



# A method for retrieving clouds with satellite infrared radiances using the particle filter

Dongmei Xu<sup>1,2</sup>, Thomas Auligné<sup>2</sup>, Gaël Descombes<sup>2</sup>, and Chris Snyder<sup>2</sup>

<sup>1</sup>Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Joint International Research Laboratory of Climate and Environment Change (ILCEC)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing 210044, China

<sup>2</sup>National Center for Atmospheric Research, Boulder, Colorado 80301, USA

Correspondence to: Dongmei Xu (xdmjolly@sina.com)

Received: 11 June 2016 – Published in Geosci. Model Dev. Discuss.: 4 July 2016

Revised: 13 September 2016 – Accepted: 10 October 2016 – Published: 2 November 2016

**Abstract.** Ensemble-based techniques have been widely utilized in estimating uncertainties in various problems of interest in geophysical applications. A new cloud retrieval method is proposed based on the particle filter (PF) by using ensembles of cloud information in the framework of Grid-point Statistical Interpolation (GSI) system. The PF cloud retrieval method is compared with the Multivariate Minimum Residual (MMR) method that was previously established and verified. Cloud retrieval experiments involving a variety of cloudy types are conducted with the PF and MMR methods with measurements of infrared radiances on multi-sensors onboard both geostationary and polar satellites, respectively. It is found that the retrieved cloud masks with both methods are consistent with other independent cloud products. MMR is prone to producing ambiguous small-fraction clouds, while PF detects clearer cloud signals, yielding closer heights of cloud top and cloud base to other references. More collections of small-fraction particles are able to effectively estimate the semi-transparent high clouds. It is found that radiances with high spectral resolutions contribute to quantitative cloud top and cloud base retrievals. In addition, a different way of resolving the filtering problem over each model grid is tested to better aggregate the weights with all available sensors considered, which is proven to be less constrained by the ordering of sensors. Compared to the MMR method, the PF method is overall more computationally efficient, and the cost of the model grid-based PF method scales more directly with the number of computing nodes.

## 1 Introduction

Modern polar orbiting and geostationary airborne instruments provide researchers unprecedented opportunities for remote sensing of the Earth with continuous flows and almost complete spectral coverage of data. The primary cloud retrieval products from satellites are cloud mask (CM), cloud height (CH), effective cloud fraction (CF) and vertical structures of clouds with larger temporal and spatial scales. These cloud retrievals provide an immense and valuable combination for better initializing hydrometeors in numerical weather prediction (NWP), (Wu and Smith, 1992; Hu et al., 2006; Bayler et al., 2000; Auligné et al., 2011) regulating the radiation budget for the planet and understanding the climate feedback mechanism (Brückner et al., 2014; Rossow and Schiffer, 1991; Rossow et al., 1993). Advanced cloud retrieval methods are able to retrieve clouds with multi-spectral techniques (Menzel et al., 1983; Platnick et al., 2003), among which the minimization methods usually directly utilize the difference between the modeled clear sky and the observed cloudy infrared (IR) radiances (e.g., the minimum residual method, Eyre and Menzel, 1989; the Minimum Local Emissivity Variance method, Huang et al., 2004; and the Multivariate Minimum Residual method, Auligné, 2014a). In particular, the Multivariate Minimum Residual (MMR) method is retrieving three-dimensional multi-layer clouds by minimizing a cost function at each field of view (FOV) (Auligné, 2014b; Xu et al., 2015). MMR has been proven to be reliable in retrieving the quantitative three-dimensional cloud fractions with infrared radiances from multiple infrared instru-

ments. However, MMR has limitations in several aspects due to its use of minimization for solution: (1) part of the control variables accounting for the cloud fraction for some certain levels are under-observed since the channels are not sensitive to the existence of clouds for those heights; (2) when clouds at different heights show opacities with the same spectral signal, MMR could lose the ability to distinguish solutions involving clouds at those levels; (3) the computational cost for the minimization procedure in MMR is rather considerable.

Ensemble-based techniques, which usually reside in short-term ensemble forecasting (Berrocal et al., 2007; Shen and Min, 2015), assembling existing model outputs (e.g., cloud retrievals) from varying algorithms (Zhao et al., 2012), or ensemble Kalman filter (EnKF) in diversified forms (Snyder and Zhang, 2003), have been widely developed in order to estimate the uncertainties of various problems in geophysical applications. To better account for the non-linearity between the observed radiance and the retrieval parameter, a novel prototype for detecting clouds and retrieving their vertical extension inspired by the particle filter (Snyder and Zhang, 2003; van Leeuwen, 2010; Shen and Tang, 2015) technique and Bayesian theory (Karlsson et al., 2015) is proposed in this study. As a competitive alternative for MMR, the PF retrieval method has same critical inputs required and cloud retrieval products as in MMR. A brief description of MMR and the new PF cloud retrieval algorithm are provided in the following section. Section 3 describes the background model, the data assimilation system, the radiative transfer models (RTMs) and the radiance observations applied in this study. Model configurations are also illustrated in Sect. 3. In Sect. 4, the single test within one FOV is conducted before the performance of PF method is assessed by comparing its cloud retrievals with those from MMR and other operational cloud products. Section 4 also discusses the computational performance for the two methods. The conclusion and anticipated future work are outlined in Sect. 5.

## 2 Methodology

Essentially, the PF cloud retrieval scheme retrieves clouds with the same critical inputs requested (i.e., clear-sky radiance from the radiative transfer model and the observed radiance) and the same cloud retrievals as outputs (i.e., three-dimensional cloud fractions, which is defined as the fraction of top of cloud as seen from a sensor) with the MMR method. Both cloud retrieval schemes consist of finding cloud fractions that allow for best fit between the cloudy radiance from model and the observation. We use  $c^1, c^2, \dots, c^K$  to denote the array of vertical effective cloud fractions for  $K$  model levels ( $c^1$  for the surface and  $c^K$  for the model top) and  $c^0$  as the fraction of clear sky with  $0 \leq c^k \leq 1$ ,  $\forall k \in [0, K]$ . The constraint for the cloud fraction is as follows:

$$\sum_{k=0}^K c^k = 1 \quad (1)$$

In this study, a cloud on one model level with a given fraction  $c^k$  is assumed to block the radiation from its lower model levels. The radiation originating from its lower levels is assumed to contribute to the top of atmosphere radiance observed by the satellites only with the residual fractions.

The MMR method is an approach to retrieve cloud fractions using the minimization technique. The residual of the modeled radiance and the observation is normalized by the observed radiance, which results in the following cost function, using  $c^k$ ,  $\forall k \in [0, K]$  as the control variables:

$$J(c^0, c^1, c^2, \dots, c^K) = \frac{1}{2} \sum_v \left[ \frac{R_v^{\text{cloud}} - R_v^{\text{obs}}}{R_v^{\text{obs}}} \right]^2, \quad (2)$$

where  $R_v^{\text{cloud}}$  is the modeled cloudy radiance, and  $R_v^{\text{obs}}$  the observed radiance at frequency  $v$ . This vertical cloud fraction  $c^1, c^2, \dots, c^K$  and  $c^0$  are control variables for the cost function, where the simulated  $R_v^{\text{cloud}}$  is defined as

$$R_v^{\text{cloud}}(c^0, c^1, c^2, \dots, c^K) = c^0 R_v^0 + \sum_{k=1}^K c^k R_v^k. \quad (3)$$

Here  $R_v^k$  is the radiance calculated assuming an overcast black cloud at the model level  $k$  and  $R_v^0$  the radiance calculated in the clear sky. Both  $R_v^k$  and  $R_v^0$  are calculated using a forward radiative transfer model with model profiles of temperature and moisture as inputs. Details of the schematic of the MMR method can be referred to in Descombes et al. (2014) and Xu et al. (2015).

Particle filter (PF) approach is one of the non-linear filters for data assimilation procedures to best estimate the initial state of a system or its parameters  $x_t$ , which describes the time evolution of the full probability density function  $p(x_t)$  conditioned by the dynamics and the observations. Similar to the study in Mechri et al. (2014), the bibliography on PF focuses on estimating the parameters, which are cloud fractions  $c^k$  in Eq. (3), in this study. While MMR retrieves the cloud fractions on each model vertical level by minimizing a cost function, PF calculates posterior weights for each ensemble member based on the observation likelihood given that member. In its simplest form, PF works by initializing a collection of cloud profiles as particles and then estimating the cloud distributions by averaging those particles with their corresponding weights. Each particle's weight is computed with the difference between the modeled cloudy radiance from the particle and the observed radiance.

As the probabilities of the cloud distribution are fully represented by the initial particles, of particular interest is to evaluate different particle initialization schemes in the PF method. Explicitly, the definition of particles corresponds to ensemble

members; i.e., one cloud profile as one of particles is corresponding to an ensemble member.

Two approaches for generating particles are initially designed; the first one is to generate the perturbed samples  $C_b^i (\forall i \in [1, n])$  from the cloud profile in the background denoted as  $C_b = (c_b^0, c_b^1, \dots, c_b^K)$  by inflating (deflating) the clouds with small magnitudes ( $C_b = \alpha \times C_b, \alpha = 50, 55, \dots, 150\%$ ) and moving upward (downward) with  $\delta z = +5, +4, \dots, -1, \dots, -5$  as the vertical magnitude, where  $n$  is the sample size. The perturbed cloud fractions are designated to replenish the ensemble by introducing the prior information of the cloud distributions from the background and to increase the ensemble spread.

Besides those perturbed particles, to represent the existence of a one-layer cloud on each model level with an even chance, another diversity set of profiles  $C_b^i (\forall i \in [1, K + 1])$  are also initialized, among which  $C_b^i$  stands for the profile with 100% cloud fraction on the model level  $i$  ( $c^i = 100\%$ ) and 0% cloud on the rest levels. In particular,  $C_b^0$  defines 100% clear ( $c^0 = 1$ ). It is also interesting to discretize the initial particles by setting the one-layer cloud with the value of  $c^i$  from 100 to 0% (e.g., 100, 90, 80, ..., 0% with 10% as the interval) and further from 100 to 0% (e.g., 100, 99, 98, 97, ..., 0% with 1% as the interval). In this cases,  $c^0 = 1 - c^i$ . For each particle  $C_b^i$ , its simulated cloudy radiance  $R_{v,i}^{\text{cloud}}$  from the model background can be obtained with Eq. (3).

A cost function  $J_o$  is defined for each particle to measure how the particle fit the observation as

$$J_o = \left( \frac{R_v^{\text{obs}} - R_{v,i}^{\text{cloud}}}{\sigma} \right)^2 \quad (4)$$

$\sigma$  is the specified observation error, which can be referred to in the first paragraph in Sect. 4.1. The weight  $w^i$  for each particle  $C_b^i$  is thus calculated by comparing the simulated  $R_{v,i}^{\text{cloud}}$  and the observation  $R_v^{\text{obs}}$  using the exponential function by accumulating the  $J_o$  for multiple frequency as

$$w^i = e^{-\sum_v \left( \frac{R_v^{\text{obs}} - R_{v,i}^{\text{cloud}}}{\sigma} \right)^2}, \quad (5)$$

$\forall i \in [1, n]$ . Here  $n$  is the particle size and  $\sigma$  is the specified observation error, which can be referred to in the first paragraph in Sect. 4.1. The final analyzed  $C_a$  is obtained by averaging the background particles  $C_b^i$  with their corresponding weight, as

$$C_a = \sum_{i=1}^p w^i C_b^i \quad (6)$$

In Eq. (6), the constraint referred to in Eq. (1) is not respected. Thus, after the analysis step for the particle filter,

the final averaged cloud fractions  $c_a^k$  are normalized by

$$c_a^k = \frac{c^k}{\sum_{k=0}^K c^k}, \quad (7)$$

where  $\forall k \in [0, K]$ .

### 3 Data and model configurations

#### 3.1 Data

The Advanced Infrared Sounder (AIRS), the Infrared Atmospheric Sounding Interferometer (IASI) and the Cross-track Infrared Sounder (CrIS) are among the most advanced hyperspectral infrared sounders and thus are applied for retrieving clouds with hundreds of channels (Blumstein et al., 2004; Aumann et al., 2003; Xu et al., 2013; Bao et al., 2015; Smith et al., 2015). The Radiance measurements from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Earth Observing System (EOS) Terra or Aqua satellites are also well suited to extracting valuable cloud information from the 36 spectral broadband frequencies in the visible, near-infrared and infrared regions at high spatial resolution (1–5 km) (Ackerman et al., 1998). Apart from the IR radiances from polar satellites, the Geostationary Operational Environmental Satellite (GOES) Imager (Menzel and Purdom, 1994) provides a continuous stream of data over the observing domain. In this study, GOES-13 (east) and GOES-15 (west) are also utilized to obtain cloud fractions over the continental United States (CONUS) domain. The GOES Imager used in this study is a five-channel (one visible, four infrared) imaging radiometer designed to sense radiant and solar reflected energy. The instrument parameters for the sensors and the setups for channel selections can be found in Xu et al. (2015).

#### 3.2 WRF, GSI and the radiative transfer model

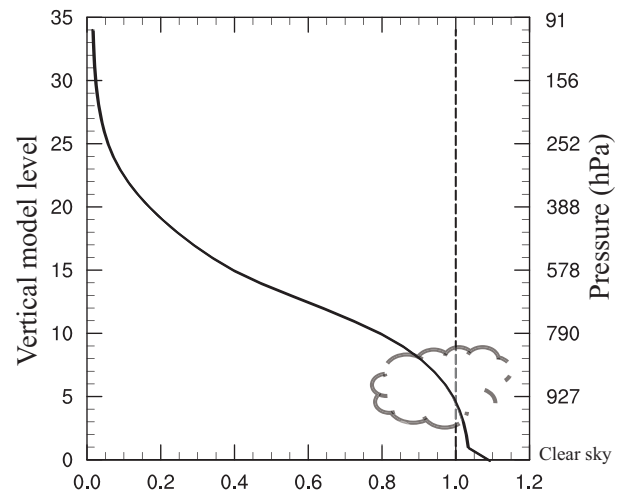
The background fields are processed running the Weather Research and Forecast (WRF) model (Skamarock et al., 2008). The MMR and PF cloud retrieval algorithms are both implemented based on the Gridpoint Statistical Interpolation (GSI) data assimilation system (Wu et al., 2002) (Kleist et al., 2009), which is a widely used data assimilation system in operations and research in NWP. GSI is capable of ingesting a large variety of satellite radiance observations and has developed capabilities for data thinning, quality control and satellite radiance bias correction. The Community Radiative Transfer Model (Liu and Weng, 2006) (Han et al., 2006) was used as the radiance forward operator for computing the clear-sky radiance and the radiance given overcast clouds at each model level.

### 3.3 Model configurations

The WRF is configured with  $415 \times 325$  horizontal grids at 15 km grid spacing, and 40 vertical levels up to 50 hPa within the single CONUS domain. The MMR and PF cloud detection schemes search the cloud top using approximately 150 hPa as the highest extent for most cloudy cases. Other clouds higher than 150 hPa, e.g., an anvil cloud in a mature thunderstorm around tropopause at low latitude region, will also be explored in future studies. Channels in the long-wave region are utilized following the channel selection scheme in (Xu et al., 2015). Since the final retrieval clouds are on model grids, the retrieved cloud fractions within one FOV are essentially extrapolated to its four neighboring model grid points. Generally, for each FOV, the retrieved cloud fractions are extrapolated to its four neighboring model grid points. For polar satellite pixels, the representative cloud fractions are extrapolated with an adaptive radius with respect to their scan positions. The cloud detecting procedure for retrieving clouds is conducted for each FOV from each individual sensor independently and sequentially. Since the clouds are retrieved FOV by FOV and the clouds on grids are referred to immediately after one FOV is completed, there is no obvious accuracy loss of radiance observations using this conservative method.

## 4 Experiments and results

The PF experiments apply two groups of particles as mentioned in Sect. 2, among which the group 2 particles contains solely 100 % one-layer clouds. To reveal how the setup of the initial particles impacts the results, apart from the MMR and PF experiments, we included another advanced experiment, denoted as Advanced PF (APF). APF requires more sampled particles including ranges of cloud fractions spanning from 0 to 100 % at intervals of 10 %. An additional experiment “APFg2”, similar to APF but excluding the perturbed particles from the background in group 1 introduced in Sect. 2, was conducted to evaluate the added values from the group 1 particles. In this section, cloud retrieval experiments for several cases containing clouds of a variety of types are conducted for comparison reason. The GOES imager retrieved products from National Aeronautics and Space Administration (NASA; Langley cloud and radiation products) are applied as a reference to validate the cloud retrieving methods for the CONUS domain with a large and uniform coverage of cloud mask. In addition, the retrieved cloud products were also compared to available CloudSat (Stephens et al., 2002) and MODIS level 2 cloud products (Platnick et al., 2003) archived by the CloudSat Data Processing Center in Colorado State and NASA, respectively.

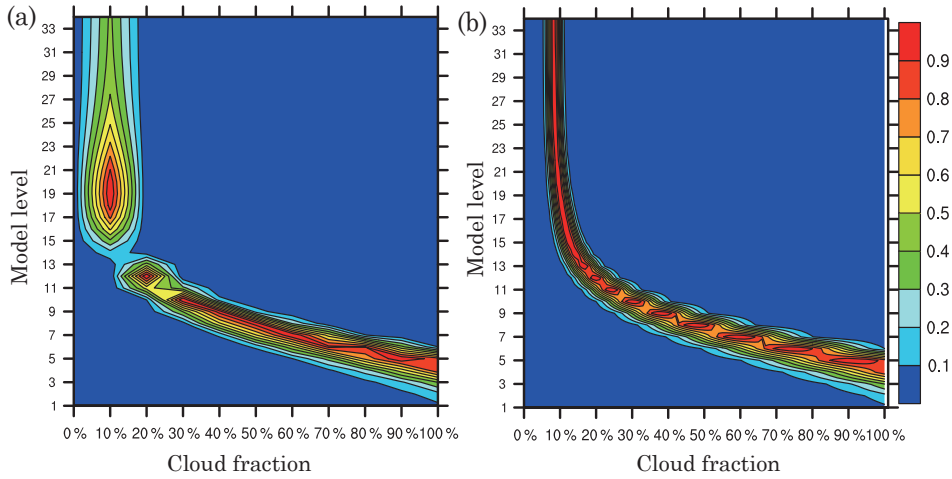


**Figure 1.** Ratio of the overcast radiances vs. the observed radiance starting from the level 1. The ratio of the clear-sky radiance normalized by the observed radiance corresponds to level 0 (see text for explanation) for the GOES-Imager for channel 5. The approximate pressures corresponding to the model levels are also denoted.

### 4.1 Single test at one field of view

The PF cloud retrieving algorithm retrieves the cloud distributions by averaging those initial particles with their weights. Before the real case experiments are carried out over the whole domain, we conduct a single cloud retrieving test at one FOV to understand what differences can be explained by the differences in the basic initial particles. In Eq. (5), the observation error  $\sigma$  can be set proportional to the observation equaling  $\frac{R_v^{\text{obs}}}{r}$ , where  $r$  is the prescribed ratio. Thus, the cloud signals on each level  $k$  are virtually determined by the extent of how close the  $\frac{R_v^k}{R_v^{\text{obs}}}$  (and  $\frac{R_v^0}{R_v^{\text{obs}}}$  for the clear part) gets to 1. An example of the ratio of the overcast radiance and the observed radiance  $\frac{R_v^k}{R_v^{\text{obs}}}$  for each model level is given in Fig. 1 of the GOES-Imager for the channel 5 ( $\sim 13.00 \mu\text{m}$ ). The clear-sky radiance normalized by the observed radiance  $\frac{R_v^0}{R_v^{\text{obs}}}$  is also shown at the level 0 (Fig. 1). It is expected that the overcast radiance from the RTM decreases with the rising of the altitude. The cloud signal is strongest around level 5, where  $R_v^k$  fits  $R_v^{\text{obs}}$  most closely. The cloud retrievals depend not only on the basic input profiles (i.e., the overcast radiance on each level from RTM normalized by the observed radiance and the clear-sky radiance from RTM normalized by the observed radiance) and but also on the algorithm applied for resolving the problem (e.g., MMR and PF in this study).

To reveal the roles of various initial particles, Fig. 2a shows the weights for different particles on the given FOV for channel 5 of the GOES-Imager for the case shown in Fig. 1. Particles in Fig. 1 include one-layer cloud in group 2 described in Sect. 2 with specified value of cloud fractions



**Figure 2.** The weights for different particles with specified cloud fractions on the  $x$  axis at one chosen model level shown on the  $y$  axis from 0 to 100 % (a) at the interval of 10 % and (b) at the interval of 1 %.

$c^k$  (on the  $x$  axis) on specified model levels  $k$  (on the  $y$  axis) from 10 to 100 % every 10 %. With a fraction  $c^k$  of one-layer cloud at a given level  $k$  and a fraction of  $c^0 = 1 - c^k$  of clear sky, the simulated cloudy radiance can be denoted as  $R_v^{\text{cloud}} = c^k R_v^k + (1 - c^k) R_v^0$ . Hence the theoretical one-layer cloud fraction is solved as  $c^k = \frac{R_v^0 - R_v^{\text{obs}}}{R_v^0 - R_v^k}$  by fitting  $R_v^{\text{cloud}}$  to  $R_v^0$ . As expected, for one-layer cloud with full fraction,  $c^5$  equals 100 %. In light of the concept that  $R_v^k > R_v^{k+1}$ , no cloud can be present below level 5 since this would imply a  $R_v^{\text{cloud}}$  larger than the observation (or a  $c^i$  larger than 100 %). It seems that clouds can be described by different possible states as particles with both large fractions and small fractions. Low clouds are easily estimated by one-layer cloud profile with large fractions (larger than 10 %). The particles with small-fraction high clouds gain some weights to retrieve high clouds. The particle with the one-layer cloud on level 13 seems to gain least weight compared to the others levels. The weights for the particles with cloud fractions from 0 to 100 % at the interval of 1 % are also presented in Fig. 2b. By including more small-fraction one-layer clouds, the clouds around level 13 can be reproduced by the group of refined particles with 1 % as the interval for approximately 10 % cloud fractions. However, changing the level of the cloud for the fixed fraction (10 %) does not seem to change the outgoing radiance much, probably due to the channel’s low weight function peak ( $\sim 750$  hPa).

The normalized  $J_o$  in Eq. (6) for different levels with a specific cloud fraction from 0 to 100 % every 10 % are shown in the bottom panel of Fig. 3, with 10 and 1 % as the intervals in Fig. 3c and d, respectively. Here,  $J_o$  can be further derived as

$$J_o = r^2 \left( 1 - c^0 \frac{R_v^0}{R_v^{\text{obs}}} - c^k \frac{R_v^k}{R_v^{\text{obs}}} \right)^2, \quad (8)$$

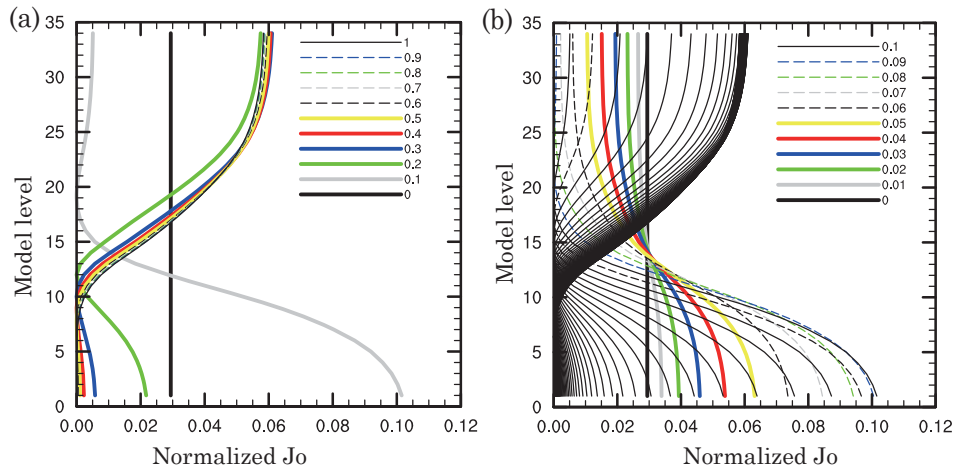
$$\text{with } \sigma = \frac{R_v^{\text{obs}}}{r} \text{ and } R_v^{\text{cloud}}(c^0, c^1, c^2, \dots, c^K) = c^0 R_v^0 + \sum_{k=1}^K c^k R_v^k.$$

From Fig. 3a, it is found that  $J_o$  is smallest around level 5 with 100 % cloud fraction (denoted as 1 in legend) for the thin black line, with respect to the fact that the overcast radiance fits the observed radiance most closely for level 5 approximately. The gray line with 10 % cloud fraction (0.1 in the legend) corresponds to the existence of a weight peak on level 19 in Fig. 2a. In addition, the gap between the gray line with 0.1 and the other lines from 0.2 to 1 explains why there’s less continuity around level 13. Figure 3b shows a similar pattern to Fig. 3a, except with densely distributed  $J_o$  values around the level 13 from 0.1 to 1 in the legend. Those contiguous black lines in Fig. 3b are associated with the set of particles with cloud fractions from 10 to 100 % at the interval of 1 %.

## 4.2 Cloud profiles

The retrieval experiments for a real case are conducted at 11:00 UTC, 3 June 2012, when AIRS measurements and the CloudSat “2B-GEOPROF” products (Mace, 2004) are available. The vertical cross sections of the cloud fraction field of a real case are illustrated to further check how different collections of initial particles impact the retrieved cloud profiles. The standard radar reflectivity profiles from the CloudSat are shown in Fig. 4a as the validation source; Fig. 4b, c and d show the cross sections of the cloud fractions along the CloudSat orbit tracks from the MMR, PF and APF experiments. The vertical structures of the clouds from MMR compare well with the radar reflectivity from CloudSat by retrieving the high clouds around 47° N and low clouds around 52° N. The PF experiment has difficulties in detecting the cloud tops appropriately. PF tends to detect a large quantity





**Figure 3.** The normalized  $J_o$  (a) at the interval of 10 % and (b) at the interval of 1 %. In (b), the normalized  $J_o$  from 0.1 to 1 are all denoted as black lines.

of low clouds; by adding a set of particles with small-fraction clouds in APF, higher clouds can be reproduced, which is consistent with the implications from Figs. 2b and 3b. APF detects clear strong cloud signals and removes the cloud fractions on near-surface levels around  $36^\circ$  N successfully. Since the existence of ground-layer radar reflectivity is likely corresponding to the strong reflection from the underlying surface of the Earth, the height of cloud bases of MMR and PF are not compared in this sub-section. The experiments with a larger size of particles including 0 to 20 % (at the interval of 1 %) plus 30 to 100 % (at the interval of 10 %) or of 0 to 100 % (at the interval of 1 %) one-layer cloud profiles (introduced in Sect. 2) yield similar results from APF but are much more costly (not shown).

The vertical profiles of the averaged cloud fractions from MMR, PF and APF are plotted in Fig. 5 at 11:00 UTC, 3 June 2012, with AIRS. Both MMR and PF experiments yield ambiguous cloud distributions, whereas APF retrieves much stronger cloud signals constrained between level 2 to level 20 (approximately from 950 to 400 hPa). More clouds around level 10 are retrieved (approximately 750 hPa) in MMR, while PF is prone to retrieving clouds near surface levels. Note that MMR retrieves much higher cloud tops and lower cloud bases compared to APF. The cloud base from PF is the lowest; the cloud top from MMR and PF is comparable. Only the APF-related methods will be further discussed in later sections, due to the absence of high clouds using PF.

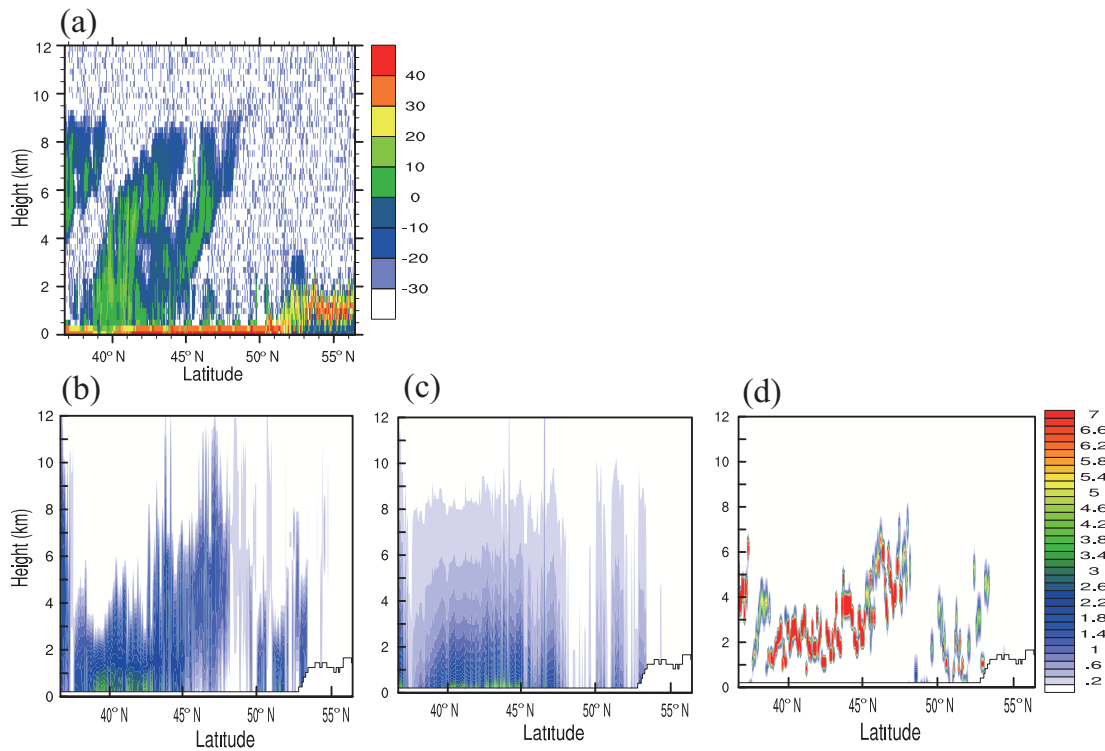
### 4.3 Cloud mask

Comparison experiments on real cases are further performed for over longer time period from 00:00 UTC, 12 December 2013, to 07:00 UTC, 12 December 2013. The cloud mask is marked as cloudy when there is a recognizable existence of cloud on any level from MMR or PF retrievals. Both the NASA GOES Imager products and the MMR-retrieved fields

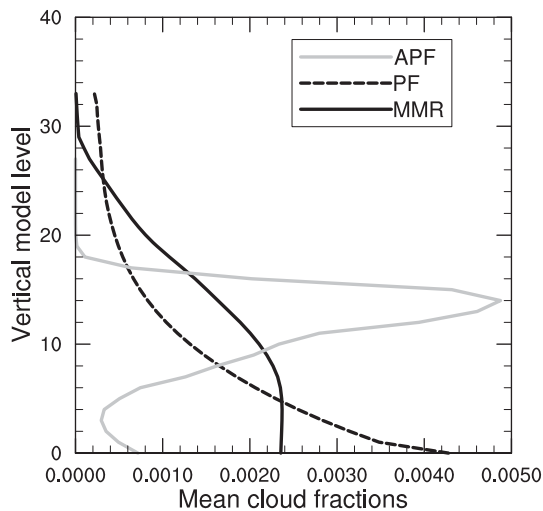
are interpolated to the same  $0.1^\circ \times 0.1^\circ$  latitude–longitude grid with 0 for clear and 1 for cloudy before the comparisons for verification. Figure 6 shows the *hits*, *false\_alarms* and *misses* locations with the use of GOES-Imager, MODIS, CrIS, AIRS and IASI radiances in the retrieval algorithms at 07:00 UTC, 12 December 2013. Note that cloud mask retrievals from both the MMR and APF hit the clear and cloudy events well in Fig. 6a and b. In most areas, the MMR experiment overestimated the cloud mask with more false alarm events compared to the APF experiment, since the MMR solution is an “overly smoothed” estimation of the true vertical profile. It seems that the accuracy of cloud detection is lower for areas with high altitude than under tropical conditions, indicating that the smaller lapse rate in the atmosphere will lead radiance less sensitivity to clouds over polar areas. Figure 6c shows the cloud mask results from the APFg2 experiment without the perturbed particles in group 1 introduced in Sect. 2. There is no large discrepancy between Fig. 6b and c, suggesting that the particles in group 2, which fully span the possibility of the cloud distributions, are more determinant in retrieving the cloud mask.

### 4.4 Cloud top and base pressure

The retrieved cloud top pressures (CTPs) and cloud bottom pressures (CBPs) from this study along with the NASA GOES cloud products are illustrated in Fig. 7. The CTPs from both methods are in good accordance with the NASA cloud products for high clouds (from 100 to 600 hPa) in Fig. 7a, c and e. The retrieved cloud top heights from MMR are overall higher than those from the NASA reference, especially for lower clouds at approximately 750–1000 hPa (e.g., between longitude  $-100$  and  $-90^\circ$ ). On the other hand, the CTPs from APF are much closer to those in the reference for both high and low clouds. APF overestimates the CBPs for some low clouds (putting the clouds too low) in Fig. 7f; the



**Figure 4.** (a) The radar reflectivity (units: dBZ) cross sections from CloudSat, (b) the MMR retrieved cloud fraction (units: %) cross sections, (c) the PF retrieved cloud fractions and (d) the APF retrieved cloud fractions valid at 11:00 UTC, 3 June 2012.



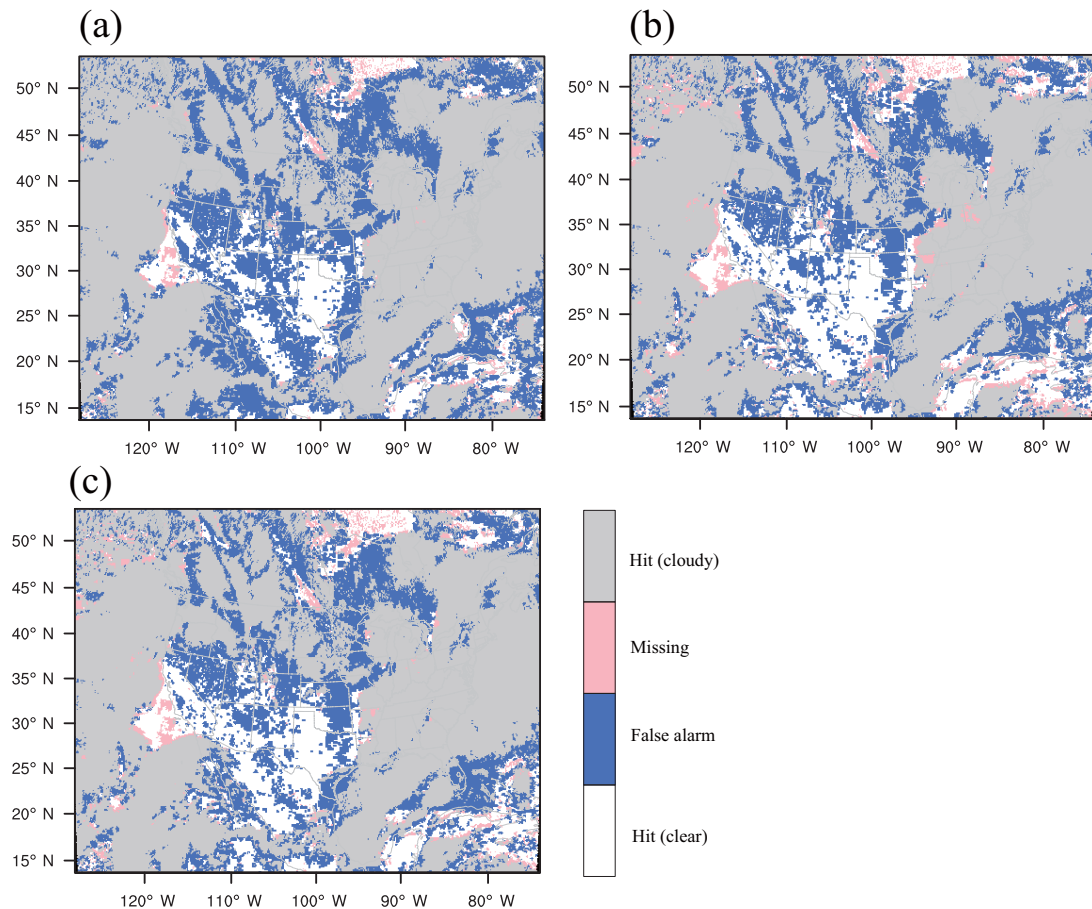
**Figure 5.** The mean cloud fraction on all model levels for the experiments MMR, PF and APF with AIRS observations valid at 11:00 UTC, 3 June 2012.

overestimation of the CBPs is even more obvious from MMR in most regions in Fig. 7d.

The CTPs from NASA GOES cloud products for more hours (03:00, 05:00, 07:00 UTC) together with the independent CTP retrievals from MODIS level 2 products

([http://modis-atmos.gsfc.nasa.gov/MOD06\\_L2/](http://modis-atmos.gsfc.nasa.gov/MOD06_L2/)) are plotted in Fig. 8. Different sub-periods of the MODIS cloud retrieval products (e.g., Fig. 8b valid at 03:20 UTC, Fig. 8c at 03:25, and Fig. 8d at 03:30 UTC) are chosen to approach the valid times in Fig. 8a, e and h. The CTPs from both cloud products agree well for both high and low clouds, confirming that NASA GOES cloud products are overall reliable for verifying the cloud retrievals and MODIS level 2 products can also be applied for validations.

Figure 9 presents the correlation coefficients and biases of the CTPs and CBPs verified against the NASA GOES and MODIS retrievals. The solid lines denote the results regarding the CTPs and CBPs vs. the NASA GOES products from 00:00 to 07:00 UTC, while the dots describe the CTP results vs. the cloud top retrievals in NASA MODIS level 2 products at 03:20, 03:25, 03:30, 05:00, 05:05, 06:35, 06:40 and 06:45 UTC. Here the negative bias means that the retrieved clouds are higher than the reference. Vice versa, the positive bias indicates the clouds are put too low. We conducted another experiment “APFing” that applies solely GOES Imager data to check the added value from the high spectral resolution radiances (such as, CrIS, AIRS and IASI). In Fig. 9a, the correlations between the retrievals from MMR and the NASA GOES retrievals are comparable with APF for most hours; APF gains overall higher correlations with the CTPs in the MODIS retrievals. From the bias in Fig. 9b, it seems that



**Figure 6.** The false alarms, misses and hits for clear and cloudy event locations with (a) the MMR method, (b) the APF method and (c) the APF method but without the group 1 particles (see text for detailed explanations) valid at 07:00 UTC, 15 December 2013.

the CTPs from MMR are underestimated (putting the clouds too high) consistently against both retrievals with GOES and MODIS radiances. Figure 9c shows that the correlations are weaker for MMR compared to others all the time. In Fig. 9d, the positive CBP biases from MMR are remarkable, whereas the CBP biases from APF are largely reduced. Generally, AP-Fimg degrades the CTP and CBP results consistently, suggesting that radiances with high spectral resolutions are able to improve the vertical descriptions of cloud profiles. It is found that the clouds retrieved with APFg2 are shrunk in terms of cloud depth with notably lower cloud top and higher cloud base compared to APF, when excluding the perturbed particles in the first group.

#### 4.5 Computational issues

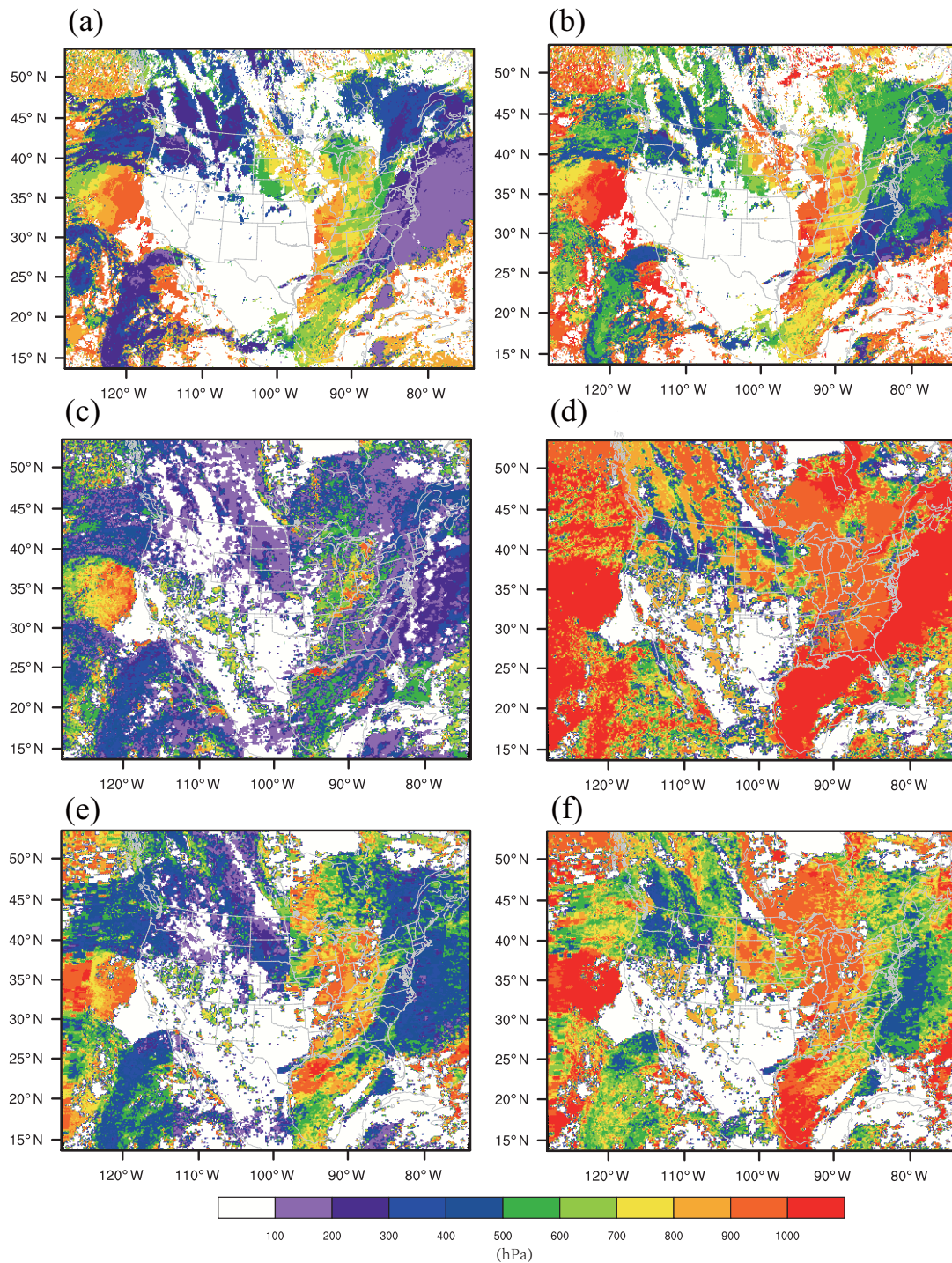
Figure 10a represents the elapsed times for the MMR and APF experiments and the counts of radiance observations in use are shown in Fig. 10b from 00:00 to 07:00 UTC, 12 December 2013. The profile of computing time in MMR is quite different from that in PF. The cost of MMR is dominated by the heavy minimization procedure, whereas APF is more as-

sociated with the processes of initializing particles and calculating weights for all the particles. The computing times were measured from cloud retrieving runs with 64 MPI (Message Passing Interface) tasks on a single computing node in an IBM iDataPlex Cluster. The measured wall clock computing times show that generally MMR is computationally more expensive for most of the time than APF. It seems the wall clock times for MMR are generally proportional to the data amount used. While for the APF experiment, the wall clock time is mostly determined by the particles size and partly affected by the channel number, such as for 2013121202 and 2013121206, when the total counts of the hyperspectral sensors (IASI, CrIS and AIRS) are large. The PF experiments using particles of the one-layer cloud with 100 % cloud fractions usually take less than 5 min for the same periods (not shown).

#### 4.6 Resolving the filtering problem on model grids

As explained in Sect. 3.3, the filtering problem is resolved in the radiance observational space at each FOV of each sensor independently and sequentially. For each FOV, the re-

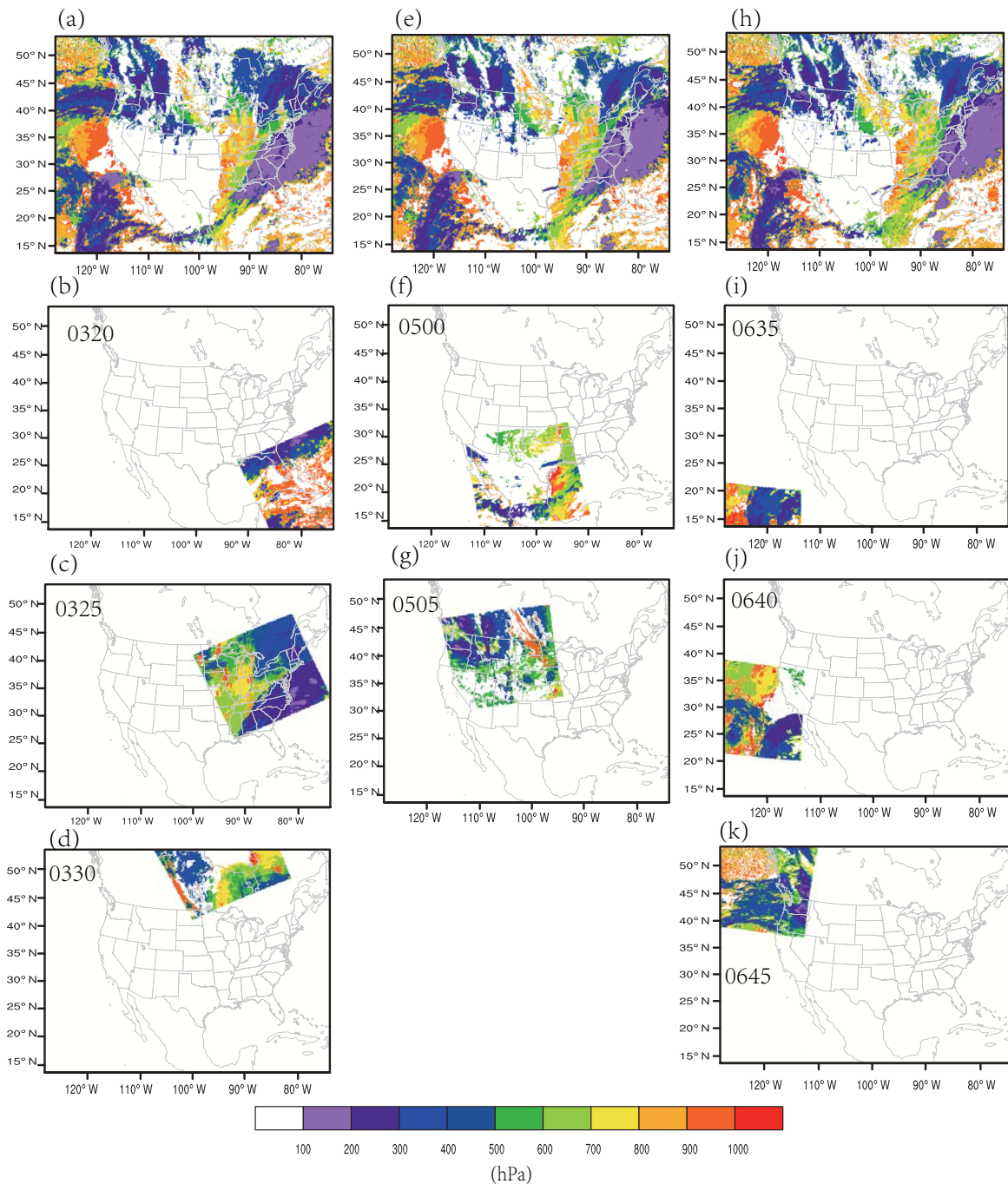




**Figure 7.** The cloud top pressure (left panels) from (a) the NASA GOES retrieval, (c) the MMR method and (e) the APF method, and the cloud bottom pressure (right panels) from (b) the NASA GOES retrieval, (d) the MMR method and (f) the APF method valid at 07:00 UTC, 15 December 2013.

trieved cloud fractions are extrapolated to its neighboring model grid points afterwards. We order the sensors in the cloud retrieving procedure as GOES-Imager, MODIS, CrIS, AIRS and IASI, aiming to optimize the vertical clouds us-

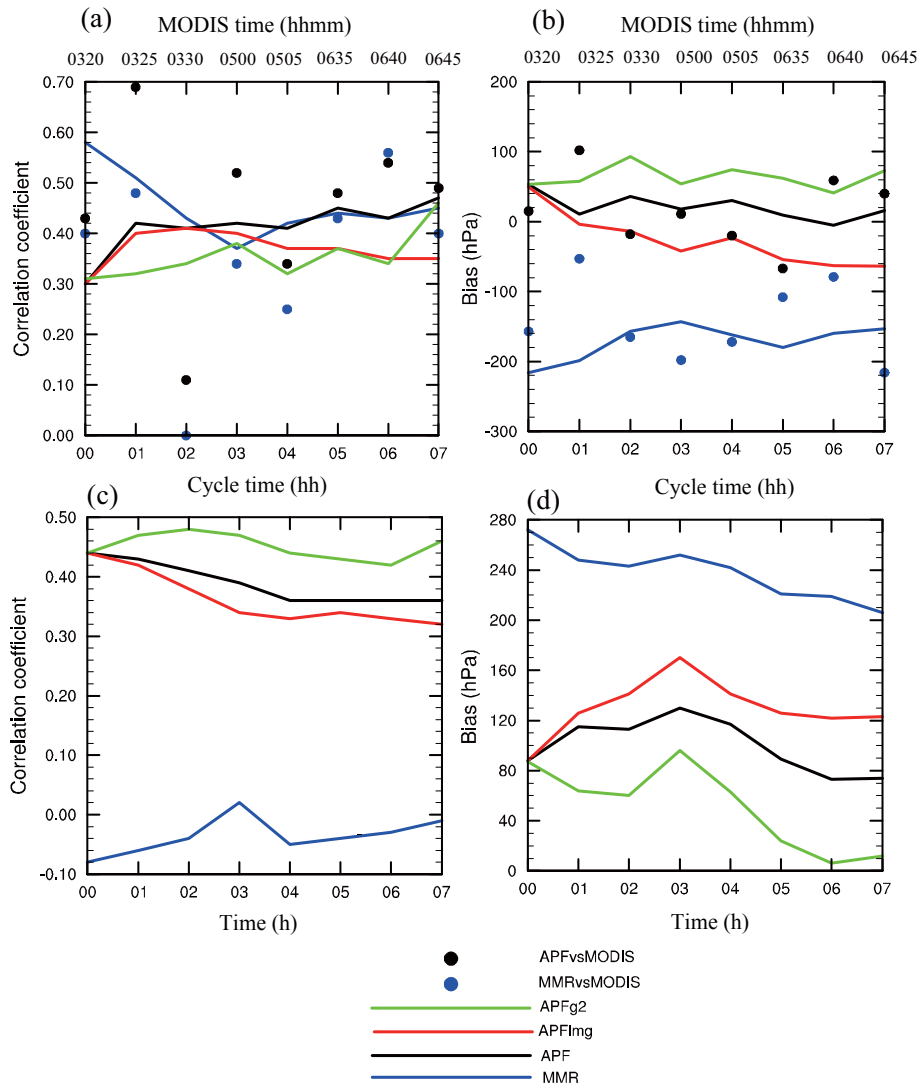
ing sensors featured with sufficient spectral resolutions. As a consequence, the retrievals from the last sensor determine the final output to the most extent, causing the cloud retrievals to be highly subjective to the ordering of the sensors. On the



**Figure 8.** The cloud top pressure for (a) 03:00 UTC from the GOES NASA retrieval; (b) 03:20 UTC, (c) 03:25 UTC and (d) 03:30 UTC from MODIS level 2 products; (e) 05:00 UTC from the GOES NASA retrieval; (f) 05:00 UTC, (g) 05:05 UTC; (h) 07:00 UTC from the GOES NASA retrieval; (i) 06:35 UTC, (j) 06:40 UTC and (k) 06:45 UTC from MODIS level 2 products.

other hand, it means the information from other prior sensors will be more or less discarded. In this section, a different way of resolving the filtering problem is preliminarily tested, in which the weights for each particle are aggregated over all available sensors by calling the forward radiative transfer model on neighboring model grids.

Figure 11 shows the clouds retrievals from the grid-based method. It is noted that the grid-based scheme yields slightly worse results of CTP and neutral results of CBP compared with those from the observation-based (FOV-based) scheme, indicating that the hyperspectral sensors probably favor the retrieved CTP and CBP in the FOV-based scheme, which are



**Figure 9.** (a) Correlation coefficient, (b) bias for the cloud top pressure, (c) correlation coefficient and (d) bias for the cloud bottom pressure vs. the NASA GOES retrievals from 06:00 UTC, 15 December 2013, to 07:00 UTC, 15 December 2013. Black and blue dots denote results vs. the MODIS level 2 cloud top pressure retrieval valid at 03:20, 03:25, 03:30, 05:00, 05:05, 06:35, 06:40 and 06:45 UTC. The valid times for the MODIS level 2 data are shown on the top of the  $x$  axis.

available for most of the time. It is worth pointing out that the ordering of different sensors has nearly no effect on the final cloud retrievals, when the weights of the particles are calculated in model space (not shown). The final cloud retrieval is no longer overwritten by the retrieval from the last sensor but is a total solution with all the sensors fairly considered, instead. The computational cost of retrieving clouds in model space is comparable or slightly heavier than that in observation space. The computational cost of the grid-based scheme scales with the number of the computing nodes more directly, compared to that of the FOV-based scheme.

### 5 Discussion and conclusion

This study presents a new cloud retrieval method based on the particle filter (PF) in the framework of GSI, as a competitive alternative to the MMR method. The behaviors of different particle initializations are demonstrated on one single field of view and the CONUS domain. Comparisons between the PF and the MMR method are conducted in terms of the features of cloud mask, cloud top, cloud base and the vertical distributions of clouds. It was found that the PF method retrieves clear cloud signals, whereas MMR is more ambiguous in detecting clouds. By adding more small-fraction particles, high clouds can be better interpreted. From the statistical results, it was found that MMR underestimates the



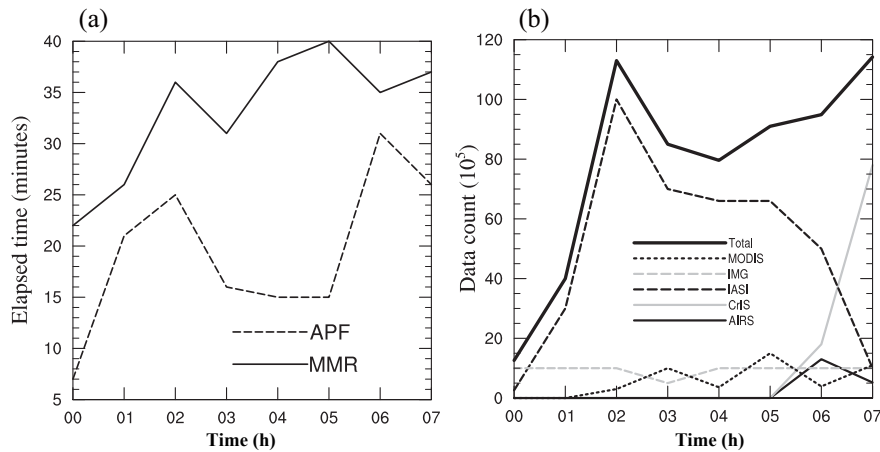


Figure 10. (a) The elapsed time and (b) the data count from 00:00 to 07:00 UTC, 15 December 2013.

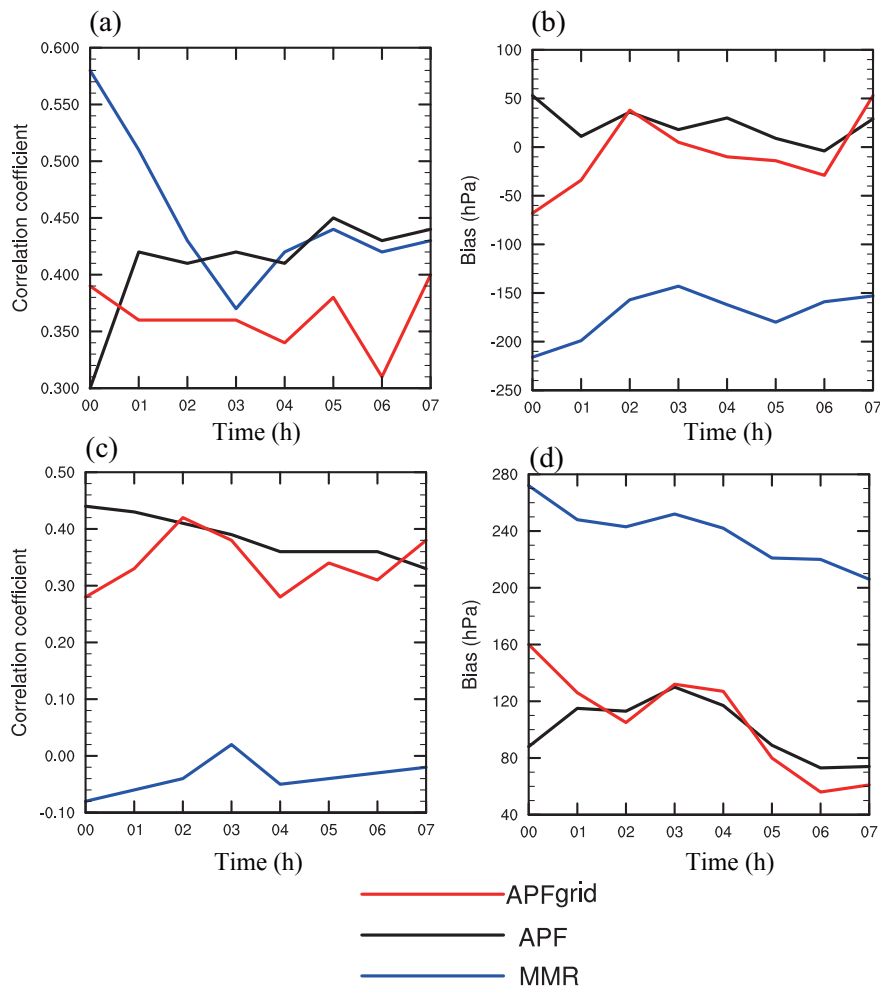


Figure 11. (a) Correlation coefficient, (b) bias for the cloud top pressure, (c) correlation coefficient and (d) bias for the cloud bottom pressure vs. the NASA GOES retrievals from 00:00 to 07:00 UTC, 15 December 2013.

cloud top pressures (put the clouds top too high) and overestimates the cloud bottom pressures (put the clouds top too low) as well. APF improves both the retrievals of cloud tops and cloud bases remarkably, especially for the cloud bases. As expected, radiances with high spectral resolutions contribute to quantitative cloud top and cloud base retrievals. In addition, a different way of resolving the filtering problem over each model grid is tested to aggregate the weights with all available sensors considered, which is proven to be less constrained by the ordering of sensors. Last but not least, the PF method is overall more computationally efficient; the cost of the model grid-based PF method scales more directly with the number of the computing nodes.

In future work, validation studies using multi-spectral imagers on geostationary satellites, spaceborne lidars (or radar) and surface site data will continue, and the results will be used to update the retrieval algorithm. Maximizing the consistency in the products across platforms and optimizing the synergistic use of multiple-source radiances in the new algorithm are important aspects. To estimate the flow-dependent uncertainties in the cloud analysis and in the forecasts, the ensemble nowcasting with three-dimensional cloud fractions via the rapid-update cycling mode is also planned. Increasing the highest extent cloudy cases will be included in future studies. Finally, the use of cloud liquid water and ice mixing ratios retrieved from the cloud fractions using multi-sensor radiances to pre-process the first guess in numerical weather forecast is another promising application.

## 6 Code and/or data availability

The MMR cloud retrieval codes can be obtained freely from (<http://www2.mmm.ucar.edu/wrf/users/wrfd/>). The other codes can be obtained via email from the authors.

*Acknowledgements.* This work was jointly sponsored by the US Air Force Weather Agency under the project “Air Force Coupled Analysis and Prediction System”, Natural Science Foundation of Jiangsu Province under grant no. BK20160954, the 973 program (grant no. 2013CB430102), the Beijing Funding from Jiangsu Research Institute of Meteorological Science (BJG201510), the National Natural Science Foundation of China (41375025) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors would like to thank Chris Davis for fruitful discussions, and to Bobbie Weaver for editing the manuscript. We greatly thank the anonymous reviewers for their valuable comments on the earlier versions of the manuscript.

Edited by: S. Remy

Reviewed by: two anonymous referees

## References

- Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, *Geophys. Res.-Atmos.*, 103, 32141–32157, 1998.
- Auligné, T.: Multivariate minimum residual method for cloud retrieval. Part I: Theoretical aspects and simulated observation experiments, *Mon. Weather Rev.*, 142, 4383–4398, 2014a.
- Auligné, T.: Multivariate minimum residual method for cloud retrieval. Part II: Real observations experiments, *Mon. Weather Rev.*, 142, 4399–4415, 2014b.
- Auligné, T., Lorenc, A., Michel, Y., Montmerle, T., Jones, A., Hu, M., and Dudhia, J.: Toward a New Cloud Analysis and Prediction System, *B. Am. Meteorol. Soc.*, 92, 207–210, 2011.
- Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M. D., Kalnay, E., McMillin, L. M., Revercomb, H., Rosenkranz, P. W., Smith, W. L., and Staelin, D. H.: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems, *Geosci. Remote Sens.*, 41, 253–264, 2003.
- Bao, Y., Xu, J., Powell Jr., A. M., Shao, M., Min, J., and Pan, Y.: Impacts of AMSU-A, MHS and IASI data assimilation on temperature and humidity forecasts with GSI-WRF over the western United States, *Atmos. Meas. Tech.*, 8, 4231–4242, doi:10.5194/amt-8-4231-2015, 2015.
- Bayler, G. M., Aune, R., and Raymond, W.: NWP cloud initialization using GOES sounder data and improved modeling of non-precipitating clouds, *Mon. Weather Rev.*, 128, 3911–3920, 2000.
- Berrocal, V. J., Raftery, A. E., and Gneiting, T.: Combining spatial statistical and ensemble information in probabilistic weather forecasts, *Mon. Weather Rev.*, 135, 1386–1402, 2007.
- Blumstein, D., Chalon, G., Carlier, T., Buil, C., Hebert, P., Maciaszek, T., Ponce, G., Phulpin, T., Tournier, B., and Simeoni, D.: IASI instrument: Technical overview and measured performances, *Proceedings of the SPIE*, 5543, 196–207, 2004.
- Brückner, M., Pospichal, B., Macke, A., and Wendisch, M.: A new multispectral cloud retrieval method for shipbased solar transmissivity measurements, *J. Geophys. Res.-Atmos.*, 119, 11338–11354, 2014.
- Descombes, G., Auligne, T., and Lin, H.-C., Xu, D., Schwartz, C. S., and Vandenberghe, F.: Multi-sensor Advection Diffusion now-Cast (MADCast) for cloud analysis and short-term prediction, *NCAR Technical Note NCAR/TN-509+STR*, 21 pp., 2014.
- Eyre, J. R. and Menzel, W. P.: Retrieval of cloud parameters from satellite sounder data: A simulation study, *J. Appl. Meteorol.*, 28, 267–275, 1989.
- Han, Y., Delst, P. V., Liu, Q., Weng, F., Yan, B., Treadon, R., and Derber, J.: JCSDA Community Radiative Transfer Model (CRTM) – Version 1, NOAA Tech. Rep. NESDIS, 122, 33 pp., 2006.
- Hu, M., Xue, M., and Brewster, K.: 3DVAR and Cloud Analysis with WSR-88D Level-II Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part I: Cloud Analysis and Its Impact, *Mon. Weather Rev.*, 134, 675–698, 2006.
- Huang, H.-L., Smith, W. L., Li, J., Antonelli, P., Wu, X., Knuteson, R. O., Huang, B., and Osborne, B. J.: Minimum local emissivity variance retrieval of cloud altitude and effective spectral emissivity-simulation and initial verification, *J. Appl. Meteorol.*, 43, 795–809, 2004.
- Karlsson, K.-G., Johansson, E., and Devasthale, A.: Advancing the uncertainty characterisation of cloud masking in passive satel-



- lite imagery: Probabilistic formulations for NOAA AVHRR data, *Remote Sens. Environ.*, 158, 126–139, 2015.
- Kleist, D. T., Parrish, D. F., Derber, J. C., Treadon, R., Wu, W. S., and Lord, S.: Introduction of the GSI into the NCEP Global Data Assimilation System, *Weather Forecast.*, 24, 1691–1705, 2009.
- Liu, Q. and Weng, F.: Advanced doubling-adding method for radiative transfer in planetary atmospheres, *J. Atmos. Sci.*, 63, 3459–3465, 2006.
- Mace, G. G.: Level 2 GEOPROF product process description and interface control document (v.3): Level 2 GEOPROF product process description and interface control document (v.3), Tech. rep., CIIRA, Colorado State University, 2004.
- Mechri, R., Ottlé, C., Pannekoucke, O., and Kallel, A.: Genetic particle filter application to land surface temperature downscaling, *J. Geophys. Res.-Atmos.*, 119, 2131–2146, 2014.
- Menzel, W., Smith, W., and Stewart, T.: Improved cloud motion wind vector and altitude assignment using VAS, *J. Clim. Appl. Meteorol.*, 22, 377–384, 1983.
- Menzel, W. P. and Purdom, J. F.: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites, *B. Am. Meteorol. Soc.*, 75, 757–781, 1994.
- Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riédi, J. C., and Frey, R. A.: The MODIS cloud products: Algorithms and examples from Terra, *Geosci. Remote Sens.*, 41, 459–473, 2003.
- Rossow, W. B. and Schiffer, R. A.: ISCCP cloud data products, *B. Am. Meteorol. Soc.*, 72, 2–20, 1991.
- Rossow, W. B., Walker, A. W., and Garder, L. C.: Comparison of ISCCP and other cloud amounts, *J. Climate*, 6, 2394–2418, 1993.
- Shen, F. F. and Min, J. Z.: Assimilating AMSU-A Radiance Data with the WRF Hybrid En3DVAR System for Track Predictions of Typhoon Megi (2010), *Adv. Atmos. Sci.*, 32, 1231–1243, 2015.
- Shen, Z. Q. and Tang, Y. M.: A modified ensemble Kalman particle filter for non-Gaussian systems with nonlinear measurement functions, *J. Adv. Model Earth Syst.*, 7, 50–66, 2015.
- Skamarock, W., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF version 3, NCAR, 113 pp., 2008.
- Smith, A., Atkinson, N., Bell, W., and Doherty, A.: An initial assessment of observations from the SuomiNPP satellite: data from the Cross-track Infrared Sounder (CrIS), *Atmos. Sci. Lett.*, 16, 260–266, 2015.
- Snyder, C. and Zhang, F. Q.: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter, *Mon. Weather Rev.*, 131, 1663–1677, 2003.
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'Connor, E. J., Rossow, W. B., and Durden, S. L.: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation, *B. Am. Meteorol. Soc.*, 83, 1771–1790, 2002.
- van Leeuwen, P. J.: Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Q. J. Roy. Meteor. Soc.*, 136, 1991–1999, 2010.
- Wu, W.-S., Purser, R. J., and Parrish, D. F.: Three-dimensional variational analysis with spatially inhomogeneous covariances, *Mon. Weather Rev.*, 130, 2905–2916, 2002.
- Wu, X. and Smith, W. L.: Assimilation of ERBE data with a nonlinear programming technique to improve cloud-cover diagnostics, *Mon. Weather Rev.*, 120, 2009–2004, 1992.
- Xu, D., Auligné, T., and Huang, X.-Y.: A Retrieval Method for 3-D Cloud Parameters Using Radiance Observations from Multiple Satellites, *Adv. Atmos. Phys.*, 32, 349–362, 2015.
- Xu, D. M., Liu, Z. Q., Huang, X. Y., Min, J. Z., and Wang, H. L.: Impact of assimilating IASI radiance observations on forecasts of two tropical cyclones, *Meteorol. Atmos. Phys.*, 122, 1–18, 2013.
- Zhao, C., Xie, S., Klein, S. A., Protat, A., Shupe, M. D., McFarlane, S. A., Comstock, J. M., Delanoë, J., Deng, M., and Dunn, M.: Toward understanding of differences in current cloud retrievals of ARM ground based measurements, *J. Geophys. Res.-Atmos.*, 117, D10206, doi:10.1029/2011JD016792, 2012.