https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

Atmospheric Measurement Techniques Discussions
Open Access

Optimal variables for retrieval products

Simone Ceccherini

Istituto di Fisica Applicata "Nello Carrara" del Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy

Correspondence to: Simone Ceccherini (S.Ceccherini@ifac.cnr.it)

Abstract. The increase of satellite instruments sounding the atmosphere will make it very likely that several instruments simultaneously measure either the same vertical profile or vertical profiles related to nearby geo-locations, and users will consult fused products rather than individual measurements. Therefore, the retrieval products should be optimized for the use in data fusion operations, rather than for the representation of the profile. This change of paradigm raises the question if a more functional representation of the retrieval products exists. New variables for the retrieval products are proposed that have several advantages with respect to the standard retrieval products. These variables, in the linear approximation of the forward model, are independent of the a priori information used in the retrieval, allow to represent the profile with any a priori information and can directly be used to perform the data fusion of a set of measurements. Furthermore, the use of these variables allows to reduce to about one third the stored data volume with respect to the use of standard retrieval products.

1 Introduction

The measurement of the vertical profile of an atmospheric parameter requires the solution of an inverse problem that is often ill-posed (Doicu et al., 2010) and in order to obtain a stable solution, some a priori information has to be added in the retrieval process. A largely used method to retrieve atmospheric measurements is the *optimal estimation method* (Rodgers, 2000), where the a priori information is represented by an a priori profile and by an a priori covariance matrix (CM) and the solution is given by the profile corresponding to the maximum a posteriori probability calculated with the Bayes theorem (Sivia and Skilling, 2006).

Since in the last years the number of satellite instruments that are sounding the atmosphere is increasing at a high rate, it is very likely that more instruments will simultaneously measure either the same vertical profile or vertical profiles corresponding to nearby geo-locations. In this case, the different retrieved profiles can be combined into a single product that includes all the available information and we refer to this combination as *data fusion* (Ceccherini et al., 2015). Accordingly, the choice of the a priori information and of the vertical grid has to take into account the possibility that the result of the retrieval will be fused with other measurements (Ceccherini et al., 2010a, Ceccherini et al., 2010b, Warner et al., 2014, Cortesi et al., 2016, Schneider et al., 2022). The data fusion approach is alternative to that of the synergistic retrieval, in which all the available observations are simultaneously used in a single retrieval: for a detailed description and comparison of the two methods see Ridolfi et al., (2022) and references therein.

In the light of the increased requirement of fused products, we consider the possibility of using new variables representing the retrieval products, with the purpose of simplifying the subsequent fusion processes. The change of the retrieval products is proposed in the perspective of developing a shared formalism, which facilitates the interface between data providers and data users, while ensuring a full exploitation of the available information. The advantages of the new variables with respect to those presently used are analyzed on a theoretical basis.

When the result of the retrieval is used in subsequent data fusion operations, the vertical grid of the fusing products should be as fine as needed for the representation of the information content of the final fused product, rather than of the information content of the individual measurement, because, as shown in Ceccherini et al. (2016), in the latter case some information is lost. This can be easily done because the use of the a priori information allows to represent the profile on a

https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

Atmospheric Measurement Techniques Discussions

vertical grid as fine as desired. Therefore, the retrieval products are chosen no longer with the objective of providing to the user a useful representation of the observed profile, but rather as the best input for the fusion process, possibly independent of a priori information. Therefore, the question arises if by removing the objective of the visual representation a more functional data transfer of the retrieval products can be considered.

Generally, in order to make a complete use of the products in further processing such as data fusion or data assimilation, the retrieval products are represented by means of the retrieved profile, the averaging kernel matrix (AKM), the retrieval CM and the a priori information used in the retrieval.

We propose new variables to represent the retrieval products that have several advantages with respect to these standard quantities. The new variables decrease the stored data volume, in the linear approximation of the forward model are independent of the a priori information used in the retrieval, can be used to represent the profile with any a priori information and are quite suitable for subsequent data fusion operations.

In section 2, we recall useful notations and equations, linearize the transfer function and introduce the new variables. In section 3, we describe the advantages of the new variables with respect to the standard retrieval products concerning representation of the profile, data fusion and reduction of the data volume. Finally, in section 4 we draw the conclusions.

2 The new variables

2.1 Recall of notations and equations

We suppose to have retrieved the vertical profile $\hat{\boldsymbol{x}}$ of an atmospheric parameter from a set of observations (radiances) $\boldsymbol{y}$ with the optimal estimation method (Rodgers, 2000), using as a priori information a profile $\boldsymbol{x}_{\mathrm{a}}$ and a CM $\mathbf{S}_{\mathrm{a}}$. We indicate with $\boldsymbol{f}(\boldsymbol{x})$ the forward model, which allows to express the observations $\boldsymbol{y}$ as a function of the true profile $\boldsymbol{x}_{\mathrm{t}}$ by the following equation:

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}_{\mathrm{t}}) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{\varepsilon}$ is the vector including the noise errors of the observations with CM $\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T \rangle$ given by $\mathbf{S}_{\mathrm{ny}}$, where $\langle \ldots \rangle$ indicates the expected value.

The sensitivity of $\hat{\boldsymbol{x}}$ to the true profile $\boldsymbol{x}_{\mathrm{t}}$ is described by the AKM $\mathbf{A} = \frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{x}_{\mathrm{t}}}$ and the retrieval errors of $\hat{\boldsymbol{x}}$ are described by the CM $\mathbf{S}$, which is the sum of the CM of the noise errors $\mathbf{S}_{\mathrm{n}}$ and of the CM of the smoothing errors $\mathbf{S}_{\mathrm{s}}$. The AKM and the CMs are given by (see Eqs. (3.28-31) in Rodgers, 2000):

$$\mathbf{A} = \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1} \mathbf{F}, \tag{2}$$

$$\mathbf{S}_{\mathrm{n}} = \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1} \mathbf{F} \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1}, \tag{3}$$

$$\mathbf{S}_{\mathrm{s}} = \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1} \mathbf{S}_{\mathrm{a}}^{-1} \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1}, \tag{4}$$

$$\mathbf{S} = \mathbf{S}_{\mathrm{n}} + \mathbf{S}_{\mathrm{s}} = \left(\mathbf{F} + \mathbf{S}_{\mathrm{a}}^{-1}\right)^{-1}, \tag{5}$$

where

$$\mathbf{F} = \mathbf{K}^T \mathbf{S}_{\mathrm{ny}}^{-1} \mathbf{K}, \tag{6}$$

https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

Atmospheric Measurement Techniques Discussions

with $\mathbf{K}$ being the Jacobian of the forward model $\boldsymbol{f}(\boldsymbol{x})$ calculated in $\hat{\boldsymbol{x}}$: $\mathbf{K} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}}$. The matrix $\mathbf{F}$ is the Fisher information matrix (Rodgers, 2000, Ceccherini et al., 2012), defined as

$$\mathbf{F} = \int P(\boldsymbol{y}\,|\,\boldsymbol{x}) \left(\frac{\partial \ln P(\boldsymbol{y}\,|\,\boldsymbol{x})}{\partial \boldsymbol{x}} \right) \left(\frac{\partial \ln P(\boldsymbol{y}\,|\,\boldsymbol{x})}{\partial \boldsymbol{x}} \right)^T d\boldsymbol{y}, \tag{7}$$

where $P(\boldsymbol{y}\,|\,\boldsymbol{x})$ is the conditional probability distribution to obtain $\boldsymbol{y}$ given $\boldsymbol{x}$, which considered as a function of $\boldsymbol{x}$ is referred to as *likelihood function* $L(\boldsymbol{x})$ (Fisher, 1921). In the case that the inverse problem can be solved without constrain ($\mathbf{S}_{\mathrm{a}}^{-1} = \mathbf{0}$), that is when we can find the solution of maximum likelihood, from Eq. (3-5) we see that $\mathbf{F}$ is equal to the inverse matrix of the CM of the retrieval errors, which coincides with the CM of the noise errors. From this consideration, we can understand that the physical meaning of $\mathbf{F}$ is that of quantifying the information provided by the observations $\boldsymbol{y}$ about the retrieved vertical profile.

The dependence of $\mathbf{F}$ on the a priori information used in the retrieval derives from how $\mathbf{K}$ depends on the profile $\boldsymbol{x}$ where it is calculated, because we calculate it in $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}$ depends on $\boldsymbol{x}_{\mathrm{a}}$ and $\mathbf{S}_{\mathrm{a}}$. Therefore, the dependence of $\mathbf{F}$ on the a priori information is due to the second order terms in the expansion of the forward model as a function of the profile $\boldsymbol{x}$ and consequently, when the linear approximation of the forward model is valid, $\mathbf{F}$ is independent of the a priori information.

2.2 Linearization of the transfer function and variables $\boldsymbol{\alpha}$

We can consider the whole measuring system, including both the observing system and the retrieval method, as an operation that transforms the true profile $\boldsymbol{x}_{\mathrm{t}}$ into the retrieved profile $\hat{\boldsymbol{x}}$ and, accordingly, define the retrieved profile $\hat{\boldsymbol{x}}$ as a function of the true profile $\boldsymbol{x}_{\mathrm{t}}$. This function is referred to as *transfer function* (Rodgers, 2000) and besides being a function of $\boldsymbol{x}_{\mathrm{t}}$ is also a function of the noise errors $\boldsymbol{\varepsilon}$ of the observations $\boldsymbol{y}$. This dependence can be seen recalling that really $\hat{\boldsymbol{x}}$ depends on $\boldsymbol{x}_{\mathrm{t}}$ through the observations $\boldsymbol{y}$, therefore, using Eq. (1) we can write $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}(\boldsymbol{y}) = \hat{\boldsymbol{x}}(\boldsymbol{f}(\boldsymbol{x}_{\mathrm{t}}) + \boldsymbol{\varepsilon})$, which we indicate as $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})$.

Expanding the transfer function at the first order around the a priori profile $\boldsymbol{x}_{\mathrm{t}} = \boldsymbol{x}_{\mathrm{a}}$ and zero errors $\boldsymbol{\varepsilon} = \boldsymbol{0}$, we obtain:

$$\hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon}) \simeq \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{0}) + \left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{x}_{\mathrm{t}}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\boldsymbol{x}_{\mathrm{a}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} (\boldsymbol{x}_{\mathrm{t}} - \boldsymbol{x}_{\mathrm{a}}) + \left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\boldsymbol{x}_{\mathrm{a}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} \boldsymbol{\varepsilon}. \tag{8}$$

Concerning the first term of the expansion, we recall that the retrieved profile obtained with the optimal estimation method in absence of errors is a weighted mean between the true profile and the a priori profile. Therefore, when the true profile coincides with the a priori profile the retrieved profile in absence of errors results in the a priori profile, that is $\hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{a}}, \boldsymbol{0}) = \boldsymbol{x}_{\mathrm{a}}$. This result is peculiar of the optimal estimation method and if we wish to extend the results of this article to retrieval methods different from the optimal estimation it is necessary to identify a linearization point for which we know the value assumed by the transfer function. This consideration also applies to the complete data fusion method (Ceccherini et al, 2015 and Ceccherini et al, 2022) and to all the methods that are based on the expansion of the transfer function. Under the approximation that the derivatives do not significantly depend on the point where they are calculated, we have $\left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{x}_{\mathrm{t}}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\boldsymbol{x}_{\mathrm{a}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} \simeq \left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{x}_{\mathrm{t}}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\hat{\boldsymbol{x}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} = \mathbf{A}$ and $\left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\boldsymbol{x}_{\mathrm{a}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} \simeq \left.\frac{\partial \hat{\boldsymbol{x}}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{\varepsilon})}{\partial \boldsymbol{y}}\right|_{\substack{\boldsymbol{x}_{\mathrm{t}}=\hat{\boldsymbol{x}} \\ \boldsymbol{\varepsilon}=\boldsymbol{0}}} = \mathbf{G}$, where $\mathbf{G}$ is the gain matrix and is given by

https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

Atmospheric Measurement Techniques Discussions

$$\mathbf{G} = \left(\mathbf{K}^T\mathbf{S}_{ny}^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{K}^T\mathbf{S}_{ny}^{-1} = \left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{K}^T\mathbf{S}_{ny}^{-1}. \quad (9)$$

On the basis of these considerations Eq. (8) becomes

$$\hat{\boldsymbol{x}} = \boldsymbol{x}_a + \mathbf{A}\left(\boldsymbol{x}_t - \boldsymbol{x}_a\right) + \mathbf{G}\boldsymbol{\varepsilon}. \quad (10)$$

Following the approach described in the complete data fusion method (Ceccherini et al, 2015 and Ceccherini et al, 2022), we define the vector $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \hat{\boldsymbol{x}} - \boldsymbol{x}_a + \mathbf{A}\boldsymbol{x}_a, \quad (11)$$

that can be calculated knowing the retrieved profile, the a priori profile and the AKM. Substituting $\hat{\boldsymbol{x}}$ from Eq. (10) into Eq. (11), we see that $\boldsymbol{\alpha}$ is equal to

$$\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{x}_t + \mathbf{G}\boldsymbol{\varepsilon} \quad (12)$$

and provides a measurement of the true profile made using the rows of $\mathbf{A}$ as weighting functions. Eq. (12), together with Eqs. (2) and (9), shows that $\boldsymbol{\alpha}$ (differently from $\hat{\boldsymbol{x}}$), in the linear approximation of the forward model, is independent of the a priori profile $\boldsymbol{x}_a$, however, through the expressions of $\mathbf{A}$ and $\mathbf{G}$, maintains the dependence on the a priori CM $\mathbf{S}_a$.

2.3 The new variables $\boldsymbol{\beta}$

We define the vector $\boldsymbol{\beta}$ as

$$\boldsymbol{\beta} = \mathbf{S}^{-1}\boldsymbol{\alpha} = \mathbf{S}^{-1}\left(\hat{\mathbf{x}} - \mathbf{x}_a + \mathbf{A}\mathbf{x}_a\right) \quad (13)$$

and using Eqs. (2, 5, 9) and (12) we obtain

$$\boldsymbol{\beta} = \mathbf{F}\mathbf{x}_t + \boldsymbol{\delta}, \quad (14)$$

where the vector $\boldsymbol{\delta}$ is given by

$$\boldsymbol{\delta} = \mathbf{K}^T\mathbf{S}_{ny}^{-1}\boldsymbol{\varepsilon}. \quad (15)$$

Eq. (14) provides the physical meaning of $\boldsymbol{\beta}$, that is the measurement of the true profile in which the weighting functions are the rows of $\mathbf{F}$ and $\boldsymbol{\delta}$ is the vector that includes the errors of this measurement. Furthermore, from Eqs. (14) and (15) we see that $\boldsymbol{\beta}$, in the linear approximation of the forward model, is uniquely determined independently of both $\boldsymbol{x}_a$ and $\mathbf{S}_a$.

Using Eq. (14), we calculate the sensitivity of $\boldsymbol{\beta}$ to the true profile, that is the AKM of $\boldsymbol{\beta}$

$$\mathbf{A}_{\beta} = \frac{\partial\boldsymbol{\beta}}{\partial\mathbf{x}_t} = \mathbf{F} \quad (16)$$

and from Eqs. (6, 14) and (15) we calculate the CM of $\boldsymbol{\beta}$

$$\mathbf{S}_{\beta} = \left\langle\left(\boldsymbol{\beta} - \langle\boldsymbol{\beta}\rangle\right)\left(\boldsymbol{\beta} - \langle\boldsymbol{\beta}\rangle\right)^T\right\rangle = \left\langle\boldsymbol{\delta}\boldsymbol{\delta}^T\right\rangle = \mathbf{K}^T\mathbf{S}_{ny}^{-1}\left\langle\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right\rangle\mathbf{S}_{ny}^{-1}\mathbf{K} = \mathbf{F}. \quad (17)$$

Therefore, both the AKM and the CM of $\boldsymbol{\beta}$ coincide with the Fisher information matrix $\mathbf{F}$.

From Eq. (13) we see that the dimensions of $\boldsymbol{\beta}$ are the inverse of the dimensions of $\hat{x}$: $[\boldsymbol{\beta}] = [\hat{x}]^{-1}$, therefore, $\boldsymbol{\beta}$ does not represent a profile of the parameter that we aim to retrieve. However, as we noticed in the introduction, this is not a problem, because the objective of the retrieval products is no longer the visual representation of the profile, but to efficiently provide all the information of the observations to subsequent data analyses.

https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

3 Advantages of the use of the variables $\boldsymbol{\beta}$

3.1 Representation of the profile using any constraint

Using Eq. (2, 5) and (13), we can obtain $\boldsymbol{\beta}$ from the retrieved profile $\hat{\boldsymbol{x}}$

$$\begin{aligned}\boldsymbol{\beta} &= \left(\mathbf{F}+\mathbf{S}_a^{-1}\right)\left[\hat{\boldsymbol{x}}-\boldsymbol{x}_a+\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)^{-1}\mathbf{F}\boldsymbol{x}_a\right]= \\ &=\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)\left[\hat{\boldsymbol{x}}-\mathbf{x}_a+\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)^{-1}\left(\mathbf{F}+\mathbf{S}_a^{-1}-\mathbf{S}_a^{-1}\right)\boldsymbol{x}_a\right]=\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)\left[\hat{\boldsymbol{x}}-\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)^{-1}\mathbf{S}_a^{-1}\mathbf{x}_a\right].\end{aligned} \tag{18}$$

120 and, multiplying on the left both sides of this equation by $\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)^{-1}$, we can derive $\hat{\boldsymbol{x}}$ from $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{x}}=\left(\mathbf{F}+\mathbf{S}_a^{-1}\right)^{-1}\left(\boldsymbol{\beta}+\mathbf{S}_a^{-1}\mathbf{x}_a\right). \tag{19}$$

Eq. (19) can be used to recover the original retrieved profile using the a priori information $\boldsymbol{x}_a$ and $\mathbf{S}_a$ used in the retrieval procedure, but since in the linear approximation of the forward model $\mathbf{F}$ and $\boldsymbol{\beta}$ are independent of the a priori information, in this approximation, Eq. (19) can be used to produce a profile with any a priori information we like.

3.2 Data fusion

125 If we suppose to have *N* independent measurements $\hat{\boldsymbol{x}}_i$ of the same vertical profile $\boldsymbol{x}_t$, obtained with the optimal estimation method and characterized by the AKMs $\mathbf{A}_i$ and CMs of the retrieval errors $\mathbf{S}_i$, we can combine these measurements in a single vertical profile that includes the information of all the *N* measurements using the complete data fusion formula (Ceccherini et al., 2022)

$$\mathbf{x}_f=\left(\sum_{i=1}^{N}\mathbf{S}_i^{-1}\mathbf{A}_i+\mathbf{S}_a^{-1}\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{S}_i^{-1}\boldsymbol{\alpha}_i+\mathbf{S}_a^{-1}\mathbf{x}_a\right), \tag{20}$$

where $\boldsymbol{x}_a$ and $\mathbf{S}_a$ are the a priori profile and CM used to constrain the fused profile $\mathbf{x}_f$ and $\boldsymbol{\alpha}_i$ are the vectors defined by

130 Eq. (11) for each measurement:

$$\boldsymbol{\alpha}_i=\hat{\boldsymbol{x}}_i-\boldsymbol{x}_{ai}+\mathbf{A}_i\boldsymbol{x}_{ai}, \tag{21}$$

with $\boldsymbol{x}_{ai}$ the a priori profile used in the retrieval of the *i*-th measurement.

Using Eqs. (2, 5) and (13) we can re-write Eq. (20) as

$$\mathbf{x}_f=\left(\sum_{i=1}^{N}\mathbf{F}_i+\mathbf{S}_a^{-1}\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{\beta}_i+\mathbf{S}_a^{-1}\mathbf{x}_a\right), \tag{22}$$

where $\boldsymbol{\beta}_i$ and $\mathbf{F}_i$ are the $\boldsymbol{\beta}$ and $\mathbf{F}$ quantities related to each one of the *N* measurements. Eq. (22) shows that the vectors $\boldsymbol{\beta}_i$ and the Fisher information matrices $\mathbf{F}_i$ are the only quantities needed to perform the data fusion of a set of measurements.

135 3.2 Reduction of the data volume

In this subsection we compare the data volume required by the standard retrieval products with that required by the new variables $\boldsymbol{\beta}$. In the case of standard retrieval products the quantities that have to be stored to allow a complete use of the products, in further processing of the data such as data fusion or data assimilation, are: $\hat{\boldsymbol{x}}$, $\mathbf{A}$, $\mathbf{S}$ and $\boldsymbol{x}_a$. $\mathbf{S}_a$ is not necessary, because it can be obtained from $\mathbf{A}$ and $\mathbf{S}$ by means of

$$\mathbf{S}_a=\left(\mathbf{I}-\mathbf{A}\right)^{-1}\mathbf{S}, \tag{23}$$

140 which is derived using Eqs. (2) and (5).

https://doi.org/10.5194/amt-2023-42
Preprint. Discussion started: 21 April 2023

Atmospheric Measurement Techniques Discussions

If we suppose that the profile has n components, then $\mathbf{A}$ is composed by n^2 values and $\mathbf{S}$ by $n(n+1)/2$ independent values (because it is a symmetric matrix). Therefore, in the case of standard retrieval products we have to store $(3n^2+5n)/2$ values. In the case of the new variables $\boldsymbol{\beta}$, the quantities that have to be stored to allow a complete use of the products are $\boldsymbol{\beta}$ and $\mathbf{F}$ (which is a symmetric matrix), therefore, the values that have to be stored are $(n^2+3n)/2$. In case that we wish give a more complete information specifying where the Jacobian $\mathbf{K}$ is calculated, we can also give $\hat{\mathbf{x}}$, and the values that have to be stored are $(n^2+5n)/2$. Since the main storage requirement is due to the square term, the use of the variables $\boldsymbol{\beta}$ allows to reduce to about one third the stored data volume with respect to the use of the standard retrieval products.

4 Conclusions

With the increasing use of the atmospheric profiles retrieved from atmospheric satellite observations in data fusion operations, the requirement that these products provide a representation of the observed quantity is less important and other features, such as completeness and compactness of the information, become more relevant. In this perspective, new retrieval variables have been proposed. These variables, referred to as $\boldsymbol{\beta}$, are the measurement of the true profile obtained using the rows of the Fisher information matrix as weighting functions. This measurement does not provide a representation of the profile, but has several useful properties: in the linear approximation of the forward model is independent of the a priori information used in the retrieval and both the AKM and the CM of $\boldsymbol{\beta}$ coincide with the Fisher information matrix. Furthermore, the variables $\boldsymbol{\beta}$ can be used to obtain the representation of the vertical profile with an a priori information selected by the user and they can be directly used to perform the data fusion of a set of measurements. For the exploitation of these products in the subsequent operations it is sufficient to provide $\boldsymbol{\beta}$ and the Fisher information matrix $\mathbf{F}$, that fully characterizes the measurement being both its AKM and its CM. Accordingly, the use of the variables $\boldsymbol{\beta}$ allows to reduce to about one third the stored data volume with respect to the use of the standard products. These properties of the variables $\boldsymbol{\beta}$ make them a perfect retrieval product when further processing is performed by the users. The communities of data providers and data users are invited to test and validate the efficiency of this new interface.

Data availability. No data sets were used in this article.

Competing interests. The author has declared that there are no competing interests.

Acknowledgments. The author is grateful to Bruno Carli for the useful discussions.

Financial support. The results reported in the article were obtained in the context of the EMM (Earth-Moon-Mars) project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 – "Education and Research", Component 2: from research to business, Investment 3.1: "Fund for the realisation of an integrated system of research and innovation infrastructures". Action 3.1.1 "Creation of new research infrastructures strengthening of existing ones and their networking for Scientific Excellence under Horizon Europe" funded by the European Union - NextGenerationEU.

References

Ceccherini, S., Carli, B., Cortesi, U., Del Bianco, S. and Raspollini, P.: Retrieval of the vertical column of an atmospheric constituent from data fusion of remote sensing measurements, J. Quant. Spectrosc. Radiat., 111, 507-514, 2010a.

Ceccherini, S., Cortesi, U., Del Bianco, S., Raspollini P., and Carli B.: IASI-METOP and MIPAS-ENVISAT data fusion, Atmos. Chem. Phys., 10, 4689-4698, 2010b.

Ceccherini, S., Carli, B., Raspollini, P.: Quality quantifier of indirect measurements, Opt. Express, 20, 5151-5167, 2012.

Ceccherini, S., Carli, B., and Raspollini, P.: Equivalence of data fusion and simultaneous retrieval, Opt. Express, 23, 8476-8488, 2015.

Ceccherini, S., Carli, B., and Raspollini, P.: Vertical grid of retrieved atmospheric profiles, J. Quant. Spectr. Rad. Trans., 174, 7-13, 2016.

Ceccherini, S., Zoppetti, N., and Carli, B.: An improved formula for the complete data fusion, Atmos. Meas. Tech.,15, 7039-7048, 2022

Cortesi, U., Del Bianco, S., Ceccherini, S., Gai, M., Dinelli, B. M., Castelli, E., Oelhaf, H., Woiwode, W., Höpfner, M., and Gerber, D.: Synergy between middle infrared and millimeter-wave limb sounding of atmospheric temperature and minor constituents, Atmos. Meas. Tech., 9, 2267–2289, https://doi.org/10.5194/amt-9- 2267-2016, 2016.

Doicu, A., Trautmann, T., and Schreier, F.: Numerical Regularization for Atmospheric Inverse Problems, Berlin, Springer Verlag, 2010.

Fisher, R. A.: On the mathematical foundation of theoretical statistics, Phil. Trans. R. Soc. Lond., A222, 309, (1921).

Ridolfi, M., Tirelli, C., Ceccherini, S., Belotti, C., Cortesi, U., and Palchetti, L.: Synergistic retrieval and complete data fusion methods applied to simulated FORUM and IASI-NG measurements, Atmos. Meas. Tech., 15, 6723–6737, https://doi.org/10.5194/amt-15-6723-2022, 2022.

Rodgers, C.D., Inverse Methods for Atmospheric Sounding: Theory and Practice, Vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, World Scientific, Singapore, 2000.

Schneider, M., Ertl, B., Tu, Q., Diekmann, C. J., Khosrawi, F., Röhling, A. N., Hase, F., Dubravica, D., García, O. E., Sepúlveda, E., Borsdorff, T., Landgraf, J., Lorente, A., Butz, A., Chen, H., Kivi, R., Laemmel, T., Ramonet, M., Crevoisier, C., Pernin, J., Steinbacher, M., Meinhardt, F., Strong, K., Wunch, D., Warneke, T., Roehl, C., Wennberg, P. O., Morino, I., Iraci, L. T., Shiomi, K., Deutscher, N. M., Griffith, D. W. T., Velazco, V. A., and Pollard, D. F.: Synergetic use of IASI profile and TROPOMI total-column level 2 methane retrieval products, Atmos. Meas. Tech., 15, 4339–4371, https://doi.org/10.5194/amt-15-4339-2022, 2022.

Sivia, D. S., and Skilling, J.: Data Analysis: A Bayesian Tutorial, Oxford University Press, 2006

Warner, J. X., Yang, R., Wei, Z., Carminati, F., Tangborn, A., Sun, Z., Lahoz, W., Attié, J.-L., El Amraoui, L., and Duncan, B.: Global carbon monoxide products from combined AIRS, TES and MLS measurements on A-train satellites, Atmos. Chem. Phys., 14, 103–114, https://doi.org/10.5194/acp-14-103-2014, 2014.

Generalization to the case when $\boldsymbol{\varepsilon}$ also includes the forward model errors

In general, $\boldsymbol{\varepsilon}$ can include both random and systematic errors, in this case it will be characterized by a bias $\langle \boldsymbol{\varepsilon} \rangle = \boldsymbol{b} \neq \mathbf{0}$ and by a CM $\left\langle (\boldsymbol{\varepsilon} - \boldsymbol{b})(\boldsymbol{\varepsilon} - \boldsymbol{b})^T \right\rangle = \mathbf{S}_{\boldsymbol{y}}$. The CM $\mathbf{S}_{\boldsymbol{y}}$ will be the sum of the CM $\mathbf{S}_{\mathrm{n}\boldsymbol{y}}$ of the noise errors plus the CM of the random components of the forward model errors. We can define new observations $\boldsymbol{y}'$ corrected for the bias:

$$\boldsymbol{y}' = \boldsymbol{y} - \boldsymbol{b} = \boldsymbol{f}\left(\boldsymbol{x}_{\mathrm{t}}\right) + \boldsymbol{\varepsilon}',$$

with

$$\boldsymbol{\varepsilon}' = \boldsymbol{\varepsilon} - \boldsymbol{b}.$$

The new errors $\boldsymbol{\varepsilon}'$ are characterized by $\langle \boldsymbol{\varepsilon}' \rangle = \mathbf{0}$ and by the CM $\left\langle \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon}'^T \right\rangle = \mathbf{S}_{\boldsymbol{y}}$.

Consequently, in the case of forward model errors, we can repeat the treatment described in the paper replacing $\boldsymbol{y}$ with $\boldsymbol{y}'$, $\boldsymbol{\varepsilon}$ with $\boldsymbol{\varepsilon}'$ and $\mathbf{S}_{\mathrm{n}\boldsymbol{y}}$ with $\mathbf{S}_{\boldsymbol{y}}$.