QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. **00**: 1–12 (2013)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.000

Monte Carlo Bayesian inference on a statistical model of sub-gridcolumn moisture variability using high-resolution cloud observations. Part II: Sensitivity tests and results

Arlindo M. da Silva[a] and Peter M. Norris[a,b]*
[a]*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[b]*Goddard Earth Sciences Technology and Research, University Space Research Association, Columbia, Maryland, USA*

Abstract: Part I presented a Monte Carlo Bayesian method for constraining a complex statistical model of GCM sub-gridcolumn moisture variability using high-resolution MODIS cloud data, thereby permitting large-scale model parameter estimation and cloud data assimilation. This part performs some basic testing of this new approach, verifying that it does indeed significantly reduce mean and standard deviation biases with respect to the assimilated MODIS cloud optical depth, brightness temperature and cloud top pressure, and that it also improves the simulated rotational-Ramman scattering cloud optical centroid pressure (OCP) against independent (non-assimilated) retrievals from the OMI instrument. Of particular interest, the Monte Carlo method does show skill in the especially difficult case where the background state is clear but cloudy observations exist. In traditional linearized data assimilation methods, a subsaturated background cannot produce clouds via any infinitesimal equilibrium perturbation, but the Monte Carlo approach allows finite jumps into regions of non-zero cloud probability. In the example provided, the method is able to restore marine stratocumulus near the Californian coast where the background state has a clear swath. This paper also examines a number of algorithmic and physical sensitivities of the new method and provides guidance for its cost-effective implementation. One obvious difficulty for the method, and other cloud data assimilation methods as well, is the lack of information content in the cloud observables on cloud vertical structure, beyond cloud top pressure and optical thickness, thus necessitating strong dependence on the background vertical moisture structure. It is found that a simple flow-dependent correlation modification due to Riishojgaard (1998) provides some help in this respect, by better honoring inversion structures in the background state. Copyright © 2013 Royal Meteorological Society

KEY WORDS Cloud data assimilation; Monte Carlo Bayesian inference; Correlation models

Received ???; Revised ???; Accepted ???

1 Introduction

In the introduction to Part I we provided motivation for this study by discussing a number of the difficulties associated with the subject of cloud data assimilation (CDA). We particularly emphasized problems associated with the mismatch between the frequently small scales of cloud variability and typical GCM gridcolumn footprints and with the strong non-linearities present in cloud processes. One of the key problems is that a subsaturated background state cannot produce clouds via any small equilibrium perturbation to moisture. We then provided a detailed description of a new Monte Carlo Bayesian CDA approach designed to address these problems, with the goal of improving poor model background states to the point where more traditional CDA approaches are able to perform more favorably with them.

This Part II now discusses the application of the new method and its performance in a number of case studies and sensitivity tests. Section 2 sets up a control cloud data assimilation experiment with the new method, making a few necessary modifications to the system described in Part I. It is verifed that the control does indeed significantly reduce mean and standard deviation biases with respect to the assimilated MODIS cloud optical depth, brightness temperature and cloud top pressure. Section 3 examines a number of algorithmic and physical sensitivities of the new method and provides guidance for its cost-effective implementation. The section also demonstates the significant utility of a simple flow-dependent correlation function modification due to Riishojgaard (1998). Section 4 presents a case study involving marine stratocumulus off the Californian coast. This case is an especially difficult test of the new method, because the observations show stratocumulus very near and sometimes right up to the coast, whilst the background state has a large clear swath in this region. The results are very encouraging. Section 5 presents an short application of the method to assimilation of geostationary cloud retrievals made for the SEVIRI instrument aboard the Meteosat-9 platform. The method is easily adaptable to assimilation of this data. A comparison against the assimilation of contemporaneous polar-orbiting Aqua MODIS data shows comparible

*Correspondence to: Dr. Peter Norris, Global Modeling and Assimilation Office, NASA/GSFC, Code 610.1, Greenbelt, MD 20771, USA. Email: peter.m.norris@nasa.gov

results for both types of observations. Section 6 presents an independent validation of the new method using a non-assimilated dataset, the so-called cloud Optical Centroid Pressure (OCP) retrieved from a rotational-Ramman scattering algorithm applied to UV spectral measurements made by the Ozone Monitoring Instrument (OMI) aboard the Aura satellite. Finally, Section 7 contains some conclusions.

2 Control Run

A number of sensitivity tests have been performed (see below) by perturbing key algorithmic and physical parameters of the Bayesian cloud assimilation system. The control run for these tests is largely defined by Part I, with the following details and modifications:

(1) $N_{\text{sim}} = N_{\bullet\text{max}} = 64$.

(2) The number of trials M per point in the MTM (Multiple-try Metropolis) chain is a factor $f_M = 1/2$ of that specified at the end of Section 2.9 of Part I, namely about half the number of effectively independent dimensions in the parameter space.

(3) The MTM proposal covariance matrix Σ_q of (17) in Part I is amplified by a factor $C = 32$, since the proported advantage of MTM over MH (Metropolis-Hastings) is that it allows for larger proposal steps.

(4) An MTM chain length of $n = 200$.

(5) A vertical correlation length scale of $L = 100\,\text{hPa}$.

(6) Despite the considerable success of an assimilation run with the above specifications, namely (1)-(5) above, an examination of the results revealed excessive cloud water path $CWP = \int \rho_c \, dz$ in the nighttime regions. In reprospect this is not surprising — at night there are no cloud optical thickness observations to contrain the vertically integrated condensate in each gridcolumn. The remaining observables, T_b and p_c, both saturate very quickly with cloud water path and so cannot constrain CWP. To resolve this issue, for nighttime gridcolumns only, we multiply the prior term $p(\boldsymbol{\alpha})$ of Part I by a Gaussian-like term

$$e^{-C_\lambda [\ln(\lambda/\lambda_B)]^2}$$

that constrains the ratio of the condensed water path to the total water path, $\lambda = CWP \,/\, TWP$, to its triangularized background state value λ_B, where $TWP = \int (\rho_v + \rho_c) \, dz$. We have found that an exponential constant $C_\lambda = 256$ works well.

(7) On closer examination of the assimilated MODIS brightness temperatures, it was found that a few were unexpectedly very cold ($\leq 220\,\text{K}$). We had thought such high clouds would use p_c rather than T_b assimilation, since CO_2-sliced cloud top pressures of $p_c \leq 550\,\text{hPa}$ are considered reliable (see Section 2.6 of Part I). Upon inspection, it was found that these particular clouds, although reporting a cloud top pressure $p_c \leq 550\,\text{hPa}$, were flagged by the MODIS algorithm as IR-window rather than CO_2-slicing derived, indicating that the CO_2-slicing retrieval had failed for some reason. This was causing our pixel gridding algorithm, described in Section 2.6 of Part I, to choose T_b rather than p_c as the observable. This was, however, inconsistent with the forward model, which always uses p_c as the observable for $p_c \leq 550\,\text{hPa}$. As a consequence, the analysis was reporting extreme negative O-A (observation minus analysis) biases in T_b at these locations, because the only forward modeled T_b were, by definition, for lower clouds. The solution was simple: to use p_c as an observable for *any* observation with a $p_c \leq 550\,\text{hPa}$, regardless of whether flagged as IR-window or CO_2-slicing derived, in order to achieve consistency with the forward model. This solution eliminated the problem with very negative O-B biases in T_b, and the p_c biases also appear to be fine (see results below).

With these modifications, the resulting run, denoted b7K64fhC32, forms the control for the sensitivity tests that follow. The run was made for the day of July 1, 2011, comprising eight three-hour cloud analyses.

Figure 1 shows global plots of cloud optical thickness τ, cloud brightness temperature T_b and cloud top pressure p_c for the control run for each of the triangularized background (B), observations (O), and result of the Bayesian analysis (A). Note that, in general, the analysis is closer to the observations than the background. We will be more quantitative shortly.

Note that in all plots and analysis presented in this Part II: (1) the use of the term "background" refers to the triangularized model background state, namely the gridcolumn of skewed triangle PDFs produced by the initialization procedure presented in Part I; (2) τ is an all-sky gridcolumn mean cloud optical thickness (zeros included for clear pixels/subcolumns), and T_b and p_c are in-cloud gridcolumn means of cloud brightness temperature and cloud top pressure, respectively; (3) T_b is the mean for only the pixels or subcolumns for which T_b is actually used, namely when $p_c > 550\,\text{hPa}$, and conversely p_c is the mean for only $p_c \leq 550\,\text{hPa}$. τ is the mean over only daytime gridcolumns.

Figure 2 shows the corresponding global plots of the O-B and O-A biases. Clearly the analysis biases are much smaller than the background biases. Figure 3 shows the PDFs of O-B and O-A biases in τ for the control run. The peaks of the O-A PDFs are all significantly narrower than for the corresponding O-B PDFs, again indicating the basic success of cloud assimilation method. Both the modal and the mean biases are also generally smaller in magnitude for O-A than O-B. The figure also shows that the analysis was quite successful in removing unobserved clouds from the background (see figure caption).

Figures 4 and 5 show the corresponding bias PDFs for the cloud brightness temperature T_b and CO_2-slicing based cloud top pressure, p_c. Similar conclusions can be drawn as from Figure 3. Note that the inset panels have a slightly different interpretation than for Figure 3. Namely, for Figure 4 the $\bullet$ refers to the existence of a T_b value, that is a "low cloud" ($p_c > 550\,\text{hPa}$), while the $\circ$ represents not only clear pixels/subcolumns, but those which do not use T_b because the cloud top is too high ($p_c \leq 550\,\text{hPa}$). Based on this interpretation, the inset illustrates that the fraction

DOI: 10.1002/qj

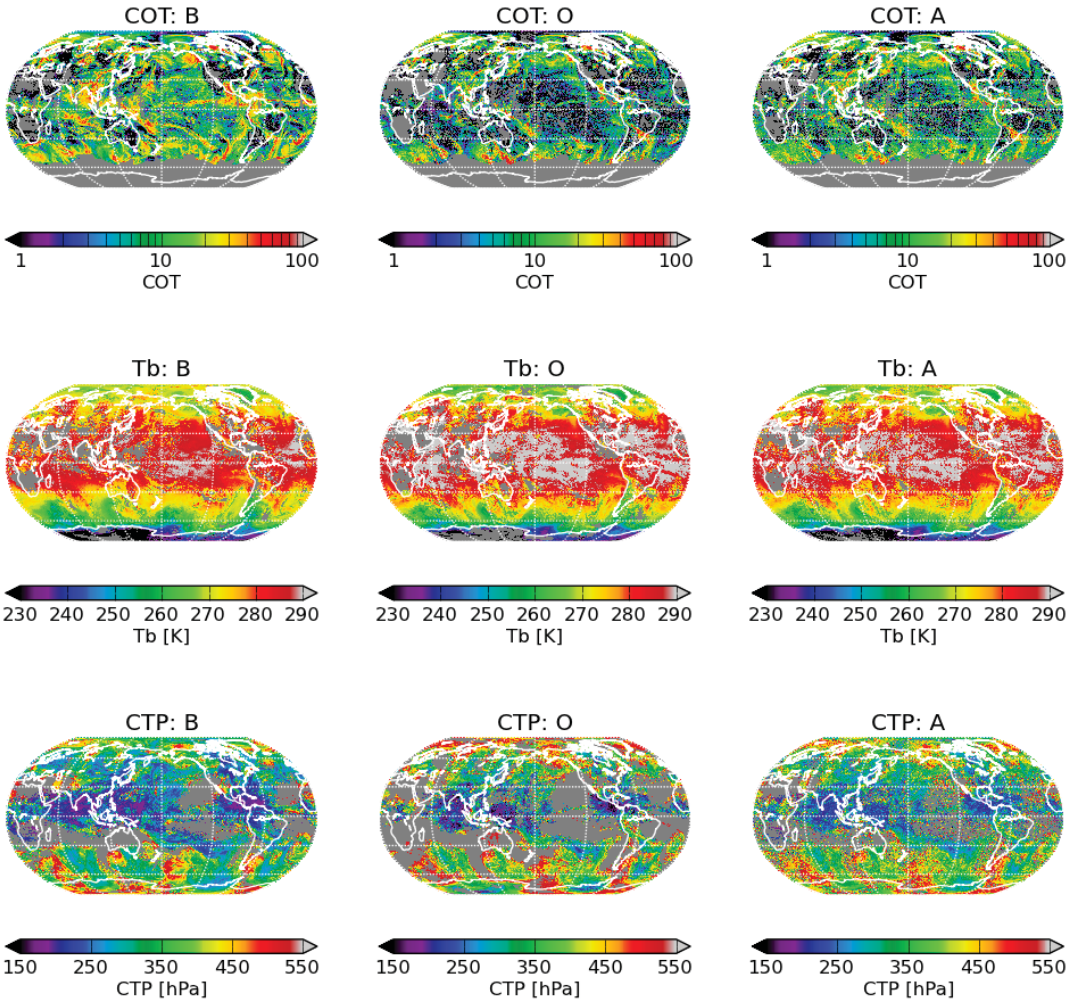

Figure 1. Global plots of all-sky gridcolumn mean cloud optical thickness τ (top) and in-cloud gridcolumn mean cloud brightness temperature T_b (mid) and cloud top pressure p_c (low) for the control run b7K64fhC32 for each of the triangularized background (B), observations (O), and analysis (A).

of modeled low clouds not seen in the observations is reduced by the analysis (transfer of upper-left red to lower-left blue) and also that the fraction of observed low clouds that are not modeled is also reduced by the analysis (transfer of lower-right red to upper-right blue). Analogous comments apply for Figure 5.

In view of the clear difference between day and night assimilation [see item (6) above] we also computed the global bias statistics for the control run separately for day and night gridcolumns. Table I shows the results. Note the following: (1) the background biases O-B are significantly smaller in the mean during the day than during the night, but very similar in standard deviation; (2) the correction of mean bias in T_b and p_c by the cloud analysis is significantly stronger during the night than during the day, and the same also applies for the standard deviation in the bias, though to a much lesser degree. In view of item (6) above, it would seem that the daytime analysis, with its additional constraint on the τ observable, achieves a reduced alignment to the T_b and p_c observables, and that addition of the nighttime prior constraint on $CWP\,/\,TWP$, although having a similar effect, is not as dominant.

Table I. Summary of the global biases for the control run b7K64fhC32, and for the control run with day and night subsetting. The format of each box is O-B $\rightarrow$ O-A, where the mean bias is reported in the upper section, and the standard deviation in the bias in the lower section. No bias for τ column is reported, since τ is only available during the day.

mean	T_b	p_c
b7K64fhC32	$1.91 \rightarrow 0.71$	$37.93 \rightarrow 8.76$
day only	$1.48 \rightarrow 0.96$	$32.17 \rightarrow 14.39$
night only	$2.21 \rightarrow 0.54$	$41.54 \rightarrow 5.25$
sdev	T_b	p_c
b7K64fhC32	$6.09 \rightarrow 3.59$	$95.41 \rightarrow 58.19$
day only	$6.04 \rightarrow 4.13$	$97.12 \rightarrow 62.43$
night only	$6.12 \rightarrow 3.16$	$94.14 \rightarrow 55.10$

DOI: 10.1002/qj

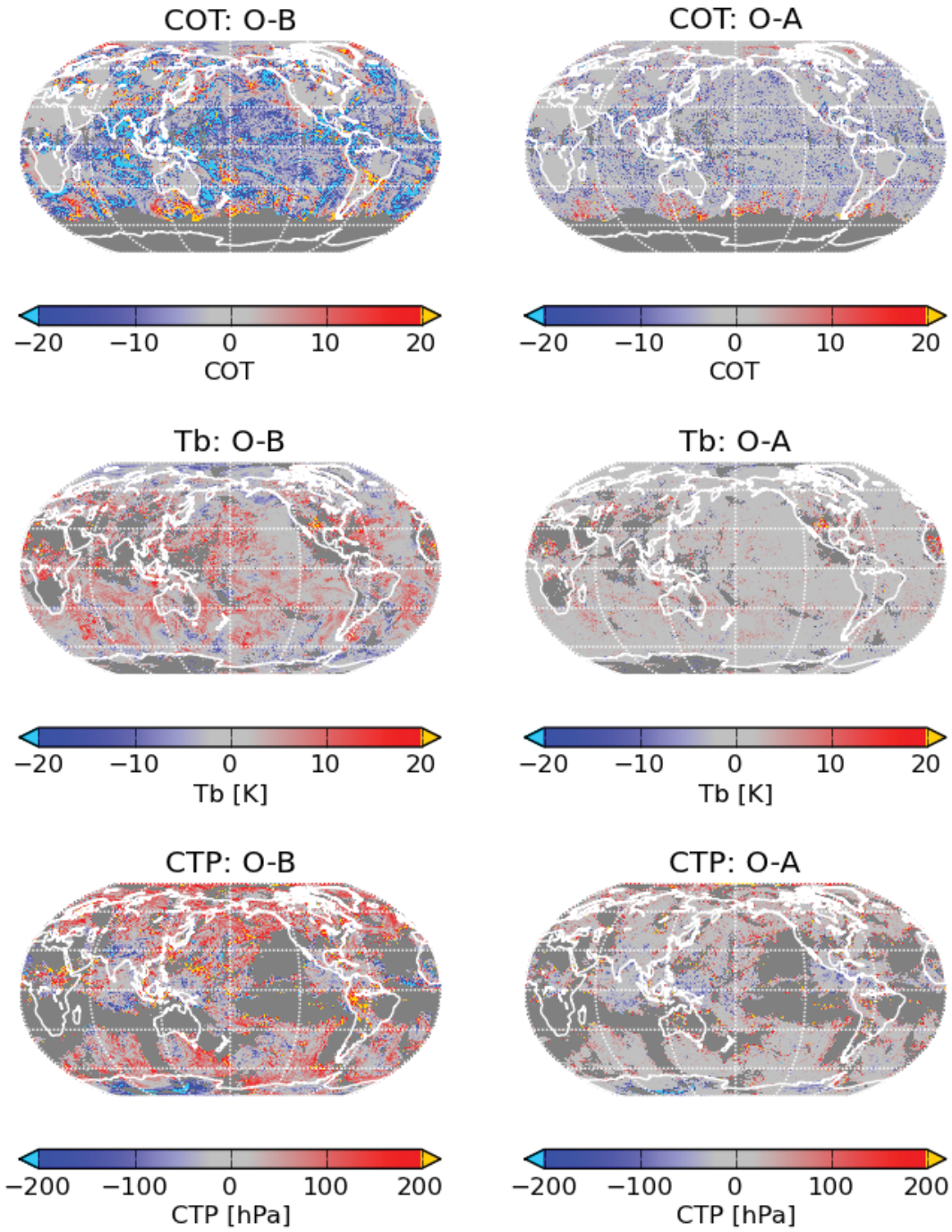

Figure 2. Global plots of O-B and O-A biases for the control run b7K64fhC32 for the observables in Figure 1. Clearly the analysis biases are much smaller than the background biases.

3 Sensitivity Tests

Table II lists the global mean biases in τ, T_b and p_c, and the standard devation (sdev) in the biases, for the control run b7K64fhC32 and for several sensitivity experiments varying key algorithmic parameters. The experiment b7K64fhC16 reduced the proposal covariance by half with $C = 16$ and produces a slightly worse mean and sdev O-A than the control. Tests with $C = 64$ (not shown) produced mixed results when compared with the control, but no significant improvement. The experiment b7K64f1C32 uses twice as many trials per MTM chain point as the control, namely $f_M = 1$. We expected this to be better, since using the same number of trial points as there are effectively independent dimensions in the parameter space seems reasonable. Indeed the results are slightly better, but certainly not enough so to justify twice

DOI: 10.1002/qj

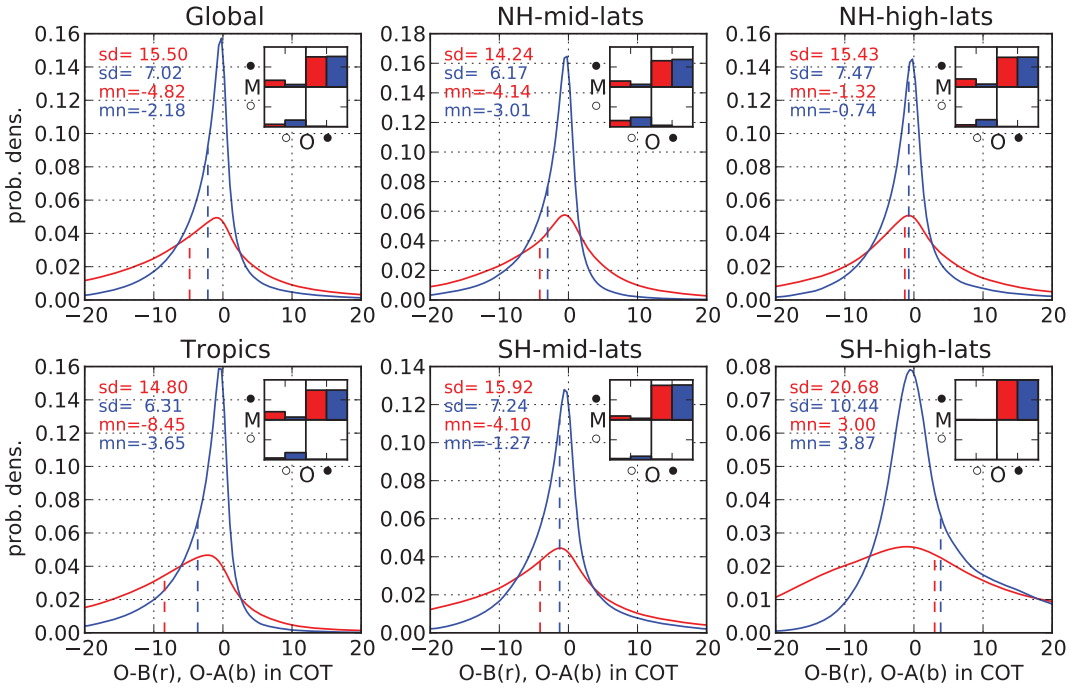

Figure 3. PDFs of the biases O-B (red) and O-A (blue) in all-sky gridcolumn mean cloud optical thickness for control run b7K64fhC32. Biases are shown for six regions, as noted, where the latitude boundaries between tropics, mid- and high-latitudes are at 22.5 and 50 degrees. The peak of the O-A PDFs are all significantly narrower than the corresponding O-B PDFs, indicating the basic success of cloud assimilation method. Both the modal biases (location of the peak) and the mean biases (dashed vertical lines) are also generally smaller in magnitude for O-A than O-B. The mean and standard deviation of the biases are also explicity printed in the upper left of each panel. Note that all these results are for the subset of gridcolumns that are non-clear (either partially cloudy or overcast) for both the observations (in the gridcolumn) and the model (B or A). This subset is also represented in the ●● quadrant of the upper right inset of each panel. This inset shows the four-way fractional split between clear (○) and non-clear (●) cases for the observations O and the model M (B or A). The fact that there is a transfer of probability from cloudy B to clear A when the observations are clear (i.e., red in upper left to blue in lower left quadrants) is also indicative of the success of the cloud assimilation method.

Table II. Summary of the variation in global biases among the different algorithmic sensitivity experiments. The format of each box is O-B → O-A, where the mean bias is reported in the upper section, and the standard deviation in the bias in the lower section.

mean	τ	T_b	p_c
b7K64fhC32	$-4.82 \rightarrow -2.18$	$1.91 \rightarrow 0.71$	$37.93 \rightarrow 8.76$
b7K64fhC16	$-4.82 \rightarrow -2.17$	$1.90 \rightarrow 0.79$	$37.96 \rightarrow 9.92$
b7K64f1C32	$-4.83 \rightarrow -1.87$	$1.91 \rightarrow 0.66$	$37.91 \rightarrow 6.89$
b7K128fhC32	$-4.78 \rightarrow -1.99$	$1.96 \rightarrow 0.67$	$37.49 \rightarrow 6.11$
b7K64fhC32_n400	$-4.82 \rightarrow -1.86$	$1.91 \rightarrow 0.64$	$37.92 \rightarrow 6.80$
b7G64fhC32	$-4.82 \rightarrow -1.94$	$1.91 \rightarrow 0.62$	$37.95 \rightarrow 10.41$
sdev	τ	T_b	p_c
b7K64fhC32	$15.50 \rightarrow 7.02$	$6.09 \rightarrow 3.59$	$95.41 \rightarrow 58.19$
b7K64fhC16	$15.51 \rightarrow 7.23$	$6.09 \rightarrow 3.81$	$95.38 \rightarrow 61.15$
b7K64f1C32	$15.51 \rightarrow 6.64$	$6.09 \rightarrow 3.49$	$95.36 \rightarrow 56.92$
b7K128fhC32	$15.40 \rightarrow 6.44$	$6.08 \rightarrow 3.39$	$95.27 \rightarrow 55.87$
b7K64fhC32_n400	$15.50 \rightarrow 6.47$	$6.09 \rightarrow 3.41$	$95.38 \rightarrow 54.93$
b7G64fhC32	$15.50 \rightarrow 6.66$	$6.09 \rightarrow 3.49$	$95.36 \rightarrow 54.21$

the expense of this experiment (doubling the number of trials approximately doubles the experiment timing). The experiment b7K128fhC32 doubles $N_{\rm sim}$ (and $N_{\bullet\rm max}$) to 128. On the whole, it is slightly better, but again, not enough so to justify twice the expense. The experiment b7K64fhC32_n400 doubles the length of the MTM chain to 400 elements. It is also slightly better than the control, but yet again not worth twice the runtime expense. All the above experiments involve algorithmic perturbations, and we conclude that the control b7K64fhC32 is satisfactory and cost effective.

One final algorithmic sensitivity experiment, b7G64fhC32, replaced the kernel density estimate (KDE) method of likelihood evaluation, described in Part I,

DOI: 10.1002/qj

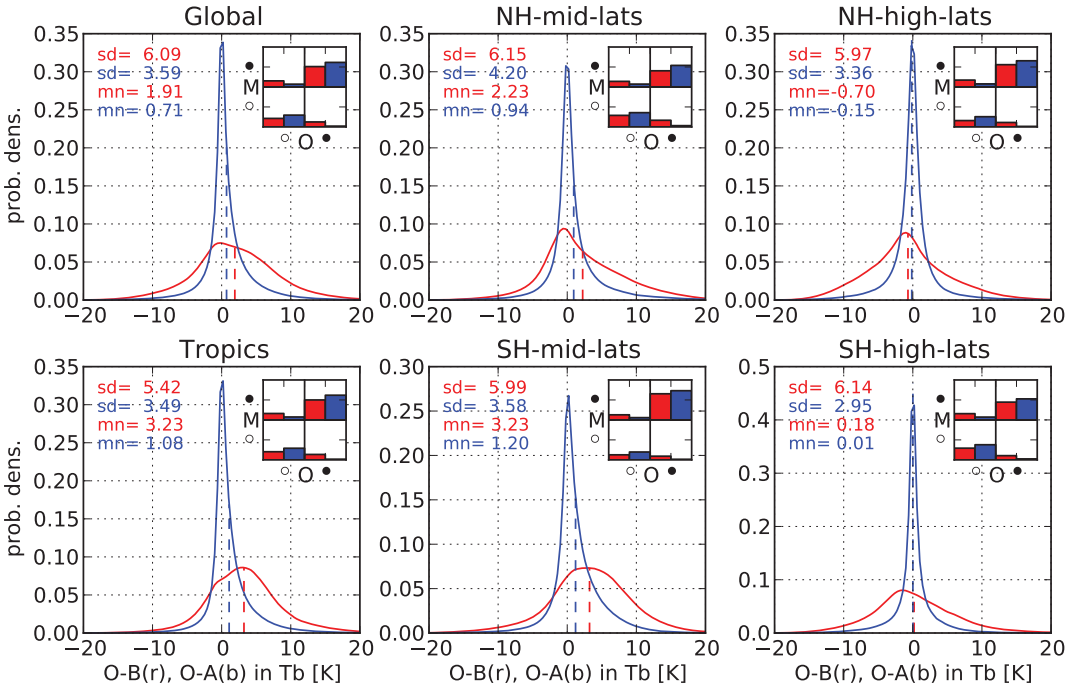

Figure 4. As in Figure 3 but for brightness temperature T_b and including nighttime as well as daytime gridcolumns. See the text for important comments on the inset, which has a different interpretation than for Figure 3.

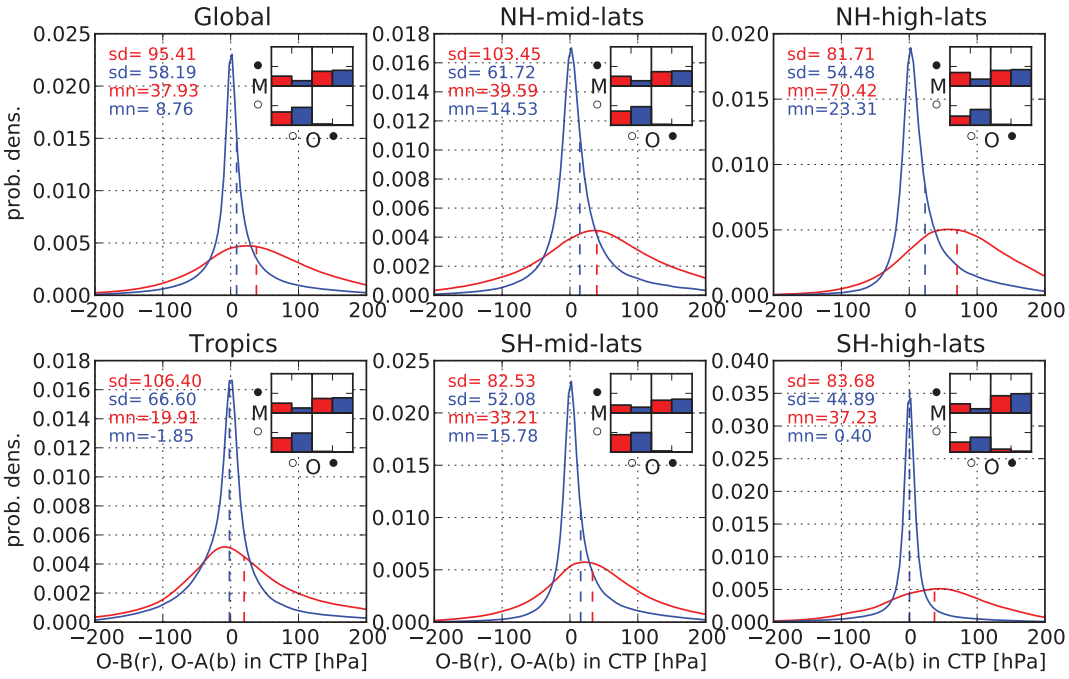

Figure 5. As in Figure 4 but for CO_2-slicing based cloud top pressure. See the text for important comments on the inset.

Section 2.8, with a simple 2D Gaussian likelihood evaluation, also discussed in that section. While this Gaussian likelihood forces the likelihood functions $\hat{p}_{\bullet p_c}((\ln\tau, p_c)|\boldsymbol{\alpha})$ and $\hat{p}_{\bullet T_b}((\ln\tau, T_b)|\boldsymbol{\alpha})$ to be unimodal and of Gaussian form (e.g., with elliptic contours), the results from Table II are actually a small improvement on the control in all measures except global p_c bias. Although we retain the KDE likelihood in the control, for its generality and all-round good performance, we intend to continue to investigate the Gaussian likelihood,

DOI: 10.1002/qj

because the simplicity of the Gaussian may allow some simplification and speed-up of the code (and perhaps a simpler integration of our CDA method with more traditional 3D- and 4D-Var approaches.) That being said, the Gaussian likelihood version is not currently faster than the KDE version because the statistical subcolumn generation dominates the evaluation of the likelihood at the observations. This might be reversed with further coding improvements. [We also plan to try a Gaussian Copula (GCOP) likelihood, a middle ground between the KDE and Gaussian likelihoods, as mentioned in Part I. This is distinct from the GCOP model of layer overlap already included in this study.]

Table III examines the sensitivity of global biases to more physical parameters. The experiment b7K64fhC32_L200 doubles the vertical correlation length scale L to 200 hPa. This is somewhat worse than the control. It is technically possible to include L in the parameter list for Bayesian inference, but we have not yet tried this. The experiment b7K64fhC32_pTrop replaces the fixed $p_{\text{ramp}} = 100$ hPa of Part I, Section 2.5, with a variable tropopause pressure, as output by the background GEOS-5 simulation, and replaces the fixed $p_{\text{lim}} = 50$ hPa with $p_{\text{ramp}} - 50$ hPa. [Note: p_{ramp} is the lowest pressure at which full parameter variability is permitted. Above this level, the parameter variability decays linearly with pressure, to zero at p_{lim}. This ramp effects both the $\bar{S}$ prior of Part I, Section 2.5, and the MTM estimated target covariance, Σ_π, discussed near the end of Part I, Section 2.9]. The results for b7K64fhC32_pTrop are very similar to the control run, with the exception that the mean cloud top pressure bias is significantly reduced (from a 77% reduction O-B $\rightarrow$ O-A to an 83% reduction). This improvement makes sense: we are constraining the analyzed cloud top pressure to approximately below the background tropopause, rather than a nominal level of 100 hPa in the control. In retrospect, this modification should have been included in the control, and will be for future work, but not for the current paper, except where noted.

The experiment b7K64fhC32_lnP uses $\ln p$ rather than p in matters relating to the specification of vertical correlation. This has some justification, since $\ln p$ is a proxy for the height above the surface, z, and it is more common to parameterize vertical correlation in terms of z than p. The implementation is as follows: in the evaluation of the vertical correlation matrix C in Part I, equation (5), $\xi_{kk'} = |p_k - p_{k'}|/L$ is replaced by $\xi_{kk'} = |\ln p_k - \ln p_{k'}|/L_{\ln}$, where $L_{\ln}$ is a "length scale" in $\ln p$, given by $L_{\ln} = \ln[(p_{\text{ref}} + \Delta p/2)/(p_{\text{ref}} - \Delta p/2)]$, such that $p_{\text{ref}} + \Delta p/2$ and $p_{\text{ref}} - \Delta p/2$ are separated by one length scale in $\ln p$, where p_{ref} is some reference pressure, set to 500 hPa for this experiment, and $\Delta p = 100$ hPa, akin to the $L = 100$ hPa of the control. This means that 450 and 550 hPa have the same decorrelation in both the control b7K64fhC32 and the b7K64fhC32_lnP experiment, but that other levels separated by 100 hPa now have different decorrelations from the control. In particular, upper levels separated by 100 hPa are relatively less correlated, since Δp higher in the atmosphere represents a relatively greater altitude difference, than the same Δp near the surface. In addition, the linear ramp in p between p_{ramp} and p_{lim}, discussed above, for near tropopause decay of parameter variability, now becomes linear in $\ln p$. Lastly, in the calculation of the number of trials M per point in the MTM chain (end of Part I, Section 2.9), $(p_0 - p_{\text{lim}})/L$ is replaced by $\ln(p_0/p_{\text{lim}})/L_{\ln}$. The performance of the b7K64fhC32_lnP experiment is at best mixed, with small improvements for p_c and degradations for T_b. The lnP experiment is also about 60% more expensive than the control, due to the increase in the number of trials M per MTM chain point, as described above $[\ln(p_0/p_{\text{lim}})/L_{\ln} \div (p_0 - p_{\text{lim}})/L = \ln(1000/50)/\ln(550/450) \div 950/100 \approx 15 \div 9.5 \approx 1.6]$. Using a p_{ref} lower in the atmosphere exacerbates this problem. The experiment b7K64fhC32_lnPtr is a combination of b7K64fhC32_pTrop and b7K64fhC32_lnP. The pTrop modification generally lowers p_{lim} within the atmosphere and therefore decreases M, resulting in a small overall decrease in computation time for b7K64fhC32_lnPtr relative to the control. However the results from Table III are not, on balance, better than the lnP experiment, with a slight improvement in p_c being offset by a slight deterioration in COT.

Finally, the experiment b7K64fhC32_RiiS1 implements a version of the simple flow-dependent correlation function proposed by Riishojgaard (1998). The idea is to replace the "pressure distance" based vertical correlation matrix C of Part I, equation (5), by the Hadamard product of itself and another correlation matrix D with elements $\mathsf{D}_{kk'} = \nu(|\theta(\boldsymbol{\alpha}^{\text{b}}_k) - \theta(\boldsymbol{\alpha}^{\text{b}}_{k'})|)$, where ν is a correlation function on R and θ is a function of the background state vector $\boldsymbol{\alpha}^{\text{b}}$ for a layer. In this study we use the Gaussian correlation function $\nu(r;\sigma) = \exp[-r^2/(2\sigma^2)]$. Riishojgaard suggests the use of some conservative function for θ, on the basis that Lagrangian transport preserves conservative properties. In effect, the D matrix would therefore decorrelate regions which have different Lagrangian origins. The use of a conservative θ is most important for horizontal correlations, since advection-based variability dominates in horizontal. In the vertical, other processes such as turbulent diffusion and radiatiave heating become significant. For our case of vertical correlations, we could use the conservative total water content q_t, but we can just as well use the control variable $\bar{S}$, the mean total saturation ratio, which has a smaller dynamic range but still captures well the sharp humidity inversion at the top of the planetary boundary layer, for example. This is really our goal — to use the D term to provide strong decorrelation across inversion features in the vertical, which are otherwise not adequately decorrelated by the fixed pressure scale L of the control C model. Thus the experiment b7K64fhC32_RiiS1 applies the above D modification, with $\theta(\boldsymbol{\alpha}^{\text{b}}_k) = \bar{S}^{\text{b}}_k$ and a standard deviation $\sigma = \kappa_S\, \sigma_{\bar{S}}$, where $\sigma_{\bar{S}} = 0.1$ is the prescribed standard deviation of $\bar{S}$ in the control (see Part I, Section 2.5 and 2.9) and κ_S is a constant, set to one for this experiment. Looking

DOI: 10.1002/qj

Table III. Summary of the variation in global biases among the sensitivity experiments examining physical parameters. The format is the same as in Table II.

mean	τ	T_b	p_c
b7K64fhC32	$-4.82 \rightarrow -2.18$	$1.91 \rightarrow 0.71$	$37.93 \rightarrow 8.76$
b7K64fhC32_L200	$-4.40 \rightarrow -3.61$	$1.79 \rightarrow 1.15$	$40.95 \rightarrow 17.40$
b7K64fhC32_pTrop	$-4.81 \rightarrow -2.16$	$1.91 \rightarrow 0.66$	$37.90 \rightarrow 6.26$
b7K64fhC32_lnP	$-4.64 \rightarrow -2.21$	$1.99 \rightarrow 1.05$	$37.81 \rightarrow 6.66$
b7K64fhC32_lnPtr	$-4.62 \rightarrow -2.34$	$1.99 \rightarrow 1.04$	$37.84 \rightarrow 5.31$
b7K64fhC32_RiiS1	$-5.12 \rightarrow -0.66$	$2.09 \rightarrow 0.35$	$35.15 \rightarrow 3.05$
sdev	τ	T_b	p_c
b7K64fhC32	$15.50 \rightarrow 7.02$	$6.09 \rightarrow 3.59$	$95.41 \rightarrow 58.19$
b7K64fhC32_L200	$15.26 \rightarrow 9.07$	$6.17 \rightarrow 4.29$	$95.13 \rightarrow 67.50$
b7K64fhC32_pTrop	$15.49 \rightarrow 7.12$	$6.09 \rightarrow 3.55$	$95.37 \rightarrow 56.94$
b7K64fhC32_lnP	$15.40 \rightarrow 7.02$	$6.14 \rightarrow 3.80$	$95.21 \rightarrow 55.61$
b7K64fhC32_lnPtr	$15.39 \rightarrow 7.29$	$6.14 \rightarrow 3.82$	$95.21 \rightarrow 54.95$
b7K64fhC32_RiiS1	$15.70 \rightarrow 6.22$	$6.05 \rightarrow 3.23$	$95.68 \rightarrow 53.65$

at the results from Table III, we see that this experiment is superior to the control in all measures. We believe that using the D matrix is better honoring the basic vertical structure of the atmosphere, as represented in the background state, and to the extent that this background state is able to capture realistic vertical structure, is producing an analyzed state more consistent with the observations. In retrospect, this modification should have been included in the control, and will be for future work, but not for the current paper, except where noted.

4 California Stratocumulus Study

Figure 6 shows a case study off the west coast of North America in which marine stratocumulus is present right up to the coast in Southern California and Baja California, but is absent in a wide swath off this coast in the background. In such cases it is typically very difficult for a cloud data assimilation system to restore equilibrium cloud to such regions because the background, subsaturated clear regions cannot produce cloudiness for any small equilibrium perturbation to the moisture field. However, the Monte Carlo Bayesian system *is* able to restore the stratocumulus, since it makes finite jumps in parameter space. The cloud brightness temperature T_b in the restored region is in good agreement with the observations, although the cloud optical thickness appears a little too high. Note that these results are for the experiment b7K64fhC32_RiiS1, that is, the control with the Riishojgaard (1998) correlation modification. While the pure control was almost as good, and certainly also restored the near coast stratocumulus, it produced a more noisy analysis with even higher COT in the restored region. It seems that the Riishojgaard modification is acting, as anticipated, to honor the background moisture inversion structure, thereby limiting the ability of the analysis to produce excessive cloud thickness by artificially raising the cloud top.

Note that our earlier experiments with this marine stratocumulus case were a failure. In those experiments we were using τ and p_c only as observables, not T_b. The failure occurred because the MODIS retieved CO_2-slicing cloud top pressure is unrealiable for low clouds (below $550\,$hPa), and was producing cloud top pressures in the range 650–$750\,$hPa, significantly above the marine inversion, which is below $900\,$hPa. The anatomy of the failure was as follows: the assimilation system was trying to produce a cloud with a top around $700\,$hPa (based on the erroneous MODIS p_c), but the background relative humidity at this height, well above the marine inversion, was only about 10%. Clearly the prior in $\bar{S}$ prevented the system from shifting the mean too far from 0.1, so the system instead produced a very positively skewed moisture PDF, with the upper tail just crossing saturation. The resulting small cloud fraction (less than 10%) near $700\,$hPa was able to raise the simulated cloud top pressure towards the incorrect MODIS value, but at the expense of changing the observed cloud fraction of close to 100% to less than 10%! This illustrates the potential problems that can arise when assimilating erroneous data, or alternatively, in using a forward model that does not well simulate the actual observable. In practise, had we made cloud fraction an additional observable with a prior constraint, then the outcome would have been different. In that case, the highly skewed, low cloud fraction solution would have been impossible, and the analysis would have remained near the background, which would have been preferrable. But as it was, we chose a better solution: we dropped the p_c observable in favor of T_b below $550\,$hPa.

5 Comparison between MODIS and SEVIRI assimilation

We ran an assimilation with a short 15 minute window centered on 13:52:30 GMT of July 1, 2011, to compare the assimilation of MODIS Aqua cloud retrievals (as above) with assimilation of cloud retrievals made by the NASA Langley Cloud and Radiation Group from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument aboard the Meteosat-9 geostationary platform (see Minnis *et al.*, 2006). [The above time is near the Aqua MODIS viewing of $0°$N, $0°$E, the sub-satellite point of Meteosat-9.] The Meteosat-9 SEVIRI retrievals have the

DOI: 10.1002/qj

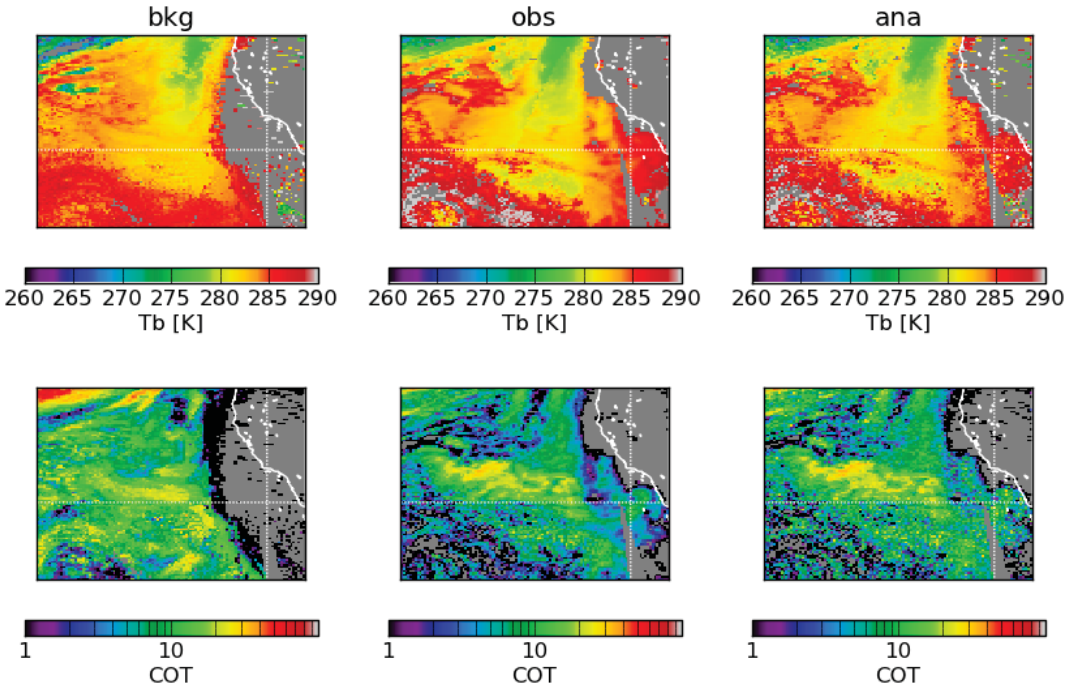

Figure 6. A case study of marine stratocumulus off the coast of California and Baja California. The upper panels show brightness temperature T_b and the lower panels all-sky cloud optical thickness (τ). The background (left) is absent the stratocumulus near the Southern or Baja Californian coast that is present in the observations (middle). The analysis (right) is able to restore this stratocumulus, albeit with a slightly higher COT than observed. The somewhat noisier appearance of the analysis comes from the fact that each gridcolumn analysis is currently independent of its neighbors.

advantage of large spatial coverage (including the whole of Europe, the Atlantic Ocean, and all of Africa) at high temporal resolution (every 15 minutes), compared with the twice-a-day overpasses of the Aqua and Terra polar-orbiting platforms. The nadir field of view is also reasonably small at 3 km, though not as small as the MODIS 1-km optical retrievals. From an assimilation point of view, the main difference is that only cloud optical thickness τ and brightness temperature T_b are assimilated from the SEVIRI retrievals, not CO_2-slicing cloud top pressure p_c as for MODIS. Thus, while the MODIS assimilation switches between T_b and p_c for cloudy pixels, depending on the value of p_c, as described earlier and in Part I, the SEVIRI assimilation always uses T_b when cloud is present.

Comparing Figures 7 and 8, it seems that assimilation of the two retrieval datasets is comparable (though the SEVIRI dataset has the obvious advantage of being available as a full-disk image every 15 minutes.) An examination of the corresponding O-A PDFs (not shown) confirms this comparability, although the ability of the assimilation algorithm to remove unobserved cloud from the background appears to be a little better for MODIS than SEVIRI assimilation. [Both experiments use the b7K64fhC32 control with the tropopause (pTrop) and Riishojgaard (RiiS1) modifications discussed above.]

6 Validation using OMI

The above results provide substantial evidence that the Bayesian cloud data assimilation system is at least doing what it was designed to do: to drive the subcolumn model to a new state more consistent with the observations. We also have evidence, from the Stratocumulus case study, that the CDA system is also able to handle particularly tricky cases where there is no linear sensitivity of the observations to the model. But we also want a broader global validation of the improvement of the analyzed state, as compared with some independent, non-assimilated dataset. For this we turn to the so-called Optical Centroid Pressure (OCP) retrieved from a rotational-Ramman scattering (RRS) algorithm (see Joiner *et al.*, 2012, hereafter J12) applied to UV spectral measurements made by the Ozone Monitoring Instrument (OMI) aboard the polar-orbiting Aura satellite (at the rear of the so-called "A-Train", 15 minutes behind Aqua).

The OCP is a measure of how far into a cloud one can see in the UV, which is much further than at IR wavelengths. Therefore, the OCP is much deeper inside a cloud than the IR-window or CO_2-slicing derived cloud-top pressure (CTP), which is typically very near the cloud top for all but very thin clouds. As such, the OCP is senstive to both the CTP and the cloud optical thickness (COT). That being said, we must realistically acknowledge that the Bayesian CDA system descibed in this paper (Part I) does not put a strong constraint

DOI: 10.1002/qj

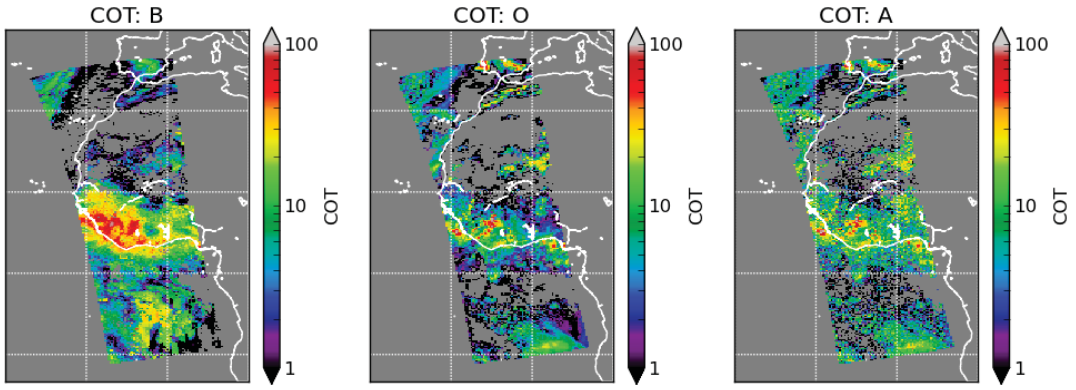

Figure 7. Cloud optical thickness for an assimilation of MODIS Aqua cloud data for a 15 minute window centered on 13:52:30 GMT of July 1, 2011. These results should be compared with the SEVIRI assimilation results in Figure 8.

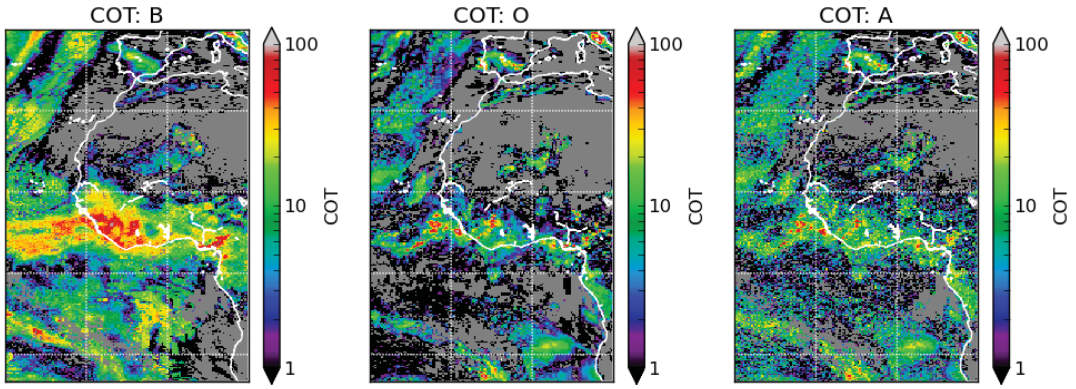

Figure 8. As in Figure 7, but for assimilation of cloud retrievals made by the NASA Langley Cloud and Radiation Group for Meteosat-9 SEVIRI data (see text.)

on the vertical structure of the cloud, but mainly on the cloud top and the overall optical thickness. This limitation is particularly true for multi-layer clouds. Thus we cannot necessarily expect to get a major improvement in OCP simulations from the analyzed cloud state, but some improvement is hoped for.

For the observations we use the OMI OMCLDRR product, which contains both the effective cloud fraction $f_{\rm eff}$ ("CloudFractionforO3") and optical centroid pressure $P_{\rm OCP}$ ("CloudPressureforO3") Scientific Datasets (SDS). The nadir FOV is approximately 12 km along track and 24 km across track. We only include FOVs satisfying the following conditions, based on guidance from J12 and the OMCLDRR README file provided with the data: (1) the OCP is valid if bits 0, 1, 2, 3, 4, 6, 7, 13, 14, and 15 of the "ProcessingQualityFlagsforO3" SDS are all zero, and if the $P_{\rm OCP}$ value itself has some positive value; (2) $f_{\rm eff}$ is valid if $P_{\rm OCP}$ is, and we further require $f_{\rm eff} \geq 0.3$, since the algorithm is very noisy for small effective cloud fractions; (3) we also exclude FOVs flagged as snow/ice by bit 5 of the above processing quality flag (pqf), since for these surface types $f_{\rm eff}$ is set to one and so $P_{\rm OCP}$ is approximate; (4) we also exclude FOVs with radiance or irradiance errors according to bits 9 and 11 of the pqf; (5) finally, we exclude FOVs possibly compromised by sea-glint under low cloud fraction conditions (namely $f_{\rm eff} < 0.3$ and bit 4 of the 'GroundPixelQualityFlags" SDS set). All FOVs satisfying these conditions are binned into the reduced longitude grid of the analysis and the mean (in-cloud) $P_{\rm OCP}$ is calculated for each accepted gridcolumn. In addition, only those gridcolumns for which the corresponding observed gridcolumn mean all-sky COT from Aqua is at least 5 are included in the analysis, since the retrieved OCP is believed to be less reliable for optically thinner clouds (see J12).

We study the day February 13, 2007 and use a modified CDA control run that only assimilates Aqua (not Terra) observations, since OMI is onboard Aura which closely follows behind Aqua. A $P_{\rm OCP}$ is simulated for the triangularized background (B) and analyzed (A) states using the fast "R^{3}S" simulator of J12. Namely $P_{\rm OCP}$ is evaluated according to equation (5) of J12: $P_{\rm OCP} = \sum_l \rho_l P_l / \sum_l \rho_l$, where P_l is the midpoint model layer pressure and ρ_l is the contribution to the cumulative UV cloud reflectance from layer l, as described in J12 (using the adding-doubling method and a conservative delta-Eddington two-stream calculation of layer reflectances and transmittances, with assymmetry factor $g = 0.8$.)

DOI: 10.1002/qj

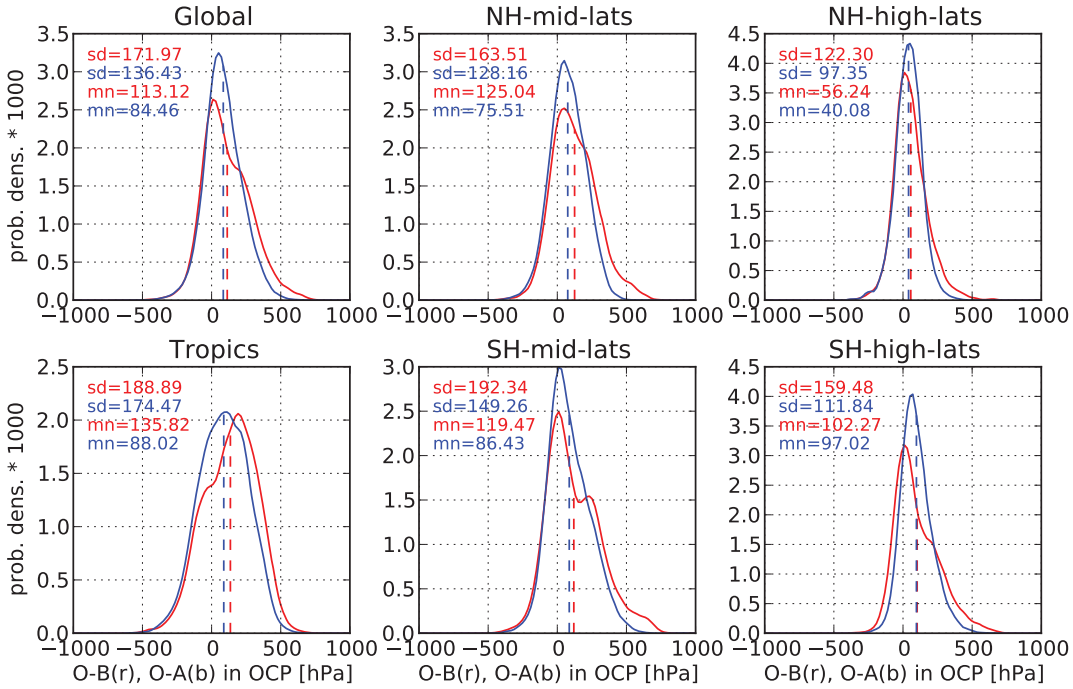

Figure 9. PDFs of the biases O-B (red) and O-A (blue) of in-cloud mean gridcolumn cloud optical centroid pressure $P_{\rm OCP}$ for a modified control run that assimilates only Aqua (not Terra) observations. Biases are shown for six regions, as noted, where the latitude boundaries between tropics, mid- and high-latitudes are at 22.5 and 50 degrees.

Figure 9 shows $P_{\rm OCP}$ bias PDFs for O-B and O-A, globally and for different regions. The biases are in the in-cloud mean $P_{\rm OCP}$ for each gridcolumn in the specified region and for each of the eight assimilation time windows in the 24-hour period. Global mean biases are reduced by about 25% and the standard deviation in the biases by about 21%. Biases also appear to be more Gaussian after assimilation, which is helpful given that these analyzed states will eventually become the background for a more traditional 3D- or 4D-variational data assimilation.

It is somewhat dissapointing that the biases were not more strongly reduced, but as stated earlier, the CDA method does not strongly constrain cloud vertical structure, and also, the fast R^{3}S OCP simulator only provides an approximation to the retieved OCP. To investigate the potential role of cloud vertical structure, we also studied the global mean bias and standard deviation reductions as a function of COT, τ. For moderately thin clouds ($5 \leq \tau \leq 10$) the global mean bias was reduced by 33% and the standard deviation by 24%. For clouds with $10 \leq \tau \leq 25$, these reductions are 27% and 23%, and for $\tau \geq 25$, only about 12% and 9%. Thus, the reductions do seem to deteriotate with increasing cloud thickness, and especially for thick clouds ($\tau \geq 25$), which is consistent with the idea that cloud vertical structure will be more important for thick clouds. That being said, the addition of the "pTrop" modification produced no significant change in the OCP bias reductions, and the Riishojgaard modification produced a small degradation, so it is clear that the issue of constraint of vertical cloud structure is not yet well understood and needs further study. It should also be noted that the Riishojgaard modification still uses the same number of MTM trials M per chain point as the control, calculated based on the standard pressure scale length method of Part I Section 2.9. This should really be modified when using the Riishojgaard correlation modification, but it is not immediately clear to us how to do that.

7 Discussion and Conclusions

Part I of this paper presented a new Monte Carlo Bayesian method for constraining a complex statistical model of GCM sub-gridcolumn moisture variability using high-resolution MODIS cloud data. The method has strong application possibilities in cloud data assimilation and cloud parameterization testing and development, since it can be used to evaluate and improve background (prior) estimates of the state of a model gridcolumn and its sub-grid parameterization parameters. We have chosen one particular gridcolumn statistical model for this study, namely skewed triangle PDFs of intra-layer moisture and a Gaussian copula to couple them in the vertical, but many other more or less complicated models can be conceived and intercompared, using this method, for fidelity in simulating observed cloud data.

Part II has performed some basic testing of the new method, verifying that it does indeed very significantly reduce mean and standard deviation biases with

DOI: 10.1002/qj

respect to assimilated MODIS cloud optical depth, brightness temperature and cloud top pressure. Of particular interest, a case study of marine stratocumulus off the Californian coast has demonstrated that the Monte Carlo method performs well in the especially difficult case where the background state is clear but cloudy observations exist. In traditional linearized data assimilation methods, a subsaturated background cannot produce clouds via any infinitesimal equilibrium perturbation, but the Monte Carlo approach allows finite jumps into regions of non-zero cloud probability. In the example provided, the method was able to restore marine stratocumulus near the Californian coast where the background state had a large clear swath.

One important application of the method is assimilation of cloud data into large-scale numerical weather prediction models. By showing that the method also easily assimilates geostationary SEVIRI retrievals, which are available every 15 minutes with huge spatial coverage, we have opened up the possibility of future significant assimilation of such retrievals to constrain the global moisture field, of which clouds are a significant marker.

One obvious difficulty for the new method (and for other CDA methods as well) is the limited information content on cloud vertical structure in the chosen cloud observables, cloud top pressure (or brightness temperature) and column optical depth. Yet even these variables do produce a significant improvement in the simulated rotational-Ramman scattering cloud optical centroid pressure (OCP) against independent (non-assimilated) retrievals from the OMI instrument. Because one can see deeper into a cloud in the UV, the OCP is sensitive not only to the cloud top pressure, but also to cloud vertical structure. Clearly the fidelity of the model's background vertical structure, including boundary layer inversions, will be important for this problem, and it is important to honor such features in the assimilation process. It was found that a simple flow-dependent correlation function due to Riishojgaard (1998) provides some help in this respect, since it permits stronger decorrelations across moisture (or other) inversions in the background state. To further improve the analyzed vertical structure of the cloud field, future work can consider the addition of new observables, with stronger vertical sensitivity, to the Bayesian problem: OCP is one such observable, but other multispectral passive observations, or active sensors, can be considered as well.

Acknowledgements

This work was supported by a NASA grant from the Modeling, Analyis, and Prediction program (solicitation NNH08ZDA001N-MAP, proposal title: "Assimilation of A-Train satellite data for constraining a new PDF-based cloud parameterization in GEOS-5", PI: Arlindo da Silva). The authors wish to thank Steven Platnick and Gala Wind for providing the MODIS cloud data and much useful insight into its properties and use; Pat Minnis and Rabindra Palikonda for the SEVIRI retrieval files used in the SEVIRI comparison and guidance on their use; and Joanna Joiner for providing the rotational-Ramman scattering OCP fast-simulator code and guidance on the use of the OMI OMCLDRR dataset.

References

Joiner J, Vasilkov AP, Gupta P, Bhartia PK, Veefkind P, Sneep M, de Haan J, Polonsky I, Spurr R. 2012. Fast simulators for satellite cloud optical centroid pressure retrievals; evaluation of OMI cloud retrievals. *Atmos. Meas. Tech.* **5**: 529–545. doi:10.5194/amt-5-529-2012.

Minnis P, Nguyen L, Smith WL, Palikonda R, Doelling DR, Ayers J, Trepte Q, Chang F. 2006. MSG SEVIRI applications for weather and climate: Cloud properties and calibrations. In *Proc. 3rd MSG RAO Workshop, Helsinki, Finland, June 15, 2006*. pp. 25–30.

Riishojgaard LP. 1998. A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus* **50A**: 42–57.