Use of Markov Chain Monte Carlo Sampling Methods to Assess and Improve Variational MODIS Cloud Retrievals

Derek Posselt*, Tristan L'Ecuyer, and Graeme Stephens
Department of Atmospheric Science, Colorado State University

1. Introduction

Many recently-proposed "physical" retrieval approaches and data assimilation techniques are based on Bayesian estimation (optimal estimation) theory, and by necessity assume mean-zero Gaussian error statistics. As a result, a Gaussian probability density distribution is imposed on the estimated state, and description of the error statistics is restricted to mean and (co)variance alone. If the estimation problem is well-posed and well-constrained such that a single, well-defined probability maximum exists, then the optimal estimation method can be quite effective. However, there is no way to know whether the PDF of the errors is truly Gaussian in form, and hence whether the variational method has produced a robust result. Given relevant observations and a range of realistic values for each estimated parameter, Markov chain Monte Carlo (MCMC) algorithms can be used to sample the true joint conditional distribution, efficiently seeking regions of relatively high probability via a quantitative measure of the fit between observations and estimated variables. Once a sufficient number of samples have been obtained, the resulting PDFs can be used to quantitatively assess the robustness of the variational solution for each estimated variable.

In this paper, the properties of the MCMC algorithm are demonstrated using a relatively straightforward two-parameter MODIS cloud retrieval. Observations consist of MODIS reflectances in visible and near-infrared channels, and a physically realistic range of possible values of cloud particle effective radius and cloud optical depth is used to define the probability space to be sampled. Results from a MCMC-based retrieval are quantitatively compared with results from an optimal estimation retrieval, and the utility of the MCMC method is explored. It is found that the PDF of optical depth obtained from MCMC is decidedly log-normal for optical depths greater than approximately 20, leading to a loss of sensitivity to large optical paths in the retrieval when a Gaussian PDF is assumed. When the variational retrieval is modified such that the natural log of optical path is estimated in place of the optical path itself, direct improvements in the variational retrieval result. We conclude that the MCMC method can provide unique and valuable information on the error characteristics of retrieved variables that can be used to improve the performance of an operational variational retrieval.

*Corresponding author address:
Derek Posselt
Colorado State University
Fort Collins, CO. 80523
derek.posselt@atmos.colostate.edu

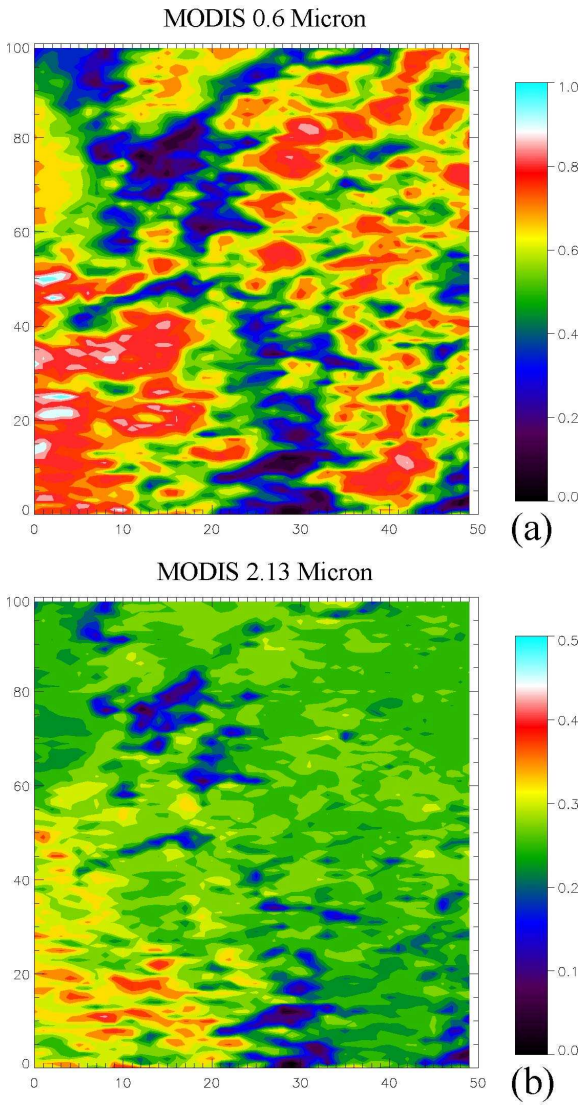

Figure 1: MODIS observed reflectance for 0.6 micron (a) and 2.13 micron (b) channels. Scene was observed over the northeast Pacific ocean at approximately 2130 UTC 4 July 2001.

2. Radiative Transfer Model and Variational Retrieval

2a. Description of the Radiant Radiative Transfer Model

The forward radiative transfer model used in the retrieval is a simplified version of the Radiant model (Gabriel et al. 2005), which is a multi-stream plane-

parallel solver applicable to both visible and infrared portions of the electromagnetic spectrum. Radiant employs a multiple scattering solution based on a combination of the well-known interaction principle and the eigenmatrix method and provides an accurate and computationally efficient solution for radiances in both cloudy and clear conditions. For our test case, we retrieve cloud optical path and effective radius from measured reflectances in 0.64 and 2.13 micron wavelengths in a manner analogous to the method described in Nakajima and King (1990). The scene of interest was observed at approximately 2130 UTC on 4 July 2001 by the MODIS instrument on EOS Terra, and consisted of low broken stratus and stratocumulus clouds over the northeast Pacific Ocean (Fig. 1).

2b. Variational Retrieval Method

Given a model F(**x**) that relates a set of observed reflectances **y** to a set of state parameters **x** we would like to estimate the true atmospheric state that produces these observations. The solution to this problem can be obtained by maximizing the conditional probability that the state is equal to the true state given the available observations and the radiative transfer model. For our purposes, the state consists of optical thickness and cloud particle effective radius, and the observations are MODIS reflectances at 0.64 and 2.13 microns. Because the state and observation error probability density distributions are often not known, it is common to assume that they take a Gaussian form. This is both useful and appropriate when only the mean and variance are known as the characteristics of the Gaussian PDF are fully determined by only two parameters (Devore 1995). Gaussian distributions are also algebraically easy to manipulate and have the added benefit that the most probable solution is identically that which minimizes the error-weighted difference between state and observations. The solution is then obtained by minimizing a cost function that consists of the negative log of the Gaussian joint probability distribution

$$\Phi(\mathbf{x}) = -2\ln P(\mathbf{x}|\mathbf{y}) = [\mathbf{F}(\mathbf{x})-\mathbf{y}]^T \mathbf{S}_\mathbf{y}^{-1}[\mathbf{F}(\mathbf{x})-\mathbf{y}] + [\mathbf{x}-\mathbf{x}_\mathbf{a}]^T \mathbf{S}_\mathbf{a}^{-1}[\mathbf{x}-\mathbf{x}_\mathbf{a}] + C \quad (1)$$

where $\mathbf{x_a}$ is the *a priori* value of the state, $\mathbf{S_y}$ and $\mathbf{S_a}$ are the observation and *a priori* error covariance matrices, respectively, and C is the normalization constant for the Gaussian probability distribution. The maximum probability (minimum variance) solution is found by minimizing the gradient of the cost function with respect to the set of unknown **x**. In this case, we set the cost function gradient equal to zero and solve for **x** through a Newtonian iterative solution to the resulting equation (Rodgers 2000)

$$\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_i + \left(\mathbf{S}_\mathbf{a}^{-1} + \mathbf{K}_i^T \mathbf{S}_\mathbf{y}^{-1}\mathbf{K}_i\right)\left[\mathbf{K}_i^T \mathbf{S}_\mathbf{y}^{-1}(\mathbf{y}-\mathbf{F}(\hat{\mathbf{x}}_i)) - \mathbf{S}_\mathbf{a}^{-1}(\hat{\mathbf{x}}_i - \mathbf{x}_\mathbf{a})\right] \quad (2)$$

Here, $\mathbf{x}_i$ is the current estimate, $\mathbf{x}_{i-1}$ is the prior estimate, **K** is the matrix of partial derivatives of the forward model with respect to **x** (the Jacobian matrix), and $\mathbf{K}^T$ is its transpose. Equation (2) makes it clear that the solution depends largely on both the specification of the form and magnitude of error covariances $\mathbf{S_y}$ and $\mathbf{S_a}$. Two key assumptions must be made (1) magnitude of the covariances and (2) imposition of a Gaussian shape. In the next section, we describe how MCMC can be used to obtain the full unapproximated conditional PDF, and how this PDF can be used to estimate errors that result from the Gaussian constraint.

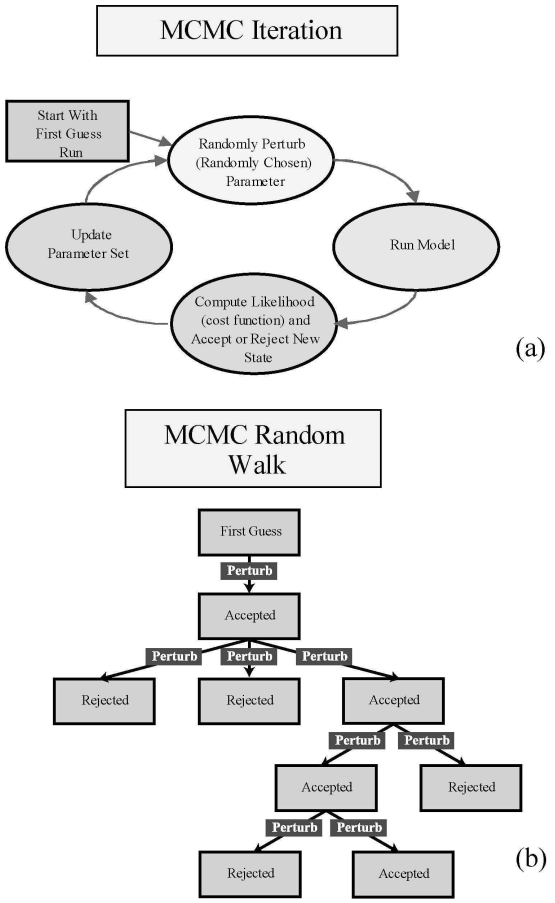

Figure 2: Schematic depiction of the MCMC iterative process. (a) A single MCMC iteration and (b) a portion of the MCMC random walk depicting the accept/reject sequence.

3. MCMC Sampling Procedure

In the retrieval process, what we desire is a search of the probability space for the most likely values of the retrieved state given a set of observations. To be able to evaluate the robustness of the retrieval, as well as the effective information content of the observations, the full joint conditional PDF for the set of retrieved state parameters is required. If we wish to retrieve only one or two parameters such that the state space is one or two dimensional, then it would be feasible to step through the range of possible values for each parameter in small increments, computing an error-weighted distance between state and observations and computing the PDF in the process. However, for state

dimensions greater than one or two, this sort of brute force sampling quickly becomes intractable. MCMC is based on the fact that most values of the state are associated with very low probability; the state space is largely empty (Tarantola 2005). In fact, as the dimensions of the estimation problem grow, the state space tends to become increasingly empty. MCMC exploits this fact, searching out areas of concentrated probability within the state space and sampling them via a guided random walk. In this paper, we use a form of MCMC similar to the well-known Metropolis-Hastings algorithm (Metropolis et al. 1953, Hastings 1970). For thorough and accessible discussions of the theoretical underpinnings of the MCMC algorithm, the reader is referred to Mosegaard and Tarantola (1995), Gelman et al. (2004) and Tarantola (2005).

The Markov chain is started from a set of state parameters drawn from a bounded Uniform PDF, with bounds set to physically realistic values for each parameter. In each MCMC search iteration a randomly-chosen state parameter is perturbed by an amount equal to a uniform random variable multiplied by the allowable parameter range and modified by a tuning factor determined during the adaptive burn-in (see discussion below).* If the newly-perturbed parameter value exceeds its allowable range, the new value is discarded and another random perturbation is applied. This process is repeated until an acceptable parameter value has been obtained, at which point new forward observations are computed and the distance between state and observations is calculated and compared to the previous value. For the current application, we employ a Gaussian (L_2) likelihood function defined identically to the Gaussian cost function in equation (2), except that MCMC needs no *a priori* term. For consistency, the observation error covariance is specified as in the optimal estimation technique according to instrument error characteristics and forward radiative transfer model errors. The new state is accepted and stored if the new likelihood exceeds the old, or if the new value is sufficiently close to the previous value as determined in a partially probabilistic manner (Fig. 2). This *probabilistic acceptance* is important in that it (1) encourages the algorithm to sample regions close to areas of relatively large probability, and (2) allows the algorithm to leave a local probability maximum. In the current application, a given state is probabilistically accepted if the ratio of the new likelihood to the old is greater than a random variable X between 0 and 1. Because we have computed the log-likelihood, the new state is accepted if

$$X < \exp(\Phi_{new} - \Phi_{old}) \qquad (3)$$

* The bounded uniform PDF is appropriate in the retrieval context as it represents the minimum possible prior information on the state, and assumes no other information aside from a reasonable range of values for each retrieved parameter.

where X is drawn from the *a priori* probability distribution. As is common practice, we use a burn-in period in which sampling is done, but no states are stored, for the purpose of "forgetting" the initial value in the chain. Similar to the method employed in Tamminem and Kyrola (2001) and later modified in Tamminem (2004), we allow the algorithm to adaptively change the step length during the burn-in period. Tuning the step length during burn-in leads to more efficient sampling of the state space once burn-in has finished. After burn-in, adaptive sampling is turned off to ensure sampling from the correct posterior distribution. It was found that a burn-in length of 2,000 samples was generally sufficient to converge to sampling a stationary PDF.

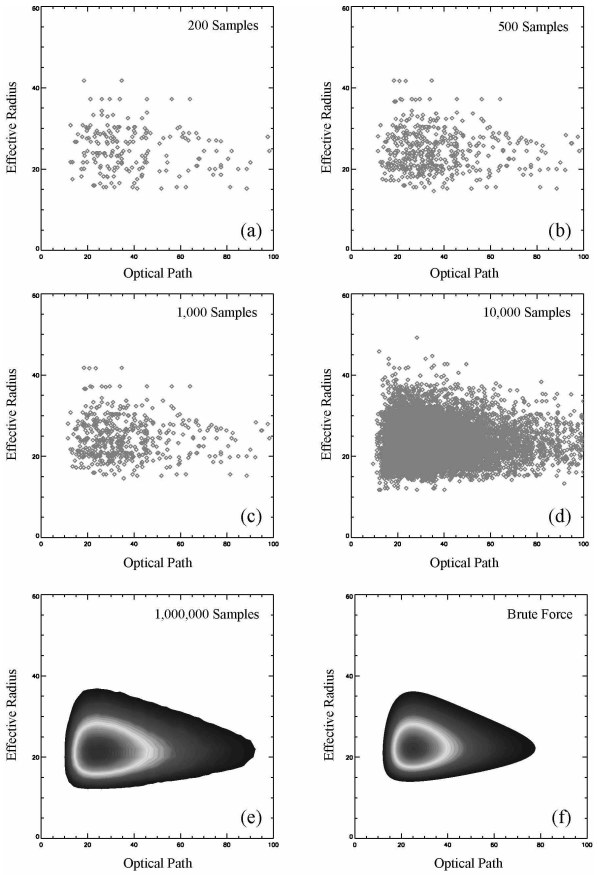

Figure 3: Example of convergence of MCMC to sampling a stationary distribution--in this case samples of the joint conditional probability of optical path and effective radius for pixel number 19,19. Progressively greater numbers of samples are shown in (a) through (e), while the result of brute force integration in increments of 0.1 in optical path and effective radius is shown in (f) for comparison.

The MCMC algorithm is run for many successive iterations to allow a thorough sampling of the state space, and a posterior probability density distribution is built for each state parameter from the set of accepted states. Based on initial tests, it was found that a sample size of 1,000 was sufficient to characterize the moments of the posterior PDF. To ensure full sampling of the PDF in each pixel, we chose to run the algorithm until 20,000 samples were accumulated for each pixel.

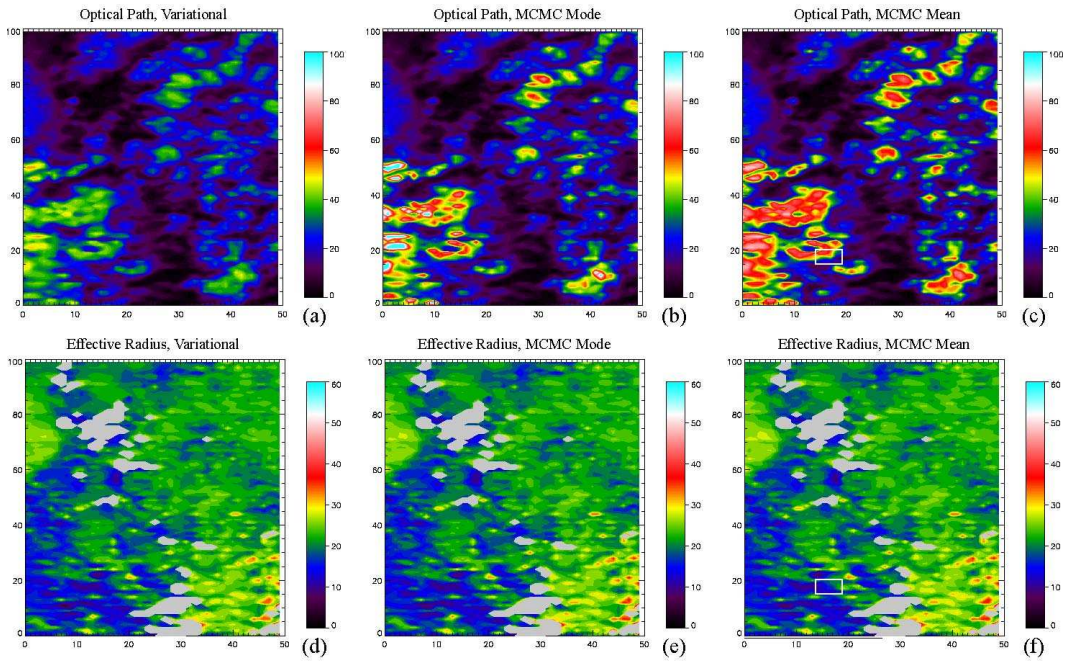

Figure 4: Optical path and effective radius for the control case from (a and b) variational retrieval, (c and d) mode of the MCMC PDF, and (e and f) mean of the MCMC PDF. Effective radius has been screened out (gray areas) where 0.64 micron reflectance is less than 0.25.

Depending on the error characteristics and the shape of the actual joint conditional PDF, it took between 35,000 and 500,000 iterations after the adaptive burn-in to obtain a 20,000 sample set.

An illustration of the convergence of the MCMC algorithm to sampling a stationary distribution is presented in figure 3, which depicts samples obtained from the joint PDF of optical path and effective radius for pixel (19,19) in the MODIS dataset. It is clear that by as early as 200 samples (Fig. 3a), the PDF is starting to take shape, and that by 10,000 samples (Fig. 3d) the PDF is well-characterized. For comparison, the PDF obtained from 1,000,000 samples (Fig. 3e) is contrasted with the result of brute force sampling across the range of values for optical path and effective radius (Fig. 3f). Brute force sampling was performed by stepping through the range of optical path and effective radius in increments of 0.1, running the forward model, and computing the likelihood function in each iteration. Though the optical path PDF has a slightly longer tail in the MCMC result, the PDFs are remarkably similar, demonstrating the effectiveness of MCMC in characterizing the true PDF.

4. Results

In this section, we present results from two different comparisons between a MCMC and a variational retrieval. Errors in each channel were specified as 5% of the observed reflectance. *A priori* optical path and effective radius for both cases were set equal to physically reasonable values for liquid stratus clouds; 20 and 10 microns, respectively (Miles et al. 2000). Error variances of 225 and 100 μm^2 were assigned to these *a priori* values for the purpose of spanning ranges of optical path and effective radius observed in nature; 0 to 100 and 0 to 60, respectively. It was assumed for simplicity that observation errors were uncorrelated between channels, as were *a priori* errors, and convergence to a solution for the scene of interest was uniformly obtained within 10 iterations. In both CTRL and 2XERR cases, we compare both retrieved values and PDFs, as well as the information content, and compare the variational retrieval with the mean and mode of the distribution obtained from MCMC. If the PDF is truly Gaussian, or at least centered and unimodal, the mean and mode retrieved values should be identical.

Retrieved optical path and effective radius are plotted for the variational retrieval in figures 4a and 4b, for the mode of the MCMC PDF in figures 4c and 4d, and for the mean of the MCMC PDF in 4e and 4f. Examination of the retrieved optical path shows immediately that, though all estimates recover the relatively thin cloud and clear areas well ($\tau < 20$), retrieval of larger values of optical depth varies significantly across the different retrievals. This is an early indication of the known loss of sensitivity in the visible channel to optical depths greater than approximately 50. Retrieved effective radius exhibits almost exact similarity across all retrievals, with the only significant differences occurring in regions of clear air, or where clouds are extremely thin.

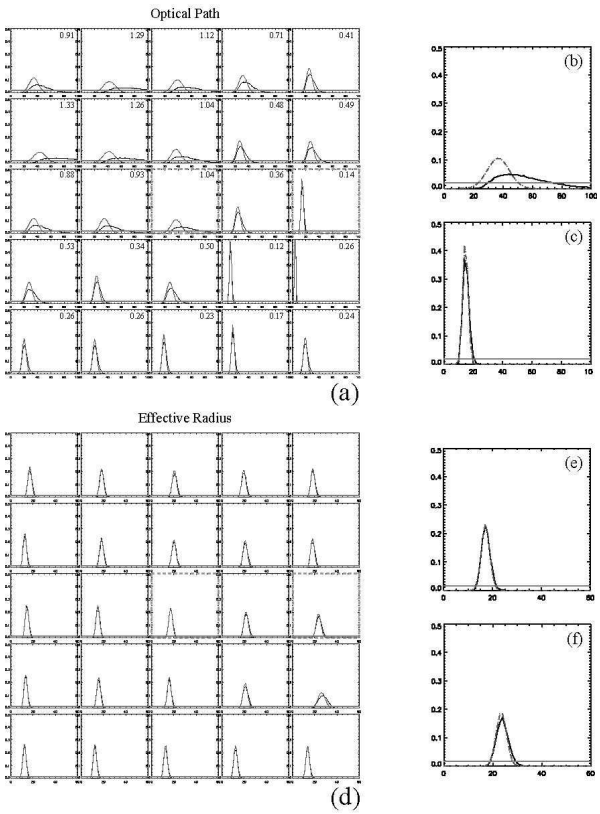

Figure 5: PDFs from MCMC retrieval of (a) optical path and (d) effective radius for the control case. Shown at right are PDFs for two selected pixels (b) and (e) relatively large optical path and small effective radius and (c) and (f) relatively small optical path and large effective radius. Uniform a priori is shown in light gray, posterior MCMC PDF is shown in black, and Gaussian variational PDF is in dashed gray. The two pixels shown in greater detail are outlined in thick dashed gray in the 5x5 matrix. The total integrated absolute difference between MCMC and Gaussian PDFs (see text for explanation) is printed in the upper right hand corner of each pixel in (a).

Retrieved optical path and effective radius PDFs from both MCMC and variational techniques are compared in figure 5 for a 5x5 pixel subregion outlined by the white box in figures 4c and 4f. In general, it can be seen that all MCMC PDFs are mono-modal, indicating that a quadratic minimization technique can be expected to perform well. In addition, the effective radius PDFs (Fig. 5d), as well as PDFs in pixels in which optical depth is less than ≈ 20, appear to be well-characterized by a Gaussian distribution. However, as optical paths increase beyond 20, the optical path PDFs increasingly depart from Gaussian. In fact, for optical paths greater than 50 (pixels (15,19), (16,19) and (16,20)), the MCMC PDF exhibits little departure from the *a priori* bounded uniform PDF, indicating that the observations have contributed very little information to the solution. For pixels with optical depths between 20 and 50, the MCMC PDFs appear to be log-normal in shape, with a mode that is increasingly displaced from the variational mode with increasing optical path. The fact that the PDFs for effective radius are not as skewed as those for optical path is consistent with the physical limitations imposed on the drop size distribution by the collision-coalescence process. As drop radii increase beyond 20 μm, the drizzle process becomes increasingly active, limiting the presence of large cloud droplets—particularly in liquid clouds over the ocean (Cooper et al. 2003).

To more clearly demonstrate the similarities and differences between variational and MCMC retrieved PDFs, τ and r_e from two selected pixels are shown enlarged at the right side of figure 5. Figures 5b and 5c depict optical path PDFs for pixels with relatively low (10-15) and relatively high (≈ 40) optical path respectively, while figures 5e and 5f depict PDFs of effective radius in these same pixels. In these figures it can be seen that the observations contribute enough information to effectively constrain the retrieval for low optical paths (Fig. 5c); the MCMC and variational solutions are nearly identical in this case. However, with increasing optical path (Fig. 5b), the solutions diverge, and the variational PDF provides a poor approximation to the solution. Though pixels with both relatively small (Fig. 5e) and large (Fig. 5f) effective radii appear to be well-characterized by a Gaussian PDF, careful examination of all four figures reveals that each of the MCMC PDFs has a tail--none are perfectly Gaussian. This tail is uniformly much smaller for the effective radius than for optical path, though it can be more clearly seen in the pixel with relatively large effective radius (Fig. 5f).

5. Quantitative Assessment and Improvement of Variational Retrieval

5a. Information Content and Assessment of Variational Retrieval

To be able to effectively retrieve certain state variables from observations, the observations must contain sufficient information on the desired state to be able to constrain the solution. It is, therefore, common to use so-called information content measures to quantify the amount of information contributed to the estimate of the state by an individual observation or sets of observations (Rodgers 2000, L'Ecuyer et al. 2005). One common measure of information is the Shannon Information Content, in which the information content of an observation or set of observations is computed as the reduction in entropy due to the addition of information from observations. Shannon information content is then defined as the difference between the entropy of the *a priori* state and the entropy of the retrieved state, which can be interpreted as the extent to which the number of allowable states is reduced by the addition of information from the measurements. For the variational case, both estimated and *a priori* PDFs are Gaussian, and the information content reduces to (Rodgers 2000, L'Ecuyer et al. 2005)

$$H = \frac{1}{2}\log_2\left|\mathbf{S_a}\mathbf{S_x}^{-1}\right| \qquad (4)$$

Since this expression is predicated on the assumption of Gaussian statistics, a quantitative assessment of the robustness of the Gaussian assumption can be obtained by comparing information content from the

variational retrieval with that computed explicitly from the MCMC PDFs.

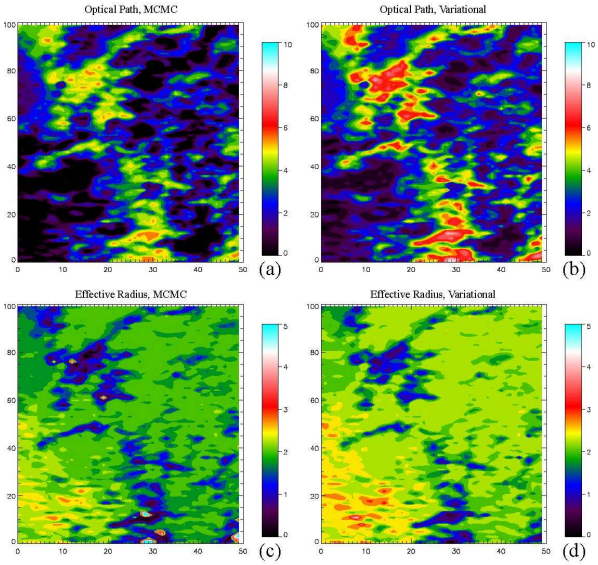

Figure 6: Comparisons of information content computed from MCMC PDFs (a and c) and variational retrievals (b and d) for optical path (a and b) and effective radius (c and d). All results are taken from the control case.

Comparison of information content between variational and MCMC retrievals (Fig. 6) indicates that the variational solution over-estimates the information content of the observations. Mean information content for optical path was 0.7 bits higher for the variational estimate, while mean information content for effective radius was 0.26 bits higher. Recall that the MCMC-retrieved PDFs of τ and r_e for the control case (Fig. 5) were decidedly log-normal for both τ and to a lesser extent r_e. This suggests that larger values of τ will have a higher probability of occurring in the MCMC retrieval. The absence of such a "tail" in the variational solution falsely leads one to believe that the observations do a better job of constraining the solution than they actually do. The fact that there is a greater discrepancy for optical path as compared to effective radius is due to the fact that MCMC-retrieved optical path PDFs exhibited a consistently greater departure from Gaussian.

A more quantitative measure of the differences between variational and MCMC solutions can be obtained by computing the integrated absolute difference (IAD) between the two retrieved PDFs. This measure is computed by calculating the absolute value of the difference between the variational and the MCMC PDF in each discrete bin, then summing the result. Since each PDF uses the same bin intervals, and since the PDFs are normalized so that they each integrate to one, this measure is independent of bin size. Since each PDF integrates to 1.0, a value of IAD of 1.0 indicates a 100% difference between the PDFs. For illustration, values of IAD for optical depth PDFs are printed in the upper right hand corner of each pixel in Figures 5a-5c. From this figure it can be seen that IAD values of 0.0 to 0.2 represent a negligible difference between variational and MCMC PDFs, while values above 0.7 represent relatively large departures.

5b. Use of MCMC Results to Modify and Improve the Variational Retrieval

The preceding section provides examples of diagnostics that can be used to quantify the quality of a Gaussian-based variational retrieval once the full PDF is known. The greater the integrated differences between PDFS (and consequent differences in information content), the more non-Gaussian the statistics are, and the poorer the retrieval will be. It is logical to ask whether this information can, in turn, be used to improve the variational retrieval, since it is still a desirable method by virtue of its computational speed and its basis in the physics of the radiative transfer problem.

In theory, visible reflectance should be sensitive to optical paths up to 50, however, the variational retrieval demonstrated loss of sensitivity at optical paths above 30 (Fig. 4). The fact that the MCMC algorithm returned log-normal PDFs for τ and r_e in both CTRL and 2XERR cases suggests that this problem could be partially mitigated by retrieving the natural log of optical path and effective radius. A log-normal PDF is thus imposed on τ and r_e, and because the natural log of a log-normally distributed random variable is distributed Gaussian, we can retain the Gaussian variational framework. The only fundamental change is that the Jacobian is computed with respect to variations in the natural log of τ and r_e, and the *a priori* error variance is specified such that the variance of the log-normal *a priori* distributions for τ and r_e is identical to the variance of the Gaussian *a priori* PDFs.

Because retrieval of the log of effective radius was essentially identical to the original retrieval effective radius (not surprising given the relatively small tails on the PDFs of r_e), we focus on the results of the retrieval of log of τ. Comparison of retrieved optical path from the original variational retrieval, and the variational retrieval of log-τ for the control case (Figs. 7a and 7d) demonstrates restored sensitivity to larger values of optical path for the log-τ retrieval. Retrieved optical path now much more closely resembles the mode of the MCMC retrieval (Fig. 4a), and the information content of the variational retrieval (Fig. 7e) now more closely approximates that of the MCMC estimate (Fig. 7b). In addition, integrated differences between variational and MCMC PDFs are reduced by approximately 40% in regions where optical path is greater than 30 (Figs. 7c and 7f). Though fairly large differences still exist in regions of relatively thick cloud, there are no differences larger than 0.8 left in the domain.

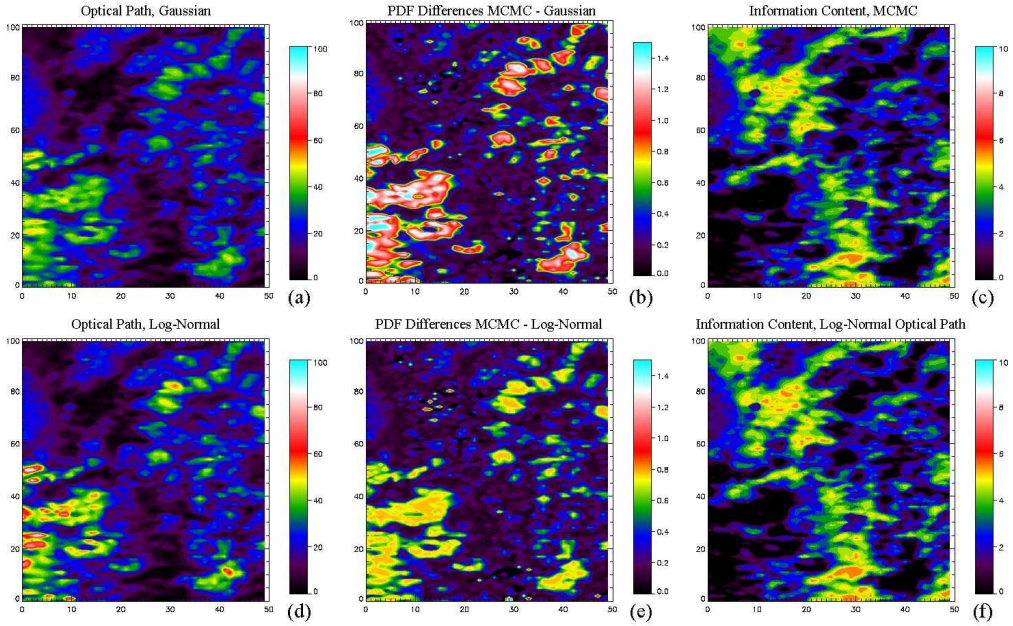

Figure 7: Plots of retrieved optical path from the control case (a) and the control case with log of optical path estimated (d). Differences between PDFs from variational retrieval and MCMC are plotted for (b) control case and (e) control case with log of optical path estimated. Optical path information content from MCMC and variational retrieval of log-tau are plotted in (e) and (f) respectively.

Improved correspondence between MCMC-retrieved τ and the log-τ retrieval is further illustrated in a scatter plot of variational optical path vs. that from the mode of the MCMC PDF (Fig. 8), in which it is clear that the variational estimate loses sensitivity at optical depths greater than 30 for the Gaussian case, while the log-τ retrieval retains a closer match to the MCMC result for larger optical paths. Note, however, that because the *a priori* optical path is still set at 20, larger values of optical path still tend to exhibit a low bias in the variational retrievals. This highlights the added benefit of the MCMC techniques; they do not rely on any assumed *a priori* information. Because the information from the visible channel saturates at values of $\tau > 50$, results are compared only for τ in the range of zero to fifty.

6. Conclusions

Given a set of observations and a model that relates observations to a set of desired state variables, Markov chain Monte Carlo methods can be used to generate the full joint conditional posterior PDF for the unknown state. In this paper we have used MCMC to obtain PDFs for optical path and effective radius retrieved from visible and near-infrared MODIS reflectances. The results have been used to evaluate a more traditional variational retrieval of τ and r_e that assumes Gaussian error characteristics. It was found that assumption of Gaussian errors leads to a loss of sensitivity in the retrieval to optical depths greater than 30. In addition, assumption of Gaussian errors leads to a consistent overestimate of the information content of the observations. Comparison of MCMC PDFs with those assumed in the variational retrieval revealed that the PDFs for both retrieved variables were log-normal in form, though the effective radius PDFs exhibited a much smaller tail compared with those from the optical path. When the variational retrieval is modified such that the natural logs of τ and r_e are retrieved, the retrieved values of effective radius change little, but sensitivity to large optical paths is restored and values of optical path up to 50 can be obtained. In addition, assumption of log-normal errors leads to better fit between the assumed PDFs in the variational retrieval and those obtained from MCMC, and information content is more consistent between the MCMC result and the variational.

The relatively straightforward application of the MCMC algorithm in this paper demonstrates the utility of MCMC for evaluating variational estimation methods. In principle, the technique can be extended to more complex estimation problems, especially ones in which we expect that the PDFs may be highly non-Gaussian. If computational difficulties can be overcome, MCMC also has obvious applications for evaluating the Gaussian assumptions used in atmospheric data assimilation systems. Because MCMC is already tractable for small-dimensional parameter spaces (e.g., < 20), initial experiments involving estimation of the PDFs of semi-empirical numerical model physics parameters should currently be possible. These PDFs would contain information on the sensitivity of the model evolution to specified constant parameters, and could provide guidance on ways to reduce model error. It should also be noted that MCMC could actually be used to *increase* the efficiency of some retrieval methods, particularly those that explicitly integrate

probabilities over a large database of solutions (e.g., Kummerow et al. 2001). In these retrievals, the difference between observations and state is computed for every solution in the database, leading to a constant and large number of numerical operations for each retrieved profile. It is possible that MCMC could be used to sample the set of solutions, returning a nearly identical solution for a much smaller number of iterations.

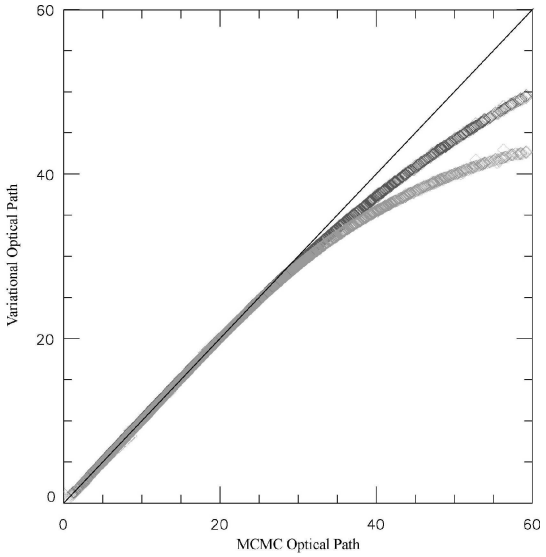

Figure 8: Scatterplot of optical path retrieved from Gaussian variational (light gray) and Log-Normal variational (dark gray) vs. optical path retrieved from the mode of the MCMC PDF. The one-to-one line is depicted in black for reference.

Acknowledgments

This work was funded by NASA NMP contract NAS1-00072. Mick Christi, Philip Gabriel, and Kyle Leesman helpfully fielded questions about the Radiant radiative transfer model, while Steve Cooper assisted in estimating the errors associated with neglecting absorption in the near-infrared. Tomi Vukicevic and Richard Davis provided guidance on the proper application of the MCMC algorithm.

References

Cooper, S. J., T. S. L'Ecuyer, and G. L. Stephens, 2003: The impact of explicit cloud boundary information on ice cloud microphysical property retrievals from infrared radiances. *J. Geophys. Res.*, **108**, DOI: 10.1029/2002JD002611

Devore, J., 1995: *Probability and Statistics for Engineering and the Sciences*, 4th Ed., Wadsworth, Belmont, CA.

Gabriel, P. M., M. Christi, and G. L. Stephens, 2005: Calculation of Jacobians for Inverse Radiative Transfer: An Efficient Hybrid Method. *J. Quant. Spect. and Rad. Trans.* In Press

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin 2004: *Bayesian Data Analysis*, 2nd Ed., Chapman and Hall/CRC, New York, NY.

Hastings, W., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrica*, **57**, 97-109.

Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G Petty, D. B. Shin, and T. T. Wilheit, 2001: The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.*, **40**, 1801-1820.

L'Ecuyer, T. S., P. M. Gabriel, K. Leesman, S. J. Cooper, and G. L. Stephens, 2005: Objective Assessment of the Information Content of Visible and Infrared Radiance Measurements for Cloud Microphysical Property Retrievals over the Global Oceans. Part I: Liquid Clouds. *J. Appl. Meteor.* In Press.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953: Equations of state on fast computing machines. *J. Chemical Physics*, **21**, 1087-1092.

Mosegaard, K., and A. Tarantola, 1995: Monte Carlo sampling of solutions to inverse problems. *J. Geophys. Res.*, **100**, 12,431-12,447.

Miller, S. D., G. L. Stephens, C. K. Drummond, A. K. Heidinger and P. T. Partain, 2000: A multisensor diagnostic satellite cloud property retrieval scheme. *J. Geophys. Res.*, **105**, 19955-19971.

Nakajima, T., and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos. Sci.*, **47**, 1878-1893.

Rodgers, C. D., 2000: *Inverse Methods for Atmospheric Sounding, Theory and Practice*, World Scientific, Singapore

Tamminem, J., and E. Kyrola, 2001: Bayesian solution for nonlinear and non-Gaussian inverse problems by Markov chain Monte Carlo method. *J. Geophys. Res.*, **106**, 14,377-14,390.

Tamminem, J., 2004: Validation of nonlinear inverse algorithms with Markov chain Monte Carlo method. *J. Geophys. Res.*, **109**, D19303, doi:10.1029/2004JD004927.

Tarantola, A. 2005: *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, PA.