Communication

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; s.ceccherini@ifac.cnr.it

Abstract: The increase in satellite instruments sounding the atmosphere will increase the frequency of several instruments simultaneously measuring 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 use in data fusion operations, rather than for the representation of the profile. This change in paradigm raises the question of whether 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 us to represent the profile with any a priori information and can be used directly to perform the data fusion of a set of measurements. Furthermore, the use of these variables allows us to reduce the stored data to about one third of its volume with respect to the use of standard retrieval products.

Keywords: retrieval products; data fusion; atmospheric vertical profiles; remote sensing

1. Introduction

The retrieval of the vertical profile of an atmospheric parameter requires the solution of an inverse problem [1–3] that is often ill-posed [4], and in order to obtain a stable solution, some a priori information has to be added in the retrieval process. A commonly used method to retrieve atmospheric parameters using remote sensing is the *optimal estimation method* [1], where the a priori information is represented by an a priori profile and by an a priori covariance matrix (CM) of the unknown parameter, and the solution is given by the profile corresponding to the maximum a posteriori probability calculated with the Bayes theorem (see e.g., [5]).

Since, in the last few years, the number of satellite instruments that are sounding the atmosphere has increased 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 of the available information, and we refer to this combination as *data fusion* [6]. 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 [7–11]. The data fusion approach is alternative to that of the synergistic retrieval [12], in which all of the available observations are simultaneously used in a single retrieval; for a detailed description and comparison of the two methods see [13] and references therein.

In 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. A change in the retrieval products is proposed in view 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 currently 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 [14], in the latter case, some information is lost. This can be easily done because the use of the a priori information allows for representing the profile on a vertical grid as fine as desired. Therefore, the retrieval products are no longer chosen with the objective of providing the user with 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 of whether by removing the objective of the graphical representation of the profile a more functional data transfer of the retrieval products can be considered.

Generally, in order to make 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 calculated starting from these standard retrieval products that are a new way to save the information provided by the measurements and have several advantages with respect to the standard quantities. In the linear approximation of the forward model, the new variables are independent of the a priori information used in the retrieval and decrease the data volume requirement. Furthermore, they 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 assume to have retrieved the vertical profile $\hat{x}$ of an atmospheric parameter from a set of observations (radiances) y with the optimal estimation method [1], using a profile x_a and a CM $\mathbf{S}_a$ as a priori information. We indicate with $f(x)$ the forward model, which allows us to express the observations y as a function of the true profile x_t by the following equation:

$$y = f(x_t) + \varepsilon, \tag{1}$$

where ε is the vector including both the noise errors of the observations and the forward model errors, due to parameter errors and physical approximations of the forward model. Generally, the forward model calculates the radiative transfer through the Earth's atmosphere and knowing the state of the atmosphere, the observation geometry (for example either limb or nadir) and the characteristics of the instrument allows us to simulate the radiances measured in the given conditions. In order to simplify the formulation, we assume that there are no forward model errors and, therefore, ε includes only the noise errors of the observations and is characterized by $\langle \varepsilon \rangle = \mathbf{0}$ and $\langle \varepsilon \varepsilon^T \rangle = \mathbf{S}_{ny}$. $\langle \ldots \rangle$ indicates the mean value, and $\mathbf{S}_{ny}$ is the CM of the noise errors of the observations. A formulation that takes into account forward model errors can be obtained defining new observations corrected for the bias of the forward model errors and replacing $\mathbf{S}_{ny}$ with the sum of $\mathbf{S}_{ny}$ and the CM of the random part of the forward model errors.

The sensitivity of $\hat{x}$ to the true profile x_t is described by the AKM $\mathbf{A} = \frac{\partial \hat{x}}{\partial x_t}$, and the retrieval errors of $\hat{x}$ are described by the CM $\mathbf{S}$, which is the sum of the CM of the noise errors $\mathbf{S}_n$ and the CM of the smoothing errors, which are due to the smoothing of the true profile caused by the averaging kernels, $\mathbf{S}_s$. The AKM and the CMs are given by (see Equations (3.28)–(31) in [1]):

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

$$\mathbf{S}_n = \left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{F}\left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}, \tag{3}$$

$$\mathbf{S}_s = \left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{S}_a^{-1}\left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}, \tag{4}$$

$$\mathbf{S} = \mathbf{S}_n + \mathbf{S}_s = \left(\mathbf{F} + \mathbf{S}_a^{-1}\right)^{-1}, \tag{5}$$

where

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

with $\mathbf{K}$ being the Jacobian of the forward model $f(x)$ calculated at $\hat{x}$: $\mathbf{K} = \left.\dfrac{\partial f(x)}{\partial x}\right|_{x=\hat{x}}$. The matrix $\mathbf{F}$ is the Fisher information matrix [1,15], defined as

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

where $P(y|x)$ is the conditional probability distribution to obtain y given x, which, considered as a function of x, is referred to as the *likelihood function* $L(x)$ [16]. In the case that the inverse problem can be solved without constrain ($\mathbf{S}_a^{-1} = \mathbf{0}$), that is when we can find the solution of maximum likelihood, from Equations (3)–(5), we see that $\mathbf{F}$ is equal to the inverse matrix of the CM of the retrieval errors ($\mathbf{S}$), which coincides with the CM of the noise errors ($\mathbf{S}_n$). From this consideration, we can understand that the physical meaning of $\mathbf{F}$ is quantifying the information provided by the observations y about the retrieved vertical profile.

$\mathbf{F}$ depends on the a priori information used in the retrieval through $\mathbf{K}$ calculated at $\hat{x}$, which depends on x_a and $\mathbf{S}_a$. 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 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 x_t into the retrieved profile $\hat{x}$ and, accordingly, defines the retrieved profile $\hat{x}$ as a function of the true profile x_t. This function is referred to as the *transfer function* [1], and besides being a function of x_t, it is also a function of the noise errors ε of the observations y. This dependence can be seen recalling that really $\hat{x}$ depends on x_t through the observations y; therefore, using Equation (1) we can write $\hat{x} = \hat{x}(y) = \hat{x}(f(x_t) + \varepsilon)$, which we indicate as $\hat{x} = \hat{x}(x_t, \varepsilon)$. We note that $\dfrac{\partial \hat{x}(x_t, \varepsilon)}{\partial \varepsilon} = \dfrac{\partial \hat{x}(y)}{\partial y}\dfrac{\partial y}{\partial \varepsilon} = \dfrac{\partial \hat{x}(y)}{\partial y}$ because from Equation (1), it results that $\dfrac{\partial y}{\partial \varepsilon}$ is the identity matrix.

Expanding the transfer function at the first order around the a priori profile $x_t = x_a$ and zero errors $\varepsilon = \mathbf{0}$, we obtain:

$$\hat{x}(x_t, \varepsilon) \simeq \hat{x}(x_a, \mathbf{0}) + \left.\frac{\partial \hat{x}(x_t, \varepsilon)}{\partial x_t}\right|_{\substack{x_t = x_a \\ \varepsilon = \mathbf{0}}} (x_t - x_a) + \left.\frac{\partial \hat{x}(x_t, \varepsilon)}{\partial \varepsilon}\right|_{\substack{x_t = x_a \\ \varepsilon = \mathbf{0}}} \varepsilon. \tag{8}$$

Concerning the first term of the expansion, we recall that the retrieved profile obtained with the optimal estimation method in the 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 the absence of errors results in the a priori profile, that

is $\hat{x}(x_a, 0) = x_a$. 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 [6,17] 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 $\left.\dfrac{\partial \hat{x}(x_t, \varepsilon)}{\partial x_t}\right|_{\substack{x_t = x_a \\ \varepsilon = 0}} \simeq \left.\dfrac{\partial \hat{x}(x_t, \varepsilon)}{\partial x_t}\right|_{\substack{x_t = \hat{x} \\ \varepsilon = 0}} = \mathbf{A}$ and

$\left.\dfrac{\partial \hat{x}(x_t, \varepsilon)}{\partial \varepsilon}\right|_{\substack{x_t = x_a \\ \varepsilon = 0}} \simeq \left.\dfrac{\partial \hat{x}(x_t, \varepsilon)}{\partial \varepsilon}\right|_{\substack{x_t = \hat{x} \\ \varepsilon = 0}} = \left.\dfrac{\partial \hat{x}(y)}{\partial y}\right|_{y=f(x_t)} = \mathbf{G}$, where $\mathbf{G}$ is the gain matrix and is given by

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

On the basis of these considerations, Equation (8) becomes

$$\hat{x} = x_a + \mathbf{A}(x_t - x_a) + \mathbf{G}\varepsilon. \tag{10}$$

Following the approach described in the complete data fusion method [6,17], we define the vector $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \hat{x} - x_a + \mathbf{A}x_a, \tag{11}$$

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

$$\boldsymbol{\alpha} = \mathbf{A}x_t + \mathbf{G}\varepsilon \tag{12}$$

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

2.3. The New Variables $\boldsymbol{\beta}$

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

$$\boldsymbol{\beta} = \mathbf{S}^{-1}\boldsymbol{\alpha} = \mathbf{S}^{-1}(\hat{x} - x_a + \mathbf{A}x_a) \tag{13}$$

and using Equations (2), (5), (9) and (10), we obtain

$$\boldsymbol{\beta} = \mathbf{F}x_t + \delta, \tag{14}$$

where the vector δ is given by

$$\delta = \mathbf{K}^T \mathbf{S}_{ny}^{-1} \varepsilon. \tag{15}$$

Equation (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 δ is the vector that includes the errors of this measurement. Furthermore, from Equations (14) and (15) we see that $\boldsymbol{\beta}$, in the linear approximation of the forward model, is uniquely determined independently of both x_a and $\mathbf{S}_a$.

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

$$\mathbf{A}_\beta = \frac{\partial \boldsymbol{\beta}}{\partial x_t} = \mathbf{F} \tag{16}$$

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

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

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

From Equation (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 graphical representation of the profile, but to efficiently provide all of the information of the observations to subsequent data analyses.

3. Advantages of the Use of the Variables $\boldsymbol{\beta}$

3.1. Representation of the Profile Using Any Constraint

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

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

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

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

Equation (19) can be used to recover the original retrieved profile using the a priori information 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, Equation (19) can be used to produce a profile with any a priori information we like.

3.2. Data Fusion

If we suppose to have N independent measurements $\hat{x}_i$ of the same vertical profile 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 of the N measurements using the complete data fusion formula [17]

$$x_f = \left(\sum_{i=1}^{N} \mathbf{S}_i^{-1} \mathbf{A}_i + \mathbf{S}_a^{-1} \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{S}_i^{-1} \alpha_i + \mathbf{S}_a^{-1} x_a \right) \tag{20}$$

where x_a and $\mathbf{S}_a$ are the a priori profile and CM used to constrain the fused profile x_f, and α_i are the vectors defined by Equation (11) for each measurement:

$$\alpha_i = \hat{x}_i - x_{ai} + \mathbf{A}_i x_{ai} \tag{21}$$

where x_{ai} is the a priori profile used in the retrieval of the i-th measurement. The use of Equation (20) is equivalent to perform the data fusion using the approach of the Kalman filter [1,18], as shown in [19]

Using Equations (2), (5) and (13), we can rewrite Equation (20) as

$$x_{\mathrm{f}} = \left(\sum_{i=1}^{N} \mathbf{F}_i + \mathbf{S}_{\mathrm{a}}^{-1} \right)^{-1} \left(\sum_{i=1}^{N} \boldsymbol{\beta}_i + \mathbf{S}_{\mathrm{a}}^{-1} x_{\mathrm{a}} \right) \tag{22}$$

where $\boldsymbol{\beta}_i$ and $\mathbf{F}_i$ are the $\boldsymbol{\beta}$ and $\mathbf{F}$ quantities related to each one of the N measurements. Equation (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.

3.3. Reduction in 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 for the complete use of the products, in further processing of the data such as data fusion or data assimilation, are: $\hat{x}$, $\mathbf{A}$, $\mathbf{S}$ and x_{a}. $\mathbf{S}_{\mathrm{a}}$ is not necessary, because it can be obtained from $\mathbf{A}$ and $\mathbf{S}$ by means of

$$\mathbf{S}_{\mathrm{a}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{S} \tag{23}$$

which is derived using Equations (2) and (5).

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 for the 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 we wish give more complete information specifying where the Jacobian $\mathbf{K}$ is calculated, we can also give $\hat{x}$, and the values that have to be stored are $(n^2 + 5n)/2$. In Table 1, we summarize the data volume of the quantities stored when using the standard retrieval products and the new variables.

Table 1. Data volume of the quantities stored when using the standard retrieval products and the new variables.

Standard Products		New Variables	
Quantities	**Number of Values**	**Quantities**	**Number of Values**
$\hat{x}$	n	$\boldsymbol{\beta}$	n
$\mathbf{A}$	n^2	$\mathbf{F}$	$n(n + 1)/2$
$\mathbf{S}$	$n(n + 1)/2$		
x_{a}	n		
Total number of values			
$(3n^2 + 5n)/2$		$(n^2 + 3n)/2$	

Since the main storage requirement is due to the square term, the use of the variables $\boldsymbol{\beta}$ allows us to reduce the stored data to about one-third of its 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, are becoming more relevant. In light of this, new retrieval variables have been proposed when the retrieval has been performed with the optimal estimation method and the first order approximation of the transfer function is appropriate. 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, it is independent of the a priori information used in the retrieval, and both the AKM and the CM of β coincide with the Fisher information matrix. Furthermore, the variables β 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 performed with different instruments. For the exploitation of these products in the subsequent operations, it is sufficient to provide β and the Fisher information matrix **F**, which fully characterizes the measurement, being both its AKM and its CM. Accordingly, the use of the variables β allows us to reduce the stored data to about one-third of its volume with respect to the use of the standard products. These properties of the variables β make them a perfect retrieval product when further processing is performed by the users and encourage the possibility of considering finer retrieval grids, possibly concerted by the scientific community rather than determined by instrumental considerations. On the other hand, the standard products have the advantage of providing a graphical representation of the measured profiles. However, it is important to notice that the possibility of a graphical representation is obtained at the cost of a constraint on the adopted retrieval grid. The retrieval grid is usually limited in extension and density of points in order to avoid a too large bias of the a priori information, and different instruments freely use different retrieval grids that complicate comparisons. A storage procedure that does not depend on the a priori information can use a retrieval grid commonly used with the other instruments and avoid these difficulties.

The communities of data providers and data users are invited to test and validate the efficiency of this new interface.

References

1. Rodgers, C.D. *Inverse Methods for Atmospheric Sounding: Theory and Practice*; Series on Atmospheric, Oceanic and Planetary Physics; World Scientific: Singapore, 2000; Volume 2. [CrossRef]
2. Menke, W. *Geophysical Data Analysis: Discrete Inverse Theory*; Academic: San Diego, CA, USA, 1984.
3. Twomay, S. *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*; Elsevier: New York, NY, USA, 1977.
4. Doicu, A.; Trautmann, T.; Schreier, F. *Numerical Regularization for Atmospheric Inverse Problems*; Springer: Berlin/Heidelberg, Germany, 2010. [CrossRef]
5. Sivia, D.S.; Skilling, J. *Data Analysis: A Bayesian Tutorial*; Oxford University Press: Oxford, UK, 2006.
6. Ceccherini, S.; Carli, B.; Raspollini, P. Equivalence of data fusion and simultaneous retrieval. *Opt. Express* **2015**, *23*, 8476–8488. [CrossRef] [PubMed]
7. Ceccherini, S.; Carli, B.; Cortesi, U.; Del Bianco, S.; Raspollini, P. Retrieval of the vertical column of an atmospheric constituent from data fusion of remote sensing measurements. *J. Quant. Spectrosc. Radiat.* **2010**, *111*, 507–514. [CrossRef]
8. Ceccherini, S.; Cortesi, U.; Del Bianco, S.; Raspollini, P.; Carli, B. IASI-METOP and MIPAS-ENVISAT data fusion. *Atmos. Chem. Phys.* **2010**, *10*, 4689–4698. [CrossRef]
9. Warner, J.X.; Yang, R.; Wei, Z.; Carminati, F.; Tangborn, A.; Sun, Z.; Lahoz, W.; Attié, J.-L.; El Amraoui, L.; Duncan, B. Global carbon monoxide products from combined AIRS, TES and MLS measurements on A-train satellites. *Atmos. Chem. Phys.* **2014**, *14*, 103–114. [CrossRef]

10. Cortesi, U.; Del Bianco, S.; Ceccherini, S.; Gai, M.; Dinelli, B.M.; Castelli, E.; Oelhaf, H.; Woiwode, W.; Höpfner, M.; Gerber, D. Synergy between middle infrared and millimeter-wave limb sounding of atmospheric temperature and minor constituents. *Atmos. Meas. Tech.* **2016**, *9*, 2267–2289. [CrossRef]

11. Schneider, M.; Ertl, B.; Tu, Q.; Diekmann, C.J.; Khosrawi, F.; Röhling, A.N.; Hase, F.; Dubravica, D.; García, O.E.; Sepúlveda, E.; et al. Synergetic use of IASI profile and TROPOMI total-column level 2 methane retrieval products. *Atmos. Meas. Tech.* **2022**, *15*, 4339–4371. [CrossRef]

12. Aires, F.; Aznay, O.; Prigent, C.; Paul, M.; Bernardo, F. Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A. *J. Geophys. Res. Atmos.* **2012**, *117*, D18. [CrossRef]

13. Ridolfi, M.; Tirelli, C.; Ceccherini, S.; Belotti, C.; Cortesi, U.; Palchetti, L. Synergistic retrieval and complete data fusion methods applied to simulated FORUM and IASI-NG measurements. *Atmos. Meas. Tech.* **2022**, *15*, 6723–6737. [CrossRef]

14. Ceccherini, S.; Carli, B.; Raspollini, P. Vertical grid of retrieved atmospheric profiles. *J. Quant. Spectr. Radiat. Transf.* **2016**, *174*, 7–13. [CrossRef]

15. Ceccherini, S.; Carli, B.; Raspollini, P. Quality quantifier of indirect measurements. *Opt. Express* **2012**, *20*, 5151–5167. [CrossRef] [PubMed]

16. Fisher, R.A. On the mathematical foundation of theoretical statistics. *Philos. Trans. R. Soc. Lond.* **1922**, *A222*, 309. [CrossRef]

17. Ceccherini, S.; Zoppetti, N.; Carli, B. An improved formula for the complete data fusion. *Atmos. Meas. Tech.* **2022**, *15*, 7039–7048. [CrossRef]

18. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng. ASME* **1960**, *82*, 35–45. [CrossRef]

19. Ceccherini, S. Comment on "Synergetic use of IASI profile and TROPOMI total-column level 2 methane retrieval products" by Schneider et al. (2022). *Atmos. Meas. Tech.* **2022**, *15*, 4407–4410. [CrossRef]