Journal of Quantitative Spectroscopy & Radiative Transfer 362 (2026) 110049

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy and Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

Retrieval of atmospheric and cloud profiles from high-spectral resolution infrared observations

Michele Martinazzo [a,*], Tiziano Maestri [a], Elisa Fabbri [a], Giuliano Liuzzi [b], Guido Masiello [b], Carmine Serio [b,c]

[a] Department of Physics, University of Bologna, Viale Berti Pichat 6/2, Bologna, 40127, Italy
[b] Scuola di Ingegneria, Università della Basilicata, Via dell'Ateneo Lucano, 10, Potenza, 85100, Italy
[c] Istituto per le Applicazioni del Calcolo "M. Picone" (IAC), Consiglio Nazionale delle Ricerche (CNR), Via Pietro Castellino 111, Napoli, 80131, Italy

ARTICLE INFO

Keywords:
Optimal estimation
IASI
Cloud microphysics retrieval
Cloud profile retrieval

ABSTRACT

This study presents an optimal estimation retrieval framework for deriving atmospheric thermodynamic state and cloud microphysical properties from infrared radiance measurements. The proposed method exploits a fast radiative transfer model, σ, capable of computing atmospheric emission spectra in all-sky conditions. The retrieval framework follows the optimal estimation approach, enhanced by principal component analysis state compression. Cloud prior information is introduced through a dedicated covariance matrix designed to generate physically plausible vertical profiles of ice content and effective dimensions and stabilize the inversion. This combination of methods offers a robust and flexible solution for retrieving atmospheric and cloud parameters from spectrally resolved infrared observations. The retrieval algorithm is tested on both synthetic and real Infrared Atmospheric Sounding Interferometer (IASI) acquisitions. The results highlight the capability of the presented scheme in retrieving cloud property profiles, such as optical depth, effective dimension, and cloud position and its feasible application to operational processors.

1. Introduction

The ability to retrieve cloud microphysical properties from satellite-based infrared sounders has direct implications for meteorology, climatology, and Earth system science [1]. State-of-the-art retrieval methodologies generally rely on optimal estimation (OE) frameworks [2]. These methods combine observations, radiative transfer models, and prior knowledge to solve the ill-posed inverse problem, producing physically consistent solutions with systematic uncertainty quantification. Several all-sky retrieval schemes have been developed for infrared sounders like the Infrared Atmospheric Sounder Interferometer (IASI) [3,4], AIRS [5], CrIS [6] or PREFIRE [7]. Existing approaches, such as those utilizing the SARTA forward model [8] or other all-sky systems, often simplify cloud representation by reducing them to one or two homogeneous slabs or by retrieving only bulk parameters like optical depth and top temperature. Other schemes are able to retrieve the entire vertical profile [9], but are constrained to specific assumptions on initial guess. Unlike previous schemes, this work achieves full ice cloud vertical profile retrieval through an ad-hoc Gaussian-like prior constraint. This approach allows for a stable characterization of the vertical distribution of ice mass and effective dimensions.

It is important to mention that, in recent years, data-driven approaches have emerged as powerful alternatives to traditional physics-based retrievals. Example of these methods are learnable pseudoinverse [10], random forests [11], deep learning architectures [12–14] and others. These machine learning algorithms are capable of exploiting non-linear relationships between measured radiances and atmospheric parameters without explicit forward modeling, making them extremely computationally efficient. However, these approaches require extensive training datasets that must sample the full state space, including extreme atmospheric conditions. The uncertainty quantification is not trivial and requires specialized architectures (e.g. variational autoencoders [15] or ensemble methods [16]). Finally, updating or adapting these models for new instruments, observational geometry, or spectral channels generally requires complete retraining, making them less flexible.

Classical OE retrievals, on the other hand, while computationally more demanding, ensure radiometric consistency between retrieved parameters and observed radiances through explicit forward modeling. The Bayesian framework allows for a straightforward characterization of retrieval errors and provides analysis tools, such as averaging

* Corresponding author.
E-mail address: michele.martinazzo2@unibo.it (M. Martinazzo).

https://doi.org/10.1016/j.jqsrt.2026.110049
Received 13 January 2026; Received in revised form 14 May 2026; Accepted 25 June 2026
Available online 30 June 2026

kernels, which facilitate interpretation of the statistical nature of the inverse problem and the retrieval results. Additionally, OE methods are extremely versatile, allowing elementary updates to forward models, changes in the instrument spectral channels, or the modification of prior constraints without fundamental algorithmic changes.

The computational cost of optimal estimation retrievals has been substantially reduced over the past decades due to advances in fast radiative transfer (RT) models [17–21]. These algorithms employ parameterizations and approximations that significantly accelerate code execution while maintaining high accuracy in radiance calculations. In addition, they are generally based on explicit physical solutions of the radiative transfer problem, thus allowing a simple implementation of new clouds or aerosol parameterizations.

In this work, an optimal estimation algorithm, based on the "σ" radiative transfer model [21], is presented. σ is a fast RT code tailored to compute the entire Earth's emission spectra in all-sky conditions. The model takes advantage of a Chou adjustment routine [22] opportunely extended to correct the radiance computation in the presence of cloud or aerosol layers in the mid and far infrared band (from 100 to 3000 cm^{-1}). Along with the pseudo-monochromatic radiances, the model returns analytical Jacobians for both the atmospheric state and the clouds or aerosols microphysics and concentration.

The optimal estimation algorithm presented here integrates the classical framework [23] with state compression through principal component analysis (PCA) [24] and introduces an idealized Gaussian representation of cloud prior properties to achieve robust convergence in atmospheric retrievals. The application of PCA to exploit low-dimensional structures in atmospheric state spaces has been extensively investigated [9,25–28], demonstrating its effectiveness in reducing computational complexity and improving the algorithm stability, while preserving essential atmospheric information. In this work, state compression is applied to all model variables, including correlations between distinct variables such as cloud microphysics and mass content. Furthermore, compression errors are explicitly characterized and incorporated into the optimal state error representation. Cloud prior knowledge is embedded through an ad-hoc covariance matrix that generates physically realistic Gaussian-like vertical profiles of cloud mass content and effective dimensions, while regularizing the profile retrieval.

The paper is organized as follows: Section 2 introduces the theoretical retrieval framework, detailing the optimal estimation approach. Section 3 describes the specific regularization strategies employed to ensure robust convergence in cloudy conditions. Section 4 presents the application of the algorithm, first validating the performance against a dataset of synthetic IASI observations, and then applying the method to a case study from the MetOp-C satellite over the Australian region. Finally, conclusions are provided in Section 5.

2. Theoretical framework

2.1. Forward model

The radiative transfer model used in this retrieval framework is σ [20,21,29]. σ presents a one-dimensional geometry, and it is capable of simulating a pseudo-monochromatic Earth's emission spectrum under all-sky conditions.

The state vector is specified by the surface temperature (T_s), the reflection property (specular or Lambertian) and the emissivity vector (EM). For the atmosphere, a set of thermodynamic profiles must be specified: the temperature (T) plus the mixing ratio profiles (q) of H_2O, HDO, O_3, CO_2, N_2O, CO, CH_4, SO_2, HNO_3, NH_3, OCS and CF_4. Additionally, the model contains also a set of fixed species that influence the simulated radiance in the forward model. When the presence of scattering layers is considered (clouds or aerosols), additional profiles are incorporated, specifically the mass mixing ratios (kg/kg) for the particles, along with the associated microphysical parameters, i.e. the effective radius (r_{eff}).

The forward model solves the radiative transfer equation in the following form:

$$I_\nu = I_\nu^s + I_\nu^a + I_\nu^r + I_\nu^i \tag{1}$$

where I_ν is the spectral radiance, decomposed in its surface term at the top of the atmosphere (I_ν^s), the contribution from the atmosphere (I_ν^a), the downward infrared radiation reflected at the surface (I_ν^r), and the solar radiation reflected at the surface I_ν^i. All these quantities depend on the wavenumber ν, and the dependence over the directional angle is made implicit.

An explicit description of each of these terms can be found in Masiello et al. [21].

The explicit solution for Eq. (1) is obtained by discretizing the atmosphere into 60 layers with fixed pressure boundaries. The general expression for the transmittance from the level j to the observer can be expressed as:

$$Tr_j = \prod_{i=j+1}^{L} e^{-(\tau_i^{gas}+\tau_i^{cloud})} \tag{2}$$

where τ_j is the layer optical depth of the jth layer for gases, clouds, and, if present, aerosols.

For each layer j the behavior of the optical depth is parametrized as a function of the layer temperature and gas concentration:

$$\tau_{i,j}^{gas} = q_{i,j}\left(c_{0,i,j} + c_{1,i,j}\Delta T + c_{2,i,j}\Delta T^2\right) \tag{3}$$

where ΔT represents the difference between the reference temperature profile and the actual equivalent temperature of the jth layer, while $q_{i,j}$ denotes the concentration of the ith gas in layer j. This parameterization applies to all gases except for water vapor, which is parameterized in the following form:

$$\tau_{H_2O,j}^{gas} = q_{H_2O,j}\left(c_{0,H_2O,j} + c_{1,H_2O,j}\Delta T + c_{2,H_2O,j}\Delta T^2 + c_{3,i,j}q_{H_2O,j}\right) \tag{4}$$

The values of the coefficients $c_{n,i,j}$ are stored by the model in a Lookup Table (LUT).

When scattering particle are present, the model exploits the theory developed by Chou et al. [30] to define an apparent absorption contribution from the layer. The optical depth contribution from clouds or aerosols is computed as follows:

$$\tau^{cloud/aerosol} = (\sigma_{abs}q + b\sigma_{sca}q)\Delta z \tag{5}$$

where q is the particle concentration, σ_{abs} and σ_{sca} represent the absorption and scattering cross-section of the scatterer, and Δz is the geometrical thickness of the layer. The parameter b is called the "backscattering coefficient", and it encapsulates information about the multiple scattering process.

Previous analysis by Martinazzo et al. [20] demonstrated that scaling methodologies introduce acceptable errors in the mid-infrared for most realistic cloud scenarios when computing upwelling spectrally resolved radiances. However, far-infrared simulations yield accurate results only for optically thin ice clouds, with Optical Depth (OD) lower than $0.1-1$. To address these limitations, σ implements the Chou adjustment scheme developed by Tang et al. [22], appropriately extended for spectral radiance calculations as described by Masiello et al. [21].

The adopted cloud model is described by Martinazzo et al. [20].

2.2. Inversion scheme

Assuming a Gaussian description of the measurement errors and the prior state, the posterior probability of the state x given the observation y, $P(x|y)$, is described by the following expression [23]:

$$-2ln(P(x|y)) = [y-F(x)]^T S_\epsilon^{-1}[y-F(x)] + [x-x_a]^T S_a^{-1}[x-x_a] + k \tag{6}$$

where $F(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the model mapping from the atmospheric state space to the observation space, S_ϵ is the covariance matrix associated to the observation (and model) errors, x_a is the prior state and S_a the associated covariance matrix. The parameter k is a constant value. Using the Maximum A Posteriori (MAP) paradigm, the best estimate $\hat{x}$ is assumed to be the most likely atmospheric state given the data and our prior knowledge of the problem. The optimal estimation problem reduces to the problem of maximizing the probability $P(x|y)$:

$$\hat{x} = \arg\max_x ln(P(x|y)) \tag{7}$$

To find the state $\hat{x}$ that minimizes equation (6), the Levenberg–Marquardt algorithm (LM) is implemented in the following iterative form:

$$\begin{aligned} x_{i+1} =& x_i + (\gamma_i S_a^{-1} + K_i^T S_\epsilon^{-1} K_i + \beta_i \mathbb{I})^{-1} \\ & (K_i^T S_\epsilon^{-1}[y - F(x_i)] - \gamma_i S_a^{-1}[x_i - x_a]) \end{aligned} \tag{8}$$

Where γ_i is a variable regularization factor that enables the algorithm to mitigate the effects of a poor initial guess, as demonstrated by Turner and Löhnert [31], Masiello et al. [26] or Carissimo et al. [32]. The term β_i is a scaling coefficient applied to the identity matrix $\mathbb{I}$, which is dynamically adjusted at each iteration to promote numerical stability and ensure convergence. Finally, K_i denotes the Jacobian matrix of the forward model $F(x)$, evaluated at $x = x_i$. The iterative scheme described in Eq. (8) is repeated until a predefined convergence criterion is satisfied [33].

The state vector x contains the profiles of the thermodynamical parameters of the atmosphere (temperature and gas concentration), aerosol/cloud microphysics and mass concentrations, and surface properties (surface temperature and emissivity profile). To enforce physical constraints, like gas concentrations being strictly non-negative, the retrieval employs a logarithmic transformation, mapping the gas concentration and the cloud/aerosol parameters to an unbounded space via a log-normal distribution. This ensures Gaussian-like behavior in the transformed variable while preserving positivity during iterative optimization. The forward model relating the log-transformed state to spectral radiance measurements is:

$$y = F(exp(\tilde{x}_1), x_2) + \epsilon \tag{9}$$

where ϵ represents the measurement noise, $\tilde{x}_1 = ln(x_1)$ represents the transformed subset of components of the vector x, while x_2 consists of the components of x for which no transformation is applied. A similar approach is used to constrain the emissivity in the $[0, 1]$ interval, using a logit transformation [34].

Atmospheric state vectors, as described previously, are inherently high dimensional, especially when the emissivity profile is required. To overcome this problem, the proposed retrieval scheme compresses the state x into a lower-dimensional latent space of principal components (PCs), which capture the main features of the original state. The PCs are derived from the set of parameters that characterize the prior distribution. Given an appropriate description of the covariance matrix S_a with maximum rank n, the following singular value decomposition (SVD) [24] can be applied:

$$S_a = U \Lambda_a V^T = V \Lambda_a V^T \tag{10}$$

Since S_a is symmetric, its left and right singular vectors coincide ($U = V$). The columns of V are the eigenvectors of S_a, and Λ_a is the diagonal matrix of its associated eigenvalues, that are real numbers sorted in decreasing order. For this symmetric (and positive) case, the eigenvalues coincide with the singular values of S_a. Under these assumptions, the projection into the first q (with $q < n$) principal components of the prior offset ($v = x - x_a$) is governed by:

$$z_q = V_q^T (x - x_a) = V_q^T v \tag{11}$$

with $V_q \in \mathbb{R}_{n\times q}$ containing the first q eigenvectors associated with the highest eigenvalues. The optimal number of principal components q to retain is determined by the intrinsic dimensionality of the residual state v. The ideal case is obtained when v in the original space satisfies $v \in span(V_q)$, or by defining the projector $P_q = V_q V_q^T$, when we have $v \simeq P_q v$ for all states that describe the physical world.

After performing the PC compression, the model applies the LM algorithm to find the optimal state in the latent space (omitting the subscript q for simplicity) as follows:

$$\begin{aligned} z_{i+1} =& z_i + (\gamma_i \Lambda_a^{-1} + \tilde{K}_i^T S_\epsilon^{-1} \tilde{K}_i + \beta_i \mathbb{I})^{-1} \\ & (\tilde{K}_i^T S_\epsilon^{-1}[y - \tilde{F}(z_i)] - \gamma_i \Lambda_a^{-1} z_i) \end{aligned} \tag{12}$$

where the forward model is now described as $\tilde{F}(z_i) = F(V z_i + x_a)$ and $\tilde{K} = KV$ is the compresses jacobian matrix. When convergence is reached, the optimal state in the latent space is transformed to the physical state $\hat{x}$ as:

$$\hat{x} = V \hat{z} + x_a \tag{13}$$

In the framework of the optimal estimation, the error covariance matrix of the solution is computed. Under the assumption that the observation error is not correlated with the background state and that the forward model is linear in the region of uncertainty, it is proved [23] that the posterior covariance matrix associated with the state $\hat{x}$ is given by:

$$\hat{S}_{\hat{x}} = (\tilde{G}K - \mathbb{I}) S_a (\tilde{G}K - \mathbb{I})^T + \tilde{G} S_\epsilon \tilde{G}^T \tag{14}$$

where the first element on the right hand side of Eq. (14) is named the smoothing error, while the second element is the retrieval noise. In this formulation, the retrieval gain matrix, $\tilde{G}$, is:

$$\tilde{G} = V(\gamma_i \Lambda_a^{-1} + \tilde{K}_i^T S_\epsilon^{-1} \tilde{K}_i)^{-1} \tilde{K}_i^T S_\epsilon^{-1} \tag{15}$$

An explicit derivation of Eq. (14) and (15) is provided in Appendix A.

Finally, the averaging kernel, or resolution matrix, is:

$$A = \tilde{G} K \tag{16}$$

The rows of A provide an estimation of the resolution in the profiles (see [23]). The features of the parameters of the state vector x that are broader than the width of the averaging kernel can be reproduced by the retrieval, while features much narrower than the width of the averaging kernel are smoothed out. Moreover, the trace of A provides the degrees of freedom for signal, quantifying the number of independent pieces of information that contribute to the reconstruction of the retrieved state $\hat{x}$.

3. Regularization strategies for robust convergence

3.1. Prior distributions and background constraints

In the presence of clouds and aerosols the non-linearity level of the radiative transfer simulations increases since multiple scattering processes must be taken into account [35]. In addition, cloud layers usually constitute a sharp discontinuity with respect to the clear sky case because of their physical and optical properties; thus, they increase the ill-posed nature of the problem. An illustrative example is a scene influenced by an optically thick, dense cloud. In such a case, an infrared sounder onboard a satellite can retrieve information only from the cloud's upper layers and the atmosphere above. Radiation emitted by the surface and by the atmospheric layers beneath the cloud is completely absorbed or scattered, preventing it from reaching the sensor. Consequently, no observational data are available for those lower regions. Because the measured top-of-atmosphere radiance depends solely on the properties of the cloud top and the overlying atmosphere, an infinite number of atmospheric state configurations could reproduce the same observed spectrum. To solve this indetermination, "virtual measurements", as suggested by Rodgers [23], can be adopted. The additional information is usually provided as Bayesian prior statistics

(x_a and S_a in Eq. (6)), and from regularization algorithms or assumptions applied to the forward model. A widely adopted practice is to set constraints on specific features of the cloud (like the geometrical thickness of the cloud) and apply the retrieval solely to a subset of cloud parameters affecting the upwelling radiance. In the present work, we focus exclusively on ice clouds; however, the same assumptions are applicable to water clouds and specific classes of aerosols [36] like transported dust, fire and volcanic plumes.

To achieve an effective regularization in the presence of clouds and aerosols, we build an ad-hoc covariance matrix representing the cloud prior state. In this process, we consider idealized clouds characterized by Gaussian-like shapes in the vertical dimension. The goal of the adopted method is to simplify cloud description by embedding information on variable correlations and profile smoothness, while allowing the model to generate physically realistic vertical profiles [37–39]. The covariance matrix for the ice cloud properties is derived from the Satellite Application Facility on Numerical Weather Prediction database (NWP SAF) [40]. This database consist of 25000 atmospheric scenarios, each including temperature, water vapor, ozone profiles, surface temperature, cloud mass content, and other relevant variables needed to define the scene. The dataset is divided into five equally sized subsets tuned to capture the natural variability of temperature, specific humidity, ozone mixing ratio, cloud condensate and precipitation, respectively.

Ice mass content profiles (in kg/kg) are extracted from the cloud condensate subset and interpolated onto the σ model grid (61 levels) using a log-linear interpolation. For each layer, we calculate an average value representative of its ice water content. This value, along with temperature, pressure, and water vapor mixing ratio, is then used to compute the effective dimension of the ice clouds particles, following the parameterization described in Sun [41]. A state vector comprising both the mass content and the effective particle size is generated, which allows to retain the relationship between cloud particle effective size and ice content.

$$x_{ice} = [IC/c_1, D_{eff}/c_2] \tag{17}$$

where c_1 and c_2 are two standardization constants, calculated as the maximum standard deviation of the two quantities, respectively. Since the goal is to obtain a simplified description of prior knowledge, only clouds characterized by a single mode vertical profile of ice content are used to generate the statistics. Among the unimodal IC profile only those with values exceeding a threshold of 0.25 g/kg at least at one level are selected to build the covariance matrix $S_{x_{ice}}$. This helps to maintain regular cloud profile shapes and to neglect complex multi-modal structures that could introduce spurious correlations in the prior covariance estimation.

An estimation for the prior state (first and second moment of state distribution) is then obtained from the arithmetic mean and the sample covariance matrix.

$$\bar{x}_{ice} = \frac{1}{N}\sum_{i}^{N} x_{ice,i} \tag{18}$$

$$S_{x_{ice}} = \frac{1}{N-1}\sum_{i}^{N}(x_{ice,i} - \bar{x}_{ice})(x_{ice,i} - \bar{x}_{ice})^T \tag{19}$$

Since the inversion scheme described here employs logarithmic transformations, to ensure the positivity of the physical quantity (as shown in (9)), the first and second moments of the distribution of $ln(x_{ice})$ are estimated by Taylor expansion [42].

A smoother prior is obtained by applying a convolution with Gaussian kernel over the covariance matrices, which introduces Gaussian correlations between physical quantities across different model levels. The optimal parameters of the Gaussian kernel are typically problem-dependent. In our study, we applied a discrete Gaussian kernel with $\sigma = 4$ (performed using the scipy.ndimage gaussian_filter function from the SciPy library [43]), which is selected based on the convergence rates.

Once the prior state is defined, additional transformations are applied to generate multiple realizations of idealized clouds. Specifically, the prior is shifted along the pressure grid to represent clouds at different altitudes, extending the retrieval's applicability across a broader range of cloud regimes.

Two pairs of $\bar{x}_{ice}$ and $S_{x_{ice}}$ are generated to represent altitude locations of idealized high (peaked near ∼290 hPa) and low clouds (peaked near ∼810 hPa). The assumption of these two distinct prior positions aims at guaranteeing a total coverage of the troposphere. Finally, a tapering is performed to remove non-physical correlations between model levels. Fig. 1 shows the covariance matrix adopted for high and low ice clouds. The quadrupolar patterns enable the algorithm to adapt the cloud's position, ice water content, and effective radius while preserving a Gaussian-like vertical structure.

Starting from the same atmospheric profile dataset, a set of prior states is also created for the H_2O, O_3, and HDO profiles, the temperature profile, and surface temperature. For these quantities, the covariance matrices are computed independently, ignoring correlations between variables, and no Gaussian filters are applied. The priors are generated for five distinct climatological regions: Polar Winter, Polar Summer, Mid-Latitude Winter, Mid-Latitude Summer, and Tropical. The covariance matrix for surface emissivity is obtained from the Huang dataset.

3.2. Dimensionality reduction and regularization of the inverse problem

A central hypothesis in this study is that it is possible to exploit low-dimensional structures and correlations in state space to reduce the dimensionality of the problem while preserving accuracy in profiles or profile-derived parameter description. When dimensionality increases, the geometric complexity of the state space makes convergence harder, a problem known as the curse of dimensionality. Dimensionality reduction techniques help tackle these issues by making the system more stable and speeding up the optimization algorithm converges [28,44]. Principal Component Analysis stands as one of the most widely used dimensionality reduction techniques. A reduced basis can be constructed by performing an eigen-decomposition of the prior covariance matrix, $S_a = V\Lambda_a V^T$, and retaining only the principal components associated with the largest eigenvalues. This approach efficiently captures the dominant patterns of variability while filtering out less significant directions. Specifically, we can assume that the atmospheric state $x \in \mathbb{R}_n$ lives in a subspace of dimension $q < n$, where again n is the maximum rank of the covariance matrix. Thus, indicating with $P_q = V_q V_q^T$ its projector, we can write:

$$x = V_q V_q^T(x - x_a) + x_a + \epsilon_c = V_q z + x_a + \epsilon_c \simeq V_q z + x_a \tag{20}$$

where ϵ_c is a small projection error. If (20) is valid, the description of the LM iterative method reported in Eq. (8) can be written in a compressed form:

$$\begin{aligned} z_{i+1} =& z_i + (\gamma_i \Lambda_a^{-1} + \tilde{K}_i^T S_\epsilon^{-1}\tilde{K}_i + \beta_i \mathbb{I})^{-1} \\ & (\tilde{K}_i^T S_\epsilon^{-1}[y - \tilde{F}(z_i)] - \gamma_i \Lambda_a^{-1} z_i) \end{aligned} \tag{21}$$

The iterative solution expressed in this form is more stable than that in the state space. Specifically, the matrices $\Lambda_a = V_q^T S_a V_q$ and $\tilde{K} = KV_q$ in the reduced latent space have a lower condition number with respect to the original ones:

$$\begin{aligned} \kappa(\Lambda_a) \le& \kappa(S_a) \\ \kappa(\tilde{K}) \le& \kappa(K) \end{aligned} \tag{22}$$

A proof is provided in Appendix B.

A well-executed dimensionality reduction (i.e., when we assume the relation described in (20) is valid) improves the stability of the inverse problem by filtering out non-physical configurations in the atmospheric state x (i.e., solutions not contained in the subspace), reducing the computational cost of the algorithm, and eliminating the risk of obtaining non-invertible matrices.

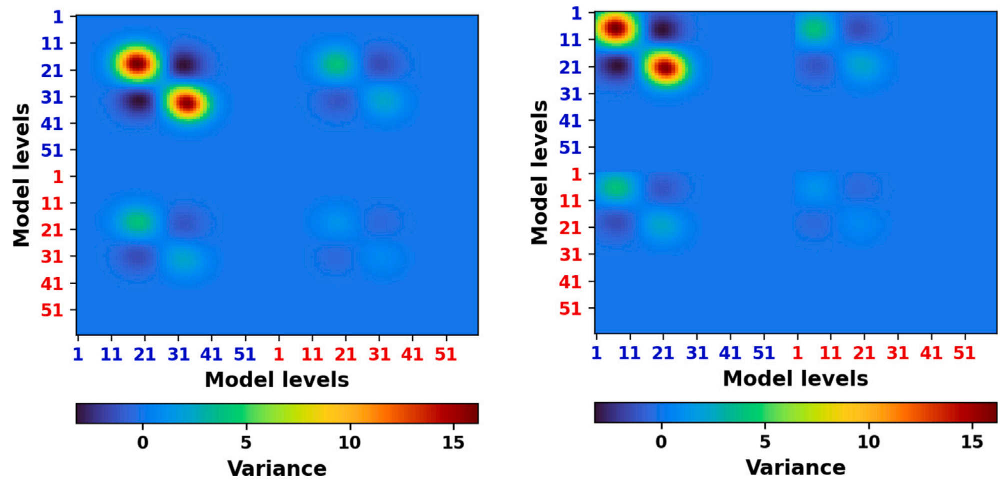

Fig. 1. Covariance matrix of $ln(x_{ice})$ for high (left panel) and low (right panel) ice clouds. The state vector components corresponding to ice water content are highlighted in blue, while those associated with effective radii are shown in red. The covariance is obtained after a 2-D convolution with Gaussian kernel.

Together with dimensionality reduction, the algorithm incorporates an additional regularization strategy by inflating the inverse of the prior covariance matrix using a scalable regularization coefficient $\gamma_i > 1$. This approach increase the robustness of the model by controlling the influence of prior information during the iterative process. In the following, we discuss methodologies for selecting an appropriate set of these coefficients.

Starting from Eq. (21), and omitting the LM coefficient β_i, we can show that the optimal state for the ill-posed non-linear problem is found by iteratively solving a sequence of linearized problems around the current estimate:

$$(\gamma_i \Lambda_a^{-1} + \tilde{K}_i^T S_\epsilon^{-1} \tilde{K}_i) z_{i+1} = \tilde{K}_i^T S_\epsilon^{-1}([y - \tilde{F}(z_i)] - \tilde{K}_i z_i) \tag{23}$$

In this form, the problem of finding the best inflation parameter γ_i is equivalent to finding the best parameter for a Tikhonov regularization [45,46] with a regularization matrix $L = \Lambda_a^{-1/2}$. When $\gamma_i = \gamma = 1$, the usual statistical regularization is obtained, as described in [23].

Different methodologies are available to estimate the optimal set of regularization parameters γ_i, including the L-curve method [47], the generalized cross-validation (GCV) [48], and maximum likelihood estimation [49].

A simpler approach is suggested by [31], which follows a practical guideline proposed by Doicu et al. [49]. At the beginning of the iterative process, large values of the regularization parameter γ_i are used to avoid local minima and to ensure that the resulting linear systems remain well-conditioned. Successively, the regularization parameter is gradually diminished during the iterations, allowing more information from the observations to be used.

In the proposed approach, the parameter γ_i is initialized with a starting inflation value at the first iteration. At each subsequent iteration, γ_i is scaled by a decreasing factor, progressively reducing its magnitude. This iterative attenuation continues until $\gamma_i = 1$ is reached, after which the parameter remains fixed at this value, until convergence is achieved.

To determine the optimal values of the initial inflation and its decreasing factor, a reduced dataset of atmospheric profiles is employed as a test set. These profiles are selected independently from those used to construct the covariance matrix, ensuring that the tuning of the two parameters remains uncorrelated with the dataset used in the main application. The retrieval is performed using the prior information described in Section 3.1, and evaluated for all combinations of initial inflation values (10, 30, 100, 300, 1000, 3000, 10000) and decreasing factors (0.1, 0.3, 0.5, 0.7, 0.9).

The average χ^2 at convergence (or at the final iteration if convergence is not reached) across the entire dataset is used as the metric

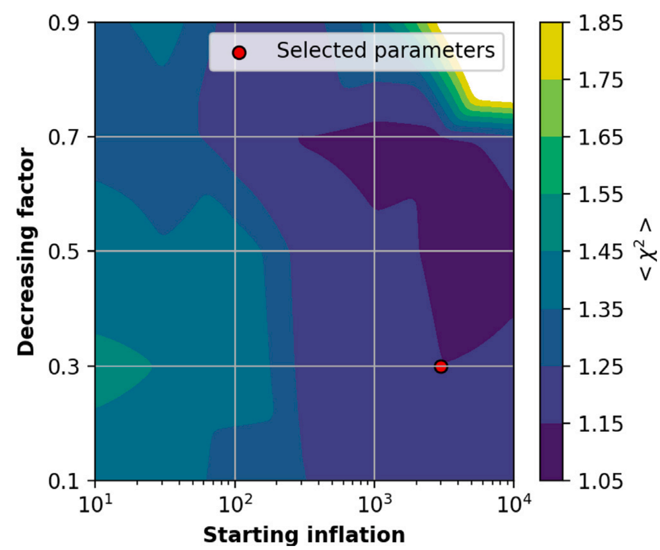

Fig. 2. Starting inflation and decreasing factor map for the optimal estimation. The colormap indicates the average chi-squared in the validation set at convergence. The selected optimal parameters are highlighted with a red dot.

for selecting the parameters. Fig. 2 shows the parameter space, where the dark blue region highlights the pairs of parameters producing the best results. Among all possible combinations, it is good practice to choose the least invasive parameters from those that provide optimal regularization. Specifically, parameters located near the left border of the dark blue region in Fig. 2. An initial inflation of 3000 and a decreasing factor of 0.3 are adopted. We note that, the choice of regularization parameters is inherently problem-dependent and generally non-unique.

4. Application to IASI observations

To assess the performance of the retrieval scheme, the model is tested against a set of synthetic observations, with focus on ice clouds. The inversion code is successively applied to real measurements acquired from MetOp-C and collocated with ECMWF analyses. The observations are taken from the Infrared Atmospheric Sounding Interferometer (IASI), developed by the Centre National d'Etudes Spatiales (CNES) in collaboration with the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). The first IASI instrument was launched on board the MetOp-A polar-orbiting satellite

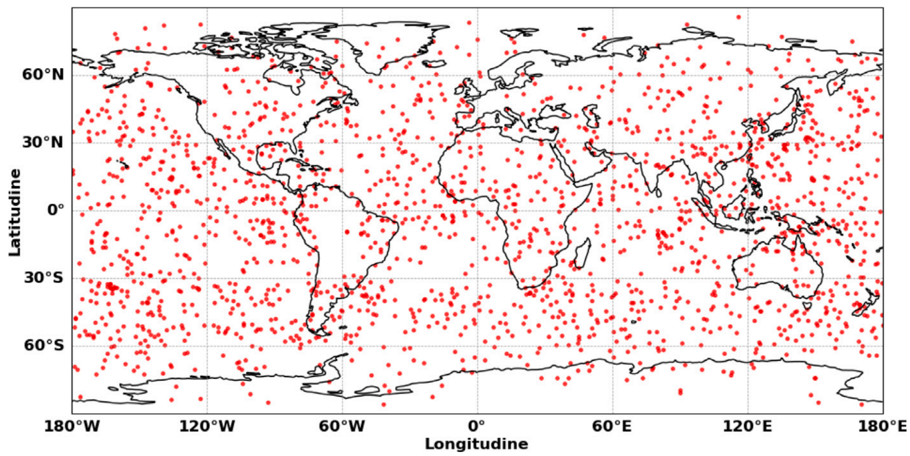

Fig. 3. Geolocation (red dots) of the scenes considered in the simulated IASI acquisitions.

in 2006, followed by two similar instruments on MetOp-B and MetOp-C at the end of 2012 and 2018, respectively. IASI consists of a Fourier transform spectrometer (FTS), based on a Michelson interferometer, designed to resolve atmospheric infrared radiation into line spectra.

The interferometer measures 8461 spectral samples between 3.62 and 15.5 μm (from 645 to 2760 cm^{-1}), with an apodized spectral resolution of 0.5 cm^{-1}. IASI scans the Earth with a swath width of ~2200 km achieving a global coverage in 12 h. The scan is across-track, and consists of 30 successive elementary fields of view (EFOV). Each element consists of a 2 × 2 circular pixel matrix (IFOV). For nadir observations the pixel's shape on the ground is circular, with a diameter of 12 km.

The spectral region between 650 and 1650 cm^{-1} is used for retrieval, as it provides sensitivity to all state vector quantities [50] while avoiding contamination from shortwave solar radiation.

4.1. Preliminary analysis on synthetic observations

To perform an initial consistency test and evaluate possible model biases, we applied the inversion routine to 1833 realistic atmospheric profiles taken from an independent subset of the NWP SAF database [40], distinct from those used to generate the prior knowledge. This dataset includes vertical profiles of temperature, H_2O, O_3, liquid and ice water content (IWC), together with surface properties, and is selected to represent the natural variability of atmospheric conditions worldwide.

For the current test, only ice clouds are considered and used to generate the scene, and the liquid water profile is ignored. The relation between the IWC and the effective dimension of the ice crystals is parameterized using the relation described by Sun [41]. The assumed cloud model is the one described by Martinazzo et al. [20], with columnar aggregates habit [51] and a gamma description of the particle size distribution.

Only scenes characterized by a total optical depth of the cloud greater than 0.01 and a surface pressure greater than 990 hPa are considered. This is done to have a similar number of model levels across different scenes and to improve the consistency between profiles when producing statistics. Fig. 3 shows the geographical distribution of the scenes selected for the analysis.

The main atmospheric absorbers in the IASI spectrum that are not present in the SAF data set (CO_2, CO, CH_4, N_2O, HNO_3, SO_2, NH_3, OCS, and CF_4) are derived from the IG2 climatological profiles [52]. The surface emissivity, assumed to be lambertian, is seasonal and spatial dependent [53]. For each scene, the IASI acquisition is simulated using the σ forward model over the spectral region between 650 and 1650 cm^{-1}, and assuming a nadir observation geometry. Finally, each synthetic spectra is perturbed accordingly with the IASI L1-C measurement noise [54]. Out of simplicity, the measurement error is assumed to be uncorrelated.

The physical inversion is applied to the entire set of atmospheric states, simultaneously retrieving profiles of temperature (T), water vapor (H_2O), ozone (O_3), heavy water (HDO), ice cloud water content (IWC), effective radius (r_{eff}), as well as surface temperature (T_S) and spectral emissivity (EM). The state vector is provided in expression (24)

$$x = [T, H_2O, O_3, HDO, IWC, r_{eff}, TS, EM] \tag{24}$$

In these retrievals, we assume that the only prior information available is derived from climatological data. No additional knowledge about cloud position is provided. For each scenario, the initial estimate is taken to be equal to the prior distribution, which is generated as described in Section 3.1. Moreover, all profiles that do not appear directly in the state vector are assumed to be known quantities during the retrieval phase. This assumption is equivalent of assuming a perfect forward model, without errors in model's parameters. Thus, the only source of error in this analysis comes from the instrument error (and the compression error), which are well described in the retrieval scheme. The algorithm selects the appropriate climatological prior according to the latitude and month of the scenario. Each retrieval is performed twice: first using the idealized high-cloud prior defined in Section 3.1, and then repeated using the idealized low-cloud prior, while all other parameters are held constant. The solution associated with the lower convergence chi-square value is taken as the optimal retrieval. This dual-run strategy mitigates the challenges arising from the assumption that no a priori information is available regarding cloud altitude. The algorithm is presented in Fig. 4. Convergence is deemed achieved if the convergence criterion is met and the final chi-squared value is below 1.1. Out of the 1833 scenarios analyzed, 1780 reached convergence, corresponding to approximately 97% of the total.

Since clouds are characterized by a range of parameter values across their vertical profile rather than a single value, three quantities are defined to quantify the quality of the retrieved cloud parameters:

- the total cloud optical depth at 900 cm^{-1}, named OD_{tot},
- the cloud effective radius at the level where the integrated cloud optical depth (from the cloud top) reaches 1, named r_1,
- the cloud position (expressed in hectopascals hPa) at the level where the integrated cloud optical depth reaches 1, named p_1.

Following the above method, for clouds with an OD lower than 1, the quantities r_1 and p_1 are not defined. Therefore, in these cases, the ice water content-weighted averages of the cloud position and effective radius are computed instead.

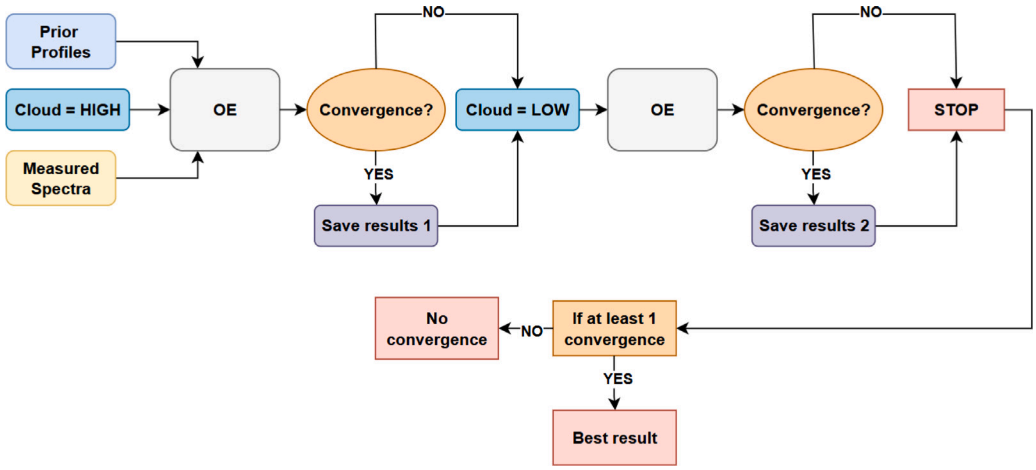

Fig. 4. Flowchart of the two-step cloud retrieval algorithm. The system performs consecutive OE retrievals assuming high- and low-altitude cloud priors, saving the output of each successful pass. If at least one step converges, the optimal atmospheric state is selected based on the lowest χ^2 convergence value.

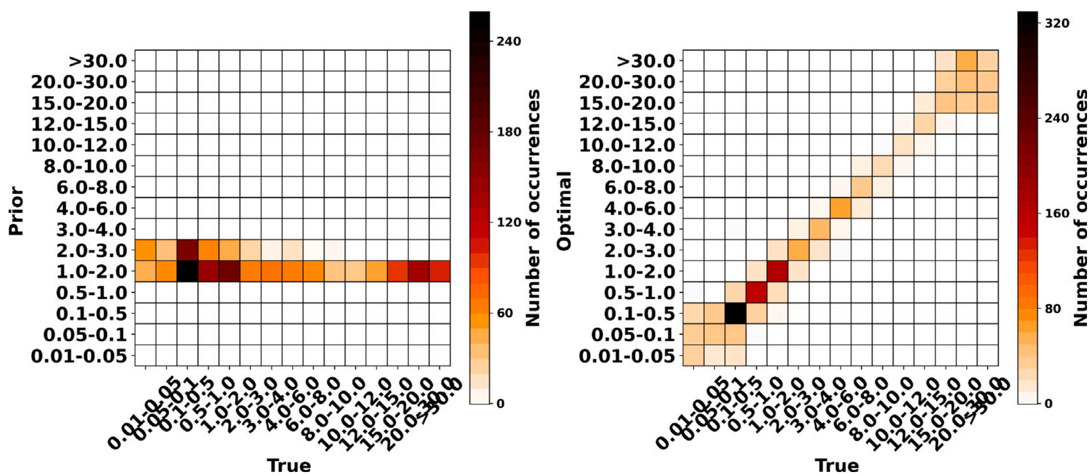

Fig. 5. Total optical depth for the ice cloud. The left image shows the distribution of the prior total optical depth versus the true total optical depth. The right image shows the same distribution for the posterior.

Table 1
Spearman's rank correlation coefficient and Pearson's correlation coefficient for prior and optimal realizations of OD_{tot}, r_1 and p_1, computed over the entire dataset.

Parameter	OD_{tot}		r_1		p_1	
	Prior	Optimal	Prior	Optimal	Prior	Optimal
Pearson	−0.31	0.62	−0.44	0.71	0.76	0.90
Spearman	−0.22	0.97	−0.12	0.75	0.67	0.95

The results obtained from the 1780 scenarios are presented in Figs. 5, 6, and 7. These figures show the dispersion diagrams for OD_{tot}, r_1 and p_1 respectively. For each dispersion diagram, both Spearman's rank correlation coefficient and Pearson's correlation coefficient are calculated. While Pearson's index measures the strength of a linear relation between two variables, Spearman's index assesses the strength of a monotonic relationship, linear or not, and is more robust to outliers. The results obtained are reported in Table 1.

In an idealized scenario, the dispersion diagram for the optimal retrieved quantity should present a clear linear relationship between true and optimal quantities, with no off-diagonal elements. However, different factors can cause the deviation from ideality, such as retrieval errors or saturation effects. The left panel of Fig. 5 shows a general excellent agreement between the retrieved optical depths and the true optical depths values. Deviation from the ideal results are observed for optical depth smaller than 0.1 and for those larger than approximately 10. In both cases, the model loses sensitivity to variations in optical depth. In case of optically thin clouds the cloud signal becomes negligible while for optically thick clouds saturation effects play a major role. These deviations influence the value of the Pearson's index (Table 1 column for OD), which is 0.62 for the optimal estimation of the total cloud OD. When limiting the analysis to clouds with retrieved total optical depth less than 10 (thus excluding the thick clouds but including the thin ice clouds) the Pearson correlation coefficients increase to 0.98 while the Spearman coefficient is 0.95, both indicating a strong linear relationship between the true and optimal values. The model also proves to be unbiased, with a Mean Bias Error and Root Mean Square Error ($MBE \pm RMSE$) of -0.32 ± 6.58 over the entire dataset (0.02 ± 0.35 for retrieved total optical depth less than 10).

In Fig. 6 the prior and retrieved r_1 (effective radii at $OD = 1$) are compared to the true values. The prior effective radii ranges from 14 to 26 μm while the retrieved ones mostly disperse on the full range covered by the true effective radii (from 6 to 60 μm). Some cases with effective radii between 26 and 40 μm are slightly underestimated by the retrieval with predicted values in the 18-30 μm range. Both Spearman and Pearson coefficient are above 0.7 (see Table 1). The posterior MBE and $RMSE$ is 0.82 ± 10.0 μm over the entire set of scenarios.

The position of the cloud are also accurately predicted by the retrieval. In Fig. 7 the prior cloud pressure at $OD = 1$ are set at

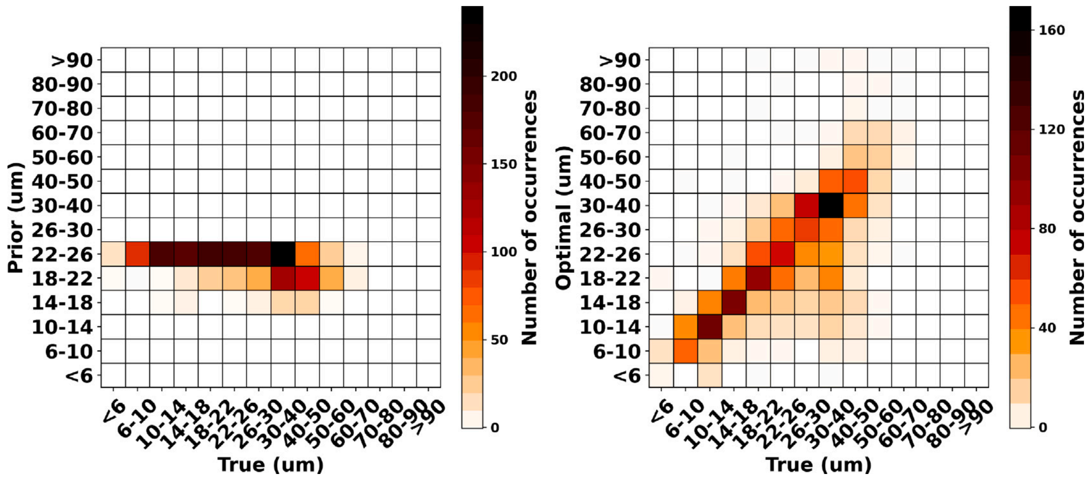

Fig. 6. As Fig. 5 but for the effective radius (μm) for the ice cloud at $OD = 1$.

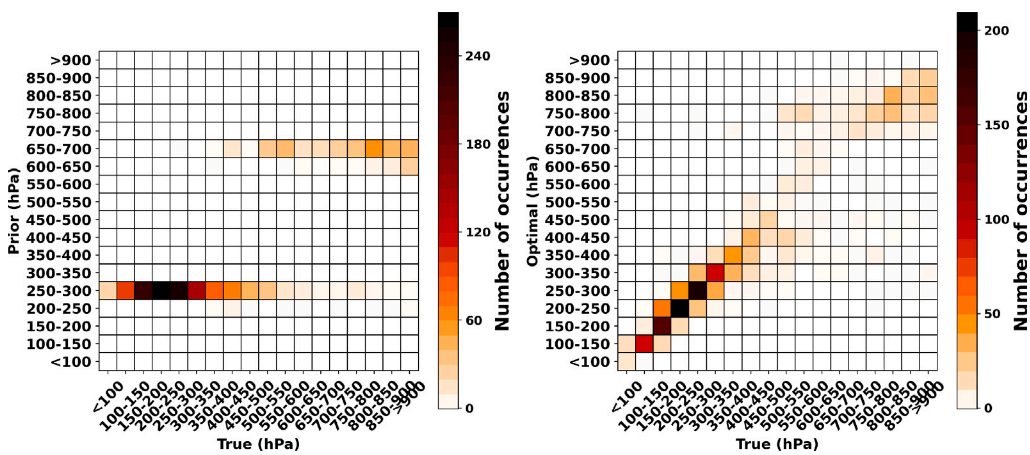

Fig. 7. As Fig. 5 but for the position (hPa) of the ice cloud at $OD = 1$.

approximately 290 and 810 hPa. The retrieved values are in agreement with the true ones especially for high clouds (pressure below 450 hPa). Some dispersion between the true and the retrieved values is observed for low level clouds. Finally, MBE and $RMSE$ over the entire dataset is 20.51 ± 111.21 hPa.

By comparing the prior and optimal dispersion diagrams, with the related indices, the retrieval model successfully extracts information on cloud OD, effective radius and cloud pressure from the measurement, updating our prior knowledge of the system. However, we note that these are derived quantities and only partially describe the full capability of the retrieval code, which is capable to extract the complete profile of cloud mass content and effective particle size. At this regard, we compute the cosine similarity for the optical depth profile for each scenario. This metric shows an average value of 0.57 for the prior profiles improving to 0.73 for the optimal state, indicating a good performance in retrieving the cloud vertical profile. It is important to point out that the cloud prior directly influences the algorithm's ability to resolve finer structures in the vertical dimension. For some specific cases, such as geometrically thin or double layer clouds, it could be useful to use different assumptions for the Gaussian kernel. To this end, the choice of the σ parameter for the filtering operation of the cloud covariance can be considered as a hyperparameter of the problem.

To conclude, the relative errors for the retrieved vertical profiles are presented. Fig. 8 highlights the average percentage error, over the entire dataset for temperature (upper left panel), water vapor (upper right), HDO (lower left) and ozone (lower right) profiles. The average values for surface temperature and emissivity are not shown for brevity. However, since all the considered cases contains clouds, the sensitivity of the retrieval to these quantities is low. Specifically, when the cloud OD approaches and exceeds values of about 5, the radiance sensitivity to atmospheric parameters below the cloud level rapidly decreases to zero.

The results are mapped over the forward model vertical grid and show a general improvement of the optimal with respect the prior. The percentage errors are usually lower in the upper and lowest layers of the atmosphere. The mid atmospheric layers correspond to the region typically associated with cloud presence. In this region the performance in retrieving the vertical profiles are lower, as the clouds mask part of the information from the profiles.

4.2. Application to real IASI acquisitions

As a final test, the retrieval code is applied to a set of collocated IASI measurements and ECMWF analyses with a temporal coincidence interval of ±15 minutes, as done by Masiello et al. [21]. The case study focuses on a subset of observations from MetOp-C orbit number 14732 over the Australian region, acquired on September 9, 2021. For each IASI spectrum, ECMWF analyses provide surface temperature, temperature profile, H_2O mixing ratio, O_3 mixing ratio, as well as ice and liquid water contents. The effective radius of ice particles is estimated using the Wyser approach [55].

Since we are interested in scenes containing exclusively ice clouds, an ice-cloud index is defined based on the collocated analysis information which is taken as the reference in the assessment of the algorithm

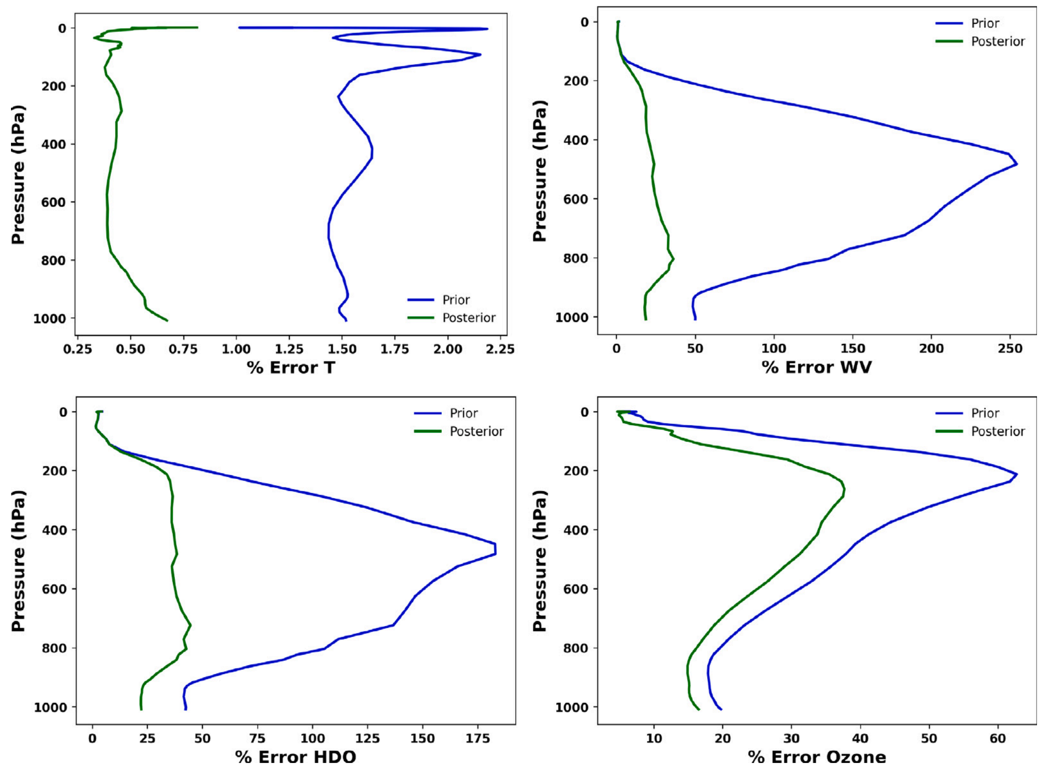

Fig. 8. Temperature (upper left panel), water vapor (upper right), HDO (lower left) and ozone (lower right) average percentage error profiles as a function of the model pressure level.

performances. The mathematical form of the index used to identify the IASI pixels containing ice clouds is described in Eq. (25):

$$\phi_{ice} = \frac{M_{iw} - M_{lw}}{M_{iw} + M_{lw}} \tag{25}$$

Where M_{iw} is the columnar integrated ice water mass (kg), while M_{iw} is the integrated liquid water mass (kg). The index value goes from -1, when only liquid water cloud is present in the scene, to 1 in the presence of ice clouds only. Fig. 9 shows the ϕ_{ice} index for the IASI acquisitions, obtained from the collocated analysis. The retrieval is run for pixels where the condition $\phi_{ice} > 0.95$ is met, resulting in a total of 410 IASI acquisitions.

The retrieval setup follows the same assumptions and parameter choices used for the simulated data discussed in Section 4.1, with the exception of the prior information. The ice-cloud prior retains the idealized formulation introduced earlier, whereas the priors for surface temperature, surface emissivity, and the atmospheric profiles of temperature, water vapor, ozone, and HDO are constructed individually for each scene from the collocated analysis. This approach ensures that the posterior estimates of the ice cloud properties derived through optimal estimation are not influenced by the corresponding ECMWF analysis, thereby preserving the independence of the retrieved cloud parameters. The IASI measurement noise used in this analysis is provided by EUMETSAT [54]. To account for potential forward model biases and correlations, the measurement noise estimate is appropriately inflated with a factor $\sqrt{2}$ [56].

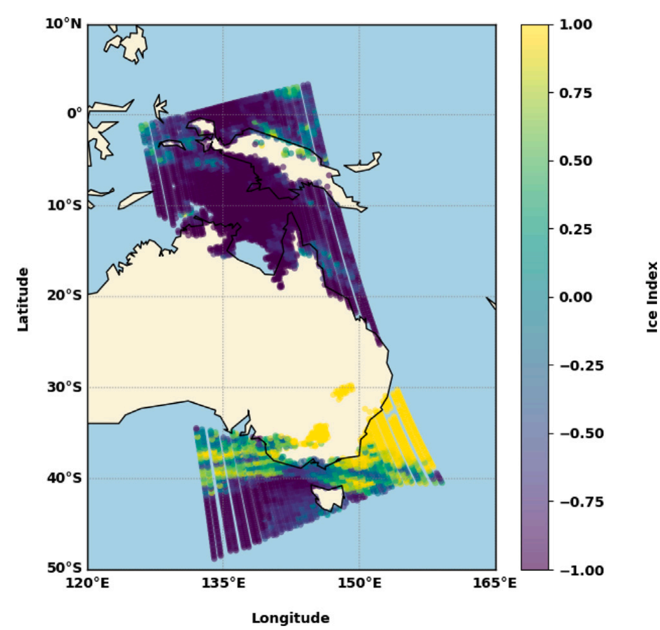

Fig. 9. Ice cloud index values from IASI acquisitions.

Fig. 10 shows the spectral residuals computed as:

$$\Delta I_x(\nu) = \frac{1}{N}\sum_{i=1}^{N}[I_{x,i}(\nu) - I_i(\nu)] \tag{26}$$

Where I_i represents the measured spectrum from IASI for the ith acquisition, while $I_{x,i}$ corresponds to the radiance calculated by the forward model using the prior state, the optimal state, and the analysis, respectively. The differences between the simulations based on the prior state and the analysis are due to the cloud parameters since the atmospheric profiles and surface properties are assumed to be identical in both cases. On average, the upwelling radiance computed using the derived parameters from the ECMWF analysis exceeds the IASI observations, likely due to an underestimation of cloud thickness. The second panel in Fig. 10 shows a zoomed-in view of the average residuals between the optimal retrieved state and the IASI acquisitions, compared with the IASI L1C measurement noise (not inflated). The analysis of the radiance residuals is interpreted in terms of possible

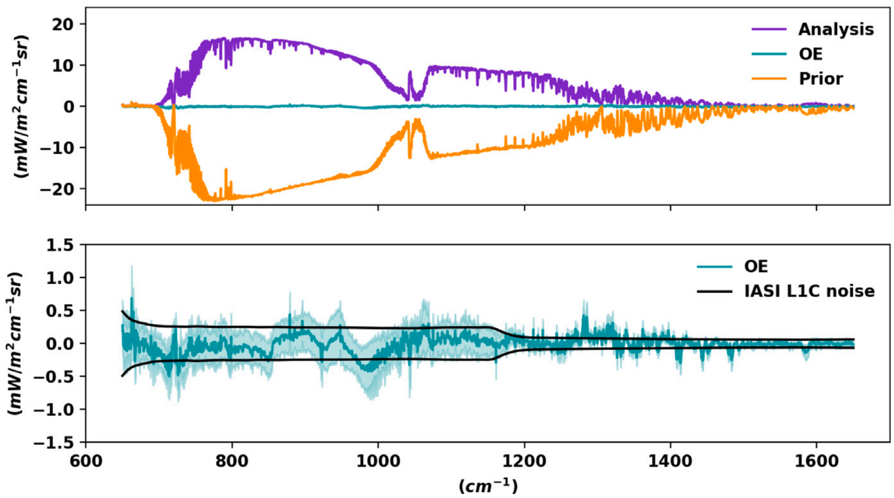

Fig. 10. The upper panel shows the average spectral residuals of the analysis (purple line), prior simulation (orange line), and optimal state (blue line) with respect to the IASI acquisitions. The second panel compares the optimal state to the IASI L1C radiance noise (solid black line) and its standard deviation (shaded area).

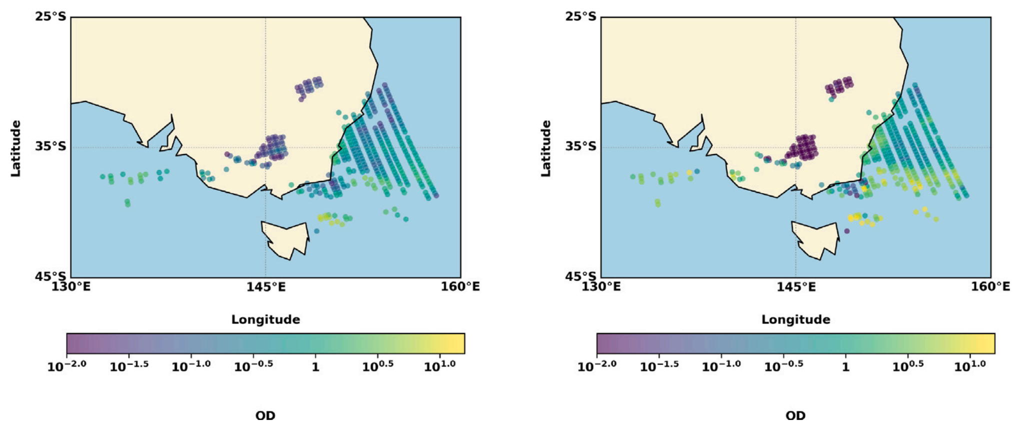

Fig. 11. Left panel: total optical depth OD_{tot} computed from ECMWF analysis ice-cloud profiles collocated at the IASI footprint. Right panel: total optical depth OD_{tot} computed from the retrieved ice-cloud profiles at the same footprints.

model inconsistencies, cloud misclassifications, unmodeled species, or spectroscopic biases. A spectral signal is noted in the atmospheric window at approximately 970 cm^{-1}. The smooth variation in the wavenumber of this bias, along with its position in the atmospheric window, is consistent with a discrepancy between the cloud model used in the forward model and the actual cloud present in the scene. In the spectral region from 1250 to 1320 cm^{-1}, the N_2O absorption band is visible, along with the HNO_3 absorption band between 1300 and 1350 cm^{-1}. These species are not included in the retrieval state vector, so their signals remain visible. In all the other spectral regions, the retrieval fits the IASI spectra within the error bar, and therefore we can conclude that the retrieval product is consistent with the observations.

Similarly to the approach used for the simulations, OD_{tot} and r_1 are used to characterize the cloud properties retrieved from real IASI acquisitions. The panels on the left sides of Figs. 11 and 12 show the spatial distributions of OD_{tot} and r_1, respectively, over the region of interest as derived from the ECMWF analysis. The same quantities are compared to those obtained from the retrievals (right panels of the same Figures). Despite using the same prior knowledge and initial guess for each IASI pixel, the retrieved spatial distributions of total optical depth and effective radius reproduce spatial patterns consistent with those shown by the ECMWF analysis. Optically thin cirrus clouds are observed over the interior of Australia, showing a smaller optical depths compared to the analysis (0.01 versus 0.05 on average) while maintaining similar effective radii. A significant portion of the offshore IASI pixels, in the region spanning from 150° to 160° longitude east and 30° to 39° latitude south, is covered by optically thick clouds. Within this area, a southern sector is characterized by higher effective radii and optical depth values (with a mean value of 3.8 obtained from the optimal estimation versus 0.9 from the analysis). We note that the effective radii in the analysis are not directly derived from the ECMWF model but obtained through the Wyser parameterization. In contrast, the retrieval's optimal state is determined pixel by pixel.

Based on the geographic distribution of cloud features and the distinct structural patterns identified in the analysis, the IASI pixels from the case study are grouped into three classes, as shown in Fig. 13. The first class (red dots) consists of optically thin continental ice clouds with low optical depth and small effective radii; the second (green dots) includes optically thicker ice clouds occurring over the ocean; and the third (blue dots) corresponds to higher-altitude, optically thinner ice clouds observed offshore. Figs. 14 and 15 show the mean retrieved ice content and effective radius profiles for each class of clouds (right panels), compared to the vertical profiles from the collocated analysis (left panels). Moreover in both Figures the cloud boundaries (top and bottom pressures), are highlighted as a shaded area spanning from the 2.5th to the 97.5th percentile of the total cloud mass content (g/m^3).

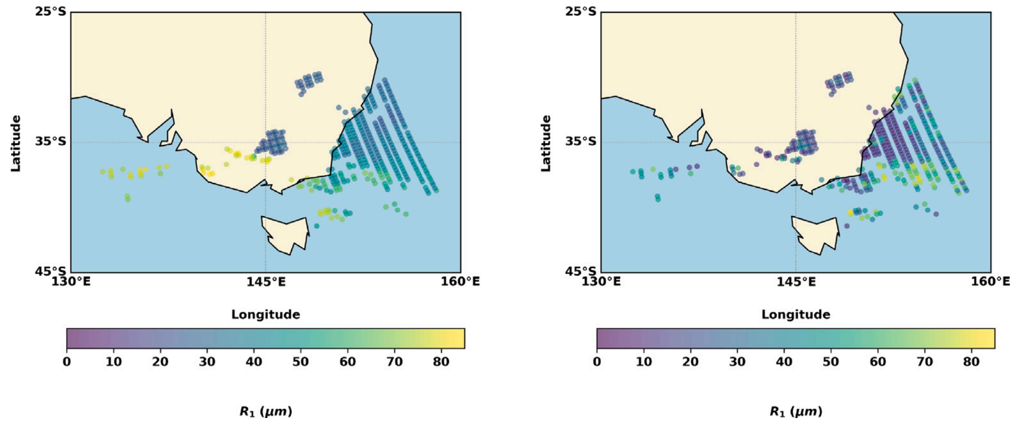

Fig. 12. Left panel: effective radius r_1 computed from ECMWF analysis ice-cloud profiles collocated at the IASI footprint. Right panel: effective radius r_1 computed from the retrieved ice-cloud profiles at the same footprints.

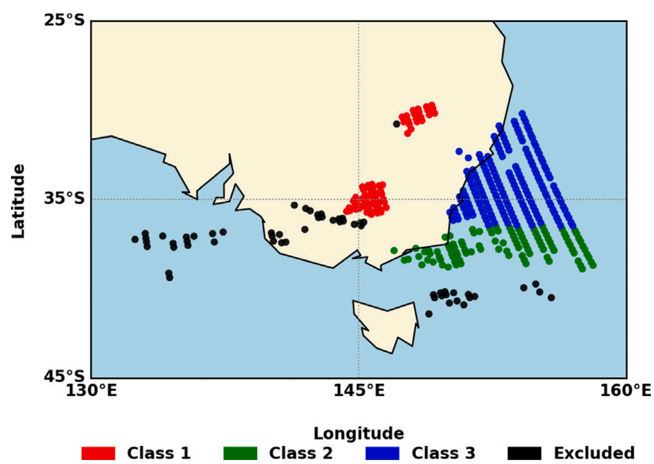

Fig. 13. Spatial classification of IASI pixels in the case study. Pixels are colored to indicate their assignment to one of three classes, which are defined by distinct optical depth structures. Black dots represent pixels excluded from the analysis due to either retrieval non-convergence or an ambiguous classification.

Elements belonging to the same class show similar vertical structures in both mass content and effective radius. Similar patterns are observed in the analysis profiles as well. Specifically, for cloud pixels classified as Class 1, the retrieval algorithm converged to an optically thin cloud, consistent with cirrus cloud formation. These clouds exhibit lower effective radius values and a cloud top near 210 hPa in both the analysis and retrieval profiles. However, the retrieved mass content is lower than in the analysis. Classes 2 and 3, by contrast, have higher retrieved optical depths and larger effective radii. Their vertical profiles are similar to those derived from the co-located analysis, with Class 3 showing a higher cloud top than Class 2. The main difference lies in the mass content, where the retrieved clouds show higher values than the analysis.

Overall, the retrieval is capable of converging to spatial structures that closely match the position and microphysical properties found in the analysis, even though it starts with constant cloud prior information and an initial guess independent of the analysis. The main differences appear in the mass content, where the retrieval tends to show, on average, higher ice cloud optical depths compared to the analysis. This increase compensates for the overestimation of upwelling radiance, as shown in Fig. 10.

5. Conclusions

This study demonstrates a robust pathway for achieving reliable cloud retrieval from high-resolution infrared observations under cloudy-sky conditions. The proposed optimal estimation framework successfully addresses the inherent ill-posed nature of cloud parameter inversion by integrating three key components: the fast radiative transfer model σ, principal component analysis-based state compression, and ad-hoc physically informed prior constraints. A crucial aspect of the forward model is the implementation of an extended Chou adjustment scheme (following the solution proposed by Tang [22]) within the σ routine. This scheme effectively extends the Chou scaling to accurately compute radiances across the mid- and far-infrared spectral range (100-3000 cm^{-1}), while maintaining an inversion time of approximately 1 min per spectrum on a standard laptop. The implementation provides both pseudo-monochromatic radiances and analytical Jacobians for atmospheric state variables and cloud microphysical properties, enabling efficient iterative convergence within the optimal estimation framework.

The dimensionality reduction strategy through PCA compression improves retrieval stability by reducing the condition number of matrices involved in the inversion. Additionally, by constraining the solution space to physically realistic atmospheric states, this approach mitigates the curse of dimensionality. Finally, the construction of ad-hoc cloud prior covariance matrices generates physically plausible Gaussian-like vertical profiles while regularizing the inversion without over-constraining the solution space.

These methodologies have been validated using synthetic IASI acquisitions from the NWP SAF database, where the algorithm achieved a convergence rate of approximately 97%, demonstrating its robustness across diverse atmospheric conditions. For retrieved cloud parameters, the framework demonstrated excellent performance. Pearson correlation coefficients for total optical depth, effective radius, and cloud position show a strong linear relationship with the true state and a clear improvement over the prior, indicating successful information extraction from the measurements. Application to real IASI acquisitions validated the retrieval scheme under operational conditions. Despite starting from constant climatological priors, independent from the analyses, the algorithm successfully converged to coherent spatial cloud structures qualitatively consistent with ECMWF analyses, demonstrating sensitivity to both optically thin cirrus over land and thicker clouds over the ocean. Retrieved vertical profiles exhibited physically consistent structures across different cloud regimes, from optically thin continental cirrus to thicker oceanic ice clouds, with spectral residuals remaining within measurement noise across most channels.

Future work should focus on extending this framework to mixed-phase and liquid-water clouds, validating it across additional hyperspectral infrared sounders (e.g., CrIS, IASI-NG, PREFIRE), and exploring

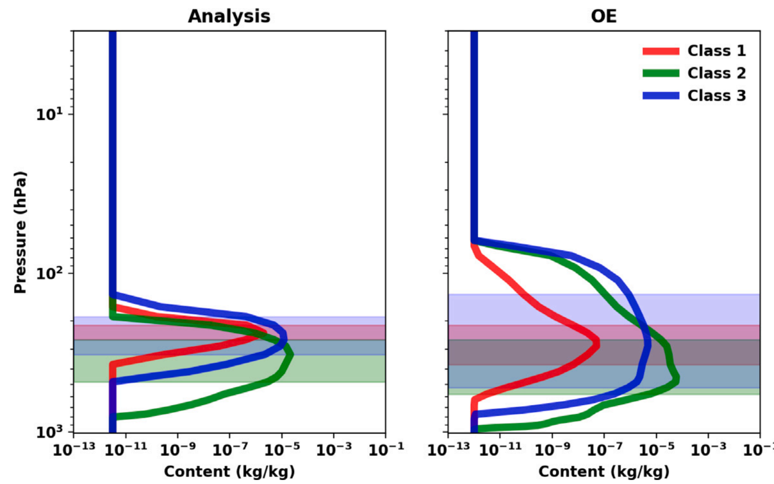

Fig. 14. Average vertical profiles of mass content (kg/kg) are shown for the IASI acquisitions of the three classes. The left panel shows profiles derived from the co-located analysis, while the right panel presents the retrieved profiles. The shaded areas highlight the regions containing 95% of the cloud mass.

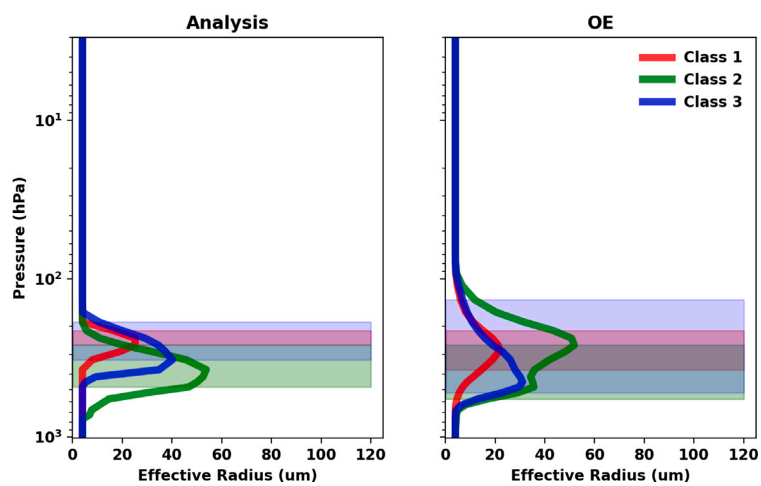

Fig. 15. Similar to Fig. 14, average vertical profiles of effective radius (μm) are presented for the same acquisitions. The shaded areas highlight the regions containing 95% of the cloud mass.

alternative cloud microphysical parameterizations. In particular, the model is well suited for integration into a processing chain that includes scene classification and can exploit additional information, used as prior knowledge or first-guess inputs, on the cloud layers present within the scene. For these reasons, the retrieval framework is appropriate for applications related to the upcoming ESA FORUM mission, which aims to deliver enhanced measurements of Earth's outgoing longwave spectral radiation in the far-infrared, a band which is particularly sensitive to cold ice clouds. Finally, the flexibility of the methodology allows for straightforward adaptation to new instruments and spectral configurations without requiring fundamental algorithmic changes, representing a significant advantage over purely data-driven approaches.

CRediT authorship contribution statement

Michele Martinazzo: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Tiziano Maestri:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis. **Elisa Fabbri:** Writing – review & editing, Software, Formal analysis. **Giuliano Liuzzi:** Writing – review & editing, Software, Data curation. **Guido Masiello:** Writing – review & editing, Software, Data curation. **Carmine Serio:** Writing – review & editing, Supervision, Software, Data curation.

Acknowledgments

The authors would like to thank the Italian Space Agency for the financial support through the project "FIT-FORUM", and the manager of the project Dr. S. Venafra (ASI).

Abbreviations

AIRS	Advanced Inertial Reference Sphere
BT	Brightness Temperature
CrIS	Cross-track Infrared Sounder
DISORT	Discrete Ordinate Radiative Transfer
ESA	European Space Agency
FIR	Far InfraRed
FORUM	Far-infrared Outgoing Radiation Understanding and Monitoring
IASI	Infrared Atmospheric Sounder Interferometer
MBE	Mean Bias Error
OD	Optical Depth
PC	Principal Component
PCA	Principal Component Analysis
PREFIRE	Polar Radiant Energy in the Far-InfraRed Experiment
RMSE	Root Mean Square Error
RT	Radiative Transfer
SVD	Singular Value Decomposition

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Michele Martinazzo reports was provided by University of Bologna Department of Physics and Astronomy 'Augusto Righi'. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Error analysis

Given a forward model $F(x)$, and a random error ϵ, the forward problem can be written as:

$$y = F(x) + \epsilon \tag{A.1}$$

And expanding at the first order the model F:

$$y \simeq F(x_a) + K(x - x_a) + \epsilon \tag{A.2}$$

Let us now assume that there exist a compression of the form

$$y \simeq F(x_a) + K(x - x_a) - \tilde{K}z + \tilde{K}z + \epsilon \tag{A.3}$$

and from the definition of $\tilde{K}$ and z

$$\begin{aligned} y &\simeq F(x_a) + \tilde{K}z + K(\mathbb{I} - VV^T)(x - x_a) + \epsilon \\ &= F(x_a) + \tilde{K}z + \epsilon_c + \epsilon \end{aligned} \tag{A.4}$$

where we define the *compression error* as $\epsilon_c = K(\mathbb{I} - VV^T)(x - x_a)$. This form is similar to the model parameter error as described by Rodgers [23]. This allows us to determine the optimal number of principal components (PC) to use in the compression process. Specifically, we can consider the minimum number of PC such that:

$$\epsilon_c \ll \epsilon \tag{A.5}$$

In order to obtain a correct error characterization in the presence of compression, we can start from Eq. (12), written here:

$$\begin{aligned} z_{i+1} =& z_i + (\gamma_i \Lambda_a^{-1} + \tilde{K}_i^T S_\epsilon^{-1} \tilde{K}_i + \beta_i \mathbb{I})^{-1} \\ & (\tilde{K}_i^T S_\epsilon^{-1} [y - \tilde{F}(z_i)] - \gamma_i \Lambda_a^{-1} z_i) \end{aligned} \tag{A.6}$$

and, following the approach proposed by Rodgers [23], we can study the behavior of this form when $i \to \infty$ and $z_i \to \hat{z}$.

$$\begin{aligned} \hat{z} =& \hat{z} + (\gamma \Lambda_a^{-1} + \tilde{K}^T S_\epsilon^{-1} \tilde{K} + \beta \mathbb{I})^{-1} \\ & (\tilde{K}^T S_\epsilon^{-1} [y - \tilde{F}(\hat{z})] - \gamma \Lambda_a^{-1} \hat{z}) \end{aligned} \tag{A.7}$$

from which

$$0 = \tilde{K}^T S_\epsilon^{-1} [y - \tilde{F}(\hat{z})] - \gamma \Lambda_a^{-1} \hat{z} \tag{A.8}$$

Now, given the expression (A.2), we can write:

$$0 = \tilde{K}^T S_\epsilon^{-1} [K(x - x_a) - \tilde{K}\hat{z} + \epsilon] - \gamma \Lambda_a^{-1} \hat{z} \tag{A.9}$$

and, after some algebra:

$$\hat{z} = [\tilde{K}^T S_\epsilon^{-1} \tilde{K} + \gamma \Lambda_a^{-1}]^{-1} \tilde{K}^T S_\epsilon^{-1} [K(x - x_a) + \epsilon] \tag{A.10}$$

Now, applying Eq. (13), and adding and subtracting the true state x, we obtain the following expression:

$$\hat{x} - x = V[\tilde{K}^T S_\epsilon^{-1} \tilde{K} + \gamma \Lambda_a^{-1}]^{-1} \tilde{K}^T S_\epsilon^{-1} [K(x - x_a) + \epsilon] + x_a - x \tag{A.11}$$

$$\hat{x} - x = (\tilde{G}K - \mathbb{I})(x_a - x) + \tilde{G}\epsilon \tag{A.12}$$

with

$$\tilde{G} = V[\tilde{K}^T S_\epsilon^{-1} \tilde{K} + \gamma \Lambda_a^{-1}]^{-1} \tilde{K}^T S_\epsilon^{-1} \tag{A.13}$$

from which

$$\hat{S}_x = (\tilde{G}K - \mathbb{I}) S_a (\tilde{G}K - \mathbb{I})^T + \tilde{G} S_\epsilon \tilde{G}^T \tag{A.14}$$

Appendix B. Condition number in the latent space

The level of ill-posedness of a linear transformation K can be evaluated through the condition number κ, whose definition is reported here below [57]:

$$\kappa(K) = \sigma_{max}(K)/\sigma_{min}(K) \tag{B.1}$$

where σ_{max} and σ_{min} are, respectively, the maximum and minimum singular values of K. A large condition number, associated with the matrices involved in the linearized problem, typically reflects the underlying ill-posed nature of the original inverse problem [49].

Let us define 2 real rectangular matrices, $K_{m\times n}$ and $O_{n\times q}$, and assume that $m \geq n \geq q$. We then assume that O is composed of orthonormal columns. The goal is to prove that, given the definition of condition number in (B.1), we have $\kappa(KO) \leq \kappa(K) \cdot \kappa(O)$. Then, from the properties of orthonormal matrices, $\kappa(O) = 1$.

Given u an unitary vector and calling $\|\cdot\|$ is the L_2 norm, we can write:

$$\begin{aligned} \frac{\|Ou\|}{\|u\|} &= \|Ou\| = \left\|U\Sigma V^T u\right\| = \left\|\Sigma V^T u\right\| \leq \\ &\leq \sigma_{max}(O) \left\|V^T u\right\| = \sigma_{max}(O)\,\|u\| = \sigma_{max}(O) \end{aligned} \tag{B.2}$$

and similarly

$$\|Ou\| \geq \sigma_{min}(O) \tag{B.3}$$

where the SVD for O is used. It is possible to reach an analogous conclusion for the matrix K applied on the vector $w = Ou$:

$$\sigma_{min}(K)\,\|w\| \leq \|Kw\| \leq \sigma_{max}(K)\,\|w\| \tag{B.4}$$

It follows from the above that:

$$\|KOu\| = \|Kw\| \leq \sigma_{max}(K)\,\|w\| \leq \sigma_{max}(K)\sigma_{max}(O) \tag{B.5}$$

and

$$\|KOu\| = \|Kw\| \geq \sigma_{min}(K)\,\|w\| \geq \sigma_{min}(K)\sigma_{min}(O) \tag{B.6}$$

From Eq. (B.5) and (B.6) we can state:

$$\max \|KOu\| = \sigma_{max}(KO) \leq \sigma_{max}(K)\sigma_{max}(O) \tag{B.7}$$

and

$$\min \|KOu\| = \sigma_{min}(KO) \geq \sigma_{min}(K)\sigma_{min}(O) \tag{B.8}$$

From the definition of condition number given in (B.1), and giving the relations (B.7) and (B.8), it is easy to prove that $\kappa(KO) \leq \kappa(K)\cdot\kappa(O)$. Finally, since O is composed of orthogonal columns, we can state that $\kappa(O) = 1$, from which it follows:

$$\kappa(KO) \leq \kappa(K) \tag{B.9}$$

Data availability

Data will be made available on request.

References

[1] Liu X, Zhou DK, Larar AM, Smith WL, Schluessel P, Newman SM, Taylor JP, Wu W. Retrieval of atmospheric profiles and cloud properties from IASI spectra using super-channels. Atmos Chem Phys 2009;9(23):9121–42. http://dx.doi.org/10.5194/acp-9-9121-2009.

[2] Maahn M, Turner DD, Löhnert U, Posselt DJ, Ebell K, Mace GG, Comstock JM. Optimal estimation retrievals and their uncertainties: What every atmospheric scientist should know. Bull Am Meteorol Soc 2020;101(9):E1512 – E1523. http://dx.doi.org/10.1175/BAMS-D-19-0027.1, URL https://journals.ametsoc.org/view/journals/bams/101/9/bamsD190027.xml.

[3] Simeoni D, Singer C, Chalon G. Infrared atmospheric sounding interferometer. Acta Astronaut 1997;40(2):113–8. http://dx.doi.org/10.1016/S0094-5765(97)00098-2, Enlarging The Scope of Space Applications.

[4] Hilton F, Armante R, August T, Barnet C, Bouchard A, Camy-Peyret C, Capelle V, Clarisse L, Clerbaux C, Coheur P-F, Collard A, Crevoisier C, Dufour G, Edwards D, Faijan F, Fourrié N, Gambacorta A, Goldberg M, Guidard V, Hurtmans D, Illingworth S, Jacquinet-Husson N, Kerzenmacher T, Klaes D, Lavanant L, Masiello G, Matricardi M, McNally A, Newman S, Pavelin E, Payan S, Péquignot E, Peyridieu S, Phulpin T, Remedios J, Schlüssel P, Serio C, Strow L, Stubenrauch C, Taylor J, Tobin D, Wolf W, Zhou D. Hyperspectral earth observation from IASI: Five years of accomplishments. Bull Am Meteorol Soc 2012;93(3):347–70. http://dx.doi.org/10.1175/BAMS-D-11-00027.1.

[5] Aumann H, Chahine M, Gautier C, Goldberg M, Kalnay E, McMillin L, Revercomb H, Rosenkranz P, Smith W, Staelin D, Strow L, Susskind J. AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. IEEE Trans Geosci Remote Sens 2003;41(2):253–64. http://dx.doi.org/10.1109/TGRS.2002.808356.

[6] Zavyalov V, Esplin M, Scott D, Esplin B, Bingham G, Hoffman E, Lietzke C, Predina J, Frain R, Suwinski L, Han Y, Major C, Graham B, Phillips L. Noise performance of the CrIS instrument. J Geophys Res: Atmos 2013;118(23):13,108–20. http://dx.doi.org/10.1002/2013JD020457, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020457.

[7] L'Ecuyer T, Drouin BJ, Anheuser J, Grames M, Henderson DS, Huang X, Kahn BH, Kay JE, Lim BH, Mateling M, Merelli A, Miller NB, Padmanabhan S, Peterson C, Schlegel N, White ML, Xie Y. The polar radiant energy in the far infrared experiment. In: BAMS. 2021, http://dx.doi.org/10.1175/BAMS-D-20-0155.1.

[8] DeSouza-Machado S, Strow LL, Tangborn A, Huang X, Chen X, Liu X, Wu W, Yang Q. Single-footprint retrievals for AIRS using a fast Two Slab cloud-representation model and the SARTA all-sky infrared radiative transfer algorithm. Atmospheric Meas Tech 2018;11(1):529–50. http://dx.doi.org/10.5194/amt-11-529-2018, URL https://amt.copernicus.org/articles/11/529/2018/.

[9] Serio C, Masiello G, Liuzzi G, Cersosimo A, Maestri T, Martinazzo M, Masin F, Proietti Pelliccia G, Venafra S, Camy-Peyret C. Demonstration of a physical inversion scheme for all-sky, day-night IASI observations and application to the analysis of the onset of the Antarctica ozone hole: Assessment of retrievals and consistency of forward modeling. J Quant Spectrosc Radiat Transfer 2024;329:109211. http://dx.doi.org/10.1016/j.jqsrt.2024.109211.

[10] Sgattoni C, Sgheri L, Chung M. A data-driven approach for fast atmospheric radiative transfer inversion. Inverse Problems 2025;41(8):085006. http://dx.doi.org/10.1088/1361-6420/adf3dd.

[11] Yan Y, Guo L, Li J, Yu Z, Sun S, Xu T, Zhao H, Guo L. Hybrid GRU-random forest model for accurate atmospheric duct detection with incomplete sounding data. Remote Sens 2024;16(22). http://dx.doi.org/10.3390/rs16224308.

[12] Guo X, Wu J, Zhang J, Han J. Deep learning for solving inversion problem of atmospheric refractivity estimation. Sustain Cities Soc 2018;43:524–31. http://dx.doi.org/10.1016/j.scs.2018.09.009.

[13] Malmgren-Hansen D, Laparra V, Aasbjerg Nielsen A, Camps-Valls G. Statistical retrieval of atmospheric profiles with deep convolutional neural networks. ISPRS J Photogramm Remote Sens 2019;158:231–40. http://dx.doi.org/10.1016/j.isprsjprs.2019.10.002.

[14] Mastro P, Masiello G, Serio C, Cimini D, Ricciardelli E, Paola FD, Hultberg T, August T, Romano F. Combined IASI-NG and MWS observations for the retrieval of cloud liquid and ice water path: A deep learning artificial intelligence approach. IEEE J Sel Top Appl Earth Obs Remote Sens 2022;15:3313–22. http://dx.doi.org/10.1109/JSTARS.2022.3166992.

[15] Kingma DP, Welling M. An introduction to variational autoencoders. Found Trends Mach Learn 2019;12:307–92.

[16] Cobb AD, Himes MD, Soboczenski F, Zorzan S, O'Beirne MD, Baydin AG, Gal Y, Domagal-Goldman SD, Arney GN, Angerhausen D. An ensemble of Bayesian neural networks for exoplanetary atmospheric retrieval. Astron J 2019;158. URL https://api.semanticscholar.org/CorpusID:166228557.

[17] Matricardi M, Saunders R. Fast radiative transfer model for simulation of infrared atmospheric sounding interferometer radiances. Appl Opt 1999;38(27):5679–91. http://dx.doi.org/10.1364/AO.38.005679.

[18] Liu X, Yang Q, Li H, Jin Z, Wu W, Kizer S, Zhou KD, Yang P. Development of a fast and accurate PCRTM radiative transfer model in the solar spectral region. Appl Opt 2016;55(29):8236–47. http://dx.doi.org/10.1364/AO.55.008236.

[19] Saunders R, Hocking J, Turner E, Rayer P, Rundle D, Brunel P, Vidot J, Roquet P, Matricardi M, Geer A, Bormann N, Lupu C. An update on the RTTOV fast radiative transfer model (currently at version 12). Geosci Model Dev 2018;11(7):2717–37. http://dx.doi.org/10.5194/gmd-11-2717-2018.

[20] Martinazzo M, Magurno D, Cossich W, Serio C, Masiello G, Maestri T. Assessment of the accuracy of scaling methods for radiance simulations at far and mid infrared wavelengths. J Quant Spectrosc Radiat Transfer 2021;271. http://dx.doi.org/10.1016/j.jqsrt.2021.107739.

[21] Masiello G, Serio C, Maestri T, Martinazzo M, Masin F, Liuzzi G, Venafra S. The new σ-IASI code for all sky radiative transfer calculations in the spectral range 10 to 2760 cm-1: σ-IASI/F2N. J Quant Spectrosc Radiat Transfer 2024;312:108814. http://dx.doi.org/10.1016/j.jqsrt.2023.108814.

[22] Tang G, Yang P, Kattawar GW, Huang X, Mlawer EJ, Baum BA, King MD. Improvement of the simulation of cloud longwave scattering in broadband radiative transfer models. J Atmos Sci 2018;75(7):2217–33. http://dx.doi.org/10.1175/JAS-D-18-0014.1.

[23] Rodgers CD. Inverse methods for atmospheric sounding. World Scientific; 2000, http://dx.doi.org/10.1142/3171, arXiv:https://www.worldscientific.com/doi/pdf/10.1142/3171.

[24] Shlens J. A tutorial on principal component analysis. 2014, arXiv:1404.1100. URL https://arxiv.org/abs/1404.1100.

[25] Masiello G, Serio C. Dimensionality-reduction approach to the thermal radiative transfer equation inverse problem. Geophys Res Lett 2004;31(11). http://dx.doi.org/10.1029/2004GL019845, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2004GL019845.

[26] Masiello G, Serio C, Antonelli P. Inversion for atmospheric thermodynamical parameters of IASI data in the principal components space. Q J R Meteorol Soc 2012;138(662):103–17. http://dx.doi.org/10.1002/qj.909, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.909.

[27] Ramanathan AK, Nguyen HM, Sun X, Mao J, Abshire JB, Hobbs JM, Braverman AJ. A singular value decomposition framework for retrievals with vertical distribution information from greenhouse gas column absorption spectroscopy measurements. Atmospheric Meas Tech 2018;11(8):4909–28. http://dx.doi.org/10.5194/amt-11-4909-2018.

[28] Fernandez-Martinez JL. Model reduction and uncertainty analysis in inverse problems. Lead Edge 2015;34(9):1006–16. http://dx.doi.org/10.1190/tle34091006.1.

[29] Liuzzi G, Masiello G, Serio C, Meloni D, Di Biagio C, Formenti P. Consistency of dimensional distributions and refractive indices of desert dust measured over Lampedusa with IASI radiances. Atmos Meas Tech 2017;10(2):599–615. http://dx.doi.org/10.5194/amt-10-599-2017.

[30] Chou M-D, Lee K-T, Tsay S-C, Fu Q. Parameterization for cloud longwave scattering for use in atmospheric models. J Climate 1999;12(1):159–69. http://dx.doi.org/10.1175/1520-0442(1999)012<0159:PFCLSF>2.0.CO;2.

[31] Turner DD, Lohnert U. Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based atmospheric emitted radiance interferometer (AERI). J Appl Meteorol Climatol 2014;53(3):752–71. http://dx.doi.org/10.1175/JAMC-D-13-0126.1.

[32] Carissimo A, De Feis I, Serio C. The physical retrieval methodology for IASI: the delta-IASI code. Environ Model Softw 2005;20(9):1111–26. http://dx.doi.org/10.1016/j.envsoft.2004.07.003.

[33] Rodgers CD. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. Rev Geophys 1976;14(4):609–24. http://dx.doi.org/10.1029/RG014i004p00609, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/RG014i004p00609.

[34] Masiello G, Serio C, Venafra S, Liuzzi G, Poutier L, Göttsche F-M. Physical retrieval of land surface emissivity spectra from hyper-spectral infrared observations and validation with in situ measurements. Remote Sens 2018;10(6). http://dx.doi.org/10.3390/rs10060976.

[35] Western LM, Rougier JC, Watson IM, Francis PN. Evaluating nonlinear maximum likelihood optimal estimation uncertainty in cloud and aerosol remote sensing. Atmos Sci Lett 2020;21(8):e980. http://dx.doi.org/10.1002/asl.980, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/asl.980.

[36] Hu Q, Liu C, Li Q, Liu T, Ji X, Zhu Y, Xing C, Liu H, Tan W, Gao M. Vertical profiles of the transport fluxes of aerosol and its precursors between Beijing and its southwest cities. Environ Pollut 2022;312:119988. http://dx.doi.org/10.1016/j.envpol.2022.119988.

[37] Robert CP. The bayesian choice: A decision-theoretic motivation. New York: Springer; 1994.

[38] Bjorkoy HB, Varagnolo D. Hyperparameters tuning in regularized system identification with nonzero prior means. In: 2024 European control conference. ECC, 2024, p. 3557–63. http://dx.doi.org/10.23919/ECC64448.2024.10590917.

[39] Veglio P, Maestri T. Statistics of vertical backscatter profiles of cirrus clouds. Atmos Chem Phys 2011;11(24):12925–43. http://dx.doi.org/10.5194/acp-11-12925-2011.

[40] Eresmaa R, McNally AP. Diverse profile datasets from the ECMWF 137-level short-range Forecasts. Tech. rep. NWP SAF report; 2014.

[41] Sun Z. Reply to comments by greg m. McFarquhar on 'Parametrization of effective sizes of cirrus-cloud particles and its verification against observations'. (October B, 1999, 125, 3037-3055). Q J R Meteorol Soc 2001;127(571):267–71. http://dx.doi.org/10.1002/qj.49712757116.

[42] Benaroya H, Han S, Nagurka M. Probability models in engineering and science. 2005, http://dx.doi.org/10.4324/9781003002314.

[43] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt S, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey C, Polat I, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy. SciPy 1.0-fundamental algorithms for scientific computing in Python. 2019, CoRR abs/1907.10121, arXiv:1907.10121.

[44] Fernandez-Martinez JL, Fernandez-Muniz Z. The curse of dimensionality in inverse problems. J Comput Appl Math 2020;369:112571. http://dx.doi.org/10.1016/j.cam.2019.112571.

[45] Tikhonov A. Solution of incorrectly formulated problems and the regularization method. Sov Math Dokl 1963;4:1035–8.

[46] Tikhonov A. Regularization of incorrectly posed problems. Sov Math Dokl 1963;4:1624–7.

[47] Hansen PC, O'Leary DP. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J Sci Comput 1993;14(6):1487–503. http://dx.doi.org/10.1137/0914086.

[48] Wahba G. Spline models for observational data. Society for Industrial and Applied Mathematics; 1990, http://dx.doi.org/10.1137/1.9781611970128, arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9781611970128.

[49] Doicu A, Trautmann T, Schreier F. In: Numerical regularization for atmospheric inverse problems, vol. 251–270, 2010, http://dx.doi.org/10.1007/978-3-642-05439-6.

[50] Blumstein D, Chalon G, Carlier T, Buil C, Hebert P, Maciaszek T, Ponce G, Phulpin T, Tournier B, Simeoni D, Astruc P, Clauss A, Kayal G, Jegou R. IASI instrument: Technical overview and measured performances. In: Strojnik M, editor. In: Infrared spaceborne remote sensing XII, vol. 5543, International Society for Optics and Photonics, SPIE; 2004, p. 196–207. http://dx.doi.org/10.1117/12.560907.

[51] Yang P, Bi L, Baum BA, Liou K-N, Kattawar GW, Mishchenko MI, Cole B. Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 μm. J Atmos Sci 2013;70(1):330–47. http://dx.doi.org/10.1175/JAS-D-12-039.1.

[52] Remedios JJ, Leigh RJ, Waterfall AM, Moore DP, Sembhi H, Parkes I, Greenhough J, Chipperfield MP, Hauglustaine D. MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical data sets. Atmos Chem Phys Discuss 2008;7(4):9973–10017.

[53] Huang X, Chen X, Zhou DK, Liu X. An observationally based global band-by-band surface emissivity dataset for climate and weather simulations. J Atmos Sci 2016;73(9):3541–55. http://dx.doi.org/10.1175/JAS-D-15-0355.1.

[54] Serio C, Masiello G, Mastro P, Tobin DC. Characterization of the observational covariance matrix of hyper-spectral infrared satellite sensors directly from measured earth views. Sensors 2020;20(5). http://dx.doi.org/10.3390/s20051492.

[55] Wyser K. The effective radius in ice clouds. J Clim 1998;11(7):1793–802. http://dx.doi.org/10.1175/1520-0442(1998)011<1793:TERIIC>2.0.CO;2.

[56] Liuzzi G, Masiello G, Serio C, Venafra S, Camy-Peyret C. Physical inversion of the full IASI spectra: Assessment of atmospheric parameters retrievals, consistency of spectroscopy and forward modelling. J Quant Spectrosc Radiat Transfer 2016;182:128–57. http://dx.doi.org/10.1016/j.jqsrt.2016.05.022.

[57] Belsley DA, Kuh E, Welsch RE. The Condition Number. New York: John Wiley & Sons; 1980, p. 100–4.