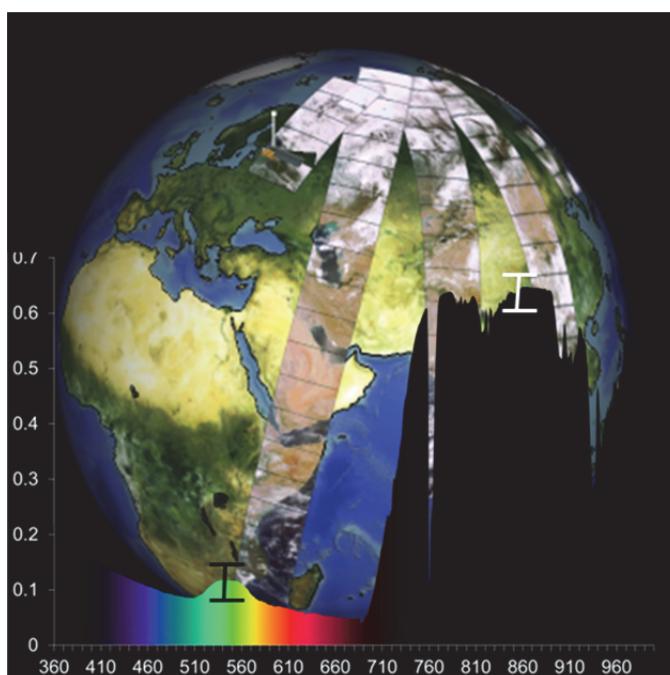


Intermediate

# Uncertainty Analysis

for **Earth Observation**

Instrument Calibration Module



## Training Course Textbook

Emma Woolliams

Andreas Hueni

Javier Gorroño



Title: Intermediate Uncertainty Analysis for Earth Observation  
(Instrument Calibration)

Reference: EMRP-ENV04-D5.2.2\_textbook

Version: 2

Date: January 2015

Dissemination level: PUBLIC

Authors: Emma Woolliams, National Physical Laboratory  
Andreas Hueni, University of Zurich  
Javier Gorroño, National Physical Laboratory

Approved by: Teresa Goodman, National Physical Laboratory

Keywords: Uncertainty analysis, Uncertainty budget, Radiometric Calibration, Imager Cal/Val

Contact: <http://www.emceoc.org/contact.html>

## About the EMRP

The European Metrology Research Programme (EMRP) is a metrology-focused European programme of coordinated R&D that facilitates closer integration of the national research programmes. The EMRP is jointly supported by the European Commission and the participating countries within the European Association of National Metrology Institutes (EURAMET e.V.). The EMRP will ensure collaboration between National Measurement Institutes, reducing duplication and increasing impact. The overall goal of the EMRP is to accelerate innovation and competitiveness in Europe whilst continuing to provide essential support to underpin the quality of our lives. See <http://www.emrponline.eu> for more information

## About the EMRP JRP “Metrology for Earth Observation and Climate” (METEOC)

This course has been produced for the MetEOC (European Metrology for Earth Observation and Climate) project, funded by the European Metrology Research Programme. MetEOC is developing new infrastructure and methods to allow higher, traceable, accuracy to be delivered to the European calibration and validation community.

This includes:

- Characterisation of the stray-light properties of an airborne hyperspectral imager using tuneable laser radiation.
- Development of a facility which allows ocean colour sensor calibrations to be performed at an NMI, at a customer site, in the field, in air and in a vacuum.
- Measurements to support the in-flight calibration of the aircraft based atmospheric limb sounding experiment GLORIA.
- The design of a set of Novel LED based (self-calibrating) radiometers enable autonomous low cost monitoring of test sites for post launch calibration and validation of satellites.
- Development of a strategy to enable ‘SI Traceability’ to be assessed and assigned to a radiation transfer model using a virtual simulation of a real, traceably calibrated, optical, geometric and mechanical, 3D target.

MetEOC was a three year project, ending in September 2014. The follow on project, MetEOC-2, began in September 2014 and minor edits to this textbook have been made within that project. For more information about the project visit: <http://www.emceoc.org/>



<b>1</b>	<b>INTRODUCTION .....</b>	<b>1</b>
<b>2</b>	<b>KEY CONCEPTS FOR UNCERTAINTY ANALYSIS.....</b>	<b>2</b>
2.1	QA4EO .....	2
2.2	THE ISO AND BIPM GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT (GUM).....	3
2.3	TRACEABILITY AND SI .....	3
2.4	CORRECTIONS, ERRORS AND UNCERTAINTIES .....	4
2.5	THE LAW OF PROPAGATION OF UNCERTAINTIES.....	5
2.6	CLASSIFICATIONS .....	6
2.6.1	<i>Random and Systematic Effects.....</i>	<i>6</i>
2.6.2	<i>Type A and Type B.....</i>	<i>8</i>
2.6.3	<i>Absolute and relative uncertainties.....</i>	<i>8</i>
2.7	WRITING ABOUT UNCERTAINTIES.....	9
<b>3</b>	<b>UNDERSTANDING THE LAW OF PROPAGATION OF UNCERTAINTIES.....</b>	<b>11</b>
3.1	SENSITIVITY COEFFICIENTS .....	11
3.2	ADDING IN QUADRATURE .....	16
3.3	TAKING AN AVERAGE OF INDEPENDENT MEASURED VALUES.....	17
3.3.1	<i>Applying the law of propagation of uncertainties .....</i>	<i>17</i>
3.3.2	<i>Averaging enough readings.....</i>	<i>19</i>
3.3.3	<i>Allan deviation/Allan variance.....</i>	<i>20</i>
3.4	TAKING AN AVERAGE OF PARTIALLY CORRELATED MEASURED VALUES .....	22
3.4.1	<i>Correlation.....</i>	<i>22</i>
3.4.2	<i>Analytical approach by modelling out the correlation: absolute uncertainties .....</i>	<i>22</i>
3.4.3	<i>Analytical approach by calculating covariance: absolute uncertainties.....</i>	<i>23</i>
3.4.4	<i>Analytical approach by modelling out the correlation: relative uncertainties.....</i>	<i>24</i>
3.5	COVARIANCE TERM IN THE LAW OF PROPAGATION OF UNCERTAINTIES .....	25
3.5.1	<i>Covariance.....</i>	<i>25</i>
3.5.2	<i>Covariance from an error model .....</i>	<i>26</i>
3.5.3	<i>Estimating the covariance from experimental and modelled data.....</i>	<i>27</i>
3.5.4	<i>Range of possibilities for correlation.....</i>	<i>28</i>
<b>4</b>	<b>THE STEPS TO AN UNCERTAINTY BUDGET .....</b>	<b>30</b>
4.1	STEPS.....	30
4.2	STEP 1: DESCRIBING THE TRACEABILITY CHAIN .....	30
4.3	STEP 2: WRITING DOWN THE CALCULATION EQUATIONS .....	32
4.4	STEP 3: CONSIDERING THE SOURCES OF UNCERTAINTY .....	33
4.5	STEPS 1 TO 3 FOR A SPACE-BORNE SENSOR.....	34
4.6	STEP 4: CREATING THE MEASUREMENT EQUATIONS .....	36
4.7	STEP 5: DETERMINING THE SENSITIVITY COEFFICIENTS.....	38
4.8	STEP 6: ASSIGNING UNCERTAINTIES.....	38
4.9	STEP 7: COMBINING AND PROPAGATING UNCERTAINTIES.....	40
4.10	STEP 8: EXPANDED UNCERTAINTIES.....	41
4.11	EXPERIMENTAL TECHNIQUES FOR TESTING AND VALIDATING AN UNCERTAINTY BUDGET .....	43
4.11.1	<i>Use a different reference.....</i>	<i>44</i>
4.11.2	<i>Use different instruments.....</i>	<i>44</i>
4.11.3	<i>Ask someone else to do the measurement.....</i>	<i>44</i>

4.11.4	<i>Change the order of the measurements</i>	45
4.11.5	<i>Use a different measurement method</i>	45
4.11.6	<i>Compare with someone else</i>	45
4.11.7	<i>Use a different artefact</i>	45
<b>5</b>	<b>CASE STUDY: APEX IMAGER CALIBRATION (SIMPLIFIED)</b>	<b>46</b>
5.1	THE APEX IMAGER CALIBRATION	46
5.2	STEP 1: DESCRIBING THE TRACEABILITY CHAIN	48
5.3	STEP 2: WRITING DOWN THE CALCULATION EQUATIONS	51
5.3.1	<i>For the calibration of the sphere</i>	51
5.3.2	<i>For the calibration of the APEX imager</i>	52
5.3.3	<i>For the user of the APEX imager to measure scene radiance</i>	52
5.4	STEP 3: CONSIDERING THE SOURCES OF UNCERTAINTY	52
5.4.1	<i>For the calibration of the SVC spectrometer against the RASTA</i>	53
5.4.2	<i>For the calibration of the sphere with the SVC spectrometer</i>	54
5.4.3	<i>For the combined source radiance from the sphere and filters</i>	55
5.4.4	<i>For the calibration of the APEX gain from the measurement of the sphere-filter source</i>	55
5.4.5	<i>For the observed scene radiance</i>	56
5.5	STEP 4: CREATING THE MEASUREMENT EQUATIONS	58
5.6	STEP 5: DETERMINING THE SENSITIVITY COEFFICIENTS	58
5.7	STEP 6: ASSIGNING UNCERTAINTIES	59
5.8	STEP 7: COMBINING AND PROPAGATING UNCERTAINTIES	61
5.9	STEP 8: EXPANDING UNCERTAINTIES	61
5.10	THIS IS JUST THE START OF THE PROCESS	61
<b>6</b>	<b>STRAIGHT LINE CALIBRATION EQUATIONS</b>	<b>63</b>
6.1	A STRAIGHT LINE CALIBRATION EQUATION	63
6.2	UNCERTAINTY ANALYSIS OVERVIEW	64
6.3	CALIBRATION DATA AND UNCERTAINTIES	66
6.4	DETERMINING THE FIT COEFFICIENTS (DOING THE FIT)	69
6.4.1	<i>Approaches to take</i>	69
6.4.2	<i>Unweighted fit and Monte Carlo</i>	69
6.4.3	<i>Weighted fit and analytical covariance</i>	72
6.4.4	<i>Rigorous analysis</i>	74
6.5	VALIDATING THE FIT	75
6.6	USING THE FIT	76
6.6.1	<i>Analytical covariance when the fit is used</i>	76
6.7	STRAIGHT LINE INTERPOLATIONS	77
<b>7</b>	<b>SPECTRAL SELECTION</b>	<b>80</b>
7.1	SPECTRAL RESPONSE FUNCTION	80
7.2	UNCERTAINTY ASSOCIATED WITH THE SPECTRAL RESPONSE FUNCTION	81
7.3	SPECTRAL INTEGRALS AND CONVOLUTION	82
7.3.1	<i>Origin of spectral integrals</i>	82
7.3.2	<i>Calculating the integral</i>	83
7.3.3	<i>The uncertainty associated with the integral</i>	84
7.4	STRAY LIGHT (OUT OF BAND)	89

7.5	SPECTRAL AND SPATIAL EFFECTS .....	90
<b>8</b>	<b>POST-LAUNCH CALIBRATION AND LEVEL-1 EO PRODUCTS RADIOMETRIC UNCERTAINTY .....</b>	<b>91</b>
8.1	IMAGER CHANGES IN ORBIT .....	91
8.2	ON-BOARD CALIBRATION SYSTEMS .....	92
8.3	VICARIOUS CAL/VAL .....	94
8.4	UNCERTAINTY ANALYSIS .....	95
8.5	EXAMPLE: PTFE DIFFUSER.....	95
8.5.1	<i>Use of the solar diffuser in orbit.....</i>	<i>96</i>
8.5.2	<i>Step 1: Describing the Traceability Chain.....</i>	<i>97</i>
8.5.3	<i>Step 2: Writing down the calculation equations.....</i>	<i>98</i>
8.5.4	<i>Step 3: Considering the sources of uncertainty.....</i>	<i>99</i>
8.5.5	<i>Step 4: Creating the measurement equation.....</i>	<i>104</i>
8.5.6	<i>Step 5: Determining the sensitivity coefficients.....</i>	<i>105</i>
8.5.7	<i>Step 6: Assigning uncertainties.....</i>	<i>106</i>
8.5.8	<i>Step 7: Combining and propagating uncertainties.....</i>	<i>107</i>
8.5.9	<i>Step 8: Expanding uncertainties.....</i>	<i>107</i>
8.6	LEVEL-1 EO PRODUCTS RADIOMETRIC UNCERTAINTY .....	107
8.6.1	<i>Radiance-to-reflectance conversion .....</i>	<i>108</i>
8.6.2	<i>Image orthorectification .....</i>	<i>109</i>
<b>9</b>	<b>CONCLUSIONS.....</b>	<b>113</b>
	<b>APPENDIX A MAKING AND USING COVARIANCE MATRICES .....</b>	<b>114</b>
A.1	HOW TO CREATE A COVARIANCE MATRIX .....	114
	<b>APPENDIX B APPENDIX: MONTE CARLO ANALYSIS .....</b>	<b>118</b>
	<b>APPENDIX C REFERENCES AND FURTHER READING .....</b>	<b>121</b>
C.1	FURTHER STUDY AND READING.....	121
C.2	FORMAL REFERENCES.....	121





## **1 Introduction**

This is the course textbook for the training course developed under the EMRP project MetEOC-1 as an intermediate level training course for uncertainty analysis with emphasis on radiometric instrument calibration and characterisation for Earth Observation.

The textbook does not follow the course lectures completely, but is designed as a stand-alone 'teach yourself' guide which is supplemented by the lectures. In particular the textbook includes topics that are not taught during the course (in particular straight line calibration equations and spectral effects) and does not include other topics that are taught in the course (most notably vicarious calibration).

## 2 Key concepts for uncertainty analysis

This intermediate-level course builds on some key ideas for uncertainty analysis that are provided in beginner level texts, for example the NPL Best Practice Guide number 11 by Stephanie Bell. This is freely downloadable at: <http://www.npl.co.uk/publications/guides/a-beginners-guide-to-uncertainty-of-measurement>. Also consider NPL's e-courses (see the Further Study section on page 121).

This section of this textbook emphasises some of the key ideas of uncertainty analysis that are assumed for the rest of the course.

### 2.1 QA4EO

The Quality Assurance Framework for Earth Observation (QA4EO) was written and endorsed by the Committee on Earth Observation Satellites (CEOS) and consists of a key principle, supported by a set of guidelines. The QA4EO Key Principle states:

*Data and derived products shall have associated with them a fully traceable indicator of their quality*

With the definitions:

#### **Quality Indicator**

*A Quality Indicator (QI) shall provide sufficient information to allow all users to readily evaluate the "fitness for purpose" of the data or derived product*

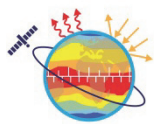
#### **Traceability**

*A QI shall be based on a documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards*

Its aim is to provide EO data users with sufficient (simple) information to enable them to evaluate the fitness-for-purpose of data/information for their applications through the assignment of Quality Indicators (QI) to data and derived products. QA4EO encourages the quantification of uncertainties and the documentation of supporting evidence.

There are seven guidelines written to aid the user in this process, providing examples, templates and details of the sort of information and evidence which should be made available. These guidelines are largely derived from existing best practice in both the EO sector and elsewhere, and are listed in Table 1.

This course text is fully compliant with QA4EO.



**Table 1 List of QA4EO Guidelines.**

QA4EO-QAEO-GEN-DQK-001	A guide to establish a Quality Indicator on a satellite sensor derived data product
QA4EO-QAEO-GEN-DQK-002	A guide to content of a documentary procedure to meet the Quality Assurance requirements of CEOS
QA4EO-QAEO-GEN-DQK-003	A guide to "reference standards" in support of Quality Assurance requirements of QA4EO
QA4EO-QAEO-GEN-DQK-004	A guide to comparisons: organisation, operation and analysis to establish measurement equivalence to underpin the Quality Assurance requirements of QA4EO
QA4EO-QAEO-GEN-DQK-005	A guide to establishing validated models, algorithms and software to underpin the Quality Assurance requirements of QA4EO
QA4EO-QAEO-GEN-DQK-006	A guide to expression of uncertainty of measurements
QA4EO-QAEO-GEN-DQK-007	A guide to establishing quantitative evidence of traceability to underpin the Quality Assurance requirements of QA4EO

## 2.2 The ISO and BIPM Guide to the Expression of Uncertainty in Measurement (GUM)

The *Guide to the Expression of Uncertainty in Measurement*, known as 'the GUM', provides guidance on how to determine, combine and express uncertainty [1]. It was developed by the JCGM (Joint Committee for Guides in Metrology), a joint committee of all the relevant standards organisations (e.g. ISO) and the BIPM (*Bureau International des Poids et Mesures*). This heritage gives the GUM authority and recognition. The JCGM continues to develop the GUM and has recently produced a number of supplements. All of these, as well as the 'VIM' (International Vocabulary of Metrology, [2]) are freely downloadable from the BIPM website<sup>1</sup>.

QA4EO Guideline 006 attempts to explain the main principles of the GUM to the EO community.

This course is fully compliant with the GUM.

## 2.3 Traceability and SI

Traceability is defined by the Committee for Earth Observation Satellites (CEOS) as the

*Property of a measurement result relating the result to a stated metrological reference (free definition and not necessarily SI) through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty.*

Traceability includes both an unbroken chain (i.e. it is calibrated against X, which was calibrated against Y, which was calibrated against Z, all the way back to SI, or, perhaps, a community reference) and the documentary evidence that each step was done in a reliable way (ideally audited, at least thoroughly peer-reviewed).

---

<sup>1</sup> <http://www.bipm.org/en/publications/guides/>

Traceability should, ideally, be to the International System of Units, known as the SI from its French name, *le Système international d'unités*. The SI units provide a coherent system of units of measurement built around seven base units and coherent derived units. A coherent system of units means that a quantity's value does not depend on how it was measured. The SI is an evolving system, with the responsibility for ensuring long term consistency with the General Conference on Weights and Measures (CGPM), run through the International Bureau of Weights and Measures, the BIPM, and maintained nationally through the National Metrology Institutes (NMIs). The Mutual Recognition Arrangement (MRA) signed in 1999 between the NMIs ensures that measurements made traceably to any NMI within the MRA are recognised by other NMIs. This is enforced by both formal international comparisons and a process of auditing and peer-reviewing statements of calibration capability. For the user, this means that traceability to SI can be achieved through any NMI within the MRA.

## 2.4 Corrections, errors and uncertainties

The terms 'error' and 'uncertainty' are not synonyms, although they are often confused. To understand the distinction, consider the result of a measurement – the measured value. The value will differ from the true value for several reasons, some of which we may know about. In these cases, we apply a **correction**. A correction is applied to a measured value to account for known differences, for example the measured value may be multiplied by a gain determined during the instrument's calibration, or a measured optical signal may have a dark reading subtracted. This correction will never be perfectly known and there will also be other effects that cannot be corrected, so after correction there will always be a residual, unknown **error** – an unknown difference between the measured value and the (unknown) true value.

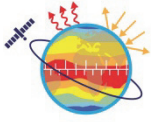
The specific error in the result of a particular measurement cannot be known, but we describe it as a draw from a probability distribution function. The **uncertainty** associated with the measured value is a measure of that probability distribution function; in particular, the **standard uncertainty** is the standard deviation of the probability distribution, and the equivalent of this for other distributions<sup>2</sup>.

There are generally several 'sources of uncertainty' that jointly contribute to the uncertainty associated with the measured value. These will include uncertainties associated with the way the measurement is set up, the values indicated by instruments, and residual uncertainties associated with corrections applied. The final (unknown) error on the measured value is drawn from the overall probability distribution described by the uncertainty associated with the measured value. This is built up from the probability distributions associated with all the different sources of uncertainty.

The use of the words 'error' and 'uncertainty' described here is consistent with paragraph 2.2.4 of the GUM. See also Section 2.6.1 and Figure 1.

---

<sup>2</sup> See the note box in section 0 .



## 2.5 The law of propagation of uncertainties

The aim of uncertainty analysis is to estimate the uncertainty associated with the measured value, which may be the result of a process involving several different parameters being controlled and set or measured, and a calculation. To obtain the final uncertainty, uncertainties due to each element in the process that affect the final result must be combined – i.e. they must be propagated through this process.

The GUM gives the Law of Propagation of Uncertainty:

$$u_c^2(y) = \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j), \quad (2.1)$$

which applies for a measurement model of the form

$$Y = f(X_1, X_2, X_3, \dots, X_i, \dots) \quad (2.2)$$

where an estimate  $x_i$  of quantity  $X_i$  has an associated uncertainty  $u(x_i)$ . The quantity  $u_c^2(y)$  is the squared standard uncertainty (standard deviation of the probability distribution) associated with the measured value  $y$  which comes from a combination of the uncertainties associated with all the different effects,  $x_i$ . The square of the standard uncertainty is also known as the **variance**.

The Law of Propagation of uncertainties is discussed in detail in Section 3. It can help to write it in terms of **sensitivity coefficients** as

$$u_c^2(y) = \sum_{i=1}^n c_i^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_i c_j u(x_i, x_j), \quad (2.3)$$

where the sensitivity coefficient  $c_i = \partial f / \partial x_i$ . The sensitivity coefficient is a 'translation' from one variable to another. It answers the question: "how sensitive is  $y$  to an uncertainty associated with  $x_i$ ?"

The law of propagation of uncertainties is written in this slightly complex notation of two parts to separate two terms:

- The first term is the sum of the squares of the standard uncertainties  $u(x_i)$  (the sum of the variances) associated with each individual effect multiplied by the relevant **sensitivity coefficient** (the partial derivative). This first term is what is meant by the description '**adding in quadrature**'.
- The second term deals with the **covariance** of correlated quantities. The covariance is a measure of how much the two quantities vary together. See also Section 3.5.

Note that the covariance term covers all pairs of different quantities, e.g.  $(x_1, x_2), (x_1, x_3), (x_2, x_3), \dots$ . Since the covariance  $u(x_1, x_2) = u(x_2, x_1)$ , the summation is

only over the combinations where  $i < j$  (i.e. only half the cases). The 2 in front of this term accounts for the opposite cases.

## 2.6 Classifications

### 2.6.1 Random and Systematic Effects

Correlation will be introduced whenever there is something in common between two measured values that will be combined (i.e. two values that will be averaged, or two quantities used in a measurement equation, or values at different wavelengths that will be combined through interpolation or integration). The simplest way to describe this is in terms of random and systematic effects.

**Random effects** are those that are not common to the multiple measurements being combined. A common example is noise: two measured values may both suffer from noise, but the effect of noise will be different for each of the two measured values (for example, if noise has increased one measured value, this provides no information about whether any other measured value is increased or decreased by that noise, nor by what extent).

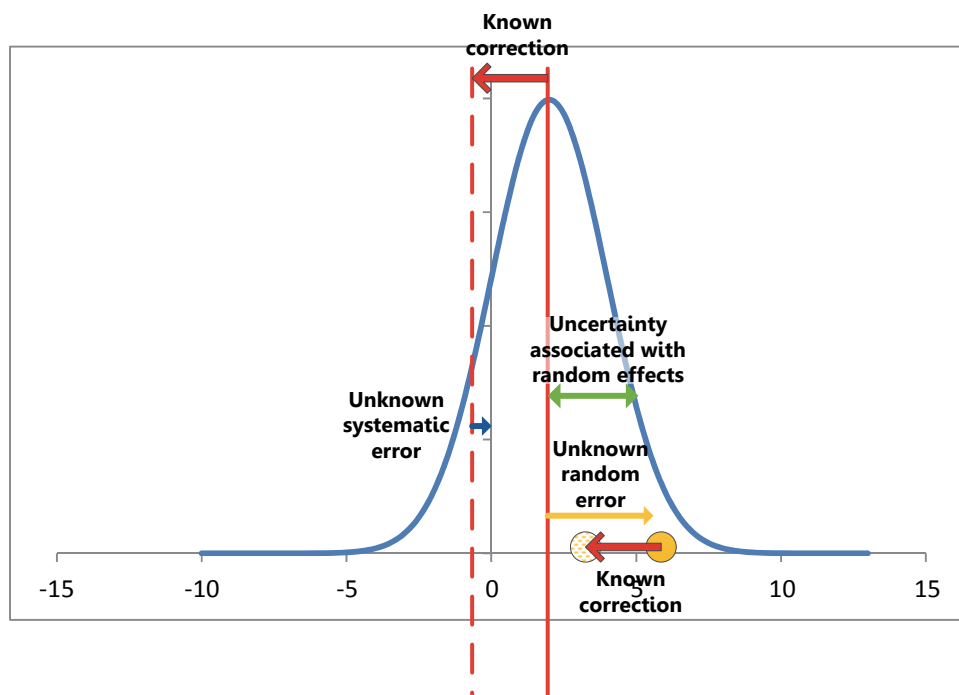
**Systematic effects** are those that are common to all measured values. If one measured value has been increased as a result of a systematic effect, then we can make a reliable prediction regarding whether any other measured value will be increased, and by how much. For example each time the distance is set for an irradiance measurement using a particular lamp, there will be a (normally small) error in that distance. This will equally affect all measurements of that lamp until the next alignment. If multiple measured values are averaged without realignment, or measured values at different wavelengths are combined in an integral, then the distance error will be common to all those measured values. This is a systematic effect.

Some effects, such as noise, are always random; other effects can be either random or systematic depending on the measurement process. For example, if three measured values of a lamp are combined in an average and the lamp is realigned between each measurement, then alignment/distance is a random effect. If the lamp is not realigned between measurements, then alignment/distance is a systematic effect.

The error in the measured value due to a random effect will change from one measured value to another. In this case the uncertainty associated with the effect may be the same for each measured value (the probability distribution for the effect is the same for each measured value), but each measured value is independent of each other measured value, as influenced by this effect. The unknown random error at each measured value is an independent draw from the probability distribution, meaning that the error due to the random effect is not only different from, but also independent of, the error at any other wavelength. The standard uncertainty associated with random effects is usually (but not always) determined by calculating the standard deviation of repeated measured values.

The error in the measured value due to a systematic effect will be the same from one measured value to another. The uncertainty associated with the effect is the same for each

measured value and the error is the same draw from the probability distribution for all measured values. The standard uncertainty associated with systematic effects cannot be determined by repeat measurements, unless the effect is intentionally altered between repeats (e.g. by realigning a source multiple times using a series of different 'extreme but acceptable' alignments<sup>3</sup> in an experiment to characterise the impact of source alignment).



**Figure 1: Representing a measurement where there is a known correction, an unknown systematic effect and random effects.**

Figure 1 represents a measurement process where there is a known correction, an unknown systematic effect and random effects. A measurement is made (obtaining the value represented by the golden circle at 6). We know of a correction – a systematic bias due to, e.g. a dark reading – and apply this correction, obtaining the value of the dotted circle – here about 3. There is still an unknown error from the true value of zero. If we make many measurements we get the probability distribution function shown in blue. The spread of this, the standard deviation of the normal distribution, is the standard uncertainty associated with random effects – those effects that change from measurement to measurement. Our measured value is a draw from this probability distribution function. If we take multiple measurements we obtain different draws. The average will tend towards the value at the peak of this distribution. When the known correction is applied, the result will be close to the true value, but differ from it by an unknown systematic error common to all the measured values. This comes from its own probability distribution function and all measured values have the same draw from that distribution (not shown in the figure, but this will take the

<sup>3</sup> See section 3.1.



form of a probability distribution centred at the true value with a standard deviation equal to the uncertainty associated with systematic effects).

## 2.6.2 Type A and Type B

The terms 'Type A' and 'Type B' are used with uncertainty analysis. This use comes from the GUM, which defines:

- **2.3.2 Type A evaluation (of uncertainty)** method of evaluation of uncertainty by the statistical analysis of series of observations
- **2.3.3 Type B evaluation (of uncertainty)** method of evaluation of uncertainty by means other than the statistical analysis of series of observations

Type A evaluation uses statistical methods to determine uncertainties. Commonly this means taking repeat measurements and determining the standard deviation of those measurements. This method can only treat uncertainties associated with random effects, for example the uncertainty associated with measurement noise.

Type B evaluation uses 'any other method' to determine the uncertainties. This can include estimates of systematic effects from previous experiments or the scientist's prior knowledge. It can also include random effects determined 'by any other method'. For example we may model room temperature by a random variable in the interval from 19 °C to 21 °C – the temperature range of the air-conditioning settings. Similarly, we may say that a voltmeter with 2 digits after the decimal place has an uncertainty associated with resolution of 0.005 V because we know the rounding range.

It is common to assume that 'Type A' evaluation is for random effects and 'Type B' evaluation is for systematic effects. This is generally, but not always, the case. For example, a 'Type A' method may be used to determine the uncertainty associated with alignment: a lamp may be realigned ten times and the standard deviation of those ten measurements used to determine an uncertainty associated with alignment<sup>4</sup>. In a later experimental set-up, measurements may be taken at multiple wavelengths and these combined in a spectral integral. For that integral, alignment is a systematic effect (the lamp is not realigned from wavelength to wavelength) even though the determination of the associated uncertainty was performed using 'Type A' methods. Similarly, the uncertainty associated with a random effect may be estimated from prior knowledge, or a measurement certificate, and thus by a 'Type B' method.

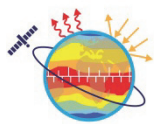
## 2.6.3 Absolute and relative uncertainties

The uncertainties given in the law of propagation of uncertainties by the symbol  $u(x_i)$  are always standard absolute uncertainties. The term **standard uncertainty** means that it is a single standard deviation of the probability distribution function associated with that

---

<sup>4</sup> If this is done, care must be taken to avoid 'double counting' any random effect due to, e.g. noise.





quantity. The term **absolute uncertainty** means that it has the same unit as the measurand. In other words, if the signal is in volts, the absolute uncertainty will also be in volts. If the distance is in metres, the absolute uncertainty will also be in metres.

It is common in radiometric calibrations to describe **relative uncertainties**, with units of per cent. The relative uncertainty is the absolute uncertainty divided by the quantity, i.e.  $u(x_i)/x_i$ .

## 2.7 Writing about uncertainties

In casual language we talk about 'averaging a set of measurements' or 'the uncertainty in the measurement is 0.5 %'. In metrology these words are defined carefully to reduce misunderstanding. We cannot 'average a set of measurements' but we can 'average the measured values' obtained from those measurements. The measurement has no uncertainty, there is an uncertainty *associated with* the measured value. For a non-specialist, such definitions can seem pedantic, as with jargon in all fields; but for a specialist, such careful use of words is a source of clarity. The words are defined through the VIM: the international vocabulary of metrology [2].

A *measurement* is made (instruments set up and value recorded) of a *measurand* (a quantity, such as radiance) to obtain a *measured value* (e.g.  $0.5 \text{ W m}^{-2} \text{ sr}^{-1} \text{ nm}^{-1}$ ) with an *associated uncertainty* (e.g. 0.5 %).

The VIM defines **measurement** as the

*process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity*

The most important word here is *process*: it defines measurement as the act of measuring. A measurement is not a quantity nor a result.

The VIM defines **measurand** as the

*quantity intended to be measured*

In turn, **quantity** is the

*property of a phenomenon, body or substance, where the property has a magnitude that can be expressed by a number and a reference.*

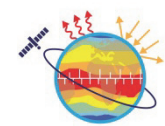
Thus quantities are things like length, mass, reflectance, irradiance, instrument gain, etc. When you measure a quantity, that quantity is the measurand of the measurement. The **measurement result** is defined by the VIM as the

*set of quantity values being attributed to a measurand together with any other available relevant information*

The "other available relevant information" refers to the associated uncertainty, perhaps expressed directly, perhaps as a probability density function, or perhaps implied by the number of digits provided with the result (the latter providing less reliable information). The **quantity value** is a

*number and reference together expressing magnitude of a quantity*

The reference usually means the unit. The **measured quantity value** (often shortened to measured value) is the quantity value that is the particular measurement result.



### 3 Understanding the Law of Propagation of Uncertainties

The Law of Propagation of Uncertainties is given in the GUM, and was provided above as Equation (2.1). This section describes some basic concepts behind this law and how to apply it. This section is reasonably theoretical. All the concepts here will be further explained in the examples and case studies of subsequent chapters.

#### 3.1 Sensitivity coefficients

Central to the law of propagation of uncertainties (Equation (2.3)) are the sensitivity coefficients, written as

$$c_i = \frac{\partial f}{\partial x_i} . \quad (3.1)$$

The sensitivity coefficient is a measure of how sensitive the measurand (the result),  $Y$ , calculated from the equation  $Y = f(X_1, X_2, X_3, \dots, X_i, \dots)$  is to the input quantity<sup>5</sup>  $X_i$ . In other words, it answers the question: "How much does this effect influence the final measured value?"

The uncertainty associated with  $Y$  due to  $X_i$  is

$$u_{y\_dueto\_x_i} = u_{y::x_i} = \frac{\partial y}{\partial x_i} u(x_i) . \quad (3.2)$$

So, for example, if we calculate a signal as a light reading minus a dark reading, we have the equation

$$V_S = V_{\text{light}} - V_{\text{dark}} . \quad (3.3)$$

The uncertainty associated with the signal due to the light reading is

$$u_{V_S\_dueto\_V_{\text{light}}} = u_{V_S::V_{\text{light}}} = \frac{\partial V_S}{\partial V_{\text{light}}} u(V_{\text{light}}) = 1 \times u(V_{\text{light}}) . \quad (3.4)$$

There are three methods for determining sensitivity coefficients and they are all equally valid and all approved by the GUM. These are:

- Mathematically (differentiating the measurement equations)
- Numerically (modelling through an instrument model in software, or changing the input parameters to the measurement equation)

---

<sup>5</sup> Note, as described in Section 2.5 that in formal mathematical notation of the Law of Propagation of Uncertainties, a capital letter is used to denote a quantity (measurand) and a small letter is used to denote a specific measured value of that measurand. This notation is not used in later stages of this book, where traditional physics notation is used, e.g. that radiance is represented by  $L$ .

- Experimentally (changing the effect in the lab and seeing how much the measured value changes)

In the development of a real-world uncertainty budget, all three will be used.

**Through differentiation (mathematically):** Where the measurement equation shows a straightforward relationship, then often the simplest method is to differentiate. This is the partial derivative term in the Law of Propagation of Uncertainties (in the GUM).

Consider the radiance of a white diffuser panel illuminated by an FEL<sup>6</sup> lamp. The radiance of the diffuser tile, viewed at an angle of 45° and for normal incidence illumination, is given by

$$L_s = \frac{E_{\text{FEL}} \beta_{0^\circ:45^\circ} d_{\text{cal}}^2}{\pi d_{\text{use}}^2} \quad (3.5)$$

where  $L_s$  is the source radiance,  $E_{\text{FEL}}$  is the lamp irradiance and  $\beta_{0^\circ:45^\circ}$  is the diffuser radiance factor. The calibration distance for the FEL lamp is  $d_{\text{cal}}$  and it is set a distance  $d_{\text{use}}$  from the diffuser. Note that here the wavelength dependence is not explicitly described, in practice  $L_s(\lambda)$ ,  $E_{\text{FEL}}(\lambda)$  and  $\beta_{0^\circ:45^\circ}(\lambda)$  are all functions of wavelength,  $\lambda$ .

Thus, the sensitivity coefficient of the diffuser radiance due to the irradiance of the lamp is

$$c_{E_{\text{FEL}}} = \frac{\partial L_s}{\partial E_{\text{FEL}}} \quad (3.6)$$

From Equation (3.5) we calculate this as

$$\begin{aligned} c_{E_{\text{FEL}}} &= \frac{\partial L_s}{\partial E_{\text{FEL}}} = \frac{\beta_{0^\circ:45^\circ} d_{\text{cal}}^2}{\pi d_{\text{use}}^2} \\ &= \frac{L_s}{E_{\text{FEL}}} \end{aligned} \quad (3.7)$$

The first line calculates the derivative. The second line shows that this can be expressed more simply. This is significant because the meaning of the sensitivity coefficient is that the uncertainty associated with the diffuser radiance, *due to the uncertainty associated with the lamp irradiance* is

$$u_{L_s:E_{\text{FEL}}} = c_{E_{\text{FEL}}} u_{E_{\text{FEL}}} \quad (3.8)$$

This means that the absolute uncertainty (units [ $\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1}$ ]) associated with the lamp-diffuser radiance *due to the lamp irradiance* is the sensitivity coefficient times the uncertainty associated with the lamp irradiance (which has units [ $\text{W m}^{-2} \text{nm}^{-1}$ ]). If the simpler, second line of (3.7) is used, this gives

---

<sup>6</sup> FEL is an ANSI standard designation denoting a specific 1 kW double-coiled tungsten halogen lamp, operating at 110 V, with a specific base. The letters are an arbitrary code and not an acronym.

$$u_{L_s:E_{FEL}} = \frac{L_s}{E_{FEL}} u_{E_{FEL}} \quad (3.9)$$

Rearranging this, gives

$$\frac{u_{L_s:E_{FEL}}}{L_s} = \frac{u_{E_{FEL}}}{E_{FEL}} \quad (3.10)$$

i.e. the relative uncertainty associated with radiance of the lamp-diffuser (usually with uncertainty expressed in %) *due to the lamp irradiance* is equal to the relative uncertainty associated with the FEL irradiance (also expressed in %).

A similar relationship can be found for all the other parameters. Note that for the distance of use (the lamp to diffuser distance)

$$\begin{aligned} c_{d_{use}} &= \frac{\partial L_s}{\partial d_{use}} = \frac{-2}{d_{use}^3} \frac{E_{FEL} \beta_{0^\circ:45^\circ} d_{cal}^2}{\pi} \\ &= \frac{-2L_s}{d_{use}}. \end{aligned} \quad (3.11)$$

And therefore

$$\frac{u_{L_s:d_{use}}}{L_s} = \frac{2u_{d_{use}}}{d_{use}} \quad (3.12)$$

The relative uncertainty associated with the radiance of the lamp-diffuser (in %) due to the lamp-diffuser distance is twice<sup>7</sup> the relative uncertainty associated with the lamp-diffuser distance (also in %).

**Through modelling (numerically):** Sometimes the most appropriate way to determine sensitivity coefficients is through modelling. This may be from a simple realisation of the measurement equation in a spreadsheet program, or it may be from a considerable piece of software that models the optical aberrations, pixel cross-talk, etc. for a multispectral imager. An example of the latter is given in the sections below.

A simple measurement equation or model may be used to determine the sensitivity coefficients numerically. This may be useful if the equation is difficult to differentiate analytically. Consider as a simple example, the equation (3.5). A spreadsheet can be set up to calculate the lamp-diffuser radiance from the input parameters. Then each parameter in turn can be altered by its uncertainty and the change in lamp-diffuser radiance recorded. It is important here that only one parameter is changed at a time. The aim is to determine the sensitivity of the calculated result to an uncertainty associated with a single parameter. If

---

<sup>7</sup> The minus sign in front of the 2 is ignored here because it will be squared. The negative sensitivity coefficients will be needed when correlation is taken into account.

multiple parameters are changed at the same time, then this would imply a correlation between them (e.g. that increasing one automatically increases the other). Changing one parameter at a time provides a sensitivity coefficient to that single parameter.

An alternative modelling approach, Monte Carlo simulation, is described in Appendix 9Appendix B. In Monte Carlo simulation all the parameters are altered simultaneously, but by a random amount, with the random number drawn from the probability distribution described by the uncertainty. Many (usually tens of thousands of) simulations are run; each simulation having different uncertainty draws. The advantages of Monte Carlo simulation are that correlations can be more easily dealt with, that statistical parameters can be determined from the end results and that the probability distribution functions do not have to be Gaussian (and therefore realistic probability distribution functions can be included). The main disadvantages are the computational time and the fact that individual sensitivity coefficients are not determined as everything is altered simultaneously. This means that the experimentalist does not easily obtain an understanding of what the most significant uncertainties are.

**Through laboratory testing (experimentally).** In some cases it is not possible to write the full measurement equation. For example, the sensitivity to a lamp alignment would have to be written as some function of the tilt and roll angles of the lamp, as well as the vertical and horizontal displacement of the lamp. Similarly, the lamp stability effect is a function of the time since calibration, whether the lamp has been transported, etc. These functions cannot be written out as explicit equations that can be differentiated or used in a mathematical model. Instead, the sensitivity is determined experimentally.

Experimental methods are based on repeat measurements where the effect is changed and the effect on the measured result is analysed. This is very good for situations where the effect can be easily controlled experimentally and there is no easy mathematical model relating the effect to the measurand.

The sensitivity to alignment, for example, may be determined by realigning the lamp ten times and comparing the standard deviation of those ten measured values with the standard deviation of ten measured values where the lamp is not realigned between measurements<sup>8</sup>. This would be an example of a Type A determination of an uncertainty (which, in the final measurements could be a systematic effect if the lamp is not realigned then, or a random effect, if the lamp is realigned).

Sometimes, a more systematic approach may be preferred. If the room temperature may vary from 19.5 °C to 20.5 °C, then it may be appropriate to determine an instrument's responsivity (e.g. spectrometer's gain) at 17 °C, 18 °C, 19 °C, 20 °C, 21 °C, 22 °C to check the sensitivity to

---

<sup>8</sup> The uncertainty associated with the alignment would then be calculated as  $u_{\text{align}} = \left( s_{\text{realigned}}^2 - s_{\text{not\_realigned}}^2 \right)^{1/2}$ , so as not to 'double count' the straightforward measurement repeatability. Note also that the measured values obtained through multiple measurements without realigning the lamp are correlated with each other because of the common alignment.

temperature and whether this is linear across a temperature range wider than that which is likely in practice. This can be done, for example, by wrapping the spectrometer with soft piping, through which water at each temperature in turn is sent. The spectrometer is set to view a stable source while its temperature is changed and the change in readings is determined.

Consider the case of estimating the sensitivity of the lamp irradiance to lamp current by experimentally varying the current. If we make too small a change in lamp current then we will not get a reasonable estimate of the effect for two reasons. First, we will not be able to control the current sufficiently well to get a clear separation of the two measured currents. For example if the current is stable to  $(8.100 \pm 0.004)$  A, then making a measurement at  $(8.104 \pm 0.004)$  A would cause too much overlap between the actually provided currents (in both situations a current of 8.102 A could naturally arise). For this reason it is sensible to change the current by at least ten times the measured effect.

Secondly, too small a change may mean any effect is hidden behind the noise due to other effects. For example, a change of 40 mA may affect the measured irradiance imperceptibly compared to the more dominant effect of detector noise. Thus, it is important that before such experimental determinations of a sensitivity coefficient are carried out, the natural repeatability of the measurements is determined. In this case it would be appropriate to make 5 consecutive measurements with a lamp current of 8.100 A, before repeating, again perhaps multiple times, with a lamp current of 8.140 A. If the change as a result of the change in current is much smaller than the noise caused by other effects, then there are several options, we could

- consider the uncertainty associated with stability in the lamp current to be an insignificant uncertainty component
- average several measurements at each of the currents used in the sensitivity coefficient investigation to reduce the impact of the noise on the measurements
- increase the current step used until an effect is seen, for example the effect may be smaller than the noise with a 40 mA change, but larger than the noise for a 100 mA change.

Any of these approaches may be valid.

As well as the risks of too small a current change, it is necessary to consider the risk of too large a current change. As well as the extreme examples, e.g. if the lamp current is increased too high the lamp filament may be destroyed or its calibration may be changed, a more modest and safe change may also provide unreliable results. The process described above: making a change ten times larger than that expected and dividing the observed change in the irradiance by ten to obtain a sensitivity coefficient, assumes that the process of change is linear. This is not the case with a lamp current: a doubling of the lamp current does not lead to a doubling of lamp irradiance. For small changes, of around 100 mA, linearity may be a reasonable approximation, but for larger changes, it is not.

It is therefore sensible to make measurements at a slightly smaller and a slightly larger current, thus providing three data points and giving a better understanding of how linear the lamp's behaviour is over the range. Similar considerations apply for all sensitivity coefficient examples. It is important to make the smallest possible change such that the observed effect can "come out of the noise", i.e. be seen as a real effect, but without changing the underlying behavioural relationship that is being investigated.

### 3.2 Adding in quadrature

From the sensitivity coefficients we determine the uncertainty associated with the measurand (the answer calculated from the measurement equation) due to each component in turn. The Law of Propagation of Uncertainties then combines these.

If there is no associated correlation (i.e. if all the components are independent of each other), then only the first half of Equation (2.3) needed

$$u_c^2(y) = \sum_{i=1}^n c_i^2 u^2(x_i) . \quad (3.13)$$

This adds uncertainties *in quadrature*. That means that the uncertainty associated with the measurand due to each component in turn is squared, they are summed and finally a square root is taken. The reason that uncertainties are added in quadrature is that it is statistically improbable that all the errors (i.e. all the draws from the probability distribution functions described by the uncertainties) are all at the extreme of the probability distribution function and in the same direction. It is more likely that some will increase and others decrease the measured value and that some errors will be smaller than their average value and others larger. Adding in quadrature provides a fair combination that is statistically robust<sup>9</sup>.

Consider two of the examples given in the previous section. First the signal calculated from a light and dark reading (equation (3.3)). The sensitivity coefficients are

$$\frac{\partial V_s}{\partial V_{\text{light}}} = 1; \quad \frac{\partial V_s}{\partial V_{\text{dark}}} = -1 . \quad (3.14)$$

Therefore,

$$u(V_s) = \sqrt{(1)^2 u^2(V_{\text{light}}) + (-1)^2 u^2(V_{\text{dark}})} = \sqrt{u^2(V_{\text{light}}) + u^2(V_{\text{dark}})} . \quad (3.15)$$

Note here that these are absolute uncertainties and all in the units of the measurand, e.g. volts (or digital numbers, or amperes).

---

<sup>9</sup> It is a principle of statistics that it is variances that are added in combining effects.



For the lamp-illuminated diffuser tile, equation (3.5), the calibration distance is assumed to have no associated uncertainty (it is given on the provided certificate), and the other terms have the relative sensitivity coefficients calculated as for equations (3.7) and (3.11), thus

$$\begin{aligned} \left( \frac{u(L)}{L} \right)^2 &= \left( \frac{u(E_{\text{FEL}})}{E_{\text{FEL}}} \right)^2 + \left( \frac{u(\beta_{0^\circ:45^\circ})}{\beta_{0^\circ:45^\circ}} \right)^2 + (-2)^2 \left( \frac{u(d_{\text{use}})}{d_{\text{use}}} \right)^2 \\ u_{\text{rel}}^2(L) &= u_{\text{rel}}^2(E_{\text{FEL}}) + u_{\text{rel}}^2(\beta_{0^\circ:45^\circ}) + 2^2 u_{\text{rel}}^2(d_{\text{use}}) \end{aligned} \quad (3.16)$$

The second line provides a version in terms of relative uncertainties, expressed in units of percent.

### 3.3 Taking an average of independent measured values

#### 3.3.1 Applying the law of propagation of uncertainties

Consider the mean,  $M$ , of three independent<sup>10</sup> measured values  $A$ ,  $B$  and  $C$ . This is calculated as

$$M = \frac{A + B + C}{3} \quad (3.17)$$

The sensitivity coefficients are therefore all

$$\frac{\partial M}{\partial A} = \frac{\partial M}{\partial B} = \frac{\partial M}{\partial C} = \frac{1}{3}. \quad (3.18)$$

We can also assume that these three measured values have the same associated uncertainty<sup>11</sup>

$$u(A) = u(B) = u(C) \equiv u(x). \quad (3.19)$$

Therefore, applying the law of propagation of uncertainties

$$\begin{aligned} u^2(M) &= \left( \frac{1}{3} \right)^2 u^2(A) + \left( \frac{1}{3} \right)^2 u^2(B) + \left( \frac{1}{3} \right)^2 u^2(C) \\ &= 3 \left( \frac{1}{3} \right)^2 u^2(x) \\ &= \left( \frac{u(x)}{\sqrt{3}} \right)^2. \end{aligned} \quad (3.20)$$

<sup>10</sup> Independent here means that the measured values are taken separately and have no associated correlation. If there are systematic effects, this should be considered separately (Section 3.4).

<sup>11</sup> If the three measurements are identical, it is a reasonable assumption that they have the same associated uncertainty. If they are not identical (taken in different ways) and there are known differences in the uncertainty associated with different methods, then it would be sensible to use a weighted mean, rather than a simple mean.

Hence, taking a mean of three readings reduces the uncertainty associated with a single reading by the square root of three.

Consider the signal calculated from the light and dark readings, but this time we assume  $N$  light readings and  $M$  dark readings. Thus

$$V_s = \frac{1}{N} \sum_{i=1}^N V_{\text{light},i} - \frac{1}{M} \sum_{j=1}^M V_{\text{dark},j} . \quad (3.21)$$

To apply the law of propagation of uncertainty to Eq. (3.21), we need to determine the sensitivity coefficients. These are

$$\begin{aligned} c_{V_{\text{light},i}} &= \frac{\partial V}{\partial V_{\text{light},i}} = \frac{1}{N}, \quad i = 1, \dots, N \\ c_{V_{\text{dark},j}} &= \frac{\partial V}{\partial V_{\text{dark},j}} = \frac{1}{M}, \quad j = 1, \dots, M. \end{aligned} \quad (3.22)$$

If we treat the  $N$  light readings as random draws from the same probability distribution function (i.e. they are independent realisations of the light reading), then we can say that for all readings,

$$u_{\text{light},i} = u_{\text{light}} ; \quad (3.23)$$

i.e. they have the same uncertainty. And similarly,

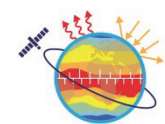
$$u_{\text{dark},j} = u_{\text{dark}} . \quad (3.24)$$

Applying the law of propagation of uncertainties to Eq. (3.21), we obtain

$$\begin{aligned} u_V^2 &= \sum_{i=1}^N \left( \frac{1}{N} \right)^2 u_{V_{\text{light},i}}^2 + \sum_{j=1}^M \left( \frac{1}{M} \right)^2 u_{V_{\text{dark},j}}^2 \\ u_V^2 &= N \left( \frac{1}{N} \right)^2 u_{V_{\text{light}}}^2 + M \left( \frac{1}{M} \right)^2 u_{V_{\text{dark}}}^2 \\ u_V^2 &= \left( \frac{u_{V_{\text{light}}}}{\sqrt{N}} \right)^2 + \left( \frac{u_{V_{\text{dark}}}}{\sqrt{M}} \right)^2 . \end{aligned} \quad (3.25)$$

In other words, in order to calculate the uncertainty associated with the measured signal, the uncertainty associated with a single light reading is divided by the square root of the number of light readings and added in quadrature with the uncertainty associated with a single dark reading divided by the square root of the number of dark readings.

Note first, that  $u_{V_{\text{light}}}$  and  $u_{V_{\text{dark}}}$  are absolute uncertainties, with units ([volts], or [Digital Numbers]). If the light signal has a value of 25000 DN  $\pm$  500 DN, and the dark signal has a value 1000 DN  $\pm$  10 DN, and assuming these uncertainties are already reduced by the square



root of the number of measurements averaged, then the signal is  $(25000 - 1000)$  DN  $\pm \sqrt{(500^2 + 10^2)}$  DN, or  $24000$  DN  $\pm 500.1$  DN. Expressed in relative terms, this uncertainty is  $500.1/24000 = 2.1\%$ .

### 3.3.2 Averaging enough readings

In order to apply the law of propagation of uncertainties, it is necessary to obtain a good estimate of  $u_{V_{\text{light}}}$  and  $u_{V_{\text{dark}}}$ . Assuming  $N$  and  $M$  are sufficiently large<sup>12</sup>, then the standard deviation of the  $N$  and  $M$  readings, gives a good estimate of  $u_{V_{\text{light}}}$  and  $u_{V_{\text{dark}}}$  respectively. Where, for example, 1000 readings are averaged using an automatic data acquisition process, then the standard deviation of those readings is a good estimate of the uncertainty associated with the individual measured values, due to random effects<sup>13</sup> (effects that vary from one reading to the next) because there are enough readings.

Sometimes, however,  $N$  and  $M$  are relatively small. This is true wherever a manual process is involved (for example where a lamp is realigned from measurement to measurement), or when a measurement is slow. When  $N$  and  $M$  are relatively small, then the standard deviation calculated from the measured values provides an unreliable estimate of the uncertainty associated with each individual measured value. The GUM deals with this through the Welch-Satterthwaite Equation (see Section 4.10). The planned revision to the GUM will, however, treat this in a different way, that is, perhaps more helpful for an intuitive understanding of whether sufficient measurements have been taken. In the GUM revision, the squared standard uncertainty associated with the mean of the light readings would be calculated from the standard deviation,  $s_{\text{light}}$ , using

$$u_{\text{light,mean}}^2 = \frac{N-1}{N-3} \left( \frac{s_{\text{light}}}{\sqrt{N}} \right)^2. \quad (3.26)$$

This equation estimates the standard variance (squared uncertainty) for the full distribution based on a squared standard deviation calculated from the  $N$  measured values. It increases it 'a bit' because a few points are may underestimate the standard deviation. Note that the  $(1/\sqrt{N})$  term here is the same term as in Eq. (3.25).

Clearly, the larger the value of  $N$ , the closer the uncertainty associated with the mean is to the standard deviation of the measurements divided by  $\sqrt{N}$ . If  $N=5$ , then the standard

---

<sup>12</sup> But not so large that drift dominates, see the discussion on Allan Variance in Section 3.3.3

<sup>13</sup> Note that 1000 readings taken in a short time period will have a standard deviation that depends only on short-term random noise. Any longer term fluctuations will not be included, nor will any effects that are common to all those 1000 readings, such as those due to e.g. current settings on a lamp used, or the alignment of a diffuser panel. It may be that even where 1000 readings are taken simultaneously, there will need to be a smaller number (say 5-10) separate averages of those 1000 readings taken over a longer time interval, or with lamps turned off and back on, or with the diffuser panel realigned. There is almost always a set of measurements whose standard deviation is determined for a very small number of measured values.

deviation is multiplied by  $\sim 1.41$ , if  $N = 10$ , then it is multiplied by  $\sim 1.13$ , and if  $N$  is 25, then it is multiplied by  $\sim 1.04$ . If  $N$  is 1000, then the standard deviation is multiplied by 1.001.

It is always preferable to use the best possible estimate of the standard deviation, from as many measurements as possible, so that the correction of Eq. (3.26) is as small as possible. It may not be practical, however, to make 25 or more measurements on every occasion. Under those circumstances, there are several things that could be done to obtain a better estimate of the standard deviation, either individually or combined:

- In the case of a spectral measurement, data from adjacent wavelength points can provide a better estimate. First, determine a standard deviation on a wavelength-by-wavelength basis. This will have some structure because of the random noise. Smooth out this structure using the data from neighbouring wavelengths. If  $N = 5$  at a single wavelength, then using data from 5 neighbouring wavelengths to 'smooth out' the standard deviation will effectively increase  $N$  to something closer<sup>14</sup> to 25.
- During a 'commissioning phase' for your instrumentation take 25 readings. Compare the standard deviation of the 5 readings taken on a single day to that of the 25 readings taken during the commissioning phase to check that it's still valid, but use the standard deviation of the 25 commissioning readings as your best estimate of the uncertainty of a single reading.
- Determine the differences of the 5 readings to the average (mean) of those five readings. Do this on 5 successive measurement days. Take the standard deviation of the differences for all five days<sup>15</sup>. Using the difference from the daily mean takes out the day-to-day variation and looks only at the variability.

Any of these options would generally be preferable to increasing the uncertainty using Equation (3.26). Note that these will give you an estimate of the standard uncertainty associated with a single reading. The uncertainty associated with the mean of today's 5 readings will be the uncertainty associated with a single reading divided by the square root of 5 (the number of measurements taken today).

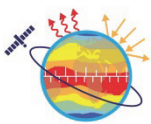
### 3.3.3 Allan deviation/Allan variance

The discussion above assumes that it is always preferable to average more and more measurements. Equation (3.25) reduces the uncertainty by the square root of the number of measurements taken. Consider an extreme example – if measurements are taken every minute for a week, would this mean the uncertainty was effectively zero? Intuitively we understand that this is not the case and there are two reasons for this. The first is that there may be systematic effects common to all the measurements. The second is because the instruments are likely to *drift* over the week. There comes a time when the noise is no longer

---

<sup>14</sup> How much closer depends on the smoothing algorithm used. Do not worry too much about the exact value of  $N$ , the point here is to make it 'sufficiently large' to give a good estimate and for the correction of Eq. (3.26) to be small.

<sup>15</sup> Again  $N = 25$  in this example.

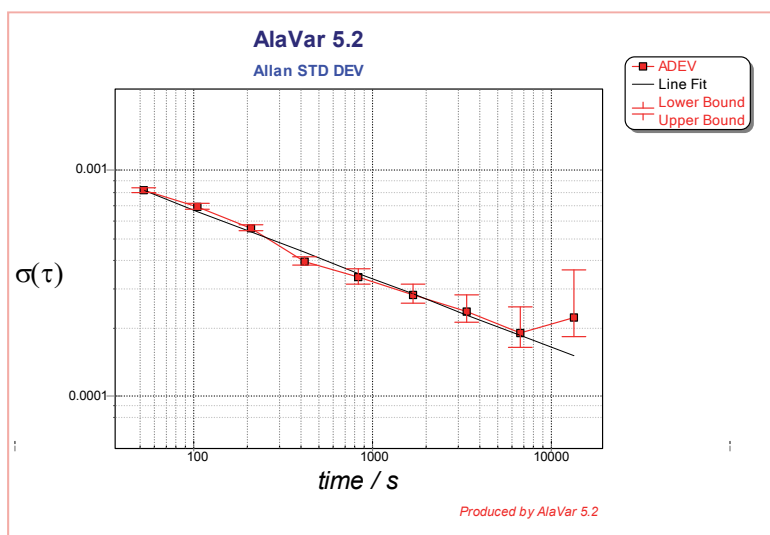


'white noise' that can be averaged, but some form of drift, the effect of which just gets worse with averaging over a longer and longer period.

The Allan Variance [3] is a means of determining whether the noise is 'white noise' or drift (or something else). Calculating the Allan Variance requires a series of measurements at regular time intervals  $\Delta t$ . The Allan Variance calculates the difference between successive data points, squares them and takes the average, before dividing the answer by two:

$$\sigma^2(\Delta t) = \frac{1}{2} \left[ \frac{1}{N/2} \left[ (y_2 - y_1)^2 + (y_4 - y_3)^2 + \dots + (y_N - y_{N-1})^2 \right] \right] \quad (3.27)$$

The pairs of data are then averaged, so now you have half as many data points with a time interval of  $2\Delta t$  and the Allan Variance is calculated again. This is repeated, each time averaging the pairs that were used before. After this has been completed, the data is plotted with the vertical axis giving the Allan Variance and the horizontal axis giving the time interval. Both axes should be represented on a log scale. For white noise, the Allan deviation lies on a straight line on this log-log plot, with a slope of -0.5, corresponding to a reduction in the uncertainty according to the square root of time. However, if the signals are accompanied by  $1/f$  noise or drift (random walk), the slope will become positive. This suggests that averaging for times greater than the Allan Deviation minimum *increases* rather than decreases uncertainty.



**Figure 3 Output plot from the Allan Variance software showing the Allan Variance of a very stable signal.**

Free software for evaluating the Allan Variance is available<sup>16</sup>. This provides plots as shown in Figure 3. In this example, the signal is stable with white noise for up to ~6000 s, after which drift dominates and further averaging makes the situation worse. From this graph we can see that if we average for 100 s the uncertainty associated with the mean is ~0.07 %, if we average for 1000 s, the uncertainty associated with the mean is ~0.03 %.

<sup>16</sup> <http://www.alamath.com/>

## 3.4 Taking an average of partially correlated measured values

### 3.4.1 Correlation

Correlation is introduced whenever there is something in common between multiple measured values due to a common effect in the measurement process. When there is something in common, then that correlation needs to be taken into account. There are two ways of dealing with that correlation in the law of propagation of uncertainties. The first is to describe the correlation explicitly in the measurement equation; the second is to use the second half of the law of propagation of uncertainties. Both are described here.

### 3.4.2 Analytical approach by modelling out the correlation: absolute uncertainties

Consider the situation where a lamp irradiance is determined by averaging three measured values of its irradiance. The measured values are taken sequentially and the lamp is not realigned between measurements. We can model the result of the  $i$ th measurement as

$$E_i = E_T + R_i + S \quad (3.28)$$

Here,  $E_T$ , is the (unknown) true irradiance of the lamp.  $R_i$  is the random error in the  $i$ th measurement. It is a draw from the probability distribution described by the uncertainty  $u(R)$ ; the uncertainty associated with the random effects. Similarly,  $S$  is the systematic error in all the measurements. It is a draw from the probability distribution described by the uncertainty  $u(S)$ ; the uncertainty associated with systematic effects. The expectation values of  $R$  and  $S$  are zero (we are not applying a correction). However there is an uncertainty associated with the measured value due to these effects.

The  $R$  error comes from all random effects – probably predominantly due to measurement noise, but also source stability, temperature fluctuations etc. The  $S$  comes from all systematic effects that didn't change between measurements; for example, alignment or the calibration of the reference lamp. If these uncertainties are determined separately, then the uncertainty associated with  $S$  can be calculated as the quadrature sum of the individual uncertainties.

The mean of the three measured values is

$$E_M = (E_1 + E_2 + E_3)/3 \quad (3.29)$$

which we can write, by applying (3.28) as

$$E_M = E_T + \frac{R_1 + R_2 + R_3}{3} + S \quad (3.30)$$

We now have an equation with four independent variables:  $R_1, R_2, R_3, S$  and, since these are independent, i.e. uncorrelated, we can apply the first half of the law of propagation of uncertainties

$$u^2(E_M) = \left( \frac{u(R)}{\sqrt{3}} \right)^2 + u^2(S) . \quad (3.31)$$

The uncertainty associated with random effects is reduced by the square root of the number of measurements; the uncertainty associated with systematic effects remains unchanged by the averaging. This is as we may intuitively expect. No matter how many measurements we make, we cannot reduce the uncertainty associated with systematic effects by averaging.

### 3.4.3 Analytical approach by calculating covariance: absolute uncertainties

In order to understand the second part of the law of propagation of uncertainties it is worth deriving Equation (3.31) using the full law of propagation of uncertainties, Equation (2.1). To do this, we need to determine the covariance  $u(E_i, E_j)$  associated with a pair of measured values, in this case a pair of irradiance values. The covariance is a measure of the uncertainty common to the two measured values. And with the model Equation (3.28) this is  $u(S)$ . Therefore

$$u(E_i, E_j) = u^2(S); \quad i \neq j . \quad (3.32)$$

Thus, including the sensitivity coefficients from (3.18), the term used in the second part of Equation (2.1) is

$$\frac{\partial E_M}{\partial E_i} \frac{\partial E_M}{\partial E_j} u(E_i, E_j) = \frac{1}{3} \times \frac{1}{3} u^2(S) . \quad (3.33)$$

The uncertainty associated with any individual irradiance value is given by

$$u^2(E_i) = u^2(R_i) + u^2(S) . \quad (3.34)$$

The law of propagation of uncertainties in full for this example is

$$\begin{aligned} u^2(E_M) = & \left( \frac{1}{3} \right)^2 u^2(E_1) + \left( \frac{1}{3} \right)^2 u^2(E_2) + \left( \frac{1}{3} \right)^2 u^2(E_3) + \\ & 2 \left( \frac{1}{3} \right)^2 u(E_1, E_2) + 2 \left( \frac{1}{3} \right)^2 u(E_1, E_3) + 2 \left( \frac{1}{3} \right)^2 u(E_2, E_3) . \end{aligned} \quad (3.35)$$

Combining (3.33), (3.34) and (3.35) gives

$$\begin{aligned}
 u^2(E_M) &= 3\left(\frac{1}{3}\right)^2 u^2(R) + 3\left(\frac{1}{3}\right)^2 u^2(S) + 6\left(\frac{1}{3}\right)^2 u^2(S) \\
 &= \left(\frac{u(R)}{\sqrt{3}}\right)^2 + 9\left(\frac{1}{3}\right)^2 u^2(S) \\
 &= \left(\frac{u(R)}{\sqrt{3}}\right)^2 + u^2(S).
 \end{aligned} \tag{3.36}$$

Which, as expected, is the same as Equation (3.31).

### 3.4.4 Analytical approach by modelling out the correlation: relative uncertainties

The example given above assumes that the systematic error is an additive effect, with the error and uncertainty having the same units as the measurand (i.e.  $[W \text{ m}^{-2} \text{ nm}^{-1}]$ ). In radiometric measurements uncertainties are more likely to be relative uncertainties in % and errors have a multiplicative effect. Therefore instead of Equation (3.28), the error model is better described by

$$E_i = E_T (1 + R_i)(1 + S) \tag{3.37}$$

where the error terms  $R_i$  and  $S$  have an expectation value of unity (one), and a relative uncertainty associated with that. In this case the mean is

$$E_M = E_T (1 + S) \left( \frac{3 + R_1 + R_2 + R_3}{3} \right). \tag{3.38}$$

Once again, we have an equation with four independent variables, so we can use the first part only of the law of propagation of uncertainties. The sensitivity coefficients are

$$\begin{aligned}
 \frac{\partial E_M}{\partial S} &= \frac{E_M}{(1 + S)} \\
 \frac{\partial E_M}{\partial R_i} &= \frac{E_T (1 + S)}{3} = \frac{E_M}{3 + R_1 + R_2 + R_3}
 \end{aligned} \tag{3.39}$$

and the law of propagation of uncertainties gives

$$u^2(E_M) = \left( \frac{E_M}{3 + R_1 + R_2 + R_3} \right)^2 (u^2(R_1) + u^2(R_2) + u^2(R_3)) + \left( \frac{E_M}{1 + S} \right)^2 u^2(S). \tag{3.40}$$

And therefore, making the reasonable assumption that  $u(R_1) = u(R_2) = u(R_3) = u(R)$ ,

$$\left( \frac{u(E_M)}{E_M} \right)^2 = \frac{3u^2(R)}{(3 + R_1 + R_2 + R_3)^2} + \left( \frac{u(S)}{1 + S} \right)^2 \tag{3.41}$$



To take this a step further, we have to understand that if the expected value of  $R_i = 0$ , then the term  $R_1 + R_2 + R_3 = 0$ . In taking the average, that is the assumption that we are making. Similarly<sup>17</sup>,  $S = 0$ ; thus

$$\left( \frac{u(E_M)}{E_M} \right)^2 = \left( \frac{u(R)}{\sqrt{3}} \right)^2 + u^2(S) \quad (3.42)$$

i.e. for a multiplicative model with relative uncertainties, the relative uncertainties behave exactly as the absolute uncertainties for an additive model in (3.31).

## 3.5 Covariance term in the law of propagation of uncertainties

### 3.5.1 Covariance

In order to apply the law of propagation of uncertainties, we need to deal with correlation in one of the following manners:

- By writing the correlation explicitly into an error model and rearranging the measurement equation so that the quantities are no longer correlated, this is the approach described in Sections 3.4.2 and 3.4.4
- By calculating the covariance from an error model. This is the approach described in Section 3.4.3 above and 3.5.2
- By determining the correlation experimentally or numerically as described in Section 3.5.3
- By obtaining a range of possibilities for the covariance as described in Section 3.5.4

Note that as with Type A and Type B evaluations of uncertainty, the first and second of these are a Type B evaluation of covariance – the covariance is estimated by an explicit measurement model. The third method is a Type A evaluation of uncertainty – statistical method are applied to the data itself to estimate the covariance. The final method does not estimate the covariance, but it does give the opportunity to consider it ‘negligible’ or ‘significant’ and to decide whether further analysis is required.

In Section 3.4 the emphasis was on averaging similar measured values (e.g. combining multiple values for the lamp irradiance). Here correlation related to whether those individual measured values were obtained under similar or different conditions, e.g. was the lamp realigned between measurements (if so, lamp alignment was a random effect with no associated correlation, if not, then lamp alignment is a systematic effect which introduces a correlation).

---

<sup>17</sup> This is the assumption we are making, that our best estimate of the systematic effect is that there isn’t one. If we had a better estimate, we would remove it with a correction. After all corrections have been applied, there is some unknown systematic error that is as likely to be positive as negative. We take it as having an expected value of 0 with an uncertainty associated with that of  $u(S)$ .

It is also important to estimate correlation between measured values from different parameters. For example, consider a spectral integral (from for example the convolution of a spectral response function and a scene radiance). Here values from different wavelengths are combined. Some components in the uncertainty budget (sources of uncertainty) will be common from one wavelength to the next, and other components will vary from wavelength to wavelength.

Correlation can also occur between effects that are different. For example if a spectrometer and a filter are both temperature sensitive, then the responsivity of the spectrometer may be correlated with the filter transmittance through the laboratory temperature. Sometimes such correlations can be explicitly described in a measurement equation (e.g. by including a term that is a function of temperature in both places), and then the Type B methods can be used for these examples too. Sometimes such a correlation can only be estimated by Type A methods, from a statistical analysis of the measurement data itself.

### 3.5.2 Covariance from an error model

The covariance can be calculated from an error model by seeing what terms are common to e.g. equations being combined or measured values being averaged. Sometimes the correlation can be written explicitly so that the terms can be treated as independent, as in Sections 3.4.2 and 3.4.4. There are occasions, however, where it is essential to describe the covariance analytically and in the form of a covariance matrix (See Appendix A). A covariance matrix is needed where data is manipulated through a modelling process, or, for example, a least squares algorithm is used. In these cases textbooks can provide information on how to manipulate a covariance matrix in, for example, the least squares algorithm. What is missing is how to form a covariance matrix from a typical experimental uncertainty budget in the first place.

To form such a covariance matrix it is necessary to form an error equation, similar to that of Equation (3.28). In such an equation,  $R_i$  represents the unknown random error which is an unknown draw from the probability distribution described by the uncertainty associated with (a combination of) random effects. To obtain that uncertainty it is necessary to add in quadrature the uncertainties associated with individual random effects (e.g. noise, source stability, etc). Similarly,  $S$  represents the unknown systematic error which is an unknown draw from the probability distribution described by the uncertainty associated with (a combination of) systematic effects. To obtain that uncertainty it is necessary to add in quadrature the uncertainties associated with individual systematic effects (e.g. system alignment, calibration certificate values etc).

For the error model of the form of (3.28) the covariance is given by (3.32). In this case the uncertainties are in the same units as the measured value. Where there is a relative model and uncertainties in percent, the covariance has to include the actual value. Thus for the relative error model (3.37) the covariance is

$$u(E_i, E_j) = u(S) E_i u(S) E_j = u^2(S) E_i E_j; \quad i \neq j. \quad (3.43)$$

i.e. the covariance is the squared relative uncertainty that is common (and in percent) multiplied by the irradiance values at the two values that are being combined.

Error models can be made as complex as required to describe the correlation between measured values. Consider, for example the more complex error model

$$E_i = E_T (1 + R_i)(1 + S)(1 + c_i F) + r_i + s \quad (3.44)$$

which has a random multiplicative error,  $R_i$ , a systematic multiplicative error,  $S$ , and a systematic multiplicate error,  $F$ , that has a measurement dependent sensitivity coefficient<sup>18</sup>  $c_i$ , as well as two additive errors – one random,  $r_i$  and one systematic,  $s$ . In this case the covariance is given by

$$u(E_i, E_j) = u^2(S) E_i E_j + u^2(F) c_i c_j E_i E_j + u^2(s); \quad i \neq j \quad (3.45)$$

The random effects are independent and do not contribute to the covariance. The relative uncertainties associated with multiplicative systematic effects are multiplied by the measured value to convert them to absolute uncertainties, and where appropriate the sensitivity coefficient. The absolute uncertainties associated with additive systematic effects are used directly.

### 3.5.3 Estimating the covariance from experimental and modelled data

In some situations it is appropriate to estimate the covariance from data. This would be the case when you assume there must be some correlation between separate variables. For example, both an instrument's responsivity and a filter's transmittance may be sensitive to room temperature, and therefore the two parameters may be correlated if they were determined at the same time. In this case, the covariance can be calculated statistically from the pairs of data.

The correlation coefficient,  $r(x, y)$ , is calculated from

$$r(x, y) = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{s_X} \right) \left( \frac{Y_i - \bar{Y}}{s_Y} \right) \quad (3.46)$$

where  $n$  is the number of data pairs  $(X_i, Y_i)$ ,  $\bar{X}$  is the mean of the  $x$  values and  $\bar{Y}$  is the mean of the  $y$  values and  $s_X, s_Y$  are the standard deviations of the  $x$  values and  $y$  values respectively. The covariance is then calculated as

$$u(x, y) = u(x)u(y)r(x, y) \quad (3.47)$$

---

<sup>18</sup> For example, if the different values are at different wavelengths, this could correspond to something like lamp current which has an associated uncertainty and an error in lamp current will affect all wavelengths in the same direction, but will have a bigger impact on the short wavelengths than on the long wavelengths.

where, as usual, the uncertainties correspond to the absolute standard uncertainties (in the same units as the measured values – if relative standard uncertainties are known, these should be multiplied by the average measured values).

This equation can also be used to determine the covariance of a pair of data points obtained through Monte Carlo simulation, see Appendix B.

### 3.5.4 Range of possibilities for correlation

Generally speaking, there are three extreme situations:

- Entirely uncorrelated data (the term  $r(x, y)$  in Eq. (3.47) is 0)
- Entirely correlated data (the term  $r(x, y)$  in Eq. (3.47) is +1)
- Entirely correlated data (the term  $r(x, y)$  in Eq. (3.47) is -1)

If it is not clear what the correlation coefficient is, it is appropriate to consider the extreme cases and say that the uncertainty lies within the range of those extremes. Where the sensitivity coefficients for both parameters are positive (i.e. a positive error in each parameter means a positive error in the result calculated from the measurement equation), or both negative (i.e. a negative error in each parameter means a positive error in the result), then correlation will always increase the uncertainty relative to uncorrelated data. Where the sensitivity coefficients are opposite signs (one is positive and the other negative), then the correlation will decrease the uncertainty relative to uncorrelated data.

Consider the very simple equation

$$x = x_{\text{meas}} + x_{\text{offset}} \quad (3.48)$$

which can represent, for example, a distance between two apertures being a measured distance plus an offset (e.g. for the thickness of the aperture if the measurement is to the back surface). Here the sensitivity coefficient is +1 for both terms ( $\partial x / \partial x_{\text{meas}} = +1$ ). What this means is that correlation between the two values represents 'making the same error twice' and it will increase the associated uncertainty. Therefore the 'worst case scenario' is for the two terms to be entirely correlated and the 'best case scenario' is for them to be entirely uncorrelated. For this equation, the full law of propagation of uncertainties is

$$\begin{aligned} u_c^2(y) &= \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \\ &= u^2(x_{\text{meas}}) + u^2(x_{\text{offset}}) + 2u(x_{\text{meas}}, x_{\text{offset}}). \end{aligned} \quad (3.49)$$

Note that,  $u(x_{\text{meas}}, x_{\text{offset}})$  is given by Equation (3.47), and since here we are assuming  $r(x_{\text{meas}}, x_{\text{offset}}) = 1$ ,  $u(x_{\text{meas}}, x_{\text{offset}}) = u(x_{\text{meas}})u(x_{\text{offset}})$ .

On the other hand, consider the simple equation

$$V_{\text{signal}} = V_{\text{light}} - V_{\text{dark}} \quad (3.50)$$

i.e. that a signal is a light reading minus a dark reading. Here, the sensitivity coefficient is +1 for the light reading and -1 for the dark reading. What this means is that correlation between the two values will be 'cancelled out' and it will decrease the associated uncertainty. This makes sense, because if, for example there was a common offset error in both the light and dark readings, subtracting the dark reading will remove the error. Thus, the full law of propagation of uncertainties is

$$\begin{aligned} u_c^2(y) &= \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \\ &= u^2(V_{\text{light}}) + u^2(V_{\text{dark}}) - 2u(V_{\text{light}}, V_{\text{dark}}), \end{aligned} \quad (3.51)$$

where the minus sign comes from the negative sensitivity coefficient. Note that,  $u(x_{\text{light}}, x_{\text{dark}})$  is given by Equation (3.47), and since here we are assuming  $r(x_{\text{light}}, x_{\text{dark}}) = 1$ ,  $u(x_{\text{light}}, x_{\text{dark}}) = u(x_{\text{light}})u(x_{\text{dark}})$ .

## 4 The steps to an uncertainty budget

### 4.1 Steps

This course considers eight steps to an uncertainty budget. Other writers have suggested somewhat different steps; however, what is common is a general framework around three areas:

- The first few steps relate to understanding the problem
- The second set of steps relate to determining the formal relationships – the measurement equation and sensitivity coefficients
- The final steps relate to 'doing the mathematics' and propagating the uncertainties

Most uncertainty textbooks and training emphasise the final steps – the Law of Propagation of Uncertainties, and how to 'do the mathematics' in calculating sensitivity coefficients and propagating uncertainties.

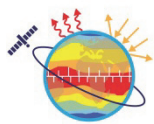
There is often very little guidance on how to do the early stages, even though they are often both the most challenging and, arguably, the most important part of uncertainty analysis. This is because the techniques are usually very specific to individual measurement fields. Fortunately, the skills required for the first step are the skills that an experimentalist who understands his or her measurement facility will naturally have. This course attempts to provide examples, questions and ways of approaching a problem that each participant can apply to his or her own laboratory.

The steps are introduced here, and then applied to case studies in the subsequent sections of this textbook.

- Understanding the problem
  - Step 1: Describing the traceability chain
  - Step 2: Writing down the calculation equations
  - Step 3: Considering the sources of uncertainty
- Determining the formal relationships
  - Step 4: Creating the measurement equation
  - Step 5: Determining the sensitivity coefficients
  - Step 6: Assigning uncertainties
- Propagating the uncertainties
  - Step 7: Combining and propagating uncertainties
  - Step 8: Expanding uncertainties

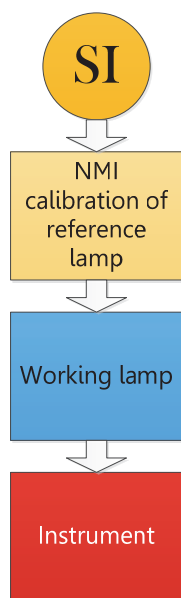
### 4.2 Step 1: Describing the traceability chain

The purpose of the first three steps is to get a very clear understanding of the uncertainty 'problem'. The emphasis should be on obtaining the overview and seeing where the sources of uncertainty come from. It is easy to become overwhelmed at this stage and to get worried about whether an uncertainty is being double counted, what might be being missed out (the



unknown-unknowns), being concerned that it's not clear which step an uncertainty component comes in at.

The first stage is to determine the traceability route. That is to show, preferably diagrammatically, the route from the measurements being performed to the SI (or, if appropriate, community reference).



**Figure 4 Simple traceability chain for an FEL lamp calibration of an instrument**

For example, consider calibrating an instrument that measures irradiance (e.g. a down welling irradiance meter for ocean colour). The instrument may be calibrated in a laboratory in comparison with an FEL lamp.

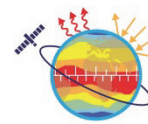
An example traceability chain is provided here. Here, a test and measurement laboratory (perhaps the instrument manufacturer) obtains an FEL lamp from a National Metrology Institute (NMI). In order to avoid over-use of the NMI-provided lamp, the test and measurement laboratory may then use it to calibrate a working standard lamp, which is then used to calibrate the instrument.

This process obtains the hierarchical calibration chain. The aim of the hierarchical approach is to work systematically from the end to the 'beginning' up all the necessary branches of the traceability tree<sup>19</sup>. How far back this process is taken depends on where, for the specific measurement purposes, the value provided by someone or something else can be trusted.

For field measurements, this point of trust may be the certificate provided by a calibration and testing laboratory (especially if that certificate is authorised by an accreditation body). In a calibration laboratory this may be the certificate provided by the national measurement institute. In a national measurement institute this may be the primary SI instrument (for radiometry most often a cryogenic radiometer). Such a reference provides the beginning of the hierarchical calibration chain and the measured value for which the uncertainty budget is being produced provides the end of the chain.

It is not necessary to make the chain go all the way back to fundamental constants, including the full traceability chain within, for example, an NMI. It is necessary to go back to the 'point of trust' – the point for which you have a calibration, with calibration certificate including a full uncertainty statement from an accredited laboratory.

<sup>19</sup> The traceability may be a 'chain' if there is a single artefact that is calibrated in comparison to a reference artefact that was calibrated by a higher tier laboratory, etc. There may be a 'tree' if for example a quantity is calculated by combining electrical, optical, thermal and dimensional measurements – where traceability will be independently to SI for each of those parameters.



**Questions to ask yourself:**

What is the traceability chain for my measurements?

What references do I use?

How were the references set up for the calibration?

Am I relying on other secondary measurements (temperature, time, distance)? How are they calibrated?

Are there intermediate steps – if I'm comparing two sources (lamps), what detector am I using? If I'm comparing two detectors, what source am I using?

How far back do I need to go before I read the answer off a certificate?

### 4.3 Step 2: Writing down the calculation equations

Each arrow in the traceability chain is likely to be calculated by combining different values in a *calculation equation*.

For example, consider the radiance of a diffuser tile illuminated by an FEL lamp. The radiance of the diffuser tile, viewed at an angle of 45° and for normal incidence illumination is given by:

$$L_s = \frac{E_{\text{FEL}} \beta_{0^\circ, 45^\circ}}{\pi} \quad (4.1)$$

where,  $L_s$  is the source radiance,  $E_{\text{FEL}}$  is the lamp irradiance and  $\beta_{0^\circ, 45^\circ}$  is the diffuser reflectance factor. Equation (3.5) forms the basis of the measurement equation and is the *calculation equation*, i.e. the equation used to calculate the 'answer' in the laboratory.

The calculation equation will be expanded to form a full *measurement equation* in Step 4. But at this stage the aim is to write down the calculation that is performed at each step explicitly.

**Questions to ask yourself:**

What is the equation used to calculate each box in my traceability chain?

Is it a single step or multiple step process?

Do I rely on other information (e.g. a distance measurement) that is not included in my traceability chain?

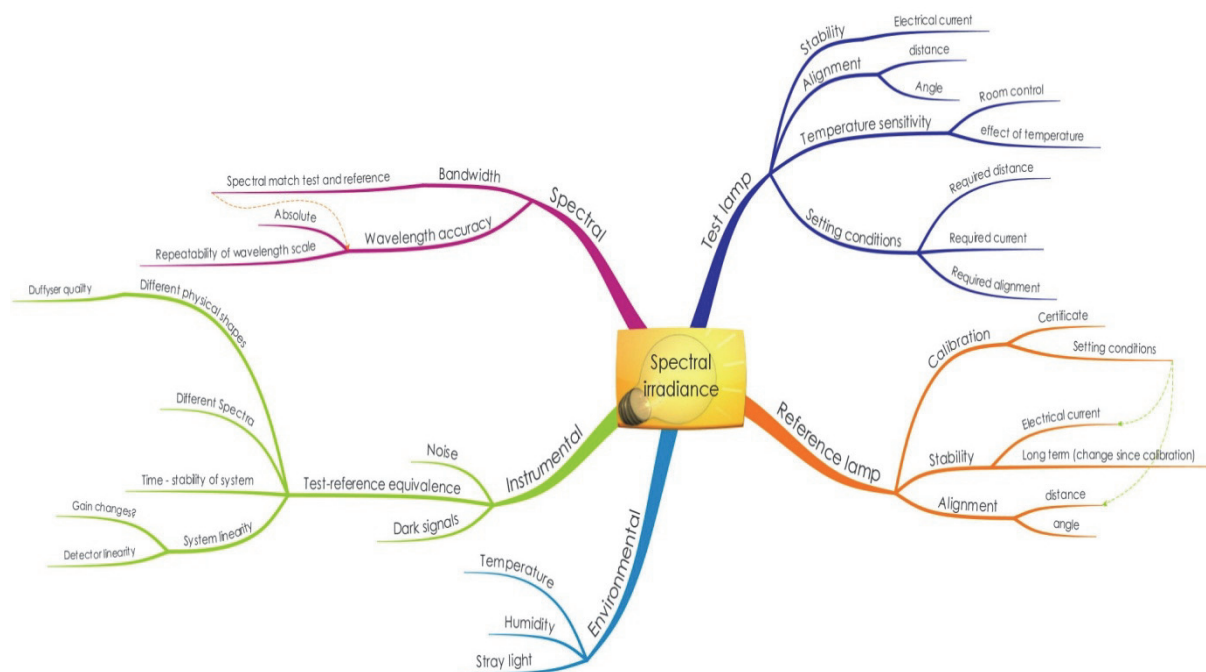
At this stage – do I need to go back to step 1 and refine my traceability chain?



#### 4.4 Step 3: Considering the sources of uncertainty

The aim here is to consider, for each step of the traceability chain, what the sources of uncertainties are. In the final uncertainty budget, these will provide the rows of the uncertainty budget.

Often a good starting point is to brainstorm – whether that is in a list, or using a more graphical approach (e.g. Figure 5). It is helpful to use other people's examples as a check-list.



**Figure 5 Brainstorm mindmap of uncertainties associated with a spectral irradiance measurement of a test lamp in comparison with a reference lamp**

Another method that can be used is to consider the calculation equation for each step of the traceability chain, as determined in Step 2. Each term in that equation will have associated uncertainties and those should be listed. For example, for the calculation equation (4.1), we immediately determine the first sources of uncertainty:

- The irradiance of the FEL lamp (as measured by, e.g. an NMI)
- The reflectance factor of the diffuser (as measured, by e.g. an NMI)

Finally, it is important to understand the underlying (often unstated) assumptions behind the calibration process. Many calibration processes involve a comparison, e.g. that the irradiance of the test lamp is compared with the irradiance of the reference lamp using a transfer spectrometer. The assumption here is that the transfer spectrometer behaves in exactly the same way to the reference lamp and test lamp. However, that may not be the case. For example:

- The spectrometer may be non-linear and the test lamp might be much higher or lower radiance than the reference lamp.
- The spectrometer may be temperature sensitive and one lamp may heat the room up more than the other, or the second lamp measurement may take place a day after the first lamp measurement and the room has changed temperature.
- The spectrometer may have changed in being moved from one source to another.
- The spectrometer may be sensitive to stray light from other wavelengths. The reference lamp (say an FEL) may emit more light at those wavelengths than the test lamp (say an LED-based source).
- The spectrometer's input optics has a diffuser on it. The diffuser is not perfectly Lambertian and the test lamp may illuminate it at a greater range of angles than the reference lamp does.
- The two sources may have different ultraviolet outputs and the diffuser may fluoresce.

All these examples are where the comparison makes underlying assumptions of equivalence, whereas the actual measurement system may not be equivalent for the two sources. At this stage, these effects should be listed.

**Questions to ask yourself (for each step of my traceability chain):**

How did I get this result? What previous information did I use? What else might have affected the result?

What is the equation I'm using to calculate the answer? Where do the values for each of the variables come from? What uncertainty is associated with each of those?

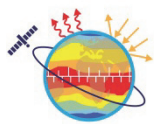
What hidden assumptions are there?

Am I doing a comparison? What am I assuming is the same for that comparison? Am I comparing apples and oranges?

## **4.5 Steps 1 to 3 for a space-borne sensor**

There is nothing fundamentally different about a sensor that is in orbit to an instrument that is calibrated for measurements on an aircraft, in a field or a laboratory. The steps described here will apply in all situations, but may need some translation and interpretation specific to the application. Probably the most significant barrier to applying this to satellite sensors is the complexity of the algorithms and the scale of the task. It is therefore important to be able to determine a simplified version of the uncertainty problem and then add in complexity in stages.

The first three steps are about understanding the problem and writing down what is known about the measurement system. There will still be a traceability chain, albeit perhaps one that is 'broken' by the launch process (see e.g. Section 8.5.2). It can help to separate the analysis into two independent sections – the first that describes the traceability of the pre-



launch calibration and characterisation of the instruments and the second that describes the process of post-launch cal/val using on-board and vicarious calibration processes.

In any calibration chain, even those of well-understood routine calibrations in an NMI laboratory, there will be aspects that are not fully understood. There are always uncertainty components that are estimated, or even guessed, based on experience, historical records or models of 'worst case scenarios'. In a space-borne instrument, there will usually be more such 'guesses' – things that cannot be known (e.g. how has the spectral response function of the instrument changed in orbit?). However, post-launch cal/val processes can put upper limits on probable changes and associate uncertainties with them.

The NMIs are well aware that there can always be systematic effects not fully understood – the "unknown unknowns" in any uncertainty budget. The NMI community attempts to determine whether such factors are present through formal blind international comparisons<sup>20</sup>. Similarly sensor-to-sensor and sensor-to-ground comparisons can give an estimate of the post-launch changes in the instruments on satellites as well as indicating any problems with the calibration (unknown unknowns). The QA4EO Guideline 4 advises on how to run similar comparisons for EO measurements – whether between sensors or between ground measurement techniques.

The first step in uncertainty analysis for a satellite sensor involves an explicit description of the in-orbit traceability chain – including both the prelaunch calibration processes and the post-launch cal/val. The second step is to write down the calculation equations – both for the calibration processes and for the data processing in orbit, e.g. how the digital numbers obtained in orbit are turned into the Level 1 product (e.g. top-of-atmosphere reflectance). This is usually done in the Algorithm Theoretical Basis Document and/or the Detailed Processing Model and should include the corrections applied to the data, including pre-launch and post-launch calibration coefficients. The third step is to write down the sources of uncertainty. This will include:

- Calibration processes in the pre-flight calibration including uncertainties associated with references, with the process of calibration and the 'hidden assumptions' in the calibration (e.g. similar to those described in Section 4.4).
- Calibration processes in the post-launch calibration/validation including uncertainties associated with references, the process of calibration and the 'hidden assumptions' in the calibration
- In-orbit degradation of all instruments and artefacts, especially those that cannot be checked by the inflight calibration/validation
- In-orbit degradation of inflight references

---

<sup>20</sup> As part of the Mutual Recognition Arrangement which ensures that measurements made traceably to SI at one NMI can be considered legally traceable to another NMI within a 'degree of equivalence' there are regular comparisons between NMIs of all significant quantities. The comparisons are organised in a formal manner to ensure rigorous impartiality and the results are published on the Key Comparison Database (<http://kcdb.bipm.org/>).

A really helpful example is given for the MERIS calibration. The document “MERIS Instrument Calibration” is freely downloadable [4]. This report reviews the MERIS calibration process. It distinguishes pre-flight and on-orbit calibration in two separate sections.

The first two sections of the MERIS calibration report are a “review of the instrument” and “the calibration principle”. The report does not explicitly provide a traceability chain. However it is possible to work it out from the provided information. Figure 12 provides the “calibration processing chain” and effectively is doing Step 2 of this uncertainty process – the “calibration processing chain” is the calculation equation. The brainstorm of sources of uncertainty of step 3 is also straightforward from the titles of the report.

#### 4.6 Step 4: Creating the measurement equations

The measurement equation is an extended version of the calculation equation that also explicitly describes the other sources of uncertainty. It is usually sensible to create a measurement equation for the each subsection of the traceability chain, although in some relatively simple cases, a full measurement equation for the whole process can be made. This will build on the ‘hidden assumptions’ described in Step 3. For example, for the calculation equation (4.1) we have the following hidden assumptions:

- The lamp-diffuser distance,  $d_{\text{use}}$  is the same as the distance at which the lamp irradiance was calibrated,  $d_{\text{cal}}$ . Assuming the FEL lamp obeys the inverse square law<sup>21</sup> [5], then there needs to be an additional term  $d_{\text{cal}}^2/d_{\text{use}}^2$  in Equation (3.5).
- The lamp has been stable since calibration and  $E_{\text{FEL}}$  is an accurate representation of the lamp irradiance at the time of the new measurements. There will actually be some variations due to:
  - Lamp stability (short term) – i.e. random fluctuations
  - Lamp stability (long term) – i.e. changes since calibration, drift
  - Lamp alignment (rotation, positioning)
  - Lamp current control: accuracy and stability
- The diffuser has been stable since calibration, and hence  $\beta_{0^\circ/45^\circ}$  is an accurate representation of the diffuser radiance factor at the time of the measurements. There may be some variations due to:
  - Diffuser stability (e.g. due to getting dirty or absorbing hydrocarbons)
- Accuracy of angles set during calibration and use (is it really a  $0^\circ/45^\circ$  geometry?)
- The irradiance patch is uniform over the field-of-view of the instrument viewing the diffuser panel<sup>22</sup>.

---

<sup>21</sup> Which it doesn’t! Actually FEL lamps, for distances greater than 500 mm, obey the ‘modified inverse square law’ and the term should be  $(d_{\text{cal}} + d_{\text{offset}})^2 / (d_{\text{use}} + d_{\text{offset}})^2$ , where the offset distance is the distance from the filament to the reference plane and is determined by testing the inverse square law behaviour using a detector with a clearly defined reference plane. For shorter distances a further term is required to account for the physical dimension of the filament. Note also that for a bulk diffuser, such as Spectralon<sup>TM</sup>, there is an additional ‘diffuser offset distance’.

These concepts should, somehow be included in the measurement equation (which is an expanded version of the calculation equation). The distance term can be directly included into the equation, expanding it to:

$$L_s = \frac{E_{\text{FEL}} \beta_{0^\circ:45^\circ}}{\pi} \frac{d_{\text{cal}}^2}{d_{\text{use}}^2} . \quad (4.2)$$

Note that if the calibration and use distances are nominally the same, the final term has an expected value of unity (one), but there will nevertheless be an uncertainty associated with taking this value of unity that must be considered.

All the terms associated with the assumptions made can be included in the equation. For example, we may write

$$L_s = \frac{E_{\text{FEL}} \beta_{0^\circ:45^\circ}}{\pi} \frac{d_{\text{cal}}^2}{d_{\text{use}}^2} K_{\text{lampstab}} K_{\text{align}} K_{\text{current}} K_{\text{diffstab}} K_{\text{unif}} . \quad (4.3)$$

Here the  $K$  terms relate to the different effects. These terms all have an expected (nominal) value  $K_i = 1$ . In other words, we are assuming that the value of each is unity<sup>23</sup>. However, there is an uncertainty associated with that assumption.

Equation (4.3) forms the full measurement equation. Note that there are still uncertainties associated with the main terms  $E_{\text{FEL}}, \beta_{0^\circ:45^\circ}$  and there are additional uncertainties associated with the assumptions in the  $K_i$  terms.

If there is correlation, then the measurement equation also provides an opportunity to describe the correlation, often in such a way that the 'writes-out' the correlation and means that the remaining terms have no associated correlation (they are independent of each other), as we saw in Section 3.4.

### **Questions to ask yourself:**

Have I included each source of uncertainty in my measurement equation?

Are these additive or multiplicative?

<sup>22</sup> Note that if the diffuser is only 500 mm from the lamp, the corners of a 300 mm by 300 mm panel are 543 mm from the centre of the filament – and from the inverse square law would be expected to be 15 % lower radiance than the centre of the panel. In practice the radiance non-uniformity can be better or worse than this, depending on the filament dimension.

<sup>23</sup> In practice, some of these may be corrections. We may make a correction for the fact that we know the lamp current was set 'wrongly' or that we know the uniformity is not perfect. In this case the terms have an assigned value (not unity), and an uncertainty associated with the correction. It can be helpful to use a different symbol, say  $C_{\text{unif}}$  for corrections.

## 4.7 Step 5: Determining the sensitivity coefficients

The sensitivity coefficient is the sensitivity of the calculated result to an error in each of the parameters of the measurement equation in turn. As described in Section 3.1, sensitivity coefficients can be determined:

- Mathematically, by differentiating the measurement equation
- Numerically, by modelling the effect of a change in that quantity using a system model
- Experimentally, by varying that parameter in the laboratory

All three methods are typically used in the development of any uncertainty budget. The method chosen depends on what information is available: if a model exists as software, it is relatively straightforward to modify that to include a numerical estimate of sensitivity coefficients. If the measurement equation has explicit relationships, then often it is most straightforward to differentiate it directly. If the measurement equation cannot be written down explicitly, then the sensitivity coefficients should be determined experimentally.

### **Questions to ask yourself:**

Can I differentiate the measurement equation?

Can I determine the sensitivity coefficient numerically / through modelling?

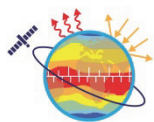
What experimental tests can I do to determine the sensitivity coefficient?

## 4.8 Step 6: Assigning uncertainties

The final aim of this section is to obtain an uncertainty budget table that lists the uncertainty components, their associated uncertainties and the uncertainty associated with the final measured value due to each of these effects in turn. Separately it is important to state clearly the correlations (or lack of correlation). For the example of the lamp-diffuser radiance, an example table is given as Table 2.

**Table 2 Example uncertainty budget for a lamp-diffuser combination**

Uncertainty component	Associated uncertainty		(relative)	Uncertainty associated with radiance due to this
	absolute	relative	Sensitivity coefficient	
Lamp irradiance (calibration)		0.30%	1	0.30%
Diffuser radiance factor (calibration)		0.30%	1	0.30%
Lamp-diffuser distance (same as calibration distance for lamp)?	1 mm in 500 mm	0.20%	2	0.40%
Stability of lamp (short term)		0.10%	1	0.10%
Stability of lamp (drift/ageing)		0.10%	1	0.10%
Alignment of lamp				0.05%
Current stability of lamp (at 350 nm)	3 mA			0.29%
Diffuser stability (ageing)		0.10%	1	0.10%
Uniformity of diffuser		0.50%	1	0.50%



Note here:

- For some effects (irradiance, reflectance factor, stability, uniformity), the estimated uncertainty associated with the parameter is a relative uncertainty in %, the sensitivity coefficient is determined mathematically (Equation (3.7) for example) and is the relative sensitivity coefficient. The final column is determined by multiplying the uncertainty associated with the component with the sensitivity coefficient.
- For the distance the same applies, here the relative sensitivity coefficient is determined mathematically (Equation (3.11), note the negative sign is ignored as this will be squared<sup>24</sup>). This requires the relative uncertainty associated with distance, which is calculated from the absolute uncertainty associated with distance.
- The alignment sensitivity for the lamp was determined experimentally. Therefore there is nothing in the first three columns as it was calculated directly from the standard deviation of multiple measurements with the lamp realigned between measurements. With experimental determinations, the sensitivity coefficient and the assigned uncertainty are often calibrated simultaneously.
- The current sensitivity was estimated to be 3 mA. This was determined to affect the lamp irradiance at 350 nm by 0.29 %. In this case this was evaluated through modelling and the result provided straight into the final column. Alternatively, it could have been done experimentally by changing the current by ten times more (see page 14). In that case a sensitivity coefficient would have been calculated from the measured values for ten times the change and then divided by ten.

It is not therefore necessary to fill in every square in this table. What matters is the final column. This section has described different methods for estimating that final column.

**Questions to ask yourself (for each source of uncertainty):**

How can I assess how large the uncertainty is?

Is it small ( $< 1/5^{\text{th}}$ ) compared with the largest uncertainty?

What is the probability distribution function?

Is it correlated with other uncertainty contributions?

---

<sup>24</sup> Keep the minus signs in if there is any correlation between this and another parameter, as the second half of the law of propagation of uncertainties does not square the sensitivity coefficients.



### Gaussian and other probability distribution functions

Note that here we have assumed that the uncertainties associated with each of the parameters are considered to describe Gaussian (normal) probability distribution functions. Thus the parameter is more likely to be close to the expected value than further from it.

There are times when a rectangular probability distribution function is more appropriate. This means that we know that the value will definitely lie between two limits, with an equal probability of being anywhere between those limits, and no probability of being outside this. This could be the case, for example, where the room temperature is controlled between 19.5 °C and 20.5 °C by the air-conditioning system and will not be outside this. It is also the case for 'rounding errors' on digital displays. If a display provides the value 3.84 it is equally likely that the value falls anywhere in the range between 3.835 and 3.845.

Rectangular probability distribution functions can be represented in uncertainty tables as standard (Gaussian) uncertainties, by dividing the half range by  $\sqrt{3}$ . Sometimes uncertainty budgets will have an additional column labelled "divisor" that will include these  $\sqrt{3}$  values for rectangular distributions.

Similarly, if a value is read off a certificate, then the certificate will probably provide expanded uncertainties (say at the 95 % confidence level) and may include a statement saying, "for  $k = 2$ ", for example. A standard uncertainty (as needed for the uncertainty budget) will be obtained by dividing the expanded uncertainty by the value for  $k$ .

## 4.9 Step 7: Combining and propagating uncertainties

Having obtained the uncertainty associated with the measured value due to each uncertainty parameter in turn it is necessary first to combine the uncertainties to obtain the uncertainty associated with the measured value due to all these uncertainty components, and then to expand that calculated standard uncertainty to an appropriate confidence level. These steps are discussed here.

Uncertainties are combined using the Law of Propagation of Uncertainties given in the GUM, and given above as Equation (2.1). If the different parameters given in the uncertainty budget table are uncorrelated, then the first half of the equation applies and if the final column in the uncertainty budget provides  $c_i u_i$ , the combined standard uncertainty is obtained by "adding the column in quadrature", i.e. taking the square root of the sum of the squares<sup>25</sup>, which for the table above is 0.84 %. If there are correlations between the input parameters, then the second half of the Law of Propagation of Uncertainties is required.

### Questions to ask yourself (for each source of uncertainty):

Will the smallest uncertainties have negligible impact?

<sup>25</sup> In Excel " =sqrt(sumsq(F10:F18)) " for example



## 4.10 Step 8: Expanded uncertainties

The *central limit theorem* states that the arithmetic mean of a sufficiently large number of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally (Gaussian) distributed<sup>26</sup>.

What this means for uncertainty analysis is that if there is a reasonably large number of input parameters, with similar enough (well-defined) uncertainties, then the probability distribution of the output parameter will be approximately Gaussian, no matter what are the probability distributions of the input parameters. This means that it is straightforward to expand uncertainties assuming that the standard uncertainty obtained by combination relates to a Gaussian distribution.

For a Gaussian distribution, the standard uncertainty represents a coverage probability of approximately 66 %, therefore the true value will be within the standard uncertainty of the measured value approximately 66 % of the time. It is more common to provide (approximately) 95 % confidence intervals. For a Gaussian distribution, these are obtained by multiplying the standard uncertainty by the coverage factor  $k = 2$ .

If the distribution is not Gaussian, then a different coverage factor is needed. A different coverage factor is also needed if the standard uncertainties in the uncertainty budget table are not sufficiently well known. Section 3.3.2 described increasing the estimate of uncertainty obtained from the standard deviation of a number of repeat readings according to Equation (3.26) to account for the fact that the standard deviation of a small number of readings is unreliable. If this is done, then it is reasonable to assume that  $k = 2$ .

Alternatively, the GUM provides the Welch-Satterthwaite Equation, which calculates the effective degrees of freedom of the combined standard uncertainty (i.e. a weighted answer to the question: "how many independent measurements are involved in estimating the uncertainty?").

The Welch-Satterthwaite equation calculates the effective degrees of freedom using the following expression

$$v_{\text{eff}} = \frac{u_c^4(y)}{\sum_{i=1}^N \frac{u_i^4(y)}{v_i}}. \quad (4.4)$$

Here,  $u_c(y)$  is the combined standard uncertainty associated with the value  $y$  and  $u_i(y) = |c_i|u(x_i)$  is the uncertainty associated with  $y$  due to one of its contributing effects,  $x_i$ , i.e. the standard uncertainty associated with  $x_i$  multiplied by the magnitude of the

---

<sup>26</sup> Provided the distribution of no one random variable dominates

sensitivity coefficient.  $\nu_i$  is the number of degrees of freedom associated with the estimate of  $u_i(y)$ .

For Type A determinations of uncertainty, i.e. where the uncertainty is determined by taking the standard deviation of multiple measurements, the degrees of freedom is the number of measurements minus one<sup>27</sup>:

$$\nu = N - 1 \quad (4.5)$$

Note that if you had a 'commissioning phase' (see Section 3.3.2) where you made large numbers of measurements to estimate the uncertainty, then the  $N$  to use here is the number of measurements combined in that commissioning phase to estimate the standard deviation and hence uncertainty. If for the routine measurements only a small number of measurements are averaged, then the degrees of freedom are still calculated from the original number of commissioning measurements, as this is about how good the uncertainty estimate is. Note however, that this uncertainty associated with a single reading will only be reduced in averaging by the square root of the number of readings actually averaged today.

For Type B determinations of uncertainty, i.e. where the uncertainty is determined 'by other means', for example prior knowledge, certificates etc, the number of degrees of freedom is harder to define. There are two ways of doing so:

1. By assuming that the Type B uncertainty is known "perfectly" and the "degree of freedom is infinite". This comes from the attitude that you're "given" a Type B uncertainty. In this example the bottom line of the Welch-Satterthwaite equation ends up with a term  $u_i^4(y)/\infty = 0$ .
2. We estimate the "uncertainty in the uncertainty". This is, of course, more of an art than a science, but if we have a feel for an uncertainty range, then the GUM provides an equation

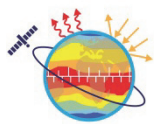
$$\nu = \frac{1}{2} \left[ \frac{u(x)}{u(u(x))} \right]^2 \quad (4.6)$$

So, for example, if we have a Type B uncertainty that is 3 %, but our experience and confidence considers that what we mean by this is that the uncertainty lies in the range from 2 % to 4 %, we may say the "uncertainty in the uncertainty" is 1 % absolute (or the uncertainty is reliable to about 0.33). And

$$\nu = \frac{1}{2} \left[ \frac{3\%}{1\%} \right]^2 = 4.5 \quad (4.7)$$

---

<sup>27</sup> The minus one comes from the fact that if a mean is calculated from  $N$  measurements, then the mean and  $N - 1$  measurements fully describe the problem: the final measurement can be calculated from the mean and the others. i.e. the mean 'uses up' one of the degrees of freedom.



The final step involved in evaluating an expanded uncertainty is to determine the value of  $k$ , i.e. the value to multiply a standard uncertainty by in order to turn it into an expanded uncertainty with a 95 % confidence interval. This is obtained from the t-distribution, and for a 95 % confidence interval the values are given in Table 3. More values are given in the GUM.

**Table 3 Degrees of freedom and value of the coverage factor for a 95 % confidence interval**

Degrees of freedom from the Welch-Sattherwaite formula	Value of $k$ required for a 95 % confidence interval
2	4.3
3	3.18
4	2.78
5	2.57
10	2.23
15	2.13
20	2.09
40	2.02
100	1.984
$\infty$	1.960

(For infinite degrees of freedom, the coverage interval for  $k = 2$  is 95.45 %).

#### **4.11 Experimental techniques for testing and validating an uncertainty budget**

We have considered above the various steps involved in calculating the uncertainty associated with any given measurement result. In summary, this involves thinking about:

- What affects the measurement result?
- How big is the uncertainty associated with each of these effects?
- How sensitive is the result to each of these effects (i.e. what is the 'sensitivity coefficient')?
- Are the effects correlated?

However no matter how thorough you are with this analysis, it is always possible that some important contribution has been missed, or that one of the contributions considered has been underestimated (or, indeed, overestimated). So the question then becomes "How do I know I have calculated the 'right' uncertainty"?

Fortunately there are a number of approaches that can be used to test or validate an uncertainty evaluation, as described below. These are all based on comparisons between results obtained in different ways, and generally involve the use of the so called  $E_N$ -ratio, defined as

$$E_N = \frac{|E_1 - E_2|}{\sqrt{U_1^2 + U_2^2}} \quad (4.8)$$

where  $E_i$  is the measurement result and  $U_i$  is the associated expanded uncertainty. A value of  $E_N < 1.0$  indicates that the two results agree with each other within the limits expected based on their associated uncertainties.

Simply repeating the measurement in exactly the same way will provide information on the impact of randomly varying variables, such as detector noise, which can be useful in terms of validating the 'Type A' contributions in the uncertainty budget. However repeat measurements, if carried out appropriately, can also yield information regarding 'Type B' (systematic) contributions. The approaches described below are all useful, but in each case it is important to remember that many of the systematic uncertainties included in the measurement uncertainty budget will be common to both sets of results involved in the comparison; these should be counted only once, or even removed completely, when calculating the  $E_N$  ratio.

#### 4.11.1 Use a different reference

The use of two or more different references is particularly useful for confirming that neither one has changed its calibration value (e.g. due to ageing effects, damage due to misuse or during transportation). Many laboratories routinely use at least two references for all measurements, while others use one on a routine basis with a second included on an occasional basis as a check on the performance of the first.

#### 4.11.2 Use different instruments.

Similarly to using a different reference, repeating the measurements in exactly the same way but using a different instrument (or instruments) can be used to confirm the performance of these instruments is consistent with the allowance made in the uncertainty budget. For example, if a DVM is used to monitor the voltage drop across a standard resistor in order to set a specified lamp current, a comparison with the results obtained using a different DVM will give useful confirmation of the associated uncertainty.

#### 4.11.3 Ask someone else to do the measurement.

This is particularly useful for testing / validating the uncertainties associated with factors that depend on human judgements, such as aligning a lamp. Often the uncertainties associated with such effects are assessed by a single operator setting what they consider to be a 'worst acceptable alignment', but this can be subject to unsuspected bias. Asking another operator to carry out the alignment may yield a different result and can be used to improve the estimated uncertainty allowance.

#### 4.11.4 Change the order of the measurements

This may appear trivial, but can yield valuable insights. For example, the environmental conditions in the laboratory may change during the day, and simply changing from making measurements in the morning to the afternoon may then give different results. Or there may be an unsuspected reflection from a baffle in a measurement set up that causes stray light to impinge on a detector placed at one position on a translation stage, but not at another; in this case, changing the positions of the detectors may result in different results being obtained.

#### 4.11.5 Use a different measurement method

Often the same measurand can be determined several different ways and comparing the results obtained using each of these is an excellent way of testing and validating the associated uncertainty budgets. A common example is spectroradiometric measurements using either an array instrument or a scanning device. In this case the majority of the systematic uncertainties will be different for each approach, but nevertheless it is important to exclude those which are common when comparing the results. For example, if the same power supply, DVM and standard resistor are used for measurement of the same lamp using both methods, then the uncertainties associated with these should be excluded when calculating the  $E_N$  ratio.

#### 4.11.6 Compare with someone else

This is often regarded as the ultimate test, and is why NMIs spend so much time and effort in carrying out international comparisons of key measurement quantities, such as spectral irradiance. Even if another laboratory uses a similar measurement approach, the actual equipment and references used, the precise method followed, and the scientists carrying out the measurements, will all be different. Comparing the results not only reveals the level of agreement between individual participants, but also provides direct validation of the claimed uncertainties; put simply, if results for individual participants do not agree with the mean result to within combined uncertainties, then that participant has underestimated their measurement uncertainty. When conducting comparisons between different laboratories, it is useful to compare not only the results and total uncertainties, but also the detailed uncertainty budgets; this can highlight, for example, uncertainties that have been overlooked, or under or over-estimated, by individual participants.

#### 4.11.7 Use a different artefact

This is an extremely valuable way of testing and validating an uncertainty budget, but is often overlooked; indeed it is often deliberately avoided, on the basis that using the 'best' artefact available will almost certainly lead to the 'best uncertainty'. The danger with always using the same type of artefact (e.g. the same type of lamp or detector) is that an unsuspected systematic bias may always be included; in other words, the results may be highly consistent, but consistently incorrect! Using a different type of artefact can help quantify any 'known unknowns' or reveal 'unknown unknowns'.



## 5 Case study: APEX imager calibration (simplified)

### 5.1 The APEX imager calibration

This case study describes a somewhat simplified version of the APEX imager calibration. Later sections consider some of the more involved aspects. It is strongly recommended that in producing an uncertainty budget, participants start with a simplified version of their own traceability and then add complexity as they become more familiar with the concepts. Although the details in this section apply to the airborne APEX instrument specifically, the concepts they introduce are more general than this and apply to all EO imagers.

ESA's Airborne Imaging Spectrometer APEX (Airborne Prism Experiment) was developed under the PRODEX (PROgramme de Développement d'EXpériences scientifiques) program by a Swiss-Belgian consortium and entered its operational phase at the end of 2010. It is an imaging spectrometer that is usually flown on a DLR aircraft to obtain high resolution images of the ground with typical pixel sizes ranging between 1.5 m to 2.5 m. APEX features up to 532 spectral bands in full spectral mode (providing spectral information from 387 nm to 2500 nm). It also has spectral programmability to achieve higher Signal-to-Noise-Ratios (SNR) by reducing the number of bands in a binned configuration. Data [6, 7] are acquired in 1000 pixels across track with a FOV of 28°.

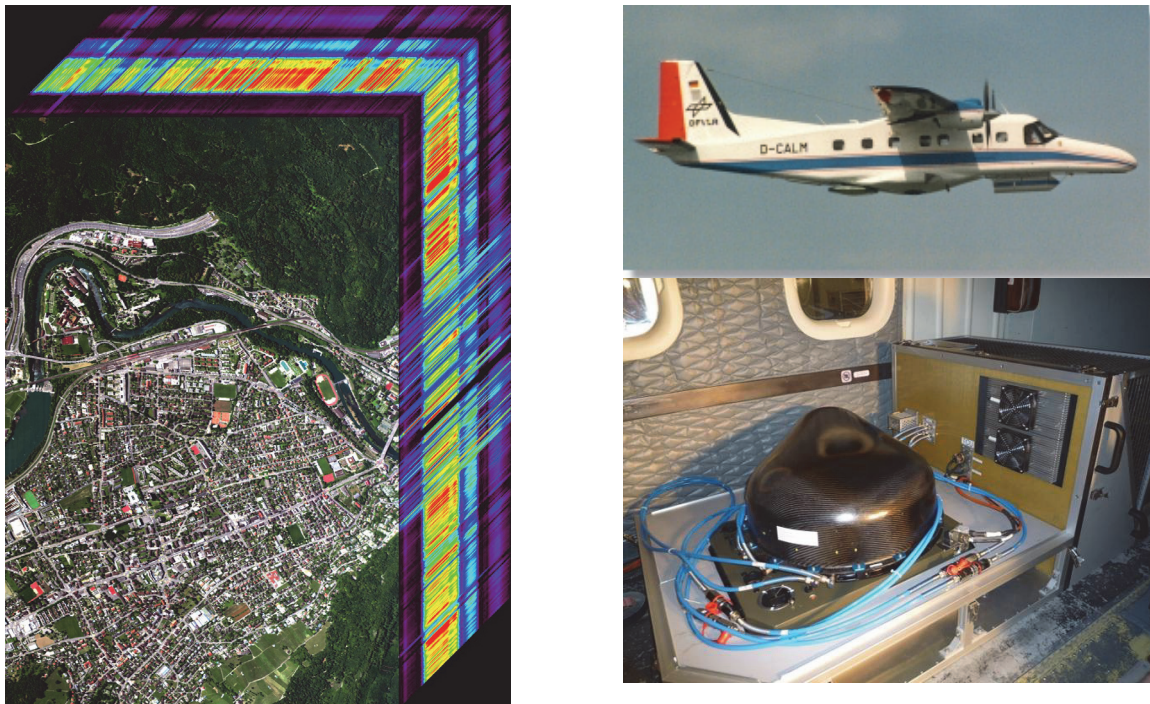
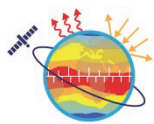


Figure 6 APEX imager on board aircraft and captured hyperspectral image<sup>28</sup>

<sup>28</sup> The data cube shown here is available for free download from the APEX web site ([www.apex-esa.org](http://www.apex-esa.org)) as APEX Open Science Dataset.



The APEX calibration is carried out at the CHB (Calibration Home Base) situated at DLR Oberpfaffenhofen [8]. The standard APEX calibration is carried out in an operational manner at least once a year, usually at the beginning of the flight season. The calibration comprises measurements on an optical bench for the geometric and spectral calibration, using a collimator-slit setup and a monochromator respectively, and measurements using a small and big integrating sphere for absolute radiometric calibration and flat-fielding respectively.

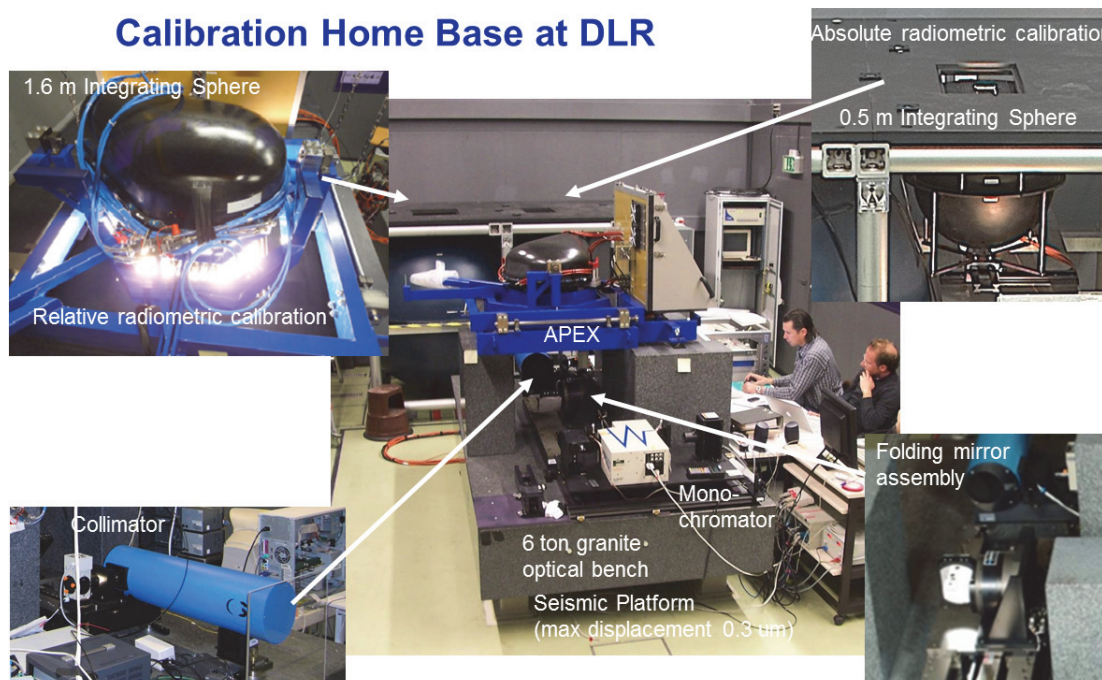


University of  
Zurich <sup>UZH</sup>

Department of Geography

RSL  
measurements | products | policy

## Calibration Home Base at DLR



**Figure 7 Calibration Home Base at DLR**

A full sensor calibration at the CHB can be achieved within 3-4 days, including sensor installation and alignment on the calibration bench. Raw calibration data are stored in the APEX Calibration Information System (CAL IS) [9] and processed to generate calibration cubes, holding calibration coefficient on a pixel per pixel basis (Figure 8).

APEX imagery is calibrated to radiance in the APEX Processing and Archiving Facility (PAF)[10], applying the calibration coefficients produced by the CAL IS to the image cubes.

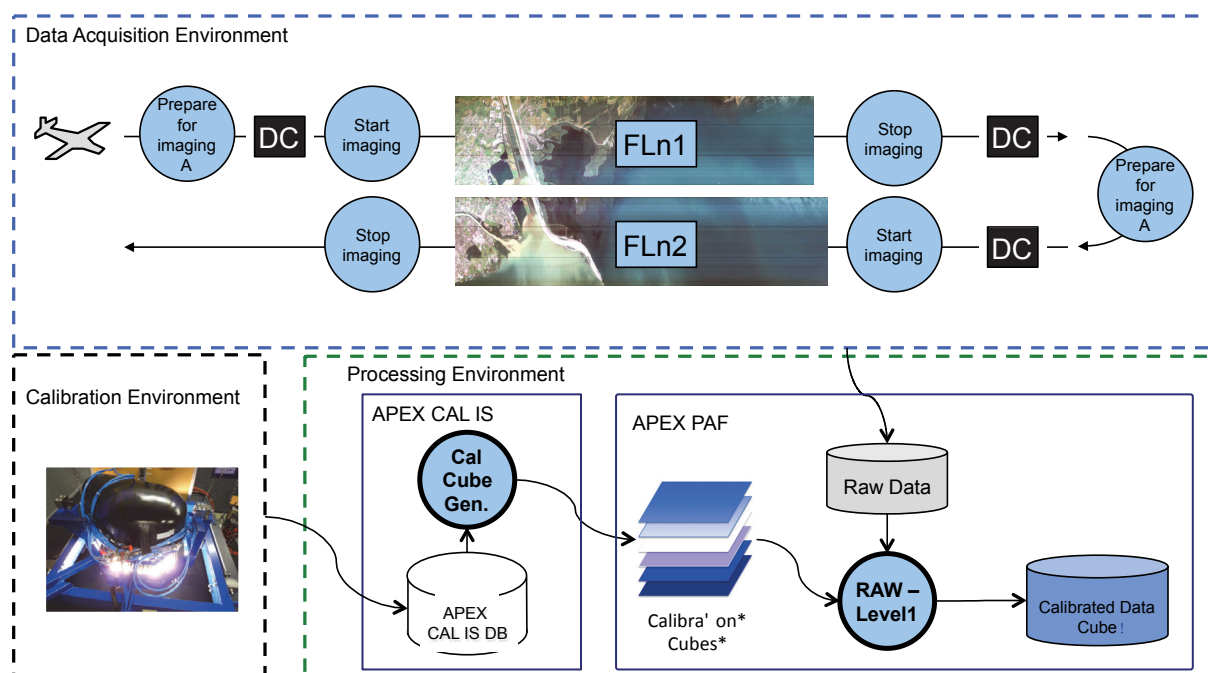


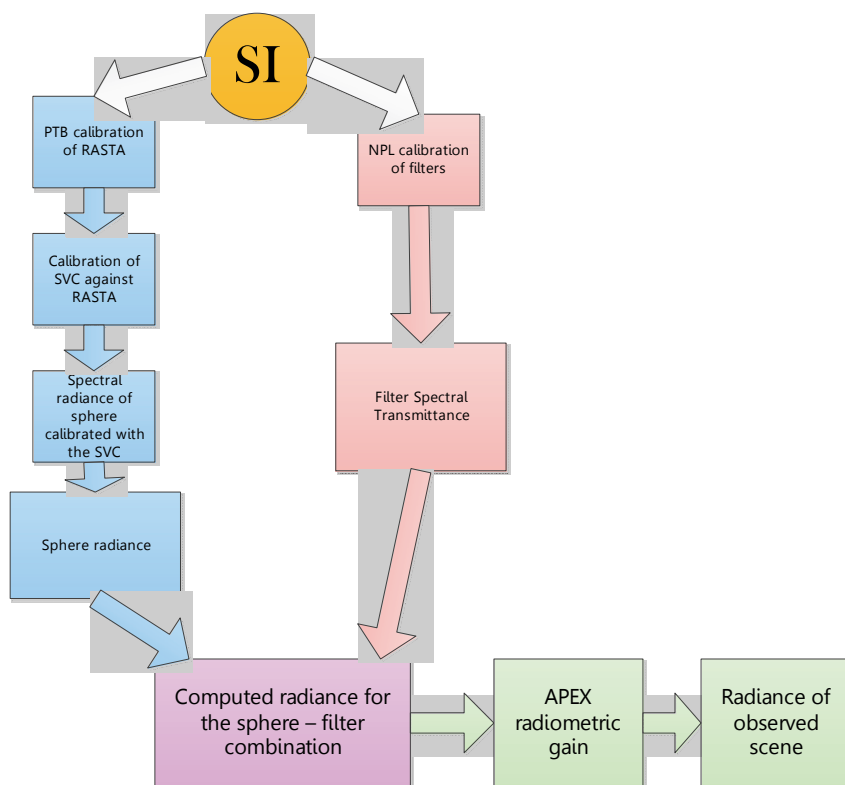
Figure 8: APEX Data Acquisition to Product Chain

## 5.2 Step 1: Describing the traceability chain

The APEX instrument is calibrated, at a few radiance levels, using a small integrating sphere source with neutral density filters placed in front of the source. By making the calibration for different radiance levels, the linearity and any bias offset for the instrument can also be determined.

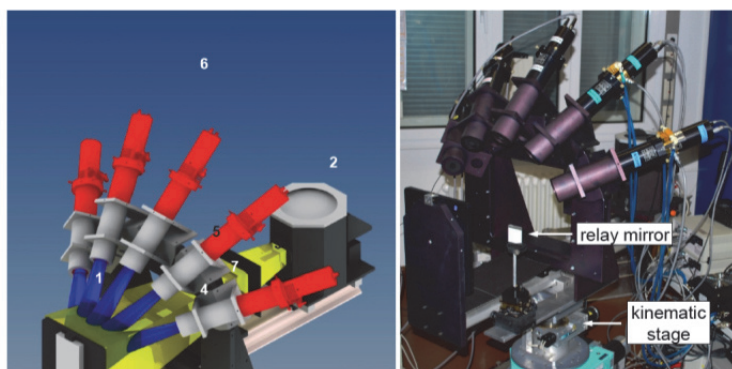
The integrating sphere source is too large to be easily calibrated at an NMI. So instead, a hand-held spectrometer is used to transfer a calibration from a source engineered by DLR and calibrated by PTB (the RASTA source) to the integrating sphere source. The neutral density filters were calibrated at NPL directly. Therefore the traceability chain, simplified, is shown in Figure 9.





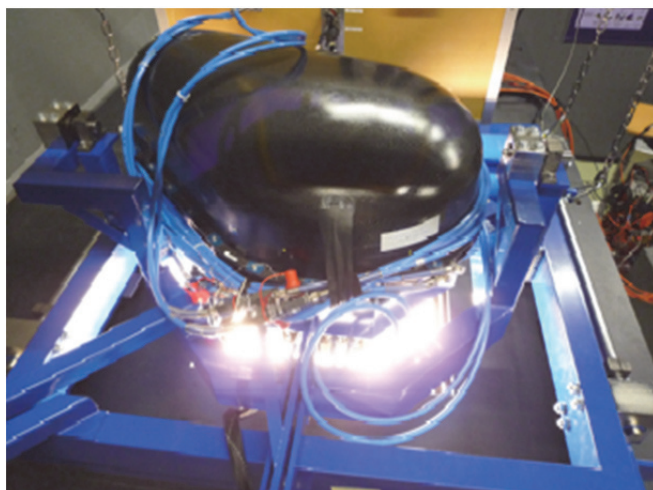
**Figure 9 Traceability chain (simplified) for the APEX calibration**

In the blue section on the left of Figure 9, the calibration is transferred from the RASTA source, to a portable spectrometer and then to the integrating sphere used for the APEX calibration. The RASTA source [11] consists of a white diffuser tile illuminated by an FEL lamp at normal incidence and viewed at 45°. The source also includes appropriate baffling to reduce and control stray light and several filter radiometers that monitor the short and long-term stability of the RASTA source. The source was calibrated at an NMI, in this case PTB, and that calibration provides the traceability to SI.



**Figure 10 From [11]: Mechanical set-up of RASTA (left). RASTA after its alignment for calibration at PTB (right). The filter radiometers are shown in red in the mechanical set-up diagram. The lamp is in the octagonal housing and the diffuser panel is mounted on the plate on the left of the diagram and picture.**

The spectrometer used is the SVC 1024 spectrometer<sup>29</sup>. This is first calibrated radiometrically against the RASTA and then used to calibrate the integrating sphere. The integrating sphere [8] is 500 mm in diameter, with a port of 40 mm × 200 mm. The sphere is operated so that the exit port is at the top of the sphere which allows APEX to be calibrated in its flight orientation (Figure 11).



**Figure 11 APEX mounted above the integrating sphere used for its calibration**

In order to perform the APEX calibration at several radiance levels, neutral density filters are used to reduce the output of the sphere. These neutral density filters were directly calibrated for transmittance at an NMI, in this case, NPL, and therefore have straightforward traceability to SI.

The purple rectangle in Figure 9 is where the computed radiance for the sphere-filter combination is calculated. This is a straightforward multiplication of the sphere radiance and the filter transmittance.

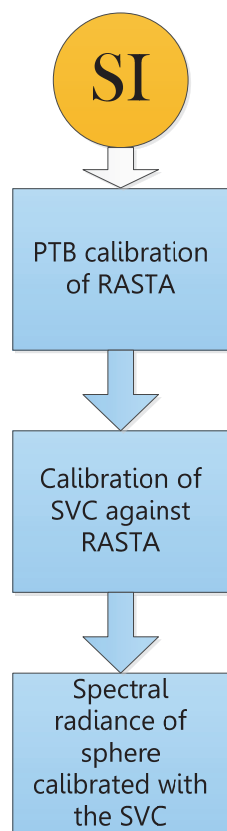
The green section transfers the calibration to the APEX instrument. First the radiometric gain of the APEX instrument (i.e. the conversion factor from its output digital numbers to radiance) is calculated from the APEX measurement of the sphere-filter source, and then the measured digital numbers of a particular scene are converted to radiance using the instrument gain.

---

<sup>29</sup> <http://www.spectravista.com/HR1024i.html>

## 5.3 Step 2: Writing down the calculation equations

### 5.3.1 For the calibration of the sphere



**Figure 12**

The traceability chain for this section is shown to the left.

In the first step PTB provides a calibration of the RASTA source (comprising a lamp and diffuser and appropriate baffling). This creates an SI-traceable calibration of the radiance of the RASTA source,  $L_{\text{RASTA}}$ . PTB provides a calibration certificate for the radiance, with an associated relative uncertainty, expressed as an expanded (95 % confidence,  $k = 2$ ) uncertainty in per cent.

In the second step the SVC transfer spectrometer views the RASTA source and obtains a spectrum (in digital numbers [DN] as a function of wavelength). This measurement is performed as the average of several light readings minus the average of several dark readings (with the entrance port of the spectrometer closed). From this, the gain of the SVC (units:  $[\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1} \text{DN}^{-1}]$ ) is determined.

In the third step, the SVC transfer spectrometer views the APEX calibration sphere source. The radiance of the APEX sphere source is calculated from the SVC gain and the measured DN for the sphere source, again an average of several light readings minus an average of several dark readings. From this the spectral radiance of the APEX calibration sphere source is determined.

**Both parts of this process can be considered comparison calibrations. In the second step of**

Figure 12, the SVC gain<sup>30</sup> is calculated from the calculation equation

$$G_{\text{SVC}} = \frac{L_{\text{RASTA}}}{DN_{\text{RASTA}}} \quad , \quad (5.1)$$

where,  $L_{\text{RASTA}}$  is the PTB-calibrated radiance of the RASTA source and  $DN_{\text{RASTA}}$  is the measured digital numbers on the SVC spectrometer when viewing the RASTA source. Note that all three quantities in Equation (5.1) are spectral quantities and are functions of wavelength. The equation is valid for a specified wavelength, but the wavelength dependence is not specifically written here to simplify the presentation.

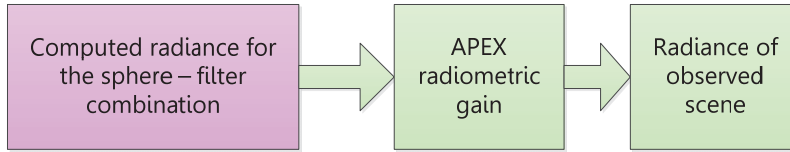
<sup>30</sup> An alternative would be to calculate the spectral radiance responsivity of the SVC instrument. The spectral radiance responsivity, with units  $[\text{DN} / (\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1})]$ , is the inverse of the gain,  $R_{\text{SVC}} = G_{\text{SVC}}^{-1}$  and is the response of the instrument per unit radiance.

The third step is similar. The radiance of the sphere is calibrated using the SVC spectrometer and is calculated using the calculation equation

$$L_{\text{sphere}} = DN_{\text{sphere}} G_{\text{SVC}} \quad , \quad (5.2)$$

where,  $L_{\text{sphere}}$  is the sphere radiance,  $DN_{\text{sphere}}$  is the measured digital numbers when the SVC views the sphere and  $G_{\text{SVC}}$  is the SVC gain, calculated from (5.1). Again, all three quantities are spectral quantities, and this equation is calculated at each wavelength in turn.

### 5.3.2 For the calibration of the APEX imager



The radiance of the sphere-filter combination is calculated as

$$L_{\text{sph-filt}} = L_{\text{sphere}} \tau_{\text{filter}} \quad (5.3)$$

where  $\tau_{\text{filter}}$  is the transmittance of the filter, and  $L_{\text{sphere}}$  is calculated in Equation (5.2). Again, these are spectral quantities and this is calculated for each wavelength in turn.

The signal on the APEX imager for a particular wavelength,  $DN_{\text{APEX,cal}}$ , in digital numbers, when viewing the sphere-filter combination is then used to calculate the APEX gain as

$$G_{\text{APEX}} = L_{\text{sph-filt}} / DN_{\text{APEX,cal}} \quad , \quad (5.4)$$

where

$$DN_{\text{APEX,cal}} = DN_{\text{APEX,cal,light}} - DN_{\text{APEX,cal,dark}} \quad . \quad (5.5)$$

### 5.3.3 For the user of the APEX imager to measure scene radiance

The radiance of an observed scene is then

$$L_{\text{scene}} = G_{\text{APEX}} DN_{\text{APEX,scene}} \quad , \quad (5.6)$$

where

$$DN_{\text{APEX,scene}} = DN_{\text{APEX,scene,light}} - DN_{\text{APEX,scene,dark}} \quad . \quad (5.7)$$

## 5.4 Step 3: Considering the sources of uncertainty

Each stage of the transfer from the PTB radiance values to the observed scene radiance is some form of comparison calibration, and therefore there are uncertainties associated with both the explicit terms in the calculation equation, and the 'hidden assumptions' that

everything is equivalent on both halves of the comparison. For each calculation equation, we can prepare an uncertainty table. These are given below,

#### 5.4.1 For the calibration of the SVC spectrometer against the RASTA

Calculation Equation (5.1) is a comparison between the values assigned to RASTA by PTB using their measurement instrument RASTA and the SVC’s measured values for the RASTA. It is important to review three sources of uncertainty: those due to  $L_{\text{RASTA}}$ , those due to  $DN_{\text{RASTA}}$  and those due to implied assumptions hidden within this comparison.

**Table 4 Uncertainties associated with Calculation Equation (5.1) and its hidden assumptions**

$u_{G_{\text{SVC}}}$	Uncertainty component	Comments
<b>Uncertainties associated with <math>L_{\text{RASTA}}</math></b>		
$L_{\text{RASTA}}$ $u_{L_{\text{RASTA}}}$	PTB calibration of $L_{\text{RASTA}}$	This will be read off the certificate. See the additional notes box on the next page.
<b>Uncertainties associated with <math>DN_{\text{RASTA}}</math></b>		
$DN_{\text{RASTA}_\ell}$ $u_{DN_{\text{RASTA}_\ell}}$	Noise in light reading	This will be obtained through ‘Type A’ methods (i.e. from a standard deviation – see Section 2.6.2.)
$DN_{\text{RASTA}_d}$ $u_{DN_{\text{RASTA}_d}}$	Noise in dark reading	This will be obtained through ‘Type A’ methods
<b>Uncertainties associated with the comparison assumptions</b>		
$K_{\text{RASTA}_{\text{age}}}$ $u_{\text{RASTA}_{\text{age}}}$	Ageing of RASTA since PTB calibration	The RASTA source has inbuilt filter radiometers that monitor the source stability. The variation of these signals can be used to estimate this uncertainty component, and, if appropriate, apply a correction.
	Stability of RASTA (short term)	This will get included in the standard deviation of the light readings, and does not need ‘double counting’ here
$K_{\text{stray}}$ $u_{\text{stray}}$	External stray light influencing the SVC calibration	This is light that contributes to the SVC signal and comes from outside its field-of-view.
	Any environmental sensitivities of RASTA (temperature, pressure, humidity)	These are assumed to be negligible (and therefore given an uncertainty of 0 %). RASTA consists of a tungsten lamp, which is insensitive to temperature, and a diffuser, which is also insensitive <sup>31</sup> to minor temperature changes.
$K_{\text{unif}}$ $u_{\text{unif}}$	Uniformity of RASTA and any differences in the field-of-view for PTB’s calibration and the SVC view	Understanding this requires a uniformity scan of the RASTA source, along with knowledge of the fields-of-view of the PTB calibration (from the measurement certificate) and the SVC source.

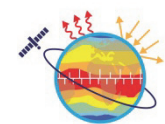
<sup>31</sup> Spectralon does show a phase transition at ~19 °C, which affects its reflectance by approximately 0.1 %. However, the lamp is likely to heat the diffuser above this temperature, and this is minor compared to other uncertainty components.

#### 5.4.2 For the calibration of the sphere with the SVC spectrometer

Calculation Equation (5.2) is a comparison between the SVC's measurement of RASTA and the SVC's measurement of the sphere. It is important to review three sources of uncertainty: those due to  $DN_{\text{sphere}}$ , those due to  $G_{\text{SVC}}$  and those due to the implied assumptions hidden within this comparison.

**Table 5 Uncertainties associated with Calculation Equation (5.2) and its hidden assumptions**

$u_{L_{\text{sphere}}}$	Uncertainty component	Comments
<b>Uncertainties associated with <math>G_{\text{SVC}}</math></b>		
$G_{\text{SVC}}$ $u_{G_{\text{SVC}}}$	As above	This will come from the previous step.
<b>Uncertainties associated with <math>DN_{\text{sphere}}</math></b>		
$DN_{\text{sphere}_\ell}$ $u_{DN_{\text{sphere}_\ell}}$	Noise in light reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation).
$DN_{\text{sphere}_d}$ $u_{DN_{\text{sphere}_d}}$	Noise in dark reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation).
<b>Uncertainties associated with the comparison assumptions</b>		
$K_{\text{SVC\_dft}}$ $u_{\text{SVC\_dft}}$	Change of SVC between measurements	We assume that the two measurements are reasonably close together in time. This means that there is no 'ageing' between one set of measurements and the next. There may be some sensitivity, however, to physically moving the SVC between the devices, for example. This 'drift' can be estimated by simulating the movement and remeasuring the same source with the SVC. This term should only be included if it is larger than the noise.
	Stability of SVC (short term)	This will get included in the standard deviation of the light readings, and does not need 'double counting' here
$K_{\text{stray}}$ $u_{\text{stray}}$	External stray light influencing the SVC during calibration against RASTA and use with the sphere	This is light that contributes to the SVC signal and comes from outside its field-of-view. What matters is the change in this between the calibration and use.
$K_{\text{stray\_in}}$ $u_{\text{stray\_in}}$	Internal stray light influencing the SVC during calibration against RASTA and use with the sphere	This is light that is scattered within the SVC onto the pixel for the wrong wavelength. This should be characterised for the SVC spectrometer and its effect will depend on the different spectral shape of the two source radiances – what matters is the difference between the RASTA and sphere source spectral radiances
$K_{\text{temp}}$ $u_{\text{temp}}$	Any environmental sensitivities of the SVC (temperature, pressure, humidity)	If the external conditions vary from the calibration of the SVC vs the RASTA to its use to calibrate the sphere, then it is necessary to account for the sensitivity of the SVC to those changes. Spectrometers can be very temperature sensitive and local room temperatures can be higher near high power tungsten sources.
$K_{\text{lin}}$ $u_{\text{lin}}$	Linearity of SVC	Spectrometers can be non-linear, either in terms of their integration time (does doubling the integration time double the signal), or in terms of their response to double the radiance. This will be a problem if the SVC is used on a different integration time for each source and/or if the sources are different radiance levels.



### 5.4.3 For the combined source radiance from the sphere and filters

Calculation equation (5.3) assumes that the radiance of the source is a product of the radiance of the sphere and the transmittance of the filter. This also has some hidden assumptions.

**Table 6 Uncertainties associated with Calculation Equation (5.3) and its hidden assumptions**

$u_{L_{\text{sph-filt}}}$	Uncertainty component	Comments
<b>Uncertainties associated with <math>L_{\text{sphere}}</math></b>		
$L_{\text{sphere}}$ $u_{L_{\text{sphere}}}$	As above	This will come from the previous step.
<b>Uncertainties associated with <math>\tau_{\text{filter}}</math></b>		
$\tau_{\text{filter}}$ $u_{\tau_{\text{filter}}}$	Transmittance of the filter	This will be obtained from NPL's calibration certificate.
<b>Uncertainties associated with the hidden assumptions</b>		
$K_{\text{reflect}}$ $u_{\text{reflect}}$	Optical interreflections between the sphere and the filters	When the filter is introduced in front of the sphere, it may be that some light from the sphere is reflected back into the sphere and this alters the radiance of the sphere. This coefficient is to account for any such interreflection effects. To estimate such effects, the distance between the sphere and filter can be varied, or the filter angled slightly so that reflections change direction (being aware that for many filters changing the angle will itself change the transmittance). These are experimental methods for estimating the sensitivity coefficient.
$K_{\text{fil\_temp}}$ $u_{\text{fil\_temp}}$	Temperature sensitivity of the filter and changes when in front of the sphere	Filters are usually temperature sensitive. The filter may be at a different temperature in front of the sphere than it was during its calibration at NPL. The temperature of the filter should be measured in front of the sphere, e.g. with a thermocouple. The filter could be recalibrated at different temperatures, or information obtained from the manufacturer on the temperature stability of the filter.
$K_{\text{sph-stab}}$ $u_{\text{sph-stab}}$	The stability of the sphere between its calibration with the SVC and its use with the filter	This can be estimated by the repeatability of multiple readings of the sphere alone over this time period
$K_{\text{filt\_age}}$ $u_{\text{filt\_age}}$	Ageing of the filters since calibration	There is a time delay between the filter calibration and their use. In this time they are irradiated with UV radiation, but even storage can alter filters. The filters should be recalibrated at regular intervals to estimate the likely ageing since calibration.

### 5.4.4 For the calibration of the APEX gain from the measurement of the sphere-filter source

Calculation Equation (5.4) is a comparison between the SVC measurement of the sphere-filter combination and the APEX instrument measuring the same source. There are, again, implied assumptions hidden within this comparison.



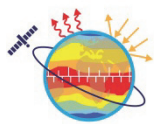
**Table 7 Uncertainties associated with Calculation Equation (5.4) and its hidden assumptions**

$u_{G_{APEX}}$	Uncertainty component	Comments
<b>Uncertainties associated with <math>L_{sph-filt}</math></b>		
$L_{sph-filt}$ $u_{L_{sph-filt}}$	As above	Calculated above.
<b>Uncertainties associated with <math>DN_{APEX,cal}</math></b>		
$DN_{APEX,cal,light}$ $u_{DN_{APEX,cal,light}}$	Noise in light reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation).
$DN_{APEX,cal,dark}$ $u_{DN_{APEX,cal,dark}}$	Noise in dark reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation).
<b>Uncertainties associated with the comparison assumptions</b>		
	Stability of sphere-filter (short term)	This will get included in the standard deviation of the light readings, and does not need 'double counting' here
$K_{stray}$ $u_{stray}$	External stray light influencing the APEX calibration	This is from light that contributes to the APEX signal and comes from outside its field-of-view.
$K_{humid}$ $u_{humid}$	Any environmental sensitivities of the sphere-filter (temperature, pressure, humidity)	Temperature and pressure effects are assumed to be negligible (and therefore given an uncertainty of 0 %). The sphere has high power lamps in it and the temperature of the sphere will be 'set' by the sphere lamps, not the room conditions. There may be some sensitivity at certain wavelengths to humidity. A sphere has a very long effective path length, and therefore shows absorption at the water absorption lines. This will change with room humidity.
$K_{unif}$ $u_{unif}$	Uniformity of the sphere-filter and any differences in the field-of-view for the SVC calibration and the APEX view	Understanding this requires a uniformity scan of the source, along with knowledge of the fields-of-view of the SVC spectrometer and the APEX spectrometer.

#### 5.4.5 For the observed scene radiance

The calculation equation (5.6) is a comparison between the calibration of APEX using the sphere and APEX's measurement of the scene. There are several assumptions.





**Table 8 Uncertainties associated with Calculation Equation (5.6) and its hidden assumptions**

$u_{I_{\text{scene}}}$	Uncertainty component	Comments
<b>Uncertainties associated with <math>G_{\text{APEX}}</math></b>		
$G_{\text{APEX}}$ $u_{G_{\text{APEX}}}$	As above	This will come from the previous step.
<b>Uncertainties associated with <math>DN_{\text{sphere}}</math></b>		
$DN_{\text{APEX,scene,light}}$ $u_{DN_{\text{APEX,scene,light}}}$	Noise in light reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation ).
$DN_{\text{APEX,scene,dark}}$ $u_{DN_{\text{APEX,scene,dark}}}$	Noise in dark reading	This will be obtained through 'Type A' methods (i.e. from a standard deviation).
<b>Uncertainties associated with the comparison assumptions</b>		
$K_{\text{APEX\_dft}}$ $u_{\text{APEX\_dft}}$	Change of APEX between measurements	We assume that the two measurements are reasonably close together in time. This means that there is no 'ageing' between one set of measurements and the next. APEX may be sensitive, however, to being transported to the aircraft, mounted in the aircraft, etc. This change can be estimated by remeasuring the same source for a second calibration after the measurement campaign (but ensure that only effects bigger than noise are included to prevent double counting).
	Stability of APEX (short term)	This will get included in the standard deviation of the light readings, and does not need 'double counting' here
$K_{\text{stray}}$ $u_{\text{stray}}$	External stray light influencing APEX during its use with the sphere	This is light that contributes to the APEX signal and comes from outside its field-of-view. This is specific to the situation in the aircraft, which is likely to be very different from in the laboratory calibration and it is the difference in stray light that is included here.
$K_{\text{stray\_in}}$ $u_{\text{stray\_in}}$	Internal stray light influencing APEX during calibration and use. Also known as cross-talk	This is light that that is scattered within the APEX onto the pixel for the wrong wavelength. This should be characterised for a spectrometer and will depend on the different spectral shape of the two source radiances. In this case there are both spatial and spectral dimensions to consider. Again, it is the change in internal stray light between the calibration source and the scene measurements that matters.
$K_{\text{temp}}$ $u_{\text{temp}}$	Any environmental sensitivities of the APEX (temperature, pressure, humidity)	The external conditions of the APEX are very different on the aircraft to in the laboratory. These changes should be simulated and the changes understood. Even if certain corrections are made, there will be a residual uncertainty.
$K_{\text{lin}}$ $u_{\text{lin}}$	Linearity of APEX	Spectrometers can be non-linear, either in terms of their integration time (does doubling the integration time double the signal), or in terms of their response to double the radiance. This is tested by making the calibration with different filters in front of the sphere

## 5.5 Step 4: Creating the measurement equations

The calculation equations need to be expanded into full measurement equations by including all the hidden assumption terms. Here we are treating all these terms as a multiplicative factors  $K_i$ . All the  $K_i$  have an expected value of unity (one), with an uncertainty associated with that, described by  $u_i$ .

Thus, calculation equation (5.1) becomes measurement equation

$$G_{SVC} = \frac{L_{RASTA} K_{RASTA\_age} K_{stray} K_{unif}}{DN_{RASTA\_l} - DN_{RASTA\_d}} . \quad (5.8)$$

Calculation equation (5.2) becomes measurement equation

$$L_{sphere} = (DN_{sphere\_l} - DN_{sphere\_d}) G_{SVC} K_{SVC\_dft} K_{stray} K_{stray\_in} K_{temp} K_{lin} . \quad (5.9)$$

Calculation equation (5.3) becomes measurement equation

$$L_{cal}(\lambda_0) = L_{sphere}(\lambda_0) \tau_{filter}(\lambda_0) K_{sph\_stab} K_{sph\_temp} K_{reflect} K_{filt\_age} K_{filt\_temp} . \quad (5.10)$$

Calculation equation (5.4) becomes measurement equation

$$G_{APEX} = \frac{L_{sph-filt} K_{stray} K_{humid} K_{unif}}{DN_{APEX,cal,light} - DN_{APEX,cal,dark}} . \quad (5.11)$$

Calculation equation (5.6) becomes measurement equation

$$L_{scene} = G_{APEX} DN_{APEX,scene} K_{APEX\_dft} K_{stray} K_{stray\_in} K_{temp} K_{lin} . \quad (5.12)$$

In all these measurement equations the terms are independent of each other, and are independent from line to line. This means that correlation does not need to be taken into account.

## 5.6 Step 5: Determining the sensitivity coefficients

For these straightforward measurement equations, the simplest way to determine the sensitivity coefficients is through differentiating. Consider first equation (5.8) for  $L_{RASTA}$

$$\frac{\partial G_{SVC}}{\partial L_{RASTA}} = \frac{K_{RASTA\_age} K_{stray} K_{unif}}{DN_{RASTA\_l} - DN_{RASTA\_d}} = \frac{G_{SVC}}{L_{RASTA}} \quad (5.13)$$

Therefore, the uncertainty associated with  $G_{SVC}$  due to  $L_{RASTA}$  is

$$u_{G_{SVC} \cdot L_{RASTA}} = \frac{G_{SVC}}{L_{RASTA}} u_{L_{RASTA}} \quad (5.14)$$

And therefore,

$$\frac{u_{G_{SVC} \cdot L_{RASTA}}}{G_{SVC}} = \frac{u_{L_{RASTA}}}{L_{RASTA}} ; \quad (5.15)$$

i.e. the *relative* uncertainty associated with the SVC gain due to the RASTA radiance is equal to the *relative* uncertainty associated with the RASTA radiance. Therefore the relative sensitivity coefficient is 1.

The same relationship holds for all the parameters in the numerator of Equation (5.8).

The denominator of Equation (5.8) is  $DN_{RASTA} = DN_{RASTA_{\ell}} - DN_{RASTA_d}$  ; i.e. a light signal minus a dark signal. This was discussed in detail in Section 3.3.1. For now, we consider this as a single quantity,  $DN_{RASTA}$  . Therefore

$$\frac{\partial G_{SVC}}{\partial DN_{RASTA}} = \frac{-L_{RASTA} K_{RASTA_{age}} K_{stray} K_{unif}}{(DN_{RASTA})^2} = \frac{-G_{SVC}}{DN_{RASTA}} \quad (5.16)$$

and the relative uncertainty associated with the gain, due to the relative uncertainty associated with the digital number signal is given by

$$\frac{u_{G_{SVC} \cdot DN_{RASTA}}}{G_{SVC}} = \frac{-u_{DN_{RASTA}}}{DN_{RASTA}} \quad (5.17)$$

and the relative sensitivity coefficient is -1.

## 5.7 Step 6: Assigning uncertainties

The final aim of this section is to obtain an uncertainty budget table that lists the uncertainty components, their associated uncertainties and the uncertainty associated with the final measured value due to each of these effects in turn.

Some methods for determining the uncertainty associated with the individual parameters were described in the tables above.

### The use of NMI calibration certificates

A calibration certificate from an NMI generally provides uncertainties with a statement similar to: "The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor  $k = 2$ , providing a coverage probability of approximately 95 %." The concept of expanded uncertainties is described in more detail in Section 4.10. When using the values, the uncertainties should be converted back to standard uncertainties. This means dividing by the provided value for  $k$  , which is generally, but not always 2.

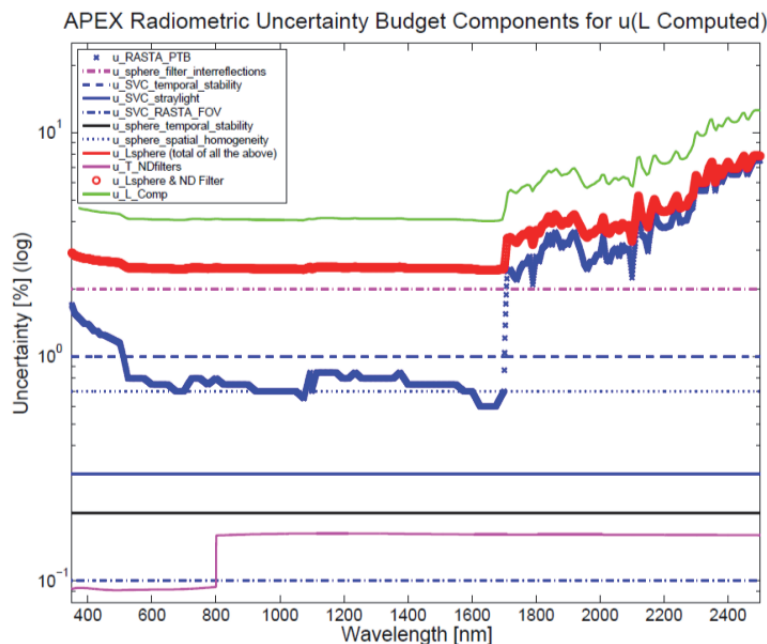
Each table in Section 5.4 can be expanded with extra columns for size of effect, (relative) sensitivity coefficient and "uncertainty associated with  $Y$  due to this effect" (here  $Y$  represents the quantity calculated by each calculation equation in turn). For example, for the

second step in the chain, the determination of the gain of the SVC, this would be as in Table 9.

**Table 9 Uncertainty budget table for the gain of the SVC**

$u_{G_{SVC}}$	Uncertainty component	Size of effect	(relative) sensitivity coefficient	Uncertainty associated with SVC gain due to this effect
<b>Uncertainties associated with <math>L_{RASTA}</math></b>				
$u_{L_{RASTA}}$	PTB calibration of $L_{RASTA}$		1	
<b>Uncertainties associated with <math>DN_{RASTA}</math></b>				
$u_{DN_{RASTA_\ell}}$	Noise in light reading		Combined, relative:	
$u_{DN_{RASTA_d}}$	Noise in dark reading		-1	
<b>Uncertainties associated with the comparison (apples and oranges)</b>				
$u_{RASTA\_age}$	Ageing of RASTA since PTB calibration		1	
	Stability of RASTA (short term)	0	1	
$u_{stray}$	External stray light influencing the SVC calibration		1	
	Any environmental sensitivities of RASTA (temperature, pressure, humidity)	0	1	
$u_{unif}$	Uniformity of RASTA and any differences in the field-of-view for PTB's calibration and the SVC view		1	

Similar tables can be prepared for the other parameters; however in all cases because of the simple measurement equations, the relative sensitivity coefficient is 1 (or occasionally -1). The parameters themselves are spectral quantities – they take a different value in each of the spectral bands. And therefore a more helpful way to present them is graphically, as in Figure 13.



**Figure 13** Uncertainty components for the APEX radiometric calibration presented graphically on a logarithmic scale.

## 5.8 Step 7: Combining and propagating uncertainties

Here the quantities are mutually independent (no correlation) and the uncertainty is propagated from one level to the next by being included as an input uncertainty at the next equation. The combined relative standard uncertainty associated with the scene radiance is therefore obtained by adding all the relative standard uncertainties in quadrature. This provides the green line in Figure 13.

## 5.9 Step 8: Expanding uncertainties

Almost always, to obtain the 95 % confidence interval, the standard uncertainty is multiplied by  $k = 2$  and this has been used here.

## 5.10 This is just the start of the process

The uncertainty analysis described here is a significant simplification of the actual calibration process for the APEX imager. It is, however, important to start with the simplification and then build in the complexity. Later sections in this report describe, for example, how the linearity measurements can be used to obtain a straight line calibration equation and how the spectral calibration of APEX is done and how it relates to the uncertainty analysis of this chapter.

As complexity is introduced it can be considered a new 'module' that improves our measurement model and understanding of the uncertainty budget. Often initially such analysis is done to understand the uncertainty to assign to a component, later, as the analysis becomes more sophisticated, it can be used to calculate a correction and associated uncertainty. For example, the linearity measurements may initially be performed to estimate

the uncertainty associated with the assumption that linearity is insignificant (i.e. that  $K_{lin}$  in Equation (5.12) takes the value unity). It may be that from those tests a more sophisticated linearity model is developed and that latter a linearity correction is applied, with an associated uncertainty.

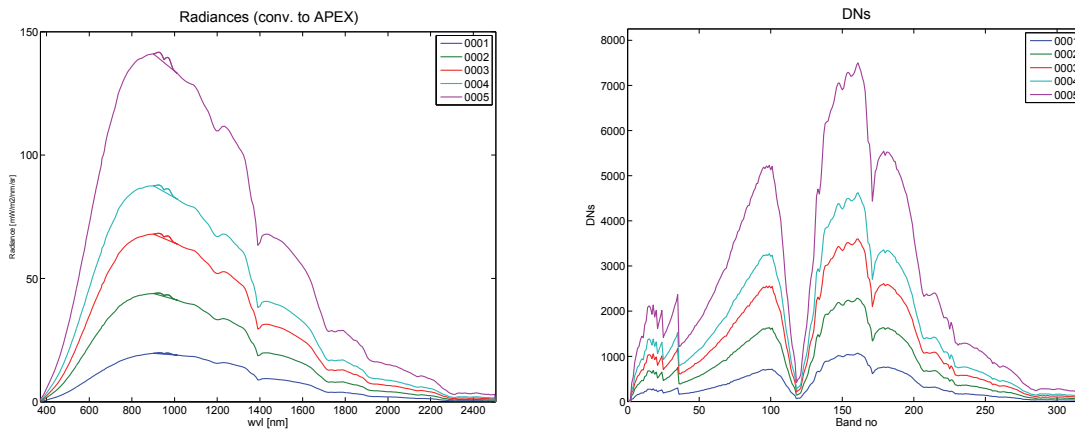
It is worth noting that because uncertainties are added in quadrature, an uncertainty that is 1/5 the size of the largest uncertainty will have a relative weighting on the uncertainty budget of 1/25<sup>th</sup>. Generally there is a point where an uncertainty becomes 'insignificant'. Usually at the end of Step 3 of the steps to an uncertainty budget it is possible to eliminate several terms as 'insignificant'. At Step 6 others may be classified in this way.

## 6 Straight line calibration equations

This chapter focuses mostly on a 'straight line calibration equation' – where the response of the instrument to different levels of input is fitted with a straight line. The final section, Section 6.6, considers interpolating experimental data by straight line interpolation.

### 6.1 A straight line calibration equation

In Equation (5.4) in Section 5.3.2, the gain of the APEX imager is determined using the integrating sphere and filter, and then used in Equation (5.6) to convert the APEX signal ( $DN$ ) when viewing a scene into scene radiance,  $L$ . This approach oversimplified the situation. In practice, the process is repeated with several different filters to obtain a response of APEX to different signal levels. The sphere-filter radiance (as a spectral quantity) and the signal response of APEX to those radiance levels are shown in Figure 14.



**Figure 14** Different radiance levels presented to APEX for different filters in front of the sphere (left) and the signal on APEX for each of these levels (right).

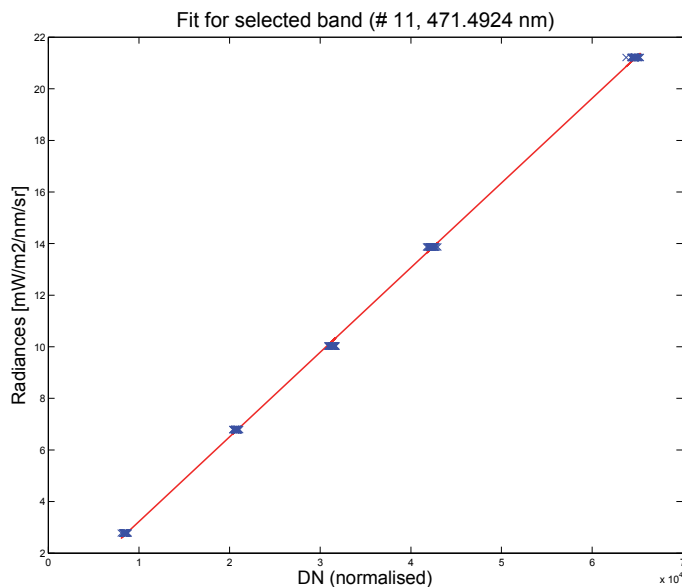
From these results, at any particular wavelength, values are obtained for the signal at different radiance levels. Example results are shown in Figure 15. A straight line calibration function is fitted to these results and from this two parameters are obtained– the slope and the origin offset. The measured values at different radiance levels provide the calibration equation

$$L = G_{\text{APEX}} \times DN + L_0 \quad (6.1)$$

where  $L$  is the measured radiance in a particular band,  $L_0$  is the offset radiance,  $G_{\text{APEX}}$  is the gain for that band and  $DN$  is the measured signal (in digital numbers) in that band. This can also be written as

$$DN = \frac{L}{G_{\text{APEX}}} + DN_0, \quad (6.2)$$

where  $DN_0 = -L_0/G_{\text{APEX}}$  is the signal offset.



**Figure 15 Example fit for a particular band of the signal on APEX and the source radiance**

This approach is similar to that described in Section 5.3.2. There a single measured value is used to determine the gain and the offset is assumed to be zero (or that  $DN_0$  is the dark reading that is subtracted from the measured signal). The straight line calibration equation expands this concept slightly by using data from other measured radiances and allowing for any offset that isn't fully removed by the dark reading.

When the instrument is then used to measure a scene, Equation (6.1) replaces (5.6). Straight line calibration equations are often used for calibrating instruments. If the instrument has a small non-linearity, then the uncertainty associated with the assumption of linearity can be predicted by the quality of such a fit. If the instrument is significantly non-linear, then higher order calibration equations may be used instead. Although the analysis below is specific for a straight line calibration equation, many of the concepts apply to higher order equations too.

## 6.2 Uncertainty analysis overview

The straight line calibration equation is a specific example of the more general problem of fitting a model to experimentally obtained data. Here the model is a straight line. For all fitting problems we must consider that:

- There is uncertainty associated with the input values (both on the horizontal and the vertical axes)
- There may be some correlation associated with input values
- The output of the fitting problem will include the model parameters (for the straight line, the slope and the intercept) and their covariance matrix (providing the uncertainty associated with each of them and the correlation between them). The



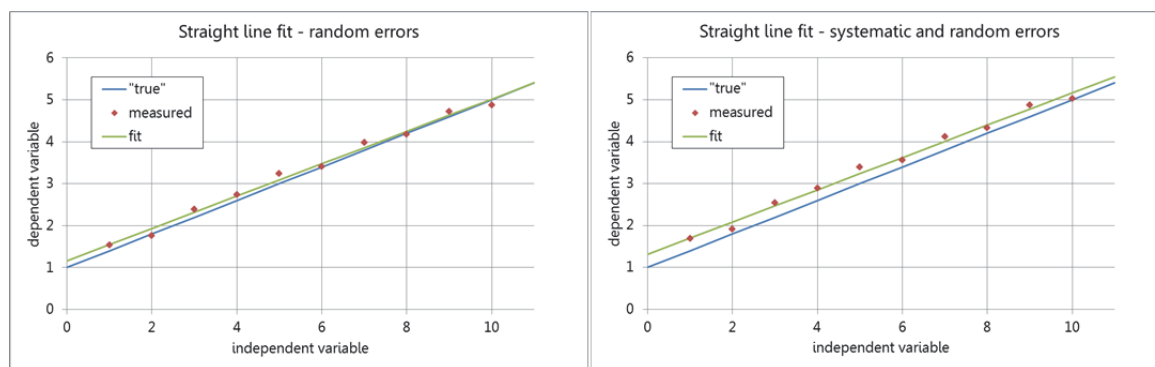
output parameters will always have some associated correlation because they were derived from the same input values.

- The uncertainty obtained from the fit assumes that the model is correct. We also need to consider any uncertainty associated with the suitability of the model used.

The straight line fit is considered in detail in the subsequent subsections. This subsection provides an 'intuitive' overview of the analysis.

If the measured data values that will be fitted by the straight line have only random errors, i.e. the uncertainty associated with those data values is associated only with random effects, then the straight line will 'average out' some of those effects. The fit will go through the data with points either side of it. Because of this, the uncertainty associated with a value determined from the straight line is smaller than the uncertainty associated with any one measured value that went into the fit (Figure 16, left). One way to understand that is that the fitted parameters are determined from all the measured data, just as in an average, the average is determined from all the values. And therefore, as in an average, the uncertainty associated with the fitted point is generally smaller than the uncertainty associated with the individual measured values.

If the measured data points that will be fitted by the straight line have systematic errors, i.e. the uncertainty associated with those data values is at least partially associated with systematic effects, then the fitted line will not correct for these common effects<sup>32</sup>. This means that this uncertainty associated with systematic effects will be 'built into' the straight line obtained, and the uncertainty obtained from points on the straight line will not be reduced through averaging (Figure 16, right).



**Figure 16 A straight line fit, on the left with random errors only. Here the fit is much closer to the 'true' value than are individual points. On the right with systematic and random errors. Here the fit shows an offset from the true value because of the systematic effect common to all the measured data points.**

<sup>32</sup> Note that systematic effects can have different values from point to point if there is a sensitivity coefficient that varies, for example a systematic stray light effect may affect higher radiance levels more than lower radiance levels, providing a 'tilt' to the straight line.

This is why it is important to understand the dependence of the input data values on random and systematic effects by assigning a covariance matrix to them. That is done for the APEX sensor as discussed in Section 6.3.

The outputs of a fitting process are the model parameters – in this case the slope and intercept of the straight line. Whether or not the input values have associated correlation, these output parameters will always be correlated because they were derived from the same input values. Parameters are derived from the fit, so in this case, when the APEX instrument is in flight, the measured scene digital numbers are converted to a scene radiance using the gain and offset (Equation (6.1)). Because the gain and offset are correlated due to the fit, the uncertainty associated with the output quantity (scene radiance) must consider the full law of propagation of uncertainties. This is discussed in Section 6.6.

The final consideration is whether the model we use, in this case the straight line, is an appropriate fit to the data. The uncertainties associated with values obtained from the model are usually surprisingly small, because of the ‘averaging’ effect described above. This may not be a ‘true’ representation of the uncertainty associated with the output, however, if the fit is not a good fit to the measured data. This is discussed in Section 6.5.

### 6.3 Calibration data and uncertainties

The overview in the previous section shows the importance of understanding the covariance of the input measured values. There are three ways to estimate this:

- If they have come from a previous step involving fitting or interpolation, the associated covariance matrix will be calculated in that step.
- If they have come from experimental data or data modelled in a Monte Carlo simulation, then the correlation can be estimated using the equations in Section 3.5.3.
- If they come from experimental data and you have a reliable instrument model, you will be able to estimate the covariance from your understanding of the processes, as described in Section 3.5.2.

The third approach is often the most useful and it is this approach that is used in the APEX example. The gain and offset for the APEX instrument (at each spectral channel analysed independently) are determined from measurements made of the response of APEX to light of different radiance levels. The different radiance levels come from introducing neutral density filters in front of the integrating sphere.

The two variables are source radiance and the signal (digital numbers) on APEX. Which is considered the  $x$ -variable and which the  $y$ -variable depends on whether Equation (6.1) or (6.2) is used. The second of these equations is more intuitive – we vary the radiance level and see the measured signal, but the former is more easily applied to a scene radiance calculation later as it directly provides the gain.

Either way, in order to fit the straight line, we need to determine the uncertainty associated with measured values of the two variables: signal and source radiance. In order to determine the covariance we need to determine what is common to the measured values and what varies from measurement to measurement.

Considering first the signal (digital numbers); as discussed in Section 5.4.4 the uncertainty here comes from noise on the light signal and noise on the dark signal. Noise, by definition, changes from measurement to measurement, and therefore the signal values are considered entirely uncorrelated<sup>33</sup> for all the signal-radiance level pairs.

The uncertainty associated with the source radiance is discussed in Section 5.4.3. It comes from the sphere radiance uncertainty (itself a result of the chain described generally in Section 5). It also comes from the uncertainty associated with the filter transmittance measurement, optical interreflections between the sphere and the filters, temperature sensitivities of the filters in front of the spheres, sphere stability and filter ageing.

The dominant uncertainty here is that associated with the sphere radiance. This is common to the measurements at all radiance levels as it is the same sphere used. The next most dominant uncertainty is the calibration of the filters by NPL. This will introduce partial correlation because NPL has measured all the filters on the same facility using the same technique. Generally speaking, measurements at an NMI will minimise the uncertainty associated with random effects, by taking repeat readings. Therefore it is likely that although this is 'partial' correlation, the correlation coefficient will be high. Interreflections are likely to be similar for all filters and since all the filters are the same type (neutral density filters), their temperature sensitivities and ageing rates may be similar. There may be some filter-dependent ageing and the sphere instability. Overall, the measured values can be considered 'highly correlated' and as a first approximation, 'fully correlated'.

Thus we have the situation where the signal levels (digital numbers) are entirely uncorrelated and the source radiance levels are entirely correlated. The covariance matrix is therefore simple to produce. The covariance matrix is square with  $2n$  rows and columns where  $n$  is the number of  $(x_i, y_i)$  measured points used for the fit. It effectively has four quarters, each square with  $n \times n$  values. The top left quarter represents the covariance between the different  $x$ -values. The bottom right quarter represents the covariance between the different  $y$ -values. The bottom left and top right quarters represent the covariance between the  $x$ -values and the  $y$ -values. These are uncorrelated and these quarters have zeroes throughout. Down the diagonal is the variance (the squared uncertainty) associated with each measured value individually.

---

<sup>33</sup> Note, that if the same dark reading is subtracted for each radiance level, then there will be a correlation introduced by this choice. Such a decision may be made if the dark signal is relatively stable, as a way of saving time. Generally, this decision is made if the dark signal is small in comparison to the light signal. And therefore the introduced correlation will also be small.

- The  $x_i$  terms – or rather, the source radiance levels,  $L_{\text{sph-filt},i}$ , are considered fully correlated. If they have an associated relative uncertainty (expressed in percent) of  $u_{\text{rel}}(L_{\text{sph-filt}})$ , then for that quarter of the covariance matrix the diagonal terms are  $u_{\text{rel}}^2(L_{\text{sph-filt}})L_{\text{sph-filt},i}^2$  (i.e. the absolute uncertainty squared) and all off-diagonal elements are  $u_{\text{rel}}^2(L_{\text{sph-filt}})L_{\text{sph-filt},i}L_{\text{sph-filt},j}$  (i.e. the product of the two absolute uncertainties).
- The  $y_i$  terms – or rather, the signal (digital number) terms  $DN_{\text{APEX,cal},i}$ , are considered uncorrelated. If they have an associated relative uncertainty of  $u_{\text{rel}}(DN_{\text{APEX,cal},i})$ , then for that segment of the covariance matrix, the diagonal terms are  $u_{\text{rel}}^2(DN_{\text{APEX,cal},i})DN_{\text{APEX,cal},i}^2$  (i.e. the absolute uncertainty squared) and all off-diagonal elements are 0.

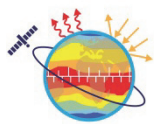
Therefore, writing for compactness  $u_{\text{rel}}^2(L_{\text{sph-filt}})L_{\text{sph-filt},i}L_{\text{sph-filt},j} = u_{s-f}^2 L_i L_j$  and  $u_{\text{rel}}^2(DN_{\text{APEX,cal},i})DN_{\text{APEX,cal},i}^2 = u_{DN,i}^2 DN_i^2$ , the full covariance matrix<sup>34</sup> is

$$\begin{array}{c}
 \begin{array}{cccccccc}
 & x_1 & x_2 & \cdots & x_n & y_1 & y_2 & \cdots & y_n \\
 \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{array} & \left[ \begin{array}{cccccccc}
 u_{s-f}^2 L_1^2 & u_{s-f}^2 L_1 L_2 & \cdots & u_{s-f}^2 L_1 L_n & 0 & 0 & \cdots & 0 \\
 u_{s-f}^2 L_2 L_1 & u_{s-f}^2 L_2^2 & \cdots & u_{s-f}^2 L_2 L_n & 0 & 0 & \cdots & 0 \\
 \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
 u_{s-f}^2 L_n L_1 & u_{s-f}^2 L_n L_2 & \cdots & u_{s-f}^2 L_n^2 & 0 & 0 & \cdots & 0 \\
 0 & 0 & \cdots & 0 & u_{DN,1}^2 DN_1^2 & 0 & \cdots & 0 \\
 0 & 0 & \cdots & 0 & 0 & u_{DN,2}^2 DN_2^2 & \cdots & 0 \\
 \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & u_{DN,n}^2 DN_n^2
 \end{array} \right]
 \end{array}
 \end{array} \quad (6.3)$$

i.e. the block in the top left corner has terms to account for the covariance. The block in the bottom right corner has values down the diagonal only, and the bottom-left and top-right corners are all zeroes.

Note that this process has to be repeated separately for each spectral band. The covariance matrix for each will be different, even if the relative uncertainties are the same at each wavelength, since the absolute value of radiance (and hence absolute uncertainties) are different.

<sup>34</sup> The red terms in Equation (6.3) are helpful labels that would not normally be included in this equation. They show how the different rows and columns represent the different  $x_i$  values and then  $y_i$  values.



## 6.4 Determining the fit coefficients (doing the fit)

### 6.4.1 Approaches to take

There are different approaches that can be taken to estimating the fit parameters and their associated uncertainty. Most of these will be some form of least squares analysis and least squares analysis can be performed at different levels of complexity. The least complex would be unweighted least squares (the input values are given equal weight and no account is made of their associated uncertainty or correlation), then weighted least squares (the input values are given different weights depending on their associated uncertainty) and finally generalised least squares (the covariance matrix of the input values is taken into account in the fit).

The associated uncertainty analysis can also be performed for these different levels of complexity. The most rigorous approach would use generalised least squares for the fit and obtain a covariance matrix for the fit parameters from this analysis. Here we consider the following levels of complexity:

1. Using an unweighted fit and calculate the associated uncertainties and covariance using Monte Carlo simulation.
2. Using a weighted fit and obtain the uncertainties and covariance from the process
3. Using a generalised fit.

The second and third of these are based on the concepts described in British Standard DD ISO/TS 28037:2010 "Determination and use of straight-line calibration functions", which describes how to deal with all these cases and provides a 'recipe' for determining both the fit parameters and their associated covariance. Matlab code to apply those recipes is available, for free download, at:

[http://www.npl.co.uk/mathematics-scientific-computing/mathematics-and-modelling-for-metrology/software-to-support-iso/ts-28037-2010\(e\)](http://www.npl.co.uk/mathematics-scientific-computing/mathematics-and-modelling-for-metrology/software-to-support-iso/ts-28037-2010(e))

### 6.4.2 Unweighted fit and Monte Carlo

#### Doing the fit

Fitting a straight line to measured data points is a common procedure that is built into most programming languages. For example, in Excel, the formula<sup>35</sup>:

=INDEX (LINEST (\$D\$5:\$D\$14, \$A\$5:\$A\$14), 1) will give you the slope of a straight line fit through the points with x-values in A5:A14 and y-values in D5:D14. The offset is given by:

=INDEX (LINEST (\$D\$5:\$D\$14, \$A\$5:\$A\$14), 2)

---

<sup>35</sup> INDEX(...,1) gives the first value in an array. INDEX(...,2) gives the second value. LINEST(y-values,x-values) returns an array where the first value is the slope and second value the offset.

In Matlab the equivalent command is written

```
c = polyfit(x, y, 1);
```

where

$x$  is a vector of measured  $x$ -values (here digital numbers) for one band at several radiance intensities. The length of  $x$  is  $N$  where  $N$  is the number of radiance levels multiplied by the number of measured values for each level (i.e. the total number of measurements).

$y$  is a vector of the measured  $y$ -values. Here a vector of the APEX convolved input radiance, of same length as  $x$  with replications of the same values for each radiance level.

$c$  is a vector with two elements holding the slope ( $1/G_{\text{APEX}}$ ) and offset ( $DN_0$ ) of the linear fit.

Both Excel and Matlab use "least squares analysis" to determine the fit, although with some minor differences in approach. Least squares analysis is a standard approach to determine the 'best' solution to a problem where there are more equations than unknowns. It determines the overall solution that minimises the sum of the squares of the residuals (the differences between the raw data and the fitted line).

So, for a straight line, described by  $y = a + bx$ , the function being minimised<sup>36</sup> is

$$S = \sum_{i=1}^N [y_i - (a + bx_i)]^2. \quad (6.4)$$

For a straight line, and for ordinary, unweighted, least squares, the solution can be calculated analytically using the following steps:

- Step 1: Calculate the means of  $x_i$  and  $y_i$  :  $x_{\text{mean}} = (\sum x_i)/N$  ,  $y_{\text{mean}} = (\sum y_i)/N$
- Step 2: Calculate the difference between each data point and the mean and sum the products:

$$T = \sum_{i=1}^N (x_i - x_{\text{mean}})(y_i - y_{\text{mean}})$$

$$B = \sum_{i=1}^N (x_i - x_{\text{mean}})^2$$

- Step 3: Calculate the slope and intercepts:

$$b = \frac{T}{B} = \frac{\sum_{i=1}^N (x_i - x_{\text{mean}})(y_i - y_{\text{mean}})}{\sum_{i=1}^N (x_i - x_{\text{mean}})^2}$$

---

<sup>36</sup> This is not exactly how matlab does it. The polyfit routine can deal with polynomials and the straight line is a simple case.

$$a = y_{\text{mean}} - bx_{\text{mean}} \quad (6.5)$$

This unsophisticated straight line fit assumes that all the data points have the same weight, that the variability is due to random noise only and there is no uncertainty associated with the  $x$  values. It also does not provide an uncertainty associated with the calculated slope and intercept.

### **Uncertainty estimate using Monte Carlo**

The unsophisticated straight line fit described above, and calculated using Equation (6.5) or the Excel or Matlab expressions given above, does not provide the uncertainties and covariance associated with the slope and intercept of the fitted straight line. While it is possible to calculate the uncertainty and covariance analytically (as described in the sections below), it is also worth<sup>37</sup> considering an alternative – the use of Monte Carlo simulation.

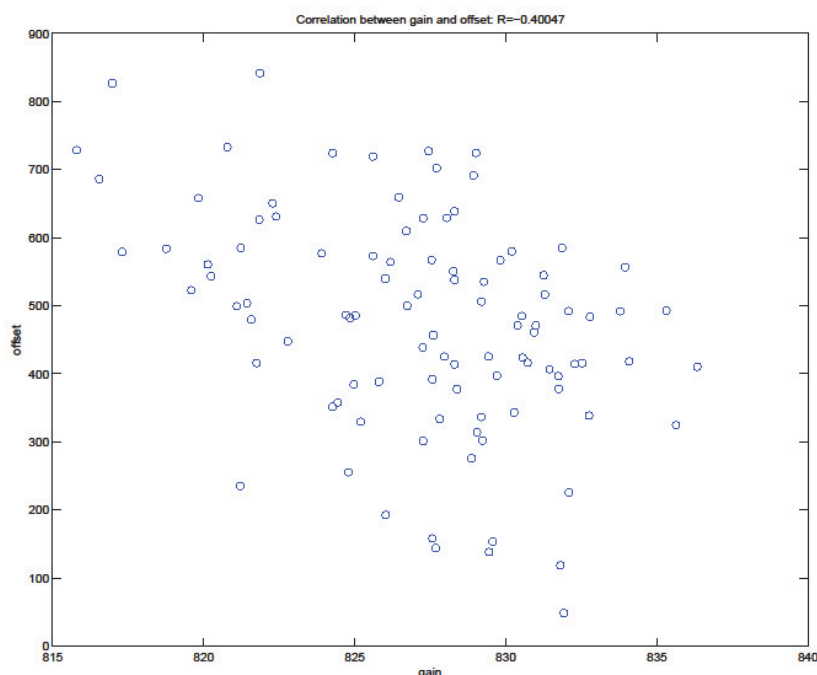
In Monte Carlo simulation, many (1000 or more) simulations are run, in each of which the input quantities are adjusted by a random variable within their uncertainties. In the example being considered here, we create a model for a perfect ‘true’ answer – by assigning an arbitrary but realistic value to the gain and offset and calculating true signal-system radiance pairs  $(x_i, y_i)$  values. We then run a loop where in each loop errors are drawn from the appropriate probability distribution functions (usually a Gaussian distribution with a standard deviation equal to the associated standard uncertainty). In this case, for the signal (which is uncorrelated from point to point), a different error is drawn for each  $y_i$  within one loop. For the system radiance (which is fully correlated) the same error is drawn for all  $x_i$  (or the same relative error is drawn and multiplied by each  $x_i$ ) within one loop. In the next loop, new random numbers are drawn from the distributions.

At the end of each loop, a fit is applied to the data, obtaining a pair for the slope and intercept of the straight line. These pairs can be plotted, as shown in Figure 17. Having run a large number of loops, the uncertainty associated with each (slope and intercept) can be determined from the standard deviation of the values for that parameter. The covariance can be calculated using Equations (3.46) and (3.47).

---

<sup>37</sup> The solution described here is pragmatic. It is often the case that some form of fitting or interpolation or other algorithm is used in a processing chain and that the algorithm has been determined with little thought for rigorous uncertainty analysis. Later an uncertainty budget is put together. In those cases, Monte Carlo provides a simple method to ‘upgrade the analysis’ to include uncertainties without rethink. The Monte Carlo analysis described here does not give the same answer as the more sophisticated approaches of the subsequent sections because the initial fit is still done with an unsophisticated unweighted least squares routine. The more sophisticated approaches take the uncertainties (and covariance) into account in the fit.





**Figure 17** Plots of offset and gain for different loops of the Monte Carlo simulation. This shows a possible mild negative correlation (high values of one parameter tend to occur with low values of the other parameter).

### 6.4.3 Weighted fit and analytical covariance

#### Doing the fit

In this case we consider fitting a straight line to measured data points with some associated uncertainties which may differ from point to point. For example, if the uncertainty associated with noise is higher at the lower radiance level measurements than for the higher radiance level measurements, then the fit should have a stronger weighting at the higher radiance level measurements. The method described here calculates both a slope and intercept for the straight line and their associated uncertainties and covariance using the uncertainties associated with the  $y$ -values. There is assumed to be no uncertainty associated with the  $x$ -values. For the example considered here, it therefore makes sense to have the  $x$ -axis the radiance levels and the  $y$ -axis the signal (digital numbers) because the different radiance levels are (assumed to be) fully correlated, whereas the signal values have uncertainties associated with random effects.

The calculation of the slope and offset is as follows:

The weights<sup>38</sup> are defined as

---

<sup>38</sup> Because this term is squared in the subsequent equation, the actual weight is inversely proportional to the square of the uncertainty.



$$w_i = \frac{1}{u(y_i)} \quad (6.6)$$

where  $u(y_i)$  is the uncertainty associated with the measured value  $y_i$  (here the signal) at the set value  $x_i$  (for a given radiance level).

The reference values are given by

$$x_0 = \frac{\sum_{i=1}^N w_i^2 x_i}{\sum_{i=1}^N w_i^2}, \quad y_0 = \frac{\sum_{i=1}^N w_i^2 y_i}{\sum_{i=1}^N w_i^2}. \quad (6.7)$$

The slope (here  $1/G_{\text{APEX}}$ ) is then calculated as

$$b = \frac{\sum_{i=1}^N w_i^2 (x_i - x_0)(y_i - y_0)}{\sum_{i=1}^N w_i^2 (x_i - x_0)^2} \quad (6.8)$$

and the intercept (here  $DN_0$ ) as

$$a = y_0 - bx_0. \quad (6.9)$$

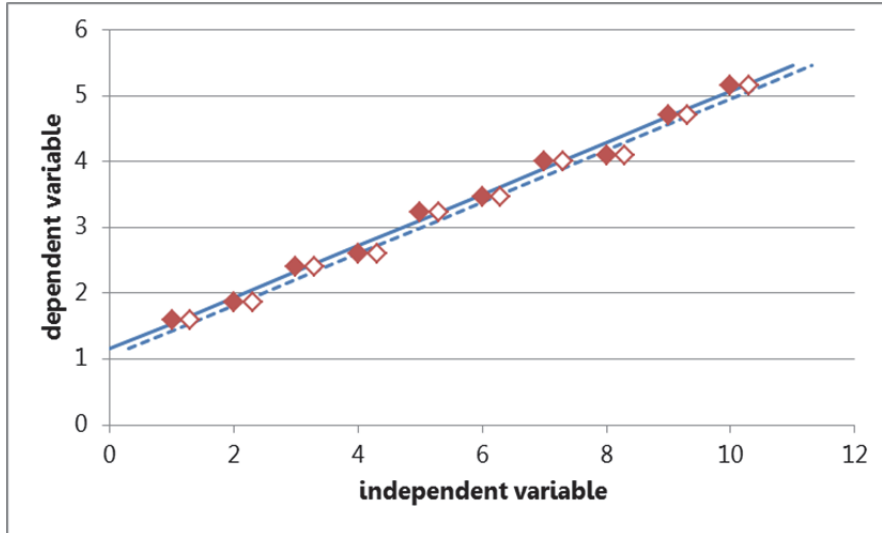
The variance (squared uncertainty) and covariance associated with the slope and intercept are given by

$$\begin{aligned} u^2(b) &= \frac{1}{\sum_{i=1}^N w_i^2 (x_i - x_0)^2}, \\ u^2(a) &= \frac{1}{\sum_{i=1}^N w_i^2} + \frac{x_0^2}{\sum_{i=1}^N w_i^2 (x_i - x_0)^2}, \\ u(a, b) &= \frac{-x_0}{\sum_{i=1}^N w_i^2 (x_i - x_0)^2}. \end{aligned} \quad (6.10)$$

### **Uncertainty analysis**

Here the uncertainty (and covariance) is given by Equation (6.10) under the initial conditions, i.e. that the uncertainties are only associated with random effects in the  $y$ -values (the signal). We did not take into account the uncertainties associated with the  $x$ -values (the radiance levels). Earlier we concluded that the radiance levels can be considered fully correlated. Therefore there is a systematic effect on the  $x$ -values. If the error is an absolute error, an additive effect in the same units, then it would have the effect of shifting the entire straight line to the left or right. This would have no effect on the slope (hence instrument gain) or its associated uncertainty. It would, however, affect the offset (Figure 18). A rigorous uncertainty analysis requires the full covariance matrix, but an indicative one can be determined by treating this as an uncertainty associated with the mean value,  $x_0$  in Equation (6.9). The uncertainty associated with the intercept in Equation (6.10) would have an

additional term  $+ b^2 u^2(x_0)$ . This is by no means a rigorous analysis, but it provides an indicative uncertainty.



**Figure 18** An error in the  $x_i$  that is common to all values will shift the fitted curve (in this case to higher values). This has no effect on the slope, but will change the offset.

#### 6.4.4 Rigorous analysis

A rigorous analysis would take into account the uncertainties and covariance associated with both the  $x_i$  and  $y_i$ . A recipe for doing this calculation is given in the British Standard Standard DD ISO/TS 28037:2010 "Determination and use of straight-line calibration functions" and Matlab code is given in the link on page 69. The most general case is given in TS28037\_GGMR1.

The software requires as input:

- The x-values in a vector
- The y-values in a vector
- A covariance matrix for the x-values (the top left quarter of the matrix in Equation (6.3))
- A covariance matrix for the y-values (the bottom right quarter of the matrix in Equation (6.3))

As output, the software provides:

- The value of the slope ( $b$ )
- The value of the intercept ( $a$ )
- The variance (squared uncertainty) of the slope  $u^2(b)$
- The variance of the intercept  $u^2(a)$
- The covariance of the slope and intercept  $u(a, b)$
- Validation information for the model in the form of a chi-squared test (see below).

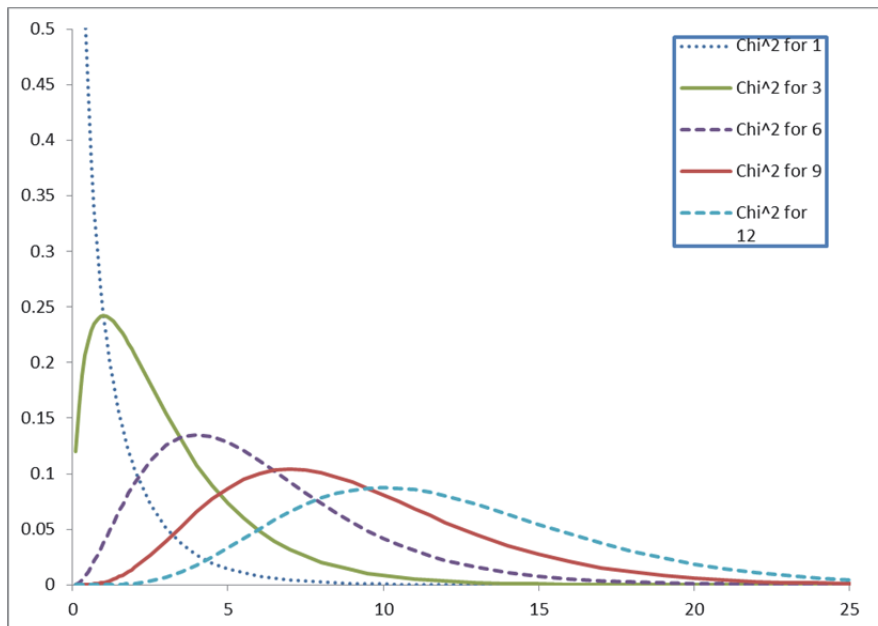
## 6.5 Validating the fit

The determination of the associated uncertainties given above assumes that the straight line fit is a 'good fit' to the measured data. This should be tested, using a statistic such as the chi-squared statistic. Applying the chi-squared test involves two parts: determining the observed chi-squared,  $\chi_{\text{obs}}^2$ , and then comparing it with an appropriate chi-squared distribution.

The observed chi-squared is calculated as the sum of squares of the weighted residuals from the measured data to the fit. For  $N$  measured values, we calculate

$$\chi_{\text{obs}}^2 = \sum_{i=1}^N \left[ \frac{y_i - (a + bx_i)}{u(y_i)} \right]^2 \quad (6.11)$$

Because it is calculated from statistical data, we expect the value of  $\chi_{\text{obs}}^2$  that we obtain to be a sample from the appropriate statistical distribution. The distribution depends on the number of 'degrees of freedom'. This is the number of measurements minus the number of fit parameters, or here  $\nu = N - 2$ . As the number of degrees of freedom increases (i.e. as more and more values were used to determine the fit), then the  $\chi_{\nu}^2$  distribution becomes more symmetric and shifts to the right. The centre of mass of the distribution is the degrees of freedom (that's the shift to the right).



**Figure 19 Chi squared distributions for different degrees of freedom**

The chi-squared test evaluates the probability that a particular observed value,  $\chi_{\text{obs}}^2$ , came from the expected distribution  $\chi_{\nu}^2$ . There are two common versions of this test. The first tests whether the value is less than the 95 % percentile of the expected distribution (i.e. it would only be failed 5 % of the time for good data). The second, the Birge ratio, compares the observed chi-squared with the expectation value (i.e. the number of degrees of freedom).

When the chi-squared test is failed it implies that the model does not adequately explain the data and a better model is required.

## 6.6 Using the fit

### 6.6.1 Analytical covariance when the fit is used

However we determine the covariance, whether<sup>39</sup> analytically or through Monte Carlo simulation, at the end of this process we have the equation  $y = a + bx$  with the uncertainties associated with  $a$  and  $b$  and their covariance. The final step is to calculate for a given, measured  $y$  the associated  $x$ . For the APEX calibration, we have determined the parameters using the provided radiance levels  $L_{\text{sph-filt},i}$  as the  $x_i$  and the measured signal  $DN_{\text{APEX,cal},i}$  for that radiance as the  $y_i$ . We now need to calculate the radiance of a scene observed by APEX from a measured signal. This is done using Equation (6.1). Consider first the generic equation:

$$\begin{aligned} y &= a + bx \\ x &= (y - a)/b \end{aligned} \quad (6.12)$$

Therefore, later, when the instrument is used, we measure a signal  $y$  (with an associated uncertainty) and convert it to the parameter  $x$ . Here we measure  $(y)$  the scene signal in digital numbers (with noise in that scene signal) and we convert it to the scene radiance  $(x)$ , using the second version of Equation (6.12). We have the sensitivity coefficients

$$\begin{aligned} c_y &= \frac{\partial x}{\partial y} = \frac{1}{b} \\ c_a &= \frac{\partial x}{\partial a} = -\frac{1}{b} \\ c_b &= \frac{\partial x}{\partial b} = \frac{-(y - a)}{b^2} \end{aligned} \quad (6.13)$$

Applying the full law of propagation of uncertainties, gives

$$\begin{aligned} u^2(x) &= c_y^2 u^2(y) + c_a^2 u^2(a) + c_b^2 u^2(b) + 2c_a c_b u(a, b) \\ &= \left( \frac{u(y)}{b} \right)^2 + \left( \frac{u(a)}{b} \right)^2 + \left( \frac{(y - a)u(b)}{b^2} \right)^2 + \frac{2(y - a)u(a, b)}{b^3}. \end{aligned} \quad (6.14)$$

It may help coding this into a language like MATLAB to write this same equation in matrix notation:

<sup>39</sup> Note that there won't be exactly the same values from these two processes, but they are likely to be similar

$$u^2(x) = \mathbf{c}^T \mathbf{V} \mathbf{c},$$

$$\mathbf{c} = \begin{bmatrix} c_a \\ c_b \\ c_y \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} u^2(a) & u(a,b) & 0 \\ u(a,b) & u^2(b) & 0 \\ 0 & 0 & u^2(y) \end{bmatrix}. \quad (6.15)$$

Translating this for the APEX calibration into its notation,  $x \equiv L_{\text{scene}}$ ,  $y \equiv DN_{\text{scene}}$ ,  $a = DN_0 = -L_0/G_{\text{APEX}}$  and  $b \equiv 1/G_{\text{APEX}}$ .

## 6.7 Straight line interpolations

The previous sections of this chapter have related to using a fit. Sometimes, however, data is interpolated rather than fitted. This is done when the data points are a reliable indication of the desired function and additional interpolated data points are required. It is often the means to obtain spectral information where measurements are made at a subset of the full set of wavelengths. The simplest interpolation is a straight line (linear) interpolation between data points (joining the dots in a graph with straight lines).

A linear interpolation will take the form

$$y_i = y_0 + (x_i - x_0) \frac{y_1 - y_0}{x_1 - x_0} \quad (6.16)$$

where the measured data points  $(x_0, y_0)$  and  $(x_1, y_1)$  have a spacing  $\Delta x = x_1 - x_0$  horizontally, and  $\Delta y = y_1 - y_0$  vertically. We are interested in the uncertainty associated with the interpolated value  $y_i$  due to uncertainties associated with  $y_0$  and  $y_1$ . We assume the horizontal scale has no associated uncertainty for simplicity.

The sensitivity coefficients are

$$\frac{\partial y_i}{\partial y_0} = 1 - \frac{(x_i - x_0)}{(x_1 - x_0)} = \frac{x_1 - x_i}{x_1 - x_0},$$

$$\frac{\partial y_i}{\partial y_1} = \frac{x_i - x_0}{x_1 - x_0}. \quad (6.17)$$

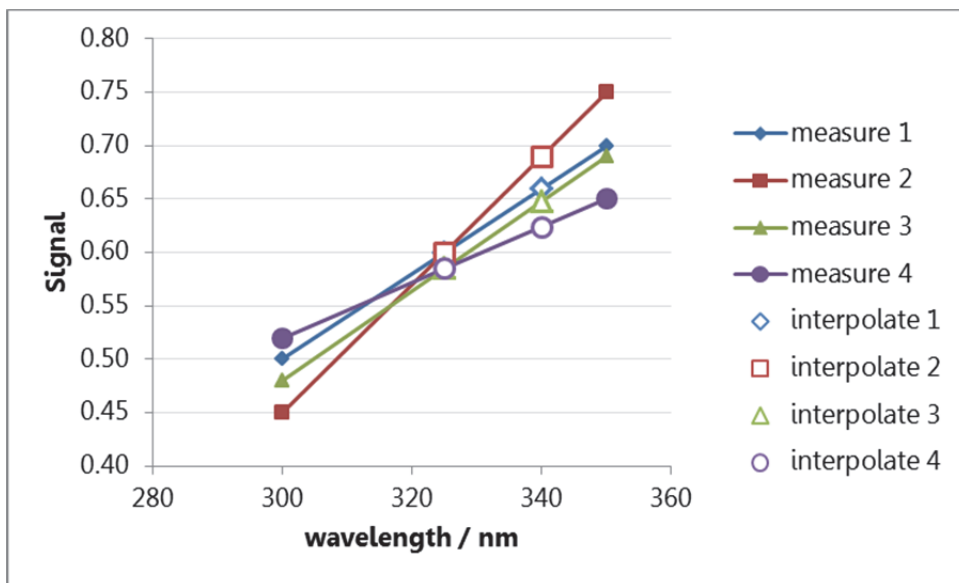
Note that both these sensitivity coefficients are fractions of the horizontal axis spacing. The first one, the sensitivity of the interpolated value to the first (left hand) point, is the fractional distance from the interpolation wavelength to the right hand point – i.e. the further it is from the right hand point, the more sensitive it is to variations in the left hand point. And vice versa for the second sensitivity coefficient.

For a point in the middle, where both sensitivity coefficients are  $1/2$ , the determined interpolated value, is equally sensitive to both ends. The uncertainty associated with the interpolated value is therefore given as

$$u(y_{\text{midpt}}) = \left[ \left( \frac{1}{2} \right)^2 u^2(y_0) + \left( \frac{1}{2} \right)^2 u^2(y_1) \right]^{1/2} \quad (6.18)$$

And, if the two measured data points have equal uncertainties, the uncertainty associated with the interpolated value is reduced by  $\sqrt{2}$ . This is clearly understood, as the interpolated value is the average of the two end points – and we have already seen that averaging reduces the uncertainty by the square root of the number of measured points.

When the interpolated point is elsewhere than the midpoint, it will be more sensitive to the measured value it is closer to, and the uncertainty reduction will be smaller (assuming both measured data points have equal uncertainties). This can be seen graphically in the following diagram. The lines and solid points represent different random measurements of the two end points. Clearly the interpolated points are closer together than the two end points. Interpolation always appears<sup>40</sup> to reduce uncertainties and the longer the distance over which the interpolation is carried out the more the reduction (the central points have smaller uncertainties than those closer to measured values).

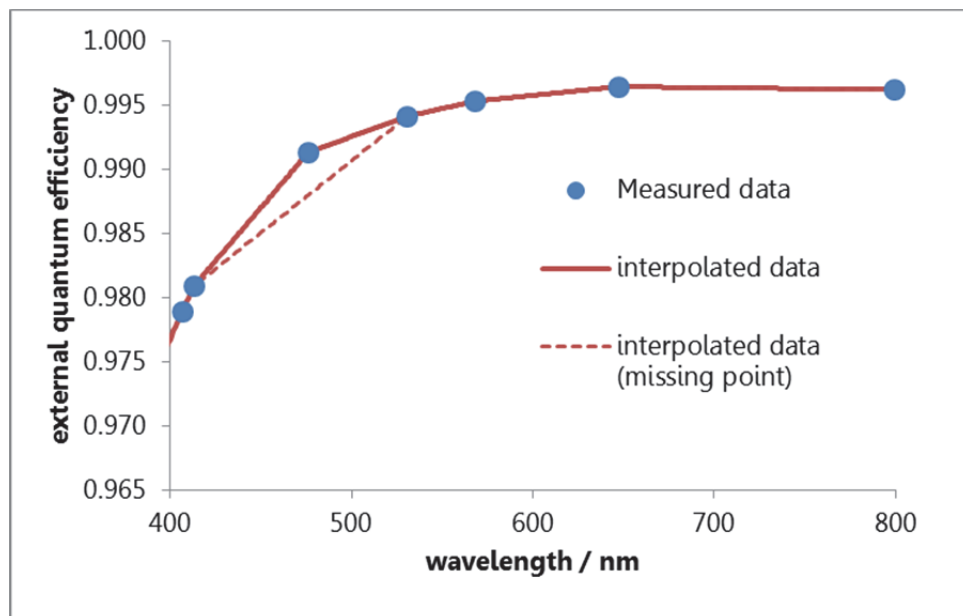


**Figure 20 A linear interpolation between measured values at 300 nm and 350 nm. The spread of values at the two ends (at 300 nm and at 350 nm) is due to measurement uncertainty. It is clear that a point in the centre experiences a smaller range of values (and hence a smaller associated uncertainty) than the original measured values.**

Intuitively though, we realise that interpolation does not really reduce uncertainties. As with fitting, a blind application of the Law of Propagation of Uncertainties without thought can falsely underestimate uncertainties. Here, also, we must consider whether a straight line fit is an appropriate fit to the measured data. Consider, for example the following diagram. This is

<sup>40</sup> See below. Interpolation reduces uncertainties from a mathematical perspective, assuming that the interpolation function is appropriate.

the spectral measured values for the external quantum efficiency of a trap detector, a standard reference silicon detector used at NPL. The blue dots are measured values, and the red line a linear fit.



**Figure 21 Measured and interpolated data for the quantum efficiency of a reference silicon trap detector. The 476 nm measured value is essential to determine the correct shape.**

If the measured value at 476 nm is removed, then the straight line fit would provide the dashed interpolation, with low uncertainties at 476 nm (as it is almost half way between the other measured data points). This uncertainty would woefully underestimate the true error at the missing data point. Again, as with fitting, some test needs to be performed to estimate the quality of the interpolation. Methods to do this include:

- Use a higher order fit as well (say a cubic spline), and compare the interpolated values for the different fit parameters
- If there is a large number of data points, interpolate taking out every other data point and compare the results with those points missing to the results with the points included
- Take additional experimental values where you intuitively feel that the fit quality may be poor (our hand drawn lines are generally closer to the true value than linear fits).

## 7 Spectral Selection

Most radiometric earth observation sensors have some spectral selection to obtain spectral information of the scene. Whether this is in the form of a few broad spectral bands, defined by a filter, or of a high resolution spectral dispersion in a spectrometer, spectral selection provides wavelength-dependent information and has its own uncertainty requirements. This section discusses the uncertainty aspects of spectral selection.

In many imaging systems a two-dimensional array provides simultaneous spectral and spatial selection. Many of the concepts described here for the spectral channels also apply to the spatial channels.

### 7.1 Spectral response function

The two most common methods for spectral selection are to place a filter in front of an individual detector, or to have a dispersive element (a grating, or occasionally prism) and an array detector. Whichever method is used, the instrument is defined by the 'spectral response function' (SRF), which is the spectral responsivity of the combined instrument (including both the transmittance – or reflectance – of the filter or dispersive element and the responsivity of the detector and spectral transmittance through any other optical element in the path e.g. telescope).

The most accurate way to determine the spectral response function is to scan a spectrally tuneable laser across the spectral response region of interest of the instrument, and a little beyond, and compare the response of the instrument at each laser wavelength to the response of a reference detector of known spectral responsibility. However, although ideal, tuneable lasers are not always readily available and sometimes the added complexity is unnecessary for the required accuracy. For a filter system a more common alternative is to obtain tuneable monochromatic radiation using a monochromator illuminated by a lamp. Here it is necessary to ensure that the bandwidth of the monochromator is sufficiently narrow not to distort the measured spectral response function. For a spectrometer-based instrument a similar process can be followed, or, making the assumption that the spectrometer's responsivity is constant over a bandwidth, the signal from adjacent spectrometer pixels can be used to determine the spectral response function.

Where a spectrometer has many pixels, for example for the APEX imager described in Section 5, which has up to 530 spectral bands for each of 1000 spatial channels, then it is not possible to determine the SRF independently for all bands (pixels). Instead the SRF is determined for several pixels using a lamp illuminated monochromator that is tuned across the spectral band of the measured pixels. A Gaussian is fitted to the measured data as the monochromator wavelength is tuned and from the Gaussian fits two parameters are obtained: the bandwidth and the centre wavelength. These are plotted as a function of pixel number and therefore Gaussian functions can be determined for all other pixels by interpolating or extrapolating the measured data. A similar method was used for the MERIS satellite imager [4].



## 7.2 Uncertainty associated with the spectral response function

The uncertainty associated with the determined spectral response function comes from several sources:

- Uncertainties relating to the spectral measurements
  - Uncertainties associated with the wavelength of the monochromatic light used (particularly the wavelength scale of the monochromator)
  - Uncertainties associated with the bandwidth of the monochromatic light used
  - Uncertainties associated with noise in the measurement of the spectral response function
- Uncertainties relating to the interpolation
  - Where a spectral response function is determined for measured data at a set of wavelengths, this relates to the interpolation between the measured data points (the monochromatic illumination used). Were sufficient wavelengths used for the illumination source to determine the full shape of the spectral response function?
  - Where a Gaussian (or similar function) is fitted to the measured spectral response function this relates to to what extent a Gaussian is a 'good fit' to the measured data points.
  - Where the SRF of only a few pixels of an array are fully characterised, and these properties are then interpolated for other pixels of the array, it relates to the quality of the SRF thus determined compared to the 'true' SRF for those intermediate pixels.
- Uncertainties relating to changes since calibration[12]
  - The SRF may change simply due to 'storage' – this is particularly true for interference-filters.
  - The SRF may change due to vibrations on transportation and launch.
  - The SRF may change due to changes of temperature, humidity and the move to vacuum.
  - The SRF may change due to both ultraviolet and high energy solar radiation damage.
  - Contamination films can change spectral and absolute levels of transmittance.

It is beyond the scope of this document to provide a rigorous analysis of all of these concepts. The size of each effect must be estimated for an instrument of interest – and this can be done through experimentation or modelling. For example, an estimate of the noise on individual SRF data values can be obtained by repeating the calibration multiple times and observing the spread of obtained results and an estimate of the effect of interpolation can be obtained by using different interpolation functions (see Section 6.6). The quality of the fit can be tested using the types of test described in Section 6.3, or by trying different models. Estimates of change since calibration can be obtained by calibration before and after use or before and after simulated exposure to a space-like environment.

Note also, that where two parameters are fitted from the same data, then there will be some correlation between the two fit parameters; this is discussed further in Section 7.3.3.

What we are really interested in estimating is the uncertainty associated with the measured value obtained by the instrument (of e.g. spectral radiance) due to the uncertainty associated with the SRF (itself a result of all these effects). To understand this we need to realise that any spectral measurement relies on a spectral integral or convolution with the spectral characteristics of the scene under observation and this is the subject of the next section.

## 7.3 Spectral integrals and convolution

### 7.3.1 Origin of spectral integrals

In Section 5, the APEX instrument calculation equation was given by Equation (5.6):

$$L_{\text{scene}} = G_{\text{APEX}} DN_{\text{APEX,scene}} .$$

If we convert this to give the measured signal on APEX as a function of the radiance of the scene we get the equation

$$DN_{\text{APEX,scene}} = L_{\text{scene}} / G_{\text{APEX}} = R_{\text{APEX}} L_{\text{scene}} . \quad (7.1)$$

where  $R_{\text{APEX}} = 1/G_{\text{APEX}}$  is the APEX responsivity. In practice, this equation oversimplifies the process because a pixel of the APEX imager does not measure a single wavelength, but has a SRF. Therefore equation (7.1) should be written as

$$DN_{\text{APEX,scene}} = \int R_{\text{SRF,APEX}}(\lambda; \lambda_0) L_{\text{scene}}(\lambda) d\lambda \quad (7.2)$$

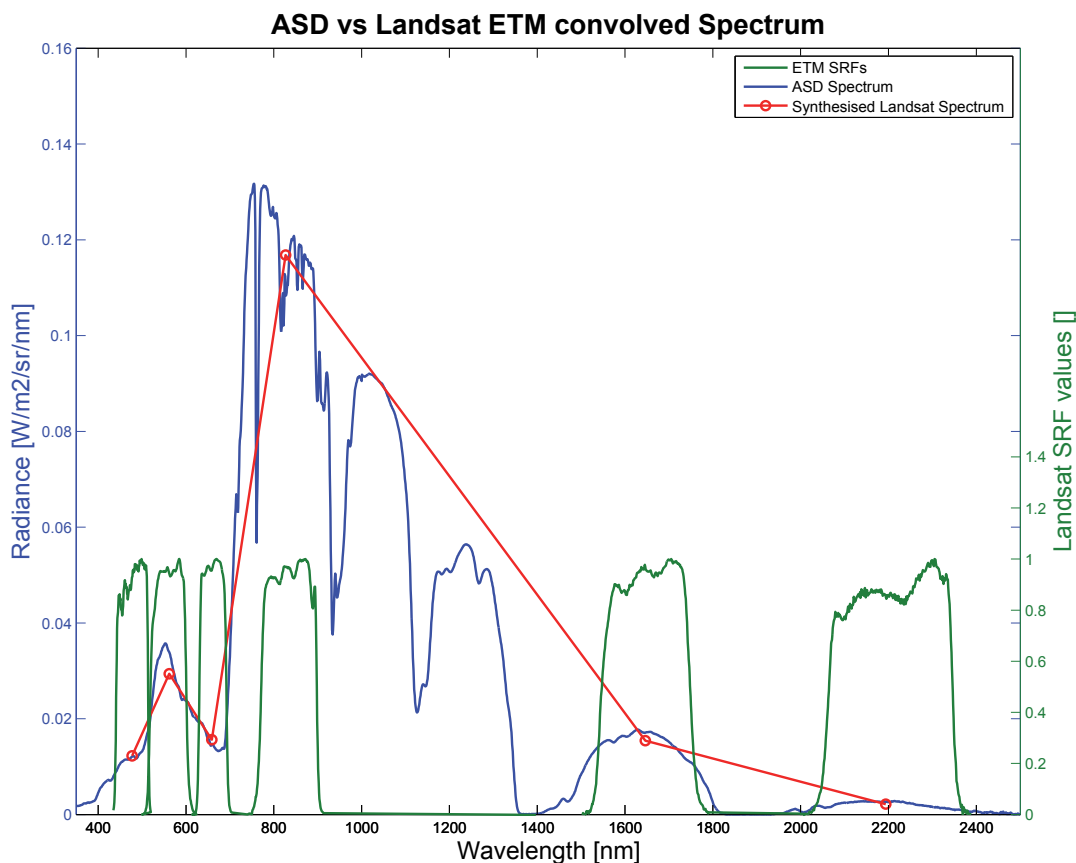
where  $R_{\text{SRF,APEX}}(\lambda; \lambda_0)$  is the spectral response function of the APEX pixel centred on wavelength  $\lambda_0$ .

Whatever the instrument, whether it uses a spectrometer or a filtered system to define the wavelength band, the real measured value will be some equation that is of the form of (7.2). The spectral response function implicitly creates a weighted integral of the scene radiance. This is sometimes described as a spectral convolution.

In some cases, where the SRF is narrow band, the subsequent analysis treats the measured value as though it were made at the centre wavelength and was truly monochromatic. In these cases, the problem is treated as one of 'instrument bandwidth' and the uncertainty analysis should consider the uncertainty associated with the assumption that the measurement is effectively made at a single wavelength.

Generally, though, for both imaging spectrometers and filtered instruments, the measured signal must be treated as a band-integrated quantity and interpretation of the results must acknowledge the implicit integral described by Equation (7.2). This interpretation usually relies on some knowledge of the spectral radiance of the scene at a higher resolution than the measurement provides. For example, in Figure 22 we model the output (red dots) of the

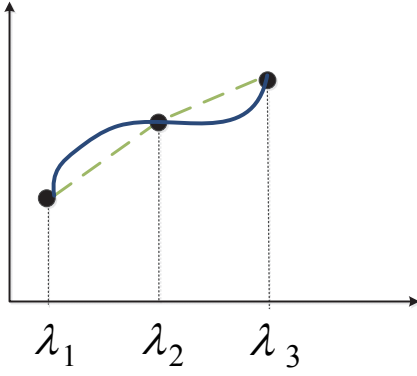
Landsat-ETM instrument by integrating a scene radiance measured with a higher resolution ASD spectrometer (blue) and the SRF of the Landsat-ETM bands (green).



**Figure 22: Example of a broadband sensor convolution, synthesising a radiance spectrum measured with an ASD spectrometer with Landsat ETM. (Note red line for indication only, only the dots have meaning – they provide band-integrated radiance values).**

### 7.3.2 Calculating the integral

In the analysis of the measured values from any sensor, there is a need to determine integrals of the form of Equation (7.2). Generally the relevant spectra are provided as discrete values at given wavelengths. A common method for evaluating such intervals is the trapezium (or trapezoidal) rule which approximates the integral by treating the integrand as varying linearly between adjacent measurement points (Figure 23).



**Figure 23 Trapezium rule diagram**

The trapezium rule replaces the integral in Equation (7.2) with a summation

$$DN_{\text{APEX,scene}} = \sum_{i=1}^N S_{\text{SRF,APEX}}(\lambda_i; \lambda_0) L_{\text{scene}}(\lambda_i) \ell_i \quad (7.3)$$

where  $\ell_i$ , which depends on the wavelength spacing either side of  $\lambda_i$ , is an appropriate weighting term. If the data are evenly spaced, such that  $\lambda_{i+1} - \lambda_i = \delta\lambda$  for all wavelength steps, then, for the trapezium rule

$$\ell_i = \begin{cases} \delta\lambda/2 & i = 1, n \\ \delta\lambda & i = 2, \dots, n-1. \end{cases} \quad (7.4)$$

Higher order rules (that fit different piecewise polynomials to the data) are also available [13] and may be suitable where the trapezium rule over-simplifies the analysis. Some care is needed if the two spectra multiplied (the scene radiance and the instrument SRF) are at different wavelength steps. In these cases at least one<sup>41</sup> of the spectra should be interpolated to the wavelengths of the other spectrum. Where one spectrum is smooth and the other 'spiky' then the smoother spectrum should be interpolated. This is a strong case for replacing the instrument SRF with a fitted Gaussian, which can then be defined at any wavelength, when it is then multiplied by a measured scene radiance with considerable structure (for example due to atmospheric or solar atomic/molecular absorption lines).

### 7.3.3 The uncertainty associated with the integral

A detailed analysis on determining the uncertainty associated with integrated quantities is available for download at [www.tinyurl.com/NPLIntegrals](http://www.tinyurl.com/NPLIntegrals). That report was written for lighting applications but many of the concepts apply equally for satellite bands<sup>42</sup>.

<sup>41</sup> There are occasions when it is appropriate to interpolate both spectra to the wavelengths of each other

<sup>42</sup> Emma Woolliams hopes to write a version of that report for satellite band applications sometime in the latter half of 2014. In the vocabulary of that report, the integral given above as Equation (7.2) is an "experimental product integral".

Section 7.2 listed different sources of uncertainty associated with the SRF itself. These different categories of uncertainty have different effects on the integral.

### Uncertainties relating to the spectral measurements

When the integral is determined using the trapezium rule from the raw measured SRF, then noise in the individual measured values of the SRF will affect the determined integral. However, the nature of noise is that it is random, therefore while one point may be determined a little higher than the true value, there is likely to be another point determined a little lower than its true value. Thus to some extent the noise will be 'averaged out' by the integration process. Integration, like averaging, reduces the effect of noise. The extent to which the noise is reduced depends on the number of measured values describing the SRF (more points increases the 'averaging out' effect) and whether the reduction is sufficient to make noise negligible depends on how noisy the data are in the first place, however, generally speaking, noise in the determination of the SRF is a tiny contribution to typical uncertainty budgets.

The uncertainty associated with the integral due to noise can be calculated either through Monte Carlo simulation (introducing noise to all data points as separate draws from a Gaussian probability distribution function) or analytically. An analytical expression is obtained from the summation in Equation (7.3). If we write the first three terms out, we get

$$DN_{\text{APEX,scene}} = R_1 L_1 \ell_1 + R_2 L_2 \ell_2 + R_3 L_3 \ell_3 + \dots \quad (7.5)$$

where  $R_1 L_1 \ell_1$  is shorthand for  $R_{\text{SRF,APEX}}(\lambda_1; \lambda_0) L_{\text{scene}}(\lambda_1) \ell_1$ . If we apply the law of propagation of uncertainties to this for an uncertainty associated with  $R_1$ , we need the sensitivity coefficient

$$c_{R1} = \frac{\partial DN_{\text{APEX,scene}}}{\partial R_1} = L_1 \ell_1 \quad (7.6)$$

And thus the law of propagation of uncertainties (considering only uncertainties associated with the  $R_i$ ) is

$$u^2(DN) = (L_1 \ell_1)^2 u^2(R_1) + (L_2 \ell_2)^2 u^2(R_2) + (L_3 \ell_3)^2 u^2(R_3) + \dots \quad (7.7)$$

Which, written in summation notation makes

$$u^2(DN_{\text{APEX,scene}}) = \sum_{i=1}^N \left[ u(R_{\text{SRF,APEX}}(\lambda_i; \lambda_0)) L_{\text{scene}}(\lambda_i) \ell_i \right]^2 \quad (7.8)$$

### Uncertainties relating to the interpolation

This category describes uncertainties associated with the original determination of the spectral response function and the assumptions made in its determination. Where discrete measured values are used to estimate the SRF this relates to whether sufficient measured

values have been made to describe the SRF<sup>43</sup>. Where the SRF is modelled by a Gaussian, this relates to the suitability of that Gaussian approximation and any loss of information from truncating the Gaussian over a finite range. Where the SRF is only determined for some bands and the SRF of intermediate bands is estimated through interpolation of, for example centre wavelength and bandwidth, this relates to the approximations introduced by those processes.

It is usually not possible to give a definitive uncertainty associated with these effects. Such a definitive uncertainty implies being able to compare the SRF used with a more accurate one – but if a more accurate one were available, it would be used. There are, however, techniques that can be used to estimate some of these uncertainties.

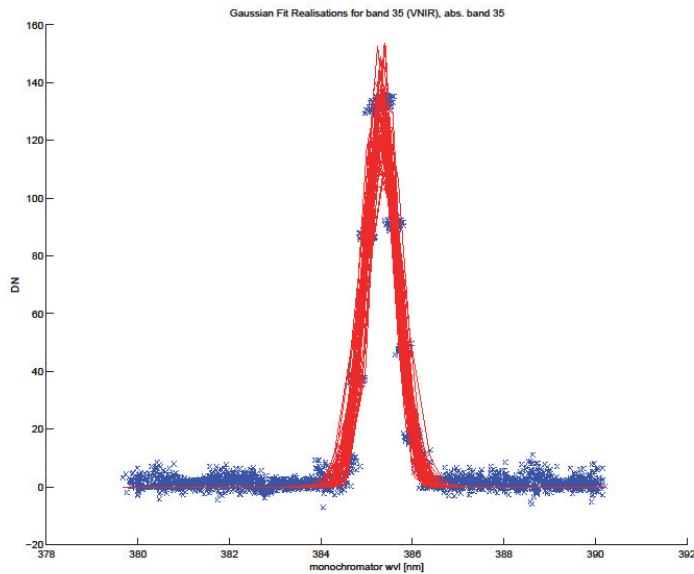
To estimate the uncertainty associated with whether sufficient measured points are available to describe the SRF and to use the trapezium rule can be estimated by comparing the trapezium rule value with the value obtained from a higher order integration rule, or by comparing a trapezium rule calculation with the raw data interpolated first using a higher order rule (such as a cubic spline interpolation). It may also be useful to recalculate the integral with half the measured data points (every other point) and compare those values. None of these methods will give a definitive answer, but they will indicate whether or not the effect can be considered negligible.

Where the SRF is modelled by a Gaussian, then an estimate of the uncertainty associated with this assumption can be made by comparing the integral calculated with the Gaussian to that calculated from the measured SRF using the trapezium rule. Because the integral involves the scene radiance, and real scene radiances are likely to be quite 'spiky', this should be done by interpolating (using e.g. linear interpolation) the measured SRF to the wavelengths at which the scene radiance is defined.

To estimate the uncertainty associated with the interpolation of the Gaussian defining parameters (bandwidth and centre wavelength) from measurements for a few pixels, it is necessary to model the process. This can be done, for example, by running a Monte Carlo simulation on the measured data. For the APEX calibration, a Monte Carlo simulation was run on the full process. First noise was added to the raw measured SRF and different Gaussians were obtained this way (Figure 24).

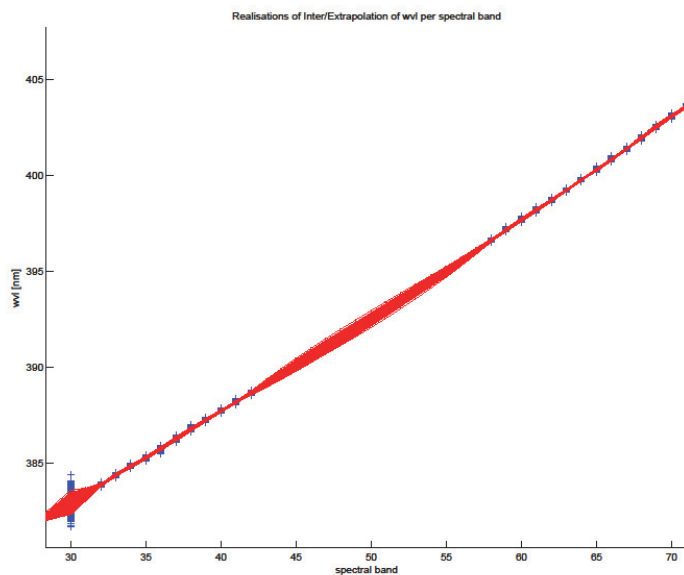
---

<sup>43</sup> Which in turn relates to whether the trapezium rule is an appropriate method for integration – because that assumes that you can join the measured points with straight lines.



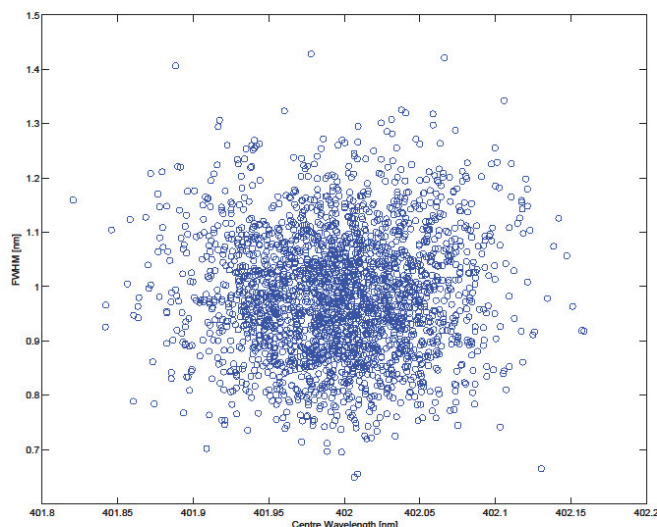
**Figure 24 Different Monte Carlo simulations of the fitted Gaussian to the raw SRF**

Then additional noise was added to account for systematic effects, and a simulation run of the interpolation to non-characterised pixels (Figure 25).



**Figure 25 Different Monte Carlo simulations to the interpolation**

From this a probability density cloud was obtained for the centre wavelength and bandwidth of one of the intermediate pixels (Figure 26). This cloud provided uncertainties associated with both the interpolated bandwidth and the interpolated centre wavelength, as well as the covariance (see Section 3.5.3).



**Figure 26 Probability density cloud of the bandwidth and central wavelength from the Monte Carlo Simulation for band 68 of the APEX VNIR detector**

From the uncertainty associated with the bandwidth and central wavelength, we need to determine the uncertainty associated with the integrated quantity calculated in Equation (7.2). This can be done either by continuing the Monte Carlo simulation and determining the integral numerically for each Gaussian obtained, or by differentiating the Gaussian analytical equation with respect to the centre wavelength and bandwidth respectively and thus obtaining the sensitivity coefficients needed for the law of propagation of uncertainties.

### Uncertainties relating to the wavelength scale of the SRF

Where the SRF is determined experimentally using a lamp-illuminated monochromator (rather than a laser), there will be uncertainties associated with the wavelength scale of the monochromator and with the bandwidth of the monochromator making the measurements. The monochromator bandwidth will have the effect of 'compressing' the peak of the SRF, making it broader and shorter.

Wavelength uncertainties fall into three<sup>44</sup> categories –

- those due to a systematic spectral offset that applies to measurements at all wavelengths (a wavelength scale error)
- those due to a wavelength dependent offset – where the same error occurs every time the monochromator is set to a given wavelength, but this error is random from one wavelength to the next
- a random effect (the reproducibility of the wavelength scale over short time periods).

---

<sup>44</sup> See also section 7.3 of the Integral Uncertainty report found at [www.tinyurl.com/NPLintegrals](http://www.tinyurl.com/NPLintegrals), which describes three types of wavelength effect.



Generally speaking, the systematic spectral offset, and some of the wavelength-dependent offset will be corrected through calibration of the wavelength scale of the monochromator. There may, however, be a residual uncertainty associated with these calibrations. When the SRF is determined on multiple occasions, then random errors in the wavelength scale will show up as random effects in the measured values and these can be treated as described above.

A systematic spectral offset (common to all wavelengths) will shift the SRF to longer or shorter wavelengths. How significant this is depends on how quickly the scene radiance changes as a function of wavelength. Consider, for example the case of a SRF that is close to an atmospheric absorption feature. A small wavelength shift could have a significant impact on the interpretation of a measured signal if the shift changes whether or not that absorption feature is considered to be within the bandpass of the sensor. In contrast, if the scene radiance is only slowly changing with wavelength, then a small wavelength shift will have almost no impact. It is best to estimate the uncertainty associated with spectral effects by modelling such a wavelength shift for typical scene spectra.

A spectral offset that varies with wavelength will almost always have a smaller impact than a systematic spectral offset especially when there is no correlation from one wavelength to the next. This is because of the averaging effect of the integral and the fact that it is reasonable to assume that some offsets will be to longer wavelengths and others to shorter wavelengths. Again, the uncertainty associated with the integral due to this effect will depend on the scene radiance itself and how that changes with wavelength. It is often reasonable to assume this effect is negligible, although modelling could be used to estimate the size of it.

### **Uncertainties related to changes since calibration**

The different effects described in Section 7.2 as uncertainties relating to changes since calibration will all affect either the absolute level of the SRF (the instrument gain), or they will change the shape of the SRF – shifting it in wavelength, broadening it or creating a spectral tilt. The uncertainty associated with changes since calibration can only be estimated from an understanding of the likely changes and modelling what effect that has on the integral[12]. The impact of these effects will be larger if the scene radiance has sharp spectral features.

## **7.4 Stray light (out of band)**

Out-of-band stray light is where light is scattered onto the 'wrong' pixel. Spectral stray light is where light of one wavelength is measured as though it were at a different (wrong) wavelength, due to scattering mechanisms within the instrument. One method to correct for out of band stray light is to use a cut-on filter to remove all light below the filter's cut-on wavelength. Signals measured at the lower wavelengths are entirely due to stray light and can be removed by subtraction, perhaps weighting the subtracted signal by the long wavelength transmission of the cut-on filter. Using a series of bandpass or blocking filters to restrict the range of wavelengths entering the spectrometer allows measurements to be made over a truncated spectral range without the influence of stray radiation at wavelengths

outside this range. If a series of such bandpass filters are used, each tailored for a given spectral region, it is possible to make measurements over a broad spectral range whilst still ensuring good stray light performance is achieved [14].

Stray light can also be characterised and corrected using a monochromatic source. A tuneable laser (or, with care, a narrow bandwidth monochromator) is used to scan sequentially through each wavelength in turn. The response at other wavelengths to this monochromatic light is used to create a stray light correction matrix which can then be used to correct stray light in any measured spectrum [15-17].

If uncorrected, stray light can cause very large errors in the spectrum measured by an array spectrometer, even errors greater than 100 % in the blue end of a solar spectrum measured using an array spectrometer with significant short wavelength stray light. Such an error will change the value of any integral calculated from that spectrum, and the size of the effect will depend on where in the spectrum the stray light occurs. If corrected, the correction method will introduce mathematical correlation to the measured spectral irradiance, which is best described using a covariance matrix.

Stray light can also come from out-of-field (spatial stray light). This can be significant, especially if the scene viewed is inhomogeneous. For example, if one pixel is viewing vegetation and an adjacent pixel a cloud or water there could be a significant effect on the observation with the bright pixel scattering light into the dim pixel. This effect can be modelled from a pre-flight determination of the Modular Transfer Function (MTF) of the instrument. Some post-launch evaluation can be carried out, particularly for high spatial resolution sensors where the impact is most pronounced, using chequer board or edge targets with sharp high contrast edges.

## **7.5 Spectral and spatial effects**

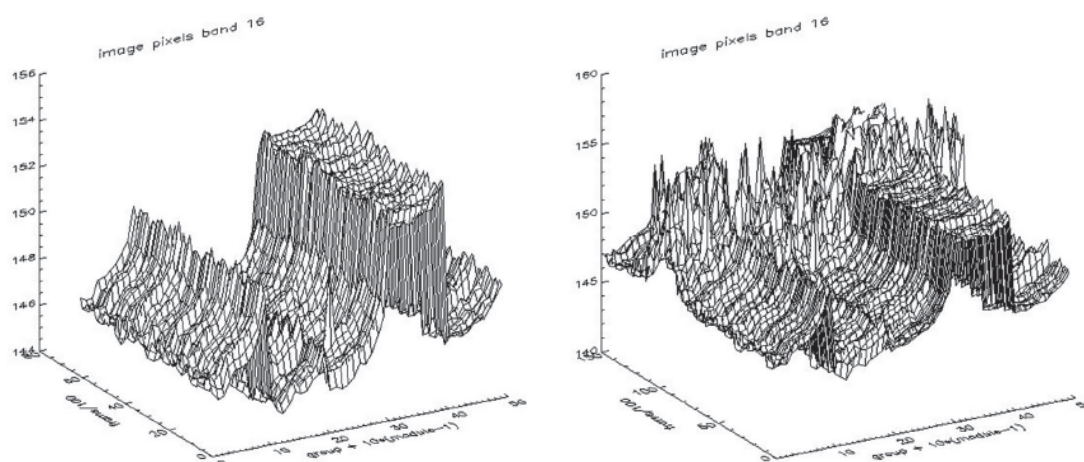
This chapter has provided an overview of some of the considerations required for uncertainty analysis for spectral and spatial effects. Generally, these effects have to be modelled with a good instrument model and the sensitivity coefficients determined numerically. This is still the subject of research, and this chapter has simply introduced some methods that can be used.

## 8 Post-launch calibration and Level-1 EO products radiometric uncertainty

### 8.1 Imager changes in orbit

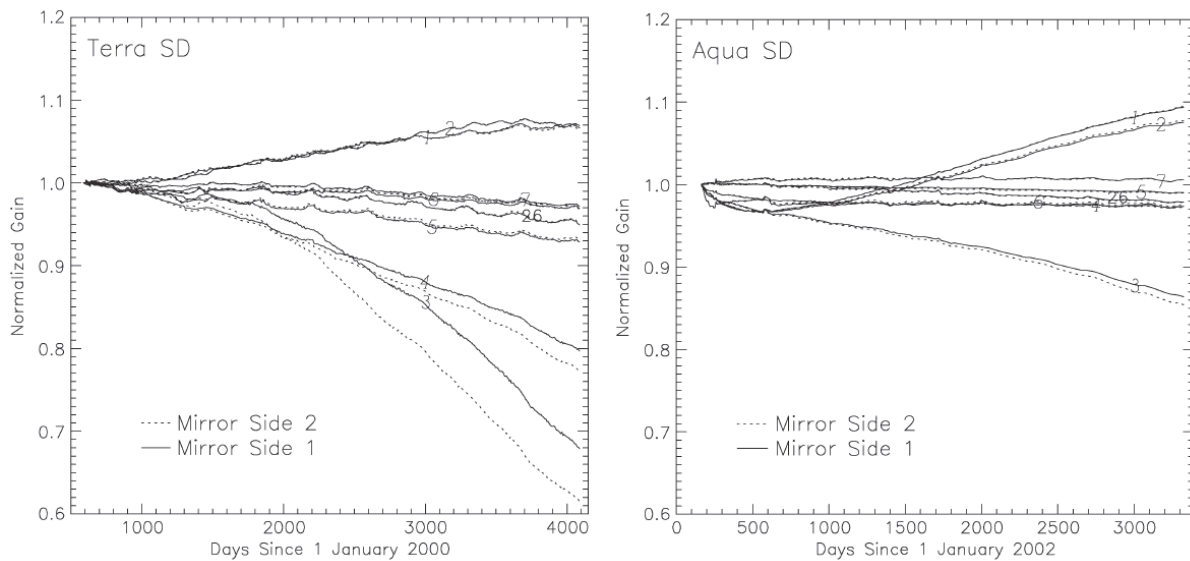
The spectral response function of a satellite sensor is likely to change on the transition to orbit (through the vibration and temperature cycles of launch, through the difference between vacuum and air operation) and once in orbit (**reversibly** due to temperature cycles, or **irreversibly** due to solarisation and other damage). Such changes affect the SRF, and similar effects will also alter the instrument gain, dark offset and linearity as components, such as mirrors or electronic systems are damaged by high energy radiation and vibration.

An example of **reversible effect** is the impact on the Dark Signal levels when measuring near an area named the South Atlantic Anomaly (SAA), Figure 27. At this area, high energy particles are confined in the inner Van Allen belt which is translated into spontaneous peaks in the readings levels as detected by the MERIS sensor in [4].



**Figure 27 MERIS Dark offset during orbit 252, OCL-ON (left), orbit 292, OCL-ON (right), band 16 (smear band), vs. pixels (front axis) and time. Spikes are clearly visible at the right image when crossing the South Atlantic Anomaly (SAA).**

However, this same effect of space radiation as well as other effects like the radiation of the sun itself, the changes of temperature and so on; can produce long term drifts in the instrument gain and spectral responses which are accounted as **irreversible effects**. The example in Figure 28, from [18] shows how the gain of MODIS Terra and Aqua gains as changed during the mission lifetime as monitored by the sun diffuser calibration system.



**Figure 28 Trends of Terra and Aqua MODIS gains for bands 1–7 and 26 obtained from the SD/SDSM measurements. The solid and dashed lines are for mirror sides 1 and 2, respectively, from [18]**

It is nevertheless true that some of these changes can be well characterised and modelled pre-flight. However, these models have a limited reliability and cannot account for all scenarios (e.g. contingencies during the operational phase) and they typically need to rely on a validation and/or update in-flight.

Such an update is done using an optics degradation model [19] which provides an exponential trend for the degradation of the optics in space environments. It was firstly used for the SeaWiFS mission and later on in the MERIS mission, in both cases it was necessary its validation and/or update using either Earth, moon or sun diffuser measurements.

The pre-flight calibration and characterisation provides a record of the instrument performance prior to launch and it is a useful indication and/or correction of the performance once in orbit. However, it is necessary to monitor the instrument performance in-flight so that the effects (reversible or irreversible) of working in a space environment are well validated and/or updated in the instrument calibration and characterisation.

## 8.2 On-board calibration systems

One method for characterising changes during flight, is to provide on-board calibration systems. For example, the on-board spectroradiometric calibration assembly for MODIS provides a characterisation of the SRF (spectral response function) and also a radiometric and spatial calibration. MODIS also has a blackbody as a prime calibration source for the mid and long-wave infrared bands ( $3.5\ \mu\text{m}$  –  $14.4\ \mu\text{m}$ ). In-flight blackbodies operating at temperatures around the instrument ambient temperature (typically<sup>45</sup> 273 K – 315 K) are commonly used as references for infrared instruments as these can be made to be

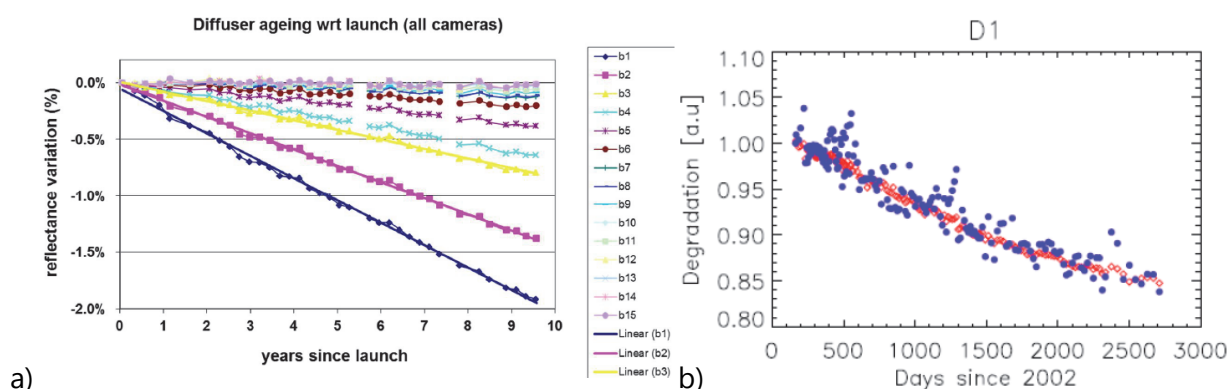
---

<sup>45</sup> although other temperatures are used depending on the application

reasonably stable. Finally MODIS, like many other earth observation satellites, has an on-board diffuser which can be placed in front of the earth imager to reflect (diffusely) sunlight. This provides a radiance-based reference, allowing top of atmosphere (TOA) radiance to be determined by reference to a value of the solar spectral irradiance<sup>46</sup>. MERIS had an additional 'pink diffuser' for spectral calibration (with absorption lines used to check the MERIS spectrometer wavelength scale) as well as the white diffuser. Some satellites have on-board lamps as radiometric references.

The most important thing to note with on-board calibration systems is that they themselves also require pre-flight radiometric characterisation and are also subject to change on transfer to orbit, and once in orbit due to the same mechanisms that affect the imager itself. Some such changes are more significant than others – for example, an on-board blackbody will usually be less liable to change during orbit than an on-board diffuser<sup>47</sup>.

One method to deal with this is to have a 'spare' on-board calibration reference or a monitoring system that is used less often than the main calibration reference. For example, MERIS had a second white diffuser plate that was deployed every three months to monitor the degradation of the frequently used plate<sup>48</sup>. A second approach is to have additional instruments to monitor the degradation. The MODIS satellite carries a "solar diffuser stability monitor" – an integrating sphere with nine filtered detectors (for nine narrow SRFs from 400 nm to 1000 nm) mounted on it that view a dark scene, direct sunlight and then the solar diffuser. The results of using these two systems to monitor the diffuser degradation are shown in Figure 29.



**Figure 29 On-board solar diffuser degradation a) for the different spectral bands of MERIS, b) for MODIS AQUA at 412 nm. Figures taken from a) MERIS 65<sup>th</sup> Cyclic Report [20]. And b) [21]**

<sup>46</sup> Usually obtained from other missions – although the choice of solar spectrum, and how to convolve it with the instrument SRF is itself the subject of some debate and results can be significantly changed when different choices are made. See also section 8.5.1

<sup>47</sup> Although care does need to be taken regarding the contact thermometers that measure its contact temperature and are used to predict the effective radiation temperature. Changes in thermometers can occur, and their physical contact with the blackbody can change, producing gradients and biases.

<sup>48</sup> This is based on the assumption that most changes in the diffuser are due to illumination of the panel by the sun, and in particular due to the ultraviolet exposure and therefore a diffuser used less frequently will be degraded much more slowly

Of course any 'spare' calibration reference and any monitor to check that reference will also degrade in orbit. The expectation is that the degradation of the spares and monitors will be significantly less than that of the primary reference and can be considered small (and ideally negligible) within the uncertainties sought by the mission.

Most of the effects described in this chapter are corrected for by the instrument science teams. For example, the raw spectral diffuser degradation for MODIS shown in the blue circles of Figure 29b, leads to a correction (the red diamonds) that is applied to the pre-flight calibration of the diffuser to obtain the values for in-flight use.

In section 8.5 the PTFE diffuser calibration will be used and generalised as much as possible to provide a simplified example of a post-launch uncertainty budget.

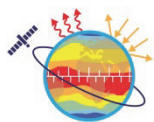
### 8.3 Vicarious cal/val

In addition to (or, in many cases, instead of) on-board calibration systems, satellites use vicarious calibration and validation (cal/val) as a means of checking the radiometric stability of their on-board instruments, as well as providing comparisons between different satellites and therefore understanding inter-satellite bias. In some cases, this vicarious calibration can provide absolute radiometric gain calibration coefficients. Validation of satellite data products is also performed – comparing the higher level satellite products (e.g. sea surface temperature, ocean colour, vegetation indices, etc.) with in-situ measurements of these products.

Radiometric calibration and validation of basic satellite products (reflectance, radiance), and the harmonisation of satellite records to correct for biases between different satellites, are performed using one of the following methods:

- Simultaneous Nadir Overpass: This is where two satellites see the same scene within a few (~5) minutes of each other anywhere on the globe.
- Bright pseudo-invariant sites: These are bright sites that are considered extremely stable over time (they are usually inaccessible deserts, for example the centre of the Sahara). Satellite measurements are compared to one another for similar view and illumination geometries (e.g. solar zenith angle) and provide satellite-to-satellite comparison or long-term stability testing for a single satellite. In this case, care also has to be taken of potential errors due to changes in the atmospheric transmittance. This is dealt with either through averaging, or by applying corrections based on external measurements of key atmospheric parameters such as aerosols, ozone and humidity.
- Natural features – such as convective clouds, Rayleigh scattering over the ocean, sun-glint etc.– can provide a stable reference. It is not always predictable exactly where and when these features are suitable for observation, but they have well-defined optical properties and can be reliably used for satellite-to-satellite comparison, and, in the case of Rayleigh scattering, absolute radiometric gain and band-to-band normalisation.





- The Moon: like the pseudo-invariant sites, the moon is stable over time<sup>49</sup> and provides a bright uniform reference source for satellite-to-satellite comparison or long-term stability testing of a single satellite.
- Instrumented reference sites: There are a small number of uniform sites which are constantly monitored by instruments on the ground to check both the ground reflectance and the optical properties of the atmosphere (aerosol optical depth, for example). These sites can be used for satellite-to-satellite comparison, using the on-the-ground instruments as references to correct for changes of the site or the atmosphere, and also for satellite-to-ground comparison. In addition, these ground measurements can be made using equipment (e.g. a field radiometer) which relies on a SI-traceable calibration. Currently a working group of CEOS-WGCV-IVOS is developing a group of these sites into a global network known as RADCALNET. The RADCALNET project is in a two-year prototype stage.

Note that all of these methods are still under development and the subject of active research. The methods themselves are being developed and full uncertainty analysis is generally not yet performed!

## 8.4 Uncertainty analysis

It is important to understand that the process of in-orbit calibration and validation, whether performed by on-board references or through vicarious calibration, will always introduce additional uncertainty components of its own. Understanding those uncertainties is, conceptually, no different from performing uncertainty analysis on ground-based calibrations and the techniques introduced in Section 4 of this textbook apply equally well to post-launch calibration. The main difference is that launch itself introduces a fundamental break in the traceability chain and many more of the uncertainty components after launch will have to be estimated in-flight and/or simply guessed. This could mean that the calibration procedure does not rigorously follow a SI-traceable chain.

The objective should be to understand as many uncertainty contributors as possible, including the origin source for each, and to link these to the measurement model. These uncertainty contributions can then be individually assessed and any that are considered to be negligible can be removed, so simplifying the uncertainty analysis.

## 8.5 Example: PTFE diffuser

As an example, some preliminary (simplified and basic) uncertainty analysis is provided for a PTFE<sup>50</sup> solar diffuser used as a reference on board an earth observation imaging satellite. This example is simplified and is not meant to be an exhaustive study of the uncertainties for such a measurement, but an example of how to apply Section 4 to a post-launch situation.

---

<sup>49</sup> Providing due care is taken to correct for lunar phases, using for example the ROLO model

<sup>50</sup> Pressed PTFE powder is available under the tradename Spectralon™ from Labsphere and under the tradename OP.DI.MA from Gigahertz Optik

### 8.5.1 Use of the solar diffuser in orbit

The solar diffuser provides a reflectance reference in orbit and is used to check the radiance calibration of the instrument. The radiance of the solar diffuser, when illuminated by the sun, is given by

$$L_{SD}(p, b) = \frac{\rho(\theta_i, \phi_i, \theta_r, \phi_r) E_{\text{sun}}(b) \cos \theta_i}{\pi d_{\text{sun}}^2} \quad (8.1)$$

where,

- $L_{SD}(p, b)$  is the radiance of the diffuser as observed by a particular 'pixel',  $p$  in a spectral band  $b$ .
- $\rho(\theta_i, \phi_i, \theta_r, \phi_r)$  is the diffuser radiance factor<sup>51</sup> for illumination at the nominal solar angle  $(\theta_i, \phi_i)$  and viewing at the pixel angles  $(\theta_r, \phi_r)$ . It is a ratio of the diffuser BRDF (characterised on-ground) with respect to the ideal lambertian BRDF  $(1/\pi)$ .
- $E_{\text{sun}}(b)$  is the band-integrated solar irradiance at a standard distance (one astronomical unit).
- $\theta_i$  is the angle from the normal axis of the diffuser to the sun-to-diffuser axis
- $d_{\text{sun}}$  is the actual satellite-to-sun distance in astronomical units

Note that this is to some extent sensitive to the viewing 'pixel' (and/or detector, depending on the instrument type) as this will determine the viewing angles to select the appropriate reflectance factor. It is also dependent on the SRF of the pixel, the instrument 'band'. The band-integrated solar irradiance is given by<sup>52</sup>

$$E_{\text{sun}}(b) = \frac{\int E_0(\lambda) S(b; \lambda) d\lambda}{\int S(b; \lambda) d\lambda} \quad (8.2)$$

where,

- $E_0(\lambda)$  is the average exo-atmospheric solar spectral irradiance<sup>53</sup> [22] at wavelength  $\lambda$

---

<sup>51</sup> This term is sometimes incorrectly referred to as 'reflectance'. *Reflectance* is the ratio of the flux reflected in a given direction to the flux in the incident beam. *Reflectance factor* converts this to a ratio between the measured reflectance and that from a 'perfect diffuser'. *Radiance factor* is both relative to a perfect diffuser and for infinitesimal angles. *BRDF* is defined as the function describing the change with angle of irradiation and angle of view, of the quotient of the radiance of a surface element in the given direction of view, by the irradiance incident on the medium from the given direction of irradiation. For a Lambertian diffuser, *BRDF* is numerically equivalent to the angular distribution (with angle of illumination and view) of the radiance factor divided by  $\pi$ . (Note all these quantities also vary with wavelength.)

<sup>52</sup> Note that if the SRF is normalised to have unit area, the denominator is unity. Often this assumption is made and the equation is written with the nominator only.

<sup>53</sup> There is considerable debate in the relevant communities as to which solar spectrum is to be used and how this integration is performed. One common method is to use the solar spectrum given by the Thuillier model, ref in text.



- $S(b; \lambda)$  is the SRF of the instrument being calibrated for band  $b$ . More generally, it could be also depend on the detector and/or pixel.

Note that equations (8.1) and (8.2) are 'calculation equations' to use the terminology of Section 4 of this textbook. The full measurement equations would take into account all the sources of uncertainty.

The diffuser radiance is used to perform, an absolute radiance responsivity calibration (or an absolute reflectance responsivity calibration, see section 8.6.1) to obtain the calibration coefficients for radiance or reflectance measurements. The radiance responsivity gain coefficient<sup>54</sup> is given by

$$G(b, p) = \frac{Q_{b,p}}{L_{sd}(b, p)} \quad (8.3)$$

where

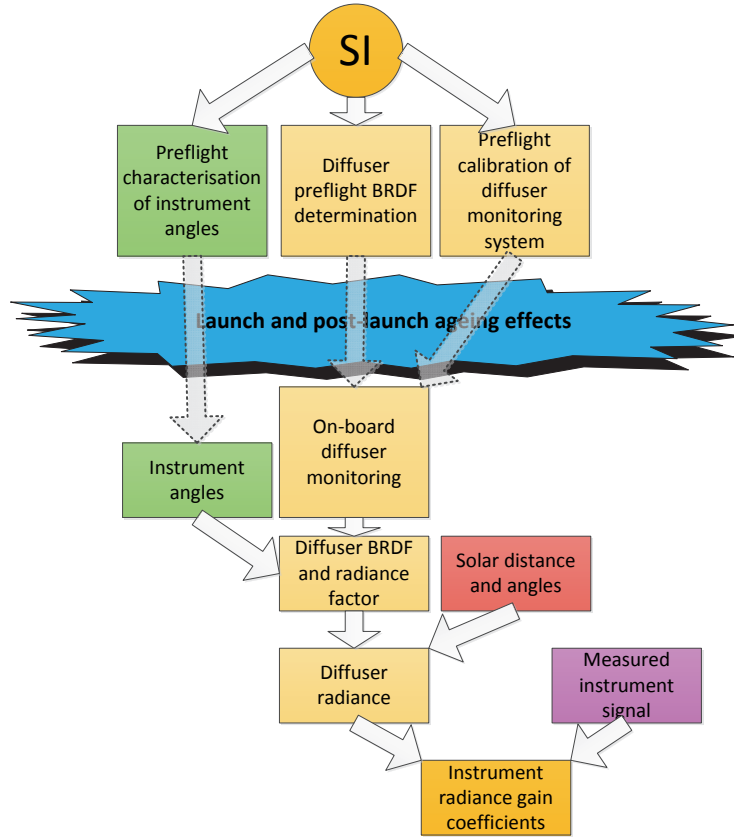
- $Q_{b,p}$  is the dark-corrected signal (digital count) measured by the instrument. Typically this signal will be the result of averaging several samples (e.g. 1000 samples) in order to reduce the instrument noise or other random effects.
- $L_{sd}(b, p)$  is the calculated solar diffuser radiance, as given in Equation (8.1).

### 8.5.2 Step 1: Describing the Traceability Chain

The traceability chain for the determination of the instrument gain (Equation (8.3)) is given in Figure 30. Note the break in the traceability chain due to launch and post-launch ageing effects. It is conceptually helpful to draw this into the traceability chain to ensure that the uncertainties associated with such changes are considered. In Figure 30, there is an assumption of some 'on-board diffuser monitoring'. This is intentionally vague to allow for several possibilities. However, for subsequent discussions in this section, we will assume that this is achieved the MERIS way, i.e. through the use of a reference diffuser that is only rarely illuminated by the sun. Thus the "preflight calibration of diffuser monitoring system" here means the pflight BRDF calibration of the reference diffuser.

---

<sup>54</sup> In practice this may be averaged across pixels within a band, for example.



**Figure 30 Traceability chain for the instrument radiance gain coefficient calibration using the inflight diffuser. Note that launch and post-launch ageing effects break the full traceability to SI.**

### 8.5.3 Step 2: Writing down the calculation equations

The calculation equations have been given above as Equations (8.3), (8.1) and (8.2).

In addition it is necessary to calculate the diffuser radiance factor from the BRDF. The BRDF is essentially the angular distribution function for the radiance/reflectance factor divided by  $\pi$ , so this is nominally a requirement to 'read off' the correct BRDF or radiance factor value for a given instrument viewing angle and solar illumination angle. In practice it may be necessary to interpolate a value from discrete measured BRDF values or fit those to a curve. Another option is to use polynomial and higher order models which are "tuned" using experimental data like hemireflectance and scattering properties of the PTFE material.

It is also necessary to have a model to account for the diffuser degradation in orbit. We assume here that the degradation in orbit is determined from a second diffuser that is used only rarely. This is the MERIS method and results obtained give the correction curve of Figure 24a. We assume that a correction is applied, such that the in-orbit radiance/reflectance factor is calculated from the pre-flight radiance factor using

$$\rho_{\text{in-orbit}}(\theta_i, \phi_i, \theta_r, \phi_r) = \rho(\theta_i, \phi_i, \theta_r, \phi_r)(1 - At) \quad (8.4)$$

where

- $-A$  is the slope of the linear function fitted to the correction curve of Figure 29a
- $t$  is the time since launch in the units of Figure 29a

#### 8.5.4 Step 3: Considering the sources of uncertainty

The aim here is to consider what all the sources of uncertainty are. This list is not intended to be exhaustive or definitive for this particular example, but is indicative of the types of effect considered.

**Preflight:**  $\rho(\theta_i, \phi_i, \theta_r, \phi_r)$

The main diffuser will have been calibrated for BRDF during the pre-flight calibration campaign as well as the reference diffuser and the instrument angles. There will be several uncertainty contributors associated with this process:

- **Uncertainty associated with the absolute BRDF characterisation:** The diffuser radiance/reflectance factor is calculated from the modelled diffuser BRDF and the known viewing angles for the relevant pixel and the solar illumination angle. The absolute diffuser BRDF characterisation will be performed by the calibration laboratory during the preflight calibration process. This laboratory will either be a national metrology institute, or the measurements will be traceable to a national metrology institute.

For example, NPL's BRDF measurements are performed on the National Reference Reflectometer (NRR) and a full uncertainty budget has been published [23]. The NRR performs radiance factor measurements using an input beam whose geometric extent is accurately defined with an aperture of known area, and measurements made (in both polarisations) of the direct beam and the light reflected from the sample. These measurements will typically be interpolated or fit into a model providing a pre-flight uncertainty of the BRDF model.

The reference diffuser (the one that is only occasionally used in orbit) will be calibrated pre-flight in a similar manner, with similar uncertainty components.

- **Uncertainty associated to the uniformity of the PTFE surface:** This is due to the surface non-uniformity of the PTFE material that can produce small variations in the BRDF depending on the position. At instrument level, each pixel will be viewing different positions of the diffuser surface across and along the FoV.

In this case we will also report the value provided by the MERIS diffuser characterisation. In that case, it was found a variability  $<0.5\%$  across the diffuser surface [24].

- **Uncertainty associated with the pixel viewing and incident angles.** The instrument and viewing angles for each pixel will be defined by characterising (or setting) the geometric arrangement preflight. There will be an uncertainty associated

with these angles due to the system itself or the influence of other systems (e.g. vibration due to an internal gyroscope system or the diffuser flatness using tomographic images). This will translate to an uncertainty associated with the radiance/reflectance factor that will depend on how rapidly the BRDF changes with angle. For a good quality diffuser, the radiance/reflectance factor is likely to change slowly with angle and therefore small uncertainties associated with the angle are likely to create negligible uncertainties associated with the BRDF.

For example, as mentioned before the micro-vibrations can be assessed pre-flight providing a Gaussian angle error distribution of e.g.  $0.1^\circ$  ( $1\sigma$ ). The uncertainty in the angle can be simply propagated by using a Monte Carlo evaluation. Simplified for this example, we assume that the radiance factor changes by  $\sim 0.3\%$  per degree and, proportionally a change of  $\sim 0.1\%$  in angle would result in an uncertainty  $0.03\%$  ( $1\sigma$ ).

- **Uncertainty associated with the suitability of the BRDF model.** As commented in step 2 (section 8.5.3), the diffuser BRDF is fitted to an analytical model (or any other technique) that is used to estimate the radiance factor at the actual angles of illumination and view. The main uncertainty component will be that due to the suitability of the model. The MERIS calibration report [4] describes the root mean square residual<sup>55</sup> (difference from model and measured values) as  $0.3\%$ . In addition to this uncertainty it has been noticed that the calculated gain (Equation(8.3)) has some sensitivity to sun azimuth angle. There is no physical process that would create such dependence, and it can be assumed that this dependence is therefore an artefact of uncorrected errors in the BRDF model. This provides a further uncertainty component at  $\sim 0.5\%$ .
- **Uncertainty associated with the spectral interpolation of the BRDF.** The BRDF values are measured at certain wavelengths only. At other wavelengths there is a need to interpolate. As discussed in Section 6.6, interpolation can have the unexpected effect of reducing the uncertainties as intermediate points are some form of average of the input data. However, that assumes that the interpolation is appropriate, a more useful measure of the uncertainty associated with interpolation is to take some additional measurements (or perform the interpolation with some measured values missing) and calculate the residual from the interpolation for the extra measured values.

---

<sup>55</sup> The peak difference is  $1\%$ . This uncertainty could be considered a Gaussian distribution with a standard deviation of  $0.3\%$  or a rectangular distribution with a half width of  $1\%$ . For a rectangular distribution to obtain the equivalent standard uncertainty it is necessary to divide by  $\sqrt{3}$  (see Section 0). This would give  $\sim 0.6\%$ . Whether  $0.3\%$  or  $0.6\%$  is used in the uncertainty budget depends on whether the residual distribution is rectangular or Gaussian.

**Postlaunch:**  $\rho_{\text{in-orbit}}(\theta_i, \phi_i, \theta_r, \phi_r)$

The post-launch diffuser BRDF is calculated from the preflight BRDF calibration and the known degradation in orbit, using Equation (8.4). There are the following sources of uncertainty for the degradation correction:

- **Reference diffuser stability.** The assumption that the reference diffuser is stable in orbit. In actual fact, it too will degrade and probably at a similar rate. The MERIS instrument calibration report has calculated that for the time scale of Figure 29a, the reference diffuser had 37 minutes of exposure (compared to 370 minutes for the main diffuser). Applying the same correction over the shorter time period, suggests a change of ~0.2 %. This can be considered the uncertainty associated with the diffuser correction due to reference diffuser stability.
- **Correction model.** The correction assumes that a linear trend reliably describes the diffuser ageing since launch based on pre-launch assessments [4]. The measured data values do not perfectly fit a linear trend, and it has been noted in the MERIS calibration report that the minor fluctuations from the linear trend are correlated with solar azimuth variation and this suggests that it is due to residual errors of the BRDF models (see below). As an initial estimate, the root-mean-square residual (deviation from the linear fit for the measured values) can be used as the uncertainty associated with the linear trend assumption (this is also ~0.2 %). It may be that for the shortest time periods an additional uncertainty component is required.

**Radiance of the diffuser**  $L_{\text{SD}}(p, b)$

The radiance of the diffuser is calculated using Equation (8.1). Just above we have described the main uncertainty sources related to the reflectance factor  $\rho(\theta_i, \phi_i, \theta_r, \phi_r)$  and its impact on the estimation of the radiance of the diffuser. The rest of components in this equation will also introduce additional uncertainties:

- $d_{\text{sun}}$  ; **The sun-satellite distance** can be considered deterministic (it is a function of the date of acquisition and the earth's orbit). The uncertainty associated with this can be considered negligible.
- $E_0(\lambda)$ ; **The sun irradiance model** for MERIS is based on SOLSPEC measurements and the paper by Thullier et al provides an uncertainty estimate with a clear uncertainty budget. There is, however, considerable debate as to which solar model to use. For TOA reflectance measurements, this will cancel out (see section 8.6.1), but for radiance it is significant. The uncertainty introduced here should include both the uncertainty associated with the model (as in the paper) and the suitability of the model choice (by comparing the outputs of different models to identify biases between different models)

For this example, we will simply provide the relative standard uncertainty reported by the Thuillier sun irradiance model of ~2.5 %

- $S(b; \lambda)$ ; **The SRF of the band** is characterised pre-launch. As discussed in Section 7 this can change in orbit and while there are some possible in-orbit methods to recharacterise it, they have their own associated uncertainties. There may also be some effects (i.e. spectral noise) from any assumption that the SRF is the same for all spatial 'pixels'. Note that the SRF is used within an integral in Equation (8.2) and therefore the analysis described in Section 7 applies. The multiplying function is the solar spectral irradiance. This means that the analysis will be most sensitive to changes in the SRF where the SRF is close to one of the solar Fraunhofer lines, and therefore the solar spectral irradiance is rapidly changing with wavelength.

To make it very simplistic just assume that the spectral noise across the pixels in a specific detector and band is  $\pm 0.5$  nm shift of central wavelength (note that a proper assessment would involve the integral over the whole band range in Equation (8.2)). The spectral response is also corrected in-flight and the stability correction residual is  $\pm 0.2$  nm ( $1\sigma$ ) for the central pixel. This leads to an associated standard uncertainty of  $\sim 0.54$  nm. Finally, the evaluation of this uncertainty in Equation (8.2) would propagate to an uncertainty on the band-integrated solar irradiance of 0.2 % (the numbers are only illustrative)

- $\cos \theta_i$ ; **lambertian term**. The same uncertainty on the incident angle knowledge that had an impact on the radiance factor will have a direct impact on the estimated diffuser radiance. Here the same angle uncertainty applies as in the previous uncertainty contributor (standard deviation of  $0.1^\circ$ ). For a diffuser measuring at a nominal incidence angle  $\theta_i = 60^\circ$ , the impact on the cosine term would be:  
 $u(\cos \theta_i) = 0.3\%$

#### Measured instrument signal: $Q_{b,p}$

The measured signal will be a light count minus a dark count. There will be uncertainties associated with:

- **Instrument noise** – this can be estimated from the standard deviation of the light measurements. Typically the measured signal will be an average of several samples that will reduce the random noise to a negligible level. For example if the instrument noise at the radiance calibration  $L_{\text{cal}}$  is 1 % for an specific band and the measured signal at calibration is averaged over 1000 samples, the noise introduced will be  $\sim 0.03$  %, which is negligible.
- **Quantisation noise** – the effect of truncating the signal for the bandwidth of the instrument communication. The radiance measured at calibration  $L_{\text{cal}}$  will be typically much higher than the minimum radiance  $L_{\text{min}}$  measured by the instrument (e.g. TOA

ocean radiance levels). With a truncation of e.g. 12 bits, the quantisation noise will be negligible at the measured signal.

- **Internal spatial stray light** – this effect will be accounted as negligible here since the input signal is largely uniform. The spread, back-reflection or any similar mechanism will “contaminate” the pixel neighbourhood. However, the same mechanism will apply for all the pixels in the FoV. Therefore, for each pixel the signal losses will be largely compensated by the signal increases from the rest of pixels meaning that the impact on the uncertainty is negligible.
- **Internal spectral stray light** – the light from other spectral wavelengths may have some effect on the signal with a spectral imager system.
- **External straylight** – during the calibration it may originate from reflections in the Sun diffuser assembly or other instrument parts. It is possible that the Earth light or other parts of the system contaminate the measurement. This light will be translated into a bias where the first component is proportional to  $L_{\text{cal}}$  and the second one is an offset of  $Q_{b,p}$ . Let's assume that this have been well characterised (e.g. using a ray tracing model) and the total bias at  $L_{\text{cal}}$  is in the order of 1 %. A correction term should be added in Equation (8.1),  $K_{\text{ext-stray}}$  that accounts for this bias. Nonetheless, the correction itself introduces a residual uncertainty of ~0.2 %. Note that the residual uncertainty will be linked not to the measured signal but to the estimated diffuser radiance (see Table 10).
- **Polarisation error** – PTFE material is not a perfect scrambler. Thus, the sun-reflected signal in the instrument could have a degree of linear polarisation e.g. of 3 %. If the instrument does not include a depolarisation stage, a typical value of polarisation sensitivity could be 5 %. Thus, the error produced by this contributor will be of 0.15 %.
- **Dark signal accuracy and stability** – the first one accounts for the reliability of the dark signal measurement (e.g. using a shutter) whereas the second one accounts for the standard deviation of the dark signal and also for the fluctuations of dark signals taken at different times. There may be some sensitivity to the instrument temperature. The radiance at calibration  $L_{\text{cal}}$  will be much higher compared to the digital signal levels and any relative uncertainty (introduced as a %) can be neglected. In addition, it is expectable that the same calibration procedure includes e.g. a dark signal measurement before and after the measurement that minimises the stability error.



## Uncertainty budget

From the discussion in this section it is possible to write the rows of the uncertainty budget and assign uncertainties to each contributor. These are given in Table 10. Note that the symbols given are for the uncertainties associated with the effect.

**Table 10 Uncertainty budget outline for post-launch radiance calibration using a diffuser**

Symbol	Effect	Uncertainty associated with this effect
<b>Diffuser Radiance Factor</b> $\rho(\theta_i, \phi_i, \theta_r, \phi_r)$		
$u_{\text{pf-BRDF}}$	Pre-flight diffuser absolute BRDF calibration	0.10 %
$u_{\text{BRDF-unif}}$	Uniformity of the PTFE surface	0.50 %
$u_{\text{BRDF-ang}}$	Angle errors and BRDF angular sensitivity	0.05 %
$u_{\text{BRDF-model}}$	BRDF model suitability	0.50 %
$u_{\text{BRDF-}\lambda\text{-int}}$	Spectral interpolation of BRDF	0.20 %
<b>Diffuser Reflectance Factor correction</b> $\rho_{\text{in-orbit}}(\theta_i, \phi_i, \theta_r, \phi_r)$		
$u_{\text{ref-stab}}$	Post-launch stability of reference diffuser	0.20 %
$u_{\text{lin-corr}}$	Suitability of post-launch correction (linear trend assumption)	0.20 %
<b>Diffuser radiance calculation</b> $L_{\text{SD}}(p, b)$		
$u_d$	Sun-satellite distance	0.00 %
$u_{E_{\text{sun}}}$	Solar irradiance model	2.50 %
$u_{\text{SRF}}$	SRF of instrument – spectral noise	0.20 %
$u(\cos \theta_i)$	Incident angle knowledge impact	0.30 %
$u(K_{\text{ext-stray}})$	External stray light correction residual	0.20 %
<b>Measured instrument signal</b> $Q_{b,p}$		
$u_{\text{noise}}$	Instrument noise	0.03 %
$u_{\text{trunc}}$	Quantisation noise	0.00 %
$u_{\text{int-stray}}$	Internal stray light	0.00 %
$u_{\text{pol}}$	Polarisation error	0.15 %
$u_{\text{dark}}$	Dark signal accuracy and stability	0.00 %

### 8.5.5 Step 4: Creating the measurement equation

The simplest way to approach the measurement equation, a method similar to that for Equation (4.3) in Section 4.6, is to use an additional symbol for each of these uncertainty



components that represent the error in the gain due to this effect. In this case, the measured signal is best treated as having additive errors, thus<sup>56</sup>

$$Q_{b,p} = Q_{b,p,\text{light}} - Q_{b,p,\text{dark}} + \delta_{\text{noise}} + \delta_{\text{trunc}} + \delta_{\text{crosstalk}} + \delta_{\text{dark}} + \delta_{\text{dark-stab}} \quad (8.5)$$

Here all the  $\delta$  terms have a nominal (expected) value of zero, but an uncertainty associated with that value that is given by the equivalent  $u$  term. Each of these parameters will have units of [digital numbers] and the associated uncertainties should also have units of [digital numbers].

For all other parameters, a similar approach is taken with a multiplicative effect, hence, for example

$$L_{\text{SD}}(p, b) = \frac{\rho(\theta_p, \phi_p) E_{\text{sun}}(b) \cos \theta_{\text{sun}}}{\pi d_{\text{sun}}^2} K_{\text{SRF}} K_{\text{SRF-Change}} K_{\text{ext-stray}} \quad (8.6)$$

Here the  $K$  terms have a nominal value of unity, and an uncertainty associated with that given by the equivalent  $u$  terms. Note that  $u_{\text{SRF}}$  is the uncertainty associated with the solar diffuser radiance due to the SRF. It is not the uncertainty associated with the SRF. Here all the terms are multiplicative and the uncertainty is a relative uncertainty, usually expressed in %.

The parameters in the calculation equation also have associated uncertainties. In the case of the diffuser radiance factor,  $\rho(\theta_p, \phi_p)$ , the uncertainty associated with this comes from all the effects listed in Table 10.

The final equation is also multiplicative,

$$G(b, p) = \frac{Q_{b,p}}{L_{\text{sd}}(b, p)} K_{\text{ext-stray}} K_{\text{int-stray}} \quad (8.7)$$

### 8.5.6 Step 5: Determining the sensitivity coefficients

With the approach described in the previous section, the equations given are simple additive (Equation (8.5)) or multiplicative (Equation (8.6), (8.7)) expressions. The sensitivity coefficients for these equations are therefore simple.

For Equation (8.5), the sensitivity coefficients for all terms are unity for uncertainties expressed in digital numbers. To use this uncertainty in Equation (8.7), the resultant absolute

---

<sup>56</sup> Here, the first two terms,  $Q_{b,p,\text{light}} - Q_{b,p,\text{dark}}$  are the nominal true value and the remaining terms are those with associated uncertainties. It is a matter of personal preference whether to treat these as uncertainties associated with the light and dark signals, respectively, and therefore write  $Q_{b,p} = Q_{b,p,\text{light}} - Q_{b,p,\text{dark}}$  as the measurement equation, with the other uncertainties inbuilt into  $u(Q_{b,p,\text{light}})$  and  $u(Q_{b,p,\text{dark}})$ , or to treat these as terms with no associated uncertainty and put all uncertainties into  $\delta$  terms.

uncertainty (in digital numbers) must be converted to a relative uncertainty (in per cent). This is achieved by dividing the resultant uncertainty by the measured signal.

Thus, from (8.5)

$$u^2(Q_{b,p}) = u_{\text{noise}}^2 + u_{\text{trunc}}^2 + u_{\text{crosstalk}}^2 + u_{\text{dark}}^2 + u_{\text{dark-stab}}^2 \quad (8.8)$$

And

$$u_{\text{rel}}^2(Q_{b,p}) = u^2(Q_{b,p}) / Q_{b,p} \quad (8.9)$$

In equation (8.6) all the terms are multiplicative. In this case it is the relative uncertainties that are added in quadrature. Treating relative uncertainties, most sensitivity coefficients are unity. The exceptions are for the  $\cos \theta_{\text{sun}}$  and  $d_{\text{sun}}^2$  terms. For the distance term the sensitivity coefficient is straightforward to derive, as in Section 3.1, equation (3.11)

$$\frac{\partial L_{\text{SD}}}{\partial d_{\text{sun}}} = -2 \frac{L_{\text{SD}}}{d_{\text{sun}}} \quad (8.10)$$

Therefore the relative sensitivity coefficient is -2. For the cosine term it is slightly more complex, as the differential of  $\cos \theta$  is  $-\sin \theta$  and for angles around  $\theta = 0$ , this suggests the sensitivity coefficient (and therefore the uncertainty contribution) is negligible. This is an example of where the in-built assumption in the Law of Propagation of Uncertainties, of linear functions, breaks down. The GUM does include advice on how to deal with this<sup>57</sup> – but often this is an example of where a sensitivity coefficient calculated numerically is the simplest and most satisfactory method for estimating it. This can be done by calculating the result of Equation (8.6) first with the nominal solar angle and then again with a slightly modified version. The propagation of a Gaussian distribution through  $\cos \theta$  would derive in a typical u-shaped function if the limits of the distribution account for a big or the whole part of the cosine period. However, in this case, the changes are so small (0.1 degrees) that its propagation is linear producing a Gaussian output.

Note that although the relative sensitivity coefficients for the  $K$  values are straightforward and nominally unity, this is because the  $K$  values have been defined as the error in the radiance (Equation (8.6)) or gain (Equation (8.7)) due to this effect. The sensitivity of radiance to the uncertainty associated with the SRF, or the sensitivity of gain to internal stray light is much more easily modelled than written in an analytical equation, and therefore the hard work of working out the sensitivity coefficients is already done in the modelling.

### 8.5.7 Step 6: Assigning uncertainties

This step involves providing the numbers in Table 10 of Section 8.5.4. Many of the methods for estimating these associated uncertainties are given in that section. It is important to be

---

<sup>57</sup> See, e.g. GUM section F.2.4.4, and the note to 5.1.2

clear in the table whether the uncertainties provided are relative (in percent) or absolute (with the same units as the quantity considered). It is also important, particularly with relative uncertainties to be clear whether this is the relative uncertainty associated with the quantity itself (e.g. distance) or the uncertainty associated with the measurand due to this effect (i.e. whether the sensitivity coefficient has already been taken into account).

Strictly, the table 'should' provide only relative or only absolute uncertainties. And strictly, the table 'should' provide the uncertainty associated with each parameter in turn, with a separate column providing the sensitivity coefficient and a final column multiplying the two together. However, as with the  $K$  factors described above, often this is not a clear cut decision and in real, experimental, uncertainty budgets this often gets messy! The most important thing is to make it clear in the wording. There should definitely be a column described as 'the uncertainty associated with the measurand due to this effect'. In some cases, this alone is filled in. In other cases this will be a product of a sensitivity coefficient and the uncertainty associated with the quantity itself.

#### 8.5.8 Step 7: Combining and propagating uncertainties

If the table provides a column labelled 'the uncertainty associated with the measurand [here gain] due to this effect' then the combination of uncertainty is simple – it requires that that column be added in quadrature.

This does make the assumption that there is no associated correlation and the second half of the law of propagation of uncertainties is not required. In this example, this can be considered the case. If, however, the reflectance (Equation (8.11), section 8.6.1) is the product of interest, then note that the scene radiance  $L$  is correlated with the band-integrated solar irradiance  $E_{\text{sun}}(b)$ . This is because it is calculated from the measured signal and gain, which is itself calculated from the solar diffuser measurement (Equation (8.3)), which depends on the band-integrated solar irradiance. It is possible to calculate the correlation coefficient between these two parameters. Alternatively, and more straightforwardly, the equations should be written out in full and the term cancelled before uncertainty analysis is carried out.

#### 8.5.9 Step 8: Expanding uncertainties

Here the uncertainty analysis has been straightforward and it is very likely that an expansion to a 95 % confidence interval is achieved by multiplying by  $k = 2$ . If any uncertainty components are determined on a Type A evaluation based on a small number of readings, then either the uncertainty should be increased using Equation (3.26) in Section 3.3.2, or the Welch-Satterthwaite Equation, Equation (4.4) should be used. For Type B evaluations it is also possible to consider, in the Welch-Satterthwaite equation the 'uncertainty in the uncertainty', as in Section 4.10.

### 8.6 Level-1 EO products radiometric uncertainty

Although the main objective of the course is about understanding the main uncertainties associated to EO instrument and its calibration (both pre-flight and post-launch), it is worth

to mention how these instrument measurements are typically disseminated to the EO users and the consequences in terms of uncertainty assessment and propagation.

The idea behind this section is not provide a detailed assessment of the specific L1 ground processing related to EO products. That would not only imply the understanding of the absolute calibration uncertainty as in section 8.5 but other ground processing corrections: DS, linearity, uniformity, temperature, etc... The objective here is to help both instrument designers and EO end-users understand of the implications that the product format has in the EO uncertainty analysis and its propagation to higher level products.

We will describe here two typical examples in EO product format that demonstrate the specific constraints associated with the uncertainty assessment and propagation: the radiance-to-reflectance conversion and the image orthorectification.

### 8.6.1 Radiance-to-reflectance conversion

The PTFE diffuser example in section 8.5 has been used to exemplify a typical post-launch calibration of EO systems. This type of calibration provides a radiance gain as indicated in Equation (8.3)

One of the main satellite level-1 products is the top-of-atmosphere reflectance. Top-of-atmosphere reflectance is calculated from the measured radiance using:

$$\rho_{\text{TOA}} = \frac{\pi d_{\text{sun\_to\_surface}}^2}{\cos \theta_{\text{sun}}(p)} \frac{L_{\text{TOA}}(b, p)}{E_{\text{sun}}(b)}. \quad (8.11)$$

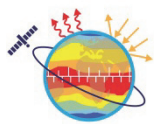
where

- $d_{\text{sun\_to\_surface}}$  represents the Sun-to-Earth surface distance in astronomical units and
- $\theta_{\text{sun}}(p)$  is the sun zenith angle that corresponds to the Earth surface projection for each pixel.

Note the similarity between Equation (8.11) and Equation (8.1). This means that for the reflectance product, in effect the reflectance of the Earth is compared to the reflectance of the panel directly and some of the uncertainty components described above are cancelled. In particular, the same integral of the product of the SRF of the instrument and the defined **solar irradiance** is used. This means that reflectance products are insensitive to errors in these parameters<sup>58</sup>. Furthermore it removes the cosine effect of different **solar zenith angles** due to time acquisition differences between images as well as the **Earth-to-sun distance** changes through the year (translation) and day (rotation).

---

<sup>58</sup> If the SRF has changed or is not the one assumed, then it will cancel out in the solar convolution integrals in Equations (8.11) and (8.1). That does not mean that changes in the SRF are insignificant, however! The calculated value  $\rho$  in equation (8.11) is the band-integrated earth top-of-atmosphere reflectance. If that product is subsequently used or compared with other products, then the band-integration must be taken into account. Because that will depend on the true SRF and the true ground spectral reflectance, then errors in the assumed SRF will have a significant effect.



Although the reflectance approach corrects for the several issues as described before, the radiance factor  $\rho_{\text{TOA}}(p, l, b)$  only accounts for a specific incident and reflected geometry. The exact viewing geometry will be different from pixel-to-pixel across the field of view, which may be different for different orbits. Some of the Earth's surfaces (e.g. desert areas) can provide BRDF characteristics "relatively" close to the ideal Lambertian one which minimises the directional dependency of the measured radiance. However, this is not the situation in most cases where the surface BRDF presents a specular component and an important anisotropy.

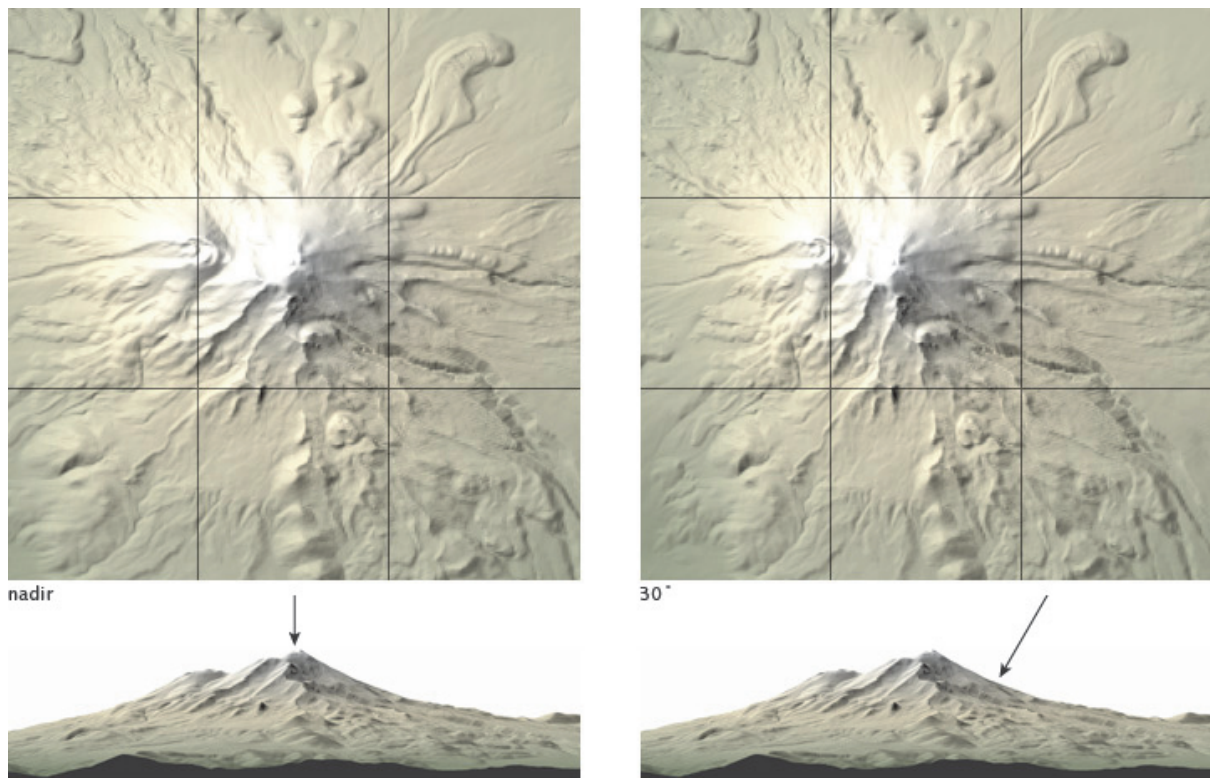
Several models have been (and are) used to characterise the different Earth surface BRDF. Recent missions, e.g. POLDER/PARASOL, estimate the directionality of the Earth surface-atmosphere (TOA) reflectance. The BRDF correction is a source of uncertainty which is accounted for in higher level products (e.g. surface reflectance products). E.g. [25]

To sum up, it is important to be clear about the quantity of measurement and assumptions in the EO products, since, for example, the uncertainties associated with reflectance are different from those associated with radiance. These differences will propagate through to higher-level products.

### 8.6.2 Image orthorectification

Providing an image at the instrument pixel viewing imposes several difficulties to the EO end users. For example, at the sensor geometry each individual pixel does not account for the surface elevation variations or the difference in projection over the surface from pixels at nadir and off-nadir as shown in Figure 31.

Typically, EO products are resampled during the ground processing (or using an external tool) that provides pixel images orthorectified to world geodetic models (e.g. WGS 84). This provides a uniform size of the pixel (pixel size provided in meters rather than mrad) while a digital elevation model (DEM) corrects (up to a certain extent) for the surface elevation variations. Although this is a benefit to EO end users it also imposes a difficulty when propagating the uncertainties from the instrument output geometry to the Earth geometry.



**Figure 31** Compared to a nadir view (looking straight down), a view from the side (30°) subtly distorts the terrain. Satellite data must be carefully corrected for even small distortions due to terrain or viewing angle if scientists want to detect changes in the landscape over time. (NASA Earth Observatory images by Robert Simmon, based on the USGS [National Elevation Dataset](http://national.elevation.dataset.gov).)  
<http://earthobservatory.nasa.gov/Features/GlobalLandSurvey/page3.php>

The difficulty arises from the fact that the pixel image does not represent anymore the focal plane geometry. That means, that differences between several detectors, filters etc. are not (a priori) traceable anymore. Nonetheless, the capacity to propagate the instrument uncertainty to the resampled products relies in a great extent to the radiometric resampling algorithm. Here we make a distinction between two typical radiometric resampling algorithms and its different impact in terms of uncertainty:

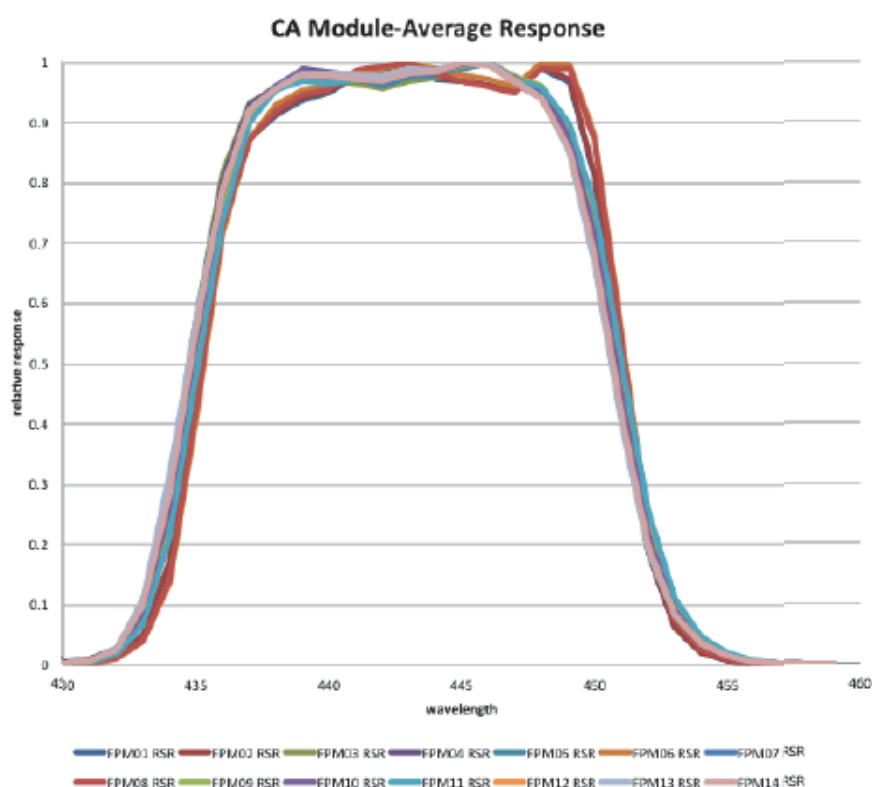
- **Cubic convolution** resampling algorithm uses a weighted average of the 16 pixels nearest to the focal cell and produces the smoothest (or most continuous) image compared to nearest neighbour resampling. It provides a great performance in terms of pixel interpolation and smoothness but limits the propagation of the instrument uncertainties to the product (no correspondence between pixels). It is also one of the most computationally intensive algorithms.

The propagation of the radiometric uncertainty from the instrument image to the resampled image becomes especially complicated for systems which include several detectors in the focal plane. In that case, it complicates the determination of the detector noise or the spectral response not only at pixel level but also at detector level.



For example, this is the case of Landsat-8 Level 1 T- Terrain Corrected products which have been resampled using a cubic interpolation algorithm. The resampled pixels have no direct equivalent pixel in the instrument image. In addition, each pixel cannot be, a priori<sup>59</sup>, directly associated to any of the 14 detector modules that the OLI instrument includes.

The Figure 32 corresponds to the per-module average spectral response of the CA band Landsat 8 OLI instrument. From the figure, it is possible to appreciate the difference in average spectral response between different modules due to different filter wafers.



**Figure 32 Landsat-8 OLI Spectral Response normalised and averaged for each one of the 14 modules of the Coastal Aerosol band [26].**

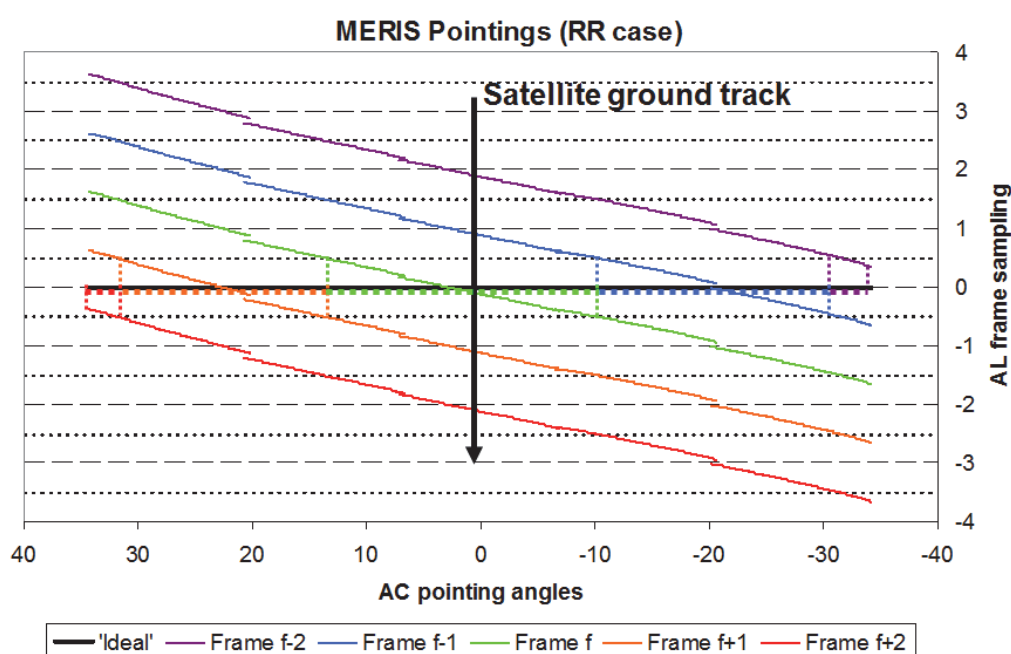
For the resampled image Level 1 T product, the per-module spectral response will not be possible to associate (a priori) to the image but the average spectral response for the whole FoV (i.e. the 14 detectors). Thus, an error would be introduced compared to the non-resampled image which would account for the difference in the convolution

<sup>59</sup> It is possible to implement mechanisms that define the footprint of the detectors in the resampled image with some limitations.



from one module to another (similarly to Equation (8.2)) and will depend on the type of scene measured<sup>60</sup>.

- **Nearest Neighbour** resampling algorithm works by matching a pixel from the original image to its corresponding position in the resized image. If no corresponding pixel is available, the pixel nearest is used instead. This type of performance has been used in several missions like ENVISAT/MERIS. It does not produce the smoothest image but reduces the processing [27] by more than 10 times compared to the cubic convolution interpolation and it is a reversible process. It means that each resampled pixel corresponds to a specific pixel of the instrument focal plane. The Figure 33 shows an example of that pixel correspondence for the ENVISAT/MERIS mission.



**Figure 33 MERIS correspondency between the resampled pixels and the focal plane pixels. Re-builds ideal swath from actual MERIS FOV: slightly misaligned plane + inter-camera dispersion [Ludovic Bourg, MERIS Level 1b processing, MERIS US Workshop, 14 July 2008]**

<sup>60</sup> Note that the level of error for this case has not been calculated. The bias introduced could be partly compensated or could have limited impact on the radiometric uncertainty budget. However, this distinction at TOA level would be of extremely benefit for the uncertainty propagation to higher level terms

## 9 Conclusions

This course textbook has provided an introduction to uncertainty analysis for earth observation instrument calibration. It has provided a suggested step-by-step methodical approach to uncertainty analysis and given some examples of how to apply this in real earth observation calibrations.

One of the first barriers many scientists and engineers have to uncertainty analysis is that it seems theoretical and complicated. Our aim in this course has been to break down that barrier and present a practical and pragmatic approach to uncertainty analysis.

Once scientists and engineers begin to apply these tools, they often reach a second barrier – the realisation of just how complex the experimental systems are and just how many sources of uncertainty they are<sup>61</sup>. Once they start to produce uncertainty budgets, they also start to see the subtle correlations – for example that if both the detector responsivity and the filter transmittance are temperature sensitive, then there is a correlation between detector responsivity and filter transmittance. This increased understanding can lead scientists and engineers to become overwhelmed with the problem and therefore put uncertainty analysis off to be tackled another day.

Our recommendation is that you consider your first uncertainty budget a simplistic one. Start with the concepts that are most critical. It is likely that you have an intuitive understanding of where your dominant uncertainty components will come from. Later this initial budget can be refined and augmented. Later it will be possible to replace what are ‘guesses’ now with formal uncertainty estimates based on experimental testing or modelling. It is Emma’s experience that initial uncertainty budgets based on ‘guesswork’ are pretty close to the formal uncertainty budget that comes after a year or more of rigorous analysis. It is also her experience that uncertainty budgets tend to increase over time, as more ‘unknown unknowns’ are recognised and included. Don’t let this put you off!

---

<sup>61</sup> This realisation is, of course, also behind the first barrier, although at that stage it is harder to be explicit about what the concern is. Once the first three steps suggested in Section 4 are completed, though, it can be quite daunting to think about all the effects that are now listed.

## Appendix A Making and using covariance matrices

When it is possible to split the uncertainty budget into systematic effects and random effects, it is not necessary to use the full form of the law of propagation of uncertainty (Equation (2.1)). In general, as discussed in Section 3.5, the systematic and random effects can separately be explicitly described in an error-model form of the measurement equation. In other words as the correlation is built into the error model, the terms within that error model are entirely uncorrelated with each other.

This approach is often the simplest, particularly when data analysis is carried out using a spreadsheet program. There are, however, occasions when such an analysis is insufficient. Sometimes correlation cannot be described by simply separating random and systematic components. Such separation is inapplicable whenever mathematical correlation has been introduced, for example through interpolation, averaging, smoothing, bandwidth correction and similar processes that combine different data points (e.g. from different wavelengths, or from different radiance levels). In these situations a covariance matrix provides the straightforward means to apply the full form of the law of propagation of uncertainty.

The law of propagation of uncertainty (equation (2.1)) is written in matrix form as

$$u^2(y) = C_y U_x C_y^T \quad (11.1)$$

where

$$C_y = \left( \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_n} \right)^T \quad (11.2)$$

is a column vector of sensitivity coefficients. (T represents transpose, i.e. that vector is written as a column, rather than as written here, for space reasons, as a row) and  $U_x$  is the covariance matrix. This form is given and used in the Supplement 2 to the GUM<sup>62</sup> as equation (3) on page 15.

Once a covariance matrix has been formed, Expression (11.1) is simple to implement in a programming language, particularly a matrix-based language such as MATLAB. This formulation can handle complicated analysis problems and partially correlated data.

### A.1 How to create a covariance matrix

A covariance matrix is a square matrix that describes the covariance of the measured values, the  $x_i$ . Each row and each column represents a different  $x_i$ . The diagonal elements are the variance,  $u^2(x_i)$ , the non-diagonal elements represent the covariance,  $u(x_i, x_j)$ . Note that

---

<sup>62</sup> Freely downloadable at <http://www.bipm.org/en/publications/guides/gum.html>

the covariance matrix is generally symmetrical as  $u(x_i, x_j) = u(x_j, x_i)$ . If two quantities have no associated covariance, (i.e. they are not correlated), then the covariance term is 0.

As an example, consider the case of data that will subsequently be integrated spectrally, the covariance matrix we need to develop is that showing the covariance associated with (e.g.) spectral irradiance values determined at different wavelengths. This example is used because it's often in spectral problems that covariance matrices are needed.

Each row (or column) of the covariance matrix will thus represent a different measurement wavelength. The diagonal terms give the variance  $u^2(E_i)$  and the off-diagonals the covariance  $u(E_i, E_j)$  between the irradiance at the row wavelength and the irradiance at the column wavelength. The covariance matrix takes the form

$$U_E = \begin{matrix} & \begin{matrix} 1 & 2 & \cdots & n \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ n \end{matrix} & \begin{bmatrix} u^2(E_1) & u(E_1, E_2) & \cdots & u(E_1, E_n) \\ u(E_2, E_1) & u^2(E_2) & \cdots & u(E_2, E_n) \\ \vdots & \vdots & \ddots & \vdots \\ u(E_n, E_1) & u(E_n, E_2) & \cdots & u^2(E_n) \end{bmatrix} \end{matrix} \quad (11.3)$$

The values shown in red simply specify the row and column numbers and would not normally be indicated.  $u^2(E_i)$  on the diagonal represents the variance: the square of the standard uncertainty associated with the spectral irradiance value at  $\lambda_i$ . This variance is the square of the combined standard uncertainty, obtained by combining in quadrature the standard uncertainties associated with all effects, whether systematic, random or mixed. The off-diagonal terms represent the covariance associated with the measured values at two different wavelengths, as explained, through examples, below.

Radiometric uncertainties are usually expressed as relative uncertainties (fractionally or in percent) rather than absolute uncertainties (with the same units as the measurand). The covariance matrix requires "absolute" variances and covariances. Thus a covariance matrix for spectral irradiance, which has units  $\text{W m}^{-2} \text{nm}^{-1}$ , will have terms with units<sup>63</sup>  $(\text{W m}^{-2} \text{nm}^{-1})^2$ . The diagonal terms can be calculated as the square of the product of the relative standard uncertainty and the spectral irradiance value at that wavelength.

In Section 3.5 we introduced an error model where the irradiance (here at one wavelength,  $\lambda_i$ ) can be expressed as a 'true irradiance' with unknown random and systematic errors (draws from the probability distribution described by the uncertainty associated with random and systematic effects, respectively)

$$E_i = E_T + R_i + S \quad \text{Repeat of (3.28)}$$

---

<sup>63</sup> This non-standard notation for the units is to aid understanding

With this model, the variances (diagonal terms in the covariance matrix) are given by

$$u^2(E(\lambda_i)) = [u^2(S) + u^2(R_i)] \times E^2(\lambda_i). \quad (11.4)$$

The off-diagonal terms, written  $u(E_i, E_j)$ , give the covariance between the spectral irradiance value determined at  $\lambda_i$  and that determined at  $\lambda_j$ . This covariance will arise only from those effects that are common to both measured values: the systematic effects. Thus for the error model in (3.28), only the term  $u(S)$  is included. The off-diagonal covariance values for this error model are

$$u(E(\lambda_i), E(\lambda_j)) = u^2(S) \times E(\lambda_i) \times E(\lambda_j). \quad (11.5)$$

Therefore the covariance matrix, expression (11.3), takes the (symmetric) form

$$\mathbf{U}_E = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ n \end{matrix} & \begin{bmatrix} [u^2(S) + u^2(R_1)] \times E_1^2 & u^2(S) E_1 E_2 & \dots & u^2(S) E_1 E_n \\ u^2(S) E_2 E_1 & [u^2(S) + u^2(R_2)] \times E_2^2 & \dots & u^2(S) E_2 E_n \\ \vdots & \vdots & \ddots & \vdots \\ u^2(S) E_n E_1 & u^2(S) E_n E_2 & \dots & [u^2(S) + u^2(R_n)] \times E_n^2 \end{bmatrix} \end{matrix}. \quad (11.6)$$

To create this matrix, it is necessary to examine each row of the uncertainty budget and consider whether that row corresponds to a systematic or random effect with wavelength. The uncertainty associated with systematic effects should be combined in quadrature to create a single uncertainty, which becomes  $u(S)$ . Similarly, on a wavelength-by-wavelength basis, the uncertainty associated with random effects should be combined to obtain a single standard uncertainty, which becomes  $u(R_i)$ .

Sometimes a more detailed error-model is needed. If there are both multiplicative and additive effects in an error model, the error model that replaces (3.28) is

$$E(\lambda_i) = E_T(\lambda_i)(1+S)(1+R_i) + \tilde{s} + \tilde{r}_i. \quad (11.7)$$

The multiplicative effect  $S$  will include terms such as distance effects, alignment effects, etc. The multiplicative effect  $R_i$  will include terms such as light-signal noise, and rapidly varying electrical current stability, temperature sensitivity, etc. The additive term  $\tilde{s}$  will be a constant offset at all wavelengths. This term may correspond to a common dark reading subtracted at all wavelengths, stray light, etc. The additive term  $\tilde{r}_i$  corresponds to a random offset: the noise in the dark signal, variations in stray light, etc.

The covariance matrix formed from the error model (11.7) has for ( $i$ th row,  $j$ th column)

$$\tilde{U}_{E,ij} = \begin{cases} E_i^2 [u^2(S) + u^2(R_i)] + u^2(\tilde{s}) + u^2(\tilde{r}_i) & (i = j) \\ E_i E_j u^2(S) + u^2(\tilde{s}) & (i \neq j). \end{cases} \quad (11.8)$$

Note that the uncertainty associated with additive effects will generally be described as an “absolute” uncertainty and does not need to be multiplied by the spectral irradiance value. It is possible to extend this concept further to account for covariance between the measured values of a test lamp and those of a reference lamp.

## Appendix B Appendix: Monte Carlo Analysis

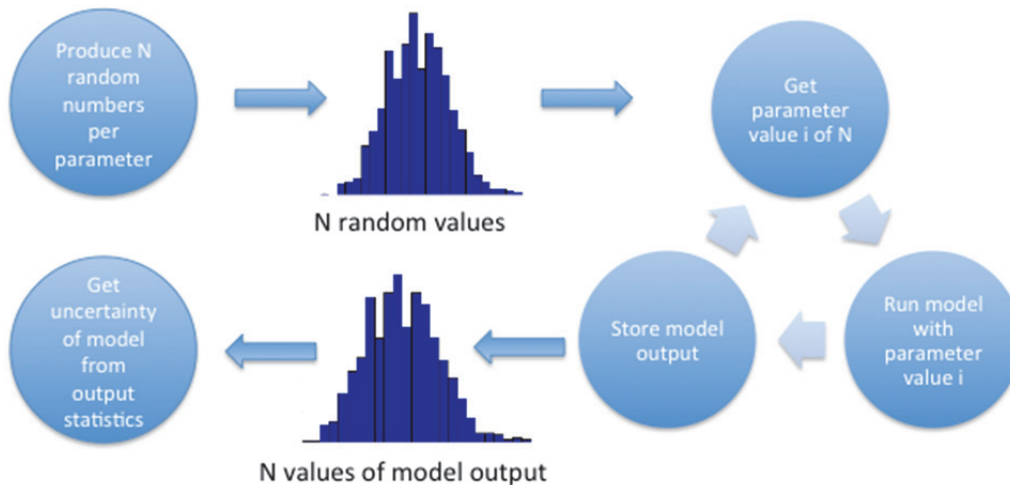
This appendix provides a recipe on how to implement a Monte Carlo simulation and uses code examples given for Matlab.

What you need before you start:

1. A model of the effect you are trying to analyse, taking a number of parameters and producing an output. This is essentially a bit of a computer code, ideally a function that encapsulates the whole model with input parameters and according output.
2. An uncertainty estimate for each parameter included in the uncertainty modelling.
3. A probability distribution function (PDF) for each parameter. Most commonly these would be a Gaussian distribution, but it may depend on the process that creates the parameter in the first place. For example, uncertainties due to sampling of a continuous variable by a digitiser will have a PDF in the form of a top hat (rectangular distribution).

The actual process of the Monte Carlo Simulation is illustrated in Figure 34 and comprises the following steps:

1. Generate  $N$  random values for each parameter as samples from the correct PDF, mean value and standard deviation. These are the realisations of the input parameter.
2. Enter a loop that will be iterated  $N$  times. For each iteration, the  $i$ -th realisation is chosen as an input parameter, the model is parameterised with the  $i$ -th value and the output stored.
3. The uncertainty of the model is estimated from the standard deviation of the model output.

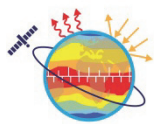


**Figure 34: Monte Carlo Simulation processes**

Table 11 introduces a Matlab code for a simple example of a Monte Carlo Simulation:

1. Realisations of an input parameter with a mean = 2 and a standard deviation, i.e. uncertainty, of 0.2 are calculated by employing the Matlab random number generator, creating 1000 random values (Figure 35). The relative uncertainty is thus  $0.2/2 \times 100 = 10\%$ .





2. These 1000 random values representing possible values the parameter could assume are fed into the model in a loop. The model is in this case the squared logarithm of the input parameter.<sup>64</sup>
3. The output of the model (Figure 36) is used to estimate the model uncertainty by computing the standard deviation of the output. In the demonstrated run the uncertainty of the model is 0.137 with a mean value of 0.48. The relative uncertainty is thus  $0.137/0.48 \times 100 = 28\%$ .

**Table 11: Matlab code for a simple Monte Carlo Simulation**

```
mu=2; % define mean value of PDF
sigma=0.2; % define standard deviation of PDF

parameter_values = (mu + sigma.*randn(1000, 1)); % generate realisations of
the parameter
% plot as histogram
figure
hist(parameter_values, 25)

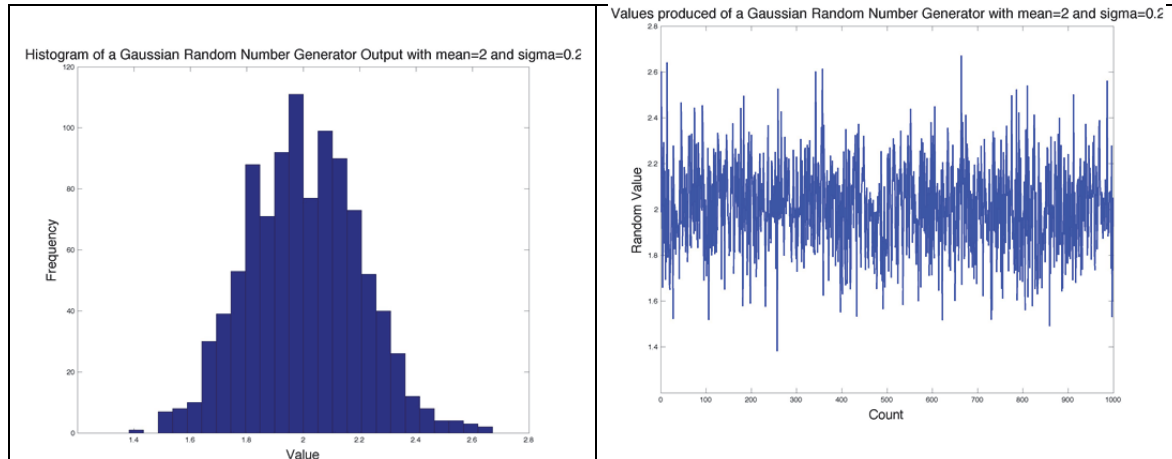
% plot vector of the random values (parameter realisations)
figure
plot(parameter_values, 'LineWidth',linewidth)

% simple model to draw from the generated parameter values
output = zeros(size(parameter_values)); % allocate output vector

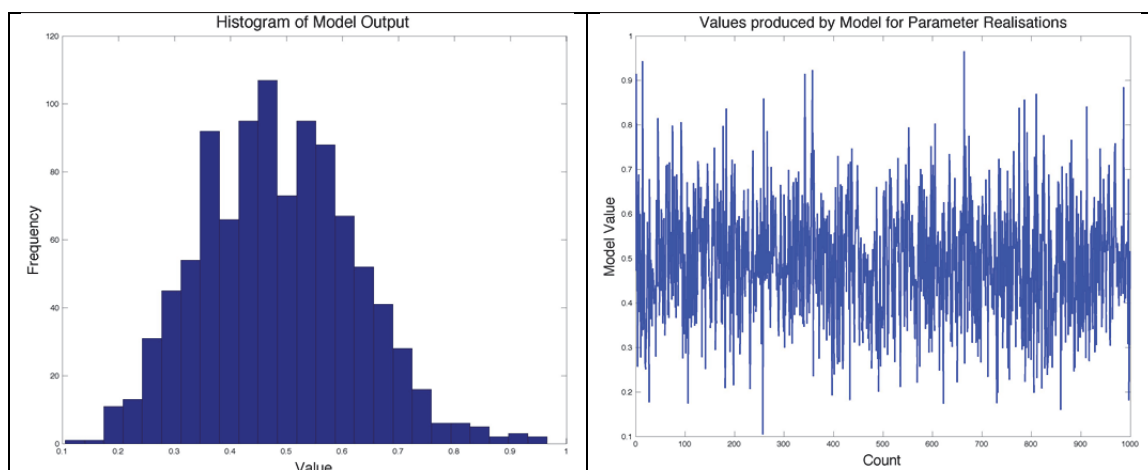
% loop over all realisations and calculate the model output
for i=1:length(parameter_values)
    parameter = parameter_values(i);
    output(i) = (log(parameter))^2;
end
% plot output as vector
figure
plot(output, 'LineWidth',linewidth)

% plot output as histogram
figure
hist(output, 25)
% get statistics of the output
mean_output = mean(output)
stddev_output = std(output)
```

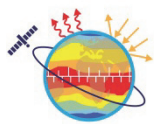
<sup>64</sup> Note that the 'for loop' is used here for educational purposes only. This simple model could of course be computed in one vector operation by `output = (log(parameter_values)).^2`. This will of course not always be possible for more complex models, in which case a loop will be required.



**Figure 35: Model input realisations as histogram (left) and value vector plot (right)**



**Figure 36: Model output as histogram (left) and value vector plot (right)**



## Appendix C References and Further Reading

### C.1 Further study and reading

Any further study on uncertainty analysis must start with the GUM itself [1]. The GUM is downloadable from <http://www.bipm.org/en/publications/guides/gum.html> and this website also contains different supplements to the GUM and an introduction to the GUM.

The best introductory textbook to the concepts of the GUM is arguably *"An introduction to uncertainty in measurement"* by Les Kirkup and Bob Frenkel. It is written in a very straightforward way and provides a good overview of the statistical concepts behind the GUM while remaining pragmatic and practical.

A slightly more advanced and detailed, but still very readable book is *"Data reduction and error analysis for the physical sciences"* by P.R. Bevington and D.K. Robinson. This book discusses the statistical basis of uncertainty analysis, and also describes Monte Carlo techniques and least square fitting.

The starting point for the application of uncertainty analysis to Earth Observation is from the QA4EO Guidelines. These are available at <http://qa4eo.org/> (see "Documentation" for the guidelines themselves and "Resources" for additional notes). There are also some case studies for the application of QA4EO to Earth Observation.

NPL offers several good practice guides on measurement and uncertainty analysis. NPL also offers a growing range of training courses – both face-to-face and e-learning. See:

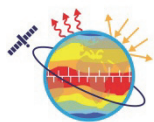
<http://www.npl.co.uk/publications/good-practice-online-modules/>.  
<http://www.npl.co.uk/learning-zone/training/>.

### C.2 Formal references

These are the numbered references in the text.

1. JCGM, JCGM 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Report, 2008.
2. JCGM, JCGM 200:2008 International vocabulary of metrology - basic and general concepts and associated terms, Report, 2008.
3. D. W. Allan, Should the classical variance be used as a basic measure in standards metrology?, IEEE Transactions on Instrumentation and Measurement, IM-36, 646-54, 1987.
4. L. Bourg and S. Delwart, MERIS Instrument Calibration, 2012.
5. P. Manninen, J. Hovila, L. Seppälä, P. Kärhä, L. Ylianttila and E. Ikonen, Determination of distance offsets of diffusers for accurate radiometric measurements, Metrologia, 43, S120-S124, 2006.
6. K. I. Itten, F. Dell'Endice, A. Hueni, M. Kneubuehler, D. Schlaepfer, D. Odermatt, F. Seidel, S. Huber, J. Schopfer, T. Kellenberger, Y. Buehler, P. D'Odorico, J. Nieke, E. Alberti and K. Meuleman, APEX - the Hyperspectral ESA Airborne Prism Experiment, Sensors, 8, 6235-6259, 2008.

7. M. Schaepman, M. Jehle, A. Hueni, K. Meuleman, T. A. Team and K. I. Itten, The 4th generation imaging spectrometer APEX and its application in Earth observation, IEEE Transactions on Geoscience and Remote Sensing, in prep., 2014.
8. P. Gege, J. Fries, P. Haschberger, P. Schoetz, H. Schwarzer, P. Strobl, B. Suhr, G. Ulbrich and W. J. Vreeling, Calibration facility for airborne imaging spectrometers, Isprs Journal of Photogrammetry and Remote Sensing, 64, 387-397, 2009.
9. A. Hueni, K. Lenhard, A. Baumgartner and M. E. Schaepman, Airborne Prism Experiment Calibration Information System, Ieee Transactions on Geoscience and Remote Sensing, 51, 5169-5180, 2013.
10. A. Hueni, J. Biesemans, K. Meuleman, F. Dell'Endice, D. Schl pfer, D. Odermatt, M. Kneubuehler, S. Adriaensen, S. Kempenaers, J. Nieke and K. I. Itten, Structure, Components, and Interfaces of the Airborne Prism Experiment (APEX) Processing and Archiving Facility, IEEE Transactions on Geoscience and Remote Sensing, 47, 29-43, 2009.
11. D. R. Taubert, J. Hollandt, P. Sperfeld, S. Pape, A. Hoepe, K. O. Hauer, P. Gege, T. Schwarzmaier, K. Lenhard and A. Baumgartner, Providing Radiometric Traceability for the Calibration Home Base of DLR by PTB, Radiation Processes in the Atmosphere and Ocean (Irs2012), 1531, 376-379, 2013.
12. D. E. Filtner and P. N. Slater, Stability of narrow-band filter radiometers in the solar-reflective range, Photogrammetric Engineering and Remote Sensing, 57, 165-171, 1999.
13. M. G. Cox, The area under a curve specified by measured values, Metrologia, 44, 365-378, 2007.
14. M. Shaw and T. Goodman, Array-based goniospectroradiometer for measurement of spectral radiant intensity and spectral total flux of light sources, Applied Optics, 47, 2637-2647, 2008.
15. Y. Zong, S. B. Brown, B. C. Johnson, K. R. Lykke and Y. Ohno, Simple spectral stray light correction method for array spectroradiometers, Applied Optics, 45, 1111-1119, 2006.
16. S. G. R. Salim, N. P. Fox, W. S. Hartree, E. R. Woolliams, T. Sun and K. T. V. Grattan, Stray light correction for diode-array-based spectrometers using a monochromator, Applied Optics, 50, 5130-5138, 2011.
17. S. Nevas, G. W bbeler, A. Sperling, C. Elster and A. Teuber, Simultaneous correction of bandpass and stray-light effects in array spectroradiometer data, Metrologia, 49, S43-S47, 2012.
18. J. Q. Sun, X. X. Xiong, A. Angal, H. D. Chen, A. S. Wu and X. Geng, Time-Dependent Response Versus Scan Angle for MODIS Reflective Solar Bands, Ieee Transactions on Geoscience and Remote Sensing, 52, 3159-3174, 2014.
19. R. A. Barnes, R. E. Eplee, G. M. Schmidt, F. S. Patt and C. R. McClain, Calibration of SeaWiFS. I. Direct techniques, Applied Optics, 40, 6682-6700, 2001.
20. [http://earth.eo.esa.int/pcs/envisat/meris/reports/cyclic/MERIS\\_CR\\_065\\_080107\\_080211.pdf](http://earth.eo.esa.int/pcs/envisat/meris/reports/cyclic/MERIS_CR_065_080107_080211.pdf)
21. X. Xiong, R. E. Eplee, Jr., J. Sun, F. S. Patt, A. Angal and C. R. McClain, Characterization of MODIS and SeaWiFS solar diffuser on-orbit degradation, Proceedings of the SPIE - The International Society for Optical Engineering, 7452, 74520Y (10 pp.)-74520Y (10 pp.), 2009.



22. G. Thuillier, M. Herse, D. Labs, T. Foujols, W. Peetermans, D. Gillotay, P. C. Simon and H. Mandel, The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the ATLAS and EURECA missions, *Solar Physics*, 214, 1-22, 2003.
23. C. J. Chunnillall, A. J. Deadman, L. Crane and E. Usadi, NPL scales for radiance factor and total diffuse reflectance, *Metrologia*, 40, S192-S195, 2003.
24. C. Olij, J. G. Schaarsberg, H. Werij, E. Zoutman, G. Baudin, B. Chommeloux, J. L. Bezy and G. Gourmelon, Spectralon diffuser calibration for MERIS, 1997.
25. J. L. Roujean, M. Leroy and P. Y. Deschamps, A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data, *Journal of Geophysical Research*, 97, 20455-68, 1992.
26. J. A. Barsi, B. L. Markham and J. A. Pedelty, The Operational Land Imager: Spectral Response and Spectral Uniformity, *Earth Observing Systems Xvi*, 8153, 2011.
27. [http://giscenter.isu.edu/research/Techpg/nasa\\_postfire/pdf/Ch3.pdf](http://giscenter.isu.edu/research/Techpg/nasa_postfire/pdf/Ch3.pdf)