

Article

Improved Atmospheric Correction for Remote Imaging Spectroscopy Missions with Accelerated Optimal Estimation

Jouni Susiluoto *, Niklas Bohn, Amy Braverman, Philip G. Brodrick, Nimrod Carmon, Michael R. Gunson, Hai Nguyen, David R. Thompson  and Michael Turmon

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

* Correspondence: jouni.i.susiluoto@jpl.nasa.gov

Abstract

Space-based imaging spectrometers that monitor the Earth's surface generate vast amounts of data, the processing of which requires fast and accurate retrieval algorithms. Estimating scientifically relevant surface properties from remotely measured radiance data typically involves first inferring spectral surface reflectance from the observed radiance, followed by discipline-specific algorithms to derive scientifically relevant properties. Probabilistic reflectance retrieval algorithms, such as the commonly used optimal estimation (OE), are computationally expensive. Furthermore, the Gaussian assumptions associated with OE have not been fully validated in the context of hyperspectral retrievals. To address these challenges, we introduce accelerated optimal estimation (AOE), a Bayesian algorithm that speeds up the OE reflectance inversion process by up to two orders of magnitude compared to a reference OE implementation (ROE), while also providing improved convergence over a number of selected test targets. We also demonstrate that, under given atmospheric conditions, Gaussian uncertainty estimates from OE-type algorithms are accurate. This is achieved by comparing the OE-type posterior distributions to non-Gaussian ones obtained with Markov chain Monte Carlo (MCMC). Finally, we demonstrate how AOE scales to a larger AVIRIS-NG scene, showcasing its ability to handle complex, large-scale data.

Keywords: imaging spectroscopy; optimal estimation; atmospheric correction



Academic Editors: Bing Zhang, Yongguang Zhang, Yinnian Liu, Liangpei Zhang, Stephan Havemann, Yuwei Chen, Jianxin Jia, Qingli Li, Mingyang Zhang, Yueming Wang and Chenchao Xiao

Received: 16 December 2024

Revised: 1 February 2025

Accepted: 12 February 2025

Published: 14 November 2025

Citation: Susiluoto, J.; Bohn, N.; Braverman, A.; Brodrick, P.G.; Carmon, N.; Gunson, M.R.; Nguyen, H.; Thompson, D.R.; Turmon, M. Improved Atmospheric Correction for Remote Imaging Spectroscopy Missions with Accelerated Optimal Estimation. *Remote Sens.* **2025**, *17*, 3719. <https://doi.org/10.3390/rs17223719>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Current and upcoming remote imaging spectroscopy missions [1–4] are able to obtain high spatial and spectral resolution radiance measurements that are, in turn, used to estimate quantities of interest, such as surface mineral composition [5], vegetation species distribution [6], snow surface albedo and radiative forcing [7], algae concentration [8,9], natural hazards [10], and even emission point sources [11,12]. Furthermore, the US National Academies' 2017 Decadal Survey recommended new observations of geophysical properties associated with surface biology and geology. The Surface Biology and Geology (SBG) mission, currently in its early phases, will incorporate measurement capabilities for both visible and shortwave infrared, as well as thermal infrared bands [4].

The massive amounts of data that the next generation of instruments will produce challenge existing retrieval algorithm implementations. In the past, high-resolution remotely sensed imaging spectroscopy data have been obtained by flying imaging spectrometers on aircraft. Examples include the Airborne Visible/Infrared Imaging Spectrometer Classic (AVIRIS-Classic) [13], and AVIRIS-NG [14]. In contrast to these airborne campaigns, the new

space-based missions will continuously observe over very large numbers of ground footprints. The Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) [8]; the Environmental Mapping and Analysis Program (EnMAP) [2]; the Plankton, Aerosol, Cloud, Ocean Ecosystem Mission (PACE) [15]; and the PRecurso IperSpettrale della Missione Operativa (PRISMA) [3] exemplify this technology. For example, EnMAP envisions making measurements at a 30 m ground sampling distance with a 30 km swath width, and with a daily swath lengths of up to 5000 km [2]. With one hundred million observations daily, a difference of just 0.1 CPU seconds in retrieval speed translates to almost four CPU-months difference in compute time for each day. Therefore, even with shortcuts and processing optimizations [16], these retrievals will be costly.

Retrieval algorithms solve an inverse problem [17], which for surface reflectance retrieval means finding the reflectances and atmospheric states when the radiance measurements are given. Radiative transfer models can solve the opposite, the forward problem: computing the at-sensor radiance for given solar irradiance, surface spectral reflectance, and limited atmospheric (and other) state information [18]. The inverse problem is generally much more difficult to solve than the forward problem. We emphasize that this work presents improvements to how the inverse problem is solved, and does not aim to analyze the correctness of the forward model. The forward function that we use to map the joint surface–atmosphere state to at-sensor radiance is a standard form that has been successfully used in the literature (e.g., [19,20]). However, the new inversion method presented in this paper easily generalizes to more complex forward models via additive linearizable corrections. Furthermore, for hyperspectral missions, finding the surface reflectance is the first part of the full retrieval problem. After solving this first inverse problem, the obtained spectra can be used as inputs to domain-specific algorithms, to obtain information about the actual quantities of interest on the surface. Following NASA nomenclature, we refer to the first stage of this procedure as the Level 2a retrieval (hereafter L2a) and to the second stage retrieval as the Level 2b (hereafter L2b) retrieval.

Traditionally, multispectral, rather than hyperspectral, two-stage algorithms have been applied to correct for the atmospheric distortion of the surface reflectance signal and retrieve surface quantities of interest [21,22]. These algorithms use a small number of strategically located bands to infer the state of the atmosphere [22–24]. A more efficient alternative is to simultaneously correct for atmospheric effects and retrieve L2b quantities. For example, ref. [25] followed this approach in the case of chlorophyll retrievals, and [26] introduced a one-step algorithm to retrieve snow grain size from MODIS radiance measurements. While using domain-specific algorithms has proven effective for multispectral data, hyperspectral data are different, as strategies that do not use all available information do not maximally constrain the quantities of interest. For this reason, some investigators have opted to perform the L2a inversion in one step by jointly retrieving both surface and atmospheric reflectance properties [5,14,19,20,27].

Rigorous probabilistic retrievals that utilize all available information, in principle, allow for accurately reporting the uncertainties of any L2 products, which is necessary for rigorous hypothesis testing [28]. Recent work has shown that accounting for uncertainty information in L2a products improves the quality of subsequent L2b retrievals [29], implying that well-founded statistical models in L2a retrievals will help us more comprehensively understand the signals in L2b data than is possible using non-probabilistic algorithms. A fully probabilistic two-stage retrieval separating L2a and L2b works well, because a full reflectance spectrum with uncertainties serves as an approximate sufficient statistic [30] for the L2b quantity of interest. This means that surface-composition-agnostic Bayesian retrievals retain the information that is relevant for the L2b inference, thereby decoupling

the L2a and L2b stages, while still fully utilizing all the information in the measurements for both stages.

To our knowledge, OE is the only probabilistic, fully Bayesian retrieval algorithm that has been extensively used and validated for surface reflectance retrievals [31]. It is a maximum a posteriori (MAP) estimation method, which finds the most probable state estimate, conditional on radiance data, and calculates a Gaussian approximation of the posterior uncertainties of this state. OE produces a joint estimate of both atmosphere and surface state components, providing a unified procedure for atmospheric correction and surface reflectance estimation [19]. A procedure that uses all available information to retrieve the full state vector is generally desirable when solving ill-posed inverse problems [32], and subsequent research has confirmed that joint estimation produces superior results [19]. The open-source ISOFIT package, version 2.5.0, [19,33] serves as our reference implementation of the OE algorithm. Hereafter, this implementation is termed ROE; we describe it in detail in Section 2.

While OE is a Bayesian inversion method, the posterior uncertainties it produces are exact only for linear forward models with Gaussian noise and priors [31]. To represent non-Gaussian posterior distributions, methods such as MCMC [34] must be used. If performing MCMC simulations is possible despite their very high computational cost, the validity of the OE posterior approximation scheme may be evaluated by comparing the OE results with MCMC retrievals. A preliminary comparison of OE with MCMC for the atmospheric part of the state was performed in [20], but the posterior uncertainty from OE has never been quantitatively compared with MCMC results, and therefore the accuracy of OE reflectance uncertainty estimates remains unknown.

There are additional difficulties in the L2a inverse problem beyond the non-linearity of the forward model. Due to the curse of dimensionality, Bayesian inference in high-dimensional spaces is more difficult than in low-dimensional spaces. This is true both for MAP estimation schemes such as OE and for sampling-based methods such as MCMC. Furthermore, the sheer volume of space-borne imaging spectroscopy measurements places constraints on the available compute time for a single retrieval. Given the downstream applications discussed above, L2a algorithms should ideally be able to utilize what is known about the input uncertainties and propagate them to the inferred surface reflectances.

In this paper, we introduce the accelerated optimal estimation (AOE) algorithm, which is a modification of the OE algorithm discussed above. The AOE algorithm simplifies the joint state estimation problem to a degree that enables probabilistic surface reflectance retrievals in one to two milliseconds on one CPU core—up to 300 times faster than our reference OE implementation. We then use Markov chain Monte Carlo (MCMC) [34] samplers, which are not subject to the linearization constraints of OE, to characterize how the true non-Gaussian posterior distributions should look. We evaluate the correctness of the OE and AOE estimates against the MCMC results using distributional distances.

This manuscript is organized as follows. Section 2 describes optimal estimation, and introduces accelerated optimal estimation and the theory behind it, followed by short descriptions of (1) Markov chain Monte Carlo (MCMC) and (2) how radiative transfer is handled in practice. Section 3 describes the data that were used in our case study, including radiance measurements, prior distributions, and instrument noise models. This is followed by the results in Section 4, which include examining the retrieval performance metrics for the selected target footprints, timing information, comparisons and analyses of MCMC results, and a demonstration of AOE with a larger scene. We discuss the numerical results further in Section 5, followed by the conclusions in Section 6.

2. Methods

Probabilistic methods for inferring surface reflectance from top-of-atmosphere radiance include both point-estimate- and sampling-based inversion methods. In this section, we first review how radiance observations relate to surface reflectance and how optimal estimation fits into the Bayesian paradigm. We then describe the accelerated OE algorithm and Markov chain Monte Carlo methods used in this paper. Finally, we go through the different retrieval configurations used in the experiments.

2.1. Reflectance Retrieval Inverse Problem and OE

2.1.1. Forward Modeling

Let $\mathbf{f}(\mathbf{x})$ be the nonlinear forward model mapping a state vector $\mathbf{x} = [\boldsymbol{\zeta}, \boldsymbol{\rho}]^T$ to the at-sensor radiance vector $\mathbf{y}$. The vectors $\boldsymbol{\zeta} \in \mathbb{R}^m$ and $\boldsymbol{\rho} \in \mathbb{R}^n$ contain the atmospheric and surface parts of the state vector. In the simulations presented in the later sections, the atmospheric part $\boldsymbol{\zeta}$ contains two parameters: columnar water vapor concentration [g cm^{-2}] and aerosol optical thickness (AOT, dimensionless), and the surface part $\boldsymbol{\rho}$ contains the hemispherical-directional reflectance factor (HDRF) [35] for each observed spectral band. This vector $\boldsymbol{\zeta}$ could, in principle, also include any other parameters that need to be inferred, such as those that control aspects of scattering, radiative transfer, or geometry. However, the parameters that are known based on prior knowledge do not need to be included.

The length of $\boldsymbol{\zeta}$ is generally low (in our case, two), whereas $\boldsymbol{\rho}$ contains hundreds of spectral HDRFs. In what follows, we use the symbol “ $:=$ ” whenever we *define* variables, and the regular “ $=$ ” to signify that the left- and right-hand sides are equal.

The reflectance retrieval inverse problem is the problem of finding $\boldsymbol{\zeta}$ and $\boldsymbol{\rho}$ given certain observations $\mathbf{y}$ and the forward model $\mathbf{f}(\mathbf{x})$. To solve this problem using probabilistic methods, we model the observed at-sensor radiance data $\mathbf{y}$ by

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\epsilon}. \quad (1)$$

The model-observation mismatch $\boldsymbol{\epsilon}$ may include terms pertaining to either $\mathbf{y}$ (observation error) or $\mathbf{f}(\mathbf{x})$ (model discrepancy). While good models for calibrating and estimating the former exist, estimating the latter requires top-down methods and is generally hard; see [36] for the only attempt at modeling discrepancy estimation in this context that we are aware of.

To keep the modeling setup simple, we only include the well-known zero-mean Gaussian measurement noise in $\boldsymbol{\epsilon}$, and write

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}). \quad (2)$$

Due to this simplification, $\boldsymbol{\epsilon}$ may now be an underestimate of the true error, and hence the posterior uncertainties may also end up being underestimated. However, this will not affect the validity of the results or the analysis below, since the forms of (1)–(2) would not be affected by including a Gaussian model discrepancy term. As written, the inverse problem is also underdetermined, since we have a larger number of parameters than observations. This is taken care of by the added constraints that the prior distribution provides, as described in Section 2.1.2.

The relationship between the at-sensor reflectance $\boldsymbol{\ell}$ and modeled at-sensor radiance $\mathbf{f}(\mathbf{x})$ is defined as

$$\ell_i := \frac{\pi \mathbf{f}(\mathbf{x})_i}{E_i \cos \theta}, \quad (3)$$

where θ is the solar zenith angle, E is the solar irradiance, and the subscripts i denote a particular channel in the spectra, $i = 1, \dots, n$. The reflectance ℓ depends on the state vector $\mathbf{x}$ by

$$\ell(\boldsymbol{\zeta}, \boldsymbol{\rho})_i := \mathbf{r}_i^\dagger(\boldsymbol{\zeta}) + \frac{\mathbf{t}_i^\dagger(\boldsymbol{\zeta})\boldsymbol{\rho}_i}{1 - \mathbf{s}_i(\boldsymbol{\zeta})\boldsymbol{\rho}_i}, \quad (4)$$

where $\mathbf{r}^\dagger$ is a random vector of path reflectances, $\mathbf{t}^\dagger$ is a random vector of atmospheric transmissivities, and $\mathbf{s}$ is a random vector of spherical albedos. All three are vector-valued functions of $\boldsymbol{\zeta}$. We absorb the constants in (3) into these vector quantities by defining $\mathbf{r}(\boldsymbol{\zeta})_i := \pi^{-1}E_i \cos \theta \mathbf{r}(\boldsymbol{\zeta})_i^\dagger$ and $\mathbf{t}(\boldsymbol{\zeta})_i := \pi^{-1}E_i \cos \theta \mathbf{t}(\boldsymbol{\zeta})_i^\dagger$, after which the forward model reads

$$\mathbf{f}_i(\mathbf{x}) := \mathbf{r}_i(\boldsymbol{\zeta}) + \frac{\mathbf{t}_i(\boldsymbol{\zeta})\boldsymbol{\rho}_i}{1 - \mathbf{s}_i(\boldsymbol{\zeta})\boldsymbol{\rho}_i}. \quad (5)$$

2.1.2. Probability Model

The inversion methods considered in this paper are based on Bayes' theorem, which describes how prior knowledge may be combined with information in the observation Equation (1) to specify the conditional probability of states $\mathbf{x}$, given observations $\mathbf{y}$,

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}. \quad (6)$$

The term on the left side of the equation is the posterior probability distribution of state $\mathbf{x}$, which we also refer to as the joint posterior density of $\boldsymbol{\zeta}$ and $\boldsymbol{\rho}$. The term $p(\mathbf{y}|\mathbf{x})$ is the likelihood, and it codifies the plausibility of making the observations $\mathbf{y}$ given our forward model and state $\mathbf{x}$. The second term in the numerator, $p(\mathbf{x})$, is the marginal distribution of the state variables and is called the prior distribution, because it codifies our prior beliefs, before seeing $\mathbf{y}$, about the state. Finally, $p(\mathbf{y})$ (evidence) is a normalization constant that does not come into play with the algorithms that are considered in this work.

The Gaussian statistics in (2) lead to the unnormalized Gaussian likelihood

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}\|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_{\mathbf{\Gamma}}^2\right), \quad (7)$$

where we use the standard notation $\|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_{\mathbf{\Gamma}}^2 := (\mathbf{f}(\mathbf{x}) - \mathbf{y})^T \mathbf{\Gamma}^{-1} (\mathbf{f}(\mathbf{x}) - \mathbf{y})$. We also prescribe a Gaussian prior for $\boldsymbol{\rho}$,

$$\begin{aligned} p(\mathbf{x}) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &\propto \exp\left(-\frac{1}{2}\|\boldsymbol{\rho} - \boldsymbol{\mu}\|_{\boldsymbol{\Sigma}}^2\right), \end{aligned} \quad (8)$$

with $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ being the prior mean and prior covariance for $\boldsymbol{\rho}$. We would like to remark that, while it is possible that heavier-tailed priors could possibly provide improved retrieval performance in some scenarios, the computation of the posterior mean with AOE depends on the Gaussian formalism, and, for this reason, we have not included them in the present work.

We use a uniform (flat) prior for the atmosphere, for which reason, $\boldsymbol{\zeta}$ does not appear in (8). With a linear $\mathbf{f}(\mathbf{x})$, this formulation would lead to a Gaussian posterior distribution $p(\mathbf{x}|\mathbf{y})$, but here $\mathbf{f}(\mathbf{x})$ is not linear in $\mathbf{x}$ and the posterior is non-Gaussian.

2.1.3. Optimal Estimation

Optimal estimation [31] has been used successfully in recent years to retrieve reflectance from radiance in imaging spectroscopy problems [19,20,27,29]. It is a form of maximum a posteriori estimation that tries to find for the most probable state

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) \quad (9)$$

using iterative optimization algorithms such as Levenberg–Marquardt or Broyden–Fletcher–Goldfarb–Shanno [37] to maximize the posterior probability with respect to $\mathbf{x}$. This requires finding the gradient of the scalar function $p(\mathbf{x}|\mathbf{y})$ with respect to the high-dimensional vector $\mathbf{x}$ at each step. Once $\hat{\mathbf{x}}$ has been found, OE computes a linearization of the model, $\mathbf{F}$, such that $\mathbf{F}_{ij} = \partial \mathbf{f}(\mathbf{x})_i / \partial \mathbf{x}_j$, after which linear theory yields the Gaussian posterior covariance approximation,

$$\text{Cov}(\mathbf{x}|\mathbf{y}) \approx \hat{\mathbf{S}} := \left(\mathbf{F}^T \mathbf{\Gamma}^{-1} \mathbf{F} + \mathbf{\Sigma}^{-1} \right)^{-1}. \quad (10)$$

In the OE literature, the Jacobian matrix $\mathbf{F}$ is sometimes denoted by $\mathbf{J}$ or $\mathbf{K}$.

The maximization in (9) is typically performed by minimizing the loss function

$$\zeta(\mathbf{x}) = -2 \log p(\mathbf{x}|\mathbf{y}), \quad (11)$$

which with the Gaussian noise and prior distributions above becomes

$$\zeta(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_{\mathbf{\Gamma}}^2 + \|\boldsymbol{\rho} - \boldsymbol{\mu}\|_{\mathbf{\Sigma}}^2 + \text{const}. \quad (12)$$

2.2. The Accelerated Optimal Estimation (AOE) Algorithm for Surface Reflectance Retrievals

Equation (9) specifies that we search for the mode of the posterior density as a function of the full state vector $\mathbf{x}$. However, this ignores a key feature of reflectance retrieval: the optimal surface parameters, which constitute the vast majority of state vector elements, are trivial to calculate for a fixed atmosphere. Our proposed algorithm exploits this structure, reorganizing the generic optimization as a nested algorithm that is far faster than the original. We note that with the specific form of $\mathbf{f}(\mathbf{x})$ in (4), we can solve for the maximum a posteriori $\boldsymbol{\rho}$ analytically if we fix the atmospheric state $\boldsymbol{\zeta}$, use a diagonal $\mathbf{\Gamma}$, and prescribe a uniform prior for all of $\mathbf{x}$. When these conditions are not met, given the form of (12), we can still perform the optimization faster than the strategy in (9) by nesting it:

$$\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\rho}} = \arg \max_{\boldsymbol{\zeta}} \left(\arg \max_{\boldsymbol{\rho}} p(\mathbf{x}|\mathbf{y}) \right). \quad (13)$$

The outer optimization is performed in a low-dimensional space, and while the inner optimization is high-dimensional, the conditional mean can be efficiently approximated in closed form.

2.2.1. AOE Inner Loop: Finding the Conditional MAP Estimate of the Surface State

Let $\tilde{\boldsymbol{\zeta}}$ denote an atmospheric state that has been fixed to a specific value, and let the radiative transfer-dependent quantities produced using this atmospheric state be $\tilde{\mathbf{r}} := \mathbf{r}(\tilde{\boldsymbol{\zeta}})$, $\tilde{\mathbf{t}} := \mathbf{t}(\tilde{\boldsymbol{\zeta}})$, and $\tilde{\mathbf{s}} := \mathbf{s}(\tilde{\boldsymbol{\zeta}})$. An approximation to the conditional mean in the inner maximization of (13) is obtained by approximating the first term in (12) with

$$\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_{\mathbf{\Gamma}}^2 \approx \|\mathbf{y} - \tilde{\mathbf{f}}(\boldsymbol{\rho})\|_{\mathbf{\Gamma}}^2 \quad (14)$$

where $\tilde{\mathbf{f}}(\boldsymbol{\rho})$ is a linearization of the forward function with elements

$$\tilde{\mathbf{f}}(\boldsymbol{\rho})_i := \tilde{\mathbf{r}}_i + \frac{\tilde{\mathbf{t}}_i \boldsymbol{\rho}_i}{1 - \tilde{\mathbf{s}}_i \boldsymbol{\rho}_i^{(k)}}. \quad (15)$$

Here, the vector $\boldsymbol{\rho}^{(k)}$ in the denominator is an estimate of the value of $\boldsymbol{\rho}$ at iteration (k) of the algorithm described in Algorithm 1 below. In this paper, we denote the algorithm iterations for all algorithms by upper indexes in parentheses.

Algorithm 1: Iterative algorithm for finding the approximate mean conditional posterior reflectance and corresponding approximate covariance.

```

1 AOE( $\tilde{\boldsymbol{\zeta}}^{(1)}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Gamma}, n_{\text{iter}}^{\text{opt}}, n_{\text{iter}}$ )
   Data: Initial guess  $\tilde{\boldsymbol{\zeta}}^{(1)}$  of the atmospheric state, reflectance prior mean  $\boldsymbol{\mu}$  and
   covariance  $\boldsymbol{\Sigma}$ , noise covariance  $\boldsymbol{\Gamma}$ , number of inner loop iterations during
   outer loop optimization  $n_{\text{iter}}^{\text{opt}}$ , number of final inner loop iterations  $n_{\text{iter}}$ ,
   and outer loop algorithm OuterLoopAlgorithm with convergence criterion
   converged and maximum number of iterations  $N$ .
   Result: Approximation of final state  $\hat{\mathbf{x}} := [\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\rho}}]^T$ , and conditional covariance
    $\mathbf{C}_{\text{AOE}}^{\boldsymbol{\mu}} = \text{Cov}(\boldsymbol{\rho} | \hat{\boldsymbol{\zeta}}, \mathbf{y})$ .

2 Initialization: Set  $\tilde{\boldsymbol{\zeta}}^{(1)}$  to initial guess
3 for  $j \leftarrow 1$  to  $N$  do
4    $\tilde{\mathbf{r}} \leftarrow \mathbf{r}(\tilde{\boldsymbol{\zeta}}^{(j)}), \tilde{\mathbf{t}} \leftarrow \mathbf{t}(\tilde{\boldsymbol{\zeta}}^{(j)}), \tilde{\mathbf{s}} \leftarrow \mathbf{s}(\tilde{\boldsymbol{\zeta}}^{(j)})$ 
5    $\tilde{\boldsymbol{\mu}}, \tilde{\mathbf{S}} \leftarrow \text{InnerLoop}(\boldsymbol{\mu}, \tilde{\mathbf{r}}, \tilde{\mathbf{t}}, \tilde{\mathbf{s}}, n_{\text{iter}}^{\text{opt}})$ 
6   Compute  $\tilde{\boldsymbol{\zeta}}([\tilde{\boldsymbol{\zeta}}^{(j)}, \tilde{\boldsymbol{\mu}}]^T)$  in (11)
7   if converged (based on previous evaluations of  $\tilde{\boldsymbol{\zeta}}$ ) then
8     break
9   else
10    Get  $\tilde{\boldsymbol{\zeta}}^{(j+1)}$  using OuterLoopAlgorithm update strategy
11  end
12 end
13  $\hat{\boldsymbol{\zeta}} \leftarrow \tilde{\boldsymbol{\zeta}}^{(j)}$ 
14  $\tilde{\mathbf{r}} \leftarrow \mathbf{r}(\hat{\boldsymbol{\zeta}}), \tilde{\mathbf{t}} \leftarrow \mathbf{t}(\hat{\boldsymbol{\zeta}})$ , and  $\tilde{\mathbf{s}} \leftarrow \mathbf{s}(\hat{\boldsymbol{\zeta}})$ 
15  $\hat{\boldsymbol{\rho}}, \mathbf{C}_{\text{AOE}}^{\boldsymbol{\mu}} \leftarrow \text{InnerLoop}(\boldsymbol{\mu}, \tilde{\mathbf{r}}, \tilde{\mathbf{t}}, \tilde{\mathbf{s}}, n_{\text{iter}})$ 
16 return  $[\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\rho}}]^T, \mathbf{C}_{\text{AOE}}^{\boldsymbol{\mu}}$ 

1 InnerLoop( $\boldsymbol{\rho}^{(1)}, \tilde{\mathbf{r}}, \tilde{\mathbf{t}}, \tilde{\mathbf{s}}, n_{\text{iter}}^{\text{inner}}$ )
2 for  $k \leftarrow 1$  to  $n_{\text{iter}}^{\text{inner}}$  do
3   Compute  $\mathbf{L}^{(k)}$  with  $\mathbf{L}_{ii}^{(k)} \leftarrow \frac{\tilde{\mathbf{t}}_i}{1 - \tilde{\mathbf{s}}_i \boldsymbol{\rho}_i^{(k)}}$ 
4   Compute  $\tilde{\mathbf{S}}^{(k)}$  and  $\tilde{\boldsymbol{\mu}}^{(k)}$  in (25) and (27)
5    $\boldsymbol{\rho}^{(k+1)} \leftarrow \tilde{\boldsymbol{\mu}}^{(k)}$ 
6 end
7 return  $\tilde{\boldsymbol{\mu}}^{(n_{\text{iter}}^{\text{inner}})}, \tilde{\mathbf{S}}^{(n_{\text{iter}}^{\text{inner}})}$ 

```

To approximate the conditional posterior density $p(\boldsymbol{\rho} | \hat{\boldsymbol{\zeta}}, \mathbf{y})$, let $\mathbf{L}^{(k)} \in \mathbb{R}^{n \times n}$ be the diagonal matrix with elements $\mathbf{L}_{ii}^{(k)} = \tilde{\mathbf{t}}_i / (1 - \tilde{\mathbf{s}}_i \boldsymbol{\rho}_i^{(k)})$, so that we can write

$$\tilde{\mathbf{f}}(\boldsymbol{\rho}) = \tilde{\mathbf{r}} + \mathbf{L}^{(k)} \boldsymbol{\rho}. \quad (16)$$

The diagonal elements of $\mathbf{L}^{(k)}$ are always positive, since both the spherical albedo $\tilde{s}$ and the surface reflectances $\rho^{(k)}$ are between zero and one. We also define

$$\tilde{\mathbf{\Gamma}}^{(k)} := \mathbf{L}^{(k)-1} \mathbf{\Gamma} \mathbf{L}^{(k)-1}. \quad (17)$$

Then, by rearranging terms in (15), the logarithm of the approximate likelihood turns out to be proportional to a quadratic form:

$$-2 \log(p(\mathbf{y}|\boldsymbol{\rho}, \tilde{\boldsymbol{\zeta}})) + \text{const.} \quad (18)$$

$$\approx \|\mathbf{y} - \tilde{\mathbf{f}}(\boldsymbol{\rho})\|_{\mathbf{\Gamma}}^2 \quad (19)$$

$$= \left\| \left(\mathbf{L}^{(k)} \right)^{-1} (\mathbf{y} - \tilde{\mathbf{f}}(\boldsymbol{\rho})) \right\|_{\tilde{\mathbf{\Gamma}}^{(k)}}^2 \quad (20)$$

$$= \left\| \left(\mathbf{L}^{(k)} \right)^{-1} \mathbf{y} - \left(\mathbf{L}^{(k)} \right)^{-1} \tilde{\mathbf{f}}(\boldsymbol{\rho}) \right\|_{\tilde{\mathbf{\Gamma}}^{(k)}}^2 \quad (21)$$

$$= \left\| \left(\mathbf{L}^{(k)} \right)^{-1} \mathbf{y} - \left(\mathbf{L}^{(k)} \right)^{-1} (\tilde{\mathbf{r}} + \mathbf{L}^{(k)} \boldsymbol{\rho}) \right\|_{\tilde{\mathbf{\Gamma}}^{(k)}}^2 \quad (22)$$

$$= \left\| \left(\mathbf{L}^{(k)} \right)^{-1} (\mathbf{y} - \tilde{\mathbf{r}}) - \boldsymbol{\rho} \right\|_{\tilde{\mathbf{\Gamma}}^{(k)}}^2. \quad (23)$$

Here, (19) is the approximate unnormalized log likelihood, in which we multiply the $\mathbf{\Gamma}^{-1}$ by identity $\left(\mathbf{L}^{(k)} \right)^{-1} \mathbf{L}^{(k)}$ on both sides in (20). Equation (21) rearranges the terms, after which we substitute the definition in (16) for $\tilde{\mathbf{f}}(\boldsymbol{\rho})$. The simplified result in (23) then yields the approximate unnormalized likelihood function,

$$\begin{aligned} p(\mathbf{y}|\boldsymbol{\rho}, \tilde{\boldsymbol{\zeta}}) \\ \propto \exp \left(-\frac{1}{2} \left\| \left(\mathbf{L}^{(k)} \right)^{-1} (\mathbf{y} - \tilde{\mathbf{r}}) - \boldsymbol{\rho} \right\|_{\tilde{\mathbf{\Gamma}}^{(k)}}^2 \right). \end{aligned} \quad (24)$$

In performing the retrieval, we use the specified Gaussian prior for $\boldsymbol{\rho}$ and a uniform prior for $\tilde{\boldsymbol{\zeta}}$ (any Gaussian prior for $\mathbf{x}$ would work, but we choose to use an uninformative prior for $\tilde{\boldsymbol{\zeta}}$). The priors for $\tilde{\boldsymbol{\zeta}}$ are furthermore truncated to only include realistic values; their limits are given in Section 3. In the inner loop optimization, where $\tilde{\boldsymbol{\zeta}}$ is kept constant, this prior specification yields the prior term $p(\mathbf{x}) \propto p(\boldsymbol{\rho}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. By using standard linear theory for solving generalized Tikhonov-regularized least squares problems, the conditional posterior covariance in the inner loop optimization over the vector $\boldsymbol{\rho}$ becomes

$$\tilde{\mathbf{S}}^{(k)} := \left(\tilde{\mathbf{\Gamma}}^{(k)-1} + \boldsymbol{\Sigma}^{-1} \right)^{-1}. \quad (25)$$

The full approximate posterior conditioned on $\tilde{\boldsymbol{\zeta}}$ is then given by

$$p(\boldsymbol{\rho}|\tilde{\boldsymbol{\zeta}}, \mathbf{y}) \approx \mathcal{N}(\tilde{\boldsymbol{\mu}}^{(k)}, \tilde{\mathbf{S}}^{(k)}), \quad (26)$$

where the mean is given by

$$\tilde{\boldsymbol{\mu}}^{(k)} := \tilde{\mathbf{S}}^{(k)} \left(\tilde{\mathbf{\Gamma}}^{(k)-1} \mathbf{L}^{(k)-1} (\mathbf{y} - \tilde{\mathbf{r}}) + \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right). \quad (27)$$

Note that the mean $\tilde{\boldsymbol{\mu}}^{(k)}$ and covariance $\tilde{\mathbf{S}}^{(k)}$ depend on a specific value of $\tilde{\boldsymbol{\zeta}}$ via quantities $\tilde{\mathbf{r}}$, $\tilde{\mathbf{t}}$, and $\tilde{s}$, which are embedded in terms in (25)–(27).

While the equations above described the Gaussian approximation for a single outer loop iteration (k) only, the Gaussian form in (26) allows finding the conditional posterior mean of the surface reflectances given the atmosphere with the fast iterative algorithm `InnerLoop()`, shown in the bottom part of Algorithm 1. A single iteration is often enough to obtain a good estimate of the mean.

2.2.2. AOE Outer Loop—Finding the MAP Atmospheric State

The outer loop optimization, ref. (13), is shown in function `AOE()` in the top part of Algorithm 1. When executing this algorithm, the best joint state is found by maximizing the probability over the atmospheric state ζ . The objective function evaluated at line 6 of `AOE()` at each outer loop iteration (j) is the $\zeta(\mathbf{x})$ in (12) with $\mathbf{x} = [\zeta^{(j)}, \tilde{\mu}]^T$, where $\tilde{\mu}$ is given by the function `InnerLoop()`, also in Algorithm 1. We note that the outer loop optimization is carried out without explicit Jacobian computations, and even though adding them would be possible, the convergence using the current setup is already very fast without them, see Sections 2.5 and 4.5.

2.2.3. Final AOE Step Yields an Accurate Gaussian Approximation of Posterior Distribution $p(\mathbf{x}|\mathbf{y})$

After the final atmospheric state $\tilde{\zeta}$ is obtained, the function `InnerLoop()` in Algorithm 1 is run one more time with a larger value of iterations to improve convergence. This produces the final estimate of the posterior mean of the surface state, $\hat{\rho}$. Depending on the application, the reported uncertainty may then be either the final conditional covariance $\mathbf{C}_{\text{AOE}}^{\mu}$, returned on line 16 of `AOE()`, in Algorithm 1, or an OE estimate can be computed with (10), possibly also taking into account the effects of ζ on the covariance. The specific normal approximation produced by the AOE algorithm in Algorithm 1 is given in the later sections, denoted by $\pi_{\mathbf{C}_{\text{AOE}}}^{\mu}$. In Section 4, we compare several different methods to estimate this covariance.

2.2.4. Tuning the Computational Cost of AOE

The AOE algorithm in Algorithm 1 consists of alternating between repeatedly finding the conditional mean of ρ , and moving in the low-dimensional atmospheric part of the state space to minimize an objective function. Varying optimization parameters provides opportunities to increase the model precision, often at a higher cost. Obvious choices include the choice of `OuterLoopAlgorithm` in Algorithm 1, the number of outer loop optimization iterations, the constant n_{iter} and convergence criteria used in the AOE algorithm, the structure of prior and noise covariances, and the spectral ranges included in the retrieval. Since the cost of computing the conditional mean in (27) is $\mathcal{O}(n^3)$, we can further make the AOE algorithm faster by using only a subset of the channels for the inversion. However, to obtain the correct atmospheric parameter values, the effects of aggressive spectral thinning need to be analyzed carefully, to ensure proper convergence. The AOE algorithm parameters for the experiments presented in this paper are shown in Table 1 and discussed in Section 2.5. The analysis of the results with respect to these trade-offs are discussed in Section 5.2.

Table 1. Retrieval algorithms studied in this paper. The number of iterations N was tuned for each algorithm to a value that produced the most compelling results in terms of accuracy and speed.

Name	N	Alg.	Description
AOE ₀	4	SLSQP	Slowest and most accurate configuration of AOE. Final $n_{\text{iter}} = 4$.
AOE ₁	3	SLSQP	Faster configuration of AOE, compromise between speed and accuracy. Final $n_{\text{iter}} = 2$. Uses half of the reflectance channels in the inner loop.
AOE ₂	0	N/A	Fastest configuration of AOE, best reflectances are found by inner loop optimization only. Final $n_{\text{iter}} = 2$.
ROE	20	TRF	Optimal estimation, with standard config from ISOFIT.
MCMC	5M	AM	MCMC for $\rho \tilde{\zeta}, \mathbf{y}$, conditioning on the MAP estimate of ζ

2.3. Markov Chain Monte Carlo for Surface Reflectance Retrievals

MCMC algorithms are used to draw from posterior distributions in situations where procedures for directly generating samples do not exist but evaluating the prior probability density and the likelihood function up to a multiplicative constant is still possible. The chain of state vectors generated by an MCMC algorithm describes the posterior distribution in a way analogous to how samples from a Gaussian distribution describe that Gaussian distribution. A precise description of non-Gaussian posterior distributions of ζ and ρ may therefore be generated using MCMC algorithms.

Most commonly used MCMC algorithms based on the Metropolis algorithm [38] generate the chain of samples by repeatedly iterating a sequence of three steps. Assuming that at a given iteration (i) of the MCMC chain is at location $\mathbf{x}^{(i)}$ in the parameter space, the first step in generating the next sample is proposing a new point $\mathbf{x}' \leftarrow \mathbf{x}^{(i)} + \delta^{(i)}$, where $\delta^{(i)} \sim \mathbf{q}(\boldsymbol{\theta})$. The distribution $\mathbf{q}(\boldsymbol{\theta})$ is called the proposal distribution and is chosen by the user. Gaussian proposal distributions are a common choice for $\mathbf{q}$, in which case, $\boldsymbol{\theta}$ determines the mean and the covariance of the proposal. After proposing the state vector $\mathbf{x}'$, the scaled posterior density, $\phi(\mathbf{x}') := p(\mathbf{y}|\mathbf{x}')p(\mathbf{x}')$, is evaluated, and then, based on an acceptance criterion, the vector $\mathbf{x}'$ is either added to the chain ($\mathbf{x}^{(i+1)} \leftarrow \mathbf{x}'$) or the vector $\mathbf{x}^{(i)}$ is repeated in the chain ($\mathbf{x}^{(i+1)} \leftarrow \mathbf{x}^{(i)}$). This sequence is repeated, possibly millions of times—the higher the dimension of $\mathbf{x}$, the more samples are needed.

In practice, the scaled posterior density is computed via evaluating the loss function, (12), which is the negative logarithm of the posterior with a constant offset. We point out that the radiative transfer outputs required for the log-likelihood evaluation are obtained from a look-up table, as described below, as sequentially running high-fidelity radiative transfer codes sufficiently many times is not practical.

In Metropolis-type algorithms with a Gaussian proposal distribution, the probability of accepting $\mathbf{x}'$ and adding it to the chain as the $(i + 1)^{\text{th}}$ member, as opposed to repeating $\mathbf{x}^{(i)}$, is the ratio of the posterior densities at the proposed and the previous parameter values, i.e., $\mathbb{P}(\text{accept } \mathbf{x}') = \min\left(1, \phi(\mathbf{x}')/\phi(\mathbf{x}^{(i)})\right)$. Hence, proposed points $\mathbf{x}'$ with a higher posterior probability than the previous point $\mathbf{x}$ are always accepted, whereas proposed points with a lower posterior probability than the previous state in the chain are sometimes accepted.

The number of samples needed in an MCMC experiment depends on the quality of the proposal distribution $\mathbf{q}$. To avoid the need to hand-tune the parameters $\boldsymbol{\theta}$, we used the adaptive Metropolis (AM) MCMC algorithm [39] for sampling, with a zero-mean Gaussian $\mathbf{q}$. AM finds optimal values of $\boldsymbol{\theta}$ as the sampling proceeds, in order to minimize the autocorrelatedness of the MCMC chain. Since AM uses Gaussian proposals in its standard configuration, the less Gaussian the target distribution, the less optimal the AM adaptation scheme may become. To ensure that we used an MCMC algorithm that was appropriate for the problem, we also performed MCMC retrievals with several alternative proposal schemes. These included using transformed proposals in radiance space, non-adaptive Metropolis sampling, component-wise Metropolis sampling, and approximate Metropolis-within-Gibbs sampling. All these methods produced results that were consistent with one another. The AM algorithm was both stable and fast, and it easily produced a more than sufficient number of uncorrelated draws from the posterior distributions. The AM algorithm is described in Algorithm 2.

Compared to earlier work in [20], we performed MCMC slightly differently, to ensure convergence and proper statistics. Most notably, we used an adaptive algorithm (AM instead of the standard Metropolis) to decrease the correlatedness of the MCMC samples, we did not restart the chains at regular intervals to follow standard practice, and the chains that we generated were substantially longer (five million vs. 20 thousand iterations).

While we performed extensive MCMC simulations over both surface and atmospheric states, the results from these joint MCMC inversions are complex and require lengthy analyses. For these reasons, we restrict the results presented in this paper to the case where the atmospheric state was fixed.

Algorithm 2: Adaptive Metropolis MCMC algorithm as used in this work. The function `UpdateProposal` updates the θ parameters of the proposal distribution q whenever the criterion `update_criterion` is satisfied. The scalar α is the acceptance probability and `Uniform()` generates uniformly random numbers from interval $[0, 1]$. The output is a set of correlated samples from the true posterior distribution.

Data: Initial state $\mathbf{x}^{(0)}$, initial proposal distribution $q(\theta)$ from which increments $\delta^{(i)}$ are drawn, likelihood function $p(\mathbf{y}|\mathbf{x})$, prior $p(\mathbf{x})$, number of MCMC iterations n_{iter} . The parameters θ determine the exact form of the proposal distribution. Write $\phi(\mathbf{x}) := p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$.

Result: Correlated samples from the posterior distribution $p(\mathbf{x}|\mathbf{y})$.

```

1 for  $i \leftarrow 1$  to  $n_{\text{iter}}$  do
2   if update_criterion = True then
3      $\theta \leftarrow \text{UpdateProposal}()$ ;
4   end
5   Draw  $\delta^{(i)}$  from  $q(\theta)$ 
6    $\mathbf{x}' \leftarrow \mathbf{x}^{(i-1)} + \delta^{(i)}$ 
7    $u \leftarrow \text{Uniform}()$ 
8    $\alpha \leftarrow \max\left(1, \frac{\phi(\mathbf{x}')}{\phi(\mathbf{x}^{(i-1)})}\right)$ 
9   if  $u < \alpha$  then
10     $\mathbf{x}^{(i)} \leftarrow \mathbf{x}'$ 
11  else
12     $\mathbf{x}^{(i)} \leftarrow \mathbf{x}^{(i-1)}$ 
13  end
14 end
```

2.4. Radiative Transfer

Inverting the surface reflectance from a measured radiance requires modeling how solar radiation travels in the atmosphere, interacts with the surface, scatters back, and finally ends up at the instrument sensor. These processes can be modeled using specialized and complex radiative transfer models. When solving the reflectance retrieval inverse problem, radiative transfer enters the model in the terms $\mathbf{r}(\zeta)$ (path reflectance), $\mathbf{t}(\zeta)$ (total transmittance including both upward and downward, diffuse and direct components), and $\mathbf{s}(\zeta)$ (spherical albedo) in the forward function, (4). In this work, we use the computationally demanding MODTRAN 6.0 [18] atmospheric radiative transfer model (RTM), which provides the radiative-transfer-dependent quantities required by the forward model in (4). The parameters in ζ are, however, given as inputs to MODTRAN, and learning those parameters can be framed as tuning the radiative transfer to specific retrieval conditions. We treat the radiative transfer model as a black box, meaning that we do not attempt to analyze how it works or was implemented here.

Since radiative transfer computations are expensive, running MODTRAN, e.g., at every MCMC iteration, is prohibitive. As is a common practice in atmospheric correction procedures, all versions of the ISOFIT software solves this problem by generating a look-up table (LUT) for the atmospheric parameters, from which $\mathbf{r}(\zeta)$, $\mathbf{t}(\zeta)$, and $\mathbf{s}(\zeta)$ are interpolated

based on the exact values in ζ [40]. This LUT acts like a fast radiative transfer emulator. We utilize this same strategy for AOE and MCMC retrievals, and use the same interpolation grids for all retrieval types for a given retrieval target. The LUT grids are described in Section 3.

The speed of LUT interpolation also plays a role in retrieval performance, and for this reason, the AOE and MCMC codes use a custom fast linear interpolator that replicates the `RegularGridInterpolator` interface in SciPy. The LUT interpolation in AOE and MCMC retrievals was cross-validated against the ROE implementation, to ensure that the results matched, and any differences in log posterior probabilities resulting from model-interpolation differences were smaller than the precision at which the results are presented in Section 4. The linear interpolation code implementations were tested to agree with the machine precision.

2.5. Retrieval Configurations

This paper compared five different retrieval methods, three of which were AOE retrievals. Of the rest, one was a conventional OE retrieval (ROE, within ISOFIT), and one was a benchmark retrieval using MCMC. The retrieval configurations are summarized in Table 1, and all retrieval types were performed for five different sites, described below in Section 3. The first column shows the name of the retrieval algorithm, the second column shows the number of optimization iterations N (outer loop for AOE), and the third column shows the optimization algorithm.

The three AOE retrievals reflected different compromises between speed and accuracy, the latter of which is defined here as the posterior probability of the retrieved state. The different AOE configurations were chosen to demonstrate the range of possibilities available within the AOE framework. AOE₀ utilizes all channels in the non-masked spectral range (see Section 3), and performs up to four outer loop iterations with the sequential least squares programming (SLSQP) algorithm [41]. This was enough to achieve convergence for all test targets. AOE₁ ran a maximum of three iterations of SLSQP, and used only half of the wavelengths, rather than the full set. This makes matrix inversion much faster, speeding up the retrieval. The last iteration of the inner loop in AOE₁ used all the channels, so that the final result was an estimate of the full state and its covariance matrix. In the last retrieval, AOE₂, no outer loop optimization was performed; the initial value of ζ was not changed and only the inner loop, shown in Algorithm 1, was executed. In Section 4, we show that even this simplified retrieval method was able to converge to good reflectance estimates.

When optimizing over the atmosphere, n_{iter} was set to 1 in both AOE₀ and AOE₁. After finding the $\hat{\zeta}$, in order to find the final $\hat{\rho}$, n_{iter} was set to 4 for AOE₀ and 2 for AOE₁ and AOE₂. While $n_{\text{iter}} = 1$, which was used when finding the best ζ , did not produce fully converged results, the ρ obtained from using $n_{\text{iter}} = 2$ were already very close to the conditional mean. The value of $n_{\text{iter}} = 4$ in AOE₀ was selected to show how well the AOE algorithm could converge in the best case—in our tests, the inner loop virtually always fully converged in four iterations. We used a standard implementation of SLSQP available in the SciPy v1.6 Python package.

We also performed OE retrievals of all the retrieval targets using ISOFIT software, version 2.5.0, here termed ROE. ROE was tested and shown to perform well in several recent papers [19,20,27,29]. The role of ROE was to demonstrate both the computational performance and MAP estimate convergence characteristics of a standard implementation. The ROE configuration is shown on line 4 of Table 1. The minimization algorithm used by ROE was the trust region reflective algorithm [42], as available in the SciPy library. The maximum number of iterations allowed in the ROE configuration was 20.

The MCMC experiments described in Table 1 used the adaptive Metropolis algorithm described in Algorithm 2. In MCMC, the atmospheric state was fixed to the best MAP estimate that was found, which happened to be the AOE₀ MAP estimate. MCMC sampling was performed over the surface reflectances only, conditioned on the best possible atmospheric state. All the MCMC chains were 5,000,000 iterations long, and in all MCMC experiments the first half of the chain was discarded as “burn-in” when computing statistics for the purpose of reporting results. This was done because, before the MCMC chains have properly mixed, the samples cannot be viewed as samples from the underlying target distribution. For convenience, we used a subsample of size 100,000 to compute all results. We thinned the chain; i.e., the iterations picked were chosen at regular intervals, to minimize autocorrelation of the samples.

In addition to the quality of the AOE and OE point estimates, we were interested in understanding how different the covariance estimate produced by the AOE inner loop was from the standard OE covariance estimate, as well as how much including atmospheric components changed the covariance estimates for the set of retrievals studied here. Three types of linearization-based covariance estimates were compared, each designated by a subscript. First, $\mathbf{C}_{\text{AOE}}$ was the estimate of the conditional covariance matrix produced by the AOE inner loop algorithm (Algorithm 1). Second, $\mathbf{C}_{\text{OE}_C}$ was the covariance estimate given by (10), where $\mathbf{F}$ is the $n \times n$ matrix containing partial derivatives of the forward model $\mathbf{f}(\mathbf{x})$ with respect to $\boldsymbol{\rho}$. The third approximation, $\mathbf{C}_{\text{OE}_F}$ was the same as the second, but using the full $\mathbf{F}$, which also contained the columns of partial derivatives with respect to the elements of $\boldsymbol{\zeta}$.

In the following sections, we use the letter π to designate estimates of posterior distributions $p(\mathbf{x}|\mathbf{y})$, $p(\boldsymbol{\rho}|\boldsymbol{\zeta}, \mathbf{y})$, and $p(\boldsymbol{\rho}|\mathbf{y})$, and superscripts and subscripts to identify the type of the particular approximation used. For instance, $\pi_{\mathbf{C}_{\text{OE}_F}}^{\mu_{\text{OE}}}$ refers to the Gaussian posterior approximation with mean set to the ROE mean estimate and covariance approximation computed as in OE using the full state. The possible means are μ_{OE} and μ_{AOE_i} with $i \in \{0, 1, 2\}$, and the possible covariances are the ones listed in the previous paragraph: $\mathbf{C}_{\text{OE}_C}$, $\mathbf{C}_{\text{OE}_F}$, and $\mathbf{C}_{\text{AOE}}$. The MCMC posteriors, describing MCMC estimates of $p(\boldsymbol{\rho}|\boldsymbol{\zeta}, \mathbf{y})$, are denoted by π_{MCMC} . We generally treated the atmospheric part $\boldsymbol{\zeta}$ as a nuisance variable, even though we also discuss AOE and ROE convergence in the atmospheric part of the state.

In Section 4.4, we compare the performance of AOE₀ to ROE for a large scene. The retrieval settings there are the same as described here and in Table 1 for AOE₀ and ROE.

3. Data

We performed the retrievals described in Section 2.5 for five targets acquired by NASA’s Next Generation Airborne Visible Infrared Imaging Spectrometer (AVIRIS-NG) and previously reported in [19,36]. They were acquired in overflights of the Jet Propulsion Laboratory and Caltech campus in Pasadena, USA, in 2014 and 2017, respectively. The relevant AVIRIS-NG flightline codes are ang20140612t215931 (177, 306, PL, and Mars) and ang20171108t184227 (BL), and these codes contain the approximate measurement times, along with the exact measurement geometries. The solar zenith angles for the retrievals were 29° for the JPL acquisitions and 52° for the Caltech/Beckman Lawn acquisition. The respective flight altitudes were 1980 m and 2350 m. The Jet Propulsion Laboratory is located at 34°12’N, 118°10’W, and the Caltech campus at 34°8’N, 118°7’W. These sites were selected because they offered a wide range of different surface types and brightness levels.

Targets 177, 306, and PL were artificial; Mars was a bare dirt surface, and BL was a vegetation target. The PL target was very dark, 306 very bright, and all the rest had brightnesses between these two extremes. Within a few days of the overflight, a field team visited each of the five locations and acquired in situ observations of upwelling radiance

under a similar solar geometry using Malvern Panalytical ASD field spectrometers. These radiance measurements were compared to those of an upward-facing reference panel, enabling a direct measurement of surface reflectance. For further information on the measurement protocol, we refer the reader to [19].

AVIRIS-NG measures the visible and shortwave infrared (VSWIR) spectrum in the range 380 nm–2500 nm with 5 nm spectral resolution [43]. We applied a channel mask to exclude the deep water absorption features, because, in practice, all radiation in these channels is absorbed by the atmosphere. After applying the mask, the channel ranges that were included for carrying out the inversions were 400 nm–1300 nm, 1450 nm–1780 nm, and 2050 nm–2450 nm for the 177, 306, PL, and Mars retrievals, and 380 nm–1300 nm, 1450 nm–1780 nm, and 1950 nm–2450 nm for the BL retrieval. This left 326 channels for the JPL retrievals, and 349 channels for the BL retrieval. Thus, the lengths of vectors ρ , y , r , t , and s were either 326 or 349, depending on the target. In the figures in Section 4, we present the data by interpolating over the masked-out sections and mention this explicitly in the captions.

Finding the maximum a posteriori estimate using (12) requires prescribing the noise covariance matrix Γ , and the prior mean μ and prior covariance matrix, Σ for the state. The instrument noise depends on photon counting statistics, which in turn depend on the brightness of the surface. Consequently, we calculated the noise covariance Γ for each observation separately. We aggregated signal-dependent and signal-independent noise sources from a first-principles understanding of the physical instrument performance. In order to speed up this process at runtime, we fit the signal/noise relationship using a concise three-parameter model, which has been shown to very accurately describe instrument characteristics [19]. This model produced a diagonal covariance structure, implying the independence of the measurement errors across wavelengths. All targets were measured several times during the overflights, and the repeated measurements were averaged to produce averaged observations with lower measurement uncertainty. The number of observations of each target is reported in the third column of Table 2.

Table 2. Retrieval targets.

Target	Abbr.	Sol. Z. $\angle$	#Obs.	Short Description
JPL/Building 177	177	29°	400	Parking lot
JPL/Building 306	306	29°	400	Office building roof, very bright
JPL/Parking Lot	PL	29°	400	Very dark target (asphalt)
JPL/Mars Yard	Mars	29°	400	Testing ground for Mars rovers, gravel
Caltech/Beckman Lawn	BL	52°	294	A grassy lawn at Caltech campus

We used a uniform prior for the atmospheric part of the state vector, ζ , meaning that there was no $p(\zeta)$ component in the formula for the posterior density (6). However, to deal with the underdetermined nature of the inverse problem, we prescribed a Gaussian prior for ρ . Estimating priors from data is a challenging task that falls outside the scope of this work. This restriction of the scope was justifiable, since prior selection does not affect the computational performance of either the ROE or the AOE algorithms. In order to obtain results that were comparable between AOE and ROE, we used the same priors for all methods for each retrieval.

The prior mean and covariance for the surface reflectances were determined in ISOFIT by trying different prior models and choosing the one that best described the signal based on a preliminary analysis. The available surface prior models were constructed from spectroscopic measurements of various materials, which were clustered to build a library of spectroscopically similar measurements, see [19]. This prior selection model was not included in the AOE code, and to make the timing results comparable across retrieval

methods, the prior selection step was also omitted for the ROE timings below. Of the prior covariance matrices, only the one used for BL was dense, whereas the priors for the JPL retrievals were mostly diagonal. These mostly diagonal priors did, however, contain dense sub-matrices for the regions 890 nm–990 nm and 1090–1190 nm, which contained water absorption features.

For radiative transfer, a pregenerated LUT was interpolated to compute the terms in the forward function that depended on the atmospheric state. The ζ vector contains two components: AOT at 550 nm and water vapor concentration, labeled as H₂O. For the JPL targets, the allowable values for AOT vary between 0.001 and 0.5, and this range was divided into three equal segments. The minimum and maximum values of the H₂O parameter were 1.310 and 1.587, and the grid was divided into four equal segments. For the Beckman Lawn target, the LUT contains radiative transfer-dependent parameters only for the lower and upper limits, between which the values were linearly interpolated. The AOT for the BL retrieval varied between 0.01 and 0.1 [–], and H₂O between 1.5 and 2 [g cm^{−2}].

In order to demonstrate how AOE works for larger scale retrievals, we also compared the performance of OE and AOE₀ for the 160,000 pixel area outlined by the red rectangle in Figure 1. The sides of the rectangle are 800 m long, which translates to a ground sampling distance of 2 m. The observations are part of the AVIRIS-NG flightline ang20171108t184829. The scene includes the Jet Propulsion Laboratory (JPL), along with its surroundings. On the south-eastern side of the JPL, the Hahamongna Watershed Park contains a small stream, trees, bushes, and grass/hay. On the northern side of the JPL, there are steep hillsides, and the south-western corner of the target area contains managed forest with some taller trees. Some of the rooftops at the JPL are very bright, and hence, in general, difficult for retrieval algorithms to handle. The buildings also cast deep shadows, and much of the space between buildings is paved with asphalt or concrete. The right-hand side of Figure 1 provides a zoomed-out view of the nearby areas for context. These areas are included in the same AVIRIS-NG flightline data, but were not retrieved in this work.



Figure 1. On the right: image of the AVIRIS-NG flightline ang20171108t184829, constructed by taking the red, green, and blue channels from the hyperspectral image. The site over which the larger-scale retrieval was performed is shown in the left part of the figure, outlined by the red rectangle. The zoomed-in part is marked by the blue rectangle in the zoomed-out image of the flightline. The site chosen for the retrieval includes the Jet Propulsion Laboratory in Pasadena, California, and consists of 160,000 pixels. The scene is heterogeneous and includes both built-up and more natural sections.

To perform these larger scale retrievals, the AOT LUT was divided into two equal segments in the range of 0.001–0.5, and the H₂O LUT into 14 equal segments in the range

of $1.04\text{--}4.505\text{ g cm}^{-1}$. The ranges of the included channels were the same as those used for the JPL test targets. However, due to a small difference in the channels' wavelengths, the regions included here contained 325 channels instead of the 326 mentioned above.

4. Results

This section presents the results from our numerical experiments and is divided into several subsections. In Section 4.1, we compare the MAP estimates of reflectances from AOE to those produced by OE using ISOFIT (ROE). Section 4.2 assesses how the point estimates of the atmospheric part of the state from AOE differed from the ones produced by ROE. In Section 4.3, we examine how well the Gaussian approximations provided by OE and AOE described the potentially non-Gaussian “true” posterior distributions from MCMC. In particular, in Section 4.3.1, we introduce and apply two summary statistics that provide overall measures of distributional agreement. Then, in Section 4.3.2, we compare the marginal distributions of reflectances from the different posterior distributions, and in Section 4.3.3, we examine correlation structures in the posterior distributions. Section 4.4 compares ROE to AOE for a scene with a large number of pixels, and finally, in Section 4.5, we look at the computational performance.

4.1. Retrieved Surface Reflectance Point Estimates

Figure 2 shows the retrieved state estimates generated by the AOE₀, AOE₂, with ROE algorithms in Table 1 for each of the five scenes in Table 2. The left panels show the reflectance point estimates and the prior mean spectra used in the retrievals, along with in situ reflectance measurements. The panels on the right show the associated radiance vectors obtained by plugging the point estimates back into the forward model, see (4). The right panels also show in situ radiance observations for comparison. For Building 177, Mars Yard, and Beckman Lawn, all the retrieved reflectance estimates were very similar. For Building 306, the AOE reflectance estimates in the visible spectrum were slightly lower than the ROE estimates, and for the Parking Lot, AOE₂ agreed with ROE but AOE₀ was slightly lower throughout the spectrum. The radiance estimates were very close to the measurements in all cases. While the in situ measurements should not be taken as representing the ground truth, for the PL target, the difference from in situ measurements to all retrieved estimates was more substantial than for the other targets. This suggests that the forward model formulation used may not be well-suited for capturing the physics of dark targets.

The difference plots in Figure 3 reveal additional interesting features of the comparisons. The left side of Figure 3 shows the differences between the retrieved reflectance estimates and the prior mean for AOE₀ and ROE. The graphic in the top-right panel shows the differences between the radiance spectra obtained by plugging in the retrieved reflectances to the forward model and remote radiance observations, also only for AOE₀ and ROE. The bottom-right graphic is mostly the same as the top-right, but it zooms in even closer on the neighborhood of zero on the y -axis and also does not include the Beckman Lawn target results. For the reflectances, the AOE₀ estimates (solid lines) in the visible wavelengths are generally closer to the prior mean than the ROE estimates (dashed lines). In the infrared (IR) region, the estimates from AOE and ROE appear nearly identical.

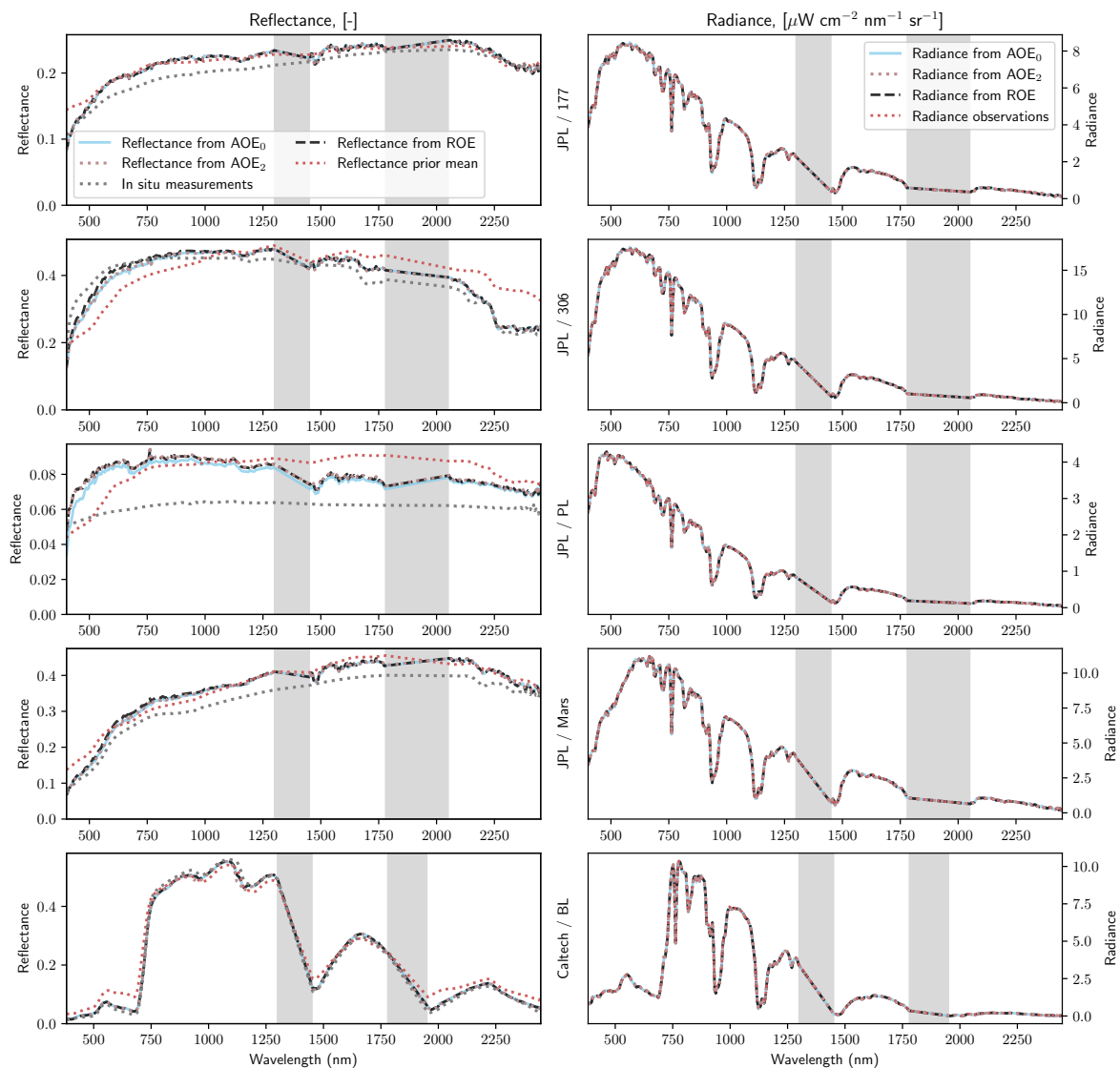


Figure 2. Retrieved reflectances (left side) and corresponding forward-modeled radiances (right side) from the slowest and fastest AOE retrievals and the ROE retrievals. For the reflectance side, prior mean spectra and in situ measurements are also shown. On the right, the radiance observations are included for comparison. The discrepancies in the visible part of the spectrum in some of the figures were due to differences in estimated aerosol contents. No observations were used in the shaded regions, for which reason all lines within them were simply interpolated.

In the top-right graphic, the differences between the implied radiances corresponding to the reflectance estimates and the observations show that the residuals from the Beckman Lawn retrieval are larger than the others. This was perhaps due to the tight, dense reflectance prior covariance matrix used in the retrieval. The bottom-right graphic shows that the AOE radiance residuals are small; much smaller than the ROE residuals. Comparing the left and the right sides of the figure, we see that AOE was able to find state vectors that were better, both in terms of the residuals to the prior mean, and also in terms of the forward-modeled distances to the radiance observations.

The in situ reflectance measurements in the left panels of Figure 2 (gray dotted lines) demonstrate how meaningful the differences between reflectance estimates from the different retrieval methods were. In all the JPL retrievals, the ground truth was much farther from the reflectance estimates than the AOE and ROE estimates were from each other, but for the Beckman Lawn target, the retrieved estimates closely tracked the in situ-measured reflectances. Since, for any given target, the difference of the retrieved reflectances to the in situ

measurements was quite similar for all the different retrievals, the source of the discrepancy was likely not due to how the optimization algorithms converged, but more probably a combination of biasing introduced by the priors and forward model misspecification.

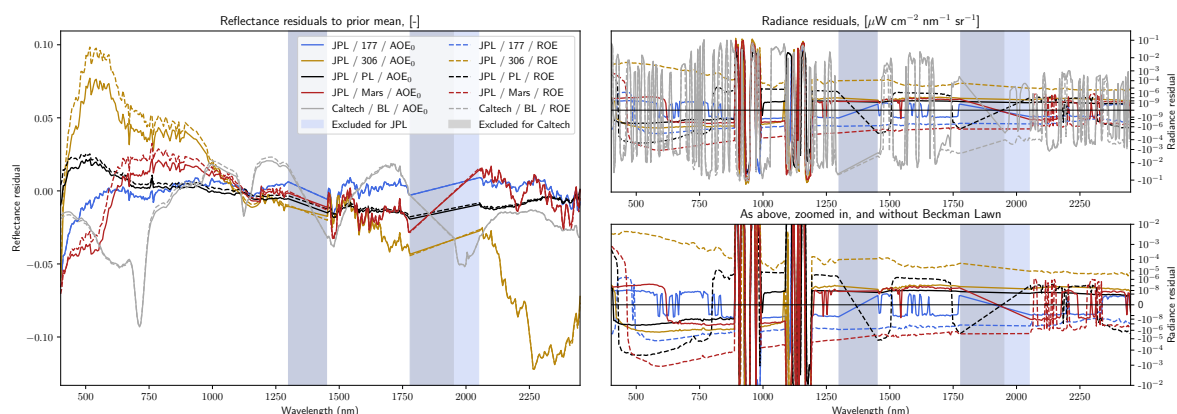


Figure 3. Differences between the retrieved reflectance estimates in Figure 2 and the prior mean (left panel), and between the implied radiances and radiance observations (right panels). AOE₀ spectra are indicated by solid lines and ROE retrievals are indicated by the dashed lines. This figure unambiguously shows that the AOE retrieval was superior, as it was able to consistently find a solution that satisfied both optimization constraints, distance to prior mean and distance to observed radiance, better than the ROE (solid lines of a given color are always closer to zero than dashed lines). In the shaded areas (water absorption features), no observational data were used, and for this reason interpolated straight lines are shown. Figure A2 shows the same data with relative instead of absolute differences.

4.2. Retrieved Atmospheric State Point Estimates

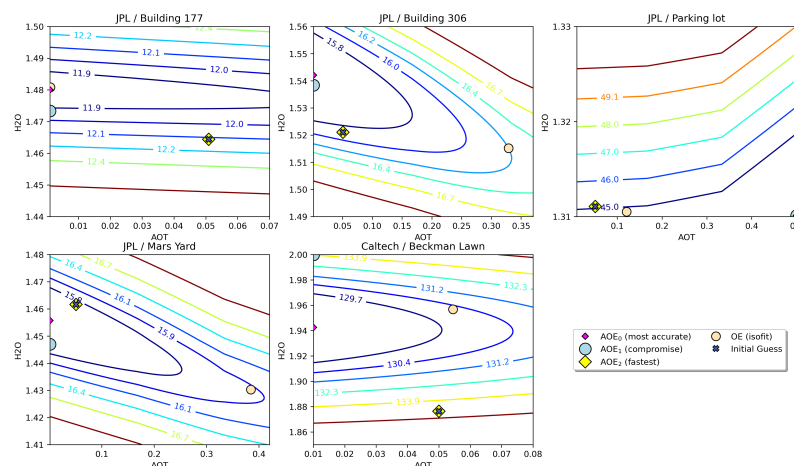
The AOE₀, AOE₁, and ROE algorithms (see Table 1) retrieved estimates of the atmospheric state in addition to surface reflectance spectra. The estimated values for these parameters for the AOE/ROE retrievals are shown in Table 3, and a graphical representation is given in Figure 4. The table and the figure show that AOE and ROE produced very similar H₂O point estimates. However, ROE had difficulty in finding the best AOT estimate in three of the five cases, while AOE₀ and AOE₁ always found the optimal AOT values. The atmospheric parameters used in AOE₂ are by definition the same as their initial values, which were determined by applying a heuristic band ratio algorithm to calculate H₂O and by setting a fixed value for AOT, and therefore, the green cross and the yellow diamond in Figure 4 are always superimposed.

The contours in Figure 4 describe the negative posterior log probabilities in each atmospheric state using AOE₂-estimated surface reflectances (conditional MAP) for the atmosphere given by the x - and y -axes. In more detail, both the H₂O and AOT parameter ranges were divided into 100 intervals, resulting in a grid of size 101×101 . For each point $\tilde{\zeta}$ in the grid, we found the conditional MAP estimate, namely $\hat{\rho} := \arg \max_{\rho} p(\rho|\tilde{\zeta}, \mathbf{y})$. The values described by the contours are then $\frac{1}{2}\tilde{\zeta} \left([\tilde{\zeta}, \hat{\rho}]^T \right)$, see (12).

By comparing the red diamonds with the contours in Figure 4, we can see that AOE₀, the slowest but most accurate AOE method, always found its way to the true MAP estimate given by the lowest point (highest probability) according to the contours. The blue circles show that AOE₁ (compromise between speed and accuracy) performed almost as well. The fastest version, AOE₂ (yellow diamonds), did not retrieve ζ . Instead, it used fixed values equal to the initial guess (black crosses). ROE performed quite well but had some difficulty finding the best AOT values.

Table 3. Point estimates of ζ : all retrieval algorithms and all test targets.

Target	AOT [-]				H ₂ O [g cm ⁻²]			
	AOE ₀	AOE ₁	AOE ₂	ROE	AOE ₀	AOE ₁	AOE ₂	ROE
177	0.001	0.001	0.051	0.001	1.48	1.473	1.464	1.481
306	0.001	0.001	0.051	0.329	1.542	1.538	1.521	1.515
PL	0.5	0.5	0.051	0.121	1.31	1.31	1.311	1.311
Mars	0.001	0.001	0.051	0.384	1.456	1.447	1.462	1.43
BL	0.01	0.01	0.05	0.054	1.943	2.0	1.876	1.957

**Figure 4.** Point estimates of atmospheric parameters in ζ for all retrieval algorithms. The contours show the unnormalized negative log probabilities of state vectors, where the atmospheric part is given by the x and y axes, and the surface reflectances are the corresponding conditional MAP estimates computed with the AOE inner loop with $n_{\text{iter}} = 2$.

Finally, we remark that the retrieved AOT values were very low—perhaps unrealistically low—for most of the retrievals. However, this does not reflect a problem with the atmospheric correction procedure itself, as the AOE algorithm managed to find the most probable states. Furthermore, the contour plots in Figure 4 reveal that the posterior marginal distribution of the AOT parameter was diffuse in all cases, meaning that any low AOT value is entirely realistic. Bias in Bayesian inversion is technically a result of the interplay of the two terms in the loss function, prior and likelihood. The latter includes, in this case, factors such as unmodeled physics, a too-sparse lookup table for radiative transfer, and inadequate forward model error description. The posterior densities can then be used to inform how the modeling setup should be developed further, in order to produce more realistic estimates.

4.3. Quality of Gaussian Approximations from AOE and OE

All the retrieval algorithms studied in this paper produced posterior distributions of the state vector given the observed radiance vector. Both ROE and AOE estimate the first two moments which, under Gaussian assumptions, fully specify the distribution. MCMC-based algorithms produce full distribution estimates, Gaussian or not. Since two different distributions can have similar first (and second) moments and still be very different, it is prudent to compare distributions more carefully than just by looking at means and covariances.

4.3.1. Posterior Probabilities of MAP States and Total Variation Distances

We used two metrics to quantify how close the two different distributions were to one another. The first was the negative log posterior density value evaluated at a distribution's

most probable state, $-\log \pi(\hat{\mathbf{x}})$. The lower this value, the higher the probability of the estimated MAP state. The second metric was the total variation distance, $D_{TV}(\pi_1, \pi_2)$, where π_1 and π_2 are the two distributions being compared:

$$D_{TV}(\pi_1, \pi_2) = \int_{\mathcal{X}} |\pi_1(\mathbf{x}) - \pi_2(\mathbf{x})| d\mathbf{x}, \quad (28)$$

where $\mathcal{X}$ is the domain of all possible values of $\mathbf{x}$. D_{TV} is a number between zero and two; a small TV distance means that the probability masses of distributions largely coincide. Numbers close to two mean that the distributions do not overlap. The computation of TV distance is somewhat involved. The details are described in Appendix C.

Table 4 shows the negative log posterior probabilities of the MAP estimates obtained from the different retrieval methods, with the best values in bold type. The Beckman Lawn retrieval in the last row had higher reported values throughout than the rest, due to the dense, more restrictive prior covariance matrix used in only that experiment and the larger number of channels used in the inversion. The sum of negative log posterior probabilities over 177, 306, PL, and Mars, describing the average convergence performance, was 84.99 for AOE₀, 85.15 for AOE₁, 88.73 for AOE₂, and 88.95 for ROE. According to this metric, all the AOE retrievals performed better than ROE, but AOE₂ and ROE were close to each other. Curiously, in the case of the PL retrieval where ROE was better than AOE₂, the AOE₂ reflectances seemed to be tracking the ROE retrieval closely according to the spectra in Figure 2, implying that evaluating convergence performance solely based on visual inspection of retrieved reflectances can be misleading. For the BL retrieval, ROE performed worse than AOE, but once again the difference was not so large that this statement could be readily made based only on looking at Figures 2 and 3.

Table 4. Negative log probabilities of the MAP estimates, and TV distances of the distributional approximations to MCMC results. The best values are shown in boldface.

Target	$-\log \pi(\cdot)$			
	$\hat{\mathbf{x}}_{\text{AOE}_0}$	$\hat{\mathbf{x}}_{\text{AOE}_1}$	$\hat{\mathbf{x}}_{\text{AOE}_2}$	$\hat{\mathbf{x}}_{\text{ROE}}$
177	11.89	11.93	12.09	11.90
306	15.43	15.45	15.84	16.15
PL	42.04	42.04	45.03	44.94
Mars	15.63	15.72	15.78	15.96
BL	128.6	133.9	135.2	141.2

Table 5 compares the various Gaussian distribution estimates with the MCMC results using the total variation distance. The best estimates values are again shown in bold type. The Gaussian approximations are centered either on the AOE₀ (first three columns) or the ROE mean (last two columns). The first column used $\mathbf{C}_{\text{AOE}}$ as the the covariance estimate, given by (25); the second and the fourth used $\mathbf{C}_{\text{OE}_C}$, and the third and the fifth used $\mathbf{C}_{\text{OE}_F}$. While both $\mathbf{C}_{\text{OE}_C}$ and $\mathbf{C}_{\text{OE}_F}$ were computed with (10), in the former case, the matrix $\mathbf{F}$ in (10) did not contain the partial derivatives with respect to the atmospheric parameters, whereas $\mathbf{C}_{\text{OE}_F}$ used an $\mathbf{F}$ that included them. For the ROE-centered estimates, the total variation distance was small for Building 177 and large for the other scenes. This was because ROE found a very good atmospheric state estimate for Building 177, but not for the others, see Figure 4. It appears that the AOE covariance estimate was the worst performer, and the conditional OE covariance was the best. In fact, $D_{TV}(\pi_{\mathbf{C}_{\text{OE}_C}}^{\mu_{\text{AOE}}}, \pi_{\text{MCMC}})$ was close to zero, suggesting that a Gaussian approximation of the posterior worked well. Furthermore, the difference between the $\pi_{\mathbf{C}_{\text{OE}_C}}^{\mu_{\text{AOE}_0}}$ and $\pi_{\mathbf{C}_{\text{OE}_F}}^{\mu_{\text{AOE}_0}}$ columns is quite small, which suggests that, in many cases, $\pi_{\mathbf{C}_{\text{OE}_C}}^{\mu_{\text{AOE}_0}}$ can serve as a good proxy for $\pi_{\mathbf{C}_{\text{OE}_F}}^{\mu_{\text{AOE}_0}}$, if this is wanted,

e.g., for computational reasons; the convergence in the atmospheric part of the state seems to be far more important. We later see that, in practice, even the $\pi_{\mathbf{C}_{\text{AOE}}}^{\mu_{\text{AOE}_0}}$ estimate, which had higher TV distances with the MCMC results, produced accurate marginal variances and covariances. We take a second look at Gaussianity and algorithm convergence in Appendix A.

Table 5. Total variation distances of distributional approximations to MCMC results. The best values are shown in boldface.

	$D_{\text{TV}}(\cdot, \pi_{\text{MCMC}})$				
	$\pi_{\text{CAOE}}^{\mu_{\text{AOE}_0}}$	$\pi_{\text{COEC}}^{\mu_{\text{AOE}_0}}$	$\pi_{\text{COEF}}^{\mu_{\text{AOE}_0}}$	$\pi_{\text{COEC}}^{\mu_{\text{OE}}}$	$\pi_{\text{COEF}}^{\mu_{\text{OE}}}$
177	0.23	0.0089	0.091	0.11	0.14
306	0.86	0.017	0.23	2.0	2.0
PL	0.34	0.013	0.11	2.0	2.0
Mars	0.42	0.0098	0.080	2.0	2.0
BL	1.89	0.31	0.43	2.0	2.0

4.3.2. Marginal Distributions

In this section, we demonstrate that all the covariance approximation methods produced marginal distributions that very closely matched the MCMC results. This included the $\mathbf{C}_{\text{AOE}}$ estimate, with which the TV distances in Table 4 were larger than with the others. Figure 5 shows, for each scene, the first, second, and third standard deviation envelopes of the posterior distributions produced by the different retrieval and covariance estimation methods for all test targets. The distributions have been centered around the AOE_0 mean reflectance estimate.

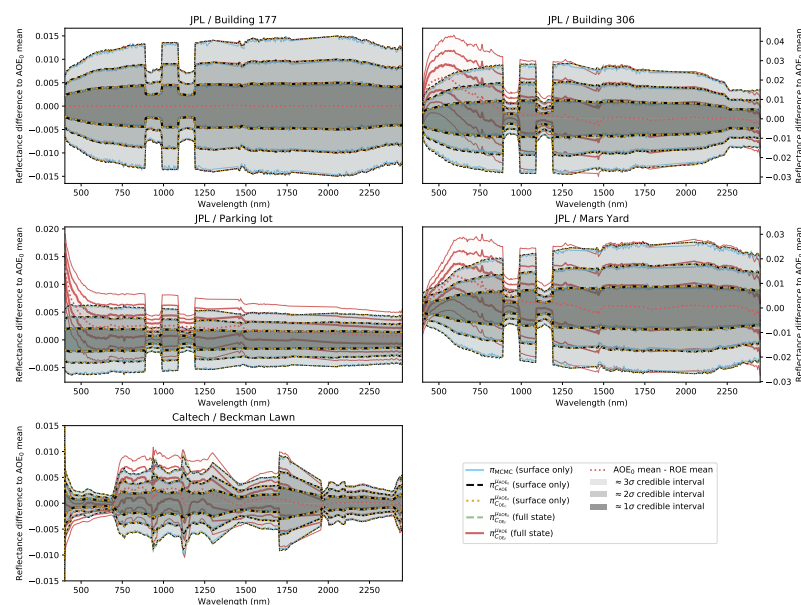


Figure 5. Marginal standard deviation envelopes for all retrieval methods and all targets. Different shadings indicate 68.3%, 95.5%, and 99.7% inter-percentile ranges (1σ , 2σ , and 3σ credible intervals in the Gaussian case). The y -axis shows differences from the AOE_0 mean state, which are the most probable states that were found for these retrievals. The difference between the ROE and AOE_0 means is shown by a dashed red line. Straight-line regions at approximately 1350 nm and 1850 nm are, as before, sections where no radiance data were used and for this reason interpolated lines are shown. The words “surface only” and “full state” in the legend refer to whether variation in the atmospheric state was accounted for in the covariance estimation. Please note that the green line is most of the time covered by the black and the yellow lines.

For target 177 in the first panel, the means and marginal uncertainties of all retrievals are practically identical. That the ROE and AOE_0 means are the same is underlined by the dotted red line, which shows their difference, being zero. For the other targets, the retrieved mean states from ROE and the other methods do not agree. The red shaded areas, which show one-, two-, and three- σ marginal uncertainties of the ROE retrieval, are translated by the difference to the AOE_0 mean. Therefore, even when the standard errors are the same, when the means disagree, the shaded areas corresponding to the AOE and ROE uncertainty ranges are not superimposed. Furthermore, covariance estimates computed at different state space locations also led to different forward model linearizations in (10), which then resulted in differing uncertainty estimates. These discrepancies can be seen for at least parts of the spectrum in all the retrievals other than 177.

For all scenes, the marginals from MCMC (blue solid line), $\pi_{\text{CAOE}}^{\mu_{\text{AOE}_0}}$ (black dashed line), $\pi_{\text{COEC}}^{\mu_{\text{AOE}_0}}$ (orange dotted line), and $\pi_{\text{COEF}}^{\mu_{\text{AOE}_0}}$ (green dashed line) agree almost exactly. In case of the BL retrieval, the larger variation in the marginal uncertainties compared to the JPL retrievals was due to the different structure of the prior.

4.3.3. Comparing the AOE Covariance Estimates to MCMC Results

The marginal envelopes in Figure 5 do not fully describe the AOE and OE distributional approximations, since they do not account for correlations between channels. Figure 6 compares the posterior correlation structures from the AOE and the MCMC sample for the BL target. The AOE result, CAOE , was based on the covariance matrix $\tilde{\mathbf{S}}$ computed in (25). The two correlation structures look very similar, which together with Figure 5 suggests that the AOE posterior covariance approximation described the cross-channel correlations accurately. The main difference is the noisiness of the MCMC correlation structure, which is to be expected, because MCMC is based on randomly sampling the posterior distribution. According to Table 4, the AOE covariance estimate for BL was the least accurate of the 25 computed estimates, and therefore the OE approximations can be expected to resemble the MCMC results even more closely.

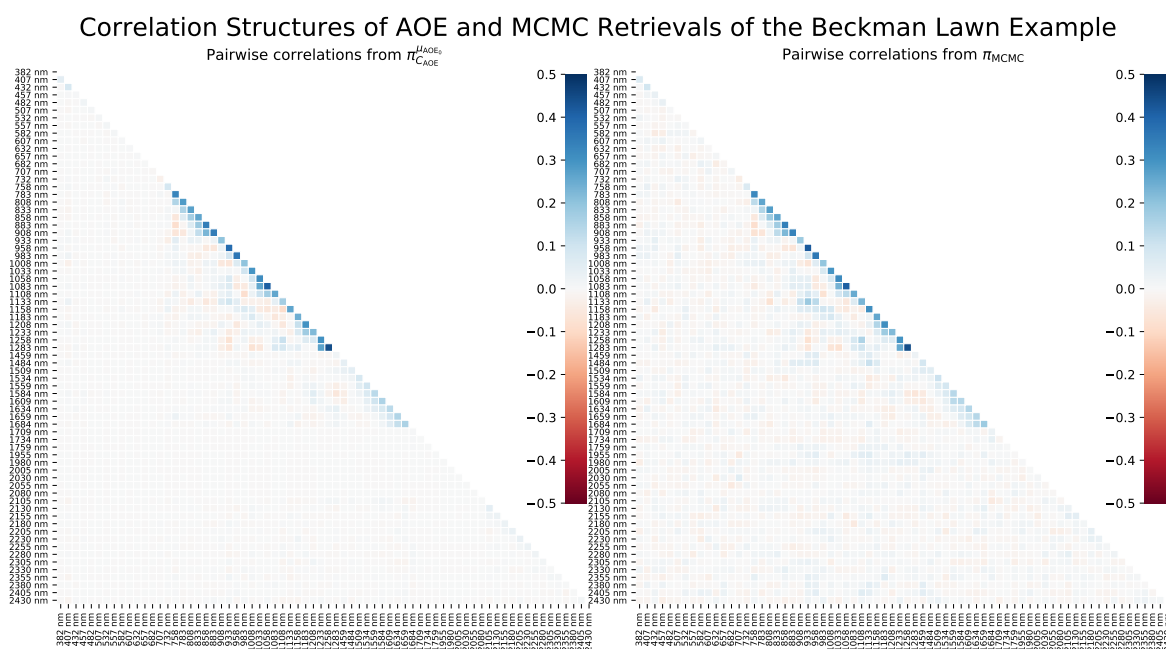


Figure 6. Correlation structures from the Beckman Lawn retrievals using AOE (left panel) and MCMC (right panel). The correlation structure on the left corresponds to the $\tilde{\mathbf{S}}$ matrix in (25) rescaled, and the one on the right describes the sample cross-correlations of the MCMC data. The covariance structures were found to be very similar, notwithstanding minor noisiness in the MCMC plot.

4.4. Retrieving Part of a Flightline with AOE

To demonstrate the scalability and real-world performance of the AOE algorithm, we retrieved a scene containing 160,000 pixels with both AOE₀ and ROE. The scene, described in Section 3 and shown in Figure 1, contains the JPL and surrounding areas. The retrieval setup was briefly described in Section 3. The results here are meant to demonstrate that AOE, even without any tuning, is a capable retrieval algorithm; for operational purposes, many of the modeling choices could be tuned further to obtain even better results.

We performed both ROE and AOE₀ retrievals and evaluated the relative convergence of the algorithms by computing the posterior log probability differences for each pixel in the scene. The initial values for both ROE and AOE were given by the ROE implementation, which also chose a prior from five possibilities for each pixel. The priors were scaled with a pixel-specific scaling factor, as well as using heuristics in ROE.

The final log probability differences are shown as a map in Figure 7. The blue color in this figure indicates that AOE₀ found a better state estimate than ROE, and the red color means that ROE converged to a more probable solution. The absolute log probability differences between the ROE and AOE₀ final states were less than five in 73.62% of the retrievals. Given the modeling assumptions for the retrieved scene, such log probability differences indicate that the reflectance estimates were similar. AOE₀ was better by more than 5 for log probability in 26.02% of the retrievals, and ROE was better than AOE₀ by at least that much for 0.36%.

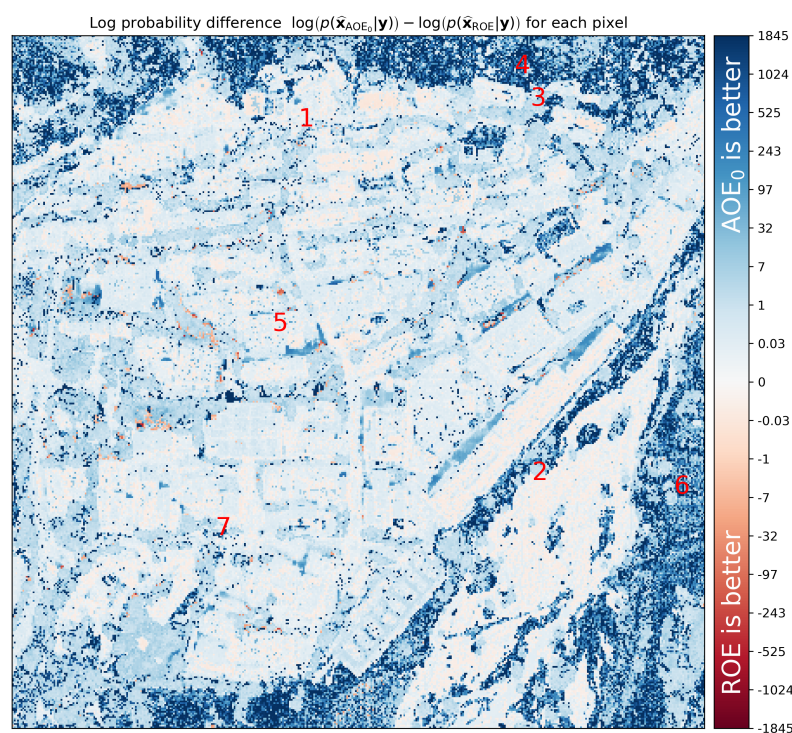


Figure 7. Log probability difference between ROE and AOE₀ retrievals. Blue colors signify a higher probability and better convergence of AOE₀, whereas red colors indicate that ROE was able to find a better state estimate than AOE. Notice that the color bar colors are scaled polynomially. The numbers 1–7 indicate where the example pixels, analyzed in the text and referred to in Table 6 and Figure 8, are located.

Table 6. Summary of the results for the test targets in the JPL scene retrieval, Figure 7. Best values are shown in boldface.

Target Number	AOT		H ₂ O		−log $\pi(\cdot)$		Max Relative $\hat{\rho}$ diff. (% of $\hat{x}_{AOE_0}$)
	AOE ₀	ROE	AOE ₀	ROE	$\hat{x}_{AOE_0}$	$\hat{x}_{ROE}$	
1	0.0	0.0	1.72	1.72	200.26	200.26	0.1
2	0.0	0.0	1.92	1.92	233.1	233.1	0.1
3	0.0	0.0	1.89	1.89	174.75	178.09	1.8
4	0.0	0.18	1.88	1.89	270.45	282.5	15.8
5	0.0	0.0	1.92	1.92	517.21	517.22	0.1
6	0.0	0.0	1.91	1.9	213.22	215.49	0.9
7	0.0	0.24	1.89	2.16	254.14	1339.8	22.0

Overall, the ROE final states had higher posterior probabilities than the AOE₀ final states in 12.04% of the retrievals, but for most of these, the difference was very small: only in 0.63% was ROE better by more than 0.05 in log probability. The 0.1, 0.5, and 0.9 quantiles of $\log p(\hat{x}_{ROE}|\mathbf{y}) - \log p(\hat{x}_{AOE_0}|\mathbf{y})$ over the whole scene were −0.00096, 0.18, and 823. In practice, AOE₀ almost always produced final states that were at least as good as those from ROE, and in a significant fraction of the retrievals, the AOE₀ final states were much better.

Figure 8 shows for seven example pixels how the retrieved reflectances from ROE and AOE₀ (left side) and the corresponding radiance residuals (right side) look. The first example is an asphalt target in the northern part of JPL. The second example is a sand target in the Hahamongna Watershed Park. The third and the sixth examples are vegetation targets, and the fourth target is a bare hillside. The fifth target is a bright rooftop of a large office building, and the seventh is a paved road. The target pixels in question correspond to the red numbers in Figure 7, where the exact locations are in the lower left corners of the numbers.

The retrieved atmospheric parameters, associated negative log probabilities, and relative maximum absolute reflectance differences are given in Table 6. For targets 1 and 2, the negative log posterior probabilities from AOE₀ and ROE were the same, and for target 5, the difference was only 0.01. The maximum absolute reflectance difference for all three retrievals was under 0.0003, and the maximum relative difference was 0.1%, which is well under 2%, which we here consider the upper limit of an acceptable deviation. With targets 6 and 3, we start to see small differences in the log posteriors, and even though the maximum absolute differences in reflectance were still only 0.001 and 0.002, the respective maximum relative errors were already 0.9 and 1.8%. For target 4, the retrieved AOT values differ, and this resulted in a log probability difference of 12 and a relative error of 15.8%, due to reflectance differences in the lower end of the spectrum, where the absolute retrieved reflectance was very low. Even though the relative error was large, Figure 8 shows that the shapes of the AOE- and ROE-retrieved reflectances were very similar. The last target describes a situation where the ROE retrieval did not converge to the best atmosphere, and this led to both a huge difference between the posterior probabilities and a large maximum relative reflectance difference of 22%. The maximum absolute reflectance difference for the seventh target was 0.044.

Overall, Figures 7 and 8 demonstrate that AOE could easily produce results that were superior to the results from ROE. The large differences between the posterior probabilities of the $\hat{x}_{AOE_0}$ and $\hat{x}_{ROE}$ states seen in Figure 7 were often due to the atmospheric states found by AOE₀ being much better than those from ROE. This is exemplified by example target 7 above. One possible approach to work around this issue with ROE would be to use multiple initial starting points for the atmosphere. While this often helps with convergence, the strategy comes with added computational costs.

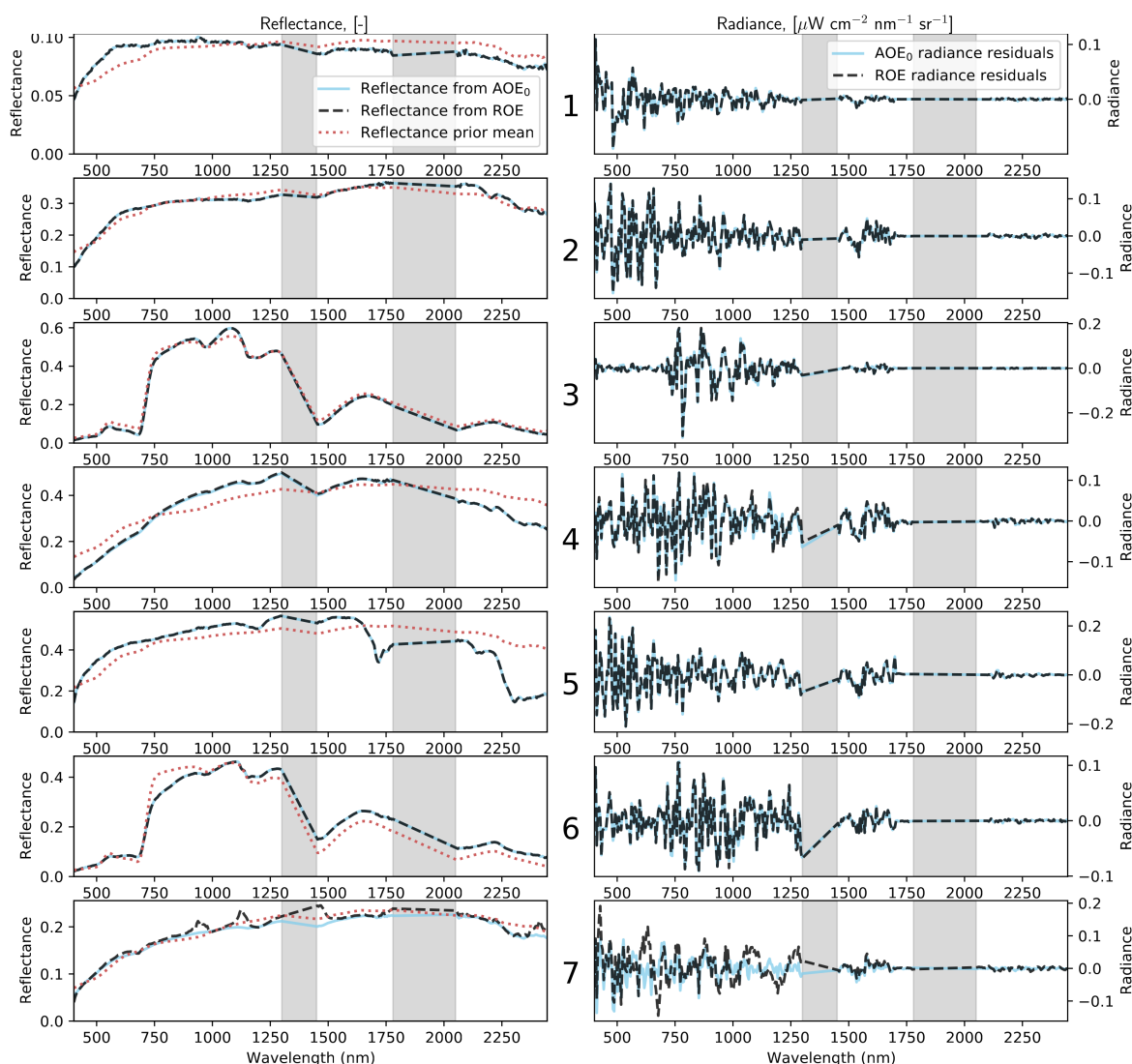


Figure 8. Reflectances and radiances for the targets numbered in Figure 7 and Table 6. The left side shows the retrieved reflectances and the prior mean, and the right side shows the final radiance residuals for ROE and AOE₀ solution states. Numbers 3 and 6 are vegetation pixels, and the rest are not. Number 1 is a pathologically dark target in deep shadow. The reflectances retrieved by AOE and ROE are very similar for 1–6, despite the log probability differences shown in Table 6. In 7, ROE failed to converge to the correct atmospheric state, while AOE₀ found a physically much more reasonable solution. As before, the sections with gray background color are the ones where no radiance measurements were used, and for this reason interpolated straight lines are shown.

4.5. Computational Performance

Table 7 shows timing information for the point estimate retrieval algorithms. The computations were performed on a single core of an Intel I7-8750H laptop processor, on a small-overhead Linux system, setting the CPU frequency governor to “performance” and not running any intensive processes in the background. The first three rows show the timings for AOE₀, the second three rows for AOE₁, and the third set of rows shows the AOE₂ and ROE timings. For the first two sets, the timings are broken down into the inversion (i.e., the outer loop optimization in (13)) and the final step, in which we computed the final conditional distributions using all channels (recall that AOE₁ used only half of the channels for the inversion). In general, the BL retrievals were slower, since they used a dense prior covariance matrix, making the matrix inversions more costly. The last section of the table

provides the speed-up factor, which shows how many times faster the AOE retrievals were compared to ROE.

The most accurate version of AOE, AOE₀, took around 10 ms to execute for 177, 306, PL, and Mars, while for BL, the inversion was slower, at 67 ms, yielding 15 retrievals/s on one core. The AOE₀ retrievals for all scenes were about 30 times faster than ROE, and as both ROE and AOE₀ used all channels for retrievals, this computational speed-up was due to the algorithmic improvements in AOE. AOE₁ was two-to-three times faster than AOE₀, due to the reduced number of channels in the inversion, with the full retrieval taking around 5 ms for the JPL targets, and 20 ms for the BL target. To sum up, these retrievals were some 60–90 times faster than ROE. In the case of AOE₂, where only surface reflectances were retrieved and the atmosphere held fixed, the whole retrieval took just over a millisecond for the JPL retrievals and 6 ms for the BL target. While Table 4 shows that the quality of these estimates was a little better than that of ROE, the AOE₂ retrievals were faster by an average factor of 279, with the actual numbers ranging between 207 and 341.

Table 7. Timings for retrieval methods in Table 1 for the test targets listed in Table 2.

	177	306	PL	Mars	BL
AOE ₀ inversion (ms)	8.2	9.3	15.0	8.3	55.3
AOE ₀ final step (ms)	2.7	1.7	1.5	2.2	11.9
AOE ₀ total time (ms)	10.8	11.0	16.5	10.5	67.2
AOE ₁ inversion (ms)	4.3	4.3	3.6	4.2	13.5
AOE ₁ final step (ms)	1.0	0.9	1.4	1.0	6.5
AOE ₁ total time (ms)	5.3	5.2	5.0	5.1	20.0
AOE ₂ total time (ms)	1.5	1.5	1.3	1.1	6.4
ROE inversion (ms)	345	317	438	369	1863
AOE ₀ faster by factor	32	29	27	35	28
AOE ₁ faster by factor	65	62	88	72	93
AOE ₂ faster by factor	230	210	330	340	290

Figure 9 synthesizes information from Tables 4 and 7. It shows the retrieval algorithms as points in a two-dimensional speed–accuracy plane. The different colors correspond to different retrieval methods. The different symbols correspond to different scenes. Probabilities are shown on the *x*-axis and are standardized so that they show the factors by which the retrieval solution states were more probable than the corresponding ROE point estimates. Computation time is shown on a logarithmic scale on the vertical axis, where the numbers indicate how many times faster than ROE a particular retrieval was (higher values are faster). As one might expect, as the computation times became faster (AOE₀ → AOE₁ → AOE₂), the quality of the solutions deteriorated. As in Table 4, Figure 9 also shows that three of the individual AOE-retrieved states, out of 15, were slightly less probable than the ROE counterparts. These differences, PL/AOE₂, 177/AOE₁, and 177/AOE₂, were small (notice log-scale in Figure 9). It also appears that the AOE speed improvement relative to ROE was close to constant for all variants of the AOE algorithm.

Regarding the performance of the JPL scene retrieval in Section 4.4, the ROE retrievals took 48 h 34 min using two CPU cores, while the AOE₀ retrievals took 2 h 48 min using a single CPU core. The average time of a ROE retrieval using two cores was 1.09 s. For AOE₀ retrievals, the average time was 56 ms. While these numbers are not directly comparable, they are in line with the 30-fold speed-up for AOE₀ in the single-CPU simulations reported in Table 7.

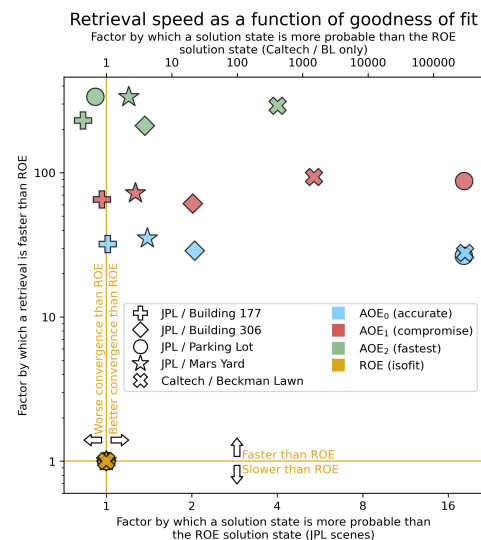


Figure 9. Retrieval performance vs. relative probability of retrieved state. ROE is used as a benchmark, so all the ROE retrievals are superimposed in the lower left. Any retrievals that are further up the y-axis were faster by the scale shown on the left, and any retrievals that are to the right were better in posterior probability by the factor given on the x-axis. Due to the large differences in log probabilities between the Beckman Lawn retrievals and the others, the BL retrieval probabilities are given on a different scale on the top of the figure. For the JPL retrievals, the probability factors are given by the x-axis scale on the bottom.

5. Discussion

The preceding sections described the AOE method and how it compared both in terms of accuracy and performance to ROE, a well-established reference implementation of optimal estimation for imaging spectroscopy retrievals. The results in Section 4 showed that both the accuracy and computational performance of ROE could be improved substantially by using a nested optimization strategy. The AOE algorithm performs inherent dimension reduction by solving the high-dimensional reflectance spectrum part of the state vector without resorting to standard high-dimensional minimization algorithms, but instead iteratively using a closed form of the conditional posterior. Based on the metrics in Section 4, this strategy proved efficient and reliable. In the following subsections, we discuss what the results in the previous section tell us about the convergence, performance, and applicability of the different forms of AOE.

5.1. How Much Does AOE Convergence Depend on the Initial ζ ?

Figure 2 shows that the retrieval bias—the distance from the ground truth to the retrieved reflectance mean spectra—does not change much from one retrieval method to another in relative terms. In practice, this means that for retrieval quality, factors other than the exact form of the minimization algorithm, such as the prior specification, RTM used, exact form of the forward model, etc., are more important.

If the initial guess for the atmospheric state is not far off the MAP state, AOE₀ generally finds the best possible surface state very quickly. If the initial guess is far from the truth, the SLSQP algorithm may converge to a good-but-not-best $\hat{\zeta}$ value. By randomly sampling the initial atmospheric state thousands of times, we found that, for the test targets, the Nelder–Mead algorithm [44] converged to the optimal atmospheric state very reliably, even when the initial guess was far from the MAP state. Even though using Nelder–Mead comes at higher computational cost than SLSQP, using AOE with it is still an order of magnitude faster than ROE. However, the JPL scene retrieval results in Section 4.4 show that the convergence of AOE₀ in the atmospheric part of the state is much better than that of ROE,

and therefore resorting to alternative outer-loop strategies like Nelder–Mead should not be necessary.

5.2. Channel Selection and Multi-Fidelity Approximations

Computing the inverses in (25) is much more expensive when the noise or prior covariances are dense matrices, as seen in the BL timings in Table 7. This may be remedied by using only a subset of the state for the inversion: the AOE₁ retrieval uses half of the channels, and in the BL case this strategy led to a four-fold speed-up for the inversion. Finding a good guess for the atmospheric state is often possible using even fewer channels, say every fifth or even tenth channel. For some targets, however, the response surface of the loss function ζ in (12) changes in a way that shifts the optimum atmospheric state to a different location in the parameter space (see the AOE₁ BL retrieval in Figure 4). Speed and accuracy could be further improved through optimal choice of just a handful of channels for use in the retrieval, such that the effect on the MAP estimates is minimal. We expect optimal channel selection to be prior-dependent, so the channel set could potentially be pre-computed for each candidate prior in advance. Research in the context of L2b retrievals has been carried out by Verrelst et al. [45].

In addition to tuning retrievals, the availability of fast, flexible retrieval methods could be used for effective multi-fidelity [46] inference for suitable estimation problems, such as learning about quantities of interest that require Monte Carlo integration. The AOE speed–accuracy tradeoff knobs (e.g., channel selection, optimization algorithm settings, LUT resolution, exact form of the forward function, included set of parameters) can be used to construct approximations at the required levels of fidelity. Furthermore, the vast quantity of remote sensing data from future missions will provide an opportunity to perform retrievals at multiple resolutions. While some of these aspects are not AOE-specific, the fast computational performance of AOE provides solutions to problems which are intractable with slow algorithms.

5.3. System-Level Performance Considerations

The AOE retrieval code is not a complete retrieval implementation, since it depends on priors and noise covariance models that require separate computation. Other overhead costs include LUT generation and data handling. However, the overall additional computational cost from these sources can be separately controlled: a comprehensive LUT could in principle be reused over and over again, and a multi-fidelity approach could be used to lower the LUT generation cost. Furthermore, at system level, the AOE and ROE implementations are distinct, in that the code paths outside the optimization loops are different, adding some uncertainty to attributing the root causes of the speed-ups. This is difficult to avoid, and for this reason, the performance gap should be understood to describe macro-level differences between two different software implementations.

Since future missions will produce hundreds of thousands of measurements per second, each of which need to be inverted, the availability of a very fast method that satisfies minimum quality requirements is essential for operations. AOE₂, in particular, offers advantages for densely-sampled spatially-adjacent retrievals: with good initial guesses for the atmosphere, the quality of AOE₂ MAP estimates surpasses ROE on average. The range of modeling choices provided by the AOE framework, as exemplified here by AOE₀, AOE₁, and AOE₂, demonstrates that AOE can be used to generate algorithms that satisfy a variety of system needs. As an example, the AOE₂ method was recently adapted for carrying out the operational reflectance retrievals for the EMIT instrument, see commit b725c in [33].

The relative speed-up given by AOE for any retrieval scheme depends on both the specific structure of the problem and the performance of the reference implementation. Regarding the performance figures in this paper, it may be that the ROE implementation is not, despite recent optimization efforts, the fastest. Then again, initial tests of porting the AOE algorithm to the Julia programming language revealed that it will be straightforward to further speed up the AOE retrievals by a large factor from the Python-based timings in Table 7.

6. Conclusions

In this paper, we showed how the widely-used optimal estimation method can be improved with the accelerated optimal estimation algorithm (AOE) in the context of imaging spectroscopy. Compared with a reference implementation of OE (ROE), we were able to either (1) improve the retrieval speed by over two orders of magnitude, while keeping the quality constant, or (2) substantially improve the quality of the retrievals, while speeding them up by a factor of 30 (Figure 9). We also showed that, based on Markov chain Monte Carlo results, both AOE and ROE produced very good conditional posterior distribution approximations of surface reflectance (e.g. Figure 5).

Computational requirements have historically been a significant obstacle to the development of imaging spectroscopy techniques [47]. Even though purely machine-learning-based methods could in principle be even more performant than the AOE algorithm, they do not in general allow for a degree of flexibility, interpretability, and uncertainty quantification comparable to that provided by AOE. As to the computational gains, the AOE algorithm, being a multiscale method, scales well. Taking an imaginary set of a million observations as an example, if just 1% of those pixels were retrieved with AOE₀ or AOE₁, by interpolating the atmospheric state for the other 99% of the pixels and then utilizing AOE₂, the full set could be inverted in under one core-hour—a comparatively modest requirement, given that we did not use any of the most recent innovations in computing to get there.

Author Contributions: J.S. designed and implemented the AOE algorithm and the sampling algorithms used, and performed the simulations, with contributions from all co-authors. D.R.T. and P.G.B. selected and provided the data for the simulation experiments. D.R.T., P.G.B. and N.C. contributed to performing the reference OE retrievals. J.S., N.B., A.B., P.G.B., N.C., M.R.G., H.N., D.R.T. and M.T. had an active role in analyzing the results and preparing the manuscript. J.S. prepared the figures. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). A portion of this work was funded by the Earth Surface Mineral Dust Source Investigation (EMIT), a NASA Earth Ventures-Instrument (EVI-4) Mission.

Data Availability Statement: AVIRIS-NG data are available on the AVIRIS-NG website, <https://avirisng.jpl.nasa.gov/>, accessed on 14 January 2025. The ISOFIT code for the ROE implementation is available at <https://github.com/isofit/isofit/>, accessed on 14 January 2025. An implementation of the AOE algorithm has been added to the ISOFIT software in commit b725c. We are currently working on making a full standalone AOE code available to the public.

Acknowledgments: Jenny and Antti Wihuri Foundation supported carrying out this research.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. A Second Look at AOE Inner Loop Convergence

In order to further verify and demonstrate that the AOE inner loop found the optimal spectral reflectances, we chose six different channels in the visible range of the spectrum,

approximately 50 nm apart from each other, and plotted how the posterior probability of the full state $\mathbf{x}$ changed when both the atmospheric parameters and all reflectance channels except for the one being perturbed were fixed to the ROE solution state. The visible range was chosen, since that is where most discrepancies between ROE and AOE retrievals were found, see Figure 2. The results from this sanity check are shown in Figure A1. In the figure, if a retrieval was successful, then its estimated reflectance (vertical lines) should be at the most probable value (highest point on the gray parabolas within table rows 1–4) for each channel.

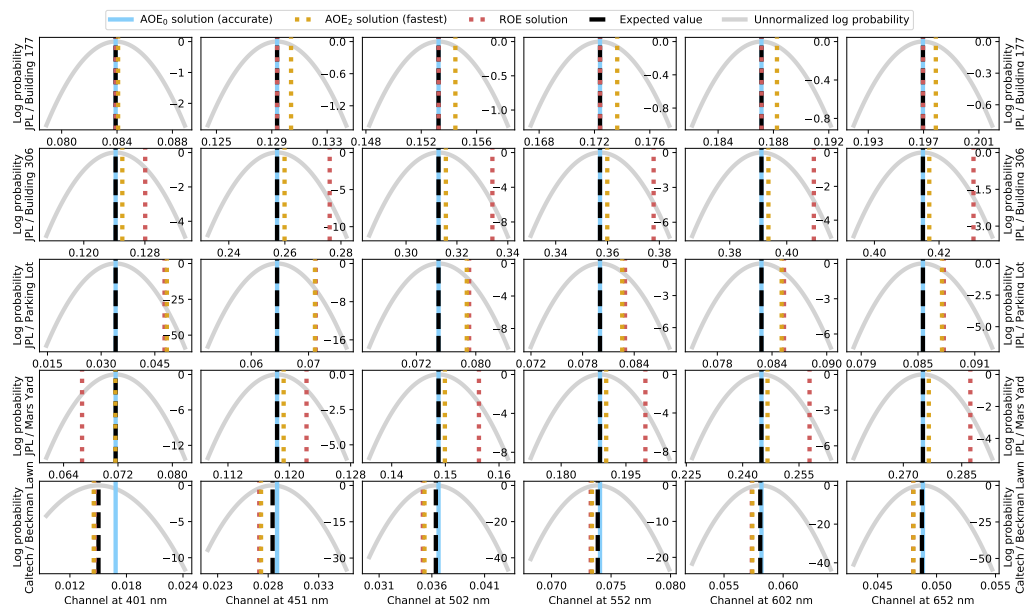


Figure A1. Log probabilities of ROE solution states as functions of perturbing the reflectance of a single channel. The x -axes show the perturbed reflectances, and the y -axes show the unnormalized log probability. Each row corresponds to a different retrieval target (shown on the sides of each row), and each column corresponds to a different choice of the channel that was perturbed (shown under each column). The blue vertical lines that mark the AOE₀ reflectance estimates for each target/channel combination are in the first four rows, exactly at the most probable values given by the gray curves. In the last row, the prior structure is different and the blue lines are not at the gray curve's peak, since the other channels were taken from the ROE solution rather than the AOE₀ solution.

The rows in the figure show the results for the different targets. The gray curves in each figure describe the unnormalized (non-Gaussian) posterior log probabilities of the states, where only the channel named at the very bottom of each column of subfigures was perturbed. The x -axes show these perturbed reflectances, and the y -axis scales show how steep the decay in probability was when the reflectance moved away from the MAP estimate. In the first four rows, the AOE₀ estimate is exactly in the middle of the gray curve, suggesting that AOE₀ did find the most probable state. The AOE₀ estimate also coincided with the expected value, which was computed with numerical integration. Numerical integration is necessary, because the gray curve is not an exact parabola, due to the non-linearity of the forward function $f(\mathbf{x})$. That the gray curves are very close to perfect parabolas is consistent with the agreement between the Gaussian approximation and the MCMC retrieval in Figure 5. Figures A1 and 5 together clarify why an inner loop iteration that resembles a second-order Newton's method iteration converged so quickly, and why the OE-type covariance approximations worked so well for the reflectance retrieval inverse problem.

For the BL retrieval (row 5), the AOE estimates are not at the peak of the gray curve. This is to be expected, since, unlike in the other retrievals, for BL the prior covariance is

a dense matrix and the channels are therefore correlated. As the other channels were set from the ROE solution rather than the AOE solution, the conditional mean of any particular channel given the ROE MAP state for the rest of the state will in general disagree with the marginal mean of the AOE state, which the blue line represents.

The non-Gaussianity of the posterior distributions is very subtly visible in Figure A1, in that the tails of the marginal posterior distributions (gray parabolas) on both sides of the reflectance mean are very slightly wider than those of the Gaussian approximation (not drawn). Furthermore, the tails on the side of lower reflectance are a tiny bit wider than those on the higher side.

Appendix B. Additional Figures

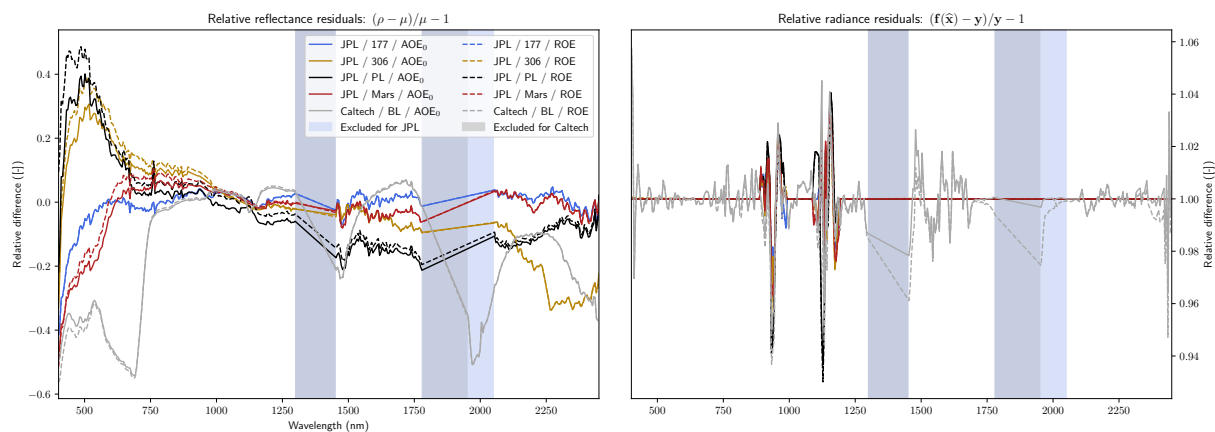


Figure A2. Relative differences from prior means and observed radiances for the test retrievals in Figures 2 and 3. See captions of those figures for additional details. Please note that the solid (AOE) and dashed (ROE) blue lines are on top of each other.

Appendix C. Total Variation Distance Between Distributions

We used total variation distance to estimate the distance between the distributions generated by MCMC and by both ROE and AOE. This is one of the most common distance metrics used for probability distributions. In what follows, the distribution p may be thought of as the MCMC estimate of the posterior distribution, and q then served the role of the Gaussian approximation from either AOE or OE.

The total variation distance between distributions $p(x)$ and $q(x)$, where $x \in \mathcal{X}$, is denoted by $D_{TV}(p, q)$ and defined by

$$D_{TV}(p, q) := \int_{\mathcal{X}} |p(x) - q(x)| dx. \quad (A1)$$

In our case, the space $\mathcal{X}$ is $\mathbb{R}^n$, where n is the number of reflectance channels, and the integration was performed with respect to the standard Lebesgue measure. Assuming that p and q are non-zero in $\mathcal{X}$, we can also integrate the absolute difference of p and q against other measures, as long as the terms are weighted appropriately. Namely,

$$\begin{aligned} D_{TV}(p, q) &:= \int_{\mathcal{X}} |p(x) - q(x)| dx \\ &= \int_{\mathcal{X}} \frac{|p(x) - q(x)|}{b(x)} d\mu_b(x), \end{aligned} \quad (A2)$$

where μ_b is the measure, with density given by some function b . We will use $b(x) = \frac{1}{2}(p(x) + q(x))$.

Assume now that we have samples $X_p := \{x_i^p : i = 1 \dots N\}$ and $X_q := \{x_i^q : i = 1 \dots N\}$ drawn from p and q , and write $X_{pq} = X_p \cup X_q$. In order to solve the discrete problem over samples in X_{pq} instead of the original continuous version, we renormalize the probabilities and sampling weights by setting

$$r_p = \sum_{x \in X_{pq}} p(x), \quad r_q = \sum_{x \in X_{pq}} q(x). \quad (\text{A3})$$

$$p_i = \frac{p(x_i)}{r_p}, \quad q_i = \frac{q(x_i)}{r_q} \quad (\text{A4})$$

$$b_i = \frac{1}{2}(p_i + q_i). \quad (\text{A5})$$

Using the samples instead of the continuous formulation in (A2), we obtain

$$D_{\text{TV}}(p, q) \approx \frac{1}{2N} \sum_{x_i \in X_{pq}} \frac{|p_i - q_i|}{b_i}, \quad (\text{A6})$$

where the leading factor comes from that there were, in total, $2N$ points in X_{pq} . This approximate total variation distance in (A6) is an importance sampling estimate of $D_{\text{TV}}(p, q)$, where b is used as the biasing distribution.

In the results in Section 4, we set N to 50,000. The exact numbers that the procedure above yields depend on the data that are sampled into X_{pq} , and for that reason the numbers in Table 4 are indicative and may vary slightly with resampling. The patterns in that table are, however, very clear and would not be affected by resampling the data.

References

1. CHIME Study Team. *Copernicus Hyperspectral Imaging Mission for the Environment—Mission Requirements Document*; Technical Report ESA-EOPSM-CHIM-MRD-3216; ESA Earth and Mission Science Division: Darmstadt, Germany, 2019.
2. Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; et al. The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sens.* **2015**, *7*, 8830–8857. [\[CrossRef\]](#)
3. Galeazzi, C.; Sacchetti, A.; Cisbani, A.; Babini, G. The PRISMA Program. In Proceedings of the IGARSS 2008–2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008. [\[CrossRef\]](#)
4. National Academies of Sciences Engineering and Medicine. *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*; The National Academies Press: Washington, DC, USA, 2018. [\[CrossRef\]](#)
5. Thompson, D.R.; Braverman, A.; Brodrick, P.G.; Candela, A.; Carmon, N.; Clark, R.N.; Connelly, D.; Green, R.O.; Kokaly, R.F.; Li, L.; et al. Quantifying uncertainty for remote spectroscopy of surface composition. *Remote Sens. Environ.* **2020**, *247*, 111898. [\[CrossRef\]](#)
6. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [\[CrossRef\]](#)
7. Painter, T.H.; Seidel, F.C.; Bryant, A.C.; Skiles, S.M.; Rittger, K. Imaging spectroscopy of albedo and radiative forcing by light-absorbing impurities in mountain snow. *J. Geophys. Res. Atmos.* **2013**, *118*, 9511–9523. [\[CrossRef\]](#)
8. Holzinger, A.; Allen, M.C.; Deheyn, D.D. Hyperspectral imaging of snow algae and green algae from aeroterrestrial habitats. *J. Photochem. Photobiol. B Biol.* **2016**, *162*, 412–420. [\[CrossRef\]](#)
9. Flores-Anderson, A.I.; Griffin, R.; Dix, M.; Romero-Oliva, C.S.; Ochaeta, G.; Skinner-Alvarado, J.; Ramirez Moran, M.V.; Hernandez, B.; Cherrington, E.; Page, B.; et al. Hyperspectral Satellite Remote Sensing of Water Quality in Lake Atitlán, Guatemala. *Front. Environ. Sci.* **2020**, *8*, 7. [\[CrossRef\]](#)
10. Ong, C.; Carrère, V.; Chabrillat, S.; Clark, R.; Hoefen, T.; Kokaly, R.; Marion, R.; Souza Filho, C.R.; Swayze, G.; Thompson, D.R. Imaging Spectroscopy for the Detection, Assessment and Monitoring of Natural and Anthropogenic Hazards. *Surv. Geophys.* **2019**, *40*, 431–470. [\[CrossRef\]](#)
11. Thompson, D.R.; Leifer, I.; Bovensmann, H.; Eastwood, M.; Fladeland, M.; Frankenberg, C.; Gerilowski, K.; Green, R.O.; Kratwurst, S.; Krings, T.; et al. Real-time remote detection and measurement for airborne imaging spectroscopy: A case study with methane. *Atmos. Meas. Tech.* **2015**, *8*, 4383–4397. [\[CrossRef\]](#)
12. Bradley, E.S.; Leifer, I.; Roberts, D.A.; Dennison, P.E.; Washburn, L. Detection of marine methane emissions with AVIRIS band ratios. *Geophys. Res. Lett.* **2011**, *38*. [\[CrossRef\]](#)

13. Green, R.O.; Eastwood, M.L.; Sarture, C.M.; Chrien, T.G.; Aronsson, M.; Chippendale, B.; Faust, J.A.; Pavri, B.E.; Chovit, C.J.; Solis, M.; et al. Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sens. Environ.* **1998**, *65*, 227–248. [\[CrossRef\]](#)
14. Thompson, D.R.; Boardman, J.W.; Eastwood, M.L.; Green, R.O.; Haag, J.M.; Mouroulis, P.; Van Gorp, B. Imaging spectrometer stray spectral response: In-flight characterization, correction, and validation. *Remote Sens. Environ.* **2018**, *204*, 850–860. [\[CrossRef\]](#)
15. Werdell, P.J.; Behrenfeld, M.J.; Bontempi, P.S.; Boss, E.; Cairns, B.; Davis, G.T.; Franz, B.A.; Gliese, U.B.; Gorman, E.T.; Hasekamp, O.; et al. The Plankton, Aerosol, Cloud, Ocean Ecosystem Mission: Status, Science, Advances. *Bull. Am. Meteorol. Soc.* **2019**, *100*, 1775–1794. [\[CrossRef\]](#)
16. Thompson, D.R.; Roberts, D.A.; Gao, B.C.; Green, R.O.; Guild, L.; Hayashi, K.; Kudela, R.; Palacios, S. Atmospheric correction with the Bayesian empirical line. *Opt. Express* **2016**, *24*, 2134–2144. [\[CrossRef\]](#)
17. Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2005.
18. Berk, A.; Hawes, F. Validation of MODTRAN®6 and its line-by-line algorithm. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *203*, 542–556. [\[CrossRef\]](#)
19. Thompson, D.R.; Natraj, V.; Green, R.O.; Helmlinger, M.C.; Gao, B.C.; Eastwood, M.L. Optimal estimation for imaging spectrometer atmospheric correction. *Remote Sens. Environ.* **2018**, *216*, 355–373. [\[CrossRef\]](#)
20. Thompson, D.R.; Babu, K.; Braverman, A.J.; Eastwood, M.L.; Green, R.O.; Hobbs, J.M.; Jewell, J.B.; Kindel, B.; Massie, S.; Mishra, M.; et al. Optimal estimation of spectral surface reflectance in challenging atmospheres. *Remote Sens. Environ.* **2019**, *232*, 111258. [\[CrossRef\]](#)
21. Painter, T.H.; Rittger, K.; McKenzie, C.; Slaughter, P.; Davis, R.E.; Dozier, J. Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sens. Environ.* **2009**, *113*, 868–879. [\[CrossRef\]](#)
22. Frouin, R.J.; Franz, B.A.; Ibrahim, A.; Knobelspiesse, K.; Ahmad, Z.; Cairns, B.; Chowdhary, J.; Dierssen, H.M.; Tan, J.; Dubovik, O.; et al. Atmospheric Correction of Satellite Ocean-Color Imagery During the PACE Era. *Front. Earth Sci.* **2019**, *7*, 145. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Vermote, E.F.; Saleous, N.Z.E.; Justice, C.O. Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sens. Environ.* **2002**, *83*, 97–111. [\[CrossRef\]](#)
24. Guanter, L.; Gómez-Chova, L.; Moreno, J. Coupled retrieval of aerosol optical thickness, columnar water vapor and surface reflectance maps from ENVISAT/MERIS data over land. *Remote Sens. Environ.* **2008**, *112*, 2898–2913. [\[CrossRef\]](#)
25. Stamnes, K.; Li, W.; Yan, B.; Eide, H.; Barnard, A.; Pegau, W.S.; Stamnes, J.J. Accurate and self-consistent ocean color algorithm: Simultaneous retrieval of aerosol optical properties and chlorophyll concentrations. *Appl. Opt.* **2003**, *42*, 939–951. [\[CrossRef\]](#)
26. Zege, E.; Katsev, I.; Malinka, A.; Prikhach, A.; Polonsky, I. New algorithm to retrieve the effective snow grain size and pollution amount from satellite data. *Ann. Glaciol.* **2008**, *49*, 139–144. [\[CrossRef\]](#)
27. Thompson, D.R.; Cawse-Nicholson, K.; Erickson, Z.; Fichot, C.G.; Frankenberg, C.; Gao, B.C.; Gierach, M.M.; Green, R.O.; Jensen, D.; Natraj, V.; et al. A unified approach to estimate land and water reflectances with uncertainties for coastal imaging spectroscopy. *Remote Sens. Environ.* **2019**, *231*, 111198. [\[CrossRef\]](#)
28. Casella, G.; Berger, R. *Statistical Inference*; Duxbury Advanced Series in Statistics and Decision Sciences; Thomson Learning: London, UK, 2002.
29. Carmon, N.; Thompson, D.R.; Bohn, N.; Susiluoto, J.; Turmon, M.; Brodrick, P.G.; Connelly, D.S.; Braverman, A.; Cawse-Nicholson, K.; Green, R.O.; et al. Uncertainty quantification for a global imaging spectroscopy surface composition investigation. *Remote Sens. Environ.* **2020**, *251*, 112038. [\[CrossRef\]](#)
30. Bickel, P.; Doksum, K. *Mathematical Statistics 2e*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2015; Volume 1.
31. Rodgers, C. *Inverse Methods for Atmospheric Sounding: Theory and Practice*; Series on Atmospheric, Oceanic and Planetary Physics; World Scientific: London, UK, 2000.
32. Mueller, J.; Siltanen, S. *Linear and Nonlinear Inverse Problems with Practical Applications*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2012. [\[CrossRef\]](#)
33. ISOFIT Team. Imaging Spectrometer Optimal FITting (ISOFIT). 2021. Available online: <https://github.com/isofit/isofit> (accessed on 13 March 2022).
34. Gamerman, D. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*; Chapman & Hall/CRC Texts in Statistical Science; Taylor & Francis: Sacramento, CA, USA, 1997.
35. Schaepman-Strub, G.; Schaepman, M.E.; Painter, T.H.; Dangel, S.; Martonchik, J.V. Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sens. Environ.* **2006**, *103*, 27–42. [\[CrossRef\]](#)
36. Thompson, D.R.; Bohn, N.; Braverman, A.; Brodrick, P.G.; Carmon, N.; Eastwood, M.L.; Fahlen, J.E.; Green, R.O.; Johnson, M.C.; Roberts, D.A.; et al. Scene invariants for quantifying radiative transfer uncertainty. *Remote Sens. Environ.* **2021**, *260*, 112432. [\[CrossRef\]](#)
37. Nocedal, J. Updating Quasi-Newton Matrices With Limited Storage. *Math. Comput.* **1980**, *35*, 773–782. [\[CrossRef\]](#)

38. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [[CrossRef](#)]
39. Haario, H.; Saksman, E.; Tamminen, J. An Adaptive Metropolis Algorithm. *Bernoulli* **2001**, *7*, 223–242. [[CrossRef](#)]
40. Guanter, L.; Richter, R.; Kaufmann, H. On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. *Int. J. Remote Sens.* **2009**, *30*, 1407–1424. [[CrossRef](#)]
41. Kraft, D. *A Software Package for Sequential Quadratic Programming*; Technical Report 88-28; DLR German Aerospace Center–Institute for Flight Mechanics: Koln, Germany, 1988.
42. Branch, M.A.; Coleman, T.F.; Li, Y. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM J. Sci. Comput.* **1999**, *21*, 1–23. [[CrossRef](#)]
43. Hamlin, L.; Green, R.; Mouroulis, P.; Eastwood, M.; Wilson, D.; Dudik, M.; Paine, C. Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments. In Proceedings of the IEEE Aerospace Conference, Big Sky, MT, USA, 5–12 March 2011; pp. 1–7.
44. Nelder, J.A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313. [[CrossRef](#)]
45. Verrelst, J.; Rivera, J.P.; Gitelson, A.; Delegido, J.; Moreno, J.; Camps-Valls, G. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 554–567. [[CrossRef](#)]
46. Peherstorfer, B.; Willcox, K.; Gunzburger, M. Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. *SIAM Rev.* **2018**, *60*, 550–591. [[CrossRef](#)]
47. Goetz, A.F. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sens. Environ.* **2009**, *113* (Suppl. 1), S5–S16. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.