Development of a radiometric uncertainty methodology for Earth Observation missions

Javier Gorroño

Submitted for the Degree of

Doctor of Philosophy

from the

University of Surrey

Surrey Space Centre

Faculty of Engineering & Physical Sciences

University of Surrey

Guildford, Surrey GU2 7XH, UK

October 2017

© Javier Gorroño 2017
Abstract

Earth Observation (EO) via remote sensing is rapidly growing in terms of satellite missions, complexity of applications and number of datasets. This situation demands that data has associated with it a quality indicator that describes the compatibility between different sensor data and suitability for particular applications.

This work describes a full end-to-end analysis of the uncertainty at a pixel level of the Top-Of-Atmosphere (TOA) radiance/reflectance factor products. It develops a methodology framework that can be adapted and reproduced by several EO missions to provide TOA radiometric uncertainty. The method is not only described but implemented as a software tool named Radiometric Uncertainty Tool (RUT) using as an example the Sentinel-2 (S2) mission.

The uncertainty methodology starts from a radiometric model, where a set of uncertainty contributors are identified and specified at a pixel level, by reviewing the pre- and post-launch sensor radiometric characterisations. These contributors are assessed using the metadata and quality information associated to the satellite products where possible. As a consequence, the uncertainty contributions are specified for the specific satellite acquisition time, scene and processing. Some of the uncertainty contributions required the use of novel estimation methods that have been specifically applied to the assessment of the uncertainty propagation produced by the image orthorectification and the radiometric impact of the spectral knowledge. The study proposes an uncertainty combination model with an important effort in using the best metrological practices as described in the ‘Guide to Expression of Uncertainty in Measurement’ (GUM) model. The assumptions in the model have been validated by comparing the results to a Monte Carlo Method (MCM), the correlation among the different uncertainty contributions has been studied, and the impact of simplifications in the combination model has been assessed. As an extension of the work towards its larger application, a methodology has been proposed and implemented to estimate the uncertainty associated to the mean of the pixels in a Region of Interest (ROI). The study considers the correlation of the pixels in the spatial, temporal and spectral dimension. As a result, the TOA radiometric uncertainty estimates can be of direct use for applications as the radiometric validation activities or product spatial binning. Further extension of the uncertainty concepts has resulted in a set of tools, algorithms and methodologies that have been used in order to estimate the radiometric uncertainty achievable for an indicative target sensor through in-flight cross-calibration using a well-calibrated hyperspectral SI-traceable reference sensor with observational characteristics such as TRUTHS (Traceable Radiometry Underpinning Terrestrial and Helio-Studies) mission. This study considers the criticality of the instrumental and observational characteristics on pixel level reflectance factors, within a defined spatial ROI within the target site. It quantifies the main uncertainty contributors in the spectral, spatial, and temporal dimension.
# Table of Contents

List of Figures ................................................................................................................. V
List of Tables .................................................................................................................. IX
List of Symbols ................................................................................................................. XI
Glossary .......................................................................................................................... XII
Acknowledgements ...................................................................................................... XVI

Chapter 1. Introduction ................................................................................................. 1
  1.1 Overview ................................................................................................................. 1
  1.2 Research motivation ............................................................................................. 1
  1.3 Definitions .............................................................................................................. 3
    1.3.1 Radiometry ....................................................................................................... 3
    1.3.2 Uncertainty analysis ......................................................................................... 4
    1.3.3 Distribution parameters ................................................................................. 7
  1.4 Motivation and aim of research .......................................................................... 7
  1.5 Research novelty ................................................................................................... 9
  1.6 Structure of the thesis and publications ............................................................... 10

Chapter 2. Literature review ....................................................................................... 12
  2.1 Introduction .......................................................................................................... 12
  2.2 Sentinel-2 Mission, MSI instrument and Level-1 products .................................. 12
  2.3 Uncertainty analysis in the EO optical missions .................................................... 15
    2.3.1 TOA Radiometric uncertainty of EO optical sensors ...................................... 15
    2.3.2 The uncertainty propagation through interpolated data .................................. 16
    2.3.3 Pixel correlation in optical radiometers ........................................................... 17
    2.3.4 Uncertainty of the biophysical and climate products ...................................... 17
    2.3.5 Uncertainty in a TOA sensor-to-sensor cross-calibration ................................. 18
  2.4 Summary of the literature review ........................................................................ 21

Chapter 3. Methodologies for a Radiometric Uncertainty tool .................................... 22
  3.1 Introduction .......................................................................................................... 22
  3.2 Sentinel 2 Level-1 Radiometric Model .................................................................. 24
  3.3 Radiometric Uncertainty Contributions ................................................................ 27
    3.3.1 Uncertainty Contributions: Identification ....................................................... 27
    3.3.2 Uncertainty Contributions: Description and Assessment .............................. 32
3.4 Model Combination and Validation........................................................................39
  3.4.1 Model Combination.........................................................................................39
  3.4.2 Discussion of the Linear Addition of Contributions.................................40
  3.4.3 Sensitivity Coefficient Impact......................................................................42
  3.4.4 Correlation between Uncertainty Contributors.............................................44
  3.4.5 Validation of the Central Limit Theorem.......................................................45
3.5 Software Implementation and Integration............................................................48
  3.5.1 Tool Integration, System Requirements and Performance.........................48
  3.5.2 Processor Documentation..............................................................................49
  3.5.3 Output Generation.........................................................................................51
  3.5.4 Radiometric Uncertainty Tool (RUT) Version 1 (v1) Case Study: Albufera Lake ..52
3.6 Conclusions and Further Work............................................................................55
Chapter 4. Novel techniques for the analysis of the TOA radiometric uncertainty........57
  4.1 Introduction........................................................................................................57
  4.2 Spectral Knowledge Uncertainty.......................................................................57
  4.3 Propagation of TOA radiometric uncertainty during orthorectification process ...65
  4.4 Conclusions and further work...........................................................................70
Chapter 5. Uncertainty for sensor-to-sensor cross-calibration ..................................72
  5.1 Introduction........................................................................................................72
  5.2 Uncertainty assessment: Spectral Domain.........................................................75
    5.2.1 Spectral domain: methodology.................................................................75
    5.2.2 Spectral domain: systematic sampling/ resolution error results..................78
    5.2.3 Spectral domain: spectral knowledge uncertainty.......................................83
    5.2.4 Spectral domain: the impact of spectral binning.........................................85
    5.2.5 Spectral domain: the TOA radiance sampling............................................87
    5.2.6 Spectral domain: the impact of the site.....................................................87
  5.3 Uncertainty assessment: Spatial Domain............................................................89
    5.3.1 Spatial domain: methodology.................................................................89
    5.3.2 Spatial domain: results for La Crau calibration site.....................................92
    5.3.3 Spatial domain: results for PICS sites.......................................................94
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>5.4.1 Temporal domain: methodology</td>
<td>98</td>
</tr>
<tr>
<td>5.4</td>
<td>5.4.2 Temporal domain: atmospheric variation and radiative transfer code impact</td>
<td>99</td>
</tr>
<tr>
<td>5.4</td>
<td>5.4.3 Temporal domain: atmospheric knowledge</td>
<td>101</td>
</tr>
<tr>
<td>5.4</td>
<td>5.4.4 Temporal dimension: atmospheric variation</td>
<td>102</td>
</tr>
<tr>
<td>5.4</td>
<td>5.4.5 Temporal domain: surface impact</td>
<td>104</td>
</tr>
<tr>
<td>5.5</td>
<td>5.5.1 Discussion: spectral domain</td>
<td>107</td>
</tr>
<tr>
<td>5.5</td>
<td>5.5.2 Discussion: spatial domain</td>
<td>108</td>
</tr>
<tr>
<td>5.5</td>
<td>5.5.3 Discussion: temporal domain</td>
<td>110</td>
</tr>
<tr>
<td>5.5</td>
<td>5.5.4 Discussion: uncertainty budget</td>
<td>111</td>
</tr>
<tr>
<td>5.6</td>
<td>Conclusion</td>
<td>113</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>The correlation of the TOA reflectance/radiance pixel measurements</td>
<td>116</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>116</td>
</tr>
<tr>
<td>6.2</td>
<td>Concept of study, the limitations and the methodology</td>
<td>117</td>
</tr>
<tr>
<td>6.2</td>
<td>6.2.1 Error, uncertainty and correlation</td>
<td>117</td>
</tr>
<tr>
<td>6.2</td>
<td>6.2.2 Uncertainty over a ROI pixel mean</td>
<td>118</td>
</tr>
<tr>
<td>6.2</td>
<td>6.2.3 Spatial, temporal and spectral dimension of the S2 MSI</td>
<td>119</td>
</tr>
<tr>
<td>6.2</td>
<td>6.2.4 An approximation method: “select/deselect”</td>
<td>120</td>
</tr>
<tr>
<td>6.2</td>
<td>6.2.5 MCM propagation</td>
<td>121</td>
</tr>
<tr>
<td>6.3</td>
<td>Qualitative assessment of the pixel-to-pixel correlation</td>
<td>122</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.1 $u_{\text{noise}}$: Instrument noise</td>
<td>122</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.2 $u_{\text{stray_sys}}$: Out-of-field Stray-light — systematic part</td>
<td>123</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.3 $u_{\text{stray_rand}}$: Out-of-field Stray-light — random part</td>
<td>124</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.4 $u_{\text{xtalk}}$: Crosstalk</td>
<td>124</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.5 $u_{\text{ADC}}$: Analog-to-digital conversion quantisation</td>
<td>124</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.6 $u_{\text{DS}}$: Dark signal stability</td>
<td>125</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.7 $u_{\text{gamma}}$: Non-linearity and non-uniformity knowledge</td>
<td>125</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.8 $u_{\text{diff_abs}}$: Diffuser reflectance absolute knowledge</td>
<td>126</td>
</tr>
<tr>
<td>6.3</td>
<td>6.3.9 $u_{\text{diff_temp}}$: Diffuser reflectance temporal knowledge</td>
<td>126</td>
</tr>
</tbody>
</table>
List of Figures

Figure 2-1 Multi-Spectral Instrument (MSI) internal configuration with (left) full instrument view (diffuser panel in yellow, telescope mirrors in dark blue) and (right) optical path construction for the splitter and SWIR/VNIR focal planes (reproduced with permission from (Gascon, Bouzinac et al. 2017)).

Figure 3-1 Illustration of the MultiSpectral Instrument (MSI) radiometric model and Level 1 (L1) ground processing with the associated parameters described in Equations (3.1)-(3.8).

Figure 3-2 Distribution of the sensitivity coefficient $c_y$ for all the pixels in the Sentinel-2 (S2) Visible and Near-InfraRed (VNIR) bands (a) B1; (b) B2; (c) B3; (d) B4; (e) B5; (f) B6; (g) B7; (h) B8 and (i) B8A.

Figure 3-3 Distribution of the sensitivity coefficient $c_y$ for all the pixels in the S2 Short-Wave InfraRed (SWIR) bands parameter $a_1$ (a) B11; (b) B12 and parameter $a_2$; (c) B11; (d) B12.

Figure 3-4 S2 L1B model validation for the S2 bands (a) B1; (b) B2; (c) B3; (d) B4; (e) B5; (f) B6; (g) B7; (h) B8; (i) B8A; (j) B11; and (k) B12. The graph shows the difference when calculating the uncertainty using the ‘Guide to Expression of Uncertainty in Measurement’ (GUM) uncertainty ($k = 1$) or the 68.27% area of the Monte Carlo Method (MCM) symmetric from its mean.

Figure 3-5 L1B MCM distribution at (left) 4.93 Wm$^{-2}$sr$^{-1}$μm$^{-1}$ and (right) 31.41 Wm$^{-2}$sr$^{-1}$μm$^{-1}$. The theoretical normal distribution of the combined standard uncertainty has been included and re-scaled to the MCM distribution peak value for comparison.

Figure 3-6 Screen-shot of the “I/O Parameters” tab of Radiometric Uncertainty Tool (RUT) tool as integrated in Sentinel Application Platform (SNAP).

Figure 3-7 Screen-shot of the “Processing Parameters” tab of RUT tool as integrated in SNAP.

Figure 3-8 Screen-shot of SNAP. It contains the pixel info and navigation panels (left); the L1C Sentinel-2 B8 image for Albufera Lake (centre) and the equivalent uncertainty $k = 1$ (right). The image is North oriented.

Figure 3-9 Screen-shot of L1C reflectance factor scale in Figure 3-8 (top); and the scale of equivalent uncertainty $k = 1$ (bottom). The L1C reflectance factor is multiplied by the quantification value of 10,000 and the uncertainty figures are given as percentages multiplied by 10.

Figure 3-10 Screen-shot of L1C Sentinel-2 B8 image of Albufera Lake and surroundings on the 12th of November 2017. Overlaid on the image, the considered ROIs under study.

Figure 4-1. Diagram describing the S2 spectral uncertainty assessment.
Figure 4-2. Simulated TOA spectral radiance for a desert case and the S2 bands in the VNIR range. 60

Figure 4-3. TOA radiance dispersion associated to S2 spectral uncertainty using cubic spline interpolation for B1 (a), B2 (b) B3 (c) B4 (d) B5 (e) B6 (f) B7 (g) B8 (h) and B8A (i). .............................. 62

Figure 4-4. TOA radiance associated to S2 spectral uncertainty results using PCHIP interpolation for B1 (a), B2 (b) B3 (c) B4 (d) B5 (e) B6 (f) B7 (g) B8 (h) and B8A (i). ................................. 63

Figure 4-5. Resampling process of the S2 L1C products (reproduced from (ESA 2017)). ........................ 65

Figure 4-6. L1C uncertainty propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels of the interpolation kernel. The settings for each panel are full random L1B uncertainty using (a) the bi-linear interpolation, (b) cubic convolution and (c) B-splines and fully correlated absolute calibration coefficient uncertainty \( A(b) \) for the L1B uncertainty using (d) the bi-linear interpolation, (e) cubic convolution and (f) B-splines. ............................... 67

Figure 4-7. L1C radiance propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels introducing a fully correlated absolute calibration coefficient uncertainty \( A(b) \) for the L1B measurements and using the bi-linear interpolation (a), cubic convolution (b) and B-splines (c). L1C uncertainty propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels of the interpolation kernel introducing a fully correlated absolute calibration coefficient uncertainty \( A(b) \) for the L1B measurements using the bi-linear interpolation (d), cubic convolution (e) and B-splines (f). The values of radiance in the L1B interpolation kernel are (values from left to right and up to down in a 4x4 kernel): \([21, 20, 10, 14], [15, 30, 24, 50], [10, 12, 30, 60], [12, 45, 36, 15]\) \( \text{Wm}^2\text{sr}^{-1}\mu\text{m}^{-1} \). ................................. 69

Figure 5-1 Illustrative method of the TOA TRUTHS spectral profile generation. The red stars are the measurements at the native spectrometer bands and the green stars are the result of the merging to a design specified bin. The merged measurements are sampled at a specified interval to obtain the reconstructed TOA as measured by TRUTHS (‘TRUTHS TOA @5x10^{-4}’). .............................. 75

Figure 5-2 (Left) Preliminary design of the TRUTHS Earth Imager spectrometer and (right) the translation of the native spectral sampling design in the instrument SRF. ................................. 76

Figure 5-3 TOA radiance as generated by MODTRANv5, resulting measurements of TOA radiance as measured by TRUTHS Earth imager and the Sentinel 2 VNIR and SWIR bands. ........................... 78

Figure 5-4 TOA error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands (above, B1-B8A) and SWIR bands (below B11 & 12) due to the TRUTHS sampling bands and preliminary resolution of the detector bands. The errors are plotted for different types of interpolation to reconstruct the TOA radiance and Sentinel 2 bands. ................................. 80
Figure 5-5 Difference between interpolations for B1 in Figure 5-4. The error at each wavelength shift position has been normalised by the mean. The legend of the plot is equivalent to Figure 5-4. 

Figure 5-6 Spectral sampling error for each S2 band considered in the cases of minimum, maximum rms array position error for a S2 SRF linear interpolation and TOA radiance cubic spline interpolation (left) and linear interpolation (right). 

Figure 5-7 Distribution of spectral sampling errors for S2 bands with an associated TRUTHS central wavelength and bandwidth knowledge uncertainty of 0.2 nm ($k = 1$). 

Figure 5-8 (Left) TOA reflectance error in estimating the Sentinel-2 MSI equivalent radiance in B1 due to the TRUTHS sampling bands and preliminary resolution of the detector bands with no spectral binning applied. (Centre) The measured TOA radiance as generated by MODTRANv5 in black colour, in red colour the resulting measurements of TOA radiance as measured by TRUTHS Earth imager and, in blue colour, the Sentinel 2 B1 band. (Right) The distribution of spectral sampling errors for S2 B1 band with an associated TRUTHS central wavelength and bandwidth knowledge uncertainty of 0.2 nm ($k = 1$) with no spectral binning applied. The legend of the plots are equivalent to Figure 5-4, Figure 5-3, and Figure 5-7 respectively. 

Figure 5-9 TOA error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands B1 (left) and B6 (right) due to the TRUTHS sampling bands and preliminary resolution of the detector bands. The legend of the plots are equivalent to Figure 5-4. 

Figure 5-10 TOA error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands (above) and SWIR bands (below) due to the TRUTHS sampling bands and preliminary resolution of the detector bands for different modelled sites. 

Figure 5-11 Methodology process for the assessment of spatial variations. 

Figure 5-12 TOA reflectance factor at the 400 × 400 m$^2$ at the LaCrau site for the considered L8 OLI and S2 MSI bands. 

Figure 5-13 TOA reflectance factor error map for the LaCrau site and the considered L8 OLI and S2 MSI bands. 

Figure 5-14 Spatial uncertainty vs. spatial offset for the LaCrau site and the considered L8 OLI and S2 MSI bands. The bands contained in each panel are: (a) S2 B1 and L8 B1, (b) S2 B8A and L8 B5, and (c) S2 B12 and L8 B7. 

Figure 5-15 TOA reflectance factor at the 20 × 20 km$^2$ at the Libya-4 site for the considered L8 OLI and S2 MSI bands. 

Figure 5-16 TOA reflectance factor error map for the Libya-4 site and the considered L8 OLI and S2 MSI bands.
Figure 5-17 Spatial uncertainty vs. spatial offset for the Libya-4 site for L8 OLI and S2 MSI bands. The bands contained in each panel are: (a) S2 B1 and L8 B1, (b) S2 B8A and L8 B5, and (c) S2 B12 and L8 B7.................................................................97

Figure 5-18 TOA reflectance factor error distribution for the Libya-4 site and L8 OLI B7 (left) and S2 MSI B12 (right)........................................................................................................................97

Figure 5-19 TOA reflectance variation for wavelengths 443 nm (left), 865 nm (centre), and 2201 nm (right) in Libya4 for the year day 173 and 355 over 30 minutes using the radiative codes of MODTRAN and 6SV1..........................................................................................................................100

Figure 5-20 Dispersion of reflectance factor errors at 30 minutes at 443 nm and yearday 173 (a), 865 nm and yearday 173 (b), 2201 nm and yearday 173 (c), 443 nm and yearday 355 (d), 865 nm and yearday 355 (e), and 2201 nm and yearday 355(f). .................................................................................................................................102

Figure 5-21 Results for TOA radiance dispersion at 30 minutes at 443 nm and yearday 173 (a), 865 nm and yearday 173 (b), 2201 nm and yearday 173 (c), 443 nm and yearday 355 (d), 865 nm and yearday 355 (e), and 2201 nm and yearday 355(f). .........................................................................................................................103

Figure 5-22 Results for surface reflectance error dispersion at 30 minutes at 443 nm and yearday 173 (a), 443 nm and yearday 355 (b), 865 nm and yearday 173 (c), and 865 nm and yearday 355 (d). ....106

Figure 6-1. VNIR focal plane schematic description. The image is reproduced with permission from (Gascon, Bouzinac et al. 2017). .................................................................................................................................119

Figure 6-2. Image of the S2-RUTv1 dialog box with the tab “Processing parameters” selected. This tab permits the selection and deselection of each uncertainty contribution.................................................121

Figure 6-3. Schematic of the MCM propagation for the ROI mean uncertainty estimate..............122

Figure 6-4. Evolution of the ROI uncertainty (k =1) with the ROI size for the RadCalNet Gobabeb site using the MCM technique....................................................................................................................134

Figure 6-5. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty (k =1) of the RadCalNet Gobabeb site. .........................134
Figure 6-6. Evolution of the ROI uncertainty ($k = 1$) with the ROI size for the Boussole site using the MCM technique. .................................................................................................................. 135

Figure 6-7. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty ($k = 1$) of the Boussole site. .......................................................... 136

Figure 6-8. Evolution of the 500m ROI uncertainty ($k = 1$) with the variation $u_{\text{stray,sys}}$ for the Boussole site using the MCM propagation technique. .................................................................................. 137

Figure 6-9. Overpass of Sentinel 2A over the Boussole site on the 28th of March 2017 using the COVE tool (Kessler, Killough et al. 2013). .................................................................................................................. 138

Figure 6-10. Evolution of the ROI uncertainty ($k = 1$) with the ROI size for the DCC site using the MCM technique. The simulation used a spectral correlation of $u_{\text{diff,k}}$ of 0.5 and $u_{\text{diff_abs}}$ of 1. .............................. 139

Figure 6-11. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty ($k = 1$) of the DCC site. The simulation used a spectral correlation of $u_{\text{diff,k}}$ of 0.5 and $u_{\text{diff_abs}}$ of 1. .................................................................................................................. 139

Figure 6-12. Evolution of the ratio of ROIs uncertainty ($k = 1$) with the variation of $u_{\text{diff_abs}}$ spectral correlation for the DCC site using the MCM technique. The simulation used a spectral correlation of $u_{\text{diff,k}}$ of 0.5. .................................................................................................................. 140

List of Tables

Table 1-1 Radiometric quantities (Palmer and Grant 2009) .................................................................................................................. 3

Table 2-1. S2 MSI spectral bands naming with associated central wavelength and spectral bandwidth. .................................................................................................................................................. 14

Table 3-1 List of Sentinel-2 L1 radiometric uncertainty contributions and their associated parameter in the radiometric processing model. The table also indicates the contributions that are considered for the S2-RUTv1 implementation (marked as “Y”), the ones that may be included in next versions of the tool (marked as “N”) and the ones that have a negligible effect (i.e., <0.1%). ........................................................................................................... 28

Table 3-2. Considered uncertainty contributors for the L1B model validation. ................................................................................. 46

Table 3-3. Statistics for the considered ROIs in Figure 3-10 and for the S2 L1C product acquired on the 12th of November 2017. .................................................................................................................. 54

Table 4-1. Parameters associated to Figure 4-3. .................................................................................................................. 61

Table 4-2. Parameters associated to Figure 4-4. .................................................................................................................. 62
Table 4-3. Parameters associated to the results in Figure 4-6........................................67
Table 4-4. RMSE error between the interpolated radiance of panels (a), (b) and (c) from Figure 10. 69
Table 5-1 Standard deviation results of the distribution of spectral sampling errors presented in Figure 5-7 ................................................................................................................................. 84
Table 5-2 S2 MSI and L8 OLI products for the spatial uncertainty assessment and validation......... 90
Table 5-3. Statistical parameters for Figure 5-20...................................................................... 101
Table 5-4. Statistical parameters for Figure 5-21...................................................................... 103
Table 5-5. Several statistical parameters for Figure 5-22 ....................................................... 106
Table 5-6. Summary of the different sources of uncertainty investigated for a cross-comparison of TRUTHS and Sentinel-2.............................................................................................................. 114
Table 6-1. Summary of pixel correlation for radiometric validation sites..............................128
List of Symbols

\( A \)  
Area

\( C \)  
Correction coefficient

\( E \)  
Irradiance

\( L \)  
Radiance

\( \bar{R}_{rs} \)  
Remote Sensing Reflectance

\( Q \)  
Radiant Energy

\( V \)  
Voltage

\( W \)  
Wind speed

\( \Phi \)  
Radiant Power

\( \Omega \)  
Projected solid angle

\( f \)  
Function

\( k \)  
Coverage factor

\( r \)  
Correlation coefficient

\( S \)  
Space

\( t \)  
Time

\( \Lambda \)  
Wavelength

\( \phi \)  
Azimuth angle

\( \mu \)  
Standard uncertainty

\( \Theta \)  
Zenith angle

\( \rho \)  
Reflectance factor

\( \omega \)  
Solid angle

\( [J] \)  
joule

\( [W] \)  
watt

\( [m] \)  
metre

\( [sr] \)  
Stereo-radian
# Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>Across-Track</td>
</tr>
<tr>
<td>ADC</td>
<td>Analog-to-Digital Conversion</td>
</tr>
<tr>
<td>ALT</td>
<td>Along-Track</td>
</tr>
<tr>
<td>AOT</td>
<td>Aerosol Optical Thickness</td>
</tr>
<tr>
<td>ATSR</td>
<td>Along-Track Scanning Radiometers</td>
</tr>
<tr>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometers</td>
</tr>
<tr>
<td>BinGO</td>
<td>BlInning patterN Generator and Optimiser</td>
</tr>
<tr>
<td>BRDF</td>
<td>Bidirectional Reflectance Distribution Function</td>
</tr>
<tr>
<td>BRF</td>
<td>Bidirectional Reflectance Factor</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge-Coupled Device</td>
</tr>
<tr>
<td>CCI</td>
<td>Climate Change Initiative</td>
</tr>
<tr>
<td>CEOS</td>
<td>Committee on Earth Observation Satellites</td>
</tr>
<tr>
<td>CERES</td>
<td>Clouds and the Earth’s Radiant Energy System</td>
</tr>
<tr>
<td>CLARREO</td>
<td>Climate Absolute Radiance and Refractivity Observatory</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary Metal–Oxide–Semiconductor</td>
</tr>
<tr>
<td>CSM</td>
<td>Calibration and Shutter Mechanism</td>
</tr>
<tr>
<td>DCC</td>
<td>Deep Convective Cloud</td>
</tr>
<tr>
<td>DN</td>
<td>Digital Number</td>
</tr>
<tr>
<td>DoP</td>
<td>Degree Of Polarisation</td>
</tr>
<tr>
<td>DPM</td>
<td>Degree Polarisation Model</td>
</tr>
<tr>
<td>DS</td>
<td>Dark Signal</td>
</tr>
<tr>
<td>ECV</td>
<td>Environmental Climate Variable</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>EO</td>
<td>Earth Observation</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FEE</td>
<td>Front-End Electronics</td>
</tr>
<tr>
<td>FOV</td>
<td>Field-of-View</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field Programmable Gate Array</td>
</tr>
</tbody>
</table>
FPN  Focal Plane Noise
FWHM Full-Width Half Maximum
GCOS  Global Climate Observing System
GIFOV Ground Instantaneous Field-Of-View
GUM  Guide to Expression of Uncertainty in Measurement
IR  Infrared
IVOS  Infrared and Visible Optical Sensors Subgroup
L1  Level-1
L7 ETM+ Landsat 7 Enhanced Thematic Mapper Plus
L8  Landsat 8
LSB  Lowest Significant Bit
LUT  Look-Up Table
MCM  Monte Carlo method
MCT  Mercure Cadmium Telluride
MERIS MEedium Resolution Imaging Spectrometer
MODIS Moderate Resolution Imaging Spectroradiometer
MODTRAN MODerate resolution atmospheric TRANsmission
MSI  Multispectral Imager
NIR  Near-Infrared
NIST  National Institute of Standards and Technology
NMI  National Measurement Institute
NPL  National Physical Laboratory
OLCI Ocean Land Colour Instrument
OLI  Operational Land Imager
PCHIP  Piecewise Cubic Hermite Interpolating Polynomials
PDF  Probability Distribution Function
PICS  Pseudo-Invariant Calibration Sites
PSF  Point Spread Function
PTFE Politetrafluoroetileno
QA4EO Quality Assurance for Earth Observation
RAA Relative Azimuth Angle
RadCalNet Radiometric Calibration Network
RMS Root Mean Square
ROI Region-Of-Interest
ROIC Readout Integrated Circuit
RPV Rahman-Pinty-Verstraete
RT Radiative Transfer
RUT Radiometric Uncertainty Tool
S3 Sentinel-3
SAR Synthetic Aperture Radar
SBAF Spectral Bandwidth Adjustment Factor
SI International System of Units; acronym from French *Système International d'Unités*
SNAP Sentinel Application Platform
SNR Signal-to-Noise Ratio
SRCA Spectroradiometric Calibration Assembly
SRF Spectral Response Function
SST Sea Surface Temperature
SWIR Short-Wave Infrared
SZA Sun Zenith Angle
TDI Time Delay Integration
TIR Thermal Infra-Red
TOA Top of Atmosphere
TMA Three-Mirror Anastigmatic
TRUTHS Traceable Radiometry Underpinning Terrestrial- and Helio- Studies
UTM Universal Transverse Mercator
UV Ultraviolet
VAA Viewing Azimuth Angle
VIIRS Visible Infrared Imaging Radiometer Suite
VIS Visible
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCU</td>
<td>Video Chain Unit</td>
</tr>
<tr>
<td>VNIR</td>
<td>Visible and Near Infrared</td>
</tr>
<tr>
<td>VZA</td>
<td>Viewing Zenith Angle</td>
</tr>
<tr>
<td>WGCV</td>
<td>Working Group on Calibration and Validation</td>
</tr>
</tbody>
</table>
Acknowledgements

First of all, I would like to thank my parents and family who supported me in the difficult moments of this PhD (which there were…).

The work here presented has been a long journey and started at the European Space Agency (ESA). I would like to thank ESA specifically for the support of their staff: Dr. Ferran Gascon who supervised my work at ESA ESRIN and has been continuously advising this work, Steven Delwart who gave me helpful technical and morale support, and Dr. Bojan Bojkov and Dr. Philippe Goryl who have continuously supported the development of the RUT tool.

I would like to thank the National Physical Laboratory where I work which has provided me with the possibility to develop this thesis. I would specifically thank Dr. Nigel Fox who has constantly supported and guided me in this work. Many thanks to the rest of my colleagues who have supported me at a technical and morale level on the day-to-day level during these years at NPL: Dr. Evangelos Theocharous, Niall Origo, Agnieszka Bialek, Dr. Andrew Banks, and the rest of the NPL ECO team.

Specific thanks to the Surrey Space Centre (SSC) of the University of Surrey for hosting me as a student. Many thanks to my supervisors Prof. Craig Underwood and Sir Prof. Martin Sweeting. Their support and guidance has been excellent and has resulted in a simplified and smooth process.

The work developed here has been implemented and made available to the users. I would like to thank Brockmann Consult GmbH since thanks to their expertise in the field have made this implementation possible. Specifically, I thank Norman Fomferra and Marco Peters since they have been key important in this process and have taught me many important issues in the field of software programming.

Thanks to the ESA, National Measurement System of the UK government’s Department for Business, Energy and Industrial Strategy, the UK space Agency’s Centre for Earth Observation Instrumentation and the European Metrology Research Programme (EMRP) for the funding of this work. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.
Chapter 1.

Introduction

1.1 Overview

Earth Observation via remote sensing provides an important source of information about the global Earth system. Satellites constantly capture observation data from the Earth. These data are processed and complemented by other auxiliary data in order to monitor natural resources, describe biophysical processes, extract geo-information or develop climate models.

The complexity of many of these applications and the growing number of records, makes necessary that these data relies in a quality indicator that describes the compatibility between different sensor data and the suitability for a certain application. The quality assured data, ideally, should be SI (International System of Units; acronym from French Système International d’Unités) traceable and accompanied with uncertainty estimates. The latter is the core of this research for which a software and scientifically rigorous solution is pursued.

1.2 Research motivation

The motivation for this research originates from the mandate of the Infrared and Visible Optical Sensors (IVOS) subgroup which is part of the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration & Validation (WGCV). The group states that: "Data and derived products shall have associated with them an indicator of quality to enable users to assess their suitability for particular applications, i.e., their “fitness for purpose”" (QA4EO 2010).

The quality indicator studied here is the uncertainty which is defined as a "parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand" (BIPM, IEC et al. 2008). Further information can be found in Section 1.3.2.

In this case, the measurand is either top-of-atmosphere (TOA) radiance or reflectance factor measured by an Earth Observation (EO) optical sensor at a pixel-level. These two quantities are defined in Section
1.3.1. Regarding the radiometric uncertainty associated to it, this must be based on a documented quantitative assessment traceable to the reference standards. The resulting uncertainty parameter must be propagated end-to-end through all steps in the data collection, processing and dissemination (QA4EO 2010).

The quantitative assessment here is based on the GUM — Guide to the expression of uncertainty in measurement (BIPM, IEC et al. 2008). This uncertainty framework — described in Section 1.3.2 — is the starting point to define a strategy and methodology that estimates the TOA radiometric uncertainty. The uncertainty assessment requires that each uncertainty contribution is linked to a mathematical representation of the TOA radiance/reflectance factor calculation. The direct application of the GUM for the combination of uncertainty contributions in TOA radiance/reflectance factor products brings out certain limitations that must be considered and, wherever possible, overcome.

By providing TOA radiometric uncertainty estimates it is possible that the users of EO products understand the “fitness for purpose” of the data to their specific application. In addition, it will be possible that the uncertainty provided can be propagated through consecutive steps of the EO processing chain being effectively the input to the higher-level products uncertainty estimates. That is, the study and rigorous evaluation of TOA radiometric uncertainty represents a key step that enables any further evaluation of uncertainty and its propagation into a final application.

An example of the benefits for end-users can be found when selecting satellite products to be ingested in a climate model. The TOA radiometric uncertainty can be used as an input and its propagation throughout the entire chain can ultimately estimate the uncertainty associated to the output (e.g., temperature rise over the century). By including uncertainty estimates in the datasets, the users can define the required datasets against their requirements — e.g. a requirement of 0.5 K ($k = 1$) — or find out the current limitations of the available products.

Additionally, it is important to mention that a rigorous uncertainty analysis may lead to a better understanding of the instrument and its radiometric performance. In better words:

“Proper estimation of uncertainties, rather than over-estimation, then leads to the increased probability of detecting systematic effects which may have been overlooked in the original analysis, which in turn leads to a better understanding of the practice of spectral radiometry.” (Gardner 2004).

The in-depth understanding of the satellite instrument and processing of TOA radiance/reflectance factor products might be helpful to identify, for example, the key limiting areas in the system. That is, a sensitivity analysis combined with a rigorous uncertainty budget can determine whether improvements of noise or calibration drifts have a positive impact on the radiometric performance and, ultimately, on the final application.
1.3 Definitions

1.3.1 Radiometry

Radiometry is defined as the measurement of optical radiant energy. This constitutes a measurement of the amplitude of the electromagnetic wave in a Maxwell physic approximation or the number of photons in quantum approach. The optical part of electromagnetic spectrum ranges from 10 nm to 1000 µm and is divided into ultraviolet (UV), visible (VIS) and Infrared (IR) regions. The spectral range of interest for the research presented in this document includes visible and infrared region with wavelengths range from 400 nm–3000 nm covering the VIS, Near Infrared (NIR) and Short Wave Infrared (SWIR) regions (the latter two sub-regions in the IR region). Table 1-1 presents a list of radiometric quantities that this document will refer to in the following sections.

Table 1-1 Radiometric quantities (Palmer and Grant 2009)

<table>
<thead>
<tr>
<th>Radiometric quantity</th>
<th>Equation and units</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiant Energy</td>
<td>$Q [J]$</td>
<td></td>
</tr>
<tr>
<td>Radiant Power</td>
<td>$\Phi = \frac{dQ}{dt} [W]$</td>
<td>Energy per unit time</td>
</tr>
<tr>
<td>Irradiance</td>
<td>$E = \frac{d\Phi}{dA} \left[ \frac{W}{m^2} \right]$</td>
<td>Power per unit area that is incident on a surface. Irradiance is measured at the detector</td>
</tr>
<tr>
<td>Solid angle</td>
<td>$\omega [sr]$</td>
<td>The plane-angle concept extended to three-dimension</td>
</tr>
<tr>
<td>Radiance</td>
<td>$L = \frac{d^2\Phi}{dA d\Omega} \left[ \frac{W}{m^2 sr} \right]$</td>
<td>Power per unit area and per unit projected solid angle.</td>
</tr>
<tr>
<td>BRDF</td>
<td>$f_r(\theta, \phi, \theta_i, \phi_i) = \frac{dL_r(\theta, \phi)}{dE_r(\theta, \phi)} \frac{dL(\theta, \phi)}{L(\theta, \phi) d\Omega} \left[ sr^{-1} \right]$</td>
<td>Differential element of reflected radiance in a specified direction per unit differential element of irradiance, also in specified direction</td>
</tr>
</tbody>
</table>

The radiant power $\Phi$ is defined as the derivative of the radiant energy $Q$ by the time $t$. The radiant power per unit of incident area $A_s$ is referred as irradiance. The radiance $L$ represents the irradiance over a projected solid angle $\Omega$. The projected solid angle is defined by the solid angle $\omega$ projected onto the plane of the observer.
The radiometric units measured in this research are the radiance and the reflectance factor $\rho$. The reflectance factor definition for optical remote sensing measurements is given as:

“...the ratio of the radiant flux reflected by a surface to that reflected into the same reflected-beam geometry and wavelength range by an ideal (lossless) and diffuse (Lambertian) standard surface, irradiated under the same conditions” (Schepman-Strub, Schepman et al. 2006).

Strictly following the naming in Nicodemus, Richmond et al. (1977) and referred in Schepman-Strub, Schepman et al. (2006), the TOA reflectance factor as measured by the optical instruments under study here, is a bi-conical reflectance factor. However, when the pixel field-of-view (FOV) is relatively small, this can be approximated as a Bidirectional Reflectance Factor (BRF). This is a desired quantity since it is not integrated over a range of angles and is defined as the ratio of the reflected radiant flux from a surface area $dA$ to the reflected radiant flux from an ideal and diffuse surface of the same area $dA$ under identical view geometry and single direction illumination:

$$\text{BRF} = \frac{d\Phi_r(\theta, \phi; \theta', \phi')}{d\Phi^d_r(\theta, \phi)} = \frac{dE_r(\theta, \phi)}{dL^d_r(\theta, \phi)} \frac{dL_r(\theta, \phi; \theta', \phi')}{dE_r(\theta, \phi)} = \pi \cdot \text{BRDF}$$ (1.1)

The BRF [unitless] of any surface can be expressed as its Bidirectional Reflectance Distribution Function (BRDF) [sr$^{-1}$] times $\pi$ approximation since an ideal Lambertian surface has a BRDF of $1/\pi$ at any viewing angle.

The relationship between the radiance and reflectance factor is based on the assumption that the Sun is an isotropic point under a lossless medium and, thus, the inverse square law of irradiance applies. In addition, it takes into account a cosine correction when the rays are not normal to the optical axis (Palmer and Grant 2009). It is written as:

$$\text{BRF}(\theta, \phi; \theta', \phi') = \pi \cdot \text{BRDF}(\theta, \phi; \theta', \phi') = \frac{\pi \cdot L(\theta, \phi; \theta', \phi')}{E_{TOA}} = \frac{\pi \cdot L(\theta, \phi; \theta', \phi')}{E_S \cdot d(t)^2 \cdot \cos(\theta_t)}$$ (1.2)

Where $E_S$ is the equivalent extra-terrestrial solar spectrum and $d(t)$ is a correction to take into account the sun-Earth distance variation.

### Uncertainty analysis

The theoretical framework for an uncertainty evaluation is defined in the GUM — Guide to the expression of uncertainty in measurement (BIPM, IEC et al. 2008).

Uncertainty ought to be quoted with appropriate coverage factor that defines how confident we are about our best estimate. It is defined as one standard deviation from the mean, assuming a normal distribution function, and expresses the confidence level that the true value falls within the 68.27% values around the estimate. The coverage factor, $k$, is a numerical factor used as a multiplier of the
combined standard uncertainty in order to specify the fraction of the probability distribution that the uncertainty represents. The expanded uncertainty is the result of the multiplication of the combined standard uncertainty by the coverage factor (BIPM, IEC et al. 2008). Most of the measurements performed at National Measurement Institutes (NMIs) use coverage factor $k=2$, that is equivalent to 95% of the measured values. Some fields of science use $k=3$ coverage factor that is defined as 99% and is mainly for risk management and medical application, were 99% confidence is essential for life saving purposes. Throughout the development of the work in this document, the default coverage factor $k=1$ is quoted. Since the goal is a tool development, the default coverage can be automatically changed to any desired value.

The association of the coverage factor, $k$, with an equivalent fraction of the probability distribution relies on the validity of the central limit theorem. For $Y = c_1X_1 + c_2X_2 + ... + c_nX_n = \sum_{i=1}^{N} c_iX_i$, even if the distributions of the $X_i$ are not normal, the distribution of $Y$ may often be approximated by a normal distribution. This theorem states that the distribution of $Y$ will be approximately normal with expectation $E(Y) = \sum_{i=1}^{N} c_iE(X_i)$ and variance $\sigma^2(Y) = \sum_{i=1}^{N} c_i^2\sigma^2(X_i)$, where $E(X_i)$ is the expectation of $X_i$ and $\sigma^2(X_i)$ is the variance of $X_i$, if the $X_i$ are independent and $\sigma^2(Y)$ is much larger than any single component $c_i^2\sigma^2(X_i)$ from a non-normally distributed $X_i$ (BIPM, IEC et al. 2008).

The GUM defines Type A and B uncertainty evaluation. Type A uncertainty evaluation method is related to the evaluation of the uncertainty based on statistical series of observations. This means that a standard uncertainty is expressed as a standard deviation. For example, the measurement of the sensor noise in a laboratory by reading temporal samples can be defined as a Type A uncertainty. Type B uncertainty evaluation is used for all other means of uncertainty estimation that are not derived from statistics. Thus, source of type B uncertainty estimation might rely on a calibration certificate of an instrument, pre-flight test information, degradation assumptions... For example, the reflectance of calibration diffusers on-board satellite are typically characterised before launch. The uncertainty associated with this characterisation can be considered as a Type B uncertainty and it is usually provided as part of a calibration certificate and other pre-flight information.

The GUM stresses the difference between an error and an uncertainty and it will be one of the important points to remark in this work. These two terms refer to different concepts. An error is the difference between the measured value and the ‘true value’. An error has random and systematic components. Random errors cannot be eliminated but can be reduced by increased number of measurements, whereas systematic errors responds to an incomplete knowledge of the quantities measured. Whenever possible and known, the systematic error must be corrected for. Nonetheless, there will be always an uncertainty associated with the correction itself. Bias is an estimate of a systematic error. These two terms will be used through the document and, in many cases, the difference between them will lead to a slightly different interpretation.
To derive a measurement output uncertainty all individual inputs uncertainty components have to be established first and then combine according to the law of propagation of uncertainty. That is based on the Taylor series approximation given by:

\[ u_c^2(y) = \sum_{i=1}^{N} \left( \dfrac{\partial f}{\partial x_i} \right)^2 u_i^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{i+1}^{N} \dfrac{\partial f}{\partial x_i} \dfrac{\partial f}{\partial x_j} u(x_i, x_j) \]  

(1.3)

Where: \( y = f(x_1, x_2, \ldots, x_j) \) is the output value and is a function of the partial derivatives, \( \dfrac{\partial f}{\partial x_i} \) also called sensitivity coefficients and \( u(x_i) \) a standard uncertainty of an input component. The second part order of the Taylor approximation needs to be calculated if the input quantities are correlated, then the term \( u(x_i, x_j) \) can be replaced by \( u(x_i)u(x_j)r(x_i, x_j) \) where \( r(x_i, x_j) \) is the correlation coefficient.

The analytical method can become difficult to apply when the function becomes too complex and/or many of the input parameters present a noticeable correlation. In those cases, the Monte Carlo Methods (MCM) for uncertainty estimation are recognised and accepted and the GUM supplement has been published (BIPM, IEC et al. 2008). This is a numerical method that requires a well-defined probability density function (PDF) of all input components to propagate its effect through the measurement equation. The MCM then runs iteratively a large number of numerical calculations of the measurement equation randomly choosing the input from the available range that is defined by the probability density function. The large number output values calculated using a different inputs values at each iteration provides the uncertainty of the output value with its PDF.

The MCM can be used as an alternative uncertainty method and also provides the possibility to validate the uncertainty results obtained applying the GUM uncertainty framework (BIPM, IEC et al. 2008). This is one of the mechanisms used in Chapter 3 in order to provide confidence in the uncertainty results obtained. Moreover, there are other validation methods that have been used for specific uncertainty contributions. For example, the method to estimate the spatial uncertainty in a cross-calibration (see Section 5.3) has been implemented using both the LandSat-8 (L8) Optical Land Imager (OLI) and Sentinel-2 (S2) Multispectral Instrument (MSI) data. The verification of the method using two independent datasets provides further confidence in the uncertainty estimates.

The analytical GUM approach is well suited for its software implementation in terms of efficient memory and processing time consumption. However, for the evaluation of uncertainty contributors with a non-linear nature, it will be more appropriate to make use of an MCM.

At the time of writing, a revision of the GUM framework is under discussion. Since the original edition of the GUM, there have been major advances in terms of software as well as the extension of the uncertainty calculation to other scientific areas. The objective is to extend the GUM uncertainty
combination framework to other metrological problems preserving highly used parts of the original guide while providing a treatment at a similar level.

1.3.3 Distribution parameters

In addition to the traditional mean and standard deviation, there are specific parameters that describe the distribution shape. The ones introduced here are the quantile information, the skewness and the kurtosis (Zwillinger and Kokoska 1999):

- **Quantiles.** They divide the range of a probability distribution into several parts (or quanta) with equal probabilities. Specific cases are:
  - **Quartiles:** split the data in four parts being the second quartile equal to the median.
  - **Deciles:** split the data into 10 parts.
  - **Percentiles:** split the data into 100 parts.

- **Skewness.** It is typically used as a lack of symmetry in a distribution. A value of skewness close to zero indicates a symmetric distribution. When the parameter is positive, it tends to indicate the effect of a distribution tail on the right and vice versa when it is negative. It is defined as the third moment of the Pearson’s distribution:

\[
g_1 = \frac{m_3}{m_2^{3/2}} = \frac{E[(X - \mu)^3]}{E[(X - \mu)^2]^{3/2}}
\]  

(1.4)

Where \(\mu\) refer to the mean, \(\sigma\) to the standard deviation, and \(E\) to the expectation operator.

- **Kurtosis** is used as an indication of the flatness vs. sharpness of the distribution. The “excess kurtosis” is calculated here by taking the standard definition of the kurtosis for each TOA distribution and subtracting three, which is the kurtosis for a normal distribution (i.e. the excess kurtosis is equal to zero when the distribution is normal). For positive values, it indicates a sharpness of the distribution — w. r. t the normal distribution — with a well-defined peak.

\[
g_2 - 3 = \frac{m_4}{m_2^2} - 3 = \frac{E[(X - \mu)^4]}{E[(X - \mu)^2]^2} - 3
\]  

(1.5)

1.4 Motivation and aim of research

In the last decades there has been a rapid increase in the demand and offer of EO data. In parallel, the demand of some of the EO applications that use these data become more challenging as, for example, the climate monitoring. The attachment of a quality indicator to these data enables the users a better assessment of the adequacy of the data for the specific application (‘fitness for purpose’). This indicator is commonly provided as a measurement uncertainty (see Section 1.3.2).
Up to date, the majority of EO data providers do not include an uncertainty indicator in their datasets or the assessment of this is only partial (see Section 2.3). This limits not only the adequacy of the dataset to its immediate application but also limits any further uncertainty propagation through the EO processing chain.

Therefore, there is a clear need for a rigorous treatment of uncertainty from the sensor itself down to all the EO processing chain. The work aims to cover and end-to-end approach to estimate the radiometric uncertainty associated to the TOA radiance/reflectance factor pixel measurements of EO satellite optical instruments. This type of product is representative (although not unique) of the products delivered to the EO community. Consequently, a rigorous uncertainty associated to these dataset can be integrated by the EO users in further processing steps.

An end-to-end approach means that the project must deal with the radiometric study of the EO optical sensors, the theory for a rigorous uncertainty assessment and its design as a software tool. These three fields must be inter-related in a coherent manner throughout this project and the trade-offs among them must be highlighted. In order to do so, the following top-line strategic goals have been proposed:

1. An end-to-end methodology that links the radiometric model, uncertainty contributors and uncertainty combination closely following the GUM guidelines (BIPM, IEC et al. 2008).
2. The discussion of each one of the uncertainty contributors and the assessment of novel uncertainty contributors as the effect of the image orthorectification and the spectral response knowledge
3. The validation of the combined standard uncertainty model vs. an MCM method, the impact of simplifications and the correlation between uncertainty contributors.
4. The research for different software strategies to implement the TOA radiometric uncertainty at a pixel level and integrate as part of EO processing chain.
5. The investigation of the pixel covariance in the spatial, spectral and temporal domain for its application and further integration in subsequent levels of EO products’ processing.
6. The development of tools and methodologies for the calculation of the uncertainty associated to a TOA satellite-to-satellite cross-calibration as an alternative to the calibration devices used on-board for the absolute radiometric calibration.

This research is focused on the S2 MSI instrument with bands in the solar-reflective spectral range (400-2500 nm) (see Section 2.2). Starting from this specific case, other sensors and solutions will be discussed. However, it is anticipated that due to the large number of potential instruments, data processing and dissemination; not all cases can be covered. Nonetheless, this research should provide a generic framework that can be adapted to other specific scenarios even for those in other spectral ranges as the thermal infrared region (TIR) or microwave radiometry.
1.5 Research novelty

The main result of this work is an end-to-end approach to calculate the radiometric uncertainty at a pixel level for the TOA products (radiance or reflectance factor) by EO optical sensors with rigorous metrological techniques and exemplify where needed the trade-offs that have been applied for its calculation, combination and implementation as a software tool. The work is essential so that the users understand the “fit for purpose” of the data used for their application and it will be the basis of subsequent research and implementation to propagate the radiometric uncertainty further to e.g. bio-physical and/or climate products.

The research explores the use of the metadata information and quality information associated to the satellite products. As a result, the methodology provides the possibility to assess the uncertainty contributions for the specific satellite acquisition time, scene and processing. The implementation as a tool also provides the capacity to define the uncertainty to a specific coverage probability — i.e. defining the coverage factor $k$ (BIPM, IEC et al. 2008)— and the sensitivity study of each of the uncertainty contributions through the selection/deselection option.

This work also relies on the rigorous treatment of the uncertainty estimation compared to a more traditional over-estimation by the community. As quoted in Section 1.2 from Gardner (2004), the rigorous treatment of uncertainty is essential to better understand the optical sensors and has the potential to arise systematic effects previously overlooked. The uncertainty combination relies on a rigorous and extensive validation as compared to previous combination models used by the EO community. The assumptions in the combination model have been validated by comparing the results to a MCM method, the correlation among the different uncertainty contributions has been studied, and the impact of simplifications in the combination model has been assessed.

The work explores areas that are expected to be key important for subsequent research and propagation of the radiometric uncertainty to higher-level products. Thus, the uncertainty estimation tries to go deep into more complex and novel contributors that contribute in a TOA radiance/reflectance factor uncertainty budget as the spectral knowledge and orthorectification. Furthermore, it also describes the main correlation sources between pixels in the spatial, temporal and spectral dimension. The study has been implemented and resulted in uncertainty estimates for the mean of the pixels in a Region-Of-Interest (ROI). This is of direct application to the radiometric validation activities and the potential spatial binning of the level-1 (L1) products.

The S2 mission has been selected as a reference for the RUT development. The nominal on-board radiometric calibration is based on a sun diffuser. Alternative methods include the use of instrumented sites and the sensor-to-sensor cross-calibration. The work here is also focused on providing rigorous metrology and software implementation for the latter case.
The research here presented can be used by a variety of EO experts and users. The direct use of the research is the RUT tool, which can be used by any S2 user to estimate the L1 product uncertainty. S2 L1 calibration engineers can also apply the results for any S2 radiometric validation activity. However, it is the methodology here presented that can be replicated by different groups working on different EO optical missions in order to provide L1 uncertainty estimates. That is, scientists and product developers of other optical missions are expected to benefit from the methodology here presented. In addition, the developers of higher-level products are also expected to benefit not only from a methodology but also from the access to L1 uncertainty images. These can be used as an input to propagate the uncertainty to further EO processing levels.

1.6 Structure of the thesis and publications

Before the enrolment in this Ph. D. programme, an initial research was developed while working at the European Space Agency (ESA). This research resulted in a preliminary version and analysis of the RUT that was presented in Gorroño and Gascon (2013).

During the development of this Ph. D. a series of articles, conferences and workshops have been developed and presented. The majority of the work presented in Chapter 3 to Chapter 6 is based on a number of publications:

- Chapter 3 is based on Gorroño, Fomferra et al. (2017). This chapter can be considered as the core of the thesis. Here the design, development and implementation of the RUT is explained in detail. The code of the RUT tool and the development can be found in a public repository (Gorroño, Fomferra et al. 2016).
- Chapter 4 is based on Gorroño, Banks et al. (2016). This chapter is an extension to the uncertainty contributions analysis in the RUT. The study seeks preliminary methods and implements the uncertainty associated spectral knowledge and orthorectication uncertainty propagation.
- Chapter 5 is based on Gorroño, Banks et al. (2017). This chapter describes the methodologies developed to account for the uncertainty in a TOA sensor-to-sensor cross-calibration in the spectral, spatial and temporal dimensions.
- Chapter 6 is based on Gorroño, Hunt et al. (2017). This chapter can be considered as an extension of the RUT development focused on the capacity to provide uncertainty estimates for the mean of pixels in a ROI. It describes the error correlation between S2 pixels, an implementation based on the RUT and its validation.

The publication Gorroño, Gascon et al. (2015) is a conference paper that describes the preliminary methodology of the RUT tool. It describes the L1 radiometric model, explores the associated uncertainty
contributors and describes the software guidelines for the RUT implementation. Furthermore, in this paper the Allan deviation technique has been implemented for the first time in order to validate the noise of an optical sensor in-flight. This is used to support some of the uncertainty contributions in Chapter 3.

Early-stage of the work related to TOA sensor-to-sensor cross-calibration was presented at conferences and can be found in Underwood, Fox et al. (2015) and Gorroño, Fox et al. (2015). The latter initially pointed out the asymmetry of the ROI reflectance pixel distribution in a PICS site. This is a topic which has been further developed and the work has been published in Gorroño, Bialek et al. (2016). Although this is not of direct applicability to this work, the concepts there expressed have been applied as part of the distribution analysis in Chapter 5.
Chapter 2.

Literature review

2.1 Introduction

This review starts by briefly describing the Sentinel-2 (S2) mission in Section 2.2. This optical mission is the reference one used in this work. It will be taken as a reference to develop the Radiometric Uncertainty Tool (RUT) (see Chapter 3) as well as an example of the sensor-to-sensor cross-calibration uncertainty (see Chapter 5). In this case, the Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (TRUTHS) mission will be also used as the reference sensor together with the S2 mission. TRUTHS is a proposed mission and modifications of the system are likely to occur during the design and development phases. Thus, rather than describing the mission in detail here, there is a brief introduction and description in Chapter 5.

Although specific missions were used to implement the concepts developed here, these are largely applicable to other EO missions. For example, efforts are on-going to implement the RUT for the Sentinel-3 (S3) Ocean Land Colour Instrument (OLCI) (Hunt and Nieke 2016). The work described in the cross-calibration is readily adapted to similar instruments. For example, the Climate Absolute Radiance and Refractivity Observatory (CLARREO) team has been using the Visible Infrared Imaging Radiometer Suite (VIIRS) or Moderate-Resolution Imaging Spectroradiometer (MODIS) as examples in the cross-calibration uncertainty assessment (Wu, Xiong et al. 2015).

Section 2.3 reviews the literature concerning the uncertainty in EO optical domain. The current available literature is very small as this field has been only explored in the last few years. Nonetheless, effort has been made to bring the best possible examples and connections between the work here presented and the rest of the literature with an emphasis on identifying the research gaps.

2.2 Sentinel-2 Mission, MSI instrument and Level-1 products

The S2 mission is currently operated by the European Space Agency (ESA), in the framework of the European Union Copernicus programme. Its mission offers an unprecedented combination of systematic global coverage of land and coastal areas, a high revisit of five days under the same viewing conditions, high spatial resolution (10 to 60 m), and a wide field of view (FOV) (295 km) for multispectral observations from 13 bands in the visible, near infrared and short wave infrared range of
the electromagnetic spectrum (Drusch, Del Bello et al. 2012). The S2 mission provides enhanced continuity to services monitoring global terrestrial surfaces and coastal waters. It is used for many specific applications, including urban planning, natural and man-made disasters management and crop monitoring. The research described in Lefebvre, Sannier et al. (2016) and Immitzer, Vuolo et al. (2016) shows a couple of the early applications of S2 data related to urban and crop monitoring.

The S2 mission is composed of two identical satellites — called Sentinel-2A (S2A) and Sentinel-2B (S2B) units¹ — and each one carrying a single imaging payload named MSI (Multi-Spectral Instrument). The orbit is Sun-synchronous at 786 km altitude (14 + 3/10 revolutions per day) with a 10:30 A.M. descending node. The S2 satellites systematically acquire observations over land and coastal areas from −56° to 84° latitude including islands larger than 100 km², European Union (EU) islands, all other islands less than 20 km from the coastline, the whole Mediterranean Sea, all inland water bodies and all closed seas. Over specific calibration sites, for example DOME-C in Antarctica, additional observations will be made. The two satellite units will work on opposite sides of the orbit. S2A launch took place in June 2015 and S2B launch took place in March 2017 (Gascon, Bouzinac et al. 2017).

Both S2A and S2B incorporate as its main payload the MultiSpectral Instrument (MSI). The instrument is based on a push-broom concept, featuring a Three-Mirror Anastigmatic (TMA) telescope feeding two focal planes — one for Visible and Near-InfraRed (VNIR) bands and another for Short-Wave InfraRed (SWIR) bands — separated by a dichroic filter. A schematic is shown in Figure 2-1:

Figure 2-1 Multi-Spectral Instrument (MSI) internal configuration with (left) full instrument view (diffuser panel in yellow, telescope mirrors in dark blue) and (right) optical path construction for the splitter and SWIR/VNIR focal planes (reproduced with permission from (Gascon, Bouzinac et al. 2017)).

¹ Most of the work was implemented for S2A satellite. Since the MSI payload is equivalent for both satellites, the results can be easily adapted from S2A to S2B. Thus, the document will refer to S2 or S2A indistinctively.
The MSI includes 12 detector modules on each focal plane that are stagger-mounted to cover altogether the 20.6° instrument FOV — swath width of 295 km on the ground across track at an altitude of 786 km. The push-broom configuration implies that each detector in the focal plane has specific characteristics such as gain, dark signal or noise. The VNIR detectors are based on monolithic Silicon complementary metal–oxide–semiconductor (CMOS) technology with 10 spectral bands integrated on a single detector (Martin-Gonthier, Magnan et al. 2010). The SWIR detectors are based on hybrid CMOS and Mercury Cadmium Telluride (MCT) technology where the latter is hybridized to a silicon Readout Integrated Circuit (ROIC). The SWIR assembly has 3 spectral bands combined on a single detector respectively and includes Time Delay Integration (TDI)(Dariel, Chorier et al. 2009).

The 13 spectral bands included in the S2 MSI are split in 4 bands at 10 m spatial resolution (blue, green, red, and near-infrared), 4 narrow bands at 20 m spatial resolution mainly used for vegetation characterization and in the red edge, 2 wider SWIR bands at 20 m spatial resolution for applications such as snow/ice/cloud detection or vegetation moisture stress, and 3 bands at 60 m spatial resolution for applications such as cloud screening and atmospheric corrections (Gascon, Bouzinac et al. 2017). The names, central wavelength and bandwidth of the S2 MSI spectral bands are provided in Table 2-1.

<table>
<thead>
<tr>
<th></th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
<th>B7</th>
<th>B8</th>
<th>B8a</th>
<th>B9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central wavelength (nm)</td>
<td>443</td>
<td>490</td>
<td>560</td>
<td>665</td>
<td>705</td>
<td>740</td>
<td>783</td>
<td>842</td>
<td>865</td>
<td>945</td>
</tr>
<tr>
<td>Bandwidth (nm)</td>
<td>20</td>
<td>65</td>
<td>35</td>
<td>30</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>115</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>B10</td>
<td>B11</td>
<td>B12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central wavelength (nm)</td>
<td>1375</td>
<td>1610</td>
<td>2190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandwidth (nm)</td>
<td>30</td>
<td>90</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The MSI instrument uses an on-board sun diffuser, which is deployed nominally every month, to update the absolute radiometric calibration and the relative gains calibration. There is no secondary diffuser on-board for checking the degradation of the calibration diffuser and, thus, it relies on a pre-determined optimised exposure based on heritage knowledge and monitoring through vicarious and cross-mission validation using terrestrial sites. Every two weeks, images acquired over ocean at night are used to update the dark signal calibration (Drusch, Del Bello et al. 2012).

The description of the S2 Level 1 (L1) processing and mathematical model can be found in Section 3.2.
2.3 Uncertainty analysis in the EO optical missions

2.3.1 TOA Radiometric uncertainty of EO optical sensors

The uncertainty analysis of EO optical sensors carried on satellite missions is traditionally circumscribed to the radiometric calibration performed by on-board systems as diffusers, irradiance lamps, blackbody reference… For example, the work described in Knight and Kvaran (2014), Morfitt, Barsi et al. (2015) specifies the Landsat-8 (L8) Operational Land Imager (OLI) pre-flight and post-launch radiometric performance in detail whereas the work in Gascon, Bouzinac et al. (2017) describes the different radiometric validation activities currently undergoing for the verification of the S2 MSI radiometric performance. In both cases, the work is described with a high level of detail and the radiometric techniques presented are very rigorous. However, the evaluation of uncertainty is not described in detail and, more importantly, the results are unconnected one from each other.

One of the most complete works for the top-of-atmosphere (TOA) radiometric uncertainty of an Earth Observation (EO) optical sensor can be found in Esposito, Xiong et al. (2004). The research method effectively links the L1 processing model with an uncertainty model combination. The study discusses and calculates each of the contributions in the L1 reflectance factors and considers the weight of each of the sources in the global budget.

The implementation as part of the MODIS L1B product is described in Toller, Isaacman et al. (2006). The “Uncertainty Index” is codified in an 8-bit unsigned integers. The four least significant bits of these integers represent the best estimate of the uncertainty in each measurement of reflectance for the reflective solar bands and radiance for the thermal emissive bands. The four most significant bits of the index are reserved for other performance metrics. The indices range of values is from 0 through 15, representing an interval of uncertainty values in which the measured uncertainty lies. A value of 15 represents uncertainties of data that cannot be calibrated, or higher values of uncertainty than possible in the measured range. A logarithmic codification of the uncertainty output has been implemented. Although the method is positive since it maximises the range of uncertainty values with a minimum number of bits, it also forces the users to post-process the uncertainty results to convert back to a linear scale.

It is common practice that the National Measurement Institutes (NMIs) provide a robust and detailed uncertainty estimates of the radiance and irradiance standards provided. For example, the work in Gardner (2004) describes in detail the uncertainty propagation through the National Institute of Standards and Technology (NIST) irradiance scale. It links the different stages in the calibration transfer starting from the primary cryogenic radiometer standard. In the process, a detailed analysis of the correlation, interpolation and ultimately uncertainty propagation is provided. Thus, it is the core goal of
this work to apply this same robustness and detail to the operational products — TOA radiance/reflectance factor measurements — that the Space Agencies disseminate to the users.

In order to achieve similar levels of detail and robustness in disseminated L1 products to the disseminated radiance/irradiance scales, several points and weaknesses will be addressed in this work:

- Previous work tends to present uncertainty analysis as a set of independent studies with no connections among them. The uncertainty analysis must link all the L1 processing and originate from a mathematical model.
- The uncertainty assessment methodology is generally that of a “conservative estimates”. Thus, the approach will try to avoid this situation and seek specific values where possible.
- Guide to the expression of Uncertainty in Measurement (GUM) guidelines have not been used (BIPM, IEC et al. 2008). No discussion of error vs. uncertainty or description of the distribution is introduced.
- No impact of several uncertainty contributors that depend on the measured scene as: the spectral response uncertainty, the orthorectification, the stray-light or the polarisation…

For the implementation of the uncertainty estimates and its use as part of a more complex EO processing chain, the points to be addressed are:

- how to use the uncertainty estimate for merged pixels or higher-level product propagation.
- how to integrate an operational tool as part of the users’ EO processing.

2.3.2 The uncertainty propagation through interpolated data

The reference work that describes the radiometric uncertainty propagation through interpolated data can be found in Gardner (2003). In this publication, Gardner describes how the effect of the radiometric interpolation type, grid used and input variations affects the resulting level of uncertainty. The work effectively describes the effects of a Lagrange and cubic-spline types of interpolation in the propagation of radiometric uncertainty. One of the examples shows the result of propagating the photometric response function \( V_\lambda \) measured at 20 nm sampling with a 1 % uncertainty \((k = 1)\) and interpolated using cubic splines at 2 nm. The results show how the uncertainty fluctuates around the level of 1 %. This fluctuation originates as a consequence of the increase of correlation due to the interpolation and is influenced by the position of the interpolation. What is more, the results also point how the use of a cubic-spline brings unstable situations when the second derivative estimation fails.

No application of this concept has been found to the orthorectification of TOA EO images — where a radiometric interpolation is commonly applied — and in general no assessment is discussed for the rest of the L1 processing. The closest example of this concept can be found for the in-flight noise evaluation
of L8 OLI in Morfitt, Barsi et al. (2015). The noise model includes a factor of 0.8 as a consequence of the cubic convolution resampling. The origin of this factor is not explained in detail and does not account for effects as correlation between the pixels in the resampling window or the position in the resampling grid.

### 2.3.3 Pixel correlation in optical radiometers

As previously discussed in Section 2.3, the effects of correlation were applied to the NIST visible spectral standards (Gardner 2004). The result for spline-interpolated irradiance values in the VIS spectral range showed correlation coefficients in the range 0.92 to 1.0. Due to the high level of correlation, both spectral irradiance values and their uncertainties provided by NIST at visible wavelengths can be simply interpolated. The near-full correlation of irradiance values is valid when the long-term drift dominates the global uncertainty budget.

The Kepler space telescope main scientific goal is to detect Earth-like planets around Sun-like stars using transit photometry. It carries on-board a 96.4-megapixel focal plane composed of 42 Charge-Coupled Devices (CCD) and 1024 × 1100 pixel arrays. As part of the mission, it also estimates uncertainties for the calibrated pixels. The implementation of this uncertainty includes the modelled read noise, the calculated shot noise, and the effective quantization noise (Clarke, Allen et al. 2010).

Despite the simplicity of the uncertainty modelling at a pixel-level, the work in Clarke, Allen et al. (2010) is mainly devoted to study and implement the cross-correlation between calibrated pixels. The concatenation of the covariance matrices along the processing produces a huge consumption of memory and storage requirements. The solution proposed is the implementation of compression techniques that reduce the memory requirements down to acceptable levels. The two largest contributors to calibration-induced correlations for the Kepler calibrated pixels are the smear level correction and the dynamic 1-dimension black correction. The resulting correlation levels can go up to several percent of the median variance but generally staying below the 1% level.

Both examples — NIST irradiance scale and Kepler mission — show the importance of studying the measurement correlation and the different levels of complexity that can be found. No examples could be found where the EO L1 processing includes an analysis of the uncertainty correlation.

### 2.3.4 Uncertainty of the biophysical and climate products

Section 2.3 briefly described the examples of TOA radiometric uncertainty. The literature provides just the example of MODIS L1B products as currently producing uncertainty estimates and delivering them embedded in the product. Similarly, biophysical and higher-level products rarely attach an uncertainty estimate with the information provided.
The best example of uncertainty estimate as part of a high-level product can be found as part of the ESA Climate Change Initiative (CCI) and particularly in the Sea Surface Temperature (SST) community. Current efforts are under development to produce and improve uncertainty estimates attached to the delivered SST products. The work in Merchant, Embury et al. (2014) describes the new SST datasets generated from satellite observations for the period 1991-2010 and intended for use in climate science applications. These datasets are generated from both observations of Along-Track Scanning Radiometers (ATSRs) and Advanced Very High Resolution Radiometers (AVHRRs) and include an attached uncertainty estimate.

The uncertainty attached to the SST datasets reflects the spatial distribution, background field used and uncertainty of the observations. Both the uncertainty associated to the AVHRR and ATSR data is included. For example the uncertainty associated to the derived skin SST comprises an estimate of the radiometric noise at pixel level propagated to the cell-mean, uncertainty in the forward modelling of the brightness temperature and global systematic uncertainty component. The 20 cm SST includes a further uncertainty contribution associated to the skin to depth adjustment.

The recent publication in Merchant, Paul et al. (2017) describes the state-of-the-art related to the uncertainty of climate data records including the SST. This review concludes with a set of recommendations that include the use of metrological concepts as “standard uncertainty”, the consideration of error correlation during the uncertainty propagation and the integration of quantitative uncertainty information as part of the datasets. Furthermore, it also warns about the lack of uncertainty information as part of the L1 products and the need of providing some more information than just “instrument noise”. Indeed, errors that may contribute in a negligible manner to the total uncertainty at L1 products might be dominant in climate data records.

### 2.3.5 Uncertainty in a TOA sensor-to-sensor cross-calibration

There are several techniques and devices currently in use for absolute radiometric calibration of EO optical missions. The devices used both for in-flight and pre-flight assessment are traditionally diffusers, lamps or blackbodies. In Section 2.3.1, a couple of examples concerning the S2 MSI and L8 OLI diffuser calibration uncertainty where mentioned. The uncertainty related to the diffuser calibration will be explained in Chapter 3 as part of the development of the RUT. However, there are other techniques that are also used for the absolute radiometric calibration of EO optical sensors and include the site vicarious calibration and the TOA sensor-to-sensor cross-calibration. The uncertainty associated to this process has not been extensively studied as compared to the on-board devices. Here the focus is on describing the state-of-the-art in the uncertainty analysis in a TOA sensor to sensor cross calibration.

A research in the literature has found that there is a global study of the uncertainty in a cross-calibration in Chander, Helder et al. (2013) whereas the CLARREO team mission has been actively studying the
uncertainty in different studies: (Wielicki, Doelling et al. 2008) as a global uncertainty study, (Roithmayr, Lukashin et al. 2014, Roithmayr, Lukashin et al. 2014) for satellite-to-satellite matching opportunities, (Wu, Xiong et al. 2015) for the spectral dimension uncertainty or (Lukashin, Wielicki et al. 2013) for the polarisation impact.

In each subsection below there is a short discussion of the current status of the uncertainty assessment associated with the spectral, spatial and temporal dimensions. These are the three dimensions studied in Chapter 5.

**Spectral dimension**

In the spectral dimension the work carried out by Chander, Helder et al. (2013) tested the impact of the spectral resolution of the reference sensor to calculate the Spectral Bandwidth Adjustment Factor (SBAF) in a cross-calibration between two sensors. The impact was assessed to fall well below an uncertainty level of 0.3% ($k = 1$). When a spectral shift in SBAF is applied, the use of filters such as those used in the Moderate Resolution Imaging Spectrometers (MODIS) (often used as a reference sensor) have suggested that worst-case tolerances/shifts of 5 nm in the bands would produce larger differences. The results for this spectral shift increase up to the 2%. However, this must be considered as a worst case assessment rather than a realistic scenario due to the tolerances used.

The work described in Wu, Xiong et al. (2015) studied the impact of the spectral dimension for the CLARREO mission. The work used a reference MODTRAN simulated TOA spectral radiance and SCHIAMACHY (from SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY) hyperspectral measurements to study the potential effect of the spectral sampling and resolution. The comparison was made between the CLARREO mission and MODIS and VIIRS spectral bands. The results provided an accuracy of within 0.16% for all surfaces and 0.1% for global average.

From the literature, it seems that the analysis of the spectral dimension has been recently developed and robust estimates have been assessed. Nonetheless, certain points need to be further developed:

- The design of CLARREO in Wu, Xiong et al. (2015) was not based on a real design. Design values and considerations must be introduced in the study.
- No spectral binning has been considered for a spectrometer design.
- The literature shows analysis based on a spectral sensitivity. Although this analysis is useful, it is insufficient and it is convenient to study the impact of the spectral knowledge. That is, how well the spectral response of the reference sensor is known.
**Spatial dimension**

The work in Wielicki, Doelling et al. (2008) describes a method that can predict the spatial matching noise at a global scale. The study used 3 months of National Oceanic and Atmospheric Administration (NOAA) 17 and NOAA 18 AVHRR visible channel data (0.65 μm) with a tight temporal and angular constrains — 1° in Viewing Zenith Angle (VZA) and Viewing Azimuth Angle (VAA) and 6 minutes overpass difference. Since both missions present almost equal spectral responses and calibration methodology, the differences are isolated to spatial errors.

The description of the method in Wielicki, Doelling et al. (2008) provides an effective assessment to study the spatial matching noise at a global scale. The method relies on the assumption that the temporal and angular effects have been sufficiently minimised. However, the limitation of the method is based on the fact that it cannot be applied to specific sites and specific mission requirements. That is, there is the need to investigate a method that provides spatial uncertainty for specific sites and that can transform the geolocation uncertainty into a TOA radiometric uncertainty.

**Temporal dimension**

Recent work in McCorkel, Thome et al. (2013) studied the effect of temporal mismatch between MODIS vs. Hyperion matches. The latter instrument was measuring in an orbit 40 minutes preceding the MODIS one until mid-2005. The orbit of Hyperion was changed from mid-2005 resulting in a rare cross-calibration between the two missions. This unusual situation triggered the possibility to compare the impact of the temporal overpass differences between coincident overpasses — within 30 to 40 minutes — and non-coincident overpass — within 30 days separation — over the Railroad Valley calibration site. The results showed that although the dispersion of the data significantly increased, the bias between the two cases was between 1-2%. To a large extent, BRDF and temporal mismatches were largely averaged out even for such a large timespan difference.

In the work presented in Roithmayr, Lukashin et al. (2014), the selection of cross-calibration matchups was set to a global scale within a 5 minutes of delay between overpass. At that time delay, the temporal noise was found to be at the 1 % level and with sufficient samples the noise reduces to <0.3% under the assumption of largely uncorrelated errors (Wielicki, Doelling et al. 2008). The results were obtained by comparing 3 months of NOAA 17 and NOAA 18 AVHRR visible channel data (0.65 μm) with orbital matches varying from 1.5 minutes to 12 minutes.

The work described in the literature describes methods for global assessment of the temporal uncertainty. Similarly to the spatial dimensions described before, there is no specific method to assess the temporal uncertainty for a specific cross-calibration at a specific site. Thus, the objective is the
estimation using metrological robust techniques of the temporal uncertainty introduced in a TOA sensor-to-sensor cross-calibration.

2.4 Summary of the literature review

The ‘state of the art’ related to the radiometric uncertainty assessment and implementation has been provided in the previous sections.

In general terms, it has been made clear that the current analysis of EO radiometric uncertainty is very limited and lacks a rigorous treatment. Most of the uncertainty estimation is based on “conservative” estimations and the model combination does not follow either the GUM guidelines or an extensive validation. Several contributions are omitted and/or not studied even when they represent an important contribution to the TOA uncertainty estimates.

At the implementation level, a precedent has been found in studies of the MODIS instrument which provide uncertainty images associated to its products. Nonetheless, the research here is focused on providing an external tool. This means that many constrains can be avoided since the uncertainty estimations will be updated to any contingency or processing update. It will be possible to describe more specific uncertainty contributors, select the confidence interval, and integrate in other routines with special emphasis in its use to the uncertainty estimates for merged pixels or higher-level product propagation.

As a result of this overview it is clear that there is a need of the EO and climate community so that the L1 products include a well-documented and reliable uncertainty. This will certainly benefit the propagation of the uncertainty and further analysis to deliver uncertainty estimates in higher-level EO products.

Finally, the literature review has been focused on the work developed in the cross-calibration uncertainty techniques. There is an important work in this field developed by the National Aeronautics and Space Administration (NASA) CLARREO team. Thus, the approach here is not in repeating previous work but in providing a complementary approach to the work already developed. Consequently, the efforts will focus on issues as: the assessment of spectral knowledge and spectral binning, the connection between the geolocation uncertainty requirements and the radiometric impact or the specific impact of the temporal uncertainty and the residual correction over a calibration site.
Chapter 3.

Methodologies for a Radiometric Uncertainty tool

3.1 Introduction

Earth Observation (EO) via remote sensing nowadays provides one of the main sources of information about the Earth system. The complexity of many of these applications—land monitoring or climate studies—and the growing number of data sets, makes it increasingly necessary that this data has associated with it a quality indicator that describes the compatibility between different sensor data and the suitability for particular applications. Indeed the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) through its Quality Assurance Framework for Earth Observation (QA4EO) explicitly states that data and derived products shall have associated with them an indicator of quality to enable users to assess the “fitness for purpose” of the data to their specific application (QA4EO 2010). Recent work described in Merchant, Embury et al. (2014) has resulted in the delivery of Sea surface temperature (SST) datasets with associated uncertainty estimates at a pixel level. The uncertainty estimates included in the SST products vary from 0.1 to 1.5 K, and they are largely dependent on the spatial and temporal distance with the satellite observations. This type of information is of a high-value to scientists working in climate models to understand the adequacy of the SST products for different observational conditions and is one example where the use of a single scene value for uncertainty could lead to misleading results. Similarly, where a user is interested in a particular localised area within a scene, unless pixel level uncertainty is available, they may not appropriately take account of potential adjacency effectives due to land cover types, coastal regions etc. and/or clouds etc.

The uncertainty provided in the EO products—Level-1 (L1) and higher—is the result of a chain that effectively links all the processing steps from the instrument down to the final product and that this chain should allow full traceability to the primary calibration through a measurement equation and account for any scene-dependent sensitivities. Ideally, TOA radiometric uncertainty provided with L1 EO products could be propagated through consecutive steps of the EO processing chain to higher level products, such as those describing biophysical parameters; through this process, uncertainties associated with L1 input parameters would effectively be one of the principal inputs to the higher-level products’
uncertainty estimates. Furthermore, a rigorous uncertainty analysis of the L1 products may lead to a better understanding of the instrument and its radiometric performance by detecting systematic effects which may have been overlooked in previous analyses (Gardner 2004). This assessment needs to consider the design of the sensor and how it makes and processes its measurements and establish a clear model which can allow any aspects that have scene-specific sensitivities. For example, detector noise, linearity, spectral characteristics, etc. can be evaluated with independent inputs and an appropriate uncertainty determined.

An example of the TOA radiometric uncertainty analysis can be found in Esposito, Xiong et al. (2004). The work proposes an uncertainty model combination and evaluates each one of the uncertainty contributions associated with the L1B algorithm of the reflectance solar bands of the Moderate Resolution Imaging Spectroradiometer (MODIS). This uncertainty resulted in an implementation of uncertainty images as part of MODIS L1B product (Xiong, Troller et al., Toller, Isaacman et al. 2006).

The aim here is to describe the efforts in designing and developing a tool and, more fundamentally, a detailed uncertainty analysis for the Sentinel-2 (S2) L1 mission products. The design and development of the tool has resulted in a first version of the tool implemented and released — S2-RUT (Sentinel-2 Radiometric Uncertainty Tool) (Gorroño, Fomferra et al. 2016) — that estimates the TOA radiometric uncertainties associated with each pixel using the S2 TOA reflectance factor images provided by the European Space Agency (ESA) as the input. This chapter describes a method for the detailed analysis following the ‘Guide to Expression of Uncertainty in Measurement’ (GUM) (BIPM, IEC et al. 2008) that can be easily adapted to other missions in the optical range. The uncertainty estimates provided by the tool can provide quality information to the users of L1 data — e.g., for instrument radiometric validation — and can be used as the input to propagate the uncertainty to higher-level products in a similar manner to that of the SST example described in (Merchant, Embury et al. 2014).

The S2 mission has been recently launched by the ESA, in the framework of the European Union Copernicus programme, and is used for many applications, including urban planning, natural and man-made disasters management and crop monitoring. The research described in Lefebvre, Sannier et al. (2016) and Immitzer, Vuolo et al. (2016) shows a couple of the early applications of S2 data related to urban and crop monitoring, respectively. S2 incorporates as its main payload the MultiSpectral Instrument (MSI) which consists of 13 Visible and Near-InfraRed (VNIR) and Short-Wave InfraRed (SWIR) bands with spatial resolutions of 10 m, 20 m and 60 m as well as a short revisit time (5 days at the equator with two satellites) and a wide field of view (290 km) (Drusch, Del Bello et al. 2012).

The MSI instrument uses an on-board sun diffuser, which is used nominally every month, to update the absolute radiometric calibration and the relative gains calibration. There is no secondary diffuser on-board for checking the degradation of the calibration diffuser and, thus, it relies on a pre-determined
optimised exposure based on heritage knowledge and monitoring through vicarious and cross-mission validation using terrestrial sites. Every two weeks, images acquired over ocean at night are used to update the dark signal calibration (Drusch, Del Bello et al. 2012).

The structure of the chapter follows the same logic applied in the design and development process of the S2-RUT tool. The first part of the design takes into account the S2 radiometric model, described in Section 3.2, and identifies each of the uncertainty contributors, described in Section 3.3, and links them to the parameters in the L1 processing model. Effort has been put into an exhaustive specification and assessment of each contributor by reviewing the results of the pre- and post-launch characterisation. The identified uncertainty contributors are combined following the guidelines in the GUM (BIPM, IEC et al. 2008). A description of the GUM can be found in Section 1.3.2. Particular attention has been given to the validation and exhaustive discussion of the uncertainty combination method in Section 3.4. The software design of the tool has emphasised the use of computationally efficient strategies to read the TOA reflectance factor images and is discussed in Section 3.5. Effort has been made in the automatisation of the uncertainty evaluation to include the option to extract the auxiliary information from the metadata in the satellite products. The first version of the tool also includes the assessment of the uncertainty at a desired coverage probability by setting a coverage factor, \( k \), and the selection/deselection of the uncertainty contributors for sensitivity studies. This initial version of the tool has been implemented — code available at Gorroño, Fomferra et al. (2016) — and integrated as part of the Sentinel Application Platform (SNAP) (ESA). The first version of the tool is referred here as S2-RUTv1. This first version focuses on describing in detail an exhaustive uncertainty methodology and a general software design that can form the basis of adoption to several other EO missions. Indeed, the methodology described in Figure 3-1 can be easily adapted to other missions with optical payloads on-board and, with some additional variations, to different type of instrumentation like the Synthetic Aperture Radar (SAR) instruments. The tool is open to the community and is under continuous evolution, with the expectation that future versions will include a refined uncertainty assessment and/or new tool features.

### 3.2 Sentinel 2 Level-1 Radiometric Model

The first step in any radiometric uncertainty analysis is, where possible, to identify the mathematical equation that describes the data processing at each step of the chain. The description of the Sentinel-2 mission and L1 processing can be found in Gascon, Cadau et al. (2014), (ESA 2017). Figure 3-1 illustrates the MSI instrument acquisition and L1 processing. In the graph, the parameters \( p, l, b \) and \( d \) correspond to the selected pixel, line, band and detector respectively.
Figure 3-1 Illustration of the MultiSpectral Instrument (MSI) radiometric model and Level 1 (L1) ground processing with the associated parameters described in Equations (3.1)-(3.8)

The MSI instrument collects the incoming radiance ($L_{TOA}$) through a three-mirror off-axis anastigmatic telescope. The incoming light is separated at the splitter into visible and near-infrared (VNIR) and short-wave infrared (SWIR). This light is filtered and detected at the instrument focal plane. At the output of the front-end electronics (FEE), the raw signal, $X$, is pre-equalised and compressed — boxes “Eq.” and “Comp.” in Figure 3-1 — by the video chain unit (VCU) and the signal $Z_{VCU}$ is sent down to Earth. Once the signal is received, a decompression ($VCU^{-1}$) is executed. This process is assumed deterministic since the signal is decompressed with the same parameters used for the on-board compression but in an inverted process. For simplification, the decompression is not included in the mathematical model.

The radiometric standard correction model for the Sentinel-2 L1B (geolocated pixels with counts proportional to Earth radiances) product is presented in (ESA 2017):

$$
Y(p,l,b,d) = X(p,l,b,d) - DS(p,l \mod 6,b,d) - PC_{masked}(l,b,d)
$$

$$
Z(p,l,b,d) = \gamma(p,b,d,Y(p,l,b,d))
$$

(3.1)

where $p$ refers to the pixel number in the selected focal plane detector, $l$ determines the specific chronogram sub-cycle, $b$ determines the selected spectral channel and $d$ identifies the selected detector in the focal plane. $Y(p,l,b,d)$ is the raw signal $X(p,l,b,d)$ of pixel $p$ corrected to allow for the dark signal $DS(p,l \mod 6,b,d)$ and the pixel contextual offset, $PC_{masked}(l,b,d)$, expressed in Least Significant Bit (LSB); $\gamma(p,b,d,Y(p,l,b,d))$ is a function that compensates for the non-linearity of the global response of the pixel $p$ and its relative behaviour with respect to other pixels; $Z(p,l,b,d)$, is the equalised signal after
relative and non-linear correction (in LSB); and \( DS(p,j,b,d) \) is the dark signal of the pixel \( p \) in channel \( b \) for chronogram sub-cycle line number \( j \) (\( j \) is in the range 1 to 6).

The initial quantisation produced at the instrument in order to digitise the analogue output of the instrument is as follows:

\[
X(p,l,b,d,L(\lambda)) = \text{trunc}[G(p,b,d,L(\lambda)) \cdot V(p,l,b,d,L(\lambda))] \tag{3.2}
\]

where \( G(p,b,d,L(\lambda)) \) represents the gain of the pixel \( p \) in channel \( b \) and detector \( d \) at the video chain and \( V(p,l,b,d,L(\lambda)) \) is the voltage recorded for the input radiance, \( L(\lambda) \).

The pixel contextual offset, \( PC_{\text{masked}} \), parameter aims to compensate for the dark signal variation due to voltage fluctuations in-orbit. It is described by:

\[
PC_{\text{masked}}(l,b,d) = \frac{1}{N_{\text{masked}}} \cdot \sum_{i=1}^{N_{\text{masked}}} (X(i,l,b,d) - DS(p,l \mod 6,b,d)) \tag{3.3}
\]

where \( DS(p,j,b,d) \) is the dark signal of the pixel \( p \) in channel \( b \), for chronogram sub-cycle line number \( j \), \( N_{\text{masked}}(b) \) is the number of masked pixels on both sides of the detector line for channel \( b \) and mod is the “modulo” function.

The relative gains, \( \gamma(p,b,d,Y(p,l,b,d)) \), correct for the non-uniformity and non-linearity of the pixels and codify them in an unsigned two-byte integer. The formulation is presented below for the VNIR bands (left) and the SWIR bands (right):

\[
Z = \text{round} \left[ g_0 + g_1 \cdot Y + g_2 \cdot Y^2 + g_3 \cdot Y^3 \right] \text{ if } Y \geq 0
\]
\[
Z = 0 \leftrightarrow Y < 0
\]

\[
Z = \text{round} \left[ a_1 \cdot (Y - Z_c) \right] \text{ if } Z_c < Y \leq Z_c + Z_s
\]
\[
Z = \text{round} \left[ a_2 \cdot (Y - Z_s - Z_c) + A_1 \cdot Z_s \right] \text{ if } Y > Z_c + Z_s \tag{3.4}
\]

The coefficients \( g0, g1, g2 \) and \( g3 \) correspond to a cubic polynomial whereas the coefficients \( a1, a2, Zs \) and \( Zc \) correspond to the coefficients of a double linear fitting equation. These values have been characterised pre-flight and are monitored in-flight by the diffuser acquisitions on a monthly basis. If the variation is sufficiently large, the curves are re-scaled as follows:

\[
\gamma_{\text{update}}(p,b,d,Y) = \gamma_{\text{ground}}^{-1}(p,b,d,\frac{Y_{\text{update}}(p,b,d)}{Y_{\text{update}}}) \cdot Y \tag{3.5}
\]

In order to obtain the LIC (orthorectified reflectance factor) product, the native pixels are re-sampled in a two-step process. First, a geometric process computes the grid that gives, for each point of the output image, its location in the focal plane and a bi-spline interpolation function calculates the radiometric quantity in the output image (Gascon, Cadau et al. 2014). The result provides an orthorectified image where \( CN_{\text{KNTD}}(i,j) \) is the equalised numeric digital count of the pixel \( (i,j) \) for the
band $k$ and NTDI, is the Number of Time Delay Integration lines. The parameters $k$ and $b$ represent both the MSI spectral bands but at L1C and L1B processing chains, respectively. That is, $k$ refers to the spectral band of the orthorectified image, whereas $b$ refers to the spectral band of the instrument focal plane image.

As a final step, the digital counts $CN_{k,NTDI}(i,j)$ are converted into reflectance factors (TOA reflectance normalised to the ideal Lambertian diffuser):

$$\rho_k(i,j) = \frac{\pi \cdot CN_{k,NTDI}(i,j)}{A_{k,NTDI} \cdot E_S \cdot d(t) \cdot \cos(\theta_s(i,j))}$$  \hspace{1cm} (3.6)

where $E_S$ is the equivalent extra-terrestrial solar spectrum and depends on the spectral response of the S2 bands, $d(t)$ is a correction for the Sun–Earth distance variation, $A_{k,NTDI}$ is the absolute calibration coefficient, and $\theta_s(i,j)$ is the per-pixel Solar Zenith Angle (SZA). The symbol of $A_{k,NTDI}$ and $A(b)$ represent the same absolute calibration coefficient but, as explained above, refer to L1C and L1B terminology respectively. They are described as:

$$A(b) = \frac{1}{N_p \cdot N_t \cdot N_d} \sum_{p,i,l} \frac{\pi \cdot d_{sun}^2(t) \cdot Y_{a}(p,l,b,d)}{K_{st} \cdot \rho(p,s,\theta_{sd}(l),\phi_{sd}(l)) \cdot E_{sun}(b) \cdot \cos \theta_{as}(l)}$$  \hspace{1cm} (3.7)

where $K_{st}$ refers to the stray-light correction in calibration, $d_{sun}(t)$ refers to the Sun–Earth distance variation in astronomical units, $N_t$ refers to the number of averaged lines $l$, and $\theta_{sd}, \phi_{sd}$ are the SZA and Viewing Zenith Angle (VZA) of the diffuser, respectively.

The correction $d(t)$ in Equation 3.6 represents the inverted square relationship with the Sun–Earth distance variation in astronomical units. That is, the effect on the diffuser irradiance of the inverse square law of irradiance as a consequence of the Sun-Earth distance variation along the orbit. The relationship is as follows:

$$d(t) = \frac{1}{d_{sun}(t)} = \frac{1}{(1-0.01673 \cdot \cos(0.0172 \cdot (t-2)))}$$  \hspace{1cm} (3.8)

### 3.3 Radiometric Uncertainty Contributions

#### 3.3.1 Uncertainty Contributions: Identification

From the radiometric equations presented in Equations 3.1–3.8, it is possible to identify the main sources of uncertainty and the sensitivity coefficients for each one of them. Table 1 identifies each of the uncertainty contributions and links them to the parameters in Equations 3.1–3.8. The table also
clarifies the contributions which are included in the S2-RUTv1 as well as the ones that have a negligible impact and the ones that are expected to be included in further versions of the tool.

Table 3-1 List of Sentinel-2 L1 radiometric uncertainty contributions and their associated parameter in the radiometric processing model. The table also indicates the contributions that are considered for the S2-RUTv1 implementation (marked as “Y”), the ones that may be included in next versions of the tool (marked as “N”) and the ones that have a negligible effect (i.e., <0.1%).

<table>
<thead>
<tr>
<th>L1B Contributor</th>
<th>Parameter</th>
<th>S2-RUTv1</th>
<th>L1C contributor</th>
<th>Parameter</th>
<th>S2-RUTv1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument noise, $u_{noise}$</td>
<td>$X(p,l,b,d)$</td>
<td>Y</td>
<td>Diffuser reflectance absolute</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>Out-of-field Stray-light—systematic part, $u_{stray_sys}$</td>
<td>$X(p,l,b,d)$</td>
<td>Y</td>
<td>Diffuser reflectance temporal</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>Out-of-field Stray-light—random part, $u_{stray_rand}$</td>
<td>$X(p,l,b,d)$</td>
<td>Y</td>
<td>Angular diffuser knowledge—BRF effect</td>
<td>$&lt;0.1%$</td>
<td></td>
</tr>
<tr>
<td>Crosstalk, $u_{c,talk}$</td>
<td>$X(p,l,b,d)$</td>
<td>Y</td>
<td>Instrument noise and dark signal during calibration</td>
<td>$Y_{ad}(p,l,b,d)$</td>
<td>&lt;0.1%</td>
</tr>
<tr>
<td>Deconvolution residual</td>
<td>$X(p,l,b,d)$</td>
<td>N</td>
<td>Solar irradiance model</td>
<td>$E_S(b)$</td>
<td>&lt;0.1%</td>
</tr>
<tr>
<td>Polarisation error</td>
<td>$X(p,l,b,d)$</td>
<td>N</td>
<td>Angular diffuser knowledge—cosine effect, $\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>ADC quantisation, $u_{ADC}$</td>
<td>$X(p,l,b,d)$</td>
<td>Y</td>
<td>Straylight in calibration mode—residual, $u_{diff,k}$</td>
<td>$K_{st}$</td>
<td>Y</td>
</tr>
<tr>
<td>Compression noise</td>
<td>$X(p,l,b,d)$</td>
<td>&lt;0.1%</td>
<td>Sun-to-satellite distance knowledge</td>
<td>$d(t)$</td>
<td>&lt;0.1%</td>
</tr>
<tr>
<td>Dark signal knowledge</td>
<td>$DS(p,j,b,d)$</td>
<td>&lt;0.1%</td>
<td>Angular observation knowledge—cosine effect</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>Dark signal stability, $u_{DOS}$</td>
<td>$PC_{mask}(l,b,d)$</td>
<td>Y</td>
<td>Orthorectification uncertainty propagation</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>Non-linearity and non-uniformity knowledge, $u_{gamma}$</td>
<td>$\gamma(p,b,d,Y)$</td>
<td>Y</td>
<td>Spectral knowledge</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>Non-uniformity spectral residual</td>
<td>$\gamma(p,b,d,Y)$</td>
<td>N</td>
<td>Geometric knowledge</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
<tr>
<td>L1B Image quantisation</td>
<td>$CN_{L1T0}(l,i,j)$</td>
<td>&lt;0.1%</td>
<td>L1C Image quantisation, $u_{ref,quant}$</td>
<td>$\rho_d(p,\theta_d(l),\phi_d(l))$</td>
<td>Y</td>
</tr>
</tbody>
</table>

The general criterion to qualify the contributors as negligible is that they have an impact of <0.1 DN or <0.1%. This criterion is set from own experience and based on the potential impact in the final uncertainty figures. That is, a value of 0.1% added in quadrature with much larger contributions virtually does not change the final result. The term Solar irradiance model is the exception to this rule. It has been classified as negligible and is indeed not included in the budget since the reflectance factor
cancels its effect. The Sun irradiance model applied is the one described in Thuillier, et al. (Thuillier, Hersé et al. 2003); this is used for both the absolute calibration calculation in Equation 3.7 and the reflectance factor conversion in Equation 3.6. However, if the product were to be converted into radiance, the uncertainty in the solar irradiance model should be included and would definitely not be negligible. A description of the estimated uncertainty for the solar model is provided in Thuillier, Hersé et al. (2003).

The following lists the negligible uncertainty contributions from Table 3-1 with a brief description of why these have been deemed negligible:

- **The dark signal knowledge** is well known to be below the 0.1 LSB level as a consequence of the averaging of several hundreds of samples. If the samples are fully uncorrelated, the standard deviation of the mean describes the associated uncertainty. When this is not the case, alternative methods such as the Allan deviation have been described and applied to the S2 dark datasets in Gorroño, Gascon et al. (2015). In that case, it was demonstrated how hundreds of samples can be considered as independent in the dark dataset for most of the pixels and bands. Thus, the averaging of hundreds of samples reduces the noise by a factor well above 10. The impact of this residual effect is very small in relative terms since the useful signal Y has values over 100 LSB for most of the measured scenes (i.e. a relative impact <0.1%).

- **The compression noise** has been optimised so that \( N_{\text{total}} \leq 1.2 \text{ NeDL} \). Here \( N_{\text{total}} \) refers to the instrument and compression noise and \( \text{NeDL} \) refers to the noise equivalent radiance without compression noise. The effect is thus limited to a worst-case situation of 20% of total noise level. The MSI instrument specifications for signal-to-noise ratio (SNR) are generally at \( L_{\text{ref}} \) and are typically between 100 and 200 for most of the bands. The term \( L_{\text{ref}} \) denotes the reference radiance for the MSI design and can be found in Drusch, Del Bello et al. (2012). The compression noise is thus a contribution that in most of the scenarios can be assessed as <0.1%. Nonetheless, this might not be true for specific cases — e.g., low radiance measurements. For example, setting the worst-case situation with an SNR of just 50 and worst-case compression rate \( (N_{\text{total}} = 1.2 \text{ NeDL}) \), the error could scale up to 0.4%. Future versions of the tool will consider more specific cases by monitoring the compression rates.

- **The L1B image quantisation** has a minimum impact since the ground processing uses a double datatype (i.e., 32 bits). The raw signal \( X \) is codified in 12 bits and its processing at 32 bits. Thus, any truncation during the ground processing has a negligible impact.

- **The Angular diffuser knowledge—BRF effect**. The vibrations and diffuser creeping limit the knowledge of the angular coordinates during calibration. The pre-flight vibration and thermal cycling tests reported a diffuser planarity of 0.13°. This angular effect has a minimum impact in the
BRF assessment due to the near-Lambertian shape of the diffuser on-board together with a low angular uncertainty. The same reasoning does not apply for the cosine correction and is further explained in Section 3.3.2.

- The *Instrument noise and dark signal during calibration* is well known to be below the 0.1% level. Similar reasoning as for *dark signal knowledge* can be inferred here. Even for a SNR of 50 as specified for B10 in Drusch, Del Bello et al. (2012), averaging over just 400 independent samples would reduce the noise below the 0.1% level.

- The *Sun-to-satellite distance knowledge* should have a negligible impact since the positioning of the satellite with respect to the Sun — which involves the Sun and Earth ephemeris, as well as satellite positioning — is declared to be known without significant error.

- The *Angular observation knowledge—cosine effect* is well known again because of the correct positioning of the satellite with respect to the Sun. In calibration mode, the micro-vibrations and diffuser creeping affect the diffuser coordinate system but does not affect the Earth surface and Sun positioning which are known without significant error. Note that this tilting — e.g., micro-vibrations of the satellite, orbit precision, etc. — is accounted for in the *Geometric knowledge* contribution.

The contributors that are not included in S2-RUTv1, marked as “N” in Table 1, are expected to be assessed in future iterations of the tool. These contributions represent a challenging assessment and are subject to interpretation. That is, several of them have a strong dependency on the TOA spectral characteristics of the scene under measurement, whereas others largely depend on the neighbouring pixels. For example, the “spectral knowledge” will largely depend on the spectral signature of the measured scene whereas the uncertainty propagation through the orthorectification and the accuracy of the resampling will depend on the radiometric uniformity of the scene. The “novel methodologies” required to assess these contributors have already been discussed in Chapter 4 which provides a preliminary discussion of the impact of the *Ortho-rectification uncertainty propagation* and *Spectral knowledge*. The refinement of these methods might result in the implementation of these contributors in further versions of the S2-RUT. Here they are briefly described:

- *Deconvolution residual*. The signal deconvolution and denoising are considered for its compensation during the ground processing. The residual of this correction — so far not assessed — would represent the uncertainty to be included in the budget (deconvolution and denoising stages in Figure 3-1). This effect would take into account the Point Spread Function (PSF) dispersion by the mirror scattering but also the potential correction of ghosting effects in the across- and along-track. Although all these contributions are similar in nature, they should not be confused with
the out-of-field stray-light—systematic part, the out-of-field stray-light—random part, and the optical crosstalk (described at Section 3.3.2).

- **Polarisation error.** This is considered not to be a major source of error for land measurement and is not corrected for during the ground processing — the polarisation sensitivity of the MSI instrument is <3% and the degree of polarisation (DoP) is generally <10% for most land measurements. Note, however, that for specific cases, the DoP could be above 10% — e.g., inner water and coastal measurements. For those cases, the error should be flagged and, if necessary, corrected.

- **Non-uniformity spectral residual.** The relative gains are updated in-flight by measuring the solar illumination-reflected from the diffuser. The disagreement of the Sun spectral signature with respect to the nominal measurements of the Earth-reflected light introduces a systematic effect. A potential method to assess this uncertainty contributor could follow a similar methodology as for Landsat-8 Operational Land Imager (OLI) (Barsi, Lee et al. 2014).

- The Ortho-rectification uncertainty propagation, involves the radiometric interpolation of the input data to transform the MSI focal plane measurements to a pre-defined grid on the Earth. Recent missions, such as S2, include a bi-spline interpolation of the data. Research in Chapter 4 proposes a preliminary implementation of the propagation of the uncertainty through the resampling process. Further assessment should also consider the accuracy of the resampling with respect to the real scene fluctuations.

- The Spectral knowledge is produced as a consequence of the limited knowledge of the pre-flight spectral calibration and the subsequent post-launch variations. The pre-flight spectral knowledge is ultimately limited by the alignment and spectral characteristics of the calibration source. The post-launch variations are produced due to in-orbit temperature variations or temporal degradation, among others. In addition, the spectral response is typically associated with the mean for all the pixels across the focal plane. The residual of this effect is accounted in Non-uniformity spectral residual contributor. See Chapter 4 for the preliminary assessment and first implementation of this contributor.

- The geometric knowledge referred to here is the impact that the geometric uncertainty has in the radiometry. That is, how limitations in the pointing and the geo-location of the sensor can lead to error in the resultant products; this will be directly dependent on the degree of radiometric non-uniformity and the size of the scene being viewed. The nature of the effect will depend on whether the user requires a single pixel measurement or a larger area, since this effect will be largely correlated in the spatio-temporal domain.
There are 12 uncertainty contributors which have been considered for the initial version of the S2-RUTv1. Each one of them presented in Section 3.3.2.

### 3.3.2 Uncertainty Contributions: Description and Assessment

#### Instrument Noise

The S2 L1C product includes, as part of the metadata, the parameters $\alpha_Z$ and $\beta_Z$ of a noise model evaluated and updated in-flight using the dark signal and diffuser measurements (Gatti and Bertolini 2016). The two parameters and the noise model are defined at pixel level as:

$$\begin{align*}
\text{Noise}_Z(p,l,b,d) &= \sqrt{\alpha_Z(p,l,b,d)^2 + \beta_Z(p,l,b,d) \cdot Z(p,l,b,d)} \\
\alpha_Z(p,l,b,d) &= \text{STD}[Z_{ds}(p,l,b,d)] \\
\beta_Z(p,l,b,d) &= \frac{\text{STD}^2[Z_{sd}(p,l,b,d)] - \alpha_Z(p,l,b,d)^2}{A(b) \cdot K_{sl} \cdot \frac{1}{N_j} \sum_i \frac{\rho(p,\theta_{sd}(l),\phi_{sd}(l))}{\pi} \frac{E_{\text{sun}}(b)}{d_{\text{sun}}^2} \cdot \cos \theta_{sd}(l)}
\end{align*} \tag{3.9}$$

where $Z_{ds}(p,l,b,d)$ and $Z_{sd}(p,l,b,d)$ are the equalised signals — as defined in Equation 3.1 — for the dark and diffuser measurements, respectively, and STD is the standard deviation operator. Other symbols are as previously defined in Section 3.2.

The noise model takes the DS standard deviation ($\alpha_Z$) as the instrument noise in the absence of light and scales it by the signal measured. The scaling factor relies on the assumption that the increase of the noise with the light intensity is produced by the photon shot noise and is linear with respect to the variance ($\text{STD}^2$).

Note that the dark and diffuser samples are taken at a different part of the orbit and thus requires the assumption that minimum noise drift — e.g., as a consequence of voltage or temperature variations — occurs during the orbit.

The S2 noise has been modelled pre-flight using a more complex model with estimations at the beginning and end of life and a per-pixel cubic fitting. This semi-empirical model relies on a complex characterisation where apart from variance estimations at different radiance levels, other noise linearity effects such as the sense node capacitance are included. However, the pre-flight model will not be used for noise estimation since it cannot be updated in-flight and the assignment of noise coefficients at pixel level requires a large memory.

Thus, calculating the noise by reading the values $\alpha$ and $\beta$ which are appended to the product metadata is a simple but effective approach since it minimises the computing requirements and it is continuously updated during the mission lifetime and during any potential dataset re-processing.
Previous work in Gorroño, Gascon et al. (2015) showed how the dark noise standard deviation ($\alpha$ parameter) was not changing significantly across the sensor array. It was ultimately limited by the Poisson distribution with approximate changes of $\pm0.1$ LSB. Nonetheless, for specific cases, the pixels present characteristic “dark spikes” of noise much higher than the rest of pixels and are typically named “hot pixels”. Note that these are not invalid pixels but pixels, for example, with different levels of impurities to other ones (Janesick, Pinter et al. 2010). These specific cases should be flagged and potentially reported to account for in future versions of the tool.

**Out-of-Field Stray-Light Systematic Part**

This contribution refers to the out-of-field stray-light measured during nominal Earth observation that is not measured during calibration due to different angular configuration. This effect of the Earth out-of-field stray-light has been analysed as $0.3\% \cdot L_{ref}$ and verified by measurements using a uniform source out of the FOV.

Contrary to the *stray-light in calibration mode* described in Section 3.3.2, this effect is not currently corrected for in the L1 processing chain. Thus, this contributor is a known systematic effect (i.e., error) and not an uncertainty contributor (see (BIPM, IEC et al. 2008)).

The error correction would apply an offset term to the absolute calibration, as shown by modifying Equation 3.6:

$$
\rho_{1}(i,j) = \pi \cdot \frac{E_{S} \cdot d(t) \cdot \cos(\theta_{i}(i,j))}{A_{E_{S},j(i)}} \left( \frac{CN_{j(i),k(i)}}{A_{E_{S},j(i)}} - 0.003 \cdot L_{ref} \right) 
$$  (3.10)

Using a constant level of $0.3\% \cdot L_{ref}$ assumes that the out-of-field stray-light comes uniformly from the whole Earth whereas it is more likely that the effect will be higher for parts of the scene closer to the limits of the FOV. Thus, the correction would have associated variations that define the knowledge of the correction and should be accounted for in the uncertainty residual.

Although the GUM recommends that corrections for known significant systematic effects must be applied to measurement results, this may not always be feasible in specific cases such as here. Since at the time of writing this effect has not been corrected for, then for the development of the S2-RUTv1, the guideline in BIPM, IEC et al. (2008) (note on 6.3.1, p. 24) has been applied. It means that this contributor will “enlarge” the expanded uncertainty estimate by adding this component linearly — see Sections 3.4.1 and 3.4.2.
**Out-of-Field Stray-Light Random Part**

This contribution is the result of the “undesired light” that scattered into the focal plane. During the prelaunch tests of light tightness, a small fraction of light was detected at VNIR focal plane. No light was detected at the SWIR focal plane (AIRBUS 2014, AIRBUS 2015).

In terms of uncertainty, this is modelled as a normal distribution due to its random nature across the detector array. Nonetheless, this modelling is to be reconsidered in further revisions of the S2-RUT for two reasons:

- The assumption of a normal distribution has not been verified and could follow a distribution other than normal.
- The effect is of random nature across the detector array but of systematic nature for an specific pixel. That is a similar offset will be expected independently of the scene for a specific pixel. The modelling as a distribution refers to the impossibility to know a specific offset for each one of the pixels in the focal plane.

**Crosstalk**

The crosstalk is subdivided into the optical and electrical effects.

The optical channel crosstalk is produced by the internal reflection between filters and detectors. As a consequence, some rays intended for a specific detector array are incident on a different detector array. Through a thorough design effort, the inter-reflections have been reduced to a negligible level (AIRBUS 2014, AIRBUS 2015).

The contributions for the electrical crosstalk arise from the detectors, the on-board electronics, and the video chain. The assessment has identified all the sources of “contamination” from one band and studied the worst case. The effect is low for the VNIR channels even at very low radiance levels, while the effect is relatively important for the SWIR bands where the detector electrical crosstalk is significant when expressed in relative terms (AIRBUS 2014, AIRBUS 2015).

This contribution is indeed a systematic effect that could be corrected for at the focal plane measurement level. For the initial version of the tool, a pragmatic approach has been utilised — the worst-case values have been used as uncertainty estimates for this contributor. The next revisions of the budget should investigate the possibility of providing a more detailed characterisation of the effect that either allows for its correction or, preferably, also determines its associated uncertainty.
**Analog-to-Digital Converter (ADC) Quantisation**

For an ideal analog-to-digital converter (ADC), the error distribution can be modelled by a rectangular distribution with an amplitude of 1/2 LSB. The S2-RUTv1 uses this information as the uncertainty for the modelling of the contributor.

Note that the quantisation noise associated with the DS signal is negligible due to the averaging of a large number of samples. For the PC\textsubscript{masked} correction in Equation 3.3 the same concept applies; however, the limited number of blind pixels could limit the averaging effect.

It is known that no ADC is ideal. Further revisions of this contributor should study the specific ADC characteristics.

**Dark Signal Stability**

The pixel contextual offset (PC\textsubscript{masked}) parameter — see Equation 3.3 — aims at compensating for the dark signal variation due to voltage fluctuations with temperature in-orbit.

The theoretical DS variations with temperature can be modelled by using the Arrhenius law (Arrhenius and pamphlets 1889). An approximation for the VNIR silicon detectors in ambient is written in the equation below (Hopkinson, Goodman et al. 2004):

\[
    DS(T_0 + \Delta T) = DS(T_0) \cdot e^{\frac{E_{\text{act}}}{kT}}
\]  

where \(K\) is the Boltzmann constant (8.602 \(\times\) 10\(^{-5}\) eV), \(T\) is the absolute temperature of the detector in Kelvin and \(E_{\text{act}}\) is the activation energy (approximately 0.63 eV in silicon detectors).

The uncertainty budget for the dark signal stability will be provided by the residual of this correction. The residual can be determined applying the log operator to both sides of the previous equation. Then, the following equivalent can be obtained:

\[
    \log(DS(T_0)) - \log(DS(T_0 + \Delta T)) \propto \frac{1}{\Delta T}
\]  

Thus, the residual can be calculated by fitting the blind pixel measurements and thermistor readings obtained at different points along the orbit to this theoretical model.

For the SWIR detectors the previous method also applies but, in addition, its pre-flight temperature characterisation has been very extensive and the detector temperature dependence has been well established, which provides a second source of comparison (Dariel, Chorier et al. 2009).
Apart from the residual in the correction, it is critical how this correction is extrapolated to other pixels. One major limitation could be the impact of “hot pixels” or “dark spikes.” These typically have a different temperature response in relative terms and its consideration, or not, in the correction could influence the residual uncertainty (Janesick, Pinter et al. 2010).

The initial approach for the uncertainty model is to provide a conservative figure of ±0.1 LSB for VNIR channels based on the pre-flight results (<0.1 LSB for 2 K variation) (Espuche, Chorvalli et al. 2014). For the SWIR, the following are allocated: ±0.24 LSB for B10, ±0.12 LSB for B11 and ±0.16 LSB for B12. These are initially modelled as a rectangular distribution assuming the likely on-orbit temperature variation. This assumption should be further verified by analysing the fitting residual as indicated above.

**Non-Linearity and Non-Uniformity Knowledge**

The gamma correction in Equation 3.1 includes the correction of both the non-linearity and non-uniformity effects. The uncertainty assessment of the first component relies on the proposed values of fitting residuals evaluated pre-flight. The second one is characterised pre-flight and updated in-flight — see Equation 3.5. The allocated value for this second component is subject to its update in-flight. Both contributions can be added in quadrature (as defined in the GUM (BIPM, IEC et al. 2008)) to provide a global figure of the relative gains uncertainty.

The uncertainty value for this contribution in the S2-RUTv1 is extracted from the L1C metadata (Gatti and Bertolini 2016). Further revisions of the uncertainty budget should reassess the fitting residual dependency on the radiance level and the non-uniformity residual after the in-flight update.

**Diffuser Reflectance Absolute Knowledge**

This contributor has been carefully assessed pre-flight as detailed in Mazy, Camus et al. (2013). The final figures regarding the associated uncertainty describe the diffuser reflectance absolute uncertainty due to the calibration, but also other secondary effects related to the angular, spatial and polarisation performance.

In the angular domain, the uncertainty includes the fitting residual of the measurement of the Bidirectional Reflectance Distribution Function (BRDF) model. In this case, the Rahman-Pinty-Verstraete (RPV) model is used, with a further cubic function to correct the relative azimuth dependency (Rahman, Pinty et al. 1993). The results reported in Mazy, Camus et al. (2013), show a standard deviation of the error between the measured and fitted values at around 0.2%. They are taken as the reference accuracy for an angular fitting of the measured values at the pre-flight results. However, these fitting residuals might vary if during the mission another fitting function or angular interpolation is applied.
The diffuser reflectance has been well characterised at different spatial positions and these are introduced as a correction depending on the specific pixel viewing of the diffuser. The fact that diffuser non-uniformity has been characterised at a fixed angular position together with the calibration relative uncertainty translates into a certain residual in the correction that should be accounted for.

Measurements of the diffuser DoP have shown a worst-case polarisation impact of 6.6% on the diffuser. With a sensitivity of the MSI instrument lower than 2.9%, this leads to an overall polarisation error of 0.19% as a worst case (AIRBUS 2014, AIRBUS 2015). This latter figure is added in quadrature in the budget.

**Diffuser Reflectance Temporal Knowledge**

This contributor presents a similar situation to the Out-of-Field Stray-Light Systematic Part described in Section 3.3.2 since it represents a systematic effect in the measurement. It means that the systematic effect is not corrected and this will “enlarge” the expanded uncertainty estimate by adding this component linearly. However, in this case, it is also a drift which is known to evolve with time and in a known direction.

The diffuser on-board S2 has undergone an extensive characterisation and test pre-flight (Mazy, Camus et al. 2013). The tests included, among others, the exposure of samples to a 95% humidity, thermal cycling between −10°C and +40°C @ 5 × 10^{-5} mbar, 2 solar hours in front of the UV lamps as well as proton and gamma radiation. None of the tests reported an effect larger than 1%. The thermal cycling results are reported in Mazy, Camus et al. (2013) with average variations slightly higher than 0.5% at different wavelength regions of the VNIR.

Despite the efforts to test the degradation pre-flight, they can only provide a verification pre-flight. The diffuser evolution in-flight is unknown and subject to issues like hydrocarbon contamination on the Polytetrafluoroethylene (PTFE) material (Stiegman, Bruegge et al. 1993). In the absence of any other information, the degradation model for the S2-RUTv1 will be based on the diffuser degradation information provided by the MERIS (MEdium Resolution Imaging Spectrometer) instrument.

The MERIS monitoring is based on the use of two on-board diffusers, using the less frequently used “Diffuser-2” to monitor the degradation of the frequently used “Diffuser-1”. These in-flight measurements permit the track of the diffuser evolution and the possibility to introduce a diffuser degradation model that corrects for this systematic effect. The ratio of the two measurements provides an estimate of the degradation of “Diffuser-1” with respect to the reference “Diffuser-2.” Note that this correction has an associated uncertainty residual due to the limitations of the system and model itself (Delwart and Bourg 2011).
The degradation trend for both missions can be approximated as linear due to the low Sun exposures. In addition, the diffuser exposure time for MERIS mission could be assumed comparable or higher due to their usage periodicity (around 15 days per MERIS compared to 30 days for S2 MSI). For S2, an optimised manufacturing and appropriate protection of the diffuser until launch expects to significantly reduce this effect. In general terms, the use of MERIS degradation rate per year can be considered as a worst-case scenario for the Sentinel 2 diffuser.

Based on the previous discussion, for the diffuser ageing contribution, the S2-RUTv1 software:

1. Uses the approximated degradation rate per-year based on MERIS for each of the S2 bands. This means that B1 = 0.15%/year, B2 = 0.09%/year, B3 = 0.04%/year, B4 = 0.02%/year and B5 = 0.01%/year. Any band above B5 is assumed to have a negligible degradation effect.

2. Based on expected linear degradation, the tool extracts the timestamp of the image to calculate the systematic effect for the specific product.

**Angular Diffuser Knowledge—Cosine Effect**

The angular knowledge effect has been well characterised pre-flight. The vibration tests and the previously mentioned thermal cycling test reported a diffuser planarity of 0.13°. This uncertainty source is propagated to the cosine term in Equation 3.7 with an estimated effect of 0.4% \((k = 1)\).

This contribution is produced by the diffuser “creeping” as a result of launch vibrations and thermal cycling. This angular diffuser knowledge is also related to the influence of micro-vibrations and shutter mechanism angular knowledge. These latter effects are not included in the budget since such random effects are minimised through the model smoothing.

**Stray-Light in Calibration Mode—Residual**

During the diffuser calibration, the specific orientation of the instrument and the shutter mechanism means that the sunlight enters the instrument through multiple reflections. This same situation does not occur during the imaging of the Earth and introduces a systematic error in the calibration coefficient.

The pre-flight analysis evaluated this error as 0.7% of the diffuser radiance and it has been corrected by introducing the term \(K_{sl}\) in Equation 3.7. It is the knowledge on this correction that needs to be accounted for in the uncertainty budget. A residual of 0.3% has been allocated for this contributor.

The determination of this contribution is a difficult task which would involve the development of techniques to evaluate e.g., the sensitivity of the ray-tracing model. The first approach brings a conservative allocation with the expectation of future refinements.
**Image Quantisation**

The entire ground processing for the S2 L1 products is performed using 32 bits. However, the final images of reflectance factors need to be codified in JPEG2000 format with a maximum number of 16 bits.

The resulting reflectance factor values need to be re-scaled to fit in the range \([0, 2^{16}]\). This is created by applying a “quantification value” — at the time of writing this is 10,000 (Gatti and Bertolini 2016). That is, the resulting reflectance factors in the image pixels are scaled by this value.

This contributor has a minimum impact for most of the measurements but has been integrated in the S2-RUTv1 since its implementation is straightforward and under very low reflectance factor values (<0.1), the effect could be slightly higher than 0.1%. In addition, any alteration of the quantification value will be accounted for.

### 3.4 Model Combination and Validation

#### 3.4.1 Model Combination

The proposed model to combine the uncertainty sources considered in the S2-RUTv1 tool (see Table 3-1) obtains an expanded uncertainty \(U(R_k(i,j))\) for an expansion coefficient, \(k\), and is the following:

\[
U(R_k(i,j))[\%] = k \cdot u(R_k(i,j)) + \sqrt{\frac{100 \cdot A_{N_DTM,i} \cdot u_{stray,sys}}{CN_{N_DTM,i}}} + u_{diff, temp} + u_{stamp, temp}
\]  

(3.13)

The equation includes two systematic contributions added linearly. There is a further discussion over this specific addition in section 3.4.2. The digital counts \(CN_{N_DTM,i}\) are obtained from an inversion of the Equation 3.6 using as input the per-pixel values of TOA reflectance factor. The term \(u(R_d(i,j))\) denotes the combined standard uncertainty and is shown in Equation 3.14.

\[
u(R_k(i,j)) = \sqrt{u_{ref,quant}^2 + u_{diff}^2 + u_{diff,temp}^2 + u_{stamp}^2 + u_{LSB}^2}
\]  

(3.14)

The combined standard uncertainty is obtained from the GUM law of propagation of uncertainty (see Equation 1.3). This equation is based on a Taylor expansion and requires the correlation between the different contributions as input. In this case, no significant correlation was found between the different contributions and is further discussed in section 3.4.4. In addition, the validity of the GUM framework is based on the validity of the central limit theorem. The combination of the different terms in Equation 3.14 are expected to result in a normal distribution. This is tested in section 3.4.5.

The term \(u_{ref,quant}\) refers to the limit of a rectangular distribution (in this case of ±0.5LSB) and its division by \(\sqrt{3}\) converts this value into the expected \(k=1\) uncertainty. The terms \(u_{diff}\), \(u_{stray}\) and \(u_{LSB}\) have been
used for simplification of the Equation 3.14. Each one of them represents several uncertainty contributions linked to the diffuser, stray-light and LSB respectively. They are specified in Equations 3.15, 3.16, and 3.17 respectively:

\[ u_{\text{diff}}[\%] = \sqrt{u_{\text{diff}_k}^2 + u_{\text{diff}_cos}^2 + u_{\text{diff}_abs}^2} \]  
(3.15)

\[ u_{\text{stay}}[\%] = \sqrt{u_{\text{stay}_\text{rand}}^2 + \left( \frac{100 \cdot A_{\text{NTDI}} \cdot u_{\text{k talk}}}{CN_{\text{NTDI}}(i,j)} \right)^2} \]  
(3.16)

\[ u_{\text{LSB}}[\%] = \sqrt{\left( \frac{100 \cdot u_{\text{noise}}}{CN_{\text{NTDI}}(i,j)} \right)^2 + u_{\text{DS}}^2 + u_{\text{ADC}}^2} \]  
(3.17)

The terms \( u'_{\text{DS}} \) and \( u'_{\text{ADC}} \) in Equation 3.17 are further calculated as:

\[ u'_{\text{ADC}}[\%] = \frac{100 \cdot \left( u_{\text{ADC}} / \sqrt{3} \right) \cdot c_Y}{CN_{\text{NTDI}}(i,j)} \]  
(3.18)

\[ u'_{\text{DS}}[\%] = \frac{100 \cdot u_{\text{DS}} \cdot c_Y}{CN_{\text{NTDI}}(i,j)} \]  
(3.19)

The sensitivity coefficients, \( c_Y \), are the derivative of the non-linearity and non-uniformity correction described in Equation 3.4 and for the VNIR bands is:

\[ c_Y = \frac{\partial Z}{\partial Y} = \left| g_1 + g_2 \cdot 2 \cdot Y + g_3 \cdot 3 \cdot Y^2 \right| \]  
(3.20)

Whereas, for the SWIR bands, is expressed as:

\[ c_Y = \frac{\partial Z}{\partial Y} = a1 \text{ if } Z_C < Y < Z_C + Z_S, \quad c_Y = \frac{\partial Z}{\partial Y} = a2 \text{ if } Y > Z_C + Z_S \]  
(3.21)

### 3.4.2 Discussion of the Linear Addition of Contributions

Although the GUM recommends that corrections for known significant systematic effects should be applied, there are two contributions in the S2-RUTv1 where this was not feasible, namely Out-of-Field Stray-Light Systematic Part and Diffuser Reflectance Temporal Knowledge detailed in Section 3.3.2. These contributions do not represent a distribution of potential values about the quantity to be measured (i.e., uncertainty) but rather a deviation from the value that is intended to be measured (i.e., error). In such cases, the guideline in (BIPM, IEC et al. 2008) (note on 6.3.1, p. 24) can be applied; thus these contributions are added linearly in the RUT and they are independent of the coverage interval. Note that if these errors where corrected for, the residual of the correction would be accounted as an uncertainty.
These two contributors “enlarge” the expanded uncertainty estimate in Equation 3.13 by adding linearly. Three different possibilities are explained and discussed here:

- **Option 1:** the combination in Equation 3.22 below is the one used in Equation 3.13 for the S2-RUTv1. It describes the addition of each systematic effect in absolute value and its addition to the expanded uncertainty. It is known that the diffuser degradation is expected to introduce a negative systematic effect (Diffuser Reflectance Temporal Knowledge in Section 3.3.2), whereas the out-of-field stray-light’s systematic part is a positive systematic effect. Thus, this approach can be considered a pessimistic approach but makes sure that the uncertainty accounts for the potential uncorrected systematic errors.

\[ U = k \cdot u + |b| + |c| \]  
(3.22)

- **Option 2:** the combination in Equation 3.23 adds linearly each of the contributors and adds the absolute value to the expanded uncertainty. This seems a logical approach from a radiometric point of view since the different sign of each systematic effect is accounted and compensated for. However, it brings a more challenging interpretation when the uncertainty associated with the knowledge of the systematic effect is large — i.e., when the uncertainty residual of a potential correction would be relatively high. This is indeed the case for the two systematic effects introduced in Equation 3.13.

\[ U = k \cdot u + |b + c| \]  
(3.23)

- **Option 3:** the combination in Equation 3.24 adds linearly the maximum of the systematic effects and adds the absolute value to the expanded uncertainty. This is a good approach when there is a single dominant systematic effect with respect to the others and it is suggested in BIPM, IEC et al. (2008). However, in Equation 3.13 this is not the case since the diffuser temporal stability depends on the instrument timestamp. That is, depending on the time, the weight of each of the two systematic effects could vary and thus also the maximum.

\[ U = k \cdot u + \max(|b|, |c|) \]  
(3.24)
3.4.3 Sensitivity Coefficient Impact

The standard uncertainty combination in Equation 3.14 requires the knowledge of the sensitivity coefficients, \( c_y \), for the contributors \( u'_{\text{DS}} \) and \( u'_{\text{ADC}} \) — Equation 3.20 for the VNIR bands and Equation 3.21 for the SWIR bands.

The decision made for the S2-RUTv1 tool is to not to calculate these sensitivity coefficients for each pixel in the focal plane. That is, the sensitivity coefficient \( c_y \) is set to a constant value of 1. Applying each one of the specific per-pixel gamma parameters and inverting the value to obtain the \( Y \) signal is not impossible, but will introduce a lot of complexity since a look-up table (LUT) would need to be included and the processing time and memory required would increase. Each of the 12 VNIR detectors in the focal plane consist of 2596 pixels for the 10m bands and 1298 pixels for the 20m bands. The 12 SWIR detectors in the focal plane consist of 1298 pixels per band. Therefore, the study presented in this section considers how the decision of not including the sensitivity coefficients impacts the global uncertainty estimates.

Figure 3-2 Distribution of the sensitivity coefficient \( c_y \) for all the pixels in the Sentinel-2 (S2) Visible and Near-InfraRed (VNIR) bands (a) B1; (b) B2; (c) B3; (d) B4; (e) B5; (f) B6; (g) B7; (h) B8 and (i) B8A.
The panels in Figure 3-2 calculate the distribution of the sensitivity coefficients for the VNIR bands for all the pixels in the S2 focal plane. The panels in Figure 3-3 repeat the same process but for the bands in the SWIR. The values g1, g2 and g3 for the VNIR and a1 and a2 for the SWIR have been obtained from the pre-flight characterization. Note that these values will be re-scaled once in-flight by using the diffuser measurements (see Equation 3.5). It is assumed that this re-scaling will not significantly change the results. The bands B9 and B10 have not been included since their main application is the cloud-screening (Drusch, Del Bello et al. 2012).

The value of $c_y$ in the VNIR (see Equation 3.20) depends on the signal $Y$. Results for VNIR have been calculated at several levels of $Y$ signal and the distribution of values in relative terms proves to be largely independent of the $Y$ signal level. Thus, Figure 3-2 and Figure 3-3 show results with $Y$ signal equivalent to a radiance level close to $L_{ref}$.

The results in Figure 3-2 and Figure 3-3 show values of standard deviation <6%. It can be said that the majority — >90% of the pixels — are in the range ±10%. There are specific bands like the B7 and B8A whose standard deviation is approximately 5% but they show an important skew of the values. For these
cases, it is possible that some pixels reach an error of 20%—i.e., sensitivity coefficients going up to 1.2 and down to 0.8.

The expected impact of this simplification in the combined standard uncertainty will be negligible. The addition in quadrature of the uncertainty contributions in Equation 3.14 minimises its impact since the uncertainty levels of these contributors—\(u'_{\text{DS}}\) and \(u'_{\text{ADC}}\)—are generally smaller if compared to other contributions in the budget.

### 3.4.4 Correlation between Uncertainty Contributors

In general terms, it is justified to add all the contributors in Equations 3.13 – 3.17 with no correlation among them. The justification relies on the fact that the different parameters of the L1 processing as \(X(p,l,b,d)\), \(DS(p,l\text{mod}6,b,d)\), or \(A(b)\) — see Equations 3.2, 3.3 and 3.7 respectively — have been measured at different times and with different methods so that no substantial relationship can be found among them. Nonetheless, some specific cases which could be more controversial are briefly discussed here:

- \(u_{\text{noise}}\) vs. \(u_{\text{DS}}\). The dark signal stability and the instrument noise could be correlated by the temperature. However, this will be important when several pixels are combined across-track. In that case, variations of temperature will produce variations of both noise and \(DS\) in a proportional way. Here the \(DS\) signal is corrected for variations due to temperature by “blind pixels\(^2\),” the residual of this correction \(u_{\text{DS}}\) is uncorrelated with \(u_{\text{noise}}\). Since, despite being measured at the same time, different pixels are used to assess the performance.

- \(u_{\text{stray_rand}}\) vs. \(u_{\text{xtalk}}\). It could be discussed whether the optical crosstalk is correlated since a similar optical mechanism produces both sources. Nonetheless, since the optical crosstalk has been minimised and the electrical crosstalk is the dominant source (see Crosstalk in Section 3.3.2), any correlation has minimal effect.

- \(u_{\text{gamma}}\) vs. \(u_{\text{diff_abs}}, u_{\text{diff_temp}}\). These effects could also be correlated since the same measurements from the diffuser are used for both the gamma correction update and the diffuser absolute calibration. However, the update of the gamma correction in Equation 3.5 is completed in-flight by using only values \(Z_{sd}\) and no conversion to radiance is necessary. It is important to note that \(u_{\text{gamma}}\) is fully correlated with the Instrument noise and dark signal during calibration reported in Table 3-1; since this contributor has a

\(^2\) Blind pixels are a small number of pixels situated at each side of the detector line array. These pixels are not illuminated. They are intended to measure the dark signal level and its variations.
negligible uncertainty level, the contributor, and hence correlation effect, is not included in the combination model.

- $u_{\text{noise}}$ vs. $u_{\text{ADC}}$. The instrument noise is propagated through the ADC quantisation. The two components can be assumed significantly uncorrelated for a noise larger than the quantisation effect. We explore this option further in Section 3.4.5 since, below a certain limit, the uniform distribution of ADC does not apply and the noise can be modulated by the digital conversion.

### 3.4.5 Validation of the Central Limit Theorem

The combined standard uncertainty in the GUM relies on the propagation of the uncertainty contributions through a linearised measurement model. The associated distribution can be approximated as normal based on the validity assumption of the central limit theorem as described in section G.2 in BIPM, IEC et al. (2008). Here we provide an initial assessment on how well the GUM method provides a valid uncertainty estimation of the L1 model by comparing the results to a Monte Carlo Method (MCM), the latter being capable of propagating the uncertainty sources through the full model (BIPM, IEC et al. 2008).

In this initial version, several uncertainty contributors concerning the L1 radiometric processing have been considered and listed in Table 3-1. The script automatically reads certain calibration parameters as $DS$ or relative gain coefficients. The version was presented in Gorroño, Gascon et al. (2015) and expanded to L1C in Gorroño, Banks et al. (2016). The effects of the contributors detailed in Out-of-Field Stray-Light Systematic Part and Diffuser Reflectance Temporal Knowledge detailed in Section 3.3.2 are not considered, since they are added linearly in Equation 3.13. The contributions of out-of-field stray-light—random part and crosstalk in Section 3.3.2 are also not considered at this stage although it is intended that their distribution and effect should be revisited in later versions and also that these contributions have a low impact on the final results. The input contributions for the MCM have been modelled as either normal or rectangular distributions. The selection of the adequate distribution is based on expert knowledge. For example, error distributions associated to noise are generally close to a normal distribution. However, quantisation error is generally close to a rectangular distribution since the truncated value is equally possible in all the precision range and; thus, the truncation error is close to a uniform probability. Table 3-2 summarises the uncertainty contributions considered, their values and associated distribution.

Figure 3-4 provides the difference between the GUM uncertainty combination ($k = 1$) and the area around the mean with approximately 68.27% of output values of the MCM uncertainty distribution. The results are presented for a range between $L_{\text{min}}$ and $L_{\text{ref}}$ approximately.
Table 3-2. Considered uncertainty contributors for the L1B model validation.

<table>
<thead>
<tr>
<th>L1B Contributor</th>
<th>Value</th>
<th>Distribution Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instrument noise, $u_{\text{noise}}$</strong></td>
<td>Calculated as in Equation 3.9. Coefficients $\alpha$ and $\beta$ extracted from the datastrip metadata (Gatti and Bertolini 2016)</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>ADC quantisation, $u_{\text{ADC}}$</strong></td>
<td>±0.5 [LSB]</td>
<td>Rectangular</td>
</tr>
<tr>
<td><strong>Dark signal knowledge</strong></td>
<td>±0.05 [LSB]</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Dark signal stability, $u_{\text{DS}}$</strong></td>
<td>±0.1 VNIR, ±0.24 B10, ±0.1 B11, ±0.16 B12 [LSB]</td>
<td>Rectangular</td>
</tr>
<tr>
<td><strong>Relative gains accuracy</strong></td>
<td>Extracted from quoted L1C metadata (0.4%) (Gatti and Bertolini 2016)</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Diffuser uncertainty</strong></td>
<td>From pre-flight characterisation (Mazy, Camus et al. 2013)</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Diffuser angle knowledge</strong></td>
<td>±0.4%</td>
<td>Normal</td>
</tr>
<tr>
<td>$K_\text{in}$ residual</td>
<td>±0.3%</td>
<td>Rectangular</td>
</tr>
</tbody>
</table>

The results in Figure 3-4 clearly show the validity of the GUM framework for the L1B product with disagreement <0.1% between both methods (GUM vs. MCM) for the majority of the radiance range. For low radiance values, the use of the GUM approach does not represent a reliable parameter to characterise the uncertainty. At this low level of radiance, the ADC noise becomes unstable and invalidates the central limit theorem.

The majority of land scenes will generally measure radiances where this effect is not applicable or negligible. However, it could be that specific scenes, such as very dark vegetation forest or case-2 water scenes, are near the radiance levels where the GUM approach becomes unreliable.
Figure 3-4 S2 L1B model validation for the S2 bands (a) B1; (b) B2; (c) B3; (d) B4; (e) B5; (f) B6; (g) B7; (h) B8; (i) B8A; (j) B11; and (k) B12. The graph shows the difference when calculating the uncertainty using the ‘Guide to Expression of Uncertainty in Measurement’ (GUM) uncertainty ($k = 1$) or the 68.27% area of the Monte Carlo Method (MCM) symmetric from its mean.

The bands B3-B8A have the largest dynamic range and, at extremely low light levels, the quantisation becomes dominant. Figure 3-5 shows an example for B6 at different radiance levels.
Figure 3-5 L1B MCM distribution at (left) 4.93 Wm$^{-2}$·sr$^{-1}$·μm$^{-1}$ and (right) 31.41 Wm$^{-2}$·sr$^{-1}$·μm$^{-1}$. The theoretical normal distribution of the combined standard uncertainty has been included and re-scaled to the MCM distribution peak value for comparison.

These two distributions agree for values close to $L_{ref}$ almost perfectly, demonstrating that the output distribution is fairly normal. However, we can see clearly the effect of the quantisation effect for low radiance values. The answer given in Carbone and Petri (1998) for a normal noise provides an alternative analytical form to understand this effect.

The validation up to L1C involves the propagation through the bi-spline radiometric interpolation. This is discussed in Chapter 4.

In summary, the effect of non-linear corrections when combining the uncertainty contributions is treated as being negligible in the S2-RUTv1. Even for low radiances the effect can be considered small (~0.1%). The unstable effects due to the ADC quantisation invalidate the central limit theorem assumption when evaluating the uncertainty at very low radiance levels. However, this situation can only occur for few bands under exceptional circumstances.

3.5 Software Implementation and Integration

3.5.1 Tool Integration, System Requirements and Performance

The development of the tool code is fully accessible in the software repository in (Gorroño, Fomferra et al. 2016). This code works as a plug-in that is embedded as part of the Sentinels Toolbox. Both SNAP and this plug-in support all major platforms, like Windows, UNIX and Mac OS. In order to support the S2 L1C data products, the S2-Toolbox needs to be installed in addition to the bare SNAP installation. Their last releases are available at (ESA). Python, as implementation language, needs to be installed and the version must be 2.7 or later. In addition, this Python installation must have the NumPy library installed.

The requirements on the hardware are very dependent on the processing operation and the source data. For S2-RUT, the requirements are not very demanding and it can run in a minimal configuration. On a
computer, with an Intel Core i7 processor and 8 GB of RAM, processing all bands of a S2 L1C product at once will roughly take 11–12 min. The computation of a single band varies from 5 to 110 seconds depending on its resolution.

The S2-RUT works as a set of routines that interface with SNAP. First, SNAP retrieves information from the s2_rut-info.xml for the creation of the user interface for the operator. When the user is satisfied with the configuration, the operator (implemented in the class S2RutOp) can run the code. As a result, the user gets the uncertainty image for the selected band(s). The class S2RutAlgo is called during the process of the uncertainty image generation as it brings the core of the uncertainty calculation. The main features of the tool software design are:

- SNAP Python libraries (snappy) for product readout.
- General tool design to accommodate other sensors.
- Maximisation of product info extraction (e.g., noise coefficients) makes it robust against re-processing, contingencies etc.
- Conservation of the geolocation information for collocation between the L1C reflectance factor and uncertainty images.

3.5.2 Processor Documentation

The S2-RUTv1 processor can be invoked in SNAP from the menu by selecting Optical->Preprocessing->Sentinel-2 Radiometric Uncertainty Tool. On the command line processor, available by means of the Graph Processing Tool gpt, which is located in SNAP bin directory, typing gpt S2Rut -h displays further information.

Selecting the Sentinel-2 Radiometric Uncertainty Tool command from the SNAP menu opens up the dialog in Figure 3-6.
In the “Source product” area, the user specifies the source product. The combo box presents a list of all products opened in SNAP. The user may select one of these or, by clicking on the button next to the combo box, choose a product from the file system. The selected product must be of type S2_MSI_Level-1C. In the “Target product” area, the user can specify a name for the generated product. As a default, the tool will automatically assign a name by adding the extension “_rut” to the Sentinel-2 L1C product selected. The user can also specify where the target product should be saved in the file system. The combo box presents a list of available file formats. The text field or the button next to it allow specification of a target directory. Finally, the option “Open in SNAP” specifies whether the target product should be opened in SNAP. When the target product is not saved, it is opened in the Sentinel Toolbox automatically.

The tool incorporates a secondary tab named “Processing Parameters.” Figure 3-7 presents a screen-shot as it appears in SNAP.
This tab is intended for expert users and allows different options for the uncertainty combination choices. Standard users have default parameters. Three main sections with multiple options are included in the S2-RUTv1:

- The option “Coverage factor” permits specifying the $k$ parameter that assigns a probability coverage to the uncertainty evaluation — $k = 1$ means 68.27% probability (BIPM, IEC et al. 2008).

- The selection of “Band names” means the uncertainty shall be computed.

- For uncertainty contribution selection, the tab includes a selection list with the uncertainty contributions included in the S2-RUTv1 (see Table 3-1). This permits the user the selection and deselection of specific contributions in order to perform sensitivity analysis and separation of random/systematic contributions, etc.

### 3.5.3 Output Generation

The uncertainty results at pixel level are codified in UINT8 (i.e., one byte). Values from 0–250 refer to uncertainty values from 0%–25% in steps of 0.1%. The values are clipped to this range so:
- 0 == Invalid uncertainty (it cannot be 0 or negative)
- 250 == Uncertainty ≥25%

This means that the tool can be helpful in managing the GML information in the product masks. The tool can convert the GML vector data in raster information that complements the uncertainty image. If the pixel status is “saturated, no data, cloud, and defect,” then the uncertainty evaluation is of course not required and the five values missing — from 251 to 255 — could be used to provide raster information of the GML masks.

Another byte image could potentially be included in future versions, and include information regarding specific flags for polarisation or stray-light events as well as information related to the correlation between pixels in the time, space and spectral dimension (see Chapter 6). To sum up, the idea is that an additional byte becomes a “quality indicator” that supports the calculation of uncertainty and their application.

3.5.4 Radiometric Uncertainty Tool (RUT) Version 1 (v1) Case Study: Albufera Lake

Figure 3-8 shows the result of running the S2-RUTv1 in SNAP for a specific S2 L1C product and band. The area shown represents Albufera Lake in Valencia (Spain) on 12th January 2016 for B8 and its associated uncertainty. The scale of values for both the L1C reflectance factor and uncertainty are shown in Figure 3-9.

The area surrounding the lake is subject to variations in the water level (rice field) and, depending on the area and overpass time, the variation of uncertainty (in relative units) can change significantly. Figure 3-8 also indicates how Albufera Lake has lower uncertainty than many of the rice fields due to the large amount of sediments.

The radiometric uncertainty — on the right of Figure 3-8 — ranges from >10% for open sea to 5%–6% in the lake body, 5%–15% in the rice fields covered by water and 2%–4% in the land areas.
Figure 3-8 Screen-shot of SNAP. It contains the pixel info and navigation panels (left); the L1C Sentinel-2 B8 image for Albufera Lake (centre) and the equivalent uncertainty $k = 1$ (right). The image is North oriented.

Figure 3-9 Screen-shot of L1C reflectance factor scale in Figure 3-8 (top); and the scale of equivalent uncertainty $k = 1$ (bottom). The L1C reflectance factor is multiplied by the quantification value of 10,000 and the uncertainty figures are given as percentages multiplied by 10.

Figure 3-10 shows another S2 L1C image of B8 over Albufera Lake and its surroundings. This one corresponds to an overpass on the 12th of November 2017. Different ROIs are overlaid on the image corresponding to different types of scenes which are:

- **lake**, which represents a section of Albufera lake;
- **rice**, which represents a small number of rice fields close to the Albufera Lake;
- **sea**, which represents an area of the Mediterranean Sea at around 20 Km from the Albufera Lake;
- **forest**, which represents an area comprised of hills covered with bushes and sparse trees;
- **fields**, which represents a small area covered by citric plantations and;
- **city**, which represents an urban area in the city centre of Valencia.
Figure 3-10 Screen-shot of L1C Sentinel-2 B8 image of Albufera Lake and surroundings on the 12th of November 2017. Overlaid on the image, the considered ROIs under study.

Table 3-3 shows the statistics for the selected ROIs in Figure 3-10 and for the S2 10 m bands (B2, B3, B4 and B8). The considered statistics are the mean value of TOA reflectance factor, the mean value of L1C uncertainty and its associated standard deviation. The uncertainty figures are given as a percentage and the scaling factor of 10 has been already applied.

In this section, the mean value of uncertainty over a ROI has been calculated. Note that this is not equivalent to the uncertainty associated to the TOA reflectance factor mean value. In order to calculate this estimate, further knowledge as the correlation is required. This is explored in detail in Chapter 6 of the thesis.

Table 3-3. Statistics for the considered ROIs in Figure 3-10 and for the S2 L1C product acquired on the 12th of November 2017.

<table>
<thead>
<tr>
<th>ROI</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean TOA</td>
<td>Mean Unc</td>
<td>Std Unc</td>
<td>Mean TOA</td>
</tr>
<tr>
<td>Lake</td>
<td>0.086</td>
<td>4.09</td>
<td>0.04</td>
<td>0.087</td>
</tr>
<tr>
<td>Rice</td>
<td>0.086</td>
<td>3.74</td>
<td>0.17</td>
<td>0.086</td>
</tr>
<tr>
<td>Sea</td>
<td>0.088</td>
<td>4.00</td>
<td>0.02</td>
<td>0.049</td>
</tr>
<tr>
<td>Forest</td>
<td>0.096</td>
<td>3.85</td>
<td>0.22</td>
<td>0.077</td>
</tr>
<tr>
<td>Fields</td>
<td>0.100</td>
<td>3.76</td>
<td>0.20</td>
<td>0.090</td>
</tr>
<tr>
<td>City</td>
<td>0.126</td>
<td>3.34</td>
<td>0.26</td>
<td>0.113</td>
</tr>
</tbody>
</table>

The results show a small variation of the uncertainty figures between the different ROIs for S2 B2. However, this variation becomes large when considering the S2 B8. That situation can be explained since the TOA bands that receive a large portion of radiation from the atmosphere usually present more
stable values of TOA radiance than those dominated by surface radiation. This is translated into a larger per-pixel variation of uncertainty for bands dominated by the surface radiation in comparison to those dominated by the atmosphere.

For the city and fields ROIs, the variation of uncertainty between the four 10 m bands is below the 1%. For the forest case, it goes up to 1%. However, for water cases — lake, rice and sea — these variations are extreme. In the case of sea, these values go from just 4 % in B2 to over 22 % for B8.

Some uncertainty contributions are not linearly dependent with the level of radiance and cannot be expressed as a fix percentage independent on the radiance level. For example, the instrument noise in section 3.3.2 has a root square-root dependence with the level of measured radiance. The ROIs covered by water — lake, rice and sea — show reduced values of TOA radiance in B8. Their associated larger values of uncertainty can be explained by the expression of the uncertainty as a relative value. In these cases, some uncertainty contributions are invariant (or non-linear variant) in absolute terms and their relative value increases when the TOA radiance decreases.

To sum up, the examples presented showed how the levels of uncertainty vary depending on, for example, the TOA band and the selected ROI. That is, this section briefly shows how the per-pixel uncertainty based on scene, time, radiance etc. is very useful to understand the ‘fitness for purpose’ of the product to the desired application.

### 3.6 Conclusions and Further Work

This chapter describes an uncertainty analysis methodology that estimates per-pixel radiometric uncertainty of TOA reflectance factor measurements. The methodology follows the metrological guidelines presented in the GUM considering its limitations and alternatives. The work here is focused on describing an exhaustive uncertainty methodology and a general software design that can be followed in subsequent versions of the tool and can be readily adapted to several other Earth Observation (EO) missions.

The uncertainty analysis effectively links the uncertainty contributors with the radiometric model. The identified uncertainty contributors are exhaustively assessed and their combination using the ‘Guide to Expression of Uncertainty in Measurement’ (GUM) model is discussed in detail, with a special effort here in validating and discussing the combination model. The uncertainty analysis has flagged and discussed the existence of systematic uncorrected effects in the S2 L1C radiometry.

The uncertainty analysis methodology has been applied to the Sentinel-2 mission and has been implemented in a tool named Sentinel 2 Radiometric Uncertainty Tool (S2-RUT) tool. This tool enables users to generate per-pixel radiometric uncertainty associated with the S2 Level 1 (L1C) products.
This initial version of the tool has been implemented — code available at Gorroño, Fomferra et al. (2016) — and integrated as part of Sentinels Application Platform (SNAP). The integration of the S2-RUT tool as part of SNAP facilitates the integration of the uncertainty products in the users’ EO processing and thus does not increase the memory and storage requirements of the operational S2 L1 processing performed at the European Space Agency (ESA). The tool software design looks for an optimal strategy to read the TOA reflectance factor images so that they can run on a “standard” computer.

The first version of the tool also includes the assessment of the uncertainty at a desired coverage probability by setting a coverage factor, $k$, and the selection/deselection of the uncertainty contributors for sensitivity studies.

Although a first version has been fully implemented, there are still several areas to study and incorporate in subsequent versions of the tool. Many of the areas that could be improved in future versions have been mentioned throughout this chapter. This would include the introduction of additional uncertainty contributions—note that this work has already started in Chapter 4—and the refinement of some of the uncertainty contributions in S2-RUTv1.

The TOA radiometric uncertainty provided by the S2-RUT is expected to be the input for its propagation through consecutive steps of the EO processing chain similarly to Merchant, Embury et al. (2014). These higher-level uncertainty estimates will require the covariance information in the spatial, temporal and spectral dimensions. The results in Chapter 6 implements an initial version of the correlation between the different pixels and shows how necessary is to include this information for combination or propagation of uncertainty estimates. The implementation of the S2-RUT as an external tool might introduce limitations, and hybrid or full implementation of the covariance might need to be included in the L1 processing as described, for example, in the Kepler mission (Clarke, Allen et al. 2010). Nonetheless, several trade-off analyses and alternatives will need to be considered for its implementation.

The software implementation in future versions would seek to include the new uncertainty contributions and update the results of a refined uncertainty analysis. In addition, it should include the processing of the quality masks’ information of the S2 L1C products as part of the uncertainty image itself (see Section 3.5.3). Use of quality flags related to polarisation and stray-light effects as specified in Section 3.5.3 or covariance information may be included as a second byte codification. The interface with other SNAP plug-ins might enhance even further the possibilities of the tool.
Chapter 4.

Novel techniques for the analysis of the TOA radiometric uncertainty

4.1 Introduction

As part of this work, the first version of the S2-RUT (Sentinel-2 Radiometric Uncertainty Tool) has been made available to the community and presented in Chapter 3. This tool estimates the radiometric uncertainty associated to each pixel using as input the Top-Of-Atmosphere (TOA) reflectance images provided by ESA. The identified uncertainty contributors are combined following the Guide to the Expression of Uncertainty in Measurement (GUM) in order to provide an estimated uncertainty (BIPM, IEC et al. 2008). This combination model is further validated by comparing the results to a multivariate Monte Carlo Method (MCM) (BIPM, IEC et al. 2008). In addition, it has been studied the correlation among the different uncertainty contributions and the impact of simplifications in the combination model.

The tool software design looks for an efficient strategy to implement the uncertainty estimates. The first version does account for the most important uncertainty contributors and effectively assess their level by modelling the uncertainty levels from the metadata, external sources etc. The tool permits, for example, the assessment of the uncertainty at a desired coverage probability and the selection/deselection of the uncertainty contributors for sensitivity studies. It provides “uncertainty images” coded in a specified user format that includes pixel geolocation.

Here we describe the recent research in order to accommodate novel uncertainty contributions to the TOA reflectance uncertainty estimates in future versions of the tool. The two contributions that we explore are the radiometric impact of the spectral knowledge in Section 4.2 and the uncertainty propagation of the resampling associated to the orthorectification process in Section 4.3.

4.2 Spectral Knowledge Uncertainty

The MSI instrument is intended to measure the radiance of a specific region of the spectrum. The effect of the several components along the optical path — telescope, splitter, filter and detectors — determines the final Spectral Response Function (SRF) that is measured at each band. The “equivalent radiance”
$L_{eq}$ — and the converted reflectance factor — is defined by the spectral convolution of the scene TOA radiance $L_{TOA}$ with the instrument SRF:

$$L_{eq}(b) = \frac{\int L_{TOA}(\lambda) SRF_{ref}(\lambda) d\lambda}{\int SRF_{ref}(\lambda) d\lambda}$$  \hspace{1cm} (4.1)

SRF$_{ref}$ here is the reference SRF that associates the measurements for each band $b$. However, SRF$_{ref}$ does not represent the actual SRF of each pixel in the focal plane, at any point in time and with a perfect knowledge. That is, it has an associated uncertainty due to effects of 1) spectral non-uniformity, 2) spectral on-orbit variations and 3) spectral calibration uncertainty. The uncertainty associated to this SRF$_{ref}$ translates into a radiometric uncertainty on $L_{eq}$. Note that here the SRF$_{ref}$ is associated to the mean of the pixels’ SRF. If each pixel had an associated SRF, the effect of spectral non-uniformity would be suppressed at the expense of an enormous increase in the processing requirements (e.g. atmospheric correction).

Previous work describes an analytical process to determine the uncertainty of the SRF at any interpolated point including the covariance between the measured points (Gardner 2003). For the study here, the analytical process is more complicated since it is necessary to propagate the SRF uncertainty through the spectral convolution. The TOA spectral radiance $L_{TOA}$ cannot be easily described as an expression that resolves the function analytically. It could be possible to investigate the problem by applying a “hybrid” approach that calculates the SRF uncertainty at any interpolated point analytically and propagates the uncertainty through the integral using the MCM approach. The limitation here exist if the algorithm wants to be regularly updated. For example, by using a different interpolation method or changing its weights. In that case a full MCM approach seems a “flexible” solution that will be beneficial if consequent revisions of the code are performed.

The intention here is to introduce an implementation that permits to assess both the radiometric dispersion as a consequence of the spectral knowledge calibration and non-uniformity and the error between the mean of that dispersion and the radiance obtained using the theoretical SRF$_{ref}$. At this first implementation, the spectral effects previously described have been tested as wavelength shifts. Further iterations of the methodology should also include changes in the spectral shape. Figure 4-1 describes the steps implemented in this first implementation of the contribution:
Figure 4-1. Diagram describing the S2 spectral uncertainty assessment.

The module sets as inputs the SRF measured at instrument level for each VNIR band and a simulated TOA radiance covering the same VNIR range. There are a total of 24 SRFs at 1 nm spectral resolution that provide an estimation of the spectral non-uniformity effect across the focal plane. There is, nonetheless, a certain limitation since not all the pixels in the focal plane have been measured but only 2 SRFs are available per detector module. Pre-flight as well as post-launch verification have shown how the spectral non-uniformity is mainly produced at the edges of the detectors due to the “etching” effects (Clerc 2016). A more detailed effect of the spectral non-uniformity could be performed if rather than the SRF at an instrument level, a SRF at component level is used as an input. However, this approach has been discarded since, among other limitations, at a focal plane level, the tele-centricity effect — variations of the optical path of 1-2° across the focal plane — is not accounted. The TOA spectral radiance has been calculated using MODTRAN 5® (MODerate resolution atmospheric TRANsmission). This version includes a minimum spectral resolution of 1 cm⁻¹ — spectral sampling <0.01 nm in the VNIR region — which uses a fine description of molecular gases and scattering processes (Berk, Anderson et al. 2004). The model used for the example here simulate conditions in a desert area with a mid-latitude summer atmospheric model and a desert default albedo. Figure 4-2 illustrates the TOA spectral radiance with the S2 VNIR bands.
Figure 4-1 also introduces the MCM process (black rectangle) that is further subdivided in the different components affecting the distribution of radiance values. The spectral non-uniformity is accounted as part of the systematic budget but in the process also the spectral calibration knowledge as well as degradation effects must be modelled.

The spectral calibration knowledge is limited by the pre-flight accuracy and precision of the measurements. These have different sources as the source short-term stability, the monochromator calibration, the setup alignment etc. The first implementation of the module divides the effects into random and systematic ones. The implementation of systematic effects associates a sample out of a normal distribution for the entire band measurements whereas the random effect associates a different sample out of a normal distribution for each 1 nm measurement. Preliminary values of 0.2 nm standard deviation for the systematic effect and 0.1 nm for the random distribution have been provided.

Figure 4-1 includes a potential effect of signed systematic errors. They refer to potential variations of the SRF during the mission lifetime. Those variations are mainly produced by temperature variations and the instrument degradation. The type of filters in the S2 MSI — ion-assisted deposition interference filters — tend to have temperature wavelength shift coefficient in the range of 0.01 nm/K (Takashashi 1995, Xiong, Che et al. 2006). Thus, the thermal variations are not expected to be a major contribution. However, the degradation could introduce an important spectral variation and needs to be accounted for. A simplified approach has used here a central wavelength shift for each band. The preliminary values used are the degradation of Terra MODIS bands for the first year in-orbit (Xiong, Che et al. 2006). The closest bands to each Sentinel-2 VNIR band are associated: -0.33(B1), -0.26(B2), 0.04(B3), -0.03(B4), -0.05(B5), -0.07(B6), 0.1(B7), 0.2(B8), -0.18(B8A). These values will be used as an example but, nonetheless, a more detailed study of the potential degradation should follow and, if possible the
degradation should specify the in-band variations for bands in the lower wavelength region (Xiong, Che et al. 2006).

The interpolation of both the TOA radiance and the SRF are performed at 0.0005 nm steps. This permits sufficient resolution so that small displacements of the wavelength range are well captured. The first interpolation method will use the well-known cubic-spline interpolation. For this interpolation, the values between partition points — knots — are represented by a polynomial of third degree and the first and second derivatives of the interpolation function are continuous in all the range (Dierckx 1993). The Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) method is used to cross-compare and verify results. This interpolation method uses Hermite interpolation conditions that define function values and derivatives are specified at each nodal point. This method provides major flexibility at the cost of some problems in the continuity of the second and higher-order derivatives (Fritsch and Carlson 1980). Lagrange interpolation or the Barycentric Lagrange interpolation cannot be directly applied here due to the high variation of both the TOA radiance and the SRF (Berrut and Trefethen 2004). Section 5.2 will reuse the same TOA spectral radiance and the results will be discussed based on the interpolation type.

The distribution of the convolved TOA radiance and associated parameters for the Sentinel-2 VNIR bands are given in Figure 4-3 and Table 4-1 using the cubic spline interpolation. The same results using the PCHIP interpolation are shown in Figure 4-4 and Table 4-2. The B9 has not been included since it is traditionally used for cloud masking. Table 4-1 and Table 4-2 also include the equivalent radiance \( L_{eq}(b) \) as described in Equation 4.1 and the error with respect to the mean of the resulting distribution.

Table 4-1. Parameters associated to Figure 4-3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B1 443nm</th>
<th>B2 490nm</th>
<th>B3 560nm</th>
<th>B4 665nm</th>
<th>B5 705nm</th>
<th>B6 740nm</th>
<th>B7 783nm</th>
<th>B8 842nm</th>
<th>B8A 865nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean [W cm(^{-2})sr(^{-1})nm(^{-1})]</td>
<td>11.423</td>
<td>11.239</td>
<td>12.081</td>
<td>10.888</td>
<td>10.008</td>
<td>10.158</td>
<td>9.642</td>
<td>7.827</td>
<td>7.953</td>
</tr>
<tr>
<td>Standard Deviation [%]</td>
<td>0.22</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.05</td>
<td>0.21</td>
<td>0.31</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td>Median [W cm(^{-2})sr(^{-1})nm(^{-1})]</td>
<td>11.424</td>
<td>11.239</td>
<td>12.081</td>
<td>10.888</td>
<td>10.008</td>
<td>10.159</td>
<td>9.639</td>
<td>7.827</td>
<td>7.953</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.36</td>
<td>0.05</td>
<td>0.34</td>
<td>-0.50</td>
<td>0.04</td>
<td>-0.19</td>
<td>0.14</td>
<td>0.08</td>
<td>-0.10</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.29</td>
<td>0.08</td>
<td>0.19</td>
<td>0.03</td>
<td>0.12</td>
<td>0.08</td>
<td>-1.30</td>
<td>-0.21</td>
<td>-0.22</td>
</tr>
<tr>
<td>SRF mean [W cm(^{-2})sr(^{-1})nm(^{-1})]</td>
<td>11.460</td>
<td>11.239</td>
<td>12.081</td>
<td>10.888</td>
<td>10.007</td>
<td>10.165</td>
<td>9.647</td>
<td>7.835</td>
<td>7.956</td>
</tr>
<tr>
<td>Mean vs SRF mean [%]</td>
<td>-0.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.07</td>
<td>-0.05</td>
<td>-0.01</td>
<td>-0.05</td>
</tr>
</tbody>
</table>
Figure 4-3. TOA radiance dispersion associated to S2 spectral uncertainty using cubic spline interpolation for B1 (a), B2 (b) B3 (c) B4 (d) B5 (e) B6 (f) B7 (g) B8 (h) and B8A (i).

Table 4-2. Parameters associated to Figure 4-4.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B1 443nm</th>
<th>B2 490nm</th>
<th>B3 560nm</th>
<th>B4 665nm</th>
<th>B5 705nm</th>
<th>B6 740nm</th>
<th>B7 783nm</th>
<th>B8 842nm</th>
<th>B8A 865nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>[W cm⁻² sr⁻¹ nm⁻¹]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.22</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.05</td>
<td>0.20</td>
<td>0.31</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td>[%]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[W cm⁻² sr⁻¹ nm⁻¹]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.29</td>
<td>0.08</td>
<td>-0.25</td>
<td>-0.53</td>
<td>-0.03</td>
<td>-0.19</td>
<td>0.14</td>
<td>0.07</td>
<td>-0.07</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.13</td>
<td>-0.00</td>
<td>0.15</td>
<td>0.03</td>
<td>-0.13</td>
<td>0.25</td>
<td>-1.30</td>
<td>-0.22</td>
<td>-0.28</td>
</tr>
<tr>
<td>[W cm⁻² sr⁻¹ nm⁻¹]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean vs SRF mean [%]</td>
<td>-0.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.08</td>
<td>-0.05</td>
<td>-0.01</td>
<td>-0.05</td>
</tr>
</tbody>
</table>
Figure 4-4. TOA radiance associated to S2 spectral uncertainty results using PCHIP interpolation for B1 (a), B2 (b) B3 (c) B4 (d) B5 (e) B6 (f) B7 (g) B8 (h) and B8A (i).

The results both using the PCHIP or cubic interpolation are quite consistent with minimum variations in the measured parameters. The difference in SD and error are minimum. The comparison of these two interpolation methods indicates the low effect of ill-conditioned interpolation situation.

The standard deviation is higher for B1 due to the high atmospheric contribution in that area combined with a 20 nm bandwidth. The B6 also introduces a relative higher standard deviation due to its narrow bandwidth of only 15 nm and its strategic situation at the re-edge. Nonetheless, for a desert site this effect is minimized. The distribution shows also an important standard deviation for B7 but, in this case, it is produced by a bi-modal distribution with approximately 0.6% difference mode-to-mode. That is the consequence of the filter manufacturing process. There are filter slices for each of the 12 detector modules that might be manufactured using different “wafers”. This is a common situation with this type of push-broom instruments as occurs, for example, with Landsat-8 Operational Land Imager (OLI) (Barsi, Lee et al. 2014). The example here shows a radiance distribution. When introducing an equalization, this could be compensated and only the residual in the reflectance factor should be taken into account (see Chapter 3).
The error introduced between the equivalent radiance $L_{eq}$ and the mean of the distribution is generally negligible except for B1 when goes over 0.3 %. This is largely the consequence of a higher degradation rate and the narrow bandwidth of this band.

Therefore, for a cross-calibration over desert areas the uncertainty introduced by the spectral calibration knowledge must be accounted in the budget but the impact is expected to be relatively small compared to other uncertainty sources (Chander, Helder et al. 2013). However, the systematic effects of non-uniformity and degradation might be considerable and minimized for specific bands. The spectral degradation of S2 is complicated to assess since there is no device on-board to track the spectral response during the mission life-time — as the Spectroradiometric Calibration Assembly (SRCA) on-board the Moderate Resolution Imaging Spectroradiometer (MODIS) mission (Xiong, Che et al. 2006) — and the degradation rate becomes unpredictable. Tests with different degradation rates on B1 showed that the error increases with the degradation rate increase but also the standard deviation as a result of changing the spectral region. This error is much lower for B2 since, although the degradation is similar to B1, its bandwidth is much larger. For the systematic effect introduced by the non-uniformity, the use of specific per-module SRF can introduce a substantial benefit in some bands as explained for B7 above.

For an operational integration of this uncertainty contribution, a look-up-table (LUT) could be generated by applying the same method to different TOA radiances. For this site and most of the bands, the output has been shown to be close to a normal distribution. Thus, it would be possible that in several situations, a theoretical approach can be applied in order to propagate the spectral uncertainty into the TOA radiance uncertainty. For other bands as B4, the distribution cannot be inferred as a normal one and a standard deviation cannot be associated with a probability area (see Section 1.3.2). In that case, a direct measurement of the distribution area can provide an interval range (i.e. uncertainty) associated to a specific probability coverage. The example here considers a desert area with a typical application to the cross-calibration or instrument monitoring. It is expected that vegetation or coastal areas the standard deviation values considerably increase for some bands. In addition, further work could include a more detailed representation of each one of the spectral uncertainty contributions to better understand the random and systematic effects. The spectral uncertainty should also include a certain knowledge on the SRF (not only the wavelength). In addition, further interpolation methods could be included. For this first exercise, the PCHIP and cubic spline interpolation were used.

For the integration in the S2-RUT, the standard deviation and error should be considered. The first one can be added in quadrature with the rest of the L1 contributors whereas the second one should be added linearly to the expanded uncertainty — the expanded uncertainty results from the multiplication of the standard uncertainty by the desired coverage factor $k$ — if no correction is introduced (BIPM, IEC et al. 2008). The development of the S2-RUTv1 already includes this situation for some uncorrected
effects as the *diffuser degradation* and the *out-of-field stray-light effect — systematic part* (Gorroño, Fomferra et al. 2016).

### 4.3 Propagation of TOA radiometric uncertainty during orthorectification process

The first version of the S2-RUT provides the radiometric uncertainty at a pixel level but does not include the effect of the uncertainty propagation over the radiometric resampling (see Chapter 3). The effect of such a resampling will generally introduce a reduction of the uncertainty level due to the random component of the uncertainty of each pixel in the resampling grid.

The first step in an orthorectification process is a geometric transformation that determines the position of the target point in the focal plane image. The target point in Sentinel-2 L1C product is the result of applying an Earth model in Universal Transverse Mercator (UTM) coordinates that includes a Digital Elevation Model (DEM). In order to determine its value, a radiometric interpolation between the focal plane points is performed. In the Sentinel-2 L1 processing, the product before orthorectification is named as L1B (see Section 3.2). The process is illustrated in Figure 4-5.

![Figure 4-5. Resampling process of the S2 L1C products (reproduced from (ESA 2017)).](image)

In Chapter 3, the combined standard uncertainty using the GUM was compared to the MCM approach to determine the validity of the combination. The MCM approach models most of the L1B uncertainty contributions as either normal or rectangular distributions and propagates them through the L1B processing chain. The validation compares the $k=1$ uncertainty produced by the methods. The resulting major limitation for the uncertainty assessment was found to be produced by the Analog-to-Digital Converter (ADC) at the low radiances. At these low radiance values, the ADC quantization becomes unstable and no longer can be modelled as a rectangular distribution (Carbone and Petri 1998). In that
scenario, the central limit theorem cannot be applied and the combined standard uncertainty is not applicable (BIPM, IEC et al. 2008).

Selecting a resampling method is a complex decision that requires the evaluation of multiple criteria. Some of them are the processing time and the accuracy of the method. Here we describe a further criterion that describes the impact on the uncertainty propagation. That is, we want to study how the radiometric uncertainty of the focal plane measurements propagates through the different radiometric interpolation schemes down to a resampled image.

The nearest-neighbour method determines the resampled pixel by assigning the value of the closest pixel in the focal plane image. This method has low computational requirements and propagates without changes to the uncertainty distribution from the focal plane to the resampled images. However, this comes at the expense of a low accuracy for non-uniform scenes.

Most of the resampling methods introduce a certain change in the radiance distribution when propagated from the focal plane to the resampled image. The implementation here propagates the resulting values of the MCM approach from the focal plane measurements to the resampled image by using three different interpolation methods. The first one is the bi-linear method which uses a $2 \times 2$ grid. The second one is the cubic convolution. The method uses a cubic interpolation through the kernel that convolves with the original measurements (Keys 1981). The last method are the B-splines interpolation that is implemented as the nominal resampling method for the S2 L1C products (Hsieh and Andrews 1978, ESA 2017). Both cubic convolution and B-splines require a $4 \times 4$ kernel grid that produces a smoother image at the expense of a more demanding computational requirement (Hsieh and Andrews 1978, Keys 1981).

Figure 4-6 shows the results for the three different interpolation schemes presented above. The panels (a), (b) and (c) assume that the L1B uncertainty distribution are fully random and, thus, the samples in the grid are uncorrelated. The panels (d), (e) and (f) model the uncertainty associated to the absolute calibration coefficient $A(b)$ (see Chapter 3) as fully correlated to all the L1B pixels in the grid. The results were obtained for a total of $20 \times 20$ positions for a kernel with a constant radiance of 30 Wm$^{-2}$sr$^{-1}$μm$^{-1}$. These positions refer to the potential place where the interpolated pixel lies in between the inner 4 focal plane pixels of the kernel. The four corners of the figure represent these 4 pixels with coordinates $(0, 0), (0, 1), (1, 0), and (1, 1)$.
Figure 4-6. L1C uncertainty propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels of the interpolation kernel. The settings for each panel are full random L1B uncertainty using (a) the bi-linear interpolation, (b) cubic convolution and (c) B-splines and fully correlated absolute calibration coefficient uncertainty $A(b)$ for the L1B uncertainty using (d) the bi-linear interpolation, (e) cubic convolution and (f) B-splines.

The results show the important variation on the propagated uncertainty depending on the collocation between the orthorectified and the focal plane grids. Table 4-3 summarizes the main parameters that describe the results of Figure 4-6.

Table 4-3. Parameters associated to the results in Figure 4-6.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Uniform and full random</th>
<th>Uniform and $A(b)$ correlated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bilinear</td>
<td>Cubic</td>
</tr>
<tr>
<td>Mean [%]</td>
<td>1.5631</td>
<td>1.8970</td>
</tr>
<tr>
<td>Standard deviation [%]</td>
<td>0.2603</td>
<td>0.2130</td>
</tr>
<tr>
<td>Minimum</td>
<td>1.1491</td>
<td>1.4750</td>
</tr>
<tr>
<td>Maximum</td>
<td>2.3165</td>
<td>2.3333</td>
</tr>
</tbody>
</table>

The lowest uncertainty results are obtained for the bi-linear interpolation. In all cases, it obtains the lowest mean and minimum values. This interpolation does provide an improved randomisation at the expense of a low “smoothness”. The B-spline obtains a better mean value for the $20 \times 20$ positions than the cubic convolution. Although the minimum value for the two methods is quite close, the distance in the mean is slightly larger. This is produced since the cubic convolution introduces a non-stable interpolation close to the kernel points that introduce slightly higher uncertainty than the focal plane.
measurements — see the maximum is 2.3333 % compared to 2.3165 % for the rest of the cases. The application of the cubic convolution in Gardner (2003) showed the same problematic at specific areas. The differences between the fully uncorrelated kernel points with the case where \( A(b) \) is fully correlated demonstrates the importance of the correlation in the uncertainty propagation. This is a topic that will be described in detail in Chapter 6. Table 4-3 reports differences in the mean value of approximately 0.1 % for the cubic convolution and B-spline interpolation and as large as 0.2 % for the bi-linear interpolation.

For its implementation in future versions of the S2-RUT, the uncertainty estimates at the L1C — orthorectified images — will be limited to assess the position in the grid and; thus the effect of the uncertainty propagation. A simple implementation would introduce a “reduction factor” of the L1B uncertainty estimates based on an analysis as above. That “reduction factor” would be based in an analysis as the one demonstrated above that compares the kernel uncertainty estimates with the potential uncertainty at different positions on the grid.

The selected approach will largely depend on the user requirements and the application. That is, if the user application involves the \( a\)-\( posteriori \) binning of the pixels, a large fraction of the uncorrelated uncertainty contributions will be reduced. In that case, the resampling uncertainty propagation will have a smaller impact to the user (see Section 6.6).

The best implementation of the uncertainty propagation would imply the assessment of the uncertainty at the L1B products and the propagation through the resampling by using pre-calculated Look-Up-Tables (LUT). Similarly to the method presented above, the LUTs can be populated for several radiance values in the grid and positions in the pixel. However, for a kernel of \( 4 \times 4 \) L1B radiance values, the number of cases could be sensibly high and a minimization of cases should be pursued. Alternatively, the fitting of a surface equation to the results in Figure 4-6 could largely simplify the approach.

Figure 4-7 and Table 4-4 show the differences in the interpolated values between the different resampling methods for a case where the radiance kernel is non-uniform. As previously commented, not only the resampling introduces a change in the uncertainty distribution but also includes an accuracy of the resampled value. That is, how well the interpolation represents the “real” scene variation. For the uniform scene previously studied, the three interpolation obviously have a perfect accuracy. In that case, a bi-linear interpolation requires less processing and can potentially better reduce the radiometric uncertainty. However, most of the scenes will not be close or near-close to a uniform TOA spectral radiance.
Figure 4-7 shows the large differences that can be obtained by using different resampling methods. In that case it is difficult to determine which the best method is since assumptions of the scene inter-pixel variations are needed. Table 4-4 quantifies these differences by calculating the root mean squared error (RMSE) between the different interpolation methods. The RMSE is considerably high between any of the three methods. If a radiance of 30 Wm⁻²sr⁻¹μm⁻¹ is taken as reference, the RMSE between the bi-linear method and the cubic convolution is approximately 5% of the reference radiance (1.6772 Wm⁻²sr⁻¹μm⁻¹). Slightly lower level for the RMSE between the bi-linear and B-spline methods (1.3559 Wm⁻²sr⁻¹μm⁻¹) whereas the RMSE is around the 2% level between the cubic convolution and B-spline methods (0.6969 Wm⁻²sr⁻¹μm⁻¹).

Table 4-4. RMSE error between the interpolated radiance of panels (a), (b) and (c) from Figure 10.

<table>
<thead>
<tr>
<th>RMSE [Wm⁻²sr⁻¹μm⁻¹]</th>
<th>Bilinear</th>
<th>Cubic</th>
<th>B-spline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinear</td>
<td>0.0000</td>
<td>1.6773</td>
<td>1.3559</td>
</tr>
<tr>
<td>Cubic</td>
<td>1.6773</td>
<td>0.0000</td>
<td>0.6969</td>
</tr>
<tr>
<td>B-spline</td>
<td>1.3559</td>
<td>0.6969</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Figure 4-7. L1C radiance propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels introducing a fully correlated absolute calibration coefficient uncertainty A(b) for the L1B measurements and using the bi-linear interpolation (a), cubic convolution (b) and B-splines (c). L1C uncertainty propagated at different positions Across-Track (ACT) and Along-Track (ALT) between neighbour L1B pixels of the interpolation kernel introducing a fully correlated absolute calibration coefficient uncertainty A(b) for the L1B measurements using the bi-linear interpolation (d), cubic convolution (e) and B-splines (f). The values of radiance in the L1B interpolation kernel are (values from left to right and up to down in a 4x4 kernel): [[21, 20, 10, 14], [15, 30, 24, 50], [10, 12, 30, 60], [12, 45, 36, 15]] Wm⁻²sr⁻¹μm⁻¹.
This is a first implementation of a complex topic in the estimation of radiometric uncertainty. Further work could include the impact of the knowledge on the geometric transformation and how this position grid knowledge influences the interpolated radiance as well as the propagated uncertainty. In addition, the resampling can be performed using a pre-determined gridding. In that case, the gridding precision should be also assessed. Finally, the effect of the resampling at the image boundaries e.g. by mirroring should be carefully analyzed.

4.4 Conclusions and further work

The work here describes the research of methodologies that estimate the spectral response uncertainty and the orthorectification uncertainty propagation. These two contributors have not been included in the uncertainty analysis presented in Chapter 3. They represent good examples on how the estimation of TOA radiometric uncertainty can be very challenging. The two uncertainty contributions — both spectral response and orthorectification — are present in most EO optical missions. These first studies show the need to account for the measurement scene properties and the non-linear properties of some transformations. Thus, further studies of TOA radiometric uncertainty should focus on providing simple and, at the same time, rigorous solutions for these two issues.

The spectral response module is capable of determining the effect of the spectral calibration knowledge, spectral non-uniformity and spectral degradation in an integrated software. The spectral calibration knowledge has been further separated into a random and systematic component and the components propagated using a MCM approach. The example used here represents a desert TOA radiance that could be typically used for instrument cross-calibration. In that case, the uncertainty introduced by the spectral calibration knowledge must be accounted in the budget but the impact is expected to be relatively small compared to other uncertainty sources (see Table 4-1). The degradation effect has shown a considerable error on B1. For the systematic effect of the non-uniformity, the use of specific per-module SRF can produce a substantial benefit in certain bands in terms of radiance as, for example, B7. However, this is likely to be reduced when the S2 MSI equalisation is performed (see Section 3.2).

The orthorectification module propagates the L1B uncertainty up to the L1C using a MCM approach in a kernel of $4 \times 4$ L1B radiance values for the cubic convolution and B-spline interpolation ($2 \times 2$ kernel for the bi-linear case). The results show the important variation on the propagated uncertainty depending on the collocation between the orthorectified and the focal plane grids. The results also show how the bi-linear interpolation effectively better reduces the radiometric uncertainty at L1C. The uncertainty reduction rate largely depends on the level of correlation between the L1B kernel points. For non-uniform scenes, the bi-linear method still better reduces the L1C uncertainty but important differences in the interpolated radiance between the three methods were found. To fully account the effect of the orthorectification, the estimation should be propagated from the L1B product. However, if
that is not possible, a rough estimation, LUT or uncertainty “surfaces” can be applied using the L1C products as an input.

Further work is needed for a full assessment and the integration of these uncertainty contributions in future versions of the S2-RUT tool. The example here primarily considers a desert area with a typical application to the cross-calibration or instrument monitoring. For an operational integration, a look-up-table (LUT) or uncertainty “surfaces” could be generated for both uncertainty contributions by applying the same method to different TOA radiance scenes. The integration of both uncertainty contributors in the S2-RUT uncertainty combination should consider both the uncertainty and the error. As previously commented, the first one can be added in quadrature with the rest of the L1 contributors whereas the second one should be added linearly to the expanded uncertainty if no correction is introduced (BIPM, IEC et al. 2008).

The spectral uncertainty module could include a detailed representation of each one of the spectral calibration uncertainty contributions and other interpolation methods. The spectral uncertainty could also include a certain knowledge on the SRF (not only the wavelength) and a detailed analysis of the potential degradation scenarios.

The orthorectification module improvement should be in an in-depth analysis of the covariance between the L1B pixels (see Chapter 6). This would determine the level of uncertainty improvement at the L1C pixels. In addition, further work must follow in the identification of the potential accuracy introduced by the interpolation method.
Chapter 5.

Uncertainty for sensor-to-sensor cross-calibration

5.1 Introduction

TRUTHS (Traceable Radiometry Underpinning Terrestrial and Helio Studies), is a proposed satellite mission led by the National Physical Laboratory (NPL), UK. This mission is designed to have sufficient accuracy to allow the unequivocal detection of trends, from a background of natural variability, in a number of key indicators of climate change in the shortest time possible, allowing verification of climate forecast models on decadal timescales (Fox, Kaiser-Weiss et al. 2011). This would be achieved by establishing a fiducial reference data set of spectrally resolved incoming and outgoing solar radiation. In terms of Earth viewing radiance, the characteristics of this data set are: spectrally-resolved — 5-10 nm Full Width Half Maximum (FWHM)) — Earth radiances, continuously sampled (spectrally and spatially) with a Ground Instantaneous Field Of View (GIFOV) of approximately 50 m over the 320-2400 nm spectral range, and the corresponding solar spectrally-resolved irradiance; both with SI-traceable radiometric uncertainties of <0.3% (Fox, Kaiser-Weiss et al. 2011). These fiducial data sets establish a high accuracy benchmark of the Earth’s spectral radiation budget in the solar spectral domain in a similar manner to its US-proposed sister mission Climate Absolute Radiance and Refractivity Observatory (CLARREO) against which future change can be detected (Wielicki, Young et al. 2013). The chosen spectral and spatial resolutions are optimum to allow the data sets to be utilised to retrieve many Essential Climate Variables (ECVs) — as defined by Global Climate Observing System (GCOS) — and facilitate detailed analysis of attribution effects (GCOS 2010) and the Earth system’s cycles and processes.

It is thus not surprising that TRUTHS’s observational specifications — spatially and spectrally — match/allow reconstruction of many of the current, and planned, solar domain EO sensors, such as Landsat-8 (L8) Operational Land Imager (OLI). However, the addition of high SI-traceable radiometric accuracy in the reference sensor, maintained throughout the mission lifetime, also provides a powerful opportunity to cross-calibrate other sensors through co-incident viewing of stable target scenes and in particular, the radiometric characterisation of Pseudo Invariant Calibration Sites (PICS). For target sensors, such as Copernicus Sentinel-2 (S2) Multispectral Imager (MSI) and Sentinel-3 (S3) Ocean and
Land Colour Instrument (OLCI), TRUTHS allows not only an assessment of performance but also a calibration upgrade towards that needed by many climate studies, and thus leads to the prospect of a space-based climate and calibration observatory as requested by the international community (Dowell, Lecomte et al. 2013).

The existing on-board calibration systems of many sensors such as Sentinel-2 and 3 have significant complementary merit, allowing assessment of any short term performance variation of the sensor over its full orbital path and between reference calibrations. In these cases, TRUTHS provides the in-flight anchor to SI units and the prospect of a regular update of the on-board monitoring systems. However, for sensors whose primary objectives do not warrant an on-board calibration system, such as the UK-DMC (Disaster Monitoring Constellation) series, similar cross-calibration activities would provide the means to achieve radiometric traceability, broadening the scope of application of such sensors, even to the point where these sensors could contribute towards climate studies and services. Following this logic, a constellation of new generation, low-cost Cube-/Nano-Sats could be envisaged, also contributing to the global observing system, radiometrically-anchored to a reference sensor such as TRUTHS.

The ideal configuration for vicarious target inter-calibration is that the two instruments should make matched measurements viewing the same target at the same time; with the same spatial and spectral responses at the same viewing geometry. Since these idealized conditions never occur in reality, there will always be some additional compensatory steps needed to allow comparison of the two instruments. The accuracy achievable by the target sensor via the inter-sensor cross-calibration is ultimately limited by the reference sensor accuracy and the inability to fully account for the differences from the ideal comparison conditions. These differences include the instrument spectral response, target site spectral signature and the radiometric properties of the selected target site for the calibration process, including effects of solar illumination and sensor view angles and any variance in the atmosphere transmittance between the observations by the two sensors. Similar conditions apply even when the reference sensor measurements are used only as an input for the radiometric characterisation of PICS. In that situation, the longer term temporal radiometric properties of the site and its atmosphere become relevant factors.

In a recent study by Chander, Helder et al. (2013), the uncertainty introduced by the main effects inherent in the cross-calibration transfer using a calibration target site was assessed to fall well below an uncertainty level of 0.3% ($k = 1$) with the exception of a spectral shift in SBAF. For this spectral shift assessment, the study found larger errors in some bands of the Moderate Resolution Imaging Spectrometers (MODIS) sensor. Nonetheless, the study considered tolerances/shifts of 5 nm in the MODIS filters. These shifts can be considered as a highly pessimistic scenario or a worst-case assessment. As a result, the uncertainty associated with the calibration of the reference sensor is now often the dominant component in the final uncertainty achieved for the test sensor.
The calibration accuracy of sensors measuring in the visible/near infrared (VNIR) and shortwave infrared (SWIR) spectral regions increased notably in the last decades. MODIS on board the Terra and Aqua satellites, or the recently launched S3 OLCI, have requirements for calibration accuracy of below 2% ($k = 1$) relative to the sun (Xiong and Barnes 2006, Donlon, Berruti et al. 2012). Instruments such as the Clouds and the Earth’s Radiant Energy System (CERES) have even more stringent calibration accuracy requirements — calibration accuracy below 1% ($k = 1$) — have highlighted the need for a reliable inter-calibration with an instrument like TRUTHS or CLARREO to overcome the data gap between the CERES mission instruments, to maintain the demanding stability requirements needed for climate (Loeb, Manalo-Smith et al. 2016). Even if these well-calibrated instruments are used for cross-calibration their accuracy levels remain the dominant contribution to the total uncertainty in the cross-calibration process compared to the ones described in Chander, Helder et al. (2013). Thus, the possibility of a reference instrument like TRUTHS or CLARREO with a radiometric uncertainty below 0.3% ($k = 2$) would be of a large benefit to reduce the total uncertainty in a cross-calibration over PICS.

This chapter addresses the uncertainty contributions affecting typical CEOS WGCV recommended land-based reference sites in its use for cross-calibration of satellite imagers in the three main domains: spectral, spatial, and temporal. The aims of this chapter are to: 1) evaluate the “inherent” uncertainty contributions with case studies 2) set up a suite of tools and methodologies useful for the exploitation and design of missions like TRUTHS or CLARREO, and 3) define the uncertainty contributions in a cross-calibration using rigorous metrology. Spectral, spatial and temporal contributors are all considered separately in Sections 5.2, 5.3, and 5.4 respectively.

For the latter point, the uncertainty propagation is based on the Monte-Carlo Method (MCM) as described in Supplement 1 to the Guide to the Expression of Uncertainty in Measurement (GUM) (BIPM, IEC et al. 2008), the use of which is explicitly encouraged in the Quality Assurance Framework for Earth Observation (QA4EO) (http://www.QA4EO.org). Thus, the cross-calibration uncertainty estimates are presented in terms of a probability distribution function (pdf) of the associated parameters. The uncertainty is reported as the interval around the best estimate that approximates a coverage of 68.27% (which is expressed as $k = 1$). The coverage factor, $k$, is a numerical factor that multiplies the combined standard uncertainty in order to specify the fraction of the probability distribution that the uncertainty represents (see Section 1.3.2).

The MCM uncertainty propagation is a well-described technique which has historically been limited by the computing resources available. The rapid development of computing capabilities in recent years has made it more accessible to the EO community. The quantification and analysis of the uncertainty contributors developed as a software tool here require access to a large amount of memory and CPU time and have thus utilised the UK’s JASMIN supercomputer facility (Lawrence, Bennett et al. 2013). The high-performance of the computer nodes permits the management of large quantities of memory,
while a cluster of virtual and physical machines sharing a dedicated network, permits the parallel processing of the MCM algorithm.

5.2 Uncertainty assessment: Spectral Domain

5.2.1 Spectral domain: methodology

This section assesses the effect of spectral mismatch between a TRUTHS-like sensor and a target sensor (Sentinel-2 MSI) in the context of the chosen test-site’s spectral properties. Specifically, it studies the capacity of a TRUTHS-like sensor to derive a continuous Top-Of-Atmosphere (TOA) reflectance factor/radiance spectrum and the effect that it introduces in a cross-calibration with a sensor like S2 MSI. The effect of such differences between the band spectral response functions (SRFs) for the reference and target sensors is traditionally compensated for using the Spectral Band Adjustment Factor (SBAF), which is calculated from the known SRFs for each sensor and the spectral radiance of the test site being measured (Chander, Mishra et al. 2013). Here, this approach has been adapted to understand the achievable accuracy of the TRUTHS sensor in a cross-calibration with a target sensor. The process is similar to that applied in Green (1998) and is illustrated in Figure 5-1: a reference TOA radiance spectrum is generated and convolved with the spectral bands of the TRUTHS sensor, the values from each band are binned as required, then used to reconstruct a hyperspectral curve via interpolation, and this reconstructed curve is then convolved with the target sensor bands.

![Illustrative method of the TOA TRUTHS spectral profile generation. The red stars are the measurements at the native spectrometer bands and the green stars are the result of the merging to a design specified bin. The merged measurements are sampled at a specified interval to obtain the reconstructed TOA as measured by TRUTHS ('TRUTHS TOA @5x10^-4').](image)

Figure 5-1 Illustrative method of the TOA TRUTHS spectral profile generation. The red stars are the measurements at the native³ spectrometer bands and the green stars are the result of the merging to a design specified bin. The merged measurements are sampled at a specified interval to obtain the reconstructed TOA as measured by TRUTHS ('TRUTHS TOA @5x10^-4').

³ We used the word “native” to indicate the measurement, spectral resolution and spectral sampling prior to any binning or post-processing.
Rather than using a specific simulation, the whole range of potential cases are studied to derive the uncertainty introduced in the spectral dimension. That means that the simulations cover different wavelength positions of the TRUTHS sensor SRF, the reconstruction is set using different interpolation techniques in Section 5.2.2, and the central wavelength and bandwidth of each spectral band have an associated uncertainty in Section 5.2.3.

In order to study the spectral error introduced, a simulated TOA radiance spectral profile was generated and used as a reference. The simulation was initiated with the following conditions: Viewing Zenith Angle (VZA) = 4.602º, Sun Zenith Angle (SZA) = 21.443º, Relative Azimuth Angle (RAA) = 179.223º for day number 173 (summer solstice), and time = 8:54:53 GMT — similar to a particular Landsat 8 OLI overpass of Lybia-4. The spectral resolution was set to the highest MODTRANv5.3.3 spectral resolution of 0.1 cm\(^{-1}\) in order to extract the maximum information (Berk, Anderson et al. 2005).

The TOA spectral radiance from MODTRAN was further interpolated to 0.0005 nm using linear interpolation over the VNIR and SWIR range. Such a low resolution — 0.002 nm at around 450 nm — can capture reasonably detailed information relating to atmospheric and solar features. The sampling is more than twice the original MODTRAN output and uses a linear interpolation meaning that oversampling does not alter the original absorption line structure. The requirement for such a fine resolution derives from the possibility of describing the instrument spectral knowledge uncertainty as a distribution of errors in Section 5.2.3. In addition, the benefit of this low resolution for the impact of sampling/resolution will be discussed in Section 5.2.2.

The sampling and binning are set to values representative of the preliminary design of the TRUTHS satellite-borne imaging spectrometer, see Figure 5-2.

![Figure 5-2](image)

**Figure 5-2** (Left) Preliminary design of the TRUTHS Earth Imager spectrometer and (right) the translation of the native spectral sampling design in the instrument SRF.

The proposed native sampling and resolution of the instrument is used to generate triangular response functions, to model the real TRUTHS spectrometer response. The very low aberration of the candidate
spectrometer means the SRF actually achieved is likely to be very close to this idealised triangular response; ultimately, the instrument spectral line characterisation will determine the precise shape. The preliminary optical design uses matching slit and pixel width dimensions; hence it is possible to approximate the pixel spectral bandwidth by the native sampling interval.

The TRUTHS SRF is then convolved with the site TOA spectral radiances, to produce an instrument “as-measured” TOA radiance. The instrument response is further binned — the binning is set by design to achieve the optimum spectral sampling and Signal to Noise Ratio (SNR) — to emulate the TOA measurement of the TRUTHS bands. The process initiates at around 400 nm and iteratively moves up to the SWIR range stopping at around 2500 nm.

In order to derive a continuous TOA radiance spectrum a fitting or an interpolation can be used. Without any further information and given an extremely high accuracy of the sensor, an interpolation is used here. Nonetheless, other sophisticated fitting methods as described in McCorkel, Thome et al. (2013) have been successfully applied and could, if prior spectral shape is accurate, reveal further information of the calibration site.

Figure 5-3 provides the “as measured” TOA radiance by TRUTHS around each of the S2A bands. The figure also overlays the MODTRANv5.3.3 TOA radiance used as a reference.
Figure 5-3 TOA radiance as generated by MODTRANv5, resulting measurements of TOA radiance as measured by TRUTHS Earth imager and the Sentinel 2 VNIR and SWIR bands.

5.2.2 Spectral domain: systematic sampling/ resolution error results

Once the TOA radiance “as-measured” is obtained, the measurements of the TRUTHS bands are interpolated at 0.0005 nm resolution to match the original TOA spectral distribution generated using MODTRANv5.3.3 as described previously. The “true” radiance spectrum, generated using the MODTRAN reference, and this “reconstructed” spectrum are then used to study the impact on the
cross-calibration of the sampling/resolution of the TRUTHS bands by convolving each with the SRFs for S2 MSI. The resulting difference between the convolved values for the “true” and “reconstructed” spectra are shown in Figure 5-4. This process is undertaken for all the S2A bands except bands 9 and 10. These are centred on the atmospheric water absorption bands but are not intended to provide accurate radiometric measurements of the water vapour level, but rather the detection of clouds in the scenes (Drusch, Del Bello et al. 2012).

The starting wavelength of the TRUTHS bands can be set at a specific spectral position — referred here as wavelength position shift from the starting wavelength of 410 nm — to simulate all the potential positions of the spectrometer bands. That is, it represents the alignment position of the detector array at the focal plane. This has an impact on the results where spectral structure is found in the observed scene. The results of simulations in Figure 5-4 show the error for all the positions in between two native spectral bands separated by the maximum sampling provided in Figure 5-1. Several interpolation method and combinations have been used in the reconstruction of the TOA radiance from the TRUTHS binned measurements and the S2A bands. Cubic spline interpolation represents the values between partition points — knots — by a polynomial of third degree with first and second derivatives of the interpolation function continuous at all points of the interpolation range (Dierckx 1993). The Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) method uses Hermite interpolation conditions that define function values and derivatives at each nodal point (Fritsch and Carlson 1980). The linear interpolation does not provide continuity in the derivatives at the interpolation knots. Thus, this selection criteria represents the three potential levels of continuity at the knots and all the potential combinations are presented in Figure 5-4.

The error remains at the 0.1%-level for all the studied bands with the exception of B1 and B6 for which the error raises up to 0.5% level and B5 which increases up to 0.2%. Bands B5 and B6 have narrower bandwidths and are placed in a spectral region largely defined by atmospheric absorption of water (720.5 nm) and oxygen (687.5 nm and 761 nm) as seen in Figure 5-2. For B6 it is shown how the use of an interpolation with continuity at the knots, provides an improvement up to 0.2 % due to the better fitting of these atmospheric absorptions. B1 is not affected by large atmospheric absorption peaks but by a large solar irradiance variation. The spectral region in 410-440 nm is one of the most challenging regions in the solar irradiance models as described in Thuillier, Hersé et al. (2003) and it contains a large spectral irradiance variability.
Figure 5-4 TOA error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands (above, B1-B8A) and SWIR bands (below B11 & 12) due to the TRUTHS sampling bands and preliminary resolution of the detector bands. The errors are plotted for different types of interpolation to reconstruct the TOA radiance and Sentinel 2 bands.

The bands show a very noticeable oscillation that remains constant for the band B5 onwards and for the different interpolation methods used here. This oscillation has an impact below 0.1 % peak-to-peak as consequence of the movement of the bands position (wavelength shift), in combination with rapid variations of TOA radiance. In addition, the sampling/resolution are not fixed values in the range of shift but have a small variation — see Figure 5-2 — and produce this interference pattern in the image.
As part of the validation process, the periodicity was found to be proportional to the slope of the sampling increase across the S2 bands. Furthermore, if the simulation keeps the same sampling/resolution across the band convolution, the systematic error remains at the same amplitude but the period is equal to the applied sampling/resolution. For the lower bands, B1 and B2, a clear periodicity of 4-5 nm can be found. At these bands, several native bands are binned and the period is dominated by the sampling period after binning. The discontinuity due to a linear interpolation in the binning requirement — see Figure 5-2 — introduces discontinuities for the lower bands B1 and B2.

Differences between all possible interpolations are small with maximum variations of around 0.2% peak-to-peak for specific bands and spectral regions such as for B1 and B6. The larger difference occurs between the TOA radiance interpolation using cubic spline and linear one whereas the PCHIP remains a middle case. For all bands but B4, the cubic spline interpolation provides the minimum error since it is able to adapt to the rapid variations of the TOA radiance. At 650 nm, there is a discontinuity of binning requirements from 2 native spectral bands to just 1. The differences between interpolations are mainly generated by the type of interpolation of TOA spectral radiance whereas the type of interpolation of the S2 SRF has a negligible impact. The assumption is that the S2 SRF sampled at 1 nm captures sufficient information regarding the spectral variations across the spectral band. The majority of the variations in the S2 spectral band — see Figure 5-4 — have a period of variation greater than twice the SRF sampling. The impact of the S2 SRF interpolation represented a small variation below 0.05%. In order to visualise the impact, the previous error of B1 in Figure 5-4 has been normalised by the mean at each wavelength shift position and plotted in Figure 5-5.

![Figure 5-5](image)

**Figure 5-5** Difference between interpolations for B1 in Figure 5-4. The error at each wavelength shift position has been normalised by the mean. The legend of the plot is equivalent to Figure 5-4

Figure 5-4 has shown that the error depends on where the focal plane array is situated. These errors are inter-dependent in between bands of the VNIR focal plane and in between bands of the SWIR focal plane. That is, the “shift” must be applied to all the bands of the focal plane since it is the whole array...
position that matters. The approach here is calculating the root mean square (RMS) error between the different bands error and for all positions of the array as follows:

\[
rm_{s} = \sqrt{\sum_{i} \left( \frac{L_{S2TRUTHS}[i,s] - L_{S2MODTRAN}[i,s]}{L_{S2MODTRAN}[i,s]} \right)^2}
\] (5.1)

\[
rm_{s} = \sqrt{\sum_{i} \left( \frac{L_{S2TRUTHS}[i,s] - L_{S2MODTRAN}[i,s]}{L_{S2MODTRAN}[i,s]} \right)^2}
\] (5.2)

where \( s \) is the position in the array and \( i \) are the S2 bands. The term \( L_{S2TRUTHS} \) refers to the TOA spectral radiance as measured by TRUTHS and convolved with the S2 band whereas \( L_{S2MODTRAN} \) refers to the TOA spectral radiance using the MODTRAN reference and convolved to the S2 band.

The result of the rms calculation produces a set of values for the different potential positions of the array. The results in Figure 5-6 show the spectral sampling error for each of the bands in the case of best and worst case rms position.

![Figure 5-6 Spectral sampling error for each S2 band considered in the cases of minimum, maximum rms array position error for a S2 SRF linear interpolation and TOA radiance cubic spline interpolation (left) and linear interpolation (right).](image)

The results in Figure 5-6 describe an error in the 0.1% range for all types of interpolations and focal plane alignment with the exception of B1, B5 and B6 bands. Potential methods to reduce the error introduced in these bands are separately studied in Sections 5.2.4 and 5.2.6 for B1, B5 and B6 respectively.

The criteria used here serves as an example of the design process and can be adapted to other scenarios. For example, the design of TRUTHS focal plane proposes a certain level of overlapping between the
VNIR and SWIR focal plane. In that context, the method described in Section 5.2.1 might need to be adjusted and it is likely that the error for bands like B8 and B8A will reduce further.

### 5.2.3 Spectral domain: spectral knowledge uncertainty

The effect relating to "knowledge" (centre wavelength/bandwidth) of the TRUTHS SRF has been studied by reconstructing the TOA radiance spectrum \( n \) times with different centre wavelength and/or bandwidth each time, before convolving it with the S2 bands. In a simplified model, intended to model likely instrumental errors, the central wavelength and bandwidth of the TRUTHS triangular bands (see Figure 5-2) are modelled as a normal distribution, with the wavelength shift constant in sign and magnitude for all wavelengths across the spectrum. This is considered as an approximation that works under the assumption that the knowledge of the smile correction and/or spectral calibration is largely correlated across each of the S2 SRF bandpass. This simulation results in a dispersion of TOA spectral radiance values for the S2 band convolution, dependant on the structure in the local TOA spectral radiance spectrum.

Figure 5-7 presents the resulting distributions for all the S2 bands with an associated central wavelength and bandwidth knowledge uncertainty of 0.2 \( \text{nm} \) \((k = 1)\) convolved with the TOA spectral radiance 10,000 times. Since the normal distribution is by definition infinite, it was decided to truncate to a maximum of 10 times the standard deviation in order to avoid out-of-range values. The values applied here of central wavelength and bandwidth knowledge can be considered as a conservative figure since previous in-flight spectral calibration exercises have proven to keep the spectral knowledge at lower levels, as described in Delwart, Preusker et al. (2007) for the Medium Resolution Imaging Spectrometer (MERIS) on-board the EnviSat mission. The starting wavelength is set to 410 \( \text{nm} \); which corresponds to a zero wavelength position shift in Figure 5-4. The selected interpolation is cubic spline and linear for the TOA spectral radiance reconstruction and linear interpolation for the S2 SRF bands. This is based on the results in Figure 5-4 which show the maximum disagreement for the TOA radiance interpolation methods and the limited impact of the S2 SRF interpolation. Note that only the spread of values is of interest here since the spectral sampling error was already reported in Figure 5-4. Thus, the error distribution has been normalised to the original central wavelength and bandwidth values.
Figure 5-7 Distribution of spectral sampling errors for S2 bands with an associated TRUTHS central wavelength and bandwidth knowledge uncertainty of 0.2 nm ($\lambda = 1$).

Table 5-1 Standard deviation results of the distribution of spectral sampling errors presented in Figure 5-7

<table>
<thead>
<tr>
<th>Statistics of TRUTHS spectral knowledge</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
<th>B7</th>
<th>B8</th>
<th>B8A</th>
<th>B11</th>
<th>B12</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOA Cubic spline std.</td>
<td>0.21</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.23</td>
<td>0.09</td>
<td>0.10</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>TOA linear interp. std.</td>
<td>0.20</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.23</td>
<td>0.08</td>
<td>0.10</td>
<td>0.01</td>
<td>0.03</td>
<td>0.04</td>
</tr>
</tbody>
</table>
The previous results in Figure 5-4 do not directly apply in this case since on the one hand, the resolution of the bands has been modified and, on the other hand, the sensitivity here is based on the wavelength and does not vary the spectral sampling requirement — see Figure 5-2 — at every shift of the bands central wavelength. Binning step was set to 0.05% for B1, B6, B7, and B8 and 0.01% was the binning for the rest of the bands.

All the results in Table 5-1 show a standard deviation below 0.1% with the exception of B1 and B6 which are around 0.2%. That is because the error variations are the most important for this band as reported in Figure 5-4. The choice of TOA spectral radiance interpolation method seems to have a negligible effect. The majority of the bands present a symmetrical distribution and thus the reported standard deviation can be reported as uncertainty at $k = 1$. However, deriving an uncertainty out of skewed distributions such as B1 or B2 is not directly feasible since this parameter cannot be directly linked to a particular probability distribution (BIPM, IEC et al. 2008).

Chapter 4 has also shown the impact of the Sentinel-2 SRF uncertainty using the same TOA radiance input as described in Section 5.2.1. The results show that assuming a spectral response uncertainty of 0.2 nm ($k = 1$) for systematic and 0.1 nm ($k = 1$) for random spectral calibration knowledge, the dispersion of the data was below 0.3%. Indeed, this maximum dispersion corresponded to B7 for which a clear bimodal distribution was found. This was a consequence of there being different sets of filters used by the Sentinel-2 MSI detectors as explained in Chapter 4. The use of the detector footprint mask embedded in the Sentinel-2 L1C products together with a specific detector SRF could significantly reduce these numbers. Alternatively, if a SRF mean of the S2 MSI is used as in Section 5.2.1, the spectral residual after the diffuser equalisation must be considered (Barsi, Lee et al. 2014).

To evaluate the effect of filter degradation, in Chapter 4 a spectral shift was added to the SRF bands of Sentinel-2. As an indicative baseline for spectral degradation, the reported degradation rates measured in-flight by Spectroradiometric Calibration Assembly (SRCA) on-board the Terra MODIS mission (Xiong, Che et al. 2006) were used. Specifically, for Sentinel-2 VNIR bands these corresponded to: -0.33 nm (B1), -0.26 nm (B2), 0.04 nm (B3), -0.03 nm (B4), -0.05 nm (B5), -0.07 nm (B6), 0.1 nm (B7), 0.2 nm (B8), -0.18 nm (B8A). The systematic error introduced by this was negligible — below 0.1% — for all the VNIR bands with the exception of B1 which rose to approximately 0.3% due to its narrower bandwidth and stronger degradation rate. The SWIR bands have not been analysed since a more dominant and variable effect is due to water condensation on the cooled detectors, leading to additional interference effects.

### 5.2.4 Spectral domain: the impact of spectral binning

As reported in Section 5.2.2, the spectral binning effect can be easily appreciated for B1 and B2. For the cross-calibration events, the binning levels reported in Figure 5-2 are not required since the specific
application requires the best estimate over a ROI. That is, the accumulation of pixels in the across and along track direction reduces the spatial and temporal uncorrelated component of the pixel noise respectively. Therefore, the SNR requirement for this specific application is comparably lower. Systems like BinGO (BInning patterN Generator and Optimiser) described by Dell’Endice, Nieke et al. (2009) can be used to re-programme the spectral binning pattern of a Field Programmable Gate Array (FPGA) card to fit specific application requirements. An example of such type of reprogramming was performed to estimate the spectral instrument response in-flight of MERIS on-board the EnviSat mission (Delwart, Preusker et al. 2007).

For a mission like TRUTHS, the binning can be adapted to further sample B1 in a specific cross-calibration event. Once the event is finished, the default setup can be restored so that the impact over other applications such as radiation budget and/or memory requirements is minimum. The previous simulations in Figure 5-4 have been run without introducing spectral binning and the results for B1 together with the convolved values at the TOA spectral radiance are shown in Figure 5-8. In addition, the results for spectral knowledge in Figure 5-7 are presented again for B1 with no spectral binning.

The results in Figure 5-8 show that the sampling/resolution largely decreases whereas the spectral knowledge uncertainty slightly improves with a standard deviation of 0.18%. Increasing the sampling introduces further averaging over the band convolution.

When spectral resolution is reduced, the bands become more sensitive to solar and atmosphere absorptions from strong absorptions such as water-vapour (Green 1998). However, what the simulation suggest is that this effect, although present at a singular native spectral band, it is largely averaged out when many bands across the S2 bandpass are considered.

![Figure 5-8](image)

Figure 5-8 (Left) TOA reflectance error in estimating the Sentinel-2 MSI equivalent radiance in B1 due to the TRUTHS sampling bands and preliminary resolution of the detector bands with no spectral binning applied. (Centre) The measured TOA radiance as generated by MODTRANv5 in black colour, in red colour the resulting measurements of TOA radiance as measured by TRUTHS Earth imager and, in blue colour, the Sentinel 2 B1 band. (Right) The distribution of spectral sampling errors for S2 B1 band with an associated TRUTHS central wavelength and bandwidth knowledge uncertainty of 0.2 nm ($\Delta = 1$) with no spectral binning applied. The legend of the plots are equivalent to Figure 5-4, Figure 5-3, and Figure 5-7 respectively.
5.2.5 Spectral domain: the TOA radiance sampling

The use of a TOA radiance with a sampling of 0.5 nm is justified in Wu, Xiong et al. (2015) on the basis that this sampling is twice or larger than the main variations in the TOA spectral radiance (i.e. is valid based on Shannon sampling theorem). Here the previous TOA spectral radiance at 0.1 cm\(^{-1}\) has been replaced by the same MODTRAN simulation at 0.5 nm for an empirical verification of the approach and the results of Figure 5-4.

![Graph showing error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands B1 (left) and B6 (right) due to the TRUTHS sampling bands and preliminary resolution of the detector bands.](image)

By comparing the results between Figure 5-4 and Figure 5-9, the results largely agree in absolute values with a negligible difference. Thus, the use of a 0.5 nm TOA radiance simulation in previous studies as in Wu, Xiong et al. (2015) is fully justified. The difference arises in the relative shape of the curves due to second order variations that cannot be captured by the 0.5 nm resolution of MODTRAN. For B6, there is a slight decrease on the ringing amplitude whereas for B1 the shape has resulted in a smoothed version of that provided in Figure 5-4. Thus, the use of a 0.1 cm\(^{-1}\) MODTRAN spectral resolution is justified in this context since it provides a much more detailed analysis of sampling/resolution error vs. the detector array position and can better help to describe the spectral knowledge as a distribution of errors as already pointed in Section 5.2.1.

5.2.6 Spectral domain: the impact of the site

The effect described in Section 5.2.2 can vary when the cross-calibration is performed under alternative test-site targets with different spectral properties and atmosphere. The snow simulations in order to be more realistic were done with a sub-arctic summer atmospheric model at 60 degrees latitude and with an associated high SZA of 65 degrees. Aerosol optical thickness (AOT) and water vapour were obtained from AERONET in Greenland. All other sites are done for mid latitude summer, June 22, same solar angles as the simulation in Section 5.2.1 but with AOT and water vapour data taken from La Crau for grassland, Ascension Island for oceanic site.
Figure 5-10 shows the error due to spectral sampling and resolution for these TOA spectral radiances as in Figure 5-4. Cubic spline and linear interpolation are used to reconstruct the TOA radiance and linear interpolation for the S2A bands.

Figure 5-10 TOA error in estimating the Sentinel-2 MSI equivalent radiance for VNIR bands (above) and SWIR bands (below) due to the TRUTHS sampling bands and preliminary resolution of the detector bands for different modelled sites.

The result shows that the impact of the different modelled TOA spectral radiances has an impact below 0.1 % peak-to-peak for the majority of bands. For band B4, the sensitivity to the grass simulation raises the error slightly above 0.1 %. For the bands B5 and B6, the error is significantly reduced when using other than a desert simulation. In the B5 case, the error is below 0.1 % at any interpolation and at any simulation other than the desert one. In the B6 case, the simulations show a considerable improvement of the error when compared to the desert case. In general, the errors can be bracketed in the range 0.2-0.5 % depending on the array position, site and interpolation type.
5.3 Uncertainty assessment: Spatial Domain

5.3.1 Spatial domain: methodology

In the spatial domain, site non-uniformity, in combination with uncertainty due to misregistration of the instrument scenes, in principle leads to a systematic uncertainty that needs to be accounted for in the total uncertainty budget. The effects of even small differences across the site can lead to a bias when it is used for radiometric calibration; therefore this needs to be carefully assessed and addressed. The results in Chander, Helder et al. (2013) showed that for a 2-pixel spatial knowledge, the potential uncertainty introduced by misregistration was at 0.1% for the VNIR bands and 0.2-0.3% for the SWIR bands of Landsat 7 Enhanced Thematic Mapper Plus (L7 ETM+).

For calibration, multiple sites and multiple observations of the same site will be used and the location knowledge will likely be known to <<1 km, so reducing significantly the uncertainty in cross calibration due to spatial co-location error. In the multi-temporal case, the TRUTHS and CLARREO orbits are defined with a 90° polar orbit and 61-day ground track repeat cycle at 609 km altitude (Roithmayr, Lukashin et al. 2014). In addition to this, the orbit is asynchronous with a different time overpass over the Equator during the year. Thus, when applying this method to several matchups, the spatial offsets will be largely independent and the effects can be reduced. This same assumption is discussed in Wielicki, Doelling et al. (2008) where a spatial matching noise below 1% is considered as a threshold to minimise the impact over the temporal aggregation of diverse satellite-to-satellite matches.

This study thus concentrates on the coarser effect of the impact of the spatial non-uniformity in a specific area and imperfect geographic location knowledge and/or lack of co-alignment between the reference sensor and the sensor under calibration. This study considers two arbitrarily chosen ROIs spatially separated within the Libya-4 site and LaCrau site and the resultant systematic uncertainty due to this separation i.e. the variance in TOA reflectance caused by surface non-uniformity. The two sites have been chosen as examples of typical calibration sites that represent different levels of uniformity.

In order to model the effect of spatial non-uniformity, a practical approach is presented using real EO data with low relative uncertainty and a sufficiently large swath that covers the area under study. L8 OLI TOA reflectance factor images have been selected with a large swath of 185 km that allows the selection of multiple ROIs across it (Irons, Dwyer et al. 2012). L8 OLI pixel-to-pixel non uniformity residual lies between 0.2% and 0.3% for the complete focal plane (Morfitt, Barsi et al. 2015). This relative uncertainty provides a sufficiently small effect, compared to the expected magnitude of ROI site variability, to allow us to use scenes of L8 OLI for this evaluation. For a high-radiance scene such as Libya-4 or La Crau, the uncertainty budget is dominated by highly correlated effects in the spatial and temporal domain, so minimising the weight of uncorrelated spatiotemporal contributions e.g.
instrument noise. The effects of such random variations are further reduced by the fact that each ROI covers several pixels, so providing some averaging of signal noise etc.

In addition to L8 OLI data, the study is repeated using S2 MSI data. This mission also provides a large swath of 295 km. Its pixel-to-pixel non uniformity has been validated by the means of diffuser and natural targets on Earth leading to values well below the specification of 0.2% (Gascon 2016).

The selected bands for the study are B1, B5 and B8 for L8 OLI — 443, 865, and 2201 nm central wavelength — and B1, B8A and B12 for S2 MSI — 443, 865, and 2190 nm central wavelength. These bands illustrate boundaries between atmosphere and surface scene composition. At 443 nm the impact from atmospheric effects is significant whereas at 865 nm and ~2200 nm the expected atmosphere transmission is above 80% for both the VNIR and SWIR respectively. In addition, by comparing very similar bands of two different missions, the method can be validated.

The products selected for Libya-4 and La Crau sites were selected based on best temporal coincidence using the CEOS COVE tool (Kessler, Killough et al. 2013). The matches found are considered optimum since the overpasses over the same site are delayed by less than 15 minutes and the cloud conditions are near zero percent for the whole product tile. The selected products are described in Table 5-2.

**Table 5-2 S2 MSI and L8 OLI products for the spatial uncertainty assessment and validation**

<table>
<thead>
<tr>
<th>Site</th>
<th>Sensor</th>
<th>Product ID</th>
<th>Timestamp centre</th>
<th>Cloud [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Libya-4</td>
<td>L8 OLI</td>
<td>LC08_L1TP_181040_20170513_20170513_01_RT</td>
<td>2017-05-13T08:54:34Z</td>
<td>0</td>
</tr>
<tr>
<td>Libya-4</td>
<td>S2 MSI</td>
<td>S2A_MSIL1C_20170513T090021_N0205_R007_T34RGS_20170513T090803.SAFE</td>
<td>2017-05-13T09:08:03Z</td>
<td>0</td>
</tr>
<tr>
<td>La Crau</td>
<td>L8 OLI</td>
<td>LC08_L1TP_196030_20170420_20170501_01_T1</td>
<td>2017-04-20T10:23:17Z</td>
<td>0.08</td>
</tr>
<tr>
<td>La Crau</td>
<td>S2 MSI</td>
<td>S2A_MSIL1C_20170420T103021_N0204_R108_T31TFJ_20170420T103454.SAFE</td>
<td>2017-04-20T10:34:54Z</td>
<td>0.5321</td>
</tr>
</tbody>
</table>

In addition to the low cloud cover, each ROI for S2 MSI and L8 OLI has been screened for the possibility of degraded, saturated or invalid pixels. That is possible by reading the quality assessment band in the L1TP L8 OLI products and the different masks integrated as part of the S2 MSI L1C product. The result is that all the pixels in the ROIs processed were valid without exception.

The L1C data from the S2 MSI is directly provided as TOA reflectance factor values scaled by a ‘quantification value’ and have been processed using free software produced by ESA and named...
Sentinel Application Platform (SNAP). The L1 Digital Number (DN) in the Landsat-8 OLI L1TP product are converted to TOA reflectance, $\rho'_i$, using Equation 5.3 according to (USGS 2015):

$$\rho'_i(i, j) = \frac{M_L \cdot DN(i, j) + A_L \cdot \cos(SZA(i, j))}{\cos(SZA(i, j))}$$  \hspace{1cm} (5.3)

where $M_L$ refers to the reflectance multiplicative scaling factor for the band and $A_L$ refers to the reflectance additive scaling factor for the band. Both values can be extracted from the product metadata.

Next, the SZA for the coordinates of each pixel at row $i$ and column $j$ is calculated using the image timestamp and lat/long position using the Pysolar library (Stafford 2015). The lat/lon at each position is obtained by a coordinate transformation of the Worldwide Reference System (WRS) path/row coordinate system using PROJ.4 libraries.

Once the ROI reflectance factors are obtained, the mean of the pixels is calculated. In an iterative mode, the ROI centre is displaced following a pre-defined grid across a defined area. This process is illustrated in Figure 5-11.

![Figure 5-11 Methodology process for the assessment of spatial variations.](image)

The results of this process is a TOA reflectance factor error map for the potential displacements over the calibration site.

The next step is the processing of the error map in order to derive the uncertainty associated with the geolocation knowledge. This process is performed by calculating the standard deviation over a growing area from the centre of the error images. Thus, for each associated geolocation knowledge, an uncertainty $k = 1$ will be associated and a curve of uncertainty vs. geolocation knowledge will be obtained. This curve can be used as a tool for either the refinement of mission design requirements,
TRUTHS or CLARREO, and/or the definition of the matching requirements in order to achieve a specified spatial systematic uncertainty.

### 5.3.2 Spatial domain: results for La Crau calibration site

The site of La Crau has been traditionally used for the calibration of sensors as SPOT (Santer, Gu et al. 1992). The calibration site is defined as 400 × 400 m² area centred at 43.556º N 4.858º E, in a 60 km² flat area composed of pebbles and sparse low vegetation.

The region studied for this example corresponds to ± 0.002º off-centre in latitude and longitude. This corresponds to approximately 0.32 × 0.44 km² rectangular spacing from the centre of the site. The TOA reflectance factor at each 400 × 400 m² pixel ROI and band considered are shown in Figure 5-12.

![TOA reflectance factor at the 400 × 400 m² at the LaCrau site for the considered L8 OLI and S2 MSI bands.](image)

Figure 5-12 TOA reflectance factor at the 400 × 400 m² at the LaCrau site for the considered L8 OLI and S2 MSI bands.
TOA reflectance factors for both missions show similar results. The S2 B1 shows a less defined pattern due to its coarser spatial resolution of 60 m. The ROI has been selected as 400 × 400 m² which can be achieved for S2 B8A and S2 B12. However for the S2 B1 and L8 bands, this distance cannot be delimited within an exact number of pixels and the final spatial dimensions are 360 × 360 m² and 390 × 390 m² respectively. This point together with the geolocation uncertainty for each sensor and band results in a displacement between bands that can be visually estimated at around 1 pixel.

Figure 5-13 shows the spatial error as a consequence of the ROI displacement ± 0.002° off-centre in latitude and longitude.

Figure 5-13 TOA reflectance factor error map for the LaCrau site and the considered L8 OLI and S2 MSI bands
The spatial error between both missions largely agree in terms of gradient and image shape. These variations, however, seem slightly shifted between the missions with the spatial variations slightly shifted towards positive values for the S2 MSI with respect to the L8 OLI. This consistency between the two missions for all three bands suggest that the small delay between overpasses and, especially, the different angular configuration is likely to be the cause of these variations.

Figure 5-14 presents the standard deviation of the spatial error pixels as growing from the centre of Figure 5-13.

![Figure 5-14 Spatial uncertainty vs. spatial offset for the LaCrau site and the considered L8 OLI and S2 MSI bands. The bands contained in each panel are: (a) S2 B1 and L8 B1, (b) S2 B8A and L8 B5, and (c) S2 B12 and L8 B7.](image)

The results presented here show a strong consistency between the two missions which provides confidence to the results. There is a clear linear dependency of the spatial uncertainty vs. the spatial knowledge of the ROI centre. The linear dependency shows a small slope variation between S2B8A and L8 B5 and S2 B12 and L8 B7. The results for S2 B1 follow a linear dependency in a more irregular way that can be attributed to the coarser spatial resolution.

These results can be easily applied to determine the spatial knowledge requirements for TRUTHS/CLARREO missions. If we were to account for a more realistic scenario with e.g. spatial uncertainty 10 times lower than the maximum displacement shown here, i.e. approximately $32 \times 44 \text{ m}^2$ — the uncertainty would be 10 times that of the maximum displayed. This means approximately 0.12% for B1, 0.27% for B5 and 0.5% for B12. This value would represent the systematic uncertainty introduced over one overpass and would be largely uncorrelated in between matches and will tend to reduce with increasing number of overpasses and match-ups.

5.3.3 Spatial domain: results for PICS sites

The ROI reference centre position is selected to be at the centre of the Libya-4 site as defined in Lacherade, Fougnie et al. (2013) — i.e. 28.55° N 23.39° E — and with a size of 20 km x 20 km. The region studied corresponds to ± 0.05° off-centre in latitude and longitude. This corresponds to approximately $10 \times 10 \text{ km}^2$ rectangular spacing from the Libya-4 centre. The centre of the ROI has been moved in 15 equidistant points in each of the directions. The reason for using a ROI size of 20 km x 20 km is that based on the results of Govaerts (2015), this is a sufficiently large area at which
the dune effect can be reasonably integrated. A smaller area would introduce a larger dispersion due to the dune effect whereas a larger area would require a bigger swath and eventually would introduce low-frequency spatial variations. The TOA reflectance factor at each 20 × 20 km² pixel ROI and example band are shown in Figure 5-15.

The values of TOA reflectance factor are very similar for both missions and no visual shift in geolocation can be identified. The size of the ROI used is sufficiently large so that the impact of a small miss-registration between missions and/or number of pixels in the ROI does not have any impact.

Figure 5-15 TOA reflectance factor at the 20 × 20 km² at the Libya-4 site for the considered L8 OLI and S2 MSI bands.

Figure 5-16 shows the spatial error as a consequence of the ROI displacement ± 0.05° off-centre in latitude and longitude.
The image error variations and levels are very similar for S2 B1 and L8 B1 and S2 B8A and L8 B5. Indeed the results for S2 bands are slightly larger than those of L8 bands in similar manner to the results for La Crau in Figure 5-13. Both images for S2 B1 and L8 B1 present an irregular pattern typical of real scene variations. However this irregular pattern cannot be found in S2 B8A and L8 B5. The pattern of errors suggests in both a dependency of TOA reflectance factors with directionality of the sun illumination. The results suggest that at these bands the errors that have been measured could be the consequence of the viewing angular variations of the push-broom sensor. The difference of results for L8 B7 and S2 B12 cannot be explained by a difference on angular configuration or scene change between overpasses.
Figure 5-17 presents the standard deviation of the spatial error pixels as growing from the centre of Figure 5-16.

Figure 5-17 Spatial uncertainty vs. spatial offset for the Libya-4 site for L8 OLI and S2 MSI bands. The bands contained in each panel are: (a) S2 B1 and L8 B1, (b) S2 B8A and L8 B5, and (c) S2 B12 and L8 B7.

The results show again an almost linear dependency of the spatial TOA uncertainty with the spatial positioning knowledge. Following the same logic as in Section 5.3.2, the impact over a realistic geolocation knowledge of a TRUTHS/CLARREO like mission would be well below 0.1%.

The results agree very well for both missions except for the S2 B12 and L8 B7. In this case the S2 B12 and L8 B7 results are largely variable and confirm the disagreement between the missions seen in Figure 5-16.

Figure 5-18 presents the error distribution for L8 B7 and S2 B12 in Figure 5-16.

Figure 5-18 TOA reflectance factor error distribution for the Libya-4 site and L8 OLI B7 (left) and S2 MSI B12 (right)

On the one hand, the results for L8 B7 present a map of error with an irregular map with a normal or similar distribution pattern. On the other hand, the results for S2 B12 show a North-South variation following the orbit overpass with a highly skewed distribution. Thus, these variations cannot be caused by the sensor angular variations or natural variability. Indeed it suggests an ACT sensor effect.
Several effects can be suggested for these variations although none of them can be fully verified here. A crosstalk effect in the SWIR bands has been detected and accordingly corrected for in the S2 L1C products. In addition, the SWIR bands undergo regular decontamination due to deposition of moisture on the top of the detectors (Gascon 2016).

The results obtained for Libya-4 are lower than those obtained in Chander, Helder et al. (2013). One reason for this disagreement could be the potential pixel non-uniformity of the L7 ETM+. This should have limited impact due to the whiskbroom design of the L7 ETM+ instrument. However, the main difference between the two studies arises from the selection of the ROI over Libya-4. The selection of a small ROI of $3 \times 3$ km$^2$ as used by Chander, Helder et al. (2013) introduces a larger dispersion due to the dune effect (Govaerts 2015).

5.4 Uncertainty assessment: Temporal Domain

5.4.1 Temporal domain: methodology

This section describes the impact on the TOA reflectance due to changes in the sun angle (azimuth and zenith) for a specified time-span after the overpass of the reference satellite (TRUTHS/CLARREO) and the resulting residual uncertainty of the correction between the reference and target satellites.

Recent work in McCorkel, Thome et al. (2013) studied the effect of temporal mismatch between MODIS vs. Hyperion matches. The latter instrument was measuring in an orbit 40 minutes preceding the MODIS one until mid-2005. The orbit of Hyperion was changed from mid-2005 resulting in a rare cross-calibration between the two missions. This unusual situation triggered the possibility to compare the impact of the temporal overpass differences between coincident overpasses — within 30 to 40 minutes — and non-coincident overpass — within 30 days separation — over the Railroad Valley calibration site. The results showed that although the dispersion of the data significantly increased, the bias between the two cases was between 1-2%. To a large extent, BRDF and temporal mismatches were largely averaged out even for such a large timespan difference.

The orbit choice of CLARREO is set as a polar 90 degrees asynchronous orbit (Roithmayr, Lukashin et al. 2014). This type of orbit permits the sparse sampling of brightness temperature over the diurnal cycle and subsequently improves the sampling error (Kirk-Davidoff, Goody et al. 2005). The climate benchmark of missions like TRUTHS or CLARREO is largely improved by this type of orbit since it assures full diurnal cycle sampling for spectral fingerprints as well as full reference inter-calibration sampling over all climate regimes and all satellite orbit thermal conditions (Wielicki, Young et al. 2013).
In the work presented in Roithmayr, Lukashin et al. (2014), the selection of cross-calibration matchups was set to a global scale within a 5 minutes of delay between overpass. At that time delay, the temporal noise was found to be at the 1% level and with sufficient samples the noise reduces to <0.3% (Wielicki, Doelling et al. 2008). Extending the concept of sparse sampling studied in Kirk-Davidoff, Goody et al. (2005) and Wielicki, Young et al. (2013), it is conceivable to assume that the temporal systematic errors are largely uncorrelated and converge to a low bias not only when a global scale is taken into account but also when a more restricted area is considered.

The approach developed here seeks to consider a complementary scenario where the surface and atmospheric conditions are considered stable in time but the inter-calibration matchups are limited to the specific locations that fulfil these conditions. Thus, the sites considered here are PICS and more specifically, the Libya-4 site due to the large amount of prior work available and its representativeness of the PICS sites. Due to the better temporal stability of these types of sites, the considered delay between overpasses can be increased and consequently the opportunities increased. Here a delay of 30 minutes will be considered as representative of the upper limit considered for the SNO cross-calibration over PICS. Furthermore, the accurate observation of missions like TRUTHS and CLARREO in conjunction with a pointing capability offer the possibility of an improved modelling of the surface BRDF models of PICS sites. The achievable uncertainty of a temporal correction using this approach will also be considered here.

5.4.2 Temporal domain: atmospheric variation and radiative transfer code impact

In this section, the TOA reflectance factor variation over a 30 minutes timespan will be evaluated as a consequence of the atmospheric solar reflected radiance due to solar angle variations. In addition, the study will also consider any discrepancy between different atmospheric correction algorithms.

Solar angles at 30 second intervals are calculated and used as inputs to the MODTRANv5 model for a parameterisation as described in Section 5.3; the starting time of the simulation is taken as a typical Landsat-8 OLI product reference timestamp for a Libya-4 overpass, specifically 8:56:32 local time for days of the year 173 and 355. The study has been undertaken at three wavelengths — 443, 865, and 2201 nm — that represent the central wavelengths of the Landsat-8 OLI bands with a high atmospheric sensitivity (B1) and with lower atmospheric sensitivity in the VNIR (B5) and SWIR (B7). The radiance is further normalised by the cosine as shown in Equation 5.3 so that a reflectance factor difference can also be calculated. Due to the large amount of time and MODTRAN runs — 400 simulations covering the VNIR and SWIR spectrum at 1 nm spectral resolution — the simulations have been setup in parallel using the JASMIN facilities (Lawrence, Bennett et al. 2013).
In addition to using MODTRAN, we have also carried out a similar analysis using the 6SV1 (second simulation of a satellite signal in the solar spectrum, vector, and version 1) radiative transfer code (Vermote, Tanre et al. 1997) and interfaced using the Py6S library (Wilson 2013). By comparing them, it is possible to assess the effect of any radiative transfer (RT) code biases. In particular, the work described in Kotchenova, Vermote et al. (2008) or Kotchenova, Vermote et al. (2006) pointed out important differences in the aerosol and molecular scattering between these two RT codes. The parameterisation follows that described in Section 5.3 for MODTRAN but using the sand model in 6SV1.

Figure 5-19 shows the reflectance factor temporal differences using both RT codes — 6SV1 and MODTRAN — at the studied wavelengths for the year day 173 and 355. These two days represent the most extreme SZA angle conditions (summer and winter solstice respectively).

The results show two anomalies at 865 nm and 2201 nm. These are produced by the MODTRANv5 simulations and a zoom to the radiance trend showed that at that point, the trend was slightly changing. That is, the assumption is that there is a software interpolation discontinuity and/or an ill-conditioned solution. This is further amplified when small errors are calculated and the SZA round off at 2 decimal digits introduces a small noise.

The graph shows the importance of the atmosphere at 443 nm due to the strong impact of scattering at this wavelength. For winter periods, the atmospheric radiance variation becomes dominant whereas in the summer period it is of the same order as the cosine effect. Similarly, the result at 865 nm shows how these two variations are largely compensated. Indeed, here the changes are so small that the different atmosphere and surface balance between the two RT codes and setup is clearly shown. Where 6S shows a minimum error increase in winter, MODTRAN does it for summer. Finally, the TOA reflectance error at 2201 nm shows an important change as a consequence of the dominant cosine effect and an almost negligible atmospheric effect.
The discrepancy between the two radiative codes can be seen at 443 nm and 865 nm. No discrepancy can be seen at 2201 nm as a result of the low atmospheric impact. The work in Kotchenova, Vermote et al. (2008) described differences between the radiative codes in the order of several percent. However, the results here demonstrate these differences are limited to approximately 0.2% at 443 nm and below 0.1% at 865 nm. This is an expected result since the constant biases are cancelled out and only the differences in temporal effect between the two radiative codes is relevant.

5.4.3 Temporal domain: atmospheric knowledge

In addition to the temporal discrepancies between radiative codes, it is important to understand the potential impact of the atmospheric knowledge as it contributes to the temporal effect. In order to study its impact, the TOA radiance calculation has been repeated nearly 1000 times with varying inputs of AOT and water vapour. The AOT and water vapour values are random samples from a normal distribution determined from a mean and standard deviation as specified in Mishra, Haque et al. (2014) as representative of Libya-4. These are 0.0858 and 0.0486 for mean and standard deviation of AOT and 2.85 and 0.7 for the water vapour mean and standard deviation. Just a very small percentage of AOT samples out of the normal distribution were negative. These samples were set to 0 for the simulations.

The resulting residual in the sun angular correction can be understood as the repetition of the error trend shown in Figure 5-19 for each simulation. This provided approximately 1000 potential error curves which will increasingly vary with length of time. The distribution of corresponding reflectance errors versus timespan is studied here. Figure 5-20 shows the results for the reflectance error dispersions for a 30 minute timespan and Table 5-3 contains the statistical parameters for the dispersions respectively.

<table>
<thead>
<tr>
<th>Day of Year 173</th>
<th>Day of Year 355</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2201 nm</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0303</td>
</tr>
<tr>
<td>Median</td>
<td>-0.0290</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.03</td>
</tr>
<tr>
<td>68.27% coverage</td>
<td>[-0.04, 0.02]</td>
</tr>
<tr>
<td>Skewness</td>
<td>-1.2034</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>2.1070</td>
</tr>
</tbody>
</table>
The results indicate the low impact that the radiative transfer code inputs have on the residual uncertainty after the sun angular correction.

### 5.4.4 Temporal dimension: atmospheric variation

Thus far, the analysis has been conducted assuming that the atmospheric parameters have remained constant over the 30 minutes timespan being studied. Similar to the description of Section 5.3, the potential atmospheric variations in this timespan are difficult to predict, although likely to be small for the types of site chosen unless an unusual weather event occurs. However, the approach taken here is to predict the worst case uncertainty in the correction and limit the potential minimum and maximum uncertainty in a temporal correction.

Figure 5-21 shows the results for the TOA radiance dispersion at a point in time as a consequence of AOT and water vapour variations—in this case 30 minutes—and Table 5-4 contains the statistical parameters for the dispersion respectively.
Figure 5-21 Results for TOA radiance dispersion at 30 minutes at 443 nm and yearday 173 (a), 865 nm and yearday 173 (b), 2201 nm and yearday 173 (c), 443 nm and yearday 355 (d), 865 nm and yearday 355 (e), and 2201 nm and yearday 355(f).

Table 5-4. Statistical parameters for Figure 5-21

<table>
<thead>
<tr>
<th>Day of Year 173</th>
<th>Day of Year 355</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B1</td>
</tr>
<tr>
<td>Mean</td>
<td>12.4844</td>
</tr>
<tr>
<td>Median</td>
<td>12.4770</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.57%</td>
</tr>
<tr>
<td>68.27% coverage interval</td>
<td>[12.41, 8.46, 0.56, 8.36, 5.42, 0.34, 12.56, 8.51, 0.58, 8.52, 5.51]</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.5784</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>-0.3525</td>
</tr>
</tbody>
</table>

The results indicate the potential dispersion of the TOA reflectance factor — the time is constant and the dispersion is indifferent of radiance or reflectance factor — are expected to be below the 1% level for 443 nm and 865 nm. For the SWIR region at 2201 nm the relative uncertainty increases due to the weak signal measured. Table 5-4 indicates standard deviation values around 0.5 % for the simulation at 443 nm. This is produced by the large impact of aerosol variations in this region. Nonetheless, the simulation here has used any potential variation in AOT and water vapour that could occur throughout the year. The variations in a 30 minute timespan are expected to be much smaller and the impact will be several times lower. In fact this uncertainty could be used as a worst case to account for aerosol variation at any time of the year and thus an uncertainty for any time gap. In addition, for a PICS site
characterisation using multiple TRUTHS overpasses, the variations of the atmosphere can be assumed random to a large extent. Since each acquisition represents a different solar angular geometry, there is no direct method to assess the random improvement over time. However, in a fully uncorrelated atmospheric situation and assuming a maximum improvement, just 10 cloudless overpasses, would reduce the uncertainty levels by a factor of over 3. This would set the uncertainty levels below the 0.5% level or close to it, without placing any limitation on time gap between sensor and test overpasses.

5.4.5 Temporal domain: surface impact

The earlier sections have evaluated the effects of the radiative code, temporal atmospheric change and atmospheric knowledge in the temporal correction. There is, however, a final factor to consider related to the knowledge of the surface reflectance. In this case, the surface reflectance model from Bouvet (2014) has been taken as a reference. This surface reflectance model is the result of an inversion of MERIS observations over the 2006-2009 period that tunes the four parameters of a Rahman-Pinty-Verstraete (RPV) model. The aerosol optical depth has been extracted from the mean year value described in Mishra, Haque et al. (2014). The selected aerosol profile is ‘continental’. The discussion in Bouvet (2014) suggested this model not because of its better representation of the aerosol profile in Libya-4 but because of its apparent better performance when combined with the RPV model. The characterisation of the surface angular variation uses this model (Rahman, Pinty et al. 1993). The model provides the reflectance, \( \rho \), defined by four parameters (\( \rho_0 \), \( k \), \( \Theta \) and \( \rho_c \)) for the viewing and illumination conditions (SZA \( \equiv \theta_s \), VZA \( \equiv \theta_V \) and RAA \( \equiv \Delta \phi \)) as follows:

\[
\rho(\theta_s, \theta_V, \Delta \phi, \rho_0, k, \Theta, \rho_c) = \rho_0 M_1(\theta_s, \theta_V, k) F_{\text{HG}}(g, \Theta) H(\rho_c, G)
\]  

(5.4)

Where each one of the terms is defined as:

\[
M_1(\theta_s, \theta_V, k) = \frac{\cos^{k-1} \theta_s \cos^{k-1} \theta_V}{(\cos \theta_s + \cos \theta_V)^{1-k}}
\]  

(5.5)

\[
F_{\text{HG}}(g, \Theta) = \frac{1 - \Theta^2}{(1 + 2 \Theta \cos g)^{3/2}}
\]  

(5.6)

\[
H(\rho_c, G) = 1 + \frac{1 - \rho_c}{1 + G}
\]  

(5.7)

\[
\cos g = \cos(\theta_s) \cos(\theta_V) + \sin(\theta_s) \sin(\theta_V) \cos(\Delta \phi)
\]  

(5.8)

\[
G = (\tan^2(\theta_s) \tan^2(\theta_V) - 2 \tan(\theta_s) \tan(\theta_V) \cos(\Delta \phi))^{1/2}
\]  

(5.9)
The terms described in Equations 5.4–5.9 represent different features of the reflectance function (Rahman, Pinty et al. 1993). The amplitude component is set by \( \rho_0 \) and then modified by the term \( M_1 \) which defines the overall shape of the angular field using the parameter \( k \). \( F_{HG} \) is a Henyey-Greenstein function that provides the balance between forward and backward scattering and is described through the parameter \( \Theta \) and \( g \) (described in Equation 5.8). \( H \) describes the hotspot effect through the parameter \( \rho_c \).

Values for \( k \), \( \Theta \) and \( \rho_c \) for the Libya-4 site have been extracted from the results obtained in Bouvet (2014) for a surface bi-directional reflectance distribution function (BRDF) model. The values have been derived from data pertaining to the whole Libya-4 ROI site and therefore describe the BRDF of large scale structures at the site such as the dunes.

The work in Bouvet (2014) also discussed the limitations of a model that cannot be traceable in-flight and discussed the possibility of providing an absolute traceable standard by using observations of missions like TRUTHS or CLARREO. The overpasses of these missions can be used in the same manner as MERIS. In addition, these two missions incorporate in their design a gimbal mechanism that can further provide different angular observations over the same site. Thus, in this section the surface reflectance model from Bouvet (2014) is tested to understand the potential surface reflectance variation and uncertainty residual correction in an overpass of TRUTHS or CLARREO after 30 minutes of a L8 OLI overpass over Libya-4.

Here, the simulation has provided a similar approach to the one described in Figure 5-20 and Table 5-3. The change in the surface reflectance over 30 minutes has been repeated 10000 times. The RPV parameters from Bouvet (2014) have been described as normal distributions with a 5\% standard deviation for which a sample out of a normal distribution is extracted for each parameter at each iteration. That is, the uncertainty of the four parameters in the RPV model have been assumed as uncorrelated. As mentioned earlier, significant upgrade in the performance of these models require accurate reference measurements of sites and surfaces. Missions like TRUTHS not only propose accurate measurement but also point to capabilities that can further tune the model as indicated in Bouvet (2014). Thus, the level of uncertainty in the surface reflectance model used for this simulation can be taken as a worst case with the expectation that the knowledge of the parameters would be much lower — i.e. the parameter \( \rho_0 \) that represents the albedo in Equation 5.4. The resulting distributions for nadir viewing and summer and winter overpass are provided in Figure 5-22 and their main statistics in Table 5-5. Only mean and standard deviation are reported due to the near-normal distribution shape (Gorroño, Bialek et al. 2016).

The results at 443 nm result in an uncertainty that oscillates between 0.2-0.3\% at any time in the year whereas the results at 865 nm range between 0.3-0.4\% at any time of the year. If we consider the
combined effect of the atmospheric impact, then the radiative code differences must also be accounted for at the bands dominated by the atmospheric scattering (see Figure 5-20). Considering an equal weight on the TOA reflectance at 443 nm, the same levels of uncertainty at 0.2-0.3% can be kept. At 865 nm the atmospheric contribution is much lower and the uncertainty in the surface reflectance variation can be considered as dominant.

The SZA variations in the summer case start at 21.2° and end with 14.8°. For the winter case, the SZA ranges from 56.1 down to 53.8 degrees. These results can be easily compared with the TOA reflectance factor dependency over Libya-4 performed by Mishra, Helder et al. (2014). Applying these SZA variations to the empirical SZA curve using TOA reflectance factor images from MODIS band 2 (841.9 nm) reveals a predicted variation of 0.88% and 0.33% respectively for the summer and winter case.

![Graphs showing surface reflectance error dispersion at 443 nm and yearday 173 (a), 443 nm and yearday 355 (b), 865 nm and yearday 173 (c), and 865 nm and yearday 355 (d).]

Figure 5-22 Results for surface reflectance error dispersion at 30 minutes at 443 nm and yearday 173 (a), 443 nm and yearday 355 (b), 865 nm and yearday 173 (c), and 865 nm and yearday 355 (d).

<table>
<thead>
<tr>
<th>Day of Year 173</th>
<th>Day of Year 355</th>
</tr>
</thead>
<tbody>
<tr>
<td>443 nm</td>
<td>865 nm</td>
</tr>
<tr>
<td>Mean</td>
<td>2.3337</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.22</td>
</tr>
</tbody>
</table>
These results are consistent with the ones obtained in Figure 5-19 for the summer and winter cases at 865 nm respectively. The predicted atmospheric variation at 865 nm in Figure 5-19 shows a small but present increase of the TOA reflectance factor at 0.1-0.2% either in the winter or summer case. If these values are considered together with the reported variations of the surface at 865 nm in Figure 5-22, the agreement with the predicted results in Mishra, Helder et al. (2014) gets very close in the summer case and just above 0.5% in the winter case.

The work in Mishra, Helder et al. (2014) obtained different TOA reflectance linear fitting curves for the rest of MODIS bands and their slopes were fitted in an exponential model. At 443 nm the predicted slope is very close to zero. Assuming that the surface represents around half of the total signal at the TOA in the spectral region around 443 nm, the impact of the surface reflectance variation reported in Figure 5-22 can be set at around -0.24% for winter and +1.16%. When combining with the results of atmospheric variation at 443 nm in Figure 5-19, the global variations are close to zero and agrees with the reported slope close to zero by Mishra, Helder et al. (2014).

To sum up, for a temporal delay of 30 minutes between a target sensor and a reference sensor the expected TOA reflectance factor variations will be at the 1%-level without further corrections. The asynchronous orbit of a mission like TRUTHS and CLARREO will be translated in an inconsistent delay pattern. That is, if these errors are largely uncorrelated, for just nine matchups over PICS, the expected temporal uncertainty will be reduced below 0.33%. Furthermore, if the temporal correction over PICS is carried out the MCM approach and radiative transfer code comparison have shown that it is possible to correct the temporal bias with an uncertainty residual in the 0.2-0.3 % range.

5.5 Discussion

This work presents the uncertainty contributions in the spectral, spatial and temporal domain of sensors of the form used in Sentinel-2 MSI or L8 OLI when cross-calibrated using TRUTHS as a reference over different cross-calibration sites. The results obtained support previous work in Chander, Helder et al. (2013) and conclude that the uncertainty from the main sources related to the radiometric properties of the site and the spectral matching of the sensors generally falls well below the 0.5% ($k = 1$) level. At this level, the uncertainty in the reference sensor becomes the main contribution in the cross-calibration uncertainty budget. In metrological terms, it means that any effort directed towards an improvement of the calibration transfer methodology will have a limited impact in terms of the overall uncertainty budget. However, the use of PICS (and other sites with similar properties e.g. snowfields) for in-orbit inter-calibration using a high accuracy sensor like TRUTHS/CLARREO presents a major opportunity to provide in-flight calibration upgrade to EO optical missions, leading to a significantly reduced uncertainty budget.
The use of specific examples, with real data has facilitated the use and description of a more rigorous methodology. In particular, the work presented has clarified the use of error and uncertainty and has discussed the implications for uncertainty improvement through use of multiple acquisitions. This type of approach is specifically relevant for an instrument like TRUTHS in asynchronous orbit with potentially different time and spatial matchings over time. An asynchronous orbit like the one proposed in the TRUTHS or CLARREO missions means that the overpasses will be slightly different in time over the year (Roithmayr, Lukashin et al. 2014). Compared to sun-synchronous missions, this represents an advantage for inter-calibration since the delay between sensors is expected to be largely uncorrelated. However, this temporal and multi-site improvement has not been extensively studied at this point. Where necessary, boundary conditions showing a best and worst case have been provided. The next steps of this study will look more carefully at the temporal correlation effects and the uncertainty improvement over several overpasses and use of multiple sites.

The effort in this study has resulted in a set of tools and methodologies that are under continuous evolution and will ultimately be useful for the operational exploration of missions like TRUTHS or CLARREO. However, at this point in time, the results are already beneficial as a feedback to the TRUTHS mission team and helpful for refinements to the mission design.

Effects due to viewing angle have not been discussed in this work but since the reference sensor is considered to be agile it can be aligned to match that of the sensor under test.

CLARREO and TRUTHS are designed to have a polarization sensitivity of less than 0.5%, \((k = 2)\) below 1000 nm, and less than 0.75% \((k = 2)\) above 1000 nm (Wielicki, Young et al. 2013). Even though the sensitivity is low, the degree of polarisation might be certainly high for certain spectral region, sites and angular configuration. Recent work for CLARREO has shown that desert areas present a degree of polarisation at the 10%-level for longer wavelengths but that can raise up to 50% at the shorter wavelengths (Sun, Baize et al. 2015). In order to account for the polarisation effect in the cross-calibration, a set of Degree Polarisation Models (DPMs) have been derived and a methodology to account for them has been proposed in Lukashin, Wielicki et al. (2013).

5.5.1 Discussion: spectral domain

Section 5.2 studies the spectral response effect. The method is similar to the one applied in Wu, Xiong et al. (2015) for the CLARREO mission. However, in this case, the data used is based on a preliminary design of the TRUTHS sensor and includes the further effect of spectral binning. For spectrally flat sites such as Libya-4 the error due to the spectral response effect for a cross-calibration of Sentinel-2 with a TRUTHS-like reference sensor is small, with values below 0.1% for most bands. For specific bands in regions with significant spectral features within the band — e.g. Sentinel-2 B1 and/or a smaller bandwidth e.g. Sentinel-2 B5 and B6 — the error has been found to rise to around the 0.5% level. The
results of CLARREO use a larger spectral resolution of about 8 nm for 4 nm spectral sampling and Gaussian shape. Although different in design and spectral simulation of the native spectrometer, both examples show the low error introduced by most of the reference sensor bands. For the cases where the error increases significantly above 0.1%, the comparison cannot be made due to the absence of similar bands between S2 bands studied here and the MODIS bands studied in Wu, Xiong et al. (2015).

The starting wavelength of the calculation represents the alignment of the detector to these optical requirements. It has been found to be significant for certain bands especially for those like S2 B1 that utilise spectral binning. This latter effect has been also found to introduce discontinuities in the results due to the linear interpolation of the binning requirements (see Figure 5-2). Due to this dependency on the alignment, a study of the rms error to minimise the impact for all the bands in the focal plane has been carried out. Since the TRUTHS design introduces separate focal planes for the ultraviolet (UV), VNIR and, SWIR regions, the rms has been calculated separately for each focal plane.

The TOA reconstruction method and S2 band SRF interpolation were evaluated with the linear, PCHIP and cubic spline interpolation. The interpolation methods showed insignificant differences for the S2 SRF interpolation and found a level of improvement up to 0.2% for the most affected bands due to its greater flexibility in capturing the relevant spectral variations. Further iterations of the tool will include more detailed spectral information at a pixel level, if available, and further discussions on the methods to reconstruct the TOA spectral radiance. Furthermore, a separate study of spectral effect of the SWIR bands should be undertaken and effects like filter contamination discussed.

In addition, a simplified uncertainty propagation of the spectral response knowledge is introduced in Section 5.2.3, where the central wavelength and bandwidth distributions of the bands are fully propagated to show the radiometric impact in the cross-calibration application. The uncertainty levels obtained are below or at 0.1% for all bands with the exception of B1 and B6 for which the lower bandwidth and high TOA spectral radiance variability compared with other bands increases the error up to 0.3% in some cases. The simulation here included only the systematic effects (the normal distribution shift is applied to all detector pixel bands). Further study should provide a better description of non-linear spectral knowledge contributions in the spectrometer. The results of TRUTHS’ spectral knowledge complement the ones studied in Section 4.2 for the impact of Sentinel-2 spectral knowledge, where similar levels of uncertainty were reported. These preliminary results also pointed to a small impact from the spectral degradation of the S2 bands. Other studies as Chander, Helder et al. (2013) have reported an impact of the spectral degradation at the level of 2% for some ETM+ bands. Although a similar desert scene was used, this is considered here a pessimistic value based on the consideration of spectral filter shift up to 5 nm in both directions. The work in Wu, Xiong et al. (2015) proposed a 0.5 nm change based on the changes observed in-flight as derived from MODIS in-flight spectral monitoring (Xiong, Che et al. 2006).
Section 5.2.4 has studied the potential improvement of the spectral sampling/resolution error if the spectral binning were not applied for the spectral region around S2 B1. The results in Figure 5-8 have shown that further sampling of the region would significantly reduce the sampling/resolution error to a level below 0.1% and also lead to an improvement in the spectral knowledge uncertainty. Recalling the results presented in Green (1998) it can be seen how the sensitivity error is largely symmetrical. That suggests that when decreasing the spectral resolution, the sensitivity increases as does the number of samples across the bands which tend to favourably balance the sensitivity error in this simulation. This approach would only increase a little the memory requirements of the mission. Systems like the one developed in Dell’Endice, Nieke et al. (2009) are suggested to be included in a mission like TRUTHS/CLARREO so that the change of spectral binning pattern in-flight can be applied to the specific application and provide further flexibility in the cross-calibration with other sensors.

The use of a MODTRAN simulation at 0.1 cm\(^{-1}\) is justified in order to derive meaningful results of the spectral knowledge impact of the TRUTHS sensor. Small spectral variations in this case are well-captured by the fine MODTRAN simulation. Section 5.2.5 has compared the results in sampling/resolution error when using a MODTRAN spectral radiance at 0.5 nm resolution. The comparison of results in Figure 5-9 for the B1 and B6 has shown the validity of using a spectral resolution of 0.5 nm as in Wu, Xiong et al. (2015). However, when the results are intended to provide an evolution of the error with the array positioning, it shows a more accurate description of the error evolution when using a narrower TOA spectral radiance.

Section 5.2.6 has studied the variability of the spectral sampling/resolution error for different types of sites. The results in Figure 5-10 reveal that the 0.1% is largely maintained for the S2 bands other than B1, B5, and B6. For B5 and B6, the desert simulation has shown to be the worst scenario with all other scene types showing improved values.

### 5.5.2 Discussion: spatial domain

Section 5.3 follows a pragmatic approach similar to the one proposed in Chander, Helder et al. (2013) to study the effect of spatial non-uniformity produced by spatial offsets. Using real EO data with low pixel-to-pixel uncertainty and displacements of the ROI — approximately ±10 km and ±0.4 km in latitude and longitude for Libya-4 and La Crau sites respectively — it has been possible to generate a map of TOA reflectance factor error from the site centre. The association of a position knowledge with a distribution of errors in the image has been processed to generate a site curve that links the TOA reflectance factor uncertainty with spatial positioning knowledge. This curve has been found to be highly linear for the studied cases and can be used as a direct input for the definition of cross-calibration requirements.
The method has been evaluated for both the S2 MSI and L8 OLI sensors in near-coincident cloud-free overpasses. The bands selected for the study are B1, B5 and B8 for L8 OLI — 443, 865, and 2201 nm central wavelength — and B1, B8A and B12 for S2 MSI — 443, 865, and 2190 nm central wavelength — which share an almost coincident SRF shape and positioning. Thus, the results can be cross-validated and provide a reliable result for different parts of the VNIR and SWIR region.

An approximate error of 0.12% for B1, 0.27% for B5 and 0.5% for B12 is calculated for a single overpass over La Crau site. The results for Libya-4 show values below 0.1% for all the studied bands. These values are less than the ones provided by Chander, Helder et al. (2013) most likely due to the impact of dune dispersion as a consequence of a much smaller ROI over the site. This method uses images with a low relative uncertainty (see earlier discussion) however, neither the solar nor the viewing angles are constant within the study area. This means that the variations across the selected area are the result of TOA reflectance changes combined — either in a constructive or destructive manner — with angular changes. These angular variations cannot be expected to be caused by the displacement over La Crau since the displacement of just 400 m represent a very small angular variation; however they could have a larger impact in the studied case of Libya-4. The TOA reflectance factor error map in Figure 5-16 suggest that for S2 B8A and L8 B5, these variations could be attributed to viewing angular variations linked to solar illumination direction. This conclusion is subject to further analyses to understand the effect of such variations. The angular information introduced by both the L8 L1TP and S2 L1C products can be ingested in a model of the site that can estimate these variations.

The agreement between the results for L8 OLI and S2 MSI is excellent with the exception of the L8 B7 and S2 B12 over Libya-4. In this case, the study of the map error in Figure 5-16 and their distribution in Figure 5-18 suggest that the S2 B12 might have an instrument related variation. The comparison of the results with the ones described in Chander, Helder et al. (2013), has found that the derived uncertainty is significantly lower. However, this difference could be largely to the consideration of an uncertainty rather than an error and the use of a significantly larger ROI that minimises the dune dispersion.

The association of a standard deviation as a proxy of an uncertainty $k = 1$ is based in the fact that it represents a 68.27% of the probability error distribution. The simulations suggest that might slightly vary from the normal distribution — see Figure 5-18 — and further analysis might be considered to provide an impact of this effect (BIPM, IEC et al. 2008).

### 5.5.3 Discussion: temporal domain

Section 5.4 studies the effect of angular changes with time and/or any potential loss of ‘knowledge’ of angle as a function of time. Calculating change in TOA reflectance that would occur over a 30 minute
period due to angular variation allows an estimation of potential error (which should be corrected) due to time delay between overpasses of a satellite under test e.g. Landsat-8 OLI and TRUTHS. Two radiative transfer codes were selected — MODTRANv5 and 6SV1 — to assess any differences that might occur due to time lapse between overpasses as a consequence of the radiative code used. The result has shown a 0.2% difference between the simulated corrections for spectral regions dominated by the atmospheric scattering i.e. shorter wavelengths. Of additional interest here was also whether this correction and associated differences could be optimised through improved parameterisation. For this, the dominant parameters, aerosol and water vapour, have been modelled as distributions making use of the values presented in Mishra, Haque et al. (2014) and propagated to the TOA radiance/reflectance factors. The results have shown a minimum impact in the correction factor with levels below 0.1%. Nonetheless, this assumes that the aerosol and water vapour knowledge are perfectly constant during the 30 minutes of the simulation. In considering the overall uncertainty due to knowledge of atmospheric parameters, an analysis of TOA reflectance distribution was performed using the full range of observed atmospheric variations over a 1 year period. The results showed a resultant maximum uncertainty below 1% for B1 and B5 and between 1% and 2% for B7. However, these values represent the uncertainty which would occur without any real correction for atmosphere (assuming worst case annual variations) and most importantly the relatively slow temporal change in atmosphere conditions. For a simultaneous nadir overpass (SNO) cross-calibration, as would be envisaged, the atmospheric variation is likely to be very small and thus has little contribution to the uncertainty. However, if we consider a characterisation of a PICS site the temporal variations of the atmosphere can be considered random providing a large uncertainty improvement over several overpasses even without any knowledge of the atmosphere. It should be noted that as a hyperspectral imager, TRUTHS will be able to make some atmospheric retrievals at the time of overpass and thus correct its own observations. The temporal module could be improved by further varying other factors such as the ozone concentration and temperature and by analysing real observations of atmospheric short-term variations. In addition, the aerosol uncertainty distribution should be further improved. A more refined model should look for a distribution of aerosols that only considers positive values and provides an expected distribution of values as e.g. a log-normal distribution.

The effect of the knowledge in the surface reflectance angular correction has been also studied in Section 5.4.5. The method employed is similar to that described for the atmospheric variation. Here we assume that over a period of 30 minutes, the surface reflectance is invariant; for a PICS site this is probably true for a much longer period, except under extreme conditions such as sand storms. The RPV model in Bouvet (2014) has been modified to introduce a 5% uncertainty on each BRDF parameter. This is a worst-case assumption that does not consider the optimisation of the model that would be possible from the TRUTHS observations. The results at 443 nm show an uncertainty that oscillates between 0.2-0.3% at any time in the year whereas the results at 865 nm range between 0.3-0.4%. If we
consider the combined effect of the atmosphere, then in the worst case, without applying corrections, the uncertainty due to temporal knowledge is well below the level of 0.5%. These variations due to solar angular change have been found to be consistent with the empirical results in Mishra, Helder et al. (2014) for Libya-4 site.

In future updates of this work, it is important that the comparison between the radiative transfer codes is extended to several other algorithms. The atmospheric variations and the impact in a correction should be extended to account for further parameterisation as the ozone content, temperature or pressure. Finally, the surface reflectance correction uncertainty should be upgraded by introducing the impact of the correlation between the different RPV model parameters.

Finally, it is important to mention the benefit of an asynchronous orbit of the reference sensor in terms of temporal effects in cross-calibration. Time delays between the sensor under test and a cross-calibration using TRUTHS or CLARREO will tend to zero as the number of match-ups increase, due to randomness and in turn reducing overall uncertainties. That is whether the systematic uncertainty is produced by the delay between overpasses or in the correction knowledge as studied in Section 5.4, the accumulation of them over different overpasses will tend to reduce the impact.

5.5.4 Discussion: uncertainty budget

Effects due to viewing angle have not been discussed in this paper but since the reference sensor is considered to be agile it can be aligned to match that of the sensor under test.

CLARREO and TRUTHS are designed to have a polarization sensitivity of less than 0.5%, \(k = 2\) below 1000 nm, and less than 0.75% \(k = 2\) above 1000 nm (Wielicki, Young et al. 2013). Even though the sensitivity is low, the degree of polarisation might be certainly high for a certain spectral region, sites and angular configuration. Recent work for CLARREO has shown that desert areas present a degree of polarisation at the 10%-level for longer wavelengths but that can raise up to 50% at the shorter wavelengths (Sun, Baize et al. 2015). In order to account for the polarisation effect in the cross-calibration, a set of Degree Polarisation Models (DPMs) have been derived and a methodology to account for them has been proposed in Lukashin, Wielicki et al. (2013).

Table 5-6 provides a summary of the sources of uncertainty and their relative importance for a range of cross-comparison scenarios using the characteristics of TRUTHS as a reference sensor and Sentinel-2 as the sensor to be calibrated. In this table, the polarisation error and viewing angle effect have not been considered.
Table 5-6. Summary of the different sources of uncertainty investigated for a cross-comparison of TRUTHS and Sentinel-2.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>Resultant Uncertainty on Sentinel-2 TOA reflectance/ %</th>
<th>Resultant uncertainty on Sentinel-2 TOA reflectance/ Mean of multiple overpasses</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best S2 bands</td>
<td>Worst S2 bands</td>
<td>Best S2 bands</td>
</tr>
<tr>
<td></td>
<td>&lt;0.1*</td>
<td>0.1-0.6*</td>
<td>&lt;0.1*</td>
</tr>
<tr>
<td>Spectral resolution/accuracy of TRUTHS</td>
<td>&lt;0.1*</td>
<td>-0.2*</td>
<td>&lt;0.1*</td>
</tr>
<tr>
<td>Spectral knowledge of TRUTHS</td>
<td>0.12</td>
<td>0.5</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Non-spatial co-alignment of TRUTHS with Sensor under test</td>
<td>&lt;0.1*</td>
<td>&lt;0.1*</td>
<td>&lt;0.1*</td>
</tr>
<tr>
<td>Error due to 30 minute difference in overpass times: solar/ view angle: atmosphere</td>
<td>&lt;0.1*</td>
<td>-0.2*</td>
<td>&lt;0.1*</td>
</tr>
<tr>
<td>Error due to 30 minute difference in overpass times: solar/ view angle: atmosphere</td>
<td>0.2-0.4*</td>
<td>0.2-0.4*</td>
<td>&lt;0.1*</td>
</tr>
<tr>
<td>Error due to lack of knowledge of surface BRF for 30 minute difference in overpass time</td>
<td>0.4-0.5</td>
<td>0.4-0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Total achievable uncertainty due to cross-comparison process for anticipated level of knowledge and conditions (uncertainty sources considered marked with asterisk)</td>
<td>0.4-0.5</td>
<td>0.4-0.7</td>
<td>0.2</td>
</tr>
</tbody>
</table>

For all sites and interpolations studied. Worst cases are B1, B5 and B6 but large variability depending on site, interpolation and spectral binning pattern.

For a 0.2 nm k=1 central wavelength and bandwidth knowledge. Worst S2 bands are B6 and B1.

For a single overpass over La Crau site with a positioning knowledge of 32 × 44 m². Best band is B1 and worst one is B12.

For a single overpass over Libya-4 site with a positioning knowledge of 50 × 50 m². Results of Aerosol path difference between 6SV1 and MODTRAN v5 calculated for mean of annual Aerosol optical thickness. Worst case is B1.

Assume mean of annual atmospheric variations for Libya 4 test site: Requires knowledge of atmosphere around time of overpass. Can be reduced further by correction of time difference. Equivalent to no knowledge of atmosphere parameters, but can be reduced by knowledge of time difference. Worst case in the SWIR bands. Model would improve with experimental data from TRUTHS and also with multiple measurements effect would tend to zero. Impact at TOA is assumed half for B1.

Reduction in uncertainty due to multiple overpasses would include multiple sites and does not take account of correlations.
5.6 Conclusion

This chapter presents a rigorous approach to evaluate the sources and quantify the uncertainty in post launch Level 1 radiometric gain obtainable from sensor-to-sensor cross-calibration. The novel approach analyses the derived probability distributions for the three main error domains: spectral; spatial; and temporal. Thus, the approach seeks to move from a sensitivity or error analysis into a description of the distribution of errors and, consequently, of uncertainty estimates. The study complements the analysis presented in Chapter 3 where an in-flight diffuser is considered as the primary method for radiometric calibration. It is common practice that a sensor-to-sensor cross-calibration or other vicarious methodologies become the primary method of radiometric calibration for several missions. Thus, it is important that where these vicarious methods are considered, rigorous uncertainty estimates are provided.

The analysis for the considered sites shows that a worst case cross-calibration uncertainty (at $k = 1$) below or at 0.5% can be achieved for a single match-up for each of the three domains for the majority of the overpasses and satellite-to-satellite matching conditions. These values indicate that missions like the proposed TRUTHS or CLARREO with an SI-traceable accuracy above 0.3% (at $k = 2$) and in an asynchronous near-polar orbit, would mean that the reference sensor calibration would no longer be the dominant source of uncertainty in sensor to sensor radiometric cross-calibration. Instead the sensor accuracy would be comparable to the spectral, spatial and temporal uncertainty contributions and could, with the right conditions and averaging over different match-ups, achieve overall uncertainties of $\leq 0.5\%$.
Chapter 6.

The correlation of the TOA reflectance/radiance pixel measurements

6.1 Introduction

The first version of the Sentinel-2 Radiometric Uncertainty Tool (S2-RUTv1) presented in Chapter 3 provides calculations of the uncertainty per pixel of the S2 Level-1C (L1C) Top-of-atmosphere (TOA) reflectance factor images derived from the Multi-Spectral Imager (MSI) on-board S2. Such pixel-level uncertainty information can be directly applied as a quality indicator of the S2 L1C products.

However, when propagating pixel-level TOA radiance or reflectance factor products to higher levels in a processing chain, this pixel-level radiometric uncertainty must be treated carefully. Many applications of higher-level products aggregate data from different pixels in space and/or time using a simple, or a weighted, mean. To determine the uncertainty associated with the mean it is not sufficient to know the uncertainty associated with a single pixel value, it is also necessary to consider whether there are systematic effects leading to common errors between different pixels. Similarly, higher level products also often involve combining data from different spectral bands. Again it is essential to understand whether there are systematic effects leading to common errors between different spectral bands.

In this chapter we consider ways of estimating the error correlation structure in spatial, temporal and spectral dimensions. This case study is based on different Regions of Interest (ROIs) used in the radiometric validation of S2. This case study was chosen because it is a direct application of L1C products and their uncertainties and is of current interest.

Radiometric validation is a process that involves comparing the instrument under test with another reference measurement or model of the TOA radiance/reflectance factor. When both the instrument under test and the reference have an associated uncertainty estimate, it is possible to validate the test

---

instrument’s uncertainty analysis using the performance criterion that the two should agree within their combined uncertainties (usually at the 95 % confidence interval).

In order to reduce the effects of noise and/or to allow comparisons of sensors to references with a different spatial pixel size, such comparisons are usually performed by averaging over a specific ROI to obtain a “ROI best estimate”. This ROI is selected on the basis of certain criteria, such as a minimum site uniformity and viewing angle dispersion.

This chapter defines a method for using the S2-RUTv1 that can provide an uncertainty estimate of the mean of the ROI for radiometric validation purposes. In Section 6.2 we describe the concepts behind estimating the uncertainty associated with the ROI mean. We consider correlation in spatial, temporal and spectral dimensions and show both a robust Monte Carlo Method (MCM; (BIPM, IEC et al. 2008)) and how an estimate of correlation can be obtained using the existing S2-RUTv1 by selecting and deselecting different uncertainty components. Note that the on-going development of the S2-RUT means that later versions are likely to include a means of providing pixel correlation information directly as discussed in Chapter 3. Section 6.3 discusses the different sources of uncertainty in turn and considers error correlation structures in spatial, temporal and spectral dimensions. There are uncertainty effects that were not included in the S2-RUTv1 (Gorroño, Fomferra et al. 2017); Section 6.4 discusses how significant these effects may be in determining the uncertainty associated with the mean value of a ROI. Finally, Section 6.5 provides an example for three locations used in radiometric validation of S2: the radiometric calibration network (RadCalNet) site at Gobabeb, the Boussole ocean buoy site and deep convective clouds (DCCs).

6.2 Concept of study, the limitations and the methodology

6.2.1 Error, uncertainty and correlation

The S2-RUTv1 provides users with the S2 L1C radiometric uncertainty per pixel. This expresses the degree of doubt around the TOA reflectance factor measured at each pixel or, an interval around the TOA reflectance factor that encompasses a certain fraction of the distribution of values that could be attributed to the measured quantity (BIPM, IEC et al. 2008). The actual error for a single pixel, that is the measured value minus the true quantity value (BIPM, IEC et al. 2012), is unknown, but is drawn from the probability distribution described by the uncertainty. In practice the uncertainty (and resultant error) is caused by the combination of individual uncertainties associated with different effects, which are combined to provide an overall uncertainty.

Although we cannot know the error associated with any given effect for a given pixel, we can evaluate whether that error is likely to be the same for different pixels, times, or spectral bands. It is this common error that creates error correlation.
6.2.2 Uncertainty over a ROI pixel mean

The direct results of the S2-RUTv1 cannot be directly used to determine the uncertainty associated with a ROI pixels’ mean. Neither the mean of the pixel uncertainties nor the standard deviation of the mean uncertainty represent a general scenario. Indeed, these two cases would correspond to the boundary scenarios where all pixels and contributions to the L1C radiometric uncertainty are positively correlated and fully uncorrelated respectively.

Consider the mean of 2 pixels that are scanned at two consecutive lines; effectively a 2 × 2 ROI. The equation to obtain the mean of the ROI pixels $\rho_{ROI}$ is:

$$\rho_{ROI} = \frac{1}{4} (\rho_A + \rho_B + \rho_C + \rho_D) \quad (6.1)$$

where each term $\rho_i$ represents a pixel indexed in the spatial (across track) and temporal (along track) dimensions.

According to the matrix-form of the Law of Propagation of Uncertainty (BIPM, IEC et al. 2008), the variance associated to the term $\rho_{ROI}$ is given by:

$$u^2_{ROI} = C^T U C \quad (6.2)$$

where the vector of sensitivity coefficients, $C$, is given by:

$$C = \left[ \frac{\partial \rho_{ROI}}{\partial \rho_A}, \frac{\partial \rho_{ROI}}{\partial \rho_B}, \frac{\partial \rho_{ROI}}{\partial \rho_C}, \frac{\partial \rho_{ROI}}{\partial \rho_D} \right]^T = \left[ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right]^T \quad (6.3)$$

and the covariance matrix $U$ is given by combining a vector of individual uncertainties $V$ with a matrix of correlation coefficients $R$, through:

$$U = VRV^T$$

$$R = \begin{bmatrix} 1 & r_{AB} & r_{AC} & r_{AD} \\ r_{BA} & 1 & r_{BC} & r_{BD} \\ r_{CA} & r_{CB} & 1 & r_{CD} \\ r_{DA} & r_{DB} & r_{DC} & 1 \end{bmatrix} \quad V = \begin{bmatrix} u_A & 0 & 0 & 0 \\ 0 & u_B & 0 & 0 \\ 0 & 0 & u_C & 0 \\ 0 & 0 & 0 & u_D \end{bmatrix} \quad (6.4)$$

The matrix operation in Equation 6.2 can be rewritten as:

$$u^2_{ROI} = \frac{1}{16} \left[ u_A^2 + u_B^2 + u_C^2 + u_D^2 \right] + \frac{2}{16} \left[ u_A u_B r_{AB} + u_A u_C r_{AC} + u_A u_D r_{AD} + u_B u_C r_{BC} + u_B u_D r_{BD} + u_C u_D r_{CD} \right] \quad (6.5)$$

118
The terms in $V$, providing the uncertainty associated to each pixel can be directly obtained by running the S2-RUTv1. However, the correlation terms, in $R$, cannot be directly obtained. Indeed, the level of correlation between pixels will depend on the scene type and/or the acquisition time.

For example, pixels measuring an ocean scene are expected to be largely uncorrelated due to the dominance of instrument noise over such low radiance scenes. On the other hand, pixels in a bright cloud scene will be highly correlated due to the dominant effect of systematic and largely correlated errors introduced by the radiometric calibration.

Thus, the pixel correlation in $R$ is a consequence of the balance between the different uncertainty contributors of the S2 L1C uncertainty. To understand common effects, and hence the correlation coefficient for a particular scene, we have to consider the instrument characteristics and ground processing.

### 6.2.3 Spatial, temporal and spectral dimension of the S2 MSI

The focal plane of the S2 MSI instrument consist of 12 detectors in staggered formation (see Section 2.2). Figure 6-1 presents a schematic of the S2A VNIR focal plane:

![Figure 6-1. VNIR focal plane schematic description. The image is reproduced with permission from (Gascon, Bouzinac et al. 2017).](image)

Each detector contains several lines of pixels, each of which have a filter on top, which defines the spectral band of each detector line. For the SWIR bands, several such detector lines are combined with Time Delayed Integration (TDI). Each detector line has a total of 1296 pixels for the 10 m bands and 2592 for the 20 m and 60 m bands. The 60 m bands are obtained by spatial binning of the original 20 m pixels.
Thus, the spatial dimension is defined by the across-track (ACT) dimension of the focal plane and the temporal dimension is the result of the successive acquisitions of the pixels in combination with the satellite motion. The spectral dimension is defined by the spectral lines on the detector.

The effects of TDI and spatial binning will not be considered further since these processes are performed during the ground processing and the performance parameters, such as the instrument noise parameters or quantisation level, already include these effects. The discontinuity between detectors has an impact on the viewing angle continuity or the correlation between pixels. The ROIs studied here comprise a small fraction of the detector and are considered to be included in just one detector.

It is important to note that the discussion here does not fully account for the orthorectification process applied to the S2 Level-1B (L1B) products. The L1C products provide radiometrically-corrected imagery with digital numbers (DN) proportional to TOA radiance values and in sensor geometry (ESA 2017). This process consists on a B-spline interpolation of the L1B DN prior to their conversion to reflectance factor. This process fits the measurements onto an Earth grid that accounts for an elevation model with equally spaced sampling in Universe Transverse Mercator (UTM) coordinates.

As a result, the S2 L1C products do not keep the original spatial and temporal focal plane dimensions. However, in the absence of more detailed information, the North-up orientation of the L1C products is used to approximate the temporal dimension (North-South) as well as of the spatial dimension (East-West). This simplification is reasonable as the S2A orbit is a near-polar orbit with a 98.62° inclination.

6.2.4 An approximation method: “select/deselect”

The “select/deselect” method uses the capability of the S2-RUTv1 to generate uncertainty images for selected uncertainty contributors to estimate the total uncertainty of the average reflectance of a ROI.

In this approach we assign each effect to being either correlated (not reduced by averaging over the ROI), or uncorrelated (reduced by averaging over the ROI) and select only the correlated effects, assuming that the uncorrelated effects become negligible at the scale of the ROI. Each effect is discussed in Section 6.3, which describes how decisions were made. Note that effects may be correlated in one dimension (e.g. spatial) and uncorrelated in another (e.g. temporal or spectral). The associated correlation should be evaluated experimentally if feasible by studying the combined variations of the quantities or using any available additional data pertaining to their interrelationship. Additionally, or in the absence of available data information based on experience and general knowledge can be utilised (BIPM, IEC et al. 2008). For the study in this chapter, the correlation will be based mostly in the latter method with the intention that, with access to further information and/or experimental data, a refinement of the values can be undertaken.
The method named “select/deselect” is intended to be as simple and quick as possible for the S2 L1C data users. In this method, the user interface incorporates a tick option to select individual effects, as shown in Figure 6-2.

![Image of the S2-RUTv1 dialog box with the tab “Processing parameters” selected. This tab permits the selection and deselection of each uncertainty contribution.](image)

The approach is very simple but has several limitations. For example, for ROIs of just a few pixels, the assumption that the random effects become insignificant may not be sufficient. Thus, this method must be tested to understand the validity of the ROI size at which this assumption is valid. The method also does not provide flexibility to cope with situations where the effect cannot be considered either perfectly correlated or perfectly uncorrelated. In such cases, the method has been adapted to produce two uncertainty images with the partially correlated contributions selected and deselected. The result is taken as the mean of the two ROI pixels for the two images.

### 6.2.5 MCM propagation

In order to understand the potential limitations of the approximate approach, a comparison was made with a more rigorous method based on MCM propagation. The MCM determines the mean TOA
reflectance factor for a ROI from the pixels over many iterations. At each iteration, the error associated with the reflectance factor is drawn from the distribution of each uncertainty contribution. If the uncertainty contribution is correlated between the pixels, the same sample is used for all the pixels in the ROI, whereas, if the uncertainty contribution is uncorrelated, a different error is drawn from the distribution for each pixel. Where there is partial correlation, two separate errors are drawn, one that is common to all pixels and one which is different from pixel to pixel. The distributions are set as normal or uniform distributions with a spread of values directly linked to the uncertainty as calculated directly from the S2-RUTv1. This uncertainty is obtained by generating an image of the specific uncertainty contribution. The method is illustrated in Figure 6-3.

![Figure 6-3. Schematic of the MCM propagation for the ROI mean uncertainty estimate.](image)

The errors $\delta_{\text{corr}_1}$ and $\delta_{\text{corr}_2}$ in Figure 6-3 represent the two extreme cases for correlated and uncorrelated pixel errors. The fully correlated case gives the same error in each pixel. Alternatively, the fully uncorrelated case gets an independent error for each pixel. Intermediate cases are also possible.

### 6.3 Qualitative assessment of the pixel-to-pixel correlation

This section describes each one of the uncertainty contributions integrated in the S2-RUTv1. A full description of how these are determined is given in Chapter 3. Here the description is focused on the correlation structure.

#### 6.3.1 $u_{\text{noise}}$: Instrument noise

The instrument noise model is characterised in-flight by the calculation of the variance of dark signal (DS) and diffuser measurements. The noise model takes the DS standard deviation as the instrument
noise in the absence of light. This model is scaled by the diffuser variance — see Section 3.3.2 — under the assumption that the increase of the noise with the light intensity is linear to the variance as if dominated by the photon shot noise.

In the temporal domain, the instrument noise can be considered completely uncorrelated between the acquisition lines. This point has been demonstrated in Gorroño, Gascon et al. (2015) for the S2 dark signal measurement using the Allan deviation (Allan 1966, Allan 1987), which disentangles higher and lower frequency components of the noise and provides an estimate of the upper bound of independent samples. The results showed that for the VNIR bands the number of independent dark samples was well above 1000. For the SWIR bands the number of independent samples could be more variable with some pixels showing values as low as 500 independent samples. For the validation activities using RadCalNet sites, ocean or deep convective clouds, the amount of temporal lines required is well below any critical limit. For example, even for the largest RadCalNet sites (1 × 1 km), the number of pixels in one dimension will be at most 100.

The VNIR detectors are monolithic complementary metal oxide semiconductor (CMOS) detectors, while the SWIR detectors are on Mercury Cadmium Telluride (MCT) detectors hybridised on a CMOS readout circuit (Drusch, Del Bello et al. 2012). This means that up to the pre-amp and voltage conversion, the noise is independent from one pixel to another. The voltage is further amplified at the front-end electronics (FEE) and video-chain unit (VCU). Thus, in the spatial and spectral dimension, the instrument noise can be considered independent between samples under the assumption that noise introduced by the post-amplification is not dominant. Note that at the FEE and VCU units, the coupling between signals may exist. This is independently accounted for in Section 6.3.4.

### 6.3.2 $u_{\text{stray}}$: Out-of-field Stray-light — systematic part

The uncertainty contribution due to out-of-field stray light is added linearly in the proposed S2 L1C uncertainty budget in Section 3.4.

The stray light generated by the Earth out-of-field contribution might vary due to the variation of the scene during the orbit. However, for any level of stray-light experienced, the result will be largely homogeneous over the VNIR and SWIR focal planes. The mirrors and splitter reflections tend to spread the stray-light entering the MSI instrument. The effect also does not arise from one source point but from a more extended source at each side apart from the focal plane, which further spreads the stray-light across the focal planes. The filters generate other non-uniform stray-light events such as ghosts that are accounted for in a case-by-case basis and are avoided or minimised for validation activities (e.g. by discarding products with a large extent of clouds in the ROI pixel vicinity).
A global figure of 0.3% of the reference Earth radiance will be used as a ROI mean error. The uncertainty combination in Section 3.4 proposed a linear combination as a result of non-corrected systematic effect. This uncertainty is considered appropriate for land validation areas where the vicinity of the swath is expected to be dominated by land. However, the error estimate might be conservative for ocean measurements if the vicinity of the swath is dominated by a land scene.

6.3.3 $u_{\text{stray.rand}}$: Out-of-field Stray-light — random part

The random component of stray light (i.e. that which varies across the focal plane) is produced by the lack of light tightness of the focal plane.

In the spatial dimension the effect is considered as fully uncorrelated. Experimental results in laboratory found out that the level was varying randomly from pixel-to-pixel in the ACT dimension.

In the spectral dimension, the effect is largely uncorrelated since the error introduce by the light tightness between two pixels in different spectral lines can be considered independent.

However, in the temporal dimension the effect is largely correlated as the radiometric validation images are taken over uniform scenes. Very similar illumination conditions apply and, thus, very similar levels of light tightness for the same pixel are expected over the temporal scan lines.

6.3.4 $u_{\text{xtalk}}$: Crosstalk

The contribution for crosstalk arises from the electrical signal coupling between spectral bands. The effect has an impact in the TOA radiometric performance for the SWIR bands. Due to the absence of a correction at the time of releasing the S2-RUTv1, the approach was to include the worst figures of contamination between bands (see Section 3.3.2). However, a correction is now applied to S2 L1C images for this effect with the deterministic nature of the correction justifying a low correction residual hence this effect is not considered further here. Furthermore, the radiometric validation exercise presented here is carried out over uniform scenes which have been chosen to avoid nearby clouds (except for DCC methods).

6.3.5 $u_{\text{ADC}}$: Analog-to-digital conversion quantisation

The analog-to-digital (ADC) conversion at the VCU units has been modelled as an error distribution with an amplitude of 1/2 Least Significant Bit (LSB) and rectangular distribution.

This uncertainty is expected to be uncorrelated in the spectral, spatial and temporal dimensions, however, there are two possible problems with this assumption. Firstly, the ADC conversion is shared at the VCU unit for several channels, however, this does not affect the uncorrelated nature of the uncertainty if the ADC does not introduce a large systematic effect. Secondly, the radiometric validation
sites are selected on the basis of a radiometric uniformity and this could result in a correlated rounding between pixels in the temporal and spatial dimensions. Fortunately, the digitisation is performed with 12 bits meaning that even small scene variations represent a large variation in terms of LSB units.

6.3.6 \textit{u}_{DS}: Dark signal stability

The VNIR focal plane is not temperature controlled and an approximation for the variations with temperatures can be described as follows (Hopkinson, Goodman et al. 2004):

\[ DS(T_0 + \Delta T) = DS(T_0) \cdot e^{\frac{E_{\text{act}}}{k \Delta T}} \]  \hspace{1cm} (6.6)

where $k$ is the Boltzmann constant (8.602.10^{-5} eV), $\Delta T$ is the variation of temperature in kelvin and $E_{\text{act}}$ is the activation energy (approximately 0.63 eV in silicon detectors).

The HgCdTe detectors are passively cooled due to the larger sensitivity of the dark signal with temperature variations (Dariel, Chorier et al. 2009). In addition, the offset level for the SWIR region could be marginally affected by the residual thermal emission.

The validation activities using RadCalNet sites, ocean or DCCs require the reading of a small fraction of pixels in the detector. In this situation, it is plausible to assume that the temperature gradient in the selected pixels will be very low and, if any thermal emission effect exists, it will be similar in the local vicinity. Thus, the dark signal stability is expected to be correlated in the spatial and temporal dimensions.

In the spectral dimension, the temperature variations cannot be assumed as perfectly correlated as the spectral lines are physically separated in the detector — or integrated in different VNIR and SWIR focal planes — and gradients of temperature along the detector could occur. In that case, a correlation of 0.5 can be considered.

6.3.7 \textit{u}_{\text{gamma}}: Non-linearity and non-uniformity knowledge

The knowledge over the non-linearity and non-uniformity correction $\gamma(p,b,d,Y(p,l,b,d))$ is given as 0.4% for the S2-RUTv1 (see Section 3.3.2).

Since all the pixels in the ROI measure a similar radiance level at calibration sites, the non-linearity estimate can be considered highly correlated over the spatial and temporal dimension. In the spectral dimension, a high correlation can be expected assuming that the pre-flight characterisation is dominated by drifts and spectrally correlated sources (Gardner 2004).

The knowledge of the non-uniformity correction is largely limited by the focal plane noise (FPN). The residual after this correction has a random nature across the field-of-view. Thus, in the spatial dimension
its effect is uncorrelated between pixels. The same parameters and similar radiance levels are measured between pixels in the temporal dimension and, thus, can be assumed as correlated error. The non-uniformity correction residual for different pixels in different spectral lines must be considered as partly correlated since the residual is not expected to be proportional on a one-to-one pixel basis but lower frequency components are expected to be largely correlated.

Depending on the validation scene (e.g. level of radiance), either the non-linearity residual or non-uniformity residual will dominate the uncertainty contribution. The final correlation figures showed in Table 6-1 are set as an intermediate case with the expectation that in future refinements a separate uncertainty for the non-linearity and non-uniformity can be assessed.

6.3.8 \[ u_{\text{diff.abs}}: \text{Diffuser reflectance absolute knowledge} \]

The diffuser reflectance factor is obtained through a complex process which involves the interpolation and/or fitting of on-ground measurements and its convolution to account for a different pupil projection of the on-ground measurements and the MSI measurements and is reported in Mazy, Camus et al. (2013).

The bidirectional reflectance factor (BRF) uncertainty accounts for several separate effects: BRF absolute characterisation knowledge, BRF spatial knowledge, polarisation effect, diffuser illumination and viewing angle knowledge, and other effects — e.g. thermal cycling impact (see Section 3.3.2).

The BRF absolute characterisation is dominated by systematic effects that are translated in a correlated nature between pixels. The propagation of the radiometric standards through the traceability chain includes contributions associated to long-term drifts. As a consequence, the BRF absolute characterisation is expected to be dominated by systematic effects that are largely correlated in the spectral domain (Gardner 2004).

For two pixels separated by a small distance in the focal plane, the pupil projections on the diffuser overlap significantly (the pupil projection is shown in Mazy, Camus et al. (2013)). Consequently, effects such as uniformity and polarisation error will be correlated between the ROI pixels. Similarly, the BRF angular knowledge is correlated due to the similarity in the viewing angle between pixels separated by a small fraction of the focal plane.

Based on the previous reasoning, this contribution is considered correlated for the ROI pixels (spatial and time dimensions) and between the bands (spectral dimension).

6.3.9 \[ u_{\text{diff.temp}}: \text{Diffuser reflectance temporal knowledge} \]

This contribution is also considered as a linear addition in Chapter 3 as for the Section 6.3.2.
In this specific case, the error estimate was based on a dedicated on-ground qualification programme (J. Nieke, personal communication, August 3, 2017). Further, in-flight experience with similar diffuser material indicates a comparable degradation estimates as described for the MERIS instrument in Delwart and Bourg (2011) allowing an interpolation to the S2 acquisition timestamp. The degradation rate for the pixels in the ROI is expected to be very similar since, again, the pupil projection of the ROI pixels is very similar and any rate gradient will be negligible. Thus, the estimated error for this contribution can be also considered as the mean error over the ROI.

6.3.10 $u_{\text{diff, cos}}$: Angular diffuser knowledge—cosine effect

The angular knowledge over the cosine correction is limited by the diffuser planarity or lack of knowledge of the diffuser angular coordinates. The error from the measured value between two pixels in the selected ROI, will be of similar level and sign since the pupil projection of two pixels in a ROI and across different bands is shared at a large extent for a small ROI (similar discussion in Section 6.3.8). Therefore, the uncertainty is largely correlated between pixels in any of the three dimensions.

6.3.11 $u_{\text{diff, k}}$: Straylight in calibration mode—residual

This contribution arises from the imperfect knowledge of a bias correction due to the multiple reflections of the Sun and Earth illumination introduced during the Sun diffuser calibration. The bias has been estimated at 0.7% of the calibration radiance level for all bands — see (Gascon, Bouzinac et al. 2017) — and a conservative residual of 0.3% has been associated to the correction knowledge (see Section 3.3.2).

The stray-light during calibration has been characterised both in absolute terms and also in terms of its uniformity level across the focal plane. Similarly to the contribution in Section 6.3.2, the combination of rays tends to homogenise the effect across the focal plane. The major source of stray-light in this situation is the Sun that can be largely considered as a point source. Nonetheless, since the contribution arises after the multiple reflections from the Calibration and Shutter Mechanism (CSM), the effect must be considered as the combination of several scattered rays entering the optical system. Nevertheless, this type of residual is not directly accounted for by the absolute calibration coefficient $A(b)$ but as part of the FPN noise at the non-uniformity correction in $\gamma(p,b,d,Y(p,l,b,d))$ (see section 6.3.7). With a conservative residual of ±0.3%, the differences in the focal plane cannot be considered the dominant limitation of the correction knowledge but the absolute characterisation of the stray-light levels. Furthermore, if the ROI used for validation is at $1 \times 1$ km or smaller, these variations will be even lower. Thus, it is justified to assume that the uncertainty residual will be largely correlated in the spatial dimension as a consequence of a “common” absolute error.
In the spectral dimension, the residual is expected to be partly correlated. The use of a common bias for any of the 13 bands implies that the residual errors will fluctuate around a normal distribution with a scale of 0.3%. Between two spectrally adjacent bands, the fluctuation over the residual will be very similar. However, the more spectrally distant the bands are — e.g. B1 and B12 —, the more uncorrelated the residual is expected to be.

In the temporal dimension, the stray-light residual is fully correlated since the effect applies to the calibration coefficient $A_{k,NDI}$ which is constant in this dimension.

### 6.3.12 $u_{\text{ref_quant}}$: L1C Image quantisation

The L1C images of reflectance factors are codified in JPEG2000 format with a maximum number of 16 bits. The rounding effect has been discussed to be very low in relative units ($<<0.1\%$) except for very low reflectance values (e.g. ocean scenes). The nature of this uncertainty contribution is to be uncorrelated on a pixel-to-pixel basis.

### 6.3.13 Summary pixel correlation for validation sites

Table 6-1. Summary of pixel correlation for radiometric validation sites

<table>
<thead>
<tr>
<th>S2-RUTv1 uncertainty contributors</th>
<th>Spatial</th>
<th>Temporal</th>
<th>Spectral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument noise, $u_{\text{noise}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Out-of-field Stray-light – systematic part, $u_{\text{stray_sys}}$</td>
<td>ROI constant error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-of-field Stray-light – random part, $u_{\text{stray_rand}}$</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Crosstalk, $u_{\text{crosstalk}}$</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>ADC quantisation, $u_{\text{ADC}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dark signal stability, $u_{\text{DS}}$</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Non-linearity and non-uniformity knowledge, $u_{\text{gamma}}$</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Diffuser reflectance absolute knowledge, $u_{\text{diff_abs}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Diffuser reflectance temporal knowledge, $u_{\text{diff_temp}}$</td>
<td>ROI constant error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angular diffuser knowledge – cosine effect, $u_{\text{diff_cos}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Straylight in calibration mode – residual, $u_{\text{diff_k}}$</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>L1C Image quantisation, $u_{\text{ref_quant}}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### 6.4 Impact of non-included contributors in a calibration site

This section considers the impact of uncertainty contributions that were not included in the S2-RUTv1 (see Chapter 3).
6.4.1 Deconvolution residual and other sources of straylight

The Point Spread Function (PSF) describes the response of the imaging system to a point source. The correction of the image for the PSF is planned as part of the S2 L1C processing but has not yet been implemented in the operational product.

The radiometrically uniform nature of validation sites largely minimises this effect because the signal that is lost by one pixel towards its neighbours is compensated by a gain from the neighbouring pixels’ optical path towards it. The same reasoning is applied for a diffuser calibration where the uniformity of the diffuser source minimises the internal stray-light and only the out-of-field component must be accounted for (see Section 6.3.11).

Ghosting effects due to the effect of crosstalk have been previously discussed in Section 6.3.4. However, a further source of ghosting can be identified due to the filter inter-reflections. This effect can be minimised when the ROI selected during validation is carefully selected to avoid images with high radiance clouds nearby.

6.4.2 Polarisation error

This is an error introduced by the sensitivity of the instrument to a difference in the light polarisation. If both the response of the instrument and the TOA signal are characterised for the polarised components of the light, this error can be corrected for. In the absence of such an information, it must be treated as an uncertainty contribution.

The polarisation sensitivity of the MSI instrument is <3% and the expected degree of polarisation (DoP) for scenes such as the RadCalNet Gobabeb site or the DCC is below 10%. However for bands dominated by larger atmospheric components such as B1, the degree of polarisation could be significantly higher (Sun, Baize et al. 2015).

For ocean scenes, the variations of DoP will be different with wavelength, angular configuration, as well as the scene characteristics such as wind speed or aerosols (Sun and Lukashin 2013). Thus, depending on the previously mentioned factors, the polarisation error can range from an almost negligible to an important effect when close to the polarisation sensitivity of the instrument.

6.4.3 Orthorectification uncertainty propagation

This processing step is introduced at the S2 L1C products and has been discussed in Section 6.2.3. The radiometric interpolation has two effects on the S2 L1C radiometric uncertainty.
First, orthorectification will reduce the uncorrelated components of the TOA radiometric uncertainty (see Section 4.3). The impact on the uncertainty of a mean ROI however is that the uncertainty values will, in real terms, converge faster to a minimum value than the predictions using the S2-RUTv1.

Second, there is an uncertainty associated with the interpolation inherent in orthorectification. This is expected to be negligible for radiometric validation sites since the radiometric variations of these sites are expected to be low.

6.4.4 Non-uniformity spectral residual

The non-uniformity correction is updated by deploying the diffuser in-flight. There is a disagreement between the Sun spectral signature and the TOA spectral signature that introduces a systematic effect. A reference to the potential impact of this contribution can be found in Barsi, Lee et al. (2014) for the Landsat-8 (L8) Optical Land Imager (OLI).

The results presented for both soil and vegetation shows a spectral residual below 0.2% for any case and a root mean square (RMS) value below the 0.1%. Thus, for a similar sensor as the S2 MSI, the impact of this contribution is to be expected low. Nonetheless, a detailed study for this mission is required that provides specific values.

6.4.5 Spectral knowledge

This contribution is the consequence of the imperfect knowledge of the spectral response characterised pre-flight and its potential variation during launch and once on orbit.

The work in Gorroño, Banks et al. (2016) presented a preliminary assessment for this effect. The results show that assuming a spectral response uncertainty of 0.2 nm \((k = 1)\) for systematic and 0.1 nm \((k = 1)\) for random spectral calibration knowledge, the dispersion of the data was below 0.3% using a desert site as example. Nonetheless, the study did not consider the equalisation (as part of the gamma correction) of the S2 L1C data and the impact of the pre-flight knowledge is expected to be significantly lower.

In the study to describe the potential impact of spectral variations, a spectral shift was added to the spectral response function (SRF) bands of S2A. As indicative values, the degradation rates measured in-flight by the Spectroradiometric Calibration Assembly (SRCA) on-board the Terra MODIS mission

\[5\] The coverage factor, \(k\), is a numerical factor used as a multiplier of the combined standard uncertainty in order to specify the fraction of the probability distribution that the uncertainty represents BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML (2008). Guide to the Expression of Uncertainty in Measurement, JCGM 100:2008.
(Xiong, Che et al. 2006) were used. The approximate values used for the S2 VNIR bands are: −0.33 nm (B1), −0.26 nm (B2), 0.04 nm (B3), −0.03 nm (B4), −0.05 nm (B5), −0.07 nm (B6), 0.1 nm (B7), 0.2 nm (B8), and −0.18 nm (B8A). The error in reflectance factor for all the bands was below the 0.1% with the exception of B1 which increased to 0.3%. The SWIR bands are not included since they represent a specific case where icing introduces an additional interference and specific evaluation is required.

As described in Section 6.4.4, the impact of these contributions is not expected to be significant for the uncertainty budget of most radiometric validation sites but they require a specific scene evaluation to provide quantitative figures and integrate them in future version of the S2-RUTv1.

### 6.4.6 Geometric knowledge

Geometric knowledge describes both the impact of geolocation accuracy and the angular dispersion of the observed pixels in a ROI.

The geolocation knowledge results in Gascon, Bouzinac et al. (2017) describe values either for refined or non-refined S2 L1C products below the 12.5 m ($k = 2$) specification. Section 5.3 describes the impact of this uncertainty as a function of the geolocation knowledge for the Libya-4 and La Crau radiometric calibration sites. With the reported geolocation knowledge of the S2A L1C products the expected impact for the calibration sites is expected to be generally below the 0.1%.

A fair approximation of the angle dispersion of the pixels in a ROI can be obtained if a linear relationship is applied to the full swath angular dispersion. For a full swath, there are 14376 pixels across 20 m bands — from Figure 6-1 (1296 pixels/detector - 98 blind pixels) × 12 detectors. If a linear relationship is assumed, the 20.6° field-of-view of the instrument, a ROI of 20 × 20 pixels of 20 m bands (the example represents a 400 m ROI) covers an estimated 0.07° peak-to-peak. Thus, the effect of this angle dispersion can be expected as negligible unless the target scene and the angular configuration is specifically affected by the BRDF hot-spot.

### 6.5 An approximation to the ROI pixel mean uncertainty using the S2-RUTv1

#### 6.5.1 Case study locations

The study is performed over three different sites that correspond to different methodologies of TOA radiometric validation: the RadCalNet site at Gobabeb, the Boussole buoy site, and DCCs. These sites have a different balance of uncertainty contributions and were chosen to determine how this changes
the level of pixel correlation for each case. Here there is no discussion of the uncertainty associated to
the validation method themselves, simply the S2 MSI uncertainty estimates.

RadCalNet, (www.radcalnet.org), once it becomes fully operational, will provide users with an
operational (routine) service for nadir-view TOA reflectance factor data from several instrumented
ground sites in the spectral region 400 nm to 1000 nm or 2500 nm, depending on available
instrumentation. The site-measured surface reflectance and atmospheric data are propagated to TOA
through a common processing chain by NASA-Goddard using MODTRAN (MODe rate resolution
atmospheric TRANsmission).

As part of the RadCalNet prototype phase, a new site is being established jointly between ESA, CNES
and NPL near to the Gobabeb research station in Namibia (Bialek, Greenwell et al. 2016). The site is
in the gravel plains at the edge of the Namib Desert and was chosen through an extensive search for a
site with high spatial uniformity and stable atmospheric conditions on a flat location (Bialek, Greenwell
et al. 2016). The ground monitoring instrument installed at Gobabeb (in July 2017) on a 10 m high
mast is an adapted CIMEL CE 318 BRDF 12-filter Sun Photometer which measures in 12 spectral
bands from 414 nm to 1640 nm. The instrument takes measurements in a pre-determined sequence,
scanning across the ground and sky. The data is processed by fitting the reflectance data to a BRDF
Roujean model (Roujean, Leroy et al. 1992) and extracting the nadir data to provide the surface
reflectance for input to the RadCalNet portal (Meygret, Santer et al. 2011).

The Boussole buoy site is a superstructure deployed in the deep waters (~2400 m) of the northwestern
Mediterranean Sea (7°54’E, 43°22’N). It is composed of radiometers at above surface, 4 m, and 9 m
depth and additional set of instruments as fluorometers or backscattering meters. All this
instrumentation provides the necessary inputs to estimate the water leaving radiance and its further
normalised water leaving reflectance. These quantities have been directly used for the vicarious
radiometric calibration of satellite ocean colour sensors but are also applicable to the validation of
Level-2 biophysical products and the long-term monitoring of ocean colour missions and site
radiometric properties (Antoine, d’Ortenzio et al. 2008, Antoine, Guevel et al. 2008).

DCCs are very vertically-extended and opaque clouds with very bright and cold tops close to the tropical
tropopause. Their reflectance spectrum, after correction of stratospheric gaseous absorption if seen from
space, is near lambertian and very spectrally flat in the VIS with amplitude primarily driven by cloud
optical thickness. Their daily occurrence within the intertropical convergence zone as well as their large
horizontal extent allow high rates of observation from remote sensing. DCCs are consequently often
used as spectral invariant targets to monitor the radiometric response degradation of reflective solar
bands of earth observation sensors (Fougnie and Bach 2009, Wang and Cao 2016, Lamquin, Bruniquel
et al. 2017). MSI products containing observations of DCCs covering few to hundreds of kilometres
can be extracted from a series of radiometric tests such as thresholding reflectance in water vapour absorption bands (especially B10, 1375 nm) to detect high opaque cloudy features (see Lamquin, Bruniquel et al. (2017) for details).

6.5.2 RadCalNet Gobabeb site case study

For the RadCalNet Gobabeb site we selected a product with minimum cloud image percentage. The product corresponds to an overpass on the 9th of June 2017 and UTM tile T33KWP.

The ROI pixels were selected with centre at lat/lon -23.6°, 15.119° with a size ranging from one pixel up to 500 m. This ROI was also checked for any potential pixels masked as cloud, cirrus, no data or defective.

The results in Figure 6-4 show the evolution of the ROI mean uncertainty as a function of the ROI size as calculated using the MCM. The first point on the left side of the figure corresponds to the per-pixel uncertainty directly obtained from the S2-RUTv1.

The decrease of the uncertainty is in the range of 0.2%-0.8%, depending on spectral band and the different correlation levels of different bands. For B1 the uncertainty only decreases by 0.1%; this is because this band has already been binned, for 3 × 3 20 m pixels, and the correlation between pixels is lower because of the low noise component — this band has reported SNR well above 1000 (Gascon, Bouzinac et al. 2017). For all bands the uncertainty levels stabilised at around ROI sizes of 200 m or less.

Figure 6-5 presents the comparison of the results in Figure 6-4 vs. the approximation method “select/deselect”.

133
Figure 6-4. Evolution of the ROI uncertainty ($k = 1$) with the ROI size for the RadCalNet Gobabeb site using the MCM technique.

Figure 6-5. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty ($k = 1$) of the RadCalNet Gobabeb site.
The results confirm an agreement between the two methods at the level of 0.1% above 200 m. A sensitivity analysis of the MCM method varied some of the contributions by adding (or not) a compensation of 0.05% to account for the truncation of images of the S2-RUTv1. This sensitivity study gave similar results to those in Figure 6-5 thus suggesting that the small level of disagreement between the methods may be produced by the truncation.

### 6.5.3 Boussole site case study

For this case study, the same criteria has been followed as for in Section 6.5.2. The product selected here corresponds to an overpass on the 28th of March 2017 and UTM tile T32TMP.

Figure 6-6 presents the uncertainty calculated by the MCM propagation for different ROI sizes at the Boussole site for the studied S2 bands. Figure 6-7 presents the corresponding agreement between the MCM propagation and the “select/deselect” approach.

![Figure 6-6](image)

*Figure 6-6. Evolution of the ROI uncertainty (k =1) with the ROI size for the Boussole site using the MCM technique.*
Figure 6-7. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty ($k = 1$) of the Boussole site.

For this site, the uncertainty decreases up to a larger ROI size compared to the RadCalNet Gobabeb site case study. Figure 6-7 indicates that the stability is reached at around 400 m for all the studied bands. This is an expected result since at such a low radiance, the pixel reflectance factors contain a much higher uncorrelated component. The rise visible for B12 for a 50 m ROI is the consequence of the S2-RUTv1 truncation. The uncertainty maximum value is 25.5% (maximum is 255 coded in a single byte) (see Chapter 3).

Even for a large ROI, the uncertainties are higher compared to the RadCalNet Gobabeb site case study. This is due to a large component from systematic out-of-field stray-light, which has been assessed as 0.3% of the reference radiance and which assumes a constant albedo of the Earth outside of the field of view. In an ocean site the radiance of the field of view can vary strongly. It is expected that the stray-light contributions of those scenes closer to the field of view have a larger impact than the ones further away. Figure 6-8 presents the sensitivity of the 500 m ROI uncertainty with variations in the out-of-field stray-light from 0% up to 0.3% of the reference radiance.
Figure 6-8. Evolution of the 500m ROI uncertainty ($\kappa =1$) with the variation $u_{\text{stray,sys}}$ for the Boussole site using the MCM propagation technique.

The values in Figure 6-8 are provided in absolute reflectance factor with mean reflectance factor of the ROI pixels as follows: 0.12 (B1), 0.085 (B2), 0.048 (B3), 0.027 (B4), 0.022 (B5), 0.019 (B6), 0.016 (B7), 0.013 (B8), 0.012 (B8A), 0.0027 (B11), and 0.0015 (B12). The uncertainty levels in Figure 6-8 are small if compared to most of the measured TOA reflectance factors over land scenes. However, these levels become important for such low reflectance factor levels measured in water scenes.

Figure 6-9 shows the S2A overpass at the Boussole site obtained by the COVE tool (Kessler, Killough et al. 2013).
Figure 6-9. Overpass of Sentinel 2A over the Boussole site on the 28th of March 2017 using the COVE tool (Kessler, Killough et al. 2013).

Figure 6-9 shows the large variation of out-of-field scene for an orbit of S2A over the Mediterranean Sea. For the Boussole site the immediate out-of-field scene is composed of water bodies however it is immediately afterwards dominated by land. The selection of the site just above or below can provide a completely different combination of out-of-field scene. In addition, the cloud coverage of the scene might further vary the levels. Thus, it is beneficial for the S2 radiometric performance over water scenes that more detailed predictions of the out-of-field stray-light are set. This means a more detailed systematic error assessment — and, if possible, correction — dependent on the out-of-field scene distribution.

6.5.4 Deep Convective Cloud case study

The selected DCC product corresponds to an overpass on the 21st of December 2015 and UTM tile T51LVH. The DCC occupies almost the entire tile size and, for this example, the selected ROI corresponds to an approximate centre of the tile (precisely -11.383, 122.617 in lat/lon degrees).

Figure 6-10 presents the uncertainty calculated by the MCM propagation for different ROI sizes at the DCC site for the studied S2 bands. Figure 6-11 presents the corresponding agreement between the MCM propagation and the “select/deselect” approach.
Figure 6-10. Evolution of the ROI uncertainty ($k = 1$) with the ROI size for the DCC site using the MCM technique. The simulation used a spectral correlation of $u_{diff, k}$ of 0.5 and $u_{diff, abs}$ of 1.

Figure 6-11. Evolution of the difference between the MCM and select/deselect technique as a function of the ROI size for the ROI uncertainty ($k = 1$) of the DCC site. The simulation used a spectral correlation of $u_{diff, k}$ of 0.5 and $u_{diff, abs}$ of 1.
The ROI uncertainty in this case converges quicker than in the previous cases and at a lower value. This is the consequence of a lower relative uncertainty — i.e. as a percentage of the ROI mean value — due to the high radiance of the scene. Specifically, the remaining uncertainty is dominated by the diffuser calibration uncertainty.

DCCs are commonly used for inter-band monitoring and radiometric validation by exploitation of the spectral flatness (or whiteness) of their spectra. One reference band is supposed well calibrated so that the signal above the DCC (i.e. TOA corrected from stratospheric gas absorption) must be comparable from one band to this reference band. Per band deviation from this expectation, which is further refined using radiative transfer modelling of the clouds predicting the supposedly-exact spectral shape, are then interpreted as inter-band calibration residuals. As an example here the S2 bands are calculated as ratios of the B4 which is one of the band used as reference for MSI in Gascon, Bouzinac et al. (2017) and Lamquin, Bruniquel et al. (2017).

Figure 6-12 presents the uncertainty associated to the ratio of each ROI-mean of the S2 bands with respect to the ROI-mean of B4. The results are presented for all ranges of spectral correlation values of the diffuser calibration uncertainty $u_{\text{diff abs}}$ (see Section 3.3.2) as this is the dominant contribution.

![Figure 6-12](image.png)

Figure 6-12. Evolution of the ratio of ROIs uncertainty ($k=1$) with the variation of $u_{\text{diff abs}}$ spectral correlation for the DCC site using the MCM technique. The simulation used a spectral correlation of $u_{\text{diff,k}}$ of 0.5.
The results show that the uncertainty associated with this ratio can be much smaller if the contribution due to the diffuser calibration is considered as largely correlated (as this is a ratio, the sensitivity coefficient is -1). For high spectral correlation of $u_{\text{diffabs}}$, the ratio uncertainty is around the 0.5% for the VNIR bands whereas it increases up to 1% for the SWIR bands. Although the spectral correlation value specified in Table 6-1 is set to 1, the actual correlation is expected to decrease the further apart the bands are spectrally. That is, spectrally closer bands (e.g. B3) might present a large spectral correlation while the correlation is lower for bands such as B11 or B12 which are in a different focal plane.

### 6.6 Discussion and conclusions

The work in Chapter 3 defines a methodology to estimate the per-pixel TOA radiometric uncertainty associated to EO products. The propagation and combination of these uncertainty estimates is not a straightforward problem. The work described in Section 4.3 showed the complexity when propagating the TOA radiometric uncertainty during the orthorectification process. The combination of several pixels is neither a simple operation. The pixel combination process is a typical processing step at TOA radiance/reflectance factor products and it is used for radiometric validation activities or the spatial binning of products. The combination of several pixels requires the study of the correlation between them and the sensitivity coefficients derived from the measurement equation.

This study has defined a method to produce an uncertainty estimate associated to the mean TOA reflectance factor of the ROI pixels by using the S2-RUTv1. The method named “select/deselect” can be directly used by the S2 L1C data users and has been designed to be as simple and quick as possible.

The method has been compared with a more robust MCM propagation approach. The results showed, in general, that for ROI pixels above ~200 m (400 m for low radiance sites) the methods agree within 0.1%.

The correlation in the spatial, temporal and spectral dimensions has been extensively discussed in Section 6.3. For several uncertainty contributions just a qualitative assessment provides sufficient description of the correlation (e.g. instrument noise). However, for some uncertainty contributions, this qualitative assessment has not been sufficient and more involved studies are required. For example, Figure 6-8 has shown the importance of understanding the out-of-field stray light, particularly for ocean scenes. Nonetheless, it is also important to recognise that quantitatively describing the correlation of some contributions might be challenging where required data are non-existent or for contributions that are complex.

In addition to the correlation assessment, it is important to analyse the effect of those uncertainty contributions that are not included in the S2-RUTv1. For the specific case of a ROI pixels under uniform sites, several of these contributions will have a small impact. For example, the ghosting effect produced
by the filters is minimal or avoided by cloud screening the ROI pixels and the neighbouring area. Other effects such as the polarisation might have a larger impact for specific bands and angular configuration, particularly for ocean sites.

The results presented here have shown the importance of considering the pixel correlation for the ROI mean uncertainty.

The method here described can be applied to support the majority of S2 radiometric validation activities and can also be applied to other processing activities. For example, this same study is useful to determine the required binning in water applications and/or can be adapted to generate uncertainty estimates associated to S2 L1C spatially binned products.

The application to radiometric activities using Pseudo-Invariant Calibration Sites (PICS) would require further study. The assumption in this work is that the ROI pixels encompass one single detector. This assumption is not usually valid for PICS monitoring where the ROI pixels will be likely to be split in between detectors. In addition, the larger ROI size required means that some of the assumptions presented in Section 6.3 might not hold validity. Therefore, for PICS more complex correlation structures are needed.

The development of the S2-RUT is an iterative process and as such, the intention of this work is to move forward towards new features and refinements from the S2-RUTv1 presented in Chapter 3. Here, the limitations, but also potential improvements, have been highlighted with the expectation that they will be revisited in future iterations. Furthermore, this work has the purpose of training the S2 users on how to use the S2-RUT uncertainty estimates in a final application. It is important that the community becomes familiar with these metrological concepts so that the tool can be successfully used and integrated in further processing applications.
Chapter 7.

Conclusions and Future Work

7.1 General overview

The research in this project describes the methodology to analyse and implement the Top-Of-Atmosphere (TOA) radiometric uncertainty of Earth Observation (EO) optical sensors at a pixel level in the form of an external tool in Chapter 3. Providing these parameters represents a challenging initiative by the complexity of the concepts and the broadness of the topic.

It is common practice that several missions provide an uncertainty based on a worst-case error scenario and just considering a ‘standard’ case. This tends to be misleading since depending on the measured pixel — scene type, acquisition time, orbit… —, a generic uncertainty value can either provide an overestimation or an underestimation of the specific pixel. Thus, the methodologies to estimate TOA radiometric uncertainty here presented can be directly used by the Sentinel-2 (S2) mission and reproduced by numerous EO missions in order to provide more realistic uncertainty estimates and adapted to each scenario. These estimates indicate the ‘fitness for purpose’ of the data associated to each specific measurement case.

From a metrological point of view, the research is interested in bringing up the rigorous metrological techniques implemented at National Metrology Institutes (NMIs) — with a special emphasis in the GUM guidelines — to the EO community. It is important a validation and discussion of the uncertainty model combination that links the radiometric model with the combination model.

From a radiometric point of view, the research looks at different strategies to analyse the different radiometric uncertainty sources, how this can be aligned with the Guide of the expression of Uncertainty in Measurement (GUM) guidelines and which are their limitations from an instrument point of view. The use of an external tool represents an advance since it provides an automatic way to adapt the uncertainty assessment to each specific scene, user requirements or processing level. The research does also seek for methodologies to estimate other novel sources as the spectral knowledge and the resampling effect during the orthorectification process. Chapter 4 has presented and discussed these methodologies with the expectation that they can be further developed and accounted for in future version of the Radiometric Uncertainty Tool (RUT).
From a software point of view, the research describes the main challenges that arise when implementing a RUT as part of an EO processing chain. Several solutions as the possibility to select or deselect the uncertainty contributions have been included with the expectation that they can be useful for their application.

The uncertainty associated to the calibration diffuser on-board S2 has been discussed in Chapter 3. However, in Chapter 5 this uncertainty evaluation is expanded to include the rigorous assessment of the uncertainty contributions in the spectral, spatial and temporal dimensions for a sensors-to-sensor cross-calibration. This calibration methodology is an alternative to the on-board calibration systems and used by several missions as a primary or secondary radiometric calibration method. The study of the sensor-to-sensor cross-calibration is a challenging one since the uncertainty assessment must be adapted to specific angular, atmospheric and surface conditions. The research has resulted in a suite of tools and methodologies useful for the exploitation and design of missions like Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (TRUTHS) or Climate Absolute Radiance and Refractivity Observatory (CLARREO) missions.

The TOA radiometric uncertainty provided by the RUT is expected to be the input for its propagation through consecutive steps of the EO processing chain similarly to Merchant, Embury et al. (2014). These higher-level uncertainty estimates will require the covariance information in the spatial, temporal and spectral dimensions. The research cannot explore the whole range of applications but provides a first implementation of the RUT to be used for radiometric validation activities and spatial binning in Chapter 6. For this further application, the correlation between pixels at the different dimensions — spatial, temporal and spectral domain — is explored and how this additional information can be incorporated in the software processing. A simple method has been proposed, tested and validated against a Monte Carlo approach.

In summary, the research here presented develops a methodology to estimate the TOA radiometric uncertainty at a pixel-level for different EO missions. The techniques and methods used in this work have been applied with the most possible rigour and represent one of the first detailed studies in the area. It is the objective of this work that the methods here presented are taken as a reference for other TOA radiometric uncertainty analyses.

Further work is independently explored for the two main areas developed here: the RUT (section 7.2) and cross-calibration uncertainty (section 7.3).

### 7.2 S2-RUT further work

The development of the RUT is an iterative process and as such, the intention in Chapter 4 and Chapter 6 is to move forward towards new features and refinements from the RUTv1 presented in Chapter 3.
Here, the limitations, but also potential improvements, have been highlighted with the expectation that they will be revisited in future iterations. In parallel with this work, it is important that the community becomes familiar with these metrological concepts so that the tool can be successfully used and integrated in further processing applications.

This first version is focused on describing an exhaustive uncertainty methodology and a general software design that can be followed in subsequent versions of the tool and can be readily adapted to several other EO missions.

In terms of uncertainty analysis, there is the need to introduce additional uncertainty contributions — note that this work has already started in Chapter 4 — and the refinement of some of the uncertainty contributions in RUTv1. The revisit of some of these uncertainty contributions would be also in line with potential changes in the Sentinel-2 (S2) mission performance. The spectral knowledge and orthorectification uncertainty propagation studied in Chapter 4 need further work for its operational integration. The spectral uncertainty module could include a detailed representation of each one of the spectral calibration uncertainty contributions and other interpolation methods. The spectral uncertainty could also include a certain relative knowledge on the Spectral Response Function (SRF) curves and a detailed analysis of the potential degradation scenarios. In the orthorectification module improvement, further work must follow in the identification of the potential accuracy introduced by the interpolation method and the update of the correlation for each of the uncertainty contributions after subsequent revisions of the work described in Chapter 6. In terms of operational implementation, a look-up-table (LUT) or uncertainty “surfaces” have been proposed and could be applied to specific TOA radiance scenes.

Chapter 6 describes the implementation of an initial version of the correlation between the different pixels and shows how necessary is to include this information for combination or propagation of uncertainty estimates. The covariance study has been mostly based on previous experience due to the little empirical information available. A refinement of the covariance assumptions is needed if there is further access to information. The application to radiometric activities using Pseudo-Invariant Calibration Sites (PICS) would require further study. The assumption in this work is that the ROI pixels encompass one single detector. This assumption is not usually valid for PICS monitoring where the ROI pixels will be likely to be split in between detectors. In addition, the larger ROI size required means that some of the assumptions presented in Section 6.3 might not hold validity. Therefore, for PICS more complex correlation structures are needed.

The software implementation in future versions would seek to include the new uncertainty contributions and update the results of a refined uncertainty analysis. In addition, it should include the processing of the quality masks’ information of the S2 L1C products as part of the uncertainty image itself. Use of
quality flags related to polarisation and stray-light effects as specified in Section 3.5.3 or covariance information may be included as a second byte codification. The interface with other Sentinel Application Platform (SNAP) plug-ins might enhance even further the possibilities of the tool.

### 7.3 Cross-calibration uncertainty further work

The modules related to spectral and spatial dimensions (see Sections 5.2 and 5.3 respectively) should be tested over several more sites. In subsequent refinements, they could be incorporated as part of operational mission tools.

The temporal module in Section 5.4 needs further refinement so that an improved atmospheric and surface modelling is accounted. It is important to test other radiative codes to provide further robustness to the temporal variations. The atmospheric variations and the impact in a correction should be extended to account for further parameterisation as the ozone content, temperature or pressure. Finally, the surface reflectance correction uncertainty should be upgraded by introducing the impact of the correlation between the different RPV model parameters.

It was mentioned in Chapter 5 the asynchronous orbit of the reference sensor and its implication in terms of temporal effects in cross-calibration. Time delays between the sensor under test and a cross-calibration using TRUTHS or CLARREO will tend to zero as the number of match-ups increase, due to randomness and in turn reducing overall uncertainties. This is well-founded assumption but has not been properly tested. Thus, next steps of this research should try to test the asynchronous impact over the error accumulation. For that purpose it will be necessary the study of the orbits in combination with different calibration sites modelling.
References


Xiong, X., G. Troller, J. Sun, B. Wenny, A. Angal and W. Barnes MODIS Level 1B Algorithm Theoretical Basis Document. N. M. C. S. Team, National Aeronautics and Space Administration