RESEARCH ARTICLE
PREPRINT - SUBMITTED TO QJRMS

Potential impact of assimilating visible and infrared satellite observations compared to radar reflectivity for convective-scale NWP

Lukas Kugler[1] | Jeffrey L. Anderson[2] | Martin Weissmann[1]

[1]Institut für Meteorologie und Geophysik, Universität Wien, Vienna, Austria

[2]Data Assimilation Research Section, Computational Information Systems Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

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

Funding information

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 paper, we estimate the potential impact of assimilating cloud-affected satellite observations of visible (0.6 µm) and near thermal infrared wavelength (6.2 µm and 7.3 µm) relative to the impact of assimilating radar reflectivity observations. We employed observing system simulation experiments (OSSE) with an identical twin nature run 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 (DART). The Weather Research and Forecasting (WRF) 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 lead to 88% of the radar impact.

Abbreviations: BT brightness temperature, CAPE convective available potential energy, CIN convective inhibition.

Under more difficult conditions, the impact of visible and infrared observations still reached 50% and 79%, respectively.

1 | INTRODUCTION

Clouds are the first area-wide observable signal of convection and heavily used in nowcasting applications. But in contrast to nowcasting, the use of cloud-affected satellite observations in data assimilation is very limited (Gustafsson et al., 2018). Clouds are detected much 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 & 7.3 µm) as well as cloud-top temperature (10.8 µm). Near-infrared channels can distinguish between ice and water clouds (1.6 µm) and detect nighttime low-level clouds and fog (3.9 µm). 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 challenges equally occur for more commonly assimilated radar observations. Firstly, the limited numerical representation of cloud processes and hydrometeors as well as simplifications of observation operators lead to systematic errors between models and observations (Geiss et al., 2021; Scheck et al., 2018), which violates a basic assumption 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. Secondly, 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 sub-optimal analysis in certain conditions.

While several studies investigated the assimilation of infrared channels, visible channels have received little attention by the research community so far. Since the fast visible operator MFASIS (Scheck et al., 2016) was published, only two studies applied the operator for data assimilation and investigated the impact of these observations on convective-scale data assimilation in an idealized and a near-operational setup: (Schröttle 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 to 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.

Most studies on the assimilation of thermal infrared satellite observations focused on wavelengths in the water vapor sensitive band (5-8.5 µm) since these wavelengths are less sensitive to surface emission, which is difficult to model accurately. (Otkin, 2012a,b) pioneered the direct assimilation on four channels between 6.2 and 8.5 µm, albeit at a resolution of 15 km that only partly resolves deep convection. 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 (Honda et al., 2018; Zhu et al., 2022; Sawada et al., 2019; Jones et al., 2016; Cintineo et al., 2016; Eure et al., 2023). A direct comparison between the impact of visible and infrared satellite channels with that of radar observations has, however, never been conducted so far.

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 potential weaknesses of each type. 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, e.g. concerning the representation of hydrometeors and related biases as well as for 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 e.g. 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 latter describes deep convection triggered at random locations and is termed ”random” case in this paper. 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. These two scenarios are used to specifically investigate:

1. if the ensemble adjustment Kalman filter (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 better understanding of the impact of satellite observations which is crucial for an 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 profiles of temperature, water vapor, moisture and wind illustrated in figure 1. It is a modified sounding from Payerne, Switzerland on 30 July 2008 and offers a highly unstable environment with 2670 J/kg CAPE and 26 J/kg CIN at 7 UTC in the nature run. In this paper, time UTC is equivalent to local solar time since the solar radiation is that of 0° 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.

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 6 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 µm infrared satellite image and figure 3 shows the same in 0.6 µm visible reflectance. At 11 UTC, five hours into the simulation, convection reached altitudes of about 10 km. Shortly after at 12 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 on the order of hours. After 16 UTC, convection decays.

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 isolated and well-organized storms 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 12 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 µm and the 0.6 µm channel. 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.

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 as well as well as the nature run in identical configurations (*identical twin*). 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 HRRR model described in (Benjamin et al., 2016) with microphysics from (Thompson et al., 2008); RRTMG

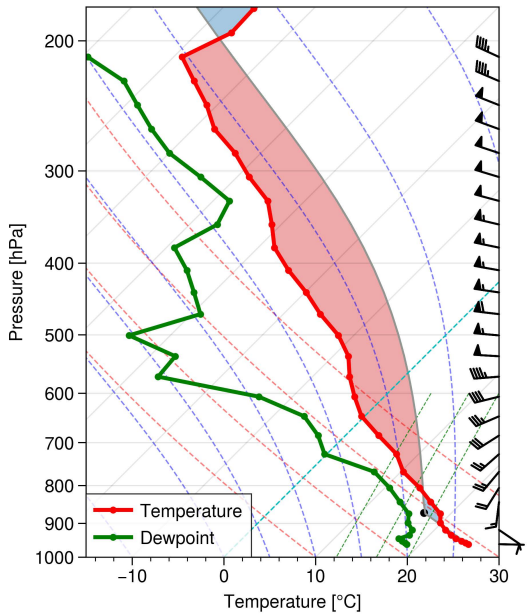


FIGURE 1 Skew *T*-log *p* diagram of the nature run initial condition for the "random" case, from domain average fields. Shown are domain-average profiles of temperature, dewpoint and a parcel lifting curve. The temperature perturbation in the warm-bubble case modifies this profile.

short- and longwave schemes (Iacono et al., 2008), MYNN 2.5 for 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.

Domain description

The domain features 200 x 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 (30 July 2008) at a latitude of 45° N.

2.3 | Initial & boundary conditions

Initial conditions of the nature run

The initial conditions of the nature run feature a highly unstable stratification with a CAPE of 2670 J/kg and a CIN of 26 J/kg at 7 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

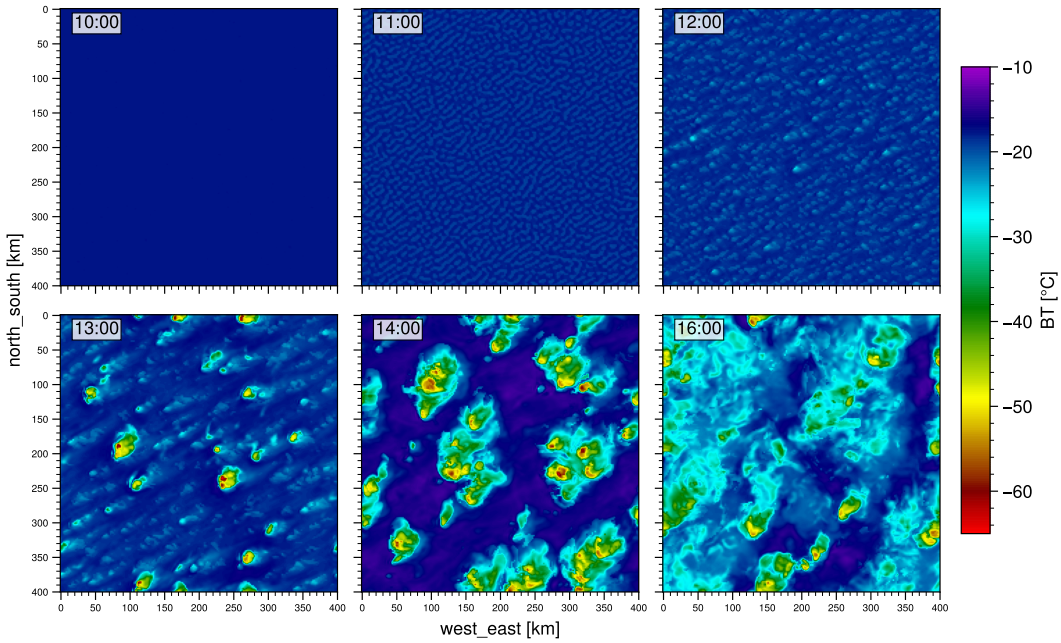


FIGURE 2 Infrared 7.3 µm satellite images of the "random" case's nature run

a bulk shear 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

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

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).

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 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 WRF's `ideal`

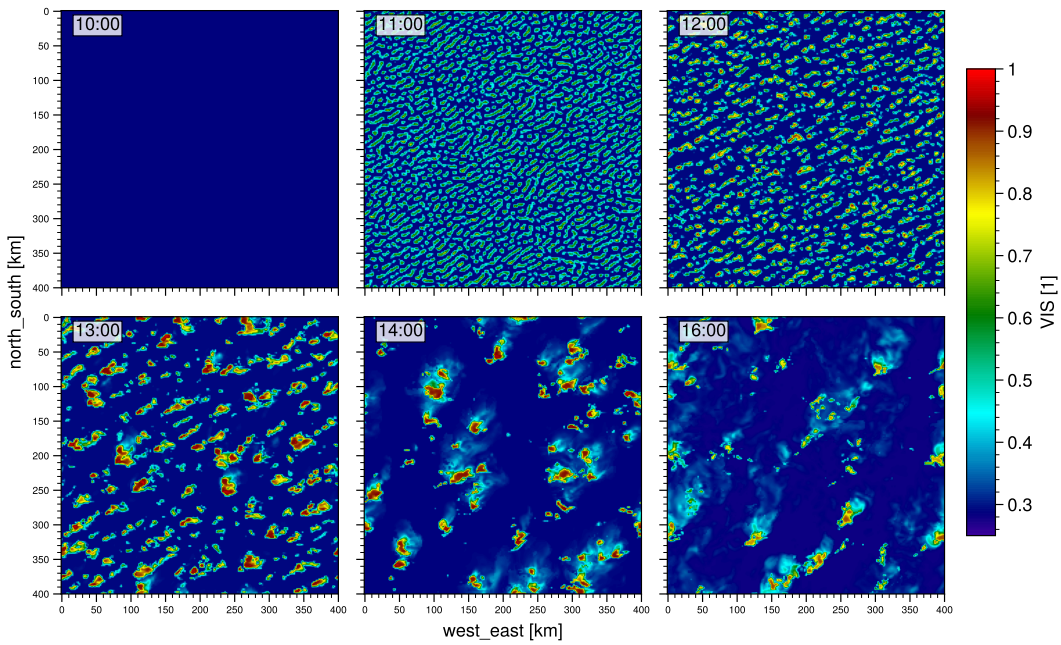


FIGURE 3 Visible 0.6 µm satellite images of the ”random” case's nature run

program, which slightly modifies the profiles 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 $x \exp((p - p_{sfc})/25)$ 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 m/s) for each column. Although the horizontal variation of temperature was this small at initialization time, the perturbations grew considerably in the following six hours, reaching a spread of 1 K in temperature, 20% in relative humidity and 2 m/s 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 h 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 m/s in zonal wind at 13 UTC (not shown).

Additional perturbations in the "warm-bubble" case

In the ”warm-bubble” case, we imposed another uncertainty about 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 ± 60 km (uniformly random) and
- the spatial extent and strength by perturbing the amplitude A by ± 1 K (uniformly random).

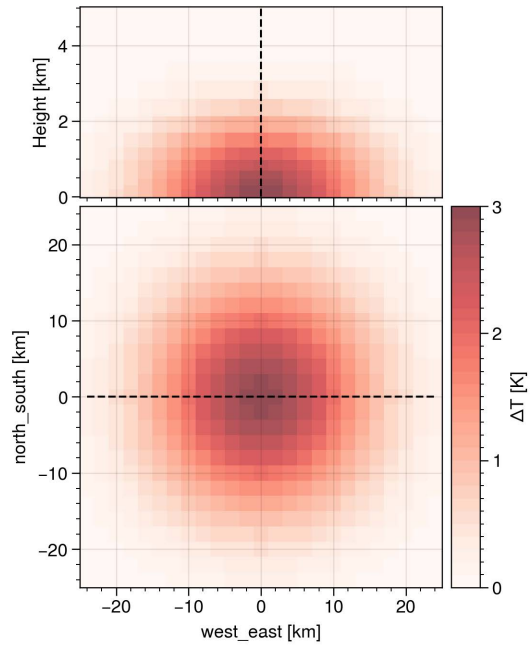


FIGURE 4 Cross-sections through the warm-bubble (temperature perturbation): vertical-horizontal slice (top) at `north_south=0` and horizontal-horizontal slice (bottom) at `Height=0`, both marked by a dashed line.

2.4 | Simulated observations

Four types of observations have been used in this study: **1)** Satellite observations of visible reflectance at a wavelength of 0.6 µm reveal how much sunlight is reflected by clouds or the surface. Contrary 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 to 1. The lowest possible value in practice is, however, determined by the surface albedo, which is around 0.27 in our setup. The observation 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 µm channel (MSG-4 SEVIRI channel 5, GOES ABI and HIMAWARI 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 observation error was chosen to be 1 K, similar to (Jones et al., 2020) who used 1.25 K.

3) Satellite observations of infrared brightness temperature of the 7.3 µm channel (MSG-4 SEVIRI channel 6, GOES ABI and HIMAWARI AHI band 10) provide similar information as channel 5/8 but are more sensitive to lower tropospheric water vapor. An observation error of 1 K was selected, which is lower than (Jones et al., 2020) who used 1.75 K (3.5 K) for non-cloudy (cloudy) observations.

4) Finally, three-dimensional radar reflectivity (10 cm) serves as a reference observation type to for evaluating the impact of satellite observations. An observation 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^o were generated using the *Data Assimilation Research Testbed* (DART) provided by (UCAR/NCAR/CISL/DAReS,

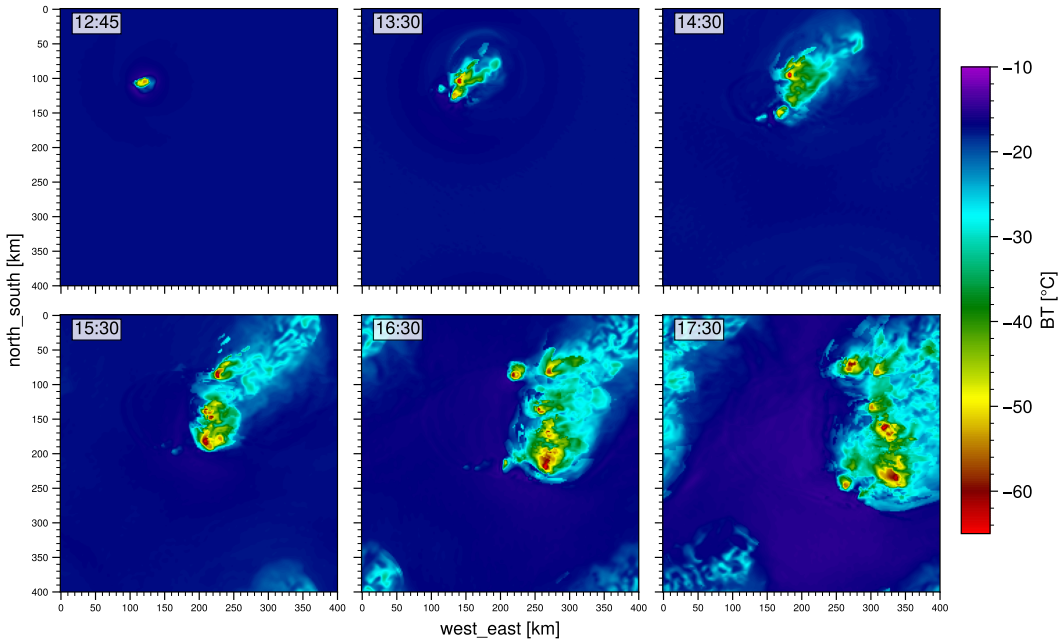


FIGURE 5 Infrared 7.3 µm satellite images of the warm-bubble case's nature run

2022). It interpolates the nature run's state x^{nat} to each observation's location and applies an observation operator $\mathcal{H}$ to the state x^{nat} before adding Gaussian observation error:

$$y^o = \mathcal{H}(x^{nat}) + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma_o) \tag{2}$$

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 and MFASIS (Scheck et al., 2016) for the visible channel, as provided in 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 µm for water droplets and 60 µm 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 latitude of 45° and longitude 0°.

2.5 | Assimilation system and settings

Our experiments applied the ensemble adjustment Kalman filter (EAKF) by (Anderson, 2001) included in the Data Assimilation Research Testbed (DART, (Anderson et al., 2009), `https://dart.ucar.edu`). The EAKF is a serial deterministic square root filter, which assimilates one observation after another. The following variables were updated:

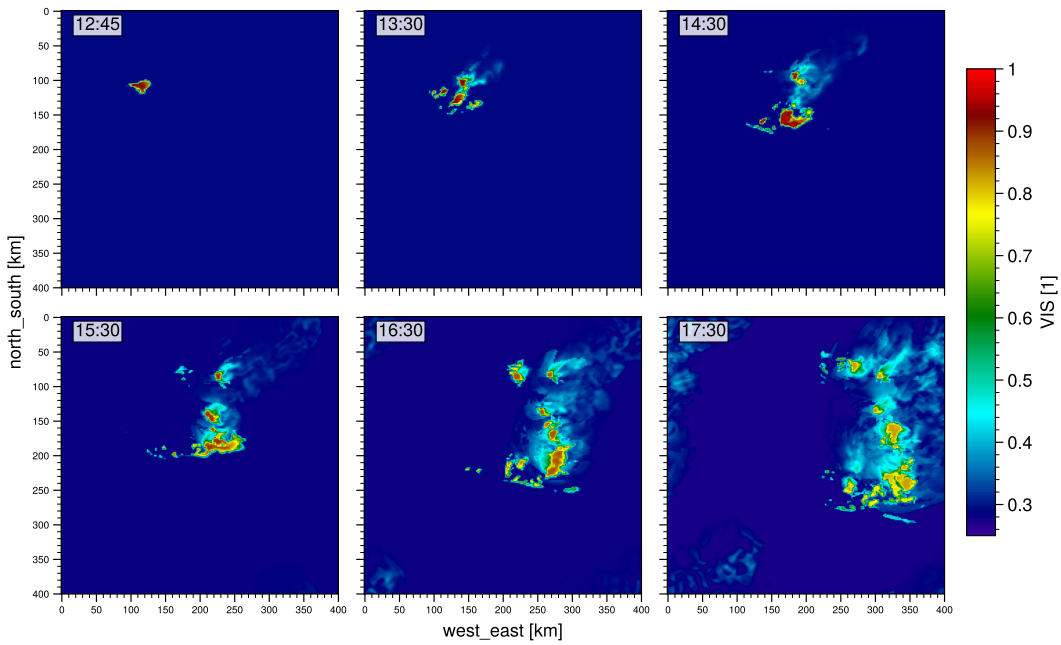


FIGURE 6 Visible 0.6 µm satellite images of the warm-bubble case's nature run

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. Meanwhile, $\alpha = 0$ 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. Although it is common in operational DA, we did not reject observation outliers as we saw relatively strong error reduction by observations with large first-guess difference.

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 150 x 150 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 to 14 km.

Superobbing can be a useful approach to assimilate high-resolution observations as it averages observations towards the resolved scale of the model. However, contrary to our first expectations, we found that superobbing 5 x 5 observations towards 10 km resolution did not generally improve forecasts. As the difference in impact was negligible, we decided not to include those experiments in this manuscript.

The assimilation of satellite observations in a Gaussian filter is sub-optimal for reasons of non-Gaussianity, like heteroscedasticity (the increase of variance with cloudiness) or boundedness. Additionally, non-linear observations

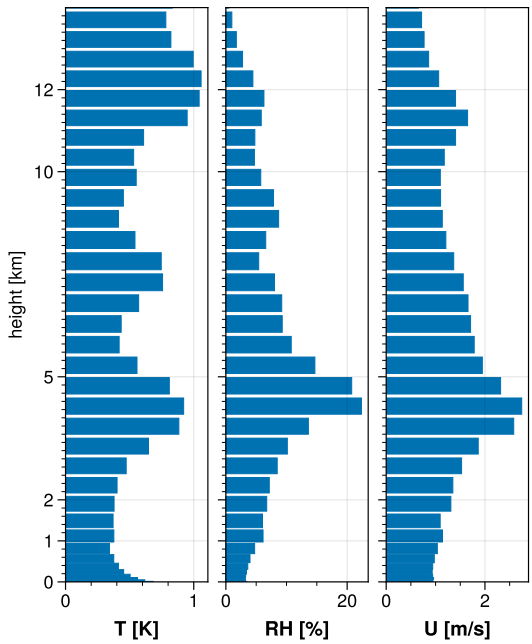


FIGURE 7 Vertical profile of ensemble spread (horizontal average) of temperature, relative humidity and the u-wind component at 13:00 prior to the first assimilation in the "random" case.

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 7 UTC and ran freely without assimilation for six hours. By 13 UTC, the model had generated a sufficient amount of spread (figure 7). From 13 to 14 UTC we assimilated five times (every 15 minutes), followed by free forecasts until 18 UTC.

In the warm-bubble case, we started to assimilate at 12:30 after a free forecast of 30 minutes. Despite of this short spin up time, the deep convection had already developed. From 12:30 to 13:30 UTC we assimilated five times, followed by free forecasts until 18 UTC. Although the assimilation window is one hour in both cases, it covers different phases of convection in each case.

Figure 9 shows the timeseries of the strongest cloud signal in each observation type, i.e. the lowest value 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.

Abbreviation	Assimilated observation type	σ generate	σ assimilate (range of tested values)
VIS	Visible reflectance 0.6 µm	0.03	0.03 (0.03-0.12)
WV62	Brightness temperature 6.2 µm	1 K	1 K (1-3)
WV73	Brightness temperature 7.3 µm	1 K	1 K (1-3)
REFL	Radar reflectivity 10 cm	2.5 dBz	2.5 dBz (2.5-7.5)
NoDA	none	-	-

TABLE 1 Experiments and their assimilated variable together with standard errors for generating observations (measurement error) and assimilating observations (assigned error). The range of tested values for the assigned error is indicated in parentheses.

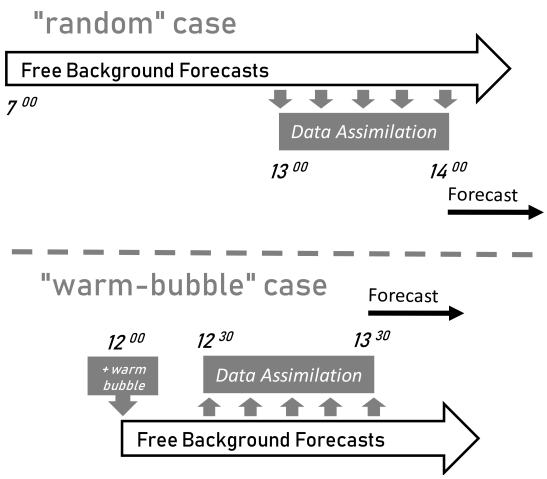


FIGURE 8 Timeline of forecasts and assimilation in the ”random” and the ”warm-bubble” case.

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 h 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 analyse 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 to 2D satellite observations in the case of random convection in section 3.3.

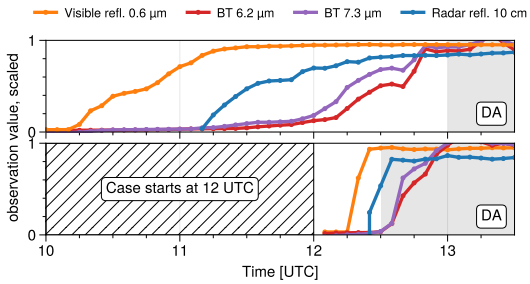


FIGURE 9 Timeseries of observations in the "random" case (top) and the "warm-bubble" case (bottom); Shown is the strongest cloud signal in each observation type, i.e. domain min of BT and max of reflectance. Values are scaled to the range 0-1 from their respective ranges: 0.27-1 for visible reflectance; 235-205 K for 6.2 µm BT; 255-205 K for 7.3 µm BT; 15-80 dBz for radar reflectivity. The assimilation time frame is shown as grey background.

3.1 | Relative potential impact

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

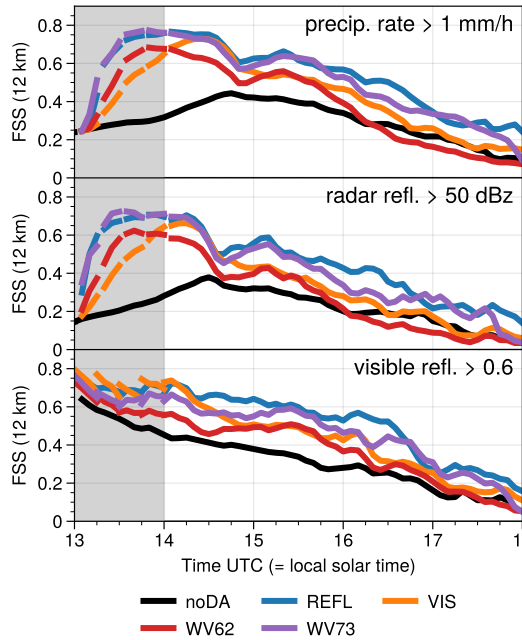
- Precipitation rate > 1 mm/h
- Radar reflectivity > 50 dBz
- Visible reflectance > 0.6

The 12 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 e.g. (Scheck et al., 2020) who calculated the FSS from the ensemble mean.

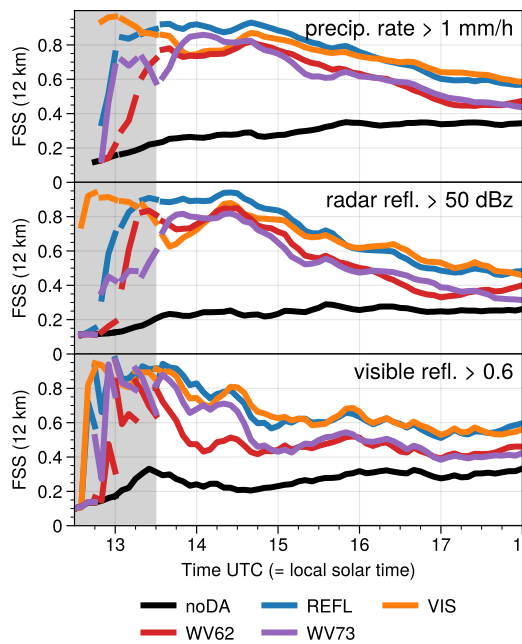
Case "random"

Figure 10a shows the impact of assimilating four different observation types in the case with deep convection randomly scattered throughout the whole domain. Compared to the REFL experiment and averaged over 14 - 17 UTC, the VIS experiment revealed an FSS improvement of 50%, the WV73 experiment 79% and the WV62 experiment 20% for the prediction of radar reflectivity > 50 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 as the VIS experiment in the first 1.5 h lead time, but provided better forecasts afterwards. The WV62 experiment's forecast skill was the lowest of all observation types. It seems that channels which see deeper into the atmosphere (visible and 7.3 µm) have a higher impact than the 6.2 µm channel which does not sense lower tropospheric vapor and clouds. Overall, forecasts in the REFL experiment were best, with 2.5 hours of skillful forecasts for light precipitation and 1.5 hours 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 RMSE of visible reflectance and 7.3 µm brightness temperature forecasts, relative to the RMSE of the noDA experiment. Specifically, we computed the RMS error of the ensemble mean forecast over all 200 x 200 grid points, while only 31 x 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 µm BT. After the analysis however, the RMSE of the WV73 experiment

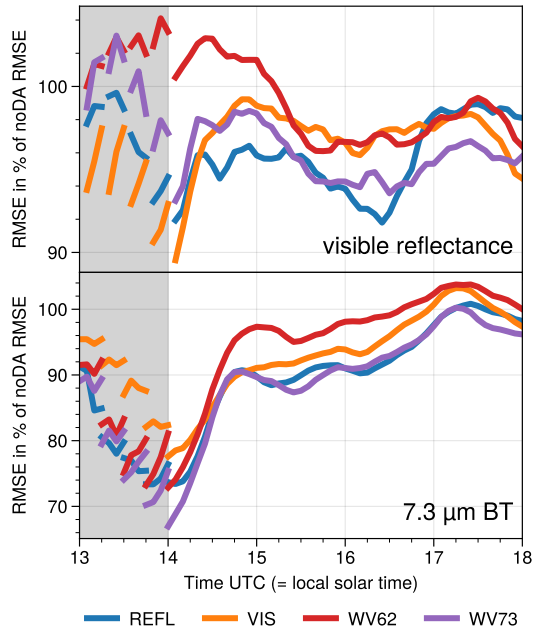


(a) random case

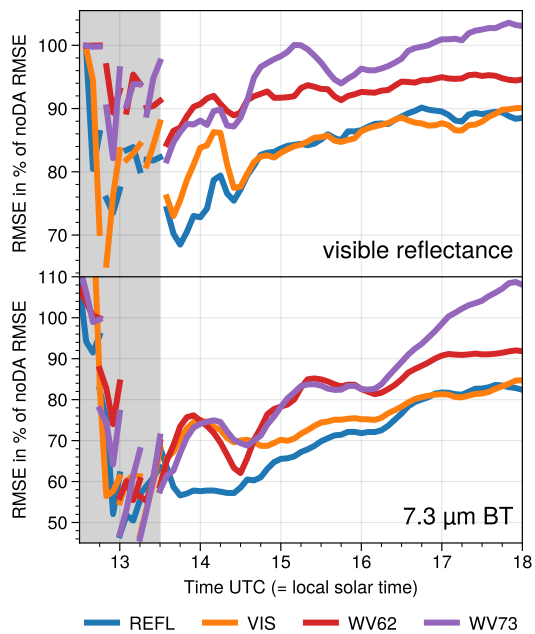


(b) warm-bubble case

FIGURE 10 Fraction skill score for precipitation > 1 mm/hr; Radar reflectivity > 50 dBz and Visible reflectance > 0.6 in (a) the "random" case and (b) the warm-bubble case. Assimilation time frame is shown in grey shade.



(a) random case



(b) warm-bubble case

FIGURE 11 RMSE of ensemble mean forecasts of the visible (upper panel) and the 7.3 µm channel (lower panel), normalized by the RMSE of the noDA control run (a) for the "random" case and (b) for the warm-bubble case; the average is taken horizontally over 200x200 grid points. Assimilation time frame is shown in grey shade.

was similar to the error of the REFL experiment. The experiments REFL, VIS and WV73 overall showed similar skill for predicting the visible channel, while the WV62 experiment had lower skill. The VIS experiment had relatively good forecasts of 7.3 µm BT and the WV73 experiment had good forecasts of visible reflectance. The WV62 experiment had less accurate forecasts of both 7.3 µm BT and visible reflectance, which is presumably related to its higher peak of its weighting function leading to smaller sensitivity to low and mid-level clouds.

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 to the noDA experiment, but some aspects should be noted: Firstly, 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. Yet, the VIS experiment overtook the REFL experiment again at around 3 hours lead time in precipitation scores. Secondly, 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 µm channel is more sensitive to higher tropospheric water vapor, while the 7.3 µm channel is more sensitive to lower tropospheric water vapor. Thirdly, 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 does not show a substantial advantage over the REFL experiment in the FSS for cloudiness, although it directly assimilates visible reflectance. For cloudiness, both the REFL and the VIS experiment gave similar performance. Compared to 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 forecasted 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. As for visible reflectance, the REFL experiment had the lowest error in 7.3 µm BT, 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 second best RMSE in 7.3 µm BT, followed by WV62 and WV73.

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 four hours (FSS>0.5) in the warm-bubble case, the random case was skillfully predicted for only 1.5 hours. 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 to the REFL experiment. Overall, satellite observations lead to a remarkable impact given that the satellite experiments assimilated only 1/7-th of the number of observations compared to 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.

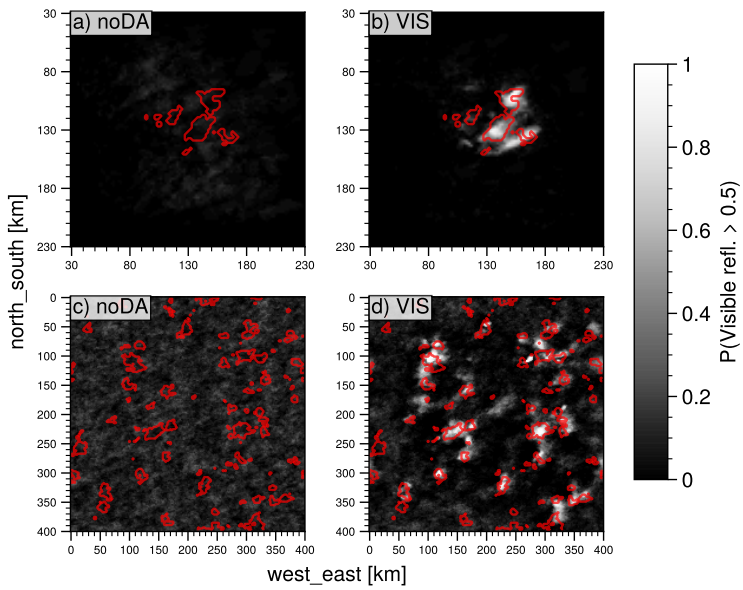


FIGURE 12 The probability for visible reflectance > 0.5, in the noDA run (a,c) and after the assimilation in the VIS experiment (b,d). The ensemble-derived probability ranges black (0) to white (1) and the nature ($\mathcal{H}(x_{nat}) > 0.5$) is shown in red contours. The warm-bubble case at 13:35 is shown in (a,b). The "random" case at 14:05 is shown in (c,d).

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 (Schröttle et al., 2020) compared the assimilation of visible (0.6 µm) and infrared observations (6.2 µm) in a convective-scale NWP model. Despite the similar setup, we clearly see more impact from assimilating visible observations than from 6.2 µm BT observations. This contrasts with (Schröttle et al., 2020) who found stronger impact from the 6.2 µm channel and less impact from visible observations. Yet, this difference might be related to overly inflated observation errors for visible observations in that study, as they assigned an up to 10 times inflated observation error for visible but not 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 µm channel compared to the 7.3 µm 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 µm 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 by assimilating 2-dimensional instead of 3-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.

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

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

3.2 | Impact on model state variables

To understand how model variables were updated by the assimilation, figure 13 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 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 14:05 in the "random" case, five minutes after the last assimilation, as mean (over 961 observed atmospheric columns) absolute error of the ensemble mean forecast.

The temperature error profile (figure 13a) shows four peaks, at the surface, at 5 km, 8 km and at 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 13b), 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 one and three kilometres.

Somewhat surprisingly, the vertical distribution of cloud water (WRF's *QCLOUD* variable) was not generally improved (figure 13c). Most layers show increased errors compared to the noDA experiment. Only the layer with the highest errors shows slightly reduced errors in the REFL experiment, that 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 (figure 11,10) 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 g/kg (40%) in the experiments assimilating 6.2 or 7.3 µm, 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 µm 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 are not very sensitive to the vertical distribution of hydrom-

eteors. This means it is not necessary for the cloud to 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 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

Why radar outperforms satellite observations

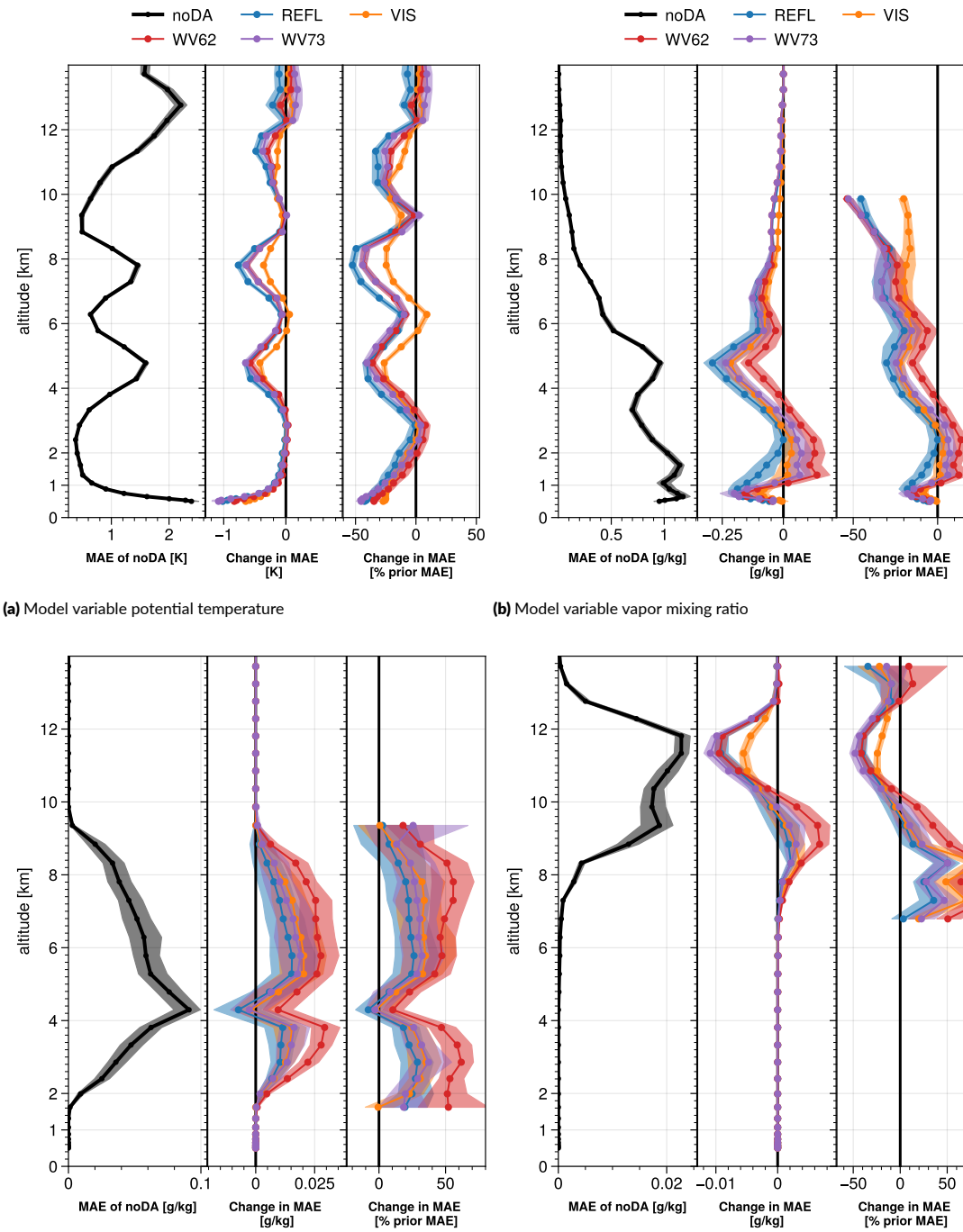
A prominent detail in the results is that radar observations have an advantage over satellite observations in the random case, but not so much in the warm-bubble case. Given the differences between the cases, we hypothesized that the advantage of radar observations comes from their vertical resolution. 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 14) shows a forecast which is relatively similar to experiments which assimilated satellite observations, indicating that vertically resolved observations are indeed crucial in this case in order to reach a high forecast skill.

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 which performed best in terms of the FSS. Figure 15 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 measurement 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 to other experiments and lead to a forecast impact that was mostly between the experiments WV73 and VIS, yet still higher than the WV62 experiment after 16.30 UTC. Cloud-affected BTs of 7.3 µm show much larger first guess deviations than in the 6.2 µm channel, a possible reason why assigning 2 K lead to better forecasts after 3 hours lead time.

Dynamic observation errors

For infrared satellite observations, the first-guess departures increase with the occurrence of clouds, mainly due 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 sub-optimal, especially for 7.3 µm 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 µm 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 might be underestimated ensemble spread for cloudy observations. To investigate this further, we compare the ensemble spread to the RMSE of the ensemble mean forecast for each observation type in figure 16. It seems that the prior error variance was well estimated in the warm-bubble case for all observation types (figure 16a). In the ”random” case (figure 16b) however, the spread was underestimated, probably due to an underestimated spread for cloudy observations, as can be seen by the spread-error relationship for the subset of cloudy observations in figure 16c. This



(a) Model variable potential temperature

(b) Model variable vapor mixing ratio

(c) Model variable cloud water mixing ratio

(d) Model variable cloud ice mixing ratio

FIGURE 13 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 to noDA. The right panel shows the change in MAE, relative to the prior MAE. The error was evaluated at 14:05 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.

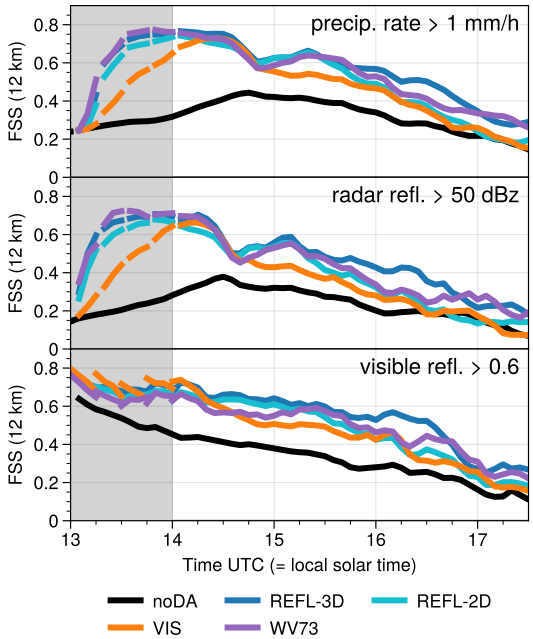


FIGURE 14 Fraction skill score of forecasts assimilating 2-dimensional radar reflectivity instead of 3-dimensional radar reflectivity in the random case. Assimilation time frame is shown in grey shade.

could explain why the experiments with the dynamic observation error model performed worse than the constant error assimilation, since the underestimated spread together with the inflated observation error lead to negligible weights for cloudy observations. 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.

4 | CONCLUSIONS

This study presents the first direct comparison of the assimilation of visible and infrared satellite observations to 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 µm visible reflectance as well as 6.2 and 7.3 µm infrared brightness temperature and radar reflectivity in an idealized perfect-model OSSE. The forecast impact was evaluated in two weather situations. Firstly, a "supercell" case in which a warm bubble of 30 km diameter initiated a single supercell storm and secondly a case where multiple deep convective cells at different stages are scattered throughout the domain. The periodic boundary domain of 400 x 400 km 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 one hour, during the growth and consecutive mature stage of convection.

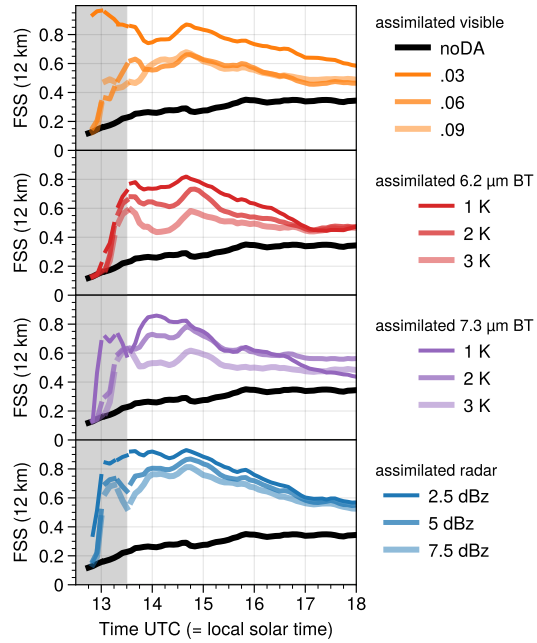


FIGURE 15 Fraction skill scores for light precipitation (>1 mm/hr) in the warm-bubble case using assigned observation errors of 1-3 times the measurement error for observations of (a) visible reflectance, (b) 7.3 µm infrared BT or (c) radar reflectivity. Assimilation time frame is shown in grey shade.

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 substantially. 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, i.e. 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 6.2 and 7.3 µm lead to 74-76% of the radar impact. In more difficult conditions, i.e. randomly located storms at different stages ("random" case), the relative impact was lower but still reached 50% for the visible observations, 20% for 6.2 µm BT and 79% for 7.3 µm 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 µm BT observations is highest when the uncertainty about the vertical structure of clouds is lowest. The vertical structure of clouds can not be retrieved from a single channel and thus is a weak spot for satellite observations. Comparing the

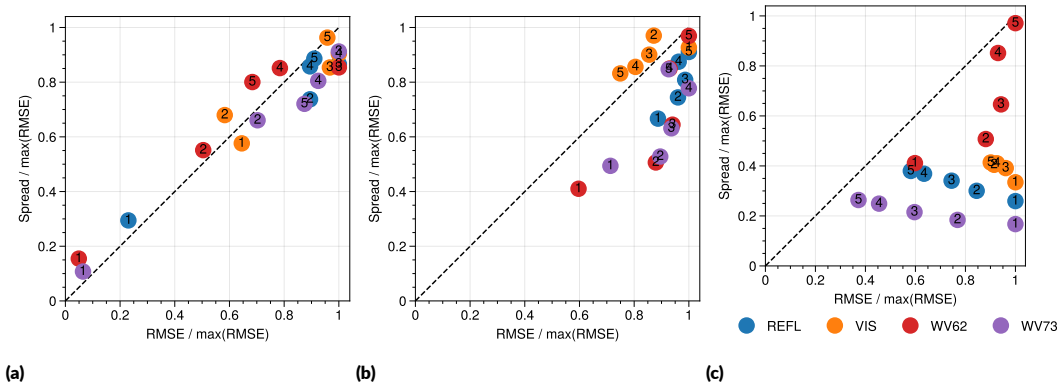


FIGURE 16 Spread error relationship for a) the warm-bubble case, b) the random case and c) the random case, considering only cloudy observations defined by visible reflectance > 0.6, BT < 240 K, radar reflectivity > 20 dBz. Spread^2 is the prior ensemble variance, averaged over observations. RMSE is $\sqrt{\langle(\overline{H(x_b)} - H(x_{nat}))^2\rangle}$ where x_b is the prior, x_{nat} nature and $\langle\cdot\rangle$ the average over observations. In order to use one axis for different observation types, we re-scaled by dividing by the maximum RMSE for each observation type.

"warm-bubble" and the "random" case, 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 which 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 µm infrared BT.

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 as well as 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 better understand their potential impact on convective-scale forecasts in the absence of all complexities of a real system. (Zhang et al., 2016) suggests 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 e.g. 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), the statistics of first-guess departures support our OSSE setup. The standard deviation of first-guess departures (to single members and not the ensemble mean) was 4.5 K for the 6.2 µm channel, 9.1 K for the 7.3 µm channel, and 0.22 for the

visible channel at 13:30. Compared to (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.

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, only very few studies 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, e.g. 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, e.g., 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. While different channels can provide information about different atmospheric levels, research on this is lacking, despite the potential of acquiring vertically resolved information on different hydrometeor types through the combination of spectral channels. For example, while the 6.2 µm channel is mostly sensitive to upper tropospheric water vapor, the 7.3 µm 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 µm 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.

Data & Software

Experiments were conducted using the python package available at `github.com/lkugler/DART-WRF`. It allows to define experiment workflows using DART and WRF and contains routines to generate DART observation sequence files, 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. 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. 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).

REFERENCES

Altuğ Aksoy, David C. Dowell, and Chris Snyder. 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, June 2009. ISSN 1520-0493, 0027-0644. doi: 10.1175/2008MWR2691.1. URL `https://journals.ametsoc.org/view/journals/mwre/137/6/2008mwr2691.1.xml`. Publisher: American Meteorological Society Section: Monthly Weather Review.

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`. 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.

Jeffrey L. Anderson. A Non-Gaussian Ensemble Filter Update for Data Assimilation. Monthly Weather Review, 138(11): 4186–4198, November 2010. ISSN 1520-0493, 0027-0644. doi: 10.1175/2010MWR3253.1. URL https://journals.ametsoc.org/view/journals/mwre/138/11/2010mwr3253.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Jeffrey L Anderson. Localization and sampling error correction in ensemble Kalman filter data assimilation. Monthly Weather Review, 140(7):2359–2371, 2012.

Jeffrey L. Anderson. A Marginal Adjustment Rank Histogram Filter for Non-Gaussian Ensemble Data Assimilation. Monthly Weather Review, 148(8):3361–3378, July 2020. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-19-0307.1. URL https://journals.ametsoc.org/view/journals/mwre/148/8/mwrD190307.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Jeffrey L. Anderson. A Quantile-Conserving Ensemble Filter Framework. Part I: Updating an Observed Variable. Monthly Weather Review, 150(5):1061–1074, May 2022. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-21-0229.1. URL http://journals.ametsoc.org/view/journals/mwre/150/5/MWR-D-21-0229.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Kevin Bachmann, Christian Keil, and Martin Weissmann. Impact of radar data assimilation and orography on predictability of deep convection. Quarterly Journal of the Royal Meteorological Society, 145(718):117–130, 2019.

Kevin Bachmann, Christian Keil, George C Craig, Martin Weissmann, and Christian 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):63–81, 2020.

Stanley G. Benjamin, Stephen S. Weygandt, John M. Brown, Ming Hu, Curtis R. Alexander, Tatiana G. Smirnova, Joseph B. Olson, Eric P. James, David C. Dowell, Georg A. Grell, Haidao Lin, Steven E. Peckham, Tracy Lorraine Smith, William R. Moninger, Jaymes S. Kenyon, and Geoffrey S. Manikin. A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. Monthly Weather Review, 144(4):1669–1694, April 2016. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-15-0242.1. URL https://journals.ametsoc.org/view/journals/mwre/144/4/mwr-d-15-0242.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Rebecca M Cintineo, Jason A Otkin, Thomas A Jones, Steven Koch, and David J Stensrud. 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, 2016.

R. H. Cuenca and M Tewari. Implementation and verification of the unified NOAH land surface model in the WRF model. 2004.

R. M. Errico and N. C. Privé. Some general and fundamental require-ments for designing observing system simulation experiments (OSSEs). Technical report, World Meteorological Organization, April 2018. URL https://community.wmo.int/wwrp-publications.

Keenan C. Eure, Paul D. Mykolajtchuk, Yunji Zhang, David J. Stensrud, Fuqing Zhang, Steven J. Greybush, and Matthew R. Kumjian. Simultaneous Assimilation of Planetary Boundary Layer Observations from Radar and All-Sky Satellite Observations to Improve Forecasts of Convection Initiation. Monthly Weather Review, -1(aop), January 2023. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-22-0188.1. URL https://journals.ametsoc.org/view/journals/mwre/aop/MWR-D-22-0188.1/MWR-D-22-0188.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

NATHAN Faggian, Belinda Roux, Peter Steinle, and Beth Ebert. Fast calculation of the fractions skill score. Mausam, 66(3): 457–466, 2015.

Alan J. Geer and Peter Bauer. Observation errors in all-sky data assimilation. Quarterly Journal of the Royal Meteorological Society, 137(661):2024–2037, 2011. ISSN 1477-870X. doi: 10.1002/qj.830. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.830. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.830.

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.

F. Harnisch, M. Weissmann, and Á Periá\~{n}ez. 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, 2016. ISSN 1477-870X. doi: 10.1002/qj.2776. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2776.

Takumi Honda, Takemasa Miyoshi, Guo-Yuan Lien, Seiya Nishizawa, Ryuji Yoshida, Sachiho A. Adachi, Koji Terasaki, Kozo Okamoto, Hirofumi Tomita, and Kotaro Bessho. Assimilating All-Sky Himawari-8 Satellite Infrared Radiances: A Case of Typhoon Soudelor (2015). Monthly Weather Review, 146(1):213–229, January 2018. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-16-0357.1. URL https://journals.ametsoc.org/view/journals/mwre/146/1/mwr-d-16-0357.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Stephan Hoyer and Joseph J. Hamman. xarray: N-D labeled arrays and datasets in Python. Journal of Open Research Software, 2017. ISSN 2049-9647. doi: 10.5334/jors.148. URL https://opensky.ucar.edu/islandora/object/articles%3A21998/.

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, n/a(n/a):e1130, 2022. ISSN 1530-261X. doi: 10.1002/asl.1130. URL http://onlinelibrary.wiley.com/doi/abs/10.1002/asl.1130.

Michael J. Iacono, Jennifer S. Delamere, Eli J. Mlawer, Mark W. Shephard, Shepard A. Clough, and William D. Collins. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. Journal of Geophysical Research: Atmospheres, 113(D13), 2008. ISSN 2156-2202. doi: 10.1029/2008JD009944. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2008JD009944.

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, Kristopher 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.

Lukas Kugler. Potential impact of assimilating visible and infrared satellite observations compared to radar reflectivity for convective-scale NWP, April 2023. URL https://doi.org/10.5281/zenodo.7840304. Version Number: v1.

Heiner Lange and George C Craig. The impact of data assimilation length scales on analysis and prediction of convective storms. Monthly Weather Review, 142(10):3781–3808, 2014.

Heiner Lange, George C Craig, and Tijana Janjić. Characterizing noise and spurious convection in convective data assimilation. Quarterly Journal of the Royal Meteorological Society, 143(709):3060–3069, 2017.

Robert A. Maddox, Jian Zhang, Jonathan J. Gourley, and Kenneth W. Howard. Weather Radar Coverage over the Contiguous United States. Weather and Forecasting, 17(4):927–934, August 2002. ISSN 1520-0434, 0882-8156. doi: 10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/wefo/17/4/1520-0434_2002_017_0927_wrcotc_2_0_co_2.xml. Publisher: American Meteorological Society Section: Weather and Forecasting.

Steven M. Martinaitis, Andrew P. Osborne, Micheal J. Simpson, Jian Zhang, Kenneth W. Howard, Stephen B. Cocks, Ami Arthur, Carrie Langston, and Brian T. Kaney. A Physically Based Multisensor Quantitative Precipitation Estimation Approach for Gap-Filling Radar Coverage. Journal of Hydrometeorology, 21(7):1485–1511, June 2020. ISSN 1525-7541, 1525-755X. doi: 10.1175/JHM-D-19-0264.1. URL https://journals.ametsoc.org/view/journals/hydr/21/7/jhmD190264.xml. Publisher: American Meteorological Society Section: Journal of Hydrometeorology.

Mikio Nakanishi and Hiroshi 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):397–407, May 2006. ISSN 1573-1472. doi: 10.1007/s10546-005-9030-8. URL https://doi.org/10.1007/s10546-005-9030-8.

Jason A. Otkin. 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, February 2012a. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR-D-11-00084.1. URL https://journals.ametsoc.org/view/journals/mwre/140/2/mwr-d-11-00084.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Jason A. Otkin. 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):n/a–n/a, October 2012b. ISSN 01480227. doi: 10.1029/2012JD017568. URL http://doi.wiley.com/10.1029/2012JD017568.

R. A. Roebeling, E. L. A. Wolters, J. F. Meirink, and H. Leijnse. 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, October 2012. ISSN 1525-7541, 1525-755X. doi: 10.1175/JHM-D-11-089.1. URL https://journals.ametsoc.org/view/journals/hydr/13/5/jhm-d-11-089_1.xml. Publisher: American Meteorological Society Section: Journal of Hydrometeorology.

Elena Saltikoff, Günther Haase, Laurent Delobbe, Nicolas Gaussiat, Maud Martet, Daniel Idziorek, Hidde Leijnse, Petr Novák, Maryna Lukach, and Klaus Stephan. OPERA the Radar Project. Atmosphere, 10(6):320, June 2019. ISSN 2073-4433. doi: 10.3390/atmos10060320. URL https://www.mdpi.com/2073-4433/10/6/320. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

Roger Saunders, James Hocking, Emma Turner, Peter Rayer, David Rundle, Pascal Brunel, Jerome Vidot, Pascale Roquet, Marco Matricardi, Alan Geer, Niels Bormann, and Cristina Lupu. An update on the RTTOV fast radiative transfer model (currently at version 12). Geoscientific Model Development, 11(7):2717–2737, July 2018. ISSN 1991-959X. doi: 10.5194/gmd-11-2717-2018. URL https://gmd.copernicus.org/articles/11/2717/2018/. Publisher: Copernicus GmbH.

Yohei Sawada, Kozo Okamoto, Masaru Kunii, and Takemasa Miyoshi. Assimilating Every-10-minute Himawari-8 Infrared Radiances to Improve Convective Predictability. Journal of Geophysical Research: Atmospheres, 124(5):2546–2561, 2019. ISSN 2169-8996. doi: https://doi.org/10.1029/2018JD029643. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029643.

Leonhard Scheck, Pascal Frèrebeau, Robert Buras-Schnell, and Bernhard Mayer. A fast radiative transfer method for the simulation of visible satellite imagery. Journal of Quantitative Spectroscopy and Radiative Transfer, 175:54–67, May 2016. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2016.02.008. URL https://www.sciencedirect.com/science/article/pii/S0022407316000285.

Leonhard Scheck, Martin Weissmann, and Bernhard Mayer. 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, 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 Atmospheric and Oceanic Technology.

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, 2020. ISSN 1477-870X. doi: https://doi.org/10.1002/qj.3840. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3840.

Josef Schröttle, 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.

Craig S. Schwartz, Glen S. Romine, and David C. Dowell. 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, April 2021. ISSN 0882-8156. doi: 10.1175/WAF-D-20-0110.1. Place: Boston Publisher: Amer Meteorological Soc WOS:000644093400002.

C. Skamarock, B. Klemp, Jimy Dudhia, O. Gill, Zhiquan Liu, Judith Berner, Wei Wang, G. Powers, G. Duda, Dale Barker, and Xiang-yu Huang. A Description of the Advanced Research WRF Model Version 4.3. Technical report, NCAR/UCAR, 2021. URL http://dx.doi.org/10.5065/1dfh-6p97.

Chris Snyder and Fuqing Zhang. Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter. Monthly Weather Review, 131(8):1663–1677, August 2003. ISSN 1520-0493, 0027-0644. doi: 10.1175//2555.1. URL https://journals.ametsoc.org/view/journals/mwre/131/8/_2555.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Gregory Thompson, Paul R. Field, Roy M. Rasmussen, and William 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):5095–5115, December 2008. ISSN 1520-0493, 0027-0644. doi: 10.1175/2008MWR2387.1. URL https://journals.ametsoc.org/view/journals/mwre/136/12/2008mwr2387.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

Mingjing Tong and Ming Xue. Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. Monthly Weather Review, 133(7):1789–1807, July 2005. ISSN 1520-0493, 0027-0644. doi: 10.1175/MWR2898.1. URL https://journals.ametsoc.org/view/journals/mwre/133/7/mwr2898.1.xml. Publisher: American Meteorological Society Section: Monthly Weather Review.

UCAR/NCAR/CISL/DAReS. The Data Assimilation Research Testbed (Version 10.5.3) [Software], 2022.

Dustan M. Wheatley, Kent H. Knopfmeier, Thomas A. Jones, and Gerald J. Creager. 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, December 2015. ISSN 1520-0434, 0882-8156. doi: 10.1175/WAF-D-15-0043.1. URL https://journals.ametsoc.org/view/journals/wefo/30/6/waf-d-15-0043_1.xml. Publisher: American Meteorological Society Section: Weather and Forecasting.

Fuqing Zhang, Masashi Minamide, and Eugene E. Clothiaux. 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, 2016. ISSN 1944-8007. doi: 10.1002/2016GL068468. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068468.

Lijian Zhu, Ming Xue, Rong Kong, and Jinzhong Min. 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, -1(aop), December 2022. ISSN 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 Review.