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

Dongmei Xu[1,2] , Thomas Auligné[2], Gaël Descombes[2], and Chris Snyder[2]


[1]*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*

[2]*National Center for Atmospheric Research, Boulder, Colorado 80301, USA*

(2016/10/13)

[*] **Corresponding Author**

Dr. Dongmei Xu

Nanjing University of Information Science & Technology, College of Atmospheric science,

Ningliu road, No. 219, Nanjing, 210044, China

E-mail: xdmjolly@sina.com

# 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 Gridpoint Statistical Interpolation system (GSI). The PF cloud retrieval method is compared with the Multivariate and 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 respectively with measurements of Infrared radiances on multi-sensors onboard both geostationary and polar satellites. 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.

Keywords: cloud retrieval methods, particle filter, GSI system, cloud height

# 1. Introduction

Modern polar orbiting and geostationary airborne instruments provide researchers unprecedented opportunities for remote sensing of 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 multispectral 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)]. Specially, 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

58  multiple infrared instruments. However, MMR has limitations in several aspects due

59  to its use of minimization for solution: 1) Part of the control variables accounting for

60  the cloud fraction for some certain levels are under-observed since the channels are

61  not sensitive to the existence of clouds for those heights. 2) When clouds at different

62  heights show opacities with the same spectral signal, MMR could lose the ability to

63  distinguish solutions involving clouds at those levels. 3) The computational cost for

64  the minimization procedure in MMR is rather considerable.

65      Ensemble-based techniques, that usually reside in short-term ensemble

66  forecasting (Berrocal et al., 2007; Shen and Min, 2015), assembling existing model

67  outputs (e. g., cloud retrievals) from varying algorithms (Zhao et al., 2012), or

68  ensemble Kalman filter (EnKF) in diversified forms (Snyder and Zhang, 2003), have

69  been widely developed in order to estimate the uncertainties of various problems in

70  geophysical applications. To better account for the non-linearity between the observed

71  radiance and the retrieval parameter, a novel prototype for detecting clouds and

72  retrieving their vertical extension inspired by the particle filter (Snyder and Zhang,

73  2003) (van Leeuwen, 2010) (Shen and Tang, 2015) technique and Bayesian theory

74  (Karlsson et al., 2015) is proposed in this study. As a competitive alternative for

75  MMR, the PF retrieval method has same critical inputs required and cloud retrieval

76  products as in MMR. A brief description of MMR and the new PF cloud retrieval

77  algorithm are provided in the following section. Section 3 describes the background

78  model, the data assimilation system, the radiative transfer models (RTMs), and the

79  radiance observations applied in this study. Model configurations are also illustrated

80   in section 3. In section 4, the single test within one FOV is conducted before the

81   performance of PF method is assessed by comparing its cloud retrievals with those

82   from MMR and other operational cloud products. Section 4 also discusses the

83   computational performance for the two methods. The conclusion and anticipated

84   future work are outlined in section 5.


## 2. Methodology

85

86      Essentially, the PF cloud retrieval scheme retrieves clouds with the same critical

87   inputs requested (i. e., clear sky radiance from the radiative transfer model and the

88   observed radiance) and the same cloud retrievals as outputs (i. e., three dimensional

89   cloud fractions, which is defined as the fraction of top of cloud as seen from a sensor)

90   with the MMR method. Both cloud retrieval schemes consist of finding cloud

91   fractions that allow best fit between the cloudy radiance from model and the

92   observation. We use $c^1, c^2, ..., c^K$ to denote the array of vertical effective cloud

93   fractions for K model levels ($c^1$ for the surface and $c^K$ for the model top) and $c^0$ as

94   the fraction of clear sky with $0 \leq c^k \leq 1, \ \forall k \in [0, K]$. The constraint for the cloud

95   fraction is as follows,

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

96                                                                                              (1)

97   In this study, a cloud on one model level with a given fraction $c^k$ is assumed to

98   block the radiation from its lower model levels. The radiation originating from its

99   lower levels is assumed to contribute to the top of atmosphere radiance observed by

100    the satellites only with the residual fractions.

101    The MMR method is an approach to retrieve cloud fractions using the

102    minimization technique. The residual of the modeled radiance and the observation is

103    normalized by the observed radiance, which results in the following cost function,

104    using $c^k$, $\forall k \in [0, \mathrm{K}]$ as the control variables,

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

106    where $R_v^{\mathrm{cloud}}$ is the modeled cloudy radiance, and $R_v^{\mathrm{obs}}$ the observed radiance at

107    frequency $v$. This vertical cloud fraction $c^1, c^2, ..., c^{\mathrm{K}}$ and $c^0$ are control variables for

108    the cost function, where the simulated $R_v^{\mathrm{cloud}}$ is defined as

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

110    Here $R_v^k$ is the radiance calculated assuming an overcast black cloud at the model

111    level $k$ and $R_v^0$ the radiance calculated in the clear sky. Both $R_v^k$ and $R_v^0$ are

112    calculated using a forward radiative transfer model with model profiles of temperature

113    and moisture as inputs. Details of the schematic of the MMR method can be referred

114    in (Descombes et al., 2014; Xu et al., 2015).

115    Particle filter (PF) approach is one of the nonlinear filters for data assimilation

116    procedures to best estimate the initial state of a system or its parameters $x_t$, which

117    describes the time evolution of the full probability density function $p(x_t)$ conditioned

118    by the dynamics and the observations. Similar to the study in (Mechri et al., 2014), the

119    bibliography on PF focuses on estimating the parameters, which are cloud fractions

120    $c^k$ in Eq. (3), in this study. While MMR retrieves the cloud fractions on each model

121  vertical level by minimizing a cost function, PF calculates posterior weights for each

122  ensemble member based on the observation likelihood given that member. In its

123  simplest form, PF works by initializing a collection of cloud profiles as particles and

124  then estimating the cloud distributions by averaging those particles with their

125  corresponding weights. Each particle's weight is computed with the difference

126  between the modeled cloudy radiance from the particle and the observed radiance.

127  As the probabilities of the cloud distribution are fully presented by the initial

128  particles, of particular interest is to evaluate different particle initialization schemes in

129  the PF method. Explicitly, the definition of particles corresponds with ensemble

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

131  member.

132  Two approaches for generating particles are firstly designed; the first one is to

133  generate the perturbed samples $C_b^i$ ( $\forall i \in [1,n]$ ) from the cloud profile in the

134  background denoted as $C_b = (c_b^0, c_b^1, ..., c_b^K)$ by inflating (deflating) the clouds with

135  small magnitudes ( $C_b = \alpha \times C_b, \alpha = 50\%, 55\%, ..., 150\%$ ) and moving upward

136  (downward) with $\delta z = +5, +4..., -1, ... - 5$ as the vertical magnitude, where n is the

137  sample size. The perturbed cloud fractions are designated to replenish the ensemble

138  by introducing the prior information of the cloud distributions from the background

139  and to increase the ensemble spread.

140  Besides those perturbed particles, to represent the existence of one-layer cloud

141  on each model level with an even chance, another diversity set of profiles $C_b^i$

142  ( $\forall i \in [1, K+1]$ ) are also initialized, among which, $C_b^i$ stands for the profile with

143  100% cloud fraction on the model level $i$ ($c^i$=100%) and 0% cloud on the rest levels.

144  In particular, $C_b^0$ defines 100% clear ($c^0$=1). It is also interesting to discretize the

145  initial particles by setting the one-layer cloud with the value of $c^i$ from 100% to 0% (e.

146  g., 100%, 90%, 80%, …, 0% with 10% as the interval) and further from 100% to 0%

147  (e. g., 100%, 99%, 98%, 97%, …, 0% with 1% as the interval). In this cases, $c^0$=1-$c^i$.

148  For each particle $C_b^i$, its simulated cloudy radiance $R_{v,i}^{cloud}$ from the model background

149  can be obtained with Eq. (3).

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

151  observation as,

152
$$J_o = (\frac{R_v^{obs} - R_{v,i}^{cloud}}{\sigma})^2. \tag{4}$$

153  $\sigma$ is the specified observation error, which can be referred in the first paragraph in

154  section 4.1. The weight $w^i$ for each particle $C_b^i$ thus is calculated by comparing the

155  simulated $R_{v,i}^{cloud}$ and the observation $R_v^{obs}$ using the exponential function by

156  accumulating the $J_o$ for multiple frequency as

157
$$w^i = e^{-\sum_v (\frac{R_v^{obs} - R_{v,i}^{cloud}}{\sigma})^2}, \tag{5}$$

158  $\forall i \in [1, n]$. Here n is the particle size and $\sigma$ is the specified observation error,

159  which can be referred in the first paragraph in section 4.1. The final analyzed $C_a$ is

160  obtained by averaging the background particles $C_b^i$ with their corresponding weight,

161  as

162
$$C_a = \sum_{i=1}^{p} w^i C_b^i. \tag{6}$$

163  In Eq. (6), the constraint referred in Eq. (1) is not respected. Thus, after the analysis

164    step for the particle filter, the final averaged cloud fractions $c_a^k$ are normalized by

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

166    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 broadbands in the visible, near infrared and infrared regions at high spatial resolution (1–5 km) (Ackerman, 1998). Apart from the IR radiances from polar satellites, the Geostationary Operational Environmental Satellites (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

183    five-channel (one visible, four infrared) imaging radiometer designed to sense

184    radiant and solar reflected energy. The instrument parameters for the sensors and the

185    setups for channel selections can be found in (Xu et al., 2015).

186    3.2 WRF, GSI and the radiative transfer model

187    The background fields are processed running the Weather Research and Forecast

188    (WRF) model (Skamarock et al., 2008). The MMR and PF cloud retrieval algorithms

189    are both implemented based on the gridpoint statistical interpolation data assimilation

190    system (GSI) (Wu et al., 2002) (Kleist et al., 2009), which is a widely used data

191    assimilation system in operations and researches in NWP. GSI is capable of ingesting

192    a large variety of satellite radiance observations and has developed capabilities for

193    data thinning, quality control, and satellite radiance bias correction. The Community

194    Radiative Transfer Model (Liu and Weng, 2006) (Han et al., 2006) was used as the

195    radiance forward operator for computing the clear-sky radiance and the radiance given

196    overcast clouds at each model level.

197    3.3 Model configurations

198    The WRF is configured with 415*325 horizontal grids at 15-km grid spacing, and

199    40 vertical levels up to 50 hPa within the single CONUS domain. The MMR and PF

200    cloud detection schemes search the cloud top using approximately 150 hPa as the

201    highest extent for most cloudy cases. Other clouds higher 150 hPa, e.g. an anvil cloud

202    in a mature thunderstorm around tropopause at low latitude region will also be

203 explored in future studies. Channels in the longwave region are utilized following the

204 channel selection scheme in (Xu et al., 2015). Since the final retrieval clouds are on

205 model grids, the retrieved cloud fractions within one FOV are essentially extrapolated

206 to its four neighboring model grid points. Generally, for each FOV, the retrieved

207 cloud fractions are extrapolated to its four neighboring model grid points. For polar

208 satellite pixels, the representative cloud fractions are extrapolated with an adaptive

209 radius with respect to their scan positions. The cloud detecting procedure for

210 retrieving clouds is conducted for each FOV from each individual sensor

211 independently and sequentially. Since the clouds are retrieved FOV by FOV and the

212 clouds on grids are referred immediately after one FOV is completed, there is no

213 obvious accuracy loss of radiance observations using this conservative method.

# 4. Experiments and results

215     The PF experiments apply two groups of particles as mentioned in section 2,

216 among which the group-2 particles contains solely 100% one-layer clouds. To reveal

217 how the setup of the initial particles impacts the results, apart from the MMR and PF

218 experiments, we included another advanced experiment, denoted as APF. APF

219 requires more sampled particles including ranges of cloud fractions spanning from 0%

220 to 100% at the interval of 10%. An additional experiment "APFg2", similar to APF

221 but excluding the perturbed particles from the background in group-1 introduced in

222 section 2, was conducted to evaluate the added values from the group-one particles. In

223 this section, cloud retrieval experiments for several cases containing clouds of a

224    variety of types are conducted for comparison reason. The GOES imager retrieved

225    products from National Aeronautics and Space Administration (NASA-Langley cloud

226    and radiation products) are applied as a reference to validate the cloud retrieving

227    methods for the CONUS domain with a large and uniform coverage of cloud mask. In

228    addition, the retrieved cloud products were also compared to available CloudSat

229    (Stephens et al., 2002) and MODIS level-2 cloud products (Platnick et al., 2003)

230    archived by the CloudSat Data Processing Center in Colorado State and NASA

231    respectively.

232    4.1 Single test at one field of view

233        The PF cloud retrieving algorithm retrieves the cloud distributions by averaging

234    those initial particles with their weights. Before the real case experiments are carried

235    out over the whole domain, we conduct a single cloud retrieving test at one FOV to

236    understand what differences can be explained by the differences in the basic initial

237    particles. In Eq. (5), the observation error $\sigma$ can be set proportional to the

238    observation, equaling to $\dfrac{R_v^{\text{obs}}}{r}$, where $r$ is the prescribed ratio. Thus, the cloud

239    signals on each level $k$ are virtually determined by the extent of how close the $\dfrac{R_v^k}{R_v^{\text{obs}}}$

240    (and $\dfrac{R_v^0}{R_v^{\text{obs}}}$ for the clear part) gets to 1. An example of the ratio of the overcast

241    radiance and the observed radiance $\dfrac{R_v^k}{R_v^{obs}}$ for each model level is given in Fig. 1 of

242    GOES-Imager for the channel 5 ($\sim 13.00\,\mu m$). The clear sky radiance normalized by

243     the observed radiance $\dfrac{R_v^0}{R_v^{\mathrm{obs}}}$ is also shown at the level 0 (Fig. 1). It is expected that

244     the overcast radiance from the RTM decrease with the rising of the altitude. The cloud

245     signal is strongest around level 5, where $R_v^k$ fits $R_v^{\mathrm{obs}}$ most closely. The cloud

246     retrievals depend not only on the basic input profiles (i.e., the overcast radiance on

247     each level from RTM normalized by the observed radiance and the clear sky radiance

248     from RTM normalized by the observed radiance) and but also on the algorithm

249     applied for resolving the problem (e.g., MMR and PF in this study).

250
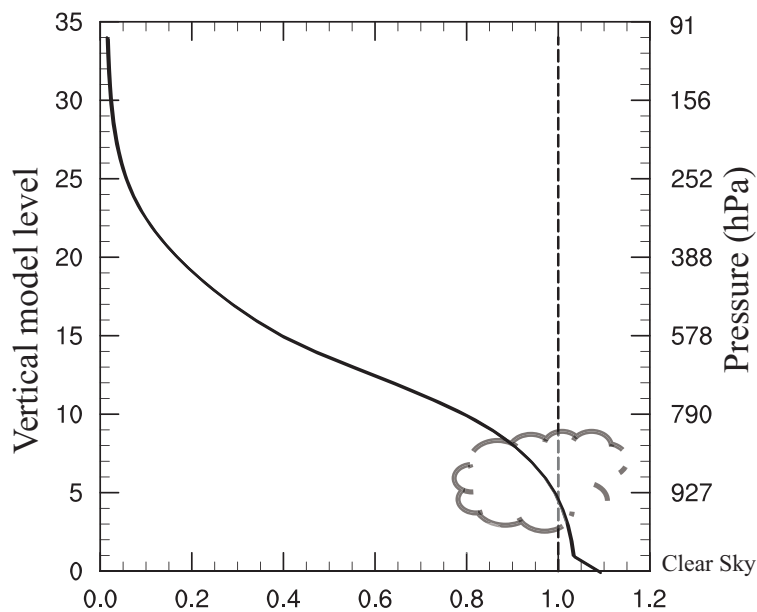


251

252     **Figure 1**. Ratio of the overcast radiances versus the observed radiance starting from the level 1.

253     The ratio of the clear sky radiance normalized by the observed radiance corresponds to the level 0

254     (see text for explanation) for GOES-Imager for the channel 5. The approximate pressures

255     corresponding to the model levels are also denoted.

256          To reveal the roles of various initial particles, Fig. 2a shows the weights for

257     different particles on the given FOV for channel 5 of GOES-Imager for the case

258    shown in Fig. 1. Particles in Fig. 1 include one-layer cloud in group 2 described in

259    section 2 with specified value of cloud fractions $c^k$ (on the x-axis) on specified model

260    levels $k$ (on the y-axis) from 10% to 100% every 10%. With a fraction $c^k$ of

261    one-cloud layer at a given level $k$ and a fraction of $c^0 = 1 - c^k$ of clear sky, the

262    simulated cloudy radiance can be denoted as $R_v^{\text{cloud}} = c^k R_v^k + (1 - c^k) R_v^0$. Hence the

263    theoretical one-layer cloud fraction is solved as $c^k = \dfrac{R_v^0 - R_v^{obs}}{R_v^0 - R_v^k}$ by fitting $R_v^{\text{cloud}}$ to $R_v^0$.

264    As expected, for one-layer cloud with full fraction, $c^5$ equals to 100% . Since with the

265    concept that $R_v^k > R_v^{k+1}$, no cloud can be present below level 5 since this would implies

266    a $R_v^{cloud}$ larger than the observation (or a $c^i$ larger than 100%). It seems that clouds

267    can be described by different possible states as particles with both large fractions and

268    small fractions. Low clouds are easily estimated by one-layer cloud profile with large

269    fractions (larger than 10%). The particles with small-fraction high clouds gain some

270    weights to retrieve high clouds. The particle with the one-layer cloud on level 13

271    seems to gain least weight compared to the others levels. The weights for the particles

272    with cloud fractions from 0% to 100% at the interval of 1% are also presented in Fig.

273    2b. By including more small-fraction one-layer clouds, the clouds around level 13 can

274    be reproduced by the group of refined particles with 1% as the interval for

275    approximate 10% cloud fractions. However, changing the level of the cloud for the

276    fixed fraction (10%) does not seem to change the outgoing radiance much, probably

277    due to the channel's low weight function peak (~750hPa).

278        The normalized $J_o$ in Eq. (6) for different levels with a specific cloud fraction

279    from 0% to 100% every 10% are shown in the bottom panel of Fig. 3, with 10% and

280    1% as the intervals in Fig. 3c and Fig. 3d respectively. Here, $J_o$ can be further derived

281    as

282
$$J_o = r^2 (1 - c^0 \frac{R_v^0}{R_v^{\text{obs}}} - c^k \frac{R_v^k}{R_v^{\text{obs}}})^2 \tag{8},$$

283    with $\sigma = \dfrac{R_v^{obs}}{r}$ and $R_v^{\text{cloud}}(c^0, c^1, c^2, ..., c^K) = c^0 R_v^0 + \sum\limits_{k=1}^{K} c^k R_v^k$ .

284        From Fig. 3a, it is found that $J_o$ is smallest around level-5 with 100% cloud

285    fraction (denoted as 1 in legend) for the thin black line, with respect to the fact that

286    the overcast radiance fits the observed radiance most closely for level-5

287    approximately. The grey line with 10% cloud fraction (0.1 in the legend) corresponds

288    to the existence of a weight peak on level 19 in Fig. 2a. In addition, the gap between

289    the grey line with 0.1 and the other lines from 0.2 to 1 explains why there's less

290    continuity around level 13. Fig. 3b shows a similar pattern to Fig. 3a, except with

291    densely-distributed $J_o$ values around the level 13 from 0.1 to 1 in the legend. Those

292    contiguous black lines in Fig. 3b are associated with the set of particles with cloud

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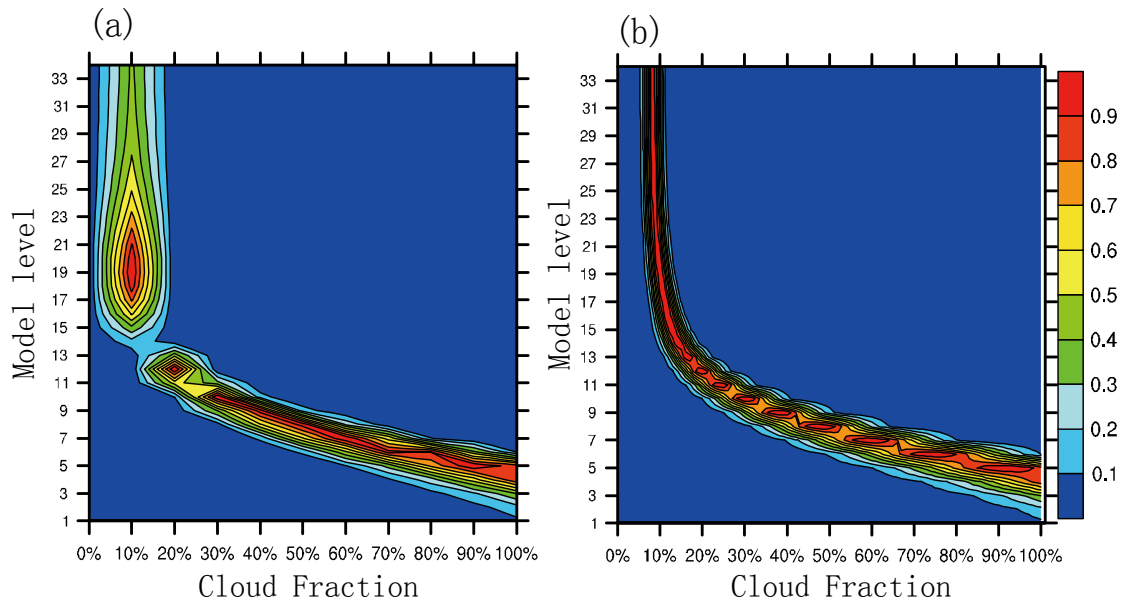
**Figure 2**. The weights for different particles with specified cloud fractions on the x-axis at one chosen model level shown on the y-axis from 0% to 100% (a) at the interval of 10% and (b) at the interval of 1%.
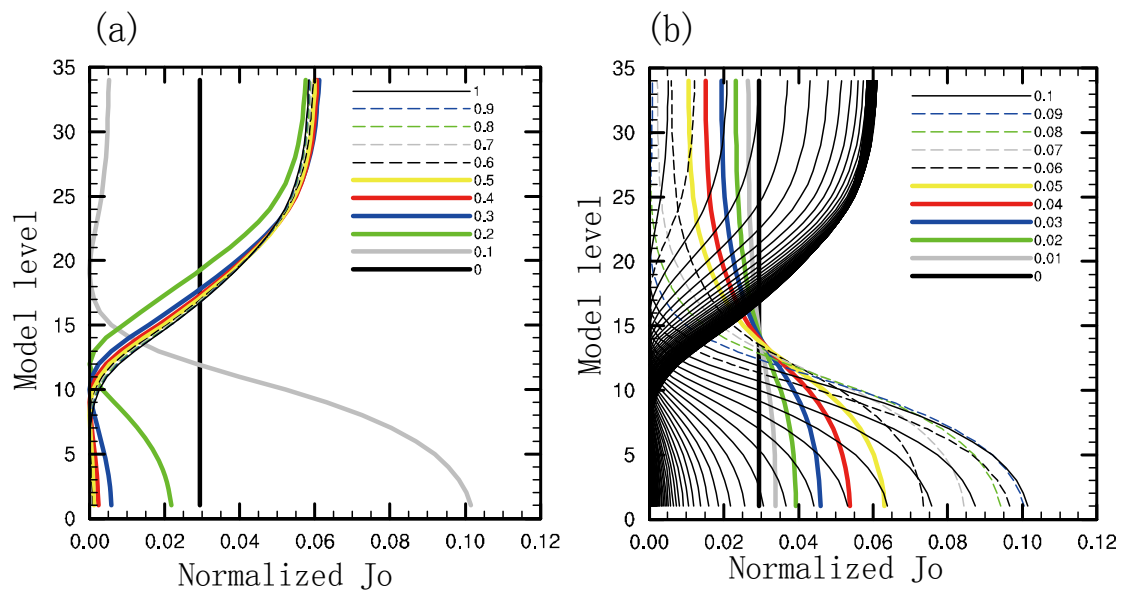


**Figure 3**. The normalized $J_o$ (a) at the interval of 10% and (b) at the interval of 1%. In (b), the normalized $J_o$ from 0.1 to 1 are all denoted as black lines.

303    4.2 Cloud profiles

304    The retrieval experiments for a real case are conducted at 1100 UTC 3 June 2012

305    when AIRS measurements and the CloudSat "2B-GEOPROF" products (Mace, 2004)

306    are available. The vertical cross sections of the cloud fraction field of a real case are

307    illustrated to further check how different collections of initial particles impact the

308    retrieved cloud profiles. The standard radar reflectivity profiles from the CloudSat are

309    shown in Fig. 4a as the validation source; Fig. 4b, Fig. 4c, and Fig. 4d show the cross

310    sections of the cloud fractions along the CloudSat orbit tracks from the MMR, PF and

311    APF experiments. The vertical structures of the clouds from MMR compare well with

312    the radar reflectivity from CloudSat by retrieving the high clouds around 47N° and

313    low clouds around 52N°. The PF experiment has difficulties in detecting the cloud

314    tops appropriately. PF tends to detect a large quantity of low clouds; by adding a set of

315    particles with small-fraction clouds in APF, higher clouds can be reproduced, which is

316    consistent with the implications from Fig. 2b and 3b. APF detects clear strong cloud

317    signals and removes the cloud fractions on near-surface levels around 36 N°

318    successfully. Since the existences of ground-layer radar reflectivity are likely

319    corresponding to the strong reflection from the underlying surface of the earth, the

320    height of cloud bases of MMR and PF are not compared in this sub-section. The

321    experiments with larger size of particles including 0% to 20% (at the interval of 1%)

322    plus 30% to 100% (at the interval of 10%) or of 0% to 100% (at the interval of 1%)

323    one-layer cloud profiles (introduced in section 2) yield similar results from APF but

324    are much more costly (not shown).

325



326
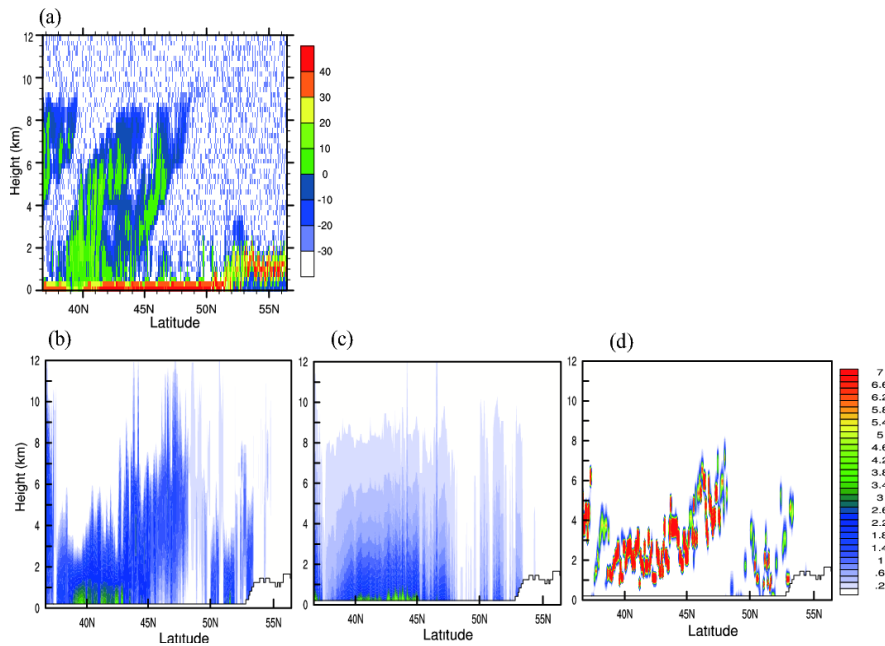
**Figure 4**. (a) The radar reflectivity (units: DBZ) cross sections from CloudSat, (b) the MMR

retrieved cloud fractions (units: %) cross sections, (c) the PF retrieved cloud fractions, and (d) the

APF retrieved cloud fractions valid at 1100 UTC 3 June 2012.

The vertical profiles of the averaged cloud fractions from MMR, PF, and APF are

plotted in Fig. 5 at 1100 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

950hPa to 400hPa). More clouds around level 10 are retrieved (approximately 750hPa)

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 lowest; the cloud top from MMR and PF is comparable. Only the

APF related methods will be further discussed in later sections owing to the missing

of high clouds using PF.
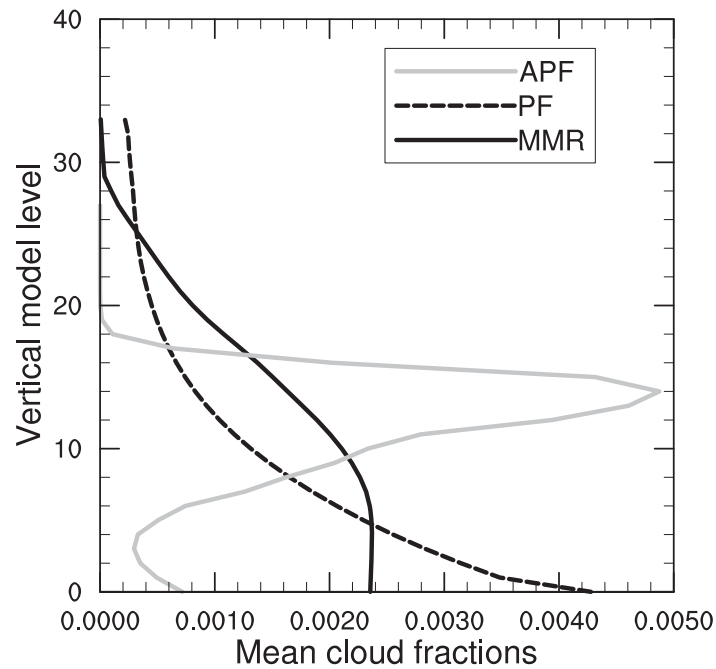
**Figure 5**. The mean cloud fraction on all model levels for the experiments MMR, PF, and APF

with AIRS observations valid at 1100 UTC 3 June 2012.

4.3 Cloud mask

Comparison experiments on real cases are further performed for over longer time

period from 0000 UTC 12 December 2013 to 0700 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°×0.1° latitude–longitude grid

with 0 for clear and 1 for cloudy before the comparisons for verification. Fig. 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 0700 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 6b. In most areas, the MMR experiment

355 overestimated the cloud mask with more false alarm events compared to the APF

356 experiment, since the MMR solution is an "overly smoothed" estimation of the true

357 vertical profile. It seems that the accuracy of cloud detection is lower for areas with

358 high altitude than under tropical conditions, indicating that the smaller lapse rate in

359 the atmosphere will lead radiance less sensitivity to clouds over polar areas. Fig. 6c

360 shows the cloud mask results from the APFg2 experiment without the perturbed

361 particles in group-1 introduced in section 2. There is no large discrepancy between

362 Fig. 6b and Fig. 6c, suggesting that the particles in group-2 that fully span the

363 possibility of the cloud distributions, are more determinant in retrieving the cloud

364 mask.

365
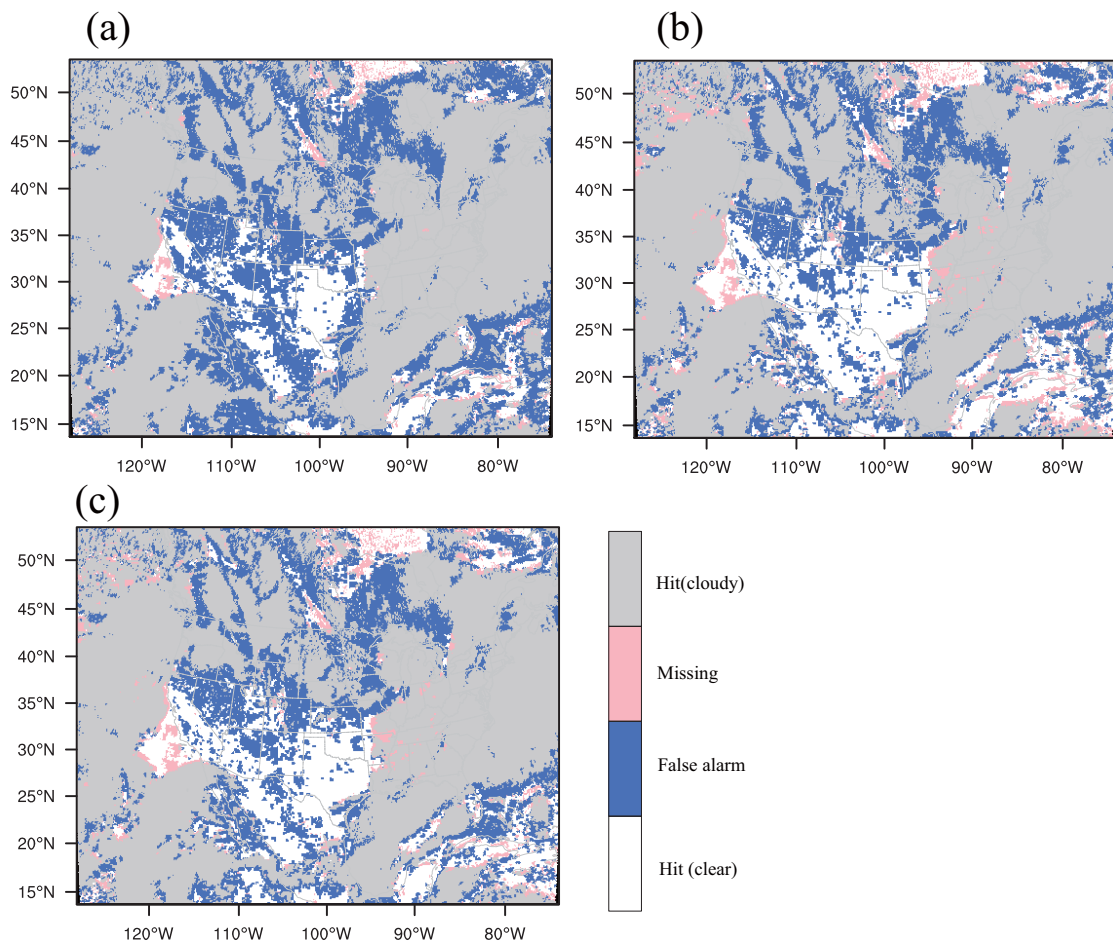


366

367 **Figure 6**. The false alarms, misses, and hits for clear and cloudy event locations with (a) the MMR

368 method, (b) the APF method, and (c) the APF method but without the group-1 particles (see text

369 for detailed explanations) valid at 0700 UTC 15 December 2013.

370 4.4 Cloud top and base pressure

371 The retrieved cloud top pressures (CTP) and cloud bottom pressures (CBP) from

372 this study along with the NASA GOES cloud products are illustrated in Fig. 7. The

373 CTPs from both methods are in good accordance with the NASA cloud products for

374 high clouds (from 100 hPa to 600 hPa) in Fig. 7a, 7c, and 7e. The retrieved cloud top

375 heights from MMR are overall higher than those from the NASA reference, especially

376 for lower clouds at approximately 750-1000 hPa (e. g., between longitude -100° and

377 -90°). On the other hand, the CTPs from APF are much closer to those in the

378 reference for both high and low clouds. APF overestimates the CBPs for some low

379 clouds (putting the clouds too low) in Fig. 7f; the overestimation of the CBP is even

380 more obvious from MMR in most regions in Fig. 7d.

381
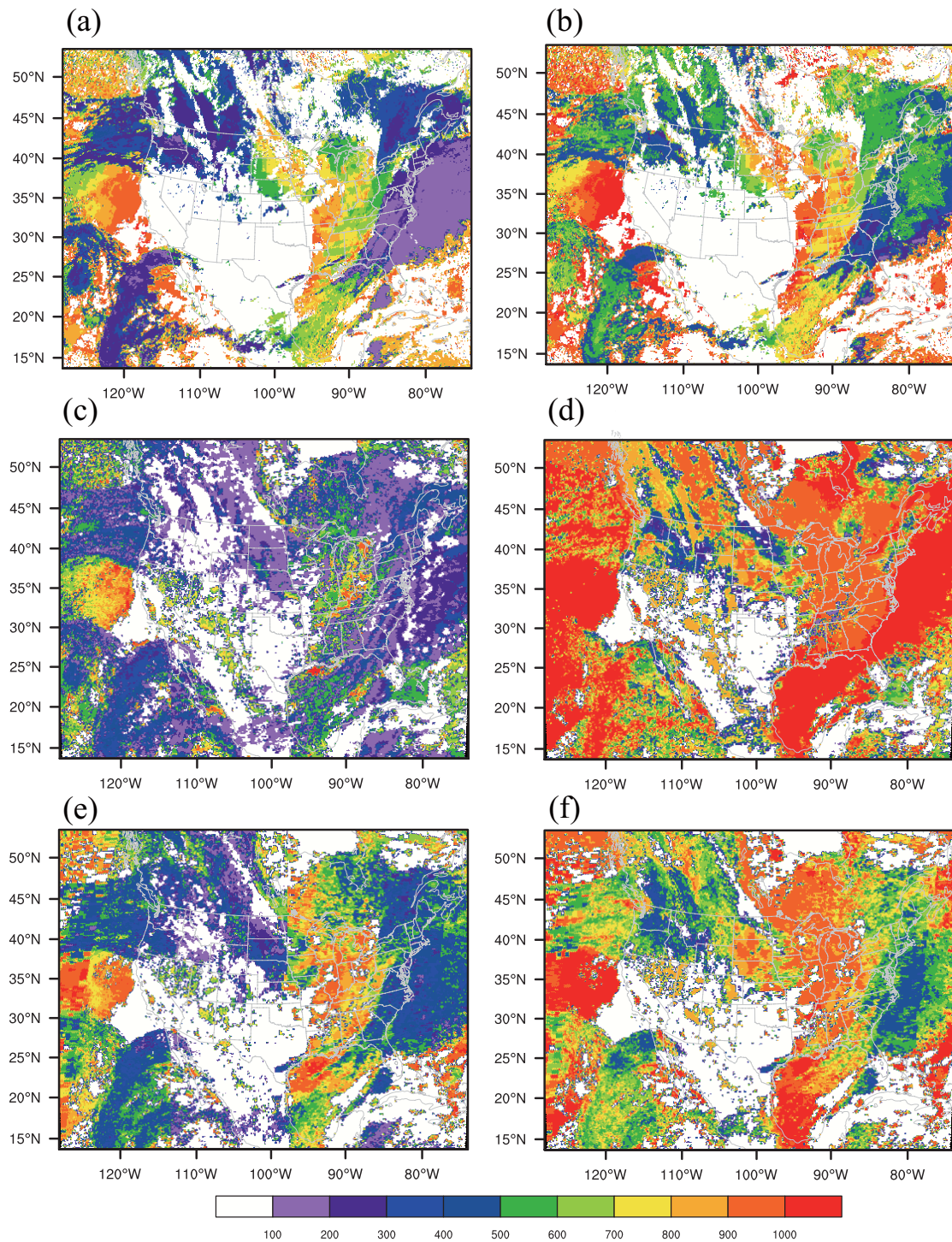
**Figure 7**. The cloud top pressure (left panels) from (a) the NASA GOES retrieval, (c) the MMR method, (e) the APF method, and the cloud bottom pressure (right panels) from (b) the NASA GOES retrieval, (d) the MMR method, (f) the APF method valid at 0700 UTC 15 December 2013.

The CTPs from NASA GOES cloud products for more hours (0300UTC, 0500UTC, 0700UTC) together with the independent CTP retrievals from MODIS

388     level-2 products (http://modis-atmos.gsfc.nasa.gov/MOD06_L2/) are plotted in Fig. 8.

389     Different sub-periods of the MODIS cloud retrieval products (e.g., Fig. 8b valid at

390     0320 UTC, Fig. 8c at 0325, and Fig. 8d at 0330 UTC) are chosen to approach the

391     valid times in Fig. 8a, Fig. 8e, and Fig. 8h respectively. The CTPs from both cloud

392     products agree well for both high and low clouds, confirming that NASA GOES cloud

393     products are overall reliable for verifying the cloud retrievals and MODIS level-2

394     products can also be applied for validations.
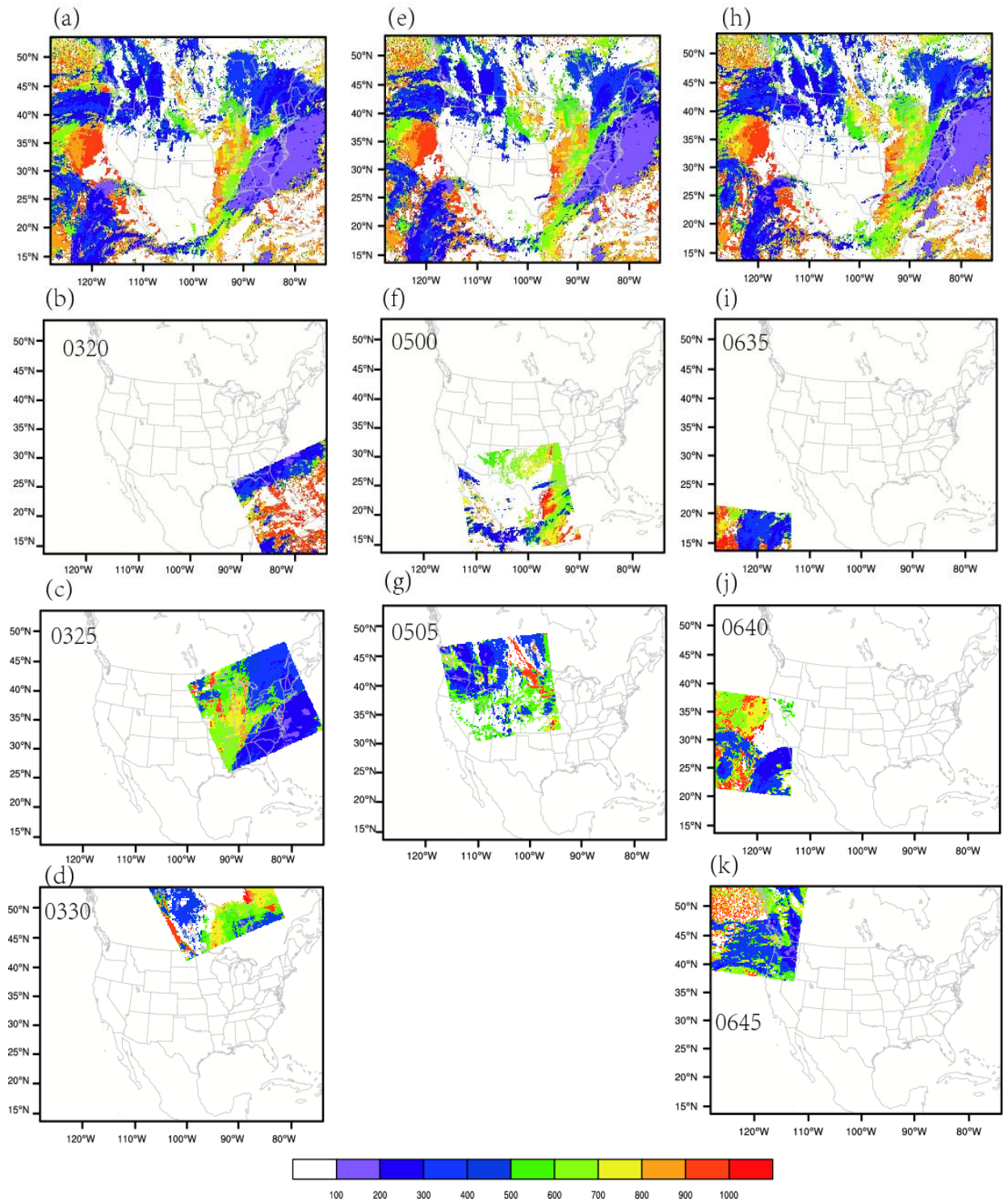
**Figure 8**. The cloud top pressure for (a) 0300UTC from the GOES NASA retrieval, (b) 0320UTC,

(c) 0325UTC, (d) 0330UTC from MODIS level-2 products; (e) 0500UTC from the GOES NASA

retrieval, (f) 0500UTC, (g) 0505UTC; (h) 0700UTC from the GOES NASA retrieval, (i)

0635UTC, (j) 0640UTC, and (k) 0645UTC from MODIS level-2 products.

400    Fig. 9 presents the correlation coefficients and biases of the CTP and CBP verified

401    against the NASA GOES and MODIS retrievals. The solid lines denote the results

402    regarding the CTP and CBP versus the NASA GOES products from 0000 UTC to

403    0700 UTC, while the dots describe the CTP results versus the cloud top retrievals in

404    NASA MODIS level-2 products at 0320UTC, 0325UTC, 0330UTC, 0500UTC,

405    0505UTC, 0635UTC, 0640UTC, and 0645UTC. Here the negative bias means that the

406    retrieved clouds are higher than the reference. Vice versa, the positive bias indicates

407    the clouds are put too low. We conducted another experiment "APFimg" that applies

408    solely GOES Imager data to check the added value from the high spectral resolution

409    radiances (such as, CrIS, AIRS, and IASI). In Fig. 9a, the correlations between the

410    retrievals from MMR and the NASA GOES retrievals are comparable with from APF

411    for most hours; APF gains overall higher correlations with the CTPs in the MODIS

412    retrievals. From the bias in Fig. 9b, it seems that the CTPs from MMR are

413    underestimated (putting the clouds too high) consistently against both retrievals with

414    GOES and MODIS radiances. Fig. 9c shows that the correlations are weaker for

415    MMR compared to others all the time. In Fig. 9d, the positive CBP biases from MMR

416    are remarkable, while the CBP biases from APF are largely reduced. Generally,

417    APFimg degrades the CTP and CBP results consistently, suggesting that radiances

418    with high spectral resolutions are able to improve the vertical descriptions of cloud

419    profiles. It is found that the clouds retrieved with APFg2 are shrunken in terms of

420    cloud depth with notably lower cloud top and higher cloud base compared to APF,

421    when excluding the perturbed particles in the first group.
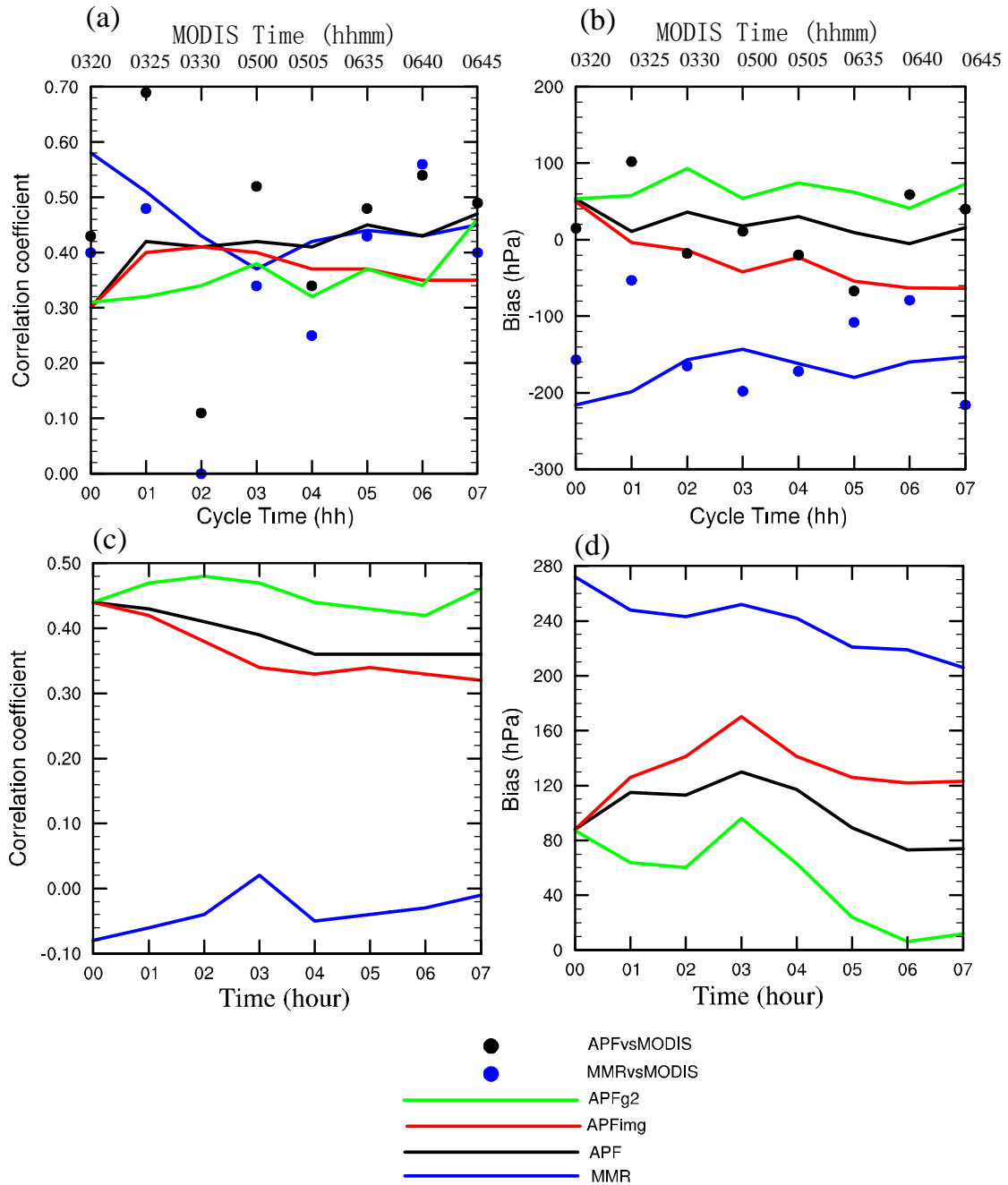
**Figure 9**. (a) Correlation coefficient, (b) bias for the cloud top pressure, (c) correlation coefficient,

and (d) bias for the cloud bottom pressure versus the NASA GOES retrievals from 0600 UTC 15

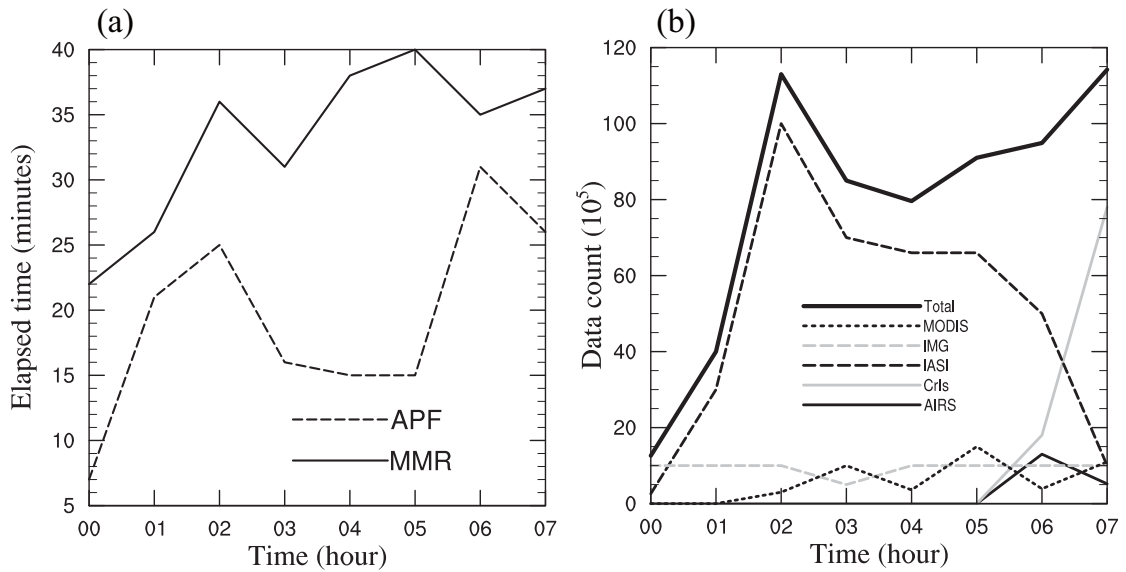December 2013 to 0700 UTC 15 December 2013. Black and blue dots denote results versus the

MODIS level-2 cloud top pressure retrieval valid at 0320UTC, 0325UTC, 0330UTC, 0500UTC,

0505UTC, 0635UTC, 0640UTC, and 0645UTC. The valid times for the MODIS level-2 data are

shown on the top of the x-axis.

429    4.5 Computational issues

430    Fig. 10a represents the elapsed times for the MMR and APF experiments and the

431    counts of radiance observations in use are shown in Fig. 10b from 0000 UTC to 0700

432    UTC 12 December 2013. The profile of computing time in MMR is quite different

433    from that in PF. The cost of MMR is dominated by the heavy minimization procedure,

434    while APF is more associated with the processes of initializing particles and

435    calculating weights for all the particles. The computing times were measured from

436    cloud retrieving runs with 64 MPI-tasks on a single computing node in an IBM

437    iDataPlex Cluster. The measured wall clock computing times show that generally

438    MMR is computationally more expensive for most of the time than APF. It seems the

439    wall clock times for MMR are generally proportional to the data amount used. While

440    for the APF experiment, the wall clock time is mostly determined by the particles size

441    and partly affected by the channel number, such as for 2013121202 and 2013121206,

442    when the total counts of the hyperspectral sensors (IASI, CrIs, and AIRS) are large.

443    The PF experiments using particles of one-layer cloud with 100% cloud fractions

444    usually take less than 5 minutes for the same periods (not shown).

**Figure 10**. (a) The elapsed time and (b) the data count from 0000 UTC to 0700 UTC 15 December 2013.

## 4.6 Resolving the filtering problem on model grids

As explained in subsection 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 retrieved 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 using 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 highly subjective to the ordering of the sensors. On the 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

460     calling the forward radiative transfer model on neighbouring model grids.

461     Fig. 11 shows the clouds retrievals from the grid-based method. It is noted that

462     the grid-based scheme yields slightly worse results of CTP and neutral results of CBP

463     compared with those from the observation-based (FOV-based) scheme, indicating that

464     the hyperspectral sensors probably favor the retrieved CTP and CBP in the

465     FOV-based scheme, which are available for most of the time. It is worth pointing out

466     that the ordering of different sensors has nearly no effect on the final cloud retrievals,

467     when the weights of the particles are calculated in model space (not shown). The final

468     cloud retrieval is no longer overwritten by the retrieval from the last sensor but is a

469     total solution with all the sensors fairly considered, instead. The computational cost of

470     retrieving clouds in model space is comparable or slightly heavier than that in

471     observation space. The computational cost of the grid-based scheme scales with the

472     number of the computing nodes more directly, compared to that of the FOV-based

473     scheme.
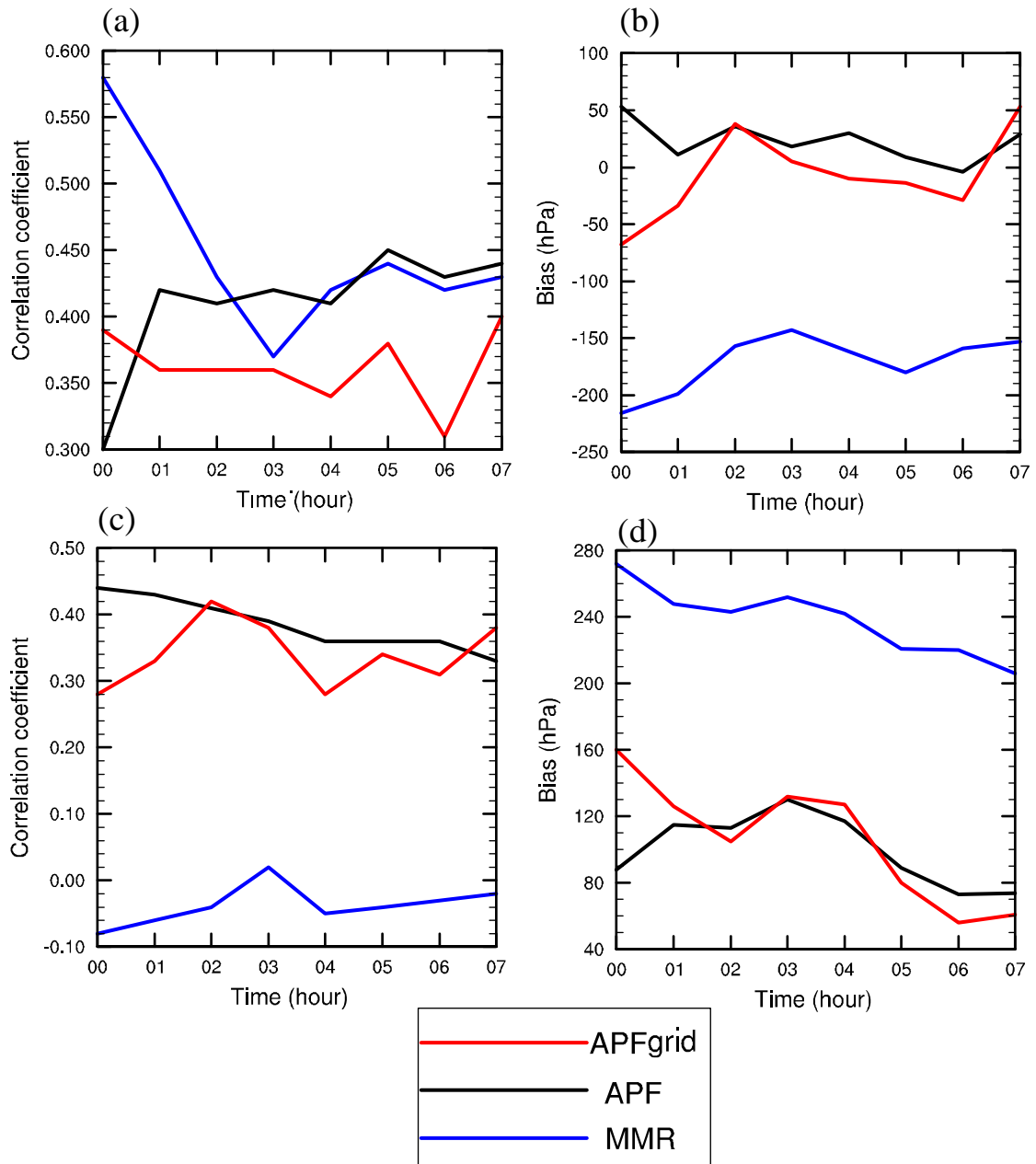
**Figure 11.** (a) Correlation coefficient, (b) bias for the cloud top pressure, (c) correlation coefficient, and (d) bias for the cloud bottom pressure versus the NASA GOES retrievals from 0000 UTC to 0700 UTC 15 December 2013.

# 5. Discussion and conclusion

479       This study presents a new cloud retrieval method based on the particle filter (PF)

480 in the framework of GSI, as a competitive alternative to the MMR method. The

481 behaviors of different particle initializations are demonstrated on one single field of

482 view and the CONUS domain respectively. Comparisons between the PF and the

483 MMR method are conducted in terms of the features of cloud mask, cloud top, cloud

484 base, and the vertical distributions of clouds. It was found that the PF method

485 retrieves clear cloud signals while MMR is more ambiguous in detecting clouds. By

486 adding more small-fraction particles, high clouds can be better interpreted. From the

487 statistical results, it was found that MMR underestimates the cloud top pressures (put

488 the clouds top too high) and overestimates the cloud bottom pressures (put the clouds

489 top too low) as well. APF improves both the retrievals of cloud tops and cloud bases

490 remarkably, especially for the cloud bases. As expected, radiances with high spectral

491 resolutions contribute to quantitative cloud top and cloud base retrievals. In addition,

492 a different way of resolving the filtering problem over each model grid is tested to

493 aggregate the weights with all available sensors considered, which is proven to be less

494 constrained by the ordering of sensors. Last but not least, the PF method is overall

495 more computationally efficient; the cost of the model grid-based PF method scales

496 more directly with the number of the computing nodes.

497      In future work, validation studies using multispectral imagers on geostationary

498 satellites, spaceborne lidars (or radar), and surface site data will continue, and the

499 results will be used to update the retrieval algorithm. Maximizing the consistency in

500 the products across platforms and optimizing the synergistic use of multiple-source

501 radiances in the new algorithm are important aspects. To estimate the flow dependent

502 uncertainties in the cloud analysis and in the forecasts, the ensemble nowcasting with

503 three dimensional cloud fractions via the rapid-update cycling mode is also planned.

504 Increasing the highest extent cloudy cases will be included in future studies. Finally,

505 the use of cloud liquid water and ice mixing ratios retrieved from the cloud fractions

506 using multi-sensor radiances to pre-process the first guess in numerical weather

507 forecast is another promising application.

508 **Code and/or data availability**

509 The MMR cloud retrieval codes can be obtained freely from

510 (http://www2.mmm.ucar.edu/wrf/users/wrfda/). The other codes can be obtained by

511 emails from the authors.

512 **Acknowledgments**

**REFERENCES**

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley: 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, Monthly Weather Review, 142, 4383-4398, 2014a.

Auligné, T.: Multivariate minimum residual method for cloud retrieval. Part II: Real observations experiments, Monthly Weather Review, 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, Geoscience and Remote Sensing, IEEE Transactions on, 41, 253-264, 2003.

Bao, Y., Xu, J., Powell, A. M., and Shao, M.: Impacts of AMSU-A/MHS and IASI data assimilation on temperature and humidity forecasts with GSI/WRF over the Western United States, Atmospheric Measurement Techniques Discussions, 8, 6439-6467, 2015.

Bayler, G. M., Aune, R., and Raymond, W.: NWP cloud initialization using GOES sounder data and improved modeling of nonprecipitating clouds, Monthly weather review, 128, 3911-3920, 2000.

Berrocal, V. J., Raftery, A. E., and Gneiting, T.: Combining spatial statistical and ensemble information in probabilistic weather forecasts, Monthly Weather Review, 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, 2004, 196-207.

Brückner, M., Pospichal, B., Macke, A., and Wendisch, M.: A new multispectral cloud retrieval method for ship‐based solar transmissivity measurements, Journal of Geophysical Research: Atmospheres, 119, 2014.

Descombes, G., Auligne, T., and Lin, H.-C., Xu, D., Schwartz, C. S., Vandenberghe, F.: Multi-sensor Advection Diffusion nowCast (MADCast) for cloud analysis and short-term prediction., NCAR Technical Note NCAR/TN-509+STR, , 2014. 21 pp., 2014.

Eyre, J. R. and Menzel, W. P.: Retrieval of cloud parameters from satellite sounder data: A simulation study, Journal of Applied Meteorology, 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, 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, Monthly Weather Review, 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, Journal of applied meteorology, 43, 795-809, 2004.

Karlsson, K.-G., Johansson, E., and Devasthale, A.: Advancing the uncertainty characterisation of cloud masking in passive satellite imagery: Probabilistic formulations for NOAA AVHRR data, Remote Sensing of Environment, 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 and Forecasting, 24, 1691-1705, 2009.

Liu, Q. and Weng, F.: Advanced doubling-adding method for radiative transfer in planetary atmospheres, Journal of the atmospheric sciences, 63, 3459-3465, 2006.

Mace, G. G., 2004: 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., CIRA, Colorado State University, 2004. 2004.

Mechri, R., Ottlé, C., Pannekoucke, O., and Kallel, A.: Genetic particle filter application to land surface temperature downscaling, Journal of Geophysical Research: Atmospheres, 119, 2131-2146, 2014.

Menzel, W., Smith, W., and Stewart, T.: Improved cloud motion wind vector and altitude assignment using VAS, Journal of Climate and Applied meteorology, 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, Geoscience and Remote Sensing, IEEE Transactions on, 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, Journal of 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), Advances in Atmospheric Sciences, 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 Sy, 7, 50-66, 2015.

Skamarock, W., C, 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

610  Research WRF version 3., NCAR, 113 pp., 2008.

611  Smith, A., Atkinson, N., Bell, W., and Doherty, A.: An initial assessment of
612  observations from the Suomi‑NPP satellite: data from the Cross‑track Infrared
613  Sounder (CrIS), Atmospheric Science Letters, 16, 260-266, 2015.

614  Snyder, C. and Zhang, F. Q.: Assimilation of simulated Doppler radar observations
615  with an ensemble Kalman filter, Monthly Weather Review, 131, 1663-1677, 2003.

616  Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z.,
617  Illingworth, A. J., O'Connor, E. J., Rossow, W. B., and Durden, S. L.: The
618  CloudSat mission and the A-Train: A new dimension of space-based observations
619  of clouds and precipitation, B Am Meteorol Soc, 83, 1771-1790, 2002.

620  van Leeuwen, P. J.: Nonlinear data assimilation in geosciences: an extremely efficient
621  particle filter, Quarterly Journal of the Royal Meteorological Society, 136,
622  1991-1999, 2010.

623  Wu, W.-S., Purser, R. J., and Parrish, D. F.: Three-dimensional variational analysis
624  with spatially inhomogeneous covariances, Monthly Weather Review, 130,
625  2905-2916, 2002.

626  Wu, X. and Smith, W. L.: Assimilation of ERBE data with a nonlinear programming
627  technique to improve cloud-cover diagnostics, 120, 2009-2004, 1992.

628  Xu, D., Auligné, T., and Huang, X.-Y.: A Retrieval Method for 3-D Cloud Parameters
629  Using Radiance Observations from Multiple Satellites, Advances in atmospheric
630  physics, 32, 349-362, 2015.

631  Xu, D. M., Liu, Z. Q., Huang, X. Y., Min, J. Z., and Wang, H. L.: Impact of
632  assimilating IASI radiance observations on forecasts of two tropical cyclones,
633  Meteorology and Atmospheric Physics, 122, 1-18, 2013.

634  Zhao, C., Xie, S., Klein, S. A., Protat, A., Shupe, M. D., McFarlane, S. A., Comstock,
635  J. M., Delanoë, J., Deng, M., and Dunn, M.: Toward understanding of differences
636  in current cloud retrievals of ARM ground‑based measurements, Journal of
637  Geophysical Research: Atmospheres, 117, 2012.

638
639