RESEARCH ARTICLE

Key Points:
- Operational satellite algorithms can retrieve cloud thermodynamic phase from a combination of shortwave and infrared observations
- We quantify phase probability using combinations of MODIS/VIIRS shortwave-only channels: 0.865, 1.64, 2.13, and 2.25 μm
- For an ice cloud of $\tau = 10$, $r_{eff} = 12$ μm probability of ice phase retrieval increases from 65 to 82% by combining 2.13/2.25 μm channels

Supporting Information:
- Supporting Information S1

Correspondence to:
O. M. Coddington,
odele.coddington@lasp.colorado.edu

Characterizing the information content of cloud thermodynamic phase retrievals from the notional PACE OCI shortwave reflectance measurements

O. M. Coddington[1] [iD], T. Vukicevic[2], K. S. Schmidt[1,3] [iD], and S. Platnick[4] [iD]

[1]Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, Colorado, USA, [2]ISED Office of Water Prediction NWS/NOAA, Tuscaloosa, Alabama, USA, [3]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado, USA, [4]Earth Science Division, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Abstract We rigorously quantify the probability of liquid or ice thermodynamic phase using only shortwave spectral channels specific to the National Aeronautics and Space Administration's Moderate Resolution Imaging Spectroradiometer, Visible Infrared Imaging Radiometer Suite, and the notional future Plankton, Aerosol, Cloud, ocean Ecosystem imager. The results show that two shortwave-infrared channels (2135 and 2250 nm) provide more information on cloud thermodynamic phase than either channel alone; in one case, the probability of ice phase retrieval increases from 65 to 82% by combining 2135 and 2250 nm channels. The analysis is performed with a nonlinear statistical estimation approach, the GEneralized Nonlinear Retrieval Analysis (GENRA). The GENRA technique has previously been used to quantify the retrieval of cloud optical properties from passive shortwave observations, for an assumed thermodynamic phase. Here we present the methodology needed to extend the utility of GENRA to a binary thermodynamic phase space (i.e., liquid or ice). We apply formal information content metrics to quantify our results; two of these (mutual and conditional information) have not previously been used in the field of cloud studies.

1. Introduction

A critical first step in useful cloud optical property retrievals (optical thickness and droplet effective radius) is the retrieval of cloud thermodynamic phase [*Marchant et al.*, 2016]. Existing operational satellite algorithms derive cloud thermodynamic phase from cloud observations in one or more discrete spectral channels where water absorbs solar and/or infrared radiation differently for liquid and ice phases. However, in many cases, the measurement information does not uniquely indicate phase [e.g., *Marchant et al.*, 2016].

Most of the satellite imagers that have contributed to individual cloud property data sets in a nearly 30 year long global record [*Stubenrauch et al.*, 2013] have used a combination of infrared measurements along with visible and near-infrared channels to derive the cloud properties [e.g., *Baum et al.*, 2012; *Menzel et al.*, 2008; *Pavolonis and Heidinger*, 2004]. For example, the National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) instrument [*King et al.*, 1992] on the Aqua and Terra platforms uses infrared channels in a weighted voting discrimination logic to help extract cloud thermodynamic phase information using trispectral infrared and cloud top temperature tests that produce an integer result, the total sum of which determines cloud phase as liquid, ice, or undetermined [*Platnick et al.*, 2014, 2017]. To a lesser extent, this is also true of the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi NPP platform where cloud brightness temperature results from a single infrared channel are used to assign a series of follow-on spectral tests (near-infrared and infrared) designed to identify cloud phase/type in five categories: liquid, super cooled mixed phase, opaque ice or deep convection, nonopaque ice, or overlapping clouds [*Pavolonis et al.*, 2005]. A future NASA mission, the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission, with a notional launch date in the early 2020s has, as part of its science goals, the generation of global cloud properties. However, unlike the imagers discussed above, the PACE mission will not have infrared measurements, which motivates us to assess the ability to discriminate cloud thermodynamic phase retrievals from shortwave channels alone.

The PACE imager, the Ocean Color Instrument (OCI), is being designed and built by NASA Goddard Space Flight Center. To meet the mission science goals, OCI is expected to be a hyperspectral instrument from 350 nm to 890 nm with 5 nm spectral resolution, plus six, discrete, shortwave spectral channels between

Table 1. Nominal Specifications for Principle PACE OCI Channels for Cloud Product Studies[a]

Central Wavelength (nm)	Band Pass (nm)	Channels in Common
665	10	MODIS, VIIRS
865[b]	40	MODIS, VIIRS
763	5	MODIS
940	25	MODIS
1240	20	MODIS, VIIRS
1378	10	MODIS, VIIRS
1640[b]	40	MODIS, VIIRS
2135[b]	50	MODIS
2250[b]	50	VIIRS

[a]The far-right column indicates the PACE shortwave channels that are in common with channels on the MODIS and VIIRS instruments.
[b]The measurement channels evaluated in this study.

940 and 2250 nm; the radiometric accuracy is currently specified at 3%. The channel centers and widths for the OCI instrument, specific to cloud product studies, are listed in Table 1 (from the PACE Science Definition Team Report, pp. xxv and xxxi [*Del Castillo*, 2012]). All of the channels will have 1 km spatial resolution at nadir.

The shortwave PACE channels listed in Table 1 are in common with channels on the MODIS and VIIRS instruments. The primary difference in the location and number of measurement channels is near 2 μm (MODIS = 2135 nm, VIIRS = 2250 nm, PACE = 2135 and 2250 nm) where water strongly absorbs and retrievals of particle size have the greatest sensitivity. The lack of infrared channels on PACE prompts this question: *Do the MODIS and VIIRS combined shortwave-infrared (SWIR) channels at 2135 and 2250 nm provide more information on cloud thermodynamic phase than each individual set of channels?* An appropriate follow-on question is *Can we rigorously quantify the probability of liquid or ice cloud phase given a set of measurements with their associated uncertainties and a set of simulated solutions from cloud radiation models with their own set of associated uncertainties?*

Previously, we rigorously quantified the information content in the retrieval of cloud optical properties *for an assumed cloud thermodynamic phase* using the GEneralized Nonlinear Retrieval Analysis (GENRA) technique, a nonlinear statistical estimation approach derived from general inverse theory [*Vukicevic et al.*, 2010; *Coddington et al.*, 2012, 2013]. Specifically, these earlier studies quantified the probability distribution of cloud optical thickness, τ, and droplet effective radius, r_{eff}, from shortwave cloud measurements. The basis of these retrievals is that differences in the absorption coefficient of liquid and ice water, which are significantly larger in the SWIR relative to the visible, result in the net cloud reflectance decreasing with particle size in the SWIR [*Pilewskie and Twomey*, 1987; *Twomey and Cocks*, 1989]. Many cloud retrieval algorithms are based on combining measurement channels insensitive to water absorption that provide information on τ with those sensitive to water absorption that provide information on both τ and r_{eff} [*Nakajima and King*, 1990; *Platnick et al.*, 2003]. Other algorithms use combinations of measurement channels of sufficient difference in sensitivity to water absorption to retrieve information on τ and r_{eff} [*Platnick et al.*, 2001; *Meyer et al.*, 2016].

In this paper, we extend the GENRA technique to a *binary thermodynamic phase parameter space* (i.e., liquid or ice) and formalize the theory necessary to rigorously quantify cloud optical property information given this additional challenge. We apply a variety of metrics to quantify the formal information in a set of measurements. The well-known Shannon information content [*Shannon and Weaver*, 1949] is a measure of the information to be gained by making a measurement. The mutual and conditional information [*Cover and Thomas*, 2006], respectively, quantify the information in a measurement that is shared between physical parameters and the degree to which information about one parameter can be gained given complete knowledge of a different physical parameter that it shares information with. The addition of mutual and conditional information to the GENRA technique is new for this work. We note that the extended GENRA methodology presented here is also relevant for retrieval studies where there are more than two equally plausible physical interpretations for a single set of measurements. While the GENRA technique has thus far only been applied to cloud studies, its generalized nature can be practically extended to any retrieval making use of a metric of best fit between measured and simulated observations.

Our results demonstrate how these information content diagnostics can be applied to evaluate cloud thermodynamic phase discrimination, to quantify the uncertainty in τ and r_{eff} retrievals, to quantify the correlations between the two retrievals, and to investigate the potential of the PACE OCI instrument in providing useful cloud property data records relative to MODIS and VIIRS. In section 2, we provide the theory of the GENRA algorithm. In section 3, we outline the implementation of simulated MODIS, VIIRS, and PACE cloud

reflectance observations in GENRA. The approach to quantify the discrimination of cloud phase follows in section 4. In section 5, we present results of the probability of retrieving the correct thermodynamic phase over a dark surface and over a broad range of τ and r_{eff} for MODIS, VIIRS, and PACE. In section 6, we show results to illustrate various entropy relationships and how these entropy relationships can be used as a visualization tool for cloud properties. Finally, in section 7, we examine the hypothetical impact of improved radiometric accuracy (0.3%, around an order of magnitude improvement from currently orbiting imagers) on retrieved cloud properties. Concluding statements are given in section 8.

2. The Theory of Generalized Inverse Problems

The mathematical theory of general stochastic inverse problems, which is similar to standard Bayesian statistical estimation theory, is used to formulate the basis of the GENRA technique as introduced by *Vukicevic et al.* [2010]. The several studies [*Vukicevic et al.*, 2010; *Coddington et al.*, 2012, 2013] that have applied GENRA to the characterization of cloud retrievals from passive shortwave (~350 to 2500 nm) remote sensing measurements were all performed for an assumed (liquid) cloud thermodynamic phase. Here we provide the mathematical theory that explicitly illustrates the utility of the GENRA algorithm when there is more than a single model that relates the measured signal to a physical quantity of interest, such as occurs when equally valid cloud reflectances occur for ice clouds and water clouds. For consistency, we adopt the notation introduced in the companion paper, *Vukicevic et al.* [2010], which is based on the formulation derived by *Mosegaard and Tarantola* [2002] and presented in *Tarantola* [2005] and *Vukicevic and Posselt* [2008].

We begin with the generalized inverse problem solution (equation (1)) [*Mosegaard and Tarantola*, 2002; *Tarantola*, 2005; *Vukicevic and Posselt*, 2008; *Vukicevic et al.*, 2010].

$$p_m(m) = \int_D \frac{1}{\gamma^*} \left[p_p(m) p_d(y) p_t(\phi(m)|m) \right] dy \tag{1}$$

The stochastic (i.e., associated with a probability density function (pdf)) solution to the generalized inverse problem is called the posterior pdf, $p_m(m)$, and it quantifies the distribution of parameters, m, based on knowledge from three sources of information: the measurements, y, a model that relates the measurements to the physical parameters of interest; $\phi(m)$; and a priori information about the parameters, if any exists. The stochastic representations of the information from the measurements (data, "d"), model (theory, "t"), and a priori ("p") information are denoted $p_d(y)$, $p_t(\phi(m)|m)$, and $p_p(m)$, respectively. An integration over the measurement space, D, removes the dependency on the observations so the posterior pdf is reported in dimensions of the parameter space, M, alone. Parameter γ^* is commonly described as a normalization constant [for example, *Vukicevic et al.*, 2010; *Coddington et al.*, 2012], and it serves dual purposes: to make the integral of the posterior pdf equal to unity over the parameter space and to ensure the property of homogeneous probability distributions in the measurement space [*Mosegaard and Tarantola*, 2002; *Tarantola*, 2005].

The role of homogeneous probability distributions is critical when making inferences from the general inverse problem solution where the representations of the parameter space are informed by more than one equally valid model solution, and those model solutions have unequal volumes in the measurement space [*Mosegaard and Tarantola*, 2002; *Tarantola*, 2005]. In equations (2)–(7), we derive the solution to the general inverse problem given in equation (1). Special attention is given to the condition of homogeneous probability distributions, and a definition of a measurement volume is provided.

Stochastic information about parameters, m, is given by the mathematical conjunction of distributions of information from a model and the observations, y, in the joint parameter space, M, and measurement space, D (i.e., the joint space, DxM as denoted in equation (2).

$$p(m,y) = \frac{1}{\gamma} \frac{p_1(m,y) p_2(m,y)}{\nu(m,y)} \tag{2}$$

In equation (2), $p_1(m, y)$ defines the joint pdf of information from the model alone, $p_2(m, y)$ defines the joint pdf given the measurements and a priori information on the parameters, $\nu(m, y)$ defines the joint

homogeneous pdf that is a pdf of unit volume in the joint *DxM* space (as demonstrated in *Tarantola* [2005]), and $p(m, y)$ is the joint posterior pdf.

The constant of proportionality, γ, forces a sum to unity in the joint posterior pdf and explicitly depends on the joint homogenous pdf as shown in equation (3).

$$\gamma = \int_{DxM} \frac{p_1(m, y) p_2(m, y)}{v(m, y)} \tag{3}$$

By applying the general statistical relationships that relate the conditional, marginal, and joint pdfs, the model-dependent joint pdf, $p_1(m, y)$, can be rewritten as shown in equation (4). In doing so, we have made the assumption that the marginal pdf in the parameter space, $p_1(m)$, is equivalent to the homogenous pdf in parameter space, $v(m, y)$, prior to conjunction and independent of prior knowledge. Note that we have also substituted the usage of $p_1(y|m)$ with the nomenclature, $p_t(\phi(m)|m)$, to denote the role of the theoretical (subscript t) mathematical model, ϕ, in relating the simulated observations, $y = \phi(m)$ to the parameters of interest, m.

$$p_1(m, y) = p_t(\phi(m)|m) v(m) \tag{4}$$

The joint pdf of the measurements and a priori information on the parameters in the absence of the model, $p_2(m, y)$, can be separated into two terms by assuming independence between the prior information in the parameters and the information in the observations (equation (5)). The subscripts, d and p, are introduced to indicate the measured data (subscript d) and a priori (subscript p) knowledge, respectively.

$$p_2(m, y) = p_d(y) p_p(m) \tag{5}$$

A similar assumption of independence is made for the joint homogeneous pdf, $v(m, y)$, allowing it to be represented as the product of two separate marginal homogeneous distributions in the measurement space and the parameter space (equation (6)).

$$v(m, y) = v(y) v(m) \tag{6}$$

In a final step, substituting equations (3)–(6) into equation (2) renders equation (1), where $\gamma^* = \gamma/v(y)$ is the normalization factor dependent upon the constant of proportionality, γ, and the marginal homogenous distribution in the measurement space, $v(y)$. The dependency upon the measurements requires that this parameter be included inside the integral shown in equation (1).

Earlier in this section, we mentioned the necessity of defining a volume in the measurement space in order to satisfy the criteria of homogeneous probability distributions; *Mosegaard and Tarantola* [2002] present the definition of a measurement volume in the general inverse problem framework. Here we use the specific example of a volume in the measurement space spanned by a grid of simulated observations where the N grid points represent specific combinations of physical parameters. In Earth remote sensing, these grids of simulated observations are commonly referred to as "look-up tables" (LUTs) and the physical parameters are "retrieved" by finding the point within the grid where the simulated observation best matches the measurement.

Each nth point (of $n = 1, 2, \ldots, N$ total grid points) within the LUT grid of parameters, m, and simulated observations, y, has a volume element of space defined by $dV(m, y) = v(m, y) \, dm \, dy$. The total volume of the LUT is obtained by computing the integral over all the simulated observations spanned by the range of parameters in the LUT: $V(m, y) = \int_{n=1}^{N} v(m, y) \, dm \, dy$. Homogeneous probability distributions are then those that ensure equal volumes for equal parameter spaces. In the example of two observational representations of the

same parameter space ($\kappa = 2$), a homogeneous probability distribution would be ensured through the use of a proportionality constant, α, defined by the following expression:

$$\alpha = \frac{V_{\kappa=1}(m, y)}{V_{\kappa=2}(m, y)}$$ (7)

In practical terms, the homogeneous probability distribution ensures that unique representations of observations that are equally valid in a physical sense are also equally weighted in a statistical sense. If there is only a single representation of the parameters, the criteria of homogeneous probability distribution are met by default and the normalization factor, γ^*, in equation (1) is simply the constant of proportionality, γ, that forces a sum to unity in the joint posterior pdf.

3. Representing Cloud Phase Discrimination as the Generalized Inverse Problem

3.1. Data

The PACE Ocean Color Instrument (OCI) is notionally a hyperspectral imager from 350 nm to more than 800 nm with six discrete shortwave spectral channels (Table 1) [*Del Castillo*, 2012]. Combinations of these channels also comprise subsets of the measurement channels used in cloud optical property retrievals (τ, r_{eff}) from the MODIS and VIIRS instruments, where the significant difference in the subsets occurs in the 2 μm window (i.e., the longest shortwave channel for MODIS is at 2135 nm and the longest shortwave channel for VIIRS is at 2250 nm).

In this study, we use the Collection 6 [*Platnick et al.*, 2017] simulated cloud reflectance data obtained with the plane-parallel discrete-ordinate radiative transfer algorithm [*Stamnes et al.*, 1988] that is used in the common MODIS/VIIRS Cloud Optical Properties product [*Platnick et al.*, 2015] to also represent simulated PACE OCI cloud reflectance measurements. The single scattering properties of the liquid phase clouds, for an assumed modified gamma droplet size distribution of spherical droplets with an effective variance of 0.1, were derived from Mie calculations [*Platnick et al.*, 2017]. For ice crystals, single scattering properties were obtained from a library of calculations based on severely roughened compact aggregates of eight solid columns [*Yang et al.*, 2013] with a gamma particle size distribution of effective variance of 0.1 [*Platnick et al.*, 2017]. The MODIS and VIIRS LUTs, separate ones for water and ice thermodynamic phase, contain the cloud reflectance as a function of spectral channel and over broad ranges in the following variables: effective radius (2–30 μm for liquid clouds and 5–60 μm for ice clouds), optical thickness (0.05 to 160 irregularly gridded, but subsequently re-gridded to a resolution of ~2), solar zenith angle, sensor zenith angle, and azimuth angle. In this study, we arbitrarily select a cosine solar zenith angle of 0.9, cosine sensor zenith angle of 0.9, and a sensor relative azimuth angle of 60° for our analysis. A black surface albedo is assumed.

3.2. Representing the pdfs for the Model, Measurement, and Prior Information

Previous studies using GENRA show that LUTs of precomputed radiative transfer calculations serve as a discretized forward model function associating the model solution in the measurement space to every value of the parameter, or combinations of parameters, in the parameter space [*Vukicevic et al.*, 2010; *Coddington et al.*, 2012, 2013]. This implies that the LUTs of cloud reflectance for $n = 1, 2, \ldots N$ combinations of τ and r_{eff} ($N = N_\tau * N_{reff}$) are used for deriving the N model pdfs, $p_t(\phi(m)|m)$, on a discrete grid of measurement values. In this work, $\kappa = 2$ unique LUTs map two simulated measurements of liquid and ice cloud reflectance to a common point on the grid of discrete values in the parameter space.

Each of the N model pdfs represents a distribution of model uncertainty in the measurement space that results, in general, from a combination of model structural deficiencies and uncertainty in model ancillary parameters. The model structural deficiencies are typically associated with approximations used when deriving the theoretical model equations and with a method of solving these equations numerically. Both systematic and random errors could result from these deficiencies, and they could be represented stochastically for the purpose of solving the parameter estimation problem expressed in equation (1).

The ancillary parameters are essential forward model inputs but are not retrieved parameters. The choices for these ancillary parameters, and how well they represent true conditions or the variability in the true

conditions, leads to uncertainties in the model results. Some examples of ancillary forward model inputs that affect the simulation of cloud reflectance include the surface albedo that determines the proportion of incident light that is reflected by a surface, the vertical profile of atmospheric molecular gases, the assumed size distribution of cloud particles in the liquid and ice cloud models, and the assumed crystal habitat in the ice cloud model. Therefore, the distributions of the model solutions in the measurement space are a statistical representation of the uncertainty in the inputs to the forward model and the model structural deficiencies. As in the previous studies we assume a relatively simple Gaussian-distributed and wavelength-independent model uncertainty of 2%, which is reasonable for establishing a baseline and for ocean albedo surfaces, but is not generally true [see, for example *Platnick et al.*, 2017; *Coddington et al.*, 2012]. MODIS cloud retrieval uncertainties due to errors in the effective variance of liquid and ice particle size distributions are on the order of 2% [*Platnick et al.*, 2017]. Estimating and using a more sophisticated stochastic representation of the model uncertainty is beyond the scope of this study.

The pdf of the stochastic measurement, $p_d(y)$, is the probability of a measurement taking discrete values between y_i and $y_i + \Delta y$ that span a range of values given the measurement uncertainty that is described by random and systematic errors. In this study, the quantity used in the actual retrieval is cloud reflectance, derived from observations of reflected cloud radiance, which is a function of satellite viewing and Sun zenith angles, and normalized by the measured downwelling irradiance. When defining $p_d(y)$ we assume that the cloud reflectance has Gaussian-distributed and wavelength-independent random errors of 3%; this assumption is based on characterization of the MODIS and VIIRS instruments and on-orbit performance monitoring [e.g., *Xiong et al.*, 2016; *Uprety and Cao*, 2015; *Xiong et al.*, 2014]. The measurement pdf is defined on the same discrete grid of measurement values as the model pdfs described above.

The pdf of prior information in the τ and r_{eff} parameters, $p_p(m)$, represents probabilities of the parameters taking values between m_k and $m_k + \Delta m$ from a range of physically plausible values that are shared between the $\kappa = 2$ representation cloud reflectance LUT's simulated using the ice cloud and liquid cloud models. In this study, we define the range of physically plausible values that are shared by ice clouds and liquid clouds as optical thickness spanning $\tau = 0.5$ to 160 and effective radius spanning $r_{eff} = 5$ to 30 μm. We are guided by statistics of global cloud properties [e.g., *King et al.*, 2013; *Platnick et al.*, 2017] when making a priori assumptions that r_{eff} values less than 5 μm occur only with liquid clouds and r_{eff} values greater than 30 μm occur only with ice clouds. In the absence of other information for the shared parameter space, the prior pdf can take uniform values (i.e., all of the values of the shared parameter range are a priori equally likely). This condition could be improved if additional information would become available from other independent measurements.

Δy corresponds to a unit discretization in the measurement space and can be interpreted as the minimum measurement error. Likewise, Δm is a unit interval in the parameter space that can be interpreted as the maximum retrieval precision [*Vukicevic et al.*, 2010; *Coddington et al.*, 2012].

3.3. Computing the Likelihood Function

The likelihood function is the probability of the observations as a function of the retrieval parameters and provides a metric of how well particular choices of model parameters describe the data [*Tarantola*, 2005]. As shown in equation (8a), for every n grid points (of N total), a convolution (i.e., a pointwise multiplication) of the nth model pdf and the measurement pdf on the discrete grid of measurement values is obtained. The convolutions are performed separately for each of the $\kappa = 2$ representations of the model pdfs.

$$p_{\text{likelihood}}^n(\kappa, m) = \int_D \frac{1}{\gamma^*} \left[p_d(y) p_t^n(\phi(m)|m) \right] dy \tag{8a}$$

We enforce the criteria for homogeneous probability functions (see equation (7)) and compute the respective volumes of the measurement space for all κ groups of likelihood functions. These volumes are then used to derive the factor, α (equation (8b), that ensures that equally valid physical representations of the model pdfs are equally weighted statistically (equation (8c)). It is acceptable to switch the numerator and denominator in

equation (8b). However, if doing so, the normalization applied in equation (8c) would then need to be applied to the $\kappa = 2$ representation instead.

$$\alpha = \frac{\sum_n \sum_y p_{\text{likelihood}}^n (\kappa = 1, m)}{\sum_n \sum_y p_{\text{likelihood}}^n (\kappa = 2, m)} \tag{8b}$$

$$p_{\text{likelihood}}(\kappa = 1, m) = \frac{p_{\text{likelihood}}(\kappa = 1, m)}{\alpha} \tag{8c}$$

3.4. Computing the Posterior Retrieval pdf

In the final step, the multiplication of the likelihood function of the homogeneous probability functions derived in equations (8a)–(8c) with the pdf of prior information about the parameters forms the posterior pdf (equation (9)). The multiplication is performed for each of the n grid points, separately for each κ, to represent the κ unique representations of the likelihood function and prior information statistics. The posterior pdf is the two-dimensional (2-D) map of probabilities in the optical thickness and effective radius parameter space, m, for each of the κ cloud thermodynamic phase possibilities. The normalization constant, γ, is used to make the integral over all dimensions of the posterior pdf space equal to unity.

$$p_m^n(\kappa, m) = \frac{1}{\gamma} p_p^n(\kappa, m) p_{\text{likelihood}}^n(\kappa, m), n \in (1, N) \tag{9}$$

The steps described in sections 3.2–3.4 are iterated for each measurement used in the retrieval. To quantify the cumulative effect of the measurements at all retrieval wavelengths, the prior pdf (beginning with the measurement at the second retrieval wavelength) would be serially updated by using the posterior pdf for the measurement at the previous retrieval wavelength introduced into the algorithm. For the measurement at the first retrieval wavelength, the prior pdf is assumed to be uniform, which means that the prior pdf is weighted equally for all physically plausible values. For independent measurement pdfs, the cumulative result from the serial processing described above would be no different from that of a batch-style processing where multiple measurements are simultaneously used to update the posterior estimate. GENRA can also be applied in an alternative approach to characterize the retrieval at each individual retrieval wavelength as opposed to the cumulative impact described above. This latter approach requires that the prior pdf is ascribed a uniform distribution at each retrieval wavelength introduced into the algorithm (i.e., the prior pdf at a subsequent iteration is not updated using the posterior pdf from the former iteration). Examples of both treatments of the prior pdf for characterizing passive shortwave cloud retrievals are shown in *Vukicevic et al.* [2010] and *Coddington et al.* [2012, 2013].

4. Characterizing Cloud Phase Discrimination Using the Posterior Retrieval pdf

The information about the possible discrete values of cloud optical thickness, droplet effective radius, and cloud thermodynamic phase contained in the posterior pdf can be used to characterize the cloud property retrievals. In this section, we discuss several standard retrieval diagnostics derived from the posterior pdf. These include the marginal probability distributions and maximum a posteriori values of the parameters and the Shannon information content [*Shannon and Weaver*, 1949] of the measurements. We also discuss two additional information content metrics that, to our knowledge, have not been previously applied to the study of cloud optical properties. These include the mutual and conditional information contents [*Cover and Thomas*, 2006; *Wang and Shen*, 2011] that, respectively, quantify the information in a measurement that is shared between parameters and the information in a measurement that remains in one parameter given complete knowledge of information in another parameter.

1. *The marginal pdfs for each parameter, the mean values, and associated error variance (i.e., retrieval precision) statistics.* The marginal pdfs are obtained by integrating the posterior pdf (equation (9)) over the parameter space. When the integration is performed over the parameter space for all $\kappa = 2$ cloud thermodynamic phase possibilities, the resulting marginal pdfs in optical thickness (equation (10a)) and effective radius (equation (10b)) represent the error variances in the retrieval parameters for the joint parameter space spanned by the liquid and ice cloud reflectances. Performing the integration over the parameter space *separately* for each cloud thermodynamic phase results in marginal pdfs in τ (equation (10c)) and

r_{eff} (equation (10d)) that represent the error variance in the retrieval parameters for each specific cloud phase alone. The mathematical sum of the marginal pdfs for a specific parameter for each κ cloud thermodynamic phase, for example, optical thickness (equation (10c)), is equivalent to the marginal pdf in optical thickness over the joint cloud thermodynamic phase space (equation (10a)).

$$p(\tau = 1, \ldots N_\tau) = \int_{\kappa=1}^{2} \int_{r_{eff}=1}^{N_{reff}} p_m^n(\kappa, \tau, r_{eff}) dr d\kappa \qquad (10a)$$

$$p(r_{eff} = 1, \ldots N_{reff}) = \int_{\kappa=1}^{2} \int_{\tau=1}^{N_\tau} p_m^n(\kappa, \tau, r_{eff}) d\tau d\kappa \qquad (10b)$$

$$p(\kappa, \tau = 1, \ldots N_\tau) = \int_{r_{eff}=1}^{N_{reff}} p_m^n(\kappa, \tau, r_{eff}) dr \qquad (10c)$$

$$p(\kappa, r_{eff} = 1, \ldots N_{reff}) = \int_{\tau=1}^{N_\tau} p_m^n(\kappa, \tau, r_{eff}) d\tau \qquad (10d)$$

The marginal pdf of cloud thermodynamic phase (equation (10e)) is obtained by integrating the posterior pdf over the space spanned by the parameter ranges in optical thickness and effective radius.

$$p(\kappa) = \int_{\tau=1}^{N_\tau} \int_{r_{eff}=1}^{N_{reff}} p_m^n(\kappa, \tau, r_{eff}) dr d\tau \qquad (10e)$$

The probability of cloud phase discrimination (equations (11a) and (11b)) is then the percent contribution of the marginal pdf for all thermodynamic phase possibilities that is explained by each of the respective κ cloud thermodynamic phase possibilities.

$$\text{probability}_{k=1}(\%) = \frac{p(\kappa=1)}{\sum_{\kappa=1}^{2} p(\kappa)} \times 100 \qquad (11a)$$

$$\text{probability}_{k=2}(\%) = \frac{p(\kappa=2)}{\sum_{\kappa=1}^{2} p(\kappa)} \times 100 \qquad (11b)$$

The statistical mean of the respective marginal pdfs can be used to represent the retrieved cloud properties. However, only when the posterior pdf is symmetrical (e.g., Gaussian distributed) will the statistical mean of the marginal pdfs be equivalent to the maximum a posteriori solution of the retrieval (discussed next). Prior studies of cloud property retrievals using general inverse theory have shown that Gaussian assumptions in the posterior pdf are not valid for regions of the parameter space where the forward model is nonlinear (i.e., the reflectance is nonlinearly related to the parameters τ and r_{eff}).

Higher-order statistics that are useful numerical metrics of the central tendency, degree of variation, and the balance of the distribution of the parameters around the center value are key strengths of using general inverse theory approaches. These metrics are computed from the joint and marginal probability distributions and are discussed in standard statistical textbooks, for example, *Wilks* [2011]. For example, the interquartile range (IQR), defined as the difference between the upper and lower quartiles (where a quartile is the midpoint between the median and the upper and lower extremes of the distribution) of the marginal pdfs, is the most common metric of dispersion in the retrieved parameters [*Wilks*, 2011], and larger values of IQR reflect a greater spread in the middle half of the data. Skewness measures a lack of symmetry in the distribution or parameter values, and a positive skewness, for example, indicates a distribution with a long (right) tail,

whereas a distribution with zero skewness is symmetric around the central value. Lastly, the joint and marginal pdfs provide key information on multimodal solutions where a nonunique relationship exists between the parameters and the observations. Skewed distributions and nonunique solutions have been observed in distributions of cloud optical properties [*Coddington et al.*, 2013] and cloud microphysical properties [*Posselt and Vukicevic*, 2010; *Posselt*, 2016].

2. *Maximum a posteriori estimate of the retrieval*. The maximum a posteriori value is the most likely value of the parameters and occurs at the maximum value of the posterior pdf,

$$(\tau^*, r_{eff}^*), \max\left[p_m^n(\kappa^*, \tau, r_{eff})\right] = p_m^n(\tau^*, r_{eff}^*)$$

where κ^* is the thermodynamic phase with the highest probability of discrimination as defined by equations (11a) and (11b). The maximum a posteriori estimate for the cloud thermodynamic phase that does not have the highest probability is also of great interest for the retrieval of cloud properties as it identifies the (τ, r_{eff}) retrieval solution when thermodynamic phase has been incorrectly identified.

3. *The Shannon information content of the measurements*. Shannon information [*Shannon and Weaver*, 1949] is the measure of information gained by making a measurement, and it is derived from the measure of entropy (disorder), H. The entropy can be computed from the joint posterior pdf (equation (12a)) and the marginal pdfs (equations (12b) and (12c)). High values of entropy equate to high levels of disorder where many parameter values are equally likely in a retrieval. Conversely, low entropy equates to low disorder, indicating that fewer parameter values are likely in a retrieval. Since we use the logarithm with base 2, the units of information are in bits.

$$H_\kappa(\tau, r_{eff}) = -\sum_{n=1}^{N=N_\tau \times N_{reff}} p_m^n(\tau, r_{eff}) log_2 p_m^n(\tau, r_{eff}) \tag{12a}$$

$$H_\kappa(\tau) = -\sum_{n=1}^{N_\tau} p^n(\tau) log_2 p^n(\tau) \tag{12b}$$

$$H_\kappa(r_{eff}) = -\sum_{n=1}^{N_{reff}} p(r_{eff}) log_2 p(r_{eff}) \tag{12c}$$

The Shannon information content, *SIC* (equation (13)), is inversely related to entropy. As entropy decreases, the Shannon information increases indicating increased retrieval precision. In equation (13), H_{prior} is the entropy of the prior pdf, p_{prior}^n, and H_{post} is the entropy of the posterior pdf, p_m^n. Using the joint representation of the posterior pdf and prior pdf when computing the SIC characterizes the information content of the retrieval for all possible combinations of (τ, r_{eff}) in each respective thermodynamic phase. Alternatively, using the marginal posterior pdfs of the parameters and marginal prior pdfs characterizes the information content in optical thickness separately from the information content in effective radius, for each respective thermodynamic phase.

$$SIC_\kappa = H_{prior} - H_{post} \tag{13}$$

4. *The conditional information of the measurements*. A single measurement may provide information about more than one parameter and the uncertainty (i.e., entropy) in the parameters can be quantified in different ways. The "conditional" entropy is the entropy of a parameter that remains (after making a measurement) when additional information is incorporated to give complete knowledge of another parameter [*Cover and Thomas*, 2006; *Wang and Shen*, 2011]. The complete knowledge of the other parameter could be obtained in different ways depending upon the application. For example, the necessary additional

information could come from in situ data, a retrieval from another platform, or making an assumption about the parameter's value. The conditional entropy of optical thickness, $H(\tau|r_{eff})$ (equation (14a)), is the entropy when conditioned on knowledge of effective radius averaged over all possible values that effective radius may take in the parameter range. The conditional entropy of effective radius $H(r_{eff}|\tau)$ can be similarly defined (equation (14b)). High values of conditional entropy represent large remaining uncertainty (i.e., low precision) in a parameter despite complete knowledge of another, correlated, parameter. Conversely, low values of conditional entropy represent that complete knowledge in the second, correlated parameter has reduced the uncertainty (i.e., higher precision) in the first parameter. In the limiting condition where one parameter is completely determined by another parameter, the conditional entropy of the first parameter is zero.

$$H_k(\tau|r_{eff}) = \sum_{n=1}^{N} p_m^n(\tau, r_{eff}) log_2 p_m^n(\tau|r_{eff}) \tag{14a}$$

$$H_k(r_{eff}|\tau) = \sum_{n=1}^{N} p_m^n(\tau, r_{eff}) log_2 p_m^n(r_{eff}|\tau) \tag{14b}$$

Through applying the relationships that relate the conditional, marginal, and joint pdfs, the conditional entropy can be equated to the difference of the Shannon entropy of the parameter (equation (12b) or (12c)) and the entropy in the measurement shared by the parameters (i.e., the mutual entropy; see next subsection, v) [for example, *Wang and Shen*, 2011].

We define the conditional information content, *CIC*, as the change in conditional entropy in the posterior pdf relative to a prior state (equations (15a) and (15b)). By this definition, the conditional information content is inversely related to conditional entropy similar to how the Shannon information content is inversely related to entropy. A reduction in conditional entropy in the posterior pdf relative to a prior state represents that ancillary knowledge in a parameter has reduced the uncertainty in the other parameter (i.e., the conditioning of one parameter effects the probability of another) and the conditional information content has increased.

$$CIC_\kappa(\tau) = H_k(\tau|r_{eff})_{prior} - H_k(\tau|r_{eff})_{post} \tag{15a}$$

$$CIC_\kappa(r_{eff}) = H_k(r_{eff}|\tau)_{prior} - H_k(r_{eff}|\tau)_{post} \tag{15b}$$

5. *The mutual information of the measurements*. The mutual entropy, *I*, quantifies how much of the information in a parameter is conveyed by another parameter [*Cover and Thomas*, 2006; *Wang and Shen*, 2011]; it is therefore a measure of how two parameters share the information from a single measurement. It is equivalent to the relative entropy (equivalent to the Kullback-Leibler distance [*Cover and Thomas*, 2006]) of the joint pdf and the product of the marginal distributions of the parameters as shown in equation (16). In the limiting condition of complete independence between the parameters, $p_m^n(\tau, r_{eff}) = p(\tau)p(r_{eff})$ and the mutual entropy is zero representing that knowledge in optical thickness does not give any information in effective radius and vice versa. When one parameter is completely determined by a second parameter, the conditional entropy of the first parameter is zero and, by extension, the mutual entropy between the parameters is a theoretical maximum defined by the entropy of the first variable alone.

$$I_\kappa(\tau; r_{eff}) = \sum_{n=1}^{N} p_m^n(\tau, r_{eff}) log_2 \frac{p_m^n(\tau, r_{eff})}{p(\tau)p(r_{eff})} \tag{16}$$

Small values of mutual entropy indicate little shared information (i.e., dependencies or correlations) between the parameters and, therefore, only a small potential for reducing retrieval uncertainty in one parameter by

gaining knowledge in another parameter. Conversely, high values of mutual entropy indicate a greater degree of shared information and larger dependencies among the parameters and, therefore, a correspondingly larger potential for reducing retrieval uncertainties in one parameter through ancillary knowledge of another parameter. Examples of ways to gain ancillary knowledge in the second parameter include the use of independent measurements or retrievals, and the making of retrieval assumptions.

We define the mutual information content, *MIC*, as the change in mutual entropy in the posterior pdf relative to a prior state (equation (17)). By this definition, increasingly positive values of the mutual information content represent an increase in shared information between the parameters by the act of making a measurement and vice versa for decreasing values of the mutual information content. In the absence of prior knowledge, we assume complete independence in the parameters in which case the mutual entropy, I_{prior}, is equal to zero by the properties of the joint distribution for independent variables.

$$MIC_\kappa = -\left[I_{prior} - I_{post}\right] \tag{17}$$

The Shannon information, mutual information, and conditional information are computed separately for each κ thermodynamic phase, and the results for the κ^* thermodynamic phase with the highest probability of discrimination would represent the respective information content metrics in τ and r_{eff} that correspond to the maximally likely retrieval. In all entropy definitions shown in this section, we have used logarithms of base 2; therefore, the resulting unit of entropy is bits.

5. The Probability of Retrieving the Correct Thermodynamic Phase

Here we provide the results from experiments that quantify the information content in optical thickness, effective radius, and thermodynamic phase from observations of simulated shortwave cloud reflectance data. The results are presented for specific cloud (τ, r_{eff}) pairs and over a broad range in cloud τ and r_{eff} using different combinations of measurement channels that are used in the operational cloud retrievals algorithms by the MODIS and VIIRS instruments and have been identified for operational cloud retrievals in the conceptual instrument design study for the future PACE imager. Since the Collection 6 MODIS and similar VIIRS cloud retrieval algorithms also incorporate information at infrared (IR) channels, we refer to our experiments using the channel combination of 865 nm, 1640, and 2130 nm as "MODIS-SW," where the "SW" refers to "shortwave." Similarly, we refer to our experiments with the channel combinations of 865 nm, 1640, and 2225 nm as "VIIRS-SW." Since the PACE imager will not have an IR sensor, experiments using the channel combinations of 865 nm, 1640 nm, 2130 nm, and 2225 nm are simply referred to as "PACE."

Unless specified otherwise, all results assume a black surface albedo, a wavelength-independent measurement uncertainty of 3%, a model uncertainty of 2%, a cosine of the sensor zenith angle of 0.9, a cosine of the solar zenith angle of 0.9, an azimuthal angle of 60°, and vertically and horizontally homogeneous clouds. By maintaining this consistency across experiments, the following results quantify the impacts of measurement channel location and number on the information content of cloud optical thickness, effective radius, and thermodynamic phase from current and future passive imagers used to provide the global record of cloud properties.

5.1. The Impacts of Number and Location of Measurement Channels

The 2-D posterior joint pdfs that correspond to the final retrieval with all wavelengths of simulated cloud reflectance measurements are shown in Figure 1 (top row) for a "true" cloud type of $\tau = 10$, $r_{eff} = 10$ μm, and phase = liquid. This particular example was chosen because at moderate cloud optical thickness values and larger ($\tau \geq 10$) passive methods to retrieve cloud optical properties have demonstrated accurate performance [*Platnick et al.*, 2017]. The results show that the posterior retrieval pdfs are not Gaussian and that there are overlapping contributions to the joint posterior pdf from the solution space identified as liquid water phase (identified by blue-green contours) and as ice phase (pink-purple contours). The percentage to which the total probability of the liquid phase solution space contributes to the total probability of the joint phase solution space quantifies the percent probability of a liquid phase retrieval. The percent probability of correct phase discrimination for this example is 35% for MODIS-SW, 63% for VIIRS-SW, and 70% for PACE. Values around 50% theoretically represent an ambiguous phase retrieval because the measurement and model
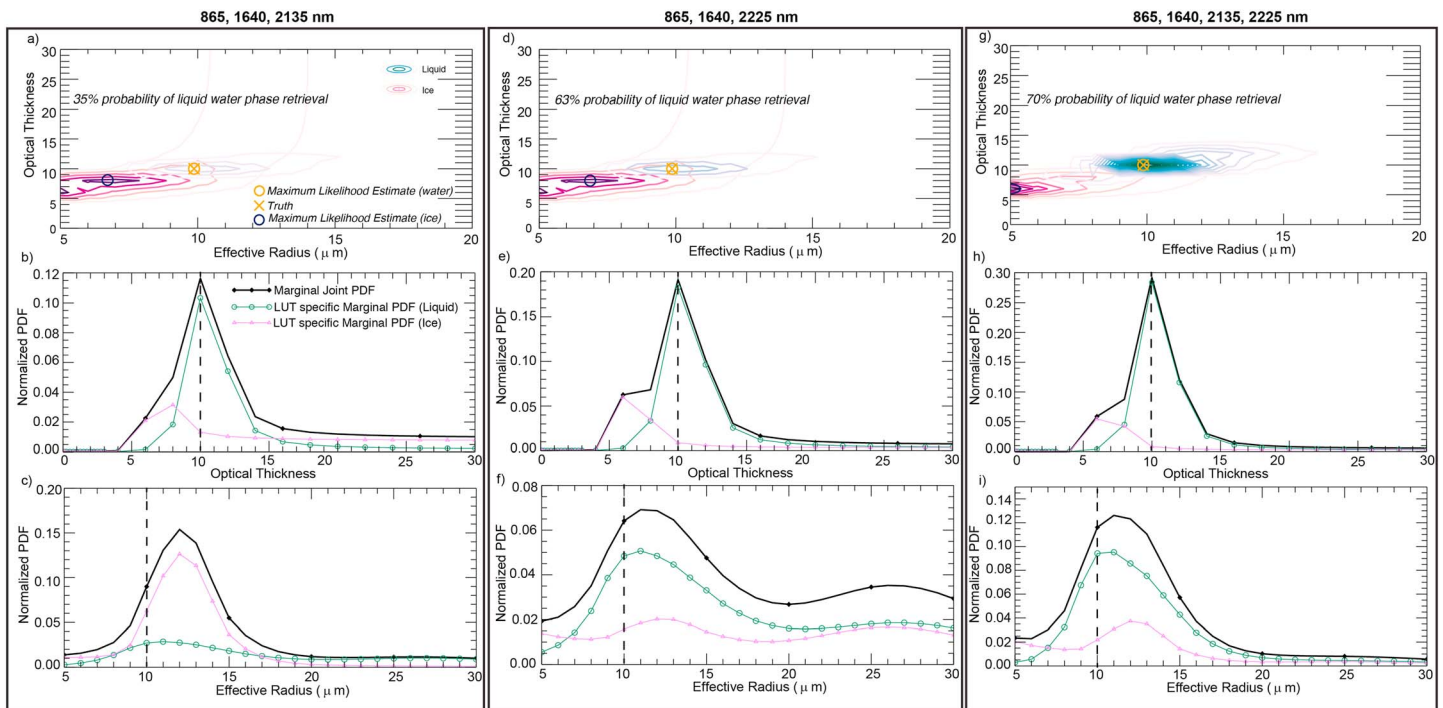
Figure 1. The final retrieval results after cumulatively ingesting information from all retrieval wavelengths into the GENRA algorithm for a "true" cloud type of $\tau = 10$, $r_{eff} = 10$ μm, and phase = liquid. (left column) Results specific to "MODIS-SW" cloud retrieval channels, (middle column) "VIIRS-SW" results, and (right column) "PACE" results (see text). (a, d, and g) The 2-D joint posterior pdf showing contributions from ice thermodynamic phase (pink contours) and liquid thermodynamic phase (blue contours). The marginal pdfs for (b, e, and h) optical thickness and (c, f, and i) effective radius, respectively, where the joint marginal pdf (black) has contributions from liquid (green) and ice (pink) thermodynamic phase. The vertical dashed lines on marginal pdf plots denote the "truth" values.

uncertainties assumed in this study are idealized uncertainties that may overestimate or underestimate the true uncertainties of a specific atmospheric state and measurement conditions. Identifying the bounds of this range is left to future work. The maximum a posteriori estimates of cloud optical thickness and effective radius for both cloud phases are annotated on the plot. The results show, given correct phase identification, that the maximum a posteriori estimates for the liquid cloud phase are centered on the "truth" for MODIS-SW, VIIRS-SW, and PACE channel combinations, which indicates a nonbiased (i.e., accurate) retrieval solution for the given simulated measurement conditions. This result is expected for the experiments with the simulated truth and verifies the accuracy of the numerical procedure in GENRA. Evaluating retrieval bias (i.e., a departure of the maximum a posteriori estimate away from the true value) is only possible when the truth is known, for example, as occurs in a theoretical study like this one or when other, independent measurements can be used to inform the truth [*Vukicevic et al.*, 2010]. In all channel combinations, inaccurately identifying cloud phase as ice would result in retrieved properties for an optically thinner ice cloud of smaller droplet size: $\tau = 8$, $r_{eff} = 7$ μm for MODIS-SW and $\tau = 6$, $r_{eff} = 5$ μm for VIIRS-SW and PACE.

The middle and bottom rows of Figure 1 show sequences of marginal pdfs of optical thickness and effective radius, respectively. "Marginal" distributions are those from a subset of the variables. For example, the "joint" marginal of τ or r_{eff} is the distribution in this respective parameter given both cloud phase solution spaces (i.e., liquid and ice). The marginal pdfs in τ or r_{eff} then further subset the respective joint marginal pdf into the distribution for a single cloud phase solution space (i.e., liquid or ice). Obtaining the marginal pdf for a parameter requires an integration over the other parameters (equations (10a)–(10a)). When the 2-D joint posterior pdf departs from a Gaussian distribution there is a nonlinear coupling between cloud optical thickness and droplet effective radius, which leads to a discrepancy due to an artifact of integration when interpreting the maximum a posteriori values obtained from the joint marginal (or marginal) pdfs relative to the 2-D joint marginal pdf [see also *Posselt*, 2016, Table 4]. The degree of discrepancy will depend upon the shape of the 2-D joint posterior pdf [*Coddington et al.*, 2013]. For these reasons, in this work we consider the "real" solution
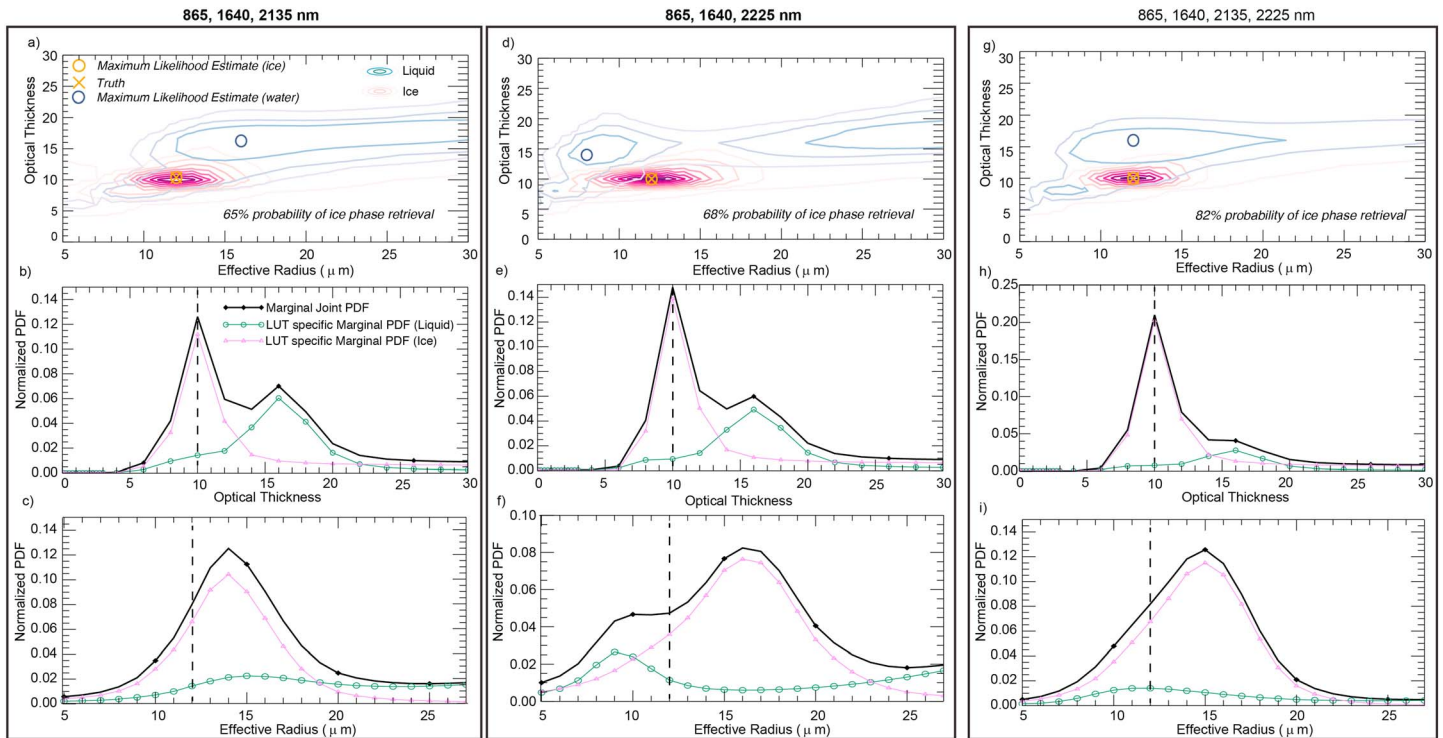
Figure 2. Same as in Figure 1 but for a "true" cloud type of of $\tau = 10$, $r_{eff} = 12$ μm, and phase = ice.

to be the 2-D joint posterior pdf. The 1-D pdf provides context and the relative contributions from the different thermodynamic phase spaces.

The middle row of Figure 1 shows the sequence of marginal pdfs of optical thickness for the MODIS-SW, VIIRS-SW, and PACE measurement channel combinations. The joint marginal pdf of optical thickness (equation (10a)) is in black, and the contributions from the liquid and ice solution spaces (equation (10c)) are shown in green and pink, respectively. The results show that the joint pdf in optical thickness is centered at the true value of $\tau = 10$ and that contributions from the liquid cloud phase solution space explain the majority of the joint marginal distribution. The contributions from the ice phase solution space broaden the joint marginal distribution to smaller optical thickness values.

The bottom row in Figure 1 is the sequence of marginal pdfs of droplet effective radius for the varying channel combinations. Here the results are more diverse. For the MODIS-SW solution, the dominant contribution to the joint marginal pdf of effective radius (equation (10b)) comes from the ice solution space, while the dominant contribution comes from the liquid solution space for the VIIRS-SW and PACE (equation (10d)). There is also more diversity in the distribution shape. For all measurement channel combinations, the peaks of the marginal pdfs (joint and LUT-specific) are biased away from the true solution of $r_{eff} = 10$ μm, nor are they centered on the maximum a posteriori values of effective radius in the 2-D joint posterior pdf just identified for the ice phase ($r_{eff} = 7$ μm for MODIS-SW or $r_{eff} = 5$ μm for VIIRS-SW and PACE). This is a consequence of the integration over τ (equation (10d)) for the LUT-specific marginal pdfs and over τ and phase for the joint marginal pdf.

Figure 2 also shows the impacts of retrieval channel number and location on the probability of retrieving the correct thermodynamic phase but for a true cloud type of $\tau = 10$, $r_{eff} = 12$ μm, and phase = ice. The 2-D posterior joint pdfs that correspond to the final retrieval with all wavelengths of simulated cloud reflectance measurements are shown in Figure 2 (top row). The 2-D joint posterior pdfs are not Gaussian-distributed and provide evidence of overlapping contributions to the joint posterior pdf from both ice and liquid water phase. The percentage to which the total probability of the ice phase solution space contributes to the total probability of the joint phase solution space quantifies the percent probability of an ice phase retrieval. The
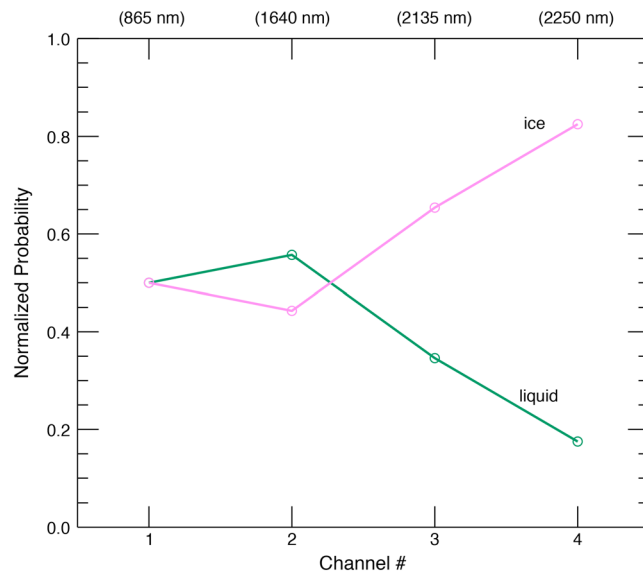
Figure 3. Cumulative probability for correctly discriminating ice cloud phase after ingesting information from the "PACE" channel set combination into the GENRA algorithm for a "true" cloud type of $\tau = 10$, $r_{eff} = 12$ μm, and phase = ice. Results at final measurement channel correspond to the percent probability of ice phase retrieval reported in Figure 2g.

percent probability of correct phase discrimination for this example is 65% for MODIS-SW, 68% for VIIRS-SW, and 82% for PACE. As for the liquid cloud case, the maximum a posteriori estimates for the ice cloud phase are centered on the truth for MODIS-SW, VIIRS-SW, and PACE channel combinations, which indicates a nonbiased (i.e., accurate) retrieval solution for the given simulated measurement conditions. Inaccurately identifying cloud phase, however, would result in very different retrieved properties for different channel combinations: an optically thicker liquid cloud of larger droplet size for MODIS-SW ($\tau = 16$, $r_{eff} = 16$ μm), an optically thicker liquid cloud of smaller droplet size for VIIRS-SW ($\tau = 14$, $r_{eff} = 8$ μm), and an optically thicker cloud of the same particle size for PACE ($\tau = 16$, $r_{eff} = 12$ μm).

The middle and bottom rows of Figure 2 show the marginal pdfs in optical thickness and droplet effective radius, respectively, for the varying channel combinations. In all cases, the contribution to the joint marginal pdfs are dominated by the ice phase, which is represented in the percent probabilities of ice phase discrimination that exceed 50% as discussed in the preceding paragraph. The contributions from the liquid phase solution space broaden the joint marginal distribution in optical thickness to larger optical thickness values for all channel combinations. For droplet effective radius, the contributions from the liquid phase solution space change the peak and breadth of the joint marginal distribution.

5.2. The Probability of Thermodynamic Phase Discrimination for MODIS, VIIRS, and PACE Over a Broad Range in Cloud Optical Thickness and Droplet Effective Radius

In this section, we extend the results from section 5.1 and Figures 1 and 2 using the same experimental setup (i.e., surface conditions, atmospheric state, and solar and sensor geometries) to a broad range of cloud optical thickness values (0.05 to 160) and droplet effective radius (5 μm to 30 μm) values, which notionally encompasses the full shared parameter space where reflectance values are equally plausible from liquid or ice cloud thermodynamic phase. Figure 3 shows the wavelength-dependent contributions to the cumulative probability of phase retrieval after ingesting information from the PACE measurement channel set into the GENRA algorithm for a thin ($\tau = 10$) ice cloud with $r_{eff} = 12$ μm. Note that, if the experiment was repeated for the same τ and r_{eff} value, but for a liquid water cloud, the probability of correctly discriminating liquid phase would not be a symmetrically reversed value. This is because the likelihood function (equation (8a)) is distributed around a measurement value that represents a choice of parameters in the liquid or ice cloud phase and encompasses a range of possible cloud parameters where the relative magnitude of each respective "solution" enveloped within the likelihood function is directly proportional to the degree of overlap between the measurement pdf and the model pdf solutions. Since clouds scatter and absorb radiation differently for liquid and ice phases, the measurement pdf, and therefore the likelihood function, for liquid or ice phase will encompass a somewhat different subset of possible model solutions.

The extension of similar analysis to that shown in Figure 3 to a broad cloud parameter space and for the final phase probability solution at the longest wavelength ingested into the GENRA algorithm are shown in Figures 4 and 5. Figure 4 is a contour plot of the total probability of the "correct" liquid phase solution
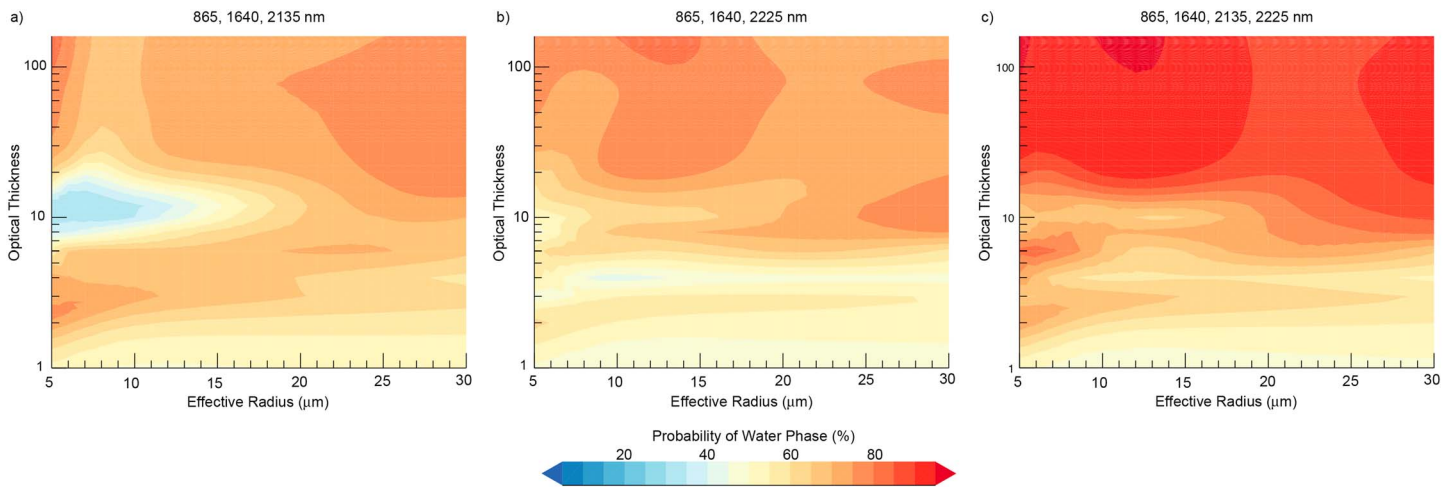
Figure 4. Contour plot of the percent probability of correctly retrieving liquid water cloud phase from the joint space spanned by ice and liquid phase solutions when the "true" cloud phase is liquid. Values around 50% indicate an ambiguous phase retrieval (see text). The subplots are specific to specific measurement channel combinations: (a) MODIS-SW, (b) VIIRS-SW, and (c) PACE.

respective to the total probability of the joint liquid and ice phase solution space. Figure 5 depicts identical results for the cases where the correct solution is the ice phase.

The results in Figures 4 and 5 show that the PACE channel combination provides significantly improved thermodynamic phase discrimination than either the MODIS-SW or VIIRS-SW channel combinations, in particular, for $\tau > 10$ and/or larger r_{eff}. The VIIRS-SW channel combination provides minor improvements for thermodynamic phase discrimination relative to MODIS-SW at liquid cloud τ values of approximately 10, ice cloud τ values of approximately 4, and r_{eff} values of 10 μm or smaller. Correctly retrieving cloud thermodynamic phase for optically thin clouds ($\tau < 10$ for liquid and $\tau < 4$ for ice) will remain problematic for any of the channel combinations shown.

Having both channels near 2 μm on the PACE imager provides the additional benefit of allowing continuity with the MODIS and VIIRS cloud data records, which use a combination of spectral channels including 660 nm, 865 nm, 1200 nm, 1640 nm, and either 2130 nm (from MODIS), or 2225 nm (from VIIRS). Future work will link these phase discrimination results to ongoing MODO6 MODIS/VIIRS uncertainty assessments [*Platnick et al.*, 2004, 2017].
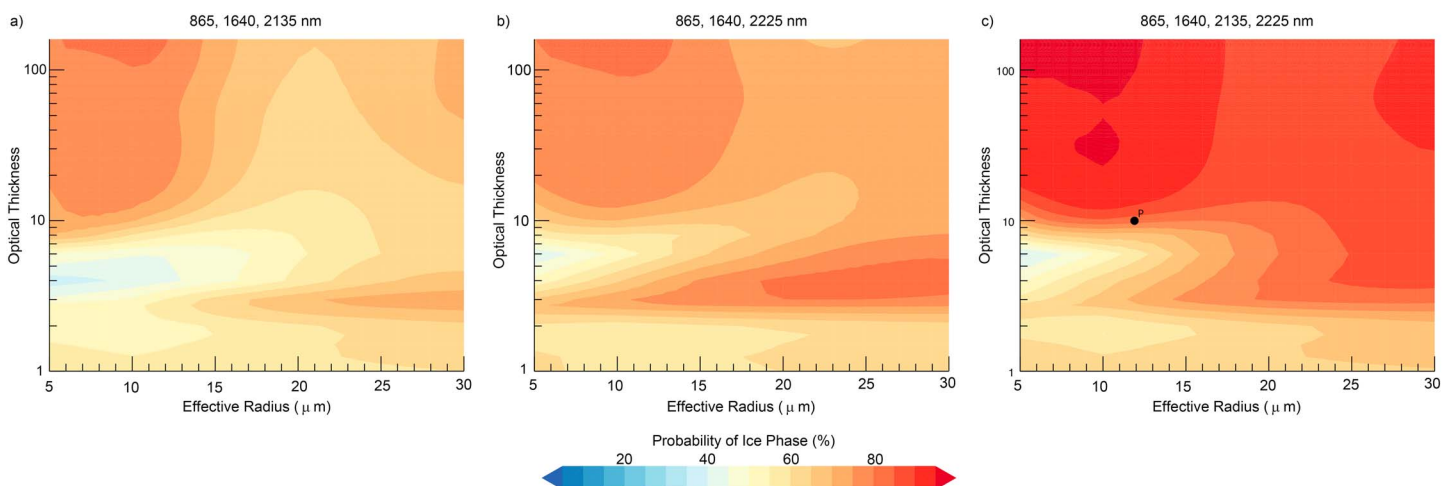

Figure 5. Same as in Figure 4 but for the ice thermodynamic phase. In Figure 5c, the black point "*P*" represents the (τ, r_{eff}) pair discussed in Figure 3.
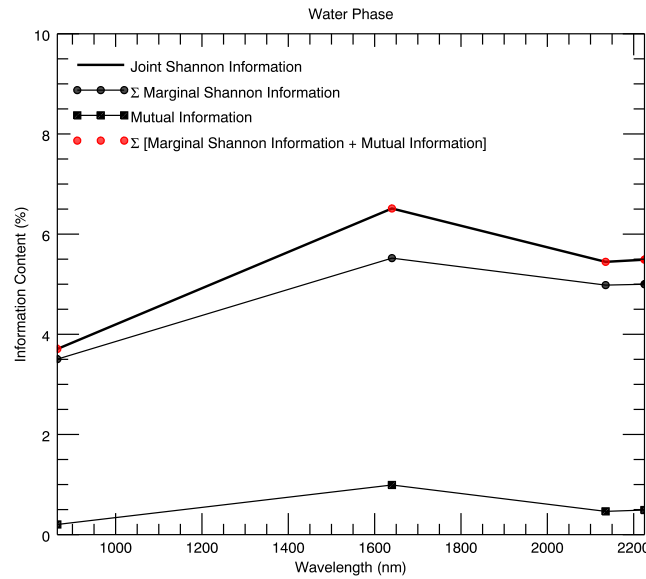
Figure 6. The normalized total (black line), marginal (dashed line with circles), and mutual information content (dashed line with squares) derived from the entropy relationships in the (τ, r_{eff}) cloud optical parameters after ingesting simulated cloud reflectance at 865 nm, 1640 nm, 2135 nm, and 2225 nm into the GENRA algorithm (with no update of the prior). The results are specific to the same cloud case described for Figure 1.

6. Entropy Relationships in Cloud Optical Properties

In section 4 and the supporting information, we described the mathematical relationships between the joint, marginal, conditional, and mutual entropy metrics. Here we graphically illustrate the relationships that occur when a measurement provides physical insight into more than one cloud optical parameter by using the total, shared, and conditional information content metrics. For example, when a particular spectral measurement provides information for both τ and r_{eff} parameters, we quantify the "shared" information in the measurement by using the mutual information content (section 4.v). Going one step further, we theoretically explore how we can exploit additional information to uniquely constrain one of the parameters (fr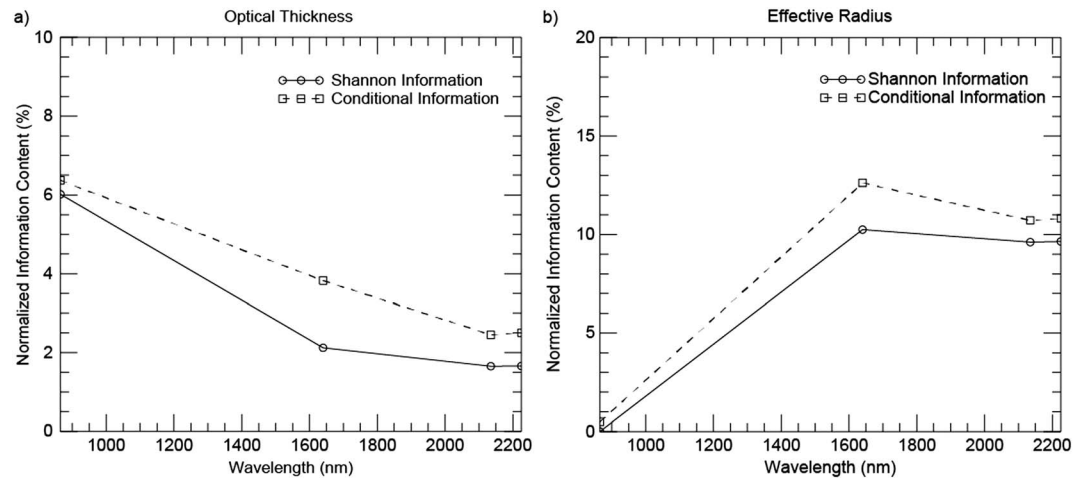om making an assumption about the parameter's value) and how this will propagate into a net information gain for the other parameter; this is called the conditional information content (section 4.iv).

6.1. The Shannon, Mutual, and Conditional Information Contents of Cloud Optical Properties

In Figure 6, we show the normalized information content (converted to percent from normalized values spanning 0–1) for the cloud case $\tau = 10$, $r_{eff} = 10$ μm, and phase = liquid (discussed in section 5.1 and Figure 1) as a function of wavelength for the PACE measurement channel combination. For this experiment, we have not updated the posterior pdf for each subsequent wavelength of measurements introduced into the GENRA algorithm using the prior information from the previous wavelength. As a result, the information content results shown in Figure 6 represent the information of the joint cloud (τ, r_{eff}) parameters for each measurement channel alone (i.e., these results do not reproduce the cumulative effect of the spectral information). The mutual information (dashed line with black squares) quantifies the information "shared" between τ and r_{eff} as measurements as each of the four PACE channels is introduced into the GENRA algorithm. For this particular cloud case, each PACE measurement channel is shown to provide some information about both τ and r_{eff} to varying degrees. The *sum* of the marginal Shannon information, identified by the dashed line with black circles, is the total of the partial information contributions gained by making a measurement when considering τ and r_{eff} independently (equations (12b) and (12c)). These partial information contributions need not sum to the maximum information provided by a spectral measurement [*Rodgers*, 1998] when that measurement provides information about both τ and r_{eff} cloud parameters; the maximum information, the joint Shannon information (equation (12a)), is identified as the solid black line. What we have demonstrated in Figure 6 is that the total information to be gained by making a spectral measurement can be broken down into a sum of the total information in τ and r_{eff} that the spectral measurement independently provides *plus* the information in the measurement that is shared by τ and r_{eff}. This is illustrated in Figure 6 where the red circles representing the sum of the marginal Shannon information content and the mutual information content lie on top of the black solid line.

In Figure 7, we use the conditional information content to quantify the total information that can be gained in a parameter (when making a spectral measurement that provides information on more than one parameter) by incorporating additional information to provide complete knowledge of the other, correlated, parameter. The source of this additional information would depend upon application, for example, in situ data, another

Figure 7. The normalized Shannon information content and conditional information content of (a) cloud optical thickness and (b) droplet effective radius for the experimental setup and implementation described for Figures 1 and 6.

ground, air, or space platform, or by making an assumption (as was done for this theoretical experiment). The results are specific to the same τ, r_{eff}, and phase experimental setup and implementation described for Figures 1 and 6.

In Figure 7a, we show the normalized Shannon information content in the marginal pdf (equation (12b)) of optical thickness as a function of PACE spectral channel (solid line with circles). This is the total information gained by making these spectral measurements when we consider that the measurements provide independent information about τ and r_{eff}. Theoretically, since the measurement provides information about both τ and r_{eff}, we have evaluated the additional information in τ that can be gained *when we assume that the effective radius value is known to complete certainty*; this is called the conditional information content (dashed line with circles). At all retrieval wavelengths, the conditional information is always greater than the marginal Shannon information in optical thickness, which is to be expected because adding information always reduces uncertainty (i.e., the Shannon information is inversely related to entropy and decreasing entropy represents decreasing uncertainty as discussed in section 4). In addition, and to be expected, the relative increase in information gained in τ by complete knowledge of r_{eff} is greatest at retrieval wavelengths where the information in a measurement that is shared by τ and r_{eff} is the largest (see mutual information; Figure 6).

Figure 7b repeats the analysis discussed in Figure 7a, but for the conditional information content in r_{eff} that can be exploited when *the optical thickness value is theoretically known to complete certainty*. Unlike the mutual information content, the conditional information contents of τ and r_{eff} are not symmetric. This simply reflects that a single measurement, while potentially providing the very real possibility of information about more than one parameter, may not equally distribute that information between the parameters (i.e., a measurement may provide more information about τ than r_{eff}, for example, and therefore, any potential change to the entropy of one parameter through theoretical knowledge of the other parameter would not be symmetric).

There are additional aspects of the information content relationships shown in Figures 6 and 7 not yet discussed that manifest themselves directly from the physical principles that govern the strong wavelength dependence of absorption and scattering of shortwave radiation by clouds. As discussed in the Introduction, absorption of radiation by cloud water droplets or ice particles is negligible at visible wavelengths and increases at longer wavelengths (and with larger particle size). The extinction (absorption + scattering) of radiation by cloud droplets depends on the particle cross section, and hence particle size. This lends to the inability to completely separate cloud optical thickness and effective radius in cloud retrievals. However, at visible wavelengths where cloud absorption is very small most, but not all, of the information is in optical thickness as shown in Figure 7a. The greatest information in particle size comes at near-infrared wavelengths where the dependency of absorption on particle size is greatest, but cannot be completely separated from information in optical thickness (Figure 7b). The mutual information and conditional information contents
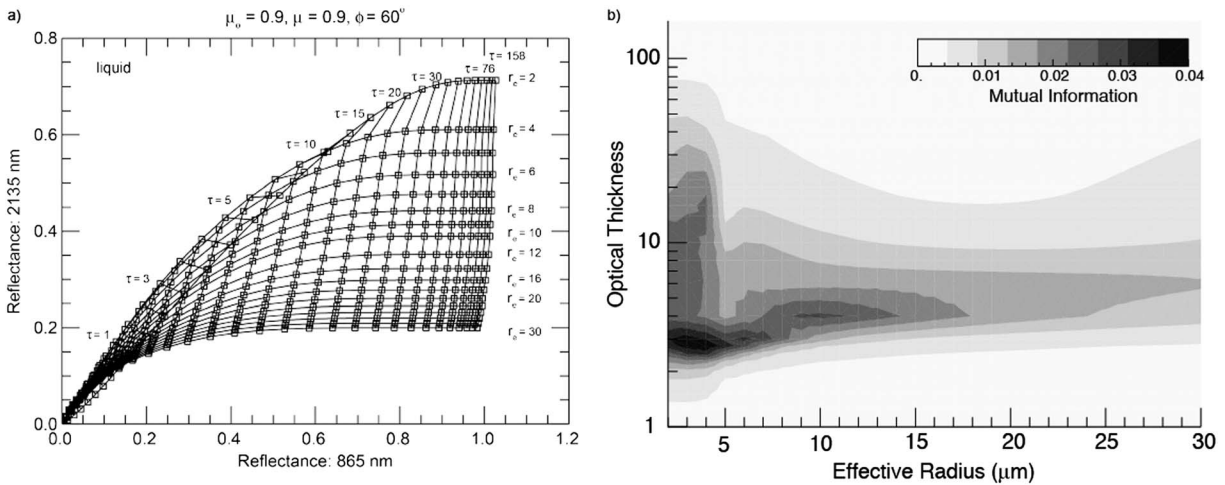
Figure 8. (a) Cloud reflectance look-up table at 865 nm and 2135 nm demonstrates near-vertical lines of constant optical thickness and near-horizontal lines of constant effective radius. (b) The mutual information in optical thickness and effective radius from measurements at 865 nm and 2135 nm.

provide the tools to quantify changes in dependencies between τ and r_{eff} that manifest themselves with changes in spectral channels. For example, such dependencies in mutual information at 2135 nm (or 2225 nm), relative to 1640 nm, can be seen in Figure 6 and the conditional information in Figure 7.

6.2. Mutual Information as a Visualization Tool for Cloud Parameters

In section 6.1, we quantified the information given measurements in different channels that is shared by τ and r_{eff} and presented formal metrics that quantify this dependency in the cloud optical properties as a function of wavelength. For many years, the physics behind two-wavelength cloud retrievals has been illustrated by plots similar to Figure 8a, which shows cloud reflectance (i.e., reflected cloud radiance normalized by downwelling irradiance) at two wavelengths spanning the very near-infrared through the near-infrared [for example, *Nakajima and King*, 1990; *Haywood et al.*, 2004; *Platnick et al.*, 2003]. At visible and very near-infrared wavelengths, such as 865 nm, the absorption of radiation by water is negligible and the magnitude of cloud reflectance is dominated by optical thickness as demonstrated by the near-vertical lines of constant optical thickness values in Figure 8a. At near-infrared wavelengths, such as 2135 nm, the absorption of radiation by water is much stronger and the magnitude of the absorbed radiation increases with particle size (i.e., cloud reflectance decreases in the near-infrared with increasing particle size as demonstrated by the near-horizontal lines of constant effective radius in Figure 8a). As optical thickness increases, the near-vertical lines of optical thickness and the near-horizontal lines of effective radius approach orthogonality (i.e., the different spectral channels provide nearly "independent" information on τ and r_{eff}).

We have shown (supporting information and Figure 6) that the mutual information content is a quantitative way to measure the degree of independence in parameters for a spectral measurement. In Figure 8b, we show results of an experiment where simulated cloud reflectance at two wavelengths (865 nm and 2135 nm) was sequentially introduced into the GENRA algorithm and the cumulative information in the retrieval was assessed for a broad range of (τ, r_{eff}) values. The experimental assumptions, setup, and implementation are kept consistent with those described for Figure 6. As expected, the mutual information content results mimic the dependencies in τ and r_{eff} depicted by the two-wavelength reflectance plot of Figure 8a. The amount of shared information is largest at small effective radius values ($r_{eff} < \sim4\ \mu m$) for all optical thickness values less than $\tau\sim50$. At small optical thicknesses ($\tau < \sim10$), a greater dependency with particle size exists for all particle sizes, but especially for those of 10 μm or smaller.

7. Assessing Impacts of Higher Measurement Accuracy on Cloud Optical Properties

Current imagers, such as MODIS and VIIRS, contributing to the global cloud data record have a radiometric accuracy of around 3% [e.g., *Xiong et al.*, 2014, 2016]. However, the Reflected Solar instrument for CLARREO pathfinder that is currently being developed will have a radiometric accuracy approximately an
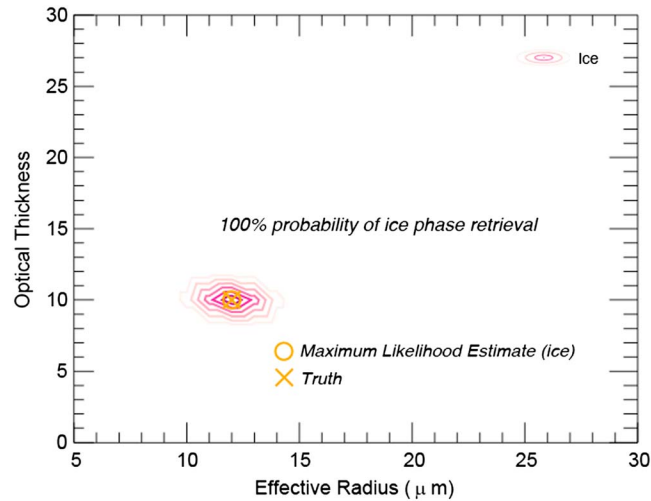
Figure 9. The 2-D joint posterior pdf for $\tau = 10$, $r_{eff} = 12$ μm, and phase = ice assuming measurement uncertainty of 0.3% and the "PACE" measurement channels (Table 1). The impacts of increased radiometric accuracy can be seen by comparing this result with the result shown in Figure 2g.

order of magnitude better (closer to 0.3%) [*Kopp et al.*, 2013]. Here we investigate the impacts of improved measurement precision on the discrimination of cloud thermodynamic phase and the retrieval of cloud optical properties. For this experiment, our true cloud type is $\tau = 10$, $r_{eff} = 12$ μm, and phase = ice. The model uncertainty is assumed to be 2%, measurement uncertainty is 0.3%, and we assume a black surface.

Figure 9 shows the 2-D posterior joint pdfs that correspond to the final retrieval after all PACE channel combinations (865 nm, 1640 nm, 2130 nm, and 2225 nm) are introduced into the GENRA algorithm and their cumulative impact evaluated (i.e., for this experiment, we have updated the posterior pdf for each subsequent wavelength of measurements using the prior information from the previous wavelength). The obvious impacts of improved measurement accuracy can be seen by comparing Figure 9 with Figure 2g (top right plot). For this cloud type (τ, r_{eff}, phase) and observational conditions, the percent probability of retrieving the correct (i.e., ice) phase is 100%; there is no overlapping contribution to the joint posterior pdf from the solution space identified as coming from the liquid water phase. The maximum a posteriori estimate of the 2-D joint posterior pdf is centered on the true (τ, r_{eff}) values. In addition, because the distribution in (τ, r_{eff}) space is more Gaussian, the maximum a posteriori estimates of optical thickness and effective radius derived from the marginal pdf distributions are also centered on the true values (not shown).
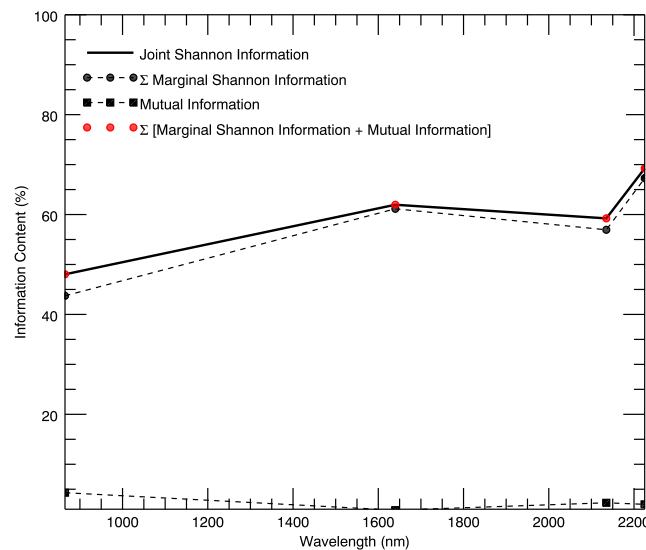

Figure 10. The normalized total (black line), marginal (dashed line with circles), and mutual information content (dashed line with squares) derived from the entropy relationships in the (τ, r_{eff}) cloud optical parameters after ingesting simulated cloud reflectance at 865 nm, 1640 nm, 2135 nm, and 2225 nm into the GENRA algorithm (prior is updated).

The reduced dependencies between optical thickness and effective radius in this experiment, reflected by the Gaussian nature of the 2-D joint posterior pdf, are also represented in the relationships in the joint, marginal, and mutual entropy metrics presented in Figure 10. In Figure 10, we demonstrate that the sum of the Shannon information content in the marginal pdfs of optical thickness and effective radius approaches the Shannon information content value of the joint (τ, r_{eff}) 2-D pdf only when the mutual information between the parameters approaches 0 (at 1640 nm, for example).

Finally, we revisit the experiment shown in Figure 8b, and repeat the simulations using a measurement uncertainty of 0.3%. The results are shown in Figure 11 and demonstrate that the dependency between optical thickness and effective radius is
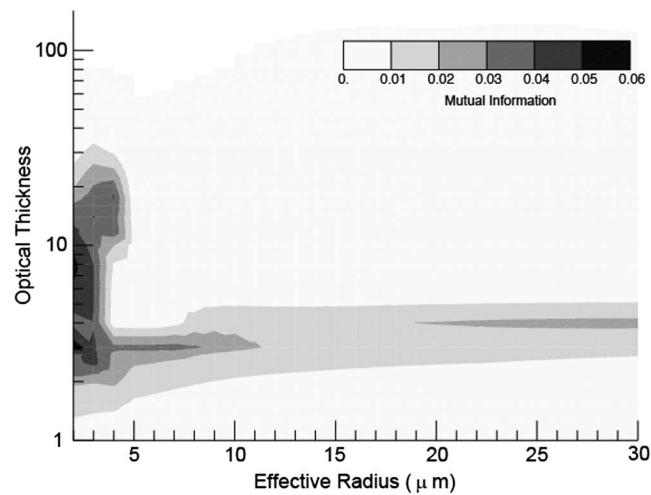
Figure 11. Same as in Figure 8b but for measurement uncertainty equal to 0.3%.

almost completely limited to optical thickness values less than $\tau \sim 10$. A practical interpretation suggests that improvements in instrument radiometric accuracy will lead to improvements in the retrieval of cloud properties *over parameter ranges for which passive shortwave images have already demonstrated retrieval "skill."* New retrieval approaches, such as spectral slopes and additional retrieval wavelengths [*McBride et al.*, 2011; *LeBlanc et al.*, 2015], the combination of observations from passive sensors [*Sourdeval et al.*, 2015], or the combination of observations from lidar and passive remote sensing methods [*Lebsock and Su*, 2014] will also be needed to make further improvements for (τ, r_{eff}) pairs that are challenging for passive sensors (for example, optically thin clouds). In addition, it would be premature to assume that only these four retrieval wavelengths, at a higher measurement accuracy, would suffice to discriminate cloud phase or retrieve cloud properties to high precision for global conditions (for example, clouds over bright snow/ice surfaces) when restricting the parameter space to $\tau > 10$.

8. Summary and Future Work

In this work, we have quantified the probability of cloud thermodynamic phase discrimination using shortwave channels alone from the MODIS, VIIRS, and future PACE imagers. The results show that the use of dual channels near 2 μm improves phase discrimination for regions of the cloud property parameter space where standard retrieval methods currently provide usable information (i.e., for moderate cloud optical thickness and larger particle sizes). In addition to quantifying the increase of information by adding retrieval channels, the GENRA toolkit has utility for comparing channel sets for differing retrievals and for selecting channels during mission development.

While our analysis was performed for simplified assumptions of measurement uncertainty (3%; wavelength-independent) and model uncertainty (2%; wavelength-independent) and for a dark, spectrally neutral surface, we believe that these results have utility in establishing a baseline and for simple cloud scenes over ocean surfaces. As part of ongoing work we are repeating the analysis approach discussed here, but for differing land surface types (i.e., snow/ice and vegetation) based on spectral measurements of spectral surface albedo from the MODIS instrument [e.g., *Moody et al.*, 2007] and from select Solar Spectral Flux Radiometer [*Pilewskie et al.*, 2003] measurements (summarized in *Coddington et al.* [2013]) because the reflected radiation from clouds is also influenced by the surface and atmosphere below the cloud, particularly for thin clouds [e.g., *Platnick et al.*, 2001]. The uncertainty in phase discrimination has not been readily available, and the best practices used for quantifying these impacts in algorithms that rely on phase information remain an open research question.

Earth's changing climate has profound implications for society [*NASA*, 2014 Strategic Plan]. The successfulness of society in adapting to and mitigating the impacts of climate change, including understanding and predicting the role of clouds in a changing climate, requires timely and accurate information. In sections 6 and 7 we applied new (to cloud retrieval studies) and existing information content metrics to define the information inherent in a spectral measurement, all with the common goal of quantifying the uncertainties in retrieved cloud properties and to improve our ability to effectively and efficiently utilize the information in current and future cloud observations.

For example, we give examples where the physical models of cloud reflectance show that a single spectral measurement gives information about both cloud optical thickness and droplet effective radius. The

conditional information content could be used to quantify the theoretical impact of how additional information about one of these parameters, possibly from a measurement or a retrieval from a different platform, may improve our knowledge of the other cloud parameter. In addition, we have shown the utility of the mutual information content in reflecting the dependencies between cloud optical thickness and droplet effective radius given spectral measurements. Historically, these dependencies have been illustrated using plots of cloud reflectance (or albedo) at two cloud retrieval wavelengths. However, illustrating these dependencies with cloud reflectance plots for anything more than two retrieval wavelengths becomes impossible to interpret with any useful physical meaning, leaving the mutual information content as a rigorous approach to reflect these dependencies in multispectral cloud retrievals.

References

Baum, B. A., W. P. Menzel, R. A. Frey, D. C. Tobin, R. E. Holz, S. A. Ackerman, A. K. Heidinger, and P. Yang (2012), MODIS cloud-top property refinements for Collection 6, *J. Appl. Meteorol. Climatol.*, *51*(6), 1145–1163, doi:10.1175/Jamc-D-11-0203.1.

Coddington, O., P. Pilewskie, and T. Vukicevic (2012), The Shannon information content of hyperspectral shortwave cloud albedo measurements: Quantification and practical applications, *J. Geophys. Res.*, *117*, D04205, doi:10.1029/2011JD016771.

Coddington, O., P. Pilewskie, K. S. Schmidt, P. J. McBride, and T. Vukicevic (2013), Characterizing a new surface-based shortwave cloud retrieval technique, based on transmitted radiance for soil and vegetated surface types, *Atmos. Basel*, *4*(1), 48–71, doi:10.3390/atmos4010048.

Cover, T. M., and J. A. Thomas (2006), *Elements of Information Theory*, 2nd ed., p. 748, John Wiley, Hoboken, N. J.

Del Castillo, C. (Chair) (2012), Pre-Aerosol, Clouds, and ocean Ecosystems (PACE) mission science definition team report, National Aeronautics and Space Administration. [Available online at https://decadal.gsfc.nasa.gov/PACE/PACE_SDT_Report_final.pdf.]

Haywood, J. M., S. R. Osborne, and S. J. Abel (2004), The effect of overlying absorbing aerosol layers on remote sensing retrievals of cloud effective radius and cloud optical depth, *Q. J. R. Meteorol. Soc.*, *130*(598), 779–800, doi:10.1256/qj.03.100.

King, M. D., Y. J. Kaufman, W. P. Menzel, and D. Tanre (1992), Remote-sensing of cloud, aerosol, and water-vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS), *IEEE Trans. Geosci. Remote*, *30*(1), 2–27, doi:10.1109/36.124212.

King, M. D., S. Platnick, W. P. Menzel, S. A. Ackerman, and P. A. Hubanks (2013), Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites, *IEEE Trans. Geosci. Remote*, *51*(7), 3826–3852, doi:10.1109/TGRS.2012.2227333.

Kopp, G., P. Pilewskie, C. Belting, Z. Castleman, G. Drake, J. Espejo, K. Heuerman, B. Lamprecht, P. Smith, and B. Vermeer (2013), Radiometric absolute accuracy improvements for imaging spectrometry with Hysics, *IEEE Int. Geosci. Remote Sens. Symp. (Igarss)*, 3518–3521, doi:10.1109/Igarss.2013.6723588.

LeBlanc, S. E., P. Pilewskie, K. S. Schmidt, and O. Coddington (2015), A spectral method for discriminating thermodynamic phase and retrieving cloud optical thickness and effective radius using transmitted solar radiance spectra, *Atmos. Meas. Tech.*, *8*(3), 1361–1383, doi:10.5194/amt-8-1361-2015.

Lebsock, M., and H. Su (2014), Application of active spaceborne remote sensing for understanding biases between passive cloud water path retrievals, *J. Geophys. Res. Atmos.*, *119*, 8962–8979, doi:10.1002/2014JD021568.

Marchant, B., S. Platnick, K. Meyer, G. T. Arnold, and J. Riedi (2016), MODIS collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP, *Atmos. Meas. Tech.*, *9*(4), 1587–1599, doi:10.5194/amt-9-1587-2016.

McBride, P. J., K. S. Schmidt, P. Pilewskie, A. S. Kittelman, and D. E. Wolfe (2011), A spectral method for retrieving cloud optical thickness and effective radius from surface-based transmittance measurements, *Atmos. Chem. Phys.*, *11*(14), 7235–7252, doi:10.5194/acp-11-7235-2011.

Menzel, W. P., R. A. Frey, H. Zhang, D. P. Wylie, C. C. Moeller, R. E. Holz, B. Maddux, B. A. Baum, K. I. Strabala, and L. E. Gumley (2008), MODIS global cloud-top pressure and amount estimation: Algorithm description and results, *J. Appl. Meteorol. Climatol.*, *47*(4), 1175–1198, doi:10.1175/2007jamc1705.1.

Meyer, K., S. Platnick, G. T. Arnold, R. E. Holz, P. Veglio, J. Yorks, and C. X. Wang (2016), Cirrus cloud optical and microphysical property retrievals from eMAS during SEAC(4)RS using bi-spectral reflectance measurements within the 1.88 mu m water vapor absorption band, *Atmos. Meas. Tech.*, *9*(4), 1743–1753, doi:10.5194/amt-9-1743-2016.

Moody, E. G., M. D. King, C. B. Schaaf, D. K. Hall, and S. Platnick (2007), Northern Hemisphere five-year average (2000–2004) spectral albedos of surfaces in the presence of snow: Statistics computed from Terra MODIS land products, *Remote Sens. Environ.*, *111*(2–3), 337–345, doi:10.1016/j.rse.2007.03.026.

Mosegaard, K., and A. Tarantola (2002), Probabilistic approach to inverse problems, in *International Handbook of Earthquake and Engineering Seismology (Part A)*, pp. 237–265, Academic, Boston, Mass.

Nakajima, T., and M. D. King (1990), Determination of the optical-thickness and effective particle radius of clouds from reflected solar-radiation measurements. 1. Theory, *J. Atmos. Sci.*, *47*(15), 1878–1893, doi:10.1175/1520-0469.

NASA (2014), Strategic Plan, [retrieved from https://smd-prod.s3.amazonaws.com/science-green/s3fs-public/atoms/files/FY2014_NASA_StrategicPlan_508c.pdf.]

Pavolonis, M. J., and A. K. Heidinger (2004), Daytime cloud overlap detection from AVHRR and VIIRS, *J. Appl. Meteorol.*, *43*(5), 762–778, doi:10.1175/2099.1.

Pavolonis, M. J., A. K. Heidinger, and T. Uttal (2005), Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparisons, *J. Appl. Meteorol.*, *44*(6), 804–826, doi:10.1175/Jam2236.1.

Pilewskie, P., and S. Twomey (1987), Discrimination of ice from water in clouds by optical remote sensing, *Atmos. Res.*, *21*, 113–122.

Pilewskie, P., J. Pommier, R. Bergstrom, W. Gore, S. Howard, M. Rabbette, B. Schmid, P. V. Hobbs, and S. C. Tsay (2003), Solar spectral radiative forcing during the southern African regional science initiative, *J. Geophys. Res.*, *108*(D13), 8486, doi:10.1029/2002JD002411.

Platnick, S., J. Y. Li, M. D. King, H. Gerber, and P. V. Hobbs (2001), A solar reflectance method for retrieving the optical thickness and droplet size of liquid water clouds over snow and ice surfaces, *J. Geophys. Res.*, *106*(D14), 15,185–15,199, doi:10.1029/2000JD900441.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey (2003), The MODIS cloud products: Algorithms and examples from Terra, *IEEE Trans. Geosci. Remote*, *41*(2), 459–473, doi:10.1109/Tgrs.2002.808301.

Platnick, S., R. Pincus, B. Wind, M. D. King, M. Gray, and P. Hubanks (2004), An initial analysis of the pixel-level uncertainties in global MODIS cloud optical thickness and effective particle size retrievals, *Passive Opt. Remote Sens. Atmos. Clouds Iv*, *5652*, 30–40, doi:10.1117/12.578353.

Platnick, S., et al. (2014), MODIS cloud optical properties: User guide for the collection 6 level-2 MOD06/MYD06 product and associated level-3 datasets, [http://modis-atmos.gsfc.nasa.gov/_docs/C6MOD06OPUserGuide.pdf.]

Platnick, S., et al. (2015), MODAWG: MODIS-VIIRS product continuity for cloud mask, Cloud-Top & Optical Properties, 2015 MODIS Science Team Meeting, [https://modis.gsfc.nasa.gov/sci_team/meetings/201505/presentations/plenary/platnick.pdf.]

Platnick, S., et al. (2017), The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua, *IEEE Trans. Geosci. Remote Sens.*, *55*, 502–525, doi:10.1109/TGRS.2016.2610522.

Posselt, D. J. (2016), A Bayesian examination of deep convective squall-line sensitivity to changes in cloud microphysical parameters, *J. Atmos. Sci.*, *73*(2), 637–665, doi:10.1175/Jas-D-15-0159.1.

Posselt, D. J., and T. Vukicevic (2010), Robust characterization of model physics uncertainty for simulations of deep moist convection, *Mon. Weather Rev.*, *138*(5), 1513–1535, doi:10.1175/2009mwr3094.1.

Rodgers, C. D. (1998), Information content and optimisation of high spectral resolution remote measurements, *Remote Sens.: Inversion Problems Nat. Hazards*, *21*(3), 361–367, doi:10.1016/S0273-1177(97)00915-0.

Shannon, C., and W. Weaver (1949), *The Mathematical Theory of Communication*, Univ. of Ill. at Urbana-Champaign, Urbana.

Sourdeval, O., L. C. Labonnote, A. J. Baran, and G. Brogniez (2015), A methodology for simultaneous retrieval of ice and liquid water cloud properties. Part I: Information content and case study, *Q. J. R. Meteorol. Soc.*, *141*(688), 870–882, doi:10.1002/qj.2405.

Stamnes, K., S. C. Tsay, W. Wiscombe, and K. Jayaweera (1988), Numerically stable algorithm for discrete-ordinate-method radiative-transfer in multiple-scattering and emitting layered media, *Appl. Opt.*, *27*(12), 2502–2509.

Stubenrauch, C. J., et al. (2013), Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, *Bull. Am. Meteorol. Soc.*, *94*(7), 1031–1049, doi:10.1175/Bams-D-12-00117.1.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, 342 pp., SIAM, Philadelphia, Pa., doi:10.1137/1.9780898717921.

Twomey, S., and T. Cocks (1989), Remote sensing of cloud parameters from spectral reflectance in the near-infrared, *Beitr. Phys. Atmos.*, *62*, 172–179.

Uprety, S., and C. Y. Cao (2015), Suomi NPP VIIRS reflective solar band on-orbit radiometric stability and accuracy assessment using desert and Antarctica Dome C sites, *Remote Sens. Environ.*, *166*, 106–115, doi:10.1016/j.rse.2015.05.021.

Vukicevic, T., and D. J. Posselt (2008), Analysis of the Impact of Model Nonlinearities in Inverse Problem Solving, *J. Atmos. Sci.*, *65*, 2803–2823.

Vukicevic, T., O. Coddington, and P. Pilewskie (2010), Characterizing the retrieval of cloud properties from optical remote sensing, *J. Geophys. Res.*, *115*, D20211, doi:10.1029/2009JD012830.

Wang, C. L., and H. W. Shen (2011), Information theory in scientific visualization, *Entropy-Switz*, *13*(1), 254–273, doi:10.3390/e13010254.

Wilks, D. S. (2011), *Statistical Methods in the Atmospheric Sciences*, 3rd ed., Elsevier Inc.

Xiong, X. X., A. Angal, J. Q. Sun, T. Y. Choi, and E. Johnson (2014), On-orbit performance of MODIS solar diffuser stability monitor, *J. Appl. Remote. Sens.*, *8*, 083514, doi:10.1117/1.Jrs.8.083514.

Xiong, X. X., J. Butler, K. F. Chiang, B. Efremova, J. Fulbright, N. Lei, J. McIntire, H. Oudrari, Z. P. Wang, and A. S. Wu (2016), Assessment of S-NPP VIIRS on-orbit radiometric calibration and performance, *Remote Sens. Basel*, *8*(2), 84, doi:10.3390/rs8020084.

Yang, P., L. Bi, B. A. Baum, K. N. Liou, G. W. Kattawar, M. I. Mishchenko, and B. Cole (2013), Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 mu m, *J. Atmos. Sci.*, *70*(1), 330–347, doi:10.1175/Jas-D-12-039.