



**HAL**  
open science

# Assimilating visible and infrared radiances in idealized simulations of deep convection

Josef Schrötle, Martin Weissmann, Leonhard Scheck, Axel Hutt

► **To cite this version:**

Josef Schrötle, Martin Weissmann, Leonhard Scheck, Axel Hutt. Assimilating visible and infrared radiances in idealized simulations of deep convection. *Monthly Weather Review*, 2020, 10.1175/MWR-D-20-0002.1 . hal-02923890

**HAL Id: hal-02923890**

**<https://inria.hal.science/hal-02923890>**

Submitted on 27 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



1 **Assimilating visible and infrared radiances in idealized**

2 **simulations of deep convection**

3 Josef Schröttele \*

4 *Meteorologisches Institut, Ludwig-Maximilians Universität, Munich, Germany and Department*  
5 *of Geophysics, School of Earth Sciences, Tel Aviv University, Israel*

6 Martin Weissmann

7 *Institut für Meteorologie und Geophysik, Universität Wien, Vienna, Austria*

8 Leonhard Scheck

9 *Hans-Ertel Zentrum für Wetterforschung, Ludwig-Maximilians Universität, Munich, Germany*

10 Axel Hutt

11 *Deutscher Wetterdienst, Offenbach, Germany and Team MIMESIS, INRIA Nancy Grand Est,*  
12 *Strasbourg, France*

13 \*Corresponding author address: Department of Geophysics, School of Earth Sciences, Tel Aviv  
14 University, Israel.

15 E-mail: josefs@mail.tau.ac.il

**Early Online Release:** This preliminary version has been accepted for publication in *Monthly Weather Review*, may be fully cited, and has been assigned DOI 10.1175/MWR-D-20-0002.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

## ABSTRACT

16 Cloud-affected radiances from geostationary satellite sensors provide the  
17 first area-wide observable signal of convection with high spatial resolution in  
18 the range of kilometers and high temporal resolution in the range of minutes.  
19 However, these observations are not yet assimilated in operational convection-  
20 resolving weather prediction models as the rapid, non-linear evolution of  
21 clouds makes the assimilation of related observations very challenging.

22 To address these challenges, we investigate the assimilation of satellite radi-  
23 ances from visible and infrared channels in idealized observing system sim-  
24 ulation experiments (OSSEs) for a day with summer-time deep convection  
25 in central Europe. This constitutes the first study assimilating a combination  
26 of all-sky observations from infrared and visible satellite channels and the  
27 experiments provide the opportunity to test various assimilation settings in  
28 an environment, where the observation forward operator and the numerical  
29 model exhibit no systematic errors.

30 The experiments provide insights into appropriate settings for the assimila-  
31 tion of cloud-affected satellite radiances in an ensemble data assimilation sys-  
32 tem and demonstrate the potential of these observations for convective-scale  
33 weather prediction. Both infrared and visible radiances individually lead to  
34 an overall forecast improvement, but best results are achieved with a com-  
35 bination of both observation types that provide complementary information  
36 on atmospheric clouds. This combination strongly improves the forecast of  
37 precipitation and other quantities throughout the whole range of 8 h lead time.

## 38 **1. Introduction**

39 Convective-scale data assimilation aims at improving forecasts of severe weather events, which  
40 are often related to deep convection. The prediction of these events requires not only an accurate  
41 initial state of the large-scale environmental conditions, but also knowledge on the location and  
42 structure of individual convective systems at the km-scale. Cloud-affected satellite observations  
43 from geostationary satellite sensors provide a promising source of information in this context as  
44 they reveal insights into dynamically active regions of the atmosphere (McNally 2002) and cover  
45 a large area with high spatial resolution in the range of kilometres and high temporal resolution  
46 in the range of minutes. Furthermore, clouds are an easily detectable signal of emerging convec-  
47 tive systems that can be observed earlier than larger precipitating hydrometeors that are seen by  
48 weather radars.

49 Observations from different satellite channels provide very complementary information for this  
50 purpose: Water vapor infrared channels are sensitive to water and ice clouds, containing informa-  
51 tion on atmospheric humidity and temperature. The brightness temperature of clouds observed  
52 in these channels provides information on the cloud top height. Due to the absorption by water  
53 vapor these channels peak fairly high, so they are only sensitive to mid- and upper-level clouds.  
54 Infrared window channels can see through the atmosphere, but low clouds are often hard to dis-  
55 tinguish from the surface in these observations. Also, high-level ice clouds are often opaque in  
56 infrared channels leading to a lack of information on water clouds beneath them. Visible channels  
57 are available only during day time. While visible channels are not sensitive to temperature, hu-  
58 midity and cloud top height and less sensitive to ice clouds, they can provide more information on  
59 low-and mid-level clouds. Visible channels allow for a clear distinction between low-level clouds  
60 and the surface (Heinze et al. 2017), unless the latter is covered by snow or ice.

61 Despite this wealth of available information, cloud-affected visible and infrared satellite obser-  
62 vations are not yet assimilated in operational convection-permitting numerical weather prediction  
63 (NWP) models (Gustafsson et al. 2018; Geer et al. 2018). Previous case studies highlighted the  
64 potential benefit of assimilating cloud-affected infrared satellite observations for the prediction of  
65 tropical cyclones (Zhang et al. 2016; Otkin et al. 2017; Honda et al. 2018) and organized convec-  
66 tion (Cintineo et al. 2016; Zhang et al. 2018, 2019) over the continental U.S. using convection-  
67 permitting models. To improve the prediction of local severe weather, infrared radiances were  
68 assimilated above the pacific (Sawada et al. 2019) with 10-min temporal resolution. Scheck et al.  
69 (2020) conducted the first study assimilating a visible satellite channel in a regional model for two  
70 cases with summertime convective precipitation over Germany. The simultaneous assimilation of  
71 visible and infrared channels has not been investigated so far. Furthermore, the impact of these  
72 two observation types on the practical predictability of precipitation has not been compared, yet.

73 The incorporation of cloud-affected microwave satellite radiances in global assimilation systems  
74 has led to significant forecast improvements in recent years (Bauer et al. 2010; Geer et al. 2010,  
75 2017, 2018), but cloud-affected infrared observations are still not assimilated directly yet and  
76 microwave channels are not available on current geostationary satellites. Polar orbiting satellites,  
77 however, do not provide sufficient temporal resolution and coverage for convective-scale data  
78 assimilation in regional models.

79 Challenges for the assimilation of cloud-affected radiances include the errors of forward op-  
80 erators (Scheck et al. 2018), correlated observation errors (Janjić et al. 2017), the non-Gaussian  
81 distribution of errors (Geer et al. 2010), systematic errors in the representation of clouds (Otkin  
82 et al. 2018) and the ambiguity of observed integrated radiation in one channel resulting from the  
83 sensitivity to various model variables (e.g. water clouds, ice clouds, humidity and temperature).  
84 Various methods have been developed to address these challenges. For instance, cloud-dependent

85 error models (Geer et al. 2010) are capable to address the non-Gaussianity of errors. Meanwhile,  
86 the error model initiated by Geer et al. (2010) has been extended for the assimilation of cloud-  
87 affected infrared radiances by Harnisch et al. (2016) and Okamoto et al. (2014). All these error  
88 models are based on error climatologies as functions of cloud impact. The error climatology typ-  
89 ically increases with cloud impact. A different approach is the error model with dynamic obser-  
90 vation error inflation developed by Minamide and Zhang (2017). Observation thinning mitigates  
91 issues due to correlated errors (see e.g. Waller et al. 2016), and recent studies tested the incorpora-  
92 tion of correlated observation errors in data assimilation (Geer 2019). Observational ambiguities  
93 may be mitigated through the combined assimilation of different channels or observation types.

94 To investigate the potential impact of satellite data assimilation and various approaches for their  
95 treatment, Houtekamer and Zhang (2016) suggested to study the optimal use of cloud-affected  
96 radiance measurements in observing system simulation experiments (OSSEs). In an OSSE, a  
97 model simulation is regarded as truth (nature run) and several data assimilation experiments with  
98 synthetic observation simulated from the nature run are conducted that aim to reproduce the nature  
99 run as closely as possible. While Zhang et al. (2016) assimilated cloud-affected radiances in  
100 the infrared with an OSSE, Cintineo et al. (2016) combined infrared and radar observations in  
101 OSSEs. The complex configuration of their OSSE includes, e.g., structured terrain and boundary  
102 conditions from global scale model ensembles.

103 To reduce the complexity and focus on a particularly challenging case with randomly located  
104 convection, we conduct a more idealized OSSE with homogeneous initial conditions and small  
105 random noise to trigger convection following studies for radar data assimilation (Lange and Craig  
106 2014; Bachmann et al. 2019, 2020). In this setup, we neglect orography and land-surface het-  
107 erogeneity. The boundary and initial conditions are perturbed randomly and the statistics of the  
108 perturbations can be reproduced for even larger ensembles without requiring boundary, or initial

109 conditions from larger scale numerical weather prediction models. Our OSSEs are based on the  
110 Payerne sounding measured over Switzerland during a day of deep convection. The convective  
111 clouds evolve throughout the the troposphere in a time scale of  $\approx 1/2$  h. Without topographic fea-  
112 tures, there is no preferential place where deep convection sets in (Bachmann et al. 2019), which  
113 makes the prediction of convection as well as the assimilation of related observations even more  
114 challenging.

115 For data assimilation, we use the local ensemble transform Kalman filter (LETKF; Hunt et al.  
116 (2007)) implemented in the km-scale ensemble data assimilation system KENDA for the opera-  
117 tional regional model COSMO (Consortium for Small-scale Modeling) of Deutscher Wetterdienst  
118 (Schraff et al. 2016). The COSMO-KENDA system is operational at Deutscher Wetterdienst and  
119 has been used for a number of assimilation studies (Schomburg et al. 2015; Necker et al. 2018;  
120 Sommer and Weissmann 2014, 2016; Hutt et al. 2020). To calculate synthetic infrared satellite  
121 observations from the model state, we simulate the cloud-affected infrared radiances with the ra-  
122 diative transfer code RTTOV (Saunders et al. 1999; Matricardi and Saunders 1999). For synthetic  
123 observations in the visible channel, we use the method MFASIS (Method for FAsT Satellite Image  
124 Simulation) recently put forward by Scheck et al. (2016, 2018), which is by now also included in  
125 RTTOV. Compared to the assimilation of conventional observations (Schraff et al. 2016; Necker  
126 et al. 2018), a larger number ( $> 6000$ ) of all-sky radiance measurements can be assimilated every  
127 hour in a model domain covering, e.g., central Europe.

128 Based on these OSSEs, we compare the impact of assimilating cloud-affected radiances from  
129 an infrared water vapor channel and from a visible channel as well as the the combination of both  
130 types. We aim to find appropriate settings for the LETKF to assimilate all-sky satellite observations  
131 for the challenging case of deep convection and address the following questions:

- 132 1. How can we efficiently assimilate cloud-affected radiances during deep convection?
- 133 2. What is the analysis and forecast impact of infrared and visible satellite radiances?
- 134 3. What is the benefit of combining the assimilation of infrared and visible radiances?

135 In the following, Section 2 describes the setup of our OSSEs. Sect. 3 discusses results from  
136 assimilating visible and infrared radiances and Sect. 4 the sensitivity of the results with respect to  
137 changes in the assimilation parameters. Conclusions are provided in Sect. 5.

## 138 **2. Observing system simulation experiments**

139 Nolan et al. (2013) discuss the complexity of simulating a nature run in OSSEs, when surface  
140 heat exchange, structured orography and boundary conditions of global scale numerical weather  
141 prediction are present. To reduce the complexity, we use an idealized setup with a flat domain  
142 and cyclic boundary conditions. This section explains the setup of our OSSEs, shows resulting  
143 fields from the nature run, and provides an impression of the simulations with a focus on synthetic  
144 satellite radiances.

### 145 *a. COSMO-KENDA in an idealized configuration with initial perturbations*

146 Our OSSE setup largely follows previous studies for radar data assimilation (Lange and Craig  
147 2014; Bachmann et al. 2019, 2020) using the COSMO model version 5.3: We initialize wind,  
148 temperature and humidity with a radiosonde profile from Payerne, Switzerland on 30 July 2007  
149 at 12 UTC and add two types of perturbation for each ensemble member to account for the uncer-  
150 tainty on smaller and larger scales (see below). The sounding is from a day with deep convection.  
151 Strong mesoscale convective systems formed on that day (Lange and Craig 2014) due to a high  
152 CAPE of  $2200 \text{ J kg}^{-1}$  and relatively low CIN in a vertical wind shear (see Fig. 1a of Bachmann



153 et al. (2020)). In contrast to the studies undertaken by Lange and Craig (2014) and Bachmann et al.  
154 (2019, 2020), the starting time of the initial forecasts corresponds to the time of the radiosonde  
155 observation. The idealized setup is homogeneous in the horizontal without vegetation or orog-  
156 raphy. The model domain covers a region of  $(L_x, L_y, L_z) = (396 \text{ km} \times 396 \text{ km} \times 22 \text{ km})$  with a  
157 horizontal resolution of  $\Delta x = \Delta y = 2 \text{ km}$ . The model integration time step is 6 s. The vertical  
158 resolution extends from 100 m in the lowest atmospheric layers to 800 m at the domain top and in-  
159 cludes 50 model levels. A Rayleigh damping is applied aloft of 15 km. The model runs with cyclic  
160 horizontal boundary conditions. The Coriolis force is neglected. During the course of the day, the  
161 radiation on the Earth's surface varies with the zenith angle of the sun. In this way, the idealized  
162 setup mimics the weather situation of a typical day with deep convection and a strong influence  
163 of the diurnal cycle. A one-moment cloud microphysics scheme similar to the one developed by  
164 Lin et al. (1983) is used, which includes cloud ice, cloud water, rain, snow and graupel hydrom-  
165 eteors and contains a simplified version of the parametrization of Seifert and Beheng (2001) for  
166 autoconversion, accretion and self-collection. Deep convection is represented explicitly and we do  
167 not apply a shallow convection scheme.

## 168 1) ENSEMBLE PERTURBATIONS AND NATURE RUN

169 To represent initial and boundary condition uncertainty of a regional ensemble system, we add  
170 two types of perturbations to the Payerne sounding to form the initial conditions for the ensemble  
171 members: A vertically correlated perturbation that depends only on height, which is meant to  
172 represent the large-scale uncertainty, and grid scale noise for the uncertainty on smaller scales.

173 As in Lange and Craig (2014), the small-scale component consists of white noise with a standard  
174 deviation of  $0.02 \text{ m s}^{-1}$  for the vertical velocity and 0.02 K for the temperature and is limited  
175 to the lowest 100 hPa. Adding this white noise triggers the development of convective cells.

176 The resulting cell-position is random and completely uncorrelated in space between ensemble  
177 members.

178 For the representation of larger scale errors, we add perturbations on the vertical profiles of the  
179 initial conditions following Bachmann et al. (2020). As the boundary conditions are cyclic, these  
180 perturbations represent both large-scale initial condition errors and boundary condition errors. We  
181 perturb the initial conditions in the vertical and add  $u'_j(z)$ ,  $v'_j(z)$  for wind,  $T'_j(z)$  for temperature,  
182 and  $rh'_j(z)$  for relative humidity for each ensemble member  $j$ . These perturbation profiles are  
183 each drawn from Gaussian random numbers without bias. The vertical correlation length be-  
184 tween the perturbations is between 1 and 3 km. The standard deviations of the perturbations are  
185  $\sigma_u = \sigma_v = 0.25 \text{ m s}^{-1}$  for wind,  $\sigma_T = 0.25 \text{ K}$  for temperature and  $\sigma_{rh} = 2 \%$  for relative humidity.  
186 These random perturbation profiles are added separately for each ensemble member to the initial  
187 conditions.

188 Due to the cyclic boundary conditions, the added random perturbations are sustained within the  
189 domain of each ensemble member and are only subject to diffusion.

190 The initial conditions for the nature run are constructed like the ones for the ensemble members,  
191 but using different random numbers. The nature run is a free forecast initialized at 12 UTC and  
192 will serve as the truth to calculate the errors of the assimilation experiments. The 40-member free  
193 ensemble forecasts (also initialized at 12 UTC) serve as the benchmark to evaluate the relative  
194 improvement by assimilating visible/infrared radiances.

195 In this simplified OSSE setup, we use both the same forecast model and forward operator for  
196 the simulated truth (nature run) and the assimilation/forecast experiments. This has the advantage  
197 to study the assimilation and potential impact of observations in the absence of systematic model,  
198 observation and operator deficiencies, which pose a severe issue for the assimilation of cloud-

199 affected observations in real-world systems. However, this also means that the achieved impact is  
200 likely significantly larger than the impact of such observations in real data assimilation systems.

## 201 2) KENDA DATA ASSIMILATION CONFIGURATION

202 The KENDA assimilation system (Schraff et al. 2016) is operational at Deutscher Wetterdienst  
203 and has been used for a number of assimilation studies (Schomburg et al. 2015; Necker et al. 2018;  
204 Sommer and Weissmann 2014, 2016; Zeng et al. 2019). It is based on a local ensemble transform  
205 Kalman filter (LETKF; Hunt et al. (2007)). As in the operational setup, we use 40 ensemble  
206 members. In the OSSEs, we only assimilate synthetic satellite observations, but no conventional  
207 and radar observations that are usually assimilated in operational assimilation systems. Synthetic  
208  $6.2\mu\text{m}$  SEVIRI images (one of the water vapour channels) are calculated from the nature run  
209 using the RTTOV package (version 10) and visible  $0.6\mu\text{m}$  images are generated using MFASIS.  
210 To represent observation errors, white noise is added to these synthetic satellite observations. This  
211 noise has a standard deviation of 3 K for brightness temperature and of 3 % for visible reflectances.  
212 In our setup, the pixels of the synthetic satellite images correspond to the cells of the horizontal  
213 model grid. While a diurnal variation of solar zenith angle (SZA) is taken into account in the  
214 internal radiative transport scheme for calculating heating rates, a fixed geometry with a SZA of  $8^\circ$ ,  
215 a satellite zenith angle of  $36^\circ$  and a scattering angle of  $152^\circ$  is used for the generation of the satellite  
216 images. Furthermore, it should be noted that we also assimilate visible observations after sunset  
217 in this idealized study, whereas these observations would be limited to daytime in real systems.  
218 We regard these simplification to be justified in this idealised setup, because we are primarily  
219 interested in fundamental properties of the observations like their information content and not in  
220 practical problems related to their systematic errors or their restricted availability. Moreover, a

221 major fraction of summertime convective precipitation does occur during daytime, where visible  
222 observations would be available.

223 The number of satellite observations is reduced by "superobbing", i.e. by averaging the satellite  
224 image on a certain length scale (see e.g. Scheck et al. 2020). For this purpose, the observation  
225 operator is called for each column of the model grid and then the results are averaged over blocks  
226 of 6 by 6 grid cells, corresponding to a superobbing scale of 12 km. Single thunderstorm cells  
227 exhibit a characteristic radius of  $\approx 10$  km during the onset of convection. The averaging area  
228 of  $12 \times 12 \text{ km}^2$  is therefore about the scale of the individual thunderstorm cells. In the standard  
229 data assimilation setup, we use a cycling period of 15 min, corresponding to the time interval be-  
230 tween full disk SEVIRI scans from the standard  $0^\circ$  METEOSAT service. A horizontal averaging  
231 of the measurements to a scale of the storm system must be in accordance with the horizontal  
232 localization (Craig and Würsch 2013). A relatively small horizontal localization ( $L_h = 32$  km) is  
233 chosen with the purpose to draw the ensemble closely to the observations as previously done for  
234 radar data assimilation (Lange and Craig 2014). For the experiments assimilating cloud-affected  
235 observations, we do not localize in the vertical as clouds reveal the convective dynamics of the  
236 whole atmospheric column. Only for the assimilation of clear-sky observations, we conducted  
237 one experiment without vertical localization and one experiment with vertical localization (a log-  
238 arithmic radius of 0.3 hPa around the observation height of 350 hPa). The observation error for  
239 the visible spectral range is set to a constant value of 0.2 in the reference experiments and to 0.3  
240 in further sensitivity experiments. For the infrared water vapor observations, a cloud-dependent  
241 dynamic error model is employed (Sect. 2.e). For the reference experiments, this leads to an as-  
242 signed observation error of 1.1 K for clear-sky observations and an assigned error between 1.5 K  
243 and 6.4 K for cloud-affected observations. Furthermore, sensitivity experiments were performed  
244 with assigned errors increased by 50 %.

245 In contrast to the assimilation experiments by Scheck et al. (2020) and Hutt et al. (2020), no  
246 multiplicative or additive inflation (Zeng et al. 2019) of the error covariance matrix is used. To  
247 conserve positivity of relative humidity, we employ saturation adjustment in the LETKF (Schraff  
248 et al. 2016). The data-assimilation begins at 20 UTC and ranges up to 5 h. We start forecasts  
249 with a lead time of 8 h for each ensemble member from the analysis after 1 h, 3 h, and 5 h of data  
250 assimilation (Fig. 1).

### 251 *b. Overview of assimilation experiments and sensitivity studies*

252 Table 1 summarizes the conducted experiments. These consist of four reference experiments  
253 that are discussed in section 3 and six further sensitivity experiments with modified settings that  
254 are discussed in section 4.

255 The first set of experiments compares the effect of assimilating different instruments and use  
256 a cycling period of 15 min: brightness temperature (BT) with standard error settings, the visi-  
257 ble channel in  $VIS_{oe=0.2}$  with an assigned observation error (OE) of 0.2, and both observation  
258 types with these settings in  $BT+VIS_{oe=0.2}$ . Experiment  $BT_{CA=0}$  assimilates clear-sky brightness  
259 temperature, only.

260 In sensitivity experiments, we increased the assigned observation errors by 50 % for brightness  
261 temperature in the experiment  $BT_{oe*1.5}$  and for visible observations in the experiments  $VIS_{oe=0.3}$   
262 and  $BT+VIS_{oe=0.3}$ . We additionally used 30 min and 60 min as cycling periods for the com-  
263 bined assimilation of brightness temperature and visible reflectance. Furthermore, only clear-sky  
264 brightness temperature was assimilated using vertical localization in experiment  $BT_{CA=0}^{loc}$ .

265 *c. Evolution of the nature run*

266 After the start of the nature run from the perturbed profile described in Sect. 2.a.1, it takes  
267 about 7 hours until the perturbations have grown sufficiently to develop into first convective cells  
268 at around 19 UTC. During this time a thin stratiform cloud layer is present, which forms right after  
269 the begin of the model run and quickly dissolves when convection sets in and air starts to descend  
270 between the convective cells. This cloud layer is probably only an artifact related to deficiencies in  
271 the model radiation and microphysics and we consider it not to be of relevance for the convective  
272 activity we are interested in.

273 In Fig. 2, hourly snapshots from the evolution of the nature run are displayed between 20 UTC  
274 and 1 UTC. The rows of Fig. 2 show brightness temperature in the  $6.2\ \mu\text{m}$  water vapor channel,  
275 visible reflectance in the  $0.6\ \mu\text{m}$  channel, column maximum of radar reflectivity, column maxi-  
276 mum of the cloud ice mixing ratio, and column maximum of the cloud water mixing ratio, respec-  
277 tively (from top to bottom). The snapshots show a representative area of the convection that occurs  
278 horizontally isotropic over the whole domain. It is obvious that the brightness temperature of the  
279 high-peaking water vapor channel is strongly correlated with the cloud ice content and that the  
280 visible reflectances mostly depend on cloud water. There is also some weak contribution from ice  
281 clouds to the visible reflectance. This contribution is much weaker than the one from cloud water,  
282 because the mass of cloud ice in the atmosphere is smaller than the one of cloud water and the ice  
283 particles are larger, which reduces their effectiveness in scattering visible light (see discussion in  
284 Scheck et al. (2020)). The radar reflectivity  $Z$  indicates precipitation and is calculated based on  
285 the prognostic fields of rain, snow, and graupel following Done et al. (2004).

286 In the first column of Fig. 2, i.e. at 20 UTC, remnants of the stratiform cloud layer are still  
287 visible in VIS and QC, but at 21 UTC the layer has completely dissolved. In all rows we see signs

288 of convective activity that increases in the first 2-3 hours and slowly decays afterwards. In BT and  
289 QI we see the increased formation of ice clouds in the first hours. The maxima in QI and the much  
290 smaller-scale structures in Z indicate the location of the cores of the convective cells. The latter are  
291 not clearly identifiable in the infrared images, as the relatively large-scale anvil clouds are opaque  
292 in this channel. In the visible channel the ice clouds are nearly transparent and smaller-scale water  
293 clouds below can be observed. It should be noted that this effect may be exaggerated by too weak  
294 anvil clouds in the model. Water clouds are not only present at the location of convective cores,  
295 but also further away, in some cases outside of the regions covered by anvil clouds. These water  
296 clouds are likely to be a result of gust fronts triggered by cold pools (Lange and Craig 2014; Lange  
297 et al. 2017).

#### 298 *d. Effect of initial perturbations on the ensemble spread*

299 Following Bachmann et al. (2020), we added vertically correlated perturbations of wind, tem-  
300 perature, and relative humidity to the initial profile to represent larger scale errors. Already during  
301 the first hour of the model integration, this leads to significant deviations of CAPE and CIN in the  
302 ensemble members.

303 The initial perturbations enhance the spread of all prognostic variables at later times: The time  
304 when deep convection sets in varies over the ensemble members as can be seen in the ensemble  
305 mean brightness temperature fields - when a cooling sets in in the mean temperature (Fig. 3).  
306 While this cooling occurred due to convection over all ensemble members within a time period  
307 of  $\pm 0.5$  h before adding perturbations to the radiosonde profile (Bachmann et al. 2019, show the  
308 variability of the onset of precipitation), the time period is now extended to  $\pm 1.5$  h with the vertical  
309 variability in the initial conditions (Fig. 3). The onset of the convection is more clearly seen in  
310 the visible channel. The mean reflectance of most members drops at 20 UTC from  $\approx 0.7$  to  $\approx 0.4$ .

311 The decrease in mean reflectance is due to the breakup of the stratus layer during the onset of  
 312 convection. As deep convective clouds form, after 20 UTC, the brightness temperature decreases  
 313 from  $\approx 236$  K to  $\leq 232$  K in all members. One ensemble member forms deep convective clouds  
 314 already earlier at 16 UTC.

315 *e. Observation error model for brightness temperature*

316 To account for the non-Gaussianity of the first guess departures mainly caused by the presence of  
 317 clouds we apply the cloud-dependent error model developed by Harnisch et al. (2016) to efficiently  
 318 assimilate cloud-affected radiances. In this approach the assigned error is increased for cases  
 319 in which the observed brightness temperature or its model equivalent is smaller than a limiting  
 320 brightness temperature  $BT_{lim}$ , which is used to distinguish between clear-sky and cloudy situations.  
 321  $BT_{lim}$  mainly depends on the satellite channel. Here we focus on the  $6.2 \mu\text{m}$  water vapour channel.

322 A number of parameters, such as limiting brightness temperature  $BT_{lim}$ , cloud impact  $C_a$ , and  
 323 dynamic error variance  $\sigma_e^2$  of the model are defined in the following. In addition, a brief overview  
 324 of the error model for assimilating cloud-affected radiances in the context of convective-scale  
 325 ensemble data assimilation is provided.

We consider the simulations for one satellite channel. The respective brightness temperature  $BT_x$  is calculated for each field-of-view (FOV), i.e., coordinate  $(\tilde{x}, \tilde{y})$ . A distribution of brightness temperatures results over all ensemble members and all FOVs. The radiative transfer model can also calculate the corresponding distribution, without the presence of clouds, i.e., without taking into account the cloud absorption and cloud induced scattering of radiation. The calculated brightness temperature for so called clear-sky radiative transfer and each field of view is  $BT_x^{clear}$ . To derive  $BT_{lim}$ , the  $BT_x$  values are grouped into classes. The member of each class  $G$  represents a certain brightness temperature  $BT$  within the respective limits  $[BT_-^G, BT_+^G]$ . We choose 0.1 K wide bins



for each class. For all members within the class, the clear-sky brightness temperature is subtracted and the mean difference is calculated:

$$\Delta BT_x = \frac{1}{M_G} \sum_{g \in G} (BT_{x,g} - BT_{x,g}^{clear}).$$

326 In this way, monotonously increasing brightness temperatures are mapped to a discrete function  
 327  $\Delta BT_x$  and  $M_G$  is the number of elements within the class  $G$ . The brightness temperature, where  
 328  $\Delta BT_x$  decreases below a certain threshold of, e.g.,  $-0.1$  K, defines  $BT_{lim}$ . Following these defini-  
 329 tions,  $BT_{lim}$  can be understood physically as the value, where clouds begin to affect the brightness  
 330 temperature over all FOVs and all ensemble members on average by less than the chosen threshold.

When the limiting brightness temperature  $BT_{lim}$  is known, the cloud impact can be calculated. The cloud impact can be defined separately as  $C_x$  for the modeled and as  $C_y$  for the observed cloud fields:

$$C_{x,ij} = \max(0, BT_{lim} - BT_{x,ij}),$$

$$C_{y,ij} = \max(0, BT_{lim} - BT_{y,ij}).$$

The combination of both values gives the cloud impact

$$C_{a,ij} = (C_{x,ij} + C_{y,ij})/2.$$

331  $i$  is a running index over each FOV, i.e., coordinate  $(\tilde{x}, \tilde{y})$ , and  $j$  is a running index over all ensemble  
 332 members.

333 The cloud-impact values range from 0 K to  $\approx 25$  K in our simulations. The resulting cloud  
 334 impact values are classified to a class  $K$  with a value of cloud impact  $C_{a,ij} \in [C_a^{K-}, C_a^{K+}]$ . The  
 335 width of each cloud-impact class is 1 K, following Harnisch et al. (2016).

The difference between measured and simulated brightness temperature values gives the so called first-guess departure (FGD) values:

$$FGD_{ij} = H(\mathbf{X}_{ij}) - Y_{ij},$$

336 where  $\mathbf{X}$  is the model state vector,  $H$  is the forward operator,  $Y$  is the observed radiance,  $i$  is  
337 mapped to a field of view as follows:  $i \mapsto (\tilde{x}, \tilde{y})$ ;  $k$  is mapped to  $i$  and an ensemble member  $j$  as  
338 follows  $k \mapsto (i, j)$ .

The variance for each class  $K$  is defined as

$$(\sigma_e^K)^2 = \frac{1}{N} \sum_{k \in K} FGD_k^2,$$

339 where  $N$  is the number of elements in the class  $K$ . A histogram over all departures results for each  
340 class  $K$ . The members of each class are normalized with the corresponding  $\sigma_e^K$ . This leads to  
341 a modified FGD histogram (Fig. 4). The resulting distributions are more Gaussian and therefore  
342 more suitable for data assimilation.

343 Notably, the FGD histograms in the idealized deep convection are wider than the ones calcu-  
344 lated by Harnisch et al. (2016) in their figure 4. We attribute this to the deep convective clouds  
345 that show a clear contrast to the warmer ground and the resulting strong FGDs at cloud edges.  
346 The distributions peak at small values, where either clear-sky or cloudy conditions occur in both  
347 the simulated observations as well as in the ensemble member. The error model leads to more  
348 Gaussian all-sky departures after the first cycle when the convection is not completely uncorre-  
349 lated anymore between ensemble members. Small clear-sky departures occur especially in early  
350 assimilation cycles during the first hour. At later times, when clouds have formed in all ensemble  
351 members, the troposphere is more mixed. The corresponding first-guess departures of clear-sky  
352 radiances exhibit a wide range of clear-sky values also following a Gaussian.

### 3. Results from assimilating visible and infrared radiances

This section focuses on the comparison of the four main assimilation experiments. The first one (BT) assimilates brightness temperatures in the infrared  $6.2\ \mu\text{m}$  channel with standard error settings, the second one ( $\text{VIS}_{oe=0.2}$ ) visible reflectance in the  $0.6\ \mu\text{m}$  channel with a constant assigned error of 0.2, and the third one ( $\text{BT}+\text{VIS}_{oe=0.2}$ ) both observation types with these error settings. Finally, experiment  $\text{BT}_{CA=0}$  assimilates clear sky brightness temperature in the infrared  $6.2\ \mu\text{m}$  channel with standard error settings, i.e., an error of 1.1 K. The discussion of further sensitivity experiments with modified settings follows in section 4.

#### *a. Impact during data assimilation cycling*

Fig. 5 shows time series of the evolution of the mean absolute error of the LETKF mean prior (15-min forecast) during the 5-h assimilation period for cloud ice (QI), cloud water (QC), water vapor (QV), meridional wind (V), and temperature (T) of the free forecast experiment and the three assimilation experiments. In this idealized setup, the zonal wind behaves similarly to the meridional wind and is not shown in the following.

The clear-sky data-assimilation experiment  $\text{BT}_{CA=0}$  assimilates 3162 observations in the first and 1257 observations in the second cycle, while all-sky experiments assimilate all available radiance observations over the whole domain. Without data assimilation, the error in all variables approximately doubles in the first 1-2 h, reaches its maximum after 1-3 h and decreases afterwards particularly for cloud water and cloud ice. This decrease is related to the decay of convection.

The three experiments with all-sky data assimilation nearly always exhibit a reduced error with respect to the free ensemble. The only exceptions are a slightly increased cloud water error in the BT experiment in the first hour and in the  $\text{VIS}_{oe=0.2}$  experiment in the second hour.

375 Generally, the BT experiment shows a more pronounced error reduction than  $VIS_{oe=0.2}$  in this  
376 situation dominated by randomly located and locally triggered deep convection. The only excep-  
377 tion is the error of cloud water in the first hour, where  $VIS_{oe=0.2}$  shows a lower error than the BT  
378 experiment. At this early state of convection, most clouds are not high enough to influence the  
379  $6.2\ \mu\text{m}$  brightness temperatures but are clearly detectable in the visible channel. The combination  
380 of both channels (BT+ $VIS_{oe=0.2}$ ) leads in most cases to an even stronger error reduction than that  
381 of the BT experiment. Overall, the BT+ $VIS_{oe=0.2}$  experiment clearly exhibits the lowest errors for  
382 all variables.

383 Vertical profiles of the mean first-guess error averaged over the 5-h assimilation period are shown  
384 (Fig. 6). The strongest reduction of wind and temperature errors occur in the upper troposphere  
385 between  $z = 6\ \text{km}$  and  $12\ \text{km}$ . Again, the BT experiment shows a clearly more pronounced error  
386 reduction than  $VIS_{oe=0.2}$  and BT+ $VIS_{oe=0.2}$  shows slightly lower errors than the BT experiment.

387 The error of cloud water peaks around  $4\ \text{km}$ , corresponding to the melting level of ice, and all  
388 three assimilation experiments show a fairly similar reduction of these errors by about 20 %. For  
389 cloud ice at upper levels, however, infrared observations are more effective in reducing the error  
390 than visible observations. Furthermore,  $VIS_{oe=0.2}$  shows a lower reduction of humidity errors in  
391 the lowest two km. As neither observation type observes humidity at this height directly, this must  
392 be related to vertical correlations and changes to surface insolation by clouds.

393 The weaker error reduction in the  $VIS_{oe=0.2}$  experiment, compared to the BT experiment, evi-  
394 dent in Figs.5 and 6 may be related to the lack of clear-sky temperature and humidity information  
395 in the visible range. Another possible explanation for this would be the ambiguity of the visible  
396 observations. BT observations are highly sensitive in clear air to the vertical profile of temperature.  
397 Visible reflectances contain no height information, so water and ice clouds can lead to the same sig-  
398 nal. In a situation where both water and ice clouds are present it is thus possible that in the LETKF

399 analysis weight is given to the ensemble members that have a cloud with the wrong phase at the  
400 right horizontal location. This ambiguity problem can be avoided when visible reflectances are  
401 assimilated together with the brightness temperatures as in the BT+VIS<sub>oe=0.2</sub> experiment. In this  
402 case the visible observations provide additional information about low clouds that is not present in  
403 the brightness temperature, leading to a further error reduction in BT+VIS<sub>oe=0.2</sub>, compared to BT.

404 During the clear-sky assimilation experiment BT<sub>CA=0</sub> the impact on hydrometeors begins to be  
405 positive for cloud ice after a few hours (Fig.5). The overall impact during the data-assimilation is  
406 neutral as can be seen in the profiles in Fig. 6, except for temperature, where the impact is positive  
407 over the height of the clear-sky weighting function. The clear-sky radiances appear to correct the  
408 phase shift of the onset of convection, but miss direct corrections of hydrometeors.

409 Mean errors ("biases") in all prognostic fields are already present in the free ensemble before  
410 data-assimilation: The errors arise from the unbiased initial perturbations due to non-linearity of  
411 the prognostic equations. To investigate if the assimilation leads to undesirable systematic effects,  
412 the evolution of the mean errors for wind, temperature and hydrometeors in the first guess during  
413 the 5 h of data assimilation and in the corresponding free forecast are compared in Figure 7. In all  
414 experiments, the mean error decreases or stays within the range of the error from the beginning of  
415 the data assimilation period - or within the range of the mean error of the free ensemble. For wind,  
416 temperature, cloud-ice, and water-vapor, the mean error decreases when BT or BT+VIS<sub>oe=0.2</sub>  
417 are assimilated. The rapid decrease in the mean error of cloud-water in the free ensemble is not  
418 reproduced sustainably in the conducted data-assimilation experiments. Assimilating clear-sky  
419 brightness temperature in BT<sub>CA=0</sub>, the mean error is overall reduced.

420 A slight degradation of the order of magnitude of the mean error occurs in the cloud-ice, when  
421 only VIS is assimilated. However, it needs to be kept in mind that the assimilation experiments  
422 are very short. Thus, it is promising that Figure 7 overall indicates no significant increase of mean

423 errors, but longer assimilation experiments over various scenarios would be required to investigate  
424 systematic effects in more detail.

#### 425 *b. Forecast impact*

426 The forecast error for cloud variables, temperature T and meridional wind V is shown in Fig. 8.  
427 Overall, the forecast error reduction is fairly consistent with the error reduction during the data  
428 assimilation period. The all-sky assimilation experiments show lower forecast errors than the free  
429 forecast for all variables. This error reduction lasts throughout the whole forecast range of 7 h  
430 with the exception of temperature errors in the  $VIS_{oe=0.2}$  experiment that become similar to the  
431 free forecast after 5.5 h. The BT experiments shows roughly twice the error reduction of  $VIS_{oe=0.2}$   
432 and BT+ $VIS_{oe=0.2}$  shows even slightly lower errors than BT. The advantage of the combined  
433 assimilation of both channels is particularly apparent for humidity, temperature and wind. For  
434 hydrometeor errors, in contrast, the differences between BT and BT+ $VIS_{oe=0.2}$  are fairly small.  
435 The clear-sky assimilation experiment  $BT_{CA=0}$  has a positive or neutral impact for all variables,  
436 except for temperature and horizontal wind after 1-2 h. There, a negative impact occurs due to the  
437 forecast at 21 UTC. In contrast, the forecast impact on temperature and horizontal wind remains  
438 positive at 23 UTC and 1 UTC (Sect. 4.b), when the temperature bias is smaller at the starting time  
439 of the forecast (Fig. 7).

440 As further metric, we employ the fractional skill score for precipitation forecasts following the  
441 evaluation of Bachmann et al. (2019) for idealized radar data assimilation OSSEs. The fractional  
442 skill score allows to derive a believable scale (sometimes referred to as skilful scale) for a precipi-  
443 tation forecasts. The results shown in (Fig. 9) are derived for a radar reflectivity threshold of 20.0  
444 dBz. The believable scale indicates a non-random overlap of precipitation fields (Mittermaier and  
445 Roberts 2010) in the forecast and nature. In all our satellite data assimilation experiments, a clear

446 reduction of the believable scale indicates improved precipitation forecasts. The believable scale  
447 increases from 100 km to 200 km in the free runs over  $\leq 7$  h until the rain decays (Fig. 9). The  
448 increase is due to a more and more scattered and random precipitation field. Similar to the evalua-  
449 tion for other forecast variables, we differentiate a clear order between the experiments from best  
450 to worst precipitation forecast as follows: Assimilating the visible channel increases the forecast  
451 skill, i.e., reduces the believable scale compared to the free background forecast at all times. The  
452 assimilation of the infrared channel leads to even better results and again, assimilating both chan-  
453 nels is best and reduces the believable scale during the first forecasting hour to 1/4, while resulting  
454 forecasts of the clear sky assimilation have a neutral or slightly negative impact. Assimilating the  
455 combination leads to the smallest believable scale in the forecast at the order of 10 km. This scale  
456 is close to the super-observation scale and effective model resolution.

457 These results are not directly comparable to the experiments for radar data assimilation by Bach-  
458 mann et al. (2020) given small differences in the setup. Nevertheless, the results overall indicated  
459 that the potential impact of satellite observation is of a similar magnitude as the impact of radar  
460 observations.

#### 4. Sensitivity experiments

In this section, we discuss the sensitivity of the data assimilation experiments to modified settings of the assigned observation error and cycling frequency.

##### *a. Sensitivity to assigned observation error*

Table 2 and Fig. 10 show the effect of increasing the assigned observation error by 50 % on the forecast error of different variables averaged over lead times of 1-8 h. The improvement of the mean absolute error as depicted before is calculated relative to the free background ensemble for cloud water  $\Delta QC / \Delta QC_{free}$ , water vapor  $\Delta QV / \Delta QV_{free}$ , cloud ice  $\Delta QI / \Delta QI_{free}$ , horizontal wind  $\Delta V / \Delta V_{free}$ , and temperature  $\Delta T / \Delta T_{free}$ . The increased observation error leads to a lower beneficial impact for all three experiments, the experiment with observations in the visible spectrum, the experiment with infrared observations and the experiment that uses both observation types. For experiments with infrared observations and the one with both observation types, however, the difference of the experiments with increased visible observation errors to the reference experiments is fairly small. Only the experiment with visible observations shows a strong difference (overall improvement of 13 % with increased error instead of 18 % improvement in the reference experiment).

Experiments with decreased assigned observation errors either led to numerical instabilities, forecast deterioration or a very small beneficial impact (not shown). This indicates that the assigned observation error of the reference experiments is a suitable choice for the assimilation. In this context, it should also be noted that the assigned errors are strongly inflated compared to the errors used for simulating the observations. Visible observations were simulated with a random error of only 3 %. Due to superobbing of 36 pixels, the actual error is reduced further by a factor of 36 for the assimilated super-observations. This discrepancy of actual and assigned errors by more



484 than a factor of ten is interesting given the absence of correlated observation errors and of repre-  
485 sentation and operator errors when using a model simulation as truth in an OSSE. We therefore  
486 speculate that the strong inflation of errors is necessary to compensate for displacement errors and  
487 other non-linear effects as well as for deficiencies of the data assimilation scheme. For infrared  
488 observations, the comparison of actual and assigned errors is a bit more complicated due to the use  
489 of the dynamic error model. Nevertheless, the assigned observation error of infrared observations  
490 is also strongly inflated compared to their actual observation error.

491 *b. Sensitivity to cycling frequency & all-sky versus clear-sky brightness temperature assimilation*

492 The cycling period was varied between 15, 30 and 60 min for the experiment assimilating the  
493 combination of infrared and visible observations with an assigned observation error of  $oe = 0.2$ .

494 All experiments with lower cycling frequency are typically evaluated hourly (referred to as  
495 "sampled hourly" in the experiment name). To study the effect of the evaluation frequency on  
496 the results, assimilation experiments with higher frequency cycling are also evaluated hourly and  
497 half-hourly. However, the error of evaluating less frequently appears to be insignificant (Fig. 11).

498 The comparison of the experiments with a cycling period of 1/4 h, 1/2 h, and 1 h (Table 3) re-  
499 veals a larger forecast improvement for higher cycling frequencies. It is therefore beneficial to  
500 assimilate the observations with higher temporal resolution. However, the differences between the  
501 experiments are rather small despite the fact that the 1 h cycling period also decreases the amount  
502 of assimilated observations by a factor of 4 compared to the experiment with a 1/4 h cycling pe-  
503 riod. Using a 1 h cycling period may therefore be a reasonable choice if the number of assimilated  
504 observation should not be too large or if other reasons restrict the cycling period.

505 Assimilating only clear-sky brightness temperature observations with or without localization  
506 leads to a clear decrease in forecast skill for all variables compared to assimilating all-sky bright-

507 ness temperature (Fig. 12). In comparison to assimilating without localization, adding localization  
508 in the clear-sky experiment  $BT_{CA=0}^{loc}$  can lead to a slight improvement in forecasting hydrometeors.

## 509 **5. Conclusions**

510 This paper investigates the potential impact of cloud-affected satellite observations in the visi-  
511 ble and infrared spectrum in idealized convective-scale observing system simulation experiments  
512 (OSSEs) with a local ensemble transform Kalman filter (LETKF) for data assimilation. We in-  
513 vestigate a particularly challenging case with locally triggered and randomly located summer-time  
514 deep convection in central Europe.

515 Observations from the visible and infrared channel provide very complementary information on  
516 atmospheric clouds with a higher sensitivity of the infrared channels to ice clouds and of the visible  
517 to water clouds. Furthermore, infrared channels provide information on cloud top heights whereas  
518 visible channels allow to distinguish low clouds from the surface. Despite these advantages, a  
519 combination of infrared and visible channels has not been used for data assimilation, yet.

520 The OSSEs demonstrate a strongly beneficial impact of satellite data assimilation on various  
521 forecast quantities for the whole forecast range of 8 h lead time. The mean relative forecast  
522 improvement ranges up to nearly 30 % for model state variables. Precipitation forecast show even  
523 more drastic improvements. The Fraction Skill Score (FSS) believable (or skilful) scale increases  
524 by up to a factor of four and means that 7-h forecasts with satellite data assimilation are better than  
525 1-h forecasts without.

526 While the results are not directly comparable to the OSSE results of Bachmann et al. (2019)  
527 and Bachmann et al. (2020) for radar data assimilation due to some differences of the setup, they  
528 indicate a comparable magnitude of the potential impact of cloud-affected satellite observations to  
529 radar observations. Both visible and infrared observations individually lead to a forecast improve-

530 ment, which is higher for infrared observations in this convective situation. Best forecast results,  
531 however, are achieved through the combined assimilation of both visible and infrared observa-  
532 tions. We assume that this is related to the reduction of ambiguities in the observations through  
533 the combination of both types.

534 It should be noted that the relative effectiveness of assimilating visible or water vapor channels  
535 can be expected to depend strongly on the weather situation. For instance, when only boundary  
536 layer clouds are present, the visible channel does not suffer from a potential confusion between  
537 water and ice clouds and the water vapor channel does not contain cloud information. Therefore,  
538 we would expect a much stronger impact from the visible channel in such a case. However, the  
539 current impact on forecasts after 22 UTC does not take into account the diurnal cycle of the sun  
540 on visible observations.

541 Sensitivity experiments with different assigned observation errors indicate that a constant error  
542 of 0.2 for visible reflectance and of 1.1 K plus an error inflation dependent on cloud-impact based  
543 on Harnisch et al. (2016) for infrared observations is an appropriate choice. This is an interesting  
544 result given that the observations were simulated using an error of only 3 % for visible reflectance  
545 and 3 K for infrared brightness temperature observations. As the assimilated observations are  
546 super-observations consisting of  $6 \times 6$  pixels, their actual error is only 1/6 of the one used for  
547 assimilating the observations. Consequently, this means that the appropriate assigned error needs  
548 to be highly inflated for the assimilation despite of the absence of correlated observation errors,  
549 representation errors, and operator errors. We assume that this strong error inflation is necessary  
550 to compensate for displacement errors and other non-linear effects as well as for deficiencies of  
551 the data assimilation scheme.

552 Furthermore, we conducted sensitivity experiments using cycling periods of 15, 30, and 60 min.  
553 These show that it is most beneficial to assimilate the observations every 15 min. However, a

554 beneficial impact is also achieved using 30-min or 60-min cycling periods and given that those  
555 experiments only assimilate half or a quarter of the observations, the forecast improvement is also  
556 remarkable. Consequently, it may as well be a suitable choice to use a cycling period of 1 h for  
557 these conditions in case of need for a reduced data amount or other operational constraints.

558 In summary, we show that an LETKF assimilation scheme is capable of using the informa-  
559 tion provided by cloud-affected satellite observations. Their assimilation strongly improves the  
560 forecast of various quantities including precipitation. While the total impact of such observations  
561 achieved in this idealized OSSE can likely not be achieved in a real NWP system, the study pro-  
562 vides important insights on the relative impact of observations. Best forecast results are achieved  
563 when assimilating both visible and infrared observations and overall, the impact is of comparable  
564 magnitude as the impact of radar observations. This strongly emphasizes the potential benefit of  
565 such observations for convective-scale NWP - especially in regions on the globe where a dense  
566 network of conventional observations or other remote sensing measurements are unavailable.

567 *Acknowledgments.* This study was carried out in the Hans-Ertel Centre for Weather Research  
568 (Weissmann et al. 2014; Simmer et al. 2016), which is a German research network of universities,  
569 research institutes and Deutscher Wetterdienst funded by the BMVI (Federal Ministry of Transport  
570 and Digital Infrastructure). We thank three anonymous referees for constructive suggestions on  
571 the manuscript. Furthermore, we acknowledge H. Lange and K. Bachmann for providing an earlier  
572 version of the experimental setup and R. Faulwetter for support with RTTOV.

## 573 **References**

574 Bachmann, K., C. Keil, G. C. Craig, M. Weissmann, and C. A. Welzbacher, 2020: Predictability  
575 of deep convection in idealized and operational forecasts: Effects of radar data assimilation,  
576 orography and synoptic weather regime. *Mon. Wea. Rev.*, **148** (1), 63–81.

- 577 Bachmann, K., C. Keil, and M. Weissmann, 2019: Impact of radar data assimilation and orography  
578 on predictability of deep convection. *Quart. J. Roy. Meteor. Soc.*, **145 (718)**, 117–130.
- 579 Bauer, P., A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky  
580 radiances. Part I: Implementation. *Q. J. R. Meteorological Society*, **136 (652)**, 1868–1885.
- 581 Cintineo, R. M., J. A. Otkin, T. A. Jones, S. Koch, and D. J. Stensrud, 2016: Assimilation of  
582 synthetic GOES-R ABI infrared brightness temperatures and WSR-88D radar observations in a  
583 high-resolution OSSE. *Mon. Wea. Rev.*, **144 (9)**, 3159–3180.
- 584 Craig, G. C., and M. Würsch, 2013: The impact of localization and observation averaging for  
585 convective-scale data assimilation in a simple stochastic model. *Quart. J. Roy. Meteor. Soc.*,  
586 **139 (671)**, 515–523.
- 587 Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: Explicit forecasts of  
588 convection using the Weather Research and Forecasting (WRF) model. *Atmos. Sci. Lett.*, **5 (6)**,  
589 110–117.
- 590 Geer, A., and Coauthors, 2017: The growing impact of satellite observations sensitive to humidity,  
591 cloud and precipitation. *Quart. J. Roy. Meteor. Soc.*, **143 (709)**, 3189–3206.
- 592 Geer, A. J., 2019: Correlated observation error models for assimilating all-sky infrared radiances.  
593 *Atmos. Meas. Tech.*, **12 (7)**, 3629–3657.
- 594 Geer, A. J., P. Bauer, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II:  
595 Assessment. *Quart. J. Roy. Meteor. Soc.*, **136 (652)**, 1886–1905.
- 596 Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather fore-  
597 casting centres. *Quart. J. Roy. Meteor. Soc.*, **144 (713)**, 1191–1217.

- 598 Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale  
599 numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*
- 600 Harnisch, F., M. Weissmann, and Á. Periañez, 2016: Error model for the assimilation of cloud-  
601 affected infrared satellite observations in an ensemble data assimilation system. *Quart. J. Roy.*  
602 *Meteor. Soc.*, **142 (697)**, 1797–1808.
- 603 Heinze, R., and Coauthors, 2017: Large-eddy simulations over Germany using ICON: a compre-  
604 hensive evaluation. *Quart. J. Roy. Meteor. Soc.*, **143 (702)**, 69–100.
- 605 Honda, T., and Coauthors, 2018: Assimilating all-sky Himawari-8 satellite infrared radiances: A  
606 case of typhoon Soudelor (2015). *Mon. Wea. Rev.*, **146 (1)**, 213–229.
- 607 Houtekamer, P., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data  
608 assimilation. *Mon. Wea. Rev.*, **144 (12)**, 4489–4532.
- 609 Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal  
610 chaos: A local ensemble transform Kalman filter. *Physica D*, **230 (1-2)**, 112–126.
- 611 Hutt, A., and Coauthors, 2020: Assimilation of SEVIRI water vapour channels with an ensemble  
612 Kalman filter on the convective scale. *Front. Earth Sci.*, **8**, 70.
- 613 Janjić, T., and Coauthors, 2017: On the representation error in data assimilation. *Quart. J. Roy.*  
614 *Meteor. Soc.*, **144 (713)**, 1257–1278.
- 615 Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and  
616 prediction of convective storms. *Mon. Wea. Rev.*, **142 (10)**, 3781–3808.
- 617 Lange, H., G. C. Craig, and T. Janjić, 2017: Characterizing noise and spurious convection in  
618 convective data assimilation. *Quart. J. Roy. Meteor. Soc.*, **143 (709)**, 3060–3069.

- 619 Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a  
620 cloud model. *J. Climate and Applied Meteorology*, **22** (6), 1065–1092.
- 621 Matricardi, M., and R. Saunders, 1999: Fast radiative transfer model for simulation of infrared  
622 atmospheric sounding interferometer radiances. *Appl. Opt.*, **38** (27), 5679–5691.
- 623 McNally, A., 2002: A note on the occurrence of cloud in meteorologically sensitive areas and the  
624 implications for advanced infrared sounders. *Quart. J. Roy. Meteor. Soc.*, **128** (585), 2551–2556.
- 625 Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky  
626 satellite radiance. *Mon. Wea. Rev.*, **145** (3), 1063–1081.
- 627 Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods:  
628 Identifying skillful spatial scales using the fractions skill score. *Weather Forecast.*, **25** (1), 343–  
629 354.
- 630 Necker, T., M. Weissmann, and S. Matthias, 2018: The importance of appropriate verification  
631 metrics for the assessment of observation impact in a convection-permitting modelling system.  
632 *Quart. J. Roy. Meteor. Soc.*, **140** (714), 1667–1680.
- 633 Nolan, D. S., R. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a  
634 hurricane nature run using the joint OSSE nature run and the wrf model. *J. Adv. Model. Earth*  
635 *Sys.*, **5** (2), 382–405.
- 636 Okamoto, K., A. McNally, and W. Bell, 2014: Progress towards the assimilation of all-sky infrared  
637 radiances: an evaluation of cloud effects. *Quart. J. Roy. Meteor. Soc.*, **140** (682), 1603–1614.
- 638 Otkin, J. A., W. E. Lewis, A. J. Lenzen, B. D. McNoldy, and S. J. Majumdar, 2017: Assessing  
639 the accuracy of the cloud and water vapor fields in the hurricane WRF (HWRF) model using  
640 satellite infrared brightness temperatures. *Mon. Wea. Rev.*, **145** (5), 2027–2046.

- 641 Otkin, J. A., R. Potthast, and A. S. Lawless, 2018: Nonlinear bias correction for satellite data  
642 assimilation using Taylor series polynomials. *Mon. Wea. Rev.*, **146** (1), 263–285.
- 643 Saunders, R., M. Matricardi, and P. Brunel, 1999: An improved fast radiative transfer model for  
644 assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, **125** (556), 1407–  
645 1425.
- 646 Sawada, Y., K. Okamoto, M. Kunii, and T. Miyoshi, 2019: Assimilating every-10-minute  
647 himawari-8 infrared radiances to improve convective predictability. *J. Geophys. Res.*, **124** (5),  
648 2546–2561.
- 649 Scheck, L., P. Frèrebeau, R. Buras-Schnell, and B. Mayer, 2016: A fast radiative transfer method  
650 for the simulation of visible satellite imagery. *J. Quantit. Spectrosc. Radiat. Transf.*, **175**, 54–67.
- 651 Scheck, L., M. Weissmann, and L. Bach, 2020: Assimilating visible satellite images for convective  
652 scale weather prediction: A case study. *Quart. J. Roy. Meteor. Soc.*, accepted, doi:[https://doi.org/  
653 10.1002/qj.3840](https://doi.org/10.1002/qj.3840).
- 654 Scheck, L., M. Weissmann, and B. Mayer, 2018: Efficient methods to account for cloud-top in-  
655 clination and cloud overlap in synthetic visible satellite images. *J. Atmos. Oceanic Technol.*,  
656 **35** (3), 665–685.
- 657 Schomburg, A., C. Schraff, and R. Potthast, 2015: A concept for the assimilation of satellite  
658 cloud information in an ensemble Kalman filter: single-observation experiments. *Quart. J. Roy.  
659 Meteor. Soc.*, **141** (688), 893–908.
- 660 Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Perriñez, and R. Potthast, 2016:  
661 Kilometre-scale ensemble data assimilation for the COSMO model (kenda). *Quart. J. Roy. Me-  
662 teor. Soc.*, **142** (696), 1453–1472.



- 663 Seifert, A., and K. D. Beheng, 2001: A double-moment parameterization for simulating autocon-  
664 version, accretion and selfcollection. *Atmos. Res.*, **59**, 265–281.
- 665 Simmer, C., and Coauthors, 2016: HErZ: The German Hans-Ertel Centre for Weather Research.  
666 *Bull. Amer. Meteor. Soc.*, **97** (6), 1057–1068.
- 667 Sommer, M., and M. Weissmann, 2014: Observation impact in a convective-scale localized en-  
668 semble transform Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **140** (685), 2672–2679.
- 669 Sommer, M., and M. Weissmann, 2016: Ensemble-based approximation of observation impact  
670 using an observation-based verification metric. *Tellus A*, **68** (1), 27 885.
- 671 Waller, J. A., S. P. Ballard, S. L. Dance, G. Kelly, N. K. Nichols, and D. Simonin, 2016: Diagnos-  
672 ing horizontal and inter-channel observation error correlations for SEVIRI observations using  
673 observation-minus-background and observation-minus-analysis statistics. *Remote Sens.*, **8** (7),  
674 581.
- 675 Weissmann, M., and Coauthors, 2014: Initial phase of the Hans-Ertel Centre for Weather  
676 Research-A virtual centre at the interface of basic and applied weather and climate research.  
677 *Meteorologische Zeitschrift*, **23**, 193–208.
- 678 Zeng, Y., T. Janjić, M. Sommer, A. de Lozar, U. Blahak, and A. Seifert, 2019: Representation  
679 of model error in convective-scale data assimilation: Additive noise based on model truncation  
680 error. *J. Adv. Model. Earth Sys.*, **11** (3), 752–770.
- 681 Zhang, F., M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky in-  
682 frared satellite radiances from goes-r on convection-permitting analysis and prediction of tropi-  
683 cal cyclones. *Geophys. Res. Lett.*, **43** (6), 2954–2963.

- 684 Zhang, Y., D. J. Stensrud, and F. Zhang, 2019: Simultaneous assimilation of radar and all-sky  
685 satellite infrared radiance observations for convection-allowing ensemble analysis and predic-  
686 tion of severe thunderstorms. *MWR*, **147** (12), 4389–4409.
- 687 Zhang, Y., F. Zhang, and D. J. Stensrud, 2018: Assimilating all-sky infrared radiances from goes-  
688 16 abi using an ensemble kalman filter for convection-allowing severe thunderstorms prediction.  
689 *Mon. Wea. Rev.*, **146** (10), 3363–3381.

690 **LIST OF TABLES**

691 **Table 1.** Overview of all data-assimilation experiments – assimilating brightness tem-  
 692 perature (BT), visible observations (VIS), and a combination of both with ob-  
 693 servation error  $oe(CA)$  depending on cloud impact in the water vapor band  
 694 (Harnisch et al. 2016) and a given constant  $oe_{vis}$  for the visible spectral range.  
 695 All experiments are assimilated for  $\geq 1$  h with a cycling period of 15 min,  
 696 30 min or 1 h. Forecasts of 8 h each can be started after 4 cycles for all 40  
 697 members from the analysis. The data-assimilation cycle in all experiments be-  
 698 gins at 20 UTC. Experiment  $BT_{CA=0}$  assimilates only clear-sky values, while  
 699  $BT_{CA=0}^{loc}$  in addition localizes the innovation around the clear-sky water vapor  
 700 weighting function for  $6.2 \mu\text{m}$ . . . . . 34

701 **Table 2.** Overview of relative improvement in percent with respect to the free back-  
 702 ground forecasts for cloud water QC, water vapor QV, cloud ice QI, meridional  
 703 wind V, temperature T, and believable scale  $Z_{BS}$  of column maximum radar re-  
 704 flectivity. The relative improvements are averaged for each experiment over the  
 705 whole forecast range of 8 h taking into account three different forecasts starting  
 706 at 21 UTC, 23 UTC, and 1 UTC. . . . . 35

707 **Table 3.** Overview of relative improvement as in Table 2, but only evaluating forecasts  
 708 starting at 1 UTC. . . . . 36

	Instrument	$\Delta t_{cycle}$ / min	$oe$	$oe_{vis}$	start times (UTC)
BT	wv 6.2 $\mu\text{m}$	15	$oe(CA)$		21, 23, 1
VIS $_{oe=0.2}$	vis 0.6 $\mu\text{m}$	15		0.2	21, 23, 1
BT+VIS $_{oe=0.2}$	wv 6.2 $\mu\text{m}$ , vis 0.6 $\mu\text{m}$	15	$oe(CA)$	0.2	21, 23, 1
BT $_{oe*1.5}$	wv 6.2 $\mu\text{m}$	15	1.5 $oe(CA)$		21, 23, 1
VIS $_{oe=0.3}$	vis 0.6 $\mu\text{m}$	15		0.3	21, 23, 1
BT+VIS $_{oe=0.3}$	wv 6.2 $\mu\text{m}$ , vis 0.6 $\mu\text{m}$	15	$oe(CA)$	0.3	21, 23, 1
BT+VIS $_{oe=0.2}^{1/2h}$	wv 6.2 $\mu\text{m}$ , vis 0.6 $\mu\text{m}$	30	$oe(CA)$	0.2	1
BT+VIS $_{oe=0.2}^{1h}$	wv 6.2 $\mu\text{m}$ , vis 0.6 $\mu\text{m}$	60	$oe(CA)$	0.2	1
BT $_{CA=0}$	wv 6.2 $\mu\text{m}$	15	1.1 K		21, 23, 1
BT $_{CA=0}^{loc}$	wv 6.2 $\mu\text{m}$	15	1.1 K		1

709 TABLE 1. Overview of all data-assimilation experiments – assimilating brightness temperature (BT), visible  
710 observations (VIS), and a combination of both with observation error  $oe(CA)$  depending on cloud impact in the  
711 water vapor band (Harnisch et al. 2016) and a given constant  $oe_{vis}$  for the visible spectral range. All experiments  
712 are assimilated for  $\geq 1$  h with a cycling period of 15 min, 30 min or 1 h. Forecasts of 8 h each can be started after  
713 4 cycles for all 40 members from the analysis. The data-assimilation cycle in all experiments begins at 20 UTC.  
714 Experiment BT $_{CA=0}$  assimilates only clear-sky values, while BT $_{CA=0}^{loc}$  in addition localizes the innovation around  
715 the clear-sky water vapor weighting function for 6.2  $\mu\text{m}$ .

	Relative improvement / %					
	QC	QV	QI	V	T	$Z_{BS}$
$VIS_{oe=0.3}$	23.0	9.4	13.0	5.3	3.1	40.8
$VIS_{oe=0.2}$	26.6	13.0	17.5	7.3	5.7	46.7
BT	35.4	20.1	31.0	17.8	13.7	77.0
$BT_{oe*1.5}$	34.6	19.4	29.5	16.5	9.4	60.0
$BT+VIS_{oe=0.3}$	36.0	22.5	32.3	20.2	17.1	80.1
$BT+VIS_{oe=0.2}$	36.1	23.5	32.4	21.1	18.5	80.0

716 TABLE 2. Overview of relative improvement in percent with respect to the free background forecasts for  
717 cloud water QC, water vapor QV, cloud ice QI, meridional wind V, temperature T, and believable scale  $Z_{BS}$  of  
718 column maximum radar reflectivity. The relative improvements are averaged for each experiment over the whole  
719 forecast range of 8 h taking into account three different forecasts starting at 21 UTC, 23 UTC, and 1 UTC.

	Relative improvement / %					
	QC	QV	QI	V	T	$Z_{BS}$
BT+VIS <sub>oe=0.2</sub>	45.7	23.7	34.6	24.4	16.0	88.0
BT+VIS <sup>1/2h</sup> <sub>oe=0.2</sub>	49.2	30.8	23.5	22.2	13.7	84.8
BT+VIS <sup>1h</sup> <sub>oe=0.2</sub>	45.1	28.3	22.0	17.1	9.6	79.0
BT	44.7	19.0	31.4	18.0	6.7	47.1
BT <sub>CA=0</sub>	4.2	3.3	3.6	2.3	0.8	-6.7
BT <sup>loc</sup> <sub>CA=0</sub>	5.9	3.8	4.1	2.7	1.1	-5.8

TABLE 3. Overview of relative improvement as in Table 2, but only evaluating forecasts starting at 1 UTC.

## LIST OF FIGURES

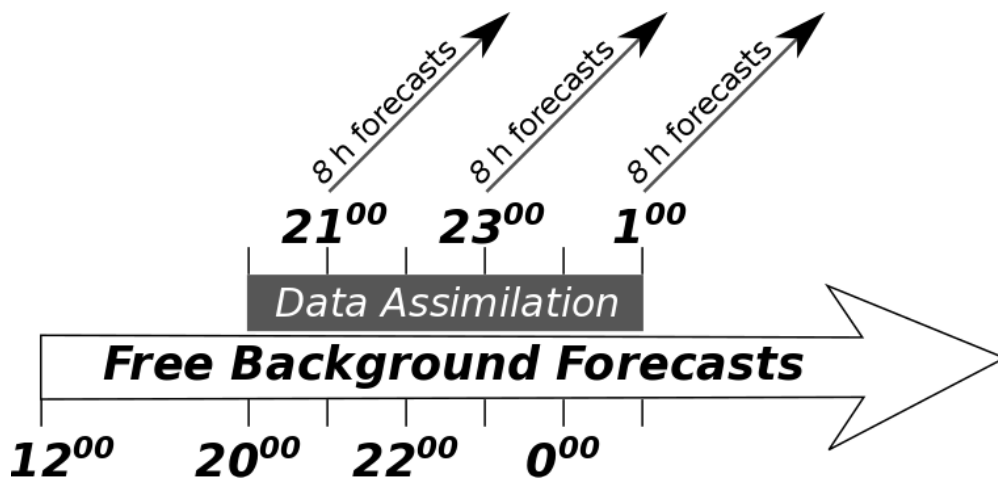
721	<b>Fig. 1.</b>	Free background forecasts start at 12 UTC. The data assimilation provides analyses from 20 UTC to 1 UTC. Forecasts of 8 hours lead time are started from the analysis at 21 UTC, 23 UTC, and 1 UTC for all ensemble members. . . . .	39
724	<b>Fig. 2.</b>	Synthetic brightness temperature (BT) in the infrared 6.2 $\mu\text{m}$ water vapor channel, reflectance in the visible 0.6 $\mu\text{m}$ channel, column maximum of synthetic radar reflectivity (Z) are plotted as time series. Corresponding time series of column maximum cloud ice (QI) and column maximum cloud water (QC) are depicted below. One fourth of the domain from the nature run is shown: the south-east corner. . . . .	40
729	<b>Fig. 3.</b>	Black lines depict horizontal means of column maximum radar reflectivity ( <b>top</b> ), visible satellite ( <b>middle</b> ), and brightness temperature field ( <b>bottom</b> ) of each ensemble member. For comparison, the values from the nature run are shown ( <i>red lines</i> ). Corresponding fields from the previous figure are shaded in gray. Brightness temperature and radar reflectivity were not stored before 16 UTC. . . . .	41
734	<b>Fig. 4.</b>	First guess departures are calculated as probability density distributions for 0.8 million visible reflectance ( <b>a</b> ), 0.8 million all-sky brightness temperature ( <b>b</b> ), and 0.5 million clear-sky brightness temperature values ( <b>c</b> ) over 5 hours of assimilation time ( <i>black lines</i> ). Additionally, brightness temperature all-sky departures for the first ( <b>d</b> ) and eighth cycle ( <b>e</b> ) are plotted. The values are normalized by corresponding observation errors. Corresponding bell curves are depicted with standard deviation $\sigma_{vis}$ and $\sigma_{BT}$ and mean $\mu_{vis}$ , $\mu_{BT} = 0$ ( <i>dashed red lines</i> ). . . . .	42
741	<b>Fig. 5.</b>	Mean absolute errors of observation minus first guess averaged over the whole domain and ensemble up to a height of 15 km are shown as time series for cloud ice QI, cloud water QC, water vapor QV, meridional wind V, and temperature T (for better readability the error of QI and QC are scaled with $10^{-2}$ ). The black line shows the error without data assimilation. Four assimilation experiments are compared: only clear-sky brightness temperature ( $BT_{CA=0}$ ), cloud-affected brightness temperature (BT), only visible reflectances ( $VIS_{oe=0.2}$ ) and a combination of both ( $BT+VIS_{oe=0.2}$ ). The observation error for the cloud-affected BTs is chosen from an error model (Harnisch et al. 2016). The observation error for visible reflectances is set constant to 0.2. The free forecast ( <i>black line</i> ) is the mean of the 40 member ensemble forecast from the experiment without data assimilation. . . . .	43
751	<b>Fig. 6.</b>	Profiles of mean absolute error are shown of the first guess during 5 h of data assimilation. QI and QC are combined in one panel. The variables and line colors correspond to the experiments in the previous figure. . . . .	44
754	<b>Fig. 7.</b>	Mean errors ("biases") of observation minus first guess are shown as time series. The variables and line colors correspond to the experiments in the previous figure. For comparison the zero mean error is indicated ( <i>thin dashed black line</i> ). . . . .	45
757	<b>Fig. 8.</b>	Mean absolute error in forecasts of cloud ice (QI), cloud water (QC), water vapor (QV), horizontal wind (V), and temperature (T) for a set of assimilation experiments. The time series are the means over all forecast times (21 UTC, 23 UTC, 1 UTC as listed in Table 1). The line colors correspond to the experiments in the previous figures. . . . .	46
761	<b>Fig. 9.</b>	Forecasts of the believable scale of column maximum radar reflectivity starting at 21 UTC ( <b>top</b> ), 23 UTC ( <b>middle</b> ), and 1 UTC ( <b>bottom</b> ). The line colors correspond to the experiments in the previous figures. . . . .	47

764 **Fig. 10.** Overall improvement  $\Sigma/\Sigma_{free}$  of mean relative error for (QI), cloud water (QC), water va-  
765 por (QV), horizontal wind (V) and temperature (T) for a set of assimilation experiments  
766 (Table 1). Presented is the improvement over the 8 h forecast starting at 1 UTC. All 5  
767 variables contribute equally to the overall improvement, i.e.,  $\Sigma/\Sigma_{free} = (\Delta QI/\Delta QI_{free} +$   
768  $\Delta QC/\Delta QC_{free} + \Delta QV/\Delta QV_{free} + \Delta V/\Delta V_{free} + \Delta T/\Delta T_{free})/5$ . Assimilation experiments  
769 with combined instruments  $BT + VIS_{oe=0.2}$  are compared for forecasts starting at **(a)**  
770 21 UTC, **(b)** 23 UTC, and **(c)** 1 UTC. . . . . 48

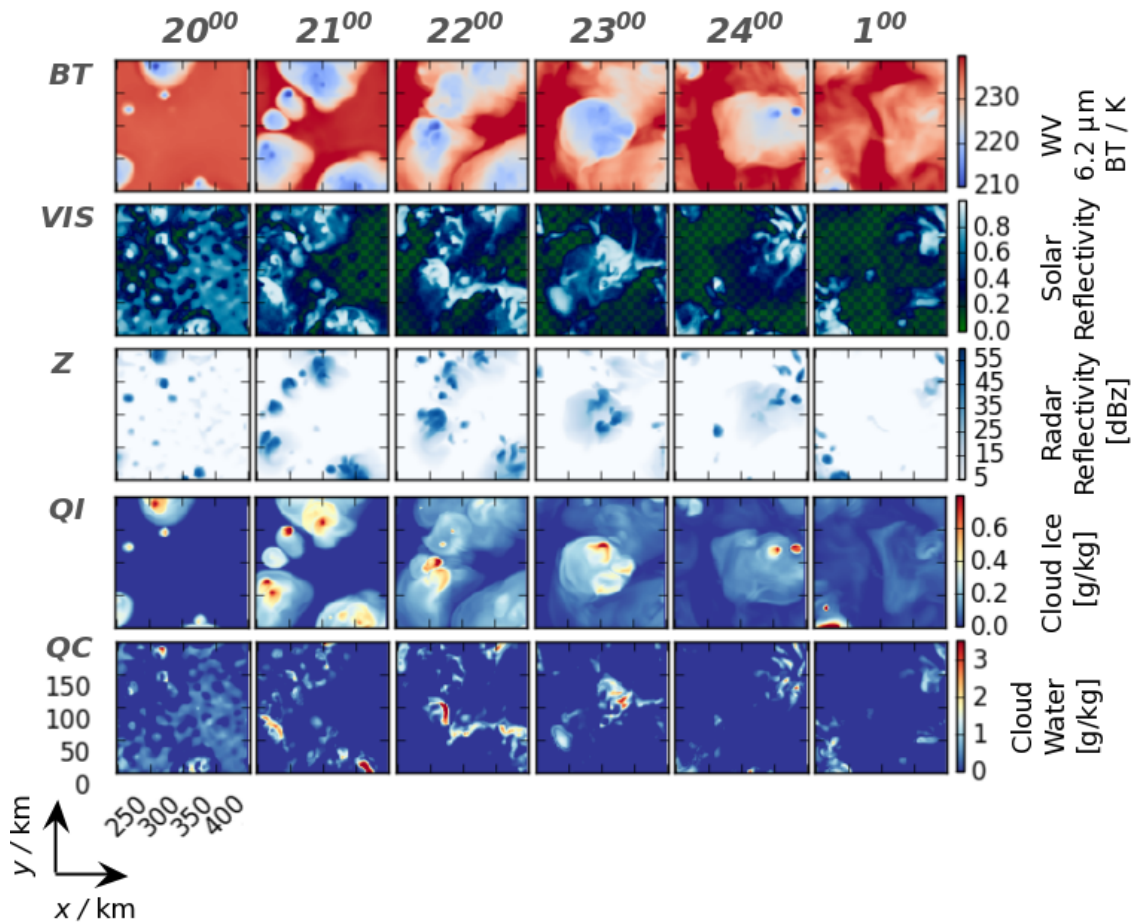
771 **Fig. 11.** As in figure 10, only varying the cycling frequency and diagnosing forecasts starting at  
772 1 UTC. . . . . 49

773 **Fig. 12.** Diagnosing forecasts starting at 1 UTC as in figure 11, after hourly assimilating all-sky and  
774 clear-sky brightness temperature in separate experiments. . . . . 50

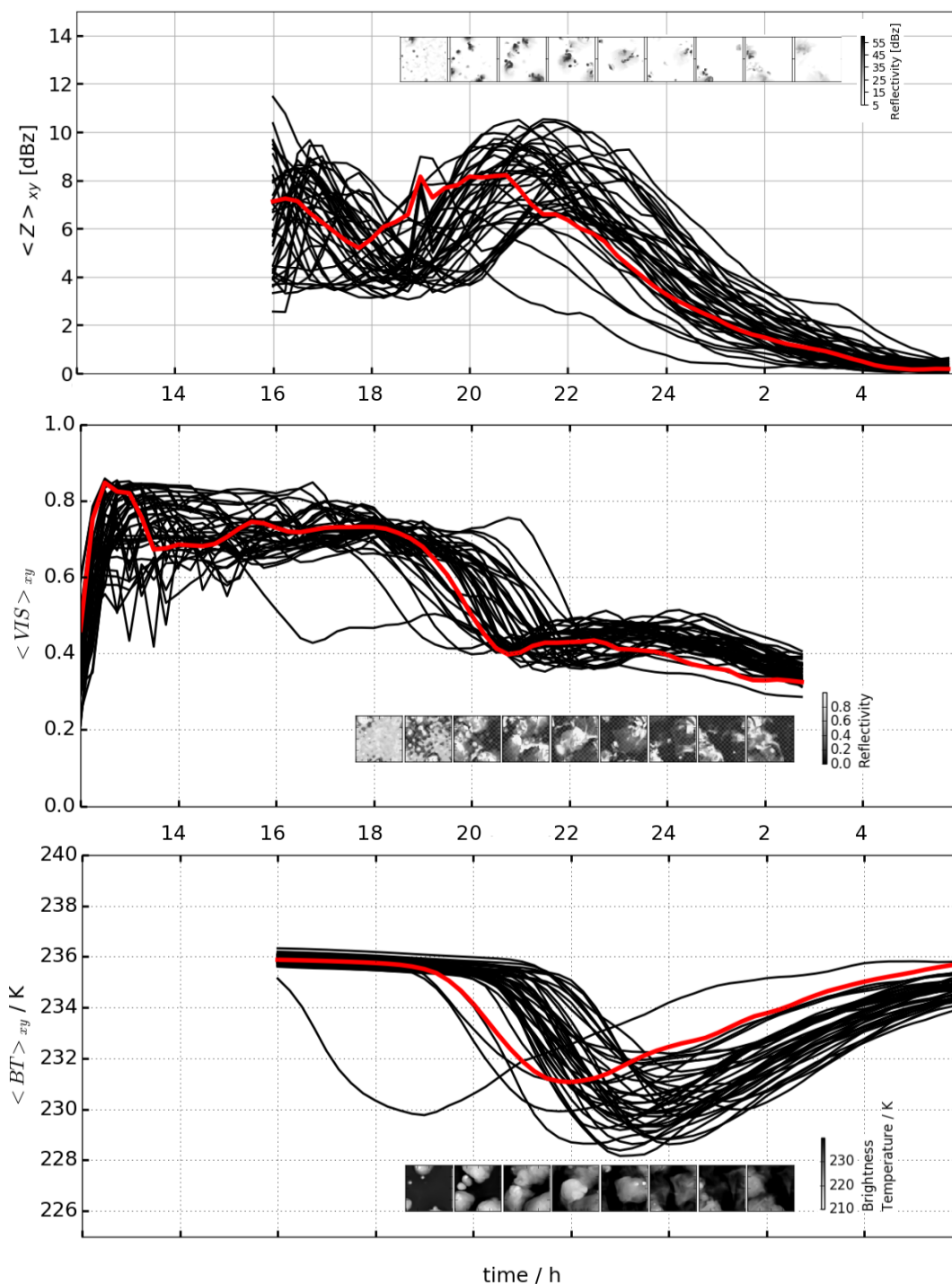




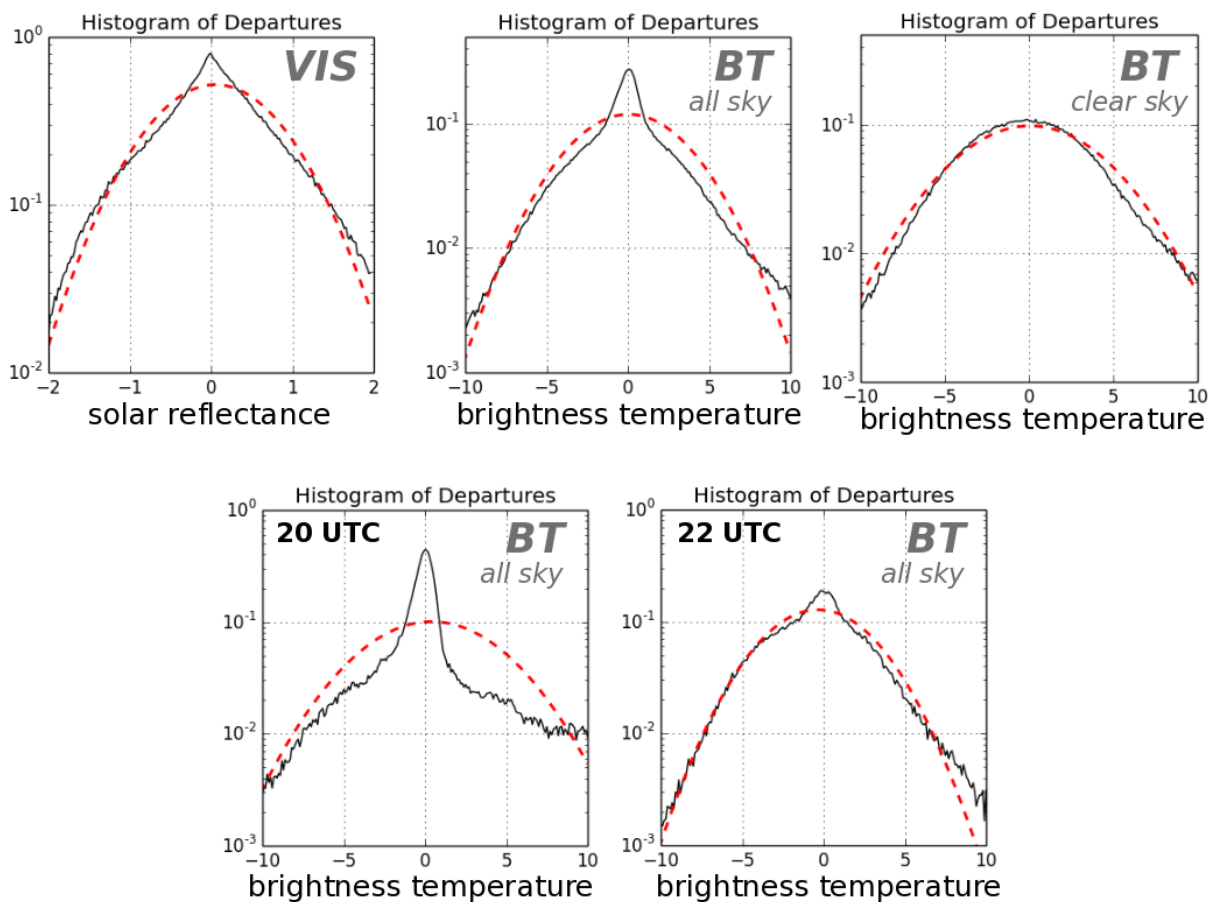
775 FIG. 1. Free background forecasts start at 12 UTC. The data assimilation provides analyses from 20 UTC  
 776 to 1 UTC. Forecasts of 8 hours lead time are started from the analysis at 21 UTC, 23 UTC, and 1 UTC for all  
 777 ensemble members.



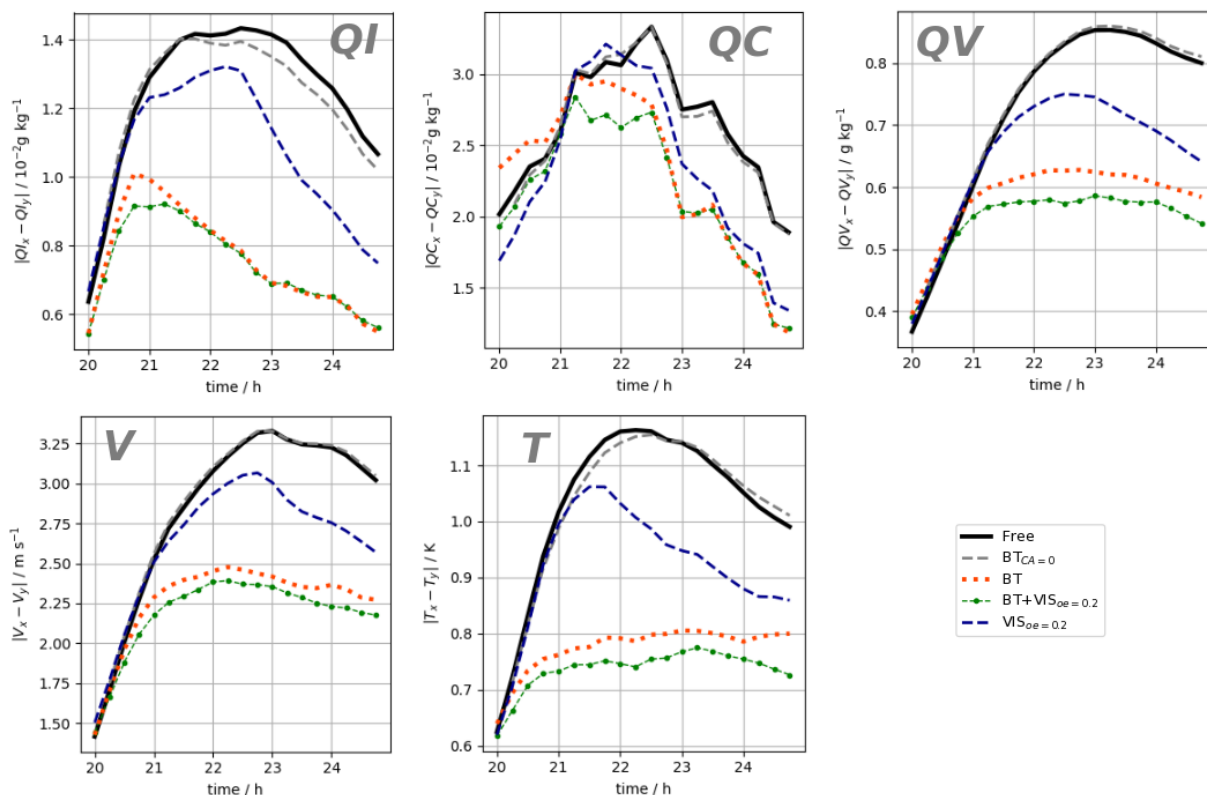
778 FIG. 2. Synthetic brightness temperature (BT) in the infrared 6.2  $\mu\text{m}$  water vapor channel, reflectance in the  
 779 visible 0.6  $\mu\text{m}$  channel, column maximum of synthetic radar reflectivity (Z) are plotted as time series. Corre-  
 780 sponding time series of column maximum cloud ice (QI) and column maximum cloud water (QC) are depicted  
 781 below. One fourth of the domain from the nature run is shown: the south-east corner.



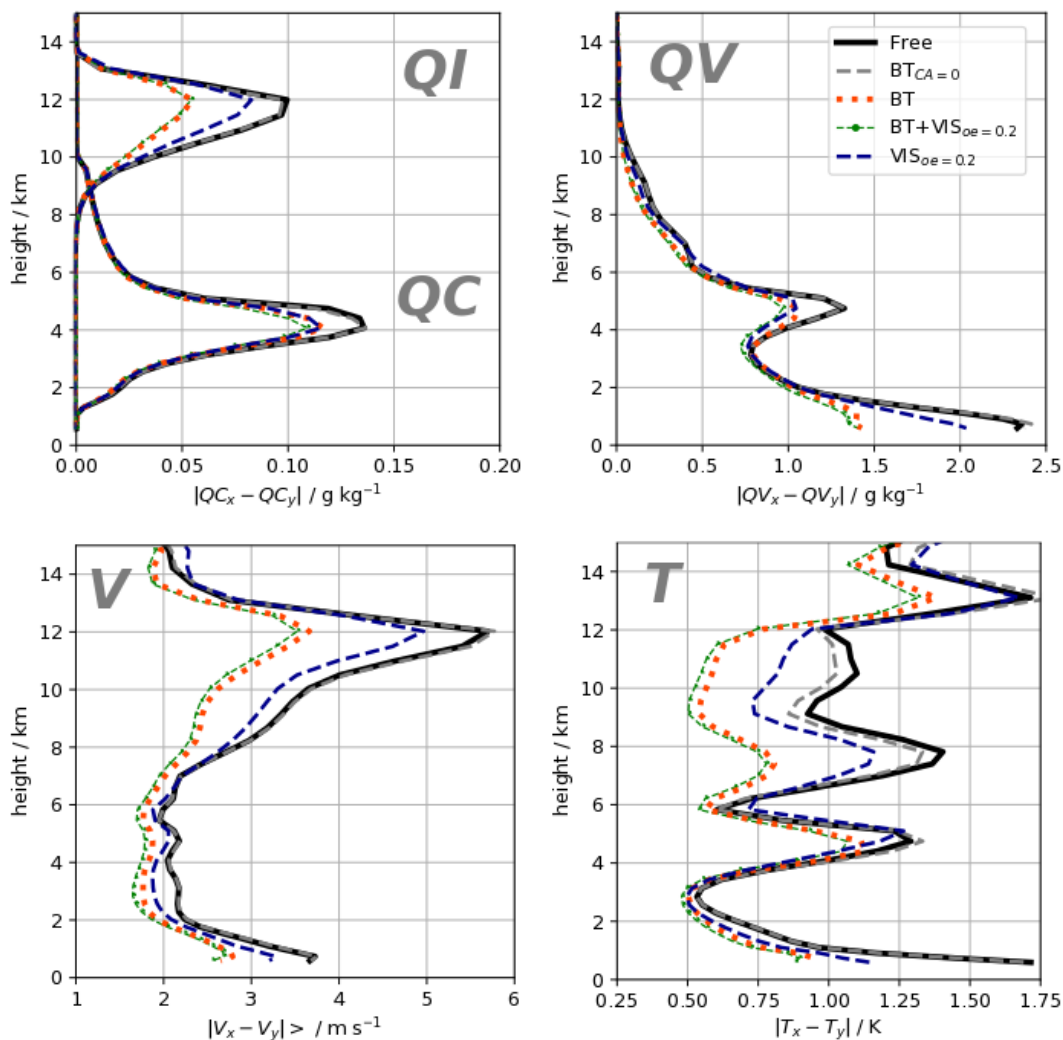
782 FIG. 3. Black lines depict horizontal means of column maximum radar reflectivity (**top**), visible satellite  
 783 (**middle**), and brightness temperature field (**bottom**) of each ensemble member. For comparison, the values  
 784 from the nature run are shown (*red lines*). Corresponding fields from the previous figure are shaded in gray.  
 785 Brightness temperature and radar reflectivity were not stored before 16 UTC.



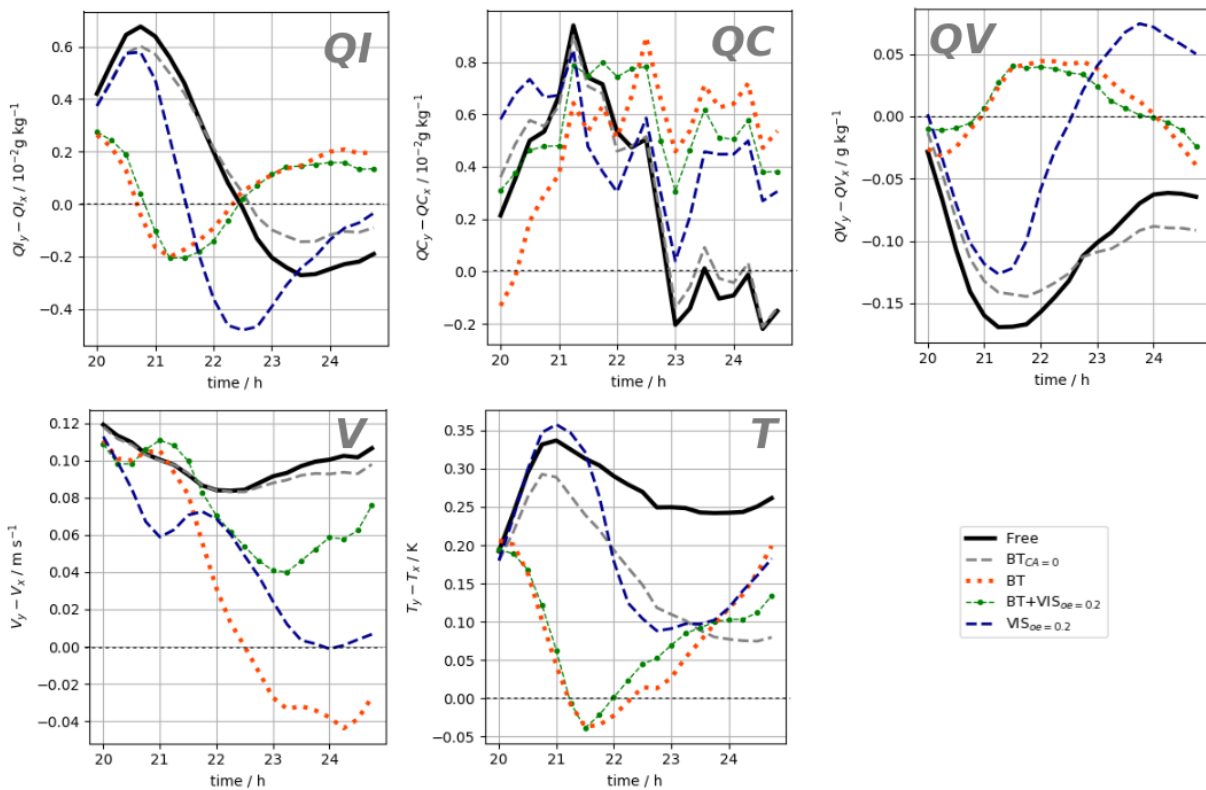
786 FIG. 4. First guess departures are calculated as probability density distributions for 0.8 million visible re-  
 787 flectance (a), 0.8 million all-sky brightness temperature (b), and 0.5 million clear-sky brightness temper-  
 788 ature values (c) over 5 hours of assimilation time (black lines). Additionally, brightness temperature all-sky depar-  
 789 tures for the first (d) and eighth cycle (e) are plotted. The values are normalized by corresponding observation  
 790 errors. Corresponding bell curves are depicted with standard deviation  $\sigma_{vis}$  and  $\sigma_{BT}$  and mean  $\mu_{vis}$ ,  $\mu_{BT} = 0$   
 791 (dashed red lines).



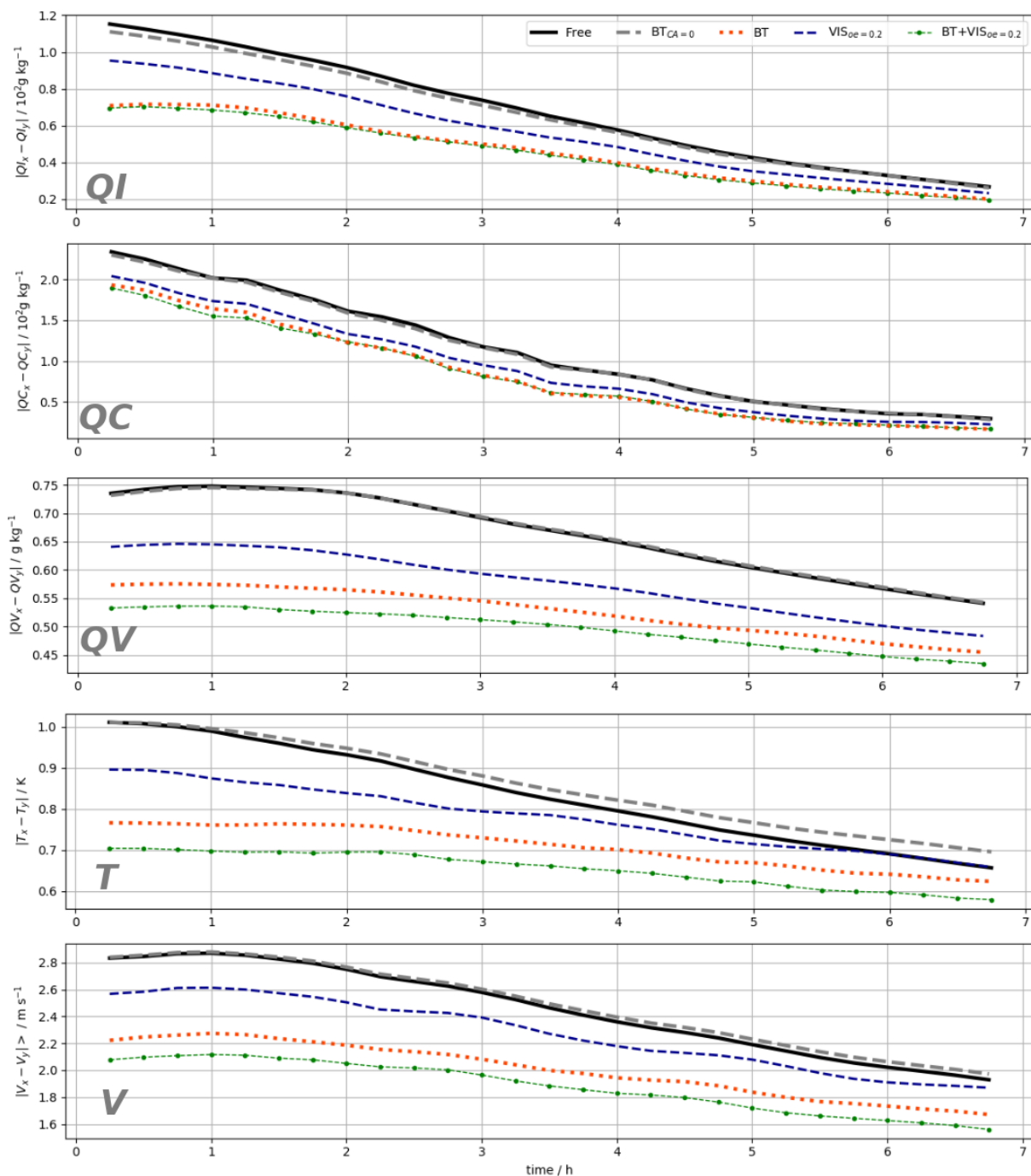
792 FIG. 5. Mean absolute errors of observation minus first guess averaged over the whole domain and ensemble  
 793 up to a height of 15 km are shown as time series for cloud ice QI, cloud water QC, water vapor QV, meridional  
 794 wind V, and temperature T (for better readability the error of QI and QC are scaled with  $10^{-2}$ ). The black line  
 795 shows the error without data assimilation. Four assimilation experiments are compared: only clear-sky bright-  
 796 ness temperature ( $BT_{CA=0}$ ), cloud-affected brightness temperature (BT), only visible reflectances ( $VIS_{oe=0.2}$ )  
 797 and a combination of both ( $BT+VIS_{oe=0.2}$ ). The observation error for the cloud-affected BTs is chosen from  
 798 an error model (Harnisch et al. 2016). The observation error for visible reflectances is set constant to 0.2. The  
 799 free forecast (*black line*) is the mean of the 40 member ensemble forecast from the experiment without data  
 800 assimilation.



801 FIG. 6. Profiles of mean absolute error are shown of the first guess during 5 h of data assimilation.  $QI$  and  $QC$   
 802 are combined in one panel. The variables and line colors correspond to the experiments in the previous figure.

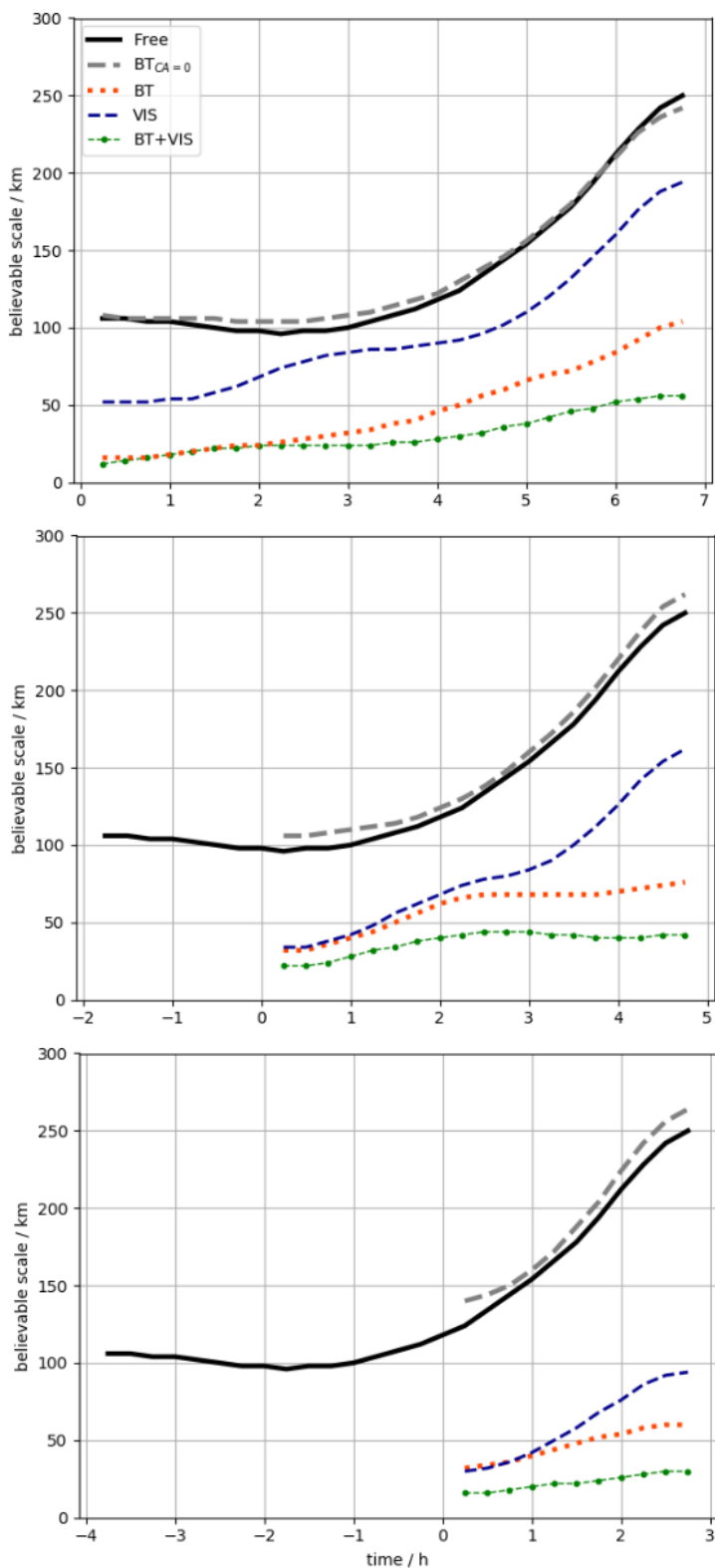


803 FIG. 7. Mean errors (“biases”) of observation minus first guess are shown as time series. The variables and  
 804 line colors correspond to the experiments in the previous figure. For comparison the zero mean error is indicated  
 805 (thin dashed black line).

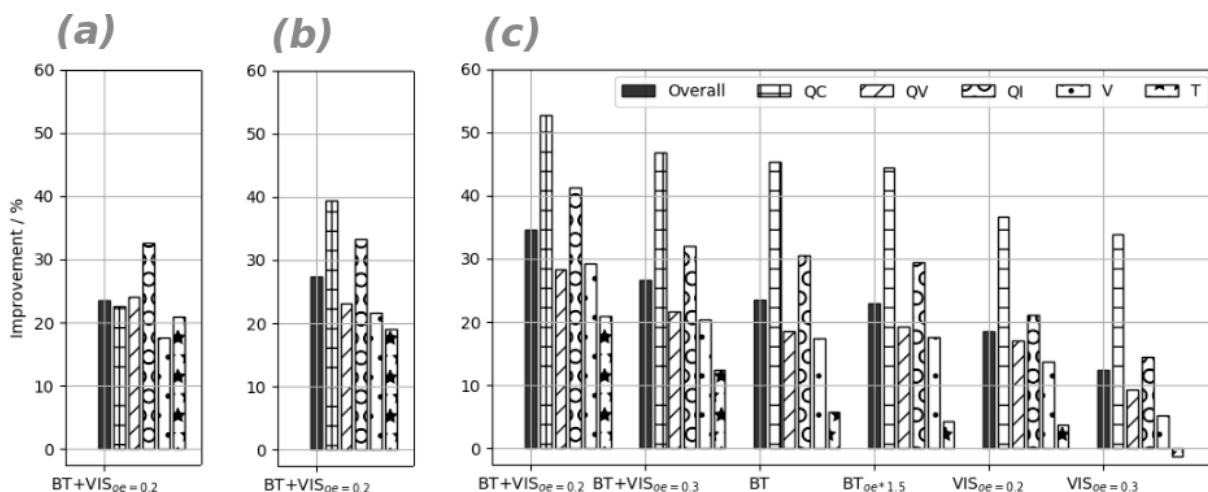


806 FIG. 8. Mean absolute error in forecasts of cloud ice (QI), cloud water (QC), water vapor (QV), horizontal  
 807 wind (V), and temperature (T) for a set of assimilation experiments. The time series are the means over all  
 808 forecast times (21 UTC, 23 UTC, 1 UTC as listed in Table 1). The line colors correspond to the experiments in  
 809 the previous figures.





810 FIG. 9. Forecasts of the believable scale of column maximum radar reflectivity starting at 21 UTC (**top**),  
 811 23 UTC (**middle**), and 1 UTC (**bottom**). The line colors correspond to the experiments in the previous figures.



812 FIG. 10. Overall improvement  $\Sigma/\Sigma_{free}$  of mean relative error for (QI), cloud water (QC), water vapor (QV),  
 813 horizontal wind (V) and temperature (T) for a set of assimilation experiments (Table 1). Presented is the im-  
 814 provement over the 8 h forecast starting at 1 UTC. All 5 variables contribute equally to the overall improvement,  
 815 i.e.,  $\Sigma/\Sigma_{free} = (\Delta QI/\Delta QI_{free} + \Delta QC/\Delta QC_{free} + \Delta QV/\Delta QV_{free} + \Delta V/\Delta V_{free} + \Delta T/\Delta T_{free})/5$ . Assimilation ex-  
 816 periments with combined instruments  $BT + VIS_{oe=0.2}$  are compared for forecasts starting at (a) 21 UTC, (b)  
 817 23 UTC, and (c) 1 UTC.

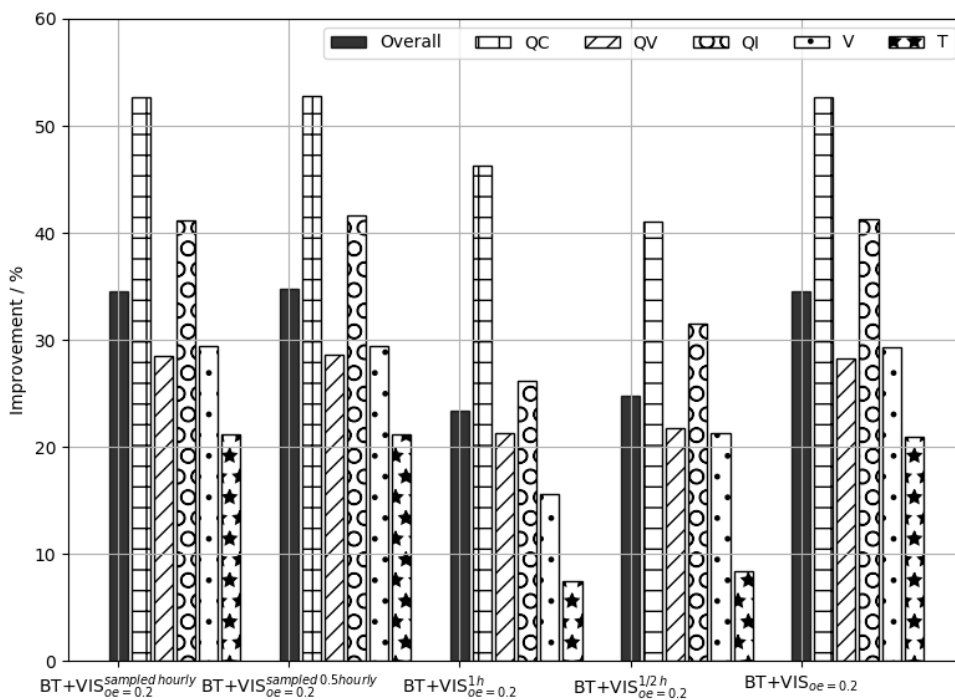
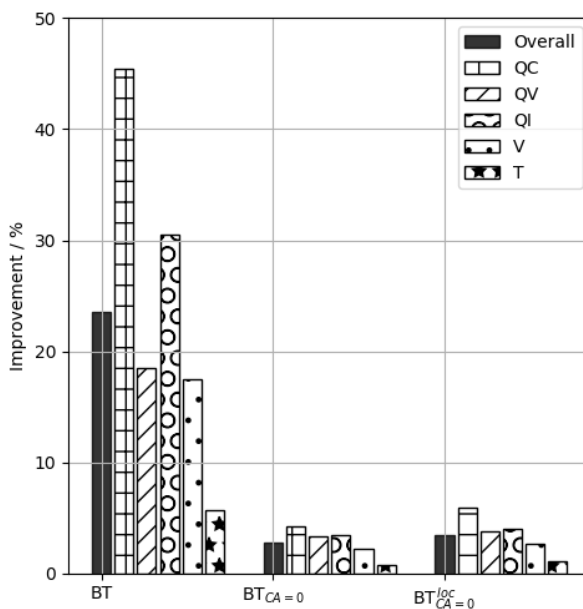


FIG. 11. As in figure 10, only varying the cycling frequency and diagnosing forecasts starting at 1 UTC.



818 FIG. 12. Diagnosing forecasts starting at 1 UTC as in figure 11, after hourly assimilating all-sky and clear-sky  
 819 brightness temperature in separate experiments.