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

TRISTAN S. L'ECUYER, PHILIP GABRIEL, KYLE LEESMAN, STEVEN J. COOPER, AND GRAEME L. STEPHENS

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 4 November 2004, in final form 9 June 2005)

ABSTRACT

The importance of accurately representing the role of clouds in climate change studies has become increasingly apparent in recent years, leading to a substantial increase in the number of satellite sensors and associated algorithms that are devoted to measuring the global distribution of cloud properties. The physics governing the radiative transfer through clouds is well understood, but the impact of uncertainties in algorithm assumptions and the true information content of the measurements in the inverse retrieval problem are generally not as clear, making it difficult to determine the best product to adopt for any particular application. This paper applies information theory to objectively analyze the problem of liquid cloud retrievals from an observing system modeled after the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument currently operating on the *Aqua* and *Terra* platforms. It is found that four diagnostics—the retrieval error covariance, the information content, the number of degrees of freedom for signal, and the effective rank of the problem—provide a rigorous test of an observing system. Based on these diagnostics, the combination of the 0.64- and 1.64-μm channels during the daytime and the 3.75- and 11.0-μm channels at night provides the most information for retrieving the properties of the wide variety of liquid clouds modeled. With an eye toward developing a coherent representation of the global distribution of cloud microphysical and radiative properties, these four channels may be integrated into a suitable multichannel inversion methodology such as the optimal estimation or Bayesian techniques to provide a common framework for cloud retrievals under varying conditions. The expected resolution of the observing system for such liquid cloud microphysical property retrievals over a wide variety of liquid cloud is also explored.

1. Introduction

Clouds play an important role in the regulation of the earth's climate. They influence both the amount of solar energy that reaches the earth's surface and the amount that is radiated back to space and, therefore, represent a critical factor governing global energy balance (Liou 1986). Furthermore, clouds play an important role in many chemical processes within the atmosphere acting as a surface for chemical reactions and providing a mechanism for the removal of aerosol particles through scavenging (Seinfeld and Pandis 1998). There is a body of evidence that suggests that human activity may affect climate by altering cloud microphysical properties as well as their vertical location and spatial distribution. Thus, quantitative global records of cloud microphysical properties are fundamental to answering many of the questions posed by the climate change community. It is not surprising, then, that worldwide inference of cloud microphysical properties continues to be a focus of a growing number of instruments, such as the Advanced Very High Resolution Radiometer (AVHRR) aboard the Geostationary Operational Environmental Satellite (GOES), the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Earth Observing System (EOS) *Aqua* and *Terra* platforms, the Polarization and Directionality of the Earth's Reflectances (POLDER) aboard the Polarization and Anisotropy of Reflectances for Atmospheric Sciences Coupled with Observations from a Lidar (PARASOL) satellite, and the Cloud Profiling Radar (CPR) and Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) aboard the soon-to-be-

Corresponding author address: Dr. Tristan L'Ecuyer, Department of Atmospheric Science, Colorado State University, Foothills Campus, Fort Collins, CO 80523-1371.
E-mail: tristan@atmos.colostate.edu

launched CloudSat and Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellites.

Given the high degree of measurement accuracy afforded by such instruments, the reliability of derived atmospheric products no longer depends as heavily on instrument calibration and noise but more so on the choice of spectral bands, the forward model, and the method of inversion. Consider, for example, optical depth and effective radius retrievals. The remote sensing literature is replete with descriptions of algorithms for inferring these cloud properties from distinct combinations of radiances at multiple wavelengths of visible and infrared radiances, some of which are summarized in Table 1 of Miller et al. (2000). The physical basis of each of these algorithms rests on the fact that water droplets, ice crystals, and the gaseous constituents of the atmosphere display different spectral signatures. Hence, microphysical and optical properties of single-layer clouds can be inferred in principle from radiances at two or more wavelengths. While many of these algorithms have successfully been applied to map clouds, few are universal in the sense that they can be applied to any scene at any time of the day independent of the background or surface. On the contrary, many can only be applied under specific conditions (e.g., during the daytime) or over a limited dynamic range (e.g., optically thin clouds) leading to unphysical discontinuities when one seeks to compile a complete database of the global distribution of clouds. Moreover, in the past, channels used in such algorithms have been selected empirically, based on prior research showing sensitivity to the properties of interest and constrained by available wavelengths fixed by satellite hardware. The use of distinct combinations of wavelengths can lead however to discrepancies between the products of different algorithms when they are applied to the same scene by virtue of subtle differences in the information provided by the measurements. Unfortunately, such differences between algorithm products are often difficult to resolve because of limited quantitative measures of the uncertainties in each.

In the case of the MODIS cloud product, the reflectance map technique of Nakajima and King (1990) is employed for daytime retrievals using an absorbing channel (e.g., 2.142 μm) and a nonabsorbing channel (e.g., 0.664 μm). In essence, the retrieval involves the solution of two nonlinear equations for two unknowns. Nighttime retrievals must resort to alternate techniques rooted in thermal emission. In this case, the CO_2 slicing method is often used to first determine the cloud-top temperature. Knowledge of the cloud temperature is then used to constrain optical depth and effective radius retrievals for thin cirrus using radiances at 3.7 and 11 μm, following a technique analogous to that introduced by Inoue (1985) and Prabhakara et al. (1988). Comparisons of these techniques under daytime conditions, however, often indicate differences of a factor of 2 or more in retrieved effective radius and optical depth. In fact, changing the particular channels used within either of these techniques can also significantly impact the results because of subtle differences in the sensitivity of the measurements to the retrieval parameters and assumptions that are required in associated radiative transfer model (RTM) calculations. Assessing the relative performance of these techniques (and others in the literature) requires a methodology that explicitly accounts for relative differences in both the sensitivity of each measured radiance to the retrieval parameters and their uncertainties including components owing to both measurement and modeling errors.

The approaches commonly adopted for assessing the information content of a system are well suited for this purpose. Interestingly though, while a casual perusal of the retrieval literature attests the emphasis placed on techniques of inverting remotely observed data, very little reference is made to the information content of the measurements themselves. For example, a sampling of the different inversion methods include the following: constrained nonlinear least squares minimization (Worden et al. 1999), neural networks (Juliette and Clerbaux 1999), principal component analyses (Tanre et al. 1996), optimal estimation using Bayesian methods (e.g., Rodgers 1976; Evans 2002), Bayesian Monte Carlo methods for non-Gaussian inverse problems (Tamminen and Kryola 2001), split-window techniques (e.g., Prabhakara et al. 1988; Suggs et al. 1998), bidirectional mapping techniques (Nakajima and King 1990; Rolland et al. 2000), regularization methods (Eriksson 2000), and discrepancy principles that extend regularization methods (Li and Huang 1999). Of the aforementioned works, however, only those of Worden et al. (1999) and Evans (2002) refer to and use theoretic information methods. The former refers to how much information is gained in a retrieval relative to the prior covariance matrix, while the latter refers to the amount of information provided by additional submillimeter microwave channels. Neither, however, attempts to relate information content to the retrieval covariance matrix.

It is, therefore, of interest to revisit the problem of cloud microphysical property retrievals from the perspective offered by information theory to determine a single optimal framework that can be applied to different satellites under as wide a variety of conditions as possible. Toward this end, this paper establishes a rig-

orous, objective methodology for determining the information content of a set of observations, selecting an optimal channel configuration, and assessing the effects of both model and instrumental uncertainties on the final inferred product. The analysis is rooted in information theoretical concepts elucidated by Shannon and Weaver (1949) and on the application of their technique to atmospheric science by Rodgers (2000). Channel selection is made objective by quantifying the amount of information contained in the spectral measurements and calculating their effective signal-to-noise ratio (SNR) in relation the desired set of retrieval parameters. This is accomplished practically through analysis of the retrieval covariance matrix, which holds the key for understanding and quantifying differences between different retrieval procedures and observational data.

To illustrate the benefits of adopting such an approach, the method is applied to the problem of retrieving cloud microphysical properties from satellite radiance observations at solar and thermal wavelengths using the MODIS channels as a baseline. For simplicity, it will be assumed that single-layer liquid and ice clouds can be discriminated from one another and multilayer cloud complexes through a combination of their radiometric signatures [e.g., through the trispectral technique explored in a series of papers by Ackerman et al. (1990), Strabala et al. (1994), and Baum et al. (2000)] and active sensors, such as the CPR or CALIOP. Under this assumption, we initially focus on the relatively straightforward problem of retrieving the parameters of a gamma distribution of water droplets in single-layer liquid clouds to facilitate illustration of the methodology and interpretation of the results. Furthermore, for this preliminary application of the technique, we focus on oceanic scenes because they represent the largest fraction of pixels a satellite will encounter. The more challenging problem of ice cloud retrievals, which is complicated by the required assumption of ice crystal habit, is analyzed in a similar manner in a companion paper by Cooper et al. (2006, hereinafter Part II). The analysis can be readily extended to other surfaces and multilayer cloud complexes but these problems are beyond the intended scope of the current study and are, therefore, left as future topics of investigation.

The channels considered in the analysis, their noise requirements, and their primary uses (adapted from the MODIS Web site, available online at http://modis.gsfc.nasa.gov/about/specifications.php) are summarized in Table 1. Note that the four channels centered on the 15-μm CO_2 band that are primarily used for determining cloud-top pressure are not analyzed for the low clouds studied here, but have been included in the analysis for ice clouds in Part II. A state-of-the-art radiative transfer model is employed to simulate radiances at these wavelengths for cloud scenes spanning a range of cloud heights, liquid water paths, effective radii, surface albedos, atmospheric pressure and temperature profiles, and solar zenith angles. The resulting simulated measurements are then fed into a series of sensitivity studies that provide the sensitivities of each channel to the retrieval parameter and rigorous estimates of the uncertainties in each because of potential errors in the assumptions required to model them. The combination of sensitivities and uncertainties completely determines the information content of the ensemble of measurements that is subsequently used to objectively determine the combination of wavelengths that provide the greatest amount of information for global microphysical property retrievals.

TABLE 1. The MODIS channels of relevance to this study and their primary uses. Solar reflectance channels (upper set) are given in nanometers; thermal emission channels (lower set) are in micrometers.

Band	Wavelength	SNR	Primary use
1	620–670	128	Land/cloud/aerosol
2	841–876	201	Boundaries
6	1628–1652	275	Land/cloud/aerosol
7	2105–2155	110	Properties
26	1360–1390	150	Cirrus clouds/water vapor
Band	**Wavelength**	**NEΔT**	**Primary use**
20	3.66–3.84	0.05	Surface/cloud
23	4.02–4.08	0.07	Temperature
27	6.535–6.895	0.05	Cirrus clouds/water vapor
29	8.4–8.7	0.05	Cloud properties
31	10.78–11.28	0.05	Surface/cloud
32	11.77–12.27	0.05	Temperature

2. Sensitivity studies

As a precursor to the more rigorous information content study that follows, it is useful to examine the sensitivity of the observations to the parameters of interest. In addition to illustrating the dominant physical processes governing the radiative transfer through clouds and providing insight into the mechanics of the retrieval problem, such analyses also form the input to the information content study itself. The required radiances are computed using an RTM called Radiant (Christi and Stephens 2002; Christi and Gabriel 2004; Gabriel et al. 2005), which is multistream, plane parallel, and accounts for multiple scattering. Atmospheric absorption is modeled using the correlated-k distributions developed for MODIS wave bands by Kratz (1995). This approach captures gaseous absorption properties to an

accuracy of 1%, while significantly reducing the computational costs incurred by explicit line-by-line calculations. The ocean surface is modeled as a Lambertian reflector with a visible albedo of 0.1 (at wavelengths less than 3 μm), consistent with the Earth Radiation Budget Experiment (ERBE) observations of Harrison et al. (1990) and an infrared emissivity of 0.99. Temperature, moisture, and gas-mixing ratios are assigned based on the McClatchey et al. (1972) tropical atmosphere while the scattering properties of cloud particles are modeled using a Mie scattering code assuming spherical droplets.

Using Radiant, radiances for the 11 channels in Table 1 were simulated for a wide variety of liquid clouds. The cloud droplets are assumed to follow a lognormal distribution

$$N(R) = \frac{N_0}{R\sqrt{2\pi}\sigma_{\log}} \exp\left\{-\frac{1}{2}\left[\frac{\ln(R/R_g)}{\sigma_{\log}}\right]^2\right\}, \quad (1)$$

where R_g is the modal radius, N_0 is the number density, and $\sigma_{\log}$ is the natural logarithm of the geometric standard deviation σ_g. The effective radius, related to the modal radius via $R_e = R_g \exp[(5/2)\sigma_{\log}^2]$, is varied between 5 and 14 μm in 1-μm increments. Here, $\sigma_{\log} = \sqrt{\overline{(\ln R - \ln R_g)^2}}$ is assumed fixed at 0.427 (Deirmendjian 1969), while number density N_0 was scaled to provide five different values of liquid water path (LWP) corresponding to visible optical depths of 5, 15, 30, 40, and 50 at an effective radius $R_e = 8$, resulting in a total of 50 test cases. The optical properties of the resulting clouds were modeled using Mie theory and they were placed between 1 and 2 km in the McClatchey midlatitude summer (MLS) atmosphere (McClatchey et al. 1972). Unless otherwise stated, all daytime calculations assume a solar zenith angle of zero, corresponding to an overhead sun. Visible optical depths corresponding to all of these cases are presented in Fig. 1 for reference.

This base set of cases is supplemented with two sets of sensitivity studies—one to establish the behavior of each channel in response to changes in the retrieval parameters themselves, and the second for use in estimating modeling uncertainties resulting from errors in those parameters are not retrieved but must be specified in order to perform the necessary radiative transfer calculations.

a. Forward model uncertainties

Regardless of the measure of information content one adopts, the result is necessarily rooted in the signal-to-noise characteristics of the retrieval system. An observation whose sensitivity to a retrieval parameter is less than the accuracy to which it can be measured cannot provide useful information. Thus, it is important to establish a rough estimate of the uncertainties in each channel that arise from a combination of random measurement and calibration errors as well as any assumptions needed to model the atmospheric radiative transfer that are not going to be explicitly retrieved. Measurement errors are modeled after the specifications for the MODIS instrument aboard *Aqua* that have been summarized in Table 1. In addition to these uncertainties, all assumptions that are required to perform radiative transfer calculations in the visible and infrared must be considered as potential sources of uncertainty in the forward model. These include the shape of the drop size distribution (DSD), cloud height and geometric thickness, surface albedo, assumed humidity and temperature profiles, the presence of aerosols, the use of plane-parallel calculations to model a cloud that is inherently inhomogeneous in the vertical and horizontal directions, the representativeness of a satellite snapshot of clouds for measuring their global distribution (i.e., sampling errors), and so on.

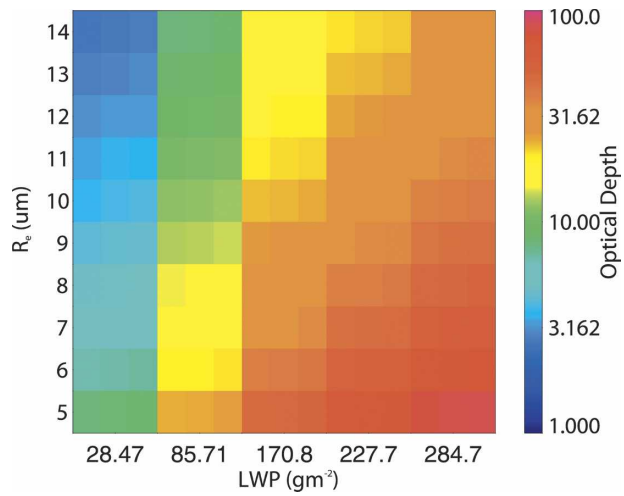

FIG. 1. Visible (0.64 μm) optical depths as a function of effective radius (vertical axis) and liquid water path (horizontal axis) for the 50 base cases modeled.

While all of these error sources are important, it is a monumental task to evaluate the contributions from all of them. In fact, a suitable methodology for assessing the global mean of 3D effects has not yet been developed and, because the analysis focuses on a MODIS-like instrument but purposely avoids defining a particular observing system, sampling errors cannot be defined. Furthermore, some of these assumptions can be imposed as soft constraints on the retrieval by allowing them to be retrieved as opposed to assigning them in advance. Allowing cloud-top height and surface albedo to vary, for example, provides additional degrees of freedom that allow the algorithm to better match the

observations. As a result, we focus only on those sources that can directly be modeled using the Radiant model and cannot realistically be retrieved from the observations, namely, uncertainties in the assumed DSD, specific humidity profile, and temperature profile. To this end, the 50 base cases described above were rerun with each of these quantities perturbed by an amount representative of their expected uncertainty.

To model errors in assumed DSD, the cloud droplet assumption is changed to follow a modified-gamma distribution,

$$N(R) = \frac{N_0}{R_e \Gamma(\nu)} \left(\frac{R}{R_e}\right)^{\nu-1} e^{-R/R_e}, \tag{2}$$

with the number density N_0 scaled to match the LWP values assumed in the base cases and the effective radius defined as noted above $R_e = R_g \exp[(5/2)\sigma_{\log}^2]$ to be consistent with the R_g values assumed in the lognormal distribution (Stephens 1994). The width parameter ν is held fixed at 3, following Deirmendjian (1969). In an operational retrieval, profiles of temperature and humidity need to be specified. A likely source of such data are numerical weather prediction (NWP) models such as that used at the European Centre for Medium-Range Weather Forecasts (ECMWF), so it is assumed that the uncertainties in ECMWF temperature and humidity predictions are representative of the level of error in the values assumed in the algorithm. Based on the sensitivity studies of Eyre (1990) and Eyre et al. (1993), then, temperatures were perturbed by 2 K at each layer and specific humidities were perturbed by 15% below 500 hPa and 30% above.

The resulting estimates of the uncertainties from each of these sources are then combined with one another and an estimate of the measurement errors to determine effective fractional errors in each channel are modeled. Once again, while no data from any particular instrument are used in the analysis, an appropriate estimate of instrument noise from the SNR and noise-equivalent temperature difference (NEΔT) requirements for MODIS measurements (summarized in Table 1) are used to represent instrument performance. Then, if it is assumed that instrument noise and each of the sources of model error are uncorrelated, the combined uncertainty resulting from all of these sources is given by the square root of the sum of the squares of each of these estimates

$$\varepsilon_i = \sqrt{\left(\frac{\delta_q}{y_i}\right)^2 + \left(\frac{\delta_T}{y_i}\right)^2 + \left(\frac{\delta_{\mathrm{DSD}}}{y_i}\right)^2 + \left(\frac{\delta_m}{y_i}\right)^2}, \tag{3}$$

where ε_i is the (dimensionless) fractional error in the ith channel (Taylor and Mohr 2004). The radiance in channel i is represented by y_i, while δ_q, δ_{DSD}, δ_T, and δ_m represent radiance uncertainties resulting from humidity, DSD, temperature, and measurement errors, respectively.

Fractional errors in six of the channels considered are presented in Fig. 2 for the range of liquid cloud R_e and LWP considered. In general, the uncertainties in the shortwave radiances are much smaller than those in the infrared because of the fact that the latter suffer from errors in the assumed temperature and humidity profiles while the former are only sensitive to errors in assumed DSD, which are typically small provided that the effective radius and liquid water path are held fixed. As a result, fractional errors at 3.75 and 11 μm range from 5% to 10% while those in the shortwave channels are ~1%–2% for all but the thinnest clouds. A subtle yet important result of this analysis is that the errors at 2.13 μm tend to be approximately 50% larger than at 1.64 μm because of differences in the strength of the water vapor absorption at the two wavelengths, which causes the 2.13-μm channel to be more sensitive to errors in the assumed humidity profile. In the analysis that follows we will see that this can have implications for determining the optimal channels for use in a retrieval algorithm. The uncertainties in 1.38-μm radiances are at least an order of magnitude larger than those in all other channels. This is a direct consequence of the strong water vapor absorption at this wavelength combined with our inability to constrain the specific humidity profile to any better than 15% using NWP model data. Last, note that the fractional errors of some channels exhibit significant scene dependence while those in others do not, suggesting that the information content and optimal channel configurations determined below are likely to depend on the region-in-state space in which the solution lies. Fractional errors at 1.38 μm, for example, increase with increasing optical depth because of enhanced cloud reflection that causes a greater fraction of the incident radiation to pass through the uncertain water vapor profile in both the downwelling and upwelling directions. Errors at 11 μm, on the other hand, are dominated by uncertainties in cloud-top temperature, which are not sensitive to the properties of the underlying cloud.

b. Sensitivities to retrieval parameters

To assess the "signal" component of the signal-to-noise ratios that drives the information content study, a second series of sensitivity studies was performed in which each of the parameters of interest was perturbed independently with sensitivities computed via

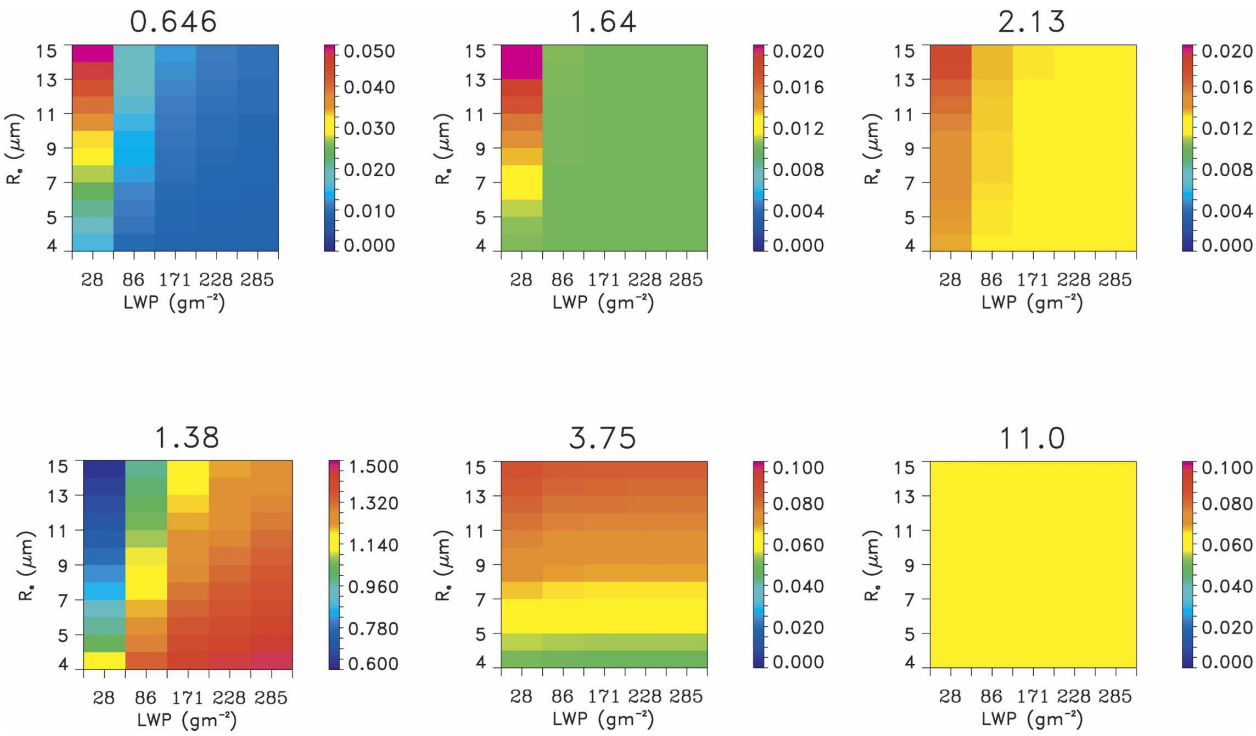

FIG. 2. Fractional errors in selected channels resulting from the combination of uncertainties in prescribed DSD, specific humidity and temperature profiles, and instrument noise.

$$S_i = \frac{\Delta I_i}{\Delta X}, \tag{4}$$

where X represents the three key retrieval parameters: R_e, LWP, and cloud-top height (C_{top}) and the ΔX are chosen to be ~5% of the value of the parameter X. The exception is C_{top}, which is perturbed by 1 km, consistent with the vertical resolution of the Radiant radiative transfer model. The resulting sensitivities can then be divided by the radiance errors in each channel to define an effective SNR for each retrieval parameter at each wavelength.

Figure 3 combines sensitivities to R_e with the uncertainty estimates from Fig. 2 to produce maps of effective SNR for effective radius retrievals defined by SNR = $S_i/\varepsilon_I I_i$. The largest effective SNR for R_e occur at 1.64 μm where they are about a factor of 2 larger than those at 2.13 μm. As indicated above, this is largely because of the fact that fractional errors at 1.64 μm are smaller than those at 2.13 μm. Even so, these two channels clearly contain the majority of the particle size information in the system. The small sensitivities at 0.64 μm owe their existence to the fact that changing R_e causes an inversely proportional change in optical depth when LWP is held fixed. Last, although they are a factor of 20 less than those at 1.64 μm, the nonnegligible effective SNRs at 3.75 μm play an important role in nighttime retrievals in the absence of a signal in the shortwave channels.

Similar results corresponding to effective SNR for LWP retrievals are presented in Fig. 4. Clearly, the visible channel exhibits the largest effective SNRs for all clouds except the optically thinnest cases (lowest LWP and largest R_e) where the largest sensitivities occur at 1.64 μm. Considering the lack of SNR in any of the infrared channels, no LWP information is expected for nighttime retrievals.

Because it is desirable to introduce as few assumptions as possible in the retrieval, it is of interest to include cloud-top height C_{top} as a retrieval parameter rather than fixing it a priori. To examine the feasibility of using the visible and infrared radiances to constrain cloud-top height, Fig. 5 summarizes the effective SNR for C_{top} retrievals. Changing cloud height affects the observed radiances through the following two mechanisms: 1) it changes the cloud-top temperature influencing emission at thermal wavelengths, and 2) it modifies the amount of water vapor above the cloud influencing the amount of radiation that gets absorbed and emitted at wavelengths corresponding to water vapor absorption bands. The first mechanism is, for example,

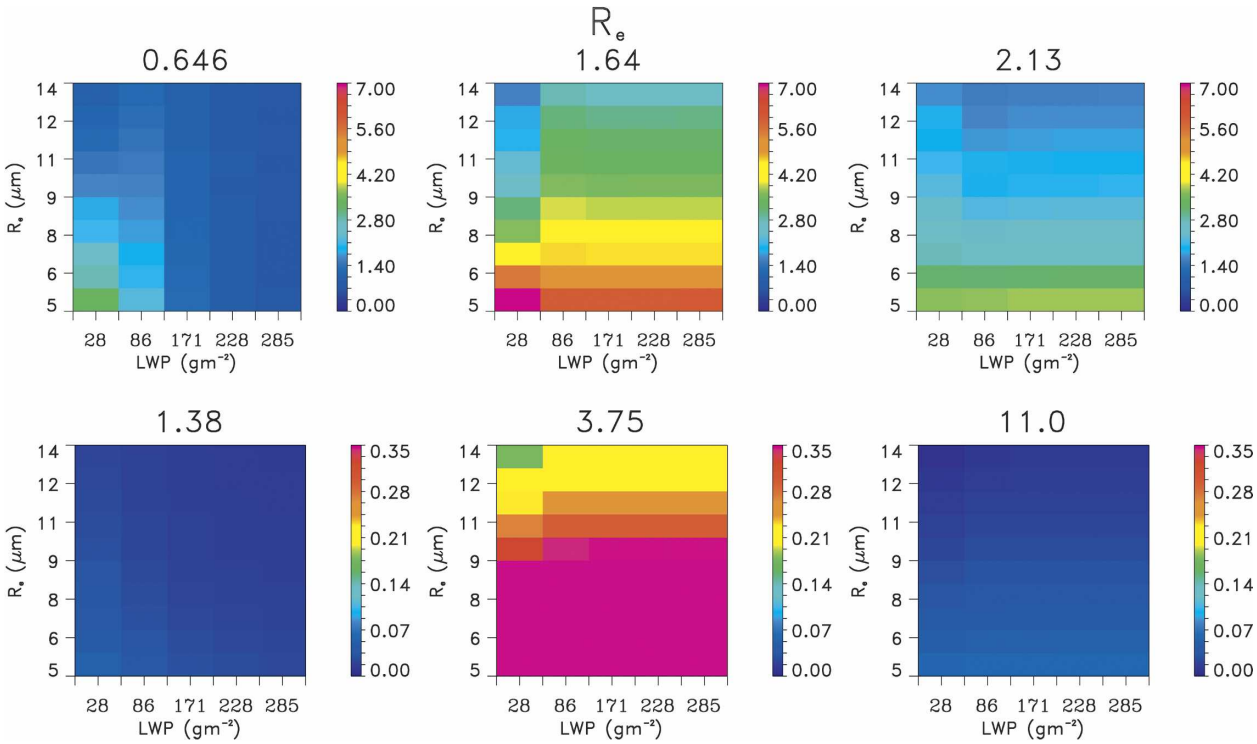

FIG. 3. Signal-to-noise ratio at selected wavelengths for effective radius perturbations. Note that the range of values on the upper panels is 20 times that on the lower panels.

responsible for the observed SNRs of ~1.2 at 11.0 μm. The latter mechanism is particularly evident at 1.38 μm, and to a lesser extent 1.64 and 2.13 μm. The extremely strong sensitivity at 1.38 μm is counterintuitive for clouds with tops as low as 2 km because of the strong water vapor absorption at this wavelength. This result is an artifact of the fact that, on paper, arbitrarily small radiances can be analyzed when in practice the absolute radiance reflected to the satellite by the cloud is so small (~0.0006 W m^{-2} sr^{-1}) that it is not likely to be detectable over the electrical noise in the sensor. In the event that a sensor could be constructed to detect radiances down to this level, however, a small increase in cloud-top height dramatically increases the amount of radiation reflected back to the satellite because the water vapor is so prominent at these altitudes. In the McClatchey MLS atmosphere assumed here, for example, changing the cloud top from 2 to 3 km gives rise to an order of magnitude increase in the modeled radiance at the top of the atmosphere. This, in turn, leads to an SNR of ~10 based on the estimated fractional uncertainties that are on the order of unity. For now, small radiances at 1.38 μm will be included in the analysis with the caveat that results involving this channel are contingent on the ability of the instrument to measure radiances as low as 0.0005 W m^{-2} sr^{-1}. In addition, it must be possible to model radiances with absolute magnitudes this small, avoiding numerical issues such as instabilities, discretization errors, and truncation errors. Regardless of these issues, the strong influence of water vapor in the 1.38-μm channel sufficiently decouples it from the other channels such that failure to model the electrical noise floor of the instrument should not impact the analysis of the remaining channels.

All three of these figures further emphasize the fact that the information content of the observing system is likely to be strongly dependent on the scene being retrieved. The sections that follow outline a framework that takes this into account and attempts to determine the subset of channels that provide the most information for the widest variety of cloud systems.

3. Information theory

The sensitivity studies described above provide insight into the response of individual measurements to the retrieval parameters, but to determine the combination of channels best suited for the retrieval, it is necessary to define a set of criteria for assessing the response of the instrument as a whole. These criteria must not only account for the sensitivities of each indi-

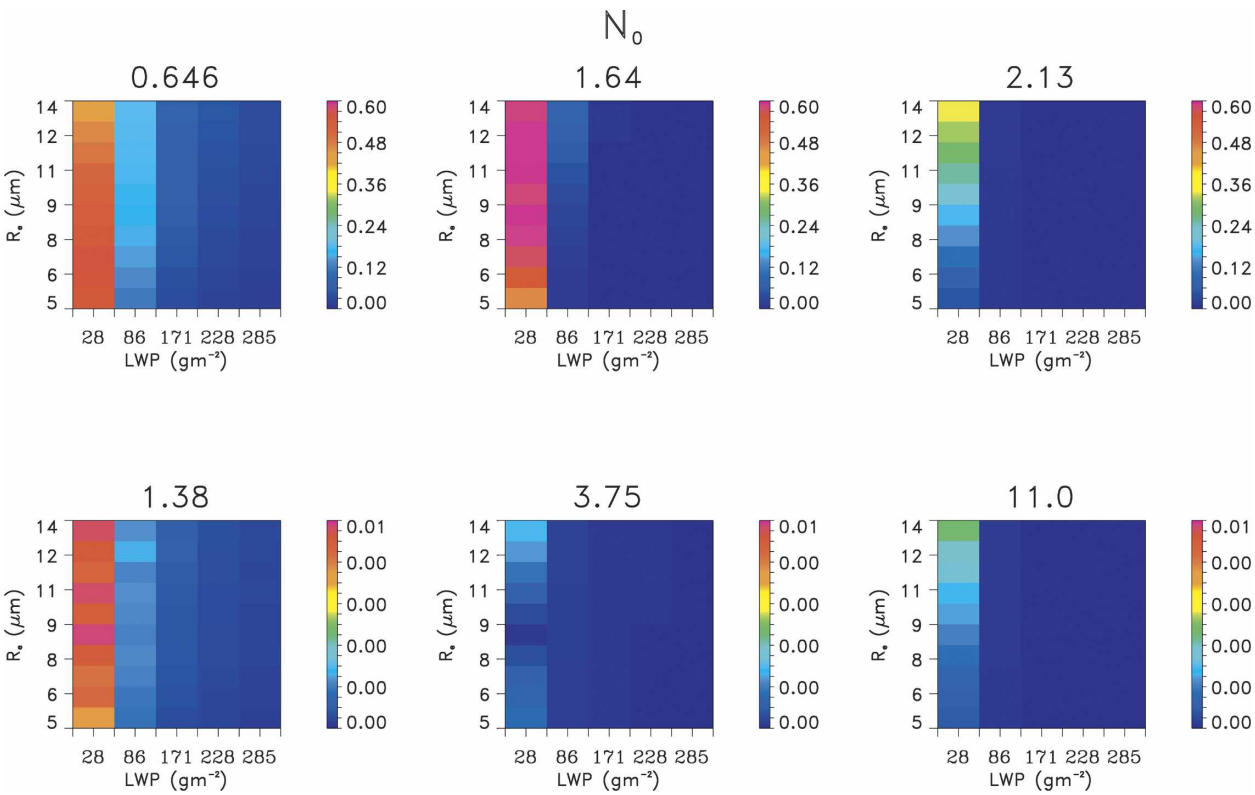

FIG. 4. As in Fig. 3, but for LWP perturbations. In this case the scale on the upper panels is 60 times that on the lower panels.

vidual channel but also correlations between them. In addition, the expected uncertainty in each measurement must be considered to accurately characterize the amount of independent information provided by the ensemble of channels relative to the level of noise inherent in the observing system. The metric adopted in this study is the information content, a tool that has been widely used in many engineering disciplines but has, to date, been underutilized in atmospheric remote sensing. In general, information content refers to the degree by which a set of observations improves our knowledge of the set of retrieval parameters (cloud height, particle size, number concentration, etc.). Put another way, one can think of information content as the factor by which the total number of distinct combinations of the retrieval parameters or "states" that satisfy our prior knowledge of the system, or the degree of nonuniqueness, is reduced by making the measurements. Because our ability to distinguish states in the retrieval system depends both on the sensitivity of the measurements to the retrieval parameters and the accuracy of those observations, the information content inherently accounts for these factors and is, therefore, well suited to analyzing the properties of an observing system.

a. Shannon information content

There is an abundance of different measures of information content (see, e.g., Kullback 1968 or Bernardo and Smith 1994, and references therein), many of which may be adapted to the current problem. We adopt the definition of Shannon and Weaver [(1949), hereinafter referred to as the Shannon information content (SIC)] that is described in detail by Rodgers (2000) in relation to the problem of atmospheric sounding from multispectral satellite radiance measurements and has recently been applied to the problem of CO_2 retrievals from infrared sounding observations by Engelen and Stephens (2004). As a measure of knowledge of the retrieval system, Shannon and Weaver (1949) adopt an analog to the thermodynamic entropy S, defined as the logarithm of the number of distinct internal states of a macroscopic system. Letting P_1 represent the PDF governing the probability of obtaining any state in the retrieval system prior to making a measurement and P_2 be the PDF after the measurement has been made, the Shannon information content is defined as the difference in entropy of these two PDFs, namely,

$$H = S(P_1) - S(P_2). \tag{5}$$

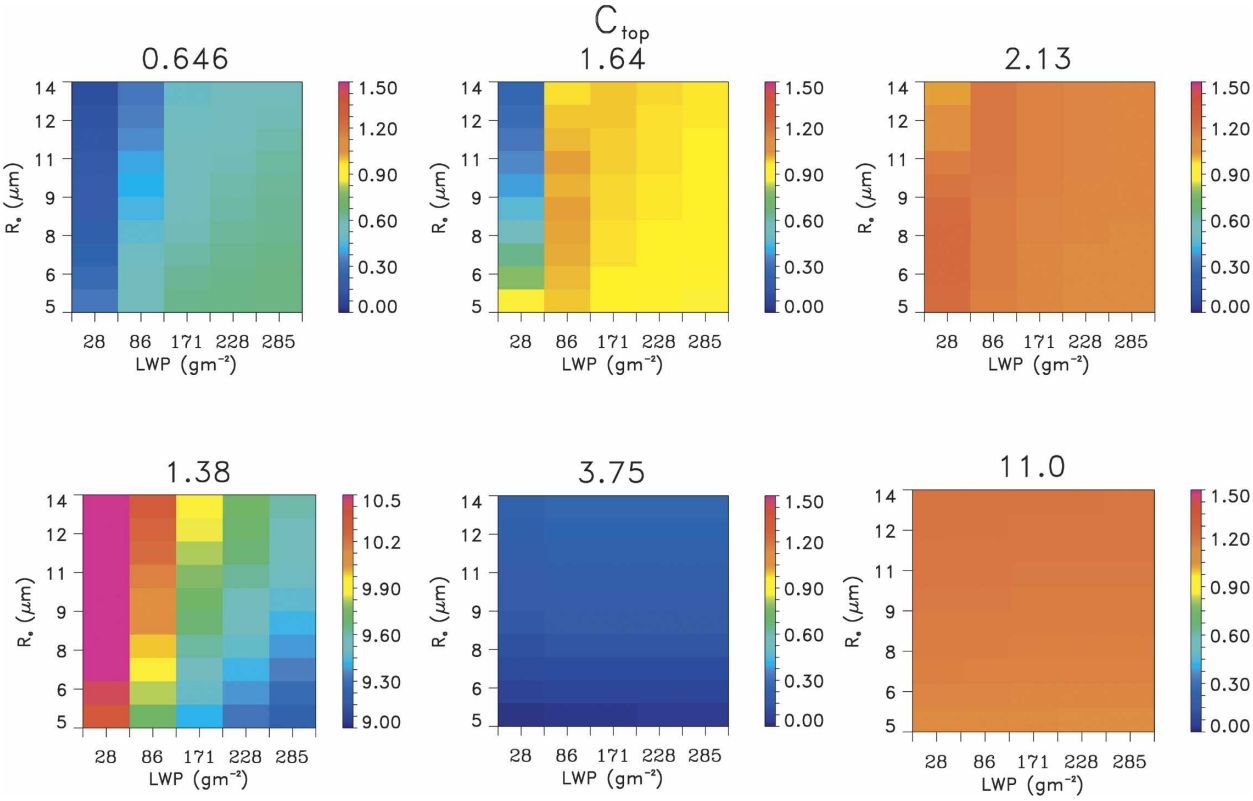

FIG. 5. As in Fig. 3, but for cloud-top height perturbations. The scale for the 1.38-μm image ranges from 0 to 10.5; those at other wavelengths range from 0 to 1.5.

The SIC thus defined measures the degree to which the addition of a measurement reduces the disorder of the retrieval state space. For convenience, we assume that both the prior and posterior PDFs follow Gaussian distributions with covariances $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively.[1] It can be demonstrated (e.g., Rodgers 2000) that the entropy of a multivariate Gaussian distribution of m variables is given by

$$S(P) = \frac{1}{2}\log_2|\mathbf{S}| + c, \tag{6}$$

where the constant $c = (m/2)\log_2(2\pi e)$. The SIC is then

$$H = \frac{1}{2}\log_2|\mathbf{S}_1\mathbf{S}_2^{-1}|. \tag{7}$$

[1] Aside from the computational benefits to assuming Gaussian distributions in the analysis, it can be shown that the Gaussian distribution maximizes the entropy or, equivalently, minimizes our assumed knowledge of the state space when only the mean and variance of the distribution of retrieval states is known (Rodgers 2000). Thus, the Gaussian distribution is, in fact, the most appropriate choice in the absence of conclusive evidence for an alternative form of PDF.

Because the covariances represent the volume-in-state space corresponding to our uncertainty in the retrieval state prior to and after making the measurement, the information content is a measure of the factor by which the measurement reduces our uncertainty in the retrieval state.

With entropy defined in this way, that is, as a logarithm to the base 2 of the total number of states, H provides the information content in bits, implying that the observations allow 2^H states to be distinguished from the prior state space. Conceptually, the observations can be thought of as refining a measuring stick by subdividing each of the prior divisions into 2^H new ones. The larger the information content H, the finer the resolution of the new tick marks on the measuring stick and the more accurately the quantity of interest can be resolved. In multidimensional problems such as the cloud property retrievals considered here, one can envision multiple measuring sticks pointed along each variable in the retrieval vector. After making an observation, each of these measuring sticks will have its own characteristic resolution that depends on the combined sensitivity of the observation vector to the parameter it represents. It will be shown that the total information

content represents the total number of divisions on the measuring sticks that lie along the set of perpendicular directions defined by the eigenvectors of $\mathbf{S}_1\mathbf{S}_2^{-1}$

We now return to the problem of retrieving cloud microphysical properties from radiance observations. Suppose $\mathbf{y}$ represents a set of observed radiances and $\mathbf{x}$ represents the vector of cloud microphysical parameters to be retrieved. Let $\mathbf{S}_a$ and $\mathbf{S}_y$ be the covariance matrices describing the state space prior to making a measurement and the measurement error, respectively. Assuming a physical relationship between $\mathbf{y}$ and $\mathbf{x}$ of the form, for example, $\mathbf{y} = F(\mathbf{x})$, it can be demonstrated that the covariance describing the posterior state space is

$$\mathbf{S}_x = (\mathbf{S}_a^{-1} + \mathbf{K}^{\mathrm{T}}\mathbf{S}_y^{-1}\mathbf{K})^{-1}, \tag{8}$$

where $\mathbf{K}$ is a linearized forward model consisting of the Jacobian of the forward model with respect to the retrieval vector with elements given by

$$K_{ij} = \frac{\partial y_i}{\partial x_j}. \tag{9}$$

The diagonal elements of $\mathbf{S}_x$ provide the variance in the retrieved products in variational retrieval techniques, such as those outlined in Rodgers (1976), Engelen and Stephens (1997), and L'Ecuyer and Stephens (2002).

As noted in Rodgers (2000), to compare the measurement error with the natural variability of the measurements across the full prior state space it is convenient to work in a basis where the measurement errors and prior variances are uncorrelated. Therefore, it is desirable to transform $\mathbf{K}$ into

$$\tilde{\mathbf{K}} = \mathbf{S}_y^{-1/2}\mathbf{K}\mathbf{S}_a^{1/2}, \tag{10}$$

which offers the added benefit of being the basis in which both the prior and measurement covariances are unit matrices. Furthermore, Rodgers (2000) demonstrates that the number of singular values of $\tilde{\mathbf{K}}$ greater than unity defines the number of independent measurements that exceed the measurement noise defining the effective rank of the problem.

Using $\mathbf{S}_a$ for the covariance of the prior state space and Eq. (8) for that of the posterior state space, the SIC becomes

$$\begin{aligned} H &= \frac{1}{2}\log_2|\mathbf{S}_a(\mathbf{K}^{\mathrm{T}}\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1})| \\ &= \frac{1}{2}\log_2|\tilde{\mathbf{K}}^{\mathrm{T}}\tilde{\mathbf{K}} + \mathbf{I}| \\ &= \frac{1}{2}\sum_{i=1}^{N}\log_2|\lambda_i^2 + 1|, \end{aligned} \tag{11}$$

where λ_is are the singular values of $\tilde{\mathbf{K}}$, $\mathbf{I}$ is the $m \times m$ identity matrix, and m is the number of retrieval parameters. In addition, Rodgers (2000) demonstrates that the number of degrees of freedom for signal can be estimated using the singular values of $\tilde{\mathbf{K}}$ via

$$d_s = \sum_i \frac{\lambda_i^2}{1 + \lambda_i^2}, \tag{12}$$

providing another important property of the observing system, namely, the number of independent measurements that can be extracted from the observations.

Thus, we have defined the following four diagnostics for assessing the capabilities of the observing system:

1) the error covariance $\mathbf{S}_x$, characterizing the width of the posterior PDF and providing a measure of the overall accuracy of the retrieval;
2) the SIC H, which measures the relative improvement to our a priori knowledge that results from the addition of the measurements;
3) the number of degrees of freedom for signal d_s, which represents the number of independent observations that can be constructed from the measurements; and
4) the number of singular values that exceed the noise level of the system that defines the effective rank N of the problem. (We can interpret N as the number of independent quantities that can be retrieved from the measurements. In this way, N is analogous to d_s, but applies in retrieval space rather than measurement space.)

Individually, these diagnostics can provide useful information concerning aspects of the retrieval problem, but in the absence of the others they are easily misinterpreted. It is, for example, important to consider the overall accuracy of the retrieval in combination with the information content, because a set of measurements may carry much information in highly underconstrained problems but can still lead to large uncertainties in retrieved products because of the ill-posed nature of the problem. Taken together, however, they provide a more or less complete quantitative description of the retrieval process from observations through to final products that allows for a critical assessment of different algorithms or even distinct platforms in an objective manner.

b. Example: Liquid cloud retrievals from shortwave reflectances

As an example, consider a reflectance-based approach to retrieving the effective radius and liquid water path analogous to that introduced by Nakajima and King (1990). The technique makes use of the fact that

reflected radiances at nonabsorbing wavelengths (referred to as conservative scattering channels) are primarily sensitive to the optical depth of a liquid cloud, while those at absorbing wavelengths (nonconservative scattering channels) are dominated by the size of its constituent cloud droplets. Thus, the combination of reflectances at a nonconservative wavelength (e.g., 2.13 μm) and a conservative wavelength (e.g., 0.64 μm) yields a two-dimensional grid in which a pair of radiance measurements can be related to geometric mean radius and liquid water path. This is illustrated in Fig. 6. Based on the sensitivity studies described above, Fig. 6 also presents the estimated uncertainties in the 0.64- and 2.13-μm channels and illustrates how these errors map into the R_e and LWP retrieval space. Note that uncertainties in the spectral radiance measurements lead to nonuniqueness in the retrieval because the relationship between a given radiance pair and the retrieval parameters becomes multivalued. This effect is particularly noticeable at the thick cloud limit where the sensitivities of the radiances are low and the error bars cover a wide range of LWP values.

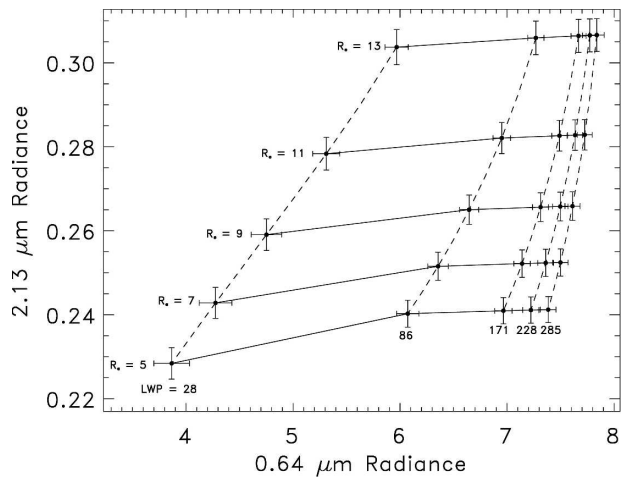

FIG. 6. The mapping of uncertainties in 0.64- and 2.13-μm radiances (W m^{-2} sr^{-1}) into effective radius (μm) and LWP (g m^{-2}) space in a reflectance-based retrieval. Filled circles represent the data points modeled, dashed lines are lines of constant liquid water path, and solid lines are lines of constant effective radius.

Making use of these error estimates and the SIC formalism outlined above, it is straightforward to compute the information content of each of these channels for retrieving R_e and LWP from an a priori range of R_e = 9 ± 4 μm and LWP = 150 ± 65 g m^{-2}, which includes a majority of nonprecipitating liquid clouds in nature (Miles et al. 2000). The results are presented in Fig. 7, which demonstrates the effect of successively adding radiance measurements in the retrieval problem. In this example, all radiance errors are assumed to be ~5% for simplicity. The upper-left-hand panel illustrates the a priori state space. The blue ellipse corresponds to the projection of a two-dimensional Gaussian PDF onto the solution space at the 2σ level, encompassing 95% of the possible solutions. In the absence of observations, any of these solutions are valid and cannot be distinguished from one another.

The upper-right panel demonstrates the impact of a 0.64-μm radiance observation corresponding to a cloud with R_e = 10 μm and LWP = 171 g m^{-2}. Given the sensitivity of this channel to the retrieval parameters and the uncertainties associated with modeling it, this observation reduces the possible solutions to the range of values centered on this combination of effective radius and LWP that is represented by the green ellipse. Following Eq. (7), the SIC in this case is 1.2, indicating that slightly more than two independent states can be resolved from within the initial state space.

The red ellipse in the lower-left panel indicates the range of allowable solutions after further adding a 2.13-μm radiance measurement. The complementary nature of the conservative and nonconservative channels leads to a significant increase in the information content of the system now allowing six independent states to be resolved from the original a priori state space. As indicated by the projections of the new posterior PDF (red ellipse), the errors in retrieved R_e and LWP are significantly reduced when both channels are included in the retrieval. Returning to the measuring stick analogy, the width of the posterior PDF can be viewed as a measure of the "resolution" of the observing system. As $\mathbf{S}_x$ decreases with the addition of more information, it is possible to measure R_e and LWP more precisely. Put simply, the finer the scale of the ruler, the greater the number of distinct states that can be measured.

Interestingly, adding all of the remaining channels from Table 1 provides only a limited amount of additional information to the retrieval. This is illustrated by the yellow ellipse in the lower-right-hand panel of Fig. 7. The range of allowable states has clearly decreased somewhat relative to the red ellipse, but it is unclear whether the increased resolution of the observing system justifies the enormous increase in computation that is required to go from a 2- to an 11-channel framework. Thus, we see that the information content provides a useful diagnostic for establishing the relative performance of different channel combinations, allowing the true value of increased algorithm complexity to be assessed.

It is important to note that the resolution of our measuring stick depends on where we are in state space. In this example, it turns out that the resolution of the

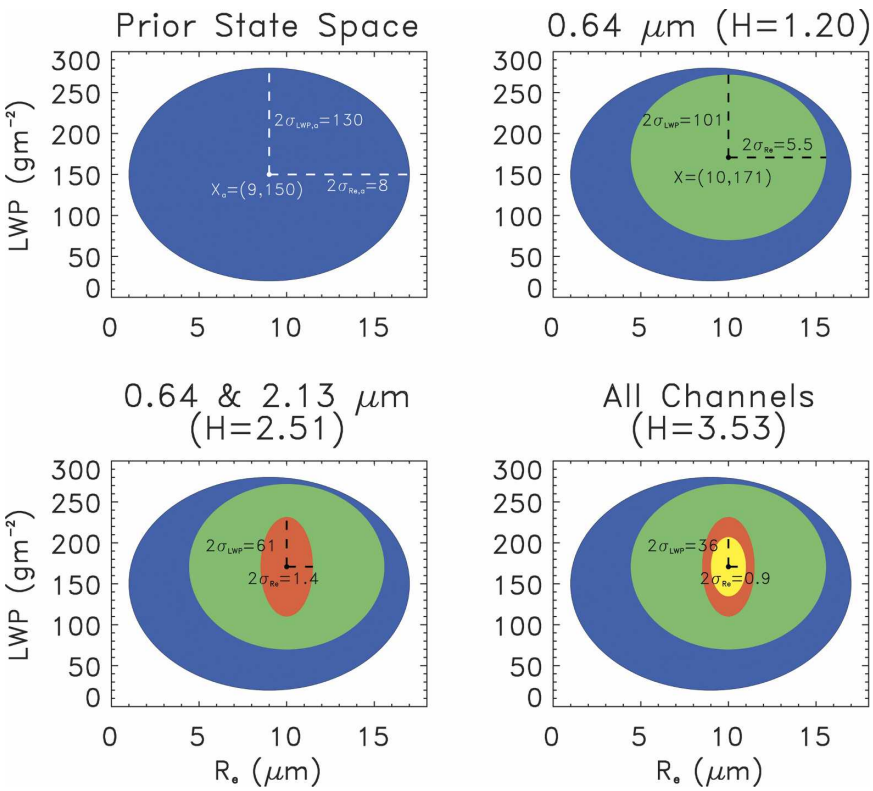

FIG. 7. Graphical representation of the impact of adding information in a two-dimensional retrieval using the retrieval of effective radius and LWP from shortwave reflectance measurements as an example. Each ellipse represents the projection of the corresponding two-dimensional posterior Gaussian PDF of solutions at the level that encompasses 95% of the possible solutions. Results apply to a liquid cloud residing between 1 and 2 km above an oceanic background.

observing system for LWP retrievals is inversely proportional to LWP and proportional to particle size, while its resolution for R_e retrievals is proportional to LWP and inversely proportional to particle size. It is also important to note that the information provided by the observations is not equally divided between the two retrieval parameters. Consider, for example, the addition of the 2.13-μm channel in Fig. 7. The uncertainty in R_e is reduced by a factor of 4 while that in LWP is reduced by only 39%.

c. *Optimal channel selection*

Rodgers (2000) describes a method for extending this framework to optimize a retrieval by objectively selecting the subset of channels that provides the greatest amount of information. To reduce the computation time required to perform the numerous matrix operations necessary for repeated application of the preceding equations, Rodgers (1998) proposes an approach based on sequential modification of the covariance matrix. The procedure first requires that we assess the information content of each individual measurement with respect to our prior knowledge of the retrieval state to create an "information spectrum." The channel with the largest amount of information is then selected and the posterior covariance matrix is adjusted accordingly to account for the information it provides. A new information spectrum for the remaining channels is then calculated with respect to this newly defined state space and a second channel is chosen that provides maximal information relative to the new covariance. This process is repeated and channels are selected sequentially until of the information in all remaining channels falls below the level of measurement noise.

Following Rodgers (1998), and letting $\mathbf{S}_i$ be the error covariance matrix for the state space after i channels have been selected, the information content of channel j of the remaining unselected channels is given by

$$H_j = \frac{1}{2}\log_2(1 + \tilde{\mathbf{k}}_j^{\mathrm{T}}\mathbf{S}_i\tilde{\mathbf{k}}_j), \tag{13}$$

where $\tilde{\mathbf{k}}_j$ is the jth row of $\tilde{\mathbf{K}}$. The H_j form the information spectrum from which the channel with the greatest information is chosen. Taking the chosen channel to be channel l, the covariance matrix is then updated prior to the next iteration via

$$\mathbf{S}_{i+1}^{-1} = \mathbf{S}_i^{-1} + \tilde{\mathbf{k}}_l \tilde{\mathbf{k}}_l^{\mathrm{T}}. \tag{14}$$

In this way channels are selected until none of the remaining channels has an information content exceeding the measurement noise.

Returning to the measuring stick analogy, the first channel chosen initially divides the stick into H_0 increments. Each subsequent selected channel further refines the divisions until all useful information is extracted from the measurements and the finest resolution possible (given the properties of the observing system) is reached. The selection procedure outlined above guarantees that the first channel will provide the greatest number of divisions, followed by the second and so on, making it possible to objectively choose the measurements that maximize our knowledge of the problem while eliminating those channels that provide redundant information.

4. Information content of MODIS measurements

The goal of this study is to assess the information content of the 11 channels summarized in Table 1 for retrieving the properties of liquid clouds. Specifically, if the cloud droplets are assumed to follow a lognormal DSD, the retrieval focuses on retrieving the geometric mean radius and the liquid water path because these parameters completely determine the DSD, provided that one assumes a value of the geometric standard deviation σ_g. In an effort to reduce the number of assumptions required in the forward radiative transfer calculations, the cloud-top height C_{top} and shortwave surface albedo are also considered retrieval parameters, although it is anticipated that these parameters may be constrained in some way using ancillary measurements from another sensor.

a. Covariance matrices

Accurate characterization of the prior and measurement error covariance matrices is central to the problem of calculating information content. In practice, $\mathbf{S}_a$ is easier to define because it merely represents the best estimate of our prior knowledge of the retrieval parameters, in this case geometric mean radius R_e, liquid water path LWP, surface albedo α, and cloud-top height C_{top}. Because our focus is on global retrievals from satellite-based radiance measurements, we anticipate little prior knowledge of the microphysical properties of the clouds, only whether or not a cloud is present by virtue of either an active or passive cloud mask. Based on the climatology of in situ observations of low-level stratiform clouds made between 1972 and 1995, presented in Miles et al. (2000), R_e varies from ~4 to 12 μm while LWP varies from ~25 to 250 g m^{-2} for such clouds. As a generous approximation of this variability, then, we will assume values of 4 and 200 for the variance in R_e and LWP, respectively.

The assumed variances for α and C_{top}, on the other hand, depend on the quality of the ancillary datasets used to define them. Two scenarios are simulated here, with the first assuming no cloud height or surface albedo information beyond climatological mean values, and the other making use of cloud boundaries inferred from radar reflectivity observations from the CloudSat Cloud Profiling Radar (Stephens et al. 2002) and the MODIS surface albedo product (Strahler et al. 1999). In the case that makes use of climatologies (hereinafter referred to as "climatological $\mathbf{x}_a$"), conservative estimates of 2.5 km and 30% are used for the standard deviation in C_{top} and α, respectively. In the case in which ancillary data are used (hereinafter referred to as "ancillary $\mathbf{x}_a$") we adopt the documented value of 10% for the accuracy in the MODIS surface albedo product (Strahler et al. 1999) as the standard deviation in α. Given the vertical resolution of the CPR, we anticipate cloud-top height to be defined with an accuracy of $\pm$250 m, which is modeled by a Gaussian distribution with a variance of 0.0625 km^2 in the second case.

It is interesting to note that using Eq. (7), the CloudSat cloud boundaries and the MODIS surface albedo product have a combined information content of $H \approx 4.9$ bits relative to the climatology-only case. The a priori state space in the second case is, therefore, a factor of $2^{4.9} = 30$ smaller than that in the first case. Thus, we anticipate that the information content of the MODIS observations will be significantly smaller in the second case than the first, even though the former represents a better-posed problem and will undoubtedly lead to more accurate retrievals (Cooper et al. 2003).

In practice $\mathbf{S}_y$ is more difficult to define because it consists of both measurement error as well as errors associated with the forward model used to map the retrieval parameters into measurement space. Two different estimates of the combined forward model and measurement errors will be tested in an effort to illustrate the importance of making rigorous uncertainty estimates in the analysis. In the first case (hereinafter referred to as "uniform measurement errors"), a constant 5% error will be assumed in the radiances at all wavelengths, while in the other (hereinafter referred to

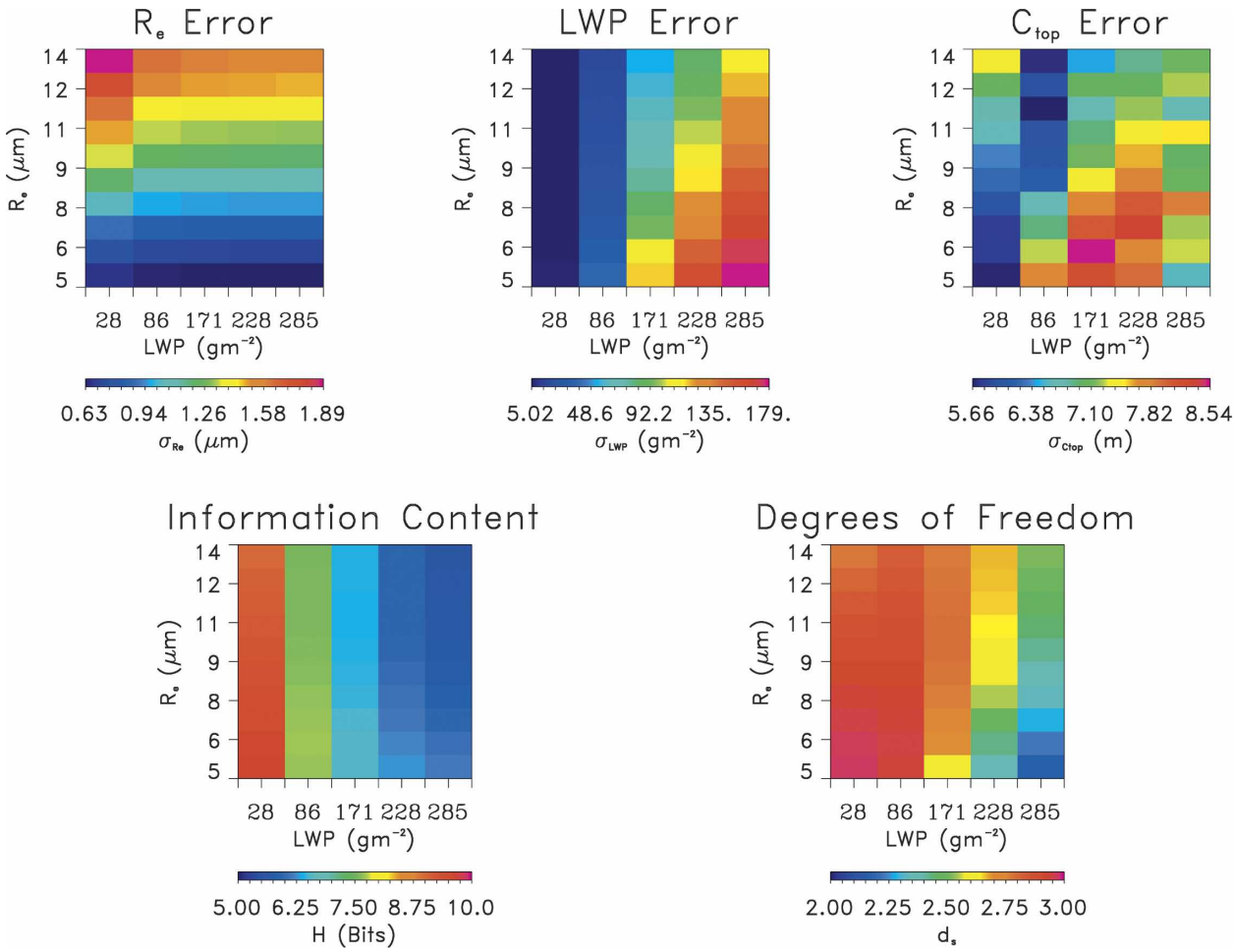

FIG. 8. Uncertainties in effective radius, LWP, and cloud-top height, total information content, and total degrees of freedom for signal over a wide range of LWP and R_e. Uncertainties in all radiances resulting from observation and forward model errors are assumed to be 5%.

as "realistic measurement errors") we will make use of the more realistic fractional uncertainties determined in section 2a to represent the measurement standard deviations with the understanding that they include a number of important sources of uncertainty but neglect others that may be significant under certain conditions.

For simplicity, both $\mathbf{S}_y$ and $\mathbf{S}_a$ will be assumed to be diagonal. While it may be reasonable to assume that measurement errors are uncorrelated, the forward model component of these uncertainties invariably causes uncertainties in channels with similar sensitivities to be correlated with one another. Furthermore, with the exception of the surface albedo, it is unlikely that the retrieval parameters themselves are uncorrelated given the microphysical processes that govern the nucleation and growth of cloud droplets and their sensitivity to temperature and humidity. It is, however, beyond the scope of this work to explore the correlations between these quantities in the depth required to accurately represent them in the covariance matrices. In fact, because including knowledge of the correlations between channels constitutes adding information to the retrieval, one runs the risk of adding spurious information that may degrade the results if care is not taken to properly estimate the off-diagonal elements of the covariance matrices. Thus, the results that follow should be viewed as a first attempt at providing ballpark estimates of the information content of the system that represent the worst-case scenario where no advance knowledge of correlations between channels is available.

b. Total information

Figure 8 presents expected retrieval errors, defined by

$$\varepsilon_i = \frac{\sqrt{S_{ii}}}{X_i} \times 100, \tag{15}$$

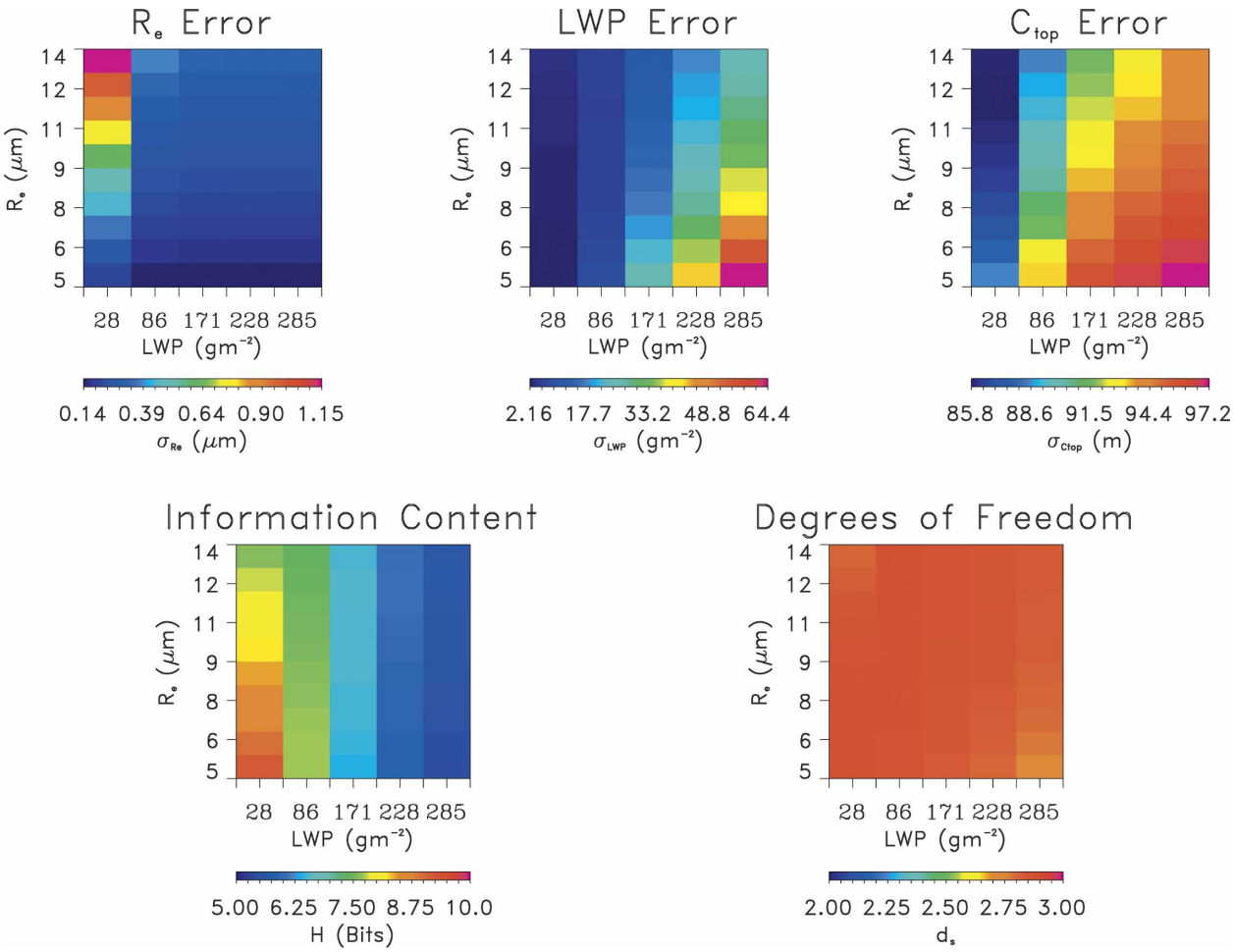

FIG. 9. As in Fig. 8, but using rigorous estimates of the uncertainties resulting from observation and forward model errors.

where S_{ii} and X_i are the diagonal element of the $\mathbf{S}_x$ and retrieval vector $\mathbf{x}$ corresponding to the ith and total information content from Eqs. (8) and (11) for 50 pairs of R_e and LWP, assuming uniform measurement errors and a climatological $\mathbf{x}_a$. The most striking artifact of failing to account for the wavelength dependence of the measurement errors is the unrealistically small uncertainties of cloud-top height that range from 5.5 to 8.5 m. This is a direct consequence of improperly accounting for the large influence of uncertainties in the assumed specific humidity profile on our ability to model radiances at 1.38 μm. When the realistic uncertainty estimates from section 3a are employed (Fig. 9), the resolution to which C_{top} can be estimated decreases by more than an order of magnitude to ~100 m in response to increasing the fractional error in the 1.38-μm channel. The loss of C_{top} information leads to a slight reduction in the total information content relative to the case with uniform measurement errors, but this is partially compensated for by improved LWP and R_e retrievals through a reduction in the uncertainties in other channels.

The top panels in Fig. 9 suggest that, provided the cloud is homogeneous across the instrument field of view and the particle DSD does not vary vertically, R_e and LWP can be retrieved with accuracies of approximately 5% and 20%, respectively over much of the range of clouds examined. This is, in part, a result of the fact that the variability in the microphysical properties of liquid clouds observed in nature is relatively small (in comparison with ice clouds, e.g.) and, in part, because of the good sensitivity of the shortwave reflectance from these clouds to the parameters of interest. In general, the information content is largest for thin clouds composed of small particles because the shortwave reflectances are most sensitive to changes in cloud properties for optically thin clouds (see Fig. 1). This leads to a factor of 3 better LWP resolution (in a fractional sense) for thin clouds than thick ones.

The effects of neglecting wavelength-dependent un-

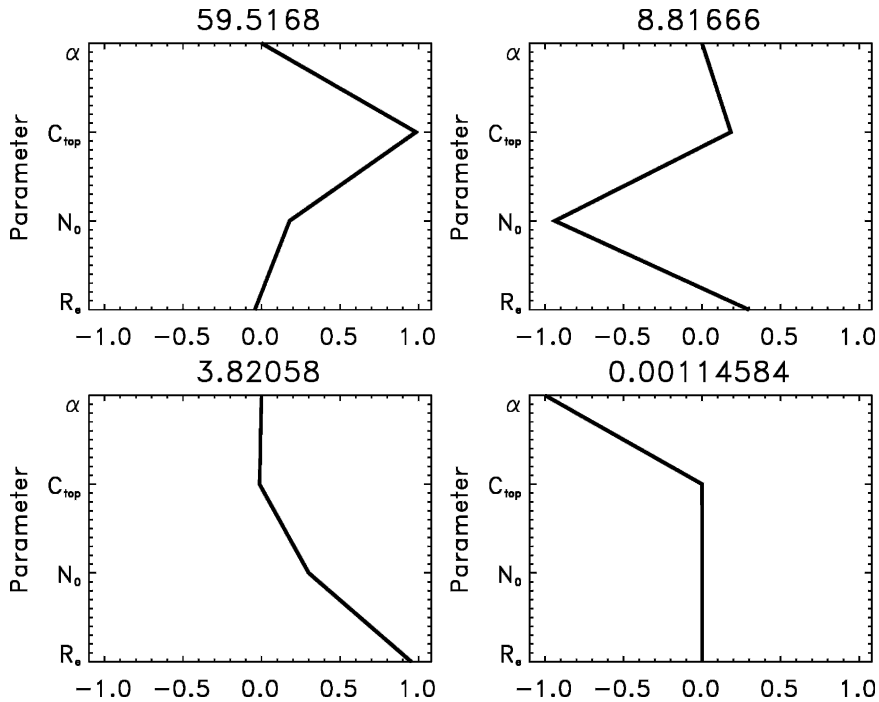

FIG. 10. Singular vectors of the $\tilde{\mathbf{K}}$ matrix assuming 5% uncertainties in all radiances resulting from observation and forward model errors. Results correspond to daytime observations of a cloud with LWP = 85.71 g m^{-2} and R_e = 8 μm residing between 1 and 2 km over an oceanic surface.

certainties are examined in greater detail in Figs. 10 and 11 where the singular vectors of the $\tilde{\mathbf{K}}$ matrix for a cloud with LWP = 85.71 g m^{-2} and R_e = 8 μm, assuming that uniform and realistic measurement errors, respectively, are presented. In both cases, three singular values exceed the noise level of 1.0 indicating that three independent parameters may be inferred from the data. Because the singular vectors indicate the linear combination of retrieval parameters that correspond to each singular value, these images confirm the shift from cloud-top height information when the 1.38-μm radiance is assumed to be unrealistically accurate to LWP and R_e when rigorous uncertainty estimates are used. In Fig. 10, the largest singular vector (corresponding to the primary information in the system) corresponds to C_{top} while LWP and R_e information are relegated to the second and third singular vectors. When more realistic errors are assumed, R_e and LWP information is coupled into the first two singular vectors while cloud-top height information corresponds to the third singular vector, which is more than an order of magnitude smaller than the other two. These results highlight the importance of accurate covariance matrix characterization in setting up an information content analysis. The results also imply that more accurate specification of the humidity profile could allow in principle cloud-top height to be retrieved very precisely. While this may not be extremely important, better estimates of cloud top may lead to small improvements in applications involving modeling the radiative impacts of such clouds. Note that in both cases the fourth singular value, corresponding to surface albedo, is much less than unity, suggesting that the clouds are thick enough to completely obscure the surface. This precludes any quantitative information regarding the surface albedo from being extracted from the data.

The potential for using infrared observations to retrieve liquid cloud properties at night is briefly explored in Fig. 12. Because of the fact that liquid clouds reside near the surface, uncertainties in nighttime retrievals approach the values of the a priori knowledge (4 μm in R_e, 200 g m^{-2} in LWP, and 250 m in C_{top}), consistent with the fact that the total information content in all of the channels sums to less than 1 in most cases, indicating that at most two distinct states can be resolved by the observing system. In general, there is no information regarding liquid water path with the exception of extremely thin optical depths. What little information there is for the remaining clouds transitions from C_{top} to particle size as R_e increases. Thus, nighttime retrievals can, at best, differentiate between overcast and clear-sky scenes, providing little modification to the cli-

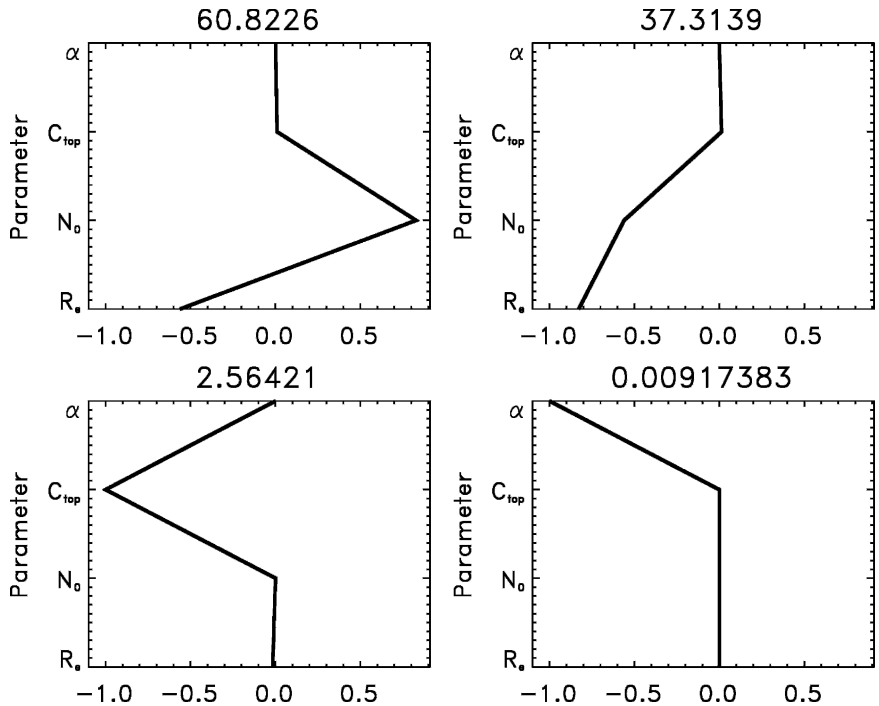

FIG. 11. As in Fig. 10, but using rigorous estimates of the uncertainties resulting from observation and forward model errors.

matological mean values of R_e and LWP used to initialize an algorithm.

c. Channel selection

Information spectra derived using Eq. (13) are presented in Fig. 13 for a cloud with LWP = 85.71 g m^{-2} and R_e = 8 μm, assuming realistic measurement errors and cloud top and surface albedo from ancillary data. The solid line traces the information content of each channel with respect to the a priori assumptions. Once the channel with the most information is selected, the solution space is reduced accordingly and the second (dotted) line traces the information content of the remaining channels with respect to this new solution space. The process is then repeated by selecting the channel with the largest information content from those that remain and readjusting the solution space to account for the information it provides, etc. The process stops as soon as none of the remaining channels exhibit an information content in excess of 0.346, representing the level of noise in the system obtained by substituting a SNR of 1 into Eq. (13). In this case, the 1.64-μm channel provides the most information relative to the a priori solution space and is, therefore, selected first. Once the solution space has been adjusted to account for the information provided by the 1.64-μm channel, there is a dramatic drop in the information content of all channels containing redundant information, such as 2.13 and 3.75 μm. This leaves the two visible channels at 0.64 and 0.88 μm with the largest information contents of those that remain. Because the information content at 0.64 μm is slightly greater than that at 0.88 μm, it is chosen second. Last, once the impact of the 0.64-μm channel has been accounted for, the only channel that provides new information to the retrieval is the 1.38-μm water vapor channel, which is sensitive to cloud-top height, so it is selected third.

Figure 14 repeats the procedure assuming uniform measurement errors. Now, the 1.38-μm channel is selected first because of the unrealistically strong information it provides regarding C_{top}. Again, 0.64 μm is picked for the information it provides in constraining LWP, but now the 3.75-μm channel is picked for particle size information rather than the 1.64-μm channel, because S_x fails to account for the order of magnitude difference in their uncertainties. This result further emphasizes the importance of rigorously modeling the wavelength dependence of model errors within the information content framework.

Applying this procedure to all 50 combinations of LWP and R_e described above the subset of the 11 channels that provide the greatest information for the widest variety of liquid cloud retrievals is identified. Table 2 summarizes the fraction of these cases for which any

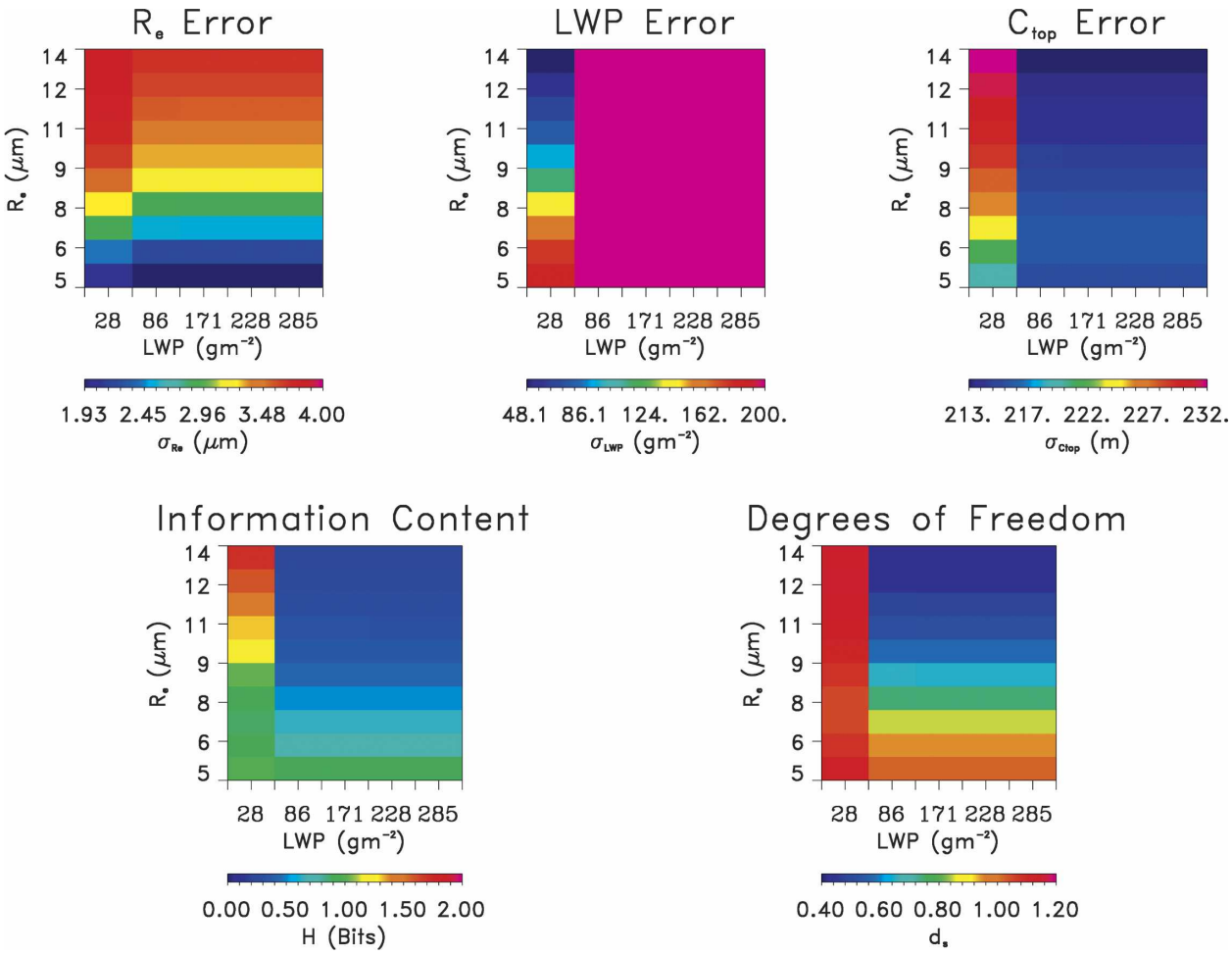

FIG. 12. As in Fig. 9, but for nighttime retrievals.

given channel is selected assuming realistic measurement errors and the existence of ancillary $\mathbf{x}_a$ data. In general, the results agree with the consensus of retrieval approaches in the literature that liquid cloud microphysical property retrievals will achieve the best results by merging the information contained in a conservative and a nonconservative shortwave scattering channel. The results also suggest that there maybe useful cloud height information in the 1.38-μm water vapor channel because it is selected for half of the retrievals (although never selected first). This result is somewhat unexpected because water vapor absorption at 1.38 μm is generally considered to completely mask low cloud, but the analysis conducted here reveals that under some conditions this channel may provide additional information in a retrieval. Such information could lead in turn to subsequent improvements in modeling the radiative impacts of low clouds in the atmosphere. Extracting it, however, requires that the sensor be capable of measuring very small absolute radiances and success is very unlikely under moist conditions such as those that prevail in the Tropics.

The analysis indicates that 1.64 and 0.64 μm are the channels of choice for R_e and LWP retrievals because of their slightly larger sensitivities and slightly lower uncertainties than 2.13 and 0.88 μm. It should, however, be noted that even though 1.64 μm is selected ~5 times more often than 2.13 μm, the information content of each of these channels with respect to the a priori is very similar, and 2.13 μm could be substituted for 1.64 μm in a retrieval with very little impact on the results. A similar argument can be made for substituting 0.88 μm for 0.64 μm under most circumstances. From a consistency standpoint, however, it is desirable to use one or the other rather than switching back and forth. Equivalent results for the case of nighttime retrievals are summarized in the last row of Table 2. The results echo the lack of information indicated by Fig. 12 because the observations provide no useful information for 60% of the liquid clouds examined. In every one of

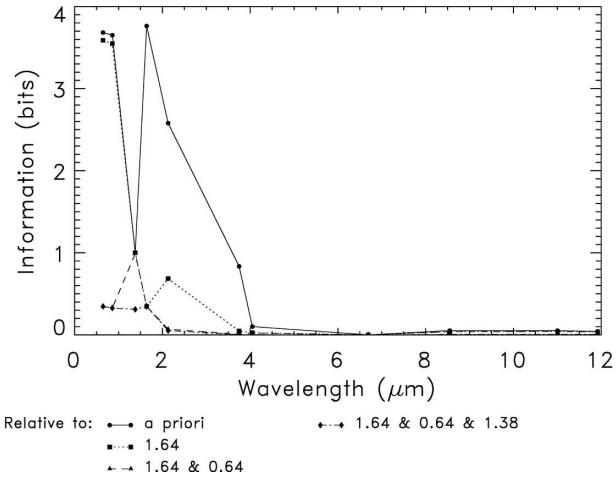

FIG. 13. Information spectra governing the selection of optimal channels for retrieving liquid cloud microphysical properties from daytime observations of a cloud with LWP = 85.71 g m^{-2} and R_e = 8 μm residing between 1 and 2 km over an oceanic surface. Results are obtained relative to the radiance uncertainty estimates from section 3a.

the remaining cases, the 3.75-μm channel provides the most information and all of the other channels are essentially redundant.

Because it is not always possible to constrain cloud boundary information using an active sensor, it is interesting to compare these results with those that correspond to retrievals in the absence of ancillary CloudSat cloud boundary information or MODIS albedo information. Similar results corresponding to retrievals employing climatological $\mathbf{x}_a$ are summarized in Table 3. Because cloud boundary information is not as important for accurate daytime retrievals, the daytime results

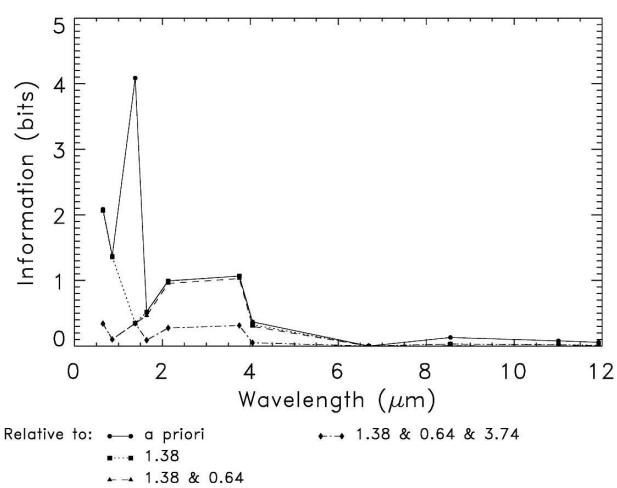

FIG. 14. As in Fig. 13, but relative to constant radiance uncertainties of 5%.

TABLE 2. Relative frequencies with which channels are selected for use in a retrieval based on their information content. The third column lists the fraction of the cases examined for which the corresponding channel provided the most information and was selected first. The fraction of cases for which the channel offered any contribution to the retrieval and was selected at any time is presented in the fourth column.

Band	Wavelength	Selected first	Selected any position
Daytime			
6	1.64	64	97
1	0.64	32	84
26	1.38	0	50
7	2.13	0	20
2	0.86	4	16
Nighttime			
20	3.75	40	40

are not extremely sensitive to the presence of ancillary data with the exception of the fact that the 11.0-μm channel now provides a nonnegligible contribution to the retrievals 50% of the time (although it is never the first channel selected). Nighttime retrievals, on the other hand, are much more promising with nonnegligible information coming from four different channels at some point over the range of clouds modeled. In general, little information is lost provided that the 3.75- and 11.0-μm channels are adopted for nighttime liquid cloud retrievals. Inspection of the singular vectors for a cloud with LWP = 85.71 g m^{-2} and R_e = 8 μm (illustrated in Fig. 15) demonstrates that all of the information from these channels is focused onto two singular vectors that derive from linear combinations of C_{top} and R_e. Thus, a nighttime retrieval using measurements at 3.75 and 11.0 μm offers the potential to retrieve some information regarding cloud height and to a lesser extend geometric mean radius in the absence of ancillary cloud boundary and albedo information.

TABLE 3. As in Table 2, but for the case in which no ancillary cloud boundary or surface albedo information is available.

Band	Wavelength	Selected first	Selected any position
Daytime			
1	0.64	28	100
6	1.64	41	96
31	11.0	0	50
26	1.38	27	50
7	2.13	0	20
2	0.86	4	12
Nighttime			
20	3.75	80	92
31	11.0	20	64
29	8.55	0	16
32	11.92	0	11

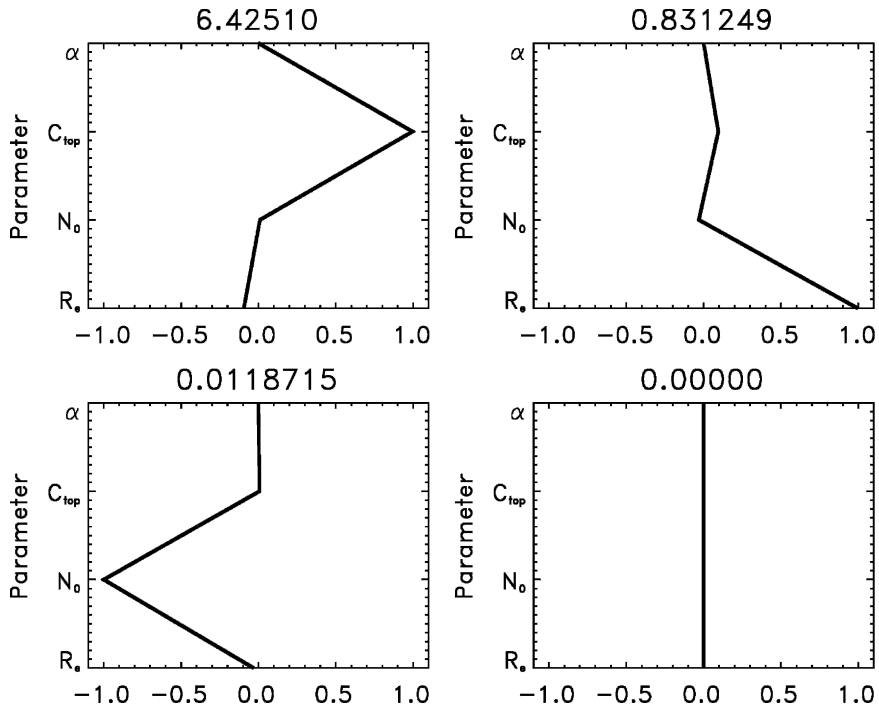

FIG. 15. As in Fig. 11, but for nighttime retrievals in the absence of cloud boundary or albedo information.

5. Conclusions

Accurate measurements of the spatial and temporal distribution of cloud microphysical properties are essential for modeling their role in global climate change. As the number of satellites and algorithms devoted to this goal increases, it is of interest to revisit the cloud retrieval to objectively assess the information content of the measurements and establish quantitative estimates of the accuracy to which various cloud properties can be retrieved. To this end, this paper introduces four diagnostics of information and uncertainty and applies them to study the problem of retrieving liquid cloud microphysical properties from visible and infrared radiances.

Based on a rigorous analysis of the errors and sensitivities of satellite measurements at 11 visible and infrared wavelengths, the 0.64- and 1.64-μm channels emerge as the optimal set for retrieving liquid water path and geometric mean radius during the daytime. There is some evidence that the 1.38-μm channel may also provide information regarding cloud-top height under certain circumstances but, because the use of this channel may put unrealistic constraints on a detector's minimum detectable absolute radiance, its utility is left as an open question at this time. At night, 3.75 and 11.0 μm are found to provide a limited amount of information particularly in the absence of ancillary data to constrain cloud height. In both cases other channels may provide a small amount of additional information under certain conditions, but generally the remaining channels supply only redundant information and do not justify the additional computational cost required to integrate them into an algorithm.

Information contents computed for 50 pairs of geometric radius and liquid water path indicate that daytime retrievals can be expected to resolve up to ~200 states at low LWP and 32 states for thicker clouds, while nighttime retrievals may distinguish at most 2–4 broad cloud states. Further analysis of retrieval error covariance matrices derived from careful consideration of a number of important sources of forward model and measurement errors demonstrate that daytime geometric mean radius retrievals can be expected to be accurate to 5%–10%, while uncertainties in retrieved LWP are expected to be ~10%–20%. Uncertainties in nighttime retrievals, on the other hand, are comparable to the ranges to which the retrieval parameters can be constrained using climatological data.

At present, these conclusions apply only over an oceanic background. Similar analyses must be conducted to assess the validity of this set of channels for retrievals over land surfaces, but such calculations are beyond the scope of this preliminary study. Furthermore only liq-

uid clouds have been modeled here but this study lays the groundwork for the application of the information content analysis to the more challenging problem of ice cloud retrievals that suffer from additional uncertainties introduced by the varying shape of their constituent crystals. This problem is addressed in Part II, which immediately follows this paper.

Last, note that the development presented here is neither contingent on the details of any forward model, nor is its scope limited to the cloud retrieval problem. In fact, to realize the full potential of the rigorous information content analyses outlined here, they should be applied in the developmental stages of future satellite missions to systematically develop instruments from the ground up. In principle, a detailed analysis of a wide variety of potential wavelengths could be performed to determine the subset that provides the most information for the desired application based on the accuracies to which they can be modeled and their sensitivities to the retrieval parameters. Then, this information could be used to design an optimal channel configuration upon which an instrument could be constructed.

Acknowledgments. This research was supported in part by NASA Research Grant NNG04GE35G, NASA Earth System Science Fellowship NGT5-30458, NASA CloudSat Mission Grant NAS5-99237, and NASA CALIPSO Mission Grant NAS1-99103. The authors thank Mick Christi and Professor Laurent Labonnote for their work in developing the Radiant model and their assistance in running it.

REFERENCES

Ackerman, S. A., W. L. Smith, J. D. Spinhirne, and H. E. Revercomb, 1990: The 27–28 October 1986 FIRE IFO cirrus case study: Spectral properties of cirrus clouds in the 8–12 μm window. *Mon. Wea. Rev.,* **118,** 2377–2388.

Baum, B. A., P. F. Soulen, K. I. Strabala, M. D. King, S. A. Ackerman, W. P. Menzel, and P. Yang, 2000: Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS 2. Cloud thermodynamic phase. *J. Geophys. Res.,* **105,** 11 781–11 789.

Bernardo, J. M., and A. F. M. Smith, 1994: *Bayesian Theory.* John Wiley and Sons, 586 pp.

Christi, M. J., and G. L. Stephens, 2002: On retrieving profiles of CO_2 in the lower atmosphere using spectroscopy in the near and far infrared: A preliminary study. Department of Atmospheric Science, Colorado State University, Paper 728, 126 pp.

——, and P. Gabriel, 2004: Radiant 2.0: A user's guide. Department of Atmospheric Science, Colorado State University, Paper 742, 39 pp.

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,** 4107, doi:10.1029/2002JD002611.

——, ——, P. Gabriel, A. J. Baran, and G. L. Stephens, 2006: Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. Part II: Ice clouds. *J. Appl. Meteor. Climatol.,* **45,** 42–62.

Deirmendjian, D., 1969: *Electromagnetic Scattering on Spherical Polydispersions.* Elsevier, 310 pp.

Engelen, R. J., and G. L. Stephens, 1997: Infrared radiative transfer in the 9.6 μm band: Application to TOVS ozone retrieval. *J. Geophys. Res.,* **102,** 6929–6940.

——, and ——, 2004: Information content of infrared satellite sounding measurements with respect to CO_2. *J. Appl. Meteor.,* **43,** 373–378.

Eriksson, P., 2000: Analysis and comparison of two linear regularization methods for passive atmospheric observations. *J. Geophys. Res.,* **105,** 18 157–18 167.

Evans, K. F., S. J. Walter, A. J. Heymsfield, and G. M. McFarquhar, 2002: Submillimeter-wave cloud ice radiometer: Simulations of retrieval algorithm performance. *J. Geophys. Res.,* **107,** 4028, doi:10.1029/2001JD000709.

Eyre, J. R., 1990: The information content of data from satellite sounding systems: A simulation study. *Quart. J. Roy. Meteor. Soc.,* **116,** 401–434.

——, G. A. Kelly, A. P. McNally, E. Anderson, and A. Persson, 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.,* **119,** 1427–1463.

Gabriel, P., M. Christi, and G. L. Stephens, 2005: Calculation of Jacobians for inverse radiative transfer: An efficient hybrid method. *J. Quant. Spectrosc. Radiat. Transfer,* in press.

Hadji-Lazaro, J., C. Clerbaux, and S. Thiria, 1999: An inversion algorithm using neural networks to retrieve atmospheric CO total columns from high-resolution nadir radiances. *J. Geophys. Res.,* **104,** 23 841–23 854.

Harrison, E. F., P. Minnis, B. R. Barkstrom, V. Ramanathan, R. D. Cess, and G. G. Gibson, 1990: Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment. *J. Geophys. Res.,* **95,** 18 687–18 703.

Inoue, T., 1985: On the temperature and effective emissivity determination of semitransparent cirrus clouds by bispectral measurements in the 10 μm window region. *J. Meteor. Soc. Japan,* **63,** 88–98.

Kratz, D., 1995: The correlated-k distribution technique as applied to the AVHRR channels. *J. Quant. Spectrosc. Radiat. Transfer,* **64,** 501–517.

Kullback, S., 1968: *Information Theory and Statistics.* John Wiley and Sons, 416 pp.

L'Ecuyer, T. S., and G. L. Stephens, 2002: An estimation-based precipitation retrieval algorithm for attenuating radars. *J. Appl. Meteor.,* **41,** 272–285.

Li, J., and H.-L. Huang, 1999: Retrieval of atmospheric profiles from satellite sounder measurements by use of the discrepancy principle. *Appl. Opt.,* **38,** 916–923.

Liou, K.-N., 1986: Influence of cirrus clouds on weather and climate processes: A global perspective. *Mon. Wea. Rev.,* **114,** 1167–1199.

McClatchey, F. A., R. W. Fenn, J. E. Selby, F. E. Volz, and J. S. Goring, 1972: Optical properties of the atmosphere. 3d ed. Air Force Cambridge Research Laboratory Tech. Rep. AFCRL-72-0497, 102 pp.

Miles, N. L., J. Verlinde, and E. E. Clothiaux, 2000: Cloud droplet

size distributions in low-level stratiform clouds. *J. Atmos. Sci.,* **57,** 295–311.

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,** 19 955–19 971.

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.

Prabhakara, C., R. S. Fraser, G. Dalu, M.-L. C. Wu, and R. J. Curran, 1988: Thin cirrus clouds: Seasonal distribution over oceans deduced from *Nimbus-4* IRIS. *J. Appl. Meteor.,* **27,** 379–399.

Rodgers, C. D., 1976: Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.,* **14,** 609–624.

——, 1998: Information content and optimisation of high spectral resolution measurements. *Adv. Space Res.,* **21,** 361–367.

——, 2000: *Inverse Methods for Atmospheric Sounding. Theory and Practice.* World Scientific, 256 pp.

Rolland, P., K.-N. Liou, M. D. King, S.-C. Tsay, and G. M. McFarquhar, 2000: Remote sensing of optical and microphysical properties of cirrus clouds using Moderate Resolution Imaging Spectroradiometer channels: Methodology and sensitivity to physical assumptions. *J. Geophys. Res.,* **105,** 11 721–11 738.

Seinfeld, J. H., and S. N. Pandis, 1998: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change.* John Wiley and Sons, 1326 pp.

Shannon, C. E., and W. Weaver, 1949: *The Mathematical Theory of Communication.* University of Illinois Press, 144 pp.

Stephens, G. L., 1994: *Remote Sensing of the Lower Atmosphere.* Oxford University Press, 544 pp.

——, and Coauthors, 2002: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bull. Amer. Meteor. Soc.,* **83,** 1771–1790.

Strabala, K. I., S. A. Ackerman, and W. P. Menzel, 1994: Cloud properties inferred from 8–12 μm data. *J. Appl. Meteor.,* **33,** 212–229.

Strahler, A. H., J.-P. Muller, and MODIS Science Team Members, 1999: MODIS BRDF/Albedo product: Algorithm theoretical basis document V. 5.0. MODIS Product MOD43, 53 pp.

Suggs, R. J., G. J. Jedlovec, and A. R. Guillory, 1998: Retrieval of geophysical parameters from GOES: Evaluation of a split-window technique. *J. Appl. Meteor.,* **37,** 1205–1227.

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

Tanre, D., M. Herman, and Y. J. Kaufman, 1996: Information on aerosol size distribution contained in solar and reflected spectral radiances. *J. Geophys. Res.,* **101,** 19 043–19 060.

Taylor, B. N., and P. J. Mohr, cited 2004: The NIST reference on constants, units, and uncertainty. NIST. [Available online at http://physics.nist.gov/cuu/index.html.]

Worden, J. R., K. W. Bowman, and D. B. Jones, 1999: Two-dimensional characterization of atmospheric profile retrievals from limb sounding observations. *J. Quant. Spectrosc. Radiat. Transfer,* **86,** 45–71.