A Deterministic Method for Profile Retrievals From Hyperspectral Satellite Measurements

Prabhat K. Koner, *Member, IEEE*, Andrew R. Harris, and Prasanjit Dash

***Abstract*—Different aspects of the operational constraints of remote sensing inverse problems are thoroughly investigated by simulation studies, using a deterministic method, namely regularized total least squares (RTLS). For demonstration purposes, water vapor profiles retrievals from simulated Suomi NPP Cross-track Infrared Souder (CrIS) hyperspectral measurements are considered. Synthetic CrIS radiances are generated using a line-by-line radiative transfer model (GENSPECT) with ∼424 realistic radiosonde profiles and US 1976 standard atmosphere as inputs. These results are also compared with those from a prevalent stochastic method. Our findings show that the stochastic method, even with additional deterministic constraints (truncated singular value decomposition) applied on top of it, is often unable to produce useful retrieval results, i.e., posterior error is more than the *a priori* error. In contrast, RTLS is able to produce deterministically unique results according to the available information content in the measurements, which could result in a paradigm shift in operational satellite inversion.**

***Index Terms*—Hyperspectral infrared sounding, ill-conditioned inverse, regularized total least squares (RTLS), Suomi NPP Cross-track Infrared Souder (CrIS).**

I. Introduction

MODERN hyperspectral sounding instruments have high information content. In particular, the information-rich Cross-track Infrared Sounder (CrIS) measurements onboard NPP Suomi satellite has a high SNR and offers great potential to rigorously characterize atmospheric composition. However, much of this potential remains unrealized primarily because of the choice of the operationally implemented inverse methods, most of which are based on stochastic approaches where errors are treated as finite information and are used as input parameters for parameter estimation. There are two prevalent schools of thought in the realm of parameter estimation [1]–[3]: 1) deterministic; and 2) stochastic. Deterministic methods assume that there is a true value for all individually retrieved parameters and each has an associated error. These methods have been derived at individual measurement instance using minimization of an objective function (requiring a functional form of the forward model) representing data misfit (measurement error) in a given norm. Stochastic methods assume that all retrieved parameter values and measurements are uncertain. Consequently, parameters are retrieved for a set of measurements instances using either Bayesian probability theory, or 1-D variational principle, or Levenberg–Marquardt (LM) optimization, or chi-square minimization, all of which are constrained by error covariances or by employing regression. Some stochastic methods require a functional form of the forward model and some do not. While stochastic methods are successfully used for parameterization of many scientific processes when mature forward models do not exist, this paper focuses on the consequence of the stochastic methods for remote sensing inverse problem where mature radiative transfer (RT) physics is available.

Most operational satellite hyperspectral profile retrievals are based on stochastic methods (e.g., [4]–[10]), where the quality of retrieval at the individual pixel level is inherently lower than optimal due to the basic assumptions in their derivation. Most of these techniques use *a priori* information to constrain the solution through atmospheric covariance statistics and *a priori* estimates of the retrieved parameters. Sometimes, ensemble average *a priori* data are used to solve the inverse problem for what is a highly dynamic atmospheric/oceanographic system. In the deterministic viewpoint, truth is effectively known when both "*a priori*" and "*a priori* error" are known at single measurement instance. Put another way, *a priori* errors specified in stochastic methods are probability distributions and will therefore generally be an overestimation or underestimation at single measurement instance. Moreover, nonlinear RT problems are prevalently solved by applying Bayesian conditional probability theory or perhaps LM with error covariance constraints, where it is implicitly assumed that the characteristics of the error distributions are perfectly known. However, in reality, it is extremely difficult to accurately estimate the errors for the inversion of satellite measurements based on RT physics because there may be contributions from many sources, namely: instrument error, forward model error, spectral error, line shape error (line overlapping, far wing effect of major molecules, line mixing, etc.), errors from minor interfering gases or unmodeled parameters, background RT error, and nonlinearity error (e.g., [11]). Moreover, it is also impossible to perform an error analysis of errors, which are input parameters to such solution procedures. Using uncertain assumptions as stated earlier, the information content of a highly nonlinear

Manuscript received April 14, 2015; revised September 24, 2015, January 28, 2016, and March 17, 2016; accepted May 5, 2016. Date of publication June 28, 2016; date of current version August 11, 2016. This work was supported in part by the National Aeronautics and Space Administration under Grant NNX14AP64A.

P. K. Koner and A. R. Harris are with Center for Satellite Applications and Research, E/RA3, NOAA's Satellite and Information Service, College Park, MD 20740 USA, and also with the Earth System Science Interdisciplinary Centre, Cooperative Institute of Climate Sciences, University of Maryland, College Park, MD 20740 USA (e-mail: prabhat.koner@noaa.gov).

P. Dash is with the Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO 80523 USA, and also with EUMETSAT, 64295 Darmstadt, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2565722

system is calculated based on linear conditional probability theory and reports that one measurement can produce more than one piece of information (e.g., [12]–[18]). A "model *minus* observation" bias correction is typically applied to satellite measurements to get some meaningful results (cf. [19]). An additional disadvantage of stochastic-based retrieval is that it needs tuning of the reference data set with an averaging kernel of the inverse model (e.g., [19]–[25]) to render more reasonable validation statistics. Consequently, advanced understanding of physics and chemistry of the atmosphere may be hampered due to lack of appropriate inverse mathematics. Mathematically ill-posed problems can be formed under several circumstances, e.g., error enhancement due to an ill-conditioned Jacobian, difficulty in constructing suitable optimization schemes due to functional complexity, mutual information separation for multiple solutions, absence of a suitable linearization scheme for some nonlinear problems, or inadequate information in system modelling. To investigate the aforementioned ambiguities in the inverse problem, we will present a comparative study with a deterministic method, for simulated profiles retrieval from satellite hyperspectral IR measurements.

Among existing deterministic methods (e.g., [26]), the proposed regularized total least squares (RTLS) is the only one which inherently determines the optimal regularization strength to be applied to the normal equation first-order Newtonian inverse using all of the noise terms embedded in the residual vector (e.g., in [11], [27]–[30]). It is worth mentioning that the RTLS method is fundamentally different and does not belong to either stochastic or Tikhonov regularization methods. It is derived from the understanding of quadratic eigenvalue analysis of matrix inversion, which is equivalent to the minimization of the Rayleigh quotient equation [28]. The family of RTLS methods has a well-established heritage in other branches of science, particularly medical imaging (e.g., [31]–[35]) but has seldom been exploited in Earth observation science to date. In contrast to RTLS, the theory of traditional stochastic methods do not explicitly include any constraint to prevent noise enhancement in the state-space parameters from the existing noise in measurement space for an inversion with a high ill-conditioned Jacobian. To stabilize the noise propagation into parameter space, a recent trend has been to use additional constraints, which have a questionable sound scientific basis, on top of the existing stochastic method (e.g., [36]–[41]). This may be regarded as something of a band-aid approach to ameliorate the effects for the chosen inverse method. This approach, where error is ambiguously treated as definite information (and is used as an input parameter) in conjunction with some deterministic constraint, is argued as a way to select the best of both approaches (stochastic and deterministic). It can overcome some of the limitations of stochastic method but that is more an indicator of flaws in the underlying assumptions. Inclusion of this additional step increases complexity, numerical calculations/noise, and ambiguities, as well as information loss due to discarding of smaller singular value components from the inversion when truncated singular value decomposition (TSVD) is applied. On the other hand, without the need to reduce the model resolution, RTLS injects additional information into the inversion by employing a Laplacian first derivative operator (LFDO) as a stabilizer. LFDO constrains the solution since the update of adjacent atmospheric parameters in a profile are close, which is less harmful than the use of an *a priori* of what are significantly dynamic atmospheric parameters. The final RTLS solution is totally independent of the IG parameters of targeted retrievals, and regularization is data driven at all iterations.

We have previously shown that satellite remote sensing retrieval problems can be uniquely solved using the RTLS method for simulated retrievals (IR and microwave) and balloon-based hyperspectral IR measurements [11], [42]. Recently, we have also successfully implemented a similar algorithm (termed modified total least squares) in an operational environment [43] for sea surface temperature (SST) retrievals, which is a comparatively low ill-conditioned problem. This implementation has been in effect since August 2013 for operational geostationary satellites at the office of the satellite product operation (OSPO), NOAA. This paradigm shift in operational inverse method is providing near-real-time high-quality SST data to the community with a 50% reduction in error, as compared with the previous stochastic (regression) method. In this paper, we will discuss water-vapor profile retrieval from simulated CrIS measurements, considering various operational constraints.

II. Inverse Model

Although the RT equation is highly nonlinear, particularly for strong absorption regions, any inverse method approximates the system as locally linear for a single iteration and solves the problems using multiple Newtonian iterations for the Hamiltonian function. A linearized inverse problem thus can be formulated using the residual between observation and model as

$$\Delta \boldsymbol{y} = \boldsymbol{K} \Delta \boldsymbol{x} \tag{1}$$

where $\boldsymbol{x}$ is a model state vector (i.e., parameters of interest), and $\boldsymbol{y}$ is a vector of instrument channel radiances. Thus, $\Delta \boldsymbol{x}$ is the difference between the IG and the true state of parameter/s (i.e., what we are trying to retrieve), and $\Delta \boldsymbol{y}$ is the difference between the observed radiances and those calculated by RTM using an IG of the full model state from ancillary data, of which the targeted parameters are usually a subset. The notation $\boldsymbol{K}$ denotes the Jacobian (partial derivatives of channel radiances with respect to the targeted parameters). The least-squares (LS) solution [44] is

$$\Delta \boldsymbol{x} = \left(\boldsymbol{K}^T \boldsymbol{K}\right)^{-1} \boldsymbol{K}^T \Delta \boldsymbol{y}. \tag{2}$$

Since the RT problem is inherently ill conditioned, LS is potentially vulnerable to noise amplification from the measurement space to the state space for any satellite retrieval problem (e.g., [45]) because measurement $(\Delta \boldsymbol{y}_\delta = \Delta \boldsymbol{y} + \delta \boldsymbol{y})$ always contains noise $(\delta \boldsymbol{y})$. However, this measurement noise is not the only source of error involved with such retrieval problems. For instance, Koner *et al.* [42] demonstrated that, for any ill-conditioned linear inversion, the amount of error in the state-space parameters is directly proportional to the condition number of the Jacobian, and all errors associated with this inversion, which can be written as

$$\|\boldsymbol{e}\| \leq \boldsymbol{\kappa} \quad \Sigma \|\boldsymbol{\delta} \mathbf{E_i}\| \tag{3}$$

where $\boldsymbol{e}$ is the realization of error in the retrieved parameters, $\boldsymbol{\kappa}$ is the condition number of the Jacobian, and $\boldsymbol{\delta}\mathbf{E_i}$ represents errors associated in the inversion including errors in Jacobian, forward model, measurements, and ancillary data. This error enhancement in retrieved space can be mitigated theoretically in two different ways.

The deterministic approach to minimize the propagation of error into the parameter space involves employing a regularization operator, using some constraints or numerical approximations, to reduce the condition number of the inverted matrix $(\boldsymbol{K}^T\boldsymbol{K})$ of (2). A large numbers of different regularization algorithms, namely, Tikhonov, regularized Gauss–Newton, LM, TSVD, etc., are available in the literature. Among the various deterministic methods, the total least square (TLS) method has a distinct advantage of being data driven to determine the regularization strength. Although TLS has first been used only in recent decades [46], [47], this fitting concept has been long referred in the statistical literature under different names, e.g., orthogonal regression or errors in variables. However, despite its long use in statistical inversion, a deterministic form of TLS has only recently been derived using linear algebra. TLS can be derived using simple linear algebra considering the error in measurement and Jacobian in (1).

The full derivation of deterministic TLS is presented in [43], and a summary of the key points is reproduced here to facilitate review. An optimization of two variables has been considered for the derivation of the TLS method as follows:

$$\frac{\min}{\|\delta\mathbf{K}\|, \|\delta\mathbf{y}\|, \mathbf{x}} \left\{\|\delta\boldsymbol{K}\|^2 + \|\delta\boldsymbol{y}\|^2\right\} \quad \text{subject to} (\boldsymbol{K} - \delta\boldsymbol{K})\boldsymbol{\Delta}x = \boldsymbol{\Delta y}_\delta - \delta\boldsymbol{y}. \tag{4}$$

Only the TLS method considers Jacobian error in its optimization, which is most appropriate for RT problems since they are inherently nonlinear. Generally, all inverse methods solve some linear equations as stated in (1) using a suitable linearization scheme. It is easily justifiable to consider the Jacobian error in accounting for nonlinearity error in the linearization point. After performing a few steps of matrix algebra on (4), in a simple form, the TLS method is given as

$$\Delta\boldsymbol{x} = (\boldsymbol{K}^T\boldsymbol{K} + \lambda\boldsymbol{I})^{-1}\boldsymbol{K}^T\Delta\boldsymbol{y_\delta} \tag{5}$$

where λ is the lowest singular value of the matrix $[\boldsymbol{K}\ \Delta\boldsymbol{y_\delta}]$ and $\boldsymbol{I}$ is the identity matrix. Since the present RT problem is highly ill-conditioned, employing the TLS method, the condition number of the inverted matrix in (5) is reduced but not sufficient to prevent noise amplification from measurement to state spaces within a permissible range for the next iteration. Thus, the regularized TLS or truncated TLS are most commonly used (e.g., [11], [27]–[35], [42], [48]) to minimize the effect of ill-conditioning on resulting solutions. The full derivation of RTLS, which will be used here, is presented in Koner and Drummond [11]. The mathematical formulation of the RTLS for a linear problem where the number of measurements are more than that of the state-space parameters (e.g., [28], [29]) is

$$\frac{\min}{\Delta\mathbf{x} \in \mathbf{X}} \phi(\mathbf{x}, \mathbf{y}) := \frac{\|\boldsymbol{K}\Delta\mathbf{x} - \Delta\mathbf{y}\|^2}{1 + \|\Delta\mathbf{x}\|^2} \quad \text{subject to} \|\boldsymbol{L}\Delta\mathbf{x}\|^2 \leq \delta^2. \tag{6}$$

The term $\boldsymbol{L}$ denotes the regularization operator, $\phi(\mathbf{x}, \mathbf{y})$ is the cost function, and δ is infinitesimal. After performing a few steps of matrix algebra, the form of RTLS can be written as

$$\Delta\boldsymbol{x} = \left(\boldsymbol{K}^T\boldsymbol{K} - g(x)\boldsymbol{I} + \boldsymbol{\alpha}\boldsymbol{L^T}\boldsymbol{L}\right)^{-1}\boldsymbol{K}^T\Delta\boldsymbol{y_\delta} \tag{7}$$

where $g(\boldsymbol{x}) = \|\boldsymbol{y}_\delta - \boldsymbol{K}\Delta\boldsymbol{x}\|^2/(1 + \|\Delta\boldsymbol{x}\|^2)$, $\boldsymbol{\alpha}$ is regularization strength, and $\boldsymbol{I}$ is the identity matrix. For the calculation of $g(\boldsymbol{x})$, an update of $\boldsymbol{x}$ is required, which is obtained as $\Delta\boldsymbol{x} = 0$ for the first iteration and the retrieved $\Delta\boldsymbol{x}$ for successive iterations. Although considering $\Delta\boldsymbol{x} = 0$ for the first iteration, the second regularization term of RTLS stabilizes the solution by the value of $\boldsymbol{\alpha}$, which is calculated using the same value of $\Delta\boldsymbol{x}$. Thus, any underestimate or overestimate of $\Delta\boldsymbol{x}$ is compensated by the value of $\boldsymbol{\alpha}$. Some literature defines the RTLS as a dual-regularized method. The success of any regularization method is dependent on the regularization strength and characteristics of the regularization operator. We use a LFDO as a stabilizer [$\boldsymbol{L}$ in (7)] that, in the case of a profile retrieval problem, provides additional information and is a better approximation compared with the regularization using the identity matrix [49]. Using LFDO, this is realized by forcing values of the update of adjacent points within a profile to be close. Moreover, in practice, it is difficult to develop a mathematical derivation for a nonlinear problem, which is what the RT equation inherently is. Thus, in this calculations, the $\boldsymbol{I}$ matrix is replaced by the $\boldsymbol{L}$ matrix in (7).

The regularization strength $(\boldsymbol{\alpha})$ of the RTLS method is data driven and is calculated from the residual vector as follows:

$$\boldsymbol{W} = \boldsymbol{L}^{-T}\left(\boldsymbol{K}^T\boldsymbol{K} - g(x)\boldsymbol{L}\right)\boldsymbol{L}^{-1}. \tag{8}$$

The lowest singular value of the matrix $\boldsymbol{W}$ has been shown to provide the optimal regularization strength [28]. For a measurement instance, the optimal regularization strength is calculated at all iterations to block the nonlinear error injection into the retrieved space, and to restrict propagation of other errors as described earlier for all measurement instances. Additionally, we used a line search method to control the step size and convergence. We also use three different criteria to reduce the required number of iterations to obtain convergence [42].

Information content can be calculated in [50] and [51] in terms of degree of freedom in retrieval (DFR) as

$$\text{DFR} = \text{trace}(\mathbf{R}_{\text{km}}) \tag{9}$$

where $\mathbf{R}_{\text{km}}$ is the regularized kernel matrix, which is basically the model resolution matrix [51]. The expression for $\mathbf{R}_{\text{km}}$ is given as $(\mathbf{K}^T\mathbf{K} - g(x)\mathbf{I} + \boldsymbol{\alpha}\mathbf{L}^T\mathbf{L})^{-1}\boldsymbol{K}^T\mathbf{K}$.

An alternative to the deterministic approach that is used to ameliorate this issue (ill-conditioned inversion) is to provide errors as inputs into the inverse model. The error propagation to the state space will inevitably be zero for any ill-conditioned inversion when total error $(\boldsymbol{\varepsilon} = \sum\|\boldsymbol{\delta}\mathbf{E_i}\|)$ is zero, as specified in (3). Since an ill-conditioned matrix, until it becomes rank deficient (e.g., condition number exceeds 10^{16} for double precision calculations) can still be accurately inverted if the errors associated with observations are accurately known, the inversion of such a problem can be simply derived as follows:

$$\Delta \boldsymbol{y}_\delta - \boldsymbol{\varepsilon} = \boldsymbol{K}\Delta \boldsymbol{x} \tag{10}$$

where $\boldsymbol{\varepsilon}$ is the total errors involved in an inverse problem. It is not feasible to determine the exact errors associated within any inverse problem. Even if the magnitudes of errors could be estimated accurately for simulation purposes, the corresponding signs cannot be determined (if they could be, there would be no need for a retrieval). To minimize these errors in an inversion, LS optimization is necessary. Using matrix algebra and employing LS minimization of (10), it can be derived as [43]

$$\Delta \boldsymbol{x} = \left(\boldsymbol{K^T}\boldsymbol{\varepsilon}^{-2}\boldsymbol{K} + \Delta \boldsymbol{x}^{-2}\right)^{-1}\boldsymbol{K^T}\boldsymbol{\varepsilon}^{-2}\Delta \boldsymbol{y_\delta}. \tag{11}$$

It may seem surprising that $\Delta \boldsymbol{x}$ is required to determine the value of $\Delta \boldsymbol{x}$ in such a formulation. In practice, per-pixel exact values of $\boldsymbol{\varepsilon}$ and $\Delta \boldsymbol{x}$ are not available to solve this problem. Therefore, if one assumes that $\Delta \boldsymbol{x}$ has a Gaussian distribution for a set of measurement, representative values of this error (e.g., full width at half maximum (FWHM) of the distribution) can be used in (11). This is what is often done in practice in stochastic methods and such approximation of *a priori* errors referred to as *a priori* covariance S_a. Similarly, representative errors in observations for the same set of measurements are also specified as measurement error covariance S_e in stochastic approaches. Although the optimal estimation method (OEM) [52] is derived using Bayes' probability theorem, (11) (derived with simple linear algebra) is effectively identical to the OEM formulation. The success of OEM retrieval depends on the accuracy of *a priori* and *a priori* covariance. However, a fundamental concern, from a deterministic point of view, is that the truth is essentially known when *a priori* and *a priori* covariance are accurately known at the pixel level.

Another uncertain outcome of OEM results arises from its reliance on the measurement and *a-priori* errors, which are specified as inputs. For example, when these two errors match well with reality, the retrievals will be accurate, but such instances are obviously left to chance and can neither be identified nor ensured. Despite the issues stated above, this is widely used in remote sensing inverse problems, perhaps because it has not been analyzed from the deterministic point of view due to overreliance on the Bayesian conditional probability concept. For example, according to the algorithm theoretical basis document (ATBD) for CrIS [38], retrieval of level-2 (L2) environmental data records (EDRs) is based on the Bayesian inverse method

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_a + \left(\boldsymbol{K}^T\boldsymbol{S_e^{-1}}\boldsymbol{K} + \boldsymbol{S}_a^{-1}\right)^{-1}\boldsymbol{K}^T\boldsymbol{S}_e^{-1} \times \left((\boldsymbol{y}_\delta - \boldsymbol{y}_i) + \boldsymbol{K}(\boldsymbol{x}_i - \boldsymbol{x}_a)\right) \tag{12}$$

where $\boldsymbol{x}_{i+1}$ is the state-space parameter at ith iteration; $\boldsymbol{K}$ is the Jacobian; $\boldsymbol{S}_e^{-1}$ is the measurement error covariance; $\boldsymbol{S_a^{-1}}$ is the *a priori* covariance; $\boldsymbol{x}_a$ is the *a priori* of the state space; y_i denotes simulated model calculation at the ith iteration, and $\boldsymbol{y}_\delta$ is the measurement.

III. Discretized Radiative Transfer Model

A measurement can be considered equivalent to a model output if a compatible model exists. The model consists of a set of parameters embedded in a mathematical framework. In this context, it is safe to state that RT modeling (RTM) for atmospheric problems using line-by-line (LBL) calculations has reached a fairly high level of maturity.

The theoretical foundation of the IR remote sensing forward model is Schwarzschild's equation of RT. In a nonscattering atmosphere under local thermodynamic equilibrium, which means that the atmosphere behaves like a gray body, the basic equation governing the transfer of emitted thermal IR radiance at nadir that reach the top of the atmosphere (TOA) (cf. [53]) at a given wavenumber ν can be described by

$$Iv = \varepsilon Bv(T_s, p_s)\tau(p_s) + \int_{ps}^{0} Bv(T,p)\Delta\tau(p) - (1-\varepsilon)\int_{ps}^{0} Bv(T,p)\Delta\tau^*(p) \tag{13}$$

where B is the Planck's function; $\boldsymbol{\varepsilon}$ is spectral emissivity; $\boldsymbol{T}_s$ is the surface skin temperature; $\boldsymbol{T}$ and p are the temperature and pressure of the atmospheric grids, respectively; $\boldsymbol{\tau}(p)$ is the upwelling layer transmittance; and $\boldsymbol{\tau}^*(p) = \boldsymbol{\tau}(p_s)^2/\boldsymbol{\tau}(p)$ is the downwelling transmittance. The reflected IR solar radiation is considered negligible for bands with wavelength longer than 4 μm during the day. The emissivity of the sea surface in the thermal IR region is generally close to unity. Under these assumptions

$$Iv = \varepsilon_{\text{eff}} Bv(T_s, p_s)\tau(p_s) + \int_{ps}^{0} Bv(T,p)\Delta\tau(p) \tag{14}$$

where

$$\varepsilon_{\text{eff}} = \varepsilon_s - \frac{1-\varepsilon_s}{Bv(T_s,p_s)}\int_{ps}^{0} Bv(T,p)\Delta\tau^*(p). \tag{15}$$

A Fourier transform spectroscopic instrument typically works in the IR spectral region and measures the radiance at a finite number of spectral points with estimated equidistant wavenumbers. Therefore, a suitable discretization process is used over the integrals (14) and (15). There are many different discretization possibilities: simple classic quadrature method, collocation points and nodes, degenerate kernel approximations (by eigenfunctions or by orthonormal systems or approximation by Taylor series or interpolation), and projection methods (Galerkin moment or least squares). We have employed GENSPECT [54] for our forward modeling, which is an LBL RTM that uses a degenerate kernel function for interpolation. The discretization process of RT equations leads to a set of nonlinear system of equations, which are in a Hilbert space.

As a test bed, we have considered the sensor specifications of CrIS hyperspectral sounder onboard Suomi NPP (http://npp.gsfc.nasa.gov/cris.html), which is based on Fourier transform spectroscopy in the IR (FTIR). For real data from CrIS, channel radiance is given by a convolution of the instrument line shape (ILS) function with the monochromatic radiance at the entrance to the interferometer. It is a common practice to suppress ringing in the spectrum by apodizing the modulation function [55], artificially reducing the abruptness of the interferogram clipping by forcing the windowing function to tend smoothly to zero at the extremes. For this simulated study, a simplified "*sinc*" ILS is considered to produce equivalent CrIS measurements by convolving with the simulated spectrum.

IV. Simulation Study

Radiative transfer equations are highly complex functions that cannot be approximated by a particular class of function (e.g., quadratic, convex, logarithmic). Thus, it is very difficult to prove theoretically (i.e., only by mathematical derivation) that RTLS is a better choice (over OEM) for RT-based hyperspectral satellite profile retrievals. It is also difficult to prove the superiority of RTLS over OEM using satellite measurements because the selection of a "correct error covariance" for OEM (e.g., residual cloud contamination may not be considered an observation error) is debatable. To overcome these hurdles, we will use simulated retrievals to demonstrate the *proof of concept* of our proposed method because we have more control over various parameters for numerical experiments. Simulations have the benefit of allowing us to either exclude or include regulated operational problems (calibration, forward model error, cloud detection, etc.) and focus on the performance of the inversion method itself.

We consider simulated water vapor (WV) profile retrievals, not least because retrieval of WV profiles from satellite measurements is a challenging problem. WV lines are ubiquitous across most of the IR region, and accurate estimation of the WV profile is a prerequisite for many climate and atmospheric studies, as well as for retrieval of other geophysical parameters from satellite measurements. Upper air estimates of WV profiles are designated by Global Climate Observation System (GCOS) as an essential climate variable (ECV), for a variety of reasons. WV is the most globally variable greenhouse gas, and its variation and evolution are sources of critical direct and indirect radiative forcing feedback, the latter via effects on cloud cover, as well as exerting a controlling influence on other key ECVs, such as precipitation. There are abundant global variations in WV profiles, and characterizing *a priori* variance of the WV profile is a challenge in itself but remains a prerequisite for implementing OEM.

A. Simulated Measurements

We have calculated the spectra using only WV at a resolution of 0.05 cm^{-1} using the GENSPECT LBL model [54], for the US 1976 standard atmospheric temperature and WV profiles, and surface temperature of 300 K and surface emissivity of one as shown in Fig. 1.

To reach a more realistic condition, we have also studied the contaminated radiance from other interfering trace gases using model profiles of 28 different trace gases from Modtran4.3 (cf. [56]). For example, a selected window (1230–1385 cm^{-1}) is shown in Fig. 2. The radiances are calculated using standard

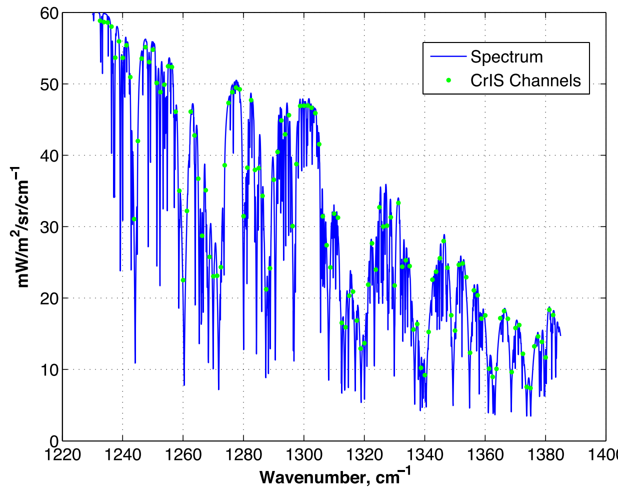

Fig. 1. Sample of simulated CrIS measurements (∗green) using LBL forward model with a spectral resolution of 0.05 cm^{-1} (Blue) and convolved with the sinc ILS.

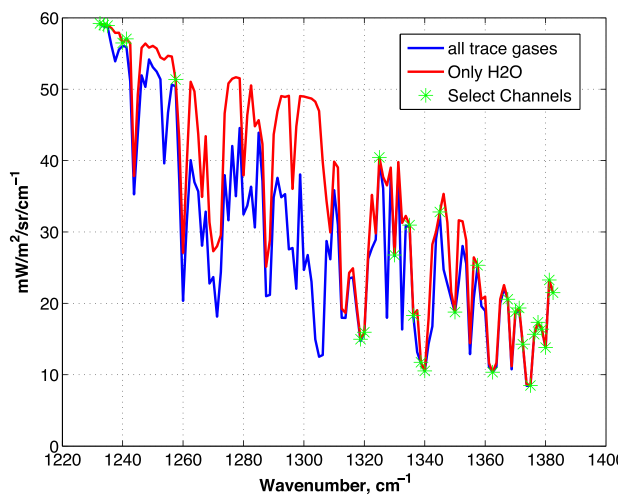

Fig. 2. Simulated spectrum of CrIS measurement with all interfering trace gases (blue) and without trace gases (red) and selection of channels (∗green), where SNR is greater than 100.

atmospheric profiles of all these gases (blue) and only using H_2O (red). It is found from this paper that many CrIS channels are subject to contamination by absorption due to various trace gases within this window. There are three alternatives in such a situation: 1) Use all interfering trace gases in the forward model simulation; 2) model all gases but the approximated error is applied at the level of error covariance to which the other gases are uncertain; and 3) discard the channels that are affected by an amount more than assumed threshold value. It is very difficult to obtain correct profiles of all interfering gases at the time of measurement and computational cost increases if the first choice is considered. From a deterministic point of view, weighted measurement using variable values of elements in error covariance, as described in the second choice above, cannot remove channel errors where interfering gases increase the measurement error relative to the target parameter. In such a process, both ambiguities in retrieval and computational cost are increased. Thus, for this experiment, we have opted for a third choice, i.e., a threshold value of 100 for SNR is considered that leaves only 29 out of 121 channels (shown in ∗green in Fig. 2) for our numerical experiment.

It is very difficult to get sufficient information to solve the WV profile of all atmospheric levels using only 29 channels in midwave IR (MWIR) of the CrIS measurement. Thus, we searched the information-rich channels for WV profile retrievals using the aforementioned approach for different bands of CrIS measurements, i.e., 40 channels from the long-wave IR region (700–960 cm^{-1}) for upper tropospheric WV information and 140 channels from the MWIR region (1230–1625 cm^{-1}) for mid tropospheric WV information. For this simulation, we did not consider any channel from the short-wave IR region to avoid additional modeling of the effect of solar scattering during daytime. A similar channel selection approach has been already used to retrieve different trace gases from balloon-based remote sensing FTIR measurements [11]. These selected channels may be somewhat incomplete for practical application but should be more than adequate for the intended purpose of comparing the inverse methods of RTLS and OEM. The present set of channels is unable to uniquely solve for the upper tropospheric region due lack of sufficient information in this region, which we will explore under a deterministic framework in future work.

B. Simulated Retrievals Using Fixed Profiles

Retrievals adding Monte Carlo noise in the simulated spectrum have been made using two different methods, i.e., RTLS and OEM, for three different true profiles (TPs), which are TP1 (realistic), TP2, and TP3 (sinusoidal profiles). TP1 is a U.S. standard WV profile for the Earth's atmosphere, and we have used two initial guess (IG) profiles: One is constant (IG1) and the other is realistic (close to TP1, IG2), shown in black and magenta, respectively (see Fig. 3). The number of altitude level is 16 for all profiles. We used three different Monte Carlo noise realizations of 1% ($\text{SNR} = 100$) in measurements, which is somewhat a conservative estimate since SNR is more than 150 for MWIR channels of CrIS (e.g., [57], [58]). Retrieval using OEM has initially been made considering a diagonal covariance matrix with the values of 100% *a priori* variance, for the profiles solved from IG1 because all of the truth is inevitably within the range of 100% variance of the IG1. For solving TP1 from IG2, *a priori* variance of 30% is used for OEM. As is typical, the IG profile is considered as the *a priori* profile, and error covariance is calculated according to the noise added into simulated measurements. All three profiles are solved using RTLS from IG1. Note that we did not guarantee that the *a priori*/IG is within the error domain of *a posterior*, and the shape of the *a priori* profile is close to truth, which are typical current practices (e.g., [22], [59], and [60]).

For more than three decades, remote sensing RT-based retrieval theories (e.g., [52] and [61]) are dominated by stochastic methods and claim that there exists no unique solution. However, recent published work on remote sensing retrieval (e.g., [62]) shows that there exist unique solutions for such problems when using deterministic inverse methods, within the limit of information content. For example, a full sinusoidal profile cannot be solved by the present retrieval model because this model has three to four independent pieces of information, and at least

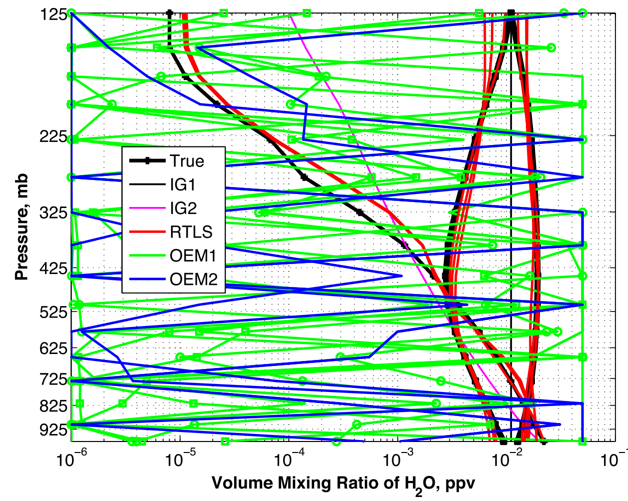

Fig. 3. Simulated retrieval using RTLS and OEM: three TPs (+solid black); RTLS solution (red solid) for all TPs from IGof IG1 (black solid); OEM solution: OEM1 (green solid) from IG1 for all TPs and OEM2 (blue) using IG2 (magenta) for realistic profile.

five pieces of information are required to solve a full sinusoidal profile. The present experiment follows the published retrieval theories on two different approaches (deterministic and stochastic). Fig. 3 shows that the RTLS method produces unique solutions for all three different profiles (TP1, TP2, and TP3) from a fixed IG profile (IG1), and it does not require *a priori* and observation error covariance matrices or, as demonstrated by use of IG1, even a representative *a priori* profile. We have used another realistic IG (IG2) for solution of only the TP1 case using OEM to verify the impact of IG in the OEM solution. All four OEM solutions, including the solution of TP1 from IG2, are outside the solution space ($> 300\%$ error), and additional boundary constraints (10^{-6} to 0.05) are required for OEM (Fig. 3). This illustrates that OEM is not an appropriate method for such problems. This can be explained from the deterministic viewpoint by arguing that using a combination of *a priori* and measurement error covariance matrices is unable to regularize the inverse problem adequately. The underlying cause is its inability to reduce the noise propagation from measurement space to state space for such an ill-conditioned inversion. Put another way, without an optimally regularized inverted matrix, even choosing *a priori* close to the truth may yield out-of-range solutions for a highly ill-conditioned inversion.

Exploring the cause of OEM failure from a deterministic point of view, regularization strength increases with an increasing value of measurement error covariance or a decreasing value of *a priori* error covariance in OEM formulation. The information from the measurement reduces with an increase in regularization strength, but this regularization is required to prevent noise propagation into the state space to obtain solutions at least within the range. We have already reported [49] results of similar profile retrievals from hyperspectral measurements using systematic (not arbitrarily) optimal regularization strength into the OEM formulation, without satisfactory success. Thus, to further demonstrate the real cause of this failure, we have conducted another experiment to increase the effective regularization strength of OEM by changing the *a priori* variance

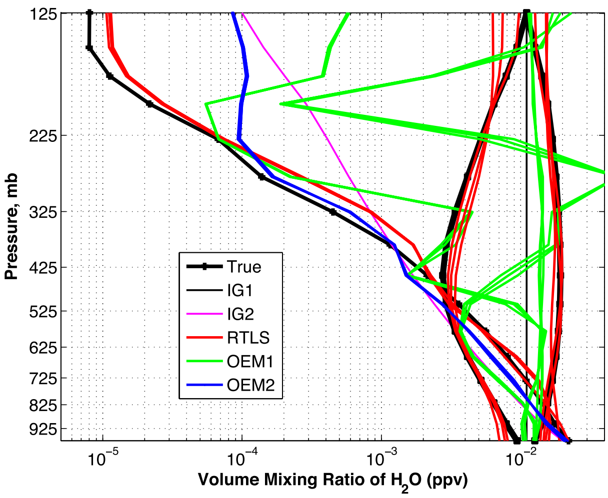

Fig. 4. Same as in Fig. 3 but for increased regularization strength (*a priori* variance of 1%) of OEM-like solution.

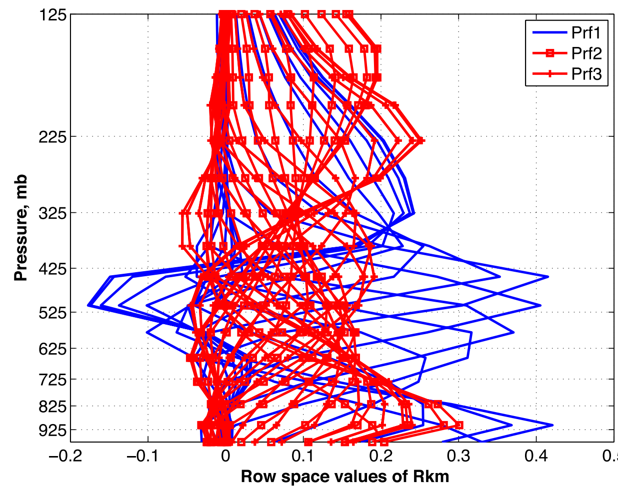

Fig. 5. Row space of regularized kernel matrix of RTLS at the last iteration for three different profiles.

arbitrarily to 1% instead of 100%, and keeping all other parameters fixed in (12), whose results are discussed next.

Fig. 4 shows that, in general, the OEM solutions are now within the domain of the solution space and the oscillations have reduced. For TP1 (from IG2 that is close to truth), OEM yields modest results at least up to an altitude of 250 mb. However, from IG1, OEM cannot satisfactorily solve any of the profiles (erratic shape with respect to the TPs). [This may be attributed to the effect of another constraint in OEM, given as $\boldsymbol{x}_a$ in the last term of (12)] This confirms that our previous solutions using OEM (see Fig. 3) were affected by low regularization. It is the high condition number of the problem that caused the oscillations. The condition number of the Jacobian for the present example (see Figs. 3 and 4) is on the order of 10^3–10^7, depending upon the choice of the retrieval grids and the shape of the profile. This confirms that a high regularization is required for such problem that is achieved by using 1% *a priori* covariance. However, an *a priori* variance of 1% for such a problem cannot be justified by any arguments in the domain of stochastic theory (e.g., OEM). On the contrary, it can be easily discussed from the deterministic viewpoint that the square root of the condition number of the inverted matrix $(K^T S_e^{-1} K + S_a^{-1})$ though reduced is still quite high ($\sim$400 from $\sim 10^5$ in original Jacobian) when 100% *a priori* covariance is used. Using (3), it is evident that 1% error in measurement can yield a state space error by 400% (condition number 400 of the inverted matrix multiplied by a measurement error of 1%). Similarly, when 1% *a priori* covariance is used the condition number of inverted matrix is $\sim$10, and the expected state space error is $\sim$10%. However, regularization error will be high for such instances of excessive regularization (1% covariance), which is discussed in the following.

Another complexity for regularization arises from the inherent properties of the chosen stabilizer. It can be confirmed (see Fig. 4) that increasing the regularization strength stabilizes the solution but concomitantly reduces the information by introducing regularization error. Thus, the type of stabilizer and optimal regularization strength for all iterative steps are the most important factors for obtaining a unique solution. RTLS inherently calculates the required strength of regularizations at all iterations using LFDO as the stabilizer, which injects very low regularization error by its own characteristics for profile retrievals and drastically reduces the condition number of the inverted matrix. We find that the value of the square root of the condition number of the RTLS inverted matrix is never more than 15. In contrast, using covariance matrices of *a priori* and measurement error as a stabilizer for OEM does not appear to be functioning effectively.

To better understand the total information content in such retrievals, DFR at the last iteration for three different profiles are calculated using (9). It is found that the information content is dependent on the shape of the profile, and the values of DFR for three different profiles are 3.88, 2.61, and 2.68, respectively. Moreover, the altitudinal information varies with the shape of the profile as shown in Fig. 5, which is the row space distribution of the regularized kernel matrix $\mathbf{R}_{\mathrm{km}}$ at the last iteration. Two peaks are observed for profile 1 at around 900 mb and 400–600 mb; however, the negative value of $\mathbf{R}_{\mathrm{km}}$ for the height of 400–600 mb raises concern about whether it is real information or pseudoinformation. In contrast, the other two profiles have very little information for height of 400–600 mb. More research is required to understand and resolve these issues.

There might be some concerns that the aforementioned numerical experiments are unfair because the OEM method is based on an *a priori* and to set up a problem with a range of 100% *a priori* variance renders the *a priori* information content meaningless. Second, according to the physics of Earth's atmosphere, the two sinusoidal profiles and *a priori* profile (IG1) do not exist in reality [note that we also have used realistic IGs and TPs]. However, we have purposefully done this simulated experiment to understand the RT inverse function, and it does not violate any limits from the point of RT physics. It is quite obvious that parameters can go beyond the boundary for a specific iteration when Newtonian iterative optimization is used in a nonlinear problem. Moreover, purposefully choosing sinusoidal profiles gives us an additional advantage to understand and analyze altitudinal information content for such measurements (see Fig. 5).

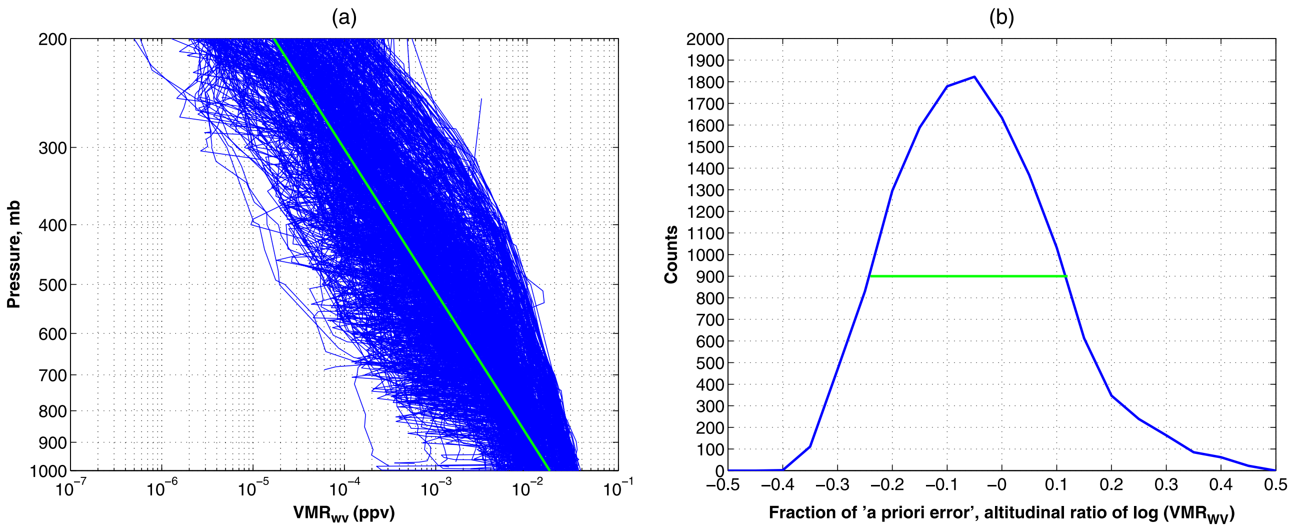

Fig. 6. (a) Shape of radiosonde profiles (blue) used in this study and assumed IGprofile (green); (b) histogram of the logarithmic of percentage of departure all truth from the IG; $\mathrm{VMR_{wv}}$ stands for volume mixing ratio of WV.

C. Simulated Retrievals Using Radiosonde Profiles

To address any potential concerns due to selection of unrealistic profiles so far, we conducted retrievals using the Forecast Systems Laboratory (FSL) radiosonde database (http://www.esrl.noaa.gov/raobs/intl/fsl_format-new.cgi) representative of the Earth's atmosphere, in conjunction with other collocated *in situ* parameters (e.g., SST). We have collected more than 400 profiles from this database to perform this simulation and assumed emissivity of one for all. A plot of all the profiles is shown in Fig. 6(a), and an approximated average profile is considered *a priori* [green line in Fig. 6(a)] for this simulated retrieval study. We have also calculated the *a priori* covariances using FWHM of the departure of these profiles from the assumed *a priori* [see Fig. 6(b)]. The simulation has been made on the grid of the individual radiosonde profiles; thus, the different atmospheric grids are considered for different profile retrievals. First, two good profiles are selected from the *in situ* database (based on arbitrary visual inspection) for WV retrievals from a fixed IG profile that lies within the domain of the realistic WV profiles.

A recent trend has developed to improve stochastic retrievals by applying additional constraints using either a finite difference operator (e.g., MIPAS [39]–[41]) or TSVD (e.g., CrIS [38]) on top of the stochastic methods to reduce the oscillations in the resultant solutions. However, the ambiguities, as we discussed earlier, for the selection of error covariances and the *a priori* dependence still exist for either case. In particular, the estimation of nonlinearity error for such method, which has to be supplied as an input parameter, is very difficult. On the contrary, as mentioned earlier, RTLS inherently determines the regularization strength for all iterations of different values depending on the associated error including nonlinearity for such inversion from the residual. TSVD is often referred to as a reduced state space inverse method and has additional issues. The primary idea is to reduce dimensionality and hence ameliorate the ill-conditioned nature of the problem by means of the reduction of the condition number of the inverted matrix. In this process, a large number of empirical orthogonal functions are discarded from the retrieval space. In contrast, RTLS keeps all of the information for state-space parameters.

The simulation study can only be used to compare the basic inverse methods, and not the current operational algorithms where multiple provisional tuning of both parameters and functions are implemented that cannot be easily replicated in a simulated study. For example, the nonlinearity error is difficult to determine for a particular iteration, but it was calculated in the EDR ATBD [38] by the multiplication of an *ad hoc* constant with the value of the "model minus observation." However, we have conducted a similar experiment using TSVD on top of OEM, where we fixed a 10% *a priori* covariance in OEM for all profiles [shown in Fig. 6(a)], and we consider only the five highest singular values of the inverted matrix $(K^T S_e^{-1} K + S_a^{-1})$. The selected 10% *a priori* covariance for OEM cannot be justified by any standard of stochastic approaches because the FWHM for this set of experiment is 36% as is shown in Fig. 6(b), but it has been considered to achieve a reasonably high regularization into OEM solution from a deterministic point of view, as we discussed before. Different combinations of *a priori* covariance, measurement error covariance, and a number of higher singular values can be made. While this may have limited scientific purpose to it, we have considered the aforementioned combination for demonstration only.

Fig. 7 shows retrieval results for the two selected profiles using "OEM *plus* TSVD" (OEM-T) and RTLS. A realistic solution is again obtained using RTLS even for the increased number of retrieval levels, whereas OEM-T retrieved profiles are unable to produce a decent solution. Fig. 7 demonstrates that after discarding some state space corresponding to small singular values, OEM-T is still not able to get reasonable results; however, RTLS is capable of managing this by virtue of its dynamic data driven regularization without discarding any information from the retrieval model. It may be possible to achieve slightly improved results by employing a closer *a priori* and/or higher measurement error covariance and/or low *a priori* covariance than expected by employing a trial and error approach as part of the OEM-T for a particular profile or

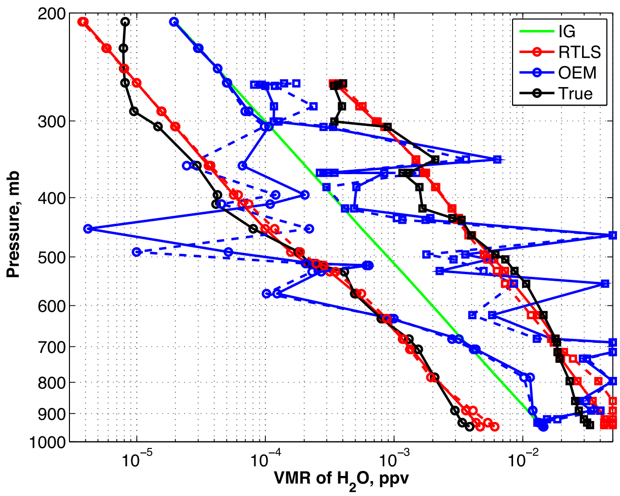

Fig. 7. Comparative simulated WV profile retrievals using RTLS (red) and 'OEM plus TSVD' (blue). The IG and *a priori* are the same (green); the truth is shown in (black). Two profiles are separated using symbols of circle and square. Perfect and perturbed ancillary data are separated by solid and dashed lens.

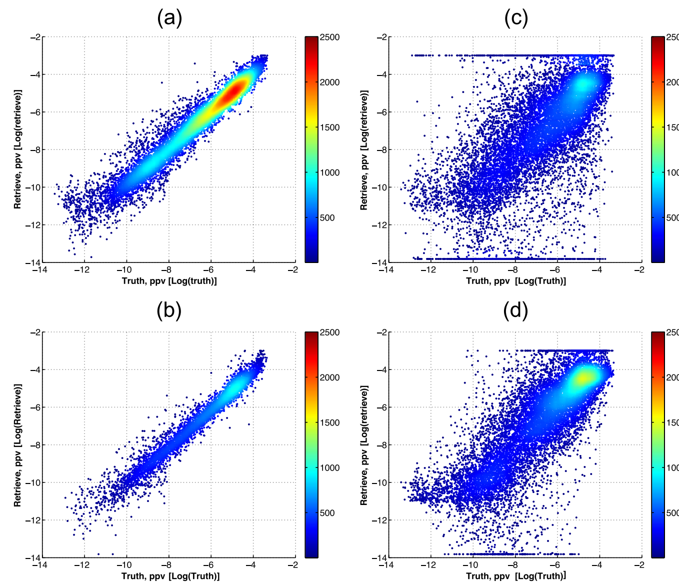

Fig. 8. Comparative density scatter plots for RTLS and OEM retrievals using 424 radiosonde profiles: (a) RTLS (finer grids), (b) RTLS (coarser grids), (c) OEM plus TSVD (finer grids with 10% *a priori* covariance), and (d) OEM plus TSVD at finer grids with 1% *a priori* covariance.

data set. However, this may fail for another case because such tuning is unlikely to be fully objective and sufficiently general in nature.

The question may arise at this point that the perfect ancillary data used here will not available in an operational environment. This issue can be addressed in two different ways: 1) simultaneous retrieval, which will be difficult sometimes due to increase of the condition number of Jacobian; 2) a part-solution, adding errors in ancillary data. To gain confidence in such cases, we perturbed the atmospheric temperature profile with 1 K and surface temperature with 0.5 K, randomly, and used those in the simulation study as the "truth" as shown in Fig. 7. The retrieval results from OEM-T are significantly different and sometimes opposite in phase from the truth, but no noticeable difference is found for RTLS solutions due to these additional perturbations. Rather simply, it is because the regularization strength alters automatically due to the increased data misfit resulting from the errors we purposely injected into the profile. In such a situation, the regularization will be higher than the original with slightly reduced information content (e.g., the value of DFR changes from 4 to 3.8) but produce a stable solution. Near to the surface, a little difference is observed that is attributed to low information content at this level. More thorough research is required to increase the information content through a choice of better channel selection under the deterministic paradigm.

In operational IR retrieval, three main factors are involved: 1) inverse method; 2) forward model; and 3) cloud detection. There are two important points to be considered here. First, in this paper, we have presented a comparative study of inverse methods considering the forward model and the cloud detection to be perfect. We have also demonstrated the sensitivity of comparative results for different inverse methods due to forward model errors. Second, in our previous work [43], [63], we have demonstrated operational geophysical parameter estimation using the same family of inverse method including the ambiguities of the forward model and the cloud detection. However, the earlier work dealt with a modestly ill-conditioned inversion, whereas this paper solves a highly ill-conditioned problem. The error propagation into retrieval space is proportional to the condition number of Jacobian for any inverse problem, which is scientifically demonstrated using a mathematical derivation [42]. Thus, the condition number of the Jacobian is a key factor for any inverse problem but not for forward model or cloud detection problem. Combining these two points, one can reasonably conclude that the RTLS method can be a viable contender for operational satellite remote sensing inverse applications.

Simulated retrievals of more than 400 profiles from the aforementioned FSL database are made using both RTLS and OEM-T from a fixed IG (as in Fig. 7), and the density scatter plot is shown in Fig. 8. As we observed in Fig. 4, the close to *a priori* and smooth profile can be solved by OEM using 1% *a priori* covariance; thus, we have added an additional experiment by reducing the *a priori* covariance to 1% (cf. the previous OEM-T). Using multiple real-life profiles extends the scope of the experiment beyond the initial theoretical study using selected profiles and strengthens our confidence in the RTLS method. Fig. 8 shows that RTLS can retrieve parameters (with associated error), but these have far fewer errors than those from OEM-T solution. One may argue here that the choice of *a priori* and *a priori* covariance in this experiment is not "favorable" for optimal working of the OEM-T. However, in reality it is also harder to get these parameters accurately because of highly dynamic atmospheric conditions, which are not at all required for obtaining good solutions using RTLS.

As discussed above, OEM-T is often unable to produce reasonable solutions due to limitations of its fundamental understanding. However, some operational validation results of CrIS WV profiles, retrieved using stochastic approach (OEM-T), have been recently published [19], [23], which report retrieval errors against radiosonde observations varying between

35–80% at different altitudes. On the other hand, our simulated retrieval error using 1% *a priori* covariance in OEM-T is higher than those results, which inspires us to understand the additional *ad-hoc* tweaking performed in the operational realm. It has been observed that OEM-T is highly popular to obtain seemingly meaningful results, and additional constraints are enforced using potentially ambiguous scientific assumptions, such as, "model *minus* observation" bias correction, and tuning either *a priori* or measurement error covariance matrices or both. From a deterministic point of view, "model *minus* observation" bias correction may result in alteration of functional physics relationships, Jacobian mismatch error, and information/residual loss [43]. In addition to all of these, if the radiance difference between the observation and calculation for a particular spectral channel does not agree to the expected value within the instrument noise plus forward model error, then an additional error term is added to the corresponding diagonal element of the measurement error covariance [19]. However, such an approach is neither rigorously justifiable nor required in our simulated retrieval experiments where all the parameters are known.

The most interesting aspect of these validation studies is that they have been made using tuned reference data sets with many ambiguous mathematical constraints and after discarding a significant number of retrievals for the sake of quality control (e.g., Chi-squares test). The primary concern here is manipulation of the reference data set: First, it has been altered for the sake of uniformity of vertical gridding between reference data sets and retrieval grids by introducing an inverse method [23]

$$\widetilde{\boldsymbol{x}} = I\left[f(\boldsymbol{x}, b), b, c\right] \tag{16}$$

where $\boldsymbol{x}$ and $\widetilde{\boldsymbol{x}}$ are the true and retrieved state vectors of radiosonde WV profile, respectively, f is the forward model with parameters b (e.g., spectroscopy), and I is the inverse model (i.e., retrieval). There is a strong possibility to introduce an error into the reference data set using another ill-conditioned inversion. Alternately, it will be less ambiguous to do this by using a simple gas law, i.e., Curtis–Godson approximation [64], [65]. One can argue that the spectroscopic measurement based on RT is vulnerable to atmospheric grid spacing, but why this is implemented for validation purposes is unclear. If the primary use of WV profile data is for climate and weather studies, the Curtis–Godson approximation using the gas law should suffice. Second, tuning of the reference data set is also performed using the *a priori* profile (x_a) and the averaging kernel $(\mathbf{A})$ that are used for the satellite-based retrieval employing OEM-T [23] as

$$\boldsymbol{x}_{mr} = \mathbf{A}(\boldsymbol{x}_r - \boldsymbol{x}_a) + \boldsymbol{x}_a$$
$$\mathbf{A} = \left(\mathbf{K}^T\mathbf{S}_e^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{K}^T\mathbf{S}_e^{-1}\mathbf{K} \tag{17}$$

where $\boldsymbol{x}_{mr}$ is the reference data set after second tuning, and $\boldsymbol{x}_r$ is the same after first tuning. For example, as observed in [59, Fig. 9], there is no scientifically justifiable mutual validity between the reference data sets after and before multiplying it with $\mathbf{A}$. There does not seem to be a firm scientific basis to alter the reference data set using (17), other than to improve validation statistics, which may impede further scientific progress. This tuning goes even one step further: A third tuning is made

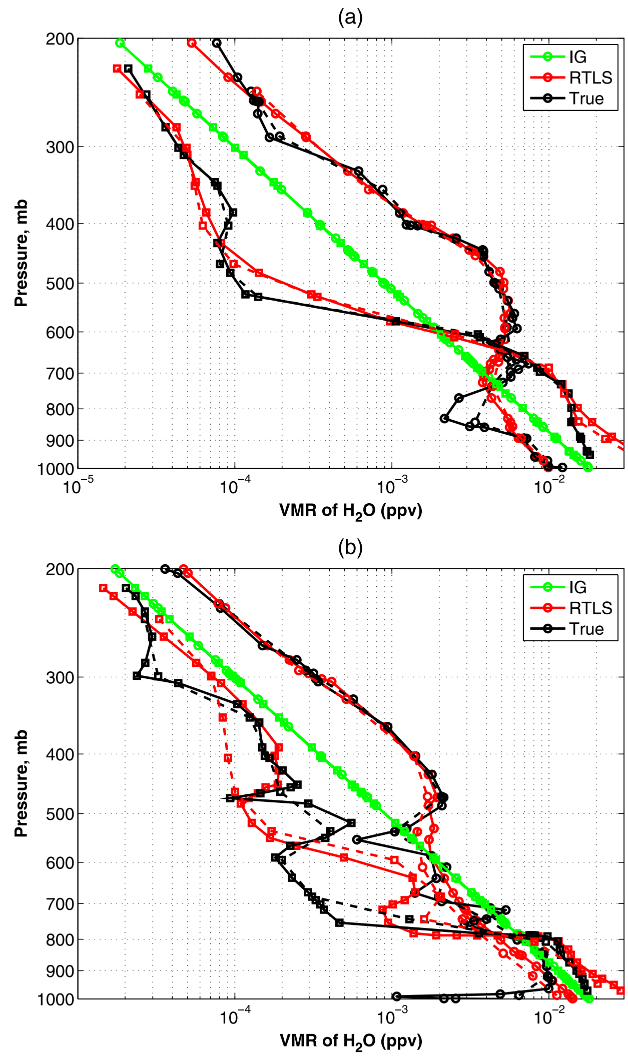

Fig. 9. (a) and (b) Comparative simulated WV profile retrievals using RTLS (red) for finer (solid) and coarser (dashed) atmospheric grids. The IG is the same (green); the truth is shown in (black). Two profiles are separated using square and circle lines, respectively.

using trapezoidal basis functions (F) to map the number of retrieval grids (m) perturbations to forward model perturbations on the finer grids $n(m < n)$ forward model layers [23]

$$\mathbf{A_c} = \mathbf{FAF}^{+}; \quad \mathbf{F}^{+} = \left(\mathbf{F}^{\boldsymbol{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\boldsymbol{T}}. \tag{18}$$

Smoothing of the correlative profiles x is then achieved for by substituting $\mathbf{A_c}$ for $\mathbf{A}$ in (17). This alteration of the reference data set ought not to be acceptable from a scientific viewpoint. A scientifically compliant alternative could be to interpolate (using the aforementioned techniques or by applying the simple gas law) the retrieved data set onto the reference grid for the purposes of validation. The main concern that we raise in this paper is that the aforementioned unobjective controls on the reference data set to obtain seemingly meaningful numbers from validation may compromise the true potential of expensive satellite missions.

The spread of RTLS solutions [see Fig. 8(a)], even if better than that of OEM-T, is still quite high. A possible reason is that it is impossible to solve all grid points of radiosonde data for altitudes where the grid spacing is less than 100 m in some profiles. It is also observed that there are many profiles that are difficult to be solved using any inverse method. To understand these issues, another numerical experiment has been performed where the simulated measurements have been constructed at finer grids of radiosonde data, but the problem is solved using coarser grids [see Fig. 8(b)]. Once the synthetic measurements (using finer grid) are obtained, for inverse purposes, the RT calculation has been made at coarser grids using Curtis–Godson approximation, which also allows us to gain confidence in RTLS retrieval in the presence of forward model error. (note that, essentially, there is no forward model error in simulated retrievals in the previously shown demonstrations because the grids are same for both forward and inverse models). The broad motive is to use satellite retrieved data to understand atmospheric phenomena that are governed mainly by thermodynamic physics, which is what the Curtis–Godson approximation is based on. Therefore, there should not be any scientific issue to use such averaging. It is observed that the average forward model error for the coarser grid is comparable to the SNR of the measurement (100). However, this process smoothes the radiosonde profiles using a simple gas law relation (Curtis–Godson approximation) as well as reducing the condition number of the Jacobian.

We have picked two other profiles as shown in Fig. 9(a), which are relatively difficult to solve accurately compared with the profiles shown in Fig. 7. The number of atmospheric grids of two different profiles shown in Fig. 9(a) is 41 and 18. We have reduced the number of grid points for both the profiles to 12 using the Curtis–Godson approximation. The forward model differences of radiances between the coarser grids and original grids in terms of the rmse are 1% and 1.2%. Despite this, the retrievals using RTLS are more or less identical in shape for either the coarse or fine grid [see Fig. 9(a)], but the error statistics are lower in the coarser grid due to the averaging of scatter points in a layer. These two profiles (close to sinusoidal) were chosen purposefully from radiosonde measurements is to gain insight into the dependence of *a priori* in the final solution. The results show that the RTLS solution is essentially independent of IG or *a priori* and grid spacing, which we already demonstrated using arbitrary profiles.

The shape of profile 1 in Fig. 9(a) (solid circle) along the altitude is approximately a full sinusoid in nature, and at least five pieces of information are needed to get a representative retrieval profile apart from the spikes. It is observed that the applied smoothing (coarser grids) does not reduce the degrees of freedom for these profiles, and the retrieved profiles for both finer and coarser grids are close to the sinusoidal. In contrast, for profile 2 in Fig. 9(b), the TP is more than one sinusoid in nature and more than five pieces of information are needed, at the least, to get a representative retrieval profile. It is almost impossible to solve this by the present retrieval model, where around three to four pieces of information are available. On the other hand, some points of profile 1 in Fig. 9(b) are abruptly distributed, and smoothing using the Curtis–Godson approximation can improve the shape of the profile as a result the reduction of maximum retrieval error or spread of the error.

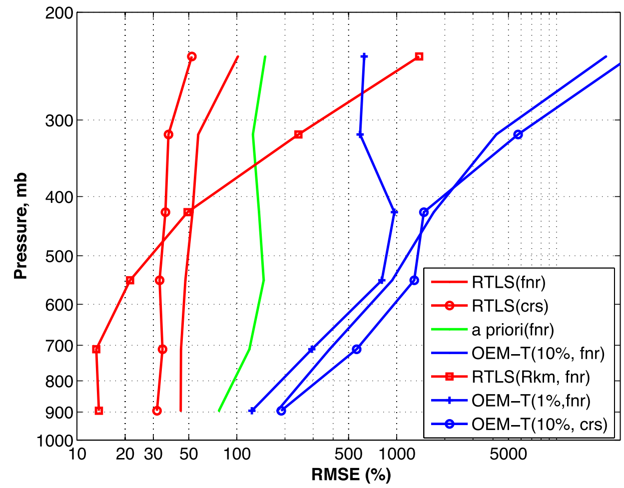

Fig. 10. RMSE along the altitude (average of 2 km grid) of more than 400 profiles for RTLS with finer grids (red solid, fnr), RTLS with coarser grids (red circle, crs), RTLS R_{km} multiplied by in situ (red square), IG with finer grids (green solid), OEM-T with finer grids with 1% *a priori* covariance (blue plus) and OEM-T finer grids with 10% *a priori* covariance (blue solid), and OEM-T coarser grids with 10% *a priori* covariance (blue circle).

In order to further understand the RTLS spread in Fig. 8(a), we performed an error calculation for two different retrievals. As the values of the WV profile varies by more than two orders of magnitude, we have considered the percentage of error as $(\delta x/x)$ for individual points as described by Maddy and Barnet [22]. The percentage of rmse for average of all grids of these retrievals [see Fig. 9(a)] using the original grids are 37% and 57%, and for coarser grids are 25% and 50%. However, the absolute maximum deviation reduces from 153% to 68% for profile 1 (circle) and from 163% to 138% for profile 2 (square) using the coarser grid. Both rmse and the spread of the RTLS solution are reduced in coarser grids for these profiles as observed in Fig. 8(b) [Note that the number of points in Fig. 8(b) is 50% lower than the same of Fig. 8(a)].

The rmse of all points for a given profile in Fig. 9(b), which are more difficult profiles to solve, are 232% and 124% for finer grids and 31% and 207% for coarser grids, for profiles 1 and 2 respectively. For profile 2, it is observed that the error for coarser grids is higher than for finer grids, as well as the spread is also increased from 485% to 560% [on the contrary, for profile 1, the spread reduced significantly from 1180% (finer grids) to 74%, as similar to Fig. 9(a)]. It is also observed that shapes of retrieved profile 2 in Fig. 9(b) are different for two different grids. However, it is observed that the retrieval errors in term of the column density for both the profiles are lower in coarser grids. The calculated column density errors for both profiles are 5% and 25% at finer grids and 5% and 13% at coarser grids, respectively.

Fig. 10 shows the 2 km average [23] rmse of RTLS and OEM-T for the finer and coarser grids without tuning the reference 'sonde data. Any form of the RTLS error is always less than the *a priori* error. The rmse values of RTLS using finer and coarser grids are ∼50% and ∼35%, respectively, which seem to be high. These numbers have no absolute meaning because these statistics can be changed by the selection data set (e.g., the number will be reduced if major profiles are similar to the profiles selected in Fig. 7, where as it will be further increased if the major profiles are similar to those selected in Fig. 9). The profiles used here are globally representative snapshot of WV profiles from radiosonde measurements. These statistics can be improved by discarding the complex profiles from this data set as bad retrievals using the total error calculation at the solution time, which is already discussed in operational SST retrieval [43], [65] publications. The objective of this paper is not the reduction of statistics; rather it is a comparative study to select the better inverse method. It is observed from this paper that the errors of coarser grids RTLS ("red circle") are $\sim$30% lower than for the finer grid ("red dot"). Despite the increased forward model error (spectroscopic error), averaging over points in the coarser grid reduces retrieval error significantly compared with the finer grid as expected. Surprisingly, error is not reduced for OEM-T using coarser grids ("blue circle") as compared with finer grids ("blue dot"). The specification of the coarser grid for this paper was arbitrarily chosen, and a detailed analysis is required for an optimal selection of atmospheric gridding for a forward model in the deterministic inverse framework.

The value of the error for finer grid RTLS can be reduced by tuning the reference data set employing (17) and replacing the A by $\mathbf{R}_{\rm km}$, which is shown ("red square") in Fig. 10(a). The RTLS error from tuned reference data is reduced from 44–47% to 13–20% ($\sim$2.5 times) up to a height of 550 mb. While such accuracies for WV retrieval are essentially unprecedented, particularly for such a generalized profile set with such high-frequency components, we would like to reiterate that such "manipulated" error calculations can only assist in reporting a "reduced number," whereas the true error of the retrieved data set remains unchanged. In practice, such approaches are widely used, the scientific basis of which we consider somewhat questionable.

The RMSE of OEM-T ("blue dot" with 10% *a priori* error) is approximately an order of magnitude more than the *a priori* error in the same scale. The OEM-T error, even using 1% *a priori* covariance for this paper that is more than *a priori* error, is very high compared to some recently published operational validation statics [23]. This can be explained as we have not applied any tuned validation approach, and as we mentioned earlier, it is unfeasible to simulate all operational conditions, e.g., dynamic variations both measurement adjustments and the values of measurement error covariance using the same "model *minus* observation," which will also be used in the retrieval (cf. [19]). From the deterministic point of view, "model *minus* observation" has one piece of information at pixel level, and the use of this value in three operations does not increase the information content. Moreover, operational algorithms use both *a priori* profile and covariance that are different for different profile retrievals, as opposed to this paper where both these parameters are fixed. Such validation studies (e.g., [18] and [22]) often consider the *a priori* profile to be very close to the truth within 2–3%, which translates to overstating the success of the stochastic method in absolute terms and may end up showing retrieval errors of 35–80% for a subselected data set [23]. Under such a circumstance, the retrieval error of OEM-T is therefore effectively 10–15 times higher than the *a priori* error as implied by the error covariance used in the inverse method, which is not dissimilar to the finding of this paper. We have also recently reported similar results that OEM retrieval error [43], [63] for operational SST (which is a fairly linear problem where the condition number of Jacobian is less than 10) is higher than the *a priori* error. Many trace gases retrieval studies show that the selected *a priori* within the domain of *a posterior* error, which implies that the *a priori* error is less than posterior error [59].

V. Conclusion

For more than two decades, operational retrievals from satellite hyperspectral IR measurements have been dominated by methods that are based on stochastic approaches where many ambiguities are ubiquitous. This paper addresses such issues, e.g., the truth is effectively known when both *a priori* and *a priori* errors are known, by analyzing the inverse problem from a deterministic point of view. To overcome these problems, we propose an alternate method, which is in the family of deterministic inverse methods, namely, RTLS.

Based on the findings of this paper by theoretical discussion and simulated experiments, we conclude that RTLS is one of the most suitable inverse methods, which can be used for any highly nonlinear remote sensing problem for extraction of quantitative information. On the other hand, stochastic inversion techniques are unable to produce unambiguous scientific information. Moreover, without implementing a series of seemingly unobjective *ad hoc* corrections, results from stochastic methods often yield more errors than are present in the initial guess.

A number of remote sensing hyperspectral instruments are onboard many research and operational satellites to understand the Earth's atmosphere and for climate studies. These instruments provide an enormous amount of scientific data. Our proposed inverse method offers a paradigm shift in operational remote sensing and may open up a new scientific frontier for unambiguously converting data to information from existing and future missions. This paper is an initial step, showing a comparative study using a relatively simple model. Solutions for hyperspectral atmospheric profiles retrievals using RTLS can be further improved by in-depth understanding of deterministic inverse methods and information content analysis of the retrieval model, as well as optimization of the atmospheric grids, spectral windows, and channel selection for discretized physical models.

Acknowledgment

The authors would like to thank two anonymous reviewers for their useful suggestions and discussions, and E. Maturi for the arrangement of facilities at NOAA to complete this work. P. Koner would like to thank Prof. J. Drummond, Dalhousie University, Canada, for the insightful discussions and inspiration to build up a paradigm shift idea in the satellite inverse problem when he was working with him. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official NOAA or U.S. Government position, policy, or decision.

REFERENCES

[1] J. L. Mead, "Parameter estimation: A new approach to weighting a priori information," *J. Inv. Ill-Posed Problems*, vol. 15, pp. 1–21, 2007.

[2] R. Zhang, M. K. Sen, S. Phan, and S. Srinivasan, "Stochastic and deterministic seismic inversion methods for thin-bed resolution," *J. Geophys. Eng.*, vol. 9, pp. 611–618, 2012.

[3] M. A. Aguiló, L. Swiler, and A. Urbina, "An overview of inverse material identification within the frameworks of deterministic and stochastic parameter estimation," *Int. J. Uncertainty Quantification*, vol. 3, pp. 289–319, 2013.

[4] "Algorithm Theoretical Basis Document for Level 2 Processing of the MTG Infra-Red Sounder Data," Doc. No. EUM/MTG/DOC/11/0188, issue: v3, EUMETSAT, 2014. or "IASI Level 2 Products Overview," EUM/OPS-EPS/MAN/04/0033: EUMETSAT; 2012.

[5] "*TES level 2 algorithm theoretical basis document*," Jet Propulsion Lab., Pasadena, CA, USA, JPL D-16474: ver. 1.15, 2002. [Online]. Available: http://eospso.nasa.gov/sites/default/files/atbd/TES.ATBD_.L2.V1.15.pdf

[6] *MIPAS Level 2 Algorithm Theoretical Baseline Document*, 2011. [Online]. Available: https://earth.esa.int/web/sppa/mission-performance/esa-missions/envisat/mipas/products-and-algorithms/products-information/

[7] P. K. Bhartia, *OMI Algorithm Theoretical Basis Document*, 2002. [Online]. Available: http://eospso.nasa.gov/sites/default/files/atbd/ATBD-OMI-02.pdf

[8] [Online]. Available: http://disc.sci.gsfc.nasa.gov/AIRS/documentation/20070301_L2_ATBD_signed.pdf

[9] M. Ridolfi *et al.*, "Optimized forward model and retrieval scheme for MIPAS near-real-time data processing," *Appl. Opt.*, vol. 39, no. 8, pp. 1323–1340, Mar. 2000.

[10] J. Susskind, C. Bamet, and J. Blaisdell, "Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 2, pp. 390–409, Feb. 2003.

[11] P. K. Koner and J. R. Drummond, "Atmospheric trace gases profile retrievals using nonlinear regularized total least squares method," *J. Quantitative Spectr. Radiative Transfer*, vol. 119, pp. 2045–2059, 2008.

[12] S. Del Bianco, B. Carli, C. Cecchi-Pestellini, B. M. Dinelli, M. Gaia, and L. Santurri, "Retrieval of minor constituents in a cloudy atmosphere with remote-sensing millimetre-wave measurements," *Quart. J. Roy. Meteorol. Soc.*, vol. 133, no. S2, pp. 163–170, Oct. 2007.

[13] D. Weidmann, W. J. Reburn, and K. M. Smith, "Retrieval of atmospheric ozone profiles from an infrared quantum cascade laser heterodyne radiometer: Results and analysis," *Appl Opt.*, vol. 46, pp. 7162–7171, 2007.

[14] C. D. Rodgers, "Information content and optimisation of high spectral resolution remote measurement," *Adv. Space Res.*, vol. 21, pp. 361–367, 1998.

[15] S. L. Tristan, P. Gabriel, K. Leesman, S. J. Cooper, and G. L. Stephens, "Objective assessment of the information content of visible and infrared radiance measurements for cloud microphysical property retrievals over the global oceans. Part I: Liquid clouds," *J. Appl. Meteorol. Climatol.*, vol. 45, pp. 20–41, 2006.

[16] H. L. Huang and R. J. Purser, "Objective measures of the information density of satellite data," *Meteorol. Atmos. Phys.*, vol. 60, pp. 105–117, 1996.

[17] M. Reuter, M. Buchwitz, O. Schneising, J. Heymann, H. Bovensmann, and J. P. Burrows, "A method for improved SCIAMACHY CO2 retrieval in the presence of optically thin clouds," *Atmos. Meas. Tech.*, vol. 3, pp. 209–232, 2010.

[18] J. Worden *et al.*, "Predicted errors of tropospheric emission spectrometer nadir retrievals from spectral window selection," *J. Geophys. Res.*, vol. 109, no. D9, 2004, Art. no. D09308, doi: 10.1029/2004JD004522.

[19] M. Divakarla *et al.*, "The CrIMSS EDR algorithm: Characterization, optimization and validation," *J. Geophys. Res. Atmos.*, vol. 119, pp. 4953–4977, 2014, doi: 10.1002/2013JD020438.

[20] H. Ohyama *et al.*, "Atmospheric temperature and water vapor retrievals from GOSAT thermal Infrared spectra and initial validation with coincident radiosonde measurements," *Sci. Online Lett. Atmos.*, vol. 9, pp. 143–147, 2013.

[21] M. G. Divakarla *et al.*, "Validation of atmospheric infrared sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts," *J. Geophys. Res.*, vol. 111, no. D9, May 2006, Art. no. D09S15, doi: 10.1029/2005JD006116.

[22] E. S. Maddy and C. D. Barnet, "Vertical resolution estimates in version 5 of AIRS operational retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2375–2384, Aug. 2008.

[23] N. R. Nalli *et al.*, "Validation of satellite sounder environmental data records: Application to the croos-track Infrared Microwave sounder suite," *J. Geophys. Res. Atmos.*, vol. 118, no. 24, pp. 628–643, Dec. 2013.

[24] X. Xiong, C. Barnet, E. S. Maddy, A. Gambacorta, T. S. King, and S. C. Wofsy, "Mid-upper tropospheric methane retrieval from IASI and its validation," *Atmos. Meas. Tech.*, vol. 6, pp. 2255–2265, 2013.

[25] N. Pougatchev *et al.*, "IASI temperature and water vapor retrievals—Error assessment and validation," *Atmos. Chem. Phys.*, vol. 9, pp. 6453–6458, 2009.

[26] A. Tikhonov, "On the solution of ill-posed problems and the method of regularization," *Dokl. Akad. NaukSSSR*, vol. 151, no. 3, pp. 501–504, 1963.

[27] A. Beck and A. Ben-tal, "On the solution of the Tikhonov regularization of the total least squares problem," *SIAM J. Optim.*, vol. 17, pp. 98–118, 2006.

[28] J. Lampe and H. Voss, "On a quadratic eigen-problem occurring in regularized total least squares," *Comput. Stat. Data Anal.*, vol. 52, pp. 1090–1102, 2007.

[29] D. M. Sima, S. VanHuffel, and G. H. Golub, "Regularized total least squares based on quadratic eigenvalue problem solvers," *Numerical Math.*, vol. 44, pp. 793–812, 2004.

[30] Markovsky and S. VanHuffel, "Overview of total least squares methods," *Signal Process.*, vol. 7, pp. 2283–2302, 2007.

[31] V. Z. Mesarovic, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized constrained total least squares image restoration," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1096–1108, Aug. 1995.

[32] W. Chen, M. Chen, and J. Zhou, "Adaptively regularized constrained total least-squares image restoration," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 588–596, Apr. 2000.

[33] N. Mastronardi, P. Lemmerling, A. Kalsi, D. P O'Leary, and S. Van Huffel, "Implementation of the regularized structured total least squares algorithms for blind image deblurring," *Linear Algebra Appl.*, vol. 391, pp. 203–221, 2004.

[34] H. Fu, M. K. Ng, and J. L. Barlow, "Structured total least squares for color image restoration," *SIAM J. Sci. Comput.*, vol. 28, pp. 1100–1119, 2006.

[35] J. Lei *et al.*, "An image reconstruction algorithm based on the regularized total least squares method for electrical capacitance tomography," *Flow Meas. Instrum.*, vol. 19, pp. 325–330, 2008.

[36] J. Susskind, C. Bamet, and J. Blaisdell, "Determination of atmospheric and surface parameters from simulated AIRS/AMSU/HSB sounding data: Retrieval and cloud clearing methodology," *Adv. Space Res.*, vol. 21, pp. 369–384, 1998.

[37] E. S. Maddy, C. D. Barnet, and A. Gambacorta, "A computationally efficient retrieval algorithm for hyperspectral sounders incorporating *A Priori* information," *IEEE Geosci. Remote Sens. lett.*, vol. 6, no. 4, pp. 802–806, Oct. 2009.

[38] *ATBD for CrIS, Vol II, EDR, GSFC JPSS CMO*, Nov. 16, 2012.

[39] M. Ridolfi and L. Sgheri, "A self-adapting and altitude-dependent regularization method for atmospheric profile retrievals," *Atmos. Chem. Phys.*, vol. 9, pp. 1883–1897, 2009.

[40] J. Steinwagner and G. Schwarz, "Shape-dependent regularization for the retrieval of atmospheric state parameter profiles," *Appl. Opt.*, vol. 45, pp. 1000–1009, 2006.

[41] M. Ridolfi and L. Sgheri, "Iterative approach to self-adapting and altitude-dependent regularization for atmospheric profile retrievals," *Opt. Exp.*, vol. 19, pp. 26 696–26 709, 2011.

[42] P. K. Koner, A. Battaglia, and C. Simmer, "A rain rate retrieval algorithm for attenuating radar measurement," *J. Appl. Meteorol. Climatol.*, vol. 49, pp. 381–393, 2010.

[43] P. K. Koner, A. R. Harris, and E. Maturi, "A physical deterministic inverse method for operational satellite remote sensing: An application for SST retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5872–5888, Nov. 2015.

[44] A. Legendre, "Nouvelles méthodes pour la détermination des orbites des comètes," in *New Methods for the Determination of the Orbits of Comets (in French)*. Paris, France: F. Didot, 1805.

[45] B. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*. New York, NY, USA: McGraw-Hill; 1978.

[46] G. Golub and C. VanLoan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, pp. 883–893, 1980.

[47] G. Golub, "Some modified matrix eigenvalue problems," *SIAM Rev.*, vol. 15, pp. 318–344, 1973.

[48] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary, "Regularization by truncated total least squares," *SIAM J. Sci. Comput.*, vol. 18, pp. 1223–1241, 1997.

[49] P. K. Koner and J. R. Drummond, "A comparison of regularization techniques for atmospheric trace gases retrieval," *J. Quantitative Spectr. Radiative Transfer*, vol. 109, pp. 514–526, 2008.

[50] T. Steck, "Methods for determining regularization for atmospheric retrieval problems," *Appl. Opt.*, vol. 41, pp. 1788–1796, 2002.

[51] W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory*. San Diego, CA, USA: Academic, 1989.

[52] C. D. Rodgers, *Inverse Methods for Atmospheric Soundings: Theory and Practice*. Singapore: World Scientific, 2000.

[53] J. Susskind, J. Rosenfield, and D. Reuter, "An accurate radiative transfer model for use in the direct physical inversion of HIRS2 and MSU temperature sounding data," *J. Geophys. Res.*, vol. 88, pp. 8550–8568, 1983.

[54] B. M. Quine and J. R. Drummond, "GENSPECT a line-by-line code with selectable interpolation error tolerance," *J. Quantitative Spectr. Radiative Transfer*, vol. 74, pp. 147–165, 2002.

[55] R. H. Norton and R. Beer, "New apodizing functions for Fourier spectrometry," *J. Opt. Soc. Amer.*, vol. 66, pp. 259–264, 1976.

[56] A. Berk *et al.*, MODTRAN4.3 Users Manual, 2003.

[57] D. Tobin *et al.*, "Suomi-NPP CrIS radiometric calibration uncertainty," *J. Geophys. Res. Atmos.*, vol. 118, no. 18, pp. 10 589–10 600, Sep. 2013.

[58] Y. Han *et al.*, "Suomi NPP CrIS measurements, sensor data record algorithm, calibration and validation activities, and record data quality," *J. Geophys. Res. Atmos.*, vol. 118, pp. 12 734–12 748, 2013.

[59] J. P. Lopez *et al.*, "TES carbon monoxide validation during two AVE campaigns using the Argus and ALIAS instruments on NASA's WB-57F," *J. Geophys. Res.*, vol. 113, 2008, Art. no. D16S47, doi: 10.1029/2007JD008811.

[60] K. W. Bowman *et al.*, "Tropospheric emission spectrometer: Retrieval method and error analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1297–1307, May 2006.

[61] J. T. Houghton, F. W. Taylor, and C. D. Rodgers, *Remote Sounding of Atmospheres*. Cambridge, U.K.: Cambridge Univ. Press: 1984.

[62] A. Doicu, T. Trautmann, and F. Schreier, *Numerical Regularization for Atmospheric Inverse Problems*. Berlin, Germany: Springer-Verlag, 2010.

[63] P. K. Koner, A. R. Harris, and E. Maturi, "Hybrid cloud and error masking to improve the quality of deterministic satellite sea surface temperature retrieval and data coverage," *Remote Sens. Environ.*, vol. 174, pp. 266–278, 2016, doi: 10.1016/j.rse.2015.12.015.

[64] A. R. Curtis, "Discussion of 'a statistical model for water vapour absorption'," *Quart. J. Roy. Meteorol. Soc.*, vol. 78, pp. 638–640, 1952.

[65] W. L. Godson, "The evaluation of infrared-radiative fluxes due to atmospheric water vapour," *Q. J. R. Meteorol. Soc.*, vol. 79, pp. 367–379, 1953.

Prabhat K. Koner (M'15) received the Ph.D. degree in solar photovoltaic (SPV) engineering from Indian Institute of Technology, New Delhi, India.

His thesis work involved the design and optimization of SPV arrays for water pumping in addition to a detailed socio/technoeconomic analysis and an analog and digital electronics controller. He has worked as a Postdoctoral Fellow with the European Union Joint Research Centre, Italy, and as a Research Scientist with Queensland University of Technology, Brisbane, QLD, Australia, on data analysis, instrument development, and quality issues of electrical grid systems. However, in the last seven years, he has been actively involved in the field of satellite remote sensing retrievals. Over the last decade, he is deeply involved in research and application of remote measurements applying physically deterministic methods. During this period, he carried his research work in different countries at reputed academic institutes such as the University of Toronto, Toronto, ON, Canada; the University of Bonn, Bonn, Germany; and Dalhousie University, Halifax, NS, Canada. He has also worked on the NASA/ESA Mars Science Orbiter mission, reviewed articles for IEEE journals and has given many invited talks in reputable universities/institutes around the world, including the IEEE Society. Since 2010, he has been jointly affiliated with the Earth System Science Interdisciplinary Centre, University of Maryland, College Park, MD, USA, and NOAA through the Center for Satellite Applications and Research E/RA3, NOAA Satellite and Information Service, College Park.

Andrew R. Harris received the B.Sc. degree in physics and astronomy and the Ph.D. degree in Earth observation from the University College London (UCL), London, U.K.

His Ph.D. study and postdoctoral work were conducted at the Mullard Space Science Laboratory, UCL, where he was involved in the development and testing of calibration and ground support equipment for the European Along-Track Scanning Radiometer (ATSR), as well as subsequent algorithm development. Since November 2002, he has been an Assistant Research Scientist with the Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA. His work at the U.K. Meteorological Office included the development of a system to account for biases in the ATSR due to surface effects, which was adopted by the world-renowned Hadley Centre for Climate Change Research. His recent work has focused on the development of methods to work with the current generation of geostationary and polar-orbiting weather satellites operated by the U.S. National Oceanographic and Atmospheric Administration, as well as the development and implementation of a new high-resolution global sea-surface temperature (SST) analysis. He is a member of the Group for High Resolution Sea Surface Temperature Science Team and the Chair of its Estimation and Retrievals Working Group. He is a member of the NASA SST Science Team. His research interests include the generation of SST products from satellite data, fundamental instrument calibration and characterization in the thermal infrared, and development of high-resolution global SST analyses.

Prasanjit Dash received the Ph.D. degree in physics (infrared radiometric applications) from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

He is a Remote Sensing Scientist with over 17 years of experience in terrestrial infrared satellite applications. He has worked on several national and international space-based projects in USA, Europe, and India. Between 2000 and 2005, He was with KIT, where he worked on the development of algorithm and validation strategy for land surface temperature. Since 2006, he has been with NOAA's Satellite and Information Service, College Park, MD, USA, as an employee of Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO, USA. Recently, he has joined EUMETSAT, Darmstadt, Germany, as a Copernicus Sentinel-3 SST Scientist for a limited duration and remains associated/affiliated with CIRA, Colorado State University and NOAA NESDIS. During the past decades, his research is focused on several aspects of satellite-borne SST: radiative transfer modeling in the infrared, analysis of spectral response function shift, sensitivity of regression algorithms to noise, cloud-mask analysis and calibration/validation. A notable contribution is the development and expansion of the NOAA SST Quality Monitor (SQUAM), in collaboration with some other team members, which routinely monitors several SSTs from different agencies. Currently, along with contribution from other colleagues at EUMETSAT, he is developing a dedicated monitoring system for Sentinel-3 SST products. He is the author of several high-impact peer-reviewed publications.

Dr. Dash has received several major awards and fellowships.