



universität  
wien

# DISSERTATION | DOCTORAL THESIS

Titel | Title

Assimilation of cloud-affected satellite observations for  
convective-scale numerical weather prediction

verfasst von | submitted by  
Lukas Kugler BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of  
Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree  
programme code as it appears on the  
student record sheet:

UA 796 605 415

Dissertationsgebiet lt. Studienblatt | Field of  
study as it appears on the student record  
sheet:

Meteorologie

Betreut von | Supervisor:

Univ.-Prof. Mag. Dr. Martin Weißmann

# Abstract

Severe convective storms cause enormous loss of life and damage to property. Despite tremendous improvements in numerical weather prediction (NWP) over the last decades, predicting hazards at the convective scale (4-40 km) remains particularly challenging. Convective-scale forecasts are sensitive to small-scale errors which conventional observations fail to resolve. Visible and infrared satellite channels observe the convective scale at sufficient resolution in time and space. However, their assimilation in operational NWP systems is still very limited as their assimilation involves nonlinear observation operators, non-Gaussian error distributions and potentially ambiguous information. This thesis tries to estimate the potential of assimilating cloud-affected infrared and visible satellite observations for the prediction of severe storms and to improve our understanding of observation operator nonlinearity.

This thesis focuses on the potential impact of satellite observations in an idealized framework, where the truth is perfectly known and systematic model and operator deficiencies are excluded. Idealized observing-system simulation experiments (OSSE) are employed, in which a model simulation is used as virtual truth. This virtual truth is simulated by integrating two selected initial conditions with the Weather Research and Forecasting (WRF) model. The first scenario represents isolated convection triggered by a warm air bubble, whose location is uncertain in the forecast ensemble. The second scenario features deep convection scattered throughout the domain with a forecast ensemble that is uncertain about the location and the life-cycle stage of convection. Observations are generated synthetically from both scenarios and assimilated into a 40-member forecast ensemble using the Ensemble Adjustment Kalman Filter.

The thesis is based on three publications. The first publication estimates the potential forecast impact of assimilating either visible ( $0.6 \mu\text{m}$ ) or infrared ( $6.2$  and  $7.3 \mu\text{m}$ ) satellite observations in the two aforementioned scenarios relative to the impact of assimilating radar reflectivity observations. Results showed that the assimilation of satellite observations was nearly as beneficial as the assimilation of three-dimensional radar reflectivity observations when the prior forecast's uncertainty was mostly in the horizontal position. When the location and the stage of convection were unknown, the impact of satellite observations was similar to that of 2D radar observations.

The second publication extends previous research on assimilating individual satellite channels to the combined assimilation of multiple channels and evaluates the vertical distribution of the resulting analysis error for different cloud conditions. It is revealed that the combined assimilation can mitigate the ambiguity of individual visible and infrared channels and improve the analysis of clouds.

In the third manuscript, the effect of observation operator nonlinearity of visible and infrared radiances, which violates a basic assumption of current data assimilation schemes,

is studied. This is done by estimating the deviation of the analysis from the expected analysis that would result from a linear observation operator. Furthermore, it is investigated if observations with small deviations from the model first-guess should be excluded from the assimilation. While these observations sometimes exhibit a detrimental impact, the results reveal that it is still beneficial to include them in the assimilation as they tell the ensemble that the first-guess is already very accurate.

This work contributes to a better understanding of the potential benefits of assimilating cloud-affected satellite observations for early and accurate convective-scale severe weather warnings.

# Zusammenfassung

Unwetter verursachen erhebliche wirtschaftliche Schäden und menschliches Leid. Trotz beeindruckender Fortschritte in der numerischen Wettervorhersage (NWP) in den letzten Jahrzehnten ist die Vorhersage von Gefahren auf konvektiver Skala (4-40 km) besonders herausfordernd. Kleinräumige Vorhersagen sind abhängig von hochauflösten Anfangsbedingungen, die den Zustand der Atmosphäre zu Vorhersagebeginn beschreiben. Konventionelle Wetterbeobachtungen bieten jedoch nicht die notwendige Auflösung, im Gegensatz zu sichtbaren und infraroten Satellitenkanälen. Die Assimilation dieser Kanäle in operationellen NWP-Systemen ist jedoch noch sehr begrenzt, da dies mit der Verwendung von nichtlinearen Beobachtungsoperatoren, nicht-gaußschen Fehlerverteilungen und potenziell mehrdeutiger Informationen verbunden ist. Das Ziel der vorliegenden Arbeit ist es, das Potenzial der Assimilation von wolkenbeeinflussten Satellitenbeobachtungen im infraroten und sichtbaren Spektrum zur Vorhersage schwerer Stürme abzuschätzen und unser Verständnis des Effektes von nichtlinearen Beobachtungsoperatoren zu verbessern.

Diese Dissertation untersucht den potenziellen Nutzen von Satellitenbeobachtungen in einem idealisierten Rahmen, in dem der wahre Zustand der Atmosphäre bekannt ist und systematische Modell- und Operatorfehler vernachlässigt werden. Der wahre Zustand wird mit dem *Weather Research and Forecasting* Vorhersagemodell simuliert. Beobachtungen werden mittels Beobachtungsoperatoren aus dem wahren Zustand der Atmosphäre generiert und zur Assimilation mittels *Ensemble Adjustment Kalman Filter* verwendet. Es werden zwei Wetterszenarien untersucht. Im ersten Szenario wird isolierte und hochreichende Konvektion durch eine Warmluftblase ausgelöst, deren Position unsicher ist. Im zweiten Fall tritt hochreichende Konvektion im Bereich der gesamten Modelldomäne auf, wobei zusätzlich zur Position auch das Entwicklungsstadium der Konvektion unsicher ist.

Die Arbeit gliedert sich in drei Teile. Der erste Teil untersucht den Vorhersagenutzen der Assimilation von entweder sichtbaren oder infraroten Satellitenmessungen in den beiden zuvor genannten Szenarien im Vergleich zum Nutzen der Assimilation von Radarreflektivitätsmessungen. Die Ergebnisse zeigten, dass die Assimilation von Satellitenmessungen einen vergleichbaren Nutzen brachte wie die Assimilation von 3D-Radarreflektivitätsmessungen, wenn die Position der Warmluftblase unsicher war. Im Fall des unbekannten Entwicklungsstadiums der Konvektion war der Nutzen von Satellitenmessungen ähnlich dem von 2D-Radarmessungen.

Im zweiten Teil wurde die Assimilation einzelner Satellitenkanäle auf die kombinierte Assimilation mehrerer Kanäle erweitert und die Analyse der vertikalen Wolkenverteilung für unterschiedliche Wolkenbedingungen untersucht. Es zeigte sich, dass die kombinierte Assimilation die Mehrdeutigkeit einzelner sichtbarer und infraroter Kanäle verringern und die Analyse der Wolken verbessern kann.

Im dritten Teil wurde der Nichtlinearitätseffekt des Beobachtungsoperators von sichtbaren

und infraroten Messungen untersucht, der eine grundlegende Annahme der Assimilation verletzt. Dazu wurde die Abweichung berechnet, die durch den nichtlinearen Beobachtungsoperator entsteht. Schlussendlich wurde beurteilt, ob Messungen assimiliert werden sollten, selbst wenn sie nur auf geringe Vorhersagefehler hinweisen. Trotz gelegentlicher negativer Auswirkungen solcher Messungen, war es dennoch vorteilhaft sie in die Assimilation einzubeziehen, da sie dem Ensemble signalisieren, dass die ursprüngliche Vorhersage bereits sehr genau war.

Diese Arbeit trägt dazu bei, ein besseres Verständnis über den Nutzen von sichtbaren und infraroten Satellitenmessungen für kleinräumige Unwettervorhersagen zu gewinnen.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Numerical weather prediction and data assimilation . . . . .	1
1.2 Satellite observations for numerical weather prediction . . . . .	3
1.3 Convective-scale prediction and assimilation . . . . .	4
1.3.1 Observations for the convective scale . . . . .	5
1.3.2 Assimilation methods for the convective scale . . . . .	8
1.4 Research goals, approach and outline . . . . .	9
1.4.1 Research goals . . . . .	9
1.4.2 Research approach and outline . . . . .	11
<b>2 Theory &amp; Methods</b>	<b>13</b>
2.1 Numerical weather prediction . . . . .	13
2.2 Data assimilation . . . . .	15
2.2.1 Foundations . . . . .	15
2.2.2 The Ensemble Kalman filter . . . . .	17
2.2.3 The Ensemble Adjustment Kalman Filter . . . . .	19
2.3 Observation operators . . . . .	20
2.3.1 Infrared radiance . . . . .	20
2.3.2 Visible reflectance . . . . .	21
<b>3 Potential impact of assimilating visible and infrared satellite observations</b>	<b>23</b>
3.1 Introduction . . . . .	25
3.2 Results . . . . .	33
3.3 Conclusions . . . . .	41
<b>4 Combined assimilation of multiple satellite channels</b>	<b>47</b>
4.1 Introduction . . . . .	49
4.2 Results . . . . .	55
4.3 Conclusions . . . . .	66
<b>5 Nonlinearity effects of assimilating visible reflectance</b>	<b>73</b>
5.1 Introduction . . . . .	75
5.2 Methods . . . . .	75

5.3	Results . . . . .	77
5.4	Conclusions . . . . .	82
<b>6</b>	<b>Conclusions</b>	<b>89</b>
6.1	Summary and discussion . . . . .	89
6.2	Concluding remarks . . . . .	92
	<b>List of Acronyms</b>	<b>95</b>
	<b>Bibliography</b>	<b>97</b>
	<b>Acknowledgements</b>	<b>105</b>

# **Chapter 1**

## **Introduction**

Severe wind and torrential rain are responsible for an enormous loss of life and damage to property. These hazards are often caused by isolated cells of deep convection or as part of larger-scale weather systems, such as cyclones and monsoons, around the globe. For example, in the United States, the National Oceanic and Atmospheric Administration (NOAA) estimated a damage of \$ 530 billion incurred mostly by severe storms from 2018 to 2022 (Smith, 2020), excluding storm surge and droughts. In Europe, the European Environment Agency (EEA, 2023) estimated 6-10 thousand deaths and damages of € 346-400 billion due to meteorological and hydrological events in its member countries in 41 years (1980-2020). In the future, strong precipitation events are expected to produce even higher rain rates, as a higher temperature increases potential precipitable water (Seneviratne et al., 2012), which underlines the importance of reliable forecasts.

Over the last 50 years, weather prediction has improved tremendously. Today, numerical weather prediction models allow early and precise warnings of synoptic-scale events like cyclones. However, the prediction of convective-scale features, such as convection and associated hazards, is still unreliable. A major source of forecast error is the uncertainty in the model initial conditions that describe the state of the atmosphere at the beginning of a forecast. Reducing the uncertainty in initial conditions is the main purpose of *data assimilation*. At regular intervals, observations are assimilated to reduce errors of the model state, which subsequently serves as initial condition for the next forecast. The increasing resolution of numerical models calls for observations that constrain the initial conditions on all scales resolved by the model. However, conventional observations, e.g., from surface weather stations, are not sufficient to infer initial conditions at the convective scale. For this scale, visible and infrared satellite channels are potentially beneficial observation types. Despite this potential, the majority of such observations were not used operationally in 2021 (Gustafsson et al., 2018). Since recent developments enabled their assimilation, this thesis aims to study the assimilation of cloud-affected satellite observations and their potential to correct initial condition errors on the convective scale and thereby improve forecasts.

### **1.1 Numerical weather prediction and data assimilation**

Today, data assimilation is a vital part of numerical weather prediction (NWP). However, 120 years ago, when the original idea of NWP was conceived, the importance of initial condition errors on the predictability of the weather was not known.

The first crucial step towards weather prediction was to conceive the atmosphere as a mathematical model. This year, 2024, marks the 120-year anniversary that Vilhelm Bjerknes (1904) laid out the concept of how the weather could be predicted using the science of physics and mathematics<sup>1</sup>. He proposed that weather could be predicted by solving a system of partial differential equations (PDE) for seven state variables: temperature, humidity, pressure, three-dimensional wind, and air density. Moreover, he correctly recognized that the solution of any PDE requires the knowledge of initial and boundary conditions. For weather prediction, this meant that the initial atmospheric state needs to be known to predict the weather at a later time. He understood that this was a great issue at that time since there was no dense network of weather observations yet. Although no amount of measurements would be enough to ascertain the initial state completely, assimilating observations improves the initial condition and thereby combines the information of observations from different times.

The second crucial step was to solve the equations of numerical weather prediction. For this, several challenges had to be solved: To solve PDEs numerically, they have to be *discretized*: In 1922, L.F. Richardson proposed discretization, which divides Central Europe into grid boxes as illustrated in Figure 1.1. Then, he computed the average temperature, pressure, and wind for these grid boxes. Another challenge was to finish the computation of a 1-day forecast in a few hours, which is necessary for the forecasts to have practical value. This became possible only in 1972 when the whole system of "primitive" equations could be solved for operational forecasting purposes (Lynch, 2008).

The third crucial step was to notice the importance of accuracy and balance in the initial conditions. Charney et al. (1950) found that geostrophically balanced initial conditions can avoid exciting spurious gravity waves. Later, Lorenz (1963) noticed that even the errors caused by the rounding of values can lead to divergent solutions. Thereby, he rediscovered *deterministic chaos* in the domain of weather prediction, by which small errors in the initial conditions can grow exponentially. Thus, the accuracy of forecasts is limited by the accuracy of initial conditions. To improve the accuracy of initial conditions, a new technique was developed. Instead of simply interpolating between observations, the idea was to combine observations and forecasts, a procedure which is called *data assimilation*. Therein, observations are integrated into a previous forecast, combining prior knowledge of older observations with the latest observations using the knowledge of physical laws (the forecast model). To quantify the uncertainty of forecasts, ensemble prediction systems (EPS) were introduced, e.g., in 1992 at the European Centre for Medium-Range Weather Forecasts (ECMWF). In an EPS, many forecasts are started from slightly different initial conditions. The resulting divergence in forecasts is then used to estimate forecast uncertainty. Nowadays, this information on forecast uncertainty is often also incorporated in the process of data assimilation by ensemble or hybrid data assimilation schemes. In summary, it became clear that data assimilation is a crucial component of numerical weather prediction.

---

<sup>1</sup>This was similarly proposed by Abbe (1901).

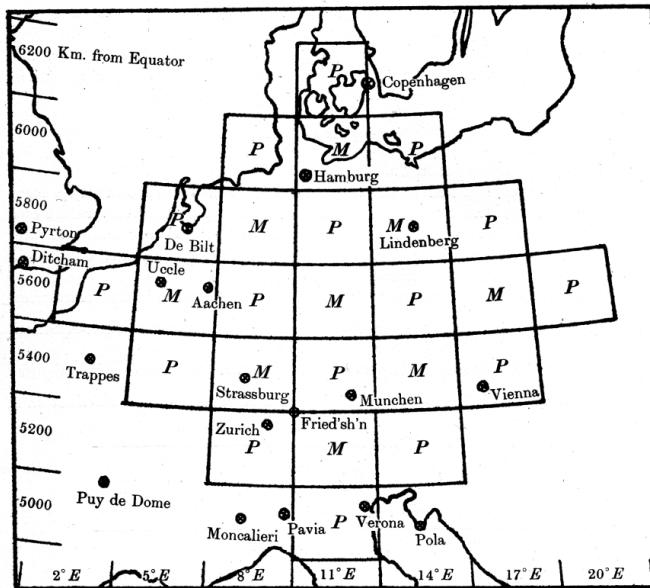


Figure 1.1: An illustration of the discretization by Richardson (1922, second edition 2007). A prediction computes the grid-box average of atmospheric variables for each grid box. Note that grid boxes are much smaller today.

## 1.2 Satellite observations for numerical weather prediction

In the 1970s, global numerical weather prediction (NWP) had been unreliable in the southern hemisphere due to insufficient observations and unsuitable assimilation methods. However, today, satellite observations with nearly global coverage are the dominant source of information for global NWP (Geer et al., 2017; Eyre et al., 2020). This success did not come over night. In the 1970s, first *sounding* instruments on polar-orbiting satellites provided vertical profiles of temperature and humidity. In the 1980s and 1990s, wind information, derived from the movement of clouds (atmospheric motion vectors, AMV), were first assimilated and introduced to operational NWP.

Moreover, the development of advanced data assimilation methods was crucial to leverage the potential of satellite observations. In the early years, satellite observations were assimilated in a sub-optimal way, in which the measured radiances were transformed into profiles of, e.g., temperature, by *retrievals*. Among other issues, these profiles imitated a high vertical resolution, which the data did not provide. Ultimately, the issues were solved by introducing a new data assimilation method, 3D-Variational data assimilation (3D-Var), which allowed the direct assimilation of radiances. It was introduced at NCEP and ECMWF in 1995/1996 (Courtier et al., 1998). Two years later, 4D-Variational data assimilation (4D-Var) was introduced, which brought another boost to global NWP (Rabier et al., 1998). 4D-Var considers that observations are taken at different times and derive increments that are consistent with model dynamics. One example is the *tracer-advection mechanism*, which can correct mislocation errors at the observation time by adjusting the wind field

before the time of observation. It acts to move the position of the feature at the time of the observation, as shown by Peubey and McNally (2009) for clear-sky radiances. They also showed that 4D-Var assimilation of clear-sky radiances is a significant source of wind information in the tropics. Regular upgrades to the 4D-Var brought a steady improvement of forecast skill in the following years and closed the forecast skill gap between the northern and the southern hemispheres. For example, a 7-day forecast in the southern hemisphere in 1997 became as skillful as a 5-day forecast in the northern hemisphere in 1987 (Figure 1.2).

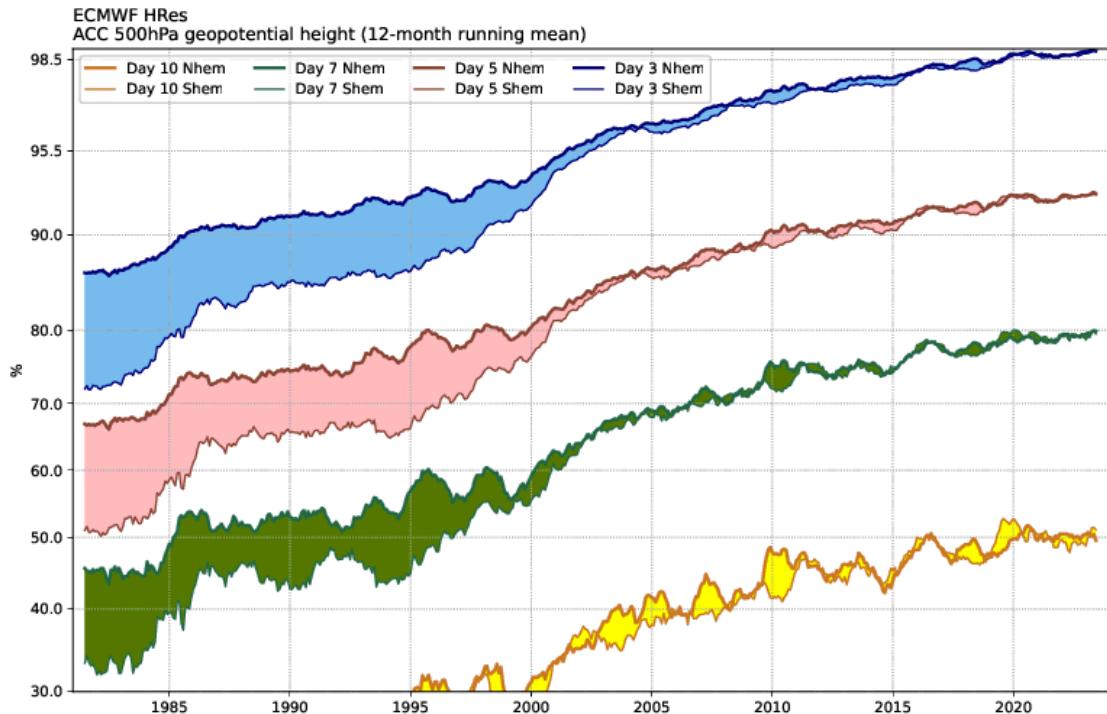


Figure 1.2: A measure of forecast skill, the correlation of 500-hPa geopotential height forecasts (3, 5, 7, 10-day forecasts) and the verifying analysis, expressed as anomalies from the climatological height. From ECMWF.

### 1.3 Convective-scale prediction and assimilation

The convective scale (also km scale or meso- $\beta$  scale, Fujita, 1981) describes meteorological features with a size between 4 and 40 km. To simulate such features at least partially, the grid resolution needs to be increased to less than 4 km, given that the effective resolution of a forecast is about 4–7 times larger than the grid resolution (Skamarock, 2004; Bierdel et al., 2012).

Convective-scale weather is typically predictable on shorter timescales than large, synoptic-scale weather. While the predictability of synoptic-scales is believed to be around 10 days,

estimates of convective-scale predictability vary, ranging from a few hours to a few days (Hohenegger and Schär, 2007; Zhang et al., 2016; Zhang et al., 2019; Žagar and Szunyogh, 2020). Since global NWP models are specialized for synoptic-scale, medium-range prediction, their horizontal resolution is limited. In contrast, regional, limited-area models, can resolve the convective scale, and with the rapidly growing computing power also global NWP models might be able to resolve the convective scale soon.

Having a sufficient numerical grid resolution, however, is not enough to make successful predictions at this scale. Successful convective-scale forecasts also require observations that provide accurate information on the convective scale for estimating the model initial state.

### 1.3.1 Observations for the convective scale

Today, a major source of uncertainty in convective-scale forecasts is the scarcity of observations to constrain the convective-scale initial conditions. First, surface observations measure temperature, humidity and wind accurately. However, especially over land, their measurements are often not representative for a larger neighborhood and are difficult to assimilate since models often do not represent the boundary layer well. Radiosonde observations provide direct in-situ measurements of the vertical profile but are very scarce. Aircraft measurements are more abundant, however, inhomogeneously distributed in space and time as fewer aircraft fly at night, aircraft often follow the same airways, they usually avoid meteorologically interesting areas of convection, and they observe the lower troposphere only close to airports. Thus, the amount of conventional observations is limited and inhomogeneously distributed, as shown in Necker et al. (2018) for the German regional model (Figure 1.3).

A large potential lies in the wide variety of indirect remote-sensing observations. Ground-based wind profilers observe scattering from turbulence-induced fluctuations in atmospheric density and humidity (Petty, 2006) from which wind profiles can be derived. Ground-based lidar can also be used to derive lower tropospheric winds, and latest research on Atmospheric Emitted Radiance Interferometers (AERI) aims to observe the lower tropospheric water vapor field (e.g. Lewis et al., 2020). Yet, observations of these instruments are limited to the small neighborhood around the observation site, and so far, the cost of these instruments is too high to install a very high number of them in the model domain.

Networks of radars can monitor a larger area almost without gaps. However, their global coverage is limited. The sea, sparsely populated regions, and low-income countries are often not covered, and data quality is often inhomogeneous across national borders. Additionally, beam-blocking by orography limits data quality in mountainous regions such as the Alps. Rain radars observe larger, precipitating particles by measuring reflected microwave radiation, allowing us to infer the thermodynamic structure within storms, including the wind field. Rain radars can, however, not sample the pre-convective environment before larger hydrometeors have formed.

Satellite-based instruments provide global coverage. Atmospheric motion vectors (AMV) are measurements of the temporal displacement of clouds, which relates to the mean wind field at the height of the cloud. Their height assignment is, however, intrinsically uncertain,

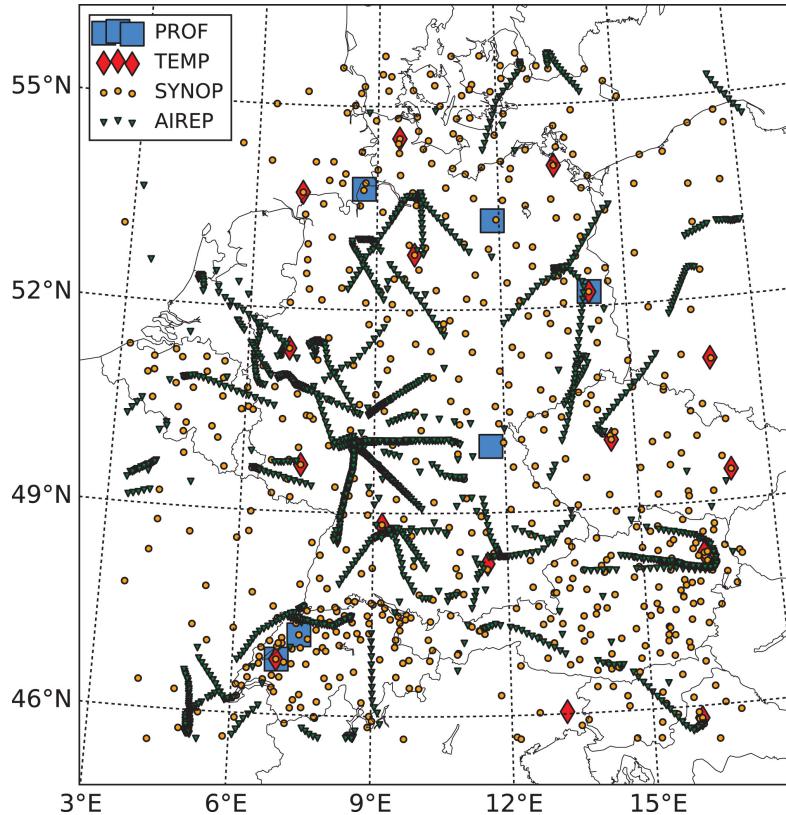


Figure 1.3: The distribution of the assimilated observations in the domain of the German regional model at 12 UTC on 29 May 2016. PROF: Wind profiler; TEMP: Radiosonde; SYNOP: surface weather station; AIREP: Aircraft observations. From Necker et al. (2018).

which leads to correlated and systematic errors (Folger and Weissmann, 2016). All-sky microwave and clear-sky infrared radiances from low Earth orbits (LEO) are of substantial value to global NWP (Geer et al., 2017). However, such observations are less suited for convective-scale NWP since a LEO satellite revisits one region only once per day, which does not allow an uninterrupted monitoring of convective initiation. GNSS total delay evaluate the signal delay between satellites and ground stations which allows to infer the vertically integrated amount of water vapor.

### **Visible and infrared observations from geostationary satellites**

Observing convective initiation and development requires a high spatial resolution and temporal frequency of observations. Instruments on geostationary satellites can provide such a high temporal ( $\leq 15$  min) and spatial ( $< 5$  km) resolution that is similar to the grid resolution of convective-scale NWP and are therefore commonly used for observation-based nowcasting of convective systems. Moreover, contrary to weather radar, their data is available over

land and sea, is homogeneous across borders, and is not affected by orographic shadowing.

The assimilation of visible and infrared channels is the focus of this thesis. In clear-sky conditions, infrared channels observe the humidity field, an essential ingredient of convection. Of all the infrared channels on Meteosat Second Generation satellites, the 6.2 and 7.3  $\mu\text{m}$  channels are unique in that they are sensitive only to emissions from water vapor and clouds, but not from the surface. This makes their assimilation easier compared to other infrared channels that are also sensitive to surface emission (e.g. window channels). In the absence of clouds, the 6.2  $\mu\text{m}$  channel observes upper, and the 7.3  $\mu\text{m}$  channel observes mid- and upper tropospheric water vapor, as indicated by the weighting function in Figure 1.4.

A large fraction of observations is, however, cloud-affected. In cloudy conditions, infrared radiances are strongly sensitive to cloud height. Infrared 7.3  $\mu\text{m}$  radiance is sensitive to mid-level clouds but less sensitive to lower clouds (Figure 1.4). The 0.6  $\mu\text{m}$  visible channel observes the reflectance of clouds in the visible spectrum during daytime. As the visible wavelength is not affected by emission or absorption, radiation from any vertical level is not attenuated and thus height-independent. Therefore, the visible channel can also detect low-level clouds. However, the assimilation of visible and infrared cloud-affected observations is challenging as they exhibit a nonlinear relationship to the model state. Figure 1.5 by Geiss et al. (2021) shows the relation between 0.6  $\mu\text{m}$  visible reflectance (left) and 10.8  $\mu\text{m}$  infrared radiance (right) to cloud liquid and ice water path, assuming an effective radius of 10  $\mu\text{m}$  and 40  $\mu\text{m}$  for liquid and ice particles, respectively.

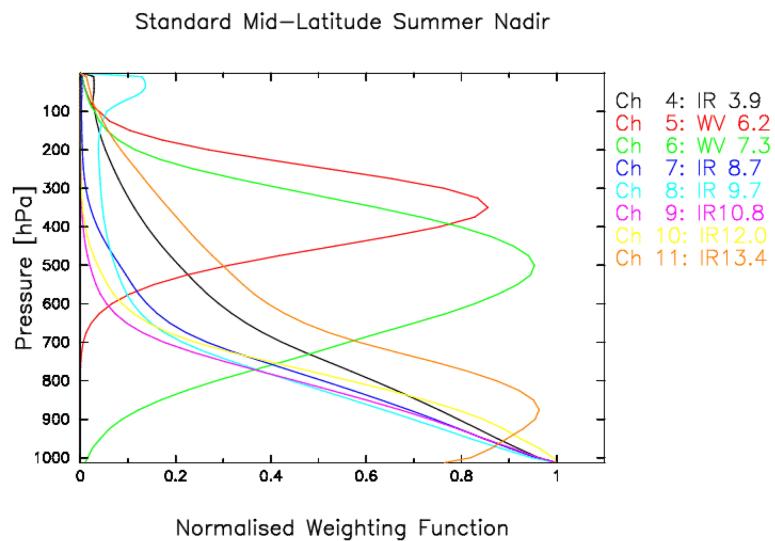


Figure 1.4: The clear-sky weighting functions of eight infrared channels of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard MSG satellites. The weighting function describes the sensitivity of radiances on temperature and humidity in different altitudes, as defined in section 2.3. This thesis used channels 5 and 6. © EUMETSAT (2022).

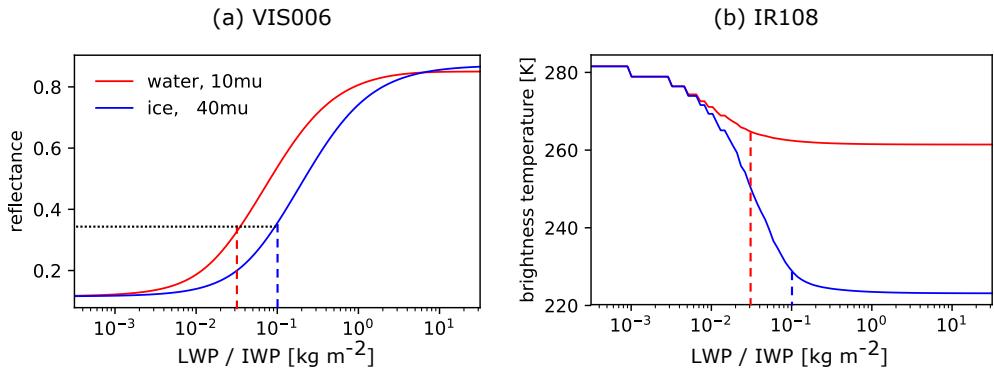


Figure 1.5: Nonlinear relationship between the observed visible reflectance (a) or  $10.8\text{ }\mu\text{m}$  infrared brightness temperature and the cloud liquid/ice water path (LWP and IWP respectively). Reproduced from Geiss et al. (2021).

### 1.3.2 Assimilation methods for the convective scale

Although global NWP models nowadays assimilate a tremendous amount and variety of satellite observations, the use of satellite observations in regional convective-scale models is still very limited. Assimilating cloud-affected satellite observations on the convective scale is still a challenging task and calls for specialized assimilation methods. Errors at the convective scale grow non-linearly, often due to moist convection. This questions the use of linearized dynamics in 4D-Var for convective-scale NWP. Hohenegger and Schär (2007) estimated that the tangent-linear approximation breaks down at 1.5 h for a 2.2 km resolution model, compared to 54 h for an 80 km resolution model.

Today, different operational centers use different assimilation algorithms for convective-scale NWP. Apart from 3D-Var and 4D-Var, Ensemble Kalman filters (EnKF) are used to assimilate observations on the convective scale. EnKF use the ensemble variance to estimate the forecast error which improves upon a constant climatological estimate of the forecast error in 3D-Var. Moreover, the ensemble covariance of the model state is used to spread the innovation in space. Innovations (or first-guess departures) describe the difference between observations and model-equivalents (see Figure 1.6). Certain implementations of the EnKF, for example the Ensemble adjustment Kalman filter (EAKF), first compute an analysis in observation space and then use the ensemble covariance between the observation and the model state to project the observation-space increments back to model state increments. Note that the use of the covariance implies a linearity assumption and avoids an explicit linearization as used in variational methods. Aside from EnKFs, nonlinear assimilation methods like particle filters do not rely on the Gaussian distributions but sample the whole probability distribution. Sampling a high-dimensional distribution, however, requires an unbearable amount of particles, which makes their practical implementation challenging.

Despite their potential positive benefit, cloud-affected satellite observations are still not assimilated in many operational centers, due to a number of challenges. First, the assimilation assumes that model-equivalents of observations can be computed without

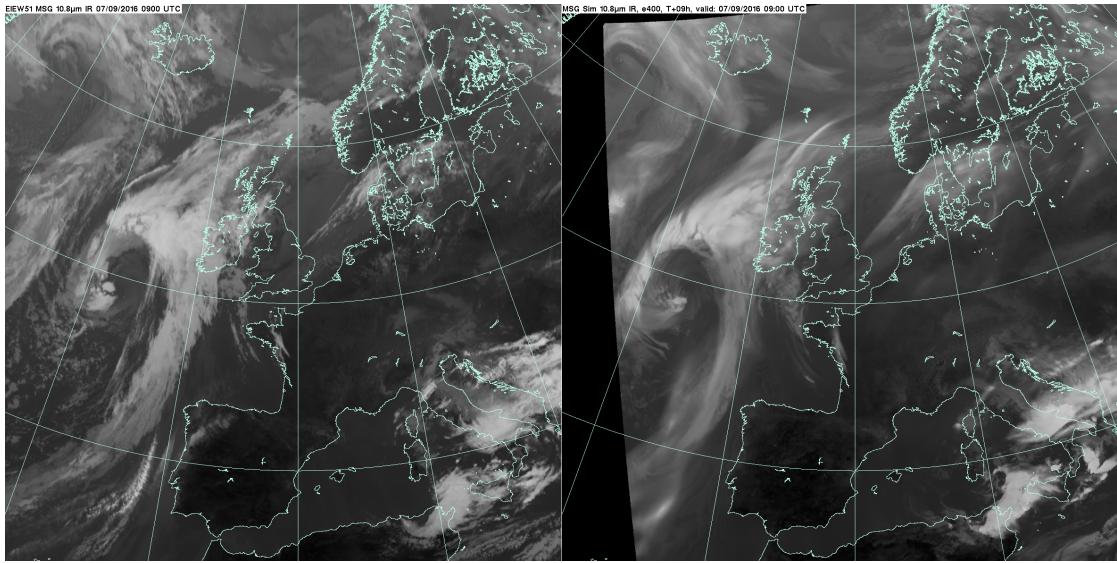


Figure 1.6: Comparison of observed (left) and forecasted (right)  $10.8\text{ }\mu\text{m}$  radiance, by a 9-hr Met Office-forecast with 4.4 km resolution. ©EUMETSAT (2024).

systematic errors compared to observations. Moreover, the model representation of cloud hydrometeors is simplified and often inconsistent with observation operators (Li et al., 2022), and the model does not resolve small-scale processes. Additionally, observation operators approximate the real radiative transfer due to computational constraints. Last but not least, assumptions of linearity and Gaussianity of observation error distributions are violated, which can make the assimilation less effective.

## 1.4 Research goals, approach and outline

### 1.4.1 Research goals

This thesis was conducted under the overarching research question,

**What are the benefits of assimilating cloud-affected satellite observations for predicting convective-scale storms?**

and addresses three related goals, which are briefly summarized below.

#### 1) Estimate the potential impact of assimilating visible and infrared satellite observations

Despite their potential benefit, visible and infrared satellite channels are still not widely assimilated in operational NWP systems. Moreover, not much was known about their benefit compared to other observations. In 2020, when the preparation for the first publication

of this thesis was started, there had been some research on the infrared assimilation, but only two studies on the assimilation of visible observations in convective-scale NWP. In a two-day case study using the near-operational setup of Deutscher Wetterdienst (DWD), Scheck et al. (2020) investigated the impact of assimilating visible observations in addition to other conventional observations but could not compare the impact to the impact of other observation types. The only comparable study to this thesis is the idealized study of Schröttle et al. (2020), which also used the assimilation framework of DWD and assimilated only visible or infrared satellite observations, but no other observations. They suggested that infrared assimilation might have a magnitude of impact similar to radar assimilation, but did not conduct a direct comparison. Thus, further studies are needed that include a direct comparison to radar assimilation, use a different NWP model and assimilation system as well as different weather scenarios. Regarding the impact of infrared assimilation, a wide range of phenomena was studied, including typhoons, mesoscale convective systems, and other severe storms (Jones et al., 2015; Jones et al., 2016; Honda et al., 2018; Sawada et al., 2019; Jones et al., 2020; Chan et al., 2020; Zhu et al., 2022). Most of these studies, however, did not allow a comparison to the impact of other observation types, such that the relative impact of infrared assimilation is still unclear. Moreover, the fact that cloud-affected infrared observations were not operationally assimilated, may indicate a lack of understanding. Chapter 3 tries to close this gap by an independent study which estimates the relative impact of assimilating visible, infrared, and radar observations.

## **2) Evaluate the combined assimilation of multiple ambiguous satellite channels**

Visible and infrared observations are ambiguous if they are considered on their own. Visible reflectance is insensitive (ambiguous) to the height of the cloud. It can show a high reflectance for low-level clouds as well as for high-level clouds. In contrast, infrared observations strongly depend on the cloud top height. Similarly, infrared observations are insensitive to cloud water or ice below an opaque cloud top. Therefore, a cumulonimbus cloud can result in the same observation as an anvil of a dissipated cumulonimbus, although the thermodynamic profile is different.

Thus, it is reasonable to believe that the assimilation of visible and infrared channels can be complementary, for example, since the infrared is sensitive to cloud height, while the visible is not. This motivated the study in Chapter 4, which investigates whether the combined assimilation of visible and infrared channels can exploit the complementary information.

## **3) Analyze the nonlinearity of visible and infrared observation operators**

Despite the nonlinearity of observation operators for visible and infrared channels, the forecast impact of these observations was found to be overall beneficial. The nonlinearity of the observation operator itself is well known as shown in Figure 1.5 for the relationship between vertically integrated cloud water content. Given the linearity assumption or linearization in current data assimilation schemes, the operator nonlinearity leads to deviations from the expected solution assuming linear operators. Scheck et al. (2020) analyzed the deviations

and found considerable effects for observations with small first-guess departures. This finding motivated a detailed investigation of operator nonlinearity of visible observations with small first-guess departures in Chapter 5, since little is known about the consequences in a complex NWP and data assimilation system. Finally, Chapter 5 quantifies the systematic and random deviations caused by operator nonlinearity for the visible and an infrared channel.

### 1.4.2 Research approach and outline

Similarly as Schröttle et al. (2020), this thesis uses observing-system simulation experiments (OSSE) to estimate the potential impact of assimilating visible, infrared and radar observation, because it allows to precisely quantify the error reduction from the assimilation given that the truth is known. The truth, also known as nature run, was simulated by running a forecast model from selected idealized initial conditions. The boundary conditions were chosen to be doubly periodic to avoid external influences. The initial conditions were carefully chosen such that two weather scenarios of summertime deep convection naturally evolve from them. The first scenario features isolated convection triggered by a warm-bubble, similar to Snyder and Zhang (2003) and Tong and Xue (2005). The other scenario (as in Bachmann et al., 2020; Schröttle et al., 2020) features deep convection scattered throughout the domain, where the location of convection is totally random, as the surface and the initial conditions are totally homogeneous except of random perturbations. Observations were generated from the nature run and assimilated into a 40-member forecast ensemble using the Ensemble adjustment Kalman filter (EAKF). The EAKF was chosen because it allows a detailed analysis due to its serial formulation. The subsequent evaluation shows a potential impact that presents an upper bound on the impact in an operational model. However, by putting the potential impact of one observation type in relation to the potential impact of another observation type, we can learn more about the relative contributions of different observation types, even in idealized experiments.

To evaluate the nonlinearity of the observation operators, the OSSE-setup was used to conduct additional experiments which show the nonlinearity effects in single-, and double-observation experiments, and many-observation experiments with cycled assimilation. The employed EAKF assimilation algorithm is ideal for this approach as due to its serial formulation, the detailed effect of every observation can be studied.

## Outline

Before going into the research studies, Chapter 2 briefly summarizes the relevant theoretical background of numerical weather prediction, data assimilation, and observation operators. Thereafter follow three manuscripts, of which the first (Chapter 3) is published and the second (Chapter 4) is under review in the Quarterly Journal of the Royal Meteorological Society (QJRMS). The third manuscript (Chapter 5) is currently in preparation for submission. Chapter 3 estimates the potential impact of assimilating either visible or infrared cloud-affected satellite observations relative to the impact of the more commonly assimilated observations of radar reflectivity. Chapter 4 extends the impact assessment to the combined assimilation of visible and infrared observations and shows their complementary value for

improved clouds. Chapter 5 evaluates the consequences of observation-operator nonlinearity for the visible and infrared channels. Finally, Chapter 6 discusses the contributions of the manuscripts to answering the overarching research question.

# Chapter 2

## Theory & Methods

This chapter introduces the methods that were applied in this thesis. First, I start by going briefly over the Numerical weather prediction (NWP) model. Second, data assimilation will be introduced more fundamentally, followed by the derivation of algorithms and a review of assumptions. Lastly, observation operators are briefly introduced.

### 2.1 Numerical weather prediction

This thesis applied the Weather Research and Forecasting (WRF) model (Skamarock et al., 2021). The prognostic variables are potential temperature, geopotential, dry-air surface pressure, inverse density, the three cartesian velocity components, and humidity, cloud water, rain, and snow mixing ratios. The model describes the atmosphere using fully compressible, Eulerian, non-hydrostatic equations. The vertical coordinate is terrain-following between a variable pressure at the surface and a constant pressure at the model top. The equations are derived from basic principles of physics, that is, conservation laws, which are applied to small volumes of air:

- The equations of motion describe the conservation of momentum (in an inertial reference frame) of air in three dimensions.
- The mass continuity equation describes the conservation of mass.
- Another budget equation ensures the conservation of water mass, considering that phase transitions lead to sources and sinks, as e.g. condensation is a source of liquid water but a sink of vapor.
- The thermodynamic equation relates temperature, pressure tendency, and the release of latent heat from phase changes of water.
- The ideal gas law relates the states of temperature, pressure, and density through the exchange of momentum of air molecules.

For a prediction, the differential equations must be solved numerically by approximating differentials with differences, a process called discretization, which leads to a discrete solution in space and time. A guide to numerical solutions is given by Durran (2010). Note that only grid-box-averages are predicted, which implies a lower bound on the size of features that the model can resolve. This thesis used a grid resolution (discretization) of 2 km, which implies that only features larger than about 8-14 km can be explicitly resolved by the model

dynamics (Skamarock, 2004; Bierdel et al., 2012). Processes on smaller scales need to be represented by so-called parametrization schemes. Parametrizations calculate the source and sinks of water vapor, liquid water, momentum, and heat due to unresolved processes as e.g. cloud microphysics, radiation, turbulence and friction.

The discretization in space partitions the spatial domain of size  $400 \times 400$  km into  $200 \times 200 \times 50$  cuboid grid boxes of length of 2 km and a depth between 50 m near the surface and 500 m in the stratosphere. A historical illustration of such discretization can be found in Figure 1.1. The prediction can be formally described as the integration of a state vector  $\mathbf{x}$  by a model  $\mathcal{M}$  from initial time  $t$  to time  $t + 1$ ,

$$\mathbf{x}_{t+1} = \mathcal{M}(\mathbf{x}_t), \quad (2.1)$$

where  $\mathbf{x}$  consists of the values of atmospheric variables at all grid boxes,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}. \quad (2.2)$$

Note that the state vector of an operational model is typically  $\mathcal{O}(10^8)$ .

### Idealized configuration

The WRF model offers two configurations, of which both require initial conditions, but they differ in their use of boundary conditions. The "real" configuration requires time-varying boundary conditions. The "ideal" configuration is used for research and allows the user to specify special boundary conditions, as for example, the one used for this thesis: the periodic boundary condition. Periodic boundary conditions have the advantage that the user can study the atmospheric dynamics inside the domain without influence from other regions. Thereby one can overcome a disadvantage of real-data regional studies, namely, to what extent does the error in the boundary conditions influence the error inside the domain. With periodic boundary conditions, however, the scale of the simulated model dynamics is limited to less than half of the domain size (e.g., a larger cyclone would decelerate itself).

For the purpose of this thesis, the model surface was assumed flat and homogeneous of type "dryland cropland and pasture" with 50% vegetation area fraction. The choice of parametrizations follows the NOAA High-Resolution Rapid Refresh (HRRR) model (Benjamin et al., 2016) with microphysics by Thompson et al. (2008), radiative transfer option RRTMG (Iacono et al., 2008), Mellor-Yamada-Nakanishi-Niino (option MYNN 2.5, Nakanishi and Niino, 2006), surface model option Noah (Tewari et al., 2004) and without convection parametrization. The Coriolis force was shut off as in Lange and Craig (2014), Bachmann et al. (2019), Bachmann et al. (2020), and Schröttle et al. (2020), because it does not change the dynamics at this timescale and because it would have led to a veering of the wind with time in a domain with periodic boundary conditions.

## 2.2 Data assimilation

### 2.2.1 Foundations

In the early days of NWP, forecasts were initialized by an interpolation of observations. However, observations are scarce in space and time, which leads to large interpolation errors. Generally, the number of observations is usually much smaller than the size of the model state vector  $\mathbf{x}$ . To overcome this limitation, the information from observations is nowadays combined with information from the latest available forecast in the process of data assimilation (DA).

Using the assumptions for the discretization of a typical convective-scale NWP model leads to more than  $10^6$  scalar temperature values, which must be specified. However, not only temperature but also other meteorological variables like humidity, pressure, geopotential, and wind need to be defined at the same resolution as temperature. The values at all gridpoints are compiled in the (model) state vector of dimension  $\mathcal{O}(10^8)$ . On the other hand, there are much fewer observations which are collected in an observation vector

$$\mathbf{y}_o = \begin{bmatrix} y_{o,1} \\ y_{o,2} \\ \vdots \\ y_{o,n} \end{bmatrix}. \quad (2.3)$$

In order to find a model state that reproduces an observation, a model state  $\mathbf{x}_b$  (the prior) needs to be converted into a first-guess  $y_b$  for the observation (model-equivalent of the observation) by an observation operator  $\mathcal{H}$ :

$$y_b = \mathcal{H}(\mathbf{x}_b). \quad (2.4)$$

By comparison of the dimension of observation and model state, we see that this relation can not be directly inverted (ill-posed problem).

### Bayes theorem

Considering that predictions are subject to uncertainty, the *Bayes theorem* of probability theory combines predictions and observations consistent with their average error statistics. The probability density function (PDF) which combines the information of observations and forecasts, is given as a product of the prior PDF and the observation "likelihood" PDF

$$\underbrace{f(\mathbf{x}|\mathbf{x}_b, \mathbf{y}_o)}_{\text{posterior PDF}} = \underbrace{f(\mathbf{x}|\mathbf{x}_b)}_{\text{prior PDF}} \cdot \underbrace{f(\mathbf{y}_o|\mathbf{x}, \mathbf{x}_b)}_{\text{likelihood PDF}} \cdot \frac{1}{f(\mathbf{y}_o|\mathbf{x}_b)} \quad (2.5)$$

where the last term normalizes the resulting probability to 1.

The most probable state is the state that maximizes the posterior PDF:

$$\mathbf{x}_a = \arg \max_x f(\mathbf{x}|\mathbf{x}_b, \mathbf{y}_o) = \arg \max_x (f(\mathbf{x}|\mathbf{x}_b) \cdot f(\mathbf{y}_o|\mathbf{x}, \mathbf{x}_b)) \quad (2.6)$$

**Assumption 1:** Although the Bayes theorem (equation 2.5) is valid for any probability distribution, operational assimilation algorithms usually assume a Gaussian PDF for computational efficiency. Assuming an  $n$ -dimensional Gaussian

$$f(\mathbf{x}) = \frac{1}{2\pi^{n/2}\sqrt{|\Sigma|}} \exp\left(\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)\right) \quad (2.7)$$

for the prior and observation likelihood PDF results in a simple solution since the PDF is completely described by the expectation  $\boldsymbol{\mu}$  and the covariance matrix  $\Sigma$ . Figure 2.1 illustrates the Bayes law applied to a Gaussian prior and observation PDF, which yields a Gaussian posterior PDF.

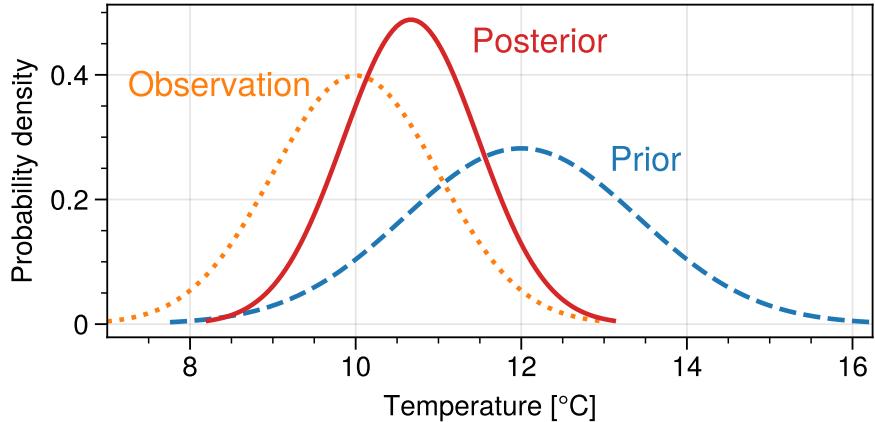


Figure 2.1: Illustration of the Bayes law (Equation 2.5) for Gaussian PDFs, combining a prior probability distribution (blue dashed line) and an observation likelihood (orange dotted line). The result is a posterior probability distribution (red solid line).

For a specific prior forecast  $\mathbf{x}_b$  and a specific observation  $\mathbf{y}_o$ , we can find the most likely atmospheric state by maximizing equation 2.6. Alternatively, the cost function

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^\top \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y}_o - \mathcal{H}(\mathbf{x}))^\top \mathbf{R}^{-1} (\mathbf{y}_o - \mathcal{H}(\mathbf{x})) \quad (2.8)$$

can be minimized with the error covariance matrices  $\mathbf{B}$  and  $\mathbf{R}$ . The result is equivalent to the Bayes theorem under Assumption 1 plus the following additional assumptions (Bátkai et al., 2016):

**Assumption 2:** One assumption is that prior forecasts are unbiased, i.e. that the prior forecast is equal to the expectation  $\mathbb{E}()$  of the true state

$$\mathbb{E}(\mathbf{X} | \mathbf{X}_b = \mathbf{x}_b) = \mathbf{x}_b. \quad (2.9)$$

This is an approximation since prior forecast can contain systematic errors.

**Assumption 3:** Moreover, model-equivalents are assumed to be unbiased, which means that the expectation of the observation, given a concrete state  $\mathbf{x}$  is equal to the forward operator applied to the state

$$\mathbb{E}(\mathbf{Y}_o | \mathbf{X} = \mathbf{x}) = \mathcal{H}(\mathbf{x}), \quad (2.10)$$

This is an approximation since forward operators may contain systematic errors.

**Assumption 4:** The error statistics of observations and first-guess forecasts are perfectly known and can be described by the respective covariance matrices  $\mathbf{R}$  and  $\mathbf{B}$ . In particular, model errors exhibit large fluctuations and are usually only crudely known.

**Assumption 5:** A linear observation operator  $\mathbf{H}$  exists so that  $\mathcal{H}(\mathbf{x}_b + (\mathbf{x} - \mathbf{x}_b)) \approx \mathcal{H}(\mathbf{x}_b) + \mathbf{H}\mathbf{x} - \mathbf{H}\mathbf{x}_b$ .

With these assumptions, we can find the minimum of the cost function by setting the derivative of the cost function to zero and resolve the equation for the analysis model state  $\mathbf{x}_a$ :

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K} [\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b)] \quad (2.11)$$

with

$$\mathbf{K} = \mathbf{B}\mathbf{H}^\top (\mathbf{H}^\top \mathbf{B}\mathbf{H} + \mathbf{R})^{-1}. \quad (2.12)$$

## 2.2.2 The Ensemble Kalman filter

This thesis applies a variant of an Ensemble Kalman filter (EnKF), which is based on the Kalman filter. Kalman (1960) derived how the forecast error covariance matrix needs to evolve over time, yet, in an impractical way for the high-dimensional problem of data assimilation and the nonlinear model equations of NWP. However, the forecast error covariance matrix can be estimated from an  $N$ -member ensemble of forecasts  $\mathbf{x}_b^i$ ,  $i = 1, \dots, N$  as introduced in (Houtekamer et al., 1996; Houtekamer and Mitchell, 1998):

$$\mathbf{B} \approx \mathbf{P}_b = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_b^i - \bar{\mathbf{x}}_b) (\mathbf{x}_b^i - \bar{\mathbf{x}}_b)^\top = \frac{1}{N-1} \mathbf{X}\mathbf{X}^\top \quad (2.13)$$

where  $\mathbf{X}$  are ensemble perturbations from the mean. Then, we can calculate the posterior covariance matrix

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H}) \mathbf{P}_b. \quad (2.14)$$

Estimating the covariance from a finite sample introduces sampling error, since computer resources only allow about  $\mathcal{O}(100)$  ensemble members. Resulting issues are insufficient ensemble variance (spread) and spurious prior correlations between observations and state variables due to the limited sample size (Anderson, 2016). The error covariance (and spread)

of an optimal analysis is given in equation 2.14. In reality, this covariance tends to underestimate the true error covariance because of violated assumptions on linearity, Gaussianity, and systematic errors or a simplification of the optimal solution due to computational restrictions. This results in a sub-optimal analysis mean state and underestimated ensemble variances. The underestimation can be mitigated by background error inflation and statistical corrections (Anderson, 2012; Anderson, 2016; Necker et al., 2020). Spurious prior correlations are caused by small ensembles. Small ensembles lead to a systematic overestimation of the magnitude of covariance between the model states and model-equivalents (Necker et al., 2020) in equation 2.12, i.e.  $\mathbf{BH} = \mathbf{X}(\mathbf{HX})^\top$ . The resulting overestimation of  $\mathbf{K}$  leads to overestimated increments (equation 2.11) and underestimated variance (equation 2.14). Additionally, the ensemble variance from equation 2.13 does not account for uncertainty due to model error (as equation 2.14 requires) unless a stochastic model is used for the integration of the ensemble in time.

### Covariance localization

Localization primarily mitigates spurious correlations. The affordable ensemble size is much smaller than the degrees of freedom of the model, which leads to sampling errors due to the small sample size. Today, the main approach is a physically-motivated distance-based localization, based on the idea that there is no sampling error in case of a zero covariance in equation 2.13 or 2.19. The physical motivation comes from knowledge of the primitive equations, where dynamical effects are local, but can propagate in space. Therefore, we expect that small-distance correlations are less affected by sampling error than long-distance correlations. To reduce the magnitude of correlations, one multiplies the sample covariance value with a factor ( $\leq 1$ ) which decreases with distance. A typical length at which the factor reaches zero would be (up to) 200 km for radiosonde observations (Schraff et al., 2016) and 40 km for radar and cloud-affected satellite observations.

### Background error inflation

Inflation is a direct way of increasing the ensemble variance. In general, inflation can be additive or multiplicative, and applied prior or posterior to the assimilation. Prior inflation tends to mitigate variance underestimation due to model error, as it is applied after model integration and before assimilation. Posterior inflation is done after assimilation and can thus mitigate sampling errors caused in the assimilation (Gharamti et al., 2019). In practice, posterior inflation also has the advantage that a potentially unphysical perturbation can evolve into a physically consistent perturbation during the model integration. This thesis applied posterior multiplicative inflation using the method "relaxation to prior spread", where the inflated posterior spread  $\sigma'_a$  is set to be a weighted average of the prior spread  $\sigma_b$  and posterior spread  $\sigma_a$ :

$$\sigma'_a \leftarrow \sigma_b(1 - \alpha) + \sigma_a\alpha \quad (2.15)$$

### 2.2.3 The Ensemble Adjustment Kalman Filter

This thesis applied the Ensemble adjustment Kalman filter (EAKF) by Anderson (2001) through a software package called Data Assimilation Research Testbed, abbreviated DART (UCAR/NCAR/CISL/DARes, 2022). The EAKF is an implementation of the EnKF, but it is a deterministic square root filter which has been implemented in a scalar formulation in DART and assimilates observations serially (one after another). The scalar and serial formulation assumes temporally (Anderson, 2001) and spatially uncorrelated observations errors (Houtekamer and Mitchell, 2001). Accounting for spatial observation error correlation in the EAKF is possible and beneficial, but requires a decomposition of the observations into multiple scale components (Ying, 2020).

The first step in the assimilation is the analysis in observation space, and equivalent to a one-dimensional optimal interpolation. The scalar algorithm, as described in Anderson (2003), Karspeck and Anderson (2007), and Anderson and Collins (2007), combines the observation  $y_o$  and the forecast first-guess ensemble mean  $\bar{y}_b$ , under consideration of their error variances,  $\sigma_o^2$  and  $\sigma_b^2$  and results in an analysis PDF that is Gaussian with mean

$$\bar{y}_a = y_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} (y_o - \bar{y}_b) \quad (2.16)$$

and variance

$$\sigma_a^2 = \frac{\sigma_b^2 \sigma_o^2}{\sigma_b^2 + \sigma_o^2}. \quad (2.17)$$

In a second step, the prior ensemble is shifted to its new mean and the ensemble perturbations are adjusted by

$$y_i^a - \bar{y}^a = \alpha (y_i^b - \bar{y}^b) \quad (2.18)$$

with the spread contraction factor  $\alpha = \sqrt{\sigma_o^2 / (\sigma_o^2 + \sigma_b^2)}$ . The posterior PDF with  $\bar{y}_a$  and  $\sigma_a^2$  is optimal.

In the third and last step, the observation increments  $\delta y_i = y_i^a - y_i^b$  for each member  $i$  are mapped into model space by state increments

$$\delta x_{ij} = \frac{\text{Cov}(x_j^b, \mathcal{H}(x_j^b))}{\sigma_{y^b}^2} \delta y_i. \quad (2.19)$$

Note that in a serial implementation, the first-guess for the next observation needs to be updated after each observation. To account for the impact of previously assimilated observations, it is necessary to either apply the observation operator on the updated state (which is inefficient) or approximate the impact using regression (efficient, see Equation 3.13 in Anderson and Collins, 2007).

#### Linearity assumptions

Although the EAKF and variational methods usually use the full observation operator  $\mathcal{H}$  for the innovation, they differ in their assumptions on linearity. Variational methods use a

linearized operator of  $\mathcal{H}$  in the Kalman gain. The EAKF does not use an explicitly linearized operator but computes the posterior in observation space and projects the increments to model space using the covariance of the nonlinear relationship between the model state variable and the model-equivalent (Equation 2.19). The EAKF thus assumes a linear relation between observations and model states in the mapping from observation increments to model space. Violating this assumption obviously introduces errors.

## 2.3 Observation operators

The following section introduces the radiative transfer models that were used to simulate the radiances that would be measured from a satellite. Both models simulate the radiative transfer without 3D effects for each atmospheric column independently.

### 2.3.1 Infrared radiance

Radiances in the infrared spectrum are determined mainly by the emission and absorption of radiation from cloud hydrometeors and greenhouse gases, mostly water vapor. The scattering in clouds is negligible for the short timescales of weather forecasts but would be relevant for climate simulations. In the infrared spectrum, we can detect the existence of clouds, their cloud top temperature, and their height. In clear-sky conditions, it is possible to infer the water vapor content and its gradient.

Infrared radiances for the assimilation were simulated using Chou-scaling (Chou et al., 1999) implemented in RTTOV (Saunders et al., 2010). Chou-scaling is a fast infrared multiple-scattering approximation which solves radiative transfer in the same form as for clear sky conditions, but with a scaling of the scattering optical depth. The scaling approximates the contribution of thermal diffuse scattered radiation. The radiative transfer equation for the infrared can be outlined as follows. An infrared radiance  $I$  passing through a layer of thickness  $dz$ , changes by the amount  $dI$  due to absorption and emission (neglecting scattering),

$$dI = \underbrace{\beta_e B_\lambda(T(z)) dz}_{\text{emission}} - \underbrace{\beta_a dz}_{\text{absorption}}, \quad (2.20)$$

with the coefficient of emission and absorption  $\beta_e = \beta_a$  and the Planck function  $B_\lambda$  which depends on the temperature  $T$ .

By integration, we find the radiance at the top of the atmosphere

$$I_\lambda(\infty) = I(0)t^* + \int_0^\infty B_\lambda(T(z))W(z)dz. \quad (2.21)$$

The weighting function (illustrated in Figure 1.4)

$$W(z) = \frac{dt(z, \infty)}{dz} = \frac{\beta_a(z)}{\mu} t(z, \infty) \quad (2.22)$$

is the derivative of the transmittance  $t$ . The transmittance from the surface to the top of the atmosphere is  $t^*$ . The radiance from the surface  $I(0)$  depends on the surface

temperature. For 6.2 and 7.3  $\mu\text{m}$ , the transmittance from the surface  $t^*$  is zero, leaving only atmospheric contributions from the second term (see Figure 1.4).

### 2.3.2 Visible reflectance

Visible reflectance measures the fraction of the sun's radiation that is reflected to the observer. This thesis investigates the assimilation of visible reflectance observations from a geostationary satellite at a wavelength of 0.6  $\mu\text{m}$ . In the absence of clouds and aerosols, the surface albedo determines the visible reflectance. In cloudy conditions, the 0.6  $\mu\text{m}$  channel is primarily sensitive to liquid water clouds and less sensitive to ice clouds. Contrary to the infrared spectrum, the height of a cloud has nearly no impact on visible reflectance since the cloud-free atmosphere hardly absorbs shortwave radiation. This height independence makes visible reflectance ambiguous regarding the vertical extent of clouds but allows the detection of low-level water clouds and shallow convection, while the infrared 6.2 and 7.3  $\mu\text{m}$  channels can hardly detect them. Contrary to infrared channels, the visible channel is less sensitive to thin ice clouds, which are usually semi-transparent in the visible channel. Thus, the visible channel can help to discriminate between cirrus and high cumulus clouds.

In this thesis, the visible reflectance was calculated using an approximation of the radiative transfer equation based on a look-up-table approach (Frèrebeau, 2014; Kostka et al., 2014; Scheck et al., 2016), since the exact solution is too resource-expensive and slow for data assimilation. MFASIS depends on eight parameters:

- vertically integrated optical depths for water and ice clouds and
- effective particle sizes for water and ice clouds,
- the surface albedo,
- the sun and satellite zenith angles and
- the scattering angle.

The dependence of visible reflectance on the liquid or ice water path is illustrated in Figure 1.5 using certain assumptions for the other parameters, e.g., the effective particle size of 10 and 40  $\mu\text{m}$  for water and ice particles, respectively. MFASIS uses 1D approximation of radiative transfer, which may be relaxed in the future. However, the approximation is less relevant for this thesis, as the operator is assumed to be perfect, i.e. the same operator is used to simulate observations and to generate model-equivalents.

For completeness, the full radiative transfer equation is given here. However, its numerical solution is too costly for operational NWP since it involves 3D cloud-scattering. The differential equation for an optical path length  $ds$  is

$$dI = -\beta_e I ds + \beta_a B(T) ds + \frac{\beta_s}{4\pi} \int_{4\pi} p(\Omega', \Omega) I(\Omega') d\omega ds. \quad (2.23)$$

where we have to integrate over all directions  $\omega$ . Radiation from any direction  $\Omega'$  can contribute to the radiation measured in the direction  $\Omega$ . The scattered direction and the scattering phase function  $p(\Omega', \Omega)$  depend on the size of the scattering particles and the shape of the particles, which can be very different for ice crystals.



# **Chapter 3**

## **Potential impact of assimilating visible and infrared satellite observations**

### **Overview**

The first publication investigates the potential benefit of visible and infrared assimilation (the first research goal in Section 1.4), relative to radar assimilation for two selected cases. It presents the first study that directly compares the assimilation of visible ( $0.6\text{ }\mu\text{m}$ ), infrared ( $6.2$  and  $7.3\text{ }\mu\text{m}$ ), and radar ( $10\text{ cm}$ ) observations. Furthermore, it represents the first application of the idealized Observing-system simulation experiments (OSSE) setup that was developed during the course of this thesis and contains a detailed description of the setup, i.e., the initial conditions and perturbations, the two cases (weather situations), the observing and the assimilation system. Systematic model and operator errors are excluded to focus on the effectiveness of assimilating nonlinear observations in an Ensemble adjustment Kalman filter (EAKF) without distraction by potential systematic model errors and systematic operator deficiencies. Therefore, the achieved absolute forecast improvement needs to be treated with caution, but the relative impact of different observation types can be expected to be largely transferable to real NWP systems as all of the compared observation types are similarly affected by such systematic errors.

### **Own contribution**

Conceptualization, formal analysis, investigation, methodology, software, visualization, writing. The author's contribution to this publication is estimated to be about 70%. In preparation to this work, the experiment setup was developed and tested, which combines the Data Assimilation Research Testbed (DART), the Weather Research and Forecasting (WRF) model and RTTOV observation operators.

### **Publication details**

Lukas Kugler, Jeffrey L. Anderson, Martin Weissmann, 2023: Potential impact of all-sky assimilation of visible and infrared satellite observations compared with radar reflectivity for convective-scale numerical weather prediction. Published in Quarterly Journal of the Royal Meteorological Society, Vol 149, pp 3623-3644. DOI: <https://doi.org/10.1002/qj.4577>.



Received: 15 April 2023 | Revised: 30 July 2023 | Accepted: 11 September 2023  
DOI: 10.1002/qj.4577

Quarterly Journal of the  
Royal Meteorological Society

RESEARCH ARTICLE

# Potential impact of all-sky assimilation of visible and infrared satellite observations compared with radar reflectivity for convective-scale numerical weather prediction

Lukas Kugler<sup>1</sup> | Jeffrey L. Anderson<sup>2</sup> | Martin Weissmann<sup>1</sup>

<sup>1</sup>Institut für Meteorologie und Geophysik,  
Universität Wien, Vienna, Austria

<sup>2</sup>Data Assimilation Research Section,  
Computational Information Systems  
Laboratory, National Center for  
Atmospheric Research, Boulder, Colorado  
USA

Correspondence

L. Kugler, Institut für Meteorologie und  
Geophysik, Universität Wien, Vienna,  
Austria.  
Email: [lukas.kugler@univie.ac.at](mailto:lukas.kugler@univie.ac.at)

## Abstract

Although cloud-affected satellite observations are heavily used for nowcasting applications, their use in regional data assimilation is very limited despite possible benefits for convective-scale forecasts. In this article, we estimate the potential impact of assimilating cloud-affected satellite observations of visible (0.6  $\mu\text{m}$ ) and near thermal infrared wavelengths (6.2 and 7.3  $\mu\text{m}$ ) relative to the impact of assimilating radar reflectivity observations. We employed observing-system simulation experiments with a perfect-model forecast for two cases of strong convective summertime precipitation. Observations are simulated using the radiative transfer model RTTOV/MFASIS and assimilated by the ensemble adjustment Kalman filter in the Data Assimilation Research Testbed. The Weather Research and Forecasting model at 2-km grid resolution was used for forecasts. Results show that satellite observations can be nearly as beneficial as three-dimensional radar reflectivity observations. Under favorable conditions, where the prior contains no error in the stage of storm development but only in horizontal position and strength, the assimilation of visible observations leads to 88% of the radar impact. Under more difficult conditions, the impact of visible and infrared observations still reached 50 and 79%, respectively.

## KEY WORDS

all-sky assimilation, convective-scale data assimilation, data assimilation, DART, infrared radiance, radar reflectivity, visible reflectance, WRF, life cycle

**Abbreviations:** BT, brightness temperature; CAPE, convective available potential energy; CIN, convective inhibition; DART, Data Assimilation Research Testbed; EAKF, ensemble adjustment Kalman filter; FSS, Fractions Skill Score; HRRR, High-Resolution Rapid Refresh; LETKF, local ensemble transform Kalman filter; NOAA, National Oceanic and Atmospheric Administration; NWP, numerical weather prediction; OSSE, observing-system simulation experiment; PBL, planetary boundary layer; RMSE, root-mean-square error; RTTOV, the RT model for the Television and Infrared Observational Satellite; TIROS, Operational Vertical Sounder; WRF, Weather Research and Forecasting model.

This 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.

© 2023 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

## 1 | INTRODUCTION

Clouds are the first area-wide observable signal of convection and are heavily used in nowcasting applications. However, in contrast to nowcasting, the use of cloud-affected satellite observations in data assimilation is very limited (Gustafsson *et al.*, 2018). Clouds are detected earlier by the visible satellite channel than by radars, which are more commonly used in regional data assimilation systems. In contrast to radar, satellite observations are available in most parts of the world, including mountainous or sparsely populated areas, and they provide homogeneous quality across borders (Maddox *et al.*, 2002; Roebeling *et al.*, 2012; Saltikoff *et al.*, 2019; Martinaitis *et al.*, 2020). Additionally, thermal infrared satellite channels observe tropospheric water-vapor content (6.2 and 7.3  $\mu\text{m}$ ) as well as cloud-top temperature (10.8  $\mu\text{m}$ ). Near-infrared channels can distinguish between ice and water clouds (1.6  $\mu\text{m}$ ) and detect nighttime low-level clouds and fog (3.9  $\mu\text{m}$ ). Additionally, visible channels can observe low-level stratus clouds during daytime, which is a major issue for regional weather forecasts (Hu *et al.*, 2022). Thus, there is a large potential for improving weather forecasts by assimilating cloud-affected satellite observations. Nevertheless, current operational regional data assimilation systems largely ignore satellite observations of clouds and thereby omit crucial information on clouds and developing storms.

The assimilation of cloud-affected satellite observations in current assimilation systems is a challenging task. However, most related challenges also apply to the more commonly assimilated radar observations. First, the limited numerical representation of cloud processes and hydrometeors, as well as simplification of observation operators, leads to systematic errors between models and observations (Scheck *et al.*, 2018; Geiss *et al.*, 2021), which violates the basic assumption of an unbiased first-guess forecast of current data assimilation schemes (Gustafsson *et al.*, 2018). These issues were avoided in this study by employing an *observing-system simulation experiment* (OSSE) with an identical model for the nature run and forecasts, as well as perfect observation operators in the forecast. Second, cloud-affected observations violate the assumption of linear observation operators and Gaussian error distributions, as the observations are nonlinear functions of model state variables and their error distributions are often non-Gaussian. In consequence, assimilating these observations violates assumptions of current data assimilation schemes and may lead to a suboptimal analysis in certain conditions.

Most studies on the assimilation of thermal infrared satellite observations focused on wavelengths in the water-vapor sensitive band (5–8.5  $\mu\text{m}$ ), since these wavelengths are less sensitive to surface emission, which

is difficult to model accurately. Otkin (2012a, 2012b) pioneered direct assimilation using four channels between 6.2 and 8.5  $\mu\text{m}$ , albeit at a resolution of 15 km that resolves deep convection only partly. In convection-permitting models, infrared observations had a positive analysis and forecast impact for the prediction of typhoons, mesoscale convective systems, and severe weather events under weak and strong large-scale forcing (Jones *et al.*, 2015, 2016, 2020; Honda *et al.*, 2018; Sawada *et al.*, 2019; Zhu *et al.*, 2022; Eure *et al.*, 2023). Comparing the assimilation of infrared and radar observations, Zhang *et al.* (2019) found that assimilating infrared observations before storm initiation can advance the warning time for mesocyclones by several tens of minutes compared with assimilating radar observations only. Similarly, assimilating visible observations is expected to advance warning times even further. A direct comparison between the impact of visible and infrared satellite channels with that of radar observations has, however, never been conducted so far.

While several studies investigated the assimilation of infrared channels in convective-scale numerical weather prediction (NWP) models, visible channels have received little attention by the research community so far. Since the fast visible forward operator method for fast satellite image simulation (MFASIS, Scheck *et al.*, 2016) was published, only two studies have investigated the benefit of visible observations for convective-scale NWP. Both these studies used the regional NWP system of Deutscher Wetterdienst, which recently also included the assimilation of visible observations in its operational configuration in March 2023. The two studies investigated the impact of visible observations in an idealized and near-operational setup: Schr  ttle *et al.* (2020) conducted an idealized OSSE and found a positive impact by assimilating infrared and visible observations, with the infrared leading to higher impact. Scheck *et al.* (2020) evaluated the impact of only visible observations in a case study with a near-operational assimilation system and found beneficial impact not only on cloud cover but also on temperature, humidity, and precipitation. Given their experimental setup, however, they could not quantify the impact in comparison with other observation types. Additionally, both previous studies used the regional model ICON and a local ensemble transform Kalman filter (LETKF) data assimilation system, whereas the impact of visible observations has not been investigated in any other convection-permitting numerical weather prediction (NWP) system. This motivated the direct comparison of the impact of different observation types in the present study and the use of a different modelling and assimilation system.

Data-denial experiments with operational NWP systems can be misleading in the way they estimate the

analysis impact of a new observation type, as the impact of additional observations may be hampered by systematic model deficiencies without extensive tuning of the assimilation and model physics settings. Additionally, increments from other observation types may conceal the impact of the newly added observations. To avoid this, we assess the forecast impact of each observation type in separate experiments. The separate assimilation of different observation types allows for a detailed analysis of the effects of each type and reveals the potential weaknesses of each one. To put the impact of satellite observations in this setup in the context of more commonly assimilated observations, we additionally conducted comparable experiments that assimilated 2D and 3D radar observations. Furthermore, current operational systems are sub-optimal in many respects: for example, concerning the representation of hydrometeors and related biases as well as the representation of related model errors. The resulting systematic differences between the forecast model and the nature run affect the analysis quality and need to be taken into account when estimating the absolute impact of observations in an OSSE (Errico and Privé, 2018). In this study, we avoid systematic model error to focus on the efficacy of assimilating cloud-related observations in an ensemble Kalman filter and only estimate the impact of observations in relative terms. Thus, we assess the forecast impact in a perfect model OSSE using the identical model configuration for the nature run and forecasts.

Convective-scale data assimilation is a challenging task full of open research questions, as for example, outlined in Hu *et al.* (2022). To gain further insights on the assimilation of additional complex observation types, researchers have studied the assimilation in weather scenarios of increasing complexity for data assimilation: (1) isolated supercells triggered from a “warm bubble” (Snyder and Zhang, 2003; Tong and Xue, 2005), (2) supercells, convective lines, and multicells (Aksoy *et al.*, 2009), and (3) chaotically triggered deep convection (Bachmann *et al.*, 2019, 2020). The third scenario describes deep convection triggered at random locations and is termed the “random” case in this article. It can be considered one of the most difficult and least predictable scenarios, as this case exhibits a high sensitivity to initial conditions and low predictability due to fast error growth and interaction between different cells.

In this study, we evaluate two cases, the less predictable “random” case and the “warm-bubble” case, and estimate the potential impact of assimilating visible and infrared satellite observations, relative to the impact of assimilating radar reflectivity. We employ the ensemble adjustment Kalman filter (EAKF), as it is a commonly used algorithm, which allows for a detailed analysis of the impact of different observations. Specifically, we investigate the following:

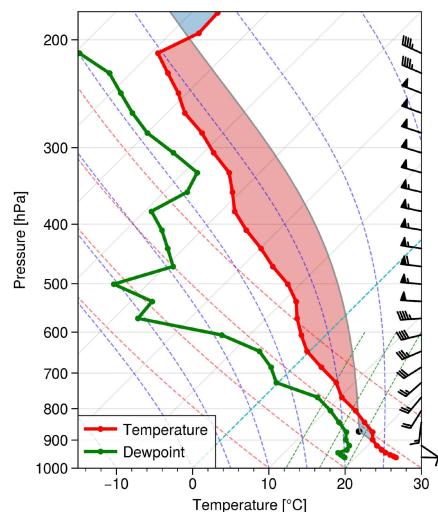
- 1 whether the EAKF is able to extract useful information from visible observations into a convective-scale model;
- 2 the relative impact of visible and infrared observations on precipitation and cloud forecasts relative to the impact of radar observations (Section 3.1);
- 3 the effect of assimilating visible and infrared observations on unobserved state variables (section 3.2);
- 4 why the assimilation of satellite observations is surprisingly beneficial in one case but less beneficial in the other (Section 3.3).

By this investigation, we intend to contribute to a better understanding of the impact of satellite observations, which is crucial for the efficient use of computational, observational, and staff resources (Gustafsson *et al.*, 2018).

## 2 | EXPERIMENTAL DESIGN

### 2.1 | Description of the cases

We estimated the potential impact of cloud-affected satellite observations in two scenarios, one isolated supercell and scattered supercells. Both cases were simulated on the same idealized domain with a homogeneous flat surface and horizontally periodic boundary conditions (see Section 2.2). Both cases share the same base-state



**FIGURE 1** Skew  $T$ -log  $p$  diagram of the nature-run initial condition, from domain average fields. Shown are domain-average profiles of temperature, dewpoint, and a parcel-lifting curve. The temperature perturbation in the warm-bubble case modifies this profile. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

profiles of temperature, water vapor, moisture, and wind illustrated in Figure 1. It is a modified sounding from Payerne, Switzerland on July 30, 2008 and offers a highly unstable environment with  $2670 \text{ J} \cdot \text{kg}^{-1}$  CAPE and  $26 \text{ J} \cdot \text{kg}^{-1}$  CIN at 0700 UTC in the nature run. In this article, time UTC is equivalent to local solar time, since the solar radiation is that of  $0^\circ$  longitude. We start with a description of the nature run. A more detailed description of initial conditions and ensemble perturbations follows in Section 2.3.

### 2.1.1 | Nature run for case “random”

In the “random” case, small random perturbations of temperature and vertical velocity (for details see Section 2.3) trigger storms at random locations scattered throughout the whole domain. The nature run is initialized at 0600 UTC (= local solar time) with the sounding profile depicted in Figure 1. Figure 2 shows the evolution of storms from the perspective of a 7.3- $\mu\text{m}$  infrared satellite image and Figure 3 shows the same in 0.6- $\mu\text{m}$  visible reflectance. At 1100 UTC, 5 hr into the simulation, convection reached altitudes of about 10 km. Shortly after, at 1200 UTC, about 10–15 cells are visible and continue to

grow while others dissipate. The resulting storms are in different stages of their development and interact dynamically, which leads to fast-growing model error and a low predictability of the order of hours. After 1600 UTC, convection decays.

### 2.1.2 | Nature run for case “warm bubble”

In the second case, a positive temperature increment (“warm bubble”) is added to the initial temperature field (see Figure 4). It triggers an isolated and well-organized supercell with  $>200 \text{ m}^2 \cdot \text{s}^{-2}$  updraft helicity (Kain *et al.*, 2008) in a confined region of the domain and suppresses convection elsewhere. Similar warm-bubbles have been used by Snyder and Zhang (2003) and Tong and Xue (2005).

This warm-bubble case was initialized at 1200 UTC with initial conditions as described in Section 2.3. Figures 5 and 6 display the evolution in simulated satellite images of the 7.3- and 0.6- $\mu\text{m}$  channels. Within a few minutes of model integration, deep convection developed in the nature run. At approximately 1235 UTC, the first precipitation developed. After 1730 UTC, the storms reached the domain boundary.

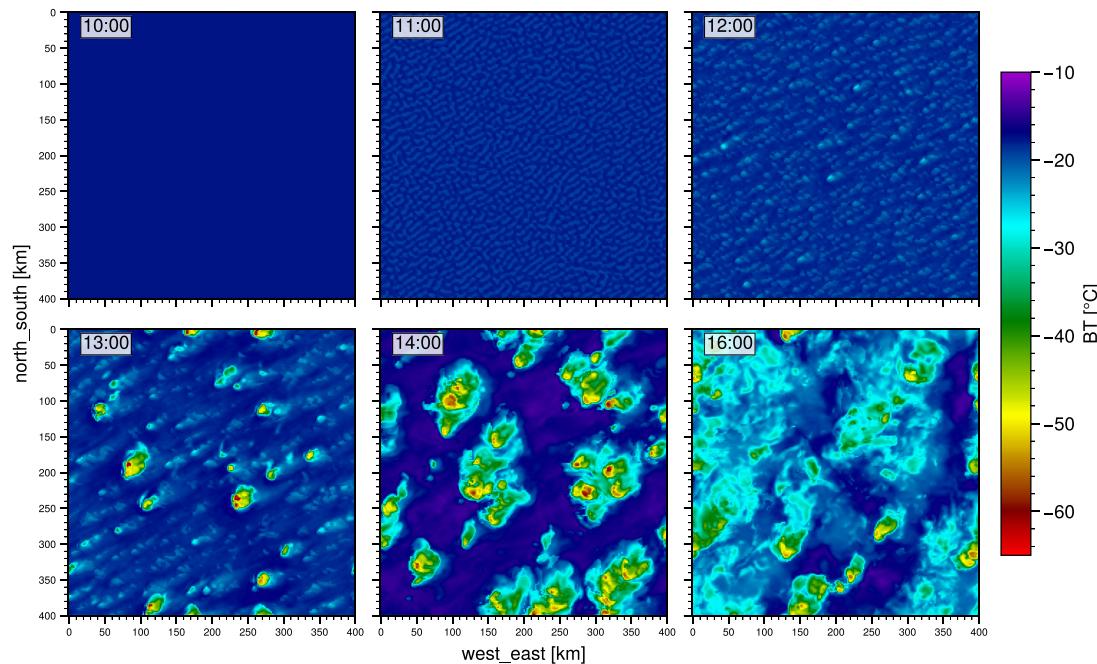


FIGURE 2 Infrared 7.3- $\mu\text{m}$  satellite images of the “random” case nature run. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

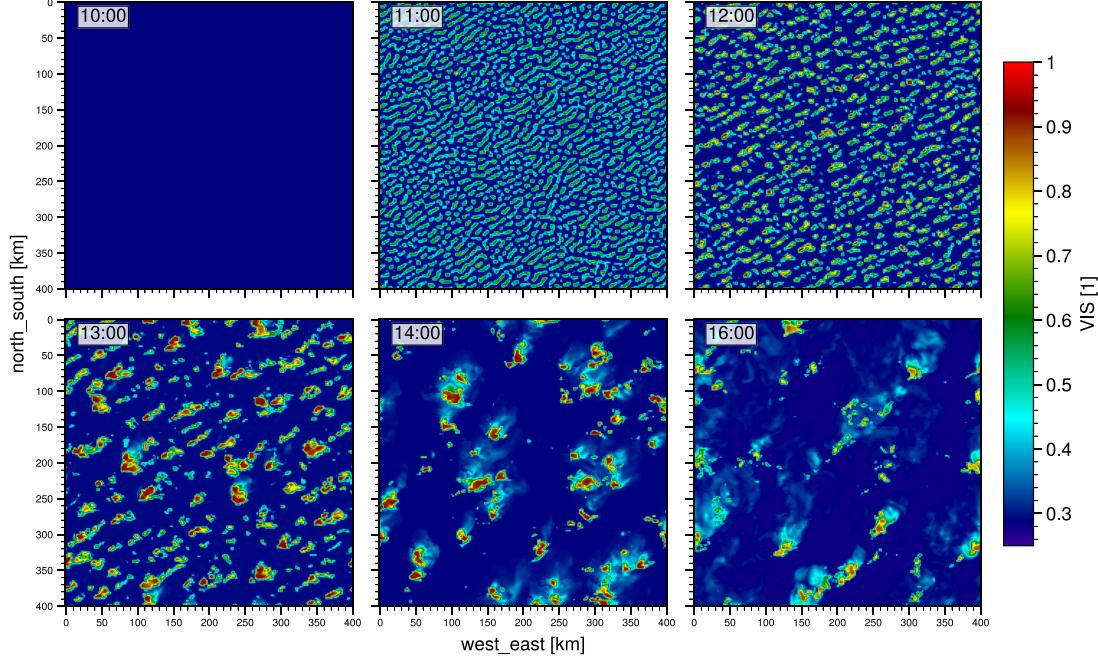


FIGURE 3 Visible 0.6- $\mu\text{m}$  satellite images of the “random” case nature run. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

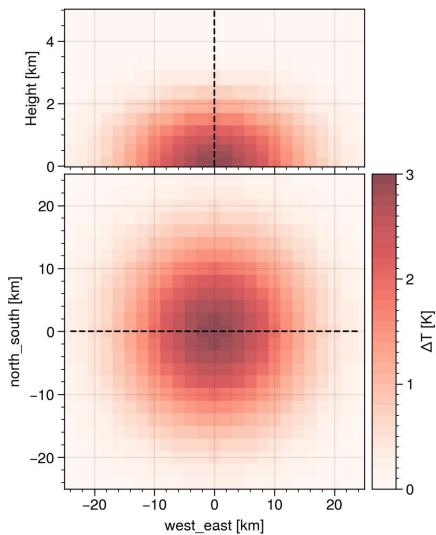
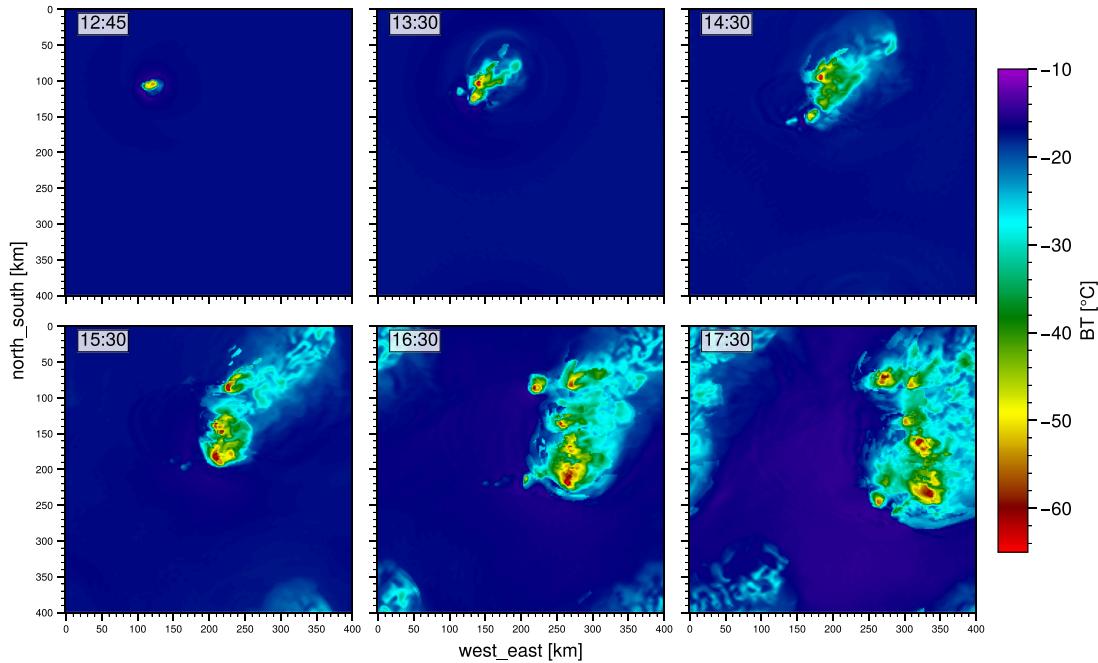


FIGURE 4 Cross-sections through the warm-bubble (temperature perturbation): Vertical–horizontal slice (top) at  $\text{north\_south}=0$  and horizontal–horizontal slice (bottom) at  $\text{Height}=0$ , both marked by a dashed line. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 2.2 | Prediction model

We used the Weather Research and Forecasting model (WRF) version 4.3 (Skamarock *et al.*, 2021) in an idealized mode for both the forecast ensemble and the nature run in identical configurations (*perfect model*). As in preceding studies (Lange and Craig, 2014; Bachmann *et al.*, 2019, 2020; Schröttle *et al.*, 2020), we neglected the Coriolis force, as it does not have a significant effect on the dynamics at this timescale. Also, it would lead to veering of the mean wind given the periodic domain. The physics parametrizations closely follow the configuration of the operational National Oceanic and Atmospheric Administration (NOAA) High-Resolution Rapid Refresh (HRRR) model described in Benjamin *et al.* (2016) with microphysics from Thompson *et al.* (2008); rapid radiative transfer model for general circulation models short- and longwave schemes (Iacono *et al.*, 2008), Mellor-Yamada-Nakanishi-Niino (MYNN 2.5 for the planetary boundary layer (PBL) and surface-layer parametrization from Nakanishi and Niino (2006), Noah land-surface model (Cuenca and Tewari, 2004), and without cumulus parametrization. Default dynamics options were used with an adaptive timestep between 6 and 16 s.



**FIGURE 5** Infrared 7.3- $\mu\text{m}$  satellite images of the warm-bubble case nature run. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 2.2.1 | Domain description

The domain features  $200 \times 200$  mass grid points with a 2-km resolution. The layer depth is 25 m at the surface and increases to about 500 m at 3 km above ground, then staying roughly constant until the model top at about 21.5 km, at staggered level 51. The upper boundary condition is a Rayleigh relaxation layer above 15 km. The surface is homogeneous and flat terrain at 489 m altitude above sea level and of type “Dryland Cropland and Pasture” (`IVGTYPE=2`) with 50% vegetation fraction on soil type “loam” (`ISLTYPE=6`). Solar radiation resembles a summer day (July 30, 2008) at a latitude of 45°N.

## 2.3 | Initial and boundary conditions

### 2.3.1 | Initial conditions of the nature run

The initial conditions of the nature run feature a highly unstable stratification with a CAPE of  $2670 \text{ J} \cdot \text{kg}^{-1}$  and a CIN of  $26 \text{ J} \cdot \text{kg}^{-1}$  at 0700 UTC, such that relatively small perturbations trigger deep convection (Figure 1). The profiles were taken from Lange *et al.* (2017), Bachmann *et al.* (2019), and Schröttle *et al.* (2020), but modified

for humidity and wind. The humidity was clipped to 80% relative humidity, which reduced the humidity in the pressure intervals 900–750 hPa and 350–200 hPa, in order to avoid stratiform clouds in the whole domain. The wind shear was increased considerably to bulk shears of 14, 38, and 61 knots in the layers 0–1, 0–3, and 0–6 km, respectively, to support long-lived supercells.

In the “warm-bubble” case, an additional temperature increment  $\Delta T(x, y, z)$  was added to the initial temperature field, defined as

$$\Delta T = A \exp \left[ -\left( \frac{r(x, y)}{c_r} \right)^2 \right] \exp \left[ -\left( \frac{z}{c_z} \right)^2 \right], \quad (1)$$

where  $r(x, y) = \sqrt{(x - x_c)^2 + (y - y_c)^2}$  is the distance from the bubble center;  $A$  is the maximum perturbation (3 K), the tuple  $(x_c, y_c)$  is the center of the bubble,  $c_r$  is the horizontal decay (15 km), and  $c_z$  is the vertical decay (2 km).

### 2.3.2 | Initial perturbations in the forecast ensemble

While a real-data experiment comes with a prior forecast ensemble, we have to make a realistic guess about

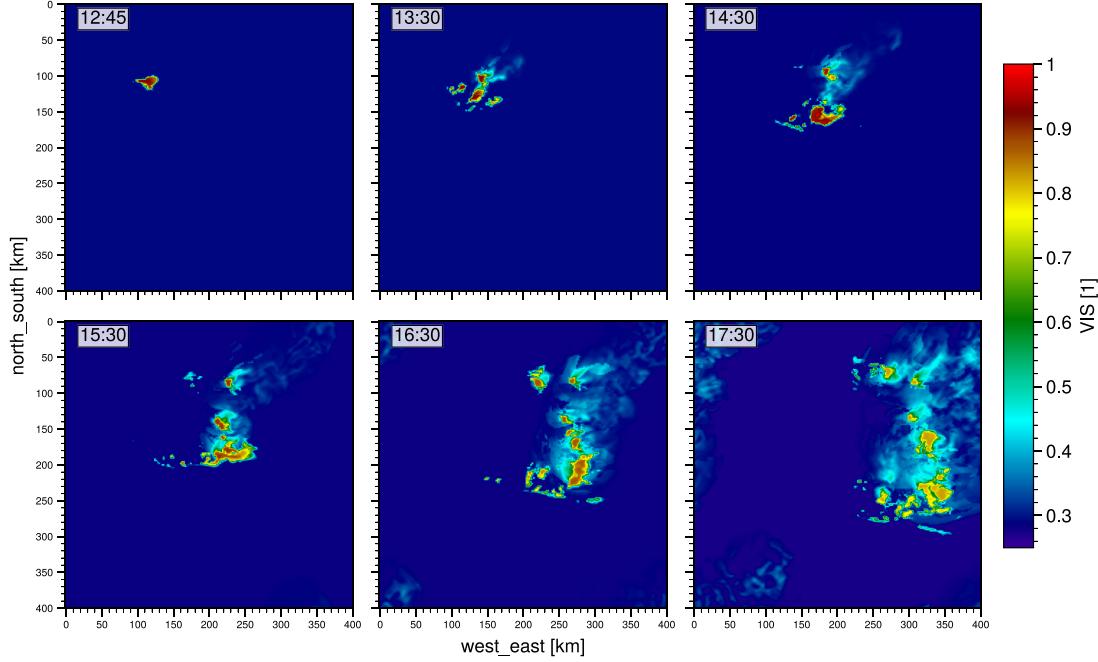


FIGURE 6 Visible 0.6- $\mu\text{m}$  satellite images of the warm-bubble case nature run. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

the prior uncertainty for this OSSE. To be consistent with prior publications, we use the approach from Schröttle *et al.* (2020) that facilitates two kinds of perturbations: (1) Vertically auto-correlated profile perturbations representing large-scale errors and (2) small-scale boundary layer noise.

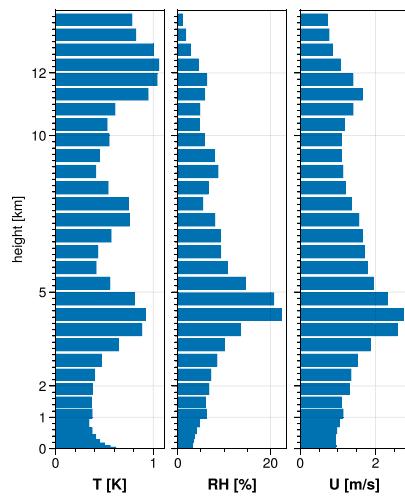
- (1) Before initializing the ensemble forecast, we perturb the vertical profile of temperature, moisture, and wind. This *inter alia* leads to modified convective stability, which delays or accelerates the evolution of deep convection. The perturbations are created by choosing one random number for every 20th vertical level of the original 200-level profile and then interpolating between them, such that we end up with a vertically auto-correlated profile of random perturbations for every ensemble member. The random numbers are created using a standard deviation of 0.25 K for temperature and 2% for relative humidity and wind. The resulting profiles are used as input profiles for the WRF initialization program (`ideal.exe`), which modifies the profiles slightly for hydrostatic balancing.
- (2) Small-scale random noise was added to the temperature and vertical velocity in the lowest levels to avoid unrealistic spatially homogeneous fields. The perturbations are relaxed toward zero with height:

for temperature following  $x \exp((p - p_{\text{sfc}})/25)$  with  $p$  in hPa, for velocity following  $x \exp((1 - k)/2)$  where  $k$  is the level number, where  $x$  was drawn from a Gaussian with  $\sigma = 0.02$  (K or  $\text{m}\cdot\text{s}^{-1}$ ) for each column. Although the horizontal variation of temperature was this small at the initialization time, the perturbations grew considerably in the following 6 hr, reaching a spread of 1 K in temperature, 20% in relative humidity, and  $2 \text{ m}\cdot\text{s}^{-1}$  in zonal wind (random case, Figure 7). In the warm-bubble case, the horizontal average spread was smaller due to the shorter spin-up time of 0.5 hr and the small fraction of the domain in which convection took place, reaching a spread of 0.5 K in temperature, 5% in relative humidity, and  $1 \text{ m}\cdot\text{s}^{-1}$  in zonal wind at 1300 UTC (not shown).

### 2.3.3 | Additional perturbations in the “warm-bubble” case

In the “warm-bubble” case, we imposed another uncertainty in two parameters (see Equation 1):

- the horizontal location of the warm bubble by perturbing the center  $(x_c, y_c)$  in the north/south and east/west direction by  $\pm 60 \text{ km}$  (uniformly random) and



**FIGURE 7** Vertical profile of ensemble spread (horizontal average) of temperature, relative humidity, and the  $u$ -wind component at 1300 prior to the first assimilation in the “random” case. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

- the spatial extent and strength by perturbing the amplitude  $A$  by  $\pm 1$  K (uniformly random).

## 2.4 | Simulated observations

Four types of observations have been used in this study, as listed below.

- (1) Satellite observations of visible reflectance at a wavelength of 0.6  $\mu\text{m}$  reveal how much sunlight is reflected by clouds or the surface. In contrast to radar reflectivity, the observations already provide information on clouds in their early stage, right after rising plumes reach the condensation level. Reflectance describes the ratio of reflected radiance to the total incoming irradiance and is therefore a dimensionless value in the range of 0–1. The lowest possible value in practice is, however, determined by the surface albedo, which is around 0.27 in our setup. The instrument error for the visible channel was chosen to be 3% following Schröttle *et al.* (2020).
- (2) Satellite observations of infrared brightness temperature of the 6.2- $\mu\text{m}$  channel (*Meteosat Second Generation (MSG) 4 Spinning Enhanced Visible and Infrared Imager (SEVIRI)* channel 5, *Geostationary Operational Environmental Satellites (GOES)* Advance ABI, and *Advanced Himawari Imager (AHI)* band 8) are specifically sensitive to upper tropospheric water vapor and clouds. For clouds, the observations mainly provide information on the cloud-top height, as can be seen by lower brightness temperatures for higher

cloud tops. The instrument error was chosen to be 1 K. In contrast, Cintineo *et al.* (2016) did not simulate instrument errors for brightness temperature (BT) observations.

- (3) Satellite observations of infrared brightness temperature of the 7.3- $\mu\text{m}$  channel (*MSG-4 SEVIRI* channel 6, *GOES* ABI, and *Himawari AHI* band 10) provide similar information to the 6.2- $\mu\text{m}$  channel but are more sensitive to lower tropospheric water vapor. An instrument error of 1 K was selected.
- (4) Finally, three-dimensional radar reflectivity (10 cm) serves as a reference observation type to for evaluating the impact of satellite observations. An instrument error of 2.5 dBz was chosen, half the error of Wheatley *et al.* (2015) and Bachmann *et al.* (2020), who used 5 dBz.

Observations  $y^0$  were generated using the Data Assimilation Research Testbed (DART) provided by UCAR/NCAR/CISL/DAReS (2022). It interpolates the nature run’s state  $x^{\text{nat}}$  to each observation’s location and applies an observation operator  $\mathcal{H}$  to the state  $x^{\text{nat}}$  before adding Gaussian instrument error:

$$y^0 = \mathcal{H}(x^{\text{nat}}) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_0). \quad (2)$$

The resolution of satellite observations was effectively grid-scale (2 km). The model equivalents of observations,  $y^b = \mathcal{H}(x^b)$ , were generated using the same observation operators to avoid systematic errors between forecast and observations.

Satellite observations were simulated using the default Chou scaling for the infrared channels (Chou *et al.*, 1999) and MFASIS Scheck *et al.*, (2016) for the visible channel, as provided in the RT model for the television and infrared observational satellite (TIROS) operational vertical sounder (RTTOV) v13.3 (Saunders *et al.*, 2018). For radar reflectivity, the operator included in the WRF Thompson microphysics module was used. The surface albedo and emissivities are given by RTTOV default values. To simulate cloudy radiances, we assumed an effective particle diameter of 20  $\mu\text{m}$  for water droplets and 60  $\mu\text{m}$  for ice crystals. For the satellite geometry, we assumed a geostationary satellite at the Equator with an azimuth of 180° and zenith of 45°. The solar angles were computed using the `pysolar` module assuming a latitude of 45° and longitude 0°.

## 2.5 | Assimilation system and settings

Our experiments applied the EAKF by Anderson (2001) included in DART Anderson *et al.* (2009)<sup>1</sup> to a 40-member

<sup>1</sup><https://dart.ucar.edu>

forecast ensemble. The EAKF is a serial deterministic square-root filter, which assimilates one observation after another. The following variables were updated: temperature, water-vapor mixing ratio, dry air mass in column, geopotential, wind components  $U, V, W$ , cloud water, and ice mixing ratio.

Posterior covariance inflation was applied, since experiments without inflation indicated that analysis ensemble spread would have been underestimated. Specifically, relaxation to prior spread (RTPS) with factor  $\alpha = 0.9$  was used to inflate ensemble perturbations. Note that a value of  $\alpha = 1$  would prevent any variance reduction and restrict updates to updates of the mean, while  $\alpha = 0$  would mean no inflation. We localized covariances in the horizontal to 20 km half-width of the Gaspari–Cohn function. Radar observations were localized to 3 km in the vertical. Satellite radiances were not localized vertically. Lastly, a sampling error correction (Anderson, 2012) was applied. We assimilated all observations, including those that were far from the first guess, as we noticed a relatively strong error reduction by these observations in our experiments. It should be noted, however, that real NWP systems might require such a first-guess check to exclude erroneous observations.

The horizontal distribution of observations was chosen to be equal for all observation types. In the horizontal, we assimilated observations every 10 km. However, we did not assimilate observations within 50 km of the domain boundary, so that only observations of the inner  $300 \times 300$  km were assimilated. This was necessary to avoid discontinuous increments at domain boundaries, since we assumed a periodic WRF domain but a limited area domain in DART. In the vertical, we assimilated radar reflectivity observations every 2 km from 2–14 km.

Superrobbing can be a useful approach to assimilate high-resolution observations, as it averages observations towards the resolved scale of the model. However, an experiment that superrobbed  $5 \times 5$  grid-scale observations towards one observation every 10 km did not generally improve forecasts. Given the perfect-model assumption, this might not be too surprising. Furthermore, we only

superrobbed observations, but not the model prior following the standard implementation in DART, which is not fully consistent. As the difference in impact was negligible, we decided not to include those experiments in this article.

The assimilation of satellite observations in a Gaussian filter is suboptimal for reasons of non-Gaussianity, like heteroscedasticity (the increase of variance with cloudiness) or boundedness. Additionally, non-linear observations operators as well as sampling error and suboptimal ensemble perturbations lead to a suboptimal analysis and ensemble spread. These effects can be mitigated by assigning inflated observation errors (Geer and Bauer, 2011), but the optimal choice of assigned observation errors often needs to be tested by sensitivity studies (see Section 3.3).

## 2.6 | Assimilation experiments

The experiments of this study are listed in Table 1. To investigate optimal assigned observation-error settings, we conducted sensitivity experiments with different assigned observation errors in Section 3.3. The resulting optimal observation errors used for the standard experiments are listed in Table 1.

The timeline of the experiments is illustrated in Figure 8. In the random case, the forecast ensemble was initialized at 0700 UTC and ran freely without assimilation for 6 hr. By 1300 UTC, the model had generated a sufficient amount of spread (Figure 7). From 1300–1400 UTC we assimilated five times (every 15 min), followed by free forecasts until 1800 UTC.

In the warm-bubble case, we started to assimilate at 1230 UTC after a free forecast of 30 min. Despite this short spin-up time, the deep convection had already developed. From 1230–1330 UTC we assimilated five times, followed by free forecasts until 1800 UTC. Although the assimilation window is 1 hr in both cases, it covers different phases of convection in each case.

Figure 9 shows the time series of the strongest cloud signal in each observation type, that is, the lowest value

**TABLE 1** Experiments and their assimilated variable together with standard errors for generating observations (instrument error) and assimilating observations (assigned error).

Abbreviation	Assimilated observation type	$\sigma_{\text{generate}}$	$\sigma_{\text{assimilate}} (\text{range of tested values})$
VIS	Visible reflectance 0.6 $\mu\text{m}$	0.03	0.03 (0.03–0.12)
WV62	Brightness temperature 6.2 $\mu\text{m}$	1 K	1 K (1–3)
WV73	Brightness temperature 7.3 $\mu\text{m}$	1 K	1 K (1–3)
REFL	Radar reflectivity 10 cm	2.5 dBz	2.5 dBz (2.5–7.5)
NoDA	None	–	–

*Note:* The range of tested values for the assigned error is indicated in parentheses.

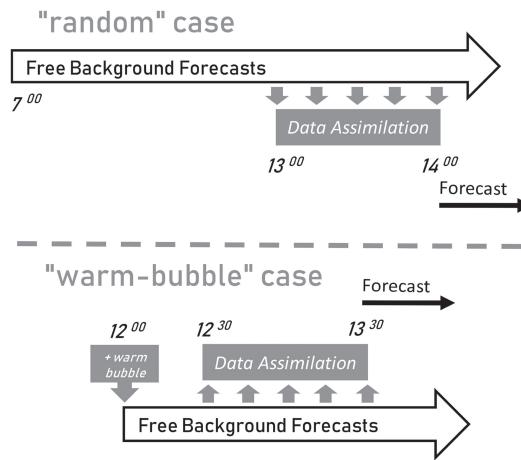


FIGURE 8 Timeline of forecasts and assimilation in the “random” and “warm-bubble” cases.

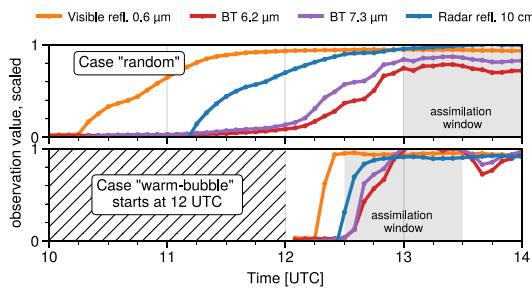


FIGURE 9 Time series of the strongest cloud signal in each observation type, defined as the average of the 40 (0.1%) largest values of reflectance and lowest values of BT. Values are scaled to the range 0–1 from their respective ranges: 0.27–1 for visible reflectance; 235–205 K for 6.2-µm BT; 255–205 K for 7.3-µm BT; 15–70 dBz for radar reflectivity. The assimilation time frame is shown as a grey background. [Colour figure can be viewed at wileyonlinelibrary.com]

for infrared BT and the highest value for visible and radar observations. The earliest stages of convection were only detected by visible observations. For radar, it took up to 60 min for convection to become apparent in the observations. In the “random” case, all observation types detected convection at the start of the assimilation window. In the warm-bubble case, however, infrared channels did not detect convection at the beginning of the assimilation, but later in the assimilation window. Overall, the warm-bubble case was more predictable. A measure of uncertainty is the time duration between earliest and latest convective initiation in the ensemble. While the time difference of

convective initiation was 1.5 hr in the “random” case, the initiation happened within 20 min in the “warm-bubble” case (not shown). This demonstrates that adding a warm bubble can act to synchronize the triggering time of convection across the ensemble, since it forces convection regardless of the stratification.

### 3 | RESULTS

The first goal in this section is to estimate how forecasts of precipitation and cloudiness benefit from assimilating cloud-affected satellite observations (Section 3.1). Subsequently, we analyze the impact on vertical profiles of state variables in Section 3.2. Lastly, we try to explain the larger impact of 3D radar observations compared with 2D satellite observations in the case of random convection in Section 3.3.

#### 3.1 | Relative potential impact

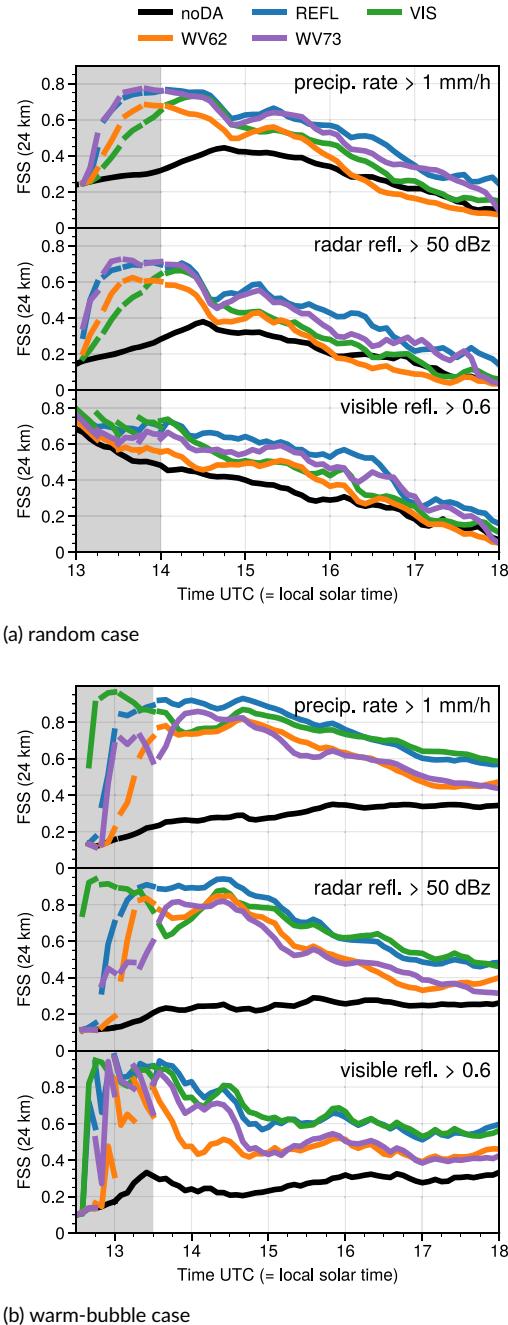
We evaluate forecasts using the Fractions Skill Score (FSS) for three quantities:

- precipitation rate  $> 1 \text{ mm}\cdot\text{hr}^{-1}$
- radar reflectivity  $> 50 \text{ dBz}$ ,
- visible reflectance  $> 0.6$ ,

The 24-km window FSS of these quantities describes how well a forecast was able to pinpoint the location of precipitation and optically thick clouds. We calculated the FSS using neighborhood ensemble probabilities after Schwartz *et al.* (2021), in contrast to, for example, Scheck *et al.* (2020), who calculated the FSS from the ensemble mean.

##### 3.1.1 | Case “random”

Figure 10a shows the impact of assimilating four different observation types in the case with deep convection scattered randomly throughout the whole domain. Compared with the REFL experiment and averaged over 1400–1700 UTC, the VIS experiment revealed an FSS improvement of 50% compared with the FSS of noDA, the WV73 experiment 79%, and the WV62 experiment 20% for the prediction of radar reflectivity  $> 50 \text{ dBz}$ . Within the first forecast hour, the VIS experiment performed nearly as well as the REFL experiment but lost impact thereafter. The WV73 experiment showed similar skill to the VIS experiment in the first 1.5 hr lead time, but provided



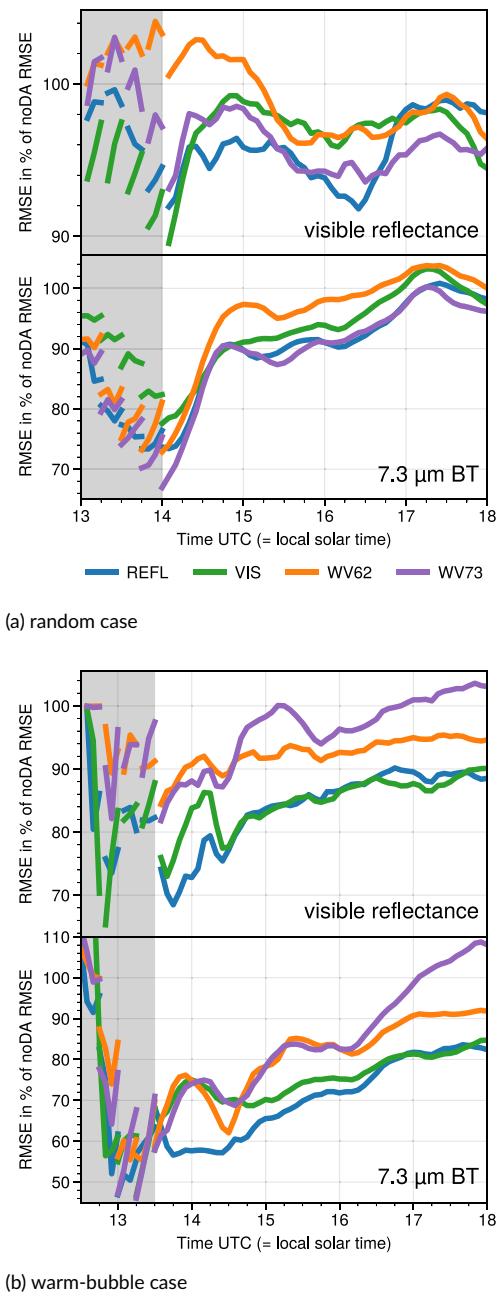
**FIGURE 10** Fraction skill score for precipitation  $> 1 \text{ mm} \cdot \text{hr}^{-1}$ ; radar reflectivity  $> 50 \text{ dBz}$  and visible reflectance  $> 0.6$  in (a) the “random” case and (b) the warm-bubble case. The assimilation time frame is marked by grey shading. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

better forecasts afterwards. The WV62 experiment’s forecast skill was the lowest of all observation types. It seems that channels that see deeper into the atmosphere (visible and  $7.3 \mu\text{m}$ ) have a higher impact than the  $6.2\text{-}\mu\text{m}$  channel, which does not sense lower tropospheric vapor and clouds. Overall, forecasts in the REFL experiment were best, with 2.5 hr of skillful forecasts for light precipitation and 1.5 hr for strong precipitation, except for the prediction of visible reflectance  $> 0.6$ , where forecasts of the VIS experiment were slightly better.

In Figure 11a, we show the root-mean-square error (RMSE) of visible reflectance and  $7.3\text{-}\mu\text{m}$  brightness temperature forecasts, relative to the RMSE of the noDA experiment. Specifically, we computed the RMSE of the ensemble mean forecast over all  $200 \times 200$  grid points, while only  $31 \times 31$  satellite observations were assimilated. At analysis time, the experiment that assimilated visible reflectance had the lowest errors in visible reflectance, as expected. The same applies to the WV73 experiment and the verification of  $7.3\text{-}\mu\text{m}$  BT. After the analysis, however, the RMSE of the WV73 experiment was similar to the error of the REFL experiment. The experiments REFL, VIS, and WV73 overall showed similar skill in predicting the visible channel, while the WV62 experiment had lower skill in the first 1.5 hr. The VIS experiment had relatively good forecasts of  $7.3\text{-}\mu\text{m}$  BT and the WV73 experiment had good forecasts of visible reflectance. The WV62 experiment had less accurate forecasts of both  $7.3\text{-}\mu\text{m}$  BT and visible reflectance, which is presumably related to the higher peak of its weighting function leading to smaller sensitivity to low and mid-level clouds.

### 3.1.2 | Case “warm bubble”

Figure 10b shows the forecast impact in terms of FSS, but now for the warm-bubble case. In general, all observation types lead to a significant FSS improvement compared with the noDA experiment, but some aspects should be noted. First, the assimilation of visible reflectance in the VIS experiment improved the FSS faster than the assimilation of infrared BT in the experiments WV62 and WV73. As visible reflectance detected convection at an early stage (Figure 9), the VIS experiment was at a clear advantage. The initially high impact in the VIS experiment deteriorated in the first forecast hour, handing over the lead to the REFL experiment. However, the VIS experiment overtook the REFL experiment again at around 3 hr lead time in precipitation scores. Second, the experiments WV62 and WV73 produced similar results, except for the FSS of cloudiness (visible reflectance  $> 0.6$ ), where most of the impact vanished within 30 min of free forecast in the WV62 experiment. Note that the  $6.2\text{-}\mu\text{m}$  channel is more



**FIGURE 11** RMSE of ensemble mean forecasts of the visible (upper panel) and 7.3-µm channel (lower panel), normalized by the RMSE of the noDA control run for (a) the “random” case and (b) the warm-bubble case; the average is taken horizontally over  $200 \times 200$  grid points. The assimilation time frame is marked by grey shading. [Colour figure can be viewed at wileyonlinelibrary.com]

sensitive to higher tropospheric water vapor, while the 7.3-µm channel is more sensitive to lower tropospheric water vapor. Third, the experiments REFL and VIS show a similar performance except for the first hour, where the skill was slightly lower for precipitation. Interestingly, the VIS experiment only shows an advantage over the REFL experiment in the FSS for cloudiness during the first 2.5 hr. For cloudiness, both the REFL and the VIS experiment gave similar performance. Compared with the REFL experiment, the experiments WV62 and WV73 showed less impact. Lastly, the REFL experiment outperformed all other observation types in the first forecast hour for light and strong precipitation but only slightly for cloudiness, where the VIS experiment was best most of the time.

Figure 11b shows the RMSE of forecasts of visible reflectance and 7.3-µm BT for the warm-bubble case. Visible reflectance was best forecast by the REFL experiment, followed by the VIS experiment with similar forecast score, except for the first forecast hour. The experiments WV62 and WV73 performed worse, as they removed less error until the last assimilation time. While the experiments REFL and VIS removed up to 30% of error, the experiments WV62 and WV73 removed only 15–20% of visible reflectance error. Also 7.3-µm BT was best forecast by the REFL experiment, removing 40% of RMSE until the last assimilation time. Other experiments removed similar amounts of error, but lost impact faster. On average, the VIS experiment had the second best RMSE in 7.3-µm BT, followed by WV62 and WV73.

### 3.1.3 | Comparison of cases

A major difference between the two cases is that the warm-bubble case is more predictable than the “random” case. While the REFL experiment skillfully predicted strong precipitation for nearly 4 hr ( $FSS > 0.5$ ) in the warm-bubble case, the random case was skillfully predicted for only 1.5 hr. The difficult forecasting conditions probably result from faster growth of errors in the “random” case, as storms interact with each other and continuously trigger new cells, leading to a chaotic environment which is very sensitive to the initial conditions.

To compare the relative impact of the observation types, Table 2 shows the relative FSS improvement over noDA of each experiment compared with the REFL experiment. Overall, satellite observations lead to a remarkable impact given that the satellite experiments assimilated only 1/7th of the number of observations compared with the REFL experiment. Table 2 also demonstrates that satellite and especially visible observations can be effectively used by the ensemble adjustment Kalman filter and lead to long-lasting forecast impact.

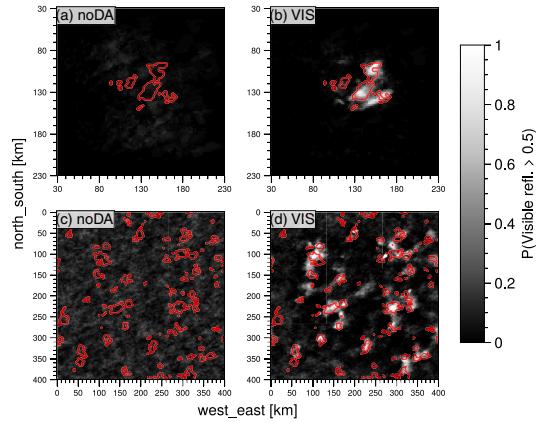
**TABLE 2** Fraction of each experiment's FSS improvement over noDA, relative to the REFL experiment in the respective case, that is,  $(FSS_{\text{exp}} - FSS_{\text{noDA}})/(FSS_{\text{REFL}} - FSS_{\text{noDA}})$  for the event of reflectivity  $> 50 \text{ dBz}$ , where FSS is averaged over the first three forecast hours.

Experiment	Case “random”	Case “warm-bubble”
REFL	100%	100%
REFL-2D	64%	-
VIS	50%	88%
WV73	79%	74%
WV62	20%	76%

Visible observations detect convection earliest and allow the filter to narrow down the location of convection much earlier than other observations can. To date, only Schr ttle *et al.* (2020) compared the assimilation of visible ( $0.6 \mu\text{m}$ ) and infrared observations ( $6.2 \mu\text{m}$ ) in a convective-scale NWP model. Despite the similar setup, we clearly see more impact from assimilating visible observations than from  $6.2\text{-}\mu\text{m}$  BT observations. This contrasts with Schr ttle *et al.* (2020) who found stronger impact from the  $6.2\text{-}\mu\text{m}$  channel and less impact from visible observations. However, this difference might be related to overly inflated observation errors for visible observations in that study, as they inflated the observation error for visible much more than for infrared observations, which presumably led to a lower weight for visible observations.

Although the WV62 experiment showed competitive forecasts of precipitation in the warm-bubble case, it performed poorly in forecasting cloudiness in both cases and precipitation in the “random” case (Figure 10). This might be due to the higher peak of the weighting function of the  $6.2\text{-}\mu\text{m}$  channel compared with the  $7.3\text{-}\mu\text{m}$  channel.

In the warm-bubble case, the uncertainty lies mostly in the warm-bubble location and strength. As the evaluation showed, these can be easily derived from satellite observations. In the “random” case, however, visible and  $6.2\text{-}\mu\text{m}$  BT observations lead to substantially less impact. We hypothesize that a possible explanation might be missing the vertically resolved information from radar observations. This hypothesis is investigated further in Section 3.3.1 by assimilating two-dimensional instead of three-dimensional radar reflectivity. Except for the higher impact of radar in the warm-bubble case, however, the experiments overall reveal the value of satellite and particularly visible observations, especially in scenarios with an uncertain location of convection. Figure 12 illustrates how assimilating visible reflectance improved the forecast of the location of clouds in the ensemble.



**FIGURE 12** The probability for visible reflectance  $> 0.5$ , (a,c) in the noDA run and (b,d) after the assimilation in the VIS experiment. The ensemble-derived probability ranges from black (0) to white (1) and the nature  $(H(x_{\text{nat}}) > 0.5)$  is shown in red contours. The warm-bubble case at 1335 is shown in (a,b). The “random” case at 1405 is shown in (c,d). [Colour figure can be viewed at wileyonlinelibrary.com]

### 3.2 | Impact on model state variables

In the OSSE framework, we can compare how different observation types impact the prior model state as the true state is perfectly known. To do that, we analyzed the vertical structure of the increments (Figure 13) and error/error reduction (Figure 14).

Figure 13 shows the vertical profiles of the absolute ensemble-mean increments, averaged spatially and over the five assimilation times. Panels 13a and 13b show the increments of the warm-bubble case and the random case, respectively. In the warm-bubble case, the temperature increments of experiments VIS and REFL are substantially lower than those of WV62 and WV73. Interestingly, WV73 shows larger increments than WV62 below 7 km, while above this the relation is reversed. In the random case, WV62 and WV73 produced very similar temperature increments. Concerning water vapor, the experiments WV62 and WV73 lead to higher increments than VIS and REFL in the warm-bubble case, but not in the random case, where most profiles are similar. The largest increments in vapor can be seen in the WV73 experiment and the lowest increments in the VIS experiment. Regarding the increments of cloud water, the VIS experiment again revealed the smallest increments in both cases, while the other experiments do not differ substantially in increment magnitude. Cloud ice increments are also similar among the experiments except for the VIS experiment. Wind increments are strongest for WV62 and WV73 and smallest

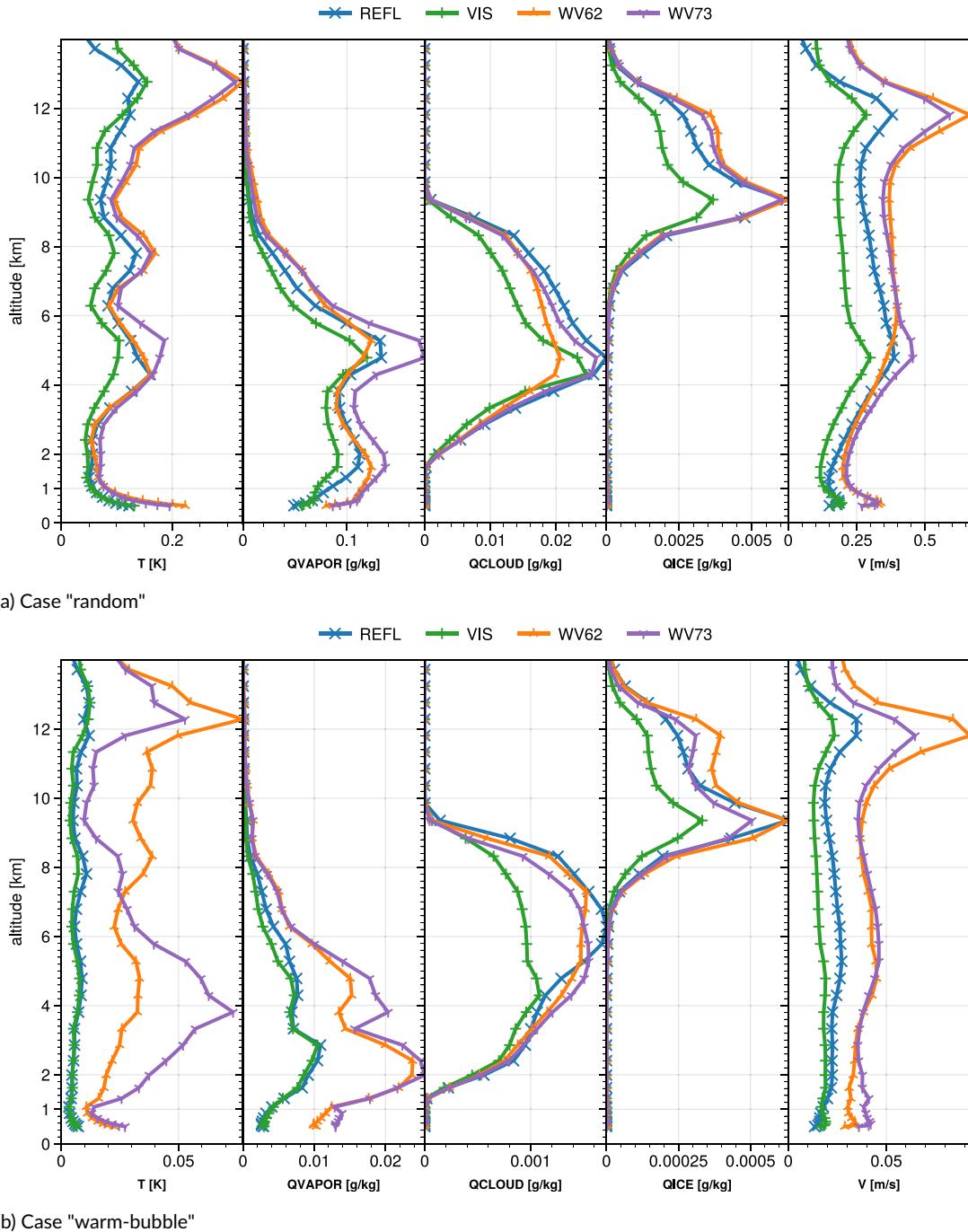
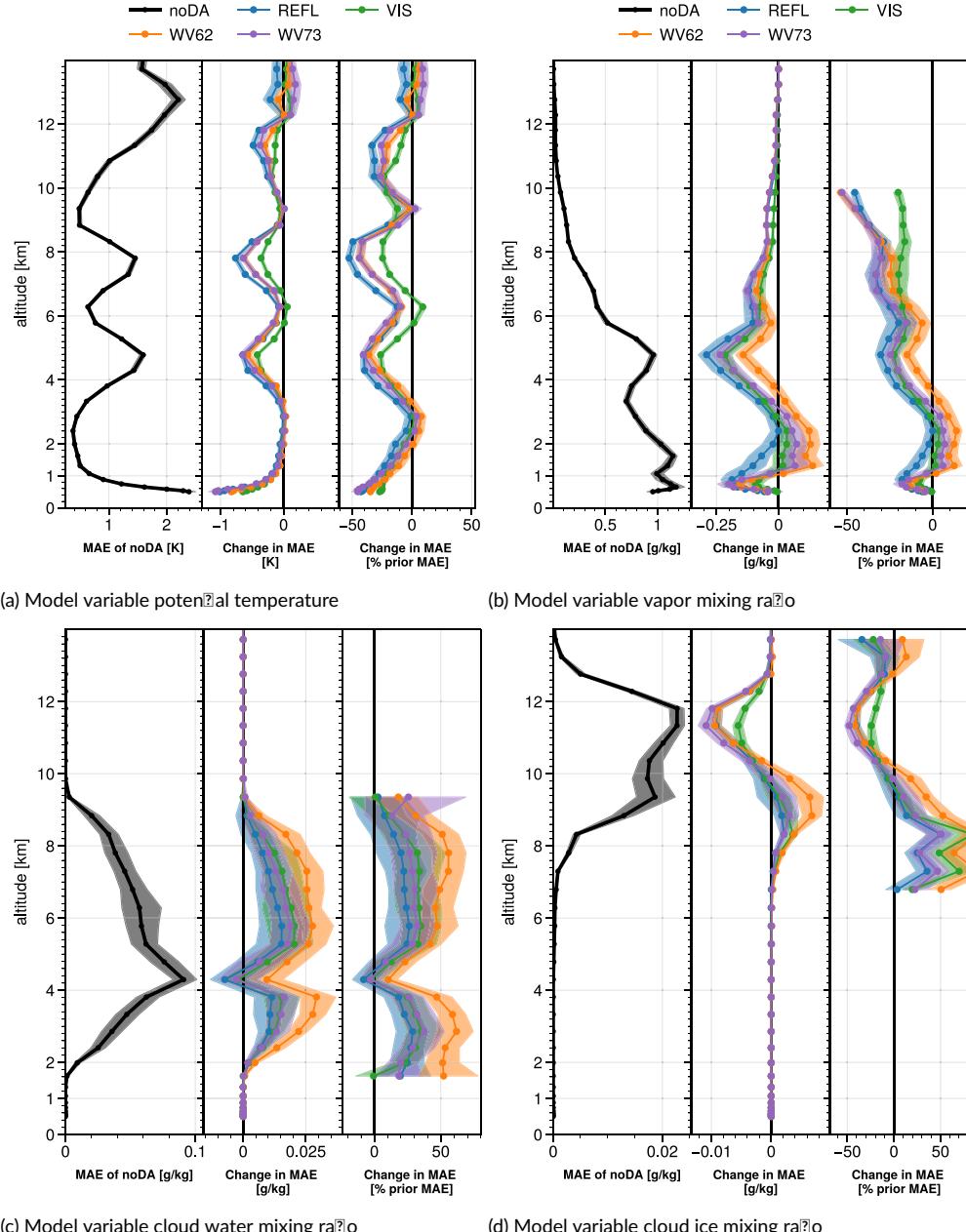


FIGURE 13 Vertical profiles of the absolute ensemble-mean increments for model variables temperature ( $T$ ), vapor mixing ratio ( $QVAPOR$ ), cloud water and ice mixing ratio ( $QCLOUD$  and  $QICE$ ), and wind speed ( $V$ ); averaged horizontally and over five assimilation times. [Colour figure can be viewed at wileyonlinelibrary.com]



**FIGURE 14** Vertical profiles for model variables temperature (top left), vapor mixing ratio (top right), cloud water, and ice mixing ratio (bottom left and right). Each panel shows the MAE of the noDA experiment (left), the MAE reduction in the experiment (center), and the relative MAE reduction in % of the noDA MAE (right). Negative values stand for lower errors in the assimilation experiments compared with noDA. The right panel shows the change in MAE, relative to the prior MAE. The error was evaluated at 1405 in the "random" case, as mean (over 961 observed atmospheric columns) absolute error of the ensemble mean forecast. Dots indicate the horizontal average, shading indicates the 95% confidence interval over 961 atmospheric columns in which observations were taken. The increments of neighboring observations were overlapping and thus not independent. [Colour figure can be viewed at wileyonlinelibrary.com]

for VIS. Generally, the increment profiles are more similar in the “random” case, while in the “warm-bubble” case, the increments of the experiments VIS and REFL differ from WV62 and WV73, especially for temperature, vapor, and wind. Furthermore, we see that the visible reflectance observations, while less sensitive to ice clouds, still lead to increments in cloud ice. Lastly, note that the magnitude of increments in the “warm-bubble” case is lower due to the large clear-sky area, where hydrometeor-related observations contain little information.

Figure 14 shows vertical profiles for temperature (top left), vapor mixing ratio (top right), cloud water, and ice mixing ratio (bottom left and right). Each panel shows the mean absolute error (MAE) of the noDA experiment (left), the MAE reduction in the experiment (center), and the relative MAE reduction in % of the noDA MAE (right). The error was evaluated at 1405, five minutes after the last assimilation in the “random” case, as mean (over 961 observed atmospheric columns) absolute error of the ensemble mean forecast. The corresponding increments can be seen in Figure 13a.

The temperature error profile (Figure 14a) shows four peaks: at the surface, 5, 8, and 13 km. The error reduction was largest in these layers in absolute and relative terms. The experiments WV62 and WV73 removed nearly as much temperature error as the REFL experiment, reaching up to 0.5 K and 40% of error. The VIS experiment also reduced the errors, but reaches only 0.2 K and 25%.

Regarding water vapor (Figure 14b), the relative error reduction was largest at altitudes with low vapor concentration, reaching 40% at 7 km but still removing 20% in the boundary layer. The experiments VIS and WV73 reduced the errors by a similar amount, except above 6 km, where the WV73 experiment shows a larger error reduction. The WV62 experiment, however, was worse than WV73 and increased the error at altitudes between 1 and 3 km.

Somewhat surprisingly, the vertical distribution of cloud water (WRF’s *QCLOUD* variable) was not generally improved (Figure 14c). Most layers show increased errors compared with the noDA experiment. Only the layer with the highest errors shows slightly reduced errors in the REFL experiment, which assimilated radar observations. Note that radar is mostly blind to cloud droplets. The largest error increase occurred for WV62, the least for REFL. Although the vertically resolved MAE of cloud water did not improve, the FSS and RMSE evaluation (Figures 11 and 10) showed that forecasts of cloudiness were improved overall when the vertical distribution of hydrometeors was not considered.

The vertical distribution of cloud ice (*QICE*) improved between 10 and 12 km. Reductions reached  $0.01 \text{ g} \cdot \text{kg}^{-1}$

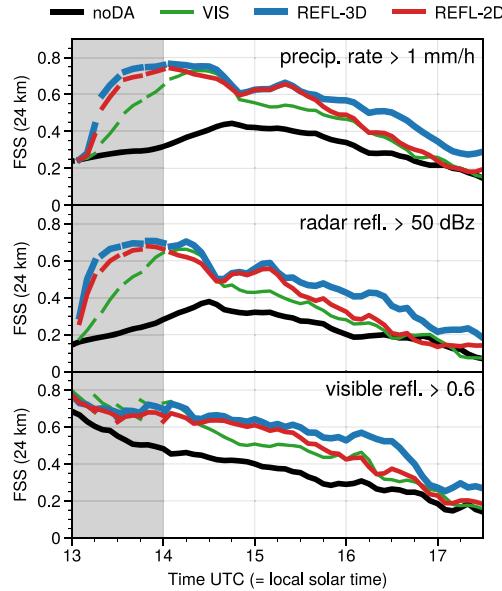
(40%) in the experiments assimilating 6.2 or 7.3  $\mu\text{m}$ , but were less in the VIS experiment. Below 10 km the errors were increased.

In summary, there are overall improvements in temperature and water vapor, except for low-level water vapour in the experiment assimilating 6.2- $\mu\text{m}$  BT. The vertical distribution of clouds was not improved despite the sensitivity of the observations to clouds and despite the improvements in terms of cloudiness revealed in the last section. Our explanation is that the assimilation improves the model equivalents but does not necessarily improve the vertical distribution of model hydrometeors, as the observations only see the uppermost part of clouds. Additionally, visible observations do not provide any information on the height of the observed cloud. Infrared observations provide information on cloud-top height, but a semi-transparent high cloud can lead to the same observed value as an optically thick cloud at a lower altitude. This means it is not necessary for the cloud to have the correct vertical structure and be at the correct height in order to reproduce observations. Instead, existing ensemble perturbations will be scaled up or down depending on the ensemble correlation between the state variable and the observation. This deficiency, however, may be overcome to some extent when multiple satellite channels with different sensitivities are assimilated together.

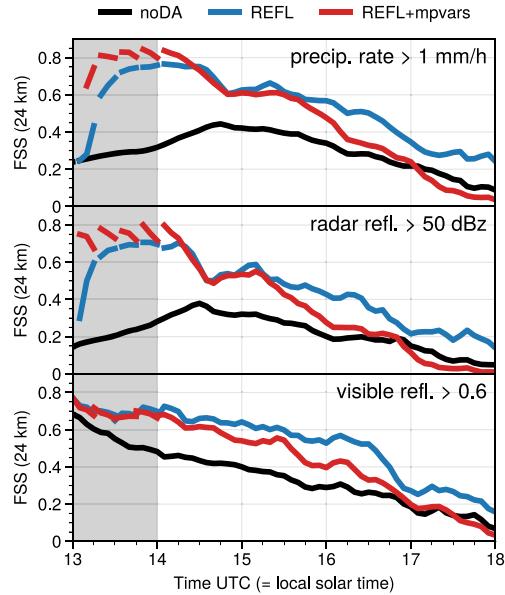
### 3.3 | Sensitivity tests

#### 3.3.1 | Why radar outperforms satellite observations

A prominent detail in the results is that radar observations have an advantage over satellite observations in the random case, but not so much in the warm-bubble case. Given the differences between the cases, we hypothesized that the advantage of radar observations comes from their vertical resolution. To test this hypothesis, we computed a two-dimensional grid of radar observations similar to the two-dimensional satellite observations by projecting the three-dimensional radar observations onto a two-dimensional grid. The projection used the maximum reflectivity of each grid column, as *maximum column reflectivity* is a common tool for operational forecasters and sometimes also used for data assimilation. The result of assimilating this two-dimensional radar (Figure 15) shows a forecast that is relatively similar to experiments that assimilated satellite observations, indicating that vertically resolved observations are indeed crucial in the “random” case in order to reach a high forecast skill.



**FIGURE 15** Fraction skill score of forecasts assimilating two-dimensional radar reflectivity instead of three-dimensional radar reflectivity in the random case. The assimilation time frame is marked by grey shading. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 16** Fraction skill score of forecasts assimilating three-dimensional radar reflectivity, with (REFL+mpvars) and without updating large hydrometeors (REFL) in the random case. The assimilation time frame is marked by grey shading. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3.3.2 | Additional update variables

Radar reflectivity is mostly sensitive to large precipitating hydrometeors, which can be found in the model variables *QRAIN*, *QGRAUP*, and *QSNOW*. Therefore, it may be necessary to update these variables in order to improve the analysis towards the observations. After the analysis, however, hydrometeors will naturally adjust given unstable and moist conditions. To reduce complexity, we chose not to update the large hydrometeor variables in the main experiments listed in Table 1, but to investigate the impact in a sensitivity experiment. As shown in Figure 16, additionally updating these variables does not generally improve forecasts for more than 20 min after the analysis. Furthermore, there is an indication that the vertical distribution of water vapor between 1 and 4 km is negatively affected by updates of large hydrometeors (not shown).

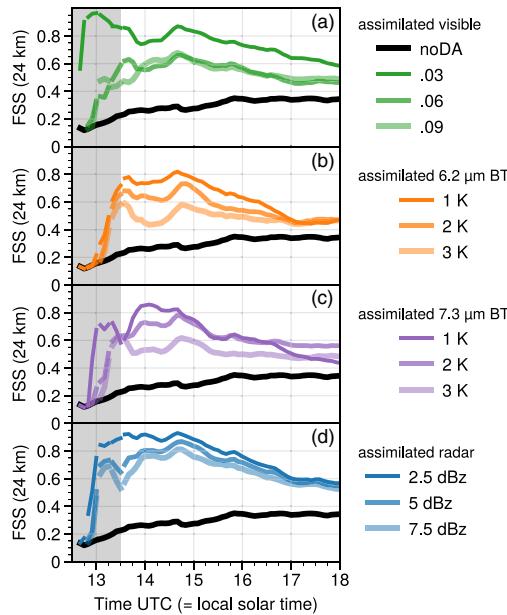
### 3.3.3 | Assigned observation-error variance

We tested a range of constant values (Table 1) for the assigned observation error in order to find the observation

errors that performed best in terms of the FSS. Figure 17 shows the sensitivity of the FSS (of light precipitation) for the increased assigned observation errors. Increasing the observation error never improved the results. Assigning the instrument error as observation error gave best results for all observation types. As there were seven times more radar observations than satellite observations due to its vertical resolution, radar had a higher combined weight in the assimilation. Nevertheless, assigning less weight (increased error) did not lead to improved forecasts. Doubling the assigned error removed its advantage compared with other experiments and led to a forecast impact that was mostly between the experiments WV73 and VIS, yet still higher than the WV62 experiment after 1630 UTC. Cloud-affected BTs of 7.3  $\mu\text{m}$  show much larger first-guess deviations than in the 6.2- $\mu\text{m}$  channel, a possible reason why assigning 2 K led to better forecasts after 3-hr lead time.

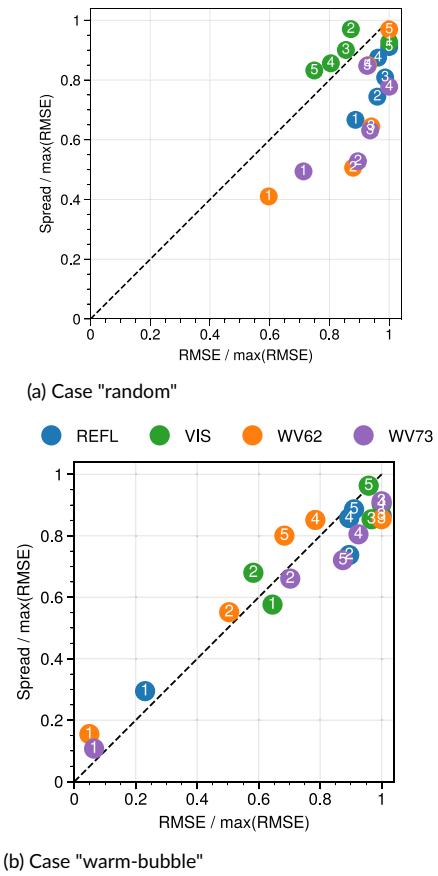
### 3.3.4 | Dynamic observation errors

For infrared satellite observations, the first-guess departures increase with the occurrence of clouds, mainly due



**FIGURE 17** Fraction skill scores for light precipitation ( $>1 \text{ mm}\cdot\text{hr}^{-1}$ ) in the warm-bubble case using assigned observation errors of 1–3 times the instrument error for observations of (a) visible reflectance, (b) 6.2- $\mu\text{m}$  infrared BT, (c) 7.3- $\mu\text{m}$  infrared BT, and (d) radar reflectivity. The assimilation time frame is marked by grey shading. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

to misplacement of clouds (Harnisch *et al.*, 2016). Following Geer and Bauer (2011), this error can be considered to be part of the observation error. Thus, assigning constant observation errors can be suboptimal, especially for 7.3- $\mu\text{m}$  BT, which shows the largest first-guess departures. We tested the dynamic observation-error model of Harnisch *et al.* (2016) but found the results to be substantially worse than using constant observation errors. This is in contrast to Schröttle *et al.* (2020), who successfully applied the dynamic model for the 6.2- $\mu\text{m}$  channel, but used the ICON model and an LETKF assimilation system instead of the WRF model with the EAKF in our study. A possible explanation could be underestimated ensemble spread together with the inflated observation error, which would lead to negligible weights for observations. However, it seems that the prior error variance was well estimated in the warm-bubble and “random” cases (Figure 18), although a small deviation from the ideal relationship can be seen. Nevertheless it should be noted that a dynamic observation-error model refined for the scenarios investigated in our study and the WRF EAKF system may still lead to higher impact of infrared observations than in our comparison.



**FIGURE 18** Spread error relationship for both cases. Spread<sup>2</sup> is the prior ensemble variance, averaged over observations. RMSE is  $\sqrt{\langle (H(x_0) - H(x_{\text{nat}}))^2 \rangle}$ , where  $x_0$  is the prior,  $x_{\text{nat}}$  nature, and  $\langle \cdot \rangle$  the average over observations. Numbers in the dots refer to the number of the cycle. In order to use one axis for different observation types, we rescaled by dividing by the maximum RMSE for each observation type. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

#### 4 | CONCLUSIONS

This study presents the first direct comparison of the assimilation of visible and infrared satellite observations with that of radar reflectivity observations and the first study assimilating visible observations using the ensemble adjustment Kalman filter (EAKF) on the convective scale. We assimilated synthetic observations of 0.6- $\mu\text{m}$  visible reflectance as well as 6.2- and 7.3- $\mu\text{m}$  infrared brightness temperature and radar reflectivity in an idealized perfect-model OSSE. The forecast impact was evaluated in two weather situations: first a “supercell” case in which a

warm bubble of 30-km diameter initiated a single supercell storm and second a case where multiple deep convective cells at different stages are scattered throughout the domain. The periodic boundary domain of  $400 \times 400 \text{ km}^2$  size was simulated using a 2-km resolution WRF model, which was used in identical configuration for the forecast as well as the nature run. The observations were assimilated five times (every 15 min) within 1 hr, during the growth and consecutive mature stage of convection.

#### 4.1 | Main findings

- (1) The EAKF is able to draw crucial information from satellite observations despite the nonlinear observation operators, and their assimilation substantially improves the subsequent forecasts of precipitation and cloudiness. Furthermore, we demonstrate that visible satellite observations can be considerably more beneficial than previously reported by Schröttle *et al.* (2020), reaching an impact of 88% of the impact of three-dimensional radar observations and also outperforming the assimilation of thermal infrared satellite observations.
- (2) Visible and infrared satellite observations can have an impact on forecasts of convective precipitation that is comparable to the impact of radar reflectivity observations. Given favorable conditions, that is, when the stage of convection is correct in the prior and only the location is uncertain (“warm-bubble” case), the assimilation of satellite observations strongly improved the precipitation forecasts: visible observations lead to 88% of the radar impact, while the vapor-sensitive channels at 6.2 and 7.3  $\mu\text{m}$  lead to 74–76% of the radar impact. In more difficult conditions, that is, randomly located storms at different stages (“random” case), the relative impact was lower but still reached 50% for visible observations, 20% for 6.2- $\mu\text{m}$  BT, and 79% for 7.3- $\mu\text{m}$  BT. Assimilating two-dimensional (max-column) radar reflectivity yielded 64% of the impact of three-dimensional radar reflectivity assimilation.
- (3) The differences between the simulated cases suggest that the impact of visible reflectance and 6.2- $\mu\text{m}$  BT observations is highest when the uncertainty about the vertical structure of clouds is lowest. The vertical structure of clouds cannot be retrieved from a single channel and thus is a weak spot for satellite observations. Comparing the “warm-bubble” and “random” cases, we noticed that, in one case, the missing vertical resolution of the assimilated satellite observations did not seem to have a detrimental effect on subsequent forecasts. We hypothesized that the uncertainty in the vertical distribution of clouds is responsible

for the reduced impact of satellite observations in the “random” case. Experiments that only assimilated 2D radar observations and withheld the vertical resolution of the radar data (Section 3.3) supported that hypothesis. This result is in agreement with Sawada *et al.* (2019), who found improved forecasts of isolated cells in case of weak large-scale forcing by assimilating observations of 7.3- $\mu\text{m}$  infrared BT.

#### 4.2 | Additional remarks

In order to generalize our results for operational numerical weather prediction, additional error sources need to be considered, which have not been included in this study: systematic model and operator errors (biases), representativeness errors, and correlated observation errors. While Errico and Privé (2018) argued for the simulation of as many error sources as possible, we refrained from that to isolate particular aspects of assimilating cloud-affected satellite observations (e.g., nonlinearity) and understand better their potential impact on convective-scale forecasts in the absence of all complexities of a real system. Zhang *et al.* (2016) suggest that the impact derived from a perfect-model OSSE may deviate from that in real systems, but the results are still very informative in a qualitative sense. The impact of the observations in operational systems is likely lower in absolute terms due to additional error sources that, for example, require an inflation of observation errors. For this reason, our study focuses on the impact of the observations relative to more commonly assimilated radar observations, which can be assumed to be less affected by the simplifications of the setup mentioned above.

The observation impact diagnosed from an OSSE depends on the choice of observation error and ensemble spread. In addition to a reasonable choice of observation error (Section 2.4) and spread (Section 2.3.2), the statistics of first-guess departures support our OSSE setup. The standard deviation of first-guess departures (for single members and not the ensemble mean) was 4.5 K for the 6.2- $\mu\text{m}$  channel, 9.1 K for the 7.3- $\mu\text{m}$  channel, and 0.22 for the visible channel at 1330. Compared with Harnisch *et al.* (2016), these values indicate that our setup features realistic departures and a case that is an even more difficult situation for numerical weather prediction.

#### 4.3 | Outlook

Our results reveal that the prediction of deep convection could strongly benefit from the assimilation of visible and infrared satellite observations. While the assimilation of infrared observations has been investigated previously,

very few studies have investigated the assimilation of visible observations up to now. Furthermore, radar observations are not available in many parts of the world or they are of limited quality, for example, due to orography that can obscure parts of the precipitation.

Despite recent progress in the effective assimilation of satellite observations, numerous open challenges still need to be addressed. The nonlinearity and non-Gaussianity of the observations and the model call for improved algorithms that allow non-Gaussian distributions, for example, as proposed by Anderson (2010, 2020, 2022), and take observation operator nonlinearity into account. Furthermore, the vertical resolution is a weak spot of visible and infrared satellite observations. Different channels provide information about different atmospheric levels as shown by their correlation structures (Zhang *et al.*, 2022). For example, while the 6.2- $\mu\text{m}$  channel is mostly sensitive to upper tropospheric water vapor, the 7.3- $\mu\text{m}$  channel sees further down into the lower troposphere. Both channels are sensitive to thin ice clouds, which makes them blind to clouds below. The 0.6- $\mu\text{m}$  visible channel can be crucial here, as thin ice clouds are mostly transparent at this wavelength (Scheck *et al.*, 2020). Lastly, cloud height information from window channels could be used to avoid assigning clouds to the wrong levels in the model. The combined assimilation of these different channels therefore has the potential to lead to a better vertical distribution of the increments and subsequently a better forecast. First experiments on this combined approach indeed revealed promising results, and a follow-on publication that investigates how the combined assimilation of visible, infrared, and radar observations affects the analysis increments and forecast is currently in preparation.

#### 4.4 | Data and Software

Experiments were conducted using the Python package available at <https://github.com/lkugler/DART-WRF>. This allows us to define experiment workflows using DART and WRF and contains routines to generate DART observation sequence files and a python-pandas interface to analyze observation sequence files. DART was used in version 10.5.3 with a slightly modified RTTOV interface with constant radii for water droplets and ice crystals, as mentioned in Section 2.4. During the course of this study, we improved the RTTOV interface in DART by removing a bug, which wrongly flipped cloud variables vertically before DART v10.1.0. Other tools: RTTOV v13.3, WRF v4.3. Python packages: xarray (Hoyer and Hamman, 2017), metpy, matplotlib, dask, proplot. Supplementary data and figures are published at <https://doi.org/10.5281/zenodo.7840304> (Kugler, 2023).

#### AUTHOR CONTRIBUTIONS

**Lukas Kugler:** conceptualization; data curation; formal analysis; investigation; methodology; software; visualization; writing – original draft. **Jeffrey L. Anderson:** methodology; software; writing – review and editing. **Martin Weissmann:** conceptualization; formal analysis; methodology; supervision; writing – review and editing.

#### ACKNOWLEDGEMENTS

We thank L. Scheck for FSS verification code based on Faggian *et al.* (2015) and T. Necker and L. Wolfgruber for a bug-fix for odd window sizes. Comments from four anonymous reviewers improved the manuscript.

#### ORCID

Lukas Kugler  <https://orcid.org/0000-0002-4537-3164>

#### TWITTER

Lukas Kugler  [@\\_lkugler](https://twitter.com/_lkugler)

#### REFERENCES

- Aksoy, A., Dowell, D.C. and Snyder, C. (2009) A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: storm-scale analyses. *Monthly Weather Review*, 137(6), 1805–1824. <https://doi.org/10.1175/2008MWR2691.1>.
- Anderson, J.L. (2001) An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129(12), 2884–2903.
- Anderson, J.L. (2012) Localization and sampling error correction in ensemble Kalman filter data assimilation. *Monthly Weather Review*, 140(7), 2359–2371.
- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R. and Avellaneda, A. (2009) The data assimilation research testbed: a community facility. *Bulletin of the American Meteorological Society*, 90(9), 1283–1296. <https://doi.org/10.1175/2009BAMS2618.1>.
- Anderson, J.L. (2010) A non-Gaussian ensemble filter update for data assimilation. *Monthly Weather Review*, 138(11), 4186–4198. <https://doi.org/10.1175/2010MWR3253.1>.
- Anderson, J.L. (2020) A marginal adjustment rank histogram filter for non-Gaussian ensemble data assimilation. *Monthly Weather Review*, 148(8), 3361–3378. <https://doi.org/10.1175/MWR-D-19-0307.1>.
- Anderson, J.L. (2022) A quantile-conserving ensemble filter framework. Part I: updating an observed variable. *Monthly Weather Review*, 150(5), 1061–1074. <https://doi.org/10.1175/MWR-D-21-0229.1>.
- Bachmann, K., Keil, C., Craig, G.C., Weissmann, M. and Welzbacher, C.A. (2020) Predictability of deep convection in idealized and operational forecasts: effects of radar data assimilation, orography, and synoptic weather regime. *Monthly Weather Review*, 148(1), 63–81.
- Bachmann, K., Keil, C. and Weissmann, M. (2019) Impact of radar data assimilation and orography on predictability of deep convection. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 117–130.

- Benjamin, S.G., Weygandt, S.S., Brown, J.M., Hu, M., Alexander, C.R., Smirnova, T.G., Olson, J.B., James, E.P., Dowell, D.C., Grell, G.A., Lin, H., Peckham, S.E., Smith, T.L., Moninger, W.R., Kenyon, J.S. and Manikin, G.S. (2016) A North American hourly assimilation and model forecast cycle: the rapid refresh. *Monthly Weather Review*, 144(4), 1669–1694. <https://doi.org/10.1175/MWR-D-15-0242.1>.
- Chou, M.-D., Lee, K.-T., Tsay, S.-C. and Qiang, F. (1999) Parameterization for cloud longwave scattering for use in atmospheric models. *Journal of Climate*, 12(1), 159–169. [https://doi.org/10.1175/1520-0442\(1999\)012<0159:PFCLSF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0159:PFCLSF>2.0.CO;2).
- Cintineo, R.M., Otkin, J.A., Jones, T.A., Koch, S. and Stensrud, D.J. (2016) Assimilation of synthetic GOES-R ABI infrared brightness temperatures and WSR-88D radar observations in a high-resolution OSSE. *Monthly Weather Review*, 144(9), 3159–3180.
- Cuenca, R.H. and Tewari, M. (2004) *Implementation and verification of the unified NOAH land surface model in the WRF model*.
- Errico, R.M. and Privé, N.C. (2018) *Some General and Fundamental Requirements for Designing Observing System Simulation Experiments (OSSEs)*. Technical Report. World Meteorological Organization. <https://community.wmo.int/wwrp-publications>.
- Eure, K.C., Mykolajchuk, P.D., Zhang, Y., Stensrud, D.J., Zhang, F., Greybush, S.J. and Kumjian, M.R. (2023) Simultaneous assimilation of planetary boundary layer observations from radar and all-sky satellite observations to improve forecasts of convection initiation. *Monthly Weather Review*, 151(3), 795–813. <https://doi.org/10.1175/MWR-D-22-0188.1>.
- Faggian, N.A.T.H.A.N., Roux, B., Steinle, P. and Ebert, B. (2015) Fast calculation of the fractions skill score. *Mausam*, 66(3), 457–466.
- Geer, A.J. and Bauer, P. (2011) Observation errors in all-sky data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 137(661), 2024–2037. <https://doi.org/10.1002/qj.830>.
- Geiss, S., Scheck, L., de Lozar, A. and Weissmann, M. (2021) Understanding the model representation of clouds based on visible and infrared satellite observations. *Atmospheric Chemistry and Physics*, 21(16), 12273–12290. <https://doi.org/10.5194/acp-21-12273-2021>.
- 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. and Fujita, T. (2018) Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713), 1218–1256.
- Harnisch, F., Weissmann, M. and Periéñez, Á. (2016) Error model for the assimilation of cloud-affected infrared satellite observations in an ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 142(697), 1797–1808. <https://doi.org/10.1002/qj.2776>.
- Honda, T., Miyoshi, T., Lien, G.-Y., Nishizawa, S., Yoshida, R., Adachi, S.A., Terasaki, K., Okamoto, K., Tomita, H. and Bessho, K. (2018) Assimilating all-sky Himawari-8 satellite infrared radiances: a case of typhoon Soudelor (2015). *Monthly Weather Review*, 146(1), 213–229. <https://doi.org/10.1175/MWR-D-16-0357.1>.
- Hoyer, S. and Hamman, J.J. (2017) Xarray: N-D labeled arrays and datasets in python. *Journal of Open Research Software*, 5, 1–6. <https://doi.org/10.5334/jors.148>.
- Hu, G., Dance, S.L., Bannister, R.N., Chipilski, H.G., Guillet, O., Macpherson, B., Weissmann, M. and Yussouf, N. (2022) Progress, challenges, and future steps in data assimilation for convection-permitting numerical weather prediction: report on the virtual meeting held on 10 and 12 November 2021. *Atmospheric Science Letters*, 21(1), e1130. <https://doi.org/10.1002/asl.1130>.
- 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: Atmospheres*, 113(D13), 1–8. <https://doi.org/10.1029/2008JD009944>.
- Jones, T.A., Knopfmeier, K., Wheatley, D., Creager, G., Minnis, P. and Palikonda, R. (2016) Storm-scale data assimilation and ensemble forecasting with the NSSL experimental warn-on-forecast system. Part II: combined radar and satellite data experiments. *Weather and Forecasting*, 31(1), 297–327.
- Jones, T.A., Skinner, P., Yussouf, N., Knopfmeier, K., Reinhardt, A., Wang, X., Bedka, K., Smith, W. and Palikonda, R. (2020) Assimilation of GOES-16 radiances and retrievals into the warn-on-forecast system. *Monthly Weather Review*, 148(5), 1829–1859. <https://doi.org/10.1175/MWR-D-19-0379.1>.
- Jones, T.A., Stensrud, D., Wicker, L., Minnis, P. and Palikonda, R. (2015) Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 may 2011. *Monthly Weather Review*, 143(1), 165–194. <https://doi.org/10.1175/MWR-D-14-00180.1>.
- Kain, J.S., Weiss, S.J., Bright, D.R., Baldwin, M.E., Levit, J.J., Carbin, G.W., Schwartz, C.S., Weisman, M.L., Droege, K.K., Weber, D.B. and Thomas, K.W. (2008) Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Weather and Forecasting*, 23(5), 931–952. <https://doi.org/10.1175/WAF2007106.1>.
- Kugler, L. (2023) *Potential impact of assimilating visible and infrared satellite observations compared to radar reflectivity for convective-scale NWP*. Version: v1. <https://doi.org/10.5281/zenodo.7840304>.
- Lange, H. and Craig, G.C. (2014) The impact of data assimilation length scales on analysis and prediction of convective storms. *Monthly Weather Review*, 142(10), 3781–3808.
- Lange, H., Craig, G.C. and Janjić, T. (2017) Characterizing noise and spurious convection in convective data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143(709), 3060–3069.
- Maddox, R.A., Zhang, J., Gourley, J.J. and Howard, K.W. (2002) Weather radar coverage over the contiguous United States. *Weather and Forecasting*, 17(4), 927–934. [https://doi.org/10.1175/1520-0434\(2002\)017<0927:WRCOTC>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2).
- Martinaitis, S.M., Osborne, A.P., Simpson, M.J., Zhang, J., Howard, K.W., Cocks, S.B., Arthur, A., Langston, C. and Kane, B.T. (2020) A physically based multisensor quantitative precipitation estimation approach for gap-filling radar coverage. *Journal of Hydrometeorology*, 21(7), 1485–1511. <https://doi.org/10.1175/JHM-D-19-0264.1>.
- 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(2), 397–407. <https://doi.org/10.1007/s10546-005-9030-8>.

- Otkin, J.A. (2012a) Assessing the impact of the covariance localization radius when assimilating infrared brightness temperature observations using an ensemble Kalman filter. *Monthly Weather Review*, 140(2), 543–561. <https://doi.org/10.1175/MWR-D-11-00084.1>.
- Otkin, J.A. (2012b) Assimilation of water vapor sensitive infrared brightness temperature observations during a high impact weather event: assimilation of infrared observations. *Journal of Geophysical Research: Atmospheres*, 117(D19), 1–16. <https://doi.org/10.1029/2012JD017568>.
- Roebeling, R.A., Wolters, E.L.A., Meirink, J.F. and Leijnse, H. (2012) Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data. *Journal of Hydrometeorology*, 13(5), 1552–1566. <https://doi.org/10.1175/JHM-D-11-089.1>.
- Saltikoff, E., Haase, G., Delobbe, L., Gaussiat, N., Martet, M., Idziorek, D., Leijnse, H., Nov k , P., Lukach, M. and Stephan, K. (2019) OPERA the radar project. *Atmosphere*, 10(6), 320. <https://doi.org/10.3390/atmos10060320>.
- Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N. and Lupu, C. (2018) An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7), 2717–2737. <https://doi.org/10.5194/gmd-11-2717-2018>.
- Sawada, Y., Okamoto, K., Kunii, M. and Miyoshi, T. (2019) Assimilating every-10-minute Himawari-8 infrared radiances to improve convective predictability. *Journal of Geophysical Research: Atmospheres*, 124(5), 2546–2561. <https://doi.org/10.1029/2018JD029643>.
- Scheck, L., Fr rebeau, P., Buras-Schnell, R. and Mayer, B. (2016) A fast radiative transfer method for the simulation of visible satellite imagery. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 175, 54–67. <https://doi.org/10.1016/j.jqsrt.2016.02.008>.
- Scheck, L., Weissmann, M. and Bach, L. (2020) Assimilating visible satellite images for convective-scale numerical weather prediction: a case-study. *Quarterly Journal of the Royal Meteorological Society*, 146(732), 3165–3186. <https://doi.org/10.1002/qj.3840>.
- Scheck, L., Weissmann, M. and Mayer, B. (2018) Efficient methods to account for cloud-top inclination and cloud overlap in synthetic visible satellite images. *Journal of Atmospheric and Oceanic Technology*, 35(3), 665–685. <https://doi.org/10.1175/JTECH-D-17-0057.1>.
- Schr ttle, J., Weissmann, M., Scheck, L. and Hutt, A. (2020) Assimilating visible and infrared radiances in idealized simulations of deep convection. *Monthly Weather Review*, 148(11), 4357–4375. <https://doi.org/10.1175/MWR-D-20-0002.1>.
- Schwartz, C.S., Romine, G.S. and Dowell, D.C. (2021) Toward unifying short-term and next-day convection-allowing ensemble forecast systems with a continuously cycling 3-km ensemble Kalman filter over the entire conterminous United States. *Weather and Forecasting*, 36(2), 379–405. <https://doi.org/10.1175/WAF-D-20-0110.1>.
- Skamarock, C., Klemp, B., Jimy Dudhia, O., Gill, Z.L., Berner, J., Wei Wang, G., Powers, G.D., Barker, D. and Huang, X.-y. (2021) *A description of the advanced research WRF model version 4.3*. Technical Report. NCAR/UCAR. <https://doi.org/10.5065/1dfh-6p97>.
- Snyder, C. and Zhang, F. (2003) Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review*, 131(8), 1663–1677. <https://doi.org/10.1175//2555.1>.
- 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(12), 5095–5115. <https://doi.org/10.1175/2008MWR2387.1>.
- Tong, M. and Xue, M. (2005) Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Monthly Weather Review*, 133(7), 1789–1807. <https://doi.org/10.1175/MWR2898.1>.
- UCAR/NCAR/CISL/DARES. (2022) *The data assimilation research testbed (Version 10.5.3)* [Software].
- Wheatley, D.M., Knopfmeier, K.H., Jones, T.A. and Creager, G.J. (2015) Storm-scale data assimilation and ensemble forecasting with the NSSL experimental warn-on-forecast system. Part I: radar data experiments. *Weather and Forecasting*, 30(6), 1795–1817. <https://doi.org/10.1175/WAF-D-15-0043.1>.
- Zhang, F., Minamide, M. and Clothiaux, E.E. (2016) Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophysical Research Letters*, 43(6), 2954–2963. <https://doi.org/10.1002/2016GL068468>.
- Zhang, Y., Clothiaux, E.E. and Stensrud, D.J. (2022) Correlation structures between satellite all-sky infrared brightness temperatures and the atmospheric state at storm scales. *Advances in Atmospheric Sciences*, 39(5), 714–732. <https://doi.org/10.1007/s00376-021-0352-3>.
- Zhang, Y., Stensrud, D.J. and Zhang, F. (2019) Simultaneous assimilation of radar and all-sky satellite infrared radiance observations for convection-allowing ensemble analysis and prediction of severe thunderstorms. *Monthly Weather Review*, 147(12), 4389–4409. <https://doi.org/10.1175/MWR-D-19-0163.1>.
- Zhu, L., Xue, M., Kong, R. and Min, J. (2022) Direct assimilation of all-sky GOES-R ABI radiances in GSI EnKF for the analysis and forecasting of a mesoscale convective system. *Monthly Weather Review*, 151(3), 737–760. <https://doi.org/10.1175/MWR-D-21-0293.1>.

**How to cite this article:** Kugler, L., Anderson, J.L. & Weissmann, M. (2023) Potential impact of all-sky assimilation of visible and infrared satellite observations compared with radar reflectivity for convective-scale numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 1–22. Available from: <https://doi.org/10.1002/qj.4577>



## **Chapter 4**

# **Combined assimilation of multiple satellite channels**

### **Overview**

The second publication extends the impact assessment of the first publication to the combined assimilation of multiple satellite channels (the second research goal in Section 1.4). It presents the first study that investigates the complementary value of the combined assimilation of 6.2  $\mu\text{m}$  and 7.3  $\mu\text{m}$  infrared as well as visible observations for the analysis of clouds and the forecasts of precipitation and other variables. Additionally, extensive sensitivity studies for assimilation parameters are included. As in the first publication, systematic model and operator errors are excluded.

### **Own contribution**

Conceptualization, formal analysis, investigation, methodology, software, visualization, writing. The author's contribution to this publication is estimated to be about 80%.

### **Reference**

Lukas Kugler and Martin Weissmann, 2024: The synergy of assimilating visible and infrared radiances and radar observations. Under review in Quarterly Journal of the Royal Meteorological Society.

---

**RESEARCH ARTICLE**

manuscript for review

# The synergy of assimilating visible and infrared radiances and radar observations

Lukas Kugler<sup>1</sup> | Martin Weissmann<sup>1</sup>

<sup>1</sup>Institut für Meteorologie und Geophysik,  
Universität Wien, Vienna, Austria

**Correspondence**

Lukas Kugler  
Email: lukas.kugler@univie.ac.at

**Funding information**

Cloud-affected satellite observations in the visible and infrared spectrum contain vast and largely complementary information on clouds and atmospheric convection and thereby constitute a promising data source for convective-scale data assimilation. Preceding studies have demonstrated that the assimilation of either one of these observation types can lead to improved convective-scale weather forecasts, but research on the combined assimilation of cloud-affected visible and infrared satellite as well as radar observations is still very scarce. In this paper, we investigate the combined assimilation of multiple infrared and visible satellite channels as well as radar observations and evaluate their analysis and forecast impact with a primary focus on clouds and precipitation. We assimilate visible ( $0.6\text{ }\mu\text{m}$ ) and thermal infrared ( $6.2\text{ }\mu\text{m}$  and  $7.3\text{ }\mu\text{m}$ ) satellite channels in observing system simulation experiments (OSSE) with a perfect-model forecast for two idealized weather scenarios. Observations are synthetically simulated and assimilated by the ensemble adjustment Kalman filter (EAKF). The forecasts used the Weather Research and Forecasting (WRF) model at 2-km grid resolution. Results show that assimilating satellite channels in addition to radar reflectivity can improve forecasts of cloudiness and precipitation while improvements in temperature, humidity and wind fields are about the same

---

**Abbreviations:** BT, brightness temperature; CAPE, convective available potential energy; CIN, convective inhibition; EAKF, ensemble adjustment Kalman filter; RMSE, root-mean-square error; CRPS, continuous ranked probability score; FSS, Fractions Skill Score.

as in the radar experiment. The evaluation of the analysis error for different cloud conditions revealed that the combined assimilation can mitigate the ambiguity of individual visible and infrared channels. Assimilating visible reflectance before infrared brightness temperature can remove pre-existing erroneous water clouds and avoid the introduction of erroneous clouds by the following assimilation of infrared channels. It follows that the combined assimilation of visible and infrared radiances could be crucial to avoid shortcomings of assimilating only visible or infrared radiances in convective-scale numerical weather prediction.

## 1 | INTRODUCTION

2 Accurate forecasts and early warnings of severe weather are the primary goal of convective-scale numerical weather  
3 prediction. To achieve this, we need to correct km-scale errors in the initial conditions using km-scale observations, es-  
4 pecially in mid-level moisture, low-level temperature and wind (Fabry, 2010; Gustafsson et al., 2018). Although direct  
5 observations of these variables on the km-scale are scarce, not all hope is lost, as indirect remote-sensing observations  
6 are abundant. Visible and infrared instruments on geostationary satellites provide observations with high temporal  
7 resolution (10-15 min) and comparably high spatial resolution (e.g. up to  $\frac{1}{2}$  km for Meteosat Third Generation). Ad-  
8 ditionally, we can infer information about the km-scale thermodynamic environment through ensemble covariances  
9 when applying an ensemble Kalman filter (EnKF). Weather radars observe precipitating cores of convective clouds.  
10 However, their coverage is limited by mountains, constrained to highly populated and financially rich regions. In con-  
11 trast, satellites provide spatially homogeneous data quality and detect convection earlier than radars. Thermal infrared  
12 satellite channels measure water vapour and cloud-top temperature. Visible channels can see low-level stratus during  
13 daytime, even through thin high-level ice clouds. It is therefore not surprising that studies found a beneficial impact of  
14 assimilating all-sky infrared satellite observations on forecasts of typhoons, mesoscale convective systems, and other  
15 severe weather events (Jones et al., 2015, 2016; Honda et al., 2018; Sawada et al., 2019; Jones et al., 2020; Chan  
16 et al., 2020; Zhu et al., 2022).

### Ambiguity motivates combined assimilation

17 Although a single-wavelength infrared radiance channel does not reveal the vertical structure of clouds, we can learn  
18 much from the complementary information of different wavelengths. A single infrared observation can be ambiguous,  
19 as cirrus clouds may lead to the same observed signal as deep convective cumulonimbus clouds. Visible observations  
20 can also be ambiguous, as a shallow convective cloud can lead to the same signal as a cumulonimbus cloud. However,  
21 the combination of visible and infrared observations can provide complementary information: Visible observations can  
22 determine whether there are low-level water clouds below the ice cloud that is seen in the infrared, while infrared  
23 channels can determine whether the cloud seen in the visible channel is shallow or deep. Similarly, satellite channels  
24 can complement the assimilation of radar reflectivity, as C/S-band radars don't detect clouds without precipitating  
25 hydrometeors. Thus, the combination of information from different observation types could potentially mitigate the  
26 ambiguity of the information from a single observation type and leverage our ability to infer details on the vertical

28 structure of clouds.

29 Combining the information of different wavelengths can, however, be challenging. The efficacy of the combined  
30 assimilation depends on the assimilation algorithm and also on its ability to account for correlated observation errors  
31 of spectrally close channels. Some preceding studies assimilated multiple infrared channels (Bormann et al., 2016;  
32 Johnson et al., 2022), but most selected only one channel (Honda et al., 2018; Okamoto et al., 2019; Zhang et al., 2019;  
33 Zhu et al., 2022) and thereby avoided inter-channel observation error correlations. However, by selecting channels in  
34 a distinctly different spectral range as e.g. visible and infrared, the inter-channel error correlations can be expected  
35 to be relatively low while the channels provide very complementary information. Infrared radiances from 6.2  $\mu\text{m}$  to  
36 7.3  $\mu\text{m}$  provide information about the vertical distribution of moisture and are sensitive to different but overlapping  
37 parts of the atmospheric state. Considering 6.2 and 7.3  $\mu\text{m}$  brightness temperature (BT), the least difficulty appears  
38 in the clear-sky assimilation as the 6.2  $\mu\text{m}$  channel is sensitive to upper tropospheric water vapour, while the 7.3  $\mu\text{m}$   
39 channel senses lower tropospheric water vapour. Moreover, low clouds (below 600 hPa) are hardly observed by the  
40 6.2  $\mu\text{m}$  channel, which leads to a distinct response in 7.3  $\mu\text{m}$  BT without a response in 6.2  $\mu\text{m}$  BT. However, mid or  
41 high clouds are sensed by both channels, making it hard to distinguish between deep convective clouds and cirrus  
42 clouds.

43 Assimilating cloud-affected visible and infrared observations in an ensemble Kalman filter is challenging, but most  
44 of the associated challenges occur similarly for the more common assimilation of radar observations. The limited  
45 numerical representation of cloud processes and hydrometeors, as well as simplifications of observation operators,  
46 can lead to systematic differences between the observations and their model equivalents (Geiss et al., 2021). We  
47 avoided these issues by using an *observing system simulation experiment* (OSSE) with an identical model for the nature  
48 run and forecasts, as well as perfect observation operators in the forecast. This allows us to focus on the remaining  
49 challenges induced by e.g. dealing with nonlinear observation operators and non-Gaussian error distributions.

50 This paper extends the study (Kugler et al., 2023), in which infrared and visible channels were assimilated sepa-  
51 rately, to the combined assimilation of visible reflectances, infrared brightness temperature (BT) and radar reflectivity  
52 observations. We investigate to which extent visible, infrared and radar observations complement each other in the  
53 assimilation to improve (i) the analysed vertical distribution of clouds and (ii) the forecast of precipitation and clouds  
54 by assimilating 0.6  $\mu\text{m}$  visible reflectance in addition to 6.2 and 7.3  $\mu\text{m}$  infrared BT and S-band radar reflectivity.

55 Section 2 reviews the setup of the observing-system simulation experiments. In Section 3, we define the veri-  
56 fication metrics and introduce the OSSE experiments. Section 4 presents the results of the experiments as well as  
57 comprehensive sensitivity-tests. Section 5 concludes the paper with a summary of the main findings.

## 58 2 | METHODS

59 To answer our research questions, we employ a set of observing system simulation experiments (OSSE) building upon  
60 the configuration of Kugler et al. (2023). In the following, we briefly review the setup.

### 61 2.1 | Observations, modelling and assimilation

#### 62 The observing system simulation experiment

63 The forecast ensemble as well as the nature run use the Weather Research and Forecasting model (WRF)  
64 version 4.3 (Skamarock et al., 2021) in identical configurations. The model domain consists of 200 x 200 x 50 mass  
65 grid points with a 2 km horizontal resolution and a layer depth of 25-500 m with horizontally periodic boundary

Observation type	Instrument error
Visible reflectance 0.6 $\mu\text{m}$ (VIS)	0.03
Brightness temperature 6.2 $\mu\text{m}$ (WV62)	1 K
Brightness temperature 7.3 $\mu\text{m}$ (WV73)	1 K
Radar reflectivity 10 cm (REFL)	2.5 dBz

**TABLE 1** Assimilated observation types and the respective instrument error standard deviation used for simulating the observations.

66 conditions. The surface is idealized with homogeneous and flat terrain at 489 m altitude above sea level and the solar  
 67 radiation resembles a summer day at 45° latitude.

#### 68 **The observations**

69 Observations were generated as in Kugler et al. (2023) and are listed in Table 1. Satellite observations were simulated  
 70 from the WRF output of the nature run using the Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV)  
 71 v13.3 (Saunders et al., 2018) with an effective particle diameter of 20  $\mu\text{m}$  for water and 60  $\mu\text{m}$  for ice particles. In-  
 72 frared and visible radiances are parametrized using Chou-scaling (Chou et al., 1999) and MFASIS (Scheck et al., 2016),  
 73 respectively. Radar reflectivity was diagnosed by WRF with Thompson microphysics. For both satellite and radar  
 74 observations, random errors were added to the simulated observations according to Table 1. Both observations and  
 75 their model equivalents were generated using the same observation operators to exclude systematic operator errors.

#### 76 **The assimilation system**

77 We used the ensemble adjustment Kalman filter (EAKF) by Anderson (2001) implemented in the Data Assimilation  
 78 Research Testbed (DART, Anderson et al. (2009), <https://dart.ucar.edu>) and a 40-member forecast ensemble to  
 79 assimilate observations serially. Following the reference configuration of Kugler et al. (2023), we updated tempera-  
 80 ture, water vapour mixing ratio, dry air mass in column, geopotential, three wind components, cloud water and ice  
 81 mixing ratio. Including larger precipitating hydrometeor variables was tested in Kugler et al. (2023), but did not have  
 82 a sustained positive impact on forecasts.

83 To retain spread in our ensemble forecasts, we applied posterior inflation, i.e. relaxation to prior spread (RTPS),  
 84 with different values for the relaxation factor  $\alpha$ . The influence of observations was localised using a Gaspari-Cohn  
 85 function with 20 km half-width in the horizontal (40 km radius of influence), similar to the 30 km correlation length  
 86 for cloudy infrared radiance found by Zhang et al. (2022). Radar observations were localised to 3 km in the verti-  
 87 cal. Satellite radiances were not localised vertically. Additionally, we applied the statistical sampling error correction  
 88 (Anderson, 2012; Necker et al., 2020).

89 We tested different values for the assigned observation error, starting with the instrumental error that was used to  
 90 simulate the observations, and then increased it. Additionally, one experiment (abbreviated OE:ID) set the observation  
 91 error variance to be the variance of first-guess departures minus ensemble variance:

$$\widehat{\sigma_o^2} = \frac{1}{T-1} \sum_{t=1}^T (\bar{y}_b - y_o)^2 - \frac{1}{T} \sum_{t=1}^T \frac{1}{N-1} \sum_{n=1}^N (y_b - \bar{y}_b)^2, \quad (1)$$

92 where  $T$  is the number of observations and  $N$  is the ensemble size.

93 In the horizontal, we assimilated observations every 10 km, except within the first 50 km from the domain bound-

94      ary, so that only observations of the inner 300 x 300 km were assimilated. In the vertical, we assimilated seven levels  
 95      of radar reflectivity observations between 2-14 km.

## 96      2.2 | Experimental set-up

97      The assimilation impact was evaluated in two weather scenarios: The "random" case, in which supercells are triggered  
 98      at random places by small random perturbations. Second, a "warm-bubble" case with localised convection and simple  
 99      uncertainties in the location and strength of convection. Both cases share a highly unstable stratification with a CAPE  
 100     of 2670 J/kg and a CIN of 26 J/kg at 7 UTC, such that relatively small perturbations trigger deep convection. For our  
 101     simulations, time in UTC is equivalent to local solar time. Sections 4.1-4.3 evaluate experiments for the random case  
 102     and section 4.4 evaluates experiments for the warm-bubble case.

### 103     Simulated cases

104     In the "random" case, storms are triggered at random locations scattered throughout the whole domain due to small  
 105     random perturbations. We add two perturbations, following Bachmann et al. (2020); Schröttle et al. (2020): (1) Per-  
 106     turbations to the vertical profile of temperature ( $\sigma=0.25$  K), wind ( $\sigma=0.25$  m/s), and humidity ( $\sigma=2\%$ ), which are auto-  
 107     correlated in the vertical and horizontally constant; (2) random noise ( $\sigma=0.02$  K or m/s) added to temperature and  
 108     vertical velocity in the lowest levels, horizontally uncorrelated. The nature run starts at 6 UTC and develops deep  
 109     convection until 11 UTC. Both in the nature and in the ensemble, small random perturbations grow rapidly. The onset  
 110     of deep convection varies by  $\pm 45$  minutes among the ensemble members.

111     In the "warm bubble" case, a bell-shaped, 30 km wide, positive temperature increment (as in Kugler et al., 2023)  
 112     is added to the initial temperature field in addition to the random perturbations. It triggers an isolated and well-  
 113     organised supercell in a confined region of the domain and suppresses convection elsewhere, which makes this case  
 114     more predictable than the "random" case. Adding the "warm bubble" at the same time across all ensemble members  
 115     acts to synchronise the convective initiation and removes uncertainty about the life-cycle of the convective cloud.

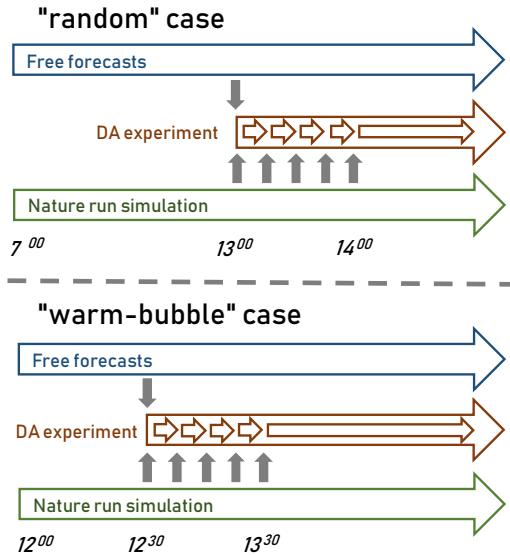
116     The timeline of assimilation and forecast is illustrated in Figure 1. In the random case, the forecast ensemble  
 117     was initialised at 7 UTC, followed by a free forecast without assimilation for six hours. From 13 to 14 UTC, we  
 118     assimilated five times (every 15 minutes), followed by free forecasts until 18 UTC. In the warm-bubble case, we started  
 119     to assimilate at 12:30 after a free forecast of 30 minutes. Despite this short spin-up time, the deep convection had  
 120     already developed. From 12:30 to 13:30 UTC, we assimilated five times, followed by free forecasts until 18 UTC. One  
 121     experiment was conducted without data assimilation (abbreviated noDA). This experiment is used as a reference and  
 122     as a scaling in the Fractions Skill Score.

## 123     3 | VERIFICATION

124     The assimilation experiments were evaluated with the four metrics described in the following.

### 125     Observation-space verification

126     To determine the fit to observations, the continuous ranked probability score (CRPS, e.g. Wilks, 2011) of visible  
 127     reflectance and 7.3  $\mu$ m BT was computed and compared with the CRPS of the noDA forecast. The CRPS compares the  
 128     distribution of the forecast ensemble with the observed value and is, thus, more suitable for verifying non-Gaussian  
 129     variables. The range of CRPS is 0- $\infty$  and smaller values are better.



**FIGURE 1** Timeline of assimilation and forecasts, reproduced from Kugler et al. (2023).

Definition	Condition	Sample size
water clouds	LWP > 300 g/m <sup>2</sup> (but none above 6 km) and IWP < 5 g/m <sup>2</sup>	89
ice clouds	LWP < 10 g/m <sup>2</sup> and IWP > 50 g/m <sup>2</sup>	235
water & ice-phase clouds	LWP > 300 g/m <sup>2</sup> and IWP > 50 g/m <sup>2</sup>	139

**TABLE 2** Definition of cloud categories; Total sample size: 1280 observed columns, separated by 20 km. Data from five assimilation times; LWP: liquid water path, IWP: ice water path.

**130 Model-space verification**

**131** To assess the accuracy of the model variables temperature and water vapour mixing ratio, the root-mean-square error  
**132** (RMSE)

$$\varepsilon(\bar{x}) = \sqrt{\frac{1}{p_{sfc} - p_{top}} \sum_k^K \Delta p_k \sum_i^M (\bar{x}_{i,k} - x_{i,k}^{nat})^2} \quad (2)$$

**133** of the ensemble-mean forecast  $\bar{x}$  to the true state  $x^{nat}$  was computed over  $M=200^2$  columns and  $K=50$  layers  
**134** (weighted by pressure  $p$ ). The relative error reduction compared to noDA is  $(\varepsilon_f - \varepsilon_{noDA})/\varepsilon_{noDA}$  such that negative  
**135** values indicate a reduced error.

**136 Verification of precipitation and cloudiness**

**137** The forecast skill was evaluated using the ensemble Fractions Skill Score (FSS) of three forecast variables over a  
**138** 24 km window: radar reflectivity > 50 dBZ (strong convection), visible reflectance > 0.6 (cloudiness), and precipitation  
**139** > 1 mm/h. The FSS was computed from neighbourhood ensemble probabilities (Schwartz et al., 2010; Necker et al.,  
**140** 2023). The relative FSS compares the FSS of an experiment to the reference experiment assimilating radar (REFL) and

141 the control experiment, which does not assimilate observations (noDA)

$$\text{FSS}_{\text{rel}} = \frac{\text{FSS}_{\text{exp}} - \text{FSS}_{\text{noDA}}}{\text{FSS}_{\text{REFL}} - \text{FSS}_{\text{noDA}}}. \quad (3)$$

142 The relative FSS is defined using 3h-averages of FSS. The range (0-100%) defined by  $\text{FSS}_{\text{noDA}}$  and  $\text{FSS}_{\text{REFL}}$  is 0.264-  
 143 0.465 for radar reflectivity > 50 dBz (strong convection); 0.345-0.569 for visible reflectance > 0.6 (cloudiness); and  
 144 0.351-0.594 for precipitation > 1 mm/h.

145 **Cloud-analysis verification**

146 To evaluate clouds in the analysis, we select the model columns which are observed from the nature run (961 obser-  
 147 vations). From these, we select 256 columns, which are separated by 20 km ( $2\times$  localisation radius) in order to reduce  
 148 the correlation in the samples. This gives a total of 1280 samples over 5 analysis times. The columns are classified into  
 149 three categories, depending on the clouds in the column: only water clouds, only ice clouds, or water and ice clouds  
 150 (as defined in Table 2). For each category, we calculate the mean absolute error (MAE) reduction due to assimilation  
 151 (posterior - prior),

$$\Delta\varepsilon(z) = \left| \overline{\{x_a\}} - \{x_{nat}\} \right| - \left| \overline{\{x_b\}} - \{x_{nat}\} \right|, \quad (4)$$

152 averaging (angle brackets) horizontally and over five different assimilation times. To avoid double penalty due to small  
 153 vertical cloud misplacement, the forecast ensemble-mean  $\bar{x}$  and nature are averaged over seven vertical levels in case  
 154 of cloud variables.

155 **3.1 | Experiments**

156 To ensure that the reported results are as generally valid as possible, we tested a range of parameter settings. In  
 157 Section 4.4, we compiled a complete list of experiments which were used in the sensitivity-testing. Sections 4.1-4.3,  
 158 however, use selected experiments. Experiments are abbreviated by the observation types they assimilated, e.g. the  
 159 experiment VIS+WV73 assimilated visible reflectance first and 7.3  $\mu\text{m}$  BT second and by the RTPS inflation parameter  
 160  $\alpha$ .

- 161 1. The experiment VIS+WV73 ( $\alpha=0.3$ ) was used to show the effect of the combined assimilation of one visible and  
 162 one infrared channel in Section 4.1;  
 163 2. the experiment VIS+WV62+WV73 ( $\alpha=0.3$ ) represents the combination of one visible and two infrared channels  
 164 (Section 4.2);  
 165 3. and the experiment VIS+WV62+WV73+REFL ( $\alpha=0.9$ ,  $\widehat{\sigma_o^2}=4\sigma_o^2$ ) assimilates all four observation types including  
 166 radar reflectivity observations (Section 4.3). The assigned observation error variance  $\widehat{\sigma_o^2}$  is given as a multiple of  
 167 the instrument error variance  $\sigma_o^2$ .

168 If not indicated otherwise, experiments used the standard configuration of RTPS( $\alpha=0.9$ ) and no observation error  
 169 inflation as in Kugler et al. (2023). The reference experiment without data assimilation is abbreviated noDA.

170 **4 | RESULTS**

171 **4.1 | Combination of visible and infrared channels**

172 **Strengths and weaknesses of individual channels**

173 At first, we revisit the strengths and weaknesses of assimilating either visible or infrared radiances to get an idea of the  
174 potential benefits of their combined assimilation. Figure 2 shows the vertical profile of MAE reduction (posterior-prior)  
175 in cloud water and ice concentration when assimilating either VIS or WV73 observations for three cloud categories.  
176 Although observation equivalents were substantially improved (shown later), the vertical distribution of clouds was not  
177 always improved, potentially due to vertical displacement errors resulting from the ambiguity of individual channels.

178 The top left panel shows that water clouds below 6 km were improved more by assimilating visible reflectance  
179 (experiment VIS) than by assimilating 7.3  $\mu\text{m}$  infrared BT (experiment WV73), which can be expected as the visible is  
180 more sensitive to liquid than ice cloud and the 7.3  $\mu\text{m}$  channel's weighting function peaks at 500 hPa. However, the  
181 assimilation of visible reflectance introduced erroneous clouds above 6 km leading to increased errors for both cloud  
182 water and ice (top left and right panel). The assimilation of 7.3  $\mu\text{m}$  infrared BT also led to some erroneous clouds above  
183 6 km, but the error increase was much smaller for both cloud ice and water than for assimilating visible observations.  
184 Presumably, the erroneous clouds in the experiment with visible observations are due to the ambiguity of the visi-  
185 ble channel, which can not discriminate between highly reflective low-level water clouds and highly reflective deep  
186 convective clouds. Therefore, the assimilation of visible reflectance introduced erroneous deep convective clouds in  
187 some instances. On the other hand, the infrared BT is capable of distinguishing high and low clouds. Correspondingly,  
188 the experiment assimilating 7.3  $\mu\text{m}$  BT (WV73) shows lower errors for high-level water and ice clouds.

189 The center left panel shows that the experiment WV73 introduced substantial errors in cloud water, while ex-  
190 periment VIS did not. It seems that the assimilation of 7.3  $\mu\text{m}$  BT introduced erroneous clouds below the observed  
191 cloud tops because the observation does not contain information about the vertical extent of the cloud such that,  
192 e.g., a cumulonimbus can not be discriminated from a cirrus cloud. The visible channel, however, sees through thin  
193 ice clouds, which probably avoids erroneous increments of cloud water in the VIS experiment.

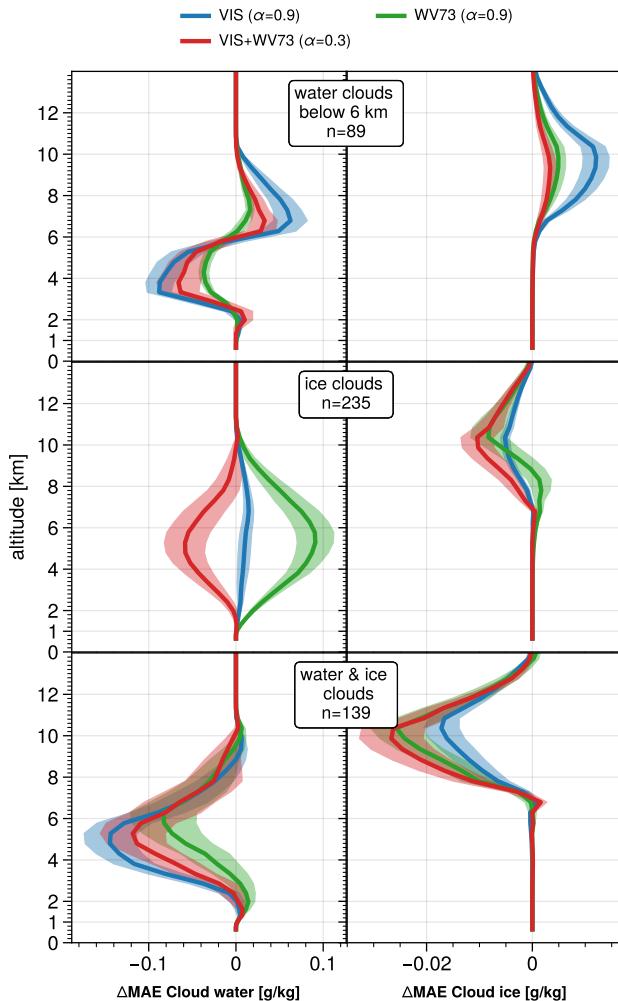
194 The lower panel shows the error reduction in situations where both water and ice clouds were present, which  
195 mostly represents the core of deep convective clouds in our simulation. The assimilation of either visible or infrared  
196 observations led to improvements in both cloud water and ice. The error reduction for cloud water was larger in the  
197 experiment that assimilated visible observations, and the error reduction for cloud ice was larger in the experiment  
198 with infrared observations, as expected.

199 Note that cloud ice increments from visible reflectance were, on average, about 35% smaller than from infrared  
200 BT (not shown), consistent with its lower sensitivity to ice clouds.

201 **Synergy of visible and infrared assimilation**

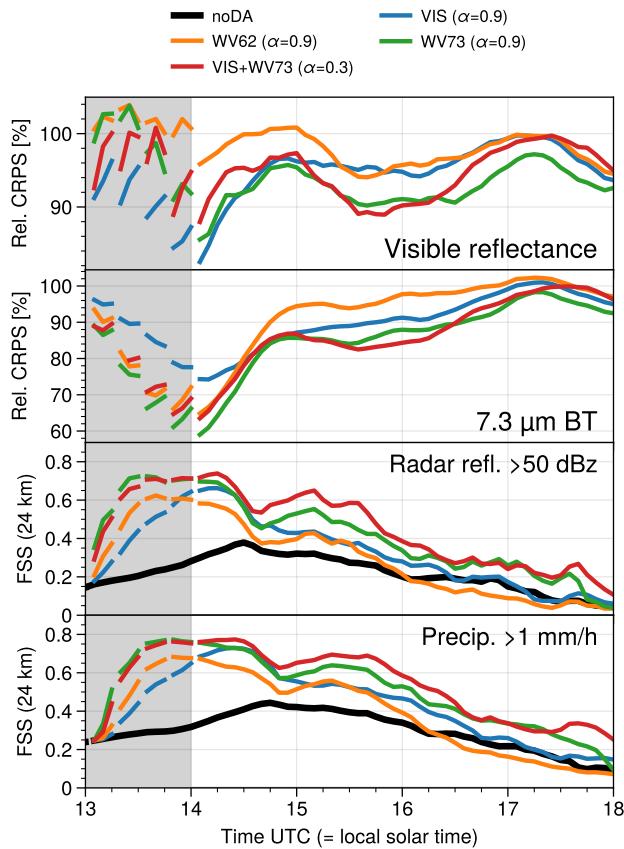
202 As discussed in the introduction, individual channels are ambiguous regarding the cloud structure, but the combination  
203 of channels provides complementary information and thereby reduces the ambiguity. To estimate the benefit of  
204 assimilating visible and infrared observations, we compare the MAE reduction of assimilating either VIS or WV73  
205 against the MAE reduction of assimilating both channels.

206 The cloud analysis for the ice-cloud category (center panel of Figure 2) shows the synergy of visible and infrared  
207 assimilation. While experiment WV73 created erroneous water clouds, assimilating visible and infrared data (exper-  
208 iment VIS+WV73) reduced such erroneous clouds. In this category, the visible channel presumably can provide the  
209 information that no water clouds are present, as water clouds usually lead to higher visible reflectance than mostly



**FIGURE 2** Cloud analysis verification. Vertical profile of MAE reduction (posterior-prior, Equation 4) of cloud water and ice concentration for experiment VIS+WV73. Categories as defined in Table 2. Negative values indicate error reduction. Shading indicates standard error.

210 semi-transparent ice clouds in this channel. This leads to the removal of spurious water clouds and a reduction of  
 211 spread in cloud water, which hinders deterioration in cloud water by the assimilation of infrared observations after-  
 212 wards. In case of water clouds below 6 km (upper panel), the analysis of VIS+WV73 reduces erroneous cloud ice  
 213 compared to the experiment with only visible or only infrared observations but still shows a slightly higher error in  
 214 cloud ice than first-guess. In case of both water and ice clouds in the column (lower panel), the cloud analysis of  
 215 VIS+WV73 is more similar to the experiment assimilating WV73, which has more accurate ice and less accurate water  
 216 clouds. This means that the improvement of cloud water in VIS+WV73 is smaller than in VIS, but the improvement in  
 217 cloud ice in VIS+WV73 is higher than in VIS and even higher than in WV73.



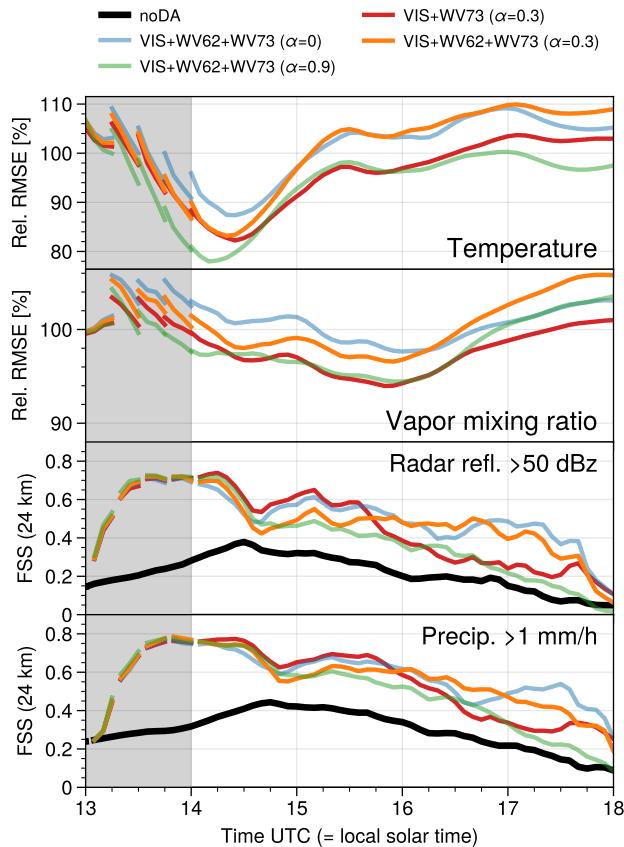
**FIGURE 3** Timeseries of forecast scores, assimilating infrared or visible observations or both.

218     Figure 3 shows the forecast impact of the combined and separate assimilation of visible and infrared channels.  
 219     Note that we used a different inflation parameter  $\alpha$  (relaxation to prior spread) for the combined experiment as this  
 220     setting led to improved scores for the combination. Further discussion on the sensitivity of the results to the relaxation  
 221     parameter will be provided in section 4.4. The combined experiment VIS+WV73 was consistently better in terms of  
 222     the FSS of precipitation and radar reflectivity (lower two panels) than each of the experiments assimilating a single  
 223     channel. This suggests that the combined serial assimilation is able to leverage synergies of the channels, and the  
 224     forecast skill is not just a mere average of the experiments VIS and WV73. Note that the combined assimilation  
 225     improved forecasts without drawing the model-equivalents closer to the observations, as shown by the observation-  
 226     space verification in the upper two panels.

## 227     4.2 | Combination of one visible and two infrared channels

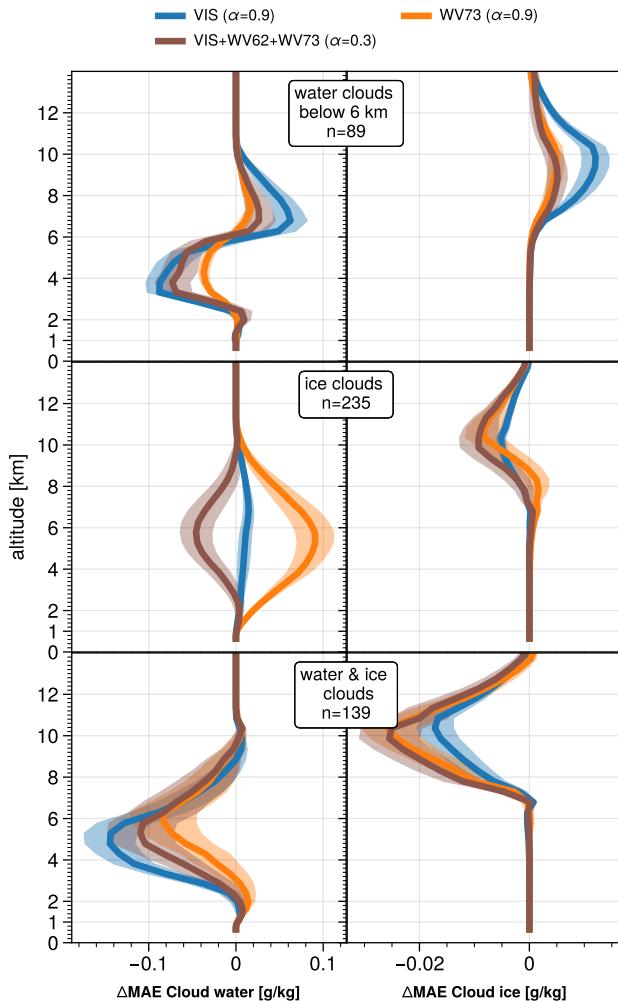
228     Figure 4 shows that the additional assimilation of 6.2  $\mu\text{m}$  BT (comparing experiments VIS+WV73 and VIS+WV62+WV73)  
 229     can improve forecasts of precipitation. However, the improvement depends on the background error inflation. With

230 a small inflation parameter ( $\alpha=0.3$ ), the inclusion of  $6.2\text{ }\mu\text{m}$  BT led to an improved FSS on average over three hours  
 231 (lower panels) but an increased error in temperature and humidity (upper panels). When the inflation parameter  
 232 was increased ( $\alpha=0.9$ ), the temperature and humidity were improved compared to VIS+WV73. However, the FSS  
 233 deteriorated due to spurious convection triggered by inflation. Turning the inflation off ( $\alpha=0$ ) improved the FSS of  
 234 precipitation and cloudiness (lower panels) but increased the error of temperature and humidity (upper panels) and  
 235 wind (not shown). Regardless of inflation, however, the combination VIS+WV62+WV73 outperformed experiments  
 236 which assimilated either VIS, WV62, or WV73 only in terms of the FSS.



**FIGURE 4** Timeseries of forecast scores, assimilating two infrared and one visible channel.

237 Assimilating multiple satellite channels generally improved the vertical distribution of clouds. Figure 5 evaluates  
 238 the vertical distribution of cloud water/ice, comparing the assimilation of two channels (VIS+WV73) with the as-  
 239 similation of three channels (VIS+WV62+WV73). A disadvantage of assimilating only the  $7.3\text{ }\mu\text{m}$  channel was the  
 240 overestimation of water clouds when ice clouds were present (center left panel). Assimilating the visible channel in  
 241 addition to both infrared channels showed a substantial reduction of mean absolute error instead of a worsening due  
 242 to the assimilation of infrared BT. Likewise, the combined assimilation cured a weakness of visible channel assimilation  
 243 in case of low water clouds (top right panel): While the assimilation of visible reflectance (experiment VIS) introduced



**FIGURE 5** Cloud analysis verification. Vertical profile of MAE reduction (posterior-prior, Equation 4) of cloud water and ice concentration when assimilating a combination of three channels. Categories as defined in Table 2. Negative values indicate error reduction. Shading indicates standard error.

244 ice clouds on average, the additional assimilation of two infrared channels mitigated this adverse effect. Mixed results  
 245 can be seen in the estimation of water clouds when only water clouds were present (upper left panel), and in the  
 246 estimation of water clouds when water and ice clouds were present (lower left panel). In these cases, the additional  
 247 assimilation of  $6.2 \mu\text{m}$  BT did not lead to improvements over the experiment VIS+WV73 in the cloud analysis (not  
 248 shown).

**249 4.3 | Combination with radar reflectivity**

250 Figure 6 shows how the additional assimilation of radar reflectivity affects the cloud analysis, comparing experiments  
 251 VIS+WV62+WV73+REFL and VIS+WV62+WV73. In case of water clouds below 6 km, the additional assimilation of  
 252 radar reflectivity partly removed erroneous clouds above 6 km (top left panel). Only radar assimilation (experiment  
 253 REFL) did not introduce erroneous ice clouds. However, the additional assimilation of radar reflectivity (experiment  
 254 VIS+WV62+WV73+REFL) could not remove the erroneous increments of ice clouds (top right panel) caused by visible  
 255 channel assimilation. In case of ice clouds (center panel), the additional assimilation of radar reflectivity led to no  
 256 substantial improvement in cloud water while assimilating satellite-only could remove erroneous water clouds. Hence,  
 257 radar assimilation introduced erroneous water clouds and led to an overall neutral influence on water clouds in this  
 258 case. In case of water and ice clouds (lower panel), there is no apparent change in MAE, given the variability of the  
 259 error reductions.

260 The forecast impact of additional radar assimilation is shown in Figure 7. The largest effect of the additional  
 261 information can be seen in the accuracy of humidity (second panel). The experiment that assimilated radar only  
 262 (REFL) or additionally (VIS+WV63+WV73+REFL  $\alpha=0.9, 4\sigma_o^2$ ) had a lower RMSE of humidity than the experiment that  
 263 did not assimilate radar but did assimilate visible and infrared channels. The RMSE of temperature (top panel) was  
 264 comparable in all experiments with the exception of the VIS+WV62+WV73 experiment without inflation ( $\alpha=0$ ), which  
 265 had a higher RMSE of temperature. Among experiments with the same inflation value ( $\alpha=0.9$ ), the FSS of precipitation  
 266 and cloudiness was better when radar was assimilated. When considering experiments with different inflation values,  
 267 the experiment without inflation and radar (VIS+WV62+WV73,  $\alpha=0$ ) resulted in a better FSS but a higher RMSE of  
 268 temperature and humidity. The effect of inflation is discussed in the next section.

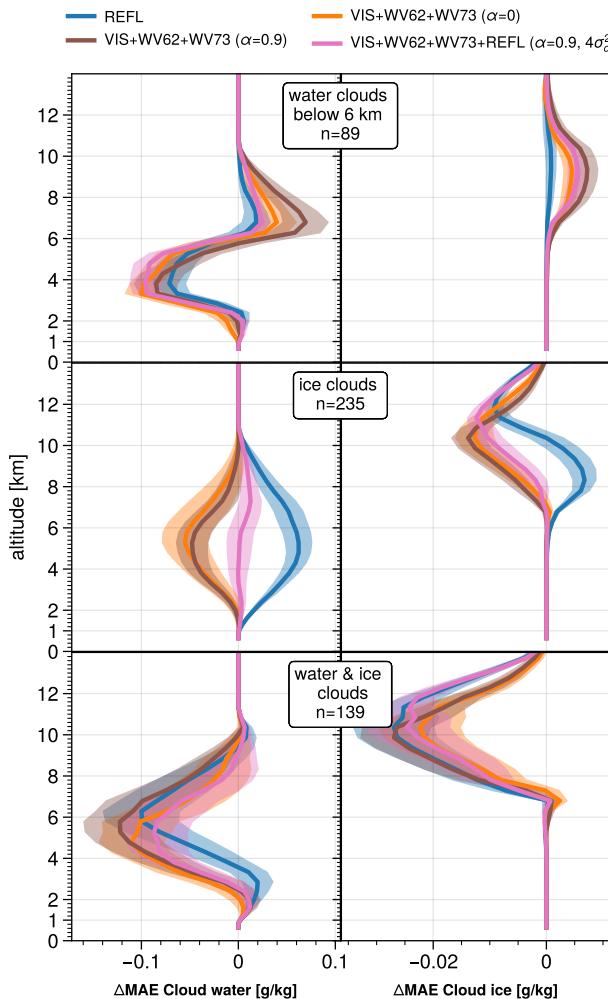
**269 4.4 | Sensitivity to parameters**

270 An overview of the assimilation impact on forecast scores in various experiments is provided by Figure 8, which shows  
 271 the analysis RMSE reduction (compared to noDA, Equation 3) of temperature, humidity, and wind, followed by 5-min  
 272 posterior observation-space CRPS reduction (compared to noDA), and the 3-hr forecast averages. The last three  
 273 columns show the time-averaged FSS of precipitation and cloudiness relative to experiments REFL and noDA. To  
 274 understand the impact of the background error inflation, the order of assimilation, and the observation error inflation,  
 275 we now discuss selected experiments of Figure 8 in detail.

**276 Inflation of the background error**

277 Inflating the background error had the largest impact among the tested parameters. Increasing the inflation to RTPS  
 278  $\alpha=0.9$  generally improved the analysis and forecast error of temperature, humidity, and wind at the cost of a deteriorated  
 279 (i.e. reduced) FSS of precipitation and cloud forecasts. Figure 9 shows the effect of increasing the background  
 280 error inflation in the experiments VIS+WV73 and VIS+WV62+WV73 on temperature RMSE and precipitation FSS.  
 281 In both experiments, the RMSE of temperature was improved with increased inflation. However, the FSS of precip-  
 282 itation and cloudiness clearly deteriorated by increasing inflation. Increasing the inflation in experiment VIS+WV73  
 283 from  $\alpha = 0.3$  over 0.5 to 0.9 decreased the relative FSS of all forecast quantities: from 106% to 82% to 64% for radar  
 284 reflectivity; from 94% to 73% to 65% in visible reflectance; and from 105% to 90% to 76% in precipitation.

285 The reason for the higher FSS in experiments with lower inflation may be seen in Figure 10 where forecasts from  
 286 experiment VIS+WV62+WV73 with  $\alpha=0$  and  $\alpha=0.9$  are compared. The experiment with less inflation (left panel)  
 287 had a higher FSS and less spurious convection than the experiment with inflation (right panel), as indicated by the

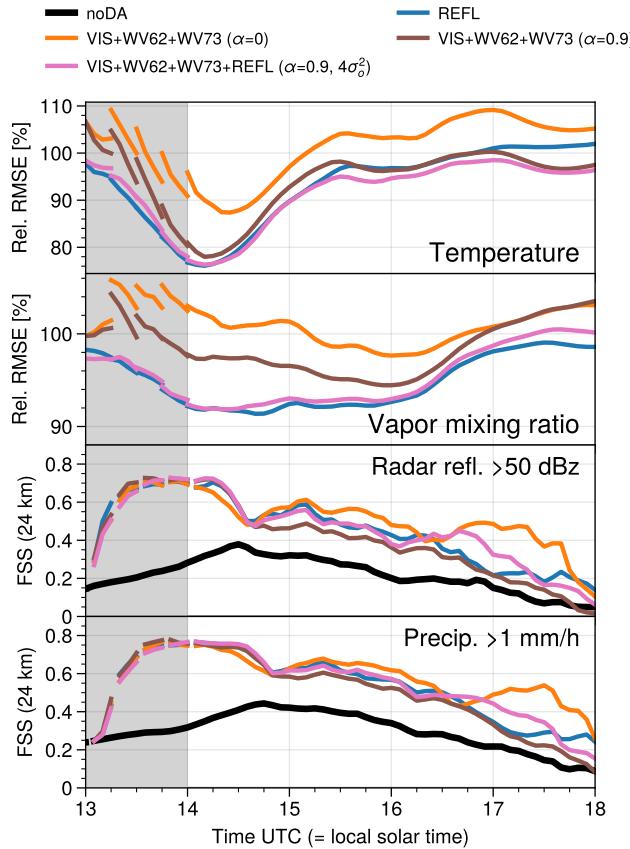


**FIGURE 6** Cloud analysis verification. Vertical profile of MAE reduction (posterior-prior, Equation 4) of cloud water and ice concentration when assimilating a combination of three satellite channels and radar reflectivity. Categories as defined in Table 2. Negative values indicate error reduction. Shading indicates standard error.

288 lower forecast probabilities (black) in regions where no rain <1 mm/h occurred. Additionally, inflation improved the  
 289 ensemble consistency (Anderson, 1997) in rank histograms (not shown) and is advisable to apply in order to achieve  
 290 accurate temperature, humidity and wind fields. However, additional inflation also led to spurious convection, while  
 291 reduced inflation improved the FSS of precipitation and cloudiness.

#### 292 The order of assimilation

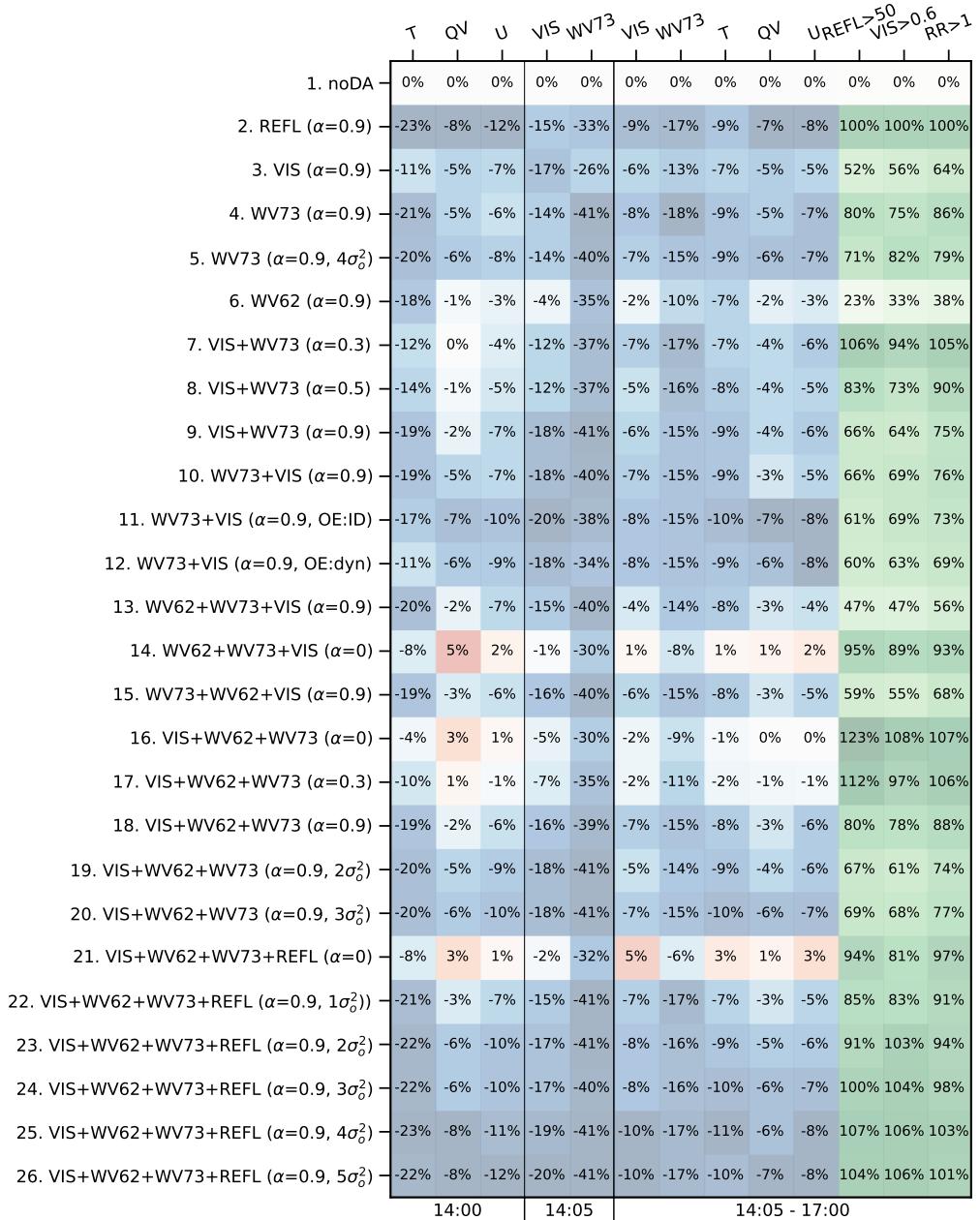
293 Changing the order of assimilation from visible-infrared to infrared-visible only caused small differences in analysis and  
 294 forecast errors (Figure 8, lines 9, 10; experiments VIS+WV73 and WV73+VIS,  $\alpha=0.9$ ). Although the analysis RMSE



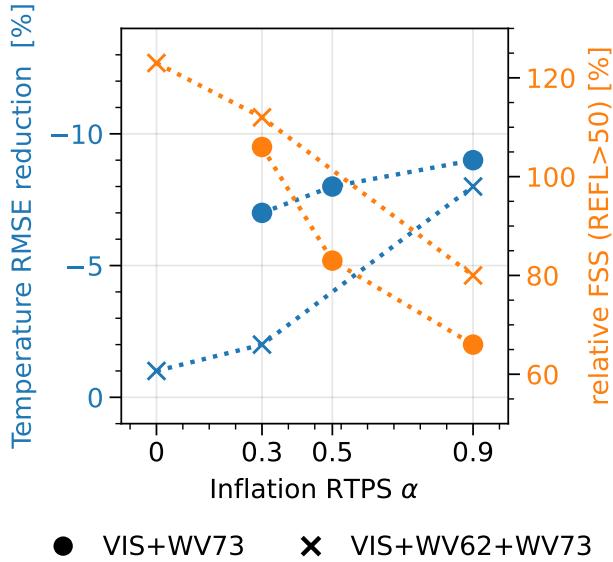
**FIGURE 7** Timeseries of forecast scores for the combined assimilation of three satellite channels and radar reflectivity. Upper panels: RMSE of visible reflectance and 7.3 µm BT. Lower panels: FSS of radar reflectivity > 50 dBz and precipitation 1 mm/h.

295 of humidity was lower when infrared observations were assimilated before visible observations (5% error reduction  
 296 in experiment WV73+VIS instead of 2% in VIS+WV73), differences in forecast errors of temperature, humidity, wind,  
 297 radar reflectivity, and precipitation were negligible.

298 Changing the order of assimilating VIS, WV62, and WV73 (Figure 8, lines 13, 15, 18;  $\alpha=0.9$ ) caused similarly small  
 299 differences (<1 %pt. difference) in the analysis and forecast errors of temperature, vapour, wind, 7.3µm BT and the  
 300 analysis errors of visible reflectance. Unsurprisingly, the forecast error reduction of visible reflectance was slightly  
 301 increased from 4% to 7% when VIS was assimilated first. Large differences, however, were caused in the FSS scores,  
 302 with the relative FSS of radar reflectivity increasing from 47% (line 13) and 59% (line 15) to 80% (line 18) when VIS was  
 303 assimilated first. The same result could be seen in experiments without inflation ( $\alpha=0$ , Figure 8, lines 14, 16). There,  
 304 assimilating visible before infrared increased the relative FSS of radar reflectivity from 95% to 123%. Closer inspection  
 305 revealed that the visible-first experiment VIS+WV62+WV73 led to a more accurate analysis of water and ice clouds  
 306 (Figure 11) while assimilating infrared-first introduced errors in the humidity below 2 km (Figure 12). This suggests



**FIGURE 8** Scorecard of experiments with error reductions compared to noDA (blue/red) and relative FSS (green). Observation types are listed in the order of assimilation.  $\alpha$ : strength of RTPS inflation, OE: observation error variance inflation factor, ID: innovation diagnostics, dyn: following Harnisch et al. (2016), T: Temperature, QV: Water vapour, U: Horizontal wind, VIS: Visible reflectance 0.6  $\mu\text{m}$ , WV62: BT 6.2  $\mu\text{m}$ , WV73: BT 7.3  $\mu\text{m}$ , REFL>50: 3D radar reflectivity > 50 dBz, RR>1: Precipitation > 1 mm/h, Columns 1-10: Error reduction compared to noDA, Columns 11-13: Relative FSS, Error: CRPS for VIS and WV73, RMSE for T and QV.



**FIGURE 9** The effect of background error inflation (RTPS  $\alpha$ ) on 3-h forecast temperature RMSE (blue) and relative FSS of REFL>50 dBz (orange).

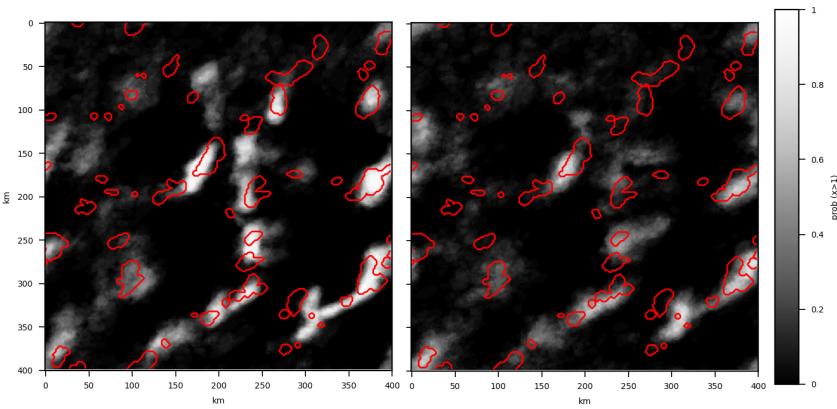
307 that assimilating visible reflectance before infrared radiance can improve cloud analysis and forecasts of precipitation  
 308 and cloudiness. Due to its insensitivity to humidity, the visible channel is only influenced by clouds (and the surface),  
 309 which allows us to correct cloud errors first before updating on other fields. Due to its ability to see through thin ice  
 310 clouds, the visible channel can ascertain the (non)-existence of clouds below the ice cloud and consequently reduce  
 311 the spread in a region to which the infrared channel may be blind.

### 312 Inflation of the observation error

313 Here, we explore assigned observation error inflation as a tuning parameter. Although non-inflated observation errors  
 314 yielded the best results with fewer observations (Kugler et al., 2023), assimilating multiple observation types multiplies  
 315 the number of observations, which can make the analysis prone to overfitting.

316 First, we assigned 1-5 times the instrument error as observation error variance in the experiment VIS+WV62+WV73+REFL  
 317 (Figure 8, lines 22-26). Increasing the assigned error decreases the weight given to an observation in the assimilation,  
 318 which would lead to larger analysis errors if the observation operator were linear. In our case, however, the inflation  
 319 improved the visible analysis while having a negligible impact on the infrared analysis. Inflating the observation error  
 320 variance by a factor 4 (Figure 8, lines 22/25) improved the forecast CRPS of visible reflectance from -8% to -10%  
 321 (compared to noDA), improved the RMSE of temperature from -7% to -11% and the RMSE of humidity from -3% to  
 322 -6%. The relative FSS was improved from 85% to 107% in radar reflectivity, from 83% to 106% in visible reflectance  
 323 and from 91% to 101% in precipitation.

324 Second, we tested an experiment with optimised assigned observation errors as derived by the innovation di-  
 325 agnostic (Equation 1) for experiment WV73+VIS, abbreviated OE:ID (Figure 8, line 11, compared to line 10). The



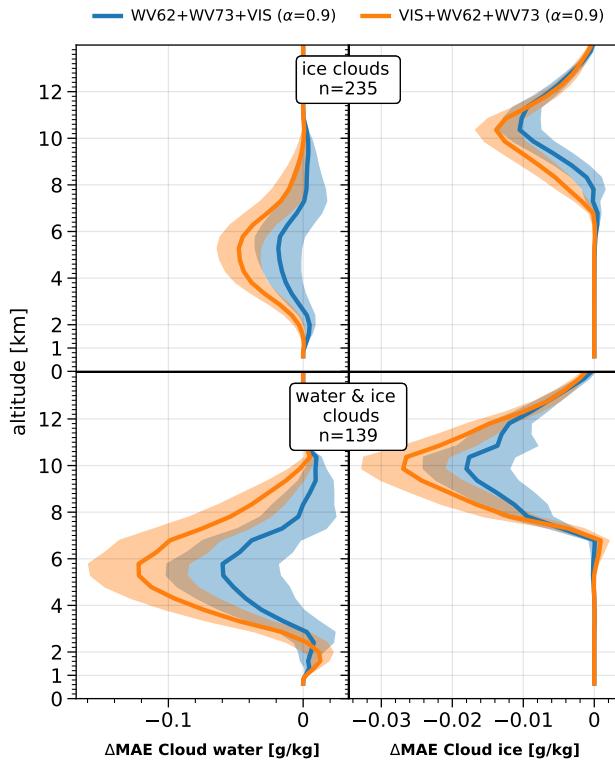
**FIGURE 10** Comparison of predicted probabilities of precipitation  $>1 \text{ mm/h}$  without (left) and with (right) background error inflation ( $\alpha=0.9$ ) for the experiment VIS+WV62+WV73. Red contours show the true regions of precipitation  $>1 \text{ mm/h}$ .

326 diagnosed and assigned error  $\sigma_o$  was approximately 0.06 for visible reflectance and 4.5 K for 7.3  $\mu\text{m}$  infrared BT. The  
 327 inflation slightly worsened the FSS of radar reflectivity (and precipitation) from 66 to 61% (76 to 73%) and had a  
 328 substantial impact on the RMSE of humidity (and wind), which improved from -3% to -7% (-5% to -8%) compared to  
 329 noDA. In observation-space, there was only a small impact on the CRPS of visible reflectance and 7.3  $\mu\text{m}$  BT (<2%pt).  
 330 This may be surprising, given the 20 times inflated error variance for infrared BT, but it may be explained by the  
 331 nonlinearity of the observation operator and the 5-minute model response to the increments. This suggests that the  
 332 observation weight  $\sigma_b^2 / (\sigma_b^2 + \sigma_o^2)$  is not necessarily a helpful metric when assimilating nonlinear observations. The  
 333 optimised constant assigned observation errors mostly improved the forecasts; however, not in all metrics.

334 Finally, we tested a dynamic observation error model, abbreviated OE:dyn in the experiment WV73+VIS (Figure  
 335 8, line 12 compared to line 10). The method of Harnisch et al. (2016) inflates the assigned observation errors for  
 336 infrared BT depending on cloud impact. Although some variables were less accurate in the analysis, the forecast  
 337 was generally better (in terms of RMSE of humidity and wind) or slightly worse (in terms of the FSS). The 7.3  $\mu\text{m}$  BT  
 338 was less accurate in the analysis (with 34% instead of 40% CRPS error reduction compared to noDA), but the 3-h  
 339 average forecast CRPS remained 15% error reduction. The RMSE of temperature was worse at analysis time (11%  
 340 error reduction compared to 19% without inflation) but better in the forecast (8% error reduction instead of 7%).  
 341 The strongest improvement was found in humidity (from 3% to 6% error reduction) and wind (from 5% to 8%). In  
 342 terms of relative FSS, the forecasts were negatively affected from 66 to 60% for radar reflectivity, from 69 to 63%  
 343 for cloudiness, and from 76 to 69% for precipitation. Similarly to the optimised constant assigned observation errors,  
 344 forecasts of certain variables can be improved by applying a dynamic observation error model. Improvements can,  
 345 however, not be expected in all metrics.

#### 346 Impact in the "warm-bubble" case

347 To test if our findings also hold in a second weather situation, we assimilated the combination in a scenario with  
 348 isolated convection. Figure 13 shows that the additional assimilation of visible and infrared channels in the warm-  
 349 bubble case reduced the FSS in the first forecast hour and increased the FSS after 2.5 hours. Satellite channels did

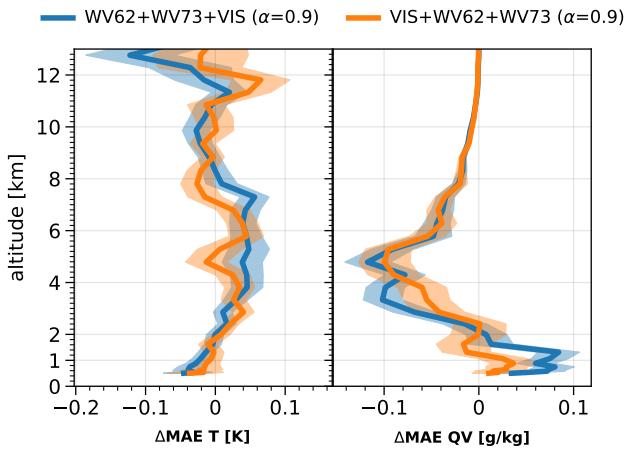


**FIGURE 11** Vertical profile of MAE reduction (Equation 4) of 7-layer vertically averaged cloud water and ice concentration. Negative values indicate error reduction. Shading indicates standard error.

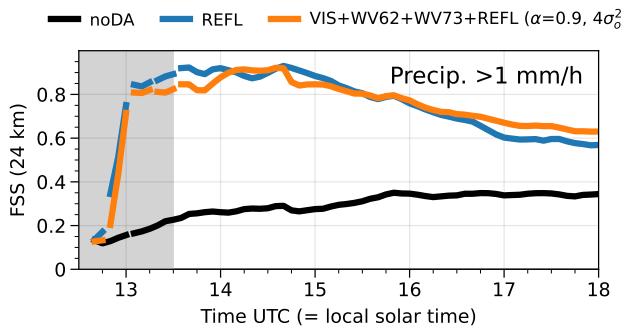
not substantially improve the spatial skill of precipitation forecasts. This might be explained by the fact that the first-guess errors were in the location and strength of the warm-bubble, which probably was well captured by the radar observations in the first place.

## 5 | CONCLUSIONS

This study extends our previous *observing-system simulation experiment* (OSSE) from the assimilation of separate assimilation to the combined assimilation of visible, infrared, and radar observations. The OSSE's nature run is taken from Kugler et al. (2023), which used a 2-km grid WRF model to simulate scattered deep moist convection. The nature run features a challenging situation for data assimilation with a highly unstable convective environment, an uncertain location of emerging deep convection, and uncertain timing of the onset of deep convection. Furthermore, we assimilate remote sensing observations with nonlinear observation operators and with ambiguous information, as these observations only indirectly observe the radiative signal of deep convection. High visible reflection, for example, can be caused by shallow clouds or cumulonimbus and low brightness temperature by high cirrus or cumulonimbus. Furthermore, rapid error growth can quickly diminish the error reduction from data assimilation. This creates one of the most



**FIGURE 12** Vertical profile of MAE reduction (Equation 4) of temperature and humidity, average over 256 gridpoints at 14 UTC. Negative values indicate error reduction. Shading indicates standard error.



**FIGURE 13** Timeseries of FSS for precipitation  $>1 \text{ mm/h}$  in the warm-bubble case.

363 difficult weather situations to predict. To avoid systematic errors and to focus on the particular characteristics of the  
 364 observation types, we used the same model and observation operators for the nature run and the forecast ensemble.  
 365 The ensemble adjustment Kalman filter assimilated visible reflectance ( $0.6 \mu\text{m}$ ), infrared brightness temperature ( $6.2$   
 366 and  $7.3 \mu\text{m}$ ) and radar ( $10 \text{ cm}$ ) reflectivity observations.

### 367 | Main results

368 Our findings can be summarised in three points:

#### 369 **1) Combined visible and infrared assimilation mitigates ambiguity**

370 The combined assimilation of visible and infrared observations mitigates the ambiguity of a single observation type  
 371 and improves the analysis of clouds as shown in section 4.1. Infrared channels can not distinguish between deep

372 convective clouds with a high cloud top (water clouds with ice clouds above) and elevated thin cirrus clouds (only ice  
373 clouds). Assimilating only infrared channels can therefore lead to erroneous water clouds through ensemble corre-  
374 lations. Assimilating visible reflectance in addition to the infrared channel can remove erroneous water clouds in the  
375 prior as visible reflectance is only weakly sensitive to ice clouds and can see through ice clouds in most instances.  
376 Moreover, the introduction of erroneous water clouds can be avoided through the additional assimilation of visible  
377 information.

378 **2) Forecast impact of the combined visible and infrared assimilation**

379 The combined assimilation of visible and infrared observations also improved forecasts of precipitation and cloudi-  
380 ness. The relative FSS for radar reflectivity  $>50 \text{ dBz}$  of the combined satellite-assimilation experiment (VIS+WV73,  
381  $\alpha=0.3$ ) reached 106% compared to the experiment that only assimilated radar (REFL), although satellite channels did  
382 not provide as much vertical resolution as radar. The errors of temperature, humidity, and wind stayed about as large  
383 as in the experiment with only one observation type.

384

385 **3) Combined assimilation of satellite and radar data**

386 Assimilating satellite channels in addition to radar reflectivity can improve forecasts of cloudiness and precipitation.  
387 Experiments which assimilated a combination of visible, infrared and radar observations had better forecasts of strong  
388 convection (FSS of radar reflectivity  $>50 \text{ dBz}$  improved to 107% compared to the experiment that only assimilated  
389 radar), cloudiness (FSS of visible reflectance  $>0.6$  improved to 106%), and light precipitation (RR  $>1 \text{ mm/h}$  improved  
390 to 103%) as indicated in Figure 8. Moreover, it was shown that the benefit of additional observations depends on  
391 background error and observation error inflation.

392

393 **Discussion**

394 Compared to Kugler et al. (2023), this follow-up study shows that the combined assimilation of visible and infrared  
395 radiances is key to achieving forecast impacts that are comparable to those from radar assimilation in a difficult-to-  
396 predict weather situation such as the "random" case of this study.

397 Our results confirm findings of Schröttle et al. (2020) using a different NWP model and assimilation algorithm.  
398 Compared to Schröttle et al. (2020), we found a higher impact of visible reflectance, which is probably related to a  
399 highly inflated assigned observation error in Schröttle et al. (2020). Moreover our results confirm their rough estimate  
400 of a comparable magnitude of potential impact of satellite and radar observations in our considered weather situation.  
401 The impact may differ for different weather situations, but satellite observations have unique advantages in case of  
402 non-precipitating clouds or in the initial stages of deep convection compared to radar. Additionally, satellites provide  
403 regularly gridded observations where radars are not available (e.g. over oceans and in less developed countries) and  
404 regardless of orographic shadowing that can limit radar coverage in complex terrain.

405 Scheck et al. (2020) showed the beneficial impact of assimilating visible observations despite observation operator  
406 nonlinearity and ambiguous information that limit the efficacy of the assimilation. As we have seen in this paper, the  
407 ambiguity can be remedied by assimilating complementary satellite observations. However, it should be noted that  
408 real-case simulations are also facing additional challenges as for example systematic model errors (Geiss et al., 2021),  
409 systematic operator errors (Scheck et al., 2018) or correlated observation errors (Okamoto et al., 2019; Hu et al., 2023)  
410 that were excluded in our study.

411 Future research should investigate the effectiveness of the combined assimilation in operational NWP centres

412 and using other assimilation algorithms such as variational or hybrid systems. Benefits can be especially expected in  
413 regions with gaps in radar coverage, such as large lakes, coastal regions and mountainous areas.

414 **Acknowledgements**

415 We thank Jeff Anderson and the Data Assimilation Research Section at UCAR/NCAR Boulder for their continued  
416 support.

417 **Data & Software**

418 Experiments were conducted using the python package available at [github.com/lkugler/DART-WRF](https://github.com/lkugler/DART-WRF). DART v10.8.3,  
419 RTTOV v13.3, WRF v4.3. Python packages: xarray (Hoyer and Hamman, 2017), metpy, matplotlib, dask, proplot.  
420 Supplementary data and code are published at <https://doi.org/10.5281/zenodo.10683151>.

421 **References**

- 422 Jeffrey Anderson, Tim Hoar, Kevin Raeder, Hui Liu, Nancy Collins, Ryan Torn, and Avelino Avellano. The Data Assimilation  
423 Research Testbed: A Community Facility. *Bulletin of the American Meteorological Society*, 90(9):1283–1296, September  
424 2009. ISSN 0003-0007, 1520-0477. doi: 10.1175/2009BAMS2618.1. URL [https://journals.ametsoc.org/view/journals/bams/90/9/2009bams2618\\_1.xml](https://journals.ametsoc.org/view/journals/bams/90/9/2009bams2618_1.xml). Publisher: American Meteorological Society Section: Bulletin of the American  
425 Meteorological Society.
- 426 Jeffrey L. Anderson. The Impact of Dynamical Constraints on the Selection of Initial Conditions for Ensemble Predictions:  
427 Low-Order Perfect Model Results. *Monthly Weather Review*, 125(11):2969–2983, November 1997. ISSN 1520-0493,  
428 0027-0644. doi: 10.1175/1520-0493(1997)125<2969:TIODCO>2.0.CO;2. URL [https://journals.ametsoc.org/view/journals/mwre/125/11/1520-0493\\_1997\\_125\\_2969\\_tiodco\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/125/11/1520-0493_1997_125_2969_tiodco_2.0.co_2.xml). Publisher: American Meteorological Society  
429 Section: Monthly Weather Review.
- 430 Jeffrey L Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly weather review*, 129(12):2884–  
431 2903, 2001. doi: [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- 432 Jeffrey L Anderson. Localization and sampling error correction in ensemble Kalman filter data assimilation. *Monthly Weather  
433 Review*, 140(7):2359–2371, 2012.
- 434 Kevin Bachmann, Christian Keil, George C Craig, Martin Weissmann, and Christian A Welzbacher. Predictability of deep  
435 convection in idealized and operational forecasts: Effects of radar data assimilation, orography, and synoptic weather  
436 regime. *Monthly Weather Review*, 148(1):63–81, 2020.
- 437 Niels Bormann, Massimo Bonavita, Rossana Dragani, Reima Eresmaa, Marco Matricardi, and Anthony McNally. Enhancing  
438 the impact of IASI observations through an updated observation-error covariance matrix. *Quarterly Journal of the Royal  
439 Meteorological Society*, 142(697):1767–1780, 2016. ISSN 1477-870X. doi: 10.1002/qj.2774. URL <http://onlinelibrary.wiley.com/doi/10.1002/qj.2774>.
- 440 Man-Yau Chan, Fuqing Zhang, Xingchao Chen, and L. Ruby Leung. Potential Impacts of Assimilating All-Sky Satellite Infrared  
441 Radiances on Convection-Permitting Analysis and Prediction of Tropical Convection. *Monthly Weather Review*, 148(8):  
442 3203–3224, July 2020. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-19-0343.1. URL <https://journals.ametsoc.org/view/journals/mwre/148/8/mwrD190343.xml>. Publisher: American Meteorological Society Section: Monthly Weather  
443 Review.
- 444 Ming-Dah Chou, Kyu-Tae Lee, Si-Chee Tsay, and Qiang Fu. Parameterization for Cloud Longwave Scattering for Use in At-  
445mospheric Models. *Journal of Climate*, 12(1):159–169, January 1999. ISSN 0894-8755, 1520-0442. doi: 10.1175/1520-0442-  
446 (1999)012<0159:PFCLSF>2.0.CO;2. URL [https://journals.ametsoc.org/view/journals/clim/12/1/1520-0442-1999\\_012\\_0159\\_pfclsf\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/clim/12/1/1520-0442-1999_012_0159_pfclsf_2.0.co_2.xml). Publisher: American Meteorological Society Section: Journal of Climate.

- 452 Frédéric Fabry. For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part  
453 II: On the Observation Signal from Different Sensors. *Monthly Weather Review*, 138(1):256–264, January 2010. ISSN  
454 1520-0493, 0027-0644. doi: 10.1175/2009MWR2884.1. URL <https://journals.ametsoc.org/view/journals/mwre/138/1/2009mwre2884.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- 456 Stefan Geiss, Leonhard Scheck, Alberto de Lozar, and Martin Weissmann. Understanding the model representation of clouds  
457 based on visible and infrared satellite observations. *Atmospheric Chemistry and Physics*, 21(16):12273–12290, August  
458 2021. ISSN 1680-7316. doi: 10.5194/acp-21-12273-2021. URL <https://acp.copernicus.org/articles/21/12273/2021/>.  
459 Publisher: Copernicus GmbH.
- 460 Nils Gustafsson, Tijana Janjić, Christoph Schraff, Daniel Leuenberger, Martin Weissmann, Hendrik Reich, Pierre Brousseau,  
461 Thibaut Montmerle, Eric Wattrelot, Antonín Bučánek, and others. Survey of data assimilation methods for convective-  
462 scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144  
463 (713):1218–1256, 2018.
- 464 F. Harnisch, M. Weissmann, and Á Periénez. Error model for the assimilation of cloud-affected infrared satellite observations in  
465 an ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 142(697):1797–1808, 2016.  
466 ISSN 1477-870X. doi: 10.1002/qj.2776. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2776>.
- 467 Takumi Honda, Takemasa Miyoshi, Guo-Yuan Lien, Seiya Nishizawa, Ryuji Yoshida, Sachio A. Adachi, Koji Terasaki, Kozo  
468 Okamoto, Hirofumi Tomita, and Kotaro Bessho. Assimilating All-Sky Himawari-8 Satellite Infrared Radiances: A Case of  
469 Typhoon Soudelor (2015). *Monthly Weather Review*, 146(1):213–229, January 2018. ISSN 1520-0493, 0027-0644. doi:  
470 10.1175/MWR-D-16-0357.1. URL <https://journals.ametsoc.org/view/journals/mwre/146/1/mwr-d-16-0357.1.xml>.  
471 Publisher: American Meteorological Society Section: Monthly Weather Review.
- 472 Stephan Hoyer and Joseph J. Hamman. xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*,  
473 2017. ISSN 2049-9647. doi: 10.5334/jors.148. URL <https://opensky.ucar.edu/islandora/object/articles%3A21998/>.
- 474 Guannan Hu, Sarah L. Dance, Ross N. Bannister, Hristo G. Chipilski, Oliver Guillet, Bruce Macpherson, Martin Weissmann, and  
475 Nusrat Yussouf. Progress, challenges, and future steps in data assimilation for convection-permitting numerical weather  
476 prediction: Report on the virtual meeting held on 10 and 12 November 2021. *Atmospheric Science Letters*, 24(1):e1130,  
477 2023. ISSN 1530-261X. doi: 10.1002/asl.1130. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/asl.1130>.
- 478 Aaron Johnson, Xuguang Wang, and Thomas Jones. Impacts of Assimilating GOES-16 ABI Channels 9 and 10 Clear Air and  
479 Cloudy Radiance Observations With Additive Inflation and Adaptive Observation Error in GSI-EnKF for a Case of Rapidly  
480 Evolving Severe Supercells. *Journal of Geophysical Research: Atmospheres*, 127(11):e2021JD036157, 2022. ISSN 2169-  
481 8996. doi: 10.1029/2021JD036157. URL <http://onlinelibrary.wiley.com/doi/abs/10.1029/2021JD036157>.
- 482 Thomas A. Jones, David Stensrud, Louis Wicker, Patrick Minnis, and Rabindra Palikonda. Simultaneous Radar and Satellite Data  
483 Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011. *Monthly Weather Review*, 143  
484 (1):165–194, January 2015. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-14-00180.1. URL <https://journals.ametsoc.org/view/journals/mwre/143/1/mwr-d-14-00180.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- 487 Thomas A Jones, Kent Knopfmeier, Dustan Wheatley, Gerald Creager, Patrick Minnis, and Rabindra Palikonda. Storm-scale  
488 data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II: Combined radar  
489 and satellite data experiments. *Weather and Forecasting*, 31(1):297–327, 2016.
- 490 Thomas A. Jones, Patrick Skinner, Nusrat Yussouf, Kent Knopfmeier, Anthony Reinhart, Xuguang Wang, Kristopher Bedka,  
491 William Smith, and Rabindra Palikonda. Assimilation of GOES-16 Radiances and Retrievals into the Warn-on-Forecast  
492 System. *Monthly Weather Review*, 148(5):1829–1859, May 2020. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-  
493 19-0379.1. URL [https://journals.ametsoc.org/view/journals/mwre/148/5/mwre148/5\\_mwre-d-19-0379.1.xml](https://journals.ametsoc.org/view/journals/mwre/148/5/mwre148/5_mwre-d-19-0379.1.xml). Publisher: American  
494 Meteorological Society Section: Monthly Weather Review.

- 495 Lukas Kugler, Jeffrey L. Anderson, and Martin Weissmann. Potential impact of all-sky assimilation of visible and infrared  
496 satellite observations compared with radar reflectivity for convective-scale numerical weather prediction. *Quarterly*  
497 *Journal of the Royal Meteorological Society*, 149(757):3623–3644, 2023. ISSN 1477-870X. doi: 10.1002/qj.4577. URL  
498 <http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.4577>.
- 499 Tobias Necker, Martin Weissmann, Yvonne Ruckstuhl, Jeffrey Anderson, and Takemasa Miyoshi. Sampling error correction  
500 evaluated using a convective-scale 1000-member ensemble. *Monthly Weather Review*, 148(3):1229–1249, 2020.
- 501 Tobias Necker, Ludwig Wolfgruber, Lukas Kugler, Martin Weissmann, Manfred Dorninger, and Stefano Serafin. The frac-  
502 tions skill score for ensemble forecast verification. 2023. URL [https://scholar.google.com/citations?view\\_op=view\\_citation&hl=en&user=ikmFqLIAAAJ&sortby=pubdate&citation\\_for\\_view=ikmFqLIAAAJ:\\_FxGoFyzp5QC](https://scholar.google.com/citations?view_op=view_ citation&hl=en&user=ikmFqLIAAAJ&sortby=pubdate&citation_for_view=ikmFqLIAAAJ:_FxGoFyzp5QC).
- 503
- 504 Kozo Okamoto, Yohei Sawada, and Masaru Kunii. Comparison of assimilating all-sky and clear-sky infrared radiances from  
505 Himawari-8 in a mesoscale system. *Quarterly Journal of the Royal Meteorological Society*, 145(719):745–766, 2019.  
506 ISSN 1477-870X. doi: 10.1002/qj.3463. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3463>.
- 507 Roger Saunders, James Hocking, Emma Turner, Peter Rayer, David Rundle, Pascal Brunel, Jerome Vidot, Pascale Roquet, Marco  
508 Matricardi, Alan Geer, Niels Bormann, and Cristina Lupu. An update on the RTTOV fast radiative transfer model (currently  
509 at version 12). *Geoscientific Model Development*, 11(7):2717–2737, July 2018. ISSN 1991-959X. doi: 10.5194/gmd-11-  
510 2717-2018. URL <https://gmd.copernicus.org/articles/11/2717/2018/>. Publisher: Copernicus GmbH.
- 511 Yohei Sawada, Kozo Okamoto, Masaru Kunii, and Takemasa Miyoshi. Assimilating Every-10-minute Himawari-8 Infrared  
512 Radiances to Improve Convective Predictability. *Journal of Geophysical Research: Atmospheres*, 124(5):2546–2561, 2019.  
513 ISSN 2169-8996. doi: <https://doi.org/10.1029/2018JD029643>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029643>.
- 514
- 515 Leonhard Scheck, Pascal Frèrebeau, Robert Buras-Schnell, and Bernhard Mayer. A fast radiative transfer method for the  
516 simulation of visible satellite imagery. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 175:54–67, May  
517 2016. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2016.02.008. URL <https://www.sciencedirect.com/science/article/pii/S0022407316000285>.
- 518
- 519 Leonhard Scheck, Martin Weissmann, and Bernhard Mayer. Efficient Methods to Account for Cloud-Top Inclination and  
520 Cloud Overlap in Synthetic Visible Satellite Images. *Journal of Atmospheric and Oceanic Technology*, 35(3):665–685,  
521 March 2018. ISSN 0739-0572, 1520-0426. doi: 10.1175/JTECH-D-17-0057.1. URL <https://journals.ametsoc.org/view/journals/atot/35/3/jtech-d-17-0057.1.xml>. Publisher: American Meteorological Society Section: Journal of At-  
522 mospheric and Oceanic Technology.
- 523
- 524 Leonhard Scheck, Martin Weissmann, and Liselotte Bach. Assimilating visible satellite images for convective-scale numerical  
525 weather prediction: A case-study. *Quarterly Journal of the Royal Meteorological Society*, 146(732):3165–3186, October  
526 2020. ISSN 0035-9009, 1477-870X. doi: 10.1002/qj.3840. URL <https://onlinelibrary.wiley.com/doi/10.1002/qj.3840>.
- 527
- 528 Josef Schröttle, Martin Weissmann, Leonhard Scheck, and Axel Hutt. Assimilating Visible and Infrared Radiances in Idealized  
529 Simulations of Deep Convection. *Monthly Weather Review*, 148(11):4357–4375, November 2020. ISSN 1520-0493,  
530 0027-0644. doi: 10.1175/MWR-D-20-0002.1. URL <https://journals.ametsoc.org/view/journals/mwre/148/11/MWR-D-20-0002.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- 531
- 532 Craig S. Schwartz, John S. Kain, Steven J. Weiss, Ming Xue, David R. Bright, Fanyou Kong, Kevin W. Thomas, Jason J. Levit,  
533 Michael C. Coniglio, and Matthew S. Wandishin. Toward Improved Convection-Allowing Ensembles: Model Physics Sen-  
534 sitivities and Optimizing Probabilistic Guidance with Small Ensemble Membership. *Weather and Forecasting*, 25(1):263–  
535 280, February 2010. ISSN 1520-0434, 0882-8156. doi: 10.1175/2009WAF2222267.1. URL [https://journals.ametsoc.org/view/journals/wefo/25/1/2009waf2222267\\_1.xml](https://journals.ametsoc.org/view/journals/wefo/25/1/2009waf2222267_1.xml). Publisher: American Meteorological Society Section: Weather  
536 and Forecasting.
- 537

- 538 C. Skamarock, B. Klemp, Jimy Dudhia, O. Gill, Zhiqian Liu, Judith Berner, Wei Wang, G. Powers, G. Duda, Dale Barker, and  
539 Xiang-yu Huang. A Description of the Advanced Research WRF Model Version 4.3. Technical report, NCAR/UCAR, 2021.  
540 URL <http://dx.doi.org/10.5065/1dfh-6p97>.
- 541 Daniel S Wilks. *Statistical methods in the atmospheric sciences*, volume 100 of *International geophysics series*. Academic  
542 press, 3rd edition, 2011.
- 543 Yunji Zhang, David J. Stensrud, and Fuqing Zhang. Simultaneous Assimilation of Radar and All-Sky Satellite Infrared Radiance  
544 Observations for Convection-Allowing Ensemble Analysis and Prediction of Severe Thunderstorms. *Monthly Weather  
545 Review*, 147(12):4389–4409, December 2019. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-19-0163.1. URL  
546 <https://journals.ametsoc.org/view/journals/mwre/147/12/mwr-d-19-0163.1.xml>. Publisher: American Meteorologi-  
547 cal Society Section: Monthly Weather Review.
- 548 Yunji Zhang, Eugene E. Clothiaux, and David J. Stensrud. Correlation Structures between Satellite All-Sky Infrared Brightness  
549 Temperatures and the Atmospheric State at Storm Scales. *Advances in Atmospheric Sciences*, 39(5):714–732, May 2022.  
550 ISSN 1861-9533. doi: 10.1007/s00376-021-0352-3. URL <https://doi.org/10.1007/s00376-021-0352-3>.
- 551 Lijian Zhu, Ming Xue, Rong Kong, and Jinzhong Min. Direct Assimilation of All-Sky GOES-R ABI Radiances in GSI EnKF for the  
552 Analysis and Forecasting of a Mesoscale Convective System. *Monthly Weather Review*, -1(aop), December 2022. ISSN  
553 1520-0493, 0027-0644. doi: 10.1175/MWR-D-21-0293.1. URL <https://journals.ametsoc.org/view/journals/mwre/aop/MWR-D-21-0293.1/MWR-D-21-0293.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather  
554 Review.

# **Chapter 5**

## **Nonlinearity effects of assimilating visible reflectance**

### **Overview**

This manuscript investigates the effects of observation operator nonlinearity on the assimilation results (the third research goal in Section 1.4). Motivated by Scheck et al. (2020), the effect of operator nonlinearity on the assimilation of visible observations with small first-guess departures is confirmed. It is tested, whether the rejection of observations with small first-guess departures can improve the assimilation of visible observations. Finally, systematic and random components in the analysis of visible and infrared observations are presented, which may be used to improve the effectiveness of assimilating visible and infrared observations in the future.

### **Own contribution**

Conceptualization, formal analysis, investigation, methodology, software, visualization, and writing. The author's contribution to this publication is estimated to be about 80%.

### **Reference**

Lukas Kugler and Martin Weissmann, 2024: The effect of nonlinear observation operators for visible and infrared radiances in ensemble data assimilation. In preparation.

# The effect of nonlinear observation operators for visible and infrared radiances in ensemble data assimilation

Lukas Kugler<sup>1</sup> | Martin Weissmann<sup>1</sup>

<sup>1</sup>Institut für Meteorologie und Geophysik  
Universität Wien, Vienna, Austria

**Correspondence**  
Lukas Kugler  
Email: lukas.kugler@univie.ac.at

**Funding information**

Data assimilation for numerical weather prediction is becoming increasingly reliant on satellite observations. Visible and infrared satellite observations are nonlinear functions of atmospheric state variables called observation operators. However, the nonlinearity of observation operators violates the linearity assumption in data assimilation, leading to sub-optimal assimilation results. This paper investigates the effect of observation operator nonlinearity for visible and infrared radiances in the Ensemble Adjustment Kalman Filter (EAKF) using observing-system simulation experiments (OSSE). First, we study the effect of nonlinearity on the assimilation of observations with small first-guess departures and show that their assimilation is beneficial. Secondly, we quantify the deviations of the ensemble mean and variance adjustment from the theoretical expectation. Deviations of linear and nonlinear adjustments can be split into systematic (average) and random (variability) components. On average, the nonlinear increments are lower than the linear increments (in absolute terms). However, in certain intervals, where the linear increment would reduce cloudiness, the nonlinear increment reverses the sign and increases cloudiness. The variance in the nonlinear posterior is never reduced by more than what would be expected with a linear observation operator. In case the prior spread is small, the variance is often increased. However, on average, the variance reduction is lower than the mean-square error reduction. Thus, the assimilation acts slightly inflationary.

## 1 | INTRODUCTION

Nonlinear observation operators are widely and successfully used in data assimilation for convective-scale numerical weather prediction (NWP) despite violating the linearity assumption in variational and ensemble assimilation methods (Gustafsson et al., 2018; Hu et al., 2023). Radar reflectivity has been assimilated for decades, however, often using a latent-heat nudging technique. Infrared radiance assimilation is a topic of ongoing research (Geer et al., 2018; Li et al., 2022) but such observations are still not operationally assimilated in operational km-scale models except at NSSL (Jones et al., 2020) and at *Deutscher Wetterdienst* (DWD). Lately, the assimilation of visible radiances has been investigated (Schröttle et al., 2020; Scheck et al., 2020; Kugler et al., 2023) and was recently introduced operationally in the km-scale model at DWD in 2023.

However, many operational centers avoid the direct assimilation of nonlinear observations: Infrared radiance assimilation with a linearized observation operator can produce detrimental prior distributions and unrealistic correlations (Zhou et al., 2023). Challenges like this may have contributed to the popularity of assimilating retrievals like cloud water paths instead of the direct assimilation of infrared radiances (Jones et al., 2015, 2016, 2020; Johnson et al., 2022). Other researchers avoid satellite data assimilation at all (Dowell et al., 2022). Instead of direct assimilation of radar reflectivity, some apply latent heat nudging (Jones and Macpherson, 1997; Weygandt et al., 2022; Leuenberger and Rossa, 2007) or assimilate retrievals (Chen et al., 2021).

The goals of this paper are to support the development of methods for an improved assimilation of nonlinear observations. In section 2, we review why the nonlinear observation operators introduce systematic deviations from the expected error reduction described by the Kalman gain and describe the model setup and data used in this study. In section 3, we investigate nonlinearity in four ways: First, we show two examples in which either a cloudy or a clear-sky observation is assimilated (single observation experiments, section 3.1). Second, we investigate a hypothesis of Scheck et al. (2020) that assimilating

observations with small first-guess (FG) departures seem to be detrimental for the analysis, but we demonstrate that this is not necessarily the case (section 3.2). Third, we investigate reasons if observations with small FG departures should be assimilated as the associated errors due to nonlinearity are large compared to the departures for these observations (section 3.2.3). Last, we analyze nonlinearity deviations of the ensemble mean and variance from the optimal posterior in observation space for visible and infrared satellite observations that are caused by observation operator nonlinearity in section 3.3.

## 2 | METHODS

In the following, we use the scalar notation of Kar speck and Anderson (2007) of the Ensemble Adjustment Kalman filter (EAKF; Anderson, 2001). The EAKF assimilates observations serially. For each observation, the EAKF adjusts the ensemble mean and variance in observation space given the assigned errors and then projects the observation-space increments to model space by using ensemble covariances.

### 2.1 | Nonlinearity deviation in observation space

The nonlinear observation operator  $\mathcal{H}$ , generates an observation

$$y^o = \mathcal{H}[x^t] + \varepsilon^o \quad (1)$$

from a model state  $x$ , adding Gaussian noise  $\varepsilon^o$ . Nonlinearity primarily refers to the relationship between an independent observed variable on a dependent variable, like liquid water path (LWP) for visible reflectance. Additional nonlinearity effects can originate from the forecast model  $x^f(t_{k+1}) = \mathcal{M}[x^f(t_k)]$ , which can cause Gaussian errors to evolve into non-Gaussian error distributions, which is, however, not the topic of this paper. Fundamentally, the optimum-interpolation ("linear posterior")

$$\overline{y_{\text{linear}}^a} = \overline{\mathcal{H}(x^b)} + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} (y^o - \overline{\mathcal{H}(x^b)}) \quad (2)$$

gives the optimal posterior given the prior, the observation, and their errors. However, the nonlinear posterior in an ensemble Kalman filter for one observation is (Kar-speck and Anderson, 2007):

$$\overline{y_{\text{nonlinear}}^a} = \overline{\mathcal{H}(x_{ij}^b + \delta x_{ij})} \quad (3)$$

with state-space increments of state variable  $x_j$  and ensemble member index  $i$

$$\delta x_{ij} = \frac{\text{Cov}(x_j^b, \mathcal{H}(x^b))}{\sigma_b^2} (y_i^a - \mathcal{H}(x_i^b)) \quad (4)$$

where the posterior in observation space is given by

$$y_i^a = \left( \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2} \right)^{-1} \left( \frac{\overline{\mathcal{H}(x^b)}}{\sigma_b^2} + \frac{y_o^a}{\sigma_o^2} \right) \quad (5)$$

$$+ \sqrt{\frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2}} \left( \mathcal{H}(x_i^b) - \overline{\mathcal{H}(x^b)} \right). \quad (6)$$

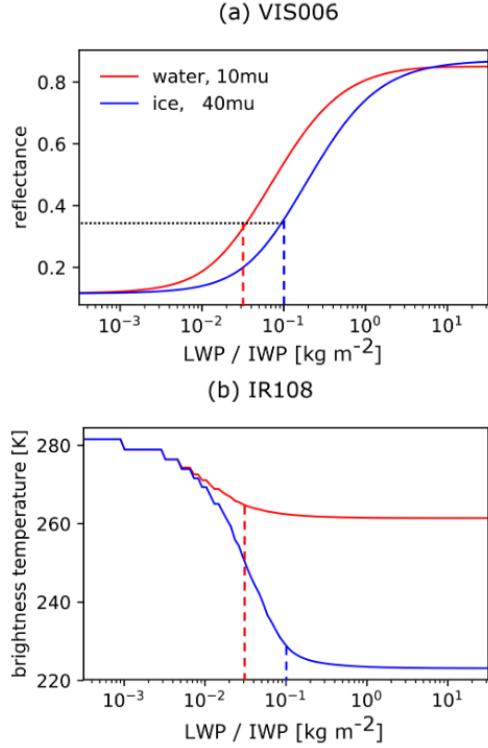
Thus, the nonlinear posterior differs from the linear posterior due to nonlinearity in the observation operator and sampling errors in the covariance. Since the linear posterior is the optimal solution in observation space, any deviation from it means that the posterior is sub-optimal.

Additionally, the ensemble variance of the linear posterior

$$\sigma_a^2 = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} \quad (7)$$

can significantly deviate from the ensemble variance of the nonlinear posterior.

As an example, suppose that  $\mathcal{H}(x)$  is the visible reflectance as a function of liquid water path as depicted in Figure 1. Then the posterior is strongly affected by an increment of  $0 \rightarrow 0.1 \text{ kg/m}^2$  of LWP, but hardly reacts to an increment  $1 \rightarrow 10 \text{ kg/m}^2$  of LWP. During the assimilation, the EAKF uses the ensemble covariance between state variables and model-equivalents of the observations  $\text{Cov}(x^b, y^b)$  to derive state variable increments from observation space increments (Anderson, 2003). In case the ensemble has little spread around a clear-sky mean, the sensitivity  $d\mathcal{H}/dx$  will be underestimated (sometimes



**FIGURE 1** Nonlinear relationship between liquid/ice water path and observed (a) visible and (b) infrared radiance. Reproduced from Geiss et al. (2021).

referred to as zero-gradient problem in variational data assimilation). If the ensemble members also sample the part of the function with larger dependency, the dependency can be overestimated. The overestimated covariance leads to an overestimated state variable increment  $\delta x$ . Applying the operator  $\mathcal{H}$ , which is sensitive to small changes in  $\delta x$  to the updated state  $x^b + \delta x$ , therefore, leads to a jump in the posterior  $\mathcal{H}(x^b + \delta x)$ , which is then well above the expected linear posterior.

To summarize, the nonlinearity deviation can be caused by the nonlinearity of the observation operator and by sampling errors of the linear regression from observation space to model space with a small sample (the ensemble).

### Definition of the nonlinearity deviation

We define a nonlinearity deviation as the difference between the linear and the nonlinear posterior ensemble mean in observation space:

$$\delta y^{NL} = \overline{y_{\text{nonlinear}}^a} - \overline{y_{\text{linear}}^a} \quad (8)$$

The nonlinearity deviation is zero  $\overline{y_{\text{nonlinear}}^a} = \overline{y_{\text{linear}}^a}$  if the observation operators are linear and if the model integration of 60 s can be neglected. In case of nonlinear operators, the regression of increments from observation space into model space introduces deviations, as this step assumes a linear relationship in the computation of the covariance (equation 4). The covariance effectively linearizes the nonlinear relationship given by pairs of state variable-observation,  $(x_{ij}, \mathcal{H}(x))$ .

Although the linear posterior is optimal if the linearity assumptions are true, it can usually not be achieved as different assumptions make the nonlinear posterior sub-optimal. The nonlinear posterior will almost always differ from the linear posterior in the case of nonlinear observation operators.

## 2.2 | Model setup and data

The data for this study was generated in an *observing-system simulation experiment* (OSSE) (Kugler et al., 2023). The weather situation features scattered deep moist convection in a highly unstable environment. The forecast model is the Weather Research and Forecasting (WRF) model in a 2-km convective-scale resolution. We evaluate the observation operator ( $\mathcal{H}$ ) nonlinearity in the all-sky assimilation of visible and infrared radiances. In order to study only the effect of the  $\mathcal{H}$ -nonlinearity, we set systematic errors to zero by assuming a perfect forecast model and no errors in the  $\mathcal{H}$ -operator. Thus, the observations were synthetically generated. Observations were assimilated into our 40-member forecast using the EAKF (Anderson, 2001) a serial deterministic square root filter, through the Data Assimilation Research Testbed (DART, Anderson et al. (2009)). The updated variables were temperature, water vapor mixing ratio, dry air mass in the column, geopotential, three wind components, cloud water

and ice mixing ratio.

The nonlinear posterior was evaluated after a 60 s forecast, similar to Scheck et al. (2020), because diagnostic cloud fractions are not available at analysis time but are computed by the model. Additionally, cloud increments can evaporate in non-saturated air. Importantly, no inflation was applied in the experiments to purely focus on the effect of the nonlinear observation operator.

Section 3.1 uses data from two single observation experiments. Observations were generated at 13 UTC from the nature run of the case "random" of Kugler et al. (2023) and assimilated with a five-member ensemble. The clear-sky observation was randomly selected. The cloudy observation was selected as the observation with the highest observed value of 961 observations of the "VIS" experiment in Kugler et al. (2023).

Sections 3.2-3.3 use data from two cycled data-denial experiments with 961 observations per assimilation time. Observations are assimilated at five times between 13 and 14 UTC. In the "control" experiment, all observations were assimilated. In the experiment "reject", only observations with a first-guess departure larger than one observation error standard deviation (0.03) were assimilated. Nature and prior were almost identical to the "random" case of Kugler et al. (2023).

In subsection 3.2.3 a data-denial experiment is repeated with only two observations, of which one is cloudy, and one is clear-sky. Both observations are approximately 10 km apart.

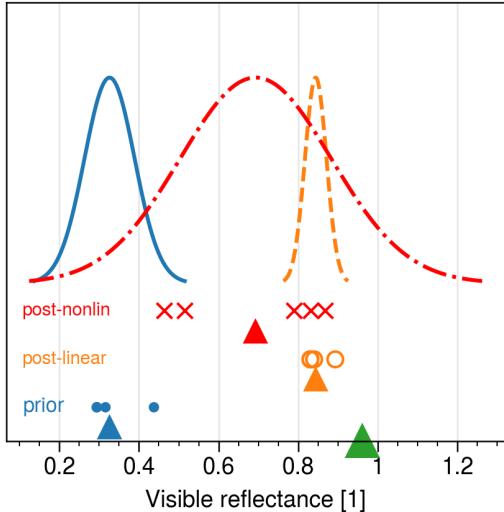
## 3 | RESULTS

### 3.1 | Single observation experiments illustrating nonlinear effects

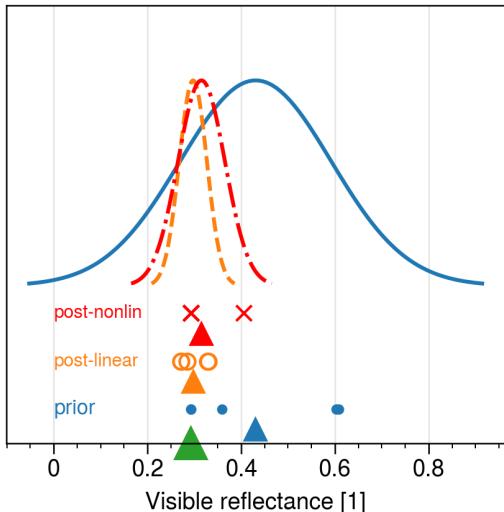
To illustrate the fundamental differences between the linear and the nonlinear analysis, we discuss two single-observation experiments.

#### Experiment 1:

We assimilate a cloudy visible satellite observation (reflectance value: 0.96) with a 5-member ensemble and with an observation error of 0.03. Figure 2a shows the



(a) Example 1 assimilating a cloudy observation



(b) Example 2 assimilating a clear-sky observation

**FIGURE 2** Two examples in which a cloudy or a clear-sky observation is assimilated with a 5-member ensemble. Shown is the observation, the prior, and the posterior ensemble in observation space with fitted PDFs for prior (solid), linear posterior (dashed), and nonlinear posterior (dash-dotted). Triangles indicate the ensemble mean or the observation on the x-axis.

observation, the prior, and the posterior ensemble in ob-

servation space with fitted PDFs. The prior has three clear-sky ensemble members (0.293), one nearly clear-sky member (0.314), and one slightly cloudy member (0.438). Thus, the prior mean is 0.326, with a spread of 0.063. Using the linear equation 2, the posterior mean is 0.84 with a spread of 0.027. Note that given the uncertainty of the prior and the observation, either the likelihood of observing such a cloud was very low or the uncertainty of the prior was heavily underestimated. Given the small ensemble, the latter is obvious.

The nonlinear posterior shows that the actual reflectance in the posterior is affected much less than the linear equation would suggest. Only three members show reflectance larger than 0.78; one member has 0.51, and one member has 0.46, resulting in an ensemble mean of 0.69 with a spread of 0.19, i.e., the posterior variance was 49 times higher than in the linear posterior. The nonlinear posterior shows much more variety in reflectance values, and thus, it seems that the nonlinearity of the observation operator mitigates the underestimation of prior variance.

## Experiment 2:

Here, we assimilate a clear-sky observation (value: 0.293) with a 5-member ensemble and an observation error of 0.03. Figure 2b illustrates the assimilation in observation space. In the prior, we had two ensemble members with clear sky (0.293), one member with 0.36, and two members close to 0.60. Thus, the prior mean is 0.431, with a spread of 0.16. The resulting linear posterior mean is 0.297, with a spread of 0.03.

In this example, the nonlinear posterior mean (0.315) is closer to the linear posterior, and the spread (0.05) is only slightly larger than expected. We conclude that in this example, the actual posterior is less affected by the nonlinearity of the observation operator. Common to both examples is that the variance of the nonlinear posterior is larger than that of the linear posterior. However, the largest deviation from the linear equations is the 49 times higher variance in example 1.

### 3.2 | Observations with small first-guess departures

#### 3.2.1 | Diagnostic results

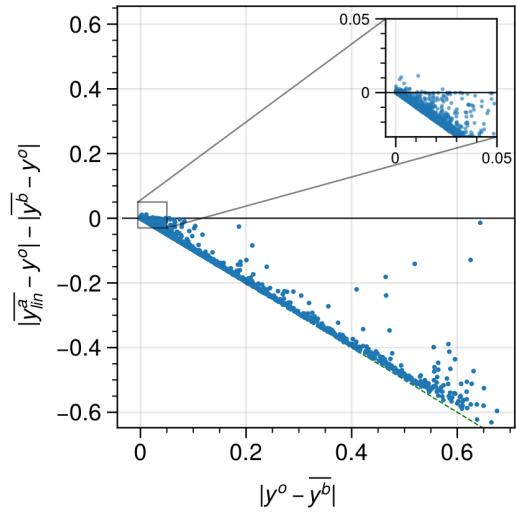
In the first study assimilating visible reflectance, Scheck et al. (2020) found that for some observations with small first-guess (FG) departures, the analysis error  $|\bar{y}^a - y^o|$  was larger than the FG error  $|\bar{y}^b - y^o|$ . Thus, it seemed that the assimilation of observations with small FG departures had a detrimental effect on the analysis. Scheck et al. (2020) concluded that "the nonlinearity of the operator thus limits the effectiveness of the visible assimilation". Therefore, one could think that rejecting those observations could improve the analysis.

To test this hypothesis, we evaluated the analysis error reduction (posterior-prior) for every observation-forecast pair, once for the linear posterior (Figure 3a) and once for the nonlinear posterior (Figure 3b). The linear posterior always reduced the error to observations. Figure 3a confirms that this is approximately the case in our experiment. However, the analysis error of the nonlinear posterior (Figure 3b) is sometimes higher than that of the prior. Especially the analysis pertaining to observations with small FG departures is often worse in the nonlinear posterior than in the prior. Consequently, it seems legitimate to reject (not assimilate) those observations with small FG departures.

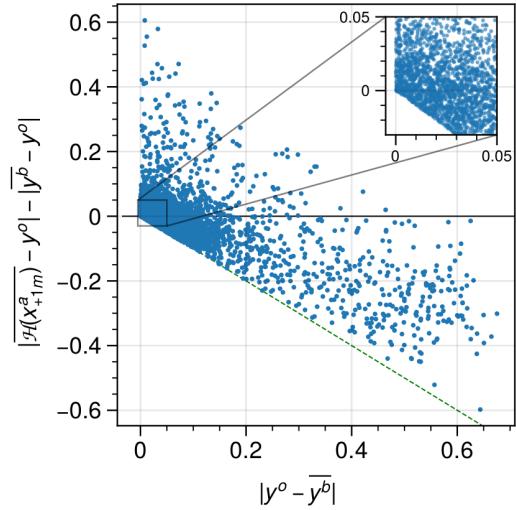
#### 3.2.2 | Cycled experiment: Impact on RMSE and bias

The diagnostic result motivated a data-denial experiments in which observations with small FG departures are rejected. In the experiment "reject", we assimilated only observations with FG departures larger than 0.03, while in the "control" experiment, we assimilated all observations.

Figure 4 compares the posterior ensemble-mean of for 104 observations with small FG departures at 13 UTC. It reveals that the posterior reflectance of the "reject" experiment was, on average, substantially higher than the posterior of the "control" experiment. Assimilating only observations with larger FG departures, thus, increased



(a) Optimal-interpolation error reduction



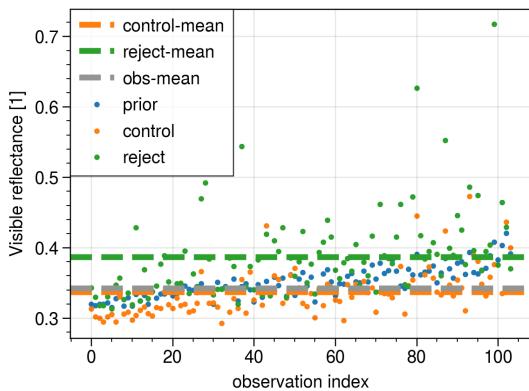
(b) Actual error reduction

**FIGURE 3** Error reduction (posterior-prior) compared to first-guess-departure for  $0.6 \mu\text{m}$  visible reflectance, both dimensionless. Orange crosses indicate that both first-guess and observation were clear-sky ( $\text{VIS} < 0.4$ ). Points on the green dashed line have 100% error reduction. Points above the 0-line mark increased the error in the posterior.

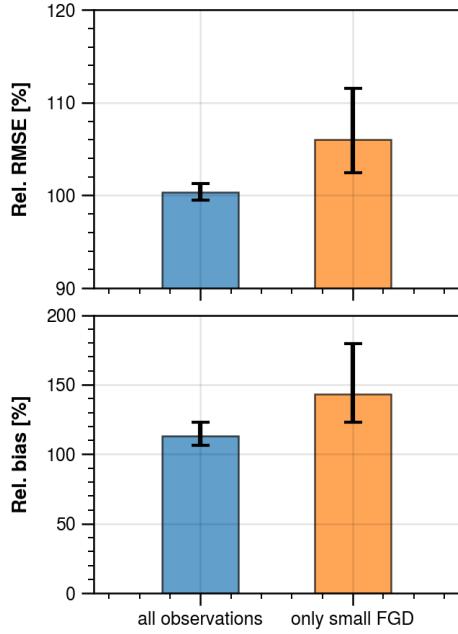
the reflectance values of these observations, on average, and "pulled the analysis from an already nearly correct prior".

While Figure 4 illustrates the effect well, it is only one snapshot in time. Thus, we also evaluate the bias and RMSE for five assimilation cycles between 13 and 14 UTC in a cycled experiment. Figure 5 shows the RMSE (upper panel) and bias (lower panel) relative to the control experiment for all observations (left bar) and for the subset with small FG departures (right bar). In the reject experiment, the bias is increased over all observations in all five assimilation cycles, and the RMSE is increased for the subset of small FG departures for all five cycles. However, when computed over all observations, the RMSE is similar in the reject and the control experiment.

We conclude that rejecting observations with small FG departures is detrimental for the analysis. Further investigation of potential reasons will follow in the subsequent section.



**FIGURE 4** The posterior ensemble mean reflectance in the "control" experiment that assimilated small FG-departures and the "reject" experiment that rejected small FG-departures. Shown are the analyses of 104 observations that had a FG-departure ( $<0.03$ ) at 1300 UTC. The posterior was evaluated 60s after analysis time.



**FIGURE 5** Error of the ensemble mean of the experiment "reject" in %, relative to the control experiment. The left column uses all data. The right column uses the subset of observations with small FG-departures ( $<0.03$ ); Bars indicate averages and whiskers indicate the range of the error over five assimilation times.

### 3.2.3 | Insights from a two-observation experiment

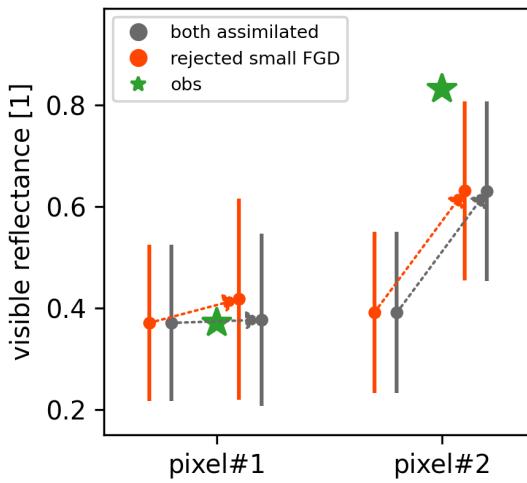
Here, we show an example of how the analysis benefits from an observation with a small first-guess departure in a two-observations-experiment.

#### Two-observations data-denial experiment

In the first experiment, we have two observations of visible reflectance: The first observes the cloud top of a deep convective cell. The second observes a cloud-free grid-box whose center is 10 km from the first observation. The prior ensemble-mean for both observations is clear-sky in most members, but cloudy in a few members ( $y^b \approx 0.4$ ). The experiment consists of a control run and a "reject" run, which does not assimilate the cloud-free observation with its small FG-departure. In the control run we assimilate the cloud-free observation first and then as-

similate the cloudy observation. In the "reject" run we do not assimilate the cloud-free observation and only assimilate the cloudy observation.

Figure 6 shows the ensemble mean and spread for the two observed grid-boxes "pixel#1" and "pixel#2". For both grid-boxes, there are three states and variances: prior, posterior-1 (after assimilating the clear-sky observation) and posterior-2 (after assimilating the cloudy observation).



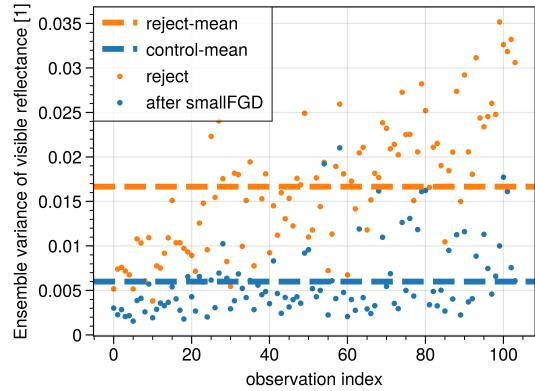
**FIGURE 6** The ensemble mean and spread of the analysis at both observation sites.

#### Spread reduction limits erroneous mean-increments

Now, we show that assimilating observations with a small FG-departure reduces the ensemble spread/variance and thereby avoids erroneous increments of the ensemble mean due to nearby observations.

Figure 7 compares the ensemble variance after assimilating (experiment "control") or rejecting (experiment "reject") observations with small FG-departure. Assimilating observations with small FG-departure reduced the variance (over the sample of small FG-departures) by a factor of three. Consequently, the variance is much higher when such observations are rejected. Then, when we assimilate close-by observations (within localization radius) with large FG-departures, these observations can modify the state (in a good or bad way). But as the prior for the

observations with small departures was quite accurate, the risk of a detrimental increment is comparably large. If observations with small FG-departures are assimilated, they reduce the ensemble variance and pin the prior at these locations so that other observations with large FG-departures can do less harm due to for example imperfect covariances. This explains why the assimilation of small FG-departures is important and why they improve the ensemble mean in Figure 4.



**FIGURE 7** Ensemble variance after assimilating observations with small FG-departures ("control") but before the assimilation of other observations, compared to the ensemble variance without assimilating observations with small FG-departures ("reject", i.e. prior variance). Both were evaluated at analysis time. Shown are the analyses of 104 observations with small FG-departures at 1300 UTC.

We have seen that the observation operator nonlinearity leads to deviations from the optimal posterior for all observations and that we can't avoid the detrimental effect for observations with small FG-departures by rejecting them. Let us now study the effect of nonlinearity on the posterior ensemble mean increment and variance reduction.

### 3.3 | Nonlinearity effects on analysis mean and variance

In this section, we quantify the effect of observation operator nonlinearity on the ensemble mean increments and

variance reduction. As shown in section 2.1, linear posterior  $\bar{y}_{\text{linear}}^a$  from Equation 2 is optimal in observation space, independent of sub-optimality in the projection to model space. The existence of a nonlinear observation operator does not change the optimality of Equation 2. Therefore, we desire that the actual (nonlinear) ensemble-mean posterior  $\bar{\mathcal{H}}(x_i^a)$  is close to the linear posterior ensemble-mean  $\bar{y}_{\text{linear}}^a$ . Similarly, the "nonlinear" posterior ensemble variance  $\text{Var}_{\text{ens}}[\bar{\mathcal{H}}(x_i^a)]$  shall be close to the "linear" variance  $\sigma_a^2 = \sigma_b^2 / (\sigma_b^2 + \sigma_o^2)$ .

### 0.6 μm visible reflectance

Before we review the deviations for visible and infrared observations, we consider the special case of a linear observation operator, a 2-m temperature observation. Figure 8a shows the nonlinear and the linear ensemble mean increment. As expected, the deviations are close to zero for the whole range. Small deviations are to be expected due to the one-minute model integration and the fact that the model state variable is potential temperature, not sensible temperature. Similarly, the deviations of the nonlinear to the linear variance adjustment are close to zero in Figure 8c.

In contrast, Figure 8b compares the increments for visible reflectance observations. It can be seen that the nonlinear increments are systematically lower than the linear increments, especially for large positive linear increments, where the observation is much larger than the prior. Similarly, for large negative linear increments, where the observation is much smaller than the prior, the nonlinear increments are smaller (in absolute terms) than the linear ones. However, for slightly negative linear increments ( $-0.2 < \Delta < .05$ ), there is a considerable number of samples showing a positive nonlinear increment, although the average of the nonlinear increments is close to zero.

The nonlinearity also introduced stochastic variability around the average deviation. Figure 8d shows the deviation of the nonlinear to the linear variance adjustment. The assimilation of visible reflectance introduced a peculiar effect, the increase of ensemble variance. Especially for small linear variance adjustments (when the prior spread was small), the nonlinear variance is often in-

creased. This effect was already illustrated in Figure 2a, where the prior was mostly clear-sky, and the assimilation yielded a much larger posterior variance than in the case of a linear operator. However, there are also cases where the nonlinear variance adjustment is close to the expectation, similar to the example in Figure 2b.

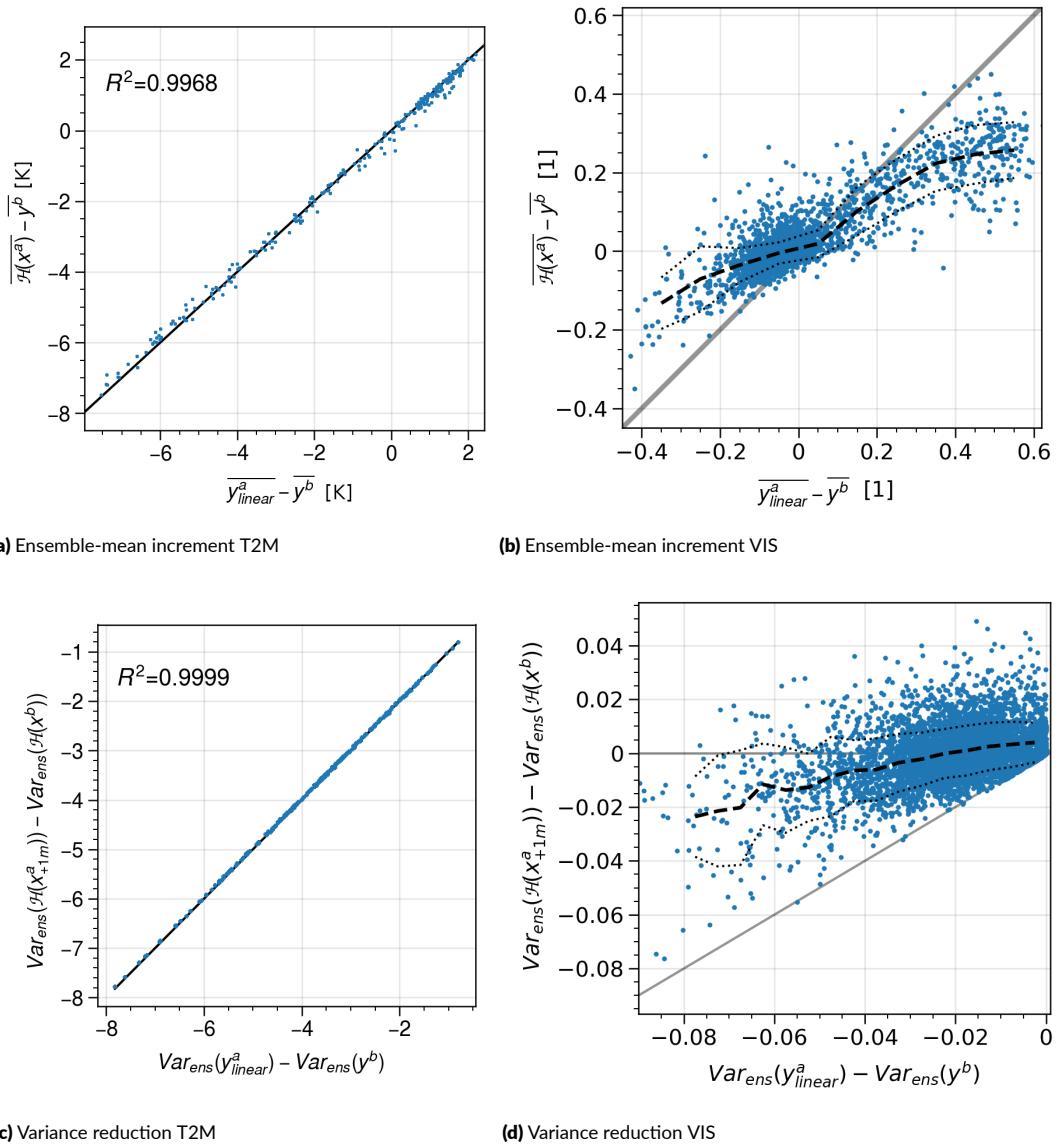
On average over all 4805 observations over five assimilation times, the mean ensemble-variance reduction was 26.9 % while the mean-square error of the ensemble mean was reduced by 40.8 %. This means that the posterior was more accurate than the ensemble variance suggested.

### 7.3 μm infrared brightness temperature

In the same way as before, we evaluate the nonlinearity deviation in Figure 9a for the ensemble mean and variance adjustment of 7.3 μm infrared brightness temperature (BT). Both the visible and infrared operators share similarities in their nonlinearity deviations. In most cases, nonlinear increments are smaller in magnitude than their linear counterparts. The deviations are smallest, on average, between -5 and +5 K. However, there are strong outliers in the range (0-10 K), where the linear increment would increase the BT (less cloudy), the nonlinear increment decreases the BT.

Figure 9b shows the deviation of nonlinear variance adjustment compared to the linear variance adjustment. In all cases, the nonlinear variance adjustment was smaller in magnitude than its expectation. Again, for small linear variance reductions, where the prior spread is small, the posterior variance was larger than the prior variance. Finally, large linear variance reductions also tend to be large in their nonlinear counterparts.

On average over 4805 observations over five assimilation times, the mean-square error of the ensemble mean was reduced by 33.1 %, while the mean ensemble-variance was reduced by 17.7 %. Thus, the posterior was improved more than the variance reduction indicated.



**FIGURE 8** Deviation of linear and nonlinear increment (posterior-prior) of the ensemble-mean when assimilating (a) 2m temperature or (b) 0.6  $\mu\text{m}$  visible reflectance. Dashed lines indicate mean and standard-deviation over bins of size 0.1. Deviation of linear and nonlinear ensemble-variance reduction when assimilating only (c) 2m temperature or (d) 0.6  $\mu\text{m}$  visible reflectance. Dashed lines indicate mean and standard-deviation over bins of size 0.005.

## 4 | CONCLUSION

This study provides insight into the systematic and random effects of assimilating observations with nonlinear observation operators, specifically visible and infrared observations. The deviations are investigated in idealized, perfect-model observing-system simulation experiments (OSSE) using a 2-km grid WRF model. Visible ( $0.6\text{ }\mu\text{m}$ ) and infrared ( $6.2$  and  $7.3\text{ }\mu\text{m}$ ) observations are synthetically generated and assimilated using the ensemble adjustment Kalman filter (EAKF). The analysis was evaluated in a summertime case with scattered deep convective storms.

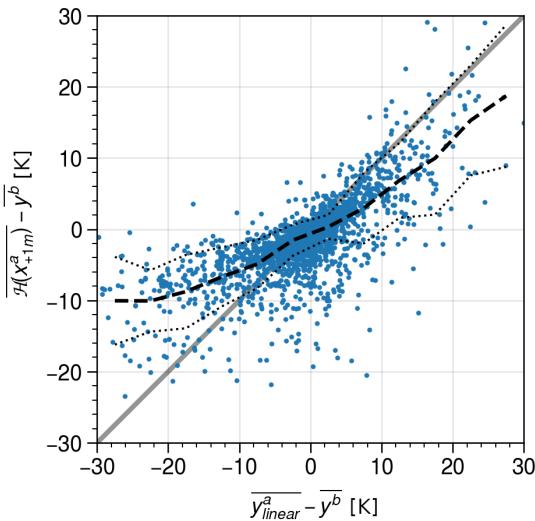
### Main findings

- 1) The nonlinearity of visible and infrared observation operators causes the posterior ensemble to deviate from what would be expected in the case of linear observation operators. Despite these deviations from linearity, the assimilation of visible and infrared observations improves the posterior ensemble mean on average. Deviations can be partitioned into systematic (average) and random (variability) components. On average, the nonlinear increments are lower than the linear increments (in absolute terms). However, in certain intervals, where the linear increment would be a decrease of reflectance (or an increase of BT), the nonlinear increment shows an increase of reflectance (or a decrease in BT). Thus, the nonlinearity reverses the sign of the increment and leads to increased cloudiness in situations where the prior already had too high reflectance (or too cold BT).
- 2) The variance in the nonlinear posterior is never reduced by more than what would be expected with a linear observation operator. Contrary to the expectation that assimilation always reduces variance, the variance can increase when the prior spread is small, as shown by the example in section 3.1. In most cases, however, the variance is reduced. Verifying the ensemble mean showed that the variance reduction is lower than the reduction of mean-square error. Thus, the assimilation acts slightly inflationary.
- 3) The assimilation of observations with small FG departures was suspected to deteriorate the analysis (Scheck et al., 2020). Although a first evaluation (section 3.2) suggested a deterioration, additional experiments showed that rejecting small first-guess departures does not reduce the overall error but increases both the bias and the mean-square error at the location of the observations with small first-guess departures. While the error at the location of observations with small departures was increased, the assimilation of them avoided that other nearby observations draw the analysis away from the accurate first guess. In summary, we therefore conclude that also observations with small first guess departures should be included in the assimilation.

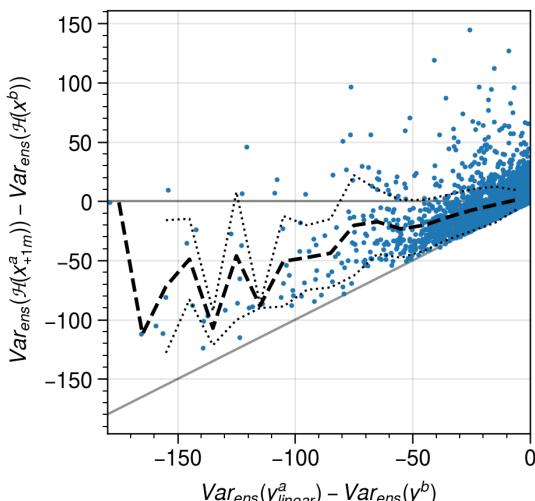
et al., 2020). Although a first evaluation (section 3.2) suggested a deterioration, additional experiments showed that rejecting small first-guess departures does not reduce the overall error but increases both the bias and the mean-square error at the location of the observations with small first-guess departures. While the error at the location of observations with small departures was increased, the assimilation of them avoided that other nearby observations draw the analysis away from the accurate first guess. In summary, we therefore conclude that also observations with small first guess departures should be included in the assimilation.

## REFERENCES

- Jeffrey Anderson, Tim Hoar, Kevin Raeder, Hui Liu, Nancy Collins, Ryan Torn, and Avelino Avellano. The Data Assimilation Research Testbed: A Community Facility. *Bulletin of the American Meteorological Society*, 90(9):1283–1296, September 2009. ISSN 0003-0007, 1520-0477. doi: 10.1175/2009BAMS2618.1. URL [https://journals.ametsoc.org/view/journals/bams/90/9/2009bams2618\\_1.xml](https://journals.ametsoc.org/view/journals/bams/90/9/2009bams2618_1.xml). Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- Jeffrey L Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly weather review*, 129(12):2884–2903, 2001. doi: [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- Jeffrey L Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003. doi: [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2).
- Haiqin Chen, Jidong Gao, Yunheng Wang, Yaodeng Chen, Tao Sun, Jacob Carlin, and Yu Zheng. Radar reflectivity data assimilation method based on background-dependent hydrometeor retrieval: Comparison with direct assimilation for real cases. *Quarterly Journal of the Royal Meteorological Society*, 147(737):2409–2428, 2021. ISSN 1477-870X. doi: 10.1002/qj.4031. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.4031>.
- David C. Dowell, Curtis R. Alexander, Eric P. James, Stephen S. Weygandt, Stanley G. Benjamin, Geoffrey S. Manikin, Benjamin T. Blake, John M. Brown, Joseph B. Olson, Ming Hu, Tatiana G. Smirnova, Terra Ladwig, Jaymes S. Kenyon, Ravan Ahmadov, David D.



(a) Ensemble-mean increment



(b) Variance reduction

**FIGURE 9** (a) Deviation of the nonlinear posterior ensemble-mean from the linear posterior ensemble-mean of 7.3  $\mu\text{m}$  infrared BT. (b) Deviation of the nonlinear posterior ensemble-variance from the linear posterior ensemble-variance of 7.3  $\mu\text{m}$  infrared BT. Dashed lines indicate mean and standard deviation over bins of size 10  $\text{K}^2$ .

Turner, Jeffrey D. Duda, and Trevor I. Alcott. The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part I: Motivation and System Description. *Weather and Forecasting*, 37(8):1371–1395, August 2022. ISSN 1520-0434, 0882-8156. doi: 10.1175/WAF-D-21-0151.1. URL <https://journals.ametsoc.org/view/journals/wefo/37/8/WAF-D-21-0151.1.xml>. Publisher: American Meteorological Society Section: Weather and Forecasting.

Alan J. Geer, Katrin Lonitz, Peter Weston, Masahiro Kazumori, Kozo Okamoto, Yanqiu Zhu, Emily Huichun Liu, Andrew Collard, William Bell, Stefano Migliorini, Philippe Chambon, Nadia Fourrié, Min-Jeong Kim, Christina Köpken-Watts, and Christoph Schraff. All-sky satellite data assimilation at operational weather forecasting centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1191–1217, 2018. ISSN 1477-870X. doi: 10.1002/qj.3202. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3202>.

Stefan Geiss, Leonhard Scheck, Alberto de Lozar, and Martin Weissmann. Understanding the model representation of clouds based on visible and infrared satellite observations. *Atmospheric Chemistry and Physics*, 21(16):12273–12290, August 2021. ISSN 1680-7316. doi: 10.5194/acp-21-12273-2021. URL <https://acp.copernicus.org/articles/21/12273/2021/>. Publisher: Copernicus GmbH.

Nils Gustafsson, Tijana Janjić, Christoph Schraff, Daniel Leuenberger, Martin Weissmann, Hendrik Reich, Pierre Brousseau, Thibaut Montmerle, Eric Wattrelot, Antonín Bučánek, and others. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1218–1256, 2018.

Guannan Hu, Sarah L. Dance, Ross N. Bannister, Hristo G. Chipilski, Oliver Guillet, Bruce Macpherson, Martin Weissmann, and Nusrat Yussouf. Progress, challenges, and future steps in data assimilation for convection-permitting numerical weather prediction: Report on the virtual meeting held on 10 and 12 November 2021. *Atmospheric Science Letters*, 24(1):e1130, 2023. ISSN 1530-261X. doi: 10.1002/asl.e1130. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/asl.e1130>.

Aaron Johnson, Xuguang Wang, and Thomas Jones. Impacts of Assimilating GOES-16 ABI Channels 9 and 10 Clear Air and Cloudy Radiance Observations With Additive Inflation and Adaptive Observation Error in GSI-EnKF for a Case of Rapidly Evolving Severe Supercells.

- Journal of Geophysical Research: Atmospheres, 127(11):e2021JD036157, 2022. ISSN 2169-8996. doi: 10.1029/2021JD036157. URL <http://onlinelibrary.wiley.com/doi/abs/10.1029/2021JD036157>.
- C D Jones and B Macpherson. A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. Meteorological Applications, 4(3):269–277, 1997. ISSN 1469-8080. doi: 10.1017/S1350482797000522. URL <https://onlinelibrary.wiley.com/doi/abs/10.1017/S1350482797000522>.
- Thomas A. Jones, David Stensrud, Louis Wicker, Patrick Minnis, and Rabindra Palikonda. Simultaneous Radar and Satellite Data Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011. Monthly Weather Review, 143(1):165–194, January 2015. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-14-00180.1. URL <https://journals.ametsoc.org/view/journals/mwre/143/1/mwr-d-14-00180.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- Thomas A Jones, Kent Knopfmeier, Dustan Wheatley, Gerald Creager, Patrick Minnis, and Rabindra Palikonda. Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II: Combined radar and satellite data experiments. Weather and Forecasting, 31(1):297–327, 2016.
- Thomas A. Jones, Patrick Skinner, Nusrat Yussouf, Kent Knopfmeier, Anthony Reinhart, Xuguang Wang, Christopher Bedka, William Smith, and Rabindra Palikonda. Assimilation of GOES-16 Radiances and Retrievals into the Warn-on-Forecast System. Monthly Weather Review, 148(5):1829–1859, May 2020. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-19-0379.1. URL <https://journals.ametsoc.org/view/journals/mwre/148/5/mwr-d-19-0379.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- Alicia R Karspeck and Jeffrey L Anderson. Experimental implementation of an ensemble adjustment filter for an intermediate ENSO model. Journal of Climate, 20(18):4638–4658, 2007.
- Lukas Kugler, Jeffrey L. Anderson, and Martin Weissmann. Potential impact of all-sky assimilation of visible and infrared satellite observations compared with radar reflectivity for convective-scale numerical weather prediction. Quarterly Journal of the Royal Meteorological Society, 149(757):3623–3644, 2023. ISSN 1477-870X. doi: 10.1002/qj.4577. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.4577>.
- D. Leuenberger and A. Rossa. Revisiting the latent heat nudging scheme for the rainfall assimilation of a simulated convective storm. Meteorology and Atmospheric Physics, 98(3):195–215, December 2007. ISSN 1436-5065. doi: 10.1007/s00703-007-0260-9. URL <https://doi.org/10.1007/s00703-007-0260-9>.
- Jun Li, Alan J. Geer, Kozo Okamoto, Jason A. Otkin, Zhiqian Liu, Wei Han, and Pei Wang. Satellite All-sky Infrared Radiance Assimilation: Recent Progress and Future Perspectives. Advances in Atmospheric Sciences, 39(1):9–21, January 2022. ISSN 1861-9533. doi: 10.1007/s00376-021-1088-9. URL <https://doi.org/10.1007/s00376-021-1088-9>.
- Leonhard Scheck, Martin Weissmann, and Liselotte Bach. Assimilating visible satellite images for convective-scale numerical weather prediction: A case-study. Quarterly Journal of the Royal Meteorological Society, 146(732):3165–3186, October 2020. ISSN 0035-9009, 1477-870X. doi: 10.1002/qj.3840. URL <https://onlinelibrary.wiley.com/doi/10.1002/qj.3840>.
- Josef Schrötle, Martin Weissmann, Leonhard Scheck, and Axel Hutt. Assimilating Visible and Infrared Radiances in Idealized Simulations of Deep Convection. Monthly Weather Review, 148(11):4357–4375, November 2020. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-20-0002.1. URL <https://journals.ametsoc.org/view/journals/mwre/148/11/MWR-D-20-0002.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.
- Stephen S. Weygandt, Stanley G. Benjamin, Ming Hu, Curtis R. Alexander, Tatiana G. Smirnova, and Eric P. James. Radar Reflectivity-Based Model Initialization Using Specified Latent Heating (Radar-LHI) within a Diabatic Digital Filter or Pre-Forecast Integration. Weather and Forecasting, 37(8):1419–1434, August 2022. ISSN 1520-0434, 0882-8156. doi: 10.1175/WAF-D-21-0142.1. URL <https://journals.ametsoc.org/view/journals/wfo/37/8/WAF-D-21-0142.1.xml>. Publisher: American Meteorological Society Section: Weather and Forecasting.
- Linfan Zhou, Lili Lei, Zhe-Min Tan, Yi Zhang, and Di Di. Impacts of Observation Forward Operator on Infrared Radiance Data Assimilation with Fine Model Resolutions. Monthly Weather Review, 151(1):163–173, January 2023. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-22-0084.1. URL <https://onlinelibrary.wiley.com/doi/abs/10.1175/MWR-D-22-0084.1>.

<https://journals.ametsoc.org/view/journals/mwre/151/1/MWR-D-22-0084.1.xml>. Publisher: American Meteorological Society Section: Monthly Weather Review.



# **Chapter 6**

## **Conclusions**

This thesis investigates the potential impact of all-sky visible and infrared assimilation for the prediction of convective-scale storms in idealized observing-system simulation experiments (OSSE). The idealized setup allows for a detailed analysis of the potential impact as well as of the consequences of observation operator nonlinearity in the absence of additional effects from systematic model and operator errors that may be present in non-idealized NWP systems. This thesis is based on three research papers, which

- 1) provide evidence for the high potential forecast impact of visible and infrared satellite observations and quantify it relative to that of radar observations,
- 2) demonstrate that the combination of infrared and visible observations reduces the ambiguity of information from individual channels, and
- 3) quantify systematic effects of the operator nonlinearity for visible and infrared observations and reveal that also observations with small deviations from the first guess should be assimilated.

Chapter 3 is the first study that provides a direct, quantitative comparison of the forecast impact of visible, infrared, and radar observations. Chapter 4 is the first study that reveals the complementary value of assimilating visible and infrared observations for the analysis of clouds and their vertical structure. Chapter 5, for the first time, analyses deviations caused by the operator nonlinearity for visible and infrared observations.

The following section discusses this thesis's main findings in relation to preceding research. Finally, the implications of the findings are discussed in a broader context in the last section.

### **6.1 Summary and discussion**

#### **Chapter 3: Potential impact of assimilating visible and infrared satellite observations**

The first research goal was to assess the potential impact of assimilating cloud-affected satellite observations for the prediction of convective storms. The publication in Chapter 3 (Kugler et al., 2023) conducted idealized experiments that neglect model and operator errors to focus on the remaining challenges for assimilating visible and infrared observations. Given these simplifications, the resulting absolute impact can likely not be realized in an operational

system. To overcome this limitation, the experiments with satellite observation are evaluated in comparison with experiments for more commonly assimilated radar observations. The resulting impact estimates relative to radar can be expected to be better transferable to real NWP systems.

Chapter 3 presents the first study directly comparing the assimilation of visible, infrared (6.2 and 7.3  $\mu\text{m}$ ), and radar observations. By interpreting the potential impact in relation to other observation types, Kugler et al. (2023) avoided the shortcomings of other assimilation impact studies that only compare the impact of one observation type in comparison to a reference experiment without these observations. This limitation of other studies makes it particularly difficult to compare preceding studies on infrared assimilation, which often revealed a positive forecast impact, but did not investigate if and to which extent the same forecast impact could have been reached by assimilating another observation type.

The first study on visible assimilation in convective-scale NWP by Scheck et al. (2020) assessed the additional forecast impact of visible observations in a near-operational setup and revealed a positive forecast impact in a case study. However, the study did not conduct any comparable experiments for radar or infrared observations. Schröttle et al. (2020) compared the forecast impact of visible and infrared 6.2  $\mu\text{m}$  assimilation to a free forecast without data assimilation in an idealized OSSE.

Kugler et al. (2023) added evidence of a high potential impact of assimilating cloud-affected satellite observations. Regarding the impact of visible observations, Kugler et al. (2023) largely agrees with Scheck et al. (2020) on a substantial potential of assimilating this observation type. However, Schröttle et al. (2020) had reported a substantially lower impact from assimilating visible observations than from infrared 6.2  $\mu\text{m}$  observations. In contrast, Kugler et al. (2023) found the impact of assimilating visible observations to be higher than that of infrared 6.2  $\mu\text{m}$  observations in a very similar idealized weather scenario. An explanation for the different results might be the overly inflated assigned observation error variance in Schröttle et al. (2020). Based on extensive sensitivity studies for parameter settings in this thesis, we conclude that the potential of infrared and visible observations is roughly of a similar magnitude. Second, Kugler et al. (2023) additionally evaluated the impact of assimilating the 7.3  $\mu\text{m}$  channel. Its impact was considerably higher than that of the 6.2  $\mu\text{m}$  channel. This is likely related to the lower peak of the weighting function of the 7.3  $\mu\text{m}$  channel, which allows to detect also clouds at lower levels compared to the 6.2  $\mu\text{m}$  channel. Third, the experiments reveal that either assimilating one infrared or visible satellite channel can be nearly as beneficial as assimilating 3D radar observations under favorable conditions and about as beneficial as 2D radar observations in a more challenging scenario for data assimilation.

## Chapter 4: Combined assimilation of multiple satellite channels

The second research goal was to analyze whether it would be possible to mitigate the ambiguity of visible and infrared observations in the assimilation by their combined use. To this end, this chapter extends the study in the previous chapter to the combined assimilation of multiple observation types. The novelty of this study is that it evaluates

the vertical distribution of analysis errors for different cloud conditions and features a more comprehensive sensitivity test for assimilation settings.

The results revealed that the combined assimilation can mitigate the ambiguity of individual visible and infrared channels and improve the analysis of the vertical distribution of clouds. Moreover, forecasts of precipitation and cloudiness were improved. The combined assimilation of visible and infrared observations led to an even higher skill score than the experiment that only assimilated radar observations. However, the stronger inflation for the radar-only experiment might have contributed to this result. Nevertheless, it is evident that combining infrared and visible observations leads to better forecasts than the assimilation of either one of them and that their combined potential is at least similar to that of radar observations.

The prior literature on a combined assimilation of visible and infrared observations is limited to a single paper. Schröttle et al. (2020) noted that the combined assimilation considerably improved the accuracy of temperature, humidity, and wind. Chapter 4 demonstrates that the combined assimilation is also beneficial for cloud variables, especially in conditions where the truth has ice clouds. In such conditions, visible observations can provide additional information on water clouds as thin ice clouds above are not opaque in this spectral range. Considering visible and infrared observations in synopsis thus allows to discriminate between thin ice clouds and deep convective clouds in many instances.

## **Chapter 5: Nonlinearity effects of assimilating visible reflectance**

The third research goal was to improve our understanding of nonlinearity effects induced by nonlinear observation operators. Although the functional relationships in the observation operator were well known, the consequences on the analysis ensemble mean/variance adjustment had been unclear. Chapter 5 presents a manuscript in preparation, which quantifies the deviations of the ensemble mean and variance from the theoretical expectation that is based on the linearity assumption. The deviations were evaluated using the same configuration as for the experiments in chapters 3 and 4. Additionally, the role of nonlinearity is evaluated for the subset of observations with small first-guess departures.

Scheck et al. (2020) found that observations with small first-guess departures are particularly affected by observation operator nonlinearity. More specifically, they noted that "the nonlinearity of the operator thus limits the effectiveness of the visible assimilation". Chapter 5 confirms the particular effect that was observed in Scheck et al. (2020). However, additional experiments showed that rejecting observations with small first-guess departures does not reduce the overall error but increases both the overall bias and the mean-square error at the location of the observations with small first-guess departures. While the observations with small departures may lead to a direct detrimental adjustment of the mean state, they also lead to a reduction of the ensemble variance (i.e. tell the ensemble that the first guess is accurate) and thereby reduce that risk that other nearby observations draw the analysis away from the accurate first guess. In summary, we therefore conclude that also observations with small first guess departures should be included in the assimilation.

Regarding the effects on the ensemble mean adjustment, this chapter reveals a systematic

deviation of the actual posterior in observation space from its theoretical expectation (assuming a linear observation operator). Moreover, the analysis ensemble variance can be much larger than expected, and the variance tends to overestimate the mean-squared error of the analysis. However, the overestimation may be beneficial in practice as it can compensate for a common underdispersiveness of convective-scale ensembles.

Chapter 5 quantifies the effect of observation operator nonlinearity of visible and infrared radiances on the assimilation results in observation space. In the future, these systematic deviations could be taken into account when assimilating these observations. Thus, the findings can be a basis for a potential correction algorithm that corrects the systematic part of the nonlinearity deviations.

## 6.2 Concluding remarks

### Synthesis of findings

Despite the violated assumptions of linearity and Gaussianity due to the visible and infrared observation operators, the all-sky assimilation of such radiances bears a substantial potential benefit for convective-scale forecasts. This thesis estimated the forecast impact in idealized studies that omit certain errors of operational NWP. To make the impact estimates better transferable to real NWP systems, the estimated impact was evaluated in relation to more commonly assimilated radar observations.

Furthermore, the estimation of nonlinearity deviations improves the understanding of nonlinearity effects and paves the way to compensate for the undesired effects of nonlinear observation operators, which could ultimately improve the effectiveness of assimilating nonlinear observations.

### Outlook

This thesis focused on the potential impact of satellite observations and the effect of operator nonlinearity in the absence of systematic model and operator errors as well as correlated observation errors. Future studies should investigate the role of these neglected additional challenges to achieve a high impact of these observations in regional NWP systems: Systematic model error can be introduced by using a higher resolution nature run. Such an endeavor is already ongoing in a companion study. Similarly, systematic operator errors could be represented through the use of different operator settings or assumptions for the experiments and the nature run (Li et al., 2022).

Second, this thesis investigated the assimilation in idealized OSSEs of the EAKF. Additionally, the benefits of the combined assimilation should be investigated with different assimilation algorithms in areas without radar coverage, such as the sea, and in cases of non-convective weather conditions, e.g., fog and low stratus.

Third, other observation types might be complementary to visible and infrared radiances. For example, in cases where visible and infrared are blind, the microwave spectrum could

still be used to observe cloud properties below the cloud top and near-infrared or infrared window channels may substitute the lack of visible information at nighttime.



# Acronyms

**3D-Var** 3D-Variational data assimilation. 3, 8

**4D-Var** 4D-Variational data assimilation. 3, 4, 8

**DART** Data Assimilation Research Testbed. 23

**EAKF** Ensemble adjustment Kalman filter. 8, 11, 19, 20, 23, 92

**ECMWF** European Centre for Medium-Range Weather Forecasts. 2–4

**EnKF** Ensemble Kalman filter. 8, 17, 19

**EUMETSAT** European Organisation for the Exploitation of Meteorological Satellites. 7, 9

**LEO** Low Earth orbit. 6

**MFASIS** Method for fast satellite image simulations. 21

**MSG** Meteosat second generation. 7

**NCEP** National Oceanic and Atmospheric Administration. 3

**NOAA** National Oceanic and Atmospheric Administration. 1, 14

**NWP** Numerical weather prediction. 1, 3, 5, 6, 8, 10, 11, 13, 15, 17, 21, 89, 90, 92

**OSSE** Observing-system simulation experiments. 11, 23, 90, 92

**PDE** Partial differential equations. 2

**RTTOV** Radiative transfer model for the Television and Infrared Observational Satellite.  
20, 23

**SEVIRI** Spinning Enhanced Visible and Infrared Imager. 7

**WRF** Weather Research and Forecasting. 13, 14, 23



# Bibliography

- Abbe, C. "The Physical Basis of long-range Weather Forecasts." *Monthly Weather Review* 29.12 (1901), pp. 551–561. DOI: 10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2.
- Anderson, J. L. "An ensemble adjustment Kalman filter for data assimilation." *Monthly weather review* 129.12 (2001), pp. 2884–2903. DOI: [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- "A local least squares framework for ensemble filtering." *Monthly Weather Review* 131.4 (2003), pp. 634–642. DOI: [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2).
  - "Localization and sampling error correction in ensemble Kalman filter data assimilation." *Monthly Weather Review* 140.7 (2012), pp. 2359–2371.
  - "Reducing Correlation Sampling Error in Ensemble Kalman Filter Data Assimilation." *Monthly Weather Review* 144.3 (2016), pp. 913–925. DOI: 10.1175/MWR-D-15-0052.1.
- Anderson, J. L. and N. Collins. "Scalable Implementations of Ensemble Filter Algorithms for Data Assimilation." *Journal of Atmospheric and Oceanic Technology* 24.8 (2007), pp. 1452–1463. DOI: 10.1175/JTECH2049.1.
- Bachmann, K., C. Keil, G. C. Craig, M. Weissmann, and C. A. Welzbacher. "Predictability of deep convection in idealized and operational forecasts: Effects of radar data assimilation, orography, and synoptic weather regime." *Monthly Weather Review* 148.1 (2020), pp. 63–81.
- Bachmann, K., C. Keil, and M. Weissmann. "Impact of radar data assimilation and orography on predictability of deep convection." *Quarterly Journal of the Royal Meteorological Society* 145.718 (2019), pp. 117–130.
- Bátkai, A., P. Csomós, I. Faragó, A. Horányi, and G. Szépszó. *Mathematical problems in meteorological modelling*. The European Consortium for Mathematics in Industry. Springer, 2016. ISBN: 978-3-319-40155-3.
- 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, H. Lin, S. E. Peckham, T. L. Smith, W. R. Moninger, J. S. Kenyon, and G. S. Manikin. "A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh." *Monthly Weather Review* 144.4 (2016), pp. 1669–1694. DOI: 10.1175/MWR-D-15-0242.1.

- Bierdel, L., P. Friederichs, and S. Bentzien. "Spatial kinetic energy spectra in the convection-permitting limited-area NWP model COSMO-DE." *Meteorologische Zeitschrift* (2012), pp. 245–258. DOI: 10.1127/0941-2948/2012/0319.
- Bjerknes, V. "Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik." *Meteorologische Zeitschrift* 21 (1904), pp. 1–7.
- Chan, M.-Y., F. Zhang, X. Chen, and L. R. Leung. "Potential Impacts of Assimilating All-Sky Satellite Infrared Radiances on Convection-Permitting Analysis and Prediction of Tropical Convection." *Monthly Weather Review* 148.8 (2020), pp. 3203–3224. DOI: 10.1175/MWR-D-19-0343.1.
- Charney, J. G., R. Fjörtoft, and J. V. Neumann. "Numerical Integration of the Barotropic Vorticity Equation." *Tellus* 2.4 (1950), pp. 237–254. DOI: 10.3402/tellusa.v2i4.8607.
- Chou, M.-D., K.-T. Lee, S.-C. Tsay, and Q. Fu. "Parameterization for Cloud Longwave Scattering for Use in Atmospheric Models." *Journal of Climate* 12.1 (1999), pp. 159–169. DOI: 10.1175/1520-0442(1999)012<0159:PFCLSF>2.0.CO;2.
- Courtier, P., E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. "The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation." *Quarterly Journal of the Royal Meteorological Society* 124.550 (1998), pp. 1783–1807. DOI: 10.1002/qj.49712455002.
- Durran, D. R. *Numerical methods for fluid dynamics : with applications to geophysics*. 2. ed. Texts in applied mathematics ; 32. New York, NY [u.a.]: Springer, 2010. ISBN: 978-1-4419-6411-3.
- EEA. *Economic losses and fatalities from weather- and climate-related events in Europe. Briefing no. 21/2021*. 2023. DOI: 10.2800/7654.
- Eyre, J. R., S. J. English, and M. Forsythe. "Assimilation of satellite data in numerical weather prediction. Part I: The early years." *Quarterly Journal of the Royal Meteorological Society* 146.726 (2020), pp. 49–68. DOI: 10.1002/qj.3654.
- Folger, K. and M. Weissmann. "Lidar-Based Height Correction for the Assimilation of Atmospheric Motion Vectors." *Journal of Applied Meteorology and Climatology* 55.10 (2016), pp. 2211–2227. DOI: 10.1175/JAMC-D-15-0260.1.
- Frèrebeau, P. "Ein schnelles Strahlungsmodell zur Simulation von solaren Satellitenbildern." PhD Thesis. Ludwig-Maximilians-Universität München, 2014.
- Fujita, T. T. "Tornadoes and Downbursts in the Context of Generalized Planetary Scales." *Journal of the Atmospheric Sciences* 38.8 (1981), pp. 1511–1534. DOI: 10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2.
- Geer, A. J., F. Baordo, N. Bormann, P. Chambon, S. J. English, M. Kazumori, H. Lawrence, P. Lean, K. Lonitz, and C. Lupu. "The growing impact of satellite observations sensitive

- to humidity, cloud and precipitation." *Quarterly Journal of the Royal Meteorological Society* 143.709 (2017), pp. 3189–3206. DOI: 10.1002/qj.3172.
- Geiss, S., L. Scheck, A. de Lozar, and M. Weissmann. "Understanding the model representation of clouds based on visible and infrared satellite observations." *Atmospheric Chemistry and Physics* 21.16 (2021), pp. 12273–12290. DOI: 10.5194/acp-21-12273-2021.
- Gharamti, M. E., K. Raeder, J. Anderson, and X. Wang. "Comparing Adaptive Prior and Posterior Inflation for Ensemble Filters Using an Atmospheric General Circulation Model." *Monthly Weather Review* 147.7 (2019), pp. 2535–2553. DOI: 10.1175/MWR-D-18-0389.1.
- Gustafsson, N., T. Janjić, C. Schraff, D. Leuenberger, M. Weissmann, H. Reich, P. Brousseau, T. Montmerle, E. Wattrelot, A. Bučánek, et al. "Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres." *Quarterly Journal of the Royal Meteorological Society* 144.713 (2018), pp. 1218–1256.
- Hohenegger, C. and C. Schär. "Atmospheric Predictability at Synoptic Versus Cloud-Resolving Scales." *Bulletin of the American Meteorological Society* 88.11 (2007), pp. 1783–1794. DOI: 10.1175/BAMS-88-11-1783.
- Honda, T., T. Miyoshi, G.-Y. Lien, S. Nishizawa, R. Yoshida, S. A. Adachi, K. Terasaki, K. Okamoto, H. Tomita, and K. Bessho. "Assimilating All-Sky Himawari-8 Satellite Infrared Radiances: A Case of Typhoon Soudelor (2015)." *Monthly Weather Review* 146.1 (2018), pp. 213–229. DOI: 10.1175/MWR-D-16-0357.1.
- Houtekamer, P. L., L. Lefavre, J. Derome, H. Ritchie, and H. L. Mitchell. "A System Simulation Approach to Ensemble Prediction." *Monthly Weather Review* 124.6 (1996), pp. 1225–1242. DOI: 10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2.
- Houtekamer, P. L. and H. L. Mitchell. "Data Assimilation Using an Ensemble Kalman Filter Technique." *Monthly Weather Review* 126.3 (1998), pp. 796–811. DOI: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- "A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation." *Monthly Weather Review* 129.1 (2001), pp. 123–137. DOI: 10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins. "Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models." *Journal of Geophysical Research: Atmospheres* 113.D13 (2008). DOI: 10.1029/2008JD009944.
- Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda. "Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II: Combined radar and satellite data experiments." *Weather and Forecasting* 31.1 (2016), pp. 297–327.

- Jones, T. A., P. Skinner, N. Yussouf, K. Knopfmeier, A. Reinhart, X. Wang, K. Bedka, W. Smith, and R. Palikonda. "Assimilation of GOES-16 Radiances and Retrievals into the Warn-on-Forecast System." *Monthly Weather Review* 148.5 (2020), pp. 1829–1859. DOI: 10.1175/MWR-D-19-0379.1.
- Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda. "Simultaneous Radar and Satellite Data Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011." *Monthly Weather Review* 143.1 (2015), pp. 165–194. DOI: 10.1175/MWR-D-14-00180.1.
- Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering* 82.1 (1960), pp. 35–45. DOI: 10.1115/1.3662552.
- Karspeck, A. R. and J. L. Anderson. "Experimental implementation of an ensemble adjustment filter for an intermediate ENSO model." *Journal of Climate* 20.18 (2007), pp. 4638–4658.
- Kostka, P. M., M. Weissmann, R. Buras, B. Mayer, and O. Stiller. "Observation Operator for Visible and Near-Infrared Satellite Reflectances." *Journal of Atmospheric and Oceanic Technology* 31 (2014).
- Kugler, L., J. L. Anderson, and M. Weissmann. "Potential impact of all-sky assimilation of visible and infrared satellite observations compared with radar reflectivity for convective-scale numerical weather prediction." *Quarterly Journal of the Royal Meteorological Society* 149.757 (2023), pp. 3623–3644. DOI: 10.1002/qj.4577.
- Lange, H. and G. C. Craig. "The impact of data assimilation length scales on analysis and prediction of convective storms." *Monthly Weather Review* 142.10 (2014), pp. 3781–3808.
- Lewis, W. E., T. J. Wagner, J. A. Otkin, and T. A. Jones. "Impact of AERI Temperature and Moisture Retrievals on the Simulation of a Central Plains Severe Convective Weather Event." *Atmosphere* 11.7 (2020), p. 729. DOI: 10.3390/atmos11070729.
- Li, J., A. J. Geer, K. Okamoto, J. A. Otkin, Z. Liu, W. Han, and P. Wang. "Satellite All-sky Infrared Radiance Assimilation: Recent Progress and Future Perspectives." *Advances in Atmospheric Sciences* 39.1 (2022), pp. 9–21. DOI: 10.1007/s00376-021-1088-9.
- Lorenz, E. N. "Deterministic nonperiodic flow." *Journal of atmospheric sciences* 20.2 (1963), pp. 130–141.
- Lynch, P. "The origins of computer weather prediction and climate modeling." *Journal of Computational Physics. Predicting weather, climate and extreme events* 227.7 (2008), pp. 3431–3444. DOI: 10.1016/j.jcp.2007.02.034.
- Nakanishi, M. and H. Niino. "An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog." *Boundary-Layer Meteorology* 119.2 (2006), pp. 397–407. DOI: 10.1007/s10546-005-9030-8.

- Necker, T., M. Weissmann, Y. Ruckstuhl, J. Anderson, and T. Miyoshi. "Sampling error correction evaluated using a convective-scale 1000-member ensemble." *Monthly Weather Review* 148.3 (2020), pp. 1229–1249.
- Necker, T., M. Weissmann, and M. Sommer. "The importance of appropriate verification metrics for the assessment of observation impact in a convection-permitting modelling system." *Quarterly Journal of the Royal Meteorological Society* 144.714 (2018), pp. 1667–1680. DOI: 10.1002/qj.3390.
- Petty, G. W. *A First Course in Atmospheric Radiation*. 2nd. Sundog Publishing LLC, 2006. ISBN: 978-0-9729033-1-8.
- Peubey, C. and A. McNally. "Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context." *Quarterly Journal of the Royal Meteorological Society* 135.644 (2009), pp. 1863–1876. DOI: 10.1002/qj.500.
- Rabier, F., J.-N. Thépaut, and P. Courtier. "Extended assimilation and forecast experiments with a four-dimensional variational assimilation system." *Quarterly Journal of the Royal Meteorological Society* 124.550 (1998), pp. 1861–1887. DOI: 10.1002/qj.49712455005.
- Richardson, L. F. *Weather Prediction by Numerical Process*. 2nd Edition. Cambridge Mathematical Library. Cambridge: New York, 2007. ISBN: 0-521-68044-1.
- Saunders, R., M. Matricardi, A. Geer, and P. Rayer. *RTTOV-9: Science and Validation Report*. Tech. rep. Met. Office, ECMWF, Edinburgh University, 2010.
- Sawada, Y., K. Okamoto, M. Kunii, and T. Miyoshi. "Assimilating Every-10-minute Himawari-8 Infrared Radiances to Improve Convective Predictability." *Journal of Geophysical Research: Atmospheres* 124.5 (2019), pp. 2546–2561. DOI: 10.1029/2018JD029643.
- Scheck, L., P. Frèrebeau, R. Buras-Schnell, and B. Mayer. "A fast radiative transfer method for the simulation of visible satellite imagery." *Journal of Quantitative Spectroscopy and Radiative Transfer* 175 (2016), pp. 54–67. DOI: 10.1016/j.jqsrt.2016.02.008.
- Scheck, L., M. Weissmann, and L. Bach. "Assimilating visible satellite images for convective-scale numerical weather prediction: A case-study." *Quarterly Journal of the Royal Meteorological Society* 146.732 (2020), pp. 3165–3186. DOI: 10.1002/qj.3840.
- Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Periéñez, and R. Potthast. "Kilometre-scale ensemble data assimilation for the COSMO model (KENDA)." *Quarterly Journal of the Royal Meteorological Society* 142.696 (2016), pp. 1453–1472. DOI: 10.1002/qj.2748.
- Schrötle, J., M. Weissmann, L. Scheck, and A. Hutt. "Assimilating Visible and Infrared Radiances in Idealized Simulations of Deep Convection." *Monthly Weather Review* 148.11 (2020), pp. 4357–4375. DOI: 10.1175/MWR-D-20-0002.1.
- Seneviratne, S. I. et al. "Changes in Climate Extremes and their Impacts on the Natural Physical Environment." *Managing the Risks of Extreme Events and Disasters to Advance*

- Climate Change Adaptation*. Ed. by C. B. Field, V. Barros, T. F. Stocker, and Q. Dahe. 1st ed. Cambridge University Press, 2012, pp. 109–230. ISBN: 978-1-107-02506-6. DOI: 10.1017/CBO9781139177245.006.
- Skamarock, C., B. Klemp, J. Dudhia, O. Gill, Z. Liu, J. Berner, W. Wang, G. Powers, G. Duda, D. Barker, and X.-y. Huang. *A Description of the Advanced Research WRF Model Version 4.3*. Tech. rep. NCAR/UCAR, 2021.
- Skamarock, W. C. “Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra.” *Monthly Weather Review* 132.12 (2004), pp. 3019–3032. DOI: 10.1175/MWR2830.1.
- Smith, A. B. *U.S. Billion-dollar Weather and Climate Disasters, 1980 - present*. NOAA National Centers for Environmental Information. 2020. DOI: <https://doi.org/10.25921/stkw-7w73>.
- Snyder, C. and F. Zhang. “Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter.” *Monthly Weather Review* 131.8 (2003), pp. 1663–1677. DOI: 10.1175//2555.1.
- Tewari, M., F. Chen, W. Wang, J. Dudhia, M. A. LeMone, K. Mitchell, M. Ek, G. Gayno, J. Wegiel, and R. H. Cuenca. “Implementation and verification of the unified NOAH land surface model in the WRF model.” 2004, pp. 11–15.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall. “Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization.” *Monthly Weather Review* 136.12 (2008), pp. 5095–5115. DOI: 10.1175/2008MWR2387.1.
- Tong, M. and M. Xue. “Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments.” *Monthly Weather Review* 133.7 (2005), pp. 1789–1807. DOI: 10.1175/MWR2898.1.
- UCAR/NCAR/CISL/DAReS. *The Data Assimilation Research Testbed (Version 10.5.3) [Software]*. Boulder, Colorado, 2022.
- Ying, Y. “Assimilating Observations with Spatially Correlated Errors Using a Serial Ensemble Filter with a Multiscale Approach.” *Monthly Weather Review* 148.8 (2020), pp. 3397–3412. DOI: 10.1175/MWR-D-19-0387.1.
- Žagar, N. and I. Szunyogh. “Comments on “What Is the Predictability Limit of Midlatitude Weather?”” *Journal of the Atmospheric Sciences* 77.2 (2020), pp. 781–785. DOI: 10.1175/JAS-D-19-0166.1.
- Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel. “What Is the Predictability Limit of Midlatitude Weather?” *Journal of the Atmospheric Sciences* 76.4 (2019), pp. 1077–1091. DOI: 10.1175/JAS-D-18-0269.1.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng. "Intrinsic Predictability of the 20 May 2013 Tornadic Thunderstorm Event in Oklahoma at Storm Scales." *Monthly Weather Review* 144.4 (2016), pp. 1273–1298. DOI: 10.1175/MWR-D-15-0105.1.

Zhu, L., M. Xue, R. Kong, and J. Min. "Direct Assimilation of All-Sky GOES-R ABI Radiiances in GSI EnKF for the Analysis and Forecasting of a Mesoscale Convective System." *Monthly Weather Review* -1.aop (2022). DOI: 10.1175/MWR-D-21-0293.1.

### **Acknowledgment of tools for writing**

Online dictionaries, translators, and other services based on convolutional neural networks and large language models have been used to correct grammatical mistakes and to get suggestions for better-fitting phrases and improved clarity: "dictionary.com/thesaurus.com" for definitions, synonyms, and antonyms; "deepl.com" for translation and writing style; "ChatGPT (3.5)" to find related phrases and to get suggestions for an improved writing style; "grammarly.com" for spelling, grammar, and punctuation correction and to get suggestions for an improved writing style.

# Acknowledgments

I would like to thank all those who supported me during the PhD project.

First and foremost, I want to thank my supervisor, Martin Weissmann, for sparking my interest in the topic, for guiding me to see the bigger picture when I was caught up in details, for allowing me to explore my own ideas, and for the extraordinary amount of time which you dedicated to supporting me over the course of four years. Moreover, I thank Philipp, Stefano, and Tobias for introducing me to the language of data assimilation and our numerous scientific discussions. I am grateful to all my office mates over the years, Anne, Kaushambi, Kathi, and Martin, and to all other colleagues in the department. I enjoyed your company very much.

Finally, I am very thankful to my family and Daria for the great moral support.