Uncertainty and Error Correlation Quantification for FIDUCEO "easy-FCDR" Products: Mathematical Recipes

Chris Merchant, Emma Woolliams and Jonathan Mittaz

University of Reading and National Physical Laboratory

20/02/2018 18:00:00 version 1.a

FIDUCEO has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no. 638822

1 Introduction

1.1 Scope

This document defines the mathematical recipes to be used to populate the uncertainty information required in the user-friendly Fundamental Climate Data Records (so-called “easy-FCDRs”) developed by the project.

Prior familiarity with the concepts of measurement-function based uncertainty analysis and FCDR effects tables developed in the FIDUCEO project is assumed. These concepts are defined in Woolliams et al. (2017) “Principles behind the FCDR effect tables”, RD.1.

The mathematical notation follows the conventions of RD.2 and prior familiarity with these conventions is also assumed. While equation contents are defined in this document, some of the information conveyed by the notation may be lost without reference to those conventions.

The scope of this document is:

- **To give a brief scientific description of the uncertainty information provided in easy-FCDRs**
- **Taking the existence of an adequate set of effects tables as given, to provide and explain the mathematical formulae for generating the uncertainty information**
- **To identify the sensor-specific decisions sensor teams must make in the implementation of these formulae**

The easy-FCDR format definition is provided in RD.3 and is not discussed here.

1.2 Version Control

Date	Person	Version	Action / Reason
20.2.18	Merchant	1.a	Release after comments on drafts

1.3 Applicable and Reference Documents

RD.1 Woolliams E, J Mittaz, C J Merchant and P Harris, 2017, Principles behind the FCDR effects table (v1.a 25/08/2017), FIDUCEO Deliverable D2-2a.

RD.2 Merchant C J and E Woolliams, 2017, Mathematical notation for FIDUCEO publications (v1a 17/11/2017), http://www.fiduceo.eu/content/mathematical-notation-fiduceo-publications.

RD.3 Block, T and S Embacher, 2017, CDR/FCDR File Format Specification. FIDUCEO deliverable D3.1., and see http://fiduceo.pbworks.com/w/page/117806913/FCDR%20File%20Format

RD.4 Merchant C J, and many authors, 2017, ***Uncertainty information in climate data records from Earth observation.*** Earth System Science Data, 9 (2). pp. 511-527. ISSN 1866-3516 doi: 10.5194/essd-9-511-2017.

2 Scientific over-view of easy-FCDR content

2.1 Context and motivation

The FIDUCEO vocabulary defines an *uncertainty-quantified fundamental climate data record* (FCDR) as:

> *A record of calibrated, geolocated, directly-measured satellite observations in geophysical units (such as radiance) in which estimates of total uncertainty (or error covariance) and/or dominant components of uncertainty (or error covariance) are provided or characterised at pixel-level (and potentially larger) scales. The FCDR should be provided with all relevant auxiliary information for the data to be meaningful, including, e.g. time of acquisition, longitude and latitude, solar and viewing angles, sensor spectral response.*

The FCDR is a long-term record of a geophysical quantity measured by a satellite with all the necessary information to interpret that record in a quantitative manner. FCDRs are produced as an initial step in a processing chain. They are used when they are converted into climate data records (CDRs) of higher level products, a process that can combine FCDR data values from different spectral channels and combine FCDR data values from different image pixels.

Uncertainty information is typically complex, particularly for the multi-variate radiance data comprising level 1 satellite imagery. Simplification and summary of uncertainty information increases its conceptual accessibility and ease of application, at the cost of reducing the scientific benefits of the uncertainty information to derived geophysical products at level 2+. In CDRs and similar geophysical products, it is recommended (RD.4):

- To include rigorous uncertainty information to support the application of the data in contexts such as policy, climate modelling, and numerical weather prediction reanalysis
- To quantify uncertainty consistently with international metrological norms
- To provide uncertainty information per datum if necessary to discriminate observations with lesser and greater uncertainty
- To quantify uncertainty across spatial scales of averaging/aggregation of data

Part of level 2+ uncertainty arises from the propagation of level 1 uncertainty through the processes of image classification, retrieval and aggregation that are typically involved in transformations to higher processing levels.

The purpose of the "easy-FCDR" products from FIDUCEO is to provide level 1 data users with sufficient radiance[1] uncertainty information to propagate uncertainty to higher-order geophysical products with adequate rigour (or to use the radiance in data assimilation with knowledge of the radiance observation error covariances). The aim is to be "as simple as possible, but not simpler".

[1] Level 1 satellite imagery may be quantified as channel-integrated spectral radiance, reflectance or brightness temperature. Throughout this document, "radiance" is used generically, encompassing all such representations of the radiant energy measured in remote sensing.

2.2 Easy-FCDR uncertainty information

Easy-FCDR products will provide users with (re)calibrated satellite radiances, viewing geometry and geolocation data in a net-CDF format as conveniently as possible (e.g., duplicates suppressed, orbits consolidated).

Additionally, there will be uncertainty data in each product. These will consist of:

- Per-pixel, per-channel magnitude of radiance uncertainty
- Per-product, per-channel length-scales of cross-element and cross-line radiance error correlation
- Per-product cross-channel radiance error correlation

The rationale and content of each of the above classes of uncertainty are presented in turn in the following sections.

2.3 Magnitude of radiance uncertainty

The measure of the magnitude of uncertainty used is the standard uncertainty (coverage factor of 1).

Across a satellite image[2], the uncertainty in radiance can (and often does) vary significantly between pixels and between channels. This in turn means the uncertainty propagated to a level 2 retrieval product varies between pixels, and thus to meet the CDR recommendations noted above, per-pixel-per-channel radiance uncertainty magnitude is provided in the easy FCDR.

We discriminate three classes of effects (error sources) here:

- **Independent errors.** Some effects (error sources) cause white noise: i.e., errors that are independent (or very nearly so) between measured radiance values for different pixels in a channel. This is referred to as “spatial independence” across the image[3]. Such errors, as well as being independent, are also random (meaning that their origin is stochastic and cannot be corrected for even in principle).
- **Structured errors.** Other effects cause errors that have spatial structure within the orbit/slot: i.e., knowledge of the size of error in one pixel would enable one to predict (fully or partially) in another pixel. Structured errors arise from both random processes and systematic effects (effects that could in principle be corrected if more or better information were available, such as an improved calibration coefficient).
- **Common errors.** Other effects cause errors that are correlated on large scales (beyond one orbit/slot, potentially across a whole mission). For a given file, these are approximated as a common error in all radiances within a given orbit/slot (i.e., approximated as errors that are fully in common within the image). The uncertainty in re-calibration of radiance by harmonisation is treated in this category, and is presently assumed to be the only case in this category.

[2] “Image” is used to mean, for example, an orbit of swath data from an across-track scanning sensor in low-Earth orbit, or the data obtained from a single acquisition slot of a sensor in geostationary orbit.

[3] Since radiances in a given channel are generally measured sequentially in time, there is also a sense in which the independence is temporal.

The total standard uncertainty in a radiance value arises from the combination of the uncertainty associated with independent, structured and common errors. However, when radiances or retrievals from radiances are averaged spatio-temporally, the reduction in uncertainty from aggregating over data is different for the independent and structured components of uncertainty, and there is no reduction of the uncertainty associated with the common component. Since, for climate and other applications, level 3 (averaged, gridded) products are common and also require per datum uncertainty information, it is necessary to distinguish independent, structured and common components of total uncertainty at level 1 to enable propagation to higher levels.

Therefore, in the easy FCDR for each channel, c, there is additionally provided:

- a data layer of the same dimension as the radiance image containing the standard uncertainty arising from the combination of all independent-error sources, u_{i}
- a data layer of the same dimension as the radiance image containing the standard uncertainty arising from the combination of all structured-error sources, u_{s}
- the single standard uncertainty from the combination of all common-error (large scale) sources, u_h.

And, from these, the total standard uncertainty for a given pixel and channel is $u = \sqrt{u_{\mathrm{i}}^2 + u_{\mathrm{s}}^2 + u_h^2}$

This multiplies the number of values in the easy-FCDR compared to only having the radiances by about 3. The data volume does not increase by the same factor if the uncertainty values can be given with lesser numerical precision. Nonetheless, the data volume increase is a significant overhead for users of the easy-FCDR. To minimise any further impact on users, the data volume of any further information about the uncertainty needs to be small; in practice this means that information about the form of the structured-error sources must be approximated.

2.4 Radiance error correlation length scales for structured-error effects

If there are n_p pixels per channel in a satellite image, the cross-pixel error covariance (or correlation) matrix is of dimension $n_p \times n_p$ per channel, which is an infeasible data volume to provide in the easy FCDR. Summary information about the spatial correlation of errors must therefore be provided, at the expense of simplifying the information available to users about the error correlation structure.

Many effects cause errors with a different correlation structure element-wise (across-track) to line-wise (along track). It is therefore reasonable to provide two scales of error correlation for each of these dimensions (although not necessary for every case, perhaps).

There are various potential variations of error correlation with pixel separation, and these may differ between the many effects that combine to give the total error in each image pixel. In order to allow users to deal with a generic characterisation for simplicity, the average dependency is assumed to be

$$r_{l,l'} = \exp\left(-\frac{|l-l'|}{\Delta_l}\right) \qquad \text{Eq. 1}$$

where Δ_l is the provided length scale for line-wise error correlation, and l and l' are the line indices for the two pixels concerned. There is an equivalent expression for element-wise separation, involving a length scale Δ_e, and for an arbitrary pair of pixels, $p = (l, e)$ and $p' = (l', e')$ it is recommended to assume the error correlation is

$$r_{p,p'} = r_{l,l'} r_{e,e'} \qquad \text{Eq. 2}$$

The length scales may often be similar between channels, but this is not assumed, and so Δ_l and Δ_e are specified per channel.

The length scales are relevant only to the structured effects: there is no cross-pixel error correlation from the independent effects by definition; and perfect cross-pixel error correlation is assumed for the common effects.

2.5 Cross-channel error correlation

Full cross-channel error correlation matrices are provided. This is affordable from the point of view of data volume because the channel dimension is generally small compared to the pixel dimensions. It is scientifically well justified because most geophysical retrievals are functions combining data from multiple channels, and rigorous propagation of uncertainty from level 1 to level 2 is desirable and computationally feasible given the cross-channel error correlation matrices, $\boldsymbol{R}_c^i$ and $\boldsymbol{R}_c^s$ together with the corresponding magnitudes of uncertainty. Indeed, in some retrieval methods (optimal estimation and data assimilation) the error covariance matrix $\boldsymbol{S}_c$ that can readily be calculated is able to be used directly within the geophysical measurement function.

It is necessary to give $\boldsymbol{R}_c^i$ because independence is defined as spatial independence between pixels in a given channel, and the independent effects in this sense do not necessarily cause errors the are channel-wise independent. However, if the independent effects are also channel-wise independent, then $\boldsymbol{R}_c^i = \boldsymbol{I}$.

2.6 Why this set of uncertainty information?

Some comments have been made about this choice of uncertainty information above, and here some additional points are made:

- The split into independent and structured is considered to be reasonably intuitive and analogous to the more familiar "random/systematic" dichotomy. (Indeed, a user treating the independent and structured terms as random and systematic spatially (ignoring the length-scale information) would be making a conservative assumption in terms of the uncertainty they would estimate for a spatial average. If the scale of averaging were small compared to the error correlation length scales this approximation would be a good approximation.)
- A tri-partite division of uncertainty components that has gained traction at level 2 is "uncorrelated/locally correlated/large-scale correlated". This division is analogous to "independent/structured/common" defined here at level 1, although the determining factors for the scales are level 2 are typically very different (arising from the retrieval process rather than propagating from level 1).
- The independent and structured errors are caused by different effects and the uncertainty magnitudes associated with independent and structured effects may vary independently of each other across an image; assuming the error correlation structures are less variable than the uncertainty magnitudes, a user can use the per-pixel uncertainty estimates with the per-image Δ_l, Δ_e, $\boldsymbol{R}_c^i$ and $\boldsymbol{R}_c^s$ values to give discrimination of more and less uncertain pixels at level 2, which is a desirable capability for CDR production.

The assumption that error correlation structures are less variable than the uncertainty magnitudes is unlikely to be universally valid, but

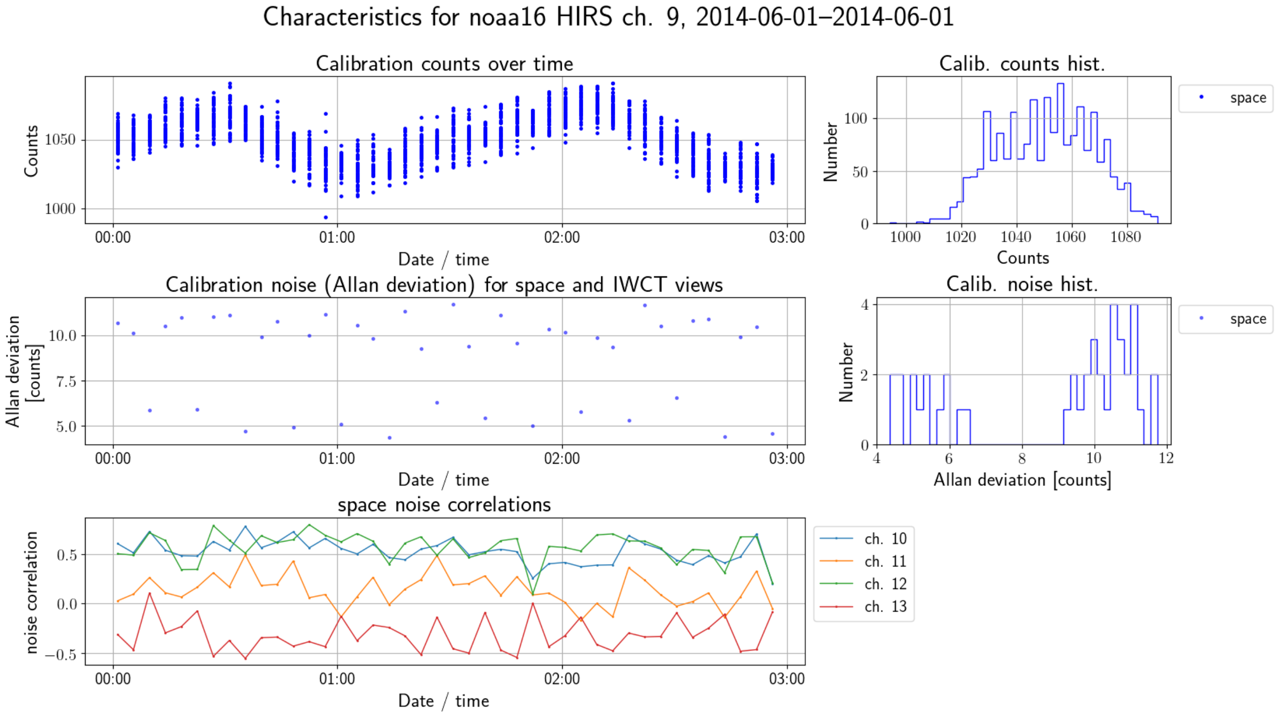

Figure 1 illustrates a case where it is a useful approach.

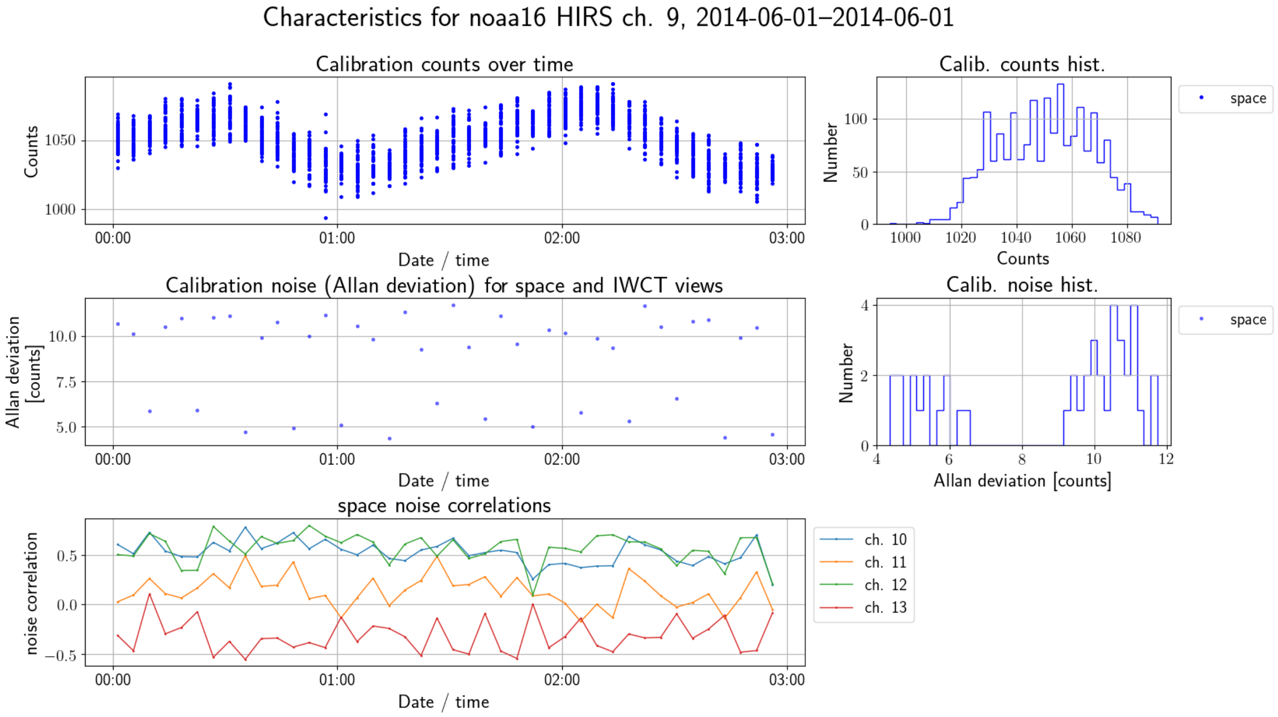

Figure 1. Timeseries (left) and histograms (right) of space-view calibration counts in Ch 9 (top), Allan deviation estimate of noise in Ch 9 (middle) and cross-pixel correlation (lower left) in four channels (legend). The noise fluctuates by a factor of two between two modes of noise behaviour, while the degree of correlation in different channels is more stable, at least for Ch 10 and Ch 12.

3 How the uncertainty information is calculated

3.1 Starting point for mathematical recipes

It is assumed that

- the measurement function, $y = f\left(x_1, \ldots, x_{n_j}, \boldsymbol{a}\right) + 0$, for radiance for a given sensor has been analysed, which means that the traceability tree for uncertainty has been constructed and error sources (effects) identified
- the sufficient set of effects for providing rigorous level 1 uncertainty information has been determined
- for each effect to be included in the radiance uncertainty modelling, an effects table has been constructed
- each effects table has been codified, including that all pre-calculated (stored) data layers have been created, and on-the-fly (virtual) variables have been codified
- the set of sufficient effects, k, describing uncertainty across all the terms x_j, has been mapped onto the independent subset, i, and structured subset, s
- harmonised calibration coefficients, $\boldsymbol{a}$, are available together with their error covariance

This in turn means that for any given effect, k, the following are available (stored or calculable on-the-fly):

- the term uncertainty, u_k, associated with the effect for any channel/pixel, (c, l, e), in an image
- the sensitivity coefficient, c_j, for the measurement function term associated with the effect, for any (c, l, e)
- the cross-pixel correlation coefficient, r_k, for any channel between any pair of pixels $(l, e; l', e')$
- the cross-channel correlation coefficient for the pair of channels (c, c') evaluated for any pixel (l, e)

The sections that follow describe the mathematics of the transformation of the above into the easy-FCDR uncertainty information.

3.2 Definition of matrices that will be used

3.2.1 Identity and all-ones matrices

The error covariance matrix between pixels for an independent effect is an example where we need the identity matrix:

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 3}$$

Many effects are fully correlated across elements of a given line, and in such a case we can identify the error covariance matrix of terms between elements as the all-ones matrix:

$$\boldsymbol{J} = \begin{bmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 4}$$

3.2.2 Cross-line term error correlation matrix for one effect

Evaluation of this matrix uses the effects table information "Error Correlation Type and Form" and "Error Correlation Scale" (from line to line) for the effect, k, under consideration. The dimension of the matrix calculated is n_l where $\max(n_l) = N_l$ where N_l is the number of lines in the image: for computational reasons it may be practical to have $n_l < N_l$, provided the choice of n_l is adequate to evaluation of the cross-line error correlation length-scale aggregated over all effects (which will be a sensor-specific judgement).

For an effect k operating on measurement function term x_j the term error correlation matrix along an element (cross-line) is evaluated for a particular element in the image e as

$$\boldsymbol{R}_l^{e,k} = \begin{bmatrix} 1 & r_k\left(x_j(1,e), x_j(2,e)\right) & \cdots \\ r_k\left(x_j(2,e), x_j(1,e)\right) & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 5}$$

where $r_k\left(x_j(2,e), x_j(1,e)\right)$ is the error correlation coefficient between x_j evaluated at pixel $(2, e)$ and at pixel $(1, e)$, etc. The matrix can vary across the elements (and lines) of the image, and evaluation for one particular element is indicated by the e superscript in $\boldsymbol{R}_l^{e,k}$. $\boldsymbol{R}_l^{e,k}$ needs to be evaluated for each channel, and the channel dependence is not made explicit here in the notation.

3.2.3 Cross-element term error correlation matrix for one effect

Evaluation of this matrix uses the effects table information "Error Correlation Type and Form" and "Error Correlation Scale" (from pixel to pixel) for the effect, k, under consideration. The dimension of the matrix calculated is n_e, the number of elements across the image.

For an effect k operating on measurement function term x_j the term error correlation matrix across a line (cross-element) is evaluated for a particular line in the image l as

$$\boldsymbol{R}_e^{l,k} = \begin{bmatrix} 1 & r_k\left(x_j(l,1), x_j(l,2)\right) & \cdots \\ r_k\left(x_j(l,2), x_j(l,1)\right) & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 6}$$

where $r_k\left(x_j(l,1), x_j(l,2)\right)$ is the error correlation coefficient between x_j evaluated at pixel $(l, 1)$ and at pixel $(l, 2)$, etc. The matrix can vary along lines of the image, and evaluation for one particular line is indicated by the l superscript in $\boldsymbol{R}_e^{l,k}$. $\boldsymbol{R}_e^{l,k}$ needs to be evaluated for each channel, and the channel dependence is not made explicit here in the notation.

3.2.4 Cross-channel term error correlation matrix for one effect

Evaluation of this matrix uses the effects table information "Channels / bands". The dimension of the matrix is n_c, the number of channels, and channels unaffected by a particular effect k (and therefore not listed as affected in the effects table) are entered with error correlations of 0.

For an effect k operating on measurement function term x_j the term error correlation matrix between channels is evaluated for a particular pixel in the image p as

$$\boldsymbol{R}_c^{p,k} = \begin{bmatrix} 1 & r_k\left(x_j(1), x_j(2)\right) & \cdots \\ r_k\left(x_j(2), x_j(1)\right) & 1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

where $r_k\left(x_j(2), x_j(1)\right)$ is the error correlation coefficient between x_j evaluated for channel 1 and for channel 2. The use of a measurement function of identical form for each channel is assumed. The matrix can vary across lines and elements of the image, and evaluation for one particular pixel is indicated by the p superscript to $\boldsymbol{R}_c^{p,k}$.

3.2.5 Cross-line term uncertainty matrix for one effect

Evaluation of this matrix uses the effects table information "Uncertainty [Magnitude]" for the effect, k, under consideration. The dimension of the matrix calculated is n_l as determined during the evaluation of the corresponding term error correlation matrix (§3.2.2). The uncertainty matrix is diagonal.

For an effect k operating on measurement function term x_j the term uncertainty matrix is

$$\boldsymbol{U}_l^{e,k} = \begin{bmatrix} u_k\left(x_j(1,e)\right) & 0 & \cdots \\ 0 & u_k\left(x_j(2,e)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 7}$$

where $u_k\left(x_j(1,e)\right)$ is the uncertainty in x_j from the considered effect evaluated at pixel $(1,e)$. The uncertainty may vary across the elements (and lines) of the image, and evaluation for one particular element is indicated by the e superscript in $\boldsymbol{U}_l^{e,k}$. $\boldsymbol{U}_l^{e,k}$ needs to be evaluated for each channel, and the channel dependence is not made explicit here in the notation.

3.2.6 Cross-element term uncertainty matrix for one effect

Evaluation of this matrix uses the effects table information "Uncertainty [Magnitude]" for the effect, k, under consideration. The dimension of the matrix calculated is n_e as determined during the evaluation of the corresponding term error correlation matrix (§3.2.2). The uncertainty matrix is diagonal.

For an effect k operating on measurement function term x_j the term uncertainty matrix is

$$\boldsymbol{U}_e^{l,k} = \begin{bmatrix} u_k\left(x_j(l,1)\right) & 0 & \cdots \\ 0 & u_k\left(x_j(l,2)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 8}$$

where $u_k\left(x_j(l,1)\right)$ is the standard uncertainty in x_j from the considered effect evaluated at pixel $(l,1)$, etc. The uncertainty can vary along the lines of the image, and evaluation for one particular line is indicated by the l superscript in $\boldsymbol{U}_e^{l,k}$. $\boldsymbol{U}_e^{l,k}$ needs to be evaluated for each channel, and the channel dependence is not made explicit here in the notation.

Additional notes on this section and §3.2.5:

- The contents of both $\boldsymbol{U}_e^{l,k}$ and $\boldsymbol{U}_l^{e,k}$ are extracts from the same array of all the term uncertainty across the image (in the given channel), $u_k\left(x_j(l,e)\right)$
- The presentation here is of the mathematics, and is not of a computational prescription: computationally, use of diagonal uncertainty matrices (which may be large particularly in the cross-line case) is unlikely to be optimum in terms of memory usage, for example
- The units of the standard uncertainty are those given in the effects table as Uncertainty [Units] which are the units of the j^{th} term of the measurement function

3.2.7 Cross-channel term uncertainty matrix for one effect

Evaluation of this matrix uses the effects table information "Uncertainty [Magnitude]" for the effect, k, under consideration. The dimension of the matrix calculated is n_c, with any channels not affected by the effect having uncertainty zero. The uncertainty matrix is diagonal. The elements are also from the arrays of all the term uncertainty values across the image, $u_k\left(x_j(l,e)\right)$, but selected for a given pixel across the channels.

For an effect k operating on measurement function term x_j the cross-channel term uncertainty matrix is evaluated at $p=(l,e)$ as

$$\boldsymbol{U}_c^{p,k} = \begin{bmatrix} u_k\left(x_j(1;l,e)\right) & 0 & \cdots \\ 0 & u_k\left(x_j(2;l,e)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 9}$$

where $u_k\left(x_j(1;l,e)\right)$ is the uncertainty in x_j from the considered effect evaluated at pixel (l,e) in channel $c=1$, etc. The uncertainty can vary between the pixels of the image, and evaluation for one particular line is indicated by the p superscript in $\boldsymbol{U}_c^{p,k}$.

3.2.8 Cross-line sensitivity matrix for one term

Evaluation of this matrix uses the effects table information "Sensitivity Coefficient" for the effect, k, under consideration, although the sensitivity coefficient is shared across all such terms affecting the uncertainty in the j^{th} term of the measurement function. The dimension of the matrix calculated is n_l as determined during the evaluation of the corresponding term error correlation matrix (§3.2.2). The sensitivity matrix is diagonal.

For a term j the cross-line sensitivity matrix is evaluated at element e of a given channel is

$$\boldsymbol{C}_l^{e,j} = \begin{bmatrix} c_j\left(x_j(1,e)\right) & 0 & \cdots \\ 0 & c_j\left(x_j(2,e)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 10}$$

where $c_j\left(x_j(1,e)\right) = \frac{\partial y(1,e)}{\partial x_j}$ is the sensitivity of radiance to x_j in the considered channel at pixel $(1,e)$, etc. The sensitivity can vary between the pixels of the image, and evaluation across lines for one particular element value is indicated by the e superscript in $\boldsymbol{C}_l^{e,j}$.

3.2.9 Cross-element sensitivity matrix for one term

Evaluation of this matrix uses the effects table information "Sensitivity Coefficient" for the effect, k, under consideration, although the sensitivity coefficient is shared across all such terms affecting the uncertainty in the j^{th} term of the measurement function. The dimension of the matrix calculated is n_e. The sensitivity matrix is diagonal.

For a term j the cross-line sensitivity matrix is evaluated at a line of a given channel is

$$\boldsymbol{C}_e^{l,j} = \begin{bmatrix} c_j\left(x_j(l,1)\right) & 0 & \cdots \\ 0 & c_j\left(x_j(l,2)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 11}$$

where $c_j\left(x_j(l,1)\right) = \frac{\partial y(l,1)}{\partial x_j}$ is the sensitivity of radiance to x_j in the considered channel at pixel $(l,1)$, etc. The sensitivity can vary between the pixels of the image, and evaluation across elements for one particular line is indicated by the l superscript in $\boldsymbol{C}_e^{l,j}$.

3.2.10 Cross-channel sensitivity matrix for one term

Evaluation of this matrix uses the effects table information "Sensitivity Coefficient" for the effect, k, under consideration, although the sensitivity coefficient is shared across all such terms affecting the uncertainty in the j^{th} term of the measurement function. The dimension of the matrix calculated is n_c, with any channels not affected by the effect having sensitivity zero. The sensitivity matrix is diagonal.

For a term j the cross-line sensitivity matrix is evaluated at a given pixel p is

$$\boldsymbol{C}_c^{p,j} = \begin{bmatrix} c_j\left(x_j(1;l,e)\right) & 0 & \cdots \\ 0 & c_j\left(x_j(2;l,e)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 12}$$

where $c_j\left(x_j(1;l,e)\right) = \frac{\partial y(1;l,e)}{\partial x_j}$ is the sensitivity of radiance to x_j in the considered channel at pixel (l,e) for channel 1, etc. The sensitivity can vary between the pixels of the image, and evaluation across pixels is indicated by the p superscript in $\boldsymbol{C}_c^{p,j}$.

Additional notes on this section, §3.2.8 and §3.2.9:

- The contents of all the sensitivity matrices are extracts from the array of all the sensitivity coefficients uncertainty across the image in different channels
- The presentation here is of the mathematics, and is not of a computational prescription: computationally, use of diagonal sensitivity matrices (which may be large particularly in the cross-line case) is unlikely to be optimum in terms of memory usage, for example

3.3 Equations for easy-FCDR contents

3.3.1 Standard uncertainty, independent and structured effects, per pixel per channel

The standard uncertainty for each channel and pixel per effect are either pre-calculated or calculated on the fly as defined in the effects tables.

Let the index of effects, k, comprise two subsets, namely, the spatially independent, i, and spatially structured, s.

The magnitudes of uncertainty from independent and structured effects are presented per pixel per channel in the easy-FCDR. Each datum is uniquely indexed by $(c;l,e)$.

The value of the independent uncertainty magnitude is:

$$u_i(c;l,e) = \sqrt{\sum_j \sum_{i|j} c_j\left(x_j(c;l,e)\right)^2 u_i\left(x_j(c;l,e)\right)^2} \quad \text{Eq. 13}$$

where $i|j$ is the subset of independent effects relevant to measurement function term j. This applies because by construction effects are independent of each other.

The value of the structured uncertainty magnitude is:

$$u_s(c;l,e) = \sqrt{\sum_j \sum_{s|j} c_j \left(x_j(c;l,e)\right)^2 u_s\left(x_j(c;l,e)\right)^2} \qquad \text{Eq. 14}$$

where $s|j$ is the subset of structured effects relevant to measurement function term j.

3.3.2 Cross-element radiance error correlation matrix, independent effects, per channel

It is not actually necessary to calculate this matrix: it is the identity matrix, since the effects considered are spatially independent. However, for understanding, we can derive this result as follows.

Let $\boldsymbol{R}_e^{l,i}$ be the error correlation matrix evaluated at any of the set of lines, l, from any of the independent effects, i. The fact that this applies to a particular channel is implicit. This matrix is defined in §3.2.3.

Now, consider the cross-element radiance error covariance associated with the errors described by $\boldsymbol{R}_e^{l,i}$ and propagation described by $\boldsymbol{C}_e^{l,j}$:

$$\boldsymbol{S}_e^{l,i} = \boldsymbol{C}_e^{l,j}\boldsymbol{U}_e^{l,i}\boldsymbol{R}_e^{l,i}\boldsymbol{U}_e^{l,i\,\mathrm{T}}\boldsymbol{C}_e^{l,j\,\mathrm{T}} \qquad \text{Eq. 15}$$

Error covariances add to make the total error covariance from a number of effects. So, the total error covariance matrix from the independent effects is:

$$\boldsymbol{S}_{e,i}^{l} = \sum_i \boldsymbol{S}_e^{l,i} = \sum_j \sum_{i|j} \boldsymbol{C}_e^{l,j}\boldsymbol{U}_e^{l,i}\boldsymbol{R}_e^{l,i}\boldsymbol{U}_e^{l,i\,\mathrm{T}}\boldsymbol{C}_e^{l,j\,\mathrm{T}} \qquad \text{Eq. 16}$$

This is still evaluated at a single line. Assume the sensor teams have identified an adequate strategy for sampling lines to form the average error covariance matrix representing an image, the averaging process across lines being notated $\langle . \rangle_l$. This gives

$$\boldsymbol{S}_{e,i} = \left\langle \boldsymbol{S}_{e,i}^{l} \right\rangle_l = \left\langle \sum_j \sum_{i|j} \boldsymbol{C}_e^{l,j}\boldsymbol{U}_e^{l,i}\boldsymbol{R}_e^{l,i}\boldsymbol{U}_e^{l,i\,\mathrm{T}}\boldsymbol{C}_e^{l,j\,\mathrm{T}} \right\rangle_l \qquad \text{Eq. 17}$$

Any covariance matrix, being positive definite and symmetric, can be represented in the form $\boldsymbol{S} = \boldsymbol{URU}^{\mathrm{T}}$. To see this, let element (m,n) of the matrix, $\boldsymbol{S}$, be $[S]_{m,n}$, etc. Make the following identifications

$$[U]_{m,n} = \delta_{m,n}\sqrt{[S]_{m,n}} \;\; ; \;\; [R]_{m,n} = \frac{[S]_{m,n}}{\sqrt{[S]_{m,m}[S]_{n,n}}} \qquad \text{Eq. 18}$$

then calculating $\boldsymbol{URU}^{\mathrm{T}}$ recovers $\boldsymbol{S}$. This is the principle on which an overall error correlation matrix can be calculated for a combination of effects. Thus, if $\boldsymbol{U}_{e,i}$ is the diagonal matrix whose elements on the diagonal are the square roots of the diagonal elements of $\boldsymbol{S}_{e,i}$, then the required matrix is

$$\boldsymbol{R}_{e,i} = \boldsymbol{U}_{e,i}^{-1}\boldsymbol{S}_{e,i}\boldsymbol{U}_{e,i}^{-1\,\mathrm{T}} \qquad \text{Eq. 19}$$

the form of which is correct irrespective of the nature of the errors. We now apply the fact that the errors are in this case independent, which means every $\boldsymbol{R}_e^{l,i} = \boldsymbol{I}$. In this case, $\boldsymbol{S}_{e,i} = \langle \sum_{\boldsymbol{i}} \boldsymbol{U}_e^{l,i}\boldsymbol{U}_e^{l,i\,\mathrm{T}} \rangle_l = \boldsymbol{U}_{e,i}\boldsymbol{U}_{e,i}^{\mathrm{T}}$, and $\boldsymbol{R}_{e,i} = \boldsymbol{U}_{e,i}^{-1}\boldsymbol{U}_{e,i}\boldsymbol{U}_{e,i}^{\mathrm{T}}\boldsymbol{U}_{e,i}^{-1\,\mathrm{T}} = \boldsymbol{II} = \boldsymbol{I}$, as expected.

Independence means no spatial error correlation length scale applies to this component of uncertainty.

3.3.3 Cross-element radiance error correlation matrix, structured effects, per channel

Using the results of the previous section, we have:

$$\boldsymbol{S}_{e,s} = \left\langle \sum_j \sum_{s|j} \boldsymbol{C}_e^{l,j} \boldsymbol{U}_e^{l,s} \boldsymbol{R}_e^{l,s} \boldsymbol{U}_e^{l,s\,\mathrm{T}} \boldsymbol{C}_e^{l,j\,\mathrm{T}} \right\rangle_l \qquad \text{Eq. 20}$$

$$\boldsymbol{U}_{e,s} = \begin{bmatrix} \sqrt{[S_{e,s}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{e,s}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 21}$$

$$\boldsymbol{R}_{e,s} = \boldsymbol{U}_{e,s}^{-1} \boldsymbol{S}_{e,s} \boldsymbol{U}_{e,s}^{-1\,\mathrm{T}} \qquad \text{Eq. 22}$$

3.3.4 Cross-element error correlation length scale, structured effects, per channel

The cross-element error correlation matrix (structured) quantifies the error correlation between every pair of elements across a line (for an average line). The concept of a single error correlation length scale implies that pairs of elements with the same separation have consistent correlation, which may be more or less true for different effects. The requirement here is to find one scale to be applied for all elements.

Consider the cross-element error correlation matrix written as $\boldsymbol{R}_{e,s} = \begin{bmatrix} 1 & r_{1,2} & r_{1,3} & r_{1,4} & r_{1,5} & \ddots \\ r_{1,2} & 1 & r_{2,3} & r_{2,4} & r_{2,5} & \ddots \\ r_{1,3} & r_{2,3} & 1 & r_{3,4} & r_{3,5} & \ddots \\ r_{1,4} & r_{2,4} & r_{3,4} & 1 & r_{4,5} & \ddots \\ r_{1,5} & r_{2,5} & r_{3,5} & r_{4,5} & 1 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$

where colours have been used to identify different minor diagonals. Interrogating the subscripts, which just refer to the pair of element numbers corresponding to each correlation coefficient, we see that each minor diagonal contains all the pixels of a given separation. Note that since the matrix is symmetric, we can from this point consider just the upper right triangle.

Form the set of data:

$$\begin{matrix} 0 & 1 \\ 1 & \langle r_{e,e+1} \rangle_e \\ 2 & \langle r_{e,e+2} \rangle_e \\ 3 & \langle r_{e,e+3} \rangle_e \\ \vdots & \vdots \\ n_e - 1 & r_{1,n_e} \end{matrix}$$

where the averages over elements of $r_{e,e+\Delta}$ are over the limits $e = 1$ to $e = n_e - \Delta$. Δ is the element separation, ranging from 0 to $n_e - 1$. The cross-element correlation length scale is found as the least-squares fit of the model $r_\Delta = \exp\left(-\frac{\Delta}{\Delta_e}\right)$ to the data, to find the optimum value of Δ_e. To estimate the error correlation for structured effects between two pixels in a line, (l, e) and (l, e'), a user would then evaluate

$$r_{e,e'} = \exp\left(-\frac{|e - e'|}{\Delta_e}\right) \qquad \text{Eq. 23}$$

3.3.5 Cross-line radiance error correlation matrix and length scale, independent effects, per channel

These are the identity matrix and 0 respectively.

3.3.6 Cross-line radiance error correlation matrix, structured effects, per channel

The equations are analogous to §3.3.3:

$$\boldsymbol{S}_{l,s} = \left\langle \sum_j \sum_{s|j} \boldsymbol{C}_l^{e,j} \boldsymbol{U}_l^{e,s} \boldsymbol{R}_l^{e,s} \boldsymbol{U}_l^{e,s\,\mathrm{T}} \boldsymbol{C}_l^{e,j\,\mathrm{T}} \right\rangle_e \quad \text{Eq. 24}$$

$$\boldsymbol{U}_{l,s} = \begin{bmatrix} \sqrt{[S_{l,s}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{l,s}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 25}$$

$$\boldsymbol{R}_{l,s} = \boldsymbol{U}_{l,s}^{-1} \boldsymbol{S}_{l,s} \boldsymbol{U}_{l,s}^{-1\,\mathrm{T}} \quad \text{Eq. 26}$$

Note that:

- It may be sufficient, as noted previously, to evaluate the correlation matrix for a subset of the full number of lines (particularly in cases where data comprise a full orbit this may be necessary).
- In such a case, the averaging across elements should also sample adequately across different starting lines for the subset. Sensor teams will need to devise an effective strategy.

3.3.7 Cross-line radiance error correlation length scale, structured effects, per channel

Analogously to §3.3.4:

- Average the minor diagonals of the upper triangle of $\boldsymbol{R}_{l,s}$ (see section 3.3.6)
- To these averages, as a function of separation in lines, fit an exponential function to determine Δ_l

3.3.8 Cross-channel radiance error correlation matrix, independent effects

"Independent effects" are classed as those spatiotemporally independent. They are not necessarily independent between channels, and therefore this error correlation matrix cannot be assumed to be the identity matrix. However, if all the matrices (see §3.2.4) $\boldsymbol{R}_c^{p,i} = \boldsymbol{I}$, then the cross-channel error correlation matrix is the identity matrix, and this may often be the case in practice.

Where the condition of all spatially independent effects being also spectrally independent is not met, the equations for the cross-channel error correlation matrix are analogous to §3.3.3:

$$\boldsymbol{S}_{c,i} = \left\langle \sum_j \sum_{i|j} \boldsymbol{C}_c^{p,j} \boldsymbol{U}_c^{p,i} \boldsymbol{R}_c^{p,i} \boldsymbol{U}_c^{p,i\,\mathrm{T}} \boldsymbol{C}_c^{p,j\,\mathrm{T}} \right\rangle_p \quad \text{Eq. 27}$$

$$\boldsymbol{U}_{c,i} = \begin{bmatrix} \sqrt{[S_{c,i}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{c,i}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \text{Eq. 28}$$

$$\boldsymbol{R}_{c,i} = \boldsymbol{U}_{c,i}^{-1} \boldsymbol{S}_{c,i} \boldsymbol{U}_{c,i}^{-1\,\mathrm{T}} \quad \text{Eq. 29}$$

Cross-channel correlation of errors may vary from pixel to pixel, and so the image-average estimate involves averaging over and adequate set of pixels (sampling across both lines and elements), according to a strategy that the sensor teams determine.

Being of tractable size, $\boldsymbol{R}_{c,i}$ is provided in the easy-FCDR.

3.3.9 Cross-channel radiance error correlation matrix, structured effects

The equations for the cross-channel error correlation matrix from structured effects are analogous to §3.3.3:

$$\boldsymbol{S}_{c,s} = \left\langle \sum_j \sum_{s|j} \boldsymbol{C}_c^{p,j} \boldsymbol{U}_c^{p,s} \boldsymbol{R}_c^{p,s} \boldsymbol{U}_c^{p,s\,\mathrm{T}} \boldsymbol{C}_c^{p,j\,\mathrm{T}} \right\rangle_p \qquad \text{Eq. 30}$$

$$\boldsymbol{U}_{c,s} = \begin{bmatrix} \sqrt{[S_{c,s}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{c,s}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{Eq. 31}$$

$$\boldsymbol{R}_{c,s} = \boldsymbol{U}_{c,s}^{-1} \boldsymbol{S}_{c,s} \boldsymbol{U}_{c,s}^{-1\,\mathrm{T}} \qquad \text{Eq. 32}$$

Cross-channel correlation of errors may vary from pixel to pixel, and so the image-average estimate involves averaging over and adequate set of pixels (sampling across both lines and elements), according to a strategy that the sensor teams determine.

Being of tractable size, $\boldsymbol{R}_{c,s}$ is provided in the easy-FCDR.

3.3.10 Uncertainty from common effects (calibration uncertainty)

Here, it is assumed that the only effect evaluated in this category is associated with uncertainty in the determination of harmonised calibration coefficients, so this could also be referred to as calibration uncertainty.

The calibration coefficients for a given channel are listed in a vector, $\boldsymbol{a}_c$, and having been estimated jointly in a minimisation process, the errors in $\boldsymbol{a}_c$ are covariant. We assume the error covariance is estimated as $\boldsymbol{S}(\boldsymbol{a}_c)$, and the sensitivity coefficients, $\boldsymbol{c}_h(c; l, e)$, can be calculated for any pixel and channel, and listed in a column vector corresponding to the ordering of the calibration coefficients.

The calculation of the calibration uncertainty (or, more generally, common uncertainty) for a given pixel is

$$\sqrt{\boldsymbol{c}_h^{\mathrm{T}} \boldsymbol{S}(\boldsymbol{a}_c) \boldsymbol{c}_h} \qquad \text{Eq. 33}$$

Where the "radiance" measurand is brightness temperature, assume that the common uncertainty will be given in the same units, kelvin, and the uncertainty is given absolutely.

If the sensitivity coefficients vary, there may be some variability of uncertainty across pixels. Despite this, this effect is approximated as causing error that is strictly in common within an orbit/slot for a given channel. To calculate the appropriate value of common uncertainty given absolutely, the above is averaged over an adequate set of pixels:

$$u_h = \left\langle \sqrt{\boldsymbol{c}_h^{\mathrm{T}} \boldsymbol{S}(\boldsymbol{a}_c) \boldsymbol{c}_h} \right\rangle_p \qquad \text{Eq. 34}$$

Where the "radiance" measurand is band radiance, a better assumption may be that the common uncertainty will be given relatively (i.e., as %).

Because of variable sensitivity coefficients, there may be some variability of uncertainty across pixels, but by expressing the uncertainty relatively, a major source of variability in uncertainty is factored out. The uncertainty expressed relatively is then approximated as being strictly in common within an orbit/slot for a given channel. To calculate the appropriate value of common uncertainty given relatively, the adapted expression averaged over an adequate set of pixels is:

$$u_h = \left\langle \frac{\sqrt{\boldsymbol{c}_h^{\mathrm{T}} \boldsymbol{S}(\boldsymbol{a}_c) \boldsymbol{c}_h}}{y_c} \right\rangle_p \times 100\% \qquad \text{Eq. 35}$$

Harmonisation is undertaken separately for each channel. Although there is potential for error correlation between channels because of common sampling effects in the harmonisation dataset, we assume this is negligible in practice.

4 How the uncertainty information may be used

It is not the objective of this report to explain how the uncertainty information may be used in detail. However, to motivate the complexity of calculation described above, it is worthwhile to give some indication of the use to which the calculated uncertainty variables could be put.

4.1 A multi-channel retrieval

A user has a retrieval of geophysical term z that operates pixel-by-pixel and uses more than one channel radiance, these being listed in column vector $\boldsymbol{y}^T = [y_1 \quad \cdots \quad y_c \quad \cdots]$. Let the retrieval be

$$z = f(\boldsymbol{y}, \boldsymbol{b})$$

where $\boldsymbol{b}$ represents retrieval terms additional to the used observations.

To find the uncertainty from propagated errors in the retrieval from the radiances at a certain pixel, $\boldsymbol{y}(p)$:

- Calculate $= \frac{\partial f}{\partial y}$, the column vector of sensitivity coefficients of the retrieval with respect to the radiances, by analytic or numerical differentiation.
- Delete the rows and columns corresponding to channels not used from the cross-channel error correlation matrices $\boldsymbol{R}_{c,i}$ and $\boldsymbol{R}_{c,s}$
- Look up and extract the independent and structured uncertainty components, $u_i(c;p)$ and $\mathrm{u_s}(c;p)$ from the easy-FCDR uncertainty data layers, along with $\overline{u_c}$.
- List the uncertainty magnitudes in matrices, namely

$$\boldsymbol{U}_{c,i} = \begin{bmatrix} u_i(1;p) & 0 & \cdots \\ 0 & u_i(2;p) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \boldsymbol{U}_{c,s} = \begin{bmatrix} u_s(1;p) & 0 & \cdots \\ 0 & u_s(2;p) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \boldsymbol{U}_{c,h} = \begin{bmatrix} \overline{u_1} & 0 & \cdots \\ 0 & \overline{u_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

- The single-pixel retrieval uncertainty *from propagation of radiance uncertainty* is then

$$u(z;p)^2 = \boldsymbol{c}^T\left(\boldsymbol{U}_{c,i}\boldsymbol{R}_{c,i}\boldsymbol{U}_{c,i}{}^{\mathrm{T}} + \boldsymbol{U}_{c,s}\boldsymbol{R}_{c,s}\boldsymbol{U}_{c,s}{}^{\mathrm{T}} + \boldsymbol{U}_{c,h}\boldsymbol{U}_{c,h}{}^{\mathrm{T}}\right)\boldsymbol{c} = \boldsymbol{c}^T\left(\boldsymbol{S}_{c,i} + \boldsymbol{S}_{c,s} + \boldsymbol{S}_{c,h}\right)\boldsymbol{c} = \boldsymbol{c}^T\boldsymbol{S}(\boldsymbol{y})\boldsymbol{c}$$

4.2 Optimal estimation

A typical maximum *a posteriori* retrieval equation is

$$\hat{\boldsymbol{z}} = \boldsymbol{z}_a + \boldsymbol{S}_a\boldsymbol{K}^T(\boldsymbol{K}\boldsymbol{S}_a\boldsymbol{K}^T + \boldsymbol{S}_\epsilon)^{-\mathbf{1}}(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{z}_a))$$

where the "measurement error covariance" is $\boldsymbol{S}_e$ and $\boldsymbol{K} = \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}}$. There is an equivalent expression in data assimilation. The measurement error covariance includes observation (radiance) error covariance (error covariance of $\boldsymbol{y}$), and also accounts for any forward model error or representativity/mismatch error in the evaluation of $\boldsymbol{F}$, the latter terms being assumed to be described by $\boldsymbol{S}_F$. The use of the radiance error covariance information is thus to set

$$\boldsymbol{S}_\epsilon = \boldsymbol{S}(\boldsymbol{y}) + \boldsymbol{S}_F$$

Thereafter, the standard equations of optimal estimation take care of the uncertainty propagation to the retrieval result (and are equivalent to the classic metrological formulation).

4.3 Grid-average clear-sky radiance

Assume that the set of pixels, p, lie within a latitude-longitude grid cell, and have been identified as clear sky radiances, of which there are n_p. The grid average clear-sky radiance is, for a given channel $f = \frac{1}{n_p}\sum_p y(p)$. The sensitivity of the average to each pixel's radiance is $\frac{1}{n_p}$, so the uncertainty in the average from radiance errors (neglecting the common uncertainty) is:

$$\frac{1}{n_p}\sqrt{\sum_p (u_i(p)^2 + u_s(p)^2) + 2\sum_{p>p\prime} u_s(p)u_s(p')\exp\left(-\frac{|e-e'|}{\Delta_e} - \frac{|l-l'|}{\Delta_l}\right)}$$