universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

High-Resolution State Estimation Using Doppler Wind LIDAR Observations, ICON-LAM and an Ensemble Kalman Filter

verfasst von | submitted by

Ludwig Wolfgruber BSc BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 614
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Meteorology
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Martin Weißmann
Mitbetreut von | Co-Supervisor:	Dott. Dott. ric. Stefano Serafin

Acknowledgements

This project could not have been completed by one on his one. I want to thank the FE11 and FE12 teams at DWD, Stefanie Hollborn, Daniela Littmann, Hendrik Reich, Thorsten Steinert, Jasmin Vural, Günther Zängl and foremost Gernot Geppert. I also want to thank my supervisors for their guidance and especially Stefano Serafin for the intense support.

I would also like to acknowledge the support of my wife Viktorija and my parents, without whom my student career would not have been possible.

Thank you!

Abstract

Data assimilation allows for the direct integration of meteorological observations into atmospheric models, significantly improving numerical weather prediction. This technique, along with the continued incorporation of new types of observational data, has led to significant advances. However, the assimilation of data and the generation of accurate forecasts in regions with complex topography, such as mountainous areas, remain particularly challenging. This thesis explores an innovative approach: the direct assimilation of Doppler LIDAR radial velocity measurements into a high-resolution weather model to improve the state estimation of the atmosphere, with a focus on valley circulation. Such a state estimation – viewed even as a local campaign data reanalysis can combine and supplement the information from campaign observations with physically consistent model simulations.

The study uses observations from the CROSSINN campaign in the Inn Valley to refine the representation of meso and micro scale alpine atmospheric processes. These thermally driven winds are difficult to accurately represent and forecast in current numerical weather prediction models. This work examines whether a direct assimilation of radial velocity can provide a more accurate state estimation of the atmosphere compared with the conversion into vertical profiles of horizontal wind using the Velocity Azimuth Display (VAD) technique. It also evaluates the effectiveness of combining LIDAR data with operationally used sources such as ground stations, radiosondes, and air reports.

The results show a complex picture. Depending on the validation method used, the direct assimilation of Doppler LIDAR data can lead to a reduction in analysis and forecast errors, thus improving the predictive accuracy of the model. In particular, the combination of multiple data sources produces the most robust results, highlighting the value of integrating diverse observational data to capture the full complexity of atmospheric conditions. This approach can not only improve the understanding and representation of thermally driven valley flow, but also provides valuable context for observations from the CROSSINN campaign and future observational efforts, potentially leading to more accurate weather forecasts in complex terrain regions.

Kurzfassung

Datenassimilation ermöglicht die direkte Integration von meteorologischen Beobachtungen in atmosphärische Modelle, wodurch die numerische Wettervorhersage erheblich verbessert wird. Diese Technik hat zusammen mit der kontinuierlichen Einbeziehung neuer Arten von Beobachtungsdaten zu erheblichen Fortschritten in der Wettervorhersage geführt. Die Assimilierung von Daten und die Erstellung genauer Vorhersagen in Regionen mit komplexer Topographie, wie z. B. in Gebirgsregionen, stellen jedoch nach wie vor eine besondere Herausforderung dar. In dieser Arbeit wird ein innovativer Ansatz untersucht: die direkte Assimilation von Doppler-LIDAR-Radialgeschwindigkeitsmessungen in ein hochauflösendes Wettermodell, um die Zustandsschätzung der Atmosphäre zu verbessern. Eine solche Zustandsschätzung - auch als lokale Reanalyse von Kampagnendaten betrachtet - kann die Informationen aus Kampagnenbeobachtungen mit physikalisch konsistenten Modellsimulationen kombinieren und ergänzen.

Die Studie nutzt Beobachtungen aus der CROSSINN-Kampagne im Inntal, um die Darstellung meso- und mikroskaliger alpiner atmosphärischer Prozesse zu verfeinern. Diese thermisch getriebenen Winde lassen sich in den aktuellen numerischen Wettervorhersagemodellen nur schwer genau darstellen und vorhersagen. Es wird untersucht, ob eine direkte Assimilation von Radialgeschwindigkeiten eine genauere Zustandsschätzung der Atmosphäre liefern kann, als die Umwandlung in vertikale Profile des horizontalen Windes unter Verwendung der Velocity Azimuth Display (VAD) Technik. Außerdem wird die Wirksamkeit der Kombination von LIDAR-Daten mit operationell genutzten Datenquellen wie Bodenstationen, Radiosonden und Luftfahrtberichten bewertet.

Die Ergebnisse zeigen ein differenziertes Bild. Je nach den verwendeten Validierungsdaten kann die direkte Assimilierung von Doppler-LIDAR-Daten zu einer Verringerung der Analyse- und Vorhersagefehler führen und damit die Vorhersagegenauigkeit des Modells verbessern. Insbesondere die Kombination mehrerer Datenquellen führt zu den robustesten Ergebnissen und unterstreicht den Wert der Integration verschiedener Beobachtungsdaten, um die gesamte Komplexität der atmosphärischen Bedingungen zu erfassen. Dieser umfassende Ansatz verbessert nicht nur das Verständnis und die Darstellung des thermischen Berg- und Talwindes, sondern bietet auch einen erweiterten Kontext für die Beobachtungen der CROSSINN-Kampagne und künftige Messkampagnen, was zu genaueren Wettervorhersagen im alpinen Raum führen kann.

Contents

1. Introduction

During a meteorological field campaign, a large amount of data is created using high-frequency remote-sensing instruments. This data is then processed and analyzed to investigate specific questions. Since the analysis is only done with a focus on the those questions, and parts of the measured timelines are filtered out due to meteorological conditions that are unsuitable to the problem, only a fraction of the available data is used. The unused data represents a lost opportunity for further insights that could be gained from the campaign. For example, the data could provide a more comprehensive understanding of atmospheric conditions. Therefore, it is useful to find ways to utilize this data effectively.

If individuals want to access and work with the data of different measurement devices separately, a lot of resources are used to at least interpolate the data if not to set it into a broader context, combining them with other information about the state of the atmosphere. This calls for an objective analysis tool which can be applied centrally, and not only prepare the data in an accessible way, but also enrich it using meteorological models.

Numerical simulations are performed to support the analysis of campaign data by providing a complete, physically consistent state estimation of the atmosphere in the form of gridded output variables with regular timing. To provide this at a suitable scale a large eddy simulation (LES) with grid lengths in the 10 to 100 meters or a high resolution simulation in a kilometer scale can be performed. Such simulations usually do not use campaign observations as just easy to assimilate observations are included. The simulation is rather nested in a model domain surrounding it and receives lateral boundary conditions. For such setups uncertainties originating from inadequate initial and boundary data can not be cleared out, since the observations have not much influence.

One weakness of numerical atmospheric modelling still is the representation of the planetary boundary layers in alpine regions (Goger et al. 2016). Over so called complex orography parametrisations for unresolved processes loose their skill as they are build on assumptions of flat terrain. The higher the resolution of a simulation, the better the model can resolve the orography and less help from parametrisations is needed. The assimilation of observations which capture the structure of the boundary layer could improve its state estimation leading to better forecasts.

The idea of this master's thesis is to assimilate high resolution campaign data in a high resolution simulation. This gives an estimate of the state of the atmosphere with the benefits and comparability of a simulation output while representing the added information of the campaign. Data which is measured with different methods and is available in different formats can be efficiently combined in an analysis. This approach allows for a more comprehensive understanding of the meteorological conditions and

for further analysis of the matter of interest of the campaign itself. When the data is processed in such a way it is easier and less time consuming to analyze it further. Additionally, the results can be shared with a broader community in an easily accessible format, making the data of the field campaign more widely available. In a way a product consisting of analysis and intermediate forecasts as state estimations could be thought of as a high resolution reanalysis as it is known from climate sciences.

While data assimilation (DA) is used operationally in NWP simulations in global and limited area models it has been rarely used with simulations with smaller scales. Miyoshi et al. (2016) assimilate spatially and temporally dense RADAR measurements using a Kalman filter to improve severe weather forecasts at temporal scales of 30 s and a grid length of 100 m. Although the context is different from the problem set here this can be viewed as a proof of concept and technical feasibility of assimilating massive amounts of spatially and temporally dense data.

In a similar approach while presenting lateral boundary adjustment to improve short time severe weather forecasting Sawada et al. 2015 assimilate LIDAR wind profiles into a sub kilometre scale simulation with an update frequency of 15 min.

Another approach combining field data and simulations comes from LASSO (Gustafson Jr et al. (2020)), an LES to reconstruct the flow around atmospheric radiation measurement sites of the U.S. Department of Energy. They operate a large eddy model nested in a mesoscale simulation. LASSO tries a so-called symbiotic simulation and observation, where both LES model in- and output and observational data can be compared to develop parametrisations for coarser atmospheric models. The LES does not get any information on the observed state of the atmosphere, so it could gain accuracy and further value for its purpose if this data would be assimilated. LASSO represents a use case of enriching LES with high-density measurement data. Similarly Neggers et al. 2012 compare LES results with observations to reproduce fast and small scale atmospheric physics in coarser atmospheric models.

More recently in the work of Pirk et al. 2022 drone measurements of meteorological variables are assimilated using different particle smother schemes to receive updated surface flux estimates. The temporal and spatial scales of both the measurement and the model realisations are of the same order of magnitude. In contrast to the goal of this work, the data are not assimilated into a LES, but the simulation is used as a forward model.

Several approaches where tested to assimilate LIDAR wind profiles into convection scale or higher resolution models. Yang et al. 2022 demonstrates that the assimilation using the LETKF outperforms nudging. How the combination of high resolution LIDAR profiles with bigger scale RADAR radial velocities works is shown by Wang et al. 2022, where a separate assimilation of the different observations at different scales is recommended. Finn et al. 2020 look at LETKF assimilation of LIDAR profiles in a high resolution simulation improving wind forecasts in the atmospheric boundary layer.

As described above there is potential usage and technical feasibility to use a Kalman filter DA to assimilate measurements that are temporally and spatially dense into a high resolution simulation. The use of Doppler LIDAR radial velocities instead of the

by now widely used vertical profiles of the horizontal wind components appears not well researched. This could be especially useful to estimate the state of the boundary layer over complex terrain as LIDARS can gather wind information from a few square kilometres around the observation site in various atmospheric layers. The main scientific question that this master's thesis aims to address is:

> Can we improve the estimation of the state of the atmosphere and its forecast over alpine terrain with a high resolution atmospheric model using Doppler LIDAR radial velocity campaign data and the LETKF?

The goal is to develop and test a methodology that allows for the assimilation of a big amount of temporally and spatially high resolution campaign data into a high resolution model, in order to obtain a more comprehensive and accurate representation of the meteorological conditions. This approach can enable the gain of additional insights from the campaign data that were previously unavailable, and thus provide a better representation and understanding of the atmosphere at smaller scales.

2. Methods and Data

2.1. CROSSINN

As a case study a subset of the CROSSINN campaign (Adler et al. 2021) is selected. The campaign was performed in the second half of 2019 in the lower Tyrolean Inn Valley. It aimed at investigating cross-valley flows and atmospheric exchange processes in the atmospheric boundary layer over mountainous terrain using Doppler LIDAR measurements amongst others. During 10 intensive operation periods (IOPs) various types of observations were performed. For this thesis the measurements from IOP 8 which took place on 13 September at the observation site Kolsass (11.6222° E, 47.3053° N) are selected. From different instruments at this site only two windcube Doppler LIDAR instruments and one sonic anemometer are considered which were operated by the University of Innsbruck. The LIDARS measure radial velocity in plan position indicator (PPI) scans. The data are available at Gohm et al. 2021 and in the i-Box data repository (Rotach et al. 2017).

2.1.1. Synoptic Situation during IOP 8 and Valley Circulation

The observation period IOP 8 is selected due to its synoptic features. A map of temperature, geopotential height and vertical wind on 2019-09-13 is displayed in Fig. 1. A low intensity high pressure system lies over western and central Europe between two more distinct low pressure systems over Scandinavia and the Atlas mountains. This creates a negligible pressure gradient across the Alps. The situation is stationary over the Alps throughout the following two days. Neither are Föhn nor clouds present.

Due to the lack of synoptic forcing the atmospheric flow in the valley atmosphere is dominated by thermally driven valley winds. These wind regimes follow a diurnal cycle and can be separated in valley winds and slope winds (Markowski et al. 2010). In general both winds flow up slope and up valley during the day, but valley winds lag behind due to their bigger inertia. Both phenomena are a challenge for numerical weather prediction with valley winds being resolved in big valleys as the Inn Valley. Slope winds are too small in scale and too local in effect so their effects are parameterised.

Fig. 2 shows the column of the u-components of the wind from model simulations and observations at Kolsass. From 0 UTC until 12 UTC (00:55 to 12:55 local solar time (LST)[1]) it is positive throughout the column. This westerly wind flows down valley in the Inn Valley. It reaches its maximum at 7 UTC in about 200 m above ground. At 12 UTC wind direction switches and the u-component is now negative throughout the

[1]At September 13 local solar time (LST) at Kolsass is 51 min (roughly one hour) ahead of UTC time.

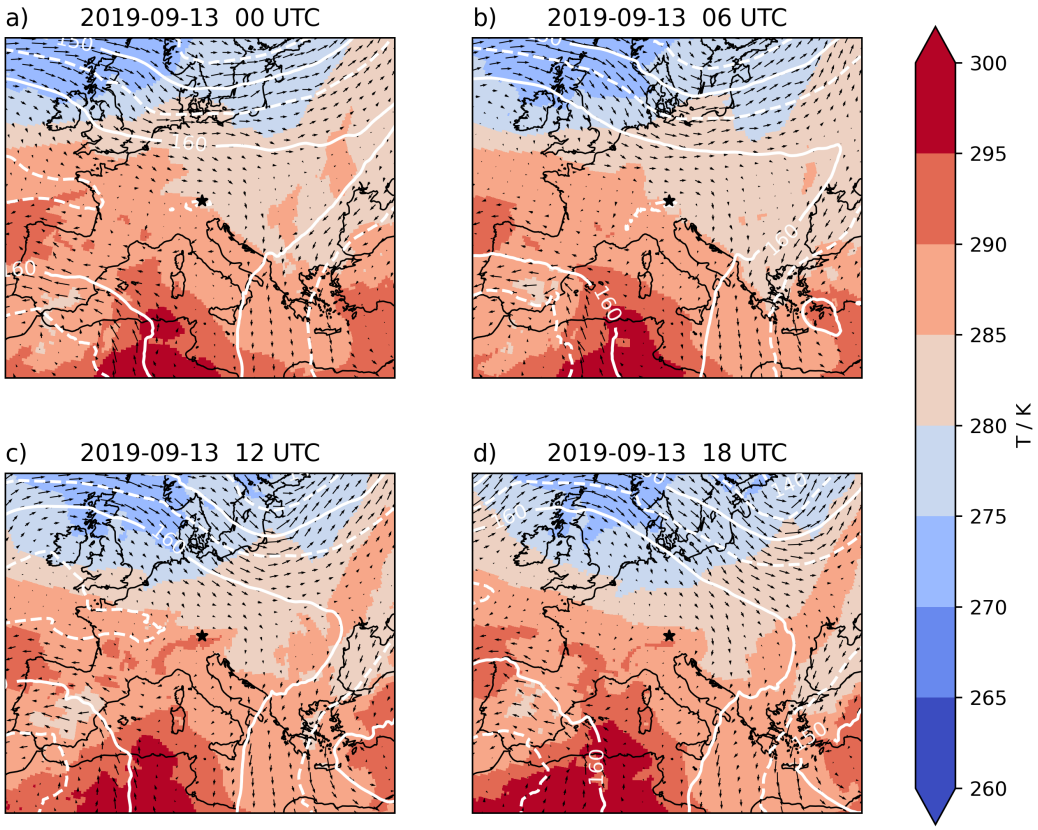


Figure 1.: Plot of temperature T in K, isolines of geopotential height z in dm and arrows of vertical wind, all in 850 hPa during CROSSINN IOP 8. $\star$ marks the location of Kolsass. The data is taken form the ERA5 reanalysis (Hersbach et al. 2023).

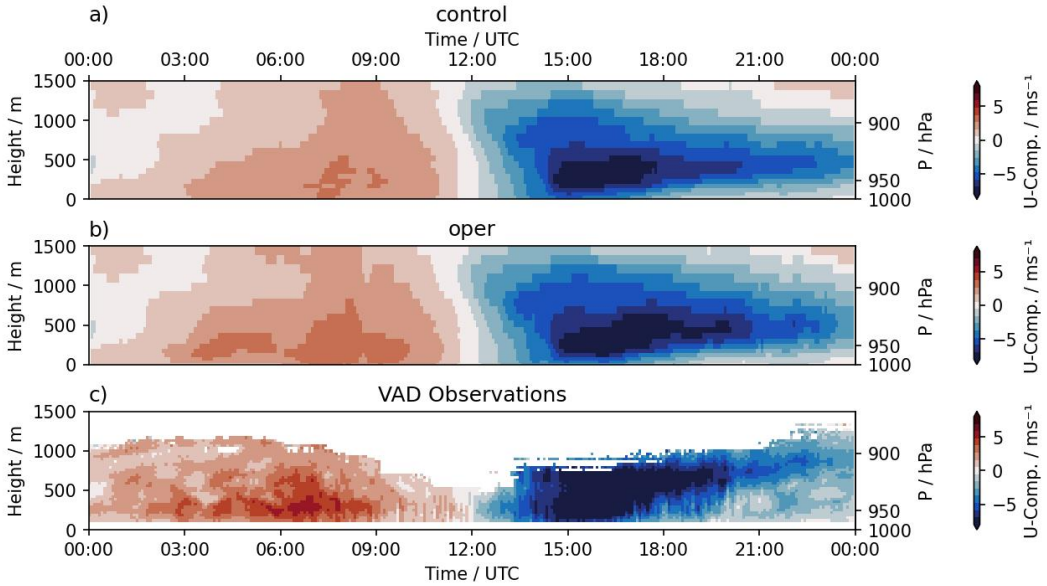


Figure 2.: Profiles of u-component at Kolsass over the full day of 2019-09-13. Comparison of (a) the CONTROL run, (b) the operational setup and (c) the VAD observations.

column, which corresponds to easterly up valley flow. As displayed in Fig. 3 the speed of this flow reaches its maximum between 16 and 17 UTC. Afterwards the flow slows down and detaches from the ground, as a stable boundary layer forms beneath it.

The results from model simulation without any DA (CONTROL) and the cycled experiment which assimilates the data operationally used by the German Weather Service (DWD) for this region (CYCLE with `oper`) are also shown in figures 2 and 3. The orientation of valley flow and the timing in its transitioning are met well by both simulations. Instead the magnitude in speed and the the growth of the nocturnal stable layer are both too low in the model simulations.

2.1.2. Devices and Data

The data from three instruments deployed during CROSSINN is used for this study: Two Doppler LIDARs and one sonic anemometer.

The LIDARs are one Streamline type and one Streamline XR type from the manufacturer 'Halo Photonics'. The first LIDAR made Plan Position Indicator (PPI) scans with an elevation angle of $\eta = 70°$. With one 360° azimuth angle scan every minute it produced one full scan of radial velocity every minute. From these scans the wind components are computed with a Velocity Azimuth Display (Browning et al. 1968). This gives profiles of the horizontal wind components which are the most used output of LIDAR measurements. Due the high reduction in data points when creating VAD-profiles there is also a potentially high loss of information. Because of the high elevation angle the dataset with radial velocity from this LIDAR will be called `radvel high`. The dataset

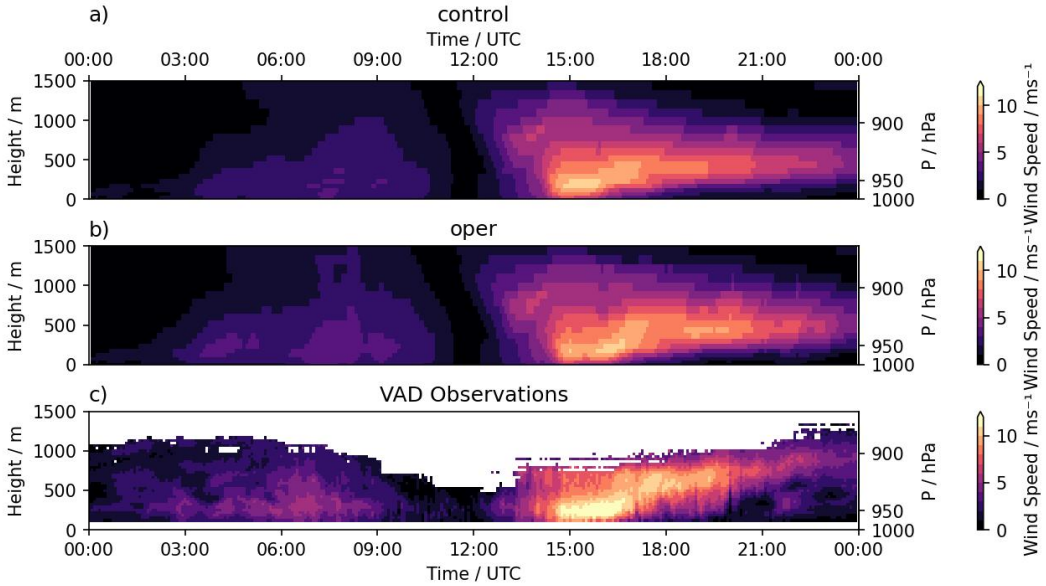


Figure 3.: Profiles of vertical wind speed at Kolsass over the full day of 2019-09-13. Comparison of (a) the CONTROL run, (b) the operational setup and (c) the VAD observations

containing the velocity components from the VAD will be called `vad`.

The second LIDAR has a more complex scanning routine. During the first half of each hour it performs vertical stare scans. Every full and half hour it makes PPI scans with an elevation angle of $\eta = 70°$. In the second half of each hour it performs PPI scans with alternating 4° and 7° elevation. We only use the PPI scans with $\eta = 4°$ and $\eta = 7°$. Due to the LIDAR range these scans give information about the low parts of the boundary layer in heights between about 10 to 240 m above ground. The dataset containing these radial velocity values will be called `radvel low`.

The combination of `radvel low` and `vad` datasets is called `all lidar`, as it combines the information of both Doppler-LIDAR instruments. The combination of `radvel low` and `radvel high` is called `all radvel`. All these datasets can be used for assimilation and validation alike.

The third type of data are horizontal wind measurements from a sonic anemometer on a mast 16 m above ground. It is part of the permanently installed i-box measurement site at Kolsass (Rotach et al. 2017). The minutely values are averaged around the closest 5 minute step. The dataset is used only for validation and is called `surface`.

For comparison with the current state of atmospheric prediction the operationally by DWD assimilated data is also used as mentioned above. It consists of measurements from surface stations (SYNOP), balloon measurements (TEMP and PILOT) and air reports (AIREP). A map of these observations is shown in Fig. 4. Note that this set not only includes wind observations, but also temperature, humidity and pressure level

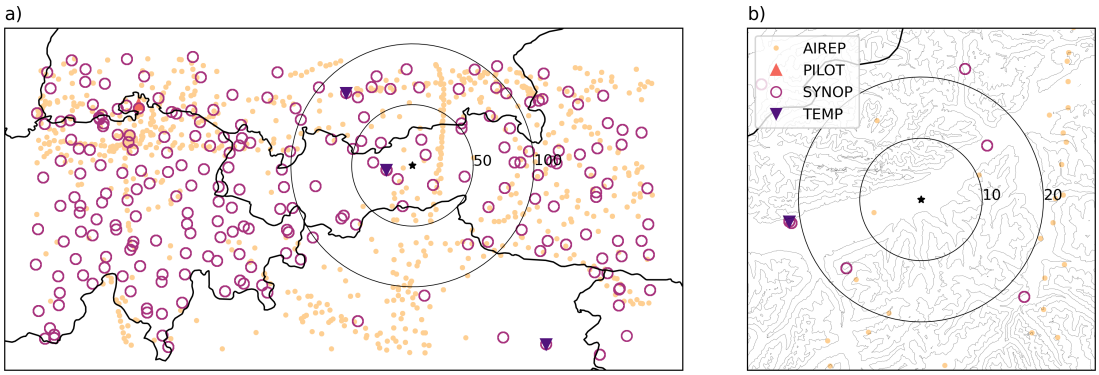


Figure 4.: Map of active `oper` observations on 2019-09-13. (a): All observations, (b): detail around the observation site Kolsass (marked with ⋆) showing 500 m isolines of topography.

height observations are assimilated. To identify observations which only include wind measurements a version of this plot showing only those observations is presented in Fig. 20 in the appendix. The assigned observation error variances for this data remain as in the operational setup. Together the dataset is called `oper`. The combination with `all lidar` is called `all` as it uses information from all sources.

2.2. Numerical Model

2.2.1. ICON Limited Area Model

As numerical weather model a limited area model (LAM) variant of the *Icosahedral Non-hydrostatic Model* (ICON), developed by DWD and the Max Planck Institute for Meteorology is available. As its name says the model is solving the non-hydrostatic equations on a grid which is based on the approximation of earth as an icosahedron (Zängl et al. 2015). This grid can be refined forming triangular grids. Triangular grids are advantageous in simulations over complex terrain as they offer a higher resolution of the orography compared to more common Cartesian grids of the same grid length.

At the DWD the ICON-LAM is used to operationally produce weather forecasts for domains covering Europe and the German speaking countries. Such simulations are performed at resolutions of 6.5 km as ICON-EU and 2.1 km as ICON-D2 respectively.

For this work a 500 m resolution grid is nested in a ICON-EU domain that provides initial and boundary conditions. On the nested grid ICON-LAM is run with most of the standard operational settings used for ICON-EU and ICON-D2. The operational assimilation of RADAR and latent heat nudging are switched off.

To produce a background error covariance matrix and be able to use the LETKF a sufficiently large ensemble size is required. For this the DWD operational ensemble size of 40 members is chosen. The single ensemble members receive their initial and lateral boundary conditions from the 40 members of the ICON-EU parent simulation.

2.3. The LETKF

Introduced by Bishop et al. 2001 as the Ensemble Transform Kalman Filter and in its localized form as the Local Ensemble Transform Kalman Filter by Hunt et al. 2007, the LETKF is a Kalman filter used in research and operational weather prediction. It combines the use of a probabilistic background error covariance and the efficiency of grid point by grid point assimilation. This is achieved by a transformation to ensemble space, making the analysis a linear combination of background ensemble members.

Following the formulation used in Diefenbach et al. 2023, the analysis ensemble mean is given by

$$\overline{x}_a = \overline{x}_b + \mathbf{X}_b \, \overline{w}_a \tag{2.1}$$

with the analysis of the i-th member being

$$x_{a,i} = \overline{x}_b + \mathbf{X}_b \left(\mathbf{W}_{a,i} + \overline{w}_a \right) \ . \tag{2.2}$$

The vector $\overline{w}_a$ which weighs the share of the ensemble perturbations to the mean increment is defined as

$$\overline{w}_a = \tilde{\mathbf{P}}_a \mathbf{Y}_b^T \mathbf{R}^{-1} \left(y_o - \overline{y}_b \right) \ . \tag{2.3}$$

Here $\mathbf{R}$ is the observation error covariance matrix representing the measurement error, the observation operator error and the representativeness error, $\mathbf{Y}_b$ is the perturbation matrix in observation space. $\overline{y}_b$ is analogous to $\overline{x}_b$ the model equivalent of the ensemble mean background and y_o are the observations. $\tilde{\mathbf{P}}_a$ is the analysis error covariance matrix in ensemble space. It is given as the inverse sum of the background error covariance and the observation error, both in ensemble space:

$$\tilde{\mathbf{P}}_a = \left[(k-1) \, \mathbf{I} + \mathbf{Y}_b^T \mathbf{R}^{-1} \mathbf{Y}_b \right]^{-1} \tag{2.4}$$

The matrix $\mathbf{W}_a$ gives the weights of the ensemble perturbations to form the analysis ensemble. It is given by the square root of the analysis error covariance:

$$\mathbf{W}_a = \left[(k-1) \, \tilde{\mathbf{P}}_a \right]^{\frac{1}{2}} \ . \tag{2.5}$$

The construction of the LETKF builds on the assumption of a linear observation operator H like

$$H \left(\overline{x}_b + \mathbf{X}_b \, w \right) \approx \overline{y}_b + \mathbf{Y}_b \, w \ . \tag{2.6}$$

2.3.1. Radial Velocity Observation Operator

The translation from the model space to the observation space, here in the from of calculating radial velocities from a model state, is done by an observation operator. In many DA algorithms the observation operator is used in the computation of the increments and thus the analysis. As this computation is done using linear algebra, the observation operator is mostly assumed to be linear and even invertible.

The LETKF is advantageous in that it only needs the model equivalents of the observation as seen in the computation of the weights $\overline{w}$ in eq. 2.3. No observation operator is used to compute the increments. Thus even if assumed in its derivation, the LETKF does not need an invertible, not even a linear observation operator.

The operators used in this work do not make full use of the possible generalisation, as they are linear, even if not invertible. To compute the radial velocity v_r, which Doppler-wind-LIDARs measure along their beam we only need to compute the wind component which is parallel to the beam $\vec{r}$:

$$v_r = \frac{\vec{v} \cdot \vec{r}}{\|\vec{r}\|} = \cos\eta \left(u \sin\alpha + v \cos\alpha\right) + w \sin\eta \tag{2.7}$$

Here the convention is used, that the radial velocity is positive if the wind vector points away from the point of reference and negative if it points towards it, so it is really $v_r = \|\mathrm{d}\vec{r}\|/\mathrm{d}t$. η is the elevation angle of the beam and α the azimuth measured clockwise from geographical north. As can be seen easily, the computation of the radial velocity v_r is linear with respect to the state vector (u, v, w).

2.3.2. BACY and KENDA

The environment for DA at the DWD is called *Basic Cycling* (BACY). It integrates the full model cycle needed for DA including the collection of boundary conditions, the running of the atmospheric and other connected models, the computation of model equivalents and the DA itself. KENDA, the *Kilometre-Scale Ensemble Data Assimilation* (KENDA) (Schraff et al. 2016) is the implementation of the LETKF is implemented in BACY for high resolution DA at the DWD.

To avoid a disruptive impact of analysis increments on the hydrostatic balance, an incremental analysis update (IAU) is used. It spreads out the increment and adds it to the background incrementally over some time interval, here 10 min.

2.4. Experimental Design

2.4.1. Simulation Domain

The simulation domain is a triangular ICON grid with a resolution of 500 m. This means the domain is resolving convection and even the effective resolution at five to ten times the actual resolution still is permitting convection.

The elevation data comes from *ASTER's Global Digital Elevation Model* (U.S./Japan ASTER Science Team 2009). At the implemented resolution the orography of the Inn Valley and its biggest tributaries are well resolved and even smaller side valleys are present. Because of technical limitations the the domain is rather big, spanning from western Switzerland to eastern Austria and from Munich to the Adria, containing most of the Alps. A map of this domain and the parent ICON-EU domain is shown in Fig. 5.

As vertical coordinates *Smooth Level Vertical Coordinates Vertical* (SLEVE) coordinates are used (Leuenberger et al. 2010). Those are terrain following coordinates where features

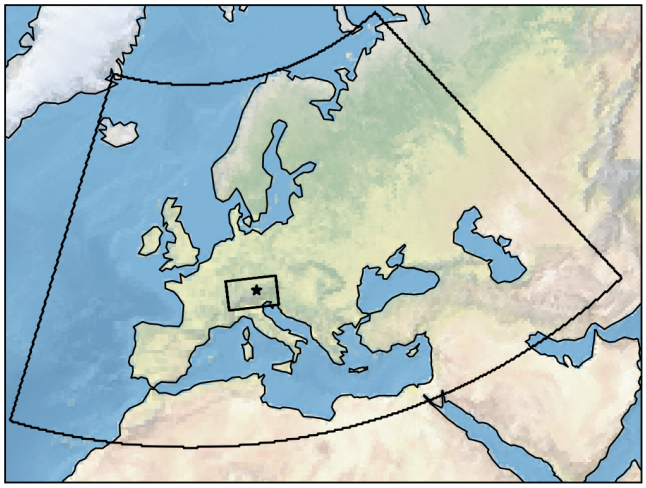


Figure 5.: Map of the outer ICON-EU domain and the nested simulation domain. $\star$ denotes the location of Kolsass.

of different scales decay separately. This makes small scale features decay faster with height. As with the ICON-EU setup the simulation domain has 60 vertical levels up to a height of 22.5 km.

2.4.2. Quality Control

The radial velocity data from the Doppler-LIDARs needs preprocessing in order to be used in the assimilation. To get the datasets from raw radial velocities several steps of quality control and spatial and temporal averaging into so called super observations are performed.

One PPI scan contains between $5 \cdot 10^3$ and $24 \cdot 10^3$ single observations, a large portion of which is faulty. With the velocity data comes the signal to noise ratio (SNR) which gives information on the quality of each measurement. Low SNR means that the observation is dominated by ambient noise while a high SNR can indicate reflection from orography. As a quality control step, only observations with a SNR between 1.0075 and 1.05 are considered. These values are obtained by identifying the removal of random noise and the back scatter from orography from the dataset.

In addition only data between the 3rd and the 55th LIDAR range gate are considered valid. The first two gates are too close to the instrument while after the 55th gate the values for Doppler velocity contain random noise between -40 and 40 ms^{-1}.

2.4.3. Super Observations

To further reduce the amount of observations and to remove fluctuations in time and space a super observation method is applied for each scan at a time. For this, each observation is assigned a (cell, level) pair. The values in the tuple correspond to grid

cells and levels of the numerical model domain. By this every observation from a scan is assigned its next neighbour in the model grid. The whole scan is also assigned a time from a set of five minute separated time steps which is closest to the mean time of the scan. All the values which are assigned to one (time, cell, level) triple are averaged and used as a super observation with the time and coordinates of the triple.

Because the quality of the LIDAR measurements varies due to environmental conditions during the day the number of grid points with super observations and the number of observations per super observation varies. To ensure comparable super observations only those are considered which contain 12 or more observations. If under this criteria a scan yields no super observation, the threshold for this scan is lowered by 2 until at least one super observation is received per scan.

In contrast to `radvel` the `vad` data is already quality controlled. It is only allocated a time, cell and level and averaged analogously to the `radvel` data.

The observation error for the LIDAR based wind speeds is only roughly estimated and copied from preexisting values from the DWD DA setup for wind measurements from tracked weather balloons. In the lower parts of the atmosphere this means $\sigma_o = 1.95$ ms^{-1}.

2.4.4. Localization

For the sake of a simple experimental design and easier interpretation of results, the horizontal adaptive localization of the DWD DA setup is turned off. Instead a constant horizontal localization length of $l_h = 10$ km is chosen. l_h denotes the half width at half maximum of the localization function.

The vertical localization in the DWD DA setup is changing linear with pressure between given values for the highest and the lowest level. For the same reasons as with the horizontal localization, the vertical localization length is set to be constant in the whole column at a value of $l_v = 0.3 \ \ln p$. This means a big increase in lower levels where l_v is otherwise set close to the lower boundary of $l_v = 0.075 \ \ln p$.

2.4.5. Experiments

Two different assimilation experiments and one background for comparison are used. The background (CONTROL) is a free forecast which starts at 0 UTC at the day of interest.

The first one which applies the assimilation at a single time followed by a 3 hour free forecast (SINGLE) is used to study the assimilation of different datasets at different times of the day. These experiments are initiated from the background state. Thus the effects of different data sources can be compared without any observations being inserted into the simulation. With these experiments the datasets `vad`, `radvel high`, `radvel low`, `all radvel`, `all lidar`, `all` or `oper` are each assimilated in the night at 2 UTC and in the afternoon at 16 UTC.

The second approach is a cycled experiment (CYCLE) starting from the background at 0 UTC and assimilating the given datasets every hour. By that the observations are fed continuously into the model. This is only done for the `vad`, `all lidar`, `all` and `oper`

Table 2.1.: Experiments and assimilated datasets

Experiment	Datasets
CONTROL	none
SINGLE	`vad`, `radvel high`, `radvel low`, `all radvel`, `all lidar`, `all`, `oper`
CYCLE	`vad`, `all lidar`, `all`, `oper`

datasets as they are the best performing in the SINGLE experiments. In that way it can be demonstrated how assimilation of combined datasets can improve valley wind systems in high resolution weather models.

All different experiments and the assimilated datasets are listed in table 2.1.

2.5. Validation Methods

To be able to assess the quality of DA one needs to be able to validate the results: The analysis and the forecast. As the LETKF takes model data in the observation space as inputs, the model equivalents are easily available for the fist guess and the analysis. Since data from several observation sources are available which might have different units, the validation dataset is split and analyzed independently by instrument and observed quantity.

As metric the root mean squared deviation (RMSD, here also root mean squared departure) is used to measure the difference between model equivalents and observations at a given assimilation time:

$$\mathrm{RMSD} = \sqrt{\left\langle (y_o - \overline{y}_b)^2 \right\rangle_{\mathrm{obs}}} = \sqrt{\left\langle d_{ob}^2 \right\rangle_{\mathrm{obs}}} \tag{2.8}$$

Here $d_{ob} = y_o - \overline{y}_b$ is the departure or increment of the ensemble mean. $\langle\cdot\rangle_{\mathrm{obs}}$ means averaging the different observations selected for verification at the given time step. At assimilation time the departure can also be computed using the analysis giving d_{ab}. This can be all observations which occurred during the last hour or only all observations at the time step. When dealing with horizontal velocities in the datasets `vad` and `surface` d is the average of the departures of the u- and v-component.

To evaluate analysis and forecast, the model equivalents are computed for and compared to the observations. In a cycling experiment this does not require any extra work as the model equivalents of the first guess are needed for the next assimilation step anyways.

To be able to evaluate at different time steps with different deviations of the background CONTROL experiment and compare an assimilation experiment with the CONTROL experiment the forecast RMSD's are divided by the one from the background experiment:

$$\mathrm{RelRMSD} = \frac{\mathrm{RMSD_e}}{\mathrm{RMSD_C}} - 1 \tag{2.9}$$

Here $\mathrm{RMSD_e}$ is the RMSD of the relevant experiment and $\mathrm{RMSD_C}$ the one of the baseline experiment.

If the validation should occur on a per observation basis dividing by the baseline experiment is not recommended as small values of $\mathrm{RMSD_C}$ may inflate the score. In this case it is better to look at the difference in absolute departures:

$$\Delta_d = |d_e| - |d_C| \tag{2.10}$$

2.5.1. Departure Statistics

Using the background or the analysis departure estimations of the observation and the background error covariance matrices can be calculated (Desroziers et al. 2005). The most interesting quantity in this case is the observation error σ_o which is a part of the observation covariance matrix $\mathbf{R}$, used by the LETKF (Eq. 2.4). In an optimal linear assimilation system, innovations d_{ob} and analysis increments d_{ob} should be compatible with $\mathbf{R}$. This leads to a consistency check for $\mathbf{R}$ by calculating the estimation value of the outer product of the analysis and the background departure:

$$\mathbf{R} = E\left[d_{oa}{d_{ob}}^T\right] \tag{2.11}$$

The main diagonal of $\mathbf{R}$ is filled with the squared observation error σ_o^2 for each observation.

3. Results

3.1. Assimilating one Hour

3.1.1. During the Night

The first suite of SINGLE experiments looks at assimilation at 2 UTC. The observations displayed in figures 2 and 3 show that at this time the down valley wind builds up in the nocturnal boundary layer. The wind speed is relatively low compared to day time and turbulence is weak.

The effects of assimilating `vad`, `radvel high` and `radvel low` at 2 UTC can be seen in Fig. 6. The maps show the increments in the u-component. The left column shows a east-northeastern cut along the valley. The right column shows the lowest model level.

The structure of all three increments look similar but the magnitudes show a big difference. The scale of the `radvel high` increment is about half the scale of the `vad` increment and the scale of the increment of `radvel low` is only about a quarter. This becomes visible when comparing the boundaries of the color-bars. From the profiles it is visible that the biggest increment happens above the instruments in a height of 1000 m above ground. There the increment is predominantly positive, meaning an increase of the already positive (down valley) u-component.

In the maps of the lowest model level positive and negative increments are visible. The positive ones stretch along the Inn Valley and the mountain ridges around Kolsass. The negative increments are found in the tributary valleys.

To evaluate these results quantitatively the metrics from section 2.5 are computed. A big difference can be seen when comparing the data sources used for assimilation and even for validation. The values for RelRMSD at assimilation time for all experiment and validation combinations are displayed in table 3.1. It appears quickly, that most values in the columns of LIDAR based validation datasets show a beneficial impact, while those in the `surface` column all show a detrimental impact on the analysis.

The assimilation of LIDAR data shows an improvement when validated against `radvel high` at 2 UTC and the metrics are shown in Fig. 7. The plots (a)–(c) show Δ_d in colored dots for the super observations of the different experiments spread in time and the azimuth angle relative to the instrument. lot (d) gives RelRMSD as a function of time. All three experiments have a beneficial to neutral impact. The assimilation of `vad` does not give the best analysis with RelRMSD barely below -10% but the forecast is able to pick up and gives the lowest value for RelRMSD for all three experiments. The assimilation of `radvel high` gives the highest improvement at assimilation time with RelRMSD $< -30\%$. This benefit is lost during the forecast time with RelRMSD quickly increasing. Assimilating `radvel low` deteriorates the analysis but has a small

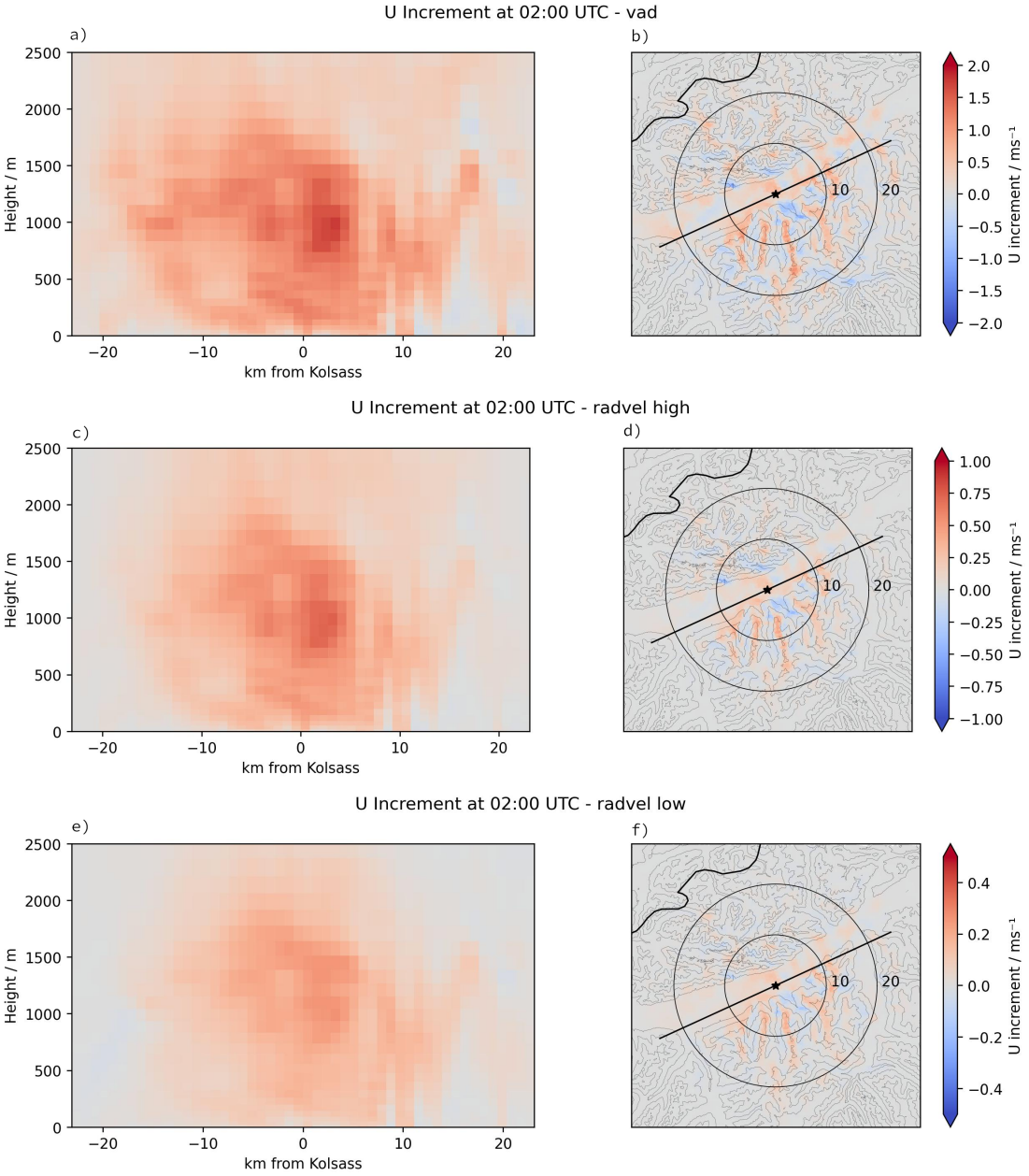


Figure 6.: Maps of increments of the assimilation of `vad`, `radvel high` and `radvel low` in the SINGLE experiments at 02:00 UTC. (a), (c), (e) show a east-northeastern cut along the line in the plots on the right. (b), (d), (f) show the lowest model level with circles indicating distance from Kolsass in km and 500 m isolines of valley topography in grey lines. Note the different scales of the color bars.

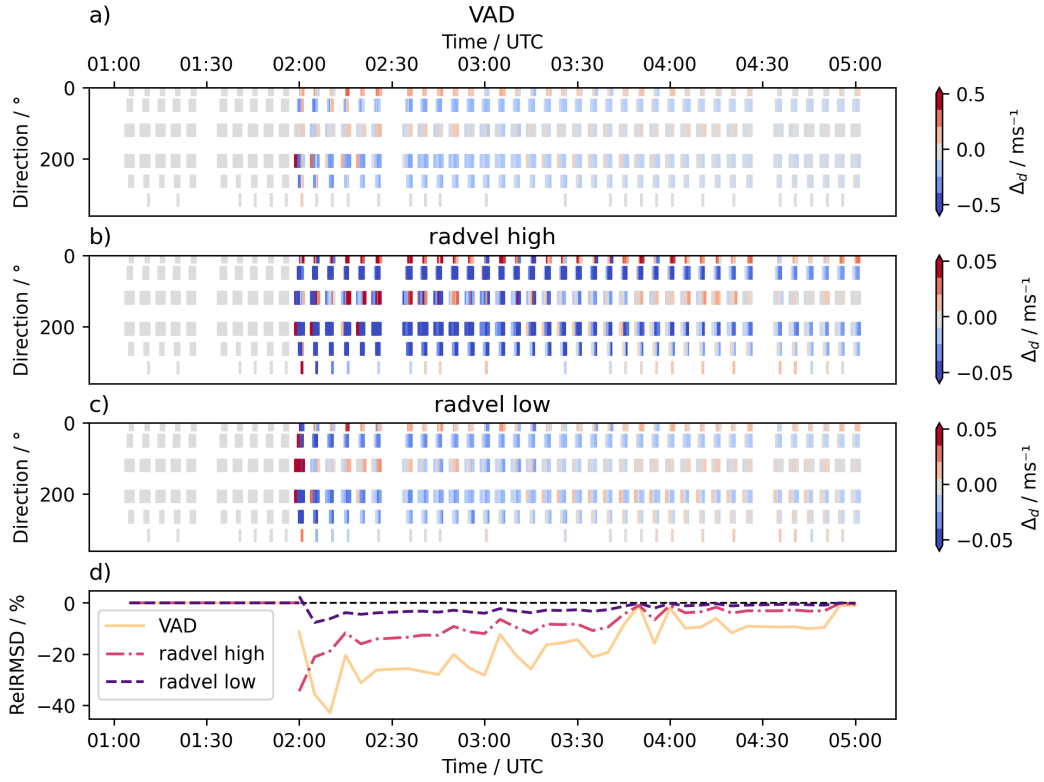


Figure 7.: Error statistics of SINGLE experiments assimilating `vad`, `radvel high` or `radvel low` scans at 2 UTC, validated against `radvel high`. Colored fields in (a)–(c) show Δ_d spread across time and azimuth angle, with increasing height from left to right in each field. Please note the different scales for different experiments. The curve in plot (d) gives RelRMSD in percent.

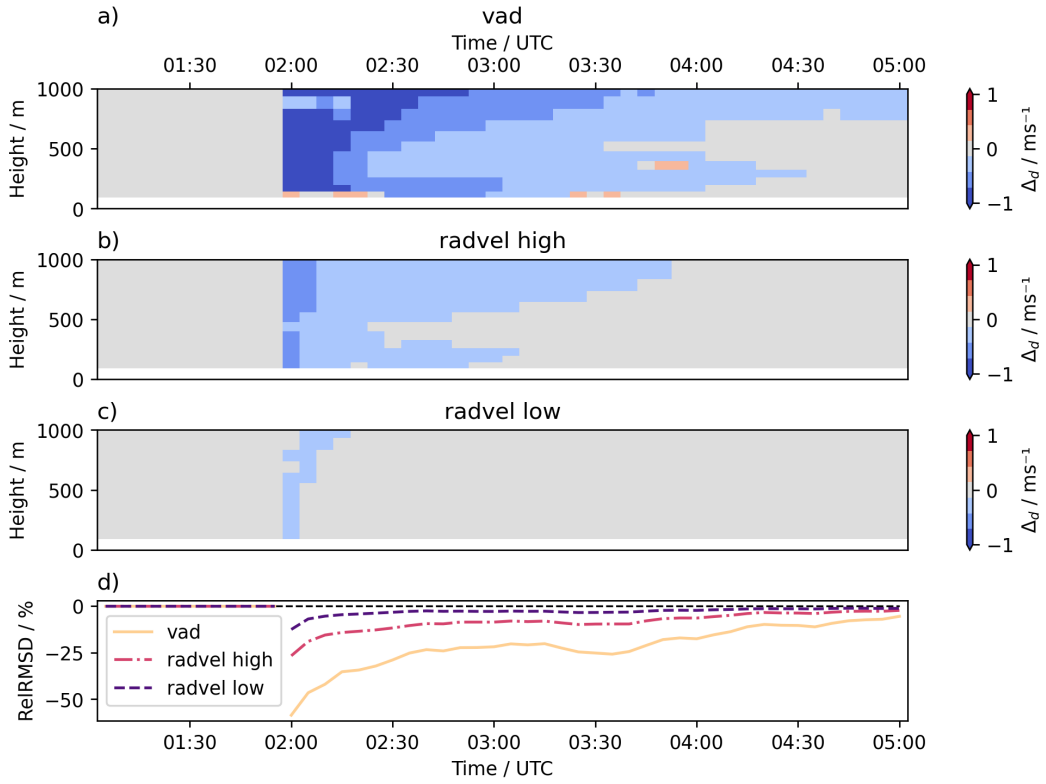


Figure 8.: Assimilation of single `vad`, `radvel high` and `radvel low` scans at 2 UTC, validated against `vad`. Colors in (a)–(c) show Δ_d for model equivalents of radial velocity measurements, the curve in plot (d) gives RelRMSD in percent.

Table 3.1.: RelRMSD in % at assimilation time, 02:00 UTC. Rows show the assimilated dataset (asm), columns show the dataset used for validation (val). The highest (worst) and lowest (best) value for each validation dataset are highlighted red and blue. The values for `radvel low` are missing because the dataset has no observations at this time to evaluate against.

asm \ val	`vad`	`radvel high`	`radvel low`	`surface`
`vad`	-58.58	-11.30	NA	67.17
`radvel high`	-26.68	-34.46	NA	23.10
`radvel low`	-12.54	2.53	NA	4.40
`all radvel`	-20.25	-27.95	NA	21.94
`all lidar`	-50.73	-44.52	NA	66.37
`all`	-53.17	-39.30	NA	55.96
`oper`	-1.32	-5.36	NA	6.64

but beneficial impact on forecast quality.

When verified against `vad` the assimilations of `vad`, `radvel high`, or `radvel low` yield an improvement for the analysis and forecast quality compared to CONTROL, as can be seen in Fig. 8. Here again the plots (a)–(c) show Δ_d for the three experiments as Hovmöller-diagrams in time and height above ground. Plot (d) gives RelRMSD as a function of time. The assimilation of `vad` scans reduces Δ_d for the analysis quite substantially for nearly the whole column, RelRMSD is down to -58.58%. The forecast is able to pick up from there and shows a value of RelRMSD $\leq -20\%$ for about two hours of forecast. Δ_d shows improvement in the forecast in two separate layers, one around 250 m and one above 750 m. The assimilation of `vad` also deteriorates the analysis in the lowest layer and in the forecast during the first half hour.

In contrast to these fairly good results the assimilation of `radvel high` and `radvel low` have a smaller impact. At analysis time Δ_d is reduced but not by a large amount throughout the column. For both experiments RelRMSD at assimilation time is at -26.68% for `radvel high` and -12.54% for radvel low, only being around half or even quarter the value of the first experiment. In the forecast RelRMSD grows quickly during the first 15 min after assimilation. It nearly stagnates afterwards, while still being slightly below zero for the three hour forecast.

The same experiments but validated against `surface` are shown in Fig. 9. The upper plot shows Δ_d as a function of time, the lower plots give the u- and v-component from the observations, the CONTROL experiment and the assimilation experiments. Here the beneficial to neutral impact is not given anymore. The assimilation of `vad` deteriorates the analysis with Δ_d above 0.5 ms^{-1}. The v-component of surface wind gets corrected, but the u-component is drawn away from the observation. In the forecast with lead times of one until two hours the quality improves with respect to CONTROL by better matching the u-component. This has to be taken with caution as the observed u is close to zero at this time and the `vad` experiment just gives a slightly smaller velocity compared to the baseline experiment. This can be generalised as the comparison of u- and v-components from the observations with the model equivalents show such big differences in their dynamics that the effect of assimilation is nearly negligible. `radvel high` and `radvel low` show little impact in the analysis and also during the three hour forecast Δ_d stays closely around zero.

The results of the SINGLE experiments for the `vad` data and the combined datasets `all lidar`, `all` and `oper` at 2 UTC are shown in Fig. 10 again as Hovmöller-plots of Δ_d for the different experiments in the plots (a)–(d) and with RelRMSD as a function of time in plot (e). The highest reduction in RelRMSD for the analysis is reached by `vad` which was already discussed above, followed by `all` and `all lidar`. In comparison to the assimilation of the `vad` data the deterioration of the analysis at the lowest scan level is mitigated by the datasets `all` and `all lidar`. The impact of the assimilation of `oper` data takes some time to reach the location at Kolsass but reduces RelRMSD by about ten percentage points after 30 min, which can be maintained during the three hour forecast.

The best forecast performance can be seen in the experiment assimilating the `all`

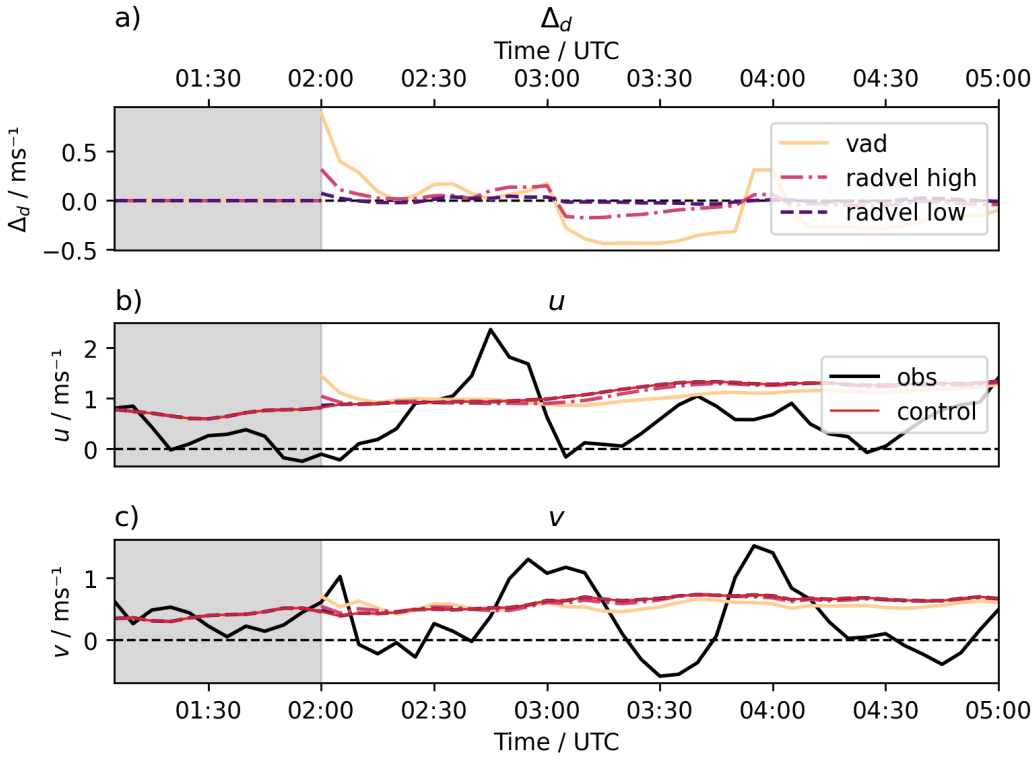


Figure 9.: Assimilation of single `vad`, `radvel high` and `radvel low` scans at 2 UTC, validated against `surface`. Plot (a) shows Δ_d, (b) and (c) show u and v with the observations in solid black and the model equivalents coloured lines.

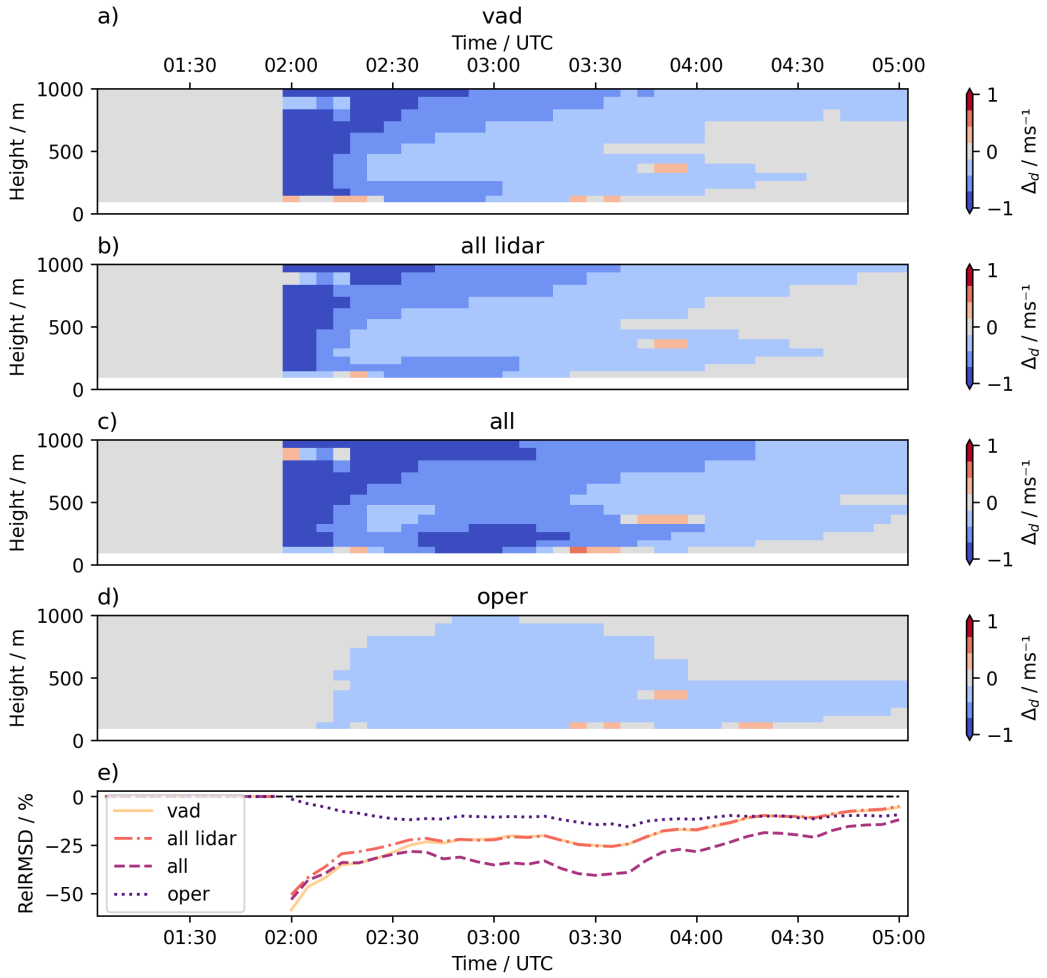


Figure 10.: Assimilation of single `vad`, `all lidar`, `all` and `oper` scans at 2 UTC, validated against `vad`. Colors in (a)–(d) show Δ_d profiles, the curve in plot (e) gives RelRMSD in percent.

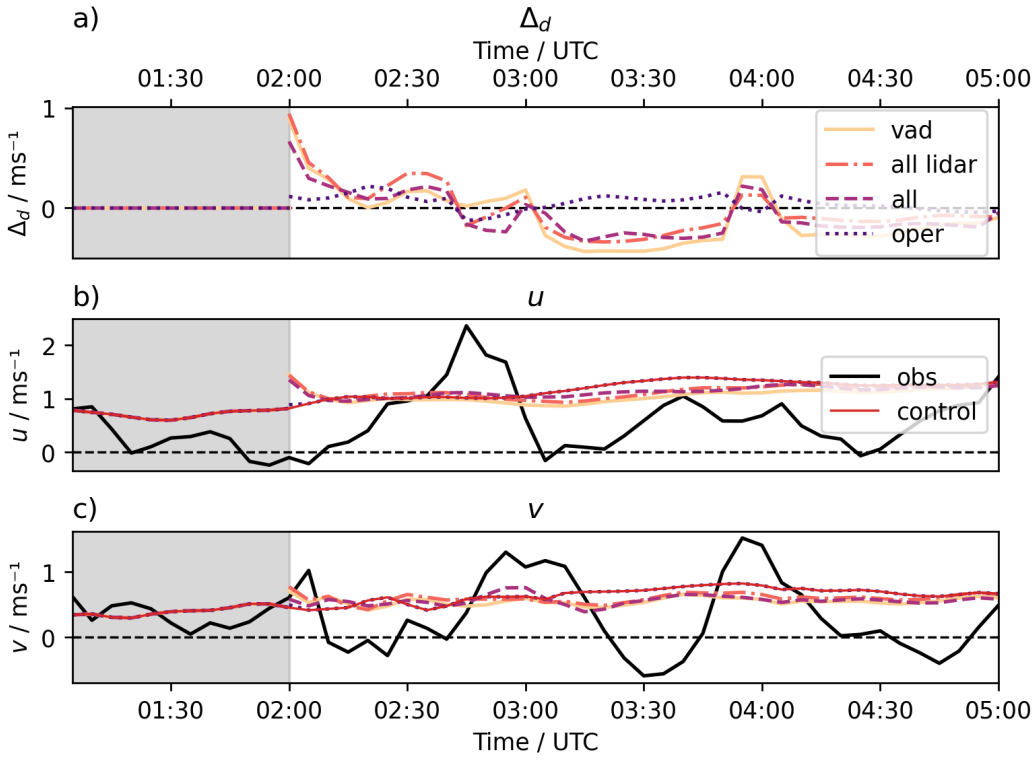


Figure 11.: Assimilation of single `vad`, `all lidar`, `all` and `oper` scans at 2 UTC, validated against `surface`. Plot (a) shows Δ_d, (b) and (c) show u and v with the observations in solid black and the model equivalents coloured lines.

dataset – a combination of `oper` and `all lidar` data. RelRMSD is below -45% % in the analysis and Δ_d shows a deviation reduction in the whole VAD column but one level. This improvement can also be passed on to the forecast which shows the lowest RelRMSD of all the experiments, staying below -20% during the three hour forecast.

The contrast to the verification against `surface` is as big as in the previously discussed experiments. Fig. 11 shows Δ_d and the u- and v-components of the observations and their model equivalents. The case for `vad`, `all lidar` and `all` is similar: Δ_d shows a deterioration in the analysis because the u-component is pushed away from the observation by about 0.5 ms^{-1} while the v-component is drawn towards the observation by only a fraction of that value. The forecast recovers from that in about 20 min and then Δ_d fluctuates around a value slightly below zero. The three experiments stay rather similar throughout the whole forecast time of three hours.

The assimilation of `oper` has very little effect on the surface analysis. As in the validation against `vad` the impact of the assimilation needs to be advected. When validated against `surface` this effect is a small positive value of Δ_d for 30 min and then Δ_d fluctuating around zero.

Table 3.2.: RelRMSD in % for a one hour forecast after assimilation time, 03:00 UTC. Rows show the assimilated dataset (asm), columns show the dataset used for validation (val). The highest (worst) and lowest (best) value for each validation dataset are highlighted red and blue.
⋆ Data for `radvel low` from 02:55 UTC.

asm \ val	`vad`	`radvel high`	`radvel low`⋆	`surface`
`vad`	-21.84	-28.24	-19.03	-6.64
`radvel high`	-8.58	-11.99	-4.67	-5.78
`radvel low`	-2.86	-4.05	-1.46	-2.06
`all radvel`	-3.69	-5.72	0.38	0.05
`all lidar`	-22.38	-28.34	-24.66	2.51
`all`	-35.32	-40.43	-37.90	-12.44
`oper`	-10.68	-13.19	-15.12	5.98

The combination of both Fig. 10 (a)–(c) and Fig. 11 shows that LIDAR based observations are not able to improve the surface layer analysis. The analysis deteriorates when validated against `surface`. Even when validated against `vad` the assimilation of `vad` does not improve the analysis at the lowest model layer.

Table 3.2 shows the values of RelRMSD for the SINGLE experiments one hour after assimilation for all four validation datasets. Here the assimilation of `all` gives the best value when validated against the LIDAR based datasets. The worst performing experiments after one hour are the one using `radvel low` and `all lidar`.

3.1.2. In the Afternoon

The picture looks more complicated for the SINGLE experiments at 16 UTC. Now the jet of the up valley flow is at its highest velocity as seen in Fig. 3. During the following hours it weakens only a little bit and separates from the valley floor, as a stable boundary layer forms beneath it. This process is not well covered by the CONTROL simulation where the stable layer grows at a much slower speed. The cause for this could be that the forming of the stable layer is not represented correctly in the model, maybe due to imprecision in the model orography or false representations of model parametrisations.

The values for RelRMSD and Δ_d validated against `vad` for this suite of SINGLE experiments is shown in Fig. 12 again as Hovmöller-plots of Δ_d for the different experiments in the plots (a)–(d) and with RelRMSD as a function of time in plot (e). Here the assimilation of `vad` yields the best analysis with both the lowest value of RelRMSD at -41.53%, but as during the night has a deteriorating effect at the lowest model level and in a height of 1000 m. The assimilation of `all` has a value of RelRMSD $= -39.09\%$ and a negative to neutral Δ_d throughout the column. The analysis of the `vad` and `all lidar` also give low values for RelRMSD but show a deterioration in the highest and for `vad` also in the lowest scan level. As in the case before `oper` has only little impact in the analysis, because the observations are not located close to Kolsass.

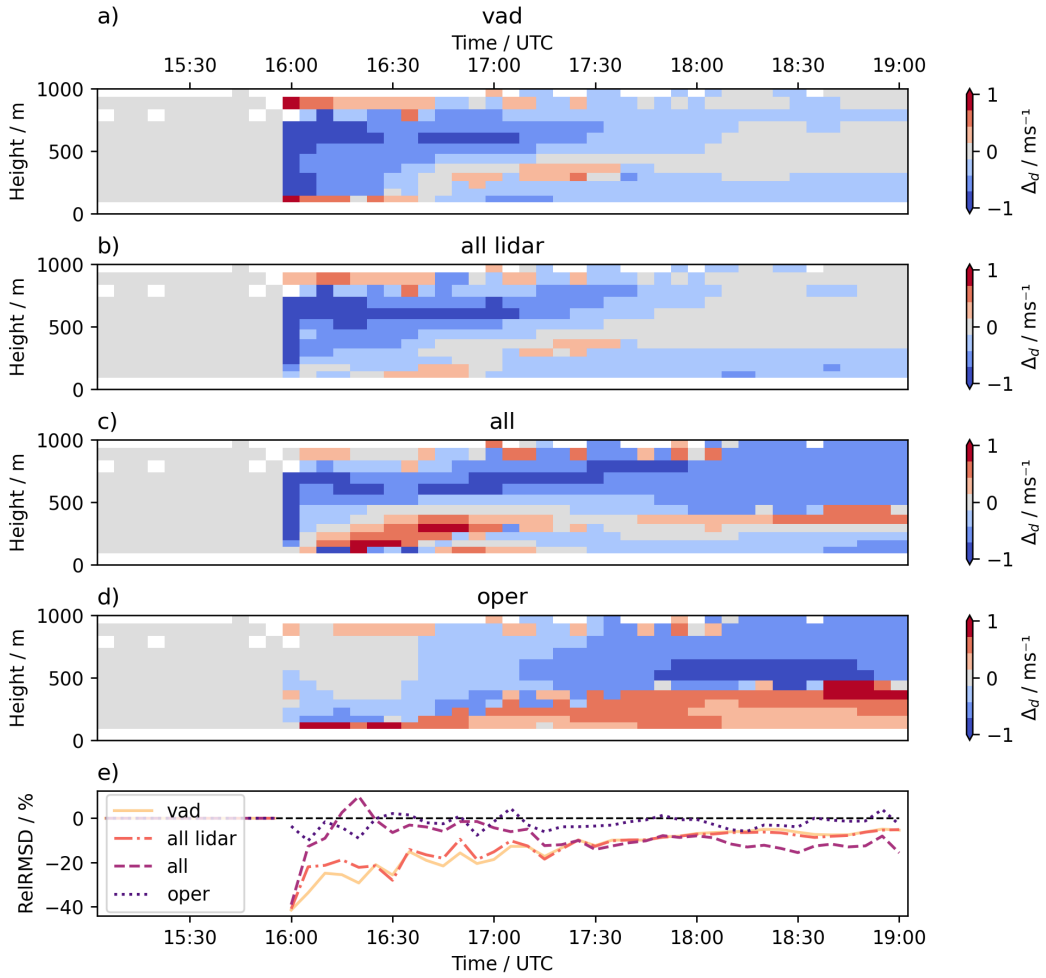


Figure 12.: Assimilation of single `vad`, `all lidar`, `all` and `oper` scans at 16 UTC, validated against `vad`. Colors in (a)–(d) show Δ_d profiles, the curve in plot (e) gives RelRMSD in percent.

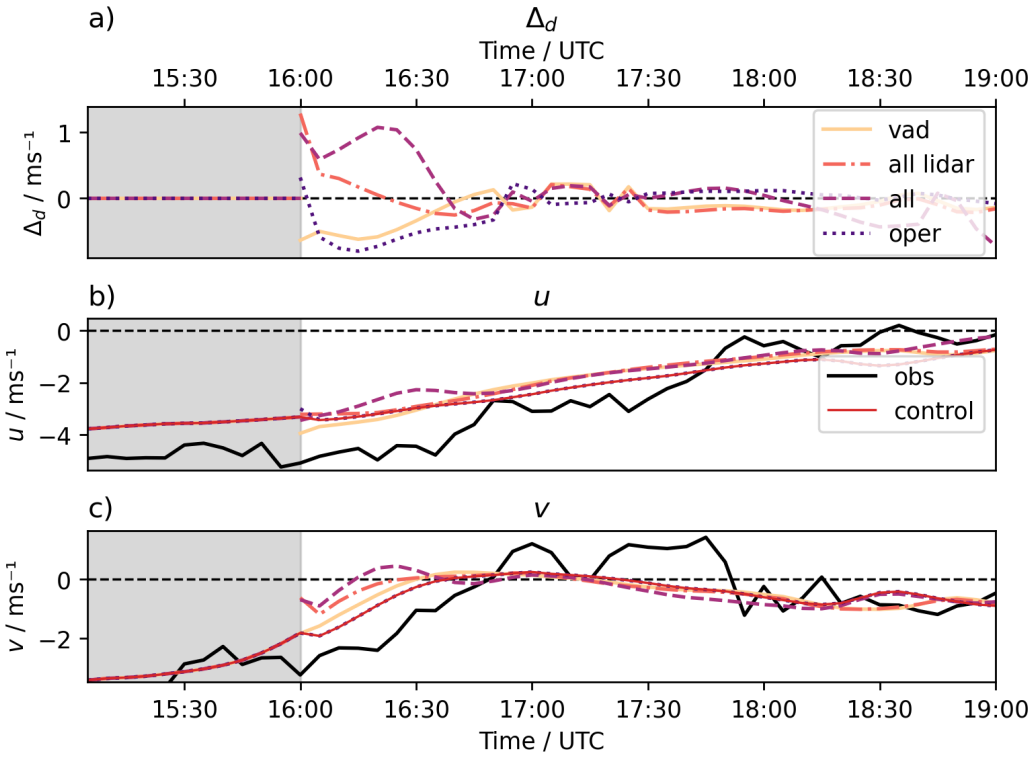


Figure 13.: Assimilation of single `vad`, `all lidar`, `all` and `oper` scans at 16 UTC, validated against `surface`. Plot (a) shows Δ_d, (b) and (c) show u and v with the observations in solid black and the model equivalents coloured lines.

The assimilation of `vad`, `all lidar` and `all` improves the forecast in a height about 600 m for two hours after assimilation. In the experiments assimilating `vad` and `all lidar` a layer of positive Δ_d, indicating a degradation of the forecast detaches from the low levels shortly after assimilation and rises to a height of 300 m above ground. Below that layer the forecast improves again. With the assimilation of `oper` the detachment of a forecast improvement is also visible, but the layer of positive Δ_d does itself not detach, but stays connected to the lowest level, forming an area of forecast deterioration which only grows with time.

This effect can be explained by the changes in forecasting the formation of stable boundary layer and thus a separation of the up valley flow from the ground mentioned above. The layer of positive Δ_d indicates an error in the different experiments when exactly and how fast the detachment happens. The zone of negative Δ_d shows that the buildup of the stable layer itself is represented better when local LIDAR data is assimilated. In contrast the assimilation of `oper` only improves the forecast above the stable boundary layer showing a worse representation of this process.

When looking at the validation against `surface` the picture changes again, as it can be

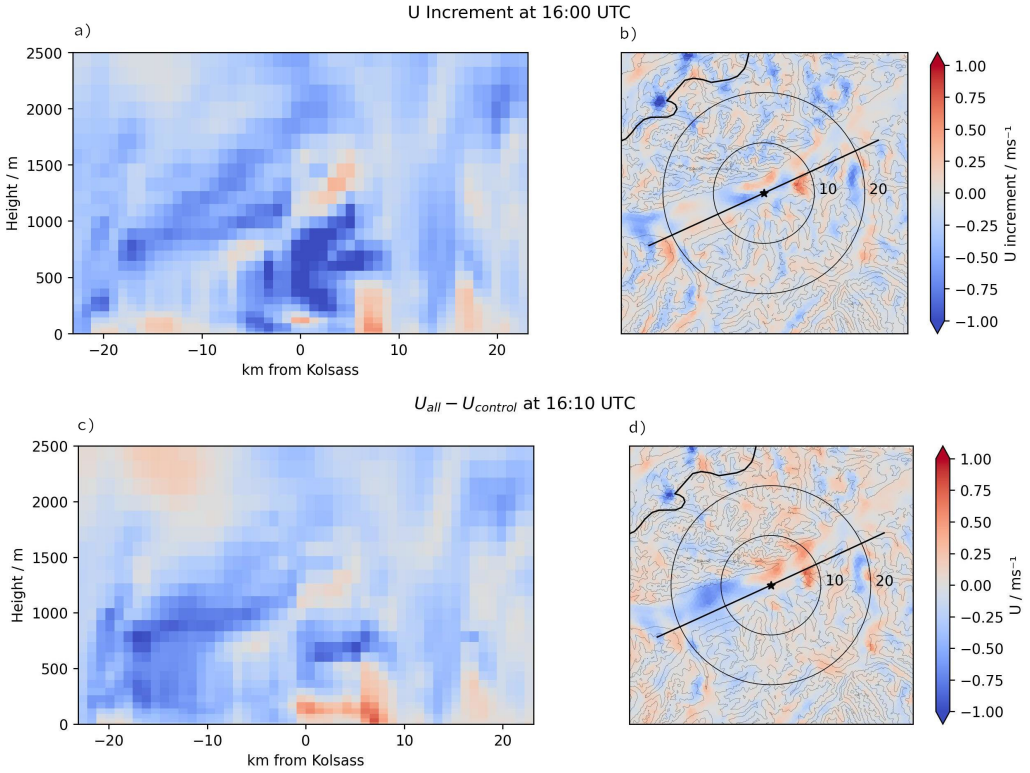


Figure 14.: The u Increment from the SINGLE `all` experiment and the difference to CONTROL at 10 minutes after assimilation time. (a) and (c) show a east-northeastern cut along the line in the plots on the right. (b) and (d) show the lowest model level with circles indicating distance from Kolsass in km.

seen in Fig. 13. The upper plot shows Δ_d as a function of time, the lower plots give the u- and v-component from the observations, the CONTROL experiment and the assimilation experiments. The analysis of the v-component of the lowest model level is worse or at least neutral for all experiments when compared to CONTROL. The assimilation of `all lidar` and `all` pulls it away by a large amount, while `vad` and `oper` have little impact. Only by successfully correcting the u-component the experiment assimilating `vad` can have a negative Δ_d at analysis time, thus reducing the absolute deviation. The other experiments have little to detrimental impact on u resulting in a positive value for Δ_d. In the forecast the experiments utilizing only LIDAR products converge after one hour and fluctuate around $\Delta_d = 0$ ms^{-1}. Only the assimilation of `oper` has a slightly beneficial impact. This can again be attributed more to fluctuations in the observations and particularly to a small offset in the right direction the forecast of the u-component that this experiment has compared to control.

Fig. 14 shows a map of the distribution of the u-increments and their development of the SINGLE assimilation of `all` dataset. The upper left side shows a vertical slice

through the Inn Valley along the path mapped in the right side. The biggest increments are found about from 250 m to 1 km above ground decreasing u by up to 1 ms^{-1}. The lower left side gives the evolution of this increment after 10 min by showing the difference of the u-component of CYCLE `all` and CONTROL. In this time most of the increment pattern directly above the site in Kolsass have stayed in place. Overall the influence of the assimilation has decreased. The right side of the Fig. shows a map of the lowest model level of the same quantities as on the left side. Isolines of topographic height are overlaid. Also here the up valley movement of negative u-increments can be observed.

Table 3.3 shows the values of RelRMSD for the SINGLE experiments in the afternoon at assimilation time 16 UTC. Here no dataset shows a good performance in multiple validations without being mediocre in another one. The experiment assimilating of `all radvel` is the best performing only when validated against `radvel high`. But it does not deteriorate the RMSD of the analysis in any validation dataset.

Table 3.3.: RelRMSD in % at assimilation time, 16:00 UTC. Rows show the assimilated dataset (asm), columns show the dataset used for validation (val). The highest (worst) and lowest (best) value for each validation dataset are highlighted red and blue.

asm \ val	`vad`	`radvel high`	`radvel low`	`surface`
`vad`	-41.53	0.87	NA	-20.30
`radvel high`	2.59	-38.27	NA	-29.40
`radvel low`	-0.08	-3.28	NA	35.97
`all radvel`	-7.62	-43.75	NA	-17.83
`all lidar`	-40.97	6.55	NA	41.41
`all`	-39.09	-1.04	NA	33.29
`oper`	-3.66	4.31	NA	11.12

Table 3.4 gives RelRMSD of the forecast one hour after assimilation for all four validation datasets. Again no single dataset performs best when validated against different validation datasets. The assimilation of `vad` and `all lidar` have a big to at least beneficial impact when validated against LIDAR data sources, but deteriorate the forecast in comparison with `surface`.

3.2. CYCLE Experiment

In the CYCLE experiment the whole day is simulated with assimilations every hour. Hovmöller-diagrams of wind speed for the experiments utilizing `all lidar` and `all` in comparison with CONTROL and VAD observations is shown in Fig. 15. During the night and the morning the cycled experiments better match the wind speed of the down valley flow. In the evening the assimilation pulls the structure of the stable boundary layer and the detached jet close to the observations but in the hour after assimilation the model pushes the flow back down. Because of that the assimilation experiments are not able to change the structure of the boundary layer. This is a sign that the assimilation

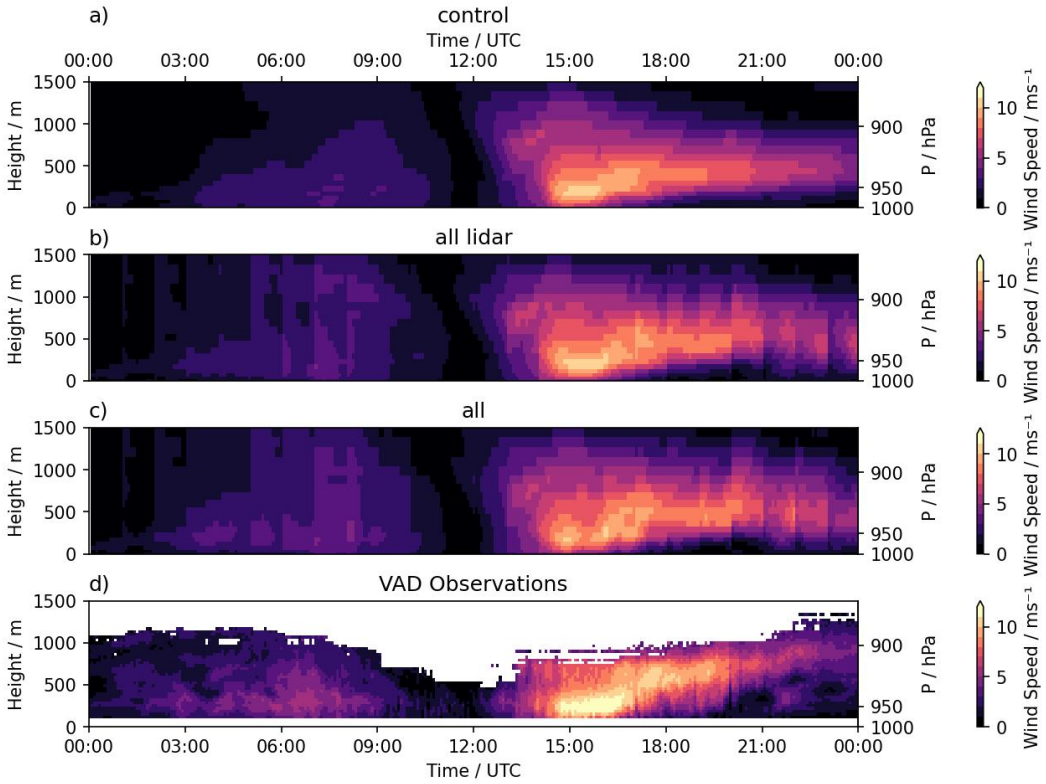


Figure 15.: (a)–(c): Vertical profiles of wind speed at Kolsass from the CYCLE experiments. For comparison VAD profiles from the same location are shown in (d).

Table 3.4.: RelRMSD in % for a one hour forecast after assimilation time, 17:00 UTC. Rows show the assimilated dataset (asm), columns show the dataset used for validation (val). The highest (worst) and lowest (best) value for each validation dataset are highlighted red and blue.
⋆ Data for `radvel low` from 16:55 UTC.

asm \ val	`vad`	`radvel high`	`radvel low`⋆	`surface`
`vad`	-18.72	-1.34	-16.06	27.37
`radvel high`	-1.13	0.15	1.28	-9.20
`radvel low`	1.98	1.17	6.93	-5.55
`all radvel`	-1.15	1.33	10.88	-8.88
`all lidar`	-15.27	-0.64	-6.83	17.52
`all`	-4.40	0.22	0.38	18.18
`oper`	-1.65	2.24	16.06	-3.07

has too little impact to change the atmospheric regime in the valley and is just able to correct it in the area around the observations.

The same assumption holds when looking at the values of Δ_d and RelRMSD validated against `vad` in Fig. 16 shown again as Hovmöller-plots of Δ_d for the different experiments in the plots (a)–(d) and with RelRMSD as a function of time in plot (e). Here especially for the experiments assimilating `vad`, `all lidar` and `all` the analysis is able to reduce deviation throughout the column. From the point of state estimation where we are not interested in a forecast this seems to be a successful result. The forecast of the following hour increases the deviation on certain levels again. This happens especially quickly during the last third of the simulated day. Overall the forecast is still an improvement compared to CONTROL during this time. This is again due to the imprecise location of the stable boundary layer growth in the model.

As in the SINGLE experiments validation against `surface`, displayed in Fig. 17, shows the challenge of improving the forecast for surface wind. The upper plot shows Δ_d as a function of time, the lower plots give the u- and v-component from the observations, the CONTROL experiment and the assimilation experiments. The curves for u and v from the assimilation experiments roughly follow the observations during the course of a day but they give no significant improvement compared to CONTROL. In the late afternoon, when the stable boundary layer forms and causes the up valley flow to separate from the ground the difference are visible because the assimilation experiments build up the stable layer below the jet a little bit quicker.

Table 3.5 gives the time averaged values of RelRMSD for the forecast of the whole 24 hours. They are computed from 01:05 UTC on and only include the forecast, ignoring the analysis. From this summary the dataset `all` appears to have the best impact on the forecast. Its value for RelRMSD is the lowest or close to the lowest for all verification cases. For verification against `surface` the assimilation of `all` still means a deterioration but it is smaller than for other assimilated datasets.

Regarding this mean forecast quality the assimilation of `oper` gives the worst results

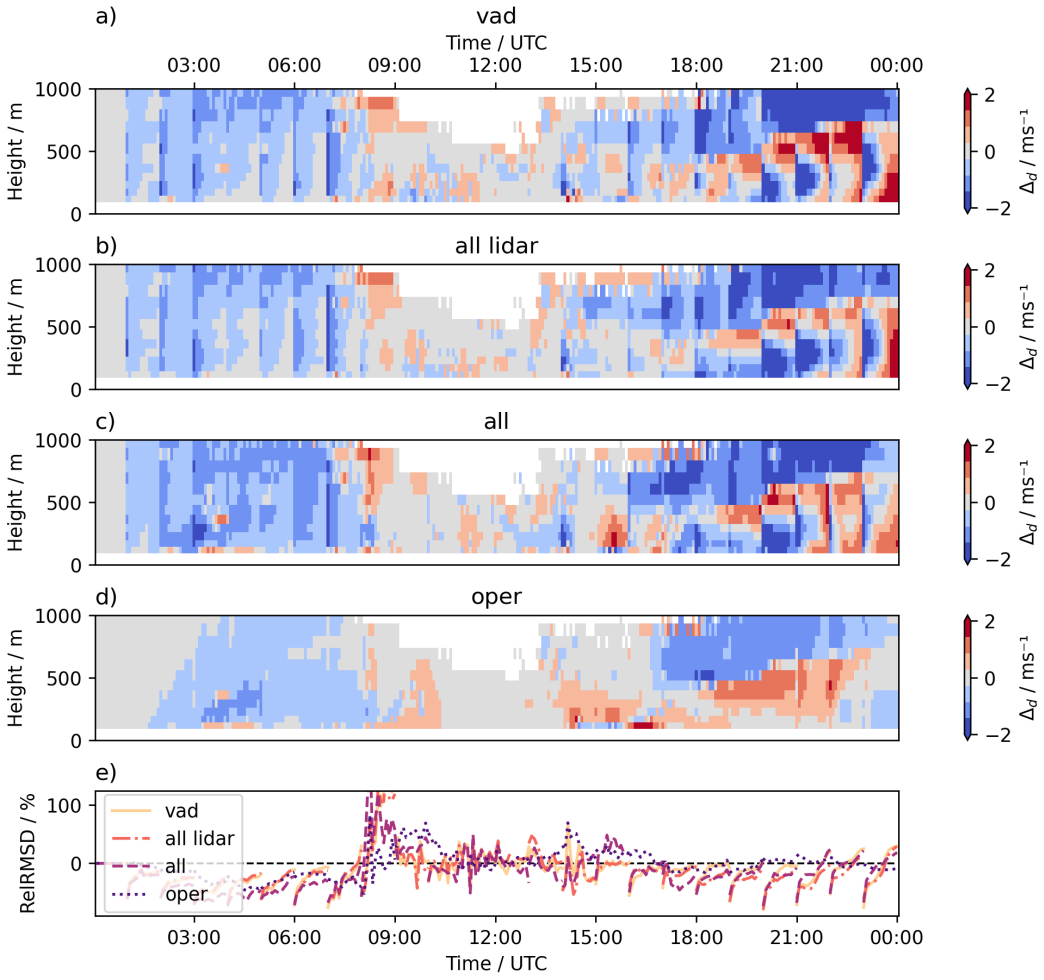


Figure 16.: Relative and absolute increment reduction for the CYCLE experiments. Plots (a)–(d) give Δ_d profiles in colors and (e) shows RelRMSD in percent.

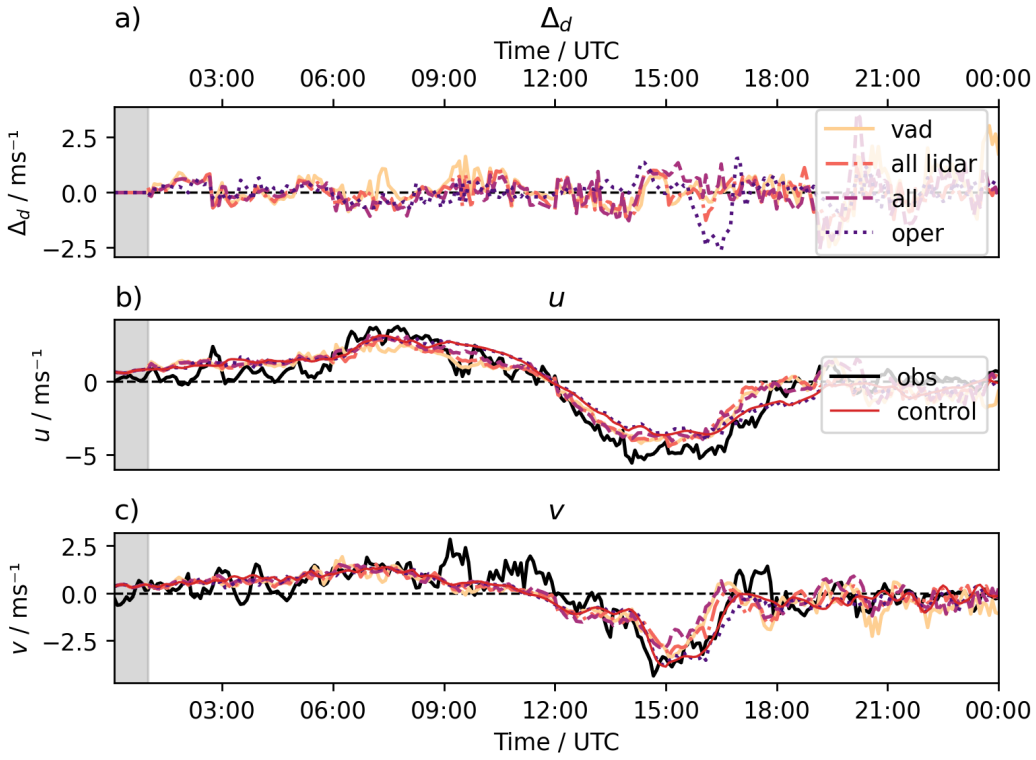
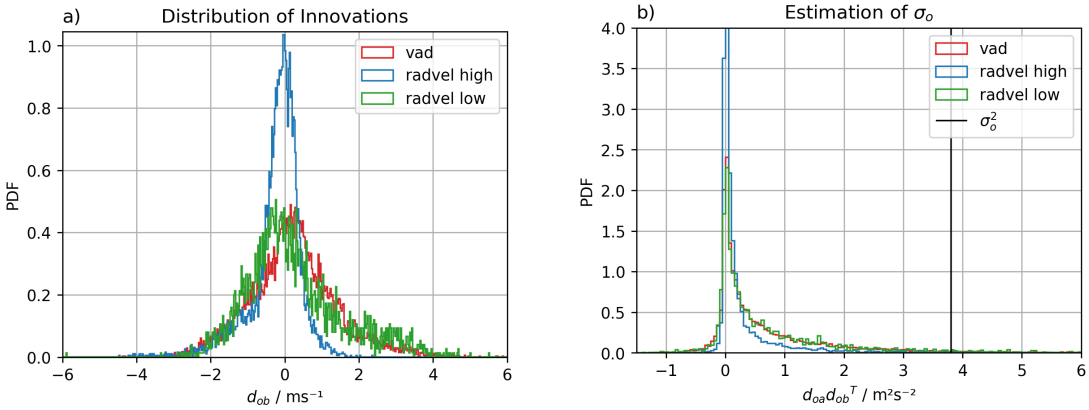


Figure 17.: CYCLE experiments validated against `surface`. Plot (a) shows Δ_d, (b) and (c) show u and v with the observations in solid black and the model equivalents coloured lines.

Table 3.5.: Average RelRMSD in % for CYCLE forecasts. Rows show the assimilated dataset (asm), columns show the dataset used for validation (val). The highest (worst) and lowest (best) value for each validation dataset are highlighted red and blue.

asm \ val	`vad`	`radvel high`	`radvel low`	`surface`
`vad`	-40.26	-4.32	-9.13	29.72
`all lidar`	-38.00	-6.59	-28.57	11.07
`all`	-38.19	-7.50	-26.36	9.06
`oper`	-3.20	-0.87	5.51	15.84



(a) Distribution of Innovations of several data sources. The data is collected from the CYCLE `all` experiment.

(b) Distribution of $d_{oa}{d_{ob}}^T$ from three different validations, see equation 2.11. The value used for σ_o^2 is indicated by the black bar.

Figure 18.: Statistics retireived from the CYCLE `all` run.

in every validation case except for `surface` where it ranges second worst. This can be explained by the distance the observations of `oper` have to Kolsass. They cannot give the necessary information to improve the valley circulation.

3.2.1. Observation Error Statistics

To check for biases in the model or observational data, a histogram of d_{ob} from the CYCLE `all` experiment is shown in Fig. 18a. The curves for `vad`, `radvel high` and `radvel low` are roughly centered around zero. Compared with a normal distribution curve the distributions of the deviations appear to be skewed. The weight of conclusions drawn from a sample of only one day is although questionable.

Fig. 18b shows the values for the diagonal of the check on the observation error covariance matrix $\mathbf{R}$ according to equation 2.11 for the same validation datasets as above. They are much lower than the value used for σ_o which is indicated by the black vertical line. This indicates that the chosen value of σ_o is to high, giving the observations little weight in comparison with the model background.

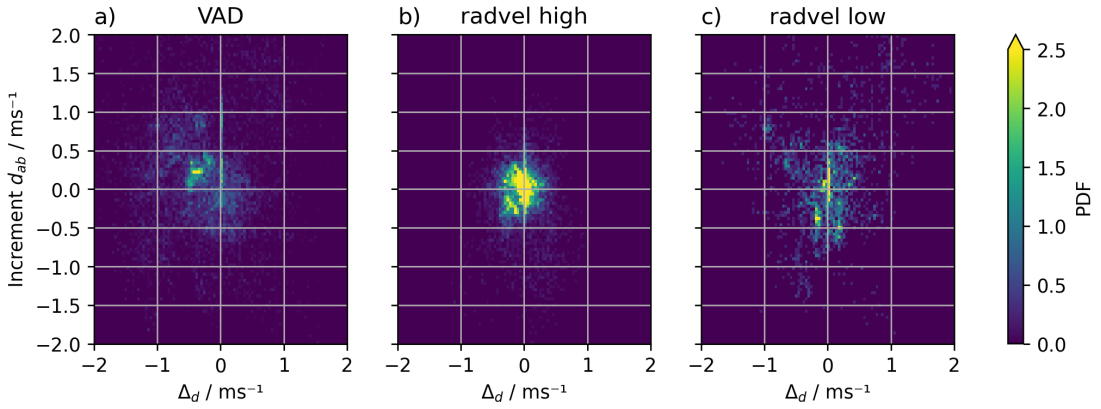


Figure 19.: 2D-Histogram of improvement Δ_d and increments d_{ab}. The data is collected from the CYCLE `all` experiment. Note that `radvel high` is not directly assimilated in this experiment. The number of observations for the data sources are `vad`: 13812, `radvel high`: 7726 and `radvel low`: 3010. Note that for `vad` u and v are counted separately whereas radial velocity has only one variable.

Fig. 19 shows a 2D-Histogram of forecast improvement Δ_d and increments d_{ab} from the same experiment. In each of the plots the left half shows increments which improve the analysis while points on the right half deteriorate it. For `vad` the biggest agglomeration is in the upper left quadrant, meaning it consists of positive increments which have a benefit on the analysis. The assimilation algorithm finds a bias to correct when looking at the `vad` data.

The distributions of the radial velocity data of `radvel high` and `radvel low` is much more centred with regards to Δ_d and d_{ab}. The values of `radvel high` are much more concentrated while `radvel low` spreads over a larger area with a few centers in or close to the lower left quadrant.

The caption of Fig. 19 also gives the number of observations assimilated in the CYCLE experiment. The `vad` dataset has about double the amount of the `radvel high` dataset and more than four times of the `radvel low` dataset. This can be an explanation for the higher impact the `vad` data gets compared to radial velocity data as it is visible in Fig. 8 and Tables 3.1 and 3.3. Because of the larger amount of data, the `vad` dataset gets a higher weight from the LETKF algorithm and thus has a greater impact.

4. Discussion

4.1. Conclusion

It is technically possible to assimilate high resolution LIDAR radial velocity data into the ICON-LAM high resolution numerical model using the LETKF. In this work both radial velocity and horizontal wind observations are assimilated and validated against themselves, independent observations at analysis time and against a forecast. The benefit and effectiveness of the assimilation depends on atmospheric conditions, the settings of the Kalman filter and the preparation and arrangement of the observational data.

The last point gives many choices to the applier. The path chosen in this work is to agglomerate and average groups of nearby observations into super observations. The positions in time and space of those super observations are selected so they overlap with the times of model output and the positions of model grid points. By that the number of observations is reduced. This means the weight given to the set of observations is also reduced if this is not compensated by adjusting the observation error accordingly. The assimilation of such prepared data shows its ability to improve the analysis, even if not as much as with other forms of data.

One drawback of the use of radial velocity especially with a high elevation angle η is its high dependency on the vertical velocity w. For $\eta = 70°$ as it is the case for `radvel high` w has a 73 % share of v_r according to equation 2.7. For `radvel low` the proportion still is 7 % to 11%. This presents two issues: Firstly w is not well represented in atmospheric simulations especially in comparison to high resolution observations. Thus the model equivalents of observed radial velocity can be influenced by this uncertain variable. Secondly as in many data assimilation environments w is not updated during assimilation in the setup used for this work. So differences in d_{ob} which arise from the vertical velocity in the background cannot be regulated by adjusting w in the analysis but may force the horizontal velocity to correct for this.

The radial velocity observations also can be processed via VAD into a column of horizontal wind components u and v on several height levels. This results in easy to interpret and also easy to assimilate data. This is visible as the assimilation of `vad` reduces the error in the analysis better or as good as `radvel high`, even when validated against `radvel high`. In the forecast it gives the best error reduction compared to the radial velocity experiments. Although the averaged impact of `vad` assimilation is quite good, the surface analysis gets deteriorated. A direct comparison with radial velocity data assimilation is not easy, as the dataset `vad` contains double or even four times the amount of observations than `radvel high` or `radvel high`.

The best results for the reduction of forecast error in comparison with CONTROL are

achieved when several data sources are combined. In combination with `radvel low` as `all lidar` the good results of the assimilation of `vad` can even be improved. Also the combination of this set with `oper` to `all` gives a higher error reduction in the forecast. Depending on the validation data set the assimilation of `all lidar` and `all` yield the best forecast results.

The assimilation of LIDAR based datasets mostly does not improve the surface analysis, as the validation against the `surface` dataset and even against the lowest level of `vad` shows. There are several reasons for this: LIDAR based observations have a technical lower height limit, they cannot retrieve information from the surface layer nor a thin boundary layer. Because of a imprecise background error covariance matrix and a large vertical localization length, this information is spread to the lower levels of the model. To avoid this, the localization length or the quality of the background error covariance matrix should be adjusted.

4.2. Shortcomings and Notes for Further Research

This sections describes a set of technical and conceptual problems that emerged during the work.

4.2.1. Technical Aspects

- Due to technical and resource-wise restrictions the dependency on the grid resolution could not be studied. With a grid length of 500 m and an effective resolution of $2.5 \cdot 10^3$ to $5 \cdot 10^3$ m the simulation sits in an intermediate spot between a high resolution LES and operational regional weather prediction. The model is not able to fully resolve the effects of orography. The analysis thus benefits from the observations because of the added accuracy. This benefit erodes over the forecast time. Reasons for this include the model does not matching the accuracy of the observations and the model being affected by systematic errors. It would be interesting to study how the analysis increments and the forecast quality changes with a higher model resolution.

- For reasons of the technical implementation the model equivalents to the radial velocity observations are not computed at the observation time as it is common practice in DA. They are computed at fixed five minute interval, to which the super observations are assigned. This introduces an timing error up to 3.5 min for single back-scatter velocity values[1]. The error introduced by this is of course negligible for showing the feasibility of the method. For simulations with higher spatial resolution and increased accuracy the computation of the model equivalents needs to be revisited. Also an increase in computational performance can be expected if

[1]The maximum timing error is estimated as the maximum distance the mean time of a scan can have to the next 5 min timestamp and the maximum time a single measurement can have to the mean time of the scan. With the scan speed of 3° s^{-1} for the low scanning LIDAR this gives 2.5 min + 1 min = 3.5 min

the model equivalents are computed 'live' and not from saved files after the model is run.

- Even if the super observations are agglomerated to fixed time steps every five minutes in the CYCLE experiment they are only assimilated hourly in batches. Because of this the analysis reduces the error also only hourly. In the following hour the error grows again. A higher frequency of assimilation can potentially keep the error low more continuously and thus increase the quality of state estimation and also the quality of the forecast.

4.2.2. Conceptual Aspects

- Quality control and the construction of radial velocity super observations are not done in an unambiguous way. To exclude reflections by fixed objects completely a blacklisting of affected regions by (azimuth, elevation, range) should be implemented. The assignment of observations to their nearest neighbour (cell, level) pair is inaccurate for values where the azimuth and elevation angle differs from the one of the (cell, level) pair. For small range distances this can mean observations are shifted by an azimuth angle of several 10°. A possible improvement here would be to rethink the computation of super-observations and possibly not confine them to (cell, level) locations of the grid. Instead of assigning the observations to fixed locations it might be advantageous to interpolate the model fields for the model equivalents as it is usually done in observation operators.

- As discussed earlier the values for the radial velocity observation error are copied from existing values of the DWD setup. Investigating the distribution and systematic error of radial velocity measurements and their averages could lead towards a more fitting and even conceptually meaningful error model. This might have a big effect on the efficiency of assimilating the raw lidar data in comparison with values from a VAD. A sensitivity study on the observation error could give valuable insights on this.

- The benefit of combination and comparison of LIDAR data with `oper` data is not without doubt in the presented case. The data was taken with the setup used for 6.5 km resolution simulations. Because of this, the observation errors for these observations might not be optimal for the use in much finer resolved simulations. Also for technical practiclity the localization lengths are set to the same values as for the LIDAR observations, also with unknown impact on the quality of `oper` assimilation.

4.3. Outlook

The result of this work stand as a proof of feasibility for the direct assimilation of Doppler-LIDAR radial velocity into a high resolution NWP model in alpine terrain.

This technique could potentially be used in future observation campaigns in complex terrain. Using a high resolution model or even an LES where all the recorded data of such a campaign is assimilated could not only combine this data in a statistically optimal way but also enrich it via a physical model which spreads the measured information in the model domain and to variables which are not directly observed. With this added value the then easy to use data can be used to perform further analysis. In the case of CROSSINN the development and timing of cross valley wind, mountain breezes and mountain valley wind systems could be studied with increased resolution and detail. ∾

Bibliography

Adler, Bianca et al. (2021). 'CROSSINN: a field experiment to study the three-dimensional flow structure in the Inn valley, Austria'. In: *Bulletin of the American Meteorological Society* 102.1, E38–E60. URL: https://doi.org/10.1175/BAMS-D-19-0283.1.

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. URL: https://journals.ametsoc.org/view/journals/mwre/129/3/1520-0493_2001_129_0420_aswtet_2.0.co_2.xml.

Browning, K. A. and R. Wexler (1968). 'The Determination of Kinematic Properties of a Wind Field Using Doppler Radar'. In: *Journal of Applied Meteorology and Climatology* 7.1, pp. 105–113. DOI: 10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/apme/7/1/1520-0450_1968_007_0105_tdokpo_2_0_co_2.xml.

Desroziers, G., L. Berre, B. Chapnik and P. Poli (2005). 'Diagnosis of observation, background and analysis-error statistics in observation space'. In: *Quarterly Journal of the Royal Meteorological Society* 131.613, pp. 3385–3396. DOI: 10.1256/qj.05.108. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1256/qj.05.108. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.108.

Diefenbach, Theresa, George Craig, Christian Keil, Leonhard Scheck and Martin Weissmann (2023). 'Partial analysis increments as diagnostic for LETKF data assimilation systems'. In: *Quarterly Journal of the Royal Meteorological Society* 149.752, pp. 740–756. DOI: 10.1002/qj.4419. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4419. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4419.

Finn, Tobias Sebastian, Gernot Geppert and Felix Ament (2020). 'Towards assimilation of wind profile observations in the atmospheric boundary layer with a sub-kilometre-scale ensemble data assimilation system'. In: *Tellus A: Dynamic Meteorology and Oceanography* 72.1, pp. 1–14. DOI: 10.1080/16000870.2020.1764307. eprint: https://doi.org/10.1080/16000870.2020.1764307. URL: https://doi.org/10.1080/16000870.2020.1764307.

Goger, Brigitta, Mathias W. Rotach, Alexander Gohm, Ivana Stiperski and Oliver Fuhrer (2016). 'Current challenges for numerical weather prediction in complex terrain: Topography representation and parameterizations'. In: pp. 890–894. DOI: 10.1109/HPCSim.2016.7568428.

Gohm, Alexander, Maren Haid and Mathias W. Rotach (Mar. 2021). *CROSSINN (Cross-valley flow in the Inn Valley investigated by dual-Doppler lidar measurements) - ACINN Doppler wind lidar data sets (SL88, SLXR142)*. Version v1. DOI: 10.5281/zenodo.4585577. URL: https://doi.org/10.5281/zenodo.4585577.

Gustafson Jr, William I et al. (2020). 'The large-eddy simulation (LES) atmospheric radiation measurement (ARM) symbiotic simulation and observation (LASSO) activity for continental shallow convection'. In: *Bulletin of the American Meteorological Society* 101.4, E462–E479. URL: https://doi.org/10.1175/BAMS-D-19-0065.1.

Hersbach, H. et al. (2023). *ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Accessed on 05-01-2024)*. DOI: 10.24381/cds.adbb2d47.

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. URL: https://doi.org/10.1016/j.physd.2006.11.008.

Leuenberger, Daniel, Marcel Koller, Oliver Fuhrer and Christoph Schär (2010). 'A Generalization of the SLEVE Vertical Coordinate'. In: *Monthly Weather Review* 138.9, pp. 3683–3689. DOI: 10.1175/2010MWR3307.1. URL: https://journals.ametsoc.org/view/journals/mwre/138/9/2010mwr3307.1.xml.

Markowski, Paul and Yvette Richardson (2010). *Mesoscale meteorology in midlatitudes*. eng. Chichester: Wiley-Blackwell. ISBN: 9780470682098. DOI: 10.1002/9780470682104.

Miyoshi, Takemasa et al. (2016). '"Big Data Assimilation" Revolutionizing Severe Weather Prediction'. In: *Bulletin of the American Meteorological Society* 97.8, pp. 1347–1354. DOI: 10.1175/BAMS-D-15-00144.1. URL: https://doi.org/10.1175/BAMS-D-15-00144.1;%20https://journals.ametsoc.org/view/journals/bams/97/8/bams-d-15-00144.1.xml.

Neggers, R. A. J., A. P. Siebesma and T. Heus (2012). 'Continuous Single-Column Model Evaluation at a Permanent Meteorological Supersite'. In: *Bulletin of the American Meteorological Society* 93.9, pp. 1389–1400. DOI: 10.1175/BAMS-D-11-00162.1. URL: https://journals.ametsoc.org/view/journals/bams/93/9/bams-d-11-00162.1.xml.

Pirk, Norbert, Kristoffer Aalstad, Sebastian Westermann, Astrid Vatne, Alouette van Hove, Lena Merete Tallaksen, Massimo Cassiani and Gabriel Katul (2022). 'Inferring surface energy fluxes using drone data assimilation in large eddy simulations'. In: *Atmospheric Measurement Techniques* 15.24, pp. 7293–7314. URL: https://doi.org/10.5194/amt-15-7293-2022.

Rotach, Mathias W., Ivana Stiperski, Oliver Fuhrer, Brigitta Goger, Alexander Gohm, Friedrich Obleitner, Gabriele Rau, Eleni Sfyri and Johannes Vergeiner (2017). 'Investigating Exchange Processes over Complex Topography: The Innsbruck Box (i-Box)'. In: *Bulletin of the American Meteorological Society* 98.4, pp. 787–805. DOI: 10.1175/BAMS-D-15-00246.1. URL: https://journals.ametsoc.org/view/journals/bams/98/4/bams-d-15-00246.1.xml.

Sawada, Masahiro, Tsuyoshi Sakai, Toshiki Iwasaki, Hiromu Seko, Kazuo Saito and Takemasa Miyoshi (2015). 'Assimilating high-resolution winds from a Doppler lidar using an ensemble Kalman filter with lateral boundary adjustment'. In: *Tellus A: Dynamic Meteorology and Oceanography* 67.1, p. 23473. DOI: 10.3402/tellusa.

v67.23473. eprint: https://doi.org/10.3402/tellusa.v67.23473. URL: https://doi.org/10.3402/tellusa.v67.23473.

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)'. In: *Quarterly Journal of the Royal Meteorological Society* 142.696, pp. 1453–1472. DOI: 10.1002/qj.2748. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2748. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2748.

U.S./Japan ASTER Science Team (2009). *ASTER Global Digital Elevation Model*. DOI: 10.5067/ASTER/ASTGTM.002.

Wang, Haoliang, Yubao Liu, Jing Duan, Yueqin Shi, Xiaofeng Lou and Jiming Li (2022). 'Assimilation of Radar Reflectivity Using a Time-Lagged Ensemble Based Ensemble Kalman Filter With the "Cloud-Dependent" Background Error Covariances'. In: *Journal of Geophysical Research: Atmospheres* 127.10. e2021JD036207 2021JD036207, e2021JD036207. DOI: 10.1029/2021JD036207. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JD036207. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JD036207.

Yang, Shu-Chih, Fang-Yi Cheng, Lian-Jie Wang, Sheng-Hsiang Wang and Chia-Hua Hsu (2022). 'Impact of lidar data assimilation on planetary boundary layer wind and PM2.5 prediction in Taiwan'. In: *Atmospheric Environment* 277, p. 119064. ISSN: 1352-2310. DOI: 10.1016/j.atmosenv.2022.119064. URL: https://www.sciencedirect.com/science/article/pii/S1352231022001297.

Zängl, Günther, Daniel Reinert, Pilar Rípodas and Michael Baldauf (2015). 'The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core'. In: *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 563–579. DOI: 10.1002/qj.2378. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2378. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2378.

A. Appendix

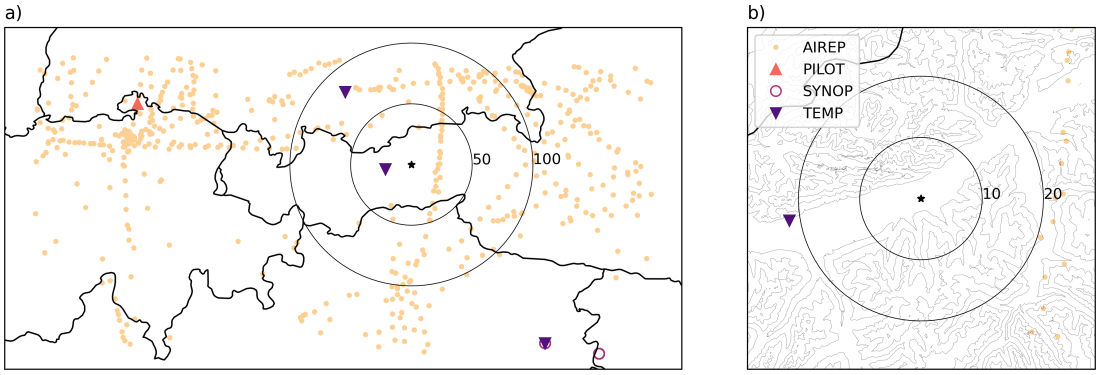


Figure 20.: Same as figure 4 but only showing actively assimilated wind observations.

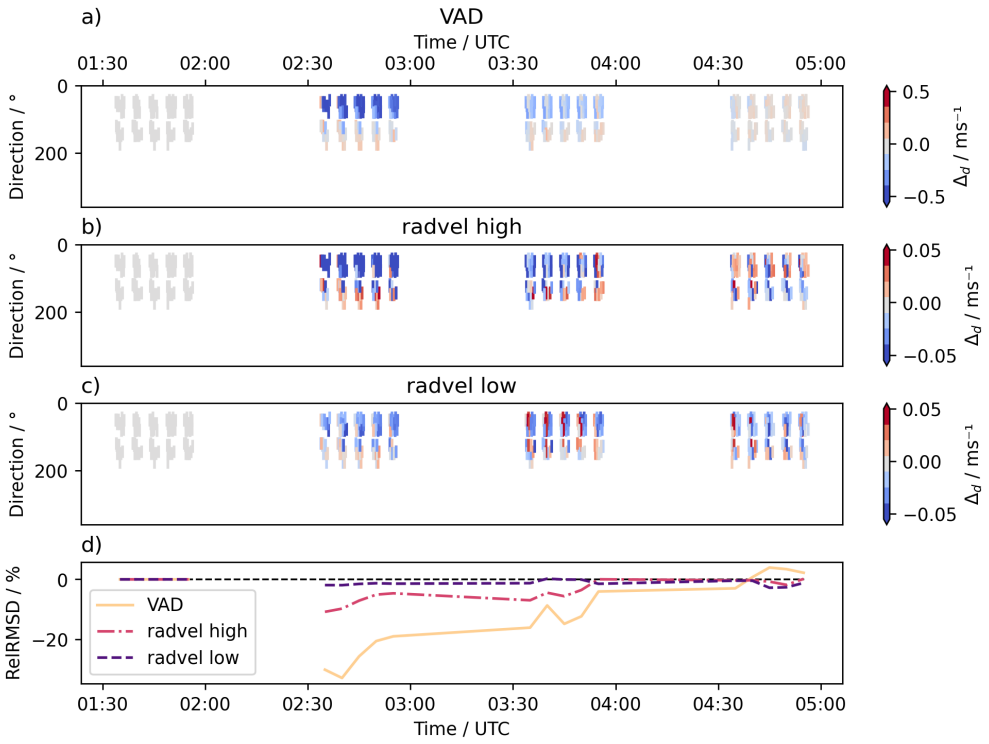


Figure 21.: Same experiments as Figure 7 but validated against `radvel low`. Colored fields show Δ_d spread across time and azimuth angle, with increasing height from left to right in each field. Please note the different scales for different experiments. The curve in the lowest plot gives RelRMSD in percent.