Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales

September 6, 2019

Ross Bannister, University of Reading

Hristo Chipilski, University of Oklahoma

Oscar Martinez-Alvarado,University of Reading

Abstract

While contemporary Numerical Weather Prediction models represent the large-scale structure of moist atmospheric processes reasonably well, they often struggle to maintain accurate forecasts of small-scale features such as convective rainfall. Even though high-resolution models resolve more of the flow, and are therefore arguably more accurate, moist convective flow becomes increasingly nonlinear and dynamically unstable. Importantly, the models' initial conditions are typically suboptimal, leaving scope to improve the accuracy of forecasts with improved data assimilation. To address issues regarding the use of atmospheric water-related observations – especially at convective scales (also known as storm scales) – this paper discusses the observation and assimilation of water-related quantities. Special emphasis is placed on background error statistics for variational and hybrid methods which need special attention for water variables.

The challenges of convective-scale data assimilation of atmospheric water information are discussed, which are more difficult to tackle than at larger scales. Some of the most important challenges include the greater degree of inhomogeneity and lower degree of smoothness of the flow, the high volume of water-related observations (e.g. from radar, microwave, and infrared instruments), the need to analyse a range of hydrometeors, the increasing importance of position errors in forecasts, the greater sophistication of forward models to allow use of indirect observations (e.g. cloud and precipitation affected observations), the need to account for the flow-dependent multivariate 'balance' between atmospheric water and both dynamical and mass fields, and the inherent non-Gaussian nature of atmospheric water variables.

---

1   Introduction

The importance of data assimilation (DA) has been appreciated for as long as numerical weather prediction (NWP) models have been used operationally [Daley, 1992], yet many obstacles related to the assimilation of atmospheric water (AW)-related observations remain. This is especially topical as there is a high (and growing) demand for observations of AW to assimilate with the latest high-resolution models, which are strongly affected by moist processes. It is well known that observations that are sensitive to AW present specific challenges to DA due to a number of reasons.

- *The complicated interpretation of the observations.* Observations are assimilated by predicting their values via observation operators and then adjusting control variables to fit them (control variables are an alternative representation of model variables suited to simplify aspects of the DA system). The observation operators may be only approximate. They may also depend on the values of variables in many locations, and on variables that are not explicitly available in the DA in the necessary detail, e.g. cloud and precipitation. Furthermore errors between different observations may be correlated.

- *The nonlinear, inhomogeneous nature of AW.* AW can influence the state evolution in a location-dependent and non-linear way, especially via changes of its phase. Moist processes are largely unresolved and therefore need to be parameterized in models. The difficulties of modelling AW impact the quality of forecasts, and lead to challenges in quantifying the uncertainties of the forecasts used as a-priori states in DA. Uncertainties are represented via probability density functions (PDFs), which in reality are multivariate, flow-dependent, and non-Gaussian. Most DA systems currently assume Gaussianity of a-priori errors.

- *The incompatibility of AW observations with the model and the DA system.* Inconsistencies between observations and simulated observations derived from the model state may have a detrimental effect on the analysis. Differences may be due to simple biases or to systematic misplacement of features in the model (horizontally and vertically – including in cloud and rain fields, and boundary layer heights). Observations measure the 'truth', yet a model's 'attractor' is likely to deviate from the truth, so the assimilation of even precise observations may lead to imbalances in the analysis.

- *The cost and capability of an NWP system.* Any extra expense needed to solve the AW-related DA problems outlined above needs to be justified and affordable.

This paper resulted from the NERC/Met Office-funded Flooding From Intense Rainfall programme, which was set up in 2013 to improve the UK's ability to predict flooding events [Clark et al., 2016, Dance et al., 2019]. The key objective of this programme was to improve rainfall predictions by investigating new and improved ways of using radar observations, and possible improvements to the error covariance matrices in the DA scheme that initialises the Met Office regional model. In view of this programme, this paper aims to review the AW DA issues discussed above. Most of this paper concerns how a-priori errors are treated in DA, and convective-scale issues, but other aspects such as the available observations of AW are also discussed.

The structure of this paper is as follows. In Sect. 2 we outline most of the operational-scale models that are in use, in Sect. 3 we discuss AW observations, in Sect. 4 we provide our view on the main research questions concerning AW-related DA, in Sect. 5 we discuss the main methods that are currently in use to introduce AW data to models, in Sect. 6 we show how problems may arise when this is not done adequately, in Sect. 7 we review options of possible AW control parameters, in Sect. 8 we pay attention to multiple phases of water, and in Sect. 9 to non-Gaussianity. In Sect. 10 we summarise the paper. For the convenience of the reader, Table 1 contains a list of the most used abbreviations.

2 High-resolution operational models

Most operational NWP centres employ limited area high-resolution NWP models (LAMs) [Clark et al., 2016, Gustafsson et al., 2018, Yano et al., 2018]. The Met Office's current operational convective-scale model over the United Kingdom (UKV, the United Kingdom Variable resolution version of the Unified Model (MetUM), Tang et al., 2013) has a core grid spacing of 1.5 km. There are two French-led high-resolution models in use – AROME (Application of Research to Operations at Mesoscale, Brousseau et al., 2016), which currently has a grid spacing of 1.3 km at Météo-France, and ALADIN (Aire Limitée Adaption Dynamique et développement InterNational, Horányi et al., 1996), which has many

AW	Atmospheric Water	MW	MicroWave
BL	Boundary Layer	NG	Non-Gaussian
CDF	Cumulative Distribution Function	NWP	Numerical Weather Prediction
(CV)T	(Control Variable) Transform	PDF	Probability Density Function
DA	Data Assimilation	PRH	Pseudo Relative Humidity
EnKF	Ensemble Kalman Filter	RH	Relative Humidity
GA	Gaussian Anamorphosis	SH	Specific Humidity
IR	InfraRed	TCWV	Total Column Water Vapour
LAM	Limited Area Model	(1/3/4D)Var	(1/3/4 Dimensional) Variational DA
LN	Log Normal	WV	Water Vapour

Table 1: A summary of the most used acronyms and abbreviations used in this paper.

configurations depending upon the agency using it. These systems are used and developed also by the HIRLAM (HIgh Resolution Limited-Area Model) and RC-LACE (Regional Co-operation for Limited Area modelling in Central Europe) consortia. The German implemented model, COSMO (Consortium for Small Scale Modelling, Baldauf et al., 2011, also used by a number of agencies in Europe) has a grid spacing of 2.8 km. The JMA's (Japan Meteorological Agency, Ito et al., 2016) highest resolution model LFM (Local Forecast Model) has a grid spacing of 2 km. Along with other models, NCEP (National Centre for Environmental Prediction) in the US uses WRF (Weather Research and Forecasting model, Schwartz and Liu, 2014). It has many configurations, e.g. with a standard grid spacing of 3 km (or higher, up to 750m in some domains), and is used as part of the HRRR (High Resolution Rapid Refresh) system [Benjamin et al., 2016]. MM5 (Fifth-Generation Penn State/NCAR Mesoscale Model, Pennsylvania State University and National Center for Atmospheric Research) is another US-developed high-resolution model designed for operational use. The CMA (China Meteorological Administration) has two high-resolution models – GRAPES (Global/Regional Assimilation and Prediction System) and AREM (Advanced Regional Eta co-ordinate Model) [Chen et al., 2013a]. GRAPES has recently been run with a grid spacing of 3km [Huang et al., 2019].

Most of the above models have grid lengths smaller than ~ 4 km, where it is believed that convective processes start to be resolved [Clark et al., 2016]. Finer grids ~ 100 m though are thought to be required to represent convection adequately [Bryan et al., 2003], which are not currently feasible for NWP. There are a number of models developed purely for research as they can be run on highly configurable grids, allowing more targeted research questions to be investigated. Some example models are the ARPS (Advanced Regional Prediction System, Xue et al., 2000, 2001), NCOMAS (National Severe Storms Laboratory Collaborative Model for Multiscale Atmospheric Simulation, Coniglio et al., 2006), and SCALE (Scalable Computing for Advanced Library and Environment, Nishizawa et al., 2015, Sato et al., 2015) systems.

We also mention the global ECMWF model here as it is used to provide lateral boundary conditions for many of the high-resolution LAMs mentioned above. Importantly the ECMWF model is the focus for many studies regarding innovative ways of assimilating cloud and precipitation information.

3 Sources of atmospheric water observations

Observations of AW are extremely important for NWP [Fabry, 2010, Saunders et al., 2015]. Mahfouf [2011] identifies three different kinds of observing system used in NWP, namely (a) conventional (in-situ) observations, (b) ground-based remotely-sensed observations, and (c) space-based remotely-sensed observations. All three potentially contain valuable information about AW, but the remotely-sensed observations require more interpretation by the DA system, which can cause problems if this is not done correctly (see Sect. 6). This is particularly difficult when observing cloudy and precipitating regions, where important diabatic processes are happening, making forecasts particularly sensitive

Instrument	Type	Platform	Measurement	Example Ref(s)	Clr	Cld	Prcp
Hygrometer	E.g. psychrometer	Surface stations	RH, dew-point temperature	Ingleby et al. [2013a]	✓	✓	✓
Hygrometer	E.g. thin-film capacitor	Radiosondes	RH, dew-point temperature	Wang et al. [2003]	✓	✓	✓
Gas sensor	Laser diode IR absorption	Commercial aircraft (e.g. AMDAR, TAMDAR)	WV mixing ratio	Hoff [2010]	✓	✓	✓
Rain gauge	E.g. tipping bucket	Surface stations	Accumulated precipitation	Burt [2012]	✓	✓	✓

Table 2: Sources of in-situ observations of AW. Acronyms not defined in Table 1are: **AMDAR** (Aircraft Meteorological DAta Relay), **TAMDAR** (Tropospheric Airborne Meteorological DAta Reporting). The last three columns indicate whether the measurements are routinely assimilated in clear, cloudy, or precipitating conditions.

to initial conditions [Ohring and Bauer, 2011, Sun and Zhang, 2016], and making observations from space difficult to use [McNally, 2002, Bauer et al., 2011a]. In-situ observations are 'point' observations, meaning that they sample sub-model-grid fields (especially true for AW fields which have short length-scales). For this reason, a representativity error contribution is added to the observation error statistics when assimilating observations that sample smaller volumes of the atmosphere than the model grid boxes (Sect. 6.5). Remotely sensed observations, on the other hand, can sample volumes larger than the model grid boxes of convective-scale models. This is the opposite problem for which observation operators should take input from a neighbourhood of model grid points (e.g. Duffourg et al. 2010). Here we look at the different observing systems available. It is outside the scope of this article to have a comprehensive list of observing systems and readers are referred to the WMO OSCAR (Observing Systems Capability Analysis and Review tool) – see references under WMO for more information. Specialist acronyms used in this section are defined in the captions Tables 1-3, and in footnotes 5 and 7.

3.1 In-situ observations

Most in-situ observations of AW measure water vapour (WV), which are in the form of relative humidity (RH) or dew-point temperature (Table 2). Ground-based observations remain important in NWP [Ingleby, 2015, Clark et al., 2016]. Humidity observations are made from some commercial aircraft via, e.g., the AMDAR system [Ingleby et al., 2013a, Petersen et al., 2016], which has recently extended the humidity observation network. Of all the data sources, in-situ measurements are thought to be the most accurate as their interpretation is relatively simple, they require the least pre-processing, and are reasonably good for resolving sharp features like inversions and cloud layers. They are not without issues though – radiosondes, e.g., may have representativeness errors that exceed the instrument error, and can report spurious saturation after passing through clouds [Kitchen, 1989, Miloshevich et al., 2001].

3.2 Ground-based remotely-sensed observations

The main ground-based remotely sensed observations are summarized in Table 3. The instruments that are used operationally by at least some centres are mentioned first, followed by research instruments with possible later operational use.
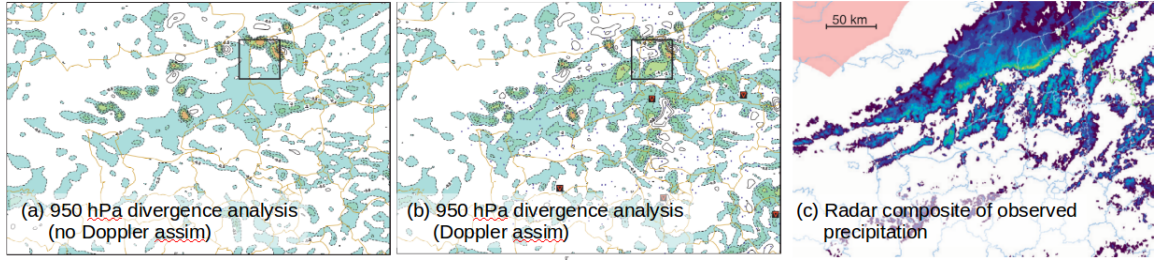
Figure 1: Low-level convergence analyses for 1800 UTC 8th November 2007 from DA experiments (a) without assimilation of Doppler radial wind and (b) with assimilation of Doppler radial wind. Dashed contours indicate horizontal convergence. The system is AROME/3DVar for the domain over France. (c) is the observed precipitation rate. ©American Meteorological Society. Used with permission. Taken from Montmerle and Faccani [2009].

3.2.1 Ground-based instruments used for operational AW analysis

Ground-based radar instruments measure the return of a radar signal scattered from condensed water particles (hydrometeors) in one or two polarizations, as a function of elevation, azimuth, and range (see e.g. Michaelides et al., 2009). The reflectivity depends on the radar wavelength, the type of hydrometeor (e.g. rain water, graupel, snow, etc.), and on the sixth power of each hydrometeor's diameter. A given volume of air will have a distribution of hydrometeor types and sizes. Most operational weather radar units operate in the S-band (10 cm wavelength) or C-band (5.6 cm), although shorter wavelength units are also used: X-band (3 cm), Ka-band (1 cm), and W-band (0.3 cm). The shorter wavelength instruments become sensitive to smaller particles like those in drizzle or fog, but are more strongly attenuated by precipitation making them less useful in heavy precipitation. An airborne radar can be especially useful for W-band radar, which are significantly attenuated by heavy precipitation [Borderies et al., 2018]. Polarimetric radars are capable of emitting radar pulses with two different polarisation directions (normally horizontal and vertical polarisations), and can distinguish the returns of each. Such measurements can provide additional information on the shape and orientation of hydrometeors, which can in turn help interpret the characteristics of the scatterers for the purposes of improved rainrate estimation (e.g. by allowing ground clutter identification, and use of an oblateness-size relationship of raindrops to help infer rainrate, see e.g. Illingworth 2004). Whether derived from single polarization, or polarimetric instruments, the wide field estimates of precipitation rate from radar make it an extremely important source of data for convective-scale initialisation (e.g. Sun, 2005b). Typical radar scans sample the atmosphere up to 250 km in the horizontal and 15 km in the vertical every 5 to 10 minutes.

Many radar instruments are of the Doppler type and observations of the Doppler shift due to hydrometeors moving radially with respect to the radar beam can measure radial winds [Lindskog et al., 2004, Doviak and Zrnić, 2006, Montmerle and Faccani, 2009]. Many Doppler radars working together can infer all wind components within precipitating air. These measurements yield indirect information about possible convection or precipitation by identifying local convergence zones. This is demonstrated in Fig. 1, which shows low-level convergence analyses (a) without and (b) with assimilation of Doppler radial winds during the passage of a frontal rain band as observed in (c). The assimilation of Doppler winds improves the convergence line, including that inside the square, which coincides with the peak in precipitation. The improved convergence structure in (b) is shown to improve the 1- and 2-hour precipitation forecasts (for details see Montmerle and Faccani, 2009). Even though measurements of wind, and not water, are made, the impact on water fields like precipitation means that we include Doppler wind in Table 3.

Measurements of the zenith total delay of navigation signals emitted by a GNSS satellite received at a ground station contain information on the refractivity of the air between the ground station and the

satellite. The refractivity depends on the profiles of pressure, temperature and partial vapour pressure, so there is information on these quantities in the delay measurements. Providing that the pressure and temperature profiles are accurately known, total column water vapour (TCWV) estimates can be made from GNSS. GNSS measurements have the benefit of being usable in all weather conditions [Bevis et al., 1992]. Even though GNSS observations rely on satellites, we refer to it as ground-based as the ground segment is where the measurements are actually taken.

Lightning strikes are associated with severe storms and can be located using a network of ground-based sensors that measure the direction of the strike relative to a particular receiver, or the time of arrival of a packet of very low frequency radio waves (3-30 kHz, called sferics) emitted by the strike. Networks of receivers include the World Wide Lightning Location Network (with around 70 stations and with lightning location errors <10 km) and the US National Lightning Detection Network (with around 100 stations and with location errors <1 km). References are provided in Table 3. Knowledge of lightning strike rates can provide valuable information on convective activity especially when other measurements, like radar, are unavailable – e.g. over oceans or mountainous regions. The challenge is to interface the measurements with NWP, which is often done via a post-DA nudging procedure (see Sect. 5.6).

3.2.2 Ground-based instruments used for research with possible future operational use

Ground-based radiometers (e.g. Blumberg et al., 2015) usually measure the microwave (MW) radiation passively emitted by the atmosphere, but infrared (IR) radiation is also measured by some instruments, e.g. the AERI (Atmospheric Emitted Radiance Interferometer, Turner and Löhnert, 2014). Amongst other quantities, these measurements contain information on the vertical moisture profile. The vertical resolution possible for moisture with a MW instrument varies, e.g. from hundreds of metres in the lower atmosphere, to many km in the upper troposphere [Liljegren et al., 2005, Caumont et al., 2016]. Blumberg et al. [2015] additionally find a better vertical resolution for their IR radiometer in the lowest 3 km of the atmosphere than for their MW radiometer. Ground-based radiometers provide the capability to give high temporal frequency, but the network of radiometers is sparse, their data are currently not widely assimilated, and the signal can be attenuated by strong precipitation (e.g. Norenberg et al., 2008).

In addition to measuring radial wind, Doppler radar instruments can measure the temporal change in the phase between the radar signal returned from pairs of closely aligned stationary ground clutter objects over a period of 20 minutes to an hour [Fabry et al., 1997, Nicol et al., 2011, 2013]. This provides information about temporal changes to near surface specific humidity (SH) and temperature. The principle is as follows. The refractivity of the air between the objects, N, depends on its pressure, temperature and partial vapour pressure [Bean and Dutton, 1968], and so ΔN depends on changes to these quantities over time. Analysis of the temporal change in the radar phase difference between these objects allows ΔN to be estimated, which a DA system can then attribute to changes in the temperature and humidity fields. This DA capability is still under development, and necessarily requires a 4D DA method since it involves temporal changes in N. Note that measurement of the absolute N – rather than ΔN – is not available with this technique as it would require knowledge of the total number of electromagnetic wave cycles of the radar beam between the two targets.

Ceilometers measure cloud base height by timing pulses of laser light to travel to and from the base of overlying cloud (e.g. Martucci et al., 2010). This is essentially the optical equivalent of the principle of radar reflectivity. Ceilometers can also measure aspects of the returned signal to give information about cloud amount. As with other ground-based systems like GNSS and surface radiometers, ceilometers can give only as much spatial resolution as the density of the network (the Met Office for instance maintains a network of over a hundred of such instruments around the UK, and is considering the assimilation of their data [Clark et al., 2016]).

Lidar instruments measure the backscatter of an upward-fired laser pulse, as a function of time (e.g. Turner et al., 2002). For example, light of 355 nm wavelength can Raman scatter to 408 nm by WV, allowing the mixing ratio profile of vapour to be measured. Unlike radar, lidar beams cannot

penetrate cloud, so are traditionally used to retrieve WV concentrations only. The vertical resolution depends on the sampling rate of the receiver, which in practice can yield profiles with tens of metres resolution. The DIAL (DIfferential Absorption Lidar) technique fires a second laser pulse at a slightly different wavelength to the WV line. The difference between the two returns contains information solely on scattering by WV and not on other processes. Lidar data have been successfully assimilated in research work (e.g. Grzeschik et al., 2008).

3.3 Space-based remotely-sensed observations

As with ground-based instruments, there are a number of remote sensing principles that are used to detect and measure AW quantities from space. These include instruments in low Earth orbits (LEOs), or in geostationary orbits. Most LEOs are sun-synchronous polar orbits, 700-850 km above the Earth's surface, where each satellite can scan a particular area typically up to twice per day. Low-inclination orbits are also used to observe the tropics (e.g. SAPHIR below). Geostationary orbits are always placed above the equator, 35786 km above the Earth's surface, where satellites can continuously scan an entire face of the Earth, although at a reduced spatial resolution than from LEOs, especially at high latitudes. Parallax effects also complicate the determination of the geographical position of clouds at high latitudes[4]. Instruments can be further categorised into passive (e.g. IR and MW imagers and sounders), and active (e.g. radar and lidar) types. As only a subset of instruments are currently used for DA, this section is divided into brief descriptions of instruments used for operational assimilation of AW information, and those used for research (and some future instruments).

3.3.1 Space-based instruments used for operational AW analysis

Table 4 lists the main instrument types used for AW components in NWP (this list is not exhaustive)[5]. Measurements of radiances in the IR and MW make up the bulk of sounding measurements from space. We note here that traditionally, the assimilation of satellite radiances has had less impact in LAMs than in global models [Hess, 2007]. Factors affecting this include the generally greater proportion of land points in LAM domains (some satellite radiances – e.g. MW imagers – are not assimilated over land at present due to the uncertainty in surface emissivity in certain channels, Bauer et al., 2010), the greater impact of position errors in features like clouds (Sect. 6.1), the 'hit-and-miss' availability of LEO observations over LAMs, and the often lower model lid in LAMs, making radiative transfer

---

[1]'No rain' observations can be assimilated.

[2]It is possible to use radar reflectivity observations from a cloud radar to retrieve liquid water content in non-precipitating cloud, e.g. Fielding et al. [2014].

[3]It is estimated that light precipitation has minimal effect of the radar refractivity signal [Fabry, 2004].

[4]A back-of-the-envelope calculation estimates that a geostationary satellite will see the apparent position of a cloud at 5 km height and at 45° latitude $\sim$ 7 km away horizontally from its correct geographical position. At 60° this increases to $\sim$ 13 km, and at 75° this is greater than 60 km. These high-latitude values are significant compared to km-scale grid boxes in covective-scale models.

[5]Acronyms in Table 4 not defined before are: **ABI** (Advanced Baseline Imager), **AHI** (Advanced Himawari Imager), **AIRS** (Atmospheric InfraRed Sounder), **AMSR** (Advanced Microwave Scanning Radiometer), **AMSU** (Advanced Microwave Sounding Unit), **ATMS** (Advanced Technology Microwave Sounder), **AVHRR** (Advanced Very High Resolution Radiometer), **CrIS** (Cross-track Infrared Sounder), **DMSP** (Defense Meteorological Satellite Programme), **DPR** (Dual frequency Precipitation Radar), **DWSS** (Defense Weather Satellite System), **FY** (Feng-Yun), **GCOM** (Global Change Observation Mission), **GLM** (Geostationary Lightning Mapper), **GMI** (GPM Microwave Imager), **GNOS** (GNSS radio Occultation Sounder), **GNSS(-RO)** (Global Navigation Satellite System(-Radio Occultation)), **GOES** (Geostationary Operational Environmental Satellite), **GPM** (Global Precipitation Measurement), **GRAS** (GNSS Receiver for Atmospheric Sounding), **(H)IRAS**} ((Hyperspectral) InfraRed Atmospheric Sounder), **HIRS** (High resolution InfraRed Sounder), **IASI(-NG)** (Infrared Atmospheric Sounding Interferometer(-New Generation)), **JPSS** (Joint Polar Satellite System), **LIS** (Lightning Imaging Sensor), **MetOp** (Meteorological Operational satellite) **MHS** (Microwave Humidity Sounder), **MISR** (Multi-angle Imaging Spectro-Radiometer), **MODIS** (MODerate resolution Imaging Spectroradiometer), **MSG** (Meteosat Second Generation), **MT** (Megha Tropiques), **MWHS** (Microwave Humidity Sounder), **NOAA** (National Oceanic and Atmospheric Administration), **SAPHIR** (Sounder for Probing Vertical Profiles of Humidity), **SEVIRI** (Spinning Enhanced Visible and Infrared Imager), **S-NPP** (Suomi-National Polar orbiting Platform), **SSM/I** (Special Sensor Microwave Imager), **SSMIS** (Special Sensor Microwave Imager/Sounder), **TCCI/W** (Total Column Cloud Ice/Water), **TRMM** (Tropical Rainfall Measuring Mission), **VIIRS** (VIsible Infrared Imaging Radiometer Suite).

7

Instrument	Measurement	Sensitive to	Example Ref(s)	Clr	Cld	Prcp
Operational instruments						
Radar	Reflectivity	Precipitation	Austin [1987], Sun and Crook [1997], Wattrelot et al. [2014], Wang et al. [2013]	✓[1]	✓[2]	✓
Radar	Doppler winds	Convergence	Tuttle and Foote [1990], Montmerle and Faccani [2009]	✗	✗	✓
GNSS	GPS zenith total delay	Total column water vapour	Bennitt and Jupp [2012]	✓	✓	✓
WWLLN, NLDN	Time of arrival / magnetic direction finding	Lightning	Dowden et al. [2002], Cummins et al. [1998]	✗	✗	✓
Research instruments						
Ground-based radiometers	IR rad.	Tropospheric temperature, and water vapour	Blumberg et al. [2015]	✓	✓	✗
Ground-based radiometers	MW rad.	Tropospheric temperature, and water vapour	Cimini et al. [2006], Caumont et al. [2016]	✓	✓	✓
Radar	Refractivity	Water vapour path changes at ground level	Fabry et al. [1997], Nicol et al. [2014]	✓	✓	✓[3]
Ceilometer	Backscatter (signal and time of return)	Cloud amount and height at cloud base	Martucci et al. [2010]	✗	✓	✓
Lidar	Backscatter (signal and time of return)	Water vapour	Turner et al. [2002]	✓	✓	✗

Table 3: Ground-based remote-sensing instruments with sensitivity to AW. Acronyms not defined in Table 1 are: **GNSS** (Global Navigation Satellite System), **GPS** (Global Positioning System), **WWLLN** (World Wide Lightning Location Network), **NLDN** ([US] National Lightning Detection Network). The last three columns indicate whether the measurements are routinely assimilated in clear, cloudy, or precipitating conditions.

modelling more erroneous. There have though been some important recent breakthroughs in the direct assimilation of radiances at convective-scales, which are discussed in Sect. 6.8.

IR instruments are divided into sounders such as AIRS, CrIS, IASI, IRAS, and HIRAS (providing information about WV in various vertical bands according to the Jacobian of each channel measured) and imagers such as those carried on geostationary platforms (for AW used mainly for cloud top height information). IR measurements are generally difficult to interpret for cloudy scenes (e.g. Pavelin et al., 2008), although this situation is improving rapidly [Zhang et al., 2016, Minamide and Zhang, 2018, Okamoto et al., 2019]. MW instruments are divided similarly: sounders such as AMSU-A/B, ATMS, MHS, SAPHIR, and MWHS-2 (capable of measuring tropospheric WV), imagers such as AMSR-2 and GMI (using MW window channels providing information on total column water vapour, and cloud and precipitation quantities), and combined imager/sounders such as the SSMIS instruments. Unlike for IR, it is possible to routinely assimilate data from MW observations in cloudy and precipitating scenes [Geer et al., 2017, Chambon and Geer, 2017, Geer et al., 2018], but accurate particle size distributions need to be specified, and multiple scattering effects need to be accounted for in the radiative transfer models [Bauer et al., 2006b, Geer and Baordo, 2014]. There are also issues with the assimilation of data from MW instruments that can see the lowest part of the atmosphere over land and sea ice. Here there can be uncertainties in the variations of the surface emissivity (e.g. Prigent et al., 2006). MW imagers are particularly prone to these problems and such data are often assimilated only over the ocean (e.g. Geer et al., 2017), although this is less of a problem in the tropics where the usually large water vapour path means that the surface is obscured [Chambon and Geer, 2017]. Water – especially cloud – is also observed with the LEO visible and IR imagers AVHRR/3, MISR, MODIS, and VIIRS, although data from these instruments is not normally assimilated, except for the purpose of tracking cloud features to estimate atmospheric wind vectors (e.g. Forsythe, 2007).

In addition to the GNSS method described in Sect. 3.2, a further use of the GPS network of satellites is made with the GNSS radio occultation (GNSS-RO) technique. GNSS-RO detects GPS signals with orbiting receivers, such as GRAS and GNOS. The technique provides information along the path of the signal which traverses the limb of the Earth's atmosphere (i.e. from the GPS satellite to the receiver, Healy and Thépaut [2006]). As the signal enters the atmosphere it changes speed and direction due to changes in the refractive index of the air. Water vapour has a negligible effect on the refractive index in the stratosphere and upper troposphere, but does affect the signal in the lower troposphere. The method is claimed to be virtually unaffected by cloud and precipitation, but it does have a low horizontal resolution [Eyre, 2007].

Information on the location and timing of lightning strikes is also available via satellite. The now defunct TRMM LEO satellite carried the LIS instrument to measure lightning strikes by making observations in the near IR (777.4 nm). A similar instrument, GLM, is carried on board the geostationary GOES-16 and 17 satellites (with plans for GOES-T and GOES-U). The data complements the existing ground-based lightning detection systems (Sect. 3.2) and can in principle be used for NWP in a similar way (Sect. 5.6).

3.3.2 Space-based instruments used for research with possible future operational use

Table 5 gives a selection of existing research instruments and some future instruments (again this list is not exhaustive)[7]. Many of the future instruments shown are planned improvements of existing ones

---

[6]Note though that there has been significant progress recently with all-sky assimilation of IR radiances, e.g. Geer et al. [2018], Minamide and Zhang [2018].

[7]Acronyms in Table 5 not defined before are: **ATLID** (ATmospheric LIDar), **CALIOP** (Cloud-Aerosol Lidar with Orthogonal Polarization), **COSMIC** (Constellation Observing System for Meteorology, Ionosphere and Climate), **CPR** (Cloud Profiling Radar), **CALIPSO** (Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations), **Earth-CARE** (Earth Clouds, Aerosol and Radiation Explorer), **FCI** (Flexible Combined Imager), **HYMS** (HYperspectral Microwave Sensor), **IASI-NG** (IASI-New Generation), **ICI** (Ice Cloud Imager), **ICESAT** (Ice, Cloud, and Land Evaluation Satellite) **IRS** (InfraRed Sounder), **LI** (Lighning Imager), **MetOp-SG** (MetOp-Second Generation), **MTG-I** (Meteosat Third Generation-Imager), **MTG-S** (Meteosat Third Generation-Sounder), **MWI** (MicroWave Imager), **MWS** (MicroWave Sounder), **RO** (Radio Occultation), **TGRS** (Triple G GNSS radio occultation Receiver System).

Instrument	Examples	Measurement	Sensitive to	Example Ref(s)	Clr	Cld	Prcp
Operational instruments							
LEO IR sounder	AIRS (Aqua), CrIS (JPSS, S-NPP), IASI (MetOp), HIRS/4 (MetOp), [H]IRAS (FY-3)	IR rad.	WV	Parkinson [2003], Guidard et al. [2011], Goldberg et al. [2013], Klaes et al. [2007]	✓	✗[6]	✗[6]
Geostationary Vis+IR imager	ABI (GOES), AHI (Himawari), SEVIRI (MSG),	Vis./IR rad.	WV, Clouds	Montmerle et al. [2007], Jin et al. [2008]	✓	✗[6]	✗[6]
LEO MW sounder	AMSU-A/B (NOAA), ATMS (JPSS, S-NPP), MHS (NOAA, MetOp), SAPHIR (MT), MWHS-2 (FY-3)	MW rad.	WV, TCCI, TCCW, TCWV, clouds, surface precip.	Klaes et al. [2007], Goldberg et al. [2013], Aguttes et al. [2000]	✓	✓	✓
LEO MW imager	AMSR-2 (GCOM), GMI (GPM-core), SSM/I (DMSP)	MW rad. (window channels)	TCWV, TCCW, TCCI, clouds, surface precip.	Draper et al. [2015], Boukabara and Garrett [2018]	✓	✓	✓
LEO MW imager/sounder	SSMIS (DMSP)	MW rad.	WV, TCCI, TCCW, TCWV, clouds, surface precip.	Bell et al. [2008], Baordo and Geer [2016]	✓	✓	✓
LEO Vis+IR imager	AVHRR/3 (MetOp), MISR (Terra), MODIS (Aqua/Terra), VIIRS (DWSS, JPSS, S-NPP)	Vis./IR rad.	Cloud attributes	Cracknell [1997], Diner et al. [1998], Salomonson et al. [2002], Lee et al. [2006]	✗	✓	✓
GNSS-RO	GRAS (MetOp), GNOS (FY-3)	Refractivity, bending angle	WV	Kursinski et al. [1997], Luntama et al. [2008], Johny and Prasad [2014]	✓	✓	✓
Radar	DPR (GPM-core)	Reflectivity	Cloud ice, cloud water, TCCI	Kojima et al. [2012]	✓	✓	✓
Lightning imager	GLM (GOES)	near IR	Lightning	Goodman et al. [2013]	✗	✗	✓

Table 4: A selection of operational space-based remote-sensing instrument types with sensitivity to AW (host platforms are given in brackets, and acronyms are defined in footnote 5). The last three columns indicate whether the measurements are routinely assimilated in clear, cloudy, or precipitating conditions.

to be carried on the next generation of platforms. This includes IASI-NG (to supersede IASI), FCI (to supersede SEVIRI), MWS (to supersede AMSU-A and MHS), and RO (to supersede GRAS). There are also new classes of instruments planned. IRS, for instance, is an IR sounder to be carried on a geostationary satellite (MeteoSat Third Generation, itself divided into separate imager and sounder platforms), and ICI is an instrument measuring the smallest MW and sub-mm wavelengths, to be carried on MetOp Second Generation (itself divided into separate MW and optical platforms). A future mm/sub-mm wavelength instrument proposed is HYMS. Unlike existing instruments it is hyperspectral (276 channels are suggested, Birman et al., 2017) allowing for more information on hydrometeors to be collected than from existing instruments (compare with the 26 channels of MWI to be carried on MetOp-SG-B).

Although not widely assimilated, satellite-borne radar and lidar measurements have been available for a number of years – e.g. the first space-based weather radar was carried on TRMM in 1997, [Raghavan, 2013], and the first space-based weather lidar was carried on ICESat in 2003, [Sun et al., 2013]. Radar and lidar instruments each emit a radiation beam into the atmosphere, part of which is returned to the satellite after being modified by the hydrometeor composition along the beam's path. The physical principles used to construct the observation operators for radar and lidar are similar [Di Michele et al., 2014a,b]. Multiple scattering from hydrometeors affects lidar and the higher frequency radar beams (e.g. CloudSat's 94 GHz radar) more than the lower frequency radar beams (e.g. C- or S-band ground radars mentioned in Sect. 3.2), but sometimes multiple scattering of higher frequency radars can be neglected in observation operators, except for scenes affected by heavy precipitation [Janisková, 2015]. The instruments DPR and CPR (on GPM-core and CloudSat/EarthCARE respectively) provide sources of radar observations, and CALIOP and ATLID (on CALIPSO and EarthCARE respectively) provide sources of lidar observations. DPR data are actually currently being assimilated by JMA (and so is included in Table 4). Radar and lidar measurements complement each other, and making both types of measurement from space simultaneously may be beneficial, which may in the future be exploited in convective-scale DA. The forthcoming EarthCARE satellite, which is due to be launched (at the time of writing) in 2021, hosts both radar and lidar instruments. The simultaneous assimilation of these observations for NWP is currently experimental (e.g. Janisková, 2015), and this activity is particularly vulnerable to cloud position errors in the background state (see Sect. 6.1). Additionally the low observation frequency over a particular location means a single satellite will be of limited benefit with LAMs.

The lightning imaging capability of the Meteosat series of geostationary satellites is planned for the forthcoming MTG-I series, which will carry the lightning imaging instrument, LI. This is a similar instrument to the GLM on board the latest GOES satellites.

4 Questions about the assimilation of atmospheric water observations

As new AW-affected observations become available for DA, and as models improve at resolving convective processes, the accompanying DA problem will need to change. The following is a list of issues, posed in the form of questions, that may arise when developing DA systems. All issues are AW related, but not exclusively.

- *Which method?* The ways of solving the DA problem have expanded over recent decades. Methods include Variational (Var, in 3D or 4D flavours), ensemble (Ens), and hybrid methods (Sects. 5.1-5.4). Many AW observations are not currently assimilated directly, so are dealt with outside of the main analysis (Sects. 5.5 and 5.6).

- *What assimilation cycling frequency?* In synoptic-scale systems the memory of AW perturbations in the model tends to be shorter than for other variables [Bengtsson and Hodges, 2005], and is expected to be very short in convection-permitting models [Sun, 2005a, Hohenegger and Schär,

Instrument	Examples	Measurement	Sensitive to	Example Ref(s)	Clr	Cld	Prcp
Research and future instruments							
LEO IR sounder	IASI-NG (MetOp-SG-A)	IR rad.	WV	Crevoisier et al. [2014]	✓	✓	✓
Geostationary Vis+IR imager	FCI (MTG-I)	IR rad.	Clouds	Rodriguez et al. [2009]	✓	✓	✓
Geostationary Vis+IR sounder	IRS (MTG-S)	IR rad.	WV, clouds	Rodriguez et al. [2009]	✓	✓	✓
LEO MW sounder	MWS (MetOp-SG-A)	MW rad.	WV, TCCI, TCCW, TCWV, clouds, surface precip.	D'Addio et al. [2014]	✓	✓	✓
LEO MW imager	MWI (MetOp-SG-B)	MW rad. (window channels)	TCWV, TCCW, TCCI, clouds, surface precip.	D'Addio et al. [2014]	✓	✓	✓
LEO mm/sub-mm imager	ICI (MetOp-SG-B), HYMS	mm/sub-mm rad.	WV, cloud ice, TCCI, precip.	Thomas et al. [2014], Birman et al. [2017]	✓	✓	✓
Radar	CPR (CloudSat, EarthCARE)	Reflectivity	Cloud ice, cloud water, TCCI	Stephens et al. [2018]	✓	✓	✓
Lidar	CALIOP (CALIPSO), ATLID (EarthCARE)	Backscatter (signal and time of return)	Cloud ice, cloud water, TCCI, TCCW	Winker et al. [2007], Stephens et al. [2018], Illingworth et al. [2015]	✗	✓	✓
GNSS-RO	RO (MetOp-SG-A/B), TGRS (COSMIC-2a/b)	Refractivity, bending angle	WV	Cook et al. [2013], Roselló et al. [2012]	✓	✓	✓
Lightning imager	LI (MTG-I)	near IR	Lightning	Rodriguez et al. [2009]	✗	✗	✓

Table 5: A selection of research and future space-based remote-sensing instruments with sensitivity to AW (acronyms are defined in footnote 7). The last three columns indicate whether the measurements are (or have the possibility of being) routinely assimilated in clear, cloudy, or precipitating conditions.

2007]. This suggests that AW information should be introduced more frequently than in large-scale systems. Rapid update cycling, where the forecast is updated frequently – e.g. a few tens of minutes [Sun et al., 2012] or one minute [Dowell et al., 2011] – can be useful for nowcasting applications. The cycle length has particular relevance for 4DVar: a short time window (e.g. sub one-hour) may allow nonlinear processes to be linearised [Fabry and Sun, 2010, Fabry, 2010], but it may not allow for imbalances introduced by the DA to stabilise without special techniques [Tong et al., 2016]. A longer time window (e.g. six hours or more) will exacerbate the nonlinearity problem mentioned above, and may even render convective-scale 4DVar invalid in such circumstances.

- *Which observations to assimilate?* In convective-scale initialisation, large numbers of observations are needed [Fischer et al., 2005], but the question of which observations to use depends on the observations' availability and timeliness (latency), their impact, and their ability to be modelled in the assimilation [Sun, 2005b, Ge et al., 2013]. Impact can be studied using, e.g., observation simulation experiments [Storto and Randriamampianina, 2010, Mahfouf et al., 2015]. The ability to model the observations requires careful consideration as the observation operator may need inputs that are unavailable, or are inaccurately diagnosed in a particular system. An example is the assimilation of cloud or precipitation affected satellite radiances. Failing to account for the cloud and precipitation contributions to the radiance calculation correctly in the observation operator will cause this error to project onto other input variables such as WV and temperature (Sects. 6.8 and 8).

- *What analysis/control parameters to use?* Control parameters are the physical variables that a DA problem is posed in (Sect. 7). In Ens systems they are usually the same as (or a subset of) the model variables, but in Var systems they are different (an exception is the study of Vukicevic et al. [2006], which used the model variables as control parameters). Many DA systems have just one AW parameter (representing WV). There is a choice of control parameter to use in Var for WV, each one having its own pros and cons (Sect. 7). There is growing need for additional AW parameters to represent cloud liquid water, ice, rain, snow, hail and graupel. Introducing these is much easier in Ens systems than in Var. Alternatively, instead of using additional parameters, cloud and precipitation may be diagnosed from the conventional variables (Sect. 8).

- *What are the error covariances between background variables?* DA systems need to have good information about the background error covariances, which affects the analysis significantly (e.g. Bannister 2008a, Houtekamer and Zhang 2016). Covariances for AW variables are less well-known than for other variables as they are particularly prone to inhomogeneity, patchiness, flow-dependency, and non-Gaussianity (e.g. Legrand et al., 2016). In Ens schemes, covariances are inferred from the ensemble and the localisation method (Sect. 5.2), but in Var they are modelled and so require special attention (Sects. 5.1 and 7). In hybrid schemes they are a combination of both (Sect. 5.4).

- *Account for correlated observation errors?* Studies have shown that errors between observations can be correlated [Weston et al., 2014, Bormann et al., 2016], often via their representativity errors [Janjic and Cohn, 2006]. This includes spatial correlations in dense observation networks, but also between channels of satellite radiance measurements. More information can be extracted from observations if these correlations are taken into account (Sect. 6.5).

- *Accompany the main DA with simpler methods for AW estimation?* Many centres perform a pre-analysis retrieval of WV, cloud, or precipitation quantities from particular observation types for assimilation as pseudo observations of humidity (Sect. 5.5). This shifts the need to rectify inadequacies from the main DA system (such as lack of adequate representation of cloud) to the retrieval system (e.g. 1DVar), which can be a much simpler platform to develop cloud/precipitation-aware solutions. There are also post-analysis procedures used such as cloud analysis methods and latent heat nudging (Sect. 5.6).

- *Account for model errors/biases?* Model errors and biases are an inevitable part of NWP, and there is a contribution from AW variables. To help alleviate this problem, operational NWP centers typically incorporate a model error scheme with the DA. For Var this involves an additional term in the cost function, e.g. Trémolet [2006], and for Ens schemes this involves modification of the ensemble by multiplicative or additive model errors, e.g. Hamill and Whitaker [2005]. Biases (i.e. the mean difference between model predictions and observations) require bias correction schemes (Sect. 6.7).

- *Reduce initialisation imbalance?* Firstly, in this paper the term 'imbalance' is used in a general sense, and does not just refer to the traditional hydrostatic and geostrophic imbalances. For instance, changes to the temperature, divergence or AW by the DA may leave the model fields overly unbalanced. This can manifest itself as spurious precipitation early in the forecast, which is known as 'spin-down' (as opposed to 'spin-up' which is the delay in initiation of precipitation, especially when then model is started from lower resolution initial conditions, Sun et al. [2014], Clark et al. [2016]). An example of spin-down from the literature is given in Fig. 2, which shows the total precipitation as a function of forecast time following various DA schemes. Many DA schemes show spurious precipitation in the first couple of hours, which is thought to be largely due to inconsistencies between the DA and the particular forecast model used [Schwartz and Liu, 2014]. Ideally the degree of imbalance introduced by the DA would be controlled by the background error covariances, but some centres additionally control residual imbalance by introducing analysis increments bit-by-bit over a time window with a method called Incremental Analysis Update (IAU, Bloom et al., 1996), or use a digital filter [Peckham et al., 2016], or (if using Var) add a third term 'J_c' to the cost function to penalise imbalance [Courtier and Talagrand, 1990, Courtier et al., 1998, Gauthier and Thépaut, 2001], which could be extended to penalise imbalances associated with AW. Some techniques that have been designed to accelerate the spin-up of models, such as the RIP (Running In Place) method of Kalnay and Yang [2010], Yang et al. [2012], may also be used to reduce the spin-down problem. RIP is suited to ensemble transform methods (e.g. ETKF), where the matrix of update weights at an analysis time is also applied to the ensemble at an earlier time in order to imitate the action of a smoother.

As an example the ways that the Met Office's convective-scale UKV system currently responds to the above questions are now outlined. The UKV adopted 4DVar DA in summer 2017 with a 1h DA window (its parent global model uses hybrid-En4DVar with a 6h window). The control parameters (Sect. 7) are streamfunction, velocity potential, unbalanced pressure, aerosol, and a non-Gaussian variable representing total water (the sum of WV, cloud liquid, and cloud ice). The UKV model itself has the ability to recognise six hydrometeors: WV, cloud water, rain water, ice, snow and graupel [Lean et al., 2008], but only four are used though as prognostic variables [Lean et al., 2008, Leoncini et al., 2013]: WV, cloud water, rain water and a single combined ice and snow variable. The balances enforced by the DA are strong hydrostatic balance, weak geostrophic balance, and a coupling between temperature and total water when reasonably close to saturation (Sect. 7.6). The Met Office also uses a latent heat nudging scheme (Sect. 5.6) to assimilate some precipitation observations.

5 Methods of assimilation of atmospheric water information

The most convenient means of assimilating AW information is in the same DA system used to assimilate dynamical/thermodynamical fields (i.e. Var, EnKF, etc.), but the nonlinear and inhomogeneous nature of cloud and precipitation means that this cannot always be done. Here we review the main methods used operationally.

5.1 Pure variational methods (Var)

Var is popular for NWP, e.g. Parrish and Derber [1992], Courtier et al. [1998], Rabier et al. [1998], Derber and Bouttier [1999], Gauthier et al. [1999], Lorenc et al. [2000], Rawlins et al. [2007], Talagrand
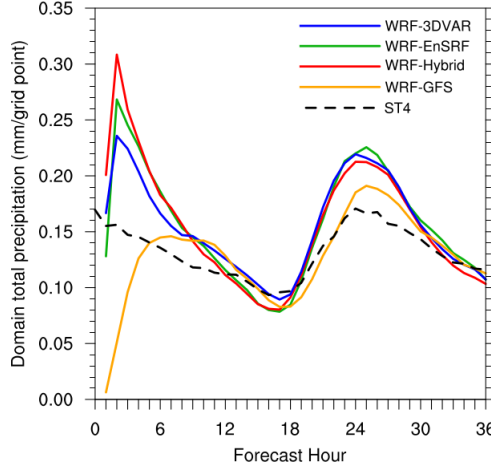
Figure 2: Forecasts of hourly accumulated total precipitation (averaged over 44 cases in May/June 2011) initialised using various DA methods as indicated in the caption, with ST4 representing the observed precipitation. The domain is the central United States and the model is WRF with a 4 km gridlength. Acronyms not defined in Table 1 are: **WRF** (Weather Research and Forecasting Model), **EnSRF** (Ensemble Square Root Kalman filter), **GFS** ([NCEP] Global Forecast System), and **ST4** ([NCEP] STage IV precipitation analysis). ©American Meteorological Society. Used with permission. Taken from Schwartz and Liu [2014].

[2010]. Its success for large-scale NWP is thought to be partly due to a background error covariance matrix (**B**) that represents realistic correlation length-scales and physical balances, and to its ability to handle a wide range of observations, even with nonlinear observation operators. Processes contributing to **B** are represented only if they are explicitly modelled (Sect. 7.1), which is a problem when the dominant balances are unknown or highly flow-dependent as in the case of AW. The **B**-matrix is full rank, but not automatically flow-dependent at the start of the DA window. 4DVar also requires linear/adjoint models, and does not work well if the window exceeds the model's linear regime [Zou et al., 1993, Xu, 1996, Zou, 1997, Errico et al., 2007a, Vukicevic, 2008, Stiller and Ballard, 2009]. 3DVar is cheaper, but 4DVar allows variables to be evolved within the DA window, which is important when model fields are strongly coupled in a flow-dependent way as 4DVar allows observations to update all relevant variables [Vukicevic et al., 2004], even if they are not directly observed. To illustrate the problem with the 3DVar **B**-matrix, panel (b) of Fig. 3 shows the analysis increment of the hail field in the ARPS/3DVar system due to the assimilation of a radar reflectivity observation made at the red dot position. The increment is spread-out due to the **B**-matrix, and is horizontally and vertically symmetric due to the way that spatial covariances are modelled. The part of the increment below the freezing level (around the level of the observation) is unrealistic – ideally the increments would match those of panel (a) which is the increment required to give the true hail in this simulation (for details see Kong et al. [2018]). This kind of problem justifies development of Ens-based DA systems.

5.2 Pure ensemble methods (EnKF)

The EnKF works without the need for linearisation or adjoint steps, or an explicit **B**-matrix [Evensen, 1994]. Instead, sensitivities of the model's version of the observations to the state are found in the space spanned by the ensemble perturbations in observation space. This is done via the $p \times N$ matrix $\mathbf{Y}' = \mathbf{HX}'$ (p observations, N ensemble perturbations), where column i of $\mathbf{Y}'$ is $\mathcal{H}(\mathbf{x}_i) - \mathcal{H}(\bar{\mathbf{x}})$. Here $\mathcal{H}(\mathbf{x}_i)$ is the non-linear observation operator acting on the ith member, and $\bar{\mathbf{x}}$ is the ensemble mean. Wherever $\mathbf{X}'$ appears in the EnKF update equation, it is always accompanied by $\mathbf{H}$ as $\mathbf{HX}'$, so $\mathbf{Y}'$ may be computed explicitly, and the transpose $(\mathbf{HX}')^{\mathrm{T}}$ then follows trivially, thus avoiding the need

15
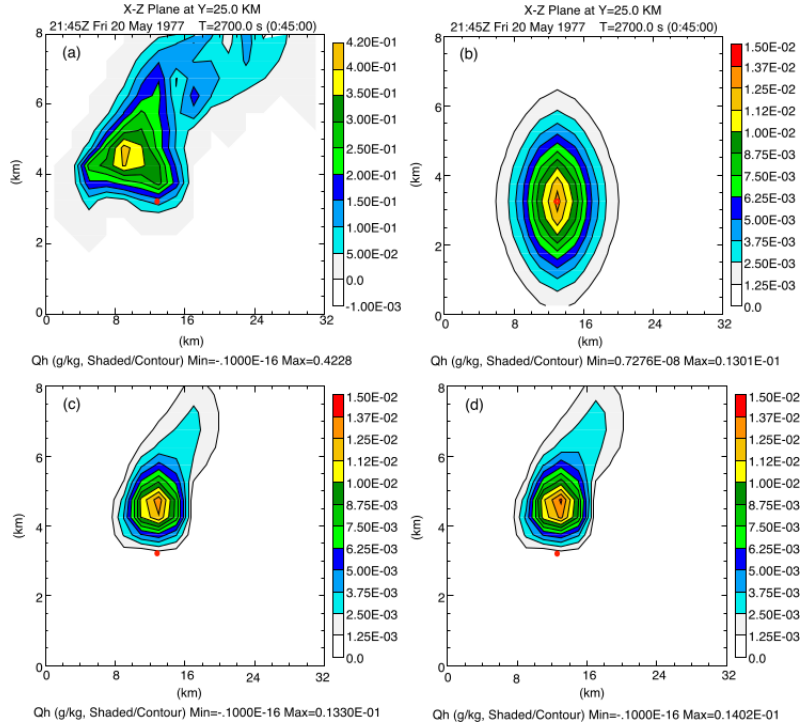
Figure 3: Increments of hail mixing ratio in a horizontal/height plane of the ARPS system (Sect. 2) with a 2 km grid length and a simulated radar reflectivity observation. Panel (a) is the true hail minus the background hail forecast (assumed 5 minute forecast). This indicates the magnitude and structure of a perfect analysis increment. The remaining panels are analysis increments due to the assimilation of a single reflectivity observation at the red dot for various DA methods: (b) 3DVar, (c) En3DVar, and (d) DfEnKF. 40 ensemble members are used in (c) and (d). ©American Meteorological Society. Used with permission. Taken from Kong et al. [2018].

for $\mathbf{H}^{\mathrm{T}}$ to be developed. The ensemble Kalman smoother avoids the need for the linear/adjoint of the model in a similar way (e.g. by including it as part of $\mathcal{H}$). A $\mathbf{B}$-matrix is not needed as the EnKF constructs this from the ensemble via an implicit evaluation of $\mathbf{B} \sim \mathbf{X}'\mathbf{X}'^{\mathrm{T}}/(N-1)$. The finite value of N though leads to rank deficiency, noisy ensemble covariances, and variances that are too small [van Leeuwen, 1999, Houtekamer and Mitchell, 2001, Ehrendorfer, 2007, Sacher and Bartello, 2008, Houtekamer and Zhang, 2016]. These problems lead to sub-optimality in the analysis and to the need for fixes such as localisation (to reduce noise in correlations), and ensemble inflation (to artificially increase the ensemble spread) [Anderson and Anderson, 1999, Hamill et al., 2001]. As an example, a recent success of the EnKF at convective-scale with radar data includes Johnson et al. [2015], which proved better resulting forecasts than a 3DVar system tested.

Environment Canada (EC) pioneered the operational implementation of the EnKF [Buehner et al., 2010a,b], and now DWD has an operational convective-scale ensemble DA system (KENDA, Kilometre-scale ENsemble Data Assimilation, Schraff et al. [2016], Bick et al. [2016]). KENDA is used with the COSMO model and utilises the Local Ensemble Transform Kalman Filter flavour of the EnKF. NCEP also runs a convective-scale system (HRRRE, High Resolution Rapid Refresh Ensemble), used with the WRF model and, as of 2019, the Ensemble Square-Root Filter (Xuguang Wang, personal communication).

5.3 Pure ensemble-variational methods (EnVar)

Ensemble-variational (EnVar) methods are increasingly being used in NWP. EnVar uses the variational machinery to give a deterministic analysis using a $\mathbf{B}$-matrix that is estimated from the forecast ensemble instead of being modelled [Lorenc, 2003, Buehner, 2005, Liu et al., 2008, Lorenc, 2013, Wang and Lei, 2014, Bannister, 2017]. The method has the ability to use the same observation operators as Var, but has access to fully flow-dependent background error statistics, as the EnKF does. EnVar does not though generate an analysis ensemble, and still relies on localisation and inflation, although rank deficiency problems can be partially alleviated by combining the $\mathbf{B}$-matrices from Var and EnVar (Sect. 5.4). To compare En3DVar[8] and the EnKF with 3DVar, panels (c) and (d) of Fig. 3 show the analysis increments of the hail field in the ARPS/En3DVar and DfEnKF[9] systems due to the assimilation of a radar reflectivity observation. The hail increments are spread in a much more realistic way than 3DVar (panel b), matching more closely the ideal increments (panel a). These kinds of properties of ensemble systems make them desirable from the covariances point of view. Humidity variables (specific humidity, SH, or relative humidity, RH) though remain difficult variables for an EnVar (and EnKF) to use. Liu et al. [2009] for instance found that humidity is particularly sensitive to sampling error, which is thought to be related to the short spatio-temporal correlation scales of AW compared to, e.g., wind variables. That said, there have been recent successes of EnVar at convective-scale with radar data, including Wang and Wang [2017].

5.4 Hybrid methods (hybrid-EnVar)

A hybrid method is one where the $\mathbf{B}$-matrix of Var is averaged with the ensemble-derived matrix of EnVar, as described in Hamill and Snyder [2000], Clayton et al. [2013], Liu and Xue [2016], Bannister [2017], Gustafsson et al. [2018]. Background error covariance modelling for $\mathbf{B}$, and localisation/inflation remain important tasks for an effective hybrid system. The shift to convective-scale DA and the inclusion of new AW-related variables in the analysis, will still demand developments to the modelled $\mathbf{B}$-matrix [Auligné et al., 2011]. These methods are now operational at many centres.

---

[8]En3DVar is the version of EnVar equivalent to 3DVar but with ensemble-derived covariances.

[9]DfEnKF is an implementation of the EnKF used by Kong et al. [2018].

5.5 Pre-assimilation retrievals

Some remotely sensed observations are not currently assimilated directly, usually due to technical and scientific reasons concerning cloud and precipitation. Instead, information can be retrieved using pre-DA procedures, which generate fields of humidity data, which can be passed to the DA system as pseudo observations.

1DVar is one such pre-DA method (e.g. Rodgers, 2000). This is a variational procedure that individually estimates vertical columns of model quantities by minimising a cost function (essentially a reduced DA problem). The cost function requires the radiance observations and a background column. The vertical background error covariances of these data need to be specified, and so many of the issues concerning the full **B**-matrix – including cloud condensate and precipitation (to be discussed in Sect. 6) – apply, although these can be simpler to treat than for the 3D/4DVar relatives. As we shall see in Sect. 8, there are two approaches to introducing cloud/precipitation information into Var. These are the diagnostic and explicit approaches. The diagnostic approach is used in the 1DVar schemes of Bauer et al. [2006a], Lopez and Bauer [2007], Janisková [2015], where cloud and precipitation changes are diagnosed from the conventional temperature and humidity control variables for input into the radiative transfer model. The relative simplicity of 1DVar though permits an explicit representation of cloud/precipitation by augmenting the conventional 1DVar control variables with separate cloud/precipitation-related variables. Example augmented 1DVar studies include Chevallier et al. [2002], who retrieved cloud cover and liquid water profiles from IR and MW data, Pavelin et al. [2008], who retrieved cloud top pressure and effective cloud fraction from IR data, Geer et al. [2008], who retrieved profiles of cloud and precipitation from MW data, and Martinet et al. [2013], who retrieved profiles of liquid and ice cloud, and cloud fraction from IR data. Although these studies show that cloud and precipitation information can be retrieved in this way, only the study of Pavelin et al. [2008] passed the retrieved cloud information to the main DA, but then only to constrain the radiative transfer model there (to facilitate cloud affected radiance assimilation) rather than to assimilate cloud information itself. The 1DVar approach also allows efficient experimentation with channel selection to minimise errors from problematic channels.

An alternative non-Var method has also been proposed by Caumont et al. [2010]. The method does not need a **B**-matrix, but instead uses the same machinery as the basic particle filter [van Leeuwen, 2009, van Leeuwen et al., 2019] by estimating the prior PDF from a population of nearby model columns $(\mathbf{x}_i, \; i = 1 \ldots N)$, $p(\mathbf{x}) = N^{-1} \sum_i \delta(\mathbf{x} - \mathbf{x}_i)$. The retrieved column is the expectation of $\mathbf{x}$ over the posterior PDF, $\mathbf{x}^{\mathrm{ret}} = \int d\mathbf{x} \, \mathbf{x} p(\mathbf{y} = \mathbf{y}^{\mathrm{o}} | \mathbf{x} = \mathbf{x}^{\mathrm{t}}) p(\mathbf{x} = \mathbf{x}^{\mathrm{t}})$, where $\mathbf{x}^{\mathrm{t}}$ is the true state. Caumont et al. [2010], Wattrelot et al. [2014] used this method successfully to retrieve RH from radar reflectivity observations; Guerbette et al. [2016], Duruisseau et al. [2019] have used the method with MW radiances; and Borderies et al. [2018] used a similar approach to validate airborne radar measurements against AROME forecasts.

Other methods exist too. Renshaw and Francis [2011] for instance describe how the Met Office computes pseudo observations of 'total RH' (like ordinary RH, but includes condensed water species as well as vapour) from SEVIRI, surface-based cloud observations, and forecasts, which are then passed to their Var system for assimilation (see references therein for methods used in other centres).

A retrieval method is often regarded as a stop-gap method until the main DA method develops to directly handle all-sky IR and MW radiances and radar/lidar reflectivities. Assimilating retrievals is sub-optimal due to the potential double use of the background state [Lopez, 2011], the absence of complementary constraints of the 3/4D problem (including all other observations), and the presumed better error characterisation of the raw data over the retrievals. We note here that, by building on the above studies, it has been demonstrated that the direct assimilation of all-sky MW and IR satellite radiances is now possible in principle. All-sky MW data are considered easier to use than all-sky IR data due to the smoother properties of the MW radiative transfer model output between cloudy and clear scenes [Bauer et al., 2011b], so these data were the first to be treated directly [Bauer et al., 2010, Zhu et al., 2016, Migliorini and Candy, 2019]. Recently though satellite IR data have also been assimilated directly [Zhang et al., 2016, Minamide and Zhang, 2018, Okamoto et al., 2019], which is

discussed further in Sects. 6.5, 6.8, and 6.9.

5.6 Post-assimilation methods

Latent heat nudging (LHN, Jones and Macpherson [1997], Sun et al. [2014]) is an example of a diabatic initialisation method, often used in lieu of a direct precipitation DA capability, for radar reflectivity data (e.g. Golding et al., 2014) or lightning flash rate data (e.g. Chang et al., 2001). LHN nudges information from either surface precipitation or 3D reflectivity observations into the evolving post-analysis forecast. The method uses the principle that the latent heat released as a consequence of precipitation is injected into under-precipitating parts of the model, and the resulting upward motion induces precipitation [Weygandt et al., 2007]. Similarly, modelled precipitation can be inhibited by reducing the latent heating (e.g. setting it to zero in areas of zero radar return, Benjamin et al., 2016).

LHN due to use of surface precipitation uses the assumption that the model's latent heat release, integrated over a vertical column of the atmosphere (levels 0 to L), is proportional to the model's precipitation rate at the bottom of that column, PR^{b}. The observed surface precipitation rate, PR^{ob}, will inevitably be different, so the method nudges-in extra temperature tendencies as the forecast proceeds (via the IAU for instance, see Sect. 4). The tendencies are consistent with the extra latent heating required to match PR^{ob}. This correction, $\dot{\theta}_k^{\mathrm{LHN}}$, is assumed to have the form $\dot{\theta}_k^{\mathrm{LHN}} \propto \dot{\theta}_k^{\mathrm{b}}$ (for level k), where $\dot{\theta}_k^{\mathrm{b}}$ is the model's own potential temperature tendency due to moist processes. The extra tendency typically has the following form:

$$\dot{\theta}_k^{\mathrm{LHN}} = \frac{PR^{\mathrm{ob}} - PR^{\mathrm{b}}}{PR^{\mathrm{b}}} \dot{\theta}_k^{\mathrm{b}}, \tag{1}$$

which is generally applied only when $PR^{\mathrm{ob}} - PR^{\mathrm{b}} > 0$. Precipitation is generally found to increase (decrease) as the temperature is increased (decreased) in the evolving model. Despite its usefulness (e.g. Dixon et al., 2009, Ballard et al., 2016), there are flaws to this procedure (e.g. the assumptions mentioned above, the neglect of horizontal precipitation advection, and the nonlinear response of the model to the extra tendency $\dot{\theta}_k^{\mathrm{LHN}}$), which motivate need for on-line solutions.

A similar nudging method can be used to encourage models to rain by pushing the model's RH towards saturation. This was done by Papadopoulos et al. [2005], Dixon et al. [2016] as a way of exploiting lightning data, which are otherwise difficult to assimilate. An alternative means of using lightning data is to modify model parameters (e.g. the trigger function in a convective parametrisation scheme) in order to encourage convection in forecasts where lightning has been observed [Mansell et al., 2007]. Other examples of diabatic initialisation are discussed in Sun et al. [2014].

Hydrometeor fields are also estimated using separate analysis procedures, such as the Rapid Refresh system of NCEP for their regional North American model domain [Benjamin et al., 2016]. This non-Var procedure estimates water and ice cloud coverage from satellite cloud top measurements, surface cloud base measurements, and model forecasts; and estimates precipitating hydrometeors from radar reflectivity measurements. This procedure can yield cloud fields that are inconsistent with those inferred by the model's WV and temperature fields, and so these model fields are adjusted accordingly. Other systems (such as Météo-France's convective-scale AROME/3DVar system, as documented by Seity et al., 2011) do not analyse some small-scale variables (like cloud and precipitation hydrometeors) leaving the model itself to spin-up these quantities.

There has been, however, progress in the direct assimilation of precipitation. Lopez [2011] shows that the direct assimilation of radar- and gauge-derived precipitation is possible, at least in a global model, suggesting that techniques like LHN can be avoided. Their paper discusses the degree of linearity of the observation operator for accumulated precipitation, and finds the surprising result that the problem appears to be more linear for a six-hour accumulation period than for a one-hour period.

6 Hazards of assimilating atmospheric water, and possible solutions

All methods of assimilation have issues that may lead to sub-optimalities or that may even make matters worse, and these raise associated research questions. Some of the major issues to do with Var, EnKF, EnVar, or hybrid-EnVar are discussed in this section and many are thought to be highly relevant to convective-scales.

6.1 Sharp features and their position errors

The smoothing effect of $\mathbf{B}$ [Bannister, 2008a] in Gaussian DA methods (Var, EnKF, and hybrid) is often beneficial, but too much smoothing can result in unphysical analyses. This is especially true in strong vertical gradients such as at the top of a boundary layer (BL) inversion or at the tropopause, and when a gradient is in the wrong place in the background [Renshaw and Francis, 2011, Ingleby et al., 2013b].

A pathological case is shown in Ingleby et al. [2013b] for the case of an RH observation near the top of the BL and a similar picture is given here (Fig. 4a). Suppose that a saturated boundary layer (true RH field, μ, black line) is forecast by the model to be too deep (μ^{b}, blue line)[10]. An RH observation (circle) made just above the true BL top, but just below the forecast BL top suggests that the model's RH field should be reduced around the observation (defined by the RH structure function of $\mathbf{B}$ associated with this observation – the yellow curve). Because μ and μ^{b} are different, and each has a sharp feature at the top of the BL, this results in a large difference between the observation and the background. Consequently the analysis increment ($\delta\mu$, yellow curve) is large (and is in this case negative). Adding μ^{b} and $\delta\mu$ gives the analysis (μ^{a}, red curve), which has a smoothed BL top and negative RH values above this, even though neither the background nor the observation values are themselves negative.

Now consider assimilating a temperature observation instead. Fig. 4b shows a true temperature maximum at the BL top (black curve), and a background forecast of the BL that is again too deep (blue curve). Assimilating the temperature observation (circle) gives increments (yellow curve) and an analysis (red). The assimilated observation has not corrected the BL top height, but has instead (in this case) reduced the static stability by decreasing dT/dz, even though the static stability was too low in the background. This degradation can conceivably go on to affect cloud and precipitation forecasts negatively. Furthermore if the $\mathbf{B}$ matrix does not couple temperature and humidity correctly then this analysis may result in moisture imbalance. In this case the analysed temperature is too low, which can lead to anomalous condensation and hence excessive precipitation in the subsequent forecast ('spin-down' as mentioned in Sect. 4). This is the mechanism proposed by Ingleby et al. [2013b] to explain excessive precipitation over the first hour of their synoptic-scale forecasts. Condensation processes (including those that are anomalous) can also impact dynamical variables in the model via the creation of potential vorticity [Chagnon et al., 2013].

Improvements to the DA system may solve this kind of problem. Coupling RH and temperature properly could reduce moisture imbalance, and using structure functions that are aware of sharp gradients could reduce the 'leaking' of information across the gradient. Piccolo and Cullen [2011, 2012] use an adaptive grid which aims to reduce the vertical covariance lengthscales around the BL, resulting in less coupling across the BL top. Fowler et al. [2010] showed that the background error covariance matrices differ greatly between those trained on data with different kinds of inversion structure (namely a 'strong' inversion, a 'weak' inversion, no inversion, or 'climatological' data), indicating the need for flow-dependency of $\mathbf{B}$. Ménétrier and Montmerle [2011] also showed clear differences between cases of clear skies and fog (in fog, temperature errors in the boundary layer have a shorter vertical lengthscale, which can help to mitigate this problem if this flow-dependence is represented in the DA, see e.g. Sect. 7.7.3).

---

[10]We take RH as a fraction ($0 \le \mu \lesssim 1$), rather than as a percentage of the saturated SH.

(a) Relative humidity around the boundary layer (b) Temperature around the boundary layer
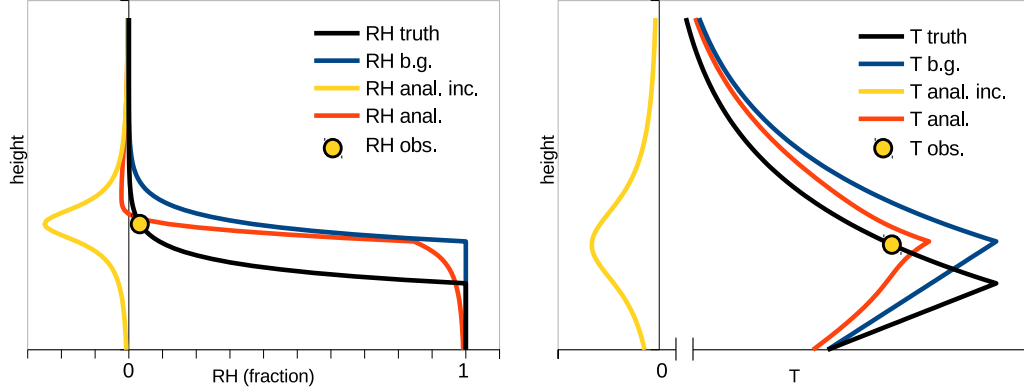
Figure 4: Schema showing the damaging effect of assimilating data close to a sharp feature – in this case (a) RH and (b) temperature fields around the top of a saturated boundary layer inversion – when the background forecasts the feature in the wrong position. Here the boundary layer inversion top is forecast too high (blue) compared to the truth (black). The DA produces negative increments (yellow) to give the analysis (red). There is a point observation of RH in (a) and temperature in (b) (circles).

Position errors also often occur in the background forecast in the horizontal dimension, especially at convective-scales and for sharp features like thunderstorms [Brewster, 2003a,b]. These problems point towards the need for feature shifting algorithms, which can avoid large increments. As well as allowing observations to correct 'amplitude' errors (as in conventional DA), such algorithms also allow correction of 'phase' errors. Nehrkorn et al. [2014] demonstrates a technique that allows a field of displacement vectors $(\delta\mathbf{i}, \delta\mathbf{j})$ to be analysed (where $\delta\mathbf{i}$ $(\delta\mathbf{j})$ is a field of displacements in the longitudinal (latitudinal) direction, subject to smoothness and other constraints). These specify how an arbitrary field, u, should be distorted, i.e. $u(\mathbf{i}, \mathbf{j}) \rightarrow u(\mathbf{i} + \delta\mathbf{i}, \mathbf{j} + \delta\mathbf{j})$, where $(\mathbf{i}, \mathbf{j})$ a grid position. This method could be used to displace features in 3D, but Nehrkorn et al. [2014] and other studies have applied this to displace fields only horizontally. Fowler et al. [2012] on the other hand used a grid distortion method to shift the BL inversion vertically. In this scheme, a new parameter, a is analysed (together with the usual amplitude adjustments), where a specifies how the usual model level heights (z^i) should be shifted to new heights $(\tilde{z}^i)$: $\tilde{z}^i = z^i + aD(z^i)$. Here $D(z^i)$ is a prescribed displacement function, which has value unity around the inversion height, and decays to zero above and below. Even without shifting algorithms, the impact of incorrectly placed features can be lessened by using observation operators that do neighbourhood averaging (Geer and Bauer, 2011, C. Gaffard, personal communication) rather like the scale-selective verification method of Roberts and Lean [2008], or by incorporating the same Bayesian estimation approach discussed in Sect. 5.5, which uses information from a neighbourhood of points in the background rather than from a single point.

6.2 Inhomogeneity and flow-dependence of error (co)variances

Forecast errors of specific humidity (SH, q) vary in the atmosphere by orders of magnitude. Most of the variability is in the vertical (e.g. over the UK q ranges from $\sim 10^{-2}\mathrm{kg\,kg}^{-1}$ at the surface to $\sim 10^{-6}\mathrm{kg\,kg}^{-1}$ in the stratosphere, and the associated error will change commensurately). This leads to the following issue, which is raised by Rabier et al. [1998] and Dee and Da Silva [2003], and is relevant when q is a DA control parameter (most Ens schemes have q as a control parameter if the model uses q, and many Var schemes choose q as a control variable, see Sect. 7.3). Consider a column with state vector $\mathbf{x} = \begin{pmatrix} q_1 & \cdots & q_L \end{pmatrix}$, and a point observation of q made at level l', $y_{l'}^{\mathrm{o}}$. The observation operator has the form $\mathbf{H} = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$, which comprises zeros, apart from 1

for level l'. The following analysis equation,

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{B}\mathbf{H}^{\mathrm{T}} \left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} \right)^{-1} \left(\mathbf{y}^{\mathrm{o}} - \mathbf{H}\mathbf{x}^{\mathrm{b}} \right),\tag{2}$$

translates to $q_l^{\mathrm{a}} = q_l^{\mathrm{b}} + \sigma_l^{\mathrm{b}}\sigma_{l'}^{\mathrm{b}}r_{ll'}(y_{l'}^{\mathrm{o}} - y_{l'}^{\mathrm{b}})/(\sigma_{l'}^{\mathrm{o}\,2} + \sigma_{l'}^{\mathrm{b}\,2})$. Here $y_{l'}^{\mathrm{b}} = \mathbf{H}\mathbf{x}^{\mathrm{b}}$ is the background's version of the observation, $r_{ll'}$ is the correlation between background errors of q_l^{b} and $q_{l'}^{\mathrm{b}}$, σ_l^{b} is the background error standard deviation of q_l^{b}, and $\sigma_{l'}^{\mathrm{o}}$ is the observation error standard deviation. This result is equivalent to performing Var or an EnKF with the single observation. The analysis update at level l, $q_l^{\mathrm{a}} - q_l^{\mathrm{b}}$, is influenced by the observation at level l' via the background error covariance $B_{ll'} = \sigma_l^{\mathrm{b}}\sigma_{l'}^{\mathrm{b}}r_{ll'}$. If l is a dry level (e.g. upper troposphere), but l' is moist, and the correlation $r_{ll'}$ is large then the large value of $\sigma_{l'}^{\mathrm{b}}r_{ll'}$ could anomalously 'flood' moisture into level l, as in Sect. 6.1.

The degree of 'flooding' is therefore affected by the (co)variances, which can be strongly flow-dependent, e.g. in the vicinity of clouds. Lorenc [2007] for instance found that the background error variances of temperature and RH are highly dependent on the presence of cloud layers, and, at the top of stratocumulus layers, variances should often be double those assumed by the DA. Montmerle and Berre [2010] find some profound differences in the background error statistics for precipitating and non-precipitating regions of the AROME model, namely that in precipitating regions all variables have shorter horizontal correlation lengthscales, SH has larger vertical correlation lengthscales, vorticity and divergence have larger variances, and SH and divergence show stronger cross-correlations. There is also a justification to increase background error variances in regions that are clear in the background but show cloud in observations [Minamide and Zhang, 2019]. The static background error correlations are also often too large across cloud-top inversions in addition to the heights of the inversions often being forecast wrong [Lorenc, 2007, Fowler et al., 2010]. When the adaptive mesh method of Piccolo and Cullen [2011, 2012] (Sect. 6.1) decreases the vertical lengthscale, the implied variance of potential temperature is automatically increased[11]. This has the potential to mitigate the mentioned variance deficiency of temperature at cloud layers, but not for moisture.

6.3 Assimilation of total column quantities in incompatible scenes

Total column water vapour (TCWV) estimates can be extracted from MW AMSU-A and SSM/I measurements and then assimilated. The observation operator for a TCWV measurement (operating on an SH state vector) has the form: $\mathbf{H} = \begin{pmatrix} \rho_1\Delta z_1 & \cdots & \rho_L\,\Delta z_L \end{pmatrix}$, where ρ_l is air density, and Δz_l is level thickness at vertical level l. Assimilating a TCWV measurement ($y_{\mathrm{TCWV}}^{\mathrm{o}}$) updates the values of q in each level according to (2):

$$q_l^{\mathrm{a}} = q_l^{\mathrm{b}} + \left(\sum_{l'=1}^{L} B_{ll'}\rho_{l'}\Delta z_{l'} \right) \left[\frac{y_{\mathrm{TCWV}}^{\mathrm{o}} - y_{\mathrm{TCWV}}^{\mathrm{b}}}{\sigma_{\mathrm{TCWV}}^{\mathrm{o}\,2} + \sum_{\ell=1}^{L} \rho_\ell\Delta z_\ell \sum_{\ell'=1}^{L} B_{\ell\ell'}\rho_{\ell'}\Delta z_{\ell'}} \right].\tag{3}$$

The notation is the same as in Sect. 6.2, and here $y_{\mathrm{TCWV}}^{\mathrm{b}} = \mathbf{H}\mathbf{x}^{\mathrm{b}}$ is the model's observation and $\sigma_{\mathrm{TCWV}}^{\mathrm{o}}$ is the observation error standard deviation. The analysis update at level l is proportional to the inner product of the structure function associated with level l (row l of $\mathbf{B}$) with $\mathbf{H}$ (round bracketed term in the above). This inner product can be large for those update levels, l, whose structure functions have significant values in the model's lower levels (where $\rho_{l'}$ is also large).

If the $\mathbf{B}$-matrix is incorrect, the assimilation of a TCWV observation would distribute the observation throughout the column wrongly. Problems can still occur though even when the $\mathbf{B}$-matrix is correct climatologically, but when the observed and forecast scenes are incompatible, e.g. when a satellite views a dry column, but the background forecasts a moist column (e.g. when the background's boundary layer is too deep as in Fig. 4a; see also Sect. 6.9). Note that MW radiances are currently

---

[11]This arises because in the Met Office, the potential temperature increment field, $\delta\theta$, is derived from the pressure increment, δp, via the hydrostatic pressure relationship, $\delta\theta \sim (\partial/\partial z)\left\{(\Pi^{\mathrm{b}}/p^{\mathrm{p}})\delta p\right\}$ (where z is height, and p^{b} (Π^{b}) is background (exner) pressure). Shorter vertical lengthscales in δp lead to larger $\delta\theta$, thus effectively increasing the implied variance of θ.
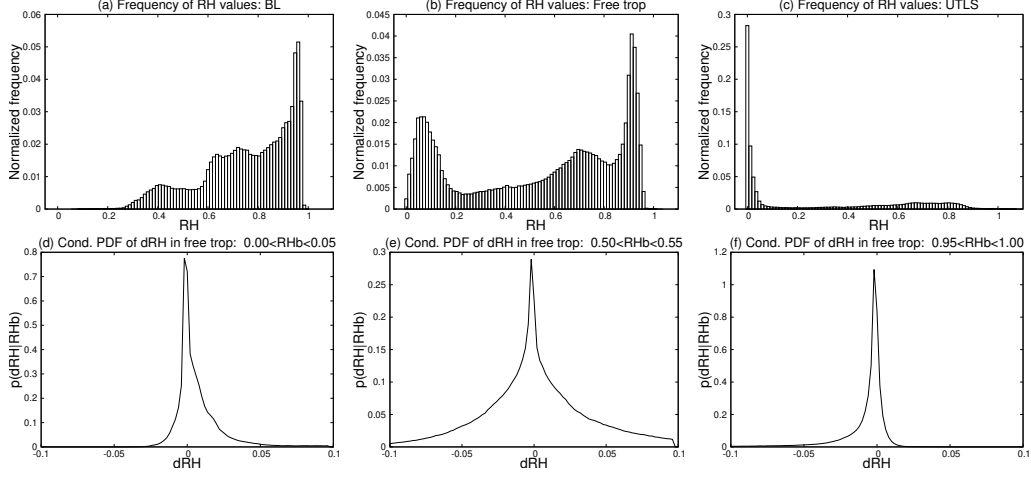
Figure 5: Panels (a)-(c): histograms of RH from three-hour forecasts (μ^{3h}) for (a) the boundary layer, (b) free troposphere, and (c) upper troposphere/lower stratosphere. Panels (d)-(f): conditional forecast error PDFs for the free troposphere. Panel (d) is for the conditioning $0 \leq \mu^{3h} < 0.05$, panel (e) is for $0.5 \leq \mu^{3h} < 0.55$, and panel (d) is for $0.95 \leq \mu^{3h} < 1$. All data are aggregated over the horizontal domain of the UKV model from 1st to 5th September 2013 (15Z). The error PDFs are computed from scaled differences of 6- and 3-hour forecasts valid at the same time $\delta\mu = (\mu^{6h} - \mu^{3h})/\sqrt{2}$. The $\sqrt{2}$ is a normalisation so that the variance of $\delta\mu$ matches that of forecast errors assuming that error statistics of μ^{6h} and μ^{3h} are identically distributed and uncorrelated, i.e. $\langle \delta\mu^2 \rangle = \left\langle \left(\mu^{6h} - \mu^{t}\right)^2/2 + \left(\mu^{3h} - \mu^{t}\right)^2/2 + \left(\mu^{6h} - \mu^{t}\right)\left(\mu^{3h} - \mu^{t}\right) \right\rangle \approx \left\langle \left(\mu^{6h} - \mu^{t}\right)^2 \right\rangle/2 + \left\langle \left(\mu^{3h} - \mu^{t}\right)^2 \right\rangle/2 \approx \left\langle \left(\mu^{3h} - \mu^{t}\right)^2 \right\rangle$, where μ^{t} is the true RH.

assimilated directly (e.g. Bell et al., 2008), rather than via TCWV, but the above issue still stands with that approach.

6.4 Humidity bounds and related non-Gaussian aspects

The physical bounds on SH (q) and RH (μ) may be expressed as $0 \leq q \lesssim q_{\text{sat}}$ and $0 \leq \mu \lesssim 1$ respectively (where q_{sat} is the saturated SH), but the issues mentioned in Sects. 6.1-6.3 could lead to these bounds being violated in the analysis. Exceeding the upper bound can create an instability (an imbalance between the AW and thermodynamic variables), which can lead to spin-down (Sect. 4, Hólm et al. 2002, Ingleby et al. 2013b).

Forecasts of μ frequently take values that are close to these bounds. Histograms of three-hour UKV forecasts of μ collected over five days of data are shown in panels a-c of Fig. 5 for three atmospheric layers. The boundary layer (panel a) has a peak at $\mu \approx 0.96$; the free troposphere (panel b) has peaks at $\mu \approx 0.05$ and $\mu \approx 0.9$; and the upper troposphere/lower stratosphere (panel c) has a peak just above $\mu = 0$. This means that there is a danger that the analysis may tip over to unphysical vales, unless the DA has a mechanism to respect these boundaries.

This issue has an obvious connection to non-Gaussianity. The PDFs of $\delta\mu$ conditioned on three intervals of the data for the free troposphere are shown in panels d-f of Fig. 5 (dry, medium and moist respectively, as explained in the caption). These error PDFs are highly non-Gaussian: skewed for the dry and moist cases, and a non-zero kurtosis for the medium case. These results are consistent with those of Legrand et al. [2016], who found negatively skewed statistics for SH in parts of the atmosphere that are expected to be close to saturation, and small or positively skewed statistics in parts that are unsaturated. The skewed cases are not surprising given their closeness to the bounds discussed above.

DA schemes generally assume that errors in background variables and observations obey Gaussian statistics, and thus cannot favour increments one way over another. Non-Gaussian PDFs are not easily compatible with the methods described in Sects. 5.1-5.4, which are concerned with only first and second moments of variables.

Substituting a non-Gaussian PDF by a Gaussian cannot therefore allow maximum information to be extracted from the data assimilated. Ways of dealing with this problem are discussed in Sect. 9. A Gaussian distributed variable has the property that its mean (the expected value), median (the point where it is equally likely to have a higher value or a lower value), and the mode (the value where the PDF is maximum) are coincident. This property is not shared with non-Gaussian PDFs, so care must be taken when interfacing a non-Gaussian variable with conventional DA when using some of the non-Gaussian methods described in Sect. 9.

6.5   (Mis)specification of the observation error covariance matrix

The observation error covariance matrix has contributions from errors in the instrument, observation operators, preprocessing (if done), and representativity. Representativity errors (e.g. Waller et al., 2014) are due to variations in the real world seen by the observations, but not represented in the state vector due to unresolved variability. This has a greater contribution for coarse grids and point observations [Waller et al., 2014], as well as for finer-scale fields like AW. In fact representativity errors in humidity sensitive channels of instruments like MHS and HIRS can make up a significant portion of the total observation error for DA in global models [Bormann and Bauer, 2010]. Geer and Bauer [2011] for instance state that for MW brightness temperatures in the ECMWF model, the representativity error can be an order of magnitude larger than instrument error. There is also justification to assimilate observations with an adaptively inflated observation error variance. Zhang et al. [2016], Minamide and Zhang [2017, 2018] for instance showed that, in their EnKF, inflating observation error variances in regions where the innovation is significantly larger than the ensemble spread improves the assimilation of all-sky satellite IR data, and Migliorini and Candy [2019] showed that inflating observation error variances in regions of large liquid water path similarly improves the assimilation of all-sky MW data. In a recent study, Bishop [2019] argues the need for state-dependent observation error variances for observations of variables that have physical bounds.

An observation error covariance matrix can be non-diagonal and failure to specify off-diagonal elements can lead to sub-optimalities in the DA. This is true for direct and indirect observations [Liu and Rabier, 2003, Bormann and Bauer, 2010]. Dense observation networks are more susceptible to observation error correlations (from whatever source) than sparse observation networks, and the correlations are rarely well-known (especially as they are situation-dependent, Janjic and Cohn, 2006, Waller et al., 2014) and so are often set to zero. Ways of mitigating for their absence include observation thinning (reducing the number of observations assimilated in data-dense regions), super-obbing (combining many observations into one, Lorenc, 1981), and observation variance inflation (to decrease their weight in the DA) – see Waller et al. [2014] and references therein. The best way to assimilate observations is to use the correct observation error covariance matrix, which may be estimated by the method of Desroziers et al. [2005], but this method makes various assumptions (e.g. that the background error covariance matrix used in the DA is correct, Waller et al., 2016). Operational progress has been made recently by accounting for cross-channel correlations of IASI data [Weston et al., 2014, Bormann et al., 2016]. These studies demonstrate the positive impact of accounting for observation error correlations, especially for WV sensitive channels.

6.6   Creation/annihilation of cloud without humidity observations

In some DA systems the RH field can be modified when there are no observations of humidity, since RH is also affected by changes in temperature [Lorenc et al., 1996]. An RH value close to 1 suggests that cloud is present, so changing RH can affect the presence of model cloud. Is the ability of a DA system to modify model cloud when no humidity observations are used potentially problematic?

On the one hand, cloud information from the background is valuable if it is the only source of cloud information. This is an argument for conserving RH in the absence of new humidity data. This was done in the Met Office's first Var scheme, which used RH as a control parameter [Lorenc et al., 2000, Ingleby, 2001]. This scheme worked by correlating temperature and SH so an increase (decrease) in temperature will create (annihilate) WV to keep RH constant (details are given later in Sect. 7.4). On the other hand this imposed correlation may still exist when air is unsaturated, which can lead to new problems, especially when temperature biases are present [Dee and Da Silva, 2003]. This is an argument against conserving RH. In this case SH is a more suitable control parameter (Sect. 7.3). More recent schemes are able to distinguish the degree of saturation, and will correlate temperature and SH only when appropriate (e.g. the Met Office's current humidity control parameter, Ingleby et al. [2013b], see Sect. 7.6).

6.7 Biases

Bias is the systematic over- or under-estimate of a quantity. It can be in the background state, in the observations, and in the observation operators, and its presence can be very damaging [Dee, 2005]. The degree of bias may be a fixed or a slowly varying component of an error, and can be revealed in observation-minus-background statistics (e.g. Lorenc et al., 1996). Sometimes persistent errors in the position or timing of features in a forecast can appear as biases and can even dominate the observation-minus-background statistics [Bauer et al., 2011a]. An example of how bias can affect moisture depending on the choice of control parameter is discussed in Sect. 7.4.

Biases can be reduced in NWP systems using a bias correction (BC) scheme, e.g. by modifying the DA's observation operator by replacing $\mathbf{h}(\mathbf{x})$, with a bias-aware one, $\tilde{\mathbf{h}}(\mathbf{x}, \boldsymbol{\beta})$:

$$\tilde{\mathbf{h}}(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{h}(\mathbf{x}) + \sum_{i=0}^{r} \beta_i \mathbf{p}^i(\mathbf{x}), \tag{4}$$

[Dee, 2005]. Here, β_i is the ith bias correction coefficient (ith component of $\boldsymbol{\beta}$, $0 \leq i \leq r$), and $\mathbf{p}^i(\mathbf{x})$ is the ith bias predictor (a function of relevant parts of the state $\mathbf{x}$ that are believed to influence the bias, e.g. the diagnosed thickness of various layers of the atmosphere relevant to atmospheric sounders as in Auligné et al., 2007). Such bias correction schemes are especially useful for satellite radiance or brightness temperature data and can affect the analysis of any quantity that the data are sensitive to. The bias correction term in (4) though is only a linear function of the predictors, and is not expected to work well for cloud affected observations [Otkin et al., 2018]. Instead, Otkin et al. [2018] propose a scheme that is a nonlinear function of the predictors. For simplicity, consider only observations from one channel of one instrument, and assume that there is a single predictor, $\mathbf{p}^1$, which may or may not be a function of the state. The following is the non-linear bias-aware operator for observation j:

$$\tilde{\mathbf{h}}_j(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{h}_j(\mathbf{x}) + f(\mathbf{p}_j^1) \approx \mathbf{h}_j(\mathbf{x}) + \sum_{i=0}^{r} \beta_i \left(\mathbf{p}_j^1 - c\right)^i, \tag{5}$$

where $f(\mathbf{p}_j^1)$ is the bias correction function, which is approximated by an rth order Taylor series, $\boldsymbol{\beta}$ (with ith component β_i) is now the set of correction coefficients for this channel/instrument, and c is the mean value of the predictor over the observations. The $\boldsymbol{\beta}$ are determined with a statistical regression from all such observations, and there is a separate $\boldsymbol{\beta}$ for each channel/instrument. The relationship between β_i and f is $\beta_i = f^{(i)}(c)/i!$, where $f^{(i)}(c)$ is the ith derivative of f evaluated at c (but f itself need not be computed in practice). Otkin et al. [2018] tests this procedure for SEVIRI 6.2μm brightness temperatures for scenes partially affected by precipitating cloud, up to $r = 3$, and for a number of predictors (observed brightness temperature, observed cloud top height, total integrated model water, and satellite zenith angle). Most of these show a significant and nonlinear bias which is virtually removed using their procedure. Their method can also be extended to multiple predictors used together.

6.8 Sensitivities of radiances to cloud and precipitation fields

It is becoming necessary to include water species in addition to WV in the DA. This is because they can play an important role in the evolution of the model atmosphere within a DA cycle (e.g. in 4DVar), and can also strongly influence many of the simulated observations. The last point particularly concerns the effective exploitation of radar reflectivity and cloud/precipitation affected satellite observations.

Many satellite MW and IR channels are sensitive to cloud and precipitation, and the lack of accurate cloud and precipitation information in the DA can make their assimilation difficult. Consequently only clear sky data are often used, thus ignoring about half of satellite observations, although, as stated in Sect. 5.5, this situation is fast improving with the advent of all-sky assimilation. As observations affected by cloud and precipitation often indicate rapid change (e.g. convective activity) [Errico et al., 2007a], there is a strong case to develop the means to assimilate them.

Early methods used to assimilate cloud-affected data did this via a pre-assimilation retrieval (Sect. 5.5). Deblonde and English [2003] for instance assimilated synthetic SSMIS cloudy and non-cloudy data with 1DVar using (i) a method that allows mainly for the WV (as $\ln q$ plus a liquid water path variable), to be varied, while penalising occurrence of supersaturation, and (ii) a retrieval method that allows WV and cloud water to be varied during the procedure (via a variable related to total water, $\ln(q + q_l)$). Temperature and surface wind speed are also retrieved in both methods. They find that only method (ii) is able to reliably retrieve temperature in cloudy conditions. Method (i) does not work well because the cloud field is not varied, which means that the retrieved temperature takes unrealistic values to fit the observations [Gong and Hólm, 2011]. This aliasing effect projects errors to other variables like temperature to compensate for the lack of treatment of cloud and precipitation. Bauer et al. [2011a] gauged the effect of cloud on the prediction of brightness temperature and showed that the difference between accounting for and not accounting for cloud in the radiative transfer model can lead to differences of more than 1K in simulated brightness temperature for AMSU-A channel 4 (52.8GHz), compared to the instrument sensitivity of 0.25K.

There are intermediate strategies to allow assimilation of cloud or precipitation affected radiances [Bauer et al., 2011a]. One option is to model the effects of cloud and precipitation, not for the purpose of determining them, but to facilitate a more accurate analysis of temperature and WV. This can be done within (e.g.) 4DVar itself, or via a pre-processing step. Pavelin et al. [2008] retrieved cloud top pressure and cloud fraction from IR radiances via 1DVar, and Pangaud et al. [2009] via a CO_2 slicing procedure. This allows some of the effects of cloud to be accounted for in the radiative transfer calculations.

More recently there have been successful studies showing how IR and MW satellite data can be assimilated in all-sky conditions using a number of direct assimilation methods. Studies include those of Minamide and Zhang [2018] and Okamoto et al. [2019] for IR data, and Zhu et al. [2016] and Migliorini and Candy [2019] for MW data. All of these studies have used some kind of observation error variance inflation scheme to account for errors due to the presence of cloud. Minamide and Zhang [2018] is of particular note as it assimilated data on a convective-scale grid (3 km). Many of these studies have built on earlier work exploring the direct assimilation of satellite radiances, e.g. Vukicevic et al. [2004, 2006], Otkin [2010, 2012], Jones et al. [2013, 2014].

Accounting for the multiple phases of water is thus becoming desirable for convective-scale DA, but questions remain on how to make the condensed water phases available to the observation operators. These fields may either be diagnosed from basic fields like WV and temperature in the DA, or may be represented explicitly. The latter is typically done in EnKF-based assimilation. In Var methods though the former option is subject to closure assumptions in the diagnosis procedure, and the latter option requires accurate background error covariances, including the covariance of each water species to other variables. This issue is discussed further in Sect. 8.

6.9 Cloud and precipitation inconsistencies between background and observations

Further to the discussion in Sects. 6.1 and 6.3, it is common to find that the cloud or precipitation properties of observations are not met by the background field (e.g. observations indicate cloud, but the background is cloud free, Vukicevic et al., 2004, Geer et al., 2018). These mismatches may be due to forecast model error (e.g. biases – Sect. 6.7), inaccurate analyses at the previous cycle, spin-up and spin-down effects (Sect. 4), or position errors in an otherwise good forecast (Sect. 6.1). The problem of allowing the DA to modify the background in these cases to match the observations can be difficult as it may require the model state to change regimes. This is particularly difficult when the background is clear, but the observations indicate cloud, and this is due to the 'zero gradient' problem [Errico et al., 2007a, Lopez, 2011, Geer et al., 2017]. In 4DVar where there are no explicit cloud or precipitation variables (Sect. 8.2), this problem happens because thermodynamic perturbations propagated through the linearised forecast model and/or observation operator in a cloud-free background state atmosphere cannot develop cloud increments (due to terms like $\partial q_c / \partial T$ being zero, where q_c is cloud condensate). This leaves no sensitivity to the cloudy observations, and hence no contribution to changes in the cost function from observations sensitive to cloud under such conditions. In ensemble schemes the equivalent issue manifests with all background ensemble members predicting clear air (i.e. a zero spread in cloud variables), meaning that cloud cannot be updated as the ensemble scheme wrongly considers the background error variance to be zero. A possible way forward in Var is to use explicit cloud and precipitation control variables with their corresponding background error statistics (see Sect. 8.2), and a possible way forward in ensemble schemes is to locally inflate the ensemble when the background is dry but the observations are cloudy. Minamide and Zhang [2019] used multiplicative inflation, although additive inflation would presumably be better when the ensemble spread is exactly zero, as otherwise the hydrometeor values cannot be updated.

7 Choice of vapour control parameter for Gaussian statistics

Control variables (CVs, $\boldsymbol{\chi}$) are the DA's representation of the state being estimated, which may – or may not – be the state $\mathbf{x}$ in the model's representation (i.e. winds, temperature, humidity variables, etc. on the model grid)[12]. The CVs for EnKF schemes are generally chosen to be the same as the model variables, but this need not be the case – a factor that has not been well explored. The CVs for pure and hybrid Var systems are generally different from $\mathbf{x}$, which mean that they minimise a cost function with respect to $\boldsymbol{\chi}$ (actually perturbations, $\delta\boldsymbol{\chi}$, in incremental schemes considered here). In general $\delta\boldsymbol{\chi}$ is related to increments in $\mathbf{x}$ ($\delta\mathbf{x}$) via a control variable transform (CVT, $\mathbf{U}$, see e.g. Bannister, 2008b), transforming $\delta\boldsymbol{\chi} \to \delta\mathbf{x}$, i.e. $\delta\mathbf{x} = \mathbf{U}\delta\boldsymbol{\chi}$. CVs are useful partly because in Var their background errors are assumed to be mutually uncorrelated and have unit variance (i.e. background error covariance $\mathbf{I}$), which means that the model-space $\mathbf{B}$-matrix does not need to be specified in the cost function when posed in terms of CVs, $J[\delta\boldsymbol{\chi}]$.

An intermediate step in the CVT involves *control parameters* ($\delta\tilde{\boldsymbol{\chi}}$), $\delta\boldsymbol{\chi} \to \delta\tilde{\boldsymbol{\chi}} \to \delta\mathbf{x}$, where $\delta\tilde{\boldsymbol{\chi}}$ comprises fields of physical parameters, one or more of which is related to AW variables. Importantly, zero cross-covariances are assumed between background errors of parameters (although each parameter can have spatially correlated errors, and these are dealt with by the remainder of the CVT). In this paper control parameters are denoted by a tilde.

One usually has choice over control parameters. Desirable properties include zero mutual correlation, and spatial homogeneity of their statistics, which approximate well to being Gaussian. Often control parameters are chosen in anticipation of minimal cross correlation, but the parameter choice representing WV is difficult as there is no known perfect choice. This section first describes the CVT method and then reviews the main WV parameters that have been used in operational Var schemes (although they could be utilised in Ens methods too). Most of this section concerns the vapour phase

---

[12]Sometimes augmented with other variables (such as the bias correction variables $\boldsymbol{\beta}$, see Sect. 6.7).

only, as we leave it to Sect. 8 to discuss other water species. There are also important choices of control parameter that account for some non-Gaussian aspects, such as logarithmic transforms, but we discuss non-Gaussianity separately in Sect. 9. Gustafsson et al. [2018] (their Table 4) provide a summary of the control variables used at a number of operational centres.

7.1 Background error covariance modelling in Var

CVTs – and hence the discussions in this section – are particularly important to make Var methods work. The relationship between $\delta\tilde{\boldsymbol{\chi}}$ and $\delta\mathbf{x}$ is called the parameter transform (or sometimes the balance operator), $\mathbf{U}_{\mathrm{p}}$:

$$\delta\mathbf{x} = \mathbf{S}^{-1}\mathbf{U}_{\mathrm{p}}\delta\tilde{\boldsymbol{\chi}}, \tag{6}$$

where $\mathbf{S}^{-1}$ is the inverse simplification operator, which increases the resolution of the fields from the analysis grid to the model grid. For simplicity this is set to $\mathbf{I}$ in this paper. The parameter transform is a powerful way of accounting for multivariate background errors of model variables, and potential improvements can be made by choosing new parameters and transforms to the standard set in a given system.

Apart from truncation differences, $\delta\mathbf{x}$, $\delta\tilde{\boldsymbol{\chi}}$, and $\delta\boldsymbol{\chi}$ are different representations of the same increment, but have different assumed statistical properties. Model space errors have covariance $\left\langle \delta\mathbf{x}\delta\mathbf{x}^{\mathrm{T}} \right\rangle_{\mathrm{b}} = \mathbf{B}$, control parameters have block diagonal covariances $\left\langle \delta\tilde{\boldsymbol{\chi}}\delta\tilde{\boldsymbol{\chi}}^{\mathrm{T}} \right\rangle_{\mathrm{b}} = \tilde{\mathbf{B}}$, and control variables are uncorrelated $\left\langle \delta\boldsymbol{\chi}\delta\boldsymbol{\chi}^{\mathrm{T}} \right\rangle_{\mathrm{b}} = \mathbf{I}$, where the 'b' subscript indicates expectation over background error samples. Posing the problem in terms of control parameters therefore results in a simpler form of $\mathbf{B}$, and implies background error covariances as follows:

$$\mathbf{B}^{\mathrm{ic}} = \mathbf{U}_{\mathrm{p}}\tilde{\mathbf{B}}\mathbf{U}_{\mathrm{p}}^{\mathrm{T}}. \tag{7}$$

The implied covariances (ic) are hoped to be (but not guaranteed to be) close to the true $\mathbf{B}$-matrix.

7.2 Balanced and unbalanced vapour parameters

$\mathbf{U}_{\mathrm{p}}$ often takes a block-lower-triangular form, e.g. in the Met Office [Lorenc et al., 2000], the ECMWF [Derber and Bouttier, 1999], Météo-France [Derber and Bouttier, 1999, Berre, 2000], and WRF [Chen et al., 2013b] systems. This form is convenient mathematically and allows for a physical interpretation of the control parameters, leading to the idea of *balanced* and *unbalanced* parameters. Consider the following system (simplified for illustration):

$$\left(\begin{array}{c} \delta\boldsymbol{\psi} \\ \delta\mathbf{p} \\ \delta\mathbf{q} \end{array} \right) = \left(\begin{array}{ccc} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{p\psi} & \mathbf{I} & \mathbf{0} \\ \mathbf{L}_{q\psi} & \mathbf{L}_{qp_{\mathrm{u}}} & \mathbf{I} \end{array} \right) \left(\begin{array}{c} \delta\tilde{\boldsymbol{\psi}} \\ \delta\tilde{\mathbf{p}}_{\mathrm{u}} \\ \delta\tilde{\mathbf{q}}_{\mathrm{u}} \end{array} \right). \tag{8}$$

This follows from (6) with $\mathbf{S}^{-1} = \mathbf{I}$, where the model variables are streamfunction ($\delta\boldsymbol{\psi}$), *full* pressure ($\delta\mathbf{p}$) and *full* SH ($\delta\mathbf{q}$) (see below for the significance of 'full'); and the control parameters are streamfunction ($\delta\tilde{\boldsymbol{\psi}}$), *unbalanced* pressure ($\delta\tilde{\mathbf{p}}_{\mathrm{u}}$) and *unbalanced* SH ($\delta\tilde{\mathbf{q}}_{\mathrm{u}}$). The parameters form a hierarchy, where the highest is $\delta\tilde{\boldsymbol{\psi}}$ and the lowest is $\delta\tilde{\mathbf{q}}_{\mathrm{u}}$. $\mathbf{L}_{p\psi}$, $\mathbf{L}_{q\psi}$, and $\mathbf{L}_{qp_{\mathrm{u}}}$ are blocks of the balance operator. These blocks may be diagonal (where the operators are local in nature) or off diagonal.

In (8) each model field is decomposed in terms of balanced and unbalanced parts whose sum is the full field (except $\delta\boldsymbol{\psi}$ which is assumed to have no unbalanced part[13]). The balanced component is the part of a field that is completely described by control parameters higher up the hierarchy, and the unbalanced component is the residual, i.e. the part that is uncorrelated with higher parameters. Line 2 of (8) for pressure is $\delta\mathbf{p} = \mathbf{L}_{p\psi}\delta\tilde{\boldsymbol{\psi}} + \delta\tilde{\mathbf{p}}_{\mathrm{u}}$ which is the sum of balanced and unbalanced parts

---

[13]This makes sense in synoptic-scale extra-tropical systems where the rotational wind is largely geostrophically balanced. In many DA schemes, geophysical balance operators are applied in the CVT, but the terms 'balanced' and 'unbalanced' have more general meanings.

respectively. Similarly line 3 for SH is $\delta\mathbf{q} = (\mathbf{L}_{q\psi}\delta\tilde{\psi} + \mathbf{L}_{qp_{\mathrm{u}}}\delta\tilde{\mathbf{p}}_{\mathrm{u}}) + \delta\tilde{\mathbf{q}}_{\mathrm{u}}$ where the balanced term is placed in parentheses. This partitioning of variables is relevant to any kind of AW variable, and not just for SH.

It is useful to note that in general we can denote the first control parameter as $\delta\tilde{\boldsymbol{\chi}}^{(1)}$ and the ith control parameter $(i > 1)$ as $\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}}^{(i)}$. The ith model variable, $\delta\mathbf{x}^{(i)}$:

$$\delta\mathbf{x}^{(i)} = \underbrace{\mathbf{L}_{i1}\delta\tilde{\boldsymbol{\chi}}^{(1)} + \sum_{j=2}^{i-1}\mathbf{L}_{ij}\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}}^{(j)} +}_{\text{balanced part of } \delta\mathbf{x}^{(i)}} \underbrace{\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}}^{(i)},}_{\text{unbalanced part of } \delta\mathbf{x}^{(i)}} \tag{9}$$

which corresponds to row i of $\mathbf{U}_{\mathrm{p}}$ in (8). By constructing an unbalanced parameter, $\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}}^{(i)}$, to represent the residual, it may be assumed to be uncorrelated with higher parameters (smaller i). This model of background error covariances allows intuitive and physically meaningful couplings to be represented in DA. Different centres tend to make different choices of (un)balanced parameters, and we now mention the most common control parameter choices that are associated with WV. Note that in many of the following examples the example control parameters are scalars rather than fields. This is done where possible for the purpose of simplicity.

7.3 Specific humidity

The *full* SH (i.e. purely *unbalanced*) has been used as a WV control parameter in the ECMWF, NCEP, Météo-France (ARPEGE), and HIRLAM DA systems. This choice assumes zero background error covariance with other parameters and so SH is effectively analysed independently. In unsaturated air, such as in downdrafts and in the stratosphere, SH is a conserved tracer, so the non-correlation property is plausible, making SH a favourable choice[14]. In saturated air though δq-δT correlations are likely to be large, due to condensation processes, making SH a potentially bad choice (Lorenc, 2007 for instance finds strong δq-δT correlations within the boundary layer). A consequence of choosing δq is that cloud can be created or annihilated inappropriately (Sect. 6.6). To see this mathematically, consider the relationship between SH (q) and RH (μ): $q = \mu q_{\mathrm{s}}(T, p)$, where q_{s} is the saturated SH which depends on temperature, T, and pressure, p (see also footnote 10). Linearisation about the background provides a relationship between $\delta\mu$, δq, δT, and δp:

$$\delta q = \mu^{\mathrm{b}}\left.\frac{\partial q_{\mathrm{s}}}{\partial T}\right|_{\mathrm{b}}\delta T + \mu^{\mathrm{b}}\left.\frac{\partial q_{\mathrm{s}}}{\partial p}\right|_{\mathrm{b}}\delta p + q_{\mathrm{s}}(T^{\mathrm{b}}, p^{\mathrm{b}})\delta\mu, \tag{10a}$$

$$= \eta\delta T + \xi\delta p + \epsilon\delta\mu, \tag{10b}$$

where η, ξ, and ϵ in (10b) are shorthand for the factors in (10a). Choosing δq as a control parameter, means that we can in this section write $\delta\tilde{q} = \delta q$. In the absence of an explicit cloud parameter in the DA, a criterion for cloud is $\mu \gtrsim 1$, so any change of μ around unity may affect the implied presence of cloud. From (10) and ignoring pressure (since its effect is relatively small), the change in RH, $\delta\mu$, given δq and δT is $\delta\mu = (1/\epsilon)\delta q - (\eta/\epsilon)\delta T$.

This highlights a potential problem with the parameter SH. Consider a background state with $\mu^{\mathrm{b}} \approx 1$ and an observation of temperature that leads to an increment $\delta T^{\mathrm{a}} > 0$ (superscript 'a' refers to analysis increment). For no WV observations, $\delta q^{\mathrm{a}} = 0$. The implied change in RH is then $\delta\mu^{\mathrm{a}} = -(\eta/\epsilon)\delta T^{\mathrm{a}}$, which can push the RH below saturation, thus destroying cloud. This was a problem in the Met Office's pre-Var system where temperature and SH were considered separately [Lorenc et al., 1996]. This system had a cold bias, and so the generally positive temperature increments systematically reduced RH (and hence cloud). The opposite case of cloud creation can happen when $\mu^{\mathrm{b}} \lesssim 1$ but $\delta T^{\mathrm{a}} < 0$. It is a reasonable question to ask whether non-WV observations should influence the amount of cloud in the system.

---

[14]This is a Lagrangian conservation property, which does not guarantee Eulerian non-correlation though. For instance downdrafts may bring simultaneous variations δq and δT at a fixed point in space. This would appear as a correlation even though they do not affect each other directly.

7.4 Relative humidity

The *full* RH (i.e. purely *unbalanced*) has been used as a control parameter, $\delta\tilde{\mu}$, in the Met Office [Lorenc et al., 2000] and NCAR [Barker et al., 2004] DA systems. This choice assumes that $\delta\tilde{\mu}$ is uncorrelated with background errors of other parameters. To understand the properties of the choice $\delta\tilde{\mu}$, consider a simple model state $\delta\mathbf{x} = \begin{pmatrix} \delta T & \delta q \end{pmatrix}^{\mathrm{T}}$ and the control parameter choice $\delta\tilde{\chi} = \begin{pmatrix} \delta\tilde{T} & \delta\tilde{\mu} \end{pmatrix}^{\mathrm{T}}$. The CVT is constructed as follows, where the last line uses the relationship in (10) (neglecting δp):

$$
\begin{pmatrix} \delta T \\ \delta q \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \eta & \epsilon \end{pmatrix} \begin{pmatrix} \delta\tilde{T} \\ \delta\tilde{\mu} \end{pmatrix}, \tag{11}
$$

The implied **B**-matrix is found from (7):

$$
\mathbf{B} = \begin{pmatrix} 1 & 0 \\ \eta & \epsilon \end{pmatrix} \begin{pmatrix} \sigma_T^{\mathrm{b}\,2} & 0 \\ 0 & \sigma_\mu^{\mathrm{b}\,2} \end{pmatrix} \begin{pmatrix} 1 & \eta \\ 0 & \epsilon \end{pmatrix} = \begin{pmatrix} \sigma_T^{\mathrm{b}\,2} & \sigma_T^{\mathrm{b}\,2}\eta \\ \sigma_T^{\mathrm{b}\,2}\eta & \sigma_q^{\mathrm{b}\,2} \end{pmatrix}, \tag{12}
$$

where $\sigma_T^{\mathrm{b}\,2}$ and $\sigma_\mu^{\mathrm{b}\,2}$ are the background error variances of $\delta\tilde{T}$ and $\delta\tilde{\mu}$ respectively, and $\sigma_q^{\mathrm{b}\,2} = \sigma_T^{\mathrm{b}\,2}\eta^2 + \sigma_\mu^{\mathrm{b}\,2}\epsilon^2$ is the background error variance of SH that is implied by using RH as a control parameter. Note that $\eta\delta\tilde{T}$ and $\epsilon\delta\tilde{\mu}$ in (11) may also be interpreted, respectively, as *balanced* and *unbalanced* SH contributions (Sect. 7.2).

Consider the analysis increment, $\delta\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}}$, using the update equation – see (2) – when (i) only T is observed and (ii) when only q is observed. In case (i) (observation T^{o}, error variance $\sigma_T^{\mathrm{o}\,2}$, and $\mathbf{H} = \begin{pmatrix} 1 & 0 \end{pmatrix}$), $\delta\mathbf{x}^{\mathrm{a}}$ is:

$$
\begin{pmatrix} \delta T^{\mathrm{a}} \\ \delta q^{\mathrm{a}} \end{pmatrix} = \begin{pmatrix} \sigma_T^{\mathrm{b}\,2} \\ \sigma_T^{\mathrm{b}\,2}\eta \end{pmatrix} \frac{T^{\mathrm{o}} - T^{\mathrm{b}}}{\sigma_T^{\mathrm{b}\,2} + \sigma_T^{\mathrm{o}\,2}}. \tag{13}
$$

Substituting these analysis increments into (10) (and $\delta p = 0$) results in $\delta\mu^{\mathrm{a}} = 0$, thus preserving cloud. In case (ii) (observation q^{o}, with error variance $\sigma_q^{\mathrm{o}2}$, and $\mathbf{H} = \begin{pmatrix} 0 & 1 \end{pmatrix}$), $\delta\mathbf{x}^{\mathrm{a}}$ is:

$$
\begin{pmatrix} \delta T^{\mathrm{a}} \\ \delta q^{\mathrm{a}} \end{pmatrix} = \begin{pmatrix} \sigma_T^{\mathrm{b}\,2}\eta \\ \sigma_q^{\mathrm{b}\,2} \end{pmatrix} \frac{q^{\mathrm{o}} - q^{\mathrm{b}}}{\sigma_q^{\mathrm{b}\,2} + \sigma_q^{\mathrm{o}\,2}}, \tag{14}
$$

which results in non-zero $\delta\mu^{\mathrm{a}}$. Case (i) shows that SH is modified even when no WV observations are made, which may be a desirable property in some situations, but not others (see Sect. 6.6 and Lorenc et al., 1996). Dee and Da Silva [2003] discuss a problem when there are no WV observations in a system that has a cold bias in the background's stratosphere in a DA system that does couple temperature and SH fields to conserve RH in the way described above. Assimilation increments of temperature tend to be positive, but the correlations act to increase SH in the stratosphere, and hence to the anomalous accumulation of WV there. This would not happen using an SH control parameter (Sect. 7.3) as SH and temperature have no implied correlation in that case.

7.5 Pseudo relative humidity

Pseudo relative humidity (PRH, $\delta\tilde{\mu}_{\mathrm{p}}$) was proposed by Dee and Da Silva [2003] to help deal with some of the difficulties found with SH and RH. It has been used in the HIRLAM [Gustafsson et al., 2011], WRF [Barker et al., 2004, Schwartz et al., 2013], and JMA [Gustafsson et al., 2018] DA systems. It may be defined as $\tilde{\mu}_{\mathrm{p}} = q/q_{\mathrm{s}}(T^{\mathrm{b}}, p^{\mathrm{b}})$, which has linearisation $\delta\tilde{\mu}_{\mathrm{p}} = \delta q/q_{\mathrm{s}}(T^{\mathrm{b}}, p^{\mathrm{b}})$ (in Dee and Da Silva [2003] it is defined with mixing ratios rather than SHs). The difference between PRH and RH is that here the q_s is evaluated at the background, and so is constant as far as the DA is concerned, making PRH a scaled and background-dependent SH-like parameter. The choice $\delta\tilde{\mu}_{\mathrm{p}}$ avoids the problem with the large variability of SH (Sect. 6.2), and the often anomalous coupling between δT and δq found when $\delta\tilde{\mu}$ is chosen (Sect. 7.4). PRH though does not automatically account for the coupling between δT and δq when coupling is appropriate.

7.6 Adaptive humidity parameter

As discussed, it seems plausible to assume that when there are no WV observations, RH should be conserved when close to saturation, but not when the air is substantially subsaturated. Neither SH nor RH satisfy both of these assumptions. Ingleby et al. [2013b] have developed a parameter that has many of these properties, which here is called the adaptive transform. There is also a nonlinear aspect in their implemented system (intended to help tackle non-Gaussianity) and is discussed in Sect. 9.4, but here we look only at the adaptive property to deal with above issues.

The simplified version of this transform shown here introduces the control parameter $\delta\tilde{\mu}_A$, which has CVT[15]:

$$\delta q = h\eta\delta T + \epsilon\delta\tilde{\mu}_A. \tag{15}$$

This transform is based on (10) with $\delta\mu \rightarrow \delta\tilde{\mu}_A$, $\delta p \rightarrow 0$, and the introduction of the coefficient h, which provides the adaptive property, namely that if $h = 0$, $\delta\tilde{\mu}_A$ is a scaled SH-like parameter (Sect. 7.3), and if $h = 1$ it is an RH-like parameter (Sect. 7.4). $\epsilon\delta\tilde{\mu}_A$ in (15) may be interpreted as an *unbalanced* SH parameter (Sect. (7.2)).

In Ingleby et al. [2013b] h comes from look-up tables, derived from training samples, which allows it to vary with a reference RH and vertical model level. For instance, h was found to be close to zero in the stratosphere, where $\delta\tilde{\mu}_A$ behaves SH-like. Although h was found to never reach unity, it increases in the troposphere for RH greater than ~ 0.8, where $\delta\tilde{\mu}_A$ becomes more RH-like[16]. This behavior of h is consistent with the arguments made in Sects. 7.3 and 7.4. For lower RH values though h goes negative and it is not clear how to interpret this in terms of SH-like or RH-like behavior.

7.7 Non-local, dynamical, and statistical parameter transforms

CVTs may be defined more generally than those shown in Sects. 7.3-7.6, e.g. by coupling information from many grid-points. Such non-local transforms for instance are used to define background error relationships between mass and wind fields (e.g. $\mathbf{L}_{p\psi}$ in (8)). This can be achieved in practice using either dynamically- or statistically-based parameter transforms, which can sometimes be extended to include moist processes.

7.7.1 Dynamical transforms

Some centres that run Var systems use dynamically-based balance operators to model mass-wind covariances. The Met Office, e.g., solves a linear balance equation [Lorenc et al., 2000], and the ECMWF and MM5 solve (linearised) nonlinear balance equations [Fisher, 2003, Barker et al., 2004]. Dynamical balance operators are naturally flow-dependent, but it is unclear how to determine an analogous operator that accounts for moist processes.

7.7.2 Statistical regressions

Systems that already use statistical regressions for AW covariances include the ALADIN and AROME [Berre, 2000], WRF [Chen et al., 2013b], and HIRLAM [Gustafsson et al., 2011] systems. Statistical regressions need to be calibrated off-line and flow-dependency has to be added with look-up tables or masks (Sect. 7.7.3), but it is generally possible to relate any two variables using statistical regressions. This approach is therefore a promising means of relating a chosen WV parameter to other variables in a physically and statistically meaningful way, instead of using a local transform based on (10).

The regressions can be efficiently formulated by regressing in the vertical only, where, depending on the design of the CVT, they can vary with horizontal location (as in Chen et al., 2013b) or with

---

[15]For reasons of simplicity and consistency with the rest of this section, there are differences between (15) and the transform of Ingleby et al. [2013b]. We use δT instead of $\delta\theta$, and δq instead of δq_T (δq_T is total water, which includes contributions from cloud water and/or ice).

[16]These are the results of Ingleby et al. [2013b] for δq_T and $\delta\theta$ parameters, but we might expect qualitatively similar results for δq and δT.

horizontal scale (as in Berre, 2000). Chen et al. [2013b] for instance produce model RH increments using

$$\delta\boldsymbol{\mu} = \mathbf{L}_{\mu\psi}\delta\tilde{\boldsymbol{\psi}} + \mathbf{L}_{\mu\chi_{\mathrm{u}}}\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}} + \mathbf{L}_{\mu T_{\mathrm{u}}}\delta\tilde{\mathbf{T}}_{\mathrm{u}} + \mathbf{L}_{\mu\mathrm{ps}_{\mathrm{u}}}\delta\tilde{\mathbf{ps}}_{\mathrm{u}} + \delta\tilde{\boldsymbol{\mu}}_{\mathrm{u}}, \tag{16}$$

where $\delta\tilde{\boldsymbol{\psi}}$ is streamfunction, $\delta\tilde{\boldsymbol{\chi}}_{\mathrm{u}}^{\mathrm{vpot}}$ is unbalanced velocity potential, $\delta\tilde{\mathbf{T}}_{\mathrm{u}}$ is unbalanced temperature, $\delta\tilde{\mathbf{ps}}_{\mathrm{u}}$ is unbalanced surface pressure, and $\delta\tilde{\boldsymbol{\mu}}_{\mathrm{u}}$ is unbalanced RH. The key aspect here is that we are dealing with field columns instead of individual gridpoints. Equation (16) is one row of the whole $\mathbf{U}_{\mathrm{p}}$ operator (and this example deals with more fields than considered so far in this paper). Extensions may be made for other parameters including hydrometeors (Sect. 8). The standard regression formula that gives the balance operators is:

$$\begin{pmatrix} \mathbf{L}_{\mu\psi} & \mathbf{L}_{\mu\chi_{\mathrm{u}}^{\mathrm{vpot}}} & \mathbf{L}_{\mu T_{\mathrm{u}}} & \mathbf{L}_{\mu\mathrm{ps}_{\mathrm{u}}} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{\mu\psi} & \mathbf{C}_{\mu\chi_{\mathrm{u}}^{\mathrm{vpot}}} & \mathbf{C}_{\mu T_{\mathrm{u}}} & \mathbf{C}_{\mu\mathrm{ps}_{\mathrm{u}}} \end{pmatrix} \times$$
$$\begin{pmatrix} \mathbf{C}_{\psi\psi} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\chi_{\mathrm{u}}^{\mathrm{vpot}}\chi_{\mathrm{u}}^{\mathrm{vpot}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_{T_{\mathrm{u}}T_{\mathrm{u}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{\mathrm{ps}_{\mathrm{u}}\mathrm{ps}_{\mathrm{u}}} \end{pmatrix}^{-1}, \tag{17}$$

where $\mathbf{C}_{ij}$ is the covariance matrix between parameters i and j, calculated from training data. Usually regressions are formulated in the vertical direction only, so the matrices are of a manageable size. If each covariance matrix is square then the expression for each $\mathbf{L}_{ij}$ is $\mathbf{L}_{ij} = \mathbf{C}_{ij}\mathbf{C}_{jj}^{-1}$ (see analogous expression in Michel et al., 2011).

7.7.3   State-dependent statistical regressions

We have already highlighted that error covariances are flow-dependent – especially for those involving AW – yet the covariances implied by (17) are fixed by the training data. The relationship between WV and temperature ($\mathbf{L}_{\mu T_{\mathrm{u}}}$) may be expected to depend strongly on the closeness to saturation, and on the presence of rainfall. Similarly, the mass-wind relationships that appear in other parts of the transform (e.g. $\mathbf{L}_{T\psi}$, $\mathbf{L}_{\mathrm{ps}\psi}$, not shown in (16)), may normally be close to geostrophic or hydrostatic balance, but not where there is strong latent heat release or precipitation [Errico et al., 2007b, Pagé et al., 2007, Caron and Fillion, 2010]. Statistical regression schemes like those described in Sect. 7.7.2 may gain flow-dependence by changing the covariance values used depending on the state using look-up tables.

In order to facilitate flow-dependency of the covariances, studies have tried to classify model regions into different categories, e.g. precipitating and non-precipitating regions [Montmerle and Berre, 2010, Montmerle, 2012], regions of different precipitation intensities [Michel et al., 2011], or foggy and clear regions [Ménétrier and Montmerle, 2011]. To introduce flow-dependency, the CVT can be written in the following way:

$$\delta\mathbf{x} = \begin{pmatrix} \mathbf{F}_1^{1/2}\mathbf{U}_1 & \cdots & \mathbf{F}_{N_{\mathrm{c}}}^{1/2}\mathbf{U}_{N_c} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{\chi}_1 \\ \vdots \\ \delta\boldsymbol{\chi}_{N_{\mathrm{c}}} \end{pmatrix}, \tag{18}$$

where $\mathbf{U}_i$ is the part of the CVT for the ith category (of N_{c} categories), $\delta\boldsymbol{\chi}_i$ is its control vector, and $\mathbf{F}_i^{1/2}$ is the geographical mask for that category (see below). Each $\mathbf{U}_i$ itself comprises a separate CVT (e.g. one that has rows of the form (16), but now with regressions that are trained for category i), and each $\delta\boldsymbol{\chi}_i$ comprises the associated balanced and unbalanced subvectors. Such a scheme can obviously become very complex. Examples of categories are heavy, light and no precipitation ($N_{\mathrm{c}} = 3$) or just precipitating and non-precipitating regions ($N_{\mathrm{c}} = 2$).

If $\mathbf{U}_i$ is modelled in physical space then each $\mathbf{F}_i^{1/2}$ is a diagonal matrix, whose values may be 1 when the category is relevant for that position, and 0 otherwise (or perhaps more pragmatically as smoothed masks, subject to $\sum_{i=1}^{N_{\mathrm{c}}} \left(\mathbf{F}_i^{1/2}\right)_{jj} = 1$ for each position index j, where $j \leq n$, n being the number of grid points). The references mentioned above model each $\mathbf{U}_i$ in spectral space, where $\mathbf{F}_i^{1/2}$
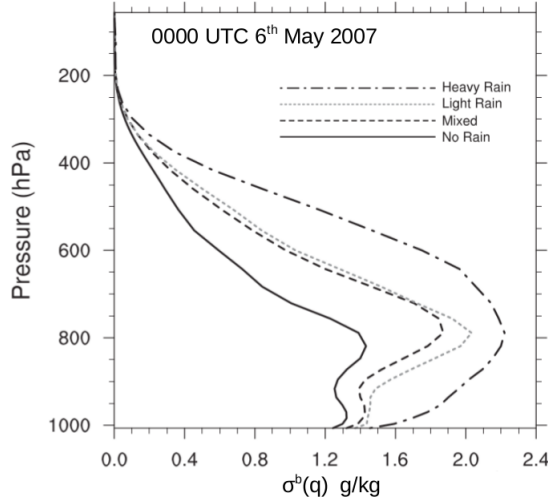
Figure 6: Background error standard deviations of specific humidity as a function of pressure level diagnosed from 30-member ensembles of 6-hour 3km WRF forecasts over the central United States for multi-cellular storm conditions. The statistics have been divided into locations that have the precipitation class indicated by the line type. ©American Meteorological Society. Used with permission. Taken from Michel et al. [2011].

is no longer diagonal, but has the form $\mathbf{F}_i^{1/2} = \mathbf{S}_\mathrm{F} \mathbf{D}_i^{1/2} \mathbf{S}_\mathrm{F}^{-1}$, where $\mathbf{S}_\mathrm{F}$ is the Fourier transform operator and $\mathbf{D}_i^{1/2}$ now takes the role of the mask in physical space. Each $\mathbf{U}_i$ is tuned for its particular category.

To illustrate how the statistics represented by the $\mathbf{U}_i$ can change with precipitation strength, Fig. 6 shows vertical profiles of background error standard deviations of q for different classes of precipitation (see Michel et al. [2011] for details). It is clear is that the standard deviations show a broad range of values for the different classes, with the no rain case generally having the smallest values. There is also great variability for other aspects of $\mathbf{U}_i$ in Michel et al.'s study (and other variables), namely horizontal and vertical correlation structures, and multivariate correlations.

Some interesting findings from Montmerle and Berre [2010] are that precipitating regions show a large increase in the degree of δq-$\delta \eta$ coupling (where η is the divergence), a large increase in wind standard deviations, a decrease in δq length-scales, and a decrease in δq standard deviations compared to non-precipitating regions. These results are perhaps not surprising (the last one may be explained by the background's q being tied to q_s during precipitation and so has less variability), and they highlight that a flow-dependent formulation like (18) is important. Montmerle [2012] in particular showed that such a formulation gave improved forecasts, including better alignment of convective cells, thought to be due to the inclusion of the enhanced δq-$\delta \eta$ covariance relationships, as though dynamical information (e.g. from Doppler radial winds) was being simultaneously assimilated. Chung et al. [2013] however warn that the local statistics may change some time before the precipitation actually starts due to the activation of some physical processes. This suggests that it may be beneficial to design a mask according to another metric such as the convective available potential energy (CAPE), or the WV flux convergence as harbingers of precipitation rather than precipitation itself.

8 Beyond water vapour

As discussed in Sect. 6.8, knowledge of all water species in the DA is useful for providing initial conditions for forecasting, and for interpreting observations like satellite radiances and radar reflectivity (e.g. Sun et al., 2014). Introducing multiple water species in DA is conceptually straightforward in schemes like the EnKF, although the usual sampling noise issues exist. In Var schemes, introducing

33

multiple water species raises new issues that are discussed here.

Consider observations that are sensitive to cloud condensate or precipitation, like from a radar instrument. Most forecast models can forecast separate water species, e.g. WV, cloud condensate and precipitation, which can be used in the background innovation calculations. Some DA systems though simplify the representation of these quantities. In incremental 4DVar for instance the linear model often simplifies the representation of the moist phases. In Met Office 4DVar, the control parameters calculate 'total water' δq_{T}, using a total water form of (15), which the linear model propagates ($q_{\mathrm{T}} = q + q_{\mathrm{c}}$ is total water – WV plus cloud condensate – and $\delta q_{\mathrm{T}} = \delta q + \delta q_{\mathrm{c}}$ is its increment) [Migliorini et al., 2017]. In this case, δq_{c} and precipitation increments need to be diagnosed from δq_{T} (and other fields) using linear parametrisations to update the inputs to the observation operators. This is a *diagnostic approach*[17]. Alternatively a DA system may have separate vapour/condensate/precipitation parameters. This is an *explicit approach* [Gong and Hólm, 2011]. Each approach has its pros and cons. Figure 7 shows various options (discussed below) for facilitating cloud-affected radiance assimilation in a generic incremental 4DVar scheme, showing the forward and adjoint steps (similar options exist for precipitation, which may also be diagnosed or represented explicitly, see Sect. 8.4).

8.1 The diagnostic cloud approach

We describe two options for doing 4DVar with a total water control parameter using the diagnostic approach, and both of these options use a partitioning operator. The partitioning operator is also sometimes called an incrementing operator [Smith, 1990, Migliorini et al., 2017], and separates δq_{T} into WV δq and condensate δq_{c}. Variations exist among other centres: the MM5/WRF systems diagnose WV, cloud, and rain from δq_{T} using 'warm rain' process assumptions (Xiao et al., 2007, Xiao and Sun, 2007, although this may need to be generalised for strong or deep convection where ice phases of water are prominent), and the ECMWF diagnoses cloud from δq using simplified versions of the stratiform and convective cloud schemes of their model [Tompkins and Janisková, 2004, Lopez and Moreau, 2005]. Temperature is also needed in the partitioning. Panels a and b in Fig. 7 show the flow of information in a generic 4DVar system assimilating a cloud-sensitive radiance at time t, $L^{\mathrm{ob}}(t)$ (the principles apply equally for other types of measurement). The two options that we discuss are as follows.

8.1.1 Option 1: If q_{c} processes are not included in the linear model

Figure 7a is relevant when micro-physics is not included in the linear model of 4DVar, meaning that δq_{T} is propagated and then partitioned to update inputs to the observation operators. The background trajectory (top line of panel a) uses the nonlinear model for the full fields, which does propagate the separate water phases, to compute the background's radiance, $L^{\mathrm{b}}(t)$. Inside 4DVar the perturbation δq_{T} is propagated and then partitioned at the observation time to produce a perturbation to $L^{\mathrm{b}}(t)$, called $\delta L(t)$ (second line of panel a). The sensitivity vector $\delta \hat{L}(t) = -\mathbf{R}^{-1} \left(L^{\mathrm{ob}}(t) - L(t) \right)$, where $L(t) = L^{\mathrm{b}}(t) + \delta L(t)$, is formed and then propagated back to $t = 0$ via the adjoints of the linearised radiative transfer, partitioning operator, and forecast models. This produces this observation's contribution to the cost function's gradient at $t = 0$ (third line of panel a), $\delta \hat{q}_{\mathrm{T}}(0) = \partial J^{\mathrm{ob}} / \partial \delta q_{\mathrm{T}}(0)$, which is needed to perform an iteration of the minimisation. This is the standard adjoint method in 4DVar.

8.1.2 Option 2: If q_{c} processes are included in the linear model

Figure 7b is relevant when micro-physics is included in the linear model, meaning that δq and δq_{c} are propagated separately by linearised processes. In this option the partitioning operator is needed only at $t = 0$. Linearisation of the strongly nonlinear micro-physics scheme introduces errors though, which

---

[17]Note that some centres use just δq rather than δq_{T} to diagnose cloud. It is also possible to partition δq_{c} into liquid and ice, but here for simplicity we keep these together as a single cloud condensate increment.
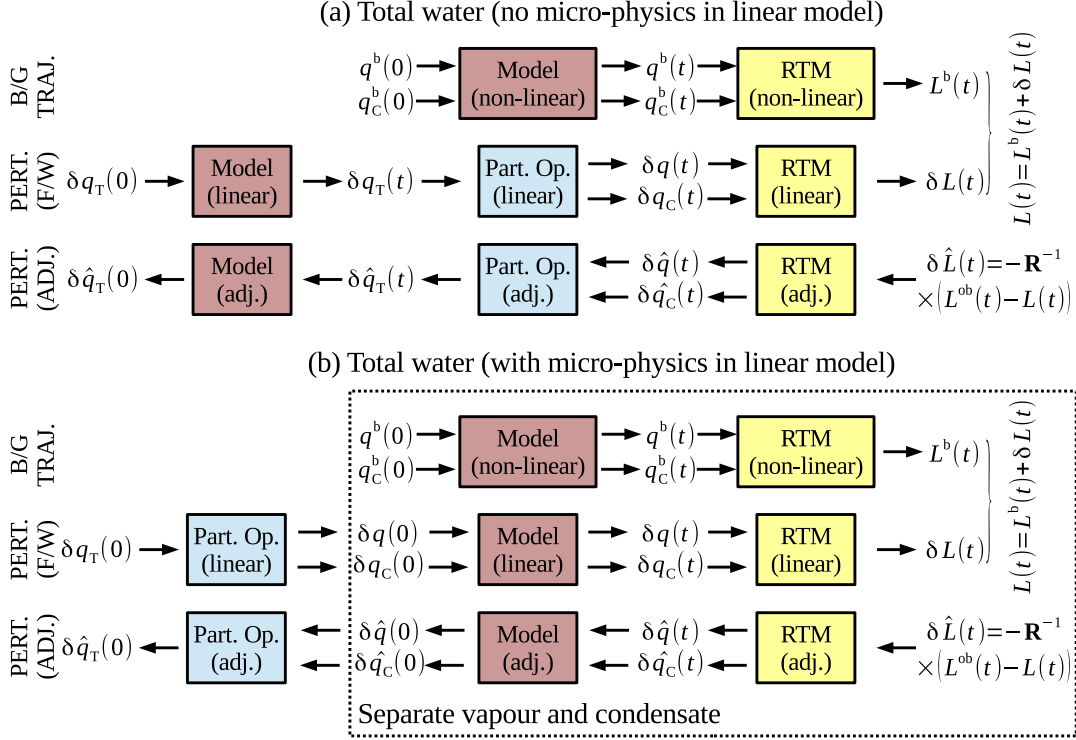
(a) Total water (no micro-physics in linear model)

B/G TRAJ.

$q^{\mathrm{b}}(0)$ → Model (non-linear) → $q^{\mathrm{b}}(t)$ → RTM (non-linear) → $L^{\mathrm{b}}(t)$
$q_{\mathrm{C}}^{\mathrm{b}}(0)$ → → $q_{\mathrm{C}}^{\mathrm{b}}(t)$ →

PERT. (F/W)

$\delta q_{\mathrm{T}}(0)$ → Model (linear) → $\delta q_{\mathrm{T}}(t)$ → Part. Op. (linear) → $\delta q(t)$ → RTM (linear) → $\delta L(t)$
→ $\delta q_{\mathrm{C}}(t)$ →

PERT. (ADJ.)

$\delta \hat{q}_{\mathrm{T}}(0)$ ← Model (adj.) ← $\delta \hat{q}_{\mathrm{T}}(t)$ ← Part. Op. (adj.) ← $\delta \hat{q}(t)$ ← RTM (adj.) ← $\delta \hat{L}(t) = -\mathbf{R}^{-1}$
← $\delta \hat{q}_{\mathrm{C}}(t)$ ← $\times \big(L^{\mathrm{ob}}(t) - L(t) \big)$

$L(t) = L^{\mathrm{b}}(t) + \delta L(t)$

(b) Total water (with micro-physics in linear model)

B/G TRAJ.

$q^{\mathrm{b}}(0)$ → Model (non-linear) → $q^{\mathrm{b}}(t)$ → RTM (non-linear) → $L^{\mathrm{b}}(t)$
$q_{\mathrm{C}}^{\mathrm{b}}(0)$ → → $q_{\mathrm{C}}^{\mathrm{b}}(t)$ →

PERT. (F/W)

$\delta q_{\mathrm{T}}(0)$ → Part. Op. (linear) → $\delta q(0)$ → Model (linear) → $\delta q(t)$ → RTM (linear) → $\delta L(t)$
→ $\delta q_{\mathrm{C}}(0)$ → → $\delta q_{\mathrm{C}}(t)$ →

PERT. (ADJ.)

$\delta \hat{q}_{\mathrm{T}}(0)$ ← Part. Op. (adj.) ← $\delta \hat{q}(0)$ ← Model (adj.) ← $\delta \hat{q}(t)$ ← RTM (adj.) ← $\delta \hat{L}(t) = -\mathbf{R}^{-1}$
← $\delta \hat{q}_{\mathrm{C}}(0)$ ← ← $\delta \hat{q}_{\mathrm{C}}(t)$ ← $\times \big(L^{\mathrm{ob}}(t) - L(t) \big)$

Separate vapour and condensate

$L(t) = L^{\mathrm{b}}(t) + \delta L(t)$

Figure 7: Flow of information within an incremental 4DVar system when a cloud-sensitive radiance measurement is made at time t, depending on the use (or not) of linear micro-physics within 4DVar. We assume that the nonlinear model has a full microphysics scheme. In (a) 4DVar has a δq_{T} (total water) control parameter but no micro-physics scheme (δq_{T} is propagated by the linear model and δq (vapour) and δq_{c} (cloud) are diagnosed from δq_{T} at each time). In (b) 4DVar has a total water control parameter but has micro-physics in the linear model (δq and δq_{c} are diagnosed at the beginning, and propagated separately). The dotted line encloses the case when δq and δq_{c} are separate control parameters (δq_{T} is then not required and so the steps outside of this box are ignored). The following abbreviations/symbols are used: 'Pert.' (perturbation), 'F/W' (forward run), 'Adj.' (adjoint), 'Traj.' (trajectory), 'Part. Op.' (partitioning operator, separating δq_{T} into δq and δq_{c}), 'RTM' (radiative transfer model), L (radiance), 'b' and 'B/G' (background), 'ob' (observation), and $\mathbf{R}$ (observation error covariance for radiance observations). Variables with a hat are adjoint variables, i.e. $\delta \hat{L}(t) = \partial J^{\mathrm{ob}} / \partial \delta L(t)$ and $\delta \hat{x} = \partial J^{\mathrm{ob}} / \partial \delta x$, where J^{ob} is the radiance observation term of the cost function).

can lead to 4DVar convergence problems and to 'contradictory regions' in phase space[18], which can in turn lead to an incorrect solution being reached [Xu, 1996, Fabry and Sun, 2010]. Furthermore, the act of linearising cloud and precipitation processes requires a significant development effort, e.g. Stiller and Ballard [2009].

In both options above, the partitioning operator is used at some point, which has advantages. The first is that error covariances of only δq_T need to be modelled. The second is that background covariances of δq_T are expected to be smooth compared to those of δq_c, so using δq_T as a control parameter rather than using separate phases can avoid complications due to discontinuities in the cloud field. There are also some notable disadvantages. One is that the partitioning operator is approximate and needs closure assumptions (e.g. the distribution of sub-grid water, see Smith [1990], Migliorini et al. [2017]). Another disadvantage, which has been mentioned already in Sect. 6.9, appears when the partitioning operator does not diagnose any condensate (especially possible if the background is cloud free). This means that observations that do detect cloud in such circumstances will have no ability to add cloud to the system since the DA has zero sensitivity of cloud condensate amount to changes in δq_T [Renshaw and Francis, 2011, Geer et al., 2017]. This is a major problem, which also arises in ensemble schemes and for radar reflectivity assimilation when the background is rain free, but the radar detects rain.

8.2 The explicit cloud approach

A third option uses separate WV and condensate control parameters throughout and does not use δq_T nor a partitioning operator (Fig. 7b only inside the dashed line). This avoids errors introduced by the partitioning operator, and the 'no background cloud' problem mentioned above. However, it shifts emphasis to providing a linear microphysics scheme and realistic background error covariances of the separate species (including cross correlations). The latter issue is complicated by the discontinuous nature of such fields and so the explicit cloud approach has not been advocated by some studies, e.g. Xiao and Sun [2007]. This issue has been studied by examination of single observation impacts (Gong and Hólm [2011], Wang et al. [2013], who reported physically reasonable increments), with real data trials (Wang et al. [2013], who reported improved precipitation forecasts), and/or with the implied covariances (Sect. 7.1) using the extra control parameters. The above issues do not need to be considered with ensemble methods as these are usually implemented in terms of the model variables themselves (see Sect. 8.5).

8.3 A combined diagnostic and explicit cloud approach

A way of combining the strengths of the diagnostic and explicit cloud approaches described in this section is to use the ideas of balanced and unbalanced parameters (Sect. 7.2). In this framework, the 'balanced' cloud condensate increment might be that diagnosed from δq_T (or whatever control parameter is used in a system) and other dynamic and thermodynamic control parameters of the system. The 'unbalanced' cloud condensate would come from a separate control parameter representing the residual condensate needed to match observations. This is essentially the procedure of Michel et al. [2011], and could help alleviate the 'zero gradient' problem mentioned in Sect. 6.9 where a DA system otherwise struggles to add cloud to a clear background state to match cloudy observations, especially when the background error covariances are enhanced with ensemble-derived information (e.g. Sect. 5.4).

It may be possible in this framework to also deal with the discontinuous nature of cloud. This could be done by modulating the climatologically determined background error standard deviation, σ_{cu}^b, of the associated unbalanced cloud parameter, with information from the diagnosed cloud (not the increment of diagnosed cloud, but the actual diagnosed cloud) – e.g. by forcing σ_{cu}^b to zero (or close

---

[18]These contradictory regions are where non-linearity is so strong that small increments to the background initial conditions that slightly improve the fit to observations do not provide useful information about the direction of the increments needed to fully minimise the cost function.

to zero) whenever the diagnosed cloud is zero. This would mean that an unbalanced cloud increment made in one location would not anomalously spread to neighbouring cloudless regions. The adoption of a log-normal unbalanced control parameter (see Sect. 9.2) could also deal with this problem differently as log-normal variables represent relative errors rather than additive ones, meaning that cloud could not be incremented from zero [Michel et al., 2011].

8.4   Precipitating hydrometeors

The above discussions about cloud condensate could also be applied to precipitating hydrometeors (e.g. rain, snow, hail and graupel, Michel et al., 2011, Xiao and Sun, 2007, Lopez, 2011). This is an essential step to facilitate the direct assimilation of radar reflectivity, and precipitation affected MW channels (see e.g. Guerbette et al. [2016] which shows brightness temperature Jacobians for a number of MW channels for snow and rain during a tropical cyclone in the Indian Ocean).

By choosing to use explicit precipitating hydrometeor control parameters in Var schemes, the extra background error covariances will need to be modelled. These will likely be flow dependent and distinct for each hydrometeor (see e.g. Amerault and Zou [2006] for data from hurricanes in the Atlantic Ocean), but it is not yet known how important accounting for flow-dependency is. Multivariate relationships within the **B**-matrix may be important to tackle moisture imbalances if the precipitation rates in the initial conditions determined by DA differ greatly from those rates that are natural for the model.

8.5   Representation of cloud/precipitation covariances with ensembles

As stated before, covariances do not need to be modelled when using Ens methods like the EnKF (Sect. 5.2), or EnVar (Sect. 5.3), and this includes covariances of the extra hydrometeors. A pioneering application of the EnKF in a convective scale model with a condensate variable was Snyder and Zhang [2003] which demonstrated the effectiveness of the EnKF with a rain water variable and simulated observations of radar radial wind and reflectivity. More recent examples that use simulated data include Otkin [2010] and Chambon et al. [2014], which are both WRF-based and include cloud and precipitating hydrometeors. Recent examples that use actual observations include Dowell et al. [2011] (WRF 1 km grid, 50 members), Yussouf et al. [2013] (WRF, 2 km grid, 45 members), Johnson et al. [2015] (WRF, 4 km grid, 40 members), Keclik et al. [2017] (WRF, 3 km grid, 40 members), Zhang et al. [2017] (WRF, 9 km grid, 32 members), Honda et al. [2018] (SCALE, 3 km grid, 50 members), and Minamide and Zhang [2018] (WRF, 3 km grid, 60 members). Again, these studies include cloud and precipitating hydrometeors.

Expanding Ens methods with multiple water species is not plain sailing though as sampling and other errors will be significant, especially for relatively small ensemble sizes, such as the studies mentioned above. Localization and ensemble inflation need to be applied carefully. The degree of localization depends on the variable and the situation. Typically the lengthscales of errors in condensate and rain variables are shorter than those of temperature and specific humidity, and these lengthscales depend on the strength of precipitation [Michel et al., 2011]. Shorter lengthscales in the error correlations generally call for tighter localization functions [Ménétrier et al., 2015b]. Ménétrier et al. [2015a] for instance use optimal linear Schur filtering theory to estimate the optimal localization factor between grid boxes i and j to be $\sim 1 - \frac{1}{N}\mathbb{E}[B^N_{ii}B^N_{jj}]/\mathbb{E}[B^{N^2}_{ij}]$ (their Eq. 64 and assuming a Gaussian distributed ensemble), where B^N_{ij} is the sample error covariance between positions i and j computed from N ensemble members ($N \gg 2$ assumed), and $\mathbb{E}[\bullet]$ indicates expectation. This is a useful formula because it can be estimated from available data. Ensemble inflation on the other hand needs to reflect the appropriate model error in these systems, it needs to compensate for filter divergence effects [Houtekamer and Mitchell, 1998], and can also be included judiciously to encourage the DA to update key variables such as clouds [Minamide and Zhang, 2019]. It may be useful to impose extra information in the covariances (such as balance) by posing the EnKF problem in terms of alternative control parameters (such as the ones reviewed here) in order to exploit the physical and statistical information present in the associated CVTs.

9 Non-Gaussianity

A convenient way of representing uncertainty in DA is to assume that errors obey Gaussian statistics, see e.g. Lorenc [1986]. Gaussian PDFs are defined entirely by a mean and a covariance, they are uni-modal, and the mean, median, and mode are coincident. Some PDFs are well approximated by a Gaussian, but others are far from Gaussian. Atmospheric water variables fall into the latter category (e.g. Sun, 2005a, Auligné et al., 2011) due to their physical bounds, nonlinear influence on the state, and the often nonlinear observation operators used in their measurement. Non-Gaussianities also appear when features are misaligned (e.g. Lawson and Hansen, 2005), and appear (e.g.) in satellite innovation statistics in the presence of cloud [Geer and Bauer, 2011]. The introduction of non-Gaussian (NG) PDFs in DA is complicated owing partly to the potential breakdown of the properties just mentioned. Any aspect of a DA system is potentially NG, namely background errors, observation errors, model errors, variational bias parameters, and analysis errors. The emphasis here is on the NG background errors of AW fields.

9.1 Consequences of assuming Gaussianity for a non-Gaussian parameter

Gaussian PDFs have infinite support, and so are not appropriate for bounded variables like AW, as they allow non-physical states (Sect. 6.4). Their uni-modality is a further inadequacy in representing PDFs of AW variables. Consider a hypothetical 'true' NG forecast PDF for q in Fig. 8 (solid black curve) comprising two well-separated peaks. This PDF may arise, for instance, from an ensemble of forecasts where some members are moistened by convective detrainment (moist mode), and others have no convection (dry mode). A similar effect is seen in Fig. 5b, but the multiple modes in that case may be due to a range of processes. The Gaussian with the same mean and variance is the long dashed black curve. Clearly this cannot properly represent the true PDF – it is too broad, it allows negative WV, and its mode lies at an improbable WV amount. Using such a Gaussian as the background error PDF could conceivably do more damage than good.

These arguments justify efforts to allow NG PDFs to be represented in DA. One way forward is to abandon the Gaussian infrastructure of methods like Var or the EnKF and use instead methods like particle filtering. Particle filters solve the general Bayesian DA problem by capturing the PDF with a set of weighted samples (particles). Particle filters are though difficult to use in NWP due to degeneracy problems [van Leeuwen, 2009], although this situation is improving [van Leeuwen et al., 2019, Potthast et al., 2019]. Sections 9.2-9.3 concern how it is possible to exploit NG PDFs by non-linear changes of variable [Bocquet et al., 2010] while still using Var and Ens-based systems and Sect. 9.4 in Var systems only. Section 9.5 concerns a different method that rigorously accounts for special cases of gamma and inverse gamma PDFs that is suitable for Ens systems.

9.2 Log normal (LN) distributions

The log normal (LN) PDF is often used for background and observation PDFs for non-negative variables. The logarithm of a log normal variable, e.g. $\ln q$, is Gaussian distributed, i.e. $\ln q = \overline{\ln q} + \delta \ln q$ where $\delta \ln q \sim \mathcal{N}(0, \sigma_{\ln q}^2)$. $\delta \ln q$ is a control parameter choice additional to those discussed in Sect. 7. Taking the exponential of each side yields $q = \exp(\overline{\ln q})\exp(\delta \ln q)$, showing that a LN PDF is equivalent to a multiplicative error scheme in q (the increment $\exp(\delta \ln q)$ multiplies $\exp(\overline{\ln q})$). The LN PDF of q is:

$$p_{\mathrm{LN}}(q) = \frac{1}{q\sqrt{2\pi\sigma_{\ln q}^2}} \exp\left[-\frac{\left(\ln q - \overline{\ln q}\right)^2}{2\sigma_{\ln q}^2} \right]. \tag{19}$$

Note that this is not simply a Gaussian with $q \to \ln q$. The mean, median, and mode of the lognormally distributed q differ: $q_{\mathrm{mean}} = \exp(\overline{\ln q} + \sigma_{\ln q}^2/2)$, $q_{\mathrm{median}} = \exp(\overline{\ln q})$, and $q_{\mathrm{mode}} = \exp(\overline{\ln q} - \sigma_{\ln q}^2)$; and the variance is $\sigma_{\mathrm{LN}}^2 = \exp(\sigma_{\ln q}^2 - 1)q_{\mathrm{mean}}^2$. Note that $\overline{\ln q}$ and $\sigma_{\ln q}^2$ are the mean and the variance of
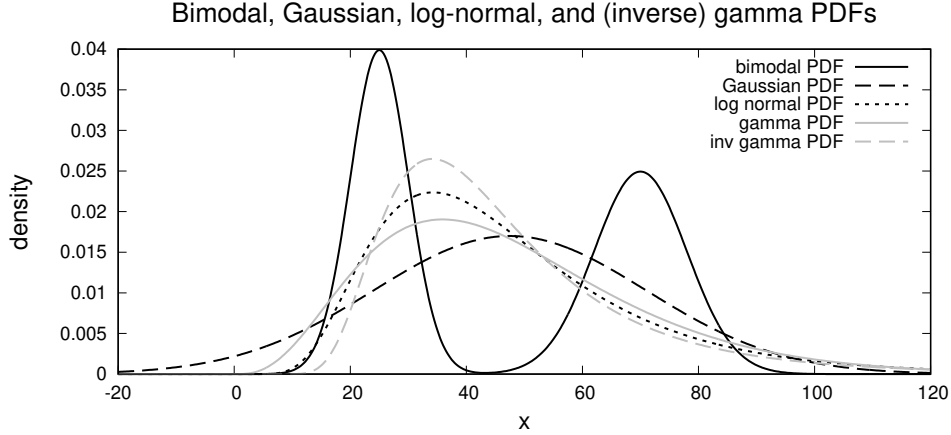
Figure 8: Schema showing example bimodal (solid black), Gaussian (long dashed black), log-normal (short dashed black), gamma (solid grey), and inverse gamma (dashed grey) PDFs. All have the same mean (47.5) and standard deviation (23.5).

the corresponding Gaussian variable $\ln q$. A LN PDF is shown in Fig. 8 (short dashed black curve), which can be compared to the Gaussian (long dashed black curve) with the same mean and variance.

Although the LN PDF has a lower bound of 0, it has no upper bound and so is not strictly suitable for some AW parameters that are capped at saturation. The LN is more difficult to use than a Gaussian owing to its different mean, median and mode [Fletcher and Zupanski, 2006], and in multivariate situations [Fletcher and Zupanski, 2007]. For instance a Gaussian PDF in $\ln q$ has its maximum at $(\ln q)_{max} = \overline{\ln q}$, i.e. $q_{max} = \exp(\overline{\ln q})$. This corresponds only to the median of the LN PDF, but in the case of application in Var, Fletcher and Zupanski [2006] show how a cost function can be modified to find the mean or the mode if required.

An issue arises when the background or the observation is assumed to be LN distributed and when conditions are very dry. In this case the analysis will remain dry even if one or other of these data values indicates a humid value [Dee and Da Silva, 2003]. This is a consequence of the multiplicative nature of errors in q when LN distributed.

Tests for LN errors are provided by Kliewer et al. [2015], who show that WV sometimes behaves as a LN variable. There is other evidence though that WV errors are generally not LN [Hólm et al., 2002, Dee and Da Silva, 2003]. A number of systems do use (or have previously used) log humidity as a control parameter, e.g. EC's global and regional systems [Gauthier et al., 2007, Fillion et al., 2010], JMA's global system [JMA, 2013], MM5 [Xiao et al., 2007], WRF [Xiao and Sun, 2007], some 1DVar systems [Poli et al., 2002, Deblonde and English, 2003, Kliewer et al., 2016], radar-based liquid water retrieval methods (e.g. Fielding et al., 2014, used with an EnKF), and ocean ecosystem models (e.g. Song et al., 2012, used in 4DVar).

9.3   Gaussian anamorphosis (GA)

A possible solution to the NG problem is Gaussian anamorphosis (GA), which has been suggested by Bertino et al. [2003] for use in ocean DA, and by Hólm et al. [2002] in atmospheric DA. It has already been used with a coupled ocean ecosystem model [Simon and Bertino, 2009], but to our knowledge it has not been implemented in an operational atmospheric DA system.

Consider an RH perturbation at a given grid-point, $\delta\mu$, which has NG PDF $p(\delta\mu|x^b)$. p is a conditional PDF here to account for flow-dependence. The GA method maps $\delta\mu$ to a new parameter $\delta\tilde{\mu}$ which has the Gaussian PDF $g(\delta\tilde{\mu}) = \mathcal{N}(0, 1)$. The same standard Gaussian is used irrespective of x^b, which is why $g(\delta\tilde{\mu})$ is not conditioned on x^b (the mapping though will depend on x^b). The GA
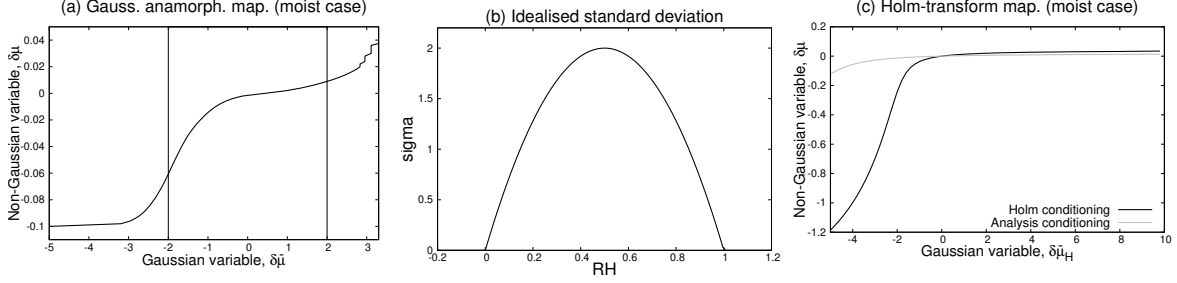
Figure 9: Panel a: example GA transform for the PDF shown in Fig. 5f ($p_\mu(\delta\mu|0.95 \le \mu < 1.0)$). Panel b: idealised standard deviation profile to demonstrate the Hólm transform. Panel c, black line: example Hólm transform for a background RH of 0.98 when the standard deviation of RH is that in panel b; grey line: as for the black line, but where the standard deviation is conditioned on the analysis (see Sect. 9.4 for details).

transformation relies on the integral probability transform theorem, e.g. Amezcua and van Leeuwen [2014], to find the mapping $\delta\mu \to \delta\tilde{\mu}$ such that $\int_{-\infty}^{\delta\mu} p(\delta\mu'|x^b)d\delta\mu' = \int_{-\infty}^{\delta\tilde{\mu}} g(\delta\tilde{\mu}')d\delta\tilde{\mu}'$, or in terms of their cumulative distribution functions (CDFs), $P(\delta\mu) = G(\delta\tilde{\mu})$. Here G and P are the CDFs of g and p respectively and we omit the conditioning notation for brevity. The forward and inverse mappings are:

$$\delta\mu = P^{-1}(G(\delta\tilde{\mu})), \tag{20a}$$
$$\delta\tilde{\mu} = G^{-1}(P(\delta\mu)). \tag{20b}$$

Sampling $\delta\tilde{\mu}$ from $\mathcal{N}(0,1)$ is equivalent to sampling $\delta\mu$ from $p(\delta\mu|x^b)$, so posing the DA problem in terms of the normal variable $\delta\tilde{\mu}$ is equivalent to the original NG problem in $\delta\mu$. Equation (20a) is needed in the DA to recover $\delta\mu$ from $\delta\tilde{\mu}$ and its Jacobian, $d\delta\mu/d\delta\tilde{\mu}$, is an extra step needed in the CVT, e.g. as part of $\mathbf{U}_p$. If $\delta\tilde{\mu}$ is chosen as a control parameter, it has Gaussian background error statistics, while $\delta\mu$ remains fully NG.

There are some potential difficulties of GA. One is that the function $P^{-1}(G(\delta\tilde{\mu}))$ is nonlinear. Another is that P can be non-invertible, which happens when of $p(\delta\mu|x^b) = 0$. A further issue is revealed by assimilating no observations. Under these circumstances the maximum a-posteriori is at $\delta\tilde{\mu} = 0$, which (20a) maps back to the median of the NG PDF, rather than the mode or the mean. If the background state does not represent the median of the NG PDF, then there will be a non-zero analysis increment for no observations. A practical solution might be to force an extra opposite offset to cancel this out. Furthermore, it is difficult to apply GA to multivariate PDFs [Amezcua and van Leeuwen, 2014], but a pragmatic way might be to first rotate the space to uncorrelated variables, and then apply the GA to each of these separately [Bocquet et al., 2010].

For illustration, Fig. 9a is the GA transform for the PDF in Fig. 5f (conditioned close to saturation). For about 95% of possible cases, the Gaussian parameter will be within two standard deviations of zero (vertical lines). Moving the Gaussian parameter from zero in the negative direction changes the NG variable significantly more than in the positive direction. This is how the skewness of the PDF in Fig. 5f is represented.

An alternative way of thinking about GA is to construct a rank-based inverse normal transformation (RBINT). Given an ordered N-member set to calibrate the GA transform, $\delta X = \{\delta\mu_1, \delta\mu_2, \ldots, \delta\mu_N\}$ – where members of the set are proxies of possible background forecast errors, e.g. as used in the NMC (National Meteorological Centre) method (e.g. Široká et al., 2003, Wang et al., 2014) – $P(\delta\mu)$ can be approximated with the empirical CDF:

$$P(\delta\mu) = p(\delta X \le \delta\mu) \approx \frac{1}{N}\sum_{i=1}^{N}\mathbb{1}\{\delta\mu_i \le \delta\mu\} = \frac{1}{N}R_{\delta\mu}, \tag{21}$$

where $\mathbb{1}\{\delta\mu_i \leq \delta\mu\}$ is the indicator function defined as

$$\mathbb{1}\{\delta\mu_i \leq \delta\mu\} = \begin{cases} 1 & \text{if } \delta\mu_i \leq \delta\mu \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

The rank $R_{\delta\mu}$ in (21) is the sum $\sum_{i=1}^{N} \mathbb{1}\{\delta\mu_i \leq \delta\mu\}$, which is the number of members that have value at or below $\delta\mu$. Complications arise when some members have equal values, but this can be dealt with by carefully devised rules [Chipilski, 2016]. Equation (21) may be used directly in (20b) to give the Gaussian control parameter, but in practice a slightly modified version is used:

$$\delta\tilde{\mu} = G^{-1}\left(\frac{R_{\delta\mu} - c}{N - 2c + 1}\right), \tag{23}$$

where the constant c is normally taken to be the value $c = 3/8$ [Beasley et al., 2009]. The equivalent of (20a) (needed for DA) involves rearranging (23):

$$R_{\delta\mu} = (N - 2c + 1)G(\delta\tilde{\mu}) + c, \tag{24}$$

which returns a rank used with the set δX to recover the value of $\delta\mu$. Note that $R_{\delta\mu}$ from (24) will not likely be an integer, but the non-integer remainder can be used to interpolate between members of δX. The RBINT approach is a straightforward implementation of GA as it only requires an ordered calibration set. It has been used in genetics studies [Beasley et al., 2009], but there is scope for its use in DA in Var and Ens schemes.

9.4 'Symmetrising' transforms

Some NG aspects of AW can be accommodated in Var using the procedure of Hólm et al. [2002]. They introduce a different kind of control parameter whose PDF is more symmetric (hence closer to a Gaussian) than the raw AW variable error's PDF, thus avoiding log-normal and GA transforms. This is the 'Hólm transform', which is used in the ECMWF 4DVar system. It is also used, with adaptations, in HIRLAM [Gustafsson et al., 2011], in Met Office Var systems [Ingleby et al., 2013b], and in an intermediate research model [Zaplotnik et al., 2018]. Since there is little literature on the Hólm transform itself, we provide some detail here. The essential steps are as follows.

1. Consider background error perturbations (e.g. RH, $\delta\mu$), such that $\mu = \mu_B + \delta\mu$. Suppose that a PDF of $\delta\mu$ can be found that has zero mean and is near Gaussian when conditioned on some variable, Φ. Let this conditional PDF be $p(\delta\mu|\Phi)$ (note that the unconditioned PDF, $p(\delta\mu)$, may itself be NG).

2. Determine the variance of each conditional PDF as a function of Φ, $\sigma^2(\Phi)$.

3. Construct a new parameter, $\delta\tilde{\mu}_H$ (subscript "H" for Hólm), which is normalised by $\sigma(\Phi)$, i.e. $\delta\tilde{\mu}_H = \delta\mu/\sigma(\Phi)$.

Given that $p(\delta\mu|\Phi)$ has zero mean and is near Gaussian for all Φ, the PDF obeyed by $\delta\tilde{\mu}_H$, $p(\delta\tilde{\mu}_H|\Phi)$, will have the same property. In addition, $p(\delta\tilde{\mu}_H|\Phi)$ will always have unit variance, which means that $p(\delta\tilde{\mu}_H|\Phi)$ will be the same for all Φ values. The parameter $\delta\tilde{\mu}_H$ is the new control parameter whose statistics have suitable properties (namely unit variance and nearly Gaussian), but imply a NG PDF in $\delta\mu$.

Point 1 can be achieved approximately by a choice of conditioning that is based on the following theorem. If two independent states are drawn from the same PDF then the PDF followed by their difference, when conditioned on their mean, is symmetric (Gaussianity is not guaranteed by this

41

theorem, but symmetry is likely to be a useful property). Let the two members be μ^1 and μ^2, so their mean is $\overline{\mu} = (\mu^1 + \mu^2)/2$, the theorem is:

$$p(\mu^2 - \mu^1|\overline{\mu}) = p(\mu^1 - \mu^2|\overline{\mu}), \tag{25}$$

which is straightforward to prove using the fact that μ^1 and μ^2 are interchangeable. The key idea is to choose $\Phi = \overline{\mu}$. Hólm et al. [2002] applies this by taking μ^1 as the background (μ^{b}), and μ^2 as an estimate of the analysis (μ^{a}). If $\delta\mu$ is the deviation from the background at a Var iteration, then an estimate of the analysis is $\mu^{\mathrm{a}} \approx \mu^{\mathrm{b}} + \delta\mu$. The above mean state is then $\Phi = \overline{\mu} = (\mu^{\mathrm{b}} + \mu^{\mathrm{a}})/2 \approx \mu^{\mathrm{b}} + \delta\mu/2$. The Hólm transform follows from step 3 above:

$$\delta\tilde{\mu}_{\mathrm{H}} = \frac{\delta\mu}{\sigma(\mu^{\mathrm{b}} + \delta\mu/2)}, \tag{26}$$

which is invoked at each Var iteration. This step has been incorporated in the Met Office in addition to the adaptive scheme described in Sect. 7.6, where the actual control variable used there is $\delta\tilde{\mu}_{\mathrm{HA}} = \delta\tilde{\mu}_{\mathrm{A}}/\sigma$ (see (15) and "HA" here being short for Hólm and adaptive). The determination of $\delta\mu$ from $\delta\tilde{\mu}_{\mathrm{H}}$ becomes part of $\mathbf{U}_{\mathrm{p}}$ in (6), and involves solving the nonlinear equation (26).

Panels b and c of Fig. 9 show how the Hólm transform constrains the analysis, but also how it can sometimes allow unphysical analyses. Panel b shows a particular choice of $\sigma(\bullet)$, which is non-zero only inside the natural bounds of its argument, $[0, 1]$. When the argument of $\sigma(\bullet)$, $\mu^{\mathrm{b}} + \delta\mu/2$, lies *outside* of the interval $[0, 1]$, $\sigma = 0$, it follows from (26) that $\delta\mu = 0$ for any $\delta\tilde{\mu}_{\mathrm{H}}$ value. However, setting $\delta\mu = 0$ means that the argument of $\sigma(\bullet)$ in (26) is $\mu = \mu^{\mathrm{b}}$, which is assumed to be *inside* the interval. This inconsistency means that $\mu^{\mathrm{b}} + \delta\mu/2$ cannot lie outside of $[0, 1]$. The relationship between $\delta\mu$ and $\delta\tilde{\mu}_{\mathrm{H}}$ satisfying (26) is plotted in panel c (black line) for a moist background, $\mu^{\mathrm{b}} = 0.98$. The asymptotic limits of validity are (i) $\mu^{\mathrm{b}} + \delta\mu/2 \to 0$ (leading to the analysis $\mu^{\mathrm{a}} = \mu^{\mathrm{b}} + \delta\mu = -\mu^{\mathrm{b}}$), and (ii) $\mu^{\mathrm{b}} + \delta\mu/2 \to 1$ ($\mu^{\mathrm{a}} = 2 - \mu^{\mathrm{b}}$). The analysis therefore does not preclude negative humidity and supersaturation. The Hólm transform though can be seen to inhibit positive increments in $\delta\mu$, and encourage negative increments for such a moist case. The opposite is found for a dry μ^{b} (not shown). We make the following notes about the Hólm transform.

- Although Hólm et al. [2002] apply this transform to RH, the approach is more general. Furthermore, the choice $\Phi = \overline{\mu}$ is not unique: Φ can be any symmetric function of μ_1 and μ_2. It may even be found by experiment that symmetry is found when Φ is a non-AW variable [Chipilski, 2016].

- The symmetry property used for (25) does not require Gaussianity, only symmetry. Furthermore, (25) applied with $\mu^1 = \mu^{\mathrm{b}}$ and $\mu^2 = \mu^{\mathrm{a}}$ is not valid (strictly speaking) as μ^{b} and μ^{a} are not samples from the same PDF.

- Useful claimed properties of the Hólm transform include the partial treatment of bias in the background PDF (this is a significant issue close to the natural bounds of humidity), and the deterrence of out-of-bounds values (discussed above). It also can prevent the DA iterations from getting stuck at small values of μ (a problem found with $\ln q$).

- As mentioned above, the Hólm transform does not always prevent increments that extend beyond the natural bounds. An alternative is to condition on $\Phi = \mu^{\mathrm{a}}$, rather than on $\Phi = \overline{\mu}$. In this case, (26) becomes $\delta\tilde{\mu} = \delta\mu/\sigma(\mu^{\mathrm{b}} + \delta\mu)$. This transform does prevent out-of-bounds analyses applied to the above case (the transform is described by the grey line in Fig. 9c) but (25) (the 'symmetry' property) no longer applies.

The changes documented by Ingleby et al. [2013b] – which prominently include the symmetrising transform – does show improvements in the Met Office global DA system, namely a reduction in the occurrence of negative humidities, and improvements in the fit to radiosondes and to most satellite data.

9.5 Gamma and inverse gamma PDFs (GIGG filter)

All of the above methods of accounting for NG require an interface to a Gaussian variable, used as a control parameter in Var or in the EnKF, and can be used in combination with the choice of transforms discussed in Sects. 7 and 8. Even though these methods do account for NG, Bishop [2016] points out that they can lead to biases in the analyses (possibly related to the finite analysis increment in the GA scheme when no observations are assimilated – see Sect. 9.3). Bishop [2016] illustrates this problem in terms of LN variables: if an analysis scheme for a LN distributed prior and likelihood is unbiased then the expectation over the analysis distribution is $\overline{\ln q^{\mathrm{a}}} = \ln q^{\mathrm{t}}$ (q^{a} is the analysis, and q^{t} is the truth). The expectation of q^{a} however is the mean of the LN PDF, $\overline{q^{\mathrm{a}}} = \exp(\overline{\ln q^{\mathrm{a}}} + \sigma_{\ln q^{\mathrm{a}}}^2/2)$ (Sect. (9.2), where $\sigma_{\ln q^{\mathrm{a}}}$ is the error standard deviation in the PDF of $\ln q^{\mathrm{a}}$). Unless $\sigma_{\ln q^{\mathrm{a}}}$ is very small, this is a biased estimate of q^{t}.

Bishop [2016] introduced a new filter that properly and efficiently accounts for special cases of NG variables, namely data with gamma (Gam) or inverse gamma (IG) PDFs in the way described below. Forms of the Gam (p_{Gam}) and IG (p_{IG}) PDFs are as follows,

$$p_{\mathrm{Gam}}(q) = \frac{\left(q\sigma_{\mathrm{r}}^{-2}/\bar{q}\right)^{\sigma_{\mathrm{r}}^{-2}}}{\Gamma(\sigma_{\mathrm{r}}^{-2})q} \exp\left(-\sigma_{\mathrm{r}}^{-2}q/\bar{q}\right), \tag{27}$$

$$p_{\mathrm{IG}}(q) = \frac{\left(\tilde{\sigma}_{\mathrm{r}}^{-2}\bar{q}\right)^{(\tilde{\sigma}_{\mathrm{r}}^{-2}+1)}}{\Gamma(\tilde{\sigma}_{\mathrm{r}}^{-2}+1)q^{(\tilde{\sigma}_{\mathrm{r}}^{-2}+2)}} \exp\left(-\tilde{\sigma}_{\mathrm{r}}^{-2}\bar{q}/q\right), \tag{28}$$

where $\bar{q}$ is the mean of each respective PDF and σ_{r}^2 and $\tilde{\sigma}_{\mathrm{r}}^2$ are the type I and II relative variances respectively. These are defined as $\sigma_{\mathrm{r}}^2 = \sigma^2/\bar{q}^2$, and $\tilde{\sigma}_{\mathrm{r}}^2 = \sigma^2/(\bar{q}^2 + \sigma^2)$ (note that $\tilde{\sigma}_{\mathrm{r}}^{-2} = \sigma_{\mathrm{r}}^{-2} + 1$). Figure 8 shows example Gam (solid grey line) and IG (dashed grey) PDFs, which can be compared to the Gaussian and LN PDFs (all PDFs in Fig. 8 have the same mean and variance). Gam PDFs are often used to describe cloud and precipitation properties (e.g. the particle size distribution in clouds, precipitation amounts, and wind speeds, Deng and Mace, 2006, Wilks, 2011). The Gam (when $\sigma_{\mathrm{r}}^{-2} > 1$) and IG PDFs are positive definite and qualitatively similar in shape to the LN PDFs. When $\sigma_{\mathrm{r}}^{-2} \leq 1$, Gam has non-zero density when $q = 0$ (positive semi-definite), and when $\sigma_{\mathrm{r}}^{-2} = 1$, the Gam takes the special form of the exponential distribution.

The text book definitions of these PDFs are usually written in terms of shape and scale parameters (e.g. for Gam, σ_{r}^{-2} is the shape parameter and $\bar{q}/\sigma_{\mathrm{r}}^{-2}$ is the scale parameter), but Bishop [2016] shows how using the above two types of relative variance allows some useful properties to be revealed when Gam/IG PDFs are used as the prior and likelihood in Bayes' Theorem for estimating a scalar quantity with a single observation. These properties are as follows.

1. The "GIG" case: if the prior is a Gam PDF (with mean $\bar{q}_{\mathrm{p}}$ and type I relative variance σ_{rp}^2), and the likelihood is an IG function (with type II relative variance $\tilde{\sigma}_{\mathrm{rl}}^2$, and observation value q^{o}), then the analysis is a Gam PDF with the following mean and type I relative variance:

$$1/\bar{q}_{\mathrm{a(GIG)}} = \frac{1}{\bar{q}^{\mathrm{p}}} + \frac{\tilde{\sigma}_{\mathrm{rp}}^2}{\tilde{\sigma}_{\mathrm{rp}}^2 + \tilde{\sigma}_{\mathrm{rl}}^2}\left(\frac{1}{q^{\mathrm{o}}} - \frac{\tilde{\sigma}_{\mathrm{rl}}^2 + 1}{\bar{q}^{\mathrm{p}}}\right), \tag{29}$$

$$1/\sigma_{\mathrm{ra(GIG)}}^2 = 1/\tilde{\sigma}_{\mathrm{rp}}^2 + 1/\tilde{\sigma}_{\mathrm{rl}}^2. \tag{30}$$

2. The "IGG" case: if the prior is an IG PDF (with mean $\bar{q}_{\mathrm{p}}$ and type II relative variance $\tilde{\sigma}_{\mathrm{rp}}^2$), and the likelihood is a Gam function (with type I relative variance σ_{rl}^2, and observation value q^{o}), then the analysis is an IG PDF with the following mean and type II relative variance:

$$\bar{q}_{\mathrm{a(IGG)}} = \bar{q}^{\mathrm{p}} + \frac{\tilde{\sigma}_{\mathrm{rp}}^2}{\tilde{\sigma}_{\mathrm{rp}}^2 + \tilde{\sigma}_{\mathrm{rl}}^2}\left(q^{\mathrm{o}} - \bar{q}^{\mathrm{p}}\right), \tag{31}$$

$$\sigma_{\mathrm{ra(IGG)}}^2 = \tilde{\sigma}_{\mathrm{rp}}^2 - \tilde{\sigma}_{\mathrm{rl}}^4/\left(\tilde{\sigma}_{\mathrm{rp}}^2 + \sigma_{\mathrm{rl}}^2\right). \tag{32}$$

Note that the likelihood functions are found from (27) or (28) by using the observed value as the random variable ($q \to q^{\circ}$), and then the possible truth as the mean ($\bar{q} \to q$). Bishop [2016] used these properties to develop a stochastic ensemble filter (the GIGG-EnKF) to properly represent the error statistics of quantities that have Gam and IG statistics in the above manner. GIGG stands for Gamma-Inverse-Gamma-Gaussian (Bishop's scheme also allows for Gaussian PDFs). These properties are demonstrated with single observations, but multiple observations can be assimilated serially in a consistent way (as long as they are mutually independent). The GIGG filter does have some limitations, e.g. for variables that are not Gam, IG, or Gaussian distributed, and for multivariate distributions. It has though has been tested successfully by Posselt and Bishop [2018] who estimated microphysical cloud parameters assumed to have a Gam prior and observations of precipitation rates assumed to have an IG likelihood. The method presents a step forward for certain classes of problems, pertinent to the analysis of AW variables.

10 Concluding remarks

Atmospheric water is difficult to measure, model, and assimilate (Sect. 1). The equations describing the behaviour of AW in models are nonlinear and concern multi-phase processes, including on sub-grid scales (Sect. 2). The aim of this paper is to summarise techniques that are used to assimilate AW observations, and to discuss some of the important issues that arise. A brief summary of the sources of observations sensitive to AW is also provided (Sect. 3) as this is helpful to understand which water species are observed and how.

The estimation of WV for large-scale NWP models by DA and other methods is done reasonably well (Sect. 5), although there are a number of problems that are known to occur (Sect. 6), which can worsen as the model resolution is increased. There is also a need for a more comprehensive dynamical representation of multiple water species in high-resolution models. This paper's key points, and some pointers to possible future directions, are now summarised.

- Data assimilation is arguably best done within a single system that considers all data together consistently. To this end, there is an effort to focus on either Var, Ens, and hybrid systems (Sects. 5.1-5.4), instead of invoking additional methods, e.g. with pre-retrieval stages, latent heat nudging, and initialisation (Sects. 4, 5.5, 5.6) for parts of the AW estimation. For pure 4DVar (Sect. 5.1), pure EnVar (Sect. 5.3), and hybrid (Sect. 5.4) systems the DA window needs to be of appropriate length to minimise error growth, but also to avoid frequent initialisation shocks (Sect. 4), and to allow for any spin-up of precipitation (e.g. Sun, 2005b). A similar issue relates to the update frequency of pure ensemble methods (Sect. 5.2).

- Models follow a different manifold to the real world. This often means that background states are incompatible with observations (Sect. 1). This can manifest itself in terms of misplaced features in forecasts, excessive smoothing of sharp features, and imbalances between the moist and (thermo)dynamic fields in analyses (Sects. 4, 6.1, 6.3). The so-called 'spin-down' problem exists in many models, where a model re-balances itself with anomalous precipitation in the early stages [Hólm et al., 2002, Dixon et al., 2009, Brousseau et al., 2016]. Multivariate background error covariance schemes that 'know' about the 'balances' between WV and dynamics (e.g. Sect. 7.7.3), and the natural bounds of AW (Sects. 6.4 and 9) may help. In the case of pure and hybrid Var systems (Sects. 5.1 and 5.4), the data used to calibrate the **B**-matrix – whether from a NMC data set [Široká et al., 2003, Wang et al., 2014], an ensemble of forecasts [Buehner, 2005], or from an ensemble of data assimilations [Fisher, 2003] – can also influence these balances. Application of grid distortion methods (Sect. 4) can reposition sharp features, but are not currently widely used. More generally, schemes to identify and reduce model errors may need to be introduced more widely, especially if they can be used to highlight problems in models that can then be addressed by modellers.

- Background errors for AW variables are inhomogeneous and flow-dependent, there are significant cross correlations with other variables which should be accounted for (Sects. 4 and 6.2), and it is difficult to obtain meaningful training samples of possible background errors for calibration of the background PDFs. The CVT method (Sect. 7.1) provides a tailored means of modelling covariances with a degree of flexibility through the broad choice of parameters for WV (Sects. 7), and other water species (below), including NG aspects (below). A promising method is to use statistical, but flow-dependent, CVTs for the AW variables (Sect. 7.7), given that there are no clear analytical balance relationships for WV that follow from the equations of motion. The flow-dependence should depend on cloud, precipitation, and arguably on pre-convective indicators of precipitation, like CAPE or WV flux convergence (Sect. 7.7.3). Similar techniques and issues apply for model errors, but these have not been emphasised in this paper.

- The CVT is traditionally applied to Var and hybrid methods, but an interesting research question is whether solving Ens-based systems in terms of different control parameters to the model variables (and localising in control parameter space) might help control spurious, unphysical cross-correlations between variables. Any CVTs used allow control over how increments should be related in addition to the natural flow-dependence from the ensemble.

- Remote sensing observations of AW quantities (such as radar reflectivity, and microwave and infrared radiances) are being increasingly assimilated (Sects. 3.2, 3.3), and new kinds of observations (such as lidar backscatter) are on the horizon. Their observation operators often need information about fields of multiple water species, but these fields are not available directly in many current DA systems. Schemes have been discussed that represent these fields either diagnostically or explicitly (Sect. 8), and how existing techniques may be used to represent their cross-covariances (Sect. 7.2). Some cloudy DA methods that rely on a diagnosis of condensed water from WV and thermodynamic fields suffer from an important problem resulting in an inability to allow cloudy observations to impact a system that has no cloud in the background state (Sects. 6.9 and 8.1).

- Non-Gaussianity is a property of PDFs of AW quantities (Sects. 6.4 and 9) but nearly all practical DA methods assume Gaussianity of the control parameters, and of observation errors. It is possible to transform NG to Gaussian variables using LN or more general GA transforms (Sects. 9.2,9.3), or to apply a 'symmetrising' transform to inhibit analysis increments that fall outside of the natural bounds of humidity (Sect. 9.4). Gaussian anamorphosis is an interesting way forward for AW, since it allows for general NG PDFs, and has not been widely explored before in atmospheric DA. The special filter Gamma-Inverse-Gamma-Gaussian-EnKF (GIGG-EnKF, Sect. 9.5) is a promising alternative to the (purely Gaussian) EnKF for a class of problems where the prior and likelihood are gamma and inverse gamma distributed (or vice-versa).

If these issues are not properly addressed, the assumptions made in classical DA (that the prominent processes are known/linear, and that all remaining uncertainties have quantifiable/Gaussian statistics) are violated to some degree leading to assimilation sub-optimalities. DA of AW quantities remains a huge and challenging problem. This article is one perspective, but our intention is to help in the scientific development of contemporary DA systems by highlighting the main options and obstacles to help better exploit observations that are sensitive to water in the atmosphere.

Acknowledgements

References

Jean Paul Aguttes, Jacques Schrive, Christophe Goldstein, Michel Rouze, and Garudachar Raju. Megha-Tropiques, a satellite for studying the water cycle and energy exchanges in the tropiques. In *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, volume 7, pages 3042–3044. IEEE, 2000.

Clark Amerault and Xiaolei Zou. Comparison of model-produced and observed microwave radiances and estimation of background error covariances for hydrometeor variables within hurricanes. *Monthly Weather Review*, 134(3):745–758, 2006.

Javier Amezcua and Peter Jan van Leeuwen. Gaussian anamorphosis in the analysis step of the EnKF: a joint state-variable/observation approach. *Tellus A*, 66, 2014.

Jeffrey L Anderson and Stephen L Anderson. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.

T Auligné, AP McNally, and DP Dee. Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133(624):631–642, 2007.

T Auligné, A Lorenc, Y Michel, T Montmerle, A Jones, M Hu, and J Dudhia. Toward a new cloud analysis and prediction system. *Bulletin of the American Meteorological Society*, 92(2):207, 2011.

Pauline M Austin. Relation between measured radar reflectivity and surface rainfall. *Monthly Weather Review*, 115(5):1053–1070, 1987.

Michael Baldauf, Axel Seifert, Jochen Förstner, Detlev Majewski, Matthias Raschendorfer, and Thorsten Reinhardt. Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, 139(12):3887–3905, 2011.

Susan P Ballard, Zhihong Li, David Simonin, and Jean-Francois Caron. Performance of 4D-Var NWP-based nowcasting of precipitation at the Met Office for summer 2012. *Quarterly Journal of the Royal Meteorological Society*, 142(694):472–487, 2016.

RN Bannister. A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143:607–633, 2017. doi: DOI:10.1002/qj.2982.

Ross N Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1951–1970, 2008a.

Ross N Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1971–1996, 2008b.

F Baordo and AJ Geer. Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval. *Quarterly Journal of the Royal Meteorological Society*, 142(700):2854–2866, 2016.

Dale M Barker, W Huang, Yong-Run Guo, AJ Bourgeois, and QN Xiao. A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Monthly Weather Review*, 132(4):897–914, 2004.

Peter Bauer, Philippe Lopez, Angela Benedetti, Deborah Salmond, and Emmanuel Moreau. Implementation of 1D+ 4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1D-Var. *Quarterly Journal of the Royal Meteorological Society*, 132(620):2277–2306, 2006a.

Peter Bauer, Emmanuel Moreau, Frédéric Chevallier, and Una O'keeffe. Multiple-scattering microwave radiative transfer for data assimilation applications. *Quarterly Journal of the Royal Meteorological Society*, 132(617):1259–1281, 2006b.

Peter Bauer, Alan J Geer, Philippe Lopez, and Deborah Salmond. Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quarterly Journal of the Royal Meteorological Society*, 136(652):1868–1885, 2010.

Peter Bauer, Thomas Auligné, William Bell, Alan Geer, Vincent Guidard, Sylvain Heilliette, Masahiro Kazumori, Min-Jeong Kim, Emily H-C Liu, Anthony P McNally, et al. Satellite cloud and precipitation assimilation at operational NWP centres. *Quarterly Journal of the Royal Meteorological Society*, 137(661):1934–1951, 2011a.

Peter Bauer, George Ohring, Chris Kummerow, and Tom Auligne. Assimilating satellite observations of clouds and precipitation into NWP models. *Bulletin of the American Meteorological Society*, 92 (6):ES25, 2011b.

BR Bean and EJ Dutton. Radio meteorology. *National Bureau of Standards Monograph*, 92(613): 435pp, 1968.

T Mark Beasley, Stephen Erickson, and David B Allison. Rank-based inverse normal transformations are increasingly used, but are they merited? *Behavior genetics*, 39(5):580–595, 2009.

William Bell, Brett Candy, Nigel Atkinson, Fiona Hilton, N Baker, Niels Bormann, Graeme Kelly, Masahiro Kazumori, William F Campbell, and Steven D Swadley. The assimilation of SSMIS radiances in numerical weather prediction models. *IEEE Transactions on Geoscience and Remote Sensing*, 46(4):884–900, 2008.

Lennart Bengtsson and Kevin I Hodges. On the impact of humidity observations in numerical weather prediction. *Tellus A*, 57(5):701–708, 2005.

Stanley G Benjamin, Stephen S Weygandt, John M Brown, Ming Hu, Curtis R Alexander, Tatiana G Smirnova, Joseph B Olson, Eric P James, David C Dowell, Georg A Grell, et al. A North American hourly assimilation and model forecast cycle: The rapid refresh. *Monthly Weather Review*, 144(4): 1669–1694, 2016.

Gemma V Bennitt and Adrian Jupp. Operational assimilation of GPS zenith total delay observations into the Met Office numerical weather prediction models. *Monthly Weather Review*, 140(8):2706–2719, 2012.

Loïk Berre. Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly Weather Review*, 128(3):644–667, 2000.

Laurent Bertino, Geir Evensen, and Hans Wackernagel. Sequential data assimilation techniques in oceanography. *International Statistical Review*, 71(2):223–241, 2003.

Michael Bevis, Steven Businger, Thomas A Herring, Christian Rocken, Richard A Anthes, and Randolph H Ware. GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, 97(D14):15787–15801, 1992.

T Bick, C Simmer, S Trömel, K Wapler, H-J Hendricks Franssen, K Stephan, U Blahak, C Schraff, H Reich, Y Zeng, et al. Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Quarterly Journal of the Royal Meteorological Society*, 142(696):1490–1504, 2016.

Camille Birman, Jean-François Mahfouf, Mathias Milz, Jana Mendrok, Stefan A Buehler, and Manfred Brath. Information content on hydrometeors from millimeter and sub-millimeter wavelengths. *Tellus A: Dynamic Meteorology and Oceanography*, 69(1):1271562, 2017.

Craig H Bishop. The GIGG-enKF: ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. *Quarterly Journal of the Royal Meteorological Society*, 142(696):1395–1412, 2016.

Craig H Bishop. Data assimilation strategies for state-dependent observation error variances. *Quarterly Journal of the Royal Meteorological Society*, 145(718):217–227, 2019.

SC Bloom, LL Takacs, AM Da Silva, and D Ledvina. Data assimilation using incremental analysis updates. *Monthly Weather Review*, 124(6):1256–1271, 1996.

WG Blumberg, DD Turner, U Löhnert, and S Castleberry. Ground-based temperature and humidity profiling using spectral infrared and microwave observations. Part II: Actual retrieval performance in clear-sky and cloudy conditions. *Journal of Applied Meteorology and Climatology*, 54(11):2305–2319, 2015.

Marc Bocquet, Carlos A Pires, and Lin Wu. Beyond Gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review*, 138(8):2997–3023, 2010.

Mary Borderies, O Caumont, C Augros, É Bresson, Julien Delanoë, V Ducrocq, N Fourrié, T Le Bastard, and M Nuret. Simulation of W-band radar reflectivity for model validation and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144(711):391–403, 2018.

Niels Bormann and Peter Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Quarterly Journal of the Royal Meteorological Society*, 136(649):1036–1050, 2010.

Niels Bormann, Massimo Bonavita, Rossana Dragani, Reima Eresmaa, Marco Matricardi, and Anthony McNally. Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, 142(697):1767–1780, 2016.

Sid-Ahmed Boukabara and Kevin Garrett. Tropospheric moisture sounding using microwave imaging channels: Application to GCOM-W1/AMSR2. *IEEE Transactions on Geoscience and Remote Sensing*, 2018.

Keith A Brewster. Phase-correcting data assimilation and application to storm-scale numerical weather prediction. part i: Method description and simulation testing. *Monthly weather review*, 131(3):480–492, 2003a.

Keith A Brewster. Phase-correcting data assimilation and application to storm-scale numerical weather prediction. part ii: Application to a severe storm outbreak. *Monthly weather review*, 131(3):493–507, 2003b.

Pierre Brousseau, Yann Seity, Didier Ricard, and Julien Léger. Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142(699):2231–2243, 2016.

G.H. Bryan, J.C. Wyngaard, and J.M. Fritsch. Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, 131:2394–2416, 2003.

Mark Buehner. Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, 131(607):1013–1043, 2005.

Mark Buehner, PL Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Monthly Weather Review*, 138(5):1550–1566, 2010a.

Mark Buehner, PL Houtekamer, Cecilien Charette, Herschel L Mitchell, and Bin He. Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Monthly Weather Review*, 138(5):1567–1586, 2010b.

Stephen Burt. *The Weather Observer's Handbook*. Cambridge University Press, 2012.

Jean-François Caron and Luc Fillion. An examination of background error correlations between mass and rotational wind over precipitation regions. *Monthly Weather Review*, 138(2):563–578, 2010.

Olivier Caumont, Véronique Ducrocq, Éric Wattrelot, Geneviève Jaubert, and Stéphanie Pradier-Vabre. 1D+ 3DVar assimilation of radar reflectivity data: a proof of concept. *Tellus A*, 62(2): 173–187, 2010.

Olivier Caumont, Domenico Cimini, Ulrich Löhnert, Lucas Alados-Arboledas, René Bleisch, Franco Buffa, Massimo Enrico Ferrario, Alexander Haefele, Thierry Huet, Fabio Madonna, et al. Assimilation of humidity and temperature observations retrieved from ground-based microwave radiometers into a convective-scale NWP model. *Quarterly Journal of the Royal Meteorological Society*, 142 (700):2692–2704, 2016.

JM Chagnon, SL Gray, and J Methven. Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1270–1282, 2013.

Philippe Chambon, Sara Q Zhang, Arthur Y Hou, Milija Zupanski, and Samson Cheung. Assessing the impact of pre-GPM microwave precipitation observations in the Goddard WRF ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1219–1235, 2014.

Phillippe Chambon and Alan J Geer. *All-sky assimilation of Megha-Tropiques/SAPHIR radiances in the ECMWF numerical weather prediction system*, volume 802. European Centre for Medium-Range Weather Forecasts, 2017.

Dong-Eon Chang, James A Weinman, Carlos A Morales, and William S Olson. The effect of spaceborne microwave and ground-based continuous lightning measurements on forecasts of the 1998 groundhog day storm. *Monthly weather review*, 129(8):1809–1833, 2001.

Jing Chen, Yu Wang, Li Li, Bin Zhao, Fajing Chen, Yinglin Li, and Yingjie Cui. A unified verification system for operational models from regional meteorological centres of China Meteorological Administration. *Meteorological Applications*, 20(2):140–149, 2013a.

Yaodeng Chen, Syed RH Rizvi, Xiang-Yu Huang, Jinzhong Min, and Xin Zhang. Balance characteristics of multivariate background error covariances and their impact on analyses and forecasts in tropical and arctic regions. *Meteorology and Atmospheric Physics*, 121(1-2):79–98, 2013b.

F Chevallier, P Bauer, J-F Mahfouf, and J-J Morcrette. Variational retrieval of cloud profile from atovs observations. *Quarterly Journal of the Royal Meteorological Society*, 128(585):2511–2525, 2002.

Hristo G Chipilski. The search for Gaussian moisture variables at the convective scale. Master's thesis, University of Reading, Department of Meteorology, United Kingdom, 2016.

Kao-Shen Chung, Weiguang Chang, Luc Fillion, and Monique Tanguay. Examination of situation-dependent background error covariances at the convective scale in the context of the ensemble Kalman filter. *Monthly Weather Review*, 141(10):3369–3387, 2013.

Domenico Cimini, Tim J Hewison, Lorenz Martin, Jürgen Güldner, Catherine Gaffard, and Frank S Marzano. Temperature and humidity profile retrievals from ground-based microwave radiometers during TUC. *Meteorologische Zeitschrift*, 15(1):45–56, 2006.

Peter Clark, Nigel Roberts, Humphrey Lean, Susan P Ballard, and Cristina Charlton-Perez. Convection-permitting models: a step-change in rainfall forecasting. *Meteorological Applications*, 23(2):165–181, 2016.

AM Clayton, Andrew C Lorenc, and Dale M Barker. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 139(675):1445–1461, 2013.

Michael C Coniglio, David J Stensrud, and Louis J Wicker. Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems. *Journal of the atmospheric sciences*, 63(4):1231–1252, 2006.

Kendra Cook, Chen-Joe Fong, Michael J Wenkel, Peter Wilczynski, Nick Yen, and GS Chang. FORMOSAT-7/COSMIC-2 GNSS radio occultation constellation mission for global weather monitoring. In *Aerospace Conference, 2013 IEEE*, pages 1–8. IEEE, 2013.

Philippe Courtier and Olivier Talagrand. Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus A*, 42(5):531–549, 1990.

Philippe Courtier, E Andersson, W Heckley, D Vasiljevic, M Hamrud, A Hollingsworth, F Rabier, M Fisher, and J Pailleux. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807, 1998.

Arthur P Cracknell. *Advanced Very High Resolution Radiometer AVHRR*. CRC Press, 1997.

Cyril Crevoisier, Cathy Clerbaux, Vincent Guidard, Thierry Phulpin, Raymond Armante, Blandine Barret, Claude Camy-Peyret, Jean-Pierre Chaboureau, Pierre-François Coheur, Laurent Crépeau, et al. Towards IASI-New Generation (IASI-NG): impact of improved spectral resolution and radiometric noise on the retrieval of thermodynamic, chemistry and climate variables. *Atmospheric Measurement Techniques*, 7:4367–4385, 2014.

Kenneth L Cummins, Martin J Murphy, Edward A Bardo, William L Hiscox, Richard B Pyle, and Alburt E Pifer. A combined TOA/MDF technology upgrade of the US National Lightning Detection Network. *Journal of Geophysical Research: Atmospheres*, 103(D8):9035–9044, 1998.

Salvatore D'Addio, Ville Kangas, Ulf Klein, Marc Loiselet, and Graeme Mason. The microwave radiometers on-board MetOp second generation satellites. In *Metrology for Aerospace (MetroAeroSpace), 2014 IEEE*, pages 599–604. IEEE, 2014.

Roger Daley. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK, 1992.

Sarah L Dance, Susan P Ballard, Ross N Bannister, Peter Clark, Hannah L Cloke, Timothy Darlington, David LA Flack, Suzanne L Gray, Lee Hawkness-Smith, Nawal Husnoo, et al. Improvements in forecasting intense rainfall: Results from the FRANC (Forecasting Rainfall exploiting new data Assimilation techniques and Novel observations of Convection) project. *Atmosphere*, 10(3):125, 2019.

Godelieve Deblonde and Stephen English. One-dimensional variational retrievals from SSMIS-simulated observations. *Journal of Applied Meteorology*, 42(10):1406–1420, 2003.

Dick P Dee. Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131 (613):3323–3343, 2005.

Dick P Dee and Arlindo M Da Silva. The choice of variable for atmospheric moisture analysis. *Monthly Weather Review*, 131(1):155–171, 2003.

Min Deng and Gerald G Mace. Cirrus microphysical properties and air motion statistics using cloud radar Doppler moments. Part I: Algorithm description. *Journal of applied meteorology and climatology*, 45(12):1690–1709, 2006.

J Derber and F Bouttier. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, 51(2):195–221, 1999.

Gérald Desroziers, Loic Berre, Bernard Chapnik, and Paul Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3385–3396, 2005.

S Di Michele, E Martins, and M Janiskova. Observation operator and observation processing for cloud lidar. Support-to-science-element (STSE) study EarthCARE assimilation. Shinfield Park, Reading, January 2014a.

S Di Michele, E Martins, and M Janiskova. Observation operator and observation processing for cloud radar. Support-to-science-element (STSE) study EarthCARE assimilation. Shinfield Park, Reading, January 2014b.

David J Diner, Jewel C Beckert, Terrence H Reilly, Carol J Bruegge, James E Conel, Ralph A Kahn, John V Martonchik, Thomas P Ackerman, Roger Davies, Siegfried AW Gerstl, et al. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1072–1087, 1998.

Ken Dixon, Clifford F Mass, Gregory J Hakim, and Robert H Holzworth. The impact of lightning data assimilation on deterministic and ensemble forecasts of convective events. *Journal of Atmospheric and Oceanic Technology*, 33(9):1801–1823, 2016.

Mark Dixon, Zhihong Li, Humphrey Lean, Nigel Roberts, and Sue Ballard. Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model. *Monthly Weather Review*, 137(5):1562–1584, 2009.

RJ Doviak and DS Zrnić. *Doppler radar and weather observations*. Dover Publications Inc., 2006.

Richard L Dowden, James B Brundell, and Craig J Rodger. Vlf lightning location by time of group arrival (toga) at multiple sites. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(7):817–830, 2002.

David C Dowell, Louis J Wicker, and Chris Snyder. Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Monthly Weather Review*, 139(1):272–294, 2011.

David W Draper, David A Newell, Frank J Wentz, Sergey Krimchansky, and Gail M Skofronick-Jackson. The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7):3452–3462, 2015.

Fanny Duffourg, Véronique Ducrocq, Nadia Fourrié, Geneviève Jaubert, and Vincent Guidard. Simulation of satellite infrared radiances for convective-scale data assimilation over the Mediterranean. *Journal of Geophysical Research: Atmospheres*, 115(D15), 2010.

Fabrice Duruisseau, Philippe Chambon, Eric Wattrelot, Marylis Barreyat, and Jean-François Mahfouf. Assimilating cloudy and rainy microwave observations from SAPHIR on board Megha Tropiques within the ARPEGE global model. *Quarterly Journal of the Royal Meteorological Society*, 145(719): 620–641, 2019.

Martin Ehrendorfer. A review of issues in ensemble-based Kalman filtering. *Meteorologische Zeitschrift*, 16(6):795–818, 2007.

Ronald M Errico, Peter Bauer, and Jean-François Mahfouf. Issues regarding the assimilation of cloud and precipitation data. *Journal of the Atmospheric Sciences*, 64(11):3785–3798, 2007a.

Ronald M Errico, George Ohring, Fuzhong Weng, Peter Bauer, Brad Ferrier, Jean-François Mahfouf, and Joe Turk. Assimilation of satellite cloud and precipitation observations in numerical weather prediction models: Introduction to the JAS special collection. *Journal of the Atmospheric Sciences*, 64(11):3737–3741, 2007b.

Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

JR Eyre. Progress achieved on assimilation of satellite data in numerical weather prediction over the last 30 years. In *ECMWF Seminar Proceedings: Recent Developments in use of satellite observations in numerical weather prediction*, pages 1–27, 2007.

Frédéric Fabry. Meteorological value of ground target measurements by radar. *Journal of Atmospheric and Oceanic Technology*, 21(4):560–573, 2004.

Frédéric Fabry. For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part II: On the observation signal from different sensors. *Monthly Weather Review*, 138(1):256–264, 2010.

Frédéric Fabry and Juanzhen Sun. For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: On the propagation of initial condition errors and their implications for data assimilation. *Monthly Weather Review*, 138(1):242–255, 2010.

Frédéric Fabry, Chuck Frush, Isztar Zawadzki, and Alamelu Kilambi. On the extraction of near-surface index of refraction using radar phase measurements from ground targets. *Journal of atmospheric and oceanic technology*, 14(4):978–987, 1997.

Mark D Fielding, J Christine Chiu, Robin J Hogan, and Graham Feingold. A novel ensemble method for retrieving properties of warm cloud in 3-D using ground-based scanning radar and zenith radiances. *Journal of Geophysical Research: Atmospheres*, 119(18), 2014.

Luc Fillion, Monique Tanguay, Ervig Lapalme, Bertrand Denis, Michel Desgagne, Vivian Lee, Nils Ek, Zhuo Liu, Manon Lajoie, Jean-François Caron, et al. The Canadian regional data assimilation and forecasting system. *Weather and Forecasting*, 25(6):1645–1669, 2010.

Claude Fischer, Thibaut Montmerle, Loïk Berre, Ludovic Auger, and Simona Ecaterina Ştefănescu. An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3477–3492, 2005.

Mike Fisher. Background error covariance modelling. In *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean*, pages 45–63, 2003.

SJ Fletcher and M Zupanski. A data assimilation method for log-normally distributed observational errors. *Quarterly Journal of the Royal Meteorological Society*, 132(621):2505–2519, 2006.

Steven J Fletcher and Milija Zupanski. Implications and impacts of transforming lognormal variables into normal variables in VAR. *Meteorologische Zeitschrift*, 16(6):755–765, 2007.

Mary Forsythe. Atmospheric motion vectors: past, present and future. In *ECMWF Annual Seminar*, 2007.

Alison Fowler, Ross Bannister, and John Eyre. A new floating model level scheme for the assimilation of boundary-layer top inversions: the univariate assimilation of temperature. *Quarterly Journal of the Royal Meteorological Society*, 138(664):682–698, 2012.

Alison M Fowler, Ross N Bannister, and John Eyre. Characterising the background errors for the boundary-layer capping inversion. *Austral. Meteorol. Oceanic J*, 59:17–24, 2010.

Pierre Gauthier and Jean-Noel Thépaut. Impact of the digital filter as a weak constraint in the preoperational 4DVar assimilation system of Météo-France. *Monthly Weather Review*, 129(8):2089–2102, 2001.

Pierre Gauthier, C Charette, L Fillion, P Koclas, and S Laroche. Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmosphere-Ocean*, 37(2):103–156, 1999.

Pierre Gauthier, Monique Tanguay, Stephane Laroche, Simon Pellerin, and Josee Morneau. Extension of 3DVar to 4DVar: Implementation of 4DVar at the Meteorological Service of Canada. *Monthly Weather Review*, 135(6):2339–2354, 2007.

Guoqing Ge, Jidong Gao, and Ming Xue. Impacts of assimilating measurements of different state variables with a simulated supercell storm and three-dimensional variational method. *Monthly Weather Review*, 141(8):2759–2777, 2013.

AJ Geer and F Baordo. Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmospheric Measurement Techniques*, 7(6):1839–1860, 2014.

AJ Geer, F Baordo, N Bormann, P Chambon, SJ English, M Kazumori, H Lawrence, P Lean, K Lonitz, and C Lupu. The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quarterly Journal of the Royal Meteorological Society*, 143(709):3189–3206, 2017. doi: DOI:10.1002/qj.3172.

Alan J Geer and Peter Bauer. Observation errors in all-sky data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 137(661):2024–2037, 2011.

Alan J Geer, Peter Bauer, and Philippe Lopez. Lessons learnt from the operational 1D+ 4D-Var assimilation of rain-and cloud-affected SSM/I observations at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 134(635):1513–1525, 2008.

Alan J Geer, Katrin Lonitz, Peter Weston, Masahiro Kazumori, Kozo Okamoto, Yanqiu Zhu, Emily Huichun Liu, Andrew Collard, William Bell, Stefano Migliorini, et al. All-sky satellite data assimilation at operational weather forecasting centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713):1191–1217, 2018. doi: DOI:10.1002/qj.3202.

Mitchell D Goldberg, Heather Kilcoyne, Harry Cikanek, and Ajay Mehta. Joint Polar Satellite System: The United States next generation civilian polar-orbiting environmental satellite system. *Journal of Geophysical Research: Atmospheres*, 118(24):13–463, 2013.

BW Golding, SP Ballard, K Mylne, N Roberts, A Saulter, C Wilson, P Agnew, LS Davis, J Trice, C Jones, D Simonin, Z Li, C Pierce, A Bennett, M Weeks, and S Moseley. Forecasting capabilities for the London 2012 Olympics. *Bulletin of the American Meteorological Society*, 95(6):883–896, 2014.

J Gong and EV Hólm. Development of cloud condensate background errors. *ECMWF Newsletter*, 128: 23–27, 2011.

Steven J Goodman, Richard J Blakeslee, William J Koshak, Douglas Mach, Jeffrey Bailey, Dennis Buechler, Larry Carey, Chris Schultz, Monte Bateman, Eugene McCaul Jr, et al. The GOES-R geostationary lightning mapper (GLM). *Atmospheric research*, 125:34–49, 2013.

Matthias Grzeschik, Hans-Stefan Bauer, Volker Wulfmeyer, Dirk Engelbart, Ulla Wandinger, Ina Mattis, Dietrich Althausen, Ronny Engelmann, Matthias Tesche, and Andrea Riede. Four-dimensional variational data analysis of water vapor Raman lidar data and their impact on mesoscale forecasts. *Journal of Atmospheric and Oceanic Technology*, 25(8):1437–1453, 2008.

Jérémy Guerbette, Jean-François Mahfouf, and Matthieu Plu. Towards the assimilation of all-sky microwave radiances from the SAPHIR humidity sounder in a limited area NWP model over tropical regions. *Tellus A*, 68, 2016.

V Guidard, N Fourrié, P Brousseau, and F Rabier. Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Quarterly Journal of the Royal Meteorological Society*, 137(661):1975–1987, 2011.

Nils Gustafsson, S Thorsteinsson, M Stengel, and E Hólm. Use of a nonlinear pseudo-relative humidity variable in a multivariate formulation of moisture analysis. *Quarterly Journal of the Royal Meteorological Society*, 137(657):1004–1018, 2011.

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

Thomas M Hamill and Chris Snyder. A hybrid ensemble Kalman filter-3D variational analysis scheme. *Monthly Weather Review*, 128(8):2905–2919, 2000.

Thomas M Hamill and Jeffrey S Whitaker. Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Monthly Weather Review*, 133 (11):3132–3147, 2005.

Thomas M Hamill, Jeffrey S Whitaker, and Chris Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129(11):2776–2790, 2001.

SB Healy and J-N Thépaut. Assimilation experiments with CHAMP GPS radio occultation measurements. *Quarterly Journal of the Royal Meteorological Society*, 132(615):605–623, 2006.

Reinhold Hess. Specific issues for the use of satellite data for limited area models. In *Proceedings of ECMWF Seminar on Recent Developments in the use of Satellite Observations in Numerical Weather Prediction*, pages 375–396, 2007.

A Hoff. Advancements in the AMDAR humidity sensing. https://amdar.noaa.gov/docs/TECO2010_AMDAR_WV_Helms-Hoff_6.pdf, 2010.

Cathy Hohenegger and Christoph Schär. Atmospheric predictability at synoptic versus cloud-resolving scales. *Bulletin of the American Meteorological Society*, 88(11):1783, 2007.

E Hólm, E Andersson, A Beljaars, P Lopez, JF Mahfouf, A Simmons, and JN Thépaut. Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. *ECMWF Technical Memorandum*, 383, 2002.

Takumi Honda, Takemasa Miyoshi, Guo-Yuan Lien, Seiya Nishizawa, Ryuji Yoshida, Sachiho A Adachi, Koji Terasaki, Kozo Okamoto, Hirofumi Tomita, and Kotaro Bessho. Assimilating all-sky Himawari-8 satellite infrared radiances: A case of Typhoon Soudelor (2015). *Monthly Weather Review*, 146(1):213–229, 2018.

András Horányi, István Ihász, and Gábor Radnóti. ARPEGE/ALADIN: A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Idojárás*, 100(4):277–301, 1996.

Peter L Houtekamer and Herschel L Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1998.

Peter L Houtekamer and Herschel L Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1):123–137, 2001.

PL Houtekamer and Fuqing Zhang. Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144(12):4489–4532, 2016.

Huijun Huang, Bin Huang, Li Yi, Chunxia Liu, Jing Tu, Guanhuan Wen, and Weikang Mao. Evaluation of the global and regional assimilation and prediction system for predicting sea fog over the South China Sea. *Advances in Atmospheric Sciences*, 36(6):623–642, 2019.

Anthony Illingworth. Improved precipitation rates and data quality by using polarimetric measurements. In *Weather radar*, pages 130–166. Springer, 2004.

Anthony J Illingworth, HW Barker, A Beljaars, Marie Ceccaldi, H Chepfer, Nicolas Clerbaux, J Cole, Julien Delanoë, C Domenech, David P Donovan, et al. The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation. *Bulletin of the American Meteorological Society*, 96(8):1311–1332, 2015.

Bruce Ingleby. Global assimilation of air temperature, humidity, wind and pressure from surface stations. *Quarterly Journal of the Royal Meteorological Society*, 141(687):504–517, 2015.

Bruce Ingleby, David Moore, Chris Sloan, and Robert Dunn. Evolution and accuracy of surface humidity reports. *Journal of Atmospheric and Oceanic Technology*, 30(9):2025–2043, 2013a.

N Bruce Ingleby. The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 127(571):209–231, 2001.

NB Ingleby, AC Lorenc, K Ngan, F Rawlins, and DR Jackson. Improved variational analyses using a nonlinear humidity control variable. *Quarterly Journal of the Royal Meteorological Society*, 139 (676):1875–1887, 2013b.

Kosuke Ito, Masaru Kunii, Takuya Kawabata, Kazuo Saito, Kazumasa Aonashi, and Le Duc. Mesoscale hybrid data assimilation system based on JMA nonhydrostatic model. *Monthly Weather Review*, 144(9):3417–3439, 2016.

M Janisková. Assimilation of cloud information from space-borne radar and lidar: experimental study using a 1D+ 4D-Var technique. *Quarterly Journal of the Royal Meteorological Society*, 141(692): 2708–2725, 2015.

Tijana Janjic and Stephen E Cohn. Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Monthly Weather Review*, 134(10):2900–2915, 2006.

Xin Jin, Jun Li, Timothy J Schmit, Jinlong Li, Mitchell D Goldberg, and James J Gurka. Retrieving clear-sky atmospheric parameters from SEVIRI and ABI infrared radiances. *Journal of Geophysical Research: Atmospheres*, 113(D15), 2008.

JMA. Chapter 2, Data assimilation systems. `http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2013-nwp/` 2013.

Aaron Johnson, Xuguang Wang, Jacob R Carley, Louis J Wicker, and Christopher Karstens. A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Monthly Weather Review*, 143 (8):3087–3108, 2015.

CJ Johny and VS Prasad. Impact of assimilation of Megha-Tropiques ROSA radio occultation refractivity by observing system simulation experiment. *Current Science*, pages 1297–1305, 2014.

CD Jones and B Macpherson. A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. *Meteorological Applications*, 4(3):269–277, 1997.

Thomas A Jones, Jason A Otkin, David J Stensrud, and Kent Knopfmeier. Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Monthly Weather Review*, 141(10):3273–3299, 2013.

Thomas A Jones, Jason A Otkin, David J Stensrud, and Kent Knopfmeier. Forecast evaluation of an observing system simulation experiment assimilating both radar and satellite data. *Monthly Weather Review*, 142(1):107–124, 2014.

Eugenia Kalnay and Shu-Chih Yang. Accelerating the spin-up of ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 136(651):1644–1651, 2010.

Alexandra M Keclik, Clark Evans, Paul J Roebber, and Glen S Romine. The influence of assimilated upstream, preconvective dropsonde observations on ensemble forecasts of convection initiation during the Mesoscale Predictability Experiment. *Monthly Weather Review*, 145(12):4747–4770, 2017.

M Kitchen. Representativeness errors for radiosonde observations. *Quarterly Journal of the Royal Meteorological Society*, 115(487):673–700, 1989.

K Dieter Klaes, Marc Cohen, Yves Buhler, Peter Schlüssel, Rosemary Munro, Juha-Pekka Luntama, Axel von Engeln, Eoin Ó Clérigh, Hans Bonekamp, Jörg Ackermann, et al. An introduction to the EUMETSAT polar system. *Bulletin of the American Meteorological Society*, 88(7):1085–1096, 2007.

AJ Kliewer, SJ Fletcher, AS Jones, and JM Forsythe. Identifying non-normal and lognormal characteristics of temperature, mixing ratio, surface pressure, and wind for data assimilation systems. *Nonlinear Processes in Geophysics Discussions*, 2:1363–1405, 2015.

AJ Kliewer, SJ Fletcher, AS Jones, and JM Forsythe. Comparison of Gaussian, logarithmic transform and mixed Gaussian–log-normal distribution based 1DVar microwave temperature–water-vapour mixing ratio retrievals. *Quarterly Journal of the Royal Meteorological Society*, 142(694):274–286, 2016.

M Kojima, T Miura, K Furukawa, Y Hyakusoku, T Ishikiri, H Kai, T Iguchi, H Hanado, and K Nakagawa. Dual-frequency precipitation radar (DPR) development on the global precipitation measurement (GPM) core observatory. In *Earth Observing Missions and Sensors: Development, Implementation, and Characterization II*, volume 8528, page 85281A. International Society for Optics and Photonics, 2012.

Rong Kong, Ming Xue, and Chengsi Liu. Development of a hybrid En3DVar data assimilation system and comparisons with 3DVar and EnKF for radar data assimilation with observing system simulation experiments. *Monthly Weather Review*, 146(1):175–198, 2018.

ER Kursinski, GA Hajj, JT Schofield, RP Linfield, and K Rer Hardy. Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, 102(D19):23429–23465, 1997.

W Gregory Lawson and James A Hansen. Alignment error models and ensemble-based data assimilation. *Monthly Weather Review*, 133(6):1687–1709, 2005.

Humphrey W Lean, Peter A Clark, Mark Dixon, Nigel M Roberts, Anna Fitch, Richard Forbes, and Carol Halliwell. Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Monthly Weather Review*, 136(9):3408–3424, 2008.

Thomas E Lee, Steven D Miller, F Joseph Turk, Carl Schueler, Richard Julian, Steve Deyo, Patrick Dills, and Sherwood Wang. The NPOESS VIIRS day/night visible sensor. *Bulletin of the American Meteorological Society*, 87(2):191–200, 2006.

R Legrand, Y Michel, and T Montmerle. Diagnosing non-Gaussianity of forecast and analysis errors in a convective-scale model. *Nonlinear Processes in Geophysics*, 23(1):1, 2016.

G Leoncini, RS Plant, SL Gray, and Peter A Clark. Ensemble forecasts of a flood-producing storm: comparison of the influence of model-state perturbations and parameter modifications. *Quarterly Journal of the Royal Meteorological Society*, 139(670):198–211, 2013.

James C Liljegren, S-A Boukabara, Karen Cady-Pereira, and Shepard A Clough. The effect of the half-width of the 22-GHz water vapor line on retrievals of temperature and water vapor profiles with a 12-channel microwave radiometer. *IEEE Transactions on Geoscience and Remote Sensing*, 43(5): 1102–1108, 2005.

Magnus Lindskog, Kirsti Salonen, Heikki Järvinen, and DB Michelson. Doppler radar wind data assimilation with HIRLAM 3DVar. *Monthly weather review*, 132(5):1081–1092, 2004.

Chengsi Liu and Ming Xue. Relationships among four-dimensional hybrid ensemble–variational data assimilation algorithms with full and approximate ensemble covariance localization. *Monthly Weather Review*, 144(2):591–606, 2016.

Chengsi Liu, Qingnong Xiao, and Bin Wang. An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Monthly Weather Review*, 136(9):3363–3373, 2008.

Chengsi Liu, Qingnong Xiao, and Bin Wang. An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with Advanced Research WRF (ARW). *Monthly Weather Review*, 137(5):1687–1704, 2009.

Z-Q Liu and F Rabier. The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quarterly Journal of the Royal Meteorological Society*, 129(594): 3013–3035, 2003.

Philippe Lopez. Direct 4D-Var assimilation of NCEP stage IV radar and gauge precipitation data at ECMWF. *Monthly Weather Review*, 139(7):2098–2116, 2011.

Philippe Lopez and Peter Bauer. 1D+ 4DVar assimilation of NCEP stage-IV radar and gauge hourly precipitation data at ECMWF. *Monthly Weather Review*, 135(7):2506–2524, 2007.

Philippe Lopez and Emmanuel Moreau. A convection scheme for data assimilation: Description and initial tests. *Quarterly Journal of the Royal Meteorological Society*, 131(606):409–436, 2005.

AC Lorenc. A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, 109(4):701–721, 1981.

AC Lorenc. A study of ob monitoring statistics from radiosondes, composited for low-level cloud layers. *Met Office NWP Forecasting Research Technical Report*, 504:1–32, 2007.

AC Lorenc, D Barker, RS Bell, B Macpherson, and AJ Maycock. On the use of radiosonde humidity observations in mid-latitude NWP. *Meteorology and Atmospheric Physics*, 60(1-3):3–17, 1996.

AC Lorenc, SP Ballard, RS Bell, NB Ingleby, PLF Andrews, DM Barker, JR Bray, AM Clayton, T Dalby, D Li, T Payne, and F Saunders. The Met. Office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000.

Andrew C Lorenc. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112(474):1177–1194, 1986.

Andrew C Lorenc. The potential of the ensemble Kalman filter for NWP: comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 129(595):3183–3203, 2003.

Andrew C Lorenc. Recommended nomenclature for EnVar data assimilation methods. *Research Activities in Atmospheric and Oceanic Modeling, WGNE*, 2013.

Juha-Pekka Luntama, Gottfried Kirchengast, Michael Borsche, Ulrich Foelsche, Andrea Steiner, Sean Healy, Axel von Engeln, Eoin O'Clerigh, and Christian Marquardt. Prospects of the EPS GRAS mission for operational atmospheric applications. *Bulletin of the American Meteorological Society*, 89(12):1863–1876, 2008.

Jean-François Mahfouf. Data assimilation of the hydrological cycle. In *Proceedings of the ECMWF Seminar on Data Assimilation for Atmosphere and Ocean*, page 109, 2011.

Jean-Francois Mahfouf, Furqan Ahmed, Patrick Moll, and Felix N Teferle. Assimilation of zenith total delays in the AROME France convective scale model: a recent assessment. *Tellus A: Dynamic Meteorology and Oceanography*, 67(1):26106, 2015.

Edward R Mansell, Conrad L Ziegler, and Donald R MacGorman. A lightning data assimilation technique for mesoscale forecast models. *Monthly weather review*, 135(5):1732–1748, 2007.

P Martinet, N Fourrié, V Guidard, F Rabier, T Montmerle, and P Brunel. Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1402–1416, 2013.

Giovanni Martucci, Conor Milroy, and Colin D O'Dowd. Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers. *Journal of Atmospheric and Oceanic Technology*, 27(2):305–318, 2010.

AP McNally. A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quarterly Journal of the Royal Meteorological Society*, 128(585):2551–2556, 2002.

Benjamin Ménétrier and Thibaut Montmerle. Heterogeneous background-error covariances for the analysis and forecast of fog events. *Quarterly Journal of the Royal Meteorological Society*, 137(661):2004–2013, 2011.

Benjamin Ménétrier, Thibaut Montmerle, Yann Michel, and Loïk Berre. Linear filtering of sample covariances for ensemble-based data assimilation. Part I: Optimality criteria and application to variance filtering and covariance localization. *Monthly Weather Review*, 143(5):1622–1643, 2015a.

Benjamin Ménétrier, Thibaut Montmerle, Yann Michel, and Loïk Berre. Linear filtering of sample covariances for ensemble-based data assimilation. Part II: Application to a convective-scale NWP model. *Monthly Weather Review*, 143(5):1644–1664, 2015b.

Silas Michaelides, V Levizzani, E Anagnostou, P Bauer, T Kasparis, and JE Lane. Precipitation: Measurement, remote sensing, climatology and modeling. *Atmospheric Research*, 94(4):512–533, 2009.

Yann Michel, Thomas Auligné, and Thibaut Montmerle. Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables. *Monthly Weather Review*, 139(9): 2994–3015, 2011.

Stefano Migliorini and Brett Candy. All-sky satellite data assimilation of microwave temperature sounding channels at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 145 (719):867–883, 2019.

Stefano Migliorini, Andrew Lorenc, and W Bell. A moisture incrementing operator for the assimilation of humidity- and cloud-sensitive observations. *Quarterly Journal of the Royal Meteorological Society*, 2017.

Larry M Miloshevich, Holger Vömel, Ari Paukkunen, Andrew J Heymsfield, and Samuel J Oltmans. Characterization and correction of relative humidity measurements from Vaisala RS80-A radiosondes at cold temperatures. *Journal of Atmospheric and Oceanic Technology*, 18(2):135–156, 2001.

Masashi Minamide and Fuqing Zhang. Adaptive observation error inflation for assimilating all-sky satellite radiance. *Monthly Weather Review*, 145(3):1063–1081, 2017.

Masashi Minamide and Fuqing Zhang. Assimilation of all-sky infrared radiances from Himawari-8 and impacts of moisture and hydrometer initialization on convection-permitting tropical cyclone prediction. *Monthly Weather Review*, 146(10):3241–3258, 2018.

Masashi Minamide and Fuqing Zhang. An adaptive background error inflation method for assimilating all-sky radiances. *Quarterly Journal of the Royal Meteorological Society*, 145(719):805–823, 2019.

Thibaut Montmerle. Optimization of the assimilation of radar data at the convective scale using specific background error covariances in precipitation. *Monthly Weather Review*, 140(11):3495–3506, 2012.

Thibaut Montmerle and Loïk Berre. Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale. *Quarterly Journal of the Royal Meteorological Society*, 136(651):1408–1420, 2010.

Thibaut Montmerle and Claudia Faccani. Mesoscale assimilation of radial velocities from Doppler radars in a preoperational framework. *Monthly Weather Review*, 137(6):1939–1953, 2009.

Thibaut Montmerle, Florence Rabier, and Claude Fischer. Relative impact of polar-orbiting and geostationary satellite radiances in the ALADIN/France numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133(624):655–671, 2007.

Thomas Nehrkorn, Bryan Woods, Thomas Auligné, and Ross N Hoffman. Application of feature calibration and alignment to high-resolution analysis: Examples using observations sensitive to cloud and water vapor. *Monthly Weather Review*, 142(2):686–702, 2014.

JC Nicol, K Bartholemew, T Darlington, and AJ Illingworth. Operational radar refractivity retrieval for numerical weather prediction. In *Proceedings of the 8th International Symposium Weather Radar and Hydrology, Exeter, UK*, pages 1–6, 2011.

JC Nicol, AJ Illingworth, T Darlington, and M Kitchen. Quantifying errors due to frequency changes and target location uncertainty for radar refractivity retrievals. *Journal of Atmospheric and Oceanic Technology*, 30(9):2006–2024, 2013.

John C Nicol, Anthony J Illingworth, and K Bartholomew. The potential of 1 h refractivity changes from an operational C-band magnetron-based radar for numerical weather prediction validation and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1209–1218, 2014.

S Nishizawa, H Yashiro, Y Sato, Y Miyamoto, and H Tomita. Influence of grid aspect ratio on planetary boundary layer turbulence in large-eddy simulations. *Geoscientific Model Development*, 8 (10):3393–3419, 2015.

D Norenberg, Susanne Crewell, U Lohnert, Th Rose, and Antonio Martellucci. A novel ground-based microwave radiometer for high precision atmospheric observations between 10 and 90 Ghz. In *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages III–919. IEEE, 2008.

George Ohring and Peter Bauer. The use of cloud and precipitation observations in data assimilation (CPDA). *Quarterly Journal of the Royal Meteorological Society*, 137(661):1933–1933, 2011.

Kozo Okamoto, Yohei Sawada, and Masaru Kunii. Comparison of assimilating all-sky and clear-sky infrared radiances from Himawari-8 in a mesoscale system. *Quarterly Journal of the Royal Meteorological Society*, 145(719):745–766, 2019.

Jason A Otkin. Clear and cloudy sky infrared brightness temperature assimilation using an ensemble Kalman filter. *Journal of Geophysical Research: Atmospheres*, 115(D19), 2010.

Jason A Otkin. Assimilation of water vapor sensitive infrared brightness temperature observations during a high impact weather event. *Journal of Geophysical Research: Atmospheres*, 117(D19), 2012.

Jason A Otkin, Roland Potthast, and Amos S Lawless. Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Monthly Weather Review*, 146(1):263–285, 2018.

Christian Pagé, Luc Fillion, and Peter Zwack. Diagnosing summertime mesoscale vertical motion: implications for atmospheric data assimilation. *Monthly Weather Review*, 135(6):2076–2094, 2007.

Thomas Pangaud, Nadia Fourrie, Vincent Guidard, Mohamed Dahoui, and Florence Rabier. Assimilation of AIRS radiances affected by mid-to low-level clouds. *Monthly Weather Review*, 137(12): 4276–4292, 2009.

Anastasios Papadopoulos, Themis G Chronis, and Emmanouil N Anagnostou. Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model. *Monthly Weather Review*, 133(7):1961–1977, 2005.

Claire L Parkinson. Aqua: An Earth-observing satellite mission to examine water and other climate variables. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2):173–183, 2003.

David F Parrish and John C Derber. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.

EG Pavelin, SJ English, and JR Eyre. The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 134(632): 737–749, 2008.

Steven E Peckham, Tatiana G Smirnova, Stanley G Benjamin, John M Brown, and Jaymes S Kenyon. Implementation of a digital filter initialization in the wrf model and its application in the rapid refresh. *Monthly Weather Review*, 144(1):99–106, 2016.

Pennsylvania State University and National Center for Atmospheric Research. Fifth-generation Penn State/NCAR Mesoscale Model. `http://www2.mmm.ucar.edu/mm5`.

Ralph Alvin Petersen, Lee Cronce, Richard Mamrosh, Randy Baker, and Patricia Pauley. On the impact and future benefits of amdar observations in operational forecasting: Part ii: water vapor observations. *Bulletin of the American Meteorological Society*, 97(11):2117–2133, 2016.

Chiara Piccolo and Mike Cullen. Adaptive mesh method in the Met Office variational data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):631–640, 2011.

Chiara Piccolo and Mike Cullen. A new implementation of the adaptive mesh transform in the Met Office 3D-Var system. *Quarterly Journal of the Royal Meteorological Society*, 138(667):1560–1570, 2012.

Paul Poli, Joanna Joiner, and E Robert Kursinski. 1DVar analysis of temperature and humidity using GPS radio occultation refractivity data. *Journal of Geophysical Research: Atmospheres*, 107(D20), 2002.

Derek J Posselt and Craig H Bishop. Nonlinear data assimilation for clouds and precipitation using a Gamma-inverse Gamma ensemble filter. *Quarterly Journal of the Royal Meteorological Society*, 2018.

Roland Potthast, Anne Walter, and Andreas Rhodin. A localized adaptive particle filter within an operational nwp framework. *Monthly Weather Review*, 147(1):345–362, 2019.

Catherine Prigent, Filipe Aires, and William B Rossow. Land surface microwave emissivities over the globe for a decade. *Bulletin of the American Meteorological Society*, 87(11):1573–1584, 2006.

F Rabier, A McNally, E Andersson, P Courtier, P Unden, J Eyre, A Hollingsworth, and F Bouttier. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1809–1829, 1998.

Soundararajan Raghavan. *Radar meteorology*, volume 27. Springer Science & Business Media, 2013.

F Rawlins, SP Ballard, KJ Bovis, AM Clayton, D Li, GW Inverarity, AC Lorenc, and TJ Payne. The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 133(623):347–362, 2007.

Richard Renshaw and Peter N Francis. Variational assimilation of cloud fraction in the operational Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, 137(661):1963–1974, 2011.

Nigel M Roberts and Humphrey W Lean. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 136(1):78–97, 2008.

Clive D Rodgers. *Inverse methods for atmospheric sounding: theory and practice*, volume 2. World scientific, 2000.

Antonio Rodriguez, Rolf Stuhlmann, Stephen Tjemkes, Donny M Aminou, Hendrik Stark, and Paul Blythe. Meteosat Third Generation: mission and system concepts. In *Infrared Spaceborne Remote Sensing and Instrumentation XVII*, volume 7453, page 74530C. International Society for Optics and Photonics, 2009.

Josep Roselló, Pierluigi Silvestrin, Roland Weigand, Salvatore d'Addio, Alberto García Rodríguez, and Gustavo López Risueño. Next generation of ESA's GNSS receivers for Earth Observation satellites. In *Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing,(NAVITEC), 2012 6th ESA Workshop on*, pages 1–8. IEEE, 2012.

William Sacher and Peter Bartello. Sampling errors in ensemble Kalman filtering. Part I: Theory. *Monthly Weather Review*, 136(8):3035–3049, 2008.

Vincent V Salomonson, William Barnes, Jack Xiong, Steve Kempler, and Ed Masuoka. An overview of the Earth observing system MODIS instrument and associated data systems performance. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 1174–1176. IEEE, 2002.

Yousuke Sato, Seiya Nishizawa, Hisashi Yashiro, Yoshiaki Miyamoto, Yoshiyuki Kajikawa, and Hirofumi Tomita. Impacts of cloud microphysics on trade wind cumulus: which cloud microphysics processes contribute to the diversity in a large eddy simulation? *Progress in Earth and Planetary Science*, 2(1):23, 2015.

R Saunders, S Crewell, R Gelaro, PJ Minnett, V-H Peuch, J Schmetz, D Turner, and C Velden. Chapter 2. Observations for global to convective scale models. In *Seamless prediction of the earth system: from minutes to months*, pages 15–36, 2015.

Christoph Schraff, Hendrik Reich, Andreas Rhodin, Annika Schomburg, Klaus Stephan, Africa Periáñez, and Roland Potthast. Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Quarterly Journal of the Royal Meteorological Society*, 142(696):1453–1472, 2016.

Craig S Schwartz and Zhiquan Liu. Convection-permitting forecasts initialized with continuously cycling limited-area 3DVar, ensemble Kalman filter, and 'hybrid' variational–ensemble data assimilation systems. *Monthly Weather Review*, 142(2):716–738, 2014.

Craig S Schwartz, Zhiquan Liu, Xiang-Yu Huang, Ying-Hwa Kuo, and Chin-Tzu Fong. Comparing limited-area 3DVar and hybrid variational-ensemble data assimilation methods for typhoon track forecasts: Sensitivity to outer loops and vortex relocation. *Monthly Weather Review*, 141(12):4350–4372, 2013.

Y Seity, P Brousseau, S Malardel, G Hello, P Bénard, F Bouttier, C Lac, and V Masson. The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3):976–991, 2011.

Ehouarn Simon and Laurent Bertino. Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment. *Ocean Science*, 5(4):495–510, 2009.

M Široká, C Fischer, V Cassé, R Brožková, and J-F Geleyn. The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteorology and Atmospheric Physics*, 82(1-4):227–244, 2003.

RNB Smith. A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, 116(492):435–460, 1990.

Chris Snyder and Fuqing Zhang. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review*, 131(8):1663–1677, 2003.

Hajoon Song, Christopher A Edwards, Andrew M Moore, and Jerome Fiechter. Incremental four-dimensional variational data assimilation of positive-definite oceanic variables using a logarithm transformation. *Ocean Modelling*, 54:1–17, 2012.

Graeme Stephens, David Winker, Jacques Pelon, Charles Trepte, Deborah Vane, Cheryl Yuhas, Tristan Ľ2019ecuyer, and Matthew Lebsock. CloudSat and CALIPSO within the A-Train: ten years of actively observing the Earth system. *Bulletin of the American Meteorological Society*, 99(3):569–581, 2018.

O Stiller and SP Ballard. Efficient moist physics schemes for data assimilation. I: Large-scale clouds and condensation. *Quarterly Journal of the Royal Meteorological Society*, 135(640):707–720, 2009.

Andrea Storto and Roger Randriamampianina. The relative impact of meteorological observations in the Norwegian regional model as determined using an energy norm-based approach. *Atmospheric Science Letters*, 11(1):51–58, 2010.

Juanzhen Sun. Convective-scale assimilation of radar data: progress and challenges. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3439–3463, 2005a.

Juanzhen Sun. Initialization and numerical forecasting of a supercell storm observed during STEPS. *Monthly Weather Review*, 133(4):793–813, 2005b.

Juanzhen Sun and N Andrew Crook. Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments. *Journal of the Atmospheric Sciences*, 54(12):1642–1661, 1997.

Juanzhen Sun, Stanley B Trier, Qingnong Xiao, Morris L Weisman, Hongli Wang, Zhuming Ying, Mei Xu, and Ying Zhang. Sensitivity of 0–12-h warm-season precipitation forecasts over the central United States to model initialization. *Weather and Forecasting*, 27(4):832–855, 2012.

Juanzhen Sun, Ming Xue, James W Wilson, Isztar Zawadzki, Sue P Ballard, Jeanette Onvlee-Hooimeyer, Paul Joe, Dale M Barker, Ping-Wah Li, Brian Golding, et al. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bulletin of the American Meteorological Society*, 95(3):409–426, 2014.

Xiaoli Sun, James B Abshire, Jan F McGarry, Gregory A Neumann, James C Smith, John F Cavanaugh, David J Harding, H Jay Zwally, David E Smith, and Maria T Zuber. Space lidar developed at the NASA Goddard Space Flight Center – The first 20 years. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3):1660–1675, 2013.

Y Qiang Sun and Fuqing Zhang. Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *Journal of the Atmospheric Sciences*, 73(3):1419–1438, 2016.

O. Talagrand. Variational assimilation. In W. Lahoz, B. Khattatov, and R. Ménard, editors, *Data Assimilation: Making Sense of Observations*, pages 41–67. Springer, Heidelberg, 2010.

Yongming Tang, Humphrey W Lean, and Jorge Bornemann. The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorological Applications*, 20(4):417–426, 2013.

Bertrand Thomas, M Brandt, A Walber, H Gibson, M Philipp, G Sonnabend, M Benzazaa, R Gonzalez, M Bergada, J Martinez, et al. Millimeter & sub-millimeter wave radiometer instruments for the next generation of polar orbiting meteorological satellites – MetOp-SG. In *Infrared, Millimeter, and Terahertz waves (IRMMW-THz), 2014 39th International Conference on*, pages 1–3. IEEE, 2014.

Adrian M Tompkins and Marta Janisková. A cloud scheme for data assimilation: Description and initial tests. *Quarterly Journal of the Royal Meteorological Society*, 130(602):2495–2517, 2004.

Wenxue Tong, Gang Li, Juanzhen Sun, Xiaowen Tang, and Ying Zhang. Design strategies of an hourly update 3DVar data assimilation system for improved convective forecasting. *Weather and Forecasting*, 31(5):1673–1695, 2016.

Yannick Trémolet. Accounting for an imperfect model in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 132(621):2483–2504, 2006.

David D Turner, RA Ferrare, LA Heilman Brasseur, WF Feltz, and TP Tooman. Automated retrievals of water vapor and aerosol profiles from an operational Raman lidar. *Journal of Atmospheric and Oceanic Technology*, 19(1):37–50, 2002.

DD Turner and U Löhnert. Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based atmospheric emitted radiance interferometer (AERI). *Journal of Applied Meteorology and Climatology*, 53(3):752–771, 2014.

John D Tuttle and G Brant Foote. Determination of the boundary layer airflow from a single Doppler radar. *Journal of Atmospheric and oceanic Technology*, 7(2):218–232, 1990.

Peter Jan van Leeuwen. Comment on 'Data assimilation using an ensemble Kalman filter technique'. *Monthly Weather Review*, 127(6):1374–1377, 1999.

Peter Jan van Leeuwen. Particle filtering in geophysical systems. *Monthly Weather Review*, 137(12): 4089–4114, 2009.

Peter Jan van Leeuwen, Hans R Künsch, Lars Nerger, Roland Potthast, and Sebastian Reich. Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal of the Royal Meteorological Society*, 2019. doi: doi:10.1002/qj.3551.

T Vukicevic, M Sengupta, AS Jones, and T Vonder Haar. Cloud-resolving satellite data assimilation: Information content of IR window observations and uncertainties in estimation. *Journal of the Atmospheric Sciences*, 63(3):901–919, 2006.

Tomislava Vukicevic. Comments on 'Issues regarding the assimilation of cloud and precipitation data'. *Journal of the Atmospheric Sciences*, 65(10):3344–3347, 2008.

Tomislava Vukicevic, T Greenwald, M Zupanski, D Zupanski, T Vonder Haar, and AS Jones. Mesoscale cloud state estimation from visible and infrared satellite radiances. *Monthly Weather Review*, 132 (12):3066–3077, 2004.

Joanne A Waller, Sarah L Dance, Amos S Lawless, Nancy K Nichols, and JR Eyre. Representativity error for temperature and humidity using the Met Office high-resolution model. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1189–1197, 2014.

Joanne A Waller, Sarah L Dance, and Nancy K Nichols. Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Quarterly Journal of the Royal Meteorological Society*, 142(694):418–431, 2016.

Hongli Wang, Juanzhen Sun, Xin Zhang, Xiang-Yu Huang, and Thomas Auligné. Radar data assimilation with WRF 4D-Var. Part I: System development and preliminary testing. *Monthly Weather Review*, 141(7):2224–2244, 2013.

Hongli Wang, Xiang-Yu Huang, Juanzhen Sun, Dongmei Xu, Man Zhang, Shuiyong Fan, and Jiqin Zhong. Inhomogeneous background error modeling for WRF-Var using the NMC method. *Journal of Applied Meteorology and Climatology*, 53(10):2287–2309, 2014.

Junhong Wang, David J Carlson, David B Parsons, Terrence F Hock, Dean Lauritsen, Harold L Cole, Kathryn Beierle, and Edward Chamberlain. Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication. *Geophysical Research Letters*, 30(16), 2003.

Xuguang Wang and Ting Lei. GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP global forecast system. *Monthly Weather Review*, 142(9):3303–3325, 2014.

Yongming Wang and Xuguang Wang. Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: Methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Monthly Weather Review*, 145(4):1447–1471, 2017.

Eric Wattrelot, Olivier Caumont, and Jean-Francois Mahfouf. Operational implementation of the 1D+ 3D-Var assimilation method of radar reflectivity data in the AROME model. *Monthly Weather Review*, 142(5):1852–1873, 2014.

PP Weston, W Bell, and JR Eyre. Accounting for correlated error in the assimilation of high-resolution sounder data. *Quarterly Journal of the Royal Meteorological Society*, 140(685):2420–2429, 2014.

Stephen S Weygandt, SG Benjamin, and J Brown. Radar reflectivity-based initialization of precipitation systems using a diabatic digital filter within the Rapid Update Cycle. In *18th Conf. Num. Wea. Pred.* Citeseer, 2007.

Daniel S Wilks. *Statistical Methods in the Atmospheric Sciences*. Elsevier, Oxford, 2011. ISBN 978-0-12-385022-5.

David M Winker, William H Hunt, and Matthew J McGill. Initial performance assessment of CALIOP. *Geophysical Research Letters*, 34(19), 2007.

WMO. OSCAR: Observing Systems Capability Analysis and Review tool. URL https://www.wmo-sat.info/oscar/spacecapabilities.

Qingnong Xiao and Juanzhen Sun. Multiple-radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002. *Monthly Weather Review*, 135 (10):3381–3404, 2007.

Qingnong Xiao, Ying-Hwa Kuo, Juanzhen Sun, Wen-Chau Lee, Dale M Barker, and Eunha Lim. An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at landfall. *Journal of applied meteorology and climatology*, 46(1):14–22, 2007.

Qin Xu. Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *Journal of the Atmospheric Sciences*, 53(8):1123–1142, 1996.

Ming Xue, Kelvin K Droegemeier, and Vince Wong. The Advanced Regional Prediction System (ARPS)–a multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteorology and atmospheric physics*, 75(3-4):161–193, 2000.

Ming Xue, Kelvin K Droegemeier, V Wong, A Shapiro, Keith Brewster, Fred Carr, D Weber, Ym Liu, and D Wang. The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorology and atmospheric physics*, 76(3-4):143–165, 2001.

Shu-Chih Yang, Eugenia Kalnay, and Takemasa Miyoshi. Accelerating the EnKF spinup for typhoon assimilation and prediction. *Weather and Forecasting*, 27(4):878–897, 2012.

Jun-Ichi Yano, Michał Z Ziemiański, Mike Cullen, Piet Termonia, Jeanette Onvlee, Lisa Bengtsson, Alberto Carrassi, Richard Davy, Anna Deluca, Suzanne L Gray, et al. Scientific challenges of convective-scale numerical weather prediction. *Bulletin of the American Meteorological Society*, 99 (4):699–710, 2018.

Nusrat Yussouf, Edward R Mansell, Louis J Wicker, Dustan M Wheatley, and David J Stensrud. The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single-and double-moment microphysics schemes. *Monthly Weather Review*, 141(10): 3388–3412, 2013.

Žiga Zaplotnik, Nedjeljka Žagar, and Nils Gustafsson. An intermediate-complexity model for four-dimensional variational data assimilation including moist processes. *Quarterly Journal of the Royal Meteorological Society*, 144(715):1772–1787, 2018.

Fuqing Zhang, Masashi Minamide, and Eugene E Clothiaux. Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophysical research letters*, 43(6):2954–2963, 2016.

Sara Q Zhang, T Matsui, S Cheung, M Zupanski, and C Peters-Lidard. Impact of assimilated precipitation-sensitive radiances on the NU-WRF simulation of the West African Monsoon. *Monthly Weather Review*, 145(9):3881–3900, 2017.

Yanqiu Zhu, Emily Liu, Rahul Mahajan, Catherine Thomas, David Groff, Paul Van Delst, Andrew Collard, Daryl Kleist, Russ Treadon, and John C Derber. All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Monthly Weather Review*, 144(12):4709–4735, 2016.

X Zou. Tangent linear and adjoint of 'on-off' processes and their feasibility for use in 4-dimensional variational data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 49(1):3–31, 1997.

Xiaolei Zou, IM Navon, and JG Sela. Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus A*, 45(5), 1993.