguang Wang (2015). "A comparison of the impacts of radiosonde and AMSU radiance observations in GSI based 3DEnsVar and 3DVar data assimilation systems for NCEP GFS". In: *Advances in Meteorology* 2015. DOI: 10.1155/2015/280546.

Lakshmivarahan, S and John M Lewis (2013). "Nudging methods: A critical overview". In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, pp. 27–57. DOI: 10.1007/978-3-642-35088-7_2.

Lange, Diego, Andreas Behrendt, and Volker Wulfmeyer (2019). "Compact operational tropospheric water vapor and temperature Raman lidar with turbulence resolution". In: *Geophysical Research Letters* 46.24, pp. 14844–14853. DOI: 10.1029/2019GL085774.

Leblanc, T, IS McDermid, and TD Walsh (2012). "Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring". In: *Atmospheric Measurement Techniques* 5.1, pp. 17–36. DOI: 10.5194/amt-5-17-2012.

Lewis, John M, Sivaramakrishnan Lakshmivarahan, and Sudarshan Dhall (2006). *Dynamic data assimilation: a least squares approach*. Vol. 13. Cambridge University Press.

Liu, Chengsi, Qingnong Xiao, and Bin Wang (2008). "An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test". In: *Monthly Weather Review* 136.9, pp. 3363–3373. DOI: 10.1175/2008MWR2312.1.

Lorenc, Andrew C (2003). "The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 129.595, pp. 3183–3203. DOI: 10.1256/qj.02.132.

Lorenc, Andrew C, Neill E Bowler, Adam M Clayton, Stephen R Pring, and David Fairbairn (2015). "Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for global NWP". In: *Monthly Weather Review* 143.1, pp. 212–229. DOI: 10.1175/MWR-D-14-00195.1.

Lorenz, Ed N (1982). "Atmospheric predictability experiments with a large numerical model". In: *Tellus* 34.6, pp. 505–513. DOI: 10.3402/tellusa.v34i6.10836.

Lynch, Peter (2008). "The origins of computer weather prediction and climate modeling". In: *Journal of computational physics* 227.7, pp. 3431–3444. DOI: 10.1016/j.jcp.2007.02.034.

Milan, Marco, Dirk Schüttemeyer, Theresa Bick, and Clemens Simmer (2014). "A sequential ensemble prediction system at convection-permitting scales". In: *Meteorology and Atmospheric Physics* 123.1, pp. 17–31. DOI: 10.1007/s00703-013-0291-3.

Montmerle, Thibaut, Yann Michel, Etienne Arbogast, Benjamin Ménétrier, and Pierre Brousseau (2018). "A 3D ensemble variational data assimilation scheme for the limited-area AROME model: Formulation and preliminary results". In: *Quarterly Journal of the Royal Meteorological Society* 144.716, pp. 2196–2215. DOI: https://doi.org/10.1002/qj.3334.

Newsom, Rob K, David D Turner, and John EM Goldsmith (2013). "Long-term evaluation of temperature profiles measured by an operational Raman lidar". In: *Journal of Atmospheric and Oceanic Technology* 30.8, pp. 1616–1634. DOI: 10.1175/JTECH-D-12-00138.1.

Nicolis, C (2007). "Dynamics of model error: The role of the boundary conditions". In: *Journal of the atmospheric sciences* 64.1, pp. 204–215. DOI: 10.1175/JAS3806.1.

Nicolis, C, Rui AP Perdigao, and S Vannitsem (2009). "Dynamics of prediction errors under the combined effect of initial condition and model errors". In: *Journal of the atmospheric sciences* 66.3, pp. 766–778. DOI: 10.1175/2008JAS2781.1.

Pan, Yujie, Kefeng Zhu, Ming Xue, Xuguang Wang, Ming Hu, Stanley G Benjamin, Stephen S Weygandt, and Jeffrey S Whitaker (2014). "A GSI-based coupled EnSRF–En3DVar

hybrid data assimilation system for the operational Rapid Refresh model: Tests at a reduced resolution". In: *Monthly Weather Review* 142.10, pp. 3756–3780. DOI: 10.1175/MWR-D-13-00242.1.

Parrish, David F and John C Derber (1992). "The National Meteorological Center's spectral statistical-interpolation analysis system". In: *Monthly Weather Review* 120.8, pp. 1747–1763. DOI: 10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Pereira, Margarida Belo and Loïk Berre (2006). "The use of an ensemble approach to study the background error covariances in a global NWP model". In: *Monthly weather review* 134.9, pp. 2466–2489. DOI: 10.1175/MWR3189.1.

Poterjoy, Jonathan, Ryan A Sobash, and Jeffrey L Anderson (2017). "Convective-scale data assimilation for the weather research and forecasting model using the local particle filter". In: *Monthly Weather Review* 145.5, pp. 1897–1918. DOI: 10.1175/MWR-D-16-0298.1.

Poterjoy, Jonathan, Louis Wicker, and Mark Buehner (2019). "Progress toward the application of a localized particle filter for numerical weather prediction". In: *Monthly Weather Review* 147.4, pp. 1107–1126. DOI: 10.1175/MWR-D-17-0344.1.

Potthast, Roland, Anne Walter, and Andreas Rhodin (2019). "A localized adaptive particle filter within an operational NWP framework". In: *Monthly Weather Review* 147.1, pp. 345–362. DOI: 10.1175/MWR-D-18-0028.1.

Radlach, Marcus, Andreas Behrendt, and V Wulfmeyer (2008). "Scanning rotational Raman lidar at 355 nm for the measurement of tropospheric temperature fields". In: *Atmospheric Chemistry and Physics* 8.2, pp. 159–169. DOI: 10.5194/acp-8-159-2008.

Reichardt, Jens, Ulla Wandinger, Volker Klein, Ina Mattis, Bernhard Hilber, and Robert Begbie (2012). "RAMSES: German Meteorological Service autonomous Raman lidar for water vapor, temperature, aerosol, and cloud measurements". In: *Applied optics* 51.34, pp. 8111–8131. DOI: 10.1364/AO.51.008111.

Repasky, Kevin S, Drew Moen, Scott Spuler, Amin R Nehrir, and John L Carlsten (2013). "Progress towards an autonomous field deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor in the lower troposphere". In: *Remote Sensing* 5.12, pp. 6241–6259. DOI: 10.3390/rs5126241.

Rotach, Mathias W, Paolo Ambrosetti, Felix Ament, Christof Appenzeller, Marco Arpagaus, Hans-Stefan Bauer, Andreas Behrendt, François Bouttier, Andrea Buzzi, Matteo Corazza, et al. (2009). "MAP D-PHASE: Real-time demonstration of weather forecast quality in the Alpine region". In: *Bulletin of the American Meteorological Society* 90.9, pp. 1321–1336. DOI: 10.1175/2009BAMS2776.1.

Saito, Kazuo, Tsukasa Fujita, Yoshinori Yamada, Jun-ichi Ishida, Yukihiro Kumagai, Kohei Aranami, Shiro Ohmori, Ryoji Nagasawa, Saori Kumagai, Chiashi Muroi, et al. (2006). "The operational JMA nonhydrostatic mesoscale model". In: *Monthly Weather Review* 134.4, pp. 1266–1298. DOI: 10.1175/MWR3120.1.

Sakai, Tetsu, Tomohiro Nagai, Toshiharu Izumi, Satoru Yoshida, and Yoshinori Shoji (2019). "Automated compact mobile Raman lidar for water vapor measurement: instrument description and validation by comparison with radiosonde, GNSS, and high-resolution objective analysis". In: *Atmospheric Measurement Techniques* 12.1, pp. 313–326. DOI: 10.5194/amt-12-313-2019.

Simmons, Adrian J and Anthony Hollingsworth (2002). "Some aspects of the improvement in skill of numerical weather prediction". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 128.580, pp. 647–677. DOI: 10.1256/003590002321042135.

Späth, Florian, Andreas Behrendt, Shravan Kumar Muppa, Simon Metzendorf, Andrea Riede, and Volker Wulfmeyer (2016). "3-D water vapor field in the atmospheric boundary layer observed with scanning differential absorption lidar". In: *Atmospheric Measurement Techniques* 9.4, pp. 1701–1720. DOI: 10.5194/amt-9-1701-2016.

Spuler, Scott M, Kevin S Repasky, Bruce Morley, D Moen, Matthew Hayman, and Amin R Nehrir (2015). "Field-deployable diode-laser-based differential absorption lidar (DIAL) for profiling water vapor". In: *Atmospheric Measurement Techniques* 8.3, pp. 1073–1087. DOI: 10.5194/amt-8-1073-2015.

Stevens, Bjorn, Sandrine Bony, David Farrell, Felix Ament, Alan Blyth, Christopher Fairall, Johannes Karstensen, Patricia K Quinn, Sabrina Speich, Claudia Acquistapace, et al. (2021). "EUREC 4 a". In: *Earth System Science Data* 13.8, pp. 4067–4119. DOI: 10.5194/essd-13-4067-2021.

Stoffelen, Ad, Jur Vogelzang, and Gert-Jan Marseille (2020). "High-resolution wind data-assimilation guide". In: URL: https://nwp-saf.eumetsat.int/site/download/documentation/scatterometer/reports/High_Resolution_Wind_Data_Assimilation_Guide_1.3.pdf.

Straaten, Chiem van, Kirien Whan, Dim Coumou, Bart van den Hurk, and Maurice Schmeits (2020). "The influence of aggregation and statistical post-processing on the subseasonal predictability of European temperatures". In: *Quarterly Journal of the Royal Meteorological Society* 146.731, pp. 2654–2670. DOI: 10.1002/qj.3810.

Tennant, Warren (2015). "Improving initial condition perturbations for MOGREPS-UK". In: *Quarterly Journal of the Royal Meteorological Society* 141.691, pp. 2324–2336. DOI: 10.1002/qj.2524.

Thundathil, Rohith, Thomas Schwitalla, Andreas Behrendt, Shravan Kumar Muppa, Stephan Adam, and Volker Wulfmeyer (2020). "Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model". In: *Journal of the Meteorological Society of Japan. Ser. II* 98.5, pp. 959–986. DOI: 10.2151/jmsj.2020-049.

Thundathil, Rohith, Thomas Schwitalla, Andreas Behrendt, and Volker Wulfmeyer (2021). "Impact of assimilating lidar water vapour and temperature profiles with a hybrid ensemble transform Kalman filter: Three-dimensional variational analysis on the convection-permitting scale". In: *Quarterly Journal of the Royal Meteorological Society* 147.741, pp. 4163–4185. DOI: 10.1002/qj.4173.

Tippett, Michael K, Jeffrey L Anderson, Craig H Bishop, Thomas M Hamill, and Jeffrey S Whitaker (2003). "Ensemble square root filters". In: *Monthly weather review* 131.7, pp. 1485–1490. DOI: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

Turner, David D and JEM Goldsmith (1999). "Twenty-four-hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement Program's 1996 and 1997 water vapor intensive observation periods". In: *Journal of Atmospheric and Oceanic Technology* 16.8, pp. 1062–1076. DOI: 10.1175/1520-0426(1999)016<1062:TFHRLW>2.0.CO;2.

Turner, DD and U Löhnert (2014). "Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI)". In: *Journal of Applied Meteorology and Climatology* 53.3, pp. 752–771. DOI: 10.1175/JAMC-D-13-0126.1.

Vannitsem, Stéphane (2017). "Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3, p. 032101. DOI: 10.1063/1.4979042.

Veselovskii, I, DN Whiteman, A Kolgotin, E Andrews, and M Korenskii (2009). "Demonstration of aerosol property profiling by multiwavelength lidar under varying relative humidity conditions". In: *Journal of Atmospheric and Oceanic Technology* 26.8, pp. 1543–1557. DOI: 10.1175/2009JTECHA1254.1.

Vié, B, G Molinié, O Nuissier, B Vincendon, V Ducrocq, F Bouttier, and E Richard (2012). "Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for Mediterranean heavy precipitating events". In: *Natural Hazards and Earth System Sciences* 12.8, pp. 2631–2645. DOI: 10.5194/nhess-12-2631-2012.

Wagner, Gerd, Andreas Behrendt, Volker Wulfmeyer, Florian Späth, and Max Schiller (2013). "High-power Ti: sapphire laser at 820 nm for scanning ground-based water–vapor differential absorption lidar". In: *Applied Optics* 52.11, pp. 2454–2469. DOI: 10.1364/AO.52.002454.

Wang, Xuguang, Dale M Barker, Chris Snyder, and Thomas M Hamill (2008a). "A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment". In: *Monthly Weather Review* 136.12, pp. 5116–5131. DOI: 10.1175/2008MWR2444.1.

Wang, Xuguang, Dale M Barker, Chris Snyder, and Thomas M Hamill (2008b). "A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments". In: *Monthly Weather Review* 136.12, pp. 5132–5147. DOI: 10.1175/2008MWR2445.1.

Wang, Xuguang and Ting Lei (2014). "GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System". In: *Monthly Weather Review* 142.9, pp. 3303–3325. DOI: 10.1175/MWR-D-13-00303.1.

Wang, Xuguang, Chris Snyder, and Thomas M Hamill (2007). "On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes". In: *Monthly Weather Review* 135.1, pp. 222–227. DOI: 10.1175/MWR3282.1.

Weckwerth, Tammy M and David B Parsons (2006). "A review of convection initiation and motivation for IHOP_2002". In: *Monthly weather review* 134.1, pp. 5–22. DOI: 10.1175/MWR3067.1.

Weckwerth, Tammy M, Kristy J Weber, David D Turner, and Scott M Spuler (2016). "Validation of a water vapor micropulse differential absorption lidar (DIAL)". In: *Journal of Atmospheric and Oceanic Technology* 33.11, pp. 2353–2372. DOI: 10.1175/JTECH-D-16-0119.1.

Weckwerth, Tammy M, Volker Wulfmeyer, Roger M Wakimoto, R Michael Hardesty, James W Wilson, and Robert M Banta (1999). "NCAR–NOAA lower-tropospheric water vapor workshop". In: *Bulletin of the American Meteorological Society* 80.11, pp. 2339–2357. URL: https://www.jstor.org/stable/26215081.

Wikle, Christopher K and L Mark Berliner (2007). "A Bayesian tutorial for data assimilation". In: *Physica D: Nonlinear Phenomena* 230.1-2, pp. 1–16. DOI: 10.1016/j.physd.2006.09.017.

Wu, Wan-Shu, David F Parrish, Eric Rogers, and Ying Lin (2017). "Regional ensemble–variational data assimilation using global ensemble forecasts". In: *Weather and Forecasting* 32.1, pp. 83–96. DOI: 10.1175/WAF-D-16-0045.1.

Wulfmeyer, Volker, Hans-Stefan Bauer, Matthias Grzeschik, Andreas Behrendt, Francois Vandenberghe, Edward V Browell, Syed Ismail, and Richard A Ferrare (2006). "Four-dimensional variational assimilation of water vapor differential absorption lidar data: The first case study within IHOP_2002". In: *Monthly weather review* 134.1, pp. 209–230. DOI: 10.1175/MWR3070.1.

Wulfmeyer, Volker, Andreas Behrendt, Christoph Kottmeier, Ulrich Corsmeier, Christian Barthlott, George C Craig, Martin Hagen, Dietrich Althausen, Fumiko Aoshima, Marco Arpagaus, et al. (2011). "The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights". In: *Quarterly Journal of the Royal Meteorological Society* 137.S1, pp. 3–30. DOI: 10.1002/qj.752.

Wulfmeyer, Volker and Jens Bösenberg (1998). "Ground-based differential absorption lidar for water-vapor profiling: assessment of accuracy, resolution, and meteorological applications". In: *Applied Optics* 37.18, pp. 3825–3844. DOI: 10.1364/AO.37.003825.

Wulfmeyer, Volker, R Michael Hardesty, David D Turner, Andreas Behrendt, Maria P Cadeddu, Paolo Di Girolamo, Peter Schlüssel, Joël Van Baelen, and Florian Zus (2015). "A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles". In: *Reviews of Geophysics* 53.3, pp. 819–895. DOI: 10.1002/2014RG000476.

Wulfmeyer, Volker, Shravan Kumar Muppa, Andreas Behrendt, Eva Hammann, Florian Späth, Zbigniew Sorbjan, David D Turner, and R Michael Hardesty (2016). "Determination of convective boundary layer entrainment fluxes, dissipation rates, and the molecular destruction of variances: Theoretical description and a strategy for its confirmation with a novel lidar system synergy". In: *Journal of the Atmospheric Sciences* 73.2, pp. 667–692. DOI: 10.1175/JAS-D-14-0392.1.

Yoshida, Satoru, Sho Yokota, Hiromu Seko, Tetsu Sakai, and Tomohiro Nagai (2020). "Observation system simulation experiments of water vapor profiles observed by Raman lidar using LETKF system". In: *Sola*. DOI: 10.2151/sola.2020-008.

Zhang, Fuqing, Y Qiang Sun, Linus Magnusson, Roberto Buizza, Shian-Jiann Lin, Jan-Huey Chen, and Kerry Emanuel (2019). "What is the predictability limit of midlatitude weather?" In: *Journal of the Atmospheric Sciences* 76.4, pp. 1077–1091. DOI: 10.1175/JAS-D-18-0269.1.

Zhang, Fuqing, Meng Zhang, and Jonathan Poterjoy (2013). "E3DVar: Coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar". In: *Monthly Weather Review* 141.3, pp. 900–917. DOI: 10.1175/MWR-D-12-00075.1.

Details of the own contribution

Publication-I

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Shravan Kumar Muppa, Stephan Adam, and Volker Wulfmeyer, 2020: "**Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model**". *Journal of the Meteorological Society of Japan*, 98(5), 959-986. DOI: 10.2151/jmsj.2020-049

Contribution: The research work leading to the publication was conceived by the doctoral candidate, Rohith Thundathil, and the supervisor, Prof. Dr Volker Wulfmeyer. The doctoral candidate provided the following contributions:

1. Design and development of the forward operator relevant for assimilating the thermodynamic profile observations obtained from the ground-based lidar systems into the numerical model.

2. Setup and configuration of the whole numerical model setup.

3. Automated design of the data assimilation (DA) architecture in the rapid update cycle (RUC) mode.

4. Conducting the simulations relevant to the research.

5. Performing the post-processing and analysis of the results.

6. Preparation and editing of the structure and final form of the manuscript.

Publication-II

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, and Volker Wulfmeyer, 2021: "**Impact of Assimilating Lidar Water Vapour and Temperature Profiles with a Hybrid Ensemble Transform Kalman Filter – Three-Dimensional Variational Analysis on the Convection Permitting Scale** ". *Quarterly Journal of the Royal Meteorological Society*, 147(741), 4163-4185. DOI: 10.1002/qj.4173

Contribution: The research study resulting in the above publication was conceived by the doctoral candidate, Rohith Thundathil, and the supervisor, Prof. Dr Volker Wulfmeyer. The doctoral candidate contributed to the following tasks:
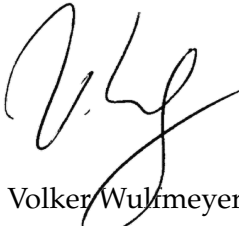
1. Design and configuration of the numerical model setup necessary for the research study. Setting up an ensemble-based hybrid **3DVar-ETKF** DA system to assimilate the lidar temperature and moisture profile observations.

2. Completely automated code design of the hybrid DA system in the RUC mode.

3. Conducting the simulations relevant to the research study.

4. Performing post-processing and analysis of the results.

5. Preparation and editing of the structure and final form of the manuscript.

Publication-III

Cyrille Flamant, Patrick Chazette, Olivier Caumont, Paolo Di Girolamo, Andreas Behrendt, Michaël Sicard, Julien Totems, Diego Lange, Nadia Fourrié, Pierre Brousseau, Clotilde Augros, Alexandre Baron, Marco Cacciani, Adolfo Comerón, Benedetto De Rosa, Veronique Ducrocq, Pascal Genau, Laurent Labatut, Constantino Muñoz-Porcar, Alejandro Rodríguez-Gómez, Donato Summa, **Rohith Thundathil** and Volker Wulfmeyer, 2021: "**A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative**". *Bulletin of Atmospheric Science and Technology*, 2,10. DOI: 10.1007/s42865-021-00037-6

Contribution: The publication was prepared and published by Dr Cyrille Flamant, detailing the initiative of a future research campaign called the Water Vapor Lidar Network Assimilation (WaLiNeAs). The paper describes the application of assimilating moisture profile data from a network of ground-based Raman lidars stationed at various locations of Southern France across the Mediterranean coast. The initiative aims to enhance the forecast of HPEs. The doctoral candidate, Rohith Thundathil, played a significant role in providing scientific knowledge transfer regarding ground-based thermodynamic lidar profile DA into numerical models for potential improvements in short-range forecasting.

Stuttgart, 15.03.2022
Place, Date

Prof. Dr Volker Wulfmeyer
Supervisor

Acknowledgement

I want to express my sincere gratitude to:

- Prof. Dr Volker Wulfmeyer for his extended support and guidance till I made the final march to the end gate of my doctoral thesis research. I could learn many problem tackling skills from his long academic career expertise. I was struck by some of his innovative ideas, which gave me a different perspective towards approaching a problem. I am honoured to have had the support of an experienced person like him throughout my doctoral thesis journey.

- Dr Thomas Schwitalla was my immediate source of help regarding research and technical problems. His knowledge of codes, high-performance computing, and diverse skills and experience in code debugging helped me a lot. I am also grateful for his valuable contribution to the editing and reviewing of my manuscripts. Personally, he has been a very good friend.

- Dr Andreas Behrendt for providing immense support during the entire course of my research. He played a crucial role in formulating research ideas, analysing results, identifying and solving research problems. Being an engineer at the core, I found many difficulties developing thoughts from a researcher point of view, and I could learn this from him. His constant guidance and support in editing, analysing, and reviewing the manuscripts was a huge help. Apart from research, he helped me familiarise myself with life in Germany and made sure I didn't face any troubles during my initial days.

- Dr Kirsten Warrach-Sagi for providing valuable knowledge in modelling research, technical and scientific support and critical comments and analysis.

- Dr Hans-Stefan Bauer for providing support in the supercomputer computing hours allocations and preparing the report and publications related to high-performance computing. I still cherish the halloween party get-togethers at his place.

- Dr Shravan Kumar Muppa for the valuable guidance in the data processing and error calculations regarding the lidar thermodynamic observations from the Raman lidar and the differential absorption lidar.

Curriculum vitae

PERSONAL DATA

Name:	Rohith Muraleedharan Thundathil
Date of birth:	10 January 1991
Place of birth:	Kodungallur, Kerala, India
Citizenship:	Indian
Email:	rohithmuraleedharan@gmail.com

EDUCATION

04/2017 - Present
Institute of Physics and Meteorology,
University of Hohenheim
Stuttgart, Germany

Research Associate and PhD Scholar

08/2014 - 05/2016
Earth System Sciences
Indian Institute of Space Science and
Technology, Thiruvananthapuram, India

Master of Technology

08/2009 - 05/2013
Electronics and Communication
Mahatma Gandhi University
Kerala, India

Bachelor of Technology

08/2008 - 05/2009
Bharatiya Vidya Bhavan, Kerala
Central Board of Secondary Education
India

Higher Secondary Education

08/2006 - 05/2007
Bharatiya Vidya Bhavan, Kerala
Central Board of Secondary Education
India

Secondary Education

RESEARCH EXPERIENCE

04/2017 - Present
Institute of Physics and Meteorology,
University of Hohenheim, Stuttgart
Germany

Research Associate Position
Convective-Scale Data Assimilation of
Thermodynamic Lidar Data into the
Weather Research and Forecasting Model

05/2015 - 05/2016
Earth System Sciences
Indian Institute of Space Science and
Technology, Thiruvananthapuram
India

Master Thesis
Comparing Limited Area Hybrid ETKF-3DVAR
and 3DVAR Data Assimilation Method for
Tropical Cyclone Track Forecasts:
Sensitivity to Model Error

06/2015 - 07/2015
National Atmospheric
Research Laboratory, Gadanki
Andhra Pradesh, India

Master Thesis Internship
Professional training in
Weather Research and Forecasting Model
and Data Assimilation Techniques

06/2012 - 07/2012
Bharath Sanchar Nigam Limited
Kerala, India

Bachelor Internship
Vocational training in

Long Distance Communication

TECHNICAL SKILLS

Programming Languages: FORTRAN, Shell script (Bash), IDL, C, Python
Meteorological Data Analysis tools: NCAR Command Language
Atmospheric model: Weather Research and Forecasting Model
Other experiences:

HPC Coding, NetCDF, GRIB, BUFR, BSCAN

PEER-REVIEWED PUBLICATIONS

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, and Volker Wulfmeyer, 2021: **"Impact of Assimilating Lidar Water Vapour and Temperature Profiles with a Hybrid Ensemble Transform Kalman Filter – Three-Dimensional Variational Analysis on the Convection Permitting Scale "**. *Quarterly Journal of the Royal Meteorological Society*, 147(741), 4163-4185. DOI: 10.1002/qj.4173

Cyrille Flamant, Patrick Chazette, Olivier Caumont, Paolo Di Girolamo, Andreas Behrendt, Michaël Sicard, Julien Totems, Diego Lange, Nadia Fourrié, Pierre Brousseau, Clotilde Augros, Alexandre Baron, Marco Cacciani, Adolfo Comerón, Benedetto De Rosa, Veronique

Ducrocq, Pascal Genau, Laurent Labatut, Constantino Muñoz-Porcar, Alejandro Rodríguez-Gómez, Donato Summa, **Rohith Thundathil** and Volker Wulfmeyer, 2021: "**A network of water vapor Raman lidars for improving heavy precipitation forecasting in southern France: introducing the WaLiNeAs initiative**". *Bulletin of Atmospheric Science and Technology*, 2,10. DOI: 10.1007/s42865-021-00037-6

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Shravan Kumar Muppa, Stephan Adam, and Volker Wulfmeyer, 2020: "**Assimilation of Lidar Water Vapour Mixing Ratio and Temperature Profiles into a Convection-Permitting Model**". *Journal of the Meteorological Society of Japan*, 98(5), 959-986. DOI: 10.2151/jmsj.2020-049

Govindan Kutty, **Rohith Thundathil**, and Amit P. Kesarkar, 2018: "**Impact of Representing Model Error in a Hybrid Ensemble-Variational Data Assimilation System for Track Forecast of Tropical Cyclones over the Bay of Bengal**". *Pure and Applied Geophysics*, 175(3), 1155-1167. DOI: 10.1007/s00024-017-1747-z

BOOK CHAPTER

Hans-Stefan Bauer, Thomas Schwitalla, Oliver Branch, **Rohith Thundathil**, Stephan Adam, and Volker Wulfmeyer, 2021: "**WRF Simulations to Investigate Processes Across Scales (WRFSCALE)**". *High Performance Computing in Science and Engineering '20*, 469-486. DOI: 10.1007/978-3-030-80602-6_31

CONFERENCE CONTRIBUTIONS

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Diego Lange, Cyrille Flamant, Olivier Caumont, David Turner, and Volker Wulfmeyer, 2022: "**The impact of Ensemble-Based Assimilation of Thermodynamic Lidar Profiles on Forecasts of the Pre-Convective Environment and Convection Initiation of Heavy Precipitation Events over the Mediterranean region: The German contribution to WaLiNeAs**". *DACH2022, Leipzig, Deutschland*, 21–25 Mar 2022, DACH2022-238. DOI: 10.5194/dach2022-238

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Diego Lange, Florain Späth, Volker Wulfmeyer, Daniel Leuenberger, Alexander Haefele, Marco Arpagaus, and Martucci Giovanni, 2021: "**Investigation of the impact of thermodynamic profiles of ground based lidar systems on short-range forecast skill by means of ETKF-hybrid 3DVAR data assimilation**". *Session GI4.2 – Lidar remote sensing of the atmosphere*, EGU21-9467, Tue, 27 Apr, 16:09–16:11, vEGU21, vPICO presentations. DOI: 10.5194/egusphere-egu21-9467

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Diego Lange, Florain Späth, Volker Wulfmeyer, Daniel Leuenberger, Alexander Haefele, Marco Arpagaus, and Martucci Giovanni, 2021: "**Impact of ground-based water vapour and temperature lidar profiles on short-range forecast skill by means of hybrid 3DVAR-ETKF data assimilation**". *THE FIRST JOINT WCRP-WWRP SYMPOSIUM ON DATA ASSIMILATION AND REANALYSIS*, Thu, 16 Sep, Poster Session. Abstract

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Volker Wulfmeyer, Daniel Leuenberger, Alexander Haefele, Marco Arpagaus, and Martucci Giovanni, 2019: "**Assimilation of water vapor mixing ratio profiles into the WRFDA System with a new forward operator**". *Ninth Symposium on Lidar Atmospheric Applications, Phoenix Convention Center, American Meteorological Society*, Thu, 10 Jan, 11:15-11:30. Abstract and Presentation

Alexander Haefele, Daniel Leuenberger, Marco Arpagaus, Martucci Giovanni, **Rohith Thundathil**, Thomas Schwitalla, Andreas Behrendt, and Volker Wulfmeyer, 2019: "**Assimilation of Temperature and Humidity Profiles from a Raman Lidar into Convective-Scale NWP Models and the Impact on the Forecast**". *Ninth Symposium on Lidar Atmospheric Applications, Phoenix Convention Center, American Meteorological Society*, Thu, 10 Jan, 11:30-11:45. Abstract and Presentation

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Shravan Kumar Muppa, Stephan Adam, and Volker Wulfmeyer, 2018: "**A new forward operator for assimilating water-vapor mixing ratio profiles into the WRF-DA system and its impact study**". *Joint WRF/MPAS Users' Workshop, Boulder, USA*, Presented by Dr. Hans Stefan Bauer, P14, Poster Session. Abstract

Rohith Thundathil, Thomas Schwitalla, Andreas Behrendt, Shravan Kumar Muppa, Stephan Adam, and Volker Wulfmeyer, 2018: "**First Direct Assimilation of Lidar Water Vapor Mixing Ratio Profiles Into the WRF-DA System**". *International Symposium on Data Assimilation 2018, Ludwig-Maxmilians University, Munich, Germany*, P5.2, Poster Session. Poster

Govindan Kutty, and **Rohith Thundathil**, 2018: "**Impact of Model Error Representation in a Hybrid Ensemble-Variational Data Assimilation System for Track Forecast of Tropical Cyclone Hudhud (2014)**". *21st Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, Washington State Convention Center, American Meteorological Society*, 632, Poster Session. Abstract and Poster

9 March 2022

Rohith Muraleedharan Thundathil

Declaration

I hereby declare that I have prepared this dissertation independently and have used only the sources and aids indicated. Verbatim or textual passages are marked as such.

Stuttgart, 09.03.2022

Rohith Muraleedharan Thundathil

138

Annex 3

Declaration in lieu of an oath on independent work

according to Sec. 18(3) sentence 5 of the University of Hohenheim's Doctoral Regulations for the Faculties of Agricultural Sciences, Natural Sciences, and Business, Economics and Social Sciences

1. The dissertation submitted on the topic

Convective-Scale Data Assimilation of Thermodynamic Lidar Data

into the Weather Research and Forecasting Model

is work done independently by me.

2. I only used the sources and aids listed and did not make use of any impermissible assistance from third parties. In particular, I marked all content taken word-for-word or paraphrased from other works.

3. I did not use the assistance of a commercial doctoral placement or advising agency.

4. I am aware of the importance of the declaration in lieu of oath and the criminal consequences of false or incomplete declarations in lieu of oath.

I confirm that the declaration above is correct. I declare in lieu of oath that I have declared only the truth to the best of my knowledge and have not omitted anything.

Stuttgart, 09.03.2022

Place, Date   Signature