Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system

Received: 30 August 2021 | Revised: 22 February 2022 | Accepted: 2 March 2022 | Published on: 26 April 2022

DOI: 10.1002/qj.4276

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system

Christine Sgoff | **Walter Acevedo** | **Zoi Paschalidi** | **Sven Ulbrich** | **Elisabeth Bauernschubert** | **Thomas Kratzsch** | **Roland Potthast**

Data Assimilation, Deutscher Wetterdienst, Offenbach, Germany

Correspondence
C. Sgoff, Deutscher Wetterdienst, Offenbach 63067, Germany.
Email: christine.sgoff@dwd.de

Funding information
BMVI (Federal Ministry of Transport and Digital Infrastructure) BMWi (Federal Ministry for Economic Affairs and Energy, Grant/Award Numbers: 19F2045B, 0350004B

Abstract

Near-surface temperature and humidity observations over Germany, coming on the one hand from the citizen weather station's network Netatmo and on the other hand from synoptic weather stations, were assimilated into the limited are mode of the Icosahedral Nonhydrostatic Model with 2-km resolution (ICON-D2). For that we use the Kilometre-Scale Ensemble Data Assimilation (KENDA) system and a bias-correction approach that improves the assimilation of the observations by taking into account the diurnal cycle of temperature and humidity variables. Our results show that the assimilation of bias-corrected observations from Netatmo stations reduces the forecast error considerably; meanwhile, the assimilation of Netatmo observations without bias correction leads to a strong warm bias with a negative impact on forecast performance. In contrast, for the assimilation of synoptic observations the usage of our bias-correction approach does not lead to any significant decrease in the forecast error, yet reduces the bias for the diurnal cycle of synoptic stations. Overall, it can be concluded that the forecast quality can gain from assimilating Netatmo data, provided an effective bias-correction approach is applied.

KEYWORDS

bias correction, crowd-sourced observations, data assimilation, numerical weather prediction, surface observations

1 | INTRODUCTION

Crowd-sourced observations (CSOs) form a set of emergent observational sources (Ghezzi *et al.*, 2017), the importance of which is growing over time, given the ever-increasing amount of privately owned environmental sensors connected to the internet. Currently, the ongoing expansion of the *Internet of Things* is leading to a revolutionary augmentation of the volume of data acquired from CSO networks, which now can be of orders of magnitude higher than the data from conventional networks, for some variables and geographical areas. However, this vast amount of available data cannot be exploited directly, as CSOs typically have lower quality standards than the standard scientific sensors, suffering from larger biases

Christine Sgoff and Walter Acevedo authors contributed equally to this work.

and higher noise levels. Thus, rigorous data-quality assessments are necessary in order to take advantage of these novel datasets (Zahumenský, 2004).

Regarding meteorological applications, the use of data coming from unconventional sources is revolutionary. The classical meteorological networks provide high-quality measurements for a series of variables, but they are not as dense as needed by ongoing human activities, which demand fast and precise weather forecasts (Moseley, 2011). The abundant amount of meteorological measurements in high temporal and spatial resolution from alternative and cost-efficient sensors can fill this gap and provide weather models with real-time datasets (Muller *et al.*, 2015). While the quality of CSOs is always questionable and a challenging issue to deal with, their quantity itself can be used to address this problem (Flanagin and Metzger, 2008; Foody *et al.*, 2013). For example, when a dense network in a very restricted geographical area agrees in measurements of a meteorological variable, it could indicate a level of measurement accuracy (See *et al.*, 2013; Bell *et al.*, 2015). The quality of the data and their error characteristics and interpolation are discussed in Williams *et al.* (2011) and Bell *et al.* (2013), while Bell (2014) and Tostrams (2017) have worked on quantifying their uncertainty. The impact of CSOs from citizen weather networks on meteorological models is an important research area and studies thus far underline their positive attribution in improving weather forecasts (Madaus *et al.*, 2014; Sobash and Stensrud, 2015; Gasperoni *et al.*, 2018; Hintz *et al.*, 2021).

CSOs collected from cars to smartphones and social media have already been utilized in a variety of applications (Hintz *et al.*, 2021), though the largest amount of meteorological CSOs worldwide comes from citizen weather stations. Citizen weather stations consist of privately owned automatic weather stations, cheap and easy to install, which resemble synoptic observational networks but provide a much denser coverage, especially over inhabited areas.

There are ongoing efforts to collect and share as much as possible of the information collected by citizen weather stations (MetOffice, 2021; US National Weather Service, 2021) and, in the last years, they have been used in numerous regional studies: for example, fine-scale analysis of hailstorms (Clark *et al.*, 2018), rainfall monitoring (Zinevich *et al.*, 2008; Fenner *et al.*, 2017; de Vos *et al.*, 2017; 2019), heat-island analysis (Steeneveld *et al.*, 2011; Wolters and Brandsma, 2012; Bell *et al.*, 2015; Chapman *et al.*, 2017; Meier *et al.*, 2017), and operational data assimilation (James and Benjamin, 2017).

Presently, an important network of citizen weather stations is provided by Netatmo (Netatmo, 2021). This company develops and sells worldwide internet-connected personal weather stations collecting several meteorological variables (i.e., temperature, wind speed, precipitation, air pressure, and humidity) in almost real time. This network, which is currently larger in Europe, has achieved such a dense coverage of temporally highly resolved meteorological observations over a large area that it is an innovation in atmospheric science (Chapman *et al.*, 2017). Its density in many urban areas has made it an important tool in urban climate studies (Meier *et al.*, 2015; de Vos *et al.*, 2017; Uteuov *et al.*, 2019). In addition, Netatmo observations have recently been used by MET Norway to post-process near-surface temperature forecasts (Lussana *et al.*, 2019; Nipen *et al.*, 2020).

The potential positive impact of Netatmo data on weather forecasting (Fenner *et al.*, 2017) has led to numerous studies for the assessment and quality improvement of Netatmo data, as the quality of all measurements cannot be checked manually. Tostrams (2017) presented a spatio-temporal analysis of Netatmo observations for dealing with the data uncertainty, as well as an error correction in order to address issues such as radiation bias. Furthermore, Bell (2014) worked on the bias of temperature sensors, showing the different bias patterns depending on the type of sensor, whereas Hintz *et al.* (2021) developed a quality control for temperature Netatmo measurements based, among other things, on comparisons with a numerical weather prediction (NWP) ensemble.

In the present work, we report a set of experiments where, besides the conventional observational sources, Netatmo temperature and humidity data are assimilated into the limited are mode of the Icosahedral Nonhydrostatic Model (ICON–LAM: Zängl *et al.*, 2015). A description of both observational sources is presented in Section 2, whereas the design of the experiments is given in Section 4. We combine various data sources in different ways and also utilize a bias-correction methodology, which is described in Section 3, together with the quality control that takes place prior to and during the assimilation, as well as the data assimilation environment itself. The performance of the weather prediction system under these different configurations is analyzed thoroughly and the results are presented in Sections 5, 6, and 7. Finally, Section 8 discusses the conclusions of this study.

2 | OBSERVATIONS

The experiments reported in this study consider the time period from September 17, 2018 to September 30, 2018 and make use of two observational datasets: (a) the so-called "conventional observations" of the German

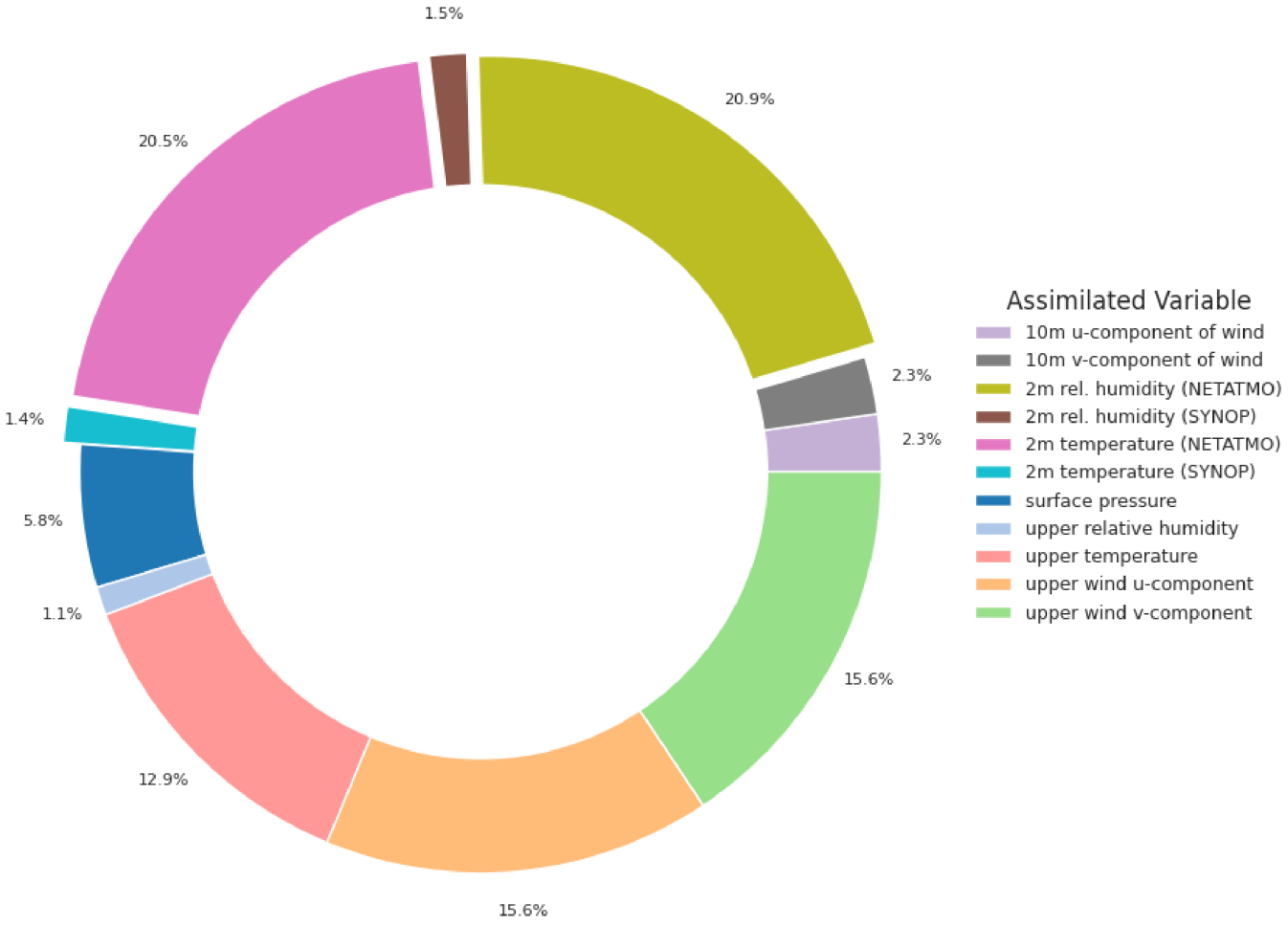

FIGURE 1 Pie chart of the number of active observations. Here we consider only one day (September 19, 2018) of our experiments, given that the distribution of observations by variable does not change considerably day by day. Notice the ratio of about 14 between Netatmo and synoptic active observations for both 2-m temperature and 2-m relative humidity. The 2-m temperature and 2-m relative humidity are assimilated over Germany, while the other observations are assimilated over the entire ICON-D2 domain [Colour figure can be viewed at wileyonlinelibrary.com]

weather service, the Deutscher Wetterdienst (DWD), which contain the synoptic weather stations, and (b) the Netatmo observations lying in German territory during the time period explored. Given the resemblance of Netatmo and synoptic weather stations, they go through the same plausibility quality control and their set-ups in the assimilation system are relatively close. Nevertheless, there are some important differences between these observation systems; accordingly, in the following we describe both datasets separately in more detail.

2.1 | Conventional observations

This dataset comprises observations coming from synoptic weather stations, as well as from aircraft, wind profilers, and radiosondes. For the time period considered, the conventional observations actually assimilated are dominated in number by the measurements of upper-air temperature and u/v wind components coming from aircraft. Regarding ground observations, the most numerous variables are 2-m temperature, 2-m humidity, 2-m dewpoint, and surface pressure (Figure 1).

2.2 | Netatmo observations

Netatmo weather stations can be found worldwide, installed by interested citizens. These devices collect time series with a period of about 5 min of temperature, humidity, pressure, and optionally of precipitation, wind speed, wind angle, wind gust speed, and wind gust angle. For the present study, we consider the dataset taken by 50,328 Netatmo stations present in Germany during the time period considered.

As our assimilation algorithm (local ensemble transform Kalman filter, LETKF) is not able to consider correlated observations, we selected 5,000 Netatmo stations

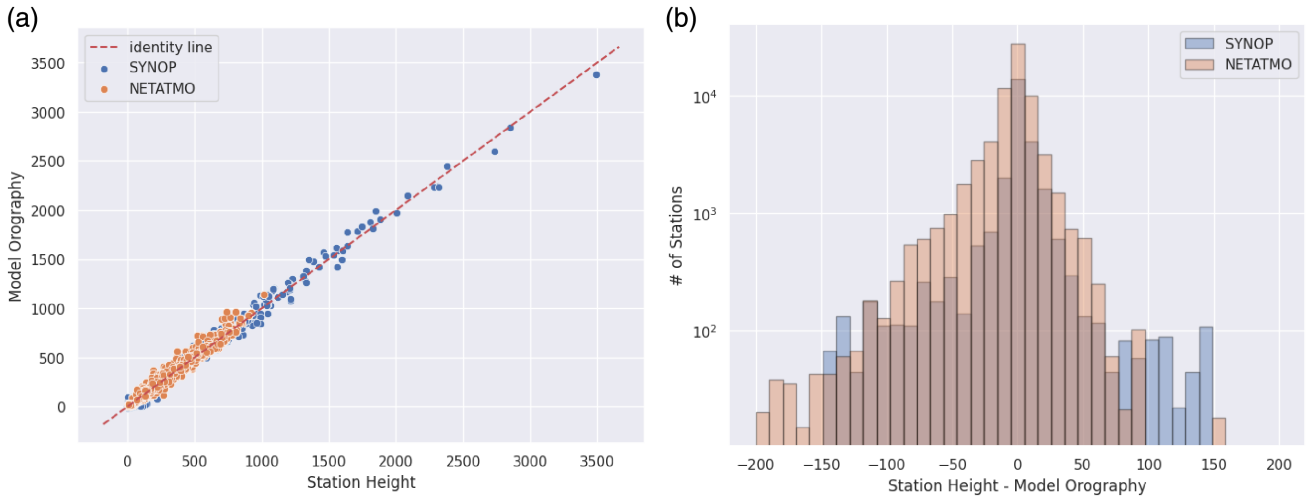

FIGURE 2 Statistical analysis of station height and model orography after quality controls: Histogram (a) of differences between station height and model orography [m] and scatter plot of (b) station height versus model orography [m]. The Netatmo stations correspond to a random selection of 5000 stations within Germany [Colour figure can be viewed at wileyonlinelibrary.com]

at random, which leads to a ratio of about 14 between Netatmo and active synoptic observations for both 2-m temperature and 2-m relative humidity. Notice that this minimalistic thinning approach assumes that all Netatmo stations provide data of similar quality, and was selected for the sake of simplicity.

Regarding altitudinal aspects, Netatmo stations display rather small differences between station heights and the corresponding model orography, represented by the orange distribution of Figure 2a, similar to the corresponding blue distribution for synoptic stations. However, it is noticeable that the Netatmo stations are located in inhabited areas, which in Germany normally correspond to lowlands, as we can see in Figure 2b, where no orange points have station heights larger than 1,000 m. Notice that the situation could be different for countries like Switzerland.

Concerning the resolution of the time series, we reduce the amount of data by dropping the first 45 min of each hour and averaging the remaining 15 min. The resulting hourly resolution is consistent with the assimilation period of our experiments. With respect to the measured variables, we assimilate only temperature and humidity observations. This selection was based on the fact that these are the only variables susceptible to our bias-correction method. Since the Netatmo stations are operated by citizens, sources of error are more numerous than is the case with synoptic stations. It can happen that stations are misplaced: for example, indoor instead of outdoor, or at a place with no radiation shield (Meier *et al.*, 2015). Observations that deviate strongly from the other measurements are sorted out in our quality control (Section 3.1). Finally, we set the observation error to 3.0 K for temperature and 30% for humidity. These values are three times higher than the corresponding ones for synoptic observations and led to reasonable results, as we will see in the following sections.

3 | METHODOLOGY

3.1 | Quality control

A series of controls before and during the assimilation processes take place, in order to guarantee the quality of the observational data for the variables being studied, that is, the 2-m temperature (T2M) and the 2-m relative humidity (RH2M).

Before any further use within the data assimilation system, both the conventional and the raw crowd-sourced Netatmo datasets are subject to an initial plausibility control. Here, the values of the raw crowd-sourced Netatmo observations were checked to be within reasonable ranges. As far as the 2-m temperature is concerned, observations with values lower than −50 °C or higher than 50 °C are discarded. Relative humidity observations lower than 0% and higher than 110% are also discarded.

For the data assimilation, in the quality control, during the derivation of the model equivalents $H(\mathbf{x})$, the synoptic observations are checked for their first guess and their station height. If the difference between observation i and its corresponding first guess $\mathbf{d}^{\mathrm{o}}_{\mathrm{fg}}(i) = \mathbf{y}(i) - H(\mathbf{x}(i))$ is too large, the observations are discarded. The threshold for the relative humidity observations is $\mathbf{d}^{\mathrm{o}}_{\mathrm{fg}}(RH2M) > 70\%$ and that for the temperature observations $\mathbf{d}^{\mathrm{o}}_{\mathrm{fg}}(T2M) > 12\,\mathrm{K}$. The height check examines the height of the observation station itself and the height difference between the measuring station and the model orography. Temperature

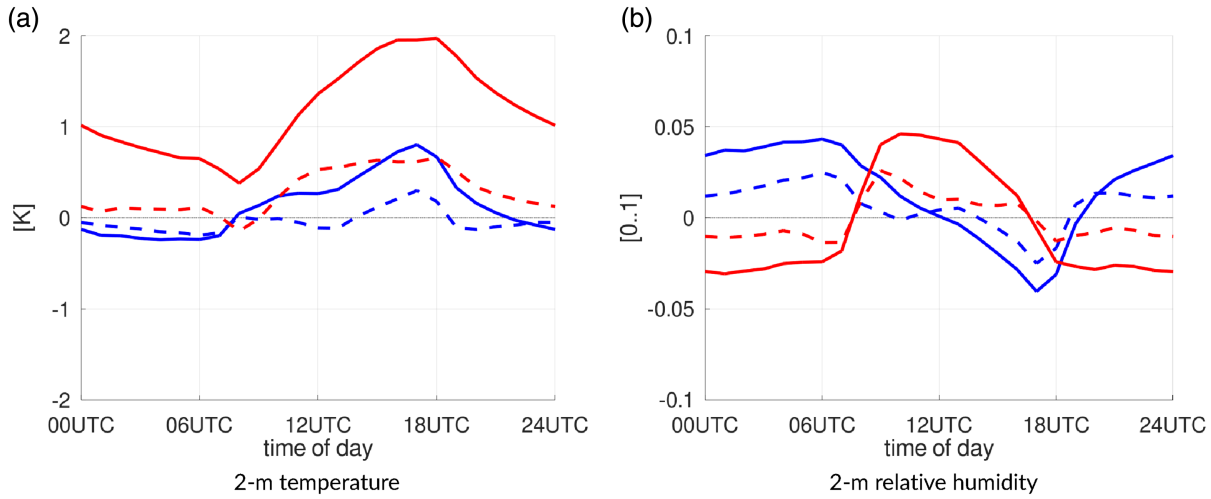

FIGURE 3 Mean diurnal cycle of difference between model and synoptic observations (blue) and Netatmo observations (red) of the September period (September 19–30, 2018) over Germany. Difference between (a) simulated and observed 2-m temperature and (b) simulated and observed 2-m relative humidity. The differences between the model and observations without bias correction (solid line) are larger than the bias-corrected differences (dashed line). In particular, the warm bias of the temperature measured by Netatmo stations is clearly reduced by the bias correction [Colour figure can be viewed at wileyonlinelibrary.com]

and relative humidity observations from stations above 5,000 m height are rejected, as well as stations with height differences larger than 150 m.

An additional quality control for synoptic and Netatmo observations that depends on the ensemble is now performed:

$$|obs - fg| > 3\sqrt{\sigma^2_{\text{obs}} + \sigma^2_{\text{model}}}, \quad (1)$$

with σ_{obs} the observational error and σ_{model} the model error defined by the ensemble spread.

3.2 | Bias and bias correction

The model equivalent of 2-m temperature is calculated from the model diagnostics and adjusted by lapse rate to the observation height. The lapse rate is calculated from the temperature and height of the lowest model levels close to the observation. Afterwards, the corrected 2-m temperature is used to calculate the model equivalent of 2-m relative humidity.

An evaluation of the observation-minus-first guess departures $\mathbf{d}^{\text{o}}_{\text{fg}}$ of synoptic and Netatmo observations, for the study period of two weeks in September, shows a diurnal bias pattern in temperature and relative humidity for both measurement systems (Figure 3). The temperature observations from synoptic weather stations show a diurnal cycle with negative bias during the night and positive bias during the daytime (Figure 3a). The temperature from Netatmo stations shows, on average, a stronger diurnal cycle with a similar shape, but has, in general, a warm bias of around 1 K. The synoptic relative humidity observations have a dry bias during the afternoon (1200–1900 UTC) and a moist bias otherwise (Figure 3b). The amplitude of the bias of relative humidity observed by Netatmo stations is similar to the amplitude of the bias of the synoptic observations (between −0.05 and 0.05), but, by contrast, the Netatmo observations have a moist bias during the day (0800–1700 UTC) and a dry bias during the night.

As far as the synoptic observations are concerned, the results are as expected, considering the strong dependence of the temperature bias on the solar radiation. During the daytime, the warm bias of temperature observations is related to the dry bias of the relative humidity, whereas during the night hours the moist bias is influenced by the cold temperature bias. However, this expected behaviour cannot totally explain the stronger diurnal cycle of temperature measured by Netatmo stations and the opposite bias behaviour of relative humidity Netatmo data. Here, the position of the Netatmo stations plays a major role, as they can be indoors or near house walls and so poor and insufficient ventilation leads to higher temperature measurements. Additionally, as highlighted in Bell (2014), the bias of temperature can also be related to the type of sensor. The use of cheap temperature sensors may give larger biases. Relative humidity measurements can be biased as well, especially if they are not recalibrated regularly (Ingleby *et al.*, 2013).

Based on the idea of Otkin *et al.* (2018), who used a Taylor series polynomial expansion of $\mathbf{d}^{\text{o}}_{\text{fg}}$ to estimate the bias of infrared brightness temperatures measured

by the Spinning Enhanced Visible and Infrared Imager (SEVIRI), we predict the bias b through a set of basis functions $\mathbf{\Phi}$, which use the daytime t and cloud cover N as predictor:

$$\left[b^{(i)}(N,t) = \mathbf{\Phi}(N,t) \cdot \mathbf{c}^{(i)}\right]_{i=1,\ldots,m} . \quad (2)$$

The index $i = 1, \ldots, m$ indicates that the bias b is estimated separately for each measuring instrument at each station. The cloud cover as a predictor is relevant, because a strong cloud cover reduces the amplitude of the diurnal bias. The set of basis functions $\mathbf{\Phi}$ is split into a set of k trigonometric basis functions U to estimate the diurnal cycle,

$$U_k(t) = \{1, \sin t\tau, \cos t\tau, \sin 2t\tau, \cos 2t\tau, \ldots\}, \quad (3)$$

and a set of l polynomial basis functions V to estimate the impact of the cloud cover,

$$V_l(N) = (9 - N)^{(l-1)} . \quad (4)$$

The cloud cover N is set to 8 okta if the station has no cloud-cover observation, which is true for several synoptic and all Netatmo stations. In these cases, the bias correction is determined by the diurnal cycle alone. The period of the trigonometric basis functions is one diurnal cycle, $\tau = 2\pi/24$. A useful estimate is given by a combination of $K = 5$ trigonometric basis functions and $L = 2$ polynomial basis functions:

$$\left[b^{(i)}(N,t) = \sum_{k=1}^{K}\sum_{l=1}^{L} U_k(t)V_l(N)c_{kl}^{(i)}\right]_{i=1,\ldots,m} , \quad (5)$$

where each measuring instrument i has its own set of coefficients $\mathbf{c}_{kl}^{(i)}$.

The system is an underdetermined system (Equation 2) and a common approach to find a solution for an ill-posed system is Tikhonov regularization (Otkin *et al.*, 2018; Nakamura and Potthast, 2015):

$$\mathbf{c}^{(i)} = (\alpha\mathbb{I} + \mathbf{\Phi}\mathbf{\Phi}^{\mathrm{T}})^{-1}\mathbf{\Phi}^{\mathrm{T}}b^{(i)}, \quad (6)$$

with a multiple of the identity matrix $\alpha\mathbb{I}$ as the Tikhonov matrix. In the following, $(\alpha\mathbb{I} + \mathbf{\Phi}\mathbf{\Phi}^{\mathrm{T}})^{-1}\mathbf{\Phi}^{\mathrm{T}}$ is summarized as weighting matrix $\mathbf{K}_{\mathrm{c}}$. The coefficients are determined dynamically during the experimental period by a coefficient update for the i separate stations and measurement instruments for each time step at which a new observation is available.

Taking the information from the previous step $t - 1$ and the actual time step t into account, the coefficients adapt to the bias behaviour of their corresponding station and observation type. The estimated bias $b^{(i)}(t)$ used to update the coefficients is derived from Equation 2 with the coefficients of the previous analysis step $\mathbf{c}(t - 1)$. With a higher value of α, the influence of the coefficients of the previous analysis steps increases. We choose $\alpha = 150$: this results in coefficients that adapt to a new weather regime within a two-day period.

The estimated bias b is used to adjust the observation with the bias correction $bc = -b$, although b is the estimation of a combined bias consisting of observation and forecast bias. The bias-correction approach introduced reduces the amplitude of the bias for both types of observation (humidity and temperature) for synoptic and Netatmo stations (Figure 3). As expected, the bias correction adapts well to the existing systematic deviations between observation and model after a few days and reduces the mean bias between 50 and 70% for the entire period. Adjusted to the measuring instruments, the mean bias correction of temperature measured by synoptic stations is smaller than 0.5 K during the whole diurnal cycle (Figure 3a). Furthermore, the mean temperature bias correction of the Netatmo stations is able to reduce the 2-K evening bias by more than 1 K.

Because our bias-correction scheme depends on the model background field, it is important to have "anchoring" observations within the assimilation framework (Eyre, 2016). The anchoring observations should prevent the model state relaxing towards its own climatology due to the bias correction. In our assimilation framework, radiosoundings, wind profiler, aircraft observations, and surface pressure from synoptic stations are used as anchoring observations.

3.3 | Data assimilation system

As assimilation system, we used the operational Kilometre-scale ENsemble Data Assimilation system (KENDA: Schraff *et al.*, 2016), which is based on the LETKF (Hunt *et al.*, 2007). The KENDA-system ensemble consists of 40 members and one additional deterministic simulation (Schraff *et al.*, 2016). The analysis of the deterministic simulation is determined by the Kalman gain matrix of the LETKF. Therefore, the deterministic simulation benefits from the background-error covariance matrix of the ensemble $\mathbf{P}$ and further can be used for the evaluation of cloud-related variables, because these are smeared in the ensemble mean by its averaging. The localization is performed at each analysis grid point by scaling the inverse observation-error covariance matrix $\mathbf{R}^{-1}$ according to the distance between the observation and the analysis grid point. The analysis grid is three times coarser than the model grid, hence the weight matrices

are interpolated to the model grid afterwards. The scaling is realized by use of the Gaspari–Cohn function (Gaspari and Cohn, 1999) with a fixed vertical localization scale and an adaptive horizontal localization scale (Schraff *et al.*, 2016). Covariance inflation is achieved by a combination of multiplicative covariance inflation (Anderson and Anderson, 1999) and relaxation to prior perturbations (Zhang *et al.*, 2004). Furthermore, the KENDA system combines the LETKF with latent-heat nudging for the assimilation of radar-derived precipitation rates (Stephan *et al.*, 2008; Schraff *et al.*, 2016).

The KENDA system uses the limited-area mode of the ICOsahedral Nonhydrostatic (ICON) model (ICON–LAM: Zängl *et al.*, 2015) with $k = 40$ ensemble members as regional weather prediction model. The experimental domain covers Germany, Switzerland, Austria, Denmark, Belgium, the Netherlands, and parts of further neighbouring countries. The ICON–LAM is performed on an unstructured triangular horizontal grid of 542,040 cells with a spatial resolution of about 2 km. The 65 vertical levels follow the terrain in the lower troposphere and become horizontally flat in the upper troposphere. The atmosphere of the model consists of dry air and water in all phases (gaseous, liquid, ice). The prognostic variables of the model are the horizontal velocity component normal to the triangle edges, the vertical wind component, the density and the virtual potential temperature. The radiation is simulated every 12 min by the Rapid Radiative Transfer Model (Mlawer *et al.*, 1997; Barker *et al.*, 2003). The ICON–LAM is driven and initialized by the European two-way nest of the global ICON. Further details can be found in Zängl *et al.* (2015) and Prill *et al.* (2020).

4 | EXPERIMENTAL SET-UP

We conducted one reference and four case experiments within the ICON–LAM–KENDA system. The experiments refer to the period from September 17–September 30, 2018. Because the first two days of this period are used as spin-up of the bias correction, the evaluation period starts at September 19, 2018. The hourly assimilation was based on a 40-member ensemble and one deterministic run. Each 6 hr, that is, at 0000, 0600, 1200, and 1800 UTC, a 24-hr forecast was initialized from the deterministic analysis. The reference experiment REF included latent-heat nudging based on radar observations and the assimilation of surface pressure and 10-m wind observations of synoptic and buoy stations, as well as temperature, relative humidity, and horizontal wind components observed by aircraft, wind profilers, and radiosondes, as mentioned in Section 2.1. That set-up corresponds mainly to the operational set-up at DWD in 2018 when three-dimensional (3D) volume radar data had not yet been assimilated, and in the following we refer to it as the CONV set-up. Within the experiments, we compare the impact of synoptic and Netatmo observations and additionally the performance of the adaptive bias-correction approach. Hence, we conducted experiments where we assimilated synoptic T2M and RH2M, one without (SNP) and one with bias correction (SNP_BC), in addition to the CONV set-up. Likewise, we performed experiments with assimilation of Netatmo observations without (NTM) and with bias correction (NTM_BC), additionally to the CONV set-up. A brief overview of the experiments can be found in Table 1.

TABLE 1 Overview of the experiments performed

Exp. name	Assimilated obs.	Bias correction
REF	CONV	No
SNP	CONV + synoptic T2M+RH2M	No
SNP_BC	CONV + synoptic T2M+RH2M	Yes
NTM	CONV + Netatmo T2M+RH2M	No
NTM_BC	CONV + Netatmo T2M+RH2M	Yes

Note: CONV includes the following observations: surface pressure and 10-m wind observations of synoptic and buoy stations, as well as temperature, relative humidity, and horizontal wind components observed by aircraft, wind profilers, and radiosondes. Additionally, latent-heat nudging is included, but not 3D volume radar, which has been operational since June 2020.
Abbreviations: T2M, 2-m temperature; RH2M, 2-m relative humidity.

5 | IMPACT ON FORECAST OF 2-M VARIABLES

The 24-hr forecasts of 2-m temperature, 2-m relative humidity, and 2-m dewpoint temperature of the reference and the four experiments are verified by observations of the synoptic observation network. In contrast to the synoptic weather stations, Netatmo observations are not quality controlled and do not follow World Meteorological Organization (WMO) standards. Furthermore, there is uncertainty about the exact location and height of the Netatmo stations. To ensure a reasonably fair comparison of the experiments, we split the synoptic observations into two subsets: one subset is used for assimilation and the other for verification. Due to the fact that the bias correction depends on the respective first guesses, it differs from experiment to experiment. Thus, we use observations without bias correction for verification.

The impact of assimilation of synoptic T2M and RH2M on the forecasts initialized from the assimilation cycle is mainly positive (Figure 4). The root-mean-square error

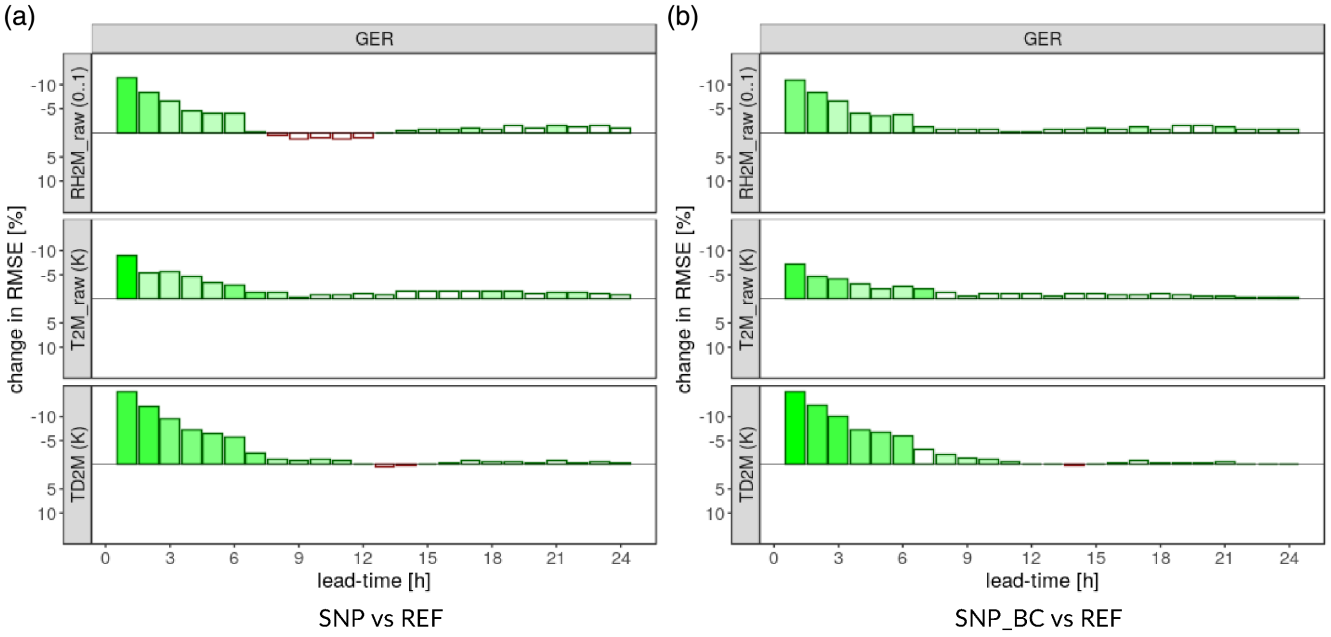

FIGURE 4 Comparison of RMSE of different variables for the different experiments for different lead times, based on temperature and humidity measurements from synoptic weather stations within Germany (GER). Shown are the 2-m relative humidity (RH2M_raw), 2-m temperature (T2M_raw), and 2-m dewpoint temperature (TD2M). The forecasts were started from September 19–30, 2018 at 0000, 0600, 1200, 1800 UTC. In (a), a positive value indicates that SNP has a reduced RMSE compared with REF, a negative value indicates the opposite. The brightness indicates the significance of the results (based on a *t*-test). (b) shows the same as (a), but for SNP_BC versus REF [Colour figure can be viewed at wileyonlinelibrary.com]

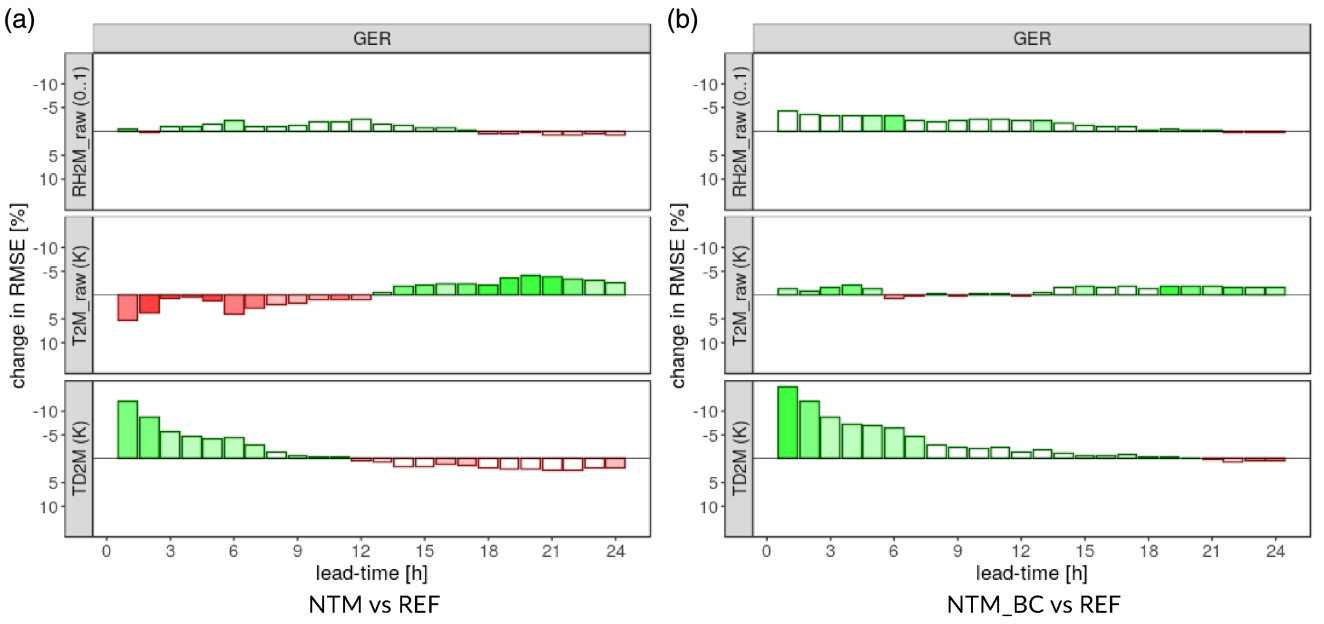

FIGURE 5 Like Figure 4 but for (a) NTM versus REF and (b) NTM_BC versus REF [Colour figure can be viewed at wileyonlinelibrary.com]

(RMSE) of the forecast of relative humidity, temperature, and dewpoint temperature at 2-m height is reduced significantly in the first six forecast hours for both experiments (SNP and SNP_BC). After the first 6 hr, the reduction of RMSE is below 3% but still positive, although less significant. An exception is the 2-m relative humidity of the experiment SNP, where a negative impact on the RMSE can be found around a lead time of 9 hr. This negative impact can be overcome by using bias correction (SNP_BC). However, within SNP_BC the assimilation has little less impact on the reduction of the RMSE of 2-m temperature.

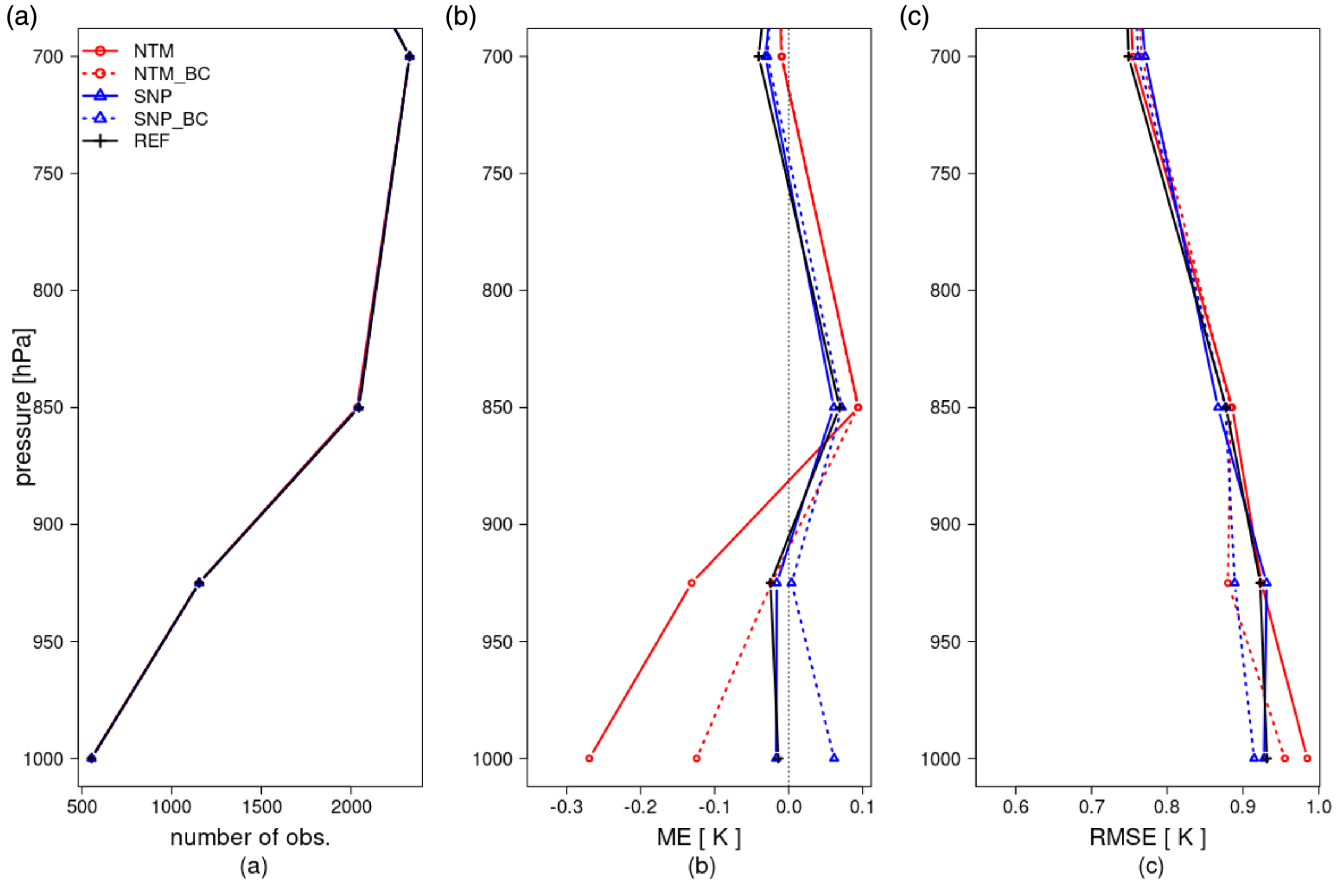

FIGURE 6 Upper-air verification at different heights, comparing the different experiment results from the assimilation cycle with observations (radiosoundings) in the experiment period (September 19, 2018, 0000 UTC–September 30, 2018, 0000 UTC). Solid lines indicate the results of first guesses without bias correction and dotted lines those with bias correction. The different experiments are depicted as different coloured lines (see legend and Table 1). Shown are (a) the number of observations used for the statistics, (b) the mean error of the temperature model equivalent, and (c) the corresponding RMSE [Colour figure can be viewed at wileyonlinelibrary.com]

The effect of assimilating temperature and relative humidity measured by Netatmo stations depends on the use of the bias correction (Figure 5). The Netatmo observations assimilated without bias correction (NTM) have a rather neutral impact on 2-m relative humidity. During short lead times, the RMSE of the 2-m temperature forecasts initialized from NTM is increased compared with the 2-m temperature RMSE of the forecasts initialized from REF. In contrast, for long lead times, the 2-m temperature forecast RMSE of NTM is reduced compared with REF. This effect could be due to cooling during the model forecasts. By assimilating the Netatmo observations, which include a warm bias, the initial model state near the surface is too warm. However, during the forecast the 2-m temperature decreases and thereby the atmospheric model state becomes closer to the observations, where in the meanwhile the atmospheric model state of REF cools down too much. The forecast of 2-m dewpoint temperature profits from the assimilation of Netatmo observations during the first 12 forecast hours; afterwards the impact is negative. The negative effects on 2-m temperature and dewpoint temperature vanish if the bias-corrected Netatmo observations are assimilated (NTM_BC). Hence, the impact of assimilating bias-corrected Netatmo observations is successful, even if it is lower than the impact of for assimilating synoptic observations.

6 | IMPACT ON FIRST GUESS OF UPPER-AIR TEMPERATURE AND HUMIDITY

In addition to the impact on the forecast of 2-m variables, the assimilation of near-surface observations has an impact on values and processes in the atmospheric boundary layer. To quantify this impact, we evaluate the hourly output of the assimilation cycle and the 24-hr forecasts. Hereby, observations of temperature and relative humidity measured by radiosondes are used for verification.

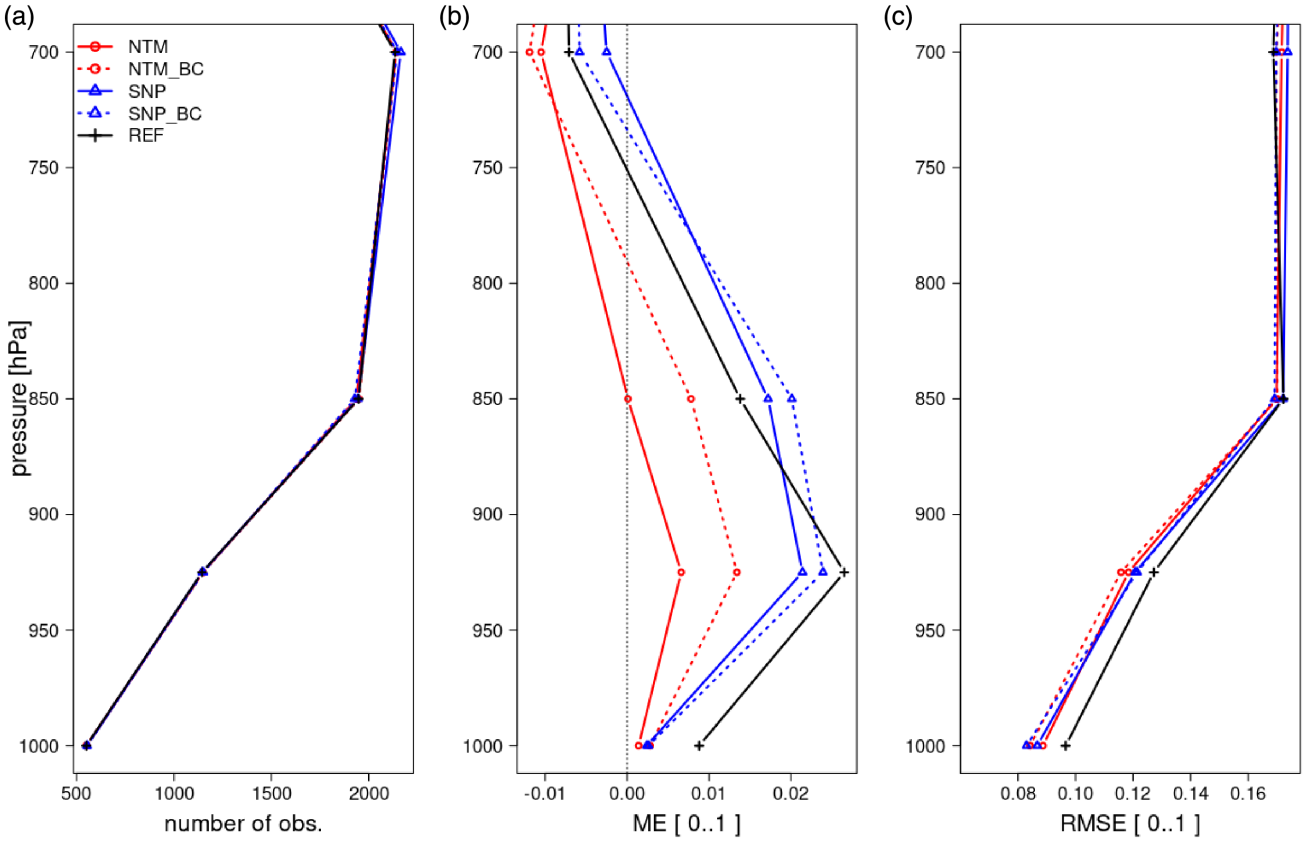

FIGURE 7 Like Figure 6, but for the upper-air relative humidity verification [Colour figure can be viewed at wileyonlinelibrary.com]

For the evaluation of the first guess, the mean error (ME) and RMSE of the experiments are calculated (Figures 6 and 7). The first-guess temperature profiles of REF and SNP have no bias near the land surface and a small cold bias at 850 hPa. The assimilation of bias-corrected synoptic T2M and RH2M introduces a small cold bias near the land surface; meanwhile, the assimilation of Netatmo T2M and RH2M introduces a warm bias, which is nearly 0.2 K stronger if no bias correction is applied. The bias correction has a cooling effect for both kinds of observation. Therefore, the bias correction has a positive impact on the Netatmo observations and reduces the near-surface temperature bias by nearly 0.2 K.

The largest temperature RMSE in the lowest level is found for the NTM experiment, with a value of almost 1 K. Compared with REF, the assimilation of Netatmo observations increases the RMSE close to the land surface, whereas the assimilation of synoptic observations reduces the RMSE. For both observation types, the bias correction improves the RMSE. In total, SNP_BC has the strongest RMSE reduction.

REF has a slight dry bias, which is reduced in each of the assimilation experiments (Figure 7b). It is noticeable that the bias is even more reduced if the observations of Netatmo stations are assimilated. The RMSE of relative humidity is reduced in each of the assimilation experiments compared with REF (Figure 7c).

To sum up, the bias correction improves the use of Netatmo observations within the assimilation cycle, while the assimilation of the bias-corrected synoptic temperature leads to a stronger cold bias compared with our reference experiment. This could be an indication that the assimilation of the bias-corrected synoptic temperature pulls the model towards its own climatology. In this case, it would be better to use the synoptic temperature as anchor observation. Overall, it can be stated that the difference between the impact of synoptic and Netatmo observations on the first guess is more defined for temperature than for relative humidity.

7 | IMPACT ON 24-HR FORECASTS OF UPPER-AIR TEMPERATURE AND HUMIDITY

The 24-hr forecasts initialized at 0000, 0600, 1200, and 1800 UTC from the deterministic analysis are verified further with radiosonde observations and aircraft measurements. The verification with radiosoundings shows that the assimilation of synoptic 2-m temperature and

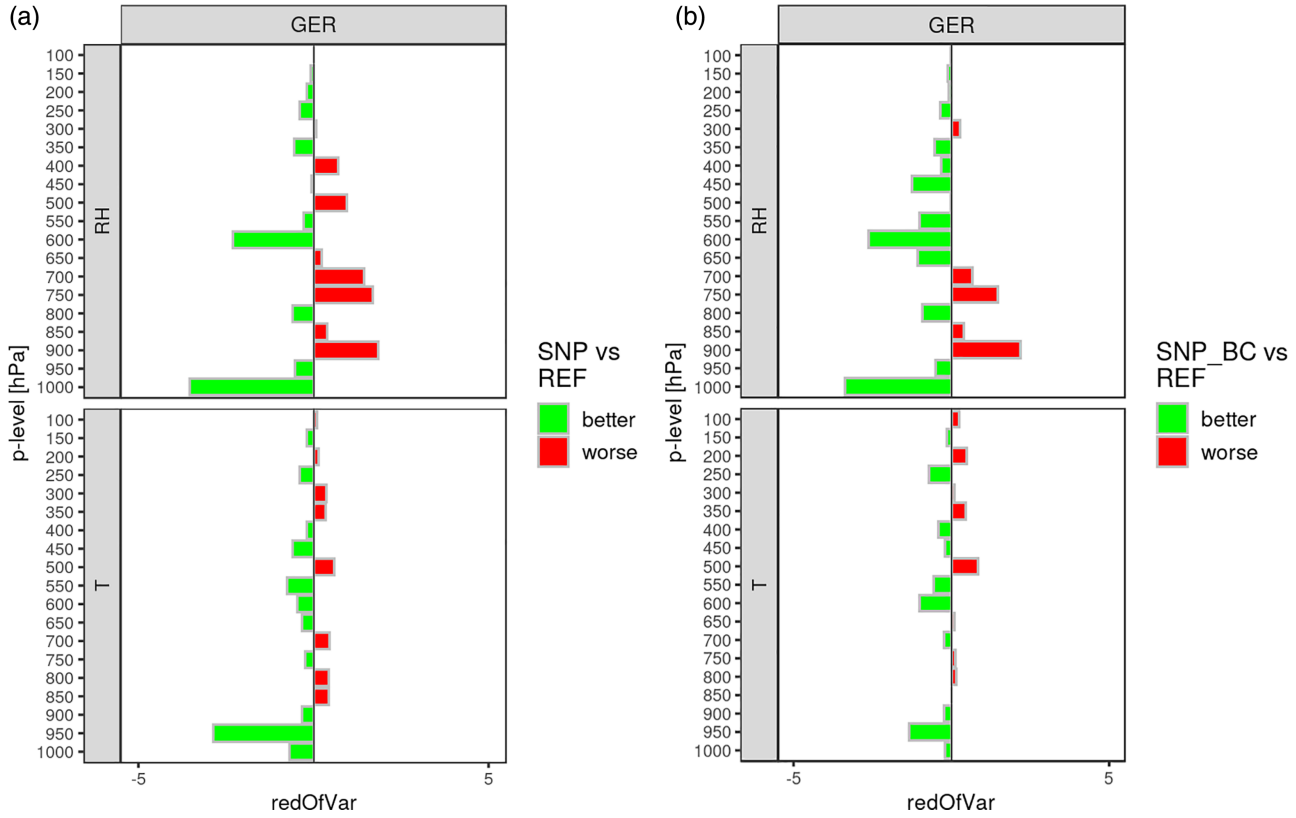

FIGURE 8 Upper-air verification above Germany (GER) at different heights, comparing the different experiment results from the 24-hr forecasts with relative humidity (RH) and temperature (T) observations of radiosoundings in the experiment period (September 19, 2018, 0000 UTC–September 30, 2018, 0000 UTC). The 24-hr forecasts are started at 0000, 0600, 1200, and 1800 UTC for the experiments REF, SNP, and SNP_BC, while REF serves as a reference (Table 1). In (a), green indicates that SNP has a reduced RMSE [%] compared with REF and red indicates the opposite. (b) shows the same as (a), but for SNP_BC versus REF [Colour figure can be viewed at wileyonlinelibrary.com]

relative humidity has a positive impact on temperature and relative humidity for the 24-hr forecasts as well. The reduction of RMSE for the near-surface levels ranges between 3 and 4%, but the bias correction reduces the positive impact (Figure 8). In contrast, the assimilation of temperature and relative humidity measured by Netatmo stations without bias correction leads to an increase of RMSE compared with the forecast initialisation from REF (Figure 9a). However, the bias correction has a positive impact and the increased temperature and relative humidity RMSE of the lower atmospheric layers (NTM) is reduced or even decreased (Figure 9b). Thereby, SNP and NTM_BC reduce the RMSE of temperature and relative humidity by a similar order of magnitude.

The verification with aircraft temperature observations shows a different image. Here, each assimilation experiment reduces the RMSE compared with REF, but the strongest reduction is achieved by the NTM experiment (Figures 10 and 11). This result may have several reasons. Firstly, the aircraft data used is composed of data from the Aircraft Meteorological Data Relay (AMDAR) and the Mode-S Enhanced Surveillance system (Mode-S), where both have a known warm bias (Ballish and Kumar, 2008; de Haan, 2011; de Haan *et al.*, 2021). Therefore, the positive impact within the aircraft verification could be due to the fact that both NTM and the aircraft observations used have a positive bias. In the recent version of Mode-S temperature, there is a correction applied that reduces the warm bias significantly. However, this was not yet active in our 2018 data. Secondly, the positive impact of Netatmo observations could be due to the location of Netatmo and aircraft measurements. The assimilated Netatmo observations are more dense within city regions with high population density (Figure 12a). Here, the observation data are likely to be warmer than the corresponding atmospheric state. Because ICON does not involve an urban canopy model, it cannot reproduce local properties related to cities. Furthermore, airports are mainly located within or close to city regions (Figure 12b). Hence, observations of the lower atmosphere by aircraft are mainly located within the same regions as the Netatmo observations. In contrast, SYNOP and TEMP observations are located outside regions with a dense population (Figure 12c,d). Thus, the bias correction is more active for the Netatmo dataset, resulting

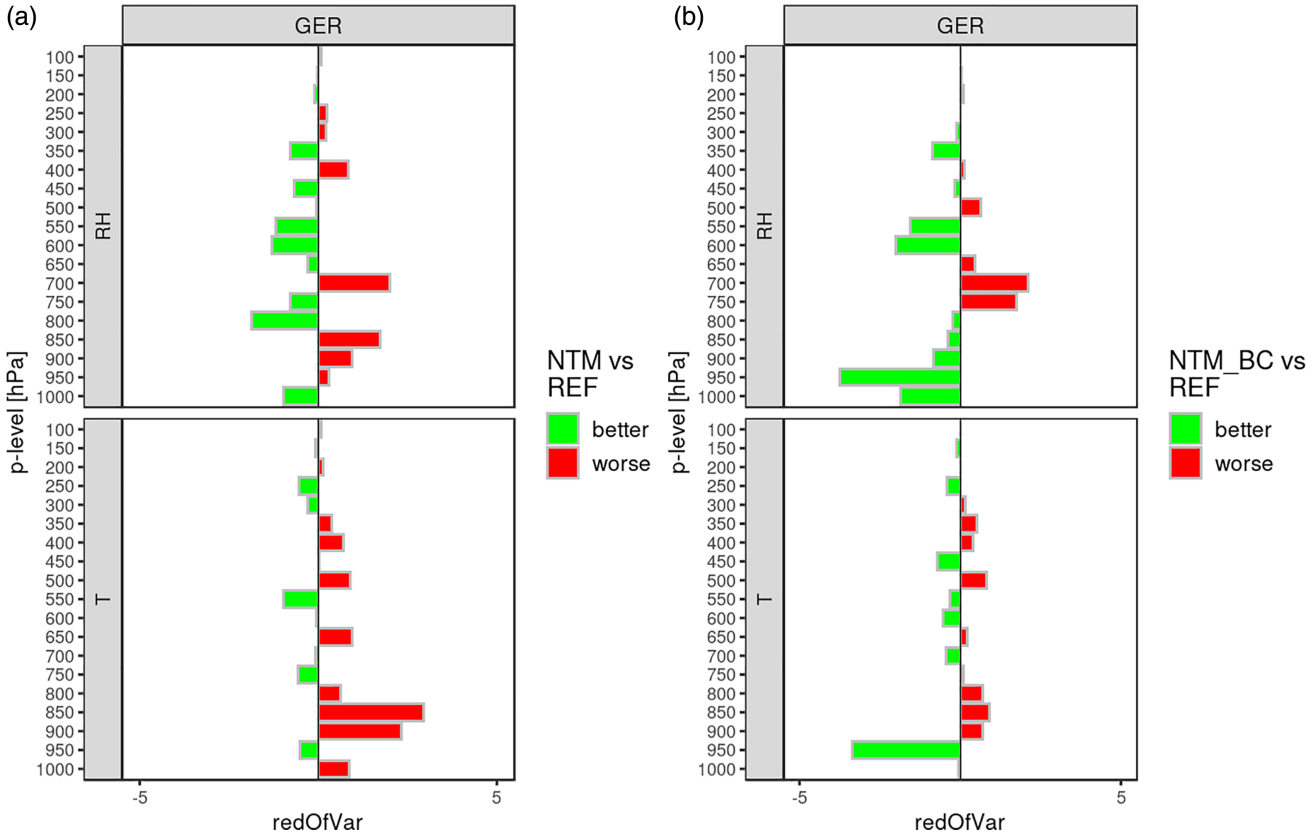

FIGURE 9 Like Figure 8, but for (a) NTM versus REF and (b) NTM_BC versus REF [Colour figure can be viewed at wileyonlinelibrary.com]

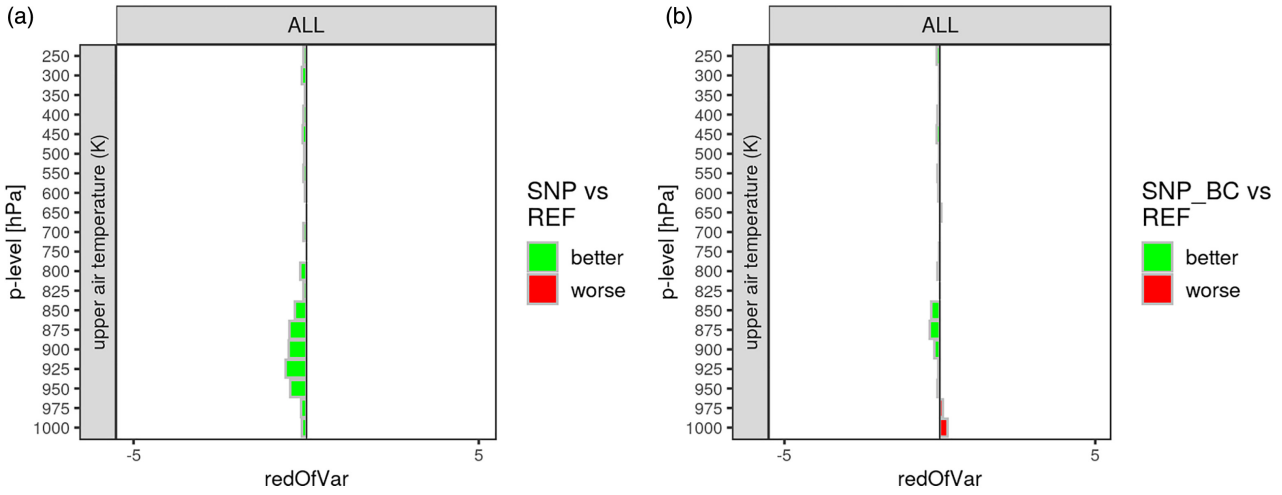

FIGURE 10 Like Figure 8, but here aircraft (AMDAR and Mode-S) observations are used for the verification [Colour figure can be viewed at wileyonlinelibrary.com]

in more stations accepted in the assimilation cycle. This indicates that some of the effects of parametrization of cities are reflected in the bias correction. This emphasizes that a parametrization of effects of cities might be useful in ICON. However, further studies are required to investigate the effects of the associated biases. It is also possible that both effects come into play and support each other.

8 | CONCLUSION

The large amount of temporal and spatial high-resolution crowd-sourced observations can be beneficial to weather prediction, if a level of quality can be ensured. Thus, in the present work, near-surface data of temperature and humidity from Netatmo stations all over Germany are

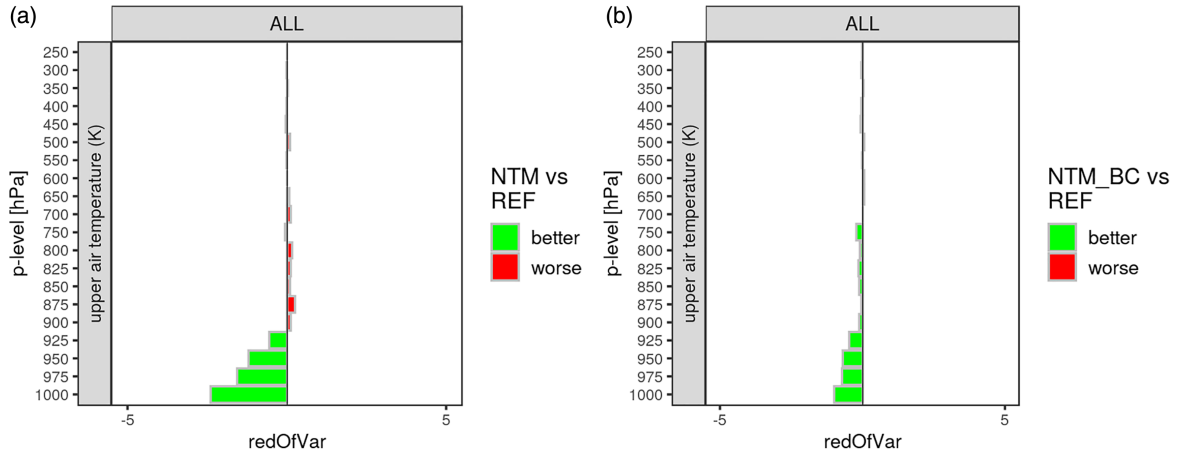

FIGURE 11 Like Figure 10, but for (a) NTM vs REF and (b) NTM_BC versus REF [Colour figure can be viewed at wileyonlinelibrary.com]

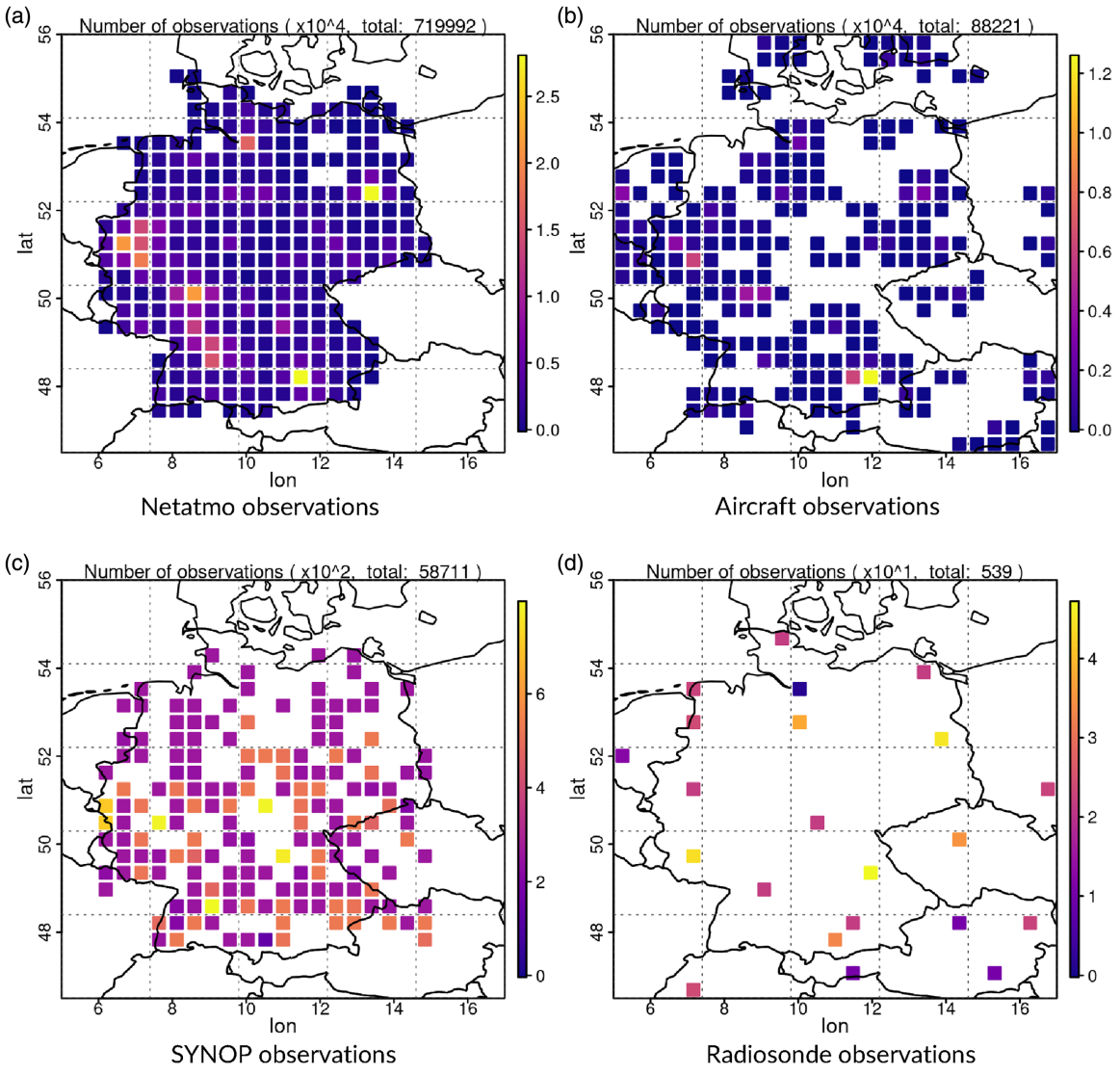

FIGURE 12 Distribution of temperature observations over Germany from September 19–30, 2018. The brighter the colour, the more observations are available. In (a), the coloured boxes indicate the number and location of Netatmo stations; (b) shows the same as (a), but for aircraft measurements between 800 and 1,000 hPa; (c) shows the same as (a), but for synoptic stations; and (d) also shows the same as (a), but for radiosonde measurements [Colour figure can be viewed at wileyonlinelibrary.com]

assimilated into the regional weather model ICON–LAM. For a complete study, the performance of Netatmo observations is compared with the assimilation of the synoptic data available for the same time period.

Firstly, the bias of the observations is studied and a bias-correction method is applied, which takes into account the diurnal cycle of temperature and humidity. Secondly, a series of different assimilation experiments is performed (Table 1), in order to assess the impact of the observations (with and without bias correction) from both sources in the model forecast. Verification of the results is performed using non-assimilated surface and upper-air observations, which can thus be regarded as independent. The results are positive and promising, showing the beneficial potential of the assimilation of crowd-sourced data for weather forecasts.

The most noticeable bias is the warm bias of the 2-m temperature data of the Netatmo stations. Because Netatmo stations are run by citizens, they are mainly located in urban areas and can suffer from poor siting and ventilation. Our bias-correction approach is able to reduce the daily temperature RMSE up to 1.5 K. The strong warm bias of 2-m temperature further affects the dew point. On the other hand, the high-quality synoptic observations show a low warm bias (less than 1 K) during the afternoon, which can be significantly reduced by assimilation with bias correction. The impact of bias correction regarding the humidity is correspondingly positive. Assimilation of the bias-corrected synoptic observations removes the negative impact that the non-bias-corrected data have on humidity.

The impact on assimilation results for both the bias- and non-bias-corrected synoptic observations is positive and stronger than the equivalent for Netatmo data, even though the Netatmo observations are more numerous. Still, verification of the assimilation of Netatmo observations against aircraft data shows a strong positive impact. Two facts play a significant role in this result: on the one hand, both Netatmo and aircraft observations (AMDAR and Mode-S) have a warm bias. On the other hand, this can be explained by the fact that both kinds of dataset are observed mostly above cities, where ICON–LAM does not include an urban canopy module.

The promising results of the current work highlight the potential of further studies with Netatmo or other crowd-sourced observations. A first step would be to extend the experiment period in order to check how the assimilation of Netatmo observations performs in different seasons. As a second step, the whole set of around 52,000 available Netatmo observations could be assimilated, in order to assess the impact of massive data on the model forecast. Here, except for application of the bias correction, a superobbing method would be necessary for dealing with the observation-error correlation issues that could arise, due to the large amount of data. Further experiments could combine the assimilation of pressure observations from the citizen network too, as well as considering the joint assimilation of synoptic and Netatmo data.

AUTHOR CONTRIBUTIONS

Christine Sgoff: formal analysis; investigation; methodology; validation; visualization; writing – original draft; writing – review and editing. **Walter Acevedo:** data curation; formal analysis; investigation; methodology; validation; visualization; writing – original draft; writing – review and editing. **Zoi Paschalidi:** data curation; visualization; writing – original draft; writing – review and editing. **Sven Ulbrich:** formal analysis; investigation; validation; writing – review and editing. **Elisabeth Bauernschubert:** investigation; methodology; software. **Thomas Kratzsch:** conceptualization; funding acquisition; project administration; supervision. **Roland Potthast:** conceptualization; funding acquisition; project administration; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

This research was carried out as part of the project FloWKar (Fleet-Weather-Maps), funded by the Federal Ministry of Transport and digital Infrastructure (BMVI), and the project Gridcast, funded by the Federal Ministry for Economic Affairs and Energy (BMWi). We gratefully acknowledge the support of Dr Felix Fundel in the visualization of the forecast evaluation. Synoptic observations, as well as wind-profiler data and radiosondes, are provided by https://www.dwd.de/EN/ourservices/opendata/opendata.html and are available on request. Open Access funding is enabled and organized by Projekt DEAL.

ORCID

Christine Sgoff https://orcid.org/0000-0003-3541-023X

REFERENCES

Anderson, J.L. and Anderson, S.L. (1999) A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127, 2741–2758.

Ballish, B.A. and Kumar, V.K. (2008) Systematic differences in aircraft and radiosonde temperatures: implications for NWP and climate studies. *Bulletin of the American Meteorological Society*, 89, 1689–1708.

Barker, H.W., Stephens, G.L., Partain, P.T., Bergman, J.W., Bonnel, B., Campana, K., Clothiaux, E.E., Clough, S., Cusack, S., Delamere, J., Edwards, J., Evans, K.F., Fouquart, Y., Freidenreich, S., Galin, V., Hou, Y., Kato, S., Li, J., Mlawer, E., Morcrette, J.-J., O'Hirok, W., Räisänen, P., Ramaswamy, V., Ritter, B., Rozanov, E., Schlesinger, M., Shibata, K., Sporyshev, P., Sun, Z., Wendisch, M., Wood, N. and Yang, F. (2003) Assessing 1D atmospheric solar radiative transfer models: interpretation and handling of unresolved clouds. *Journal of Climate*, 16, 2676–2699.

Bell, S. (2014). Quantifying uncertainty in citizen weather data. PhD thesis, Birmingham, Aston University.

Bell, S., Cornford, D. and Bastin, L. (2013) The state of automated amateur weather observations. *Weather*, 68, 36–41.

Bell, S., Cornford, D. and Bastin, L. (2015) How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70, 75–84.

Chapman, L., Bell, C. and Bell, S. (2017) Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, 37, 3597–3605.

Clark, M.R., Webb, J.D.C. and Kirk, P.J. (2018) Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations. *Meteorological Applications*, 25, 472–492.

Eyre, J.R. (2016) Observation bias-correction schemes in data assimilation systems: a theoretical study of some of their properties. *Quarterly Journal of the Royal Meteorological Society*, 142, 2284–2291.

Fenner, D., Meier, F., Bechtel, B., Otto, M. and Scherer, D. (2017) Intra and inter 'local climate zone' variability of air temperature as observed by crowdsourced citizen weather stations in Berlin, Germany. *Meteorologische Zeitschrift*, 26, 525–547.

Flanagin, A.J. and Metzger, M.J. (2008) The credibility of volunteered geographic information. *GeoJournal*, 72, 137–148.

Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C. and Boyd, D.S. (2013) Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. *Transactions in GIS*, 17, 847–860.

Gaspari, G. and Cohn, S.E. (1999) Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757.

Gasperoni, N.A., Wang, X., Brewster, K.A. and Carr, F.H. (2018) Assessing impacts of the high-frequency assimilation of surface observations for the forecast of convection initiation on 3 April 2014 within the Dallas–Fort Worth test bed. *Monthly Weather Review*, 146, 3845–3872.

Ghezzi, A., Gabelloni, D., Martini, A. and Natalicchio, A. (2017) Crowdsourcing: a review and suggestions for future research. *International Journal of Management Reviews*, 20, 343–363.

de Haan, S. (2011) High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar. *Journal of Geophysical Research: Atmospheres*, 116, D10111.

de Haan, S., de Jong, P. and van der Meulen, J. (2022) Characterizing and correcting the warm bias observed in Aircraft Meteorological Data Relay (AMDAR) temperature observations. *Atmospheric Measurement Techniques*, 15, 811–818.

Hintz, K.S., McNicholas, C., Randriamampianina, R., Williams, H.T., Macpherson, B., Mittermaier, M., Onvlee-Hooimeijer, J. and Szintai, B. (2021) Crowd-sourced observations for short-range numerical weather prediction: report from EWGLAM/SRNWP Meeting 2019. *Atmospheric Science Letters*, 22, e1031.

Hunt, B., Kostelich, E. and Szunyogh, I. (2007) Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230, 112–126.

Ingleby, B., Moore, D., Sloan, C. and Dunn, R. (2013) Evolution and accuracy of surface humidity reports. *Journal of Atmospheric and Oceanic Technology*, 30, 2025–2043.

James, E.P. and Benjamin, S.G. (2017) Observation system experiments with the hourly updating rapid refresh model using GSI hybrid ensemble-variational data assimilation. *Monthly Weather Review*, 145, 2897–2918.

Lussana, C., Seierstad, I.A., Nipen, T.N. and Cantarello, L. (2019) Spatial interpolation of two-metre temperature over Norway based on the combination of numerical weather prediction ensembles and in situ observations. *Quarterly Journal of the Royal Meteorological Society*, 145, 3626–3643.

Madaus, L.E., Hakim, G.J. and Mass, C.F. (2014) Utility of dense pressure observations for improving mesoscale analyses and forecasts. *Monthly Weather Review*, 142, 2398–2413.

Meier, F., Fenner, D., Grassmann, T., Jänicke, B., Otto, M. and Scherer, D. (2015) Challenges and benefits from crowdsourced atmospheric data for urban climate research using Berlin, Germany, as testbed. In:, *ICUC9 – 9th International Conference on Urban Climate jointly with 12th Symposium on the Urban Environment Challenges*. Toulouse, France. http://www.meteo.fr/icuc9/LongAbstracts/nomtm6-2-6171335_a.pdf.

Meier, F., Fenner, D., Grassmann, T., Otto, M. and Scherer, D. (2017) Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19, 170–191.

MetOffice (2021) *Weather Observations Website – WOW*. https://wow.metoffice.gov.uk/. accessed 22 February 2022.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J. and Clough, S.A. (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, 102, 16663–16682.

Moseley, S. (2011) From observations to forecasts – part 12: getting the most out of model data. *Weather*, 66, 272–276.

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A. and Leigh, R. (2015) Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35, 3185–3203.

Nakamura, G. and Potthast, R. (2015) *Programming of Numerical Algorithms and Useful Tools*. Bristol, UK: IOP Publishing, pp. 6–1 to 6–18.

Netatmo (2021) *Netatmo Website*. https://www.netatmo.com/product/weather/; accessed 22 February 2022.

Nipen, T.N., Seierstad, I.A., Lussana, C., Kristiansen, J. and Hov, O. (2020) Adopting citizen observations in operational weather prediction. *Bulletin of the American Meteorological Society*, 101, E43–E57.

Otkin, J.A., Potthast, R. and Lawless, A.S. (2018) Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Monthly Weather Review*, 146, 263–285.

Prill, F., Reinert, D., Rieger, D. and Zängl, G. (2020) Working with the ICON Model ICON Tutorial. Offenbach, Germany: Deutscher Wetterdienst.

Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A. and Potthast, R. (2016) Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quarterly Journal of the Royal Meteorological Society*, 142, 1453–1472.

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., Schill, C., McCallum, I., Kraxner, F. and Obersteiner, M. (2013) Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS One*, 8, 1–11.

Sobash, R.A. and Stensrud, D.J. (2015) Assimilating surface mesonet observations with the EnKF to improve ensemble forecasts of

convection initiation on 29 May 2012. *Monthly Weather Review*, 143, 3700–3725.

Steeneveld, G.J., Koopmans, S., Heusinkveld, B.G., Van Hove, L.W. and Holtslag, A.A. (2011) Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research Atmospheres*, 116, D20129.

Stephan, K., Klink, S. and Schraff, C. (2008) Assimilation of radar-derived rain rates into the convective-scale model COSMO-DE at DWD. *Quarterly Journal of the Royal Meteorological Society*, 134, 1315–1326.

Tostrams, L. (2017) *Spatio-temporal analysis and error correction of consumer weather station data*. PhD thesis, Nijmegen, Radbound University.

US National Weather Service (2021) *Citizen Weather Observer Program*. https://www.weather.gov/cle/CWOP. accessed 22 February 2022.

Uteuov, A., Kalyuzhnaya, A. and Boukhanovsky, A. (2019) The cities weather forecasting by crowdsourced atmospheric data. *Procedia Computer Science*, 156, 347–356.

de Vos, L.W., Leijnse, H., Overeem, A. and Uijlenhoet, R. (2017) The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, 21, 765–777.

de Vos, L.W., Leijnse, H., Overeem, A. and Uijlenhoet, R. (2019) Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. *Geophysical Research Letters*, 46, 8820–8829.

Williams, M., Cornford, D., Bastin, L., Jones, R. and Parker, S. (2011) Automatic processing, quality assurance and serving of real-time weather data. *Computers and Geosciences*, 37, 353–362.

Wolters, D. and Brandsma, T. (2012) Estimating the urban heat island in residential areas in the Netherlands using observations by weather amateurs. *Journal of Applied Meteorology and Climatology*, 51, 711–721.

Zahumenský, I. (2017) Guidelines on quality control procedures for data from automatic weather stations. *World Meteorological Organization*.

Zhang, F., Snyder, C. and Sun, J. (2004) Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Monthly Weather Review*, 132, 1238–1253.

Zinevich, A., Alpert, P. and Messer, H. (2008) Estimation of rainfall fields using commercial microwave communication networks of variable density. *Advances in Water Resources*, 31, 1470–1480.

Zängl, G., Reinert, D., Rípodas, P. and Baldauf, M. (2015) The ICON (ICOsahedral Nonhydrostatic) modelling framework of DWD and MPI-M: description of the nonhydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141, 563–579.

How to cite this article: Sgoff, C., Acevedo, W., Paschalidi, Z., Ulbrich, S., Bauernschubert, E., Kratzsch, T. *et al*. (2022) Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 148(745), 1752–1767. Available from: https://doi.org/10.1002/qj.4276