LUNAR IRRADIANCE MODEL ALGORITHM AND THEORETICAL BASIS DOCUMENT

Version 3.3

ABSTRACT

This document describes the algorithmic and theoretic basis of the lunar irradiance model and the Degree and Angle of Linear Polarization model.

Adriaensen Stefan (1), Pieter De Vis (2)

VITO (1), NPL (2)

02/11/2025

This document was produced as part of the ESA-funded project “Lunar spectral irradiance measurement and modelling for absolute calibration of EO optical sensors” and “Improving the lunar irradiance model of ESA” under ESA contract numbers: 4000121576/17/NL/AF/hh and 4000136003/21/I-DT-lr

Signatures and version history

	Name	Organisation	Date
Written by	Stefan Adriaensen	VITO	26-Feb-2018
Reviewed by (consortium)	Emma Woolliams	NPL	9-Dec-2019
Approved by (ESA)	Marc Bouvet	ESA	12-Dec-2019
Updated by	Stefan Adriaensen Pieter De Vis	VITO, NPL	19-Dec-2023
Reviewed by (consortium)	Agnieszka Bialek	NPL	25-Jan-2024
Approved by (ESA)	Marc Bouvet	ESA	

Version history

Version	Date	Publicly available or private to consortium?
0.1	19/03/2018	created
0.2	28/02/2019	Add description model regression and rearrange the document
0.3	24/05/2019	Include Polarization (chapter 4)
	24/05/2019	Include Uncertainty estimation procedure (chapter 5) : source notes Emma Woolliams
0.4	08/07/2019	All sections, added figures uncertainty
0.5	16/08/2019	Added uncertainty section, Added section on model output
0.6	09/09/2019	Updated sections except part 3
0.7	4/11/2019	updated section on uncertainty, showing new results with the updated measurements and derived models
0.8	18/11/2019	updated all sections
0.9	07/12/2019	Reviewed by Emma Woolliams
1	12/12/2019	Approved by Marc Bouvet
1.1	18/12/2019	update with new measurements : • New model coefficients • Measurements geometry overview • Update uncertainty analysis
1.2	03/12/2020	update with new measurements : • New model coefficients • Measurements geometry overview

		• Update uncertainty analysis
1.3	31/02/2022	update with new measurements : • New model coefficients • Measurements geometry overview • Update uncertainty analysis • Update DOLP model (2.7) • Fig 3,4,5,7,8,17 updated
2.0	29/06/2022	• Added section on residual based spline regression in section 2.6.1
2.1	28/04/2023	Updates interpolation and regression and uncertainties propagation
2.1	07/06/2023	Re-arrange
2.1	26/06/2023	
2.1	01/08/2023	Update interpolation algorithms and uncertainty propagation
2.8	11/11/2023	Update on DoLP algorithm
2.8	24/11/2023	Update sections layout
2.8	19/12/2023	Accept changes, clean document
2.9	01/12/2024	Update DoLP algorithm
3.0	28/04/2025	Model update
3.0	28/04/2025	AoLP/DoLP description update
3.1	19/05/2025	Update : • uncertainty plots • Dolp / Aolp description
3.2	08/06/2025	Update Dolp and Aolp coeff tables with wavelengths Add [RD14]
3.3	02/11/2025	Update model coefficients and plots with 10/2025 measurments

Contents

1 Introduction

1.1 Purpose and Scope

This document describes the algorithmic and theoretic basis of the lunar irradiance model. It includes the method to adjust the model parameters based on irradiance measurements. It also describes how to propagate the uncertainties throughout the entire model chain.

1.2 Applicable and reference documents

1.2.1 Applicable Documents

The following applicable documents are those specification, standards, criteria, etc. used to define the requirements of this representative task order.

Number	Reference
[AD1]	ESA-TECEEP-SOW-002720. Lunar spectral irradiance measurement and modelling for absolute calibration of EO optical sensors.
[AD2]	Deliverable-1: LIME 1 project
[AD3]	Deliverable-2: LIME 1 project
[AD4]	Deliverable-3: LIME 1 project
[AD5]	Deliverable-4: LIME 1 project
[AD6]	Deliverable-6: LIME 1 project
[AD7]	Deliverable-1: LIME 2 project

1.2.2 Reference Documents

Reference documents are those documents included for information purposes; they provide insight into the operation, characteristics, and interfaces, as well as relevant background information.

Number	Reference
[RD1]	H.H. Kieffer and T.C. Stone. The Spectral Irradiance of the Moon. 2005. The American Astronomical Society. DOI:10.1086/430185.
[RD2]	Apollo 16 Samples from http://www.planetary.brown.edu/
[RD3]	Numerical Recipes in C, William H. Press . . . [et al.]. – 2nd ed.
[RD4]	SPICE : https://naif.jpl.nasa.gov/naif/documentation.html
[RD5]	Optical measurements of the Moon as a tool to study its surface, Y. Shkuratov et al., 2011
[RD6]	MODIS and SeaWIFS On-Orbit Lunar Calibration, Sun J. et All, 2008
[RD7]	On-orbit Radiometric Calibration Over Time and Between Spacecraft Using the Moon, Kieffer H. et al., 2003
[RD8]	Lunar Calibration Of MSG/SEVIRI Solar Channels, Viticchie et al., 2014
[RD9]	Correction of a lunar-irradiance model for aerosol optical depth retrieval and comparison with a star photometer, Roman et al., 2020
[RD10]	Coddington, O. M., Richard, E. C., Harber, D., Pilewskie, P., Woods, T. N., Chance, K., Liu, X., and Sun, K.: The TSIS-1 Hybrid Solar Reference Spectrum, Geophys. Res. Lett., 48, e2020GL091709, https://doi.org/10.1029/2020GL091709, 2021.
[RD11]	Astronomical Algorithms – 2nd edition, Jean Meeus, ISBN 0-943396-1
[RD12]	Multiband Spectropolarimetry of Lunar Maria, Pyroclastics, Fresh Craters, and Swirl Materials, Wöhler C. et Al, The Astronomical Journal, 167:187 (12pp), 2024
[RD13]	Optical measurements of the Moon as a tool to study its surface, Shkuratov Y. Et al, Planetary and Space Science,2011.
[RD14]	Wavelength Dependence of Polarization. XX. The Integrated Disk of the Moon, Coyne et al, 1969

1.3 Glossary

1.3.1 Abbreviations

Abbreviation	Stands For	Notes
ESA	European Space Agency	Project customer
NPL	National Physical Laboratory	Project partner
J2000	Celestial reference frame for coordinates	
JPL	Jet Propulsion Laboratory	
NAIF	Navigation and Ancillary Information Facility	
ROLO	Robotic Lunar Observatory	
SPICE	Spacecraft Planet Instrument C-matrix Events	
SWIR	Short-Wave InfraRed	
USGS	U. S. Geological Survey	
UVa	University of Valladolid	Project partner
VITO	Flemish Institute for Technological Research *(Vlaamse Instelling voor Technologisch Onderzoek)*	Project partner
VNIR	Visual and Near InfraRed	

2 Lunar irradiance model derivation

2.1 Rationale

The moon has already been observed for thousands of years. These observations have, in recent years, evolved into detailed and automated radiometric measurements. Different measurements from various locations are being carried out: ground-based sensors, satellite sensors and even lunar orbiting sensors.

With the Robotic Lunar Observatory (ROLO), USGS has acquired 85000+ images of the moon, during a period of almost 10 years [KIEFER and STONE,2005]. The moon-disk-integrated irradiance was measured in 32 bands of which 23 are VNIR and 9 SWIR. About 1000 images per spectral band were used to fit the lunar spectral reflectance ROLO model. This reflectance model can be used to simulate any lunar irradiance up to 90 degrees phase angle. Many inter-comparisons between the ROLO model and e.g. space-born sensors have revealed a possible discrepancy in absolute levels of the model ([RD6], [RD7], [RD8]). These studies have shown a possible underestimation of the ROLO model by 5 % to 15 % in the VNIR and SWIR with respect to the satellite-based lunar observations. Some studies also indicate that there is a model dependency on the phase angle [RD8].

In this project, a new lunar irradiance model has been developed, based upon lunar measurements acquired with the CE318-TP9 instrument (also referred to as the 1088 instrument). The development of the new model is based on the analytical formulation of the ROLO model with new estimates of the calibration parameters. However, where needed, the formulation was adapted.
At the UVa institute in Izaña (Tenerife), a second CIMEL instrument has been used to measure the lunar irradiance (also referred to as 933). This instrument was deployed during the period 2016 and 2017. New measurements with the 933 instrument have been conducted in the period 2018 and 2019, ensuring overlap between the two instruments. These extra measurements have been very useful in the early definition of the new model.

The measurements of this instrument were used in this study to investigate the feasibility of the derivation of a new lunar irradiance model. The current LIME model however is based on the 1088 instrument measurements only. At the time of writing, measurements from the 1088 instrument are available for 7 years, from 03/2018 until 12/2024. About 570 irradiance measurements have been recorded. After filtering, only allowing low uncertainty measurements, about 310 measurements have been used for model derivation.

In addition, with the CIMEL dataset, an ASD campaign was organized to capture 4 phases of the moon. The data is used to interpolate the reflectance in between the wavelengths of the CIMEL. The measurements are described in [AD7].

2.2 Lunar model definition

The model is based on the lunar irradiance measurements from the CE318-TP9 “1088” instrument (see D2, D3, D4 and D6 from the previous study)

The model is detailed in equation 1. It is a slightly modified version of the USGS ROLO lunar model [RD1].

The only difference is that for each spectral band in the model an independent set of *c*-coefficients has been defined, while in the original model, the *c*-coefficients are identical for all bands.

$$\ln(A_k) = \sum_{i=0}^{3} a_{ik} g^i + \sum_{i=1}^{3} b_{ik} \Phi^{2i-1} + c_{1k}\theta + c_{2k}\phi + c_{3k}\Phi\theta + c_{4k}\Phi\phi + d_{1k} e^{-\frac{gC}{p_1}} + d_{2k} e^{-\frac{gC}{p_2}} + d_{3k}\, cos\left(\frac{gC - p_3}{p_4}\right)$$

k is model spectral band,

A is the lunar reflectance, ln(*A*) the natural logarithm of *A*,

g is the absolute phase angle [radians],

θ selenographic latitude observer [degrees],

ϕ selenographic longitude of the observer [degrees],

Φ selenographic longitude of the Sun [radians],

$C = \frac{180}{\pi}$ conversion radians to degrees.

The reflectance model can be split-up in four different sections.

The **basic photometric function** is represented by the first polynomial depending solely on the phase angle. It is a wavelength-dependent third-degree polynomial, described with the a_k^i coefficients.

The variations of the reflectance of the moon due to changes in the actual area **of the Moon illuminated by the sun** and driven by changes in the distribution of maria and highlands, is expressed in the second polynomial. This polynomial is depending only on the solar selenographic longitude Φ. Fourth order coefficients b_{ik} are defined for every wavelength.

The third section, with four wavelength dependent coefficients c_{ik}, represents the **visible part of the Moon** and how it is illuminated (topographic libration).

The last part of the equation is a set of parameterized exponential and cosine functions modulated by a set of d_{ik} coefficient: it is an **empirical iterative least square fitting** of **non-linear residuals** in the irradiance, with respect to the phase angle.

The output reflectance of the model, with varying phase angle is shown in Figure 1.

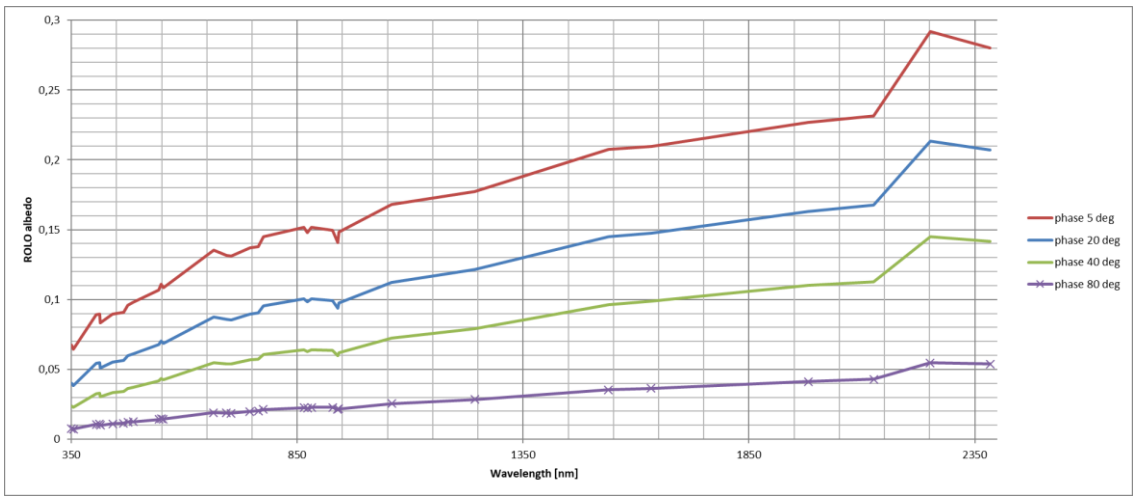

Figure 1: ROLO model reflectance spectrum output for different phase angles.

2.3 Geometric Calculations

2.3.1 Definition

Although the geometric calculations for Moon, Sun and observer are not explicitly part of the model, it is useful to mention that they are done using the JPL NAIF spice library. The software is available from: https://naif.jpl.nasa.gov/naif/

From the observation timestamp and exact topographic position it is possible to calculate in the J2000 celestial frame the position of observer, sun and moon.

With these positions, all inputs for the lunar reflectance model are defined: g, θ, ϕ, Φ.

The JPL NAIF spice library provides a set of kernels to define the position and motion vectors of the different celestial bodies involved in the geometric calculation:

List of kernels used to configure the spice library:

Table 1: SPICE kernel list used in geometric calculations.

Kernel name	Kernel type
naif0010.tls	Leap-seconds (for UTC)
pck00010.tpc	Planetary constants
earth_000101_130520_130227.bpc	Planetary constants - earth
earth_070425_370426_predict.bpc	Planetary constants - earth
moon_pa_de421_1900-2050.bpc	Planetary constants - moon
de421.bsp	Ephemeris - earth
earth_assoc_itrf93.tf	Reference frame ITRF93
moon_080317.tf	Reference frame Moon
moon_assoc_me.tf	

2.3.2 Geometric coverage

There are many periodic cycles that apply to the Moon, Earth and Sun geometry. The cycle with the longest period is called the Saros cycle and its duration is 223 synodic months, which is 18 years, 11 days, and 8 hours. After this cycle, Earth, Moon and Sun return to the same relative geometry.

The shortest cycle is the variation in phase angle which takes about 28 days between two full moons. The cycle for the distance between Sun and Earth/Moon takes about one year.

The complete Saros cycle covers all possible relative positions between Moon, Earth and Sun. Ideally, measurements need to be done for the entire cycle to get a complete coverage of the libration between all three bodies. This is not feasible within the scope of a project like this. Fortunately, about 6 years of daily measurements from an Earth fixed position are sufficient to homogeneously sample the space of possible selenographic latitude/longitude and phase angles occurring during the lunar cycle. In Figure 2 the corresponding selenographic latitude and longitude and phase angles are displayed for a period of six years (simulated with SPICE [RD 2]). Phase angles above 90 degrees absolute are discarded. The regression of the model interpolates between the measured librations.

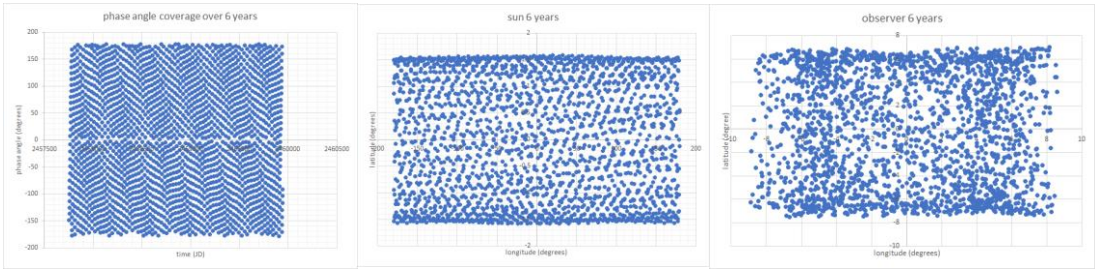

Figure 2: Simulated nightly phase angle, solar and observer selenographic longitude and latitude coverage for 6 years continuous observations.

Plots of geometry of the actual 1088 measurements show a limited coverage, compared to the previous plot. However, the plot hereafter obtained after 3.5 years of measurements, illustrates that the good sampling of the phase angle and libration angles space that can be obtained with just few years of measurements.

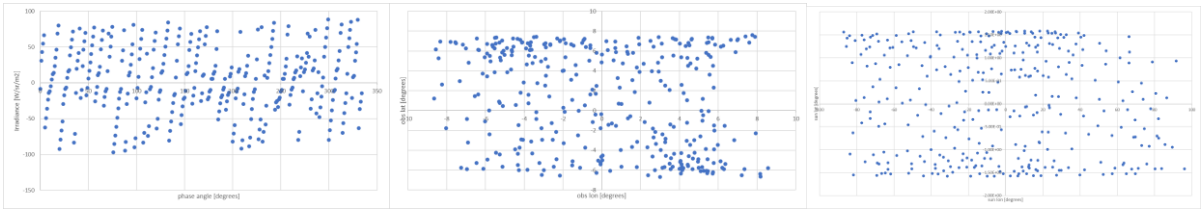

Figure 3: Phase angle, solar and observer selenographic longitude and latitude coverage for the 1088 instrument after 3.5 years of measurements.

2.4 Lunar Measurements

Figure 4 is a plot of lunar irradiance measurements with the 440 nm spectral band relative to the phase angle for the period starting in March 2018 until October 2024.

The irradiance is normalized for distances between sun, moon and observer at the time of observation. The irradiance is converted to reflectance values using the following formula. All further model derivation is performed on the disc equivalent reflectance.

$$A_\lambda = \frac{I_\lambda \pi}{\Omega_M E_\lambda}$$

A_λ lunar reflectance for a wavelength λ,

I_λ measured irradiance,

E_λ extra-atmospheric solar irradiance,

Ω_M solid angle of the Moon (6.4177×10^{-5} sr).

The solar irradiance spectrum used is TSIS-1 from Coddington et al. Replacement with other irradiance standards is possible, but the same model needs to be applied when converting back from reflectance to irradiance (i.e. when comparing with other Lunar irradiances), to maintain the CIMEL absolute level of irradiance.

Figure 4 is the plot of all irradiance measurements used in the derivation of the model (instrument 440 nm band). 3-sigma filtering is applied to the original measurements.

Close to full moon there is an apparent increase in scattering in the lower phase angles. However, in Figure 5, after removal of outliers, the relative residuals between the measurements and the model appear to be independent of the phase angle.

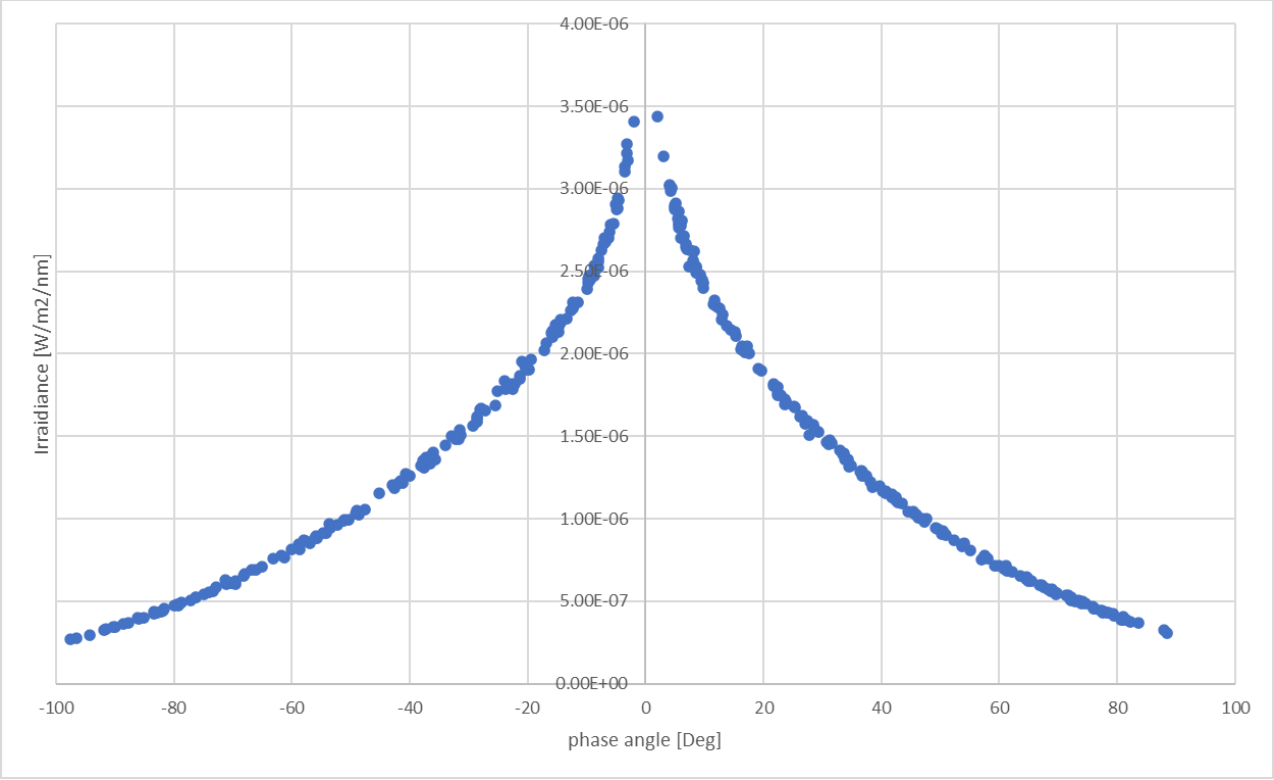

Figure 4: Lunar irradiance measurements for 440 nm.

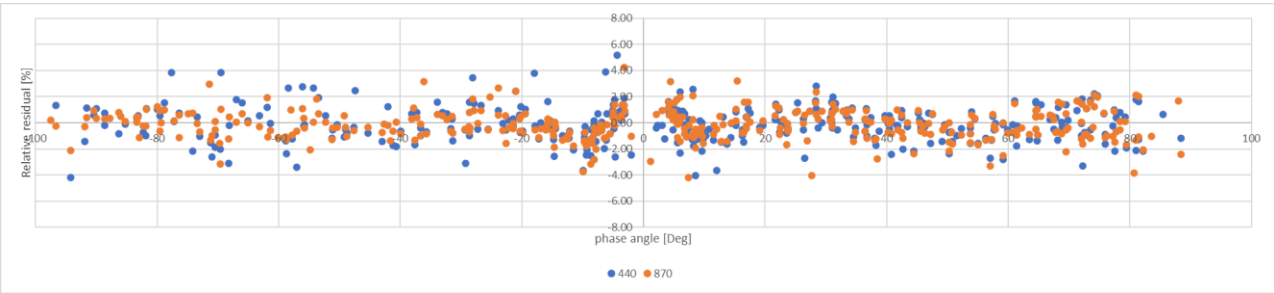

Figure 5: % Relative residuals between measurements and model for 440 nm and 870 nm.

The measurements are quality checked at the input of the model regression with an uncertainty on the stability of the Langley regression. This information is also applied to the Monte Carlo derivation of the LIME uncertainties.

2.5 Lunar irradiance model coefficient regression

The equation 1 models the natural logarithm variations of the disc equivalent reflectance. The parameters a_{ik}, $b_{ik,}$ c_{ik} and d_{ik}can be derived from direct irradiance measurements with the 1088 instrument, for all spectral bands.

The model equation can be split in two main components, representing a linear part and a non-linear part. The regression strategy is also separate for each part. There is also a distinction between band specific coefficients and the ones that are fitted for all six spectral bands.

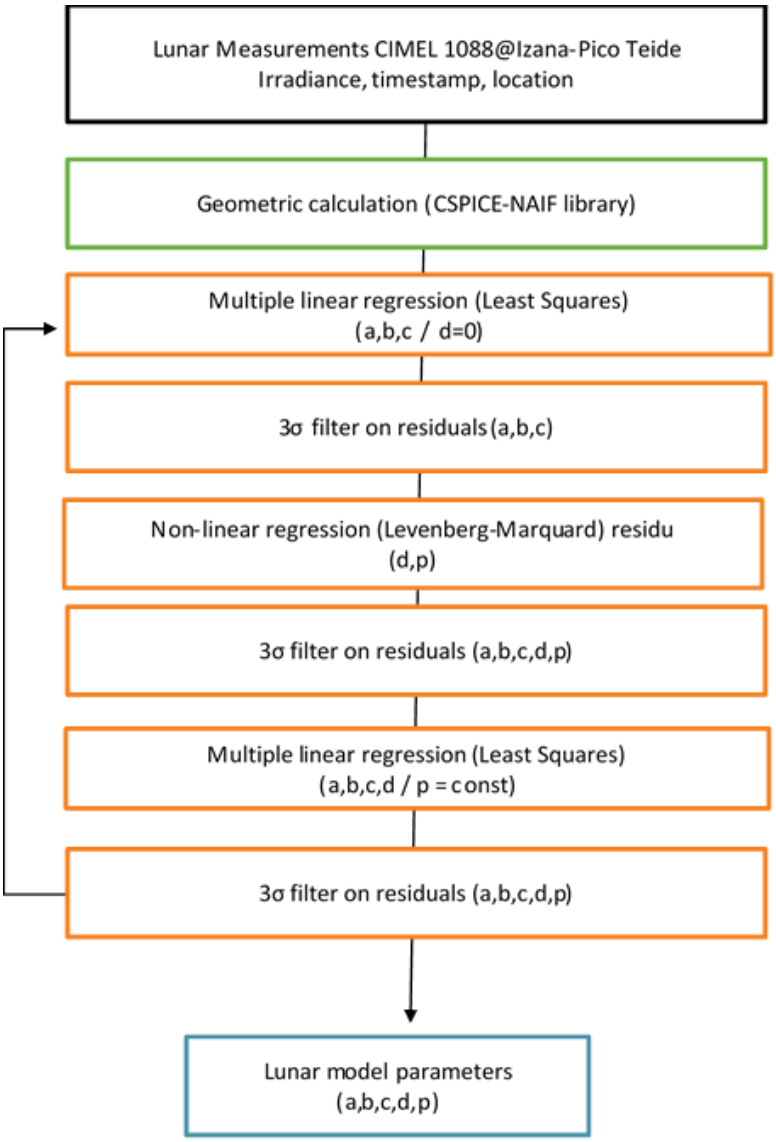

Figure 6: Lunar model coefficients regression algorithm.

First, a least-squares fit on the linear part of the model is calculated, by putting all *d*-parameters in equation 1 to zero. First, the *a*, *b* and *c* band specific coefficients are derived. With this set of coefficients, (*a*, *b* and *c*), a first 3 sigma outlier removal is done.

Then, a regression is performed on the non-linear part of the equation, using the Levenberg-Marquardt method. The *d* and *p* parameters are calculated using the residuals of all bands (Figure 8). The *p* parameters are then used in further regression and outlier removals.

Finally, again a full linear fitting is performed on the entire equation, keeping the previously derived non-linear parameters constant (*p*-parameters).

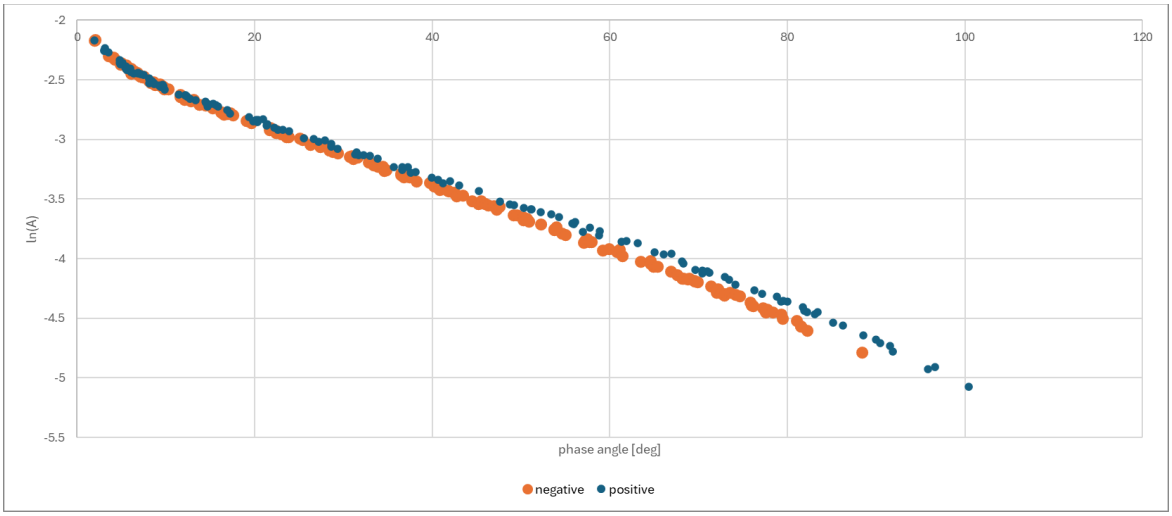

Figure 7: Natural logarithm of lunar reflectance measurement against absolute phase angle [degrees]. Negative and positive angles plotted separately.

2.5.1 Linear model fitting procedure

The first step in the regression approach is regression on the linear part of the model: in Figure 7, the natural logarithm of the measured reflectance is taken before doing the first regression.

Multivariate linear regression is performed for the first three sets of coefficients, using the matrix approach. The linear part of the model is described by the first part of the model formula:

$$\ln(A_k) = \sum_{i=0}^{3} a_{ik} g^i + \sum_{i=1}^{3} b_{ik} \Phi^{2i-1} + c_{1k}\theta + c_{2k}\phi + c_{3k}\Phi\theta + c_{4k}\Phi\phi$$

The regression is calculated per band: one set of *a*, *b* and *c* coefficients, a total of 11 parameters, is determined simultaneously for each band using the following formulation:

$$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \dots \\ h_p \end{pmatrix}$$

The *X* matrix dimension in this formula is *n* * *p*:

- *n* number of measurements,
- *p* number of coefficients to be fitted.

Definition of the parameters in the matrices:

- y_n the natural logarithm disc reflectance $ln(A_i)$,
- x_{np} predicted p, calculated for measurement n,
- h_p coefficient p.

The *y* values are the natural logarithm for every measured reflectance *A*, the *x* values are all calculated predictor values. They are calculated using phase angle, solar selenographic longitude, observer selenographic longitude and latitude. Practically, the matrix is constructed with every factor the of the lunar model calculated, as if all *a*, *b* and *c* parameters are equal to 1.0 (see also X-matrix below).

The *h*-matrix represents the coefficients to be fitted and *e* the resulting fitting error. Rewriting the regression formula gives following matrix equation:

$$\mathbf{Y} = \mathbf{Xh}$$

After converting the formula, the solution for matrix ***h*** can be found (***X***′ is the transposed ***X*** matrix. $\boldsymbol{X}^{-1}$ is the inverse of ***X***.

$$\mathbf{h} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

Steps to calculate the resulting ***h*** matrix:

First you calculate ***X***′, then you multiply ***X***′ and ***X***, which results in a squared matrix of dimension $n \times n$. The inverse of this new matrix ***A*** is calculated using LU decomposition using Crout's algorithm with partial pivoting.

Construction of the matrix ***X*** is done by filling in the predictors at their matrix position, for every measurement. For band *k* and *n* measurements, the construction of the matrix is as follows:

$$\begin{matrix}
1 & g_{1k}^1 & g_{1k}^2 & g_{1k}^3 & \Phi_{1k}^1 & \Phi_{1k}^3 & \Phi_{1k}^5 & \theta_{1k} & \phi_{1k} & \Phi\theta_{1k} & \Phi\phi_{1k} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
1 & g_{mk}^1 & g_{mk}^2 & g_{mk}^3 & \Phi_{mk}^1 & \Phi_{mk}^3 & \Phi_{mk}^5 & \theta_{mk} & \phi_{mk} & \Phi\theta_{mk} & \Phi\phi_{mk} \\
1 & g_{(m+1)k}^1 & g_{(m+1)k}^2 & g_{(m+1)k}^2 & \Phi_{(m+1)k}^1 & \Phi_{(m+1)k}^3 & \Phi_{(m+1)k}^5 & \theta_{(m+1)k} & \phi_{(m+1)k} & \Phi\theta_{(m+1)k} & \Phi\phi_{(m+1)k} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
1 & g_{nk}^1 & g_{nk}^2 & g_{nk}^3 & \Phi_{nk}^1 & \Phi_{nk}^3 & \Phi_{nk}^5 & \theta_{nk} & \theta_{nk} & \theta_{nk} & \Phi\phi_{nk}
\end{matrix}$$

The result is a $n \times 11$ matrix, for *n* measurements and 11 coefficients.

Similarly, for the construction of the ***Y*** matrix, each irradiance measurement is first converted to reflectance *A* and the natural logarithm resulting in the $n \times 1$ matrix:

$$\begin{matrix} \ln(A_1)_k \\ \cdots \\ \ln(A_m)_k \\ \ln(A_{m+1})_k \\ \cdots \\ \ln(A_n)_k \end{matrix}$$

After obtaining the coefficients, the remaining residuals represent the nonlinear part of the model. The procedure for retrieving the *d* and *p* coefficients is described in the next section. **After the nonlinear regression**, the linear regression procedure is repeated, using the selected measurements. The ***X*** matrix is expanded with factors for the *d* coefficients. All 14 linear coefficients are then fitted again for the remainder of the measurements.

$$\begin{matrix} e^{-\frac{g}{p_1}}{}_k & e^{-\frac{g}{p_2}}{}_k & cos\left(\frac{g-p_3}{p_4}\right)_k \\ \cdots & \cdots & \cdots \\ e^{-\frac{g}{p_1}}{}_{nk} & e^{-\frac{g}{p_2}}{}_{nk} & cos\left(\frac{g-p_3}{p_4}\right)_{nk} \end{matrix}$$

2.5.2 Non - Linear fitting procedure

These residuals (Figure 8) will be used to fit the non-linear part of the model. To obtain the band independent parameters (p), regression will be done based on the Levenberg-Marquardt [RD3] method.

$$Res = d_{1k}e^{-\frac{g}{p_1}} + d_{2k}e^{-\frac{g}{p_2}} + d_{3k}\cos\left(\frac{g-p_3}{p_4}\right) \quad (1)$$

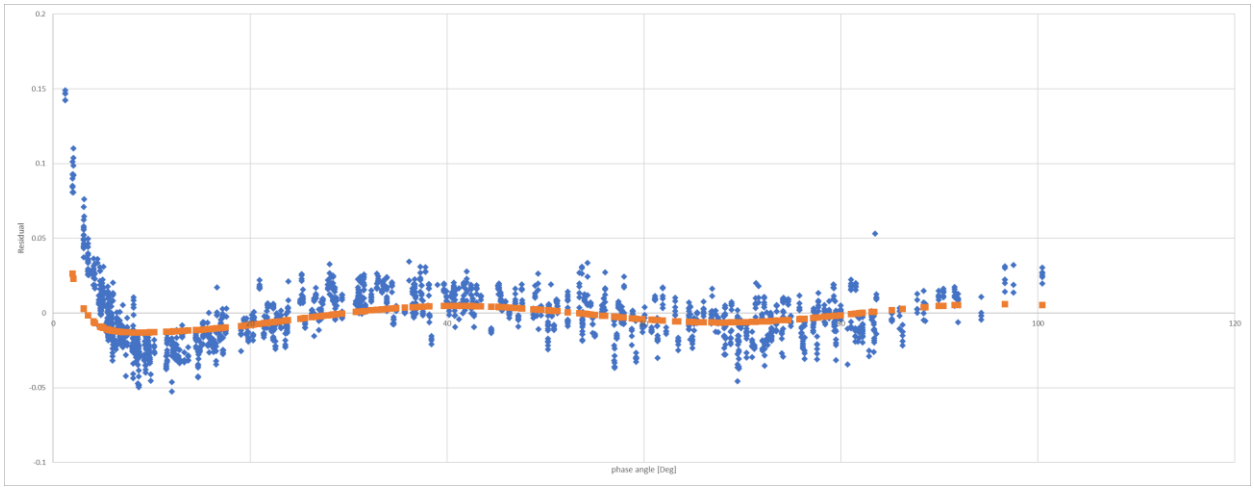

Figure 8: Measurement residual (blue) after linear fitting process for all bands, in orange model fit.

The non-linear part of the lunar reflectance model depends on the measurement phase angle. From the residuals calculated with previous steps, the next non-linear relationship is fitted. For convenience all 7 fitting parameters *d* and *p*. In the first iteration, all parameters are be fitted against residuals of

all bands. In the second iteration the *p*-parameters are adopted from the first fitting. From that point, the *d* parameters are re-fitted, but band specific and in a linear least square fitting with all *a*, *b* and *c* parameters.

Based on the calculation of the merit-function. The model must be fitted:

Model:

$$y = y(x; a)$$

Merit-function:

$$\mathcal{X}^2(\mathbf{a}) = \sum_{i=1}^{N} \left[\frac{y_i - y(x_i;\ \mathbf{a})}{\sigma_i}\right]^2$$

For every $\mathbf{a_k}$ one can calculate the derivative of the merit function: $k = 1 \ldots M$ with M the number of parameters (7 in our case).

$$\frac{\partial \mathcal{X}^2}{\partial a_k} = -2 \sum_{i=1}^{N} \left[\frac{y_i - y(x_i;\ \mathbf{a})}{{\sigma_i}^2}\right] \frac{\partial y(x_i;\ \mathbf{a})}{\partial a_k}$$

And additional partial derivatives:

$$\frac{\partial^2 \partial \mathcal{X}^2}{\partial a_k \partial a_l} = 2 \sum_{i=1}^{n} \frac{1}{{\sigma_i}^2} \left[\frac{\partial y(x_i;\ \mathbf{a})}{\partial a_k} \frac{\partial y(x_i;\ \mathbf{a})}{\partial a_l} - \left[y_i - y(x_i;\ \mathbf{a})\right] \frac{\partial^2 y(x_i;\ \mathbf{a})}{\partial a_l \partial a_k}\right]$$

This can be rewritten as a set of linear equations:

$$\sum_{l=1}^{M} \alpha_{kl} \delta a_l = \beta_k$$

With

$$\beta_k \equiv -\frac{1}{2} \frac{\partial \mathcal{X}^2}{\partial a_k}$$

and

$$\alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \partial \mathcal{X}^2}{\partial a_k \partial a_l}$$

2.5.3 Residual outlier removal approach

After removing all measurements outside the filter interval [-90;90] degrees, a filter approach as applied to the residuals of the measurements. The residuals are calculated based on the model parameters that have been derived at

3 sigma filter procedure:

res_i residuals for i^{th} measurement,

N total number of measurements (i),

$$res_{mean} = \frac{\sum res_i}{N}$$

$$s(res) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(res_i - res_{\text{mean}})^2}$$

res_{mean} is the mean residual and *s(res)* is the residuals standard deviation. A 3-sigma filter (99.7% confidence interval) is applied, to remove outlier measurements. This filter is applied multiple times during the coefficient regression procedure.

$$res_{\text{mean}} - 3s(res) < res_i < res_{\text{mean}} + 3s(res)$$

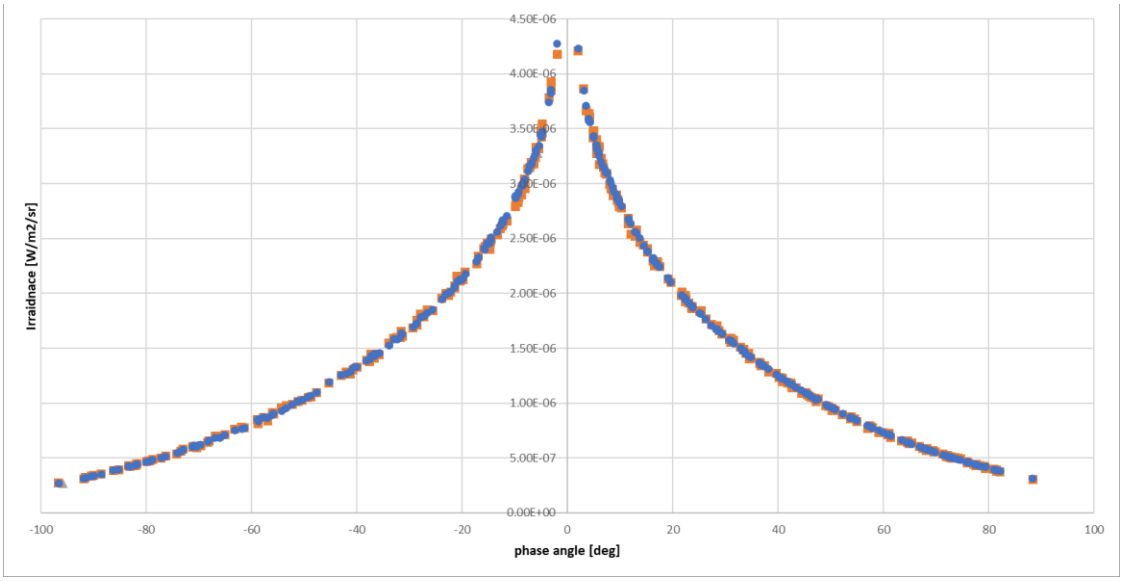

Figure 9: Filtered vs unfiltered irradiance (restricted number of measurements for illustration).

As an example, one can see the filtering result in Figure 9. Blue dots are filtered out after applying the 3-sigma filter to the residuals.

2.5.4 Iterative regression procedure

Outliers influence the results of the regressions quite significantly, so they need to be removed from the measurement population as much as possible. After all regression steps, the 3-sigma filter is applied.

Regression sequence:

- Fit the linear coefficients *a*, *b,* and *c* (all *d* coefficients equal to zero)
- Remove outlier 3-sigma from the residuals:
    - Residual = (measurement irradiance – model as is),
- Perform non-linear regression for d-coefficients and p-coefficients,
- Remove outlier 3-sigma from the residuals,
- Perform fitting over all linear coefficients: a, b, c, d,
    - non-linear p coefficients previous step,
- Remove outliers based on residuals full model.

After the first iteration, a second full iteration is performed.

2.5.5 Model coefficients 1088 instrument

Within the period 03/2018 until 11/2022,about 572 lunar irradiance measurements have been performed (depend per spectral band) and after filtering the data, based upon the quality of the Langley plots, between 440 and 485 retrievals have been used to derive the model parameters in Table 2.

Table 2: Model coefficients

wl[nm]	a0	a1	a2	a3	b1	b2	b3
440	-2.845826	-0.638992	-0.356838	-0.018674	0.044132	0.017323	-0.007338
500	-2.835119	-0.256738	-0.669409	0.070965	0.045100	0.017509	-0.008108
675	-2.488881	-0.371945	-0.538320	0.030280	0.036971	0.028278	-0.011261
870	-2.401018	-0.144913	-0.711386	0.079849	0.042253	0.027430	-0.011410
1020	-2.323384	-0.129630	-0.713925	0.081491	0.044992	0.032692	-0.013497
1640	-1.852121	-0.268270	-0.543445	0.027159	0.041477	0.021095	-0.008795
wl[nm]	**c1**	**c2**	**c3**	**c4**	**d1**	**d2**	**d3**
440	0.0005468	-0.0010941	0.0010326	0.0003428	0.5303836	0.6239650	-0.0039817
500	0.0004583	-0.0013349	0.0011645	0.0003892	0.3989398	0.7587000	-0.0025670
675	0.0004844	-0.0011166	0.0010767	0.0004950	0.4117637	0.6268670	-0.0020087
870	0.0003814	-0.0012480	0.0012339	0.0005315	0.4127379	0.6766222	-0.0008233
1020	0.0004653	-0.0012740	0.0012905	0.0006455	0.4148628	0.6541983	-0.0013789
1640	0.0002790	-0.0010727	0.0012639	0.0005961	0.4021542	0.4805239	-0.0018533
	p1	**p2**	**p3**	**p4**			
all	1.306236	18.771380	12.315492	8.973327			

With these coefficients, the lunar reflectance is calculated for every model wavelength.

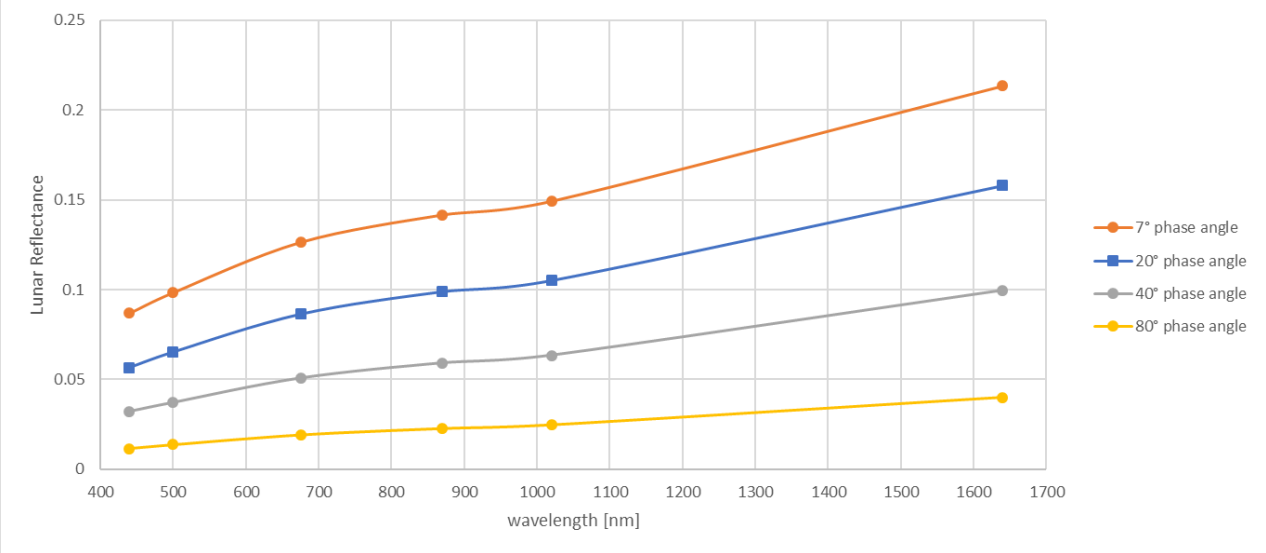

Figure 10: Lunar reflectance per wavelength for different phase angles. Interpolated values only for visual guidance, not the actual spectral interpolation.

2.6 Spectral model adjustment

With the model parameters tabulated in Table 2, the lunar reflectance can be calculated for the 1088 instrument spectral bands (Figure 10). The model output wavelengths correspond to the lunar photometer spectral bands at 6 wavelengths. The reflectance values obtained from the model are subsequently used as absolute references to radiometrically scale a hyperspectral lunar reflectance reference spectrum.

Within the LIME-TBX, there are two spectral reference spectra that can be used. The first is the Apollo spectrum, calculated from 16 Samples taken from the lunar surface during the Apollo missions (data available from http://www.planetary.brown.edu/).

The second spectral reference dataset comes from ASD measurements obtained over a 3 lunar (monthly) cycle campaign. From this campaign, a set of phase dependant ASD reference spectra were derived.

2.6.1 ASD data used as spectral model

ASD data was obtained during a three-month campaign at Izana Observatory in Tenerife (at same location as CIMEL measurements used for fitting LIME model). The Langley methodology was used to extrapolate these measurements to TOA.

2.6.1.1 Masking

Within atmospheric absorption features, this methodology returns too noisy results, and we discarded these data. The masked regions and the reason for discarding them are provided in Table 3

Table 3: masked regions in the ASD lunar data

Wavelengths masked	Reason
300 nm – 400 nm	ASD spectra too noisy
680 nm – 690 nm	O_2 absorption feature (Fraunhofer B)
713 nm - 740 nm	H_2O absorption feature
757 nm – 769 nm	O_2 absorption feature (Fraunhofer A)
809 nm – 840 nm	H_2O absorption feature
890 nm – 1000 nm	H_2O absorption feature
1080 nm – 1230 nm	H_2O absorption feature
1295 nm – 1540 nm	H_2O absorption feature
1700 nm – 2080 nm	H_2O absorption feature
2345 nm – 2500 nm	H_2O absorption feature

In addition, in the 2025 data, some significant outliers were detected outside these atmospheric features. The exact cause of the offsets is not clear, but they happen in particular for spectra with low signal-to-noise ratios. To remove these outliers, the following masking routine is followed for each ASD spectrum (i.e. independently between measurements at different phase angles).

- For each ASD spectrum, three ASD anchor points were defined using spectral regions that are unaffected by any of the outliers. These spectral regions were centered at 500, 1250 and 2150 nm. For each of these, the reflectances were averaged for all wavelengths within 20nm of the central wavelength.
- A reference spectrum was then calculated by scaling the Apollo spectrum to go through these anchor points, using the same interpolation methodology as in Section 2.6.2.

- The ASD spectrum was then compared to this reference spectrum, and any data that is more than 10% different from the reference is masked.
- This mask is combined with the mask from the atmospheric features.

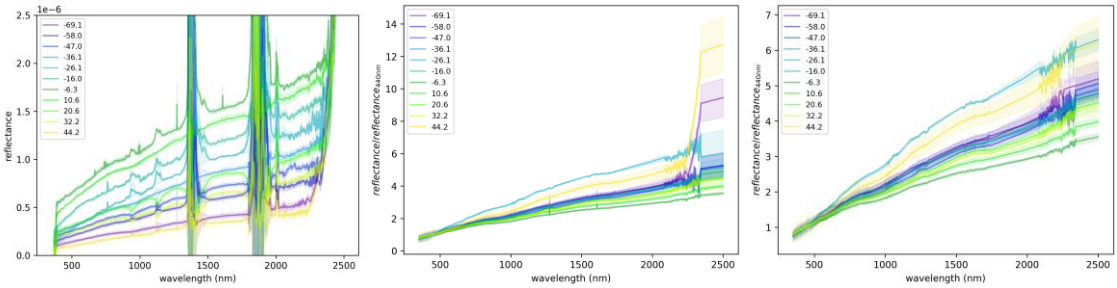

Figure 11: Example plots illustrating the masking procedure for the June 2025 data. Left: Binned ASD reflectance without any masking. Centre: Binned ASD reflectances with atmospheric masking, normalised at 440nm. Right: Binned ASD reflectances with all masking, normalised at 440nm.

2.6.1.2 Interpolating ASD data

The masked data was replaced with data from the apollo spectrum, using the same spectral interpolation method described in the next section.

In order to obtain the best ASD measurements for a given phase angle, the ASD measurements were first divided in bins of 10 degrees spanning from -95 to 95 degrees in phase angle (19 bins in total) as shown in (left). The means of all lunar reflectance observations in each bin was calculated (as well as the standard deviation, which provides the random uncertainty, see Section 3.5). These binned lunar reflectances are then linearly interpolated with phase angle to the phase angle required within the LIME-TBX (based on user inputs). Interpolated lunar reflectances for all phase angles between -90 and 90 in steps of 1 nm are shown in (right).

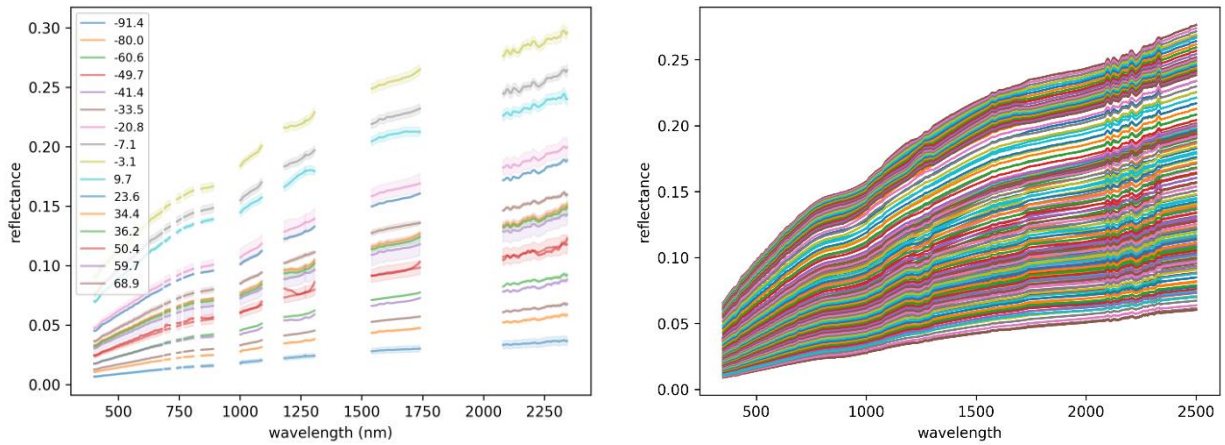

Figure 12: Left: ASD lunar reflectance data in 10-degree bins of phase angle from -95 degrees to 95 degrees (legend shows mean phase angle in each bin) shaded region shows the standard deviation in each bin. Right: interpolated ASD lunar reflectances with one line for each 1-degree step of phase angle (between -90 and 90).

2.6.2 Spectral interpolation of CIMEL residuals with uncertainty propagation

When applying the LAD method for the spectral interpolation as described in RD1, the residuals between model reflectance and apollo spectral reflectance at CIMEL wavelengths are used to perform a regression. The outcome is a linear curve applied to expand the spectral range. However, tests have shown that the linear scaling is not always a good approximation of the residuals between model and apollo spectrum.

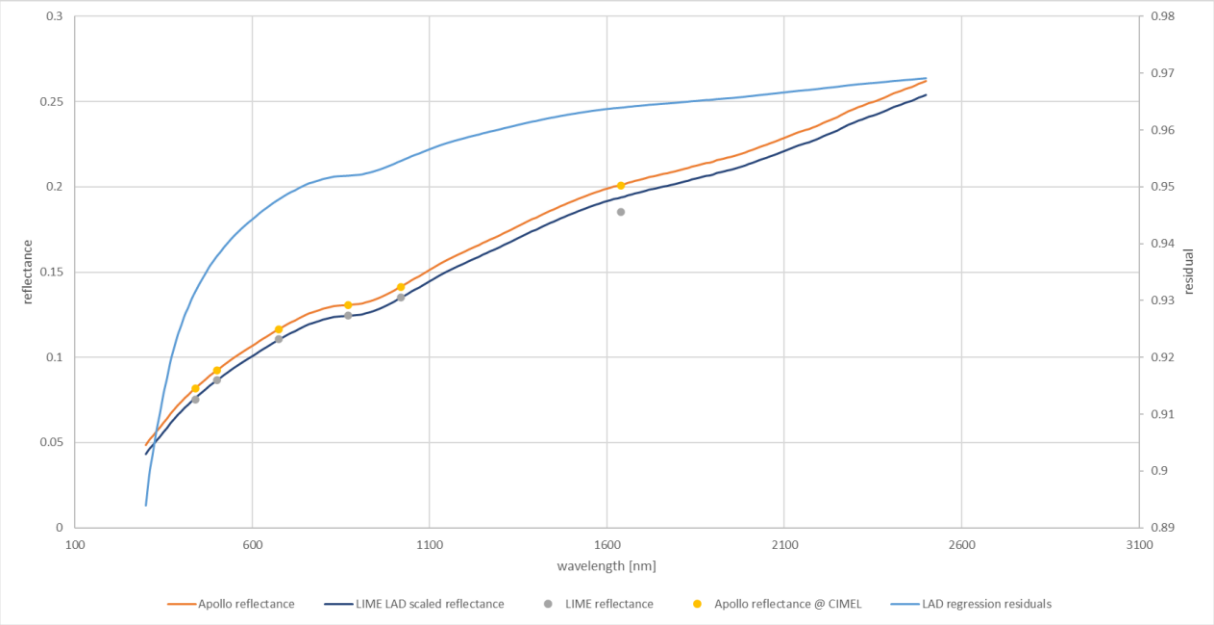

Figure 13: overview of the residuals between LIME and Apollo reflectances.

When calculating the actual residuals at the CIMEL wavelengths between both LIME and Apollo reflectance at the model wavelengths, one can observe a different shape of the curve (). The green line is shown for clarity in the plot, it is not the actual result of a regression analysis.

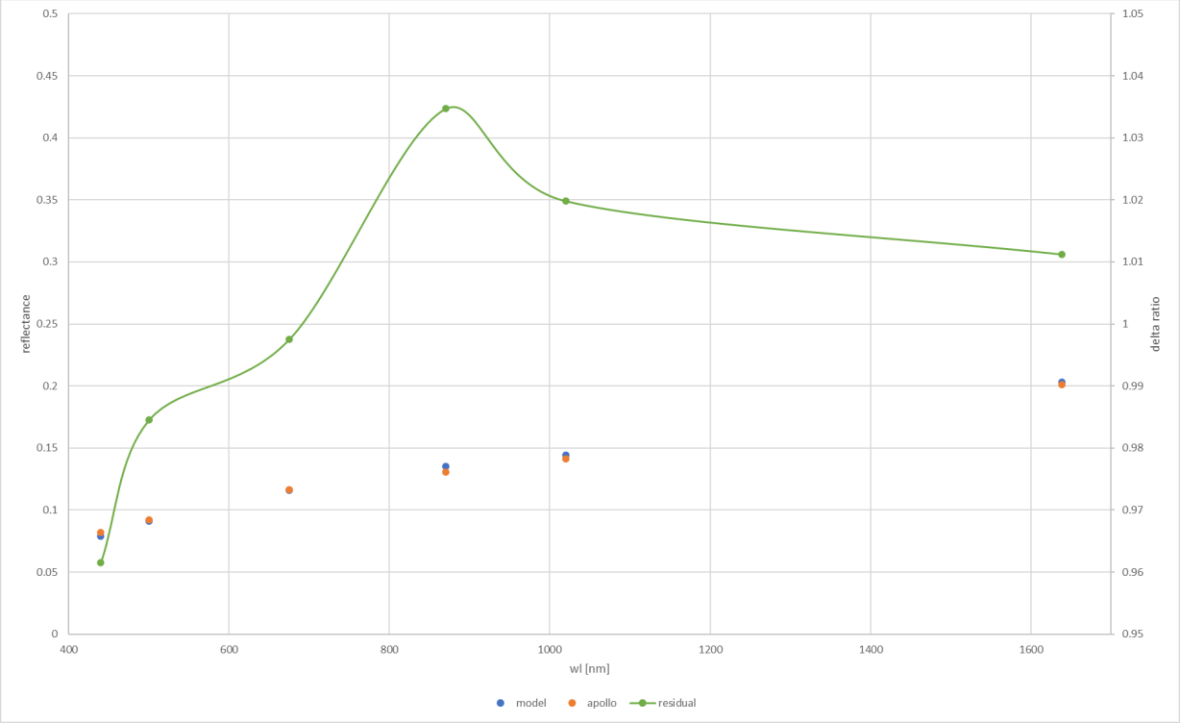

Figure 14: Residuals between Apollo and LIME reflectances at CIMEL wavelengths (green plot is filled between points for clarity only).

We want to interpolate between low spectral resolution LIME data points (based on CIMEL data with a very good absolute calibration). Between these low-resolution data points, we want to curve to follow the shape of a high-resolution example with a good relative calibration (though the absolute calibration is of less importance). Here the high-resolution data from the ASD measurements of the moon taken at al Teide in Tenerife were used.

In short, we can do an interpolation by first calculating the relative residuals by dividing the LIME reflectances by the ASD reflectances at the same wavelengths, then performing a classical interpolation to interpolate those residuals to all required wavelengths, and finally multiplying the residuals by the ASD reflectances again to get the interpolated reflectances. This methodology is implemented in the comet_maths module of the CoMet toolkit (For a full description see https://comet-maths.readthedocs.io/en/latest/content/interpolation_atbd.html). In the remainder of this section, we write this down in a mathematically rigorous manner applied to the lunar use case.

Assuming a function f that describes how the true lunar reflectance ρ_{lunar} varies with wavelength λ .

$$\rho_{lunar} = f(\lambda)$$

We want to calculate the model lunar reflectances ρ_{LIME} that approach the true ρ_{lunar} as close as possible.

$$\rho_{LIME} = \rho_{lunar} + \Delta(\lambda)$$

Where $\Delta(\lambda)$ is the error between the true and LIME reflectances, which we want to keep as small as possible.

We can define a function $\hat{f}$ that gives the interpolation function for LIME:

$$\rho_{LIME} = \hat{f}(\lambda)$$

$$\rho_{lunar} = \hat{f}(\lambda) + \Delta(\lambda)$$

This interpolation function $\hat{f}$ depends not only on wavelength, but also on the low-resolution data points from the output of the LIME reflectance model at CIMEL channels $\{(\lambda_i, \rho_i)\}$ and the high-resolution ASD data points $\{(\lambda_{\mathrm{HR,i}}, \rho_{\mathrm{HR,i}})\}$. We can thus also write $\hat{f}$ more completely as:

$$\rho_{LIME} = \hat{f}(\lambda) = \hat{f}_{\{(\lambda_i,\rho_i)\};\{(\lambda_{\mathrm{HR,i}},\rho_{\mathrm{HR,i}})\}}(\lambda)$$

To define our algorithm for the interpolation function $\hat{f}$, we start by first defining an interpolation function $\hat{h}$ that interpolates between the high-resolution ASD data points. This interpolation can be done with classical interpolation methods. In our case, we simply use a cubic spline interpolation to interpolate to any wavelength between the ASD wavelengths.

$$\rho_{HR}(\lambda) = \hat{h}(\lambda) = \hat{h}_{\{(\lambda_{\mathrm{HR,i}},\rho_{\mathrm{HR,i}})\}}(\lambda)$$

Here we again use the notation $\hat{h}$ instead of h, to indicate these are model values, which have an error compared to the true spectral dependency of h.

Once our high-resolution function $\hat{h}$ is defined, we can define our interpolation function $\hat{f}$ as:

$$\hat{f}\,(\lambda) = \hat{h}\,(\lambda) \times \hat{p}\,(\lambda)$$

We label $\hat{p}$ as the residual function. Given we already know the function $\hat{h}$ is a simple cubic interpolation algorithm, we only need to define the residual function $\hat{p}$, in order to have a full description of our interpolation function $\hat{f}$.

For the CIMEL wavelengths λ_i, we assume that the function $\hat{f}$ goes through the output of the LIME reflectance model at CIMEL channels, after applying a correction to take into account the CIMEL SRF (see Section 2.6.2.1) $\rho_i{}'$. We thus have:

$$\rho_i{}' = \hat{f}\,(\lambda_i) = \hat{h}\,(\lambda_i) \times \hat{p}\,(\lambda_i)$$

Or:

$$\hat{p}\,(\lambda_i) = \frac{\rho_i{}'}{\hat{h}(\lambda_i)}$$

For those wavelengths, the residuals thus give the ratio of the output of the LIME reflectance model at CIMEL channels and the ASD reflectance at those wavelengths. If the offset between the LIME model output and ASD data is fairly constant and/or smooth, there will be relatively little variation among the residuals at the CIMEL wavelengths. This means we can interpolate the residual function between these wavelengths without introducing too much error and define $\hat{p}$ as:

$$\hat{p}(\lambda) = \hat{p}_{\left\{\left(\lambda_i, \frac{\rho_i{}'}{\hat{h}(\lambda_i)}\right)\right\}}(\lambda)$$

Where again this interpolation can be done using classical methods such as cubic spline.

Once $\hat{p}$ has been defined, it can be multiplied with $\hat{h}$ to obtain $\hat{f}$. This way we have an interpolation function $\hat{f}$ that goes through the CIMEL datapoints and follows the shape of the high-resolution data as close as possible.

Uncertainties can be propagated through each of the steps by propagating the uncertainties on the input quantities (ρ_i and $\rho_{\mathrm{HR,i}}$) and estimating the uncertainty associated with the differences between $\hat{h}$ and h and between $\hat{p}$ and p (e.g. through comparing various interpolation methods).

As one can see in , extrapolating before and beyond the CIMEL wavelengths, the values tend be underestimated, which might introduce large errors. It is proposed to not perform extrapolation, but constantly propagate the first residual value (440nm) and the last value (1640nm) to the start and the end of the model wavelengths.

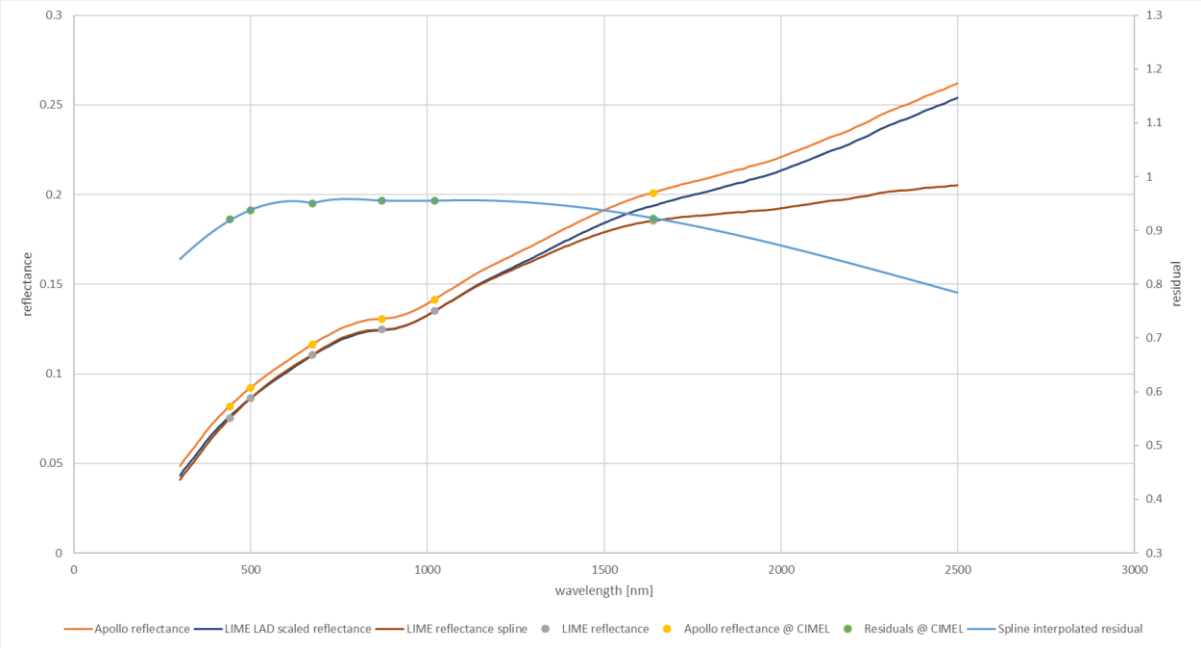

Figure 15: Comparison scaled reflectances with spline interpolated residuals.

2.6.2.1 Accounting for spectral response functions

The lunar reflectances are spectrally smooth, and thus the effects of applying different spectral response functions (SRF) are minimal. However, the CIMEL SRF is quite wide, and thus could have an effect on the observed lunar reflectance. Therefore, prior to the interpolation process described in the previous section, a correction $corr_{SRF}$ is applied to the LIME reflectances at the CIMEL wavelengths. The correction is calculated as the difference between the ASD spectrum at those wavelengths, and the CIMEL-convolved ASD reflectances at the same wavelengths.

$$corr_{SRF}(\lambda_i) = \rho_{\mathrm{HR}}(\lambda_i) - \frac{\sum_{j=\lambda_0}^{\lambda_n} \rho_{\mathrm{HR}}(\lambda_j) * RSR_{i,j} * \lambda_j}{\sum_{j=\lambda_0}^{\lambda_n} RSR_{i,j} * \lambda_j}$$

Where $\rho_{\mathrm{HR}}(\lambda_i)$ is the ASD reflectanes at the CIMEL wavelengths λ_i, $RSR_{i,j}$ is the relative spectral response of the CIMEL band i, sampled at wavelengths λ_j, and $\rho_{\mathrm{HR}}(\lambda_j)$ are the ASD reflectances at the wavelengths of the spectral response function.

The output of the LIME reflectance model at CIMEL channels, after applying a correction to take into account the CIMEL ρ_i' thus becomes:

$$\rho_i'(\lambda_i) = \rho_i(\lambda_i) - corr_{SRF}(\lambda_i)$$

where $\rho_i(\lambda_i)$ are the output of the LIME reflectance model at CIMEL channels, prior to any correction.

2.7 Lunar irradiance model output

The radiometrically rescaled hyperspectral lunar reflectance spectrum calculated in the previous section can be converted to irradiance and then convolved with at a given remote sensing instrument spectral response curves in order to simulate remote sensing instrument observations.

2.7.1 Lunar irradiance model

The lunar irradiances $I_k(\lambda)$ for wavelength λ can be calculated as (see Eq 3 in Roman et al. 2020 [RD9]):

$$I_k(\lambda) = \frac{\rho_k(\lambda) * \Omega_M * E_k(\lambda)}{\pi} \frac{1}{d_{sm}^2} \frac{d_{em}^2}{d_{om}^2}$$

Where:

$\rho_k(\lambda)$ is the lunar reflectance, for wavelength λ with hyperspectral SRF k,

Ω_M solid angle of the moon (6.4177×10^{-5} sr),

$E_k(\lambda)$ Exo-atmospheric solar irradiance, for wavelength λ with hyperspectral SRF k,

d_{sm} is the sun-moon distance, expressed in AU,

d_{em} is the earth-moon distance, expressed in km,

d_{om} is the observer-moon distance, expressed in km.

Some of these parameters are specific to a given hyperspectral SRF k. Within the LIME-TBX, there are three hyperspectral SRF options:

- A Gaussian SRF with 1nm spectral sampling and 3nm width,
- A Triangular SRF with 1 nm spectral sampling and 3nm width,
- The ASD SRF (Gaussian SRF with wavelength dependent width).

The $E_k(\lambda)$ for each of these SRF have been precalculated (with uncertainties) in the LIME-TBX based on the TSIS-1 exo-atmospheric solar irradiance, spectrally integrated with the above SRF.

The lunar reflectances $\rho(\lambda)$, calculated as the output of the spectral interpolation discussed in the previous section as spectrally smooth. Hence, spectrally integrating them over the SRF listed above makes a negligible difference and we assume $\rho_k(\lambda) = \rho(\lambda)$.

2.7.2 Simulating lunar irradiance from the lunar irradiance model

The hyperspectral irradiance model calculated in the previous section can then be convolved with a given remote sensing instrument relative spectral response (e.g: a multi-spectral instrument). By means of linear interpolation, the reflectance spectrum is resampled to the relative spectral response wavelengths and then convolved with the model simulated lunar irradiance spectrum.

For every wavelength point of the response curve, the linear interpolation is calculated. Then the integral normalized by the integral of the Relative Spectral Response (RSR) of the sensor is calculated.

$$I_l = \frac{\sum_{i=\lambda_0}^{\lambda_n} I_i * RSR_{l,i} * \lambda_i}{\sum_{i=\lambda_0}^{\lambda_n} RSR_i * \lambda_i}$$

I_l integrated lunar model irradiance (in W m^{-2}) for spectral band l,

I_i interpolated lunar spectral irradiance spectrum (in W m^{-2} nm^{-1}) at the sensor RSR wavelengths,

RSR_l band spectral response for spectral band l,

λ_i wavelengths at which the sensor RSR are defined.

This model irradiance value is the input for comparison with lunar acquisitions done with the band represented by the RSR provided to the model.

2.8 Degree and Angle of Linear Polarization

2.8.1 Introduction

The linear polarization of lunar reflected light has been studied and the phase angle dependency is described in [RD5]. The lunar polarization shows negative degree of linear polarisation at phase angles smaller than the inversion angle. This is the phase angle at which the polarisation becomes negative (around 22° absolute phase angle) [RD5].

The CE318-TP9 lunar photometer (1088 instrument) measures the polarization of incoming light. With these measurements, it is possible to derive a simplified model, based on linear regression. The model is defined per band and consists of 2 separate 5th degree polynomial functions, one for positive and one for negative phase angles.

For polarized light, the angle of polarization is angle of the polarization ellipse major axis wrt. reference plane of calculation. In this section we attempt to model the phase angle dependency of this parameter.

2.8.2 Degree and Angle of linear polarization

The degree of linear polarization (DoLP) of a signal is a measure for the fraction of the incoming light that is linearly polarized. *I, Q, U* and *V* are the Stokes parameters that describe the polarization of electromagnetic radiation. *I* is a measure for the unpolarized intensity, while Q and U describe the state of linear polarization. The circular polarization component *V* is ignored in this study.

The design of the 1088 instrument prevents the Stokes parameters from being measured directly, but the DoLP can be calculated from the different instrument filter outputs. This implies however, that it is not possible to measure directly negative degree of linear polarization and that the measurements Stokes parameters need conversion. The way to convert the output of the instrument to Stokes parameters is described in the following section.

The instrument observes the intensity through a triplet of polarizers for each wavelength. The polarizers are separated with an angle of 60 degrees between the polarizer transmittance axes. Intensity is subsequently measured at 0, 60 and 120 degrees. The intensity can be calculated directly from the instrument readout by calculating:

$$I_{ij} = \frac{1}{2} * C_{ij} * S_{ij}$$

$i = 1,2,3$ for the three axis (0°,60° and 120°) and j is the channel index. S is the instrument readout and I is the polarized radiance. C is the polarizer calibration coefficient. The readout $I(\varphi_i)$ signal intensities are the result of the application of the 3 polarizing directions to the actual preferred polarizer states (0.0,45,90) degrees.

$$\begin{pmatrix} I(\varphi_1) \\ I(\varphi_2) \\ I(\varphi_3) \end{pmatrix} = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\varphi_1) & \sin(2\varphi_1) \\ 1 & \cos(2\varphi_2) & \sin(2\varphi_2) \\ 1 & \cos(2\varphi_3) & \sin(2\varphi_3) \end{bmatrix} \begin{pmatrix} I_0 \\ U_0 \\ Q_0 \end{pmatrix}$$

As mentioned, the circularly polarization part of the light is ignored (V_0). As the 3 internal polarizers contribute to the polarization measurement, a 3x3 matrix is constructed, taking the three angles of rotation (0,60,120 degrees) into account.

$$M_c = \frac{1}{2}\begin{bmatrix} 1 & \cos(2\varphi_1) & \sin(2\varphi_1) \\ 1 & \cos(2\varphi_2) & \sin(2\varphi_2) \\ 1 & \cos(2\varphi_3) & \sin(2\varphi_3) \end{bmatrix}$$

The inverted matrix can be applied M_c^{-1} can be applied directly to calculate the Stokes parameters at the instrument local coordinate plane.

$$\begin{pmatrix} I_0 \\ U_0 \\ Q_0 \end{pmatrix} = M_c^{-1} * \begin{pmatrix} I(\varphi_1) \\ I(\varphi_2) \\ I(\varphi_3) \end{pmatrix}$$

I_0, U_0, Q_0 are stokes vectors in the instrument plane, after the application of the inverted instrument Mueller matrix M_c. The matrix implies rotation of the internal angles to the reference angles 0 and 90 degrees.

The instrument Stokes parameters are referenced to the internal coordinate system, measuring the components of the incident polarized light. The instrument remains oriented towards the incoming light in a certain way. This orientation needs to be considered when measuring light from a polarizing target (like the Moon). Therefore, a second rotation of the instrument coordinate system towards the coordinate system of the polarizing body is required. In our specific instrument setup, the instrument remains pointed towards the moon all the time, but the orientation/rotation angle α of the instrument plane changes with time.

$$\begin{pmatrix} I \\ Q \\ U \end{pmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(2\alpha) & -\sin(2\alpha) \\ 0 & \cos(2\alpha) & \sin(2\alpha) \end{bmatrix}\begin{pmatrix} I_0 \\ U_0 \\ Q_0 \end{pmatrix}$$

From these converted Stokes parameters, the Degree (P) and Angle of Linear Polarization (χ) can be computed immediately applying the next formula:

$$P = \frac{Q}{I}$$

$$\chi = \frac{1}{2} * \operatorname{atan}\left(\frac{U}{Q}\right)$$

The sign of the polarization is defined by the sign of Q, as the total intensity of the signal I in the reference plane should always be positive.

All measurements performed in this project are performed with the 1088 instrument. This means for the period of about 1 year, more than 120000 measurements of lunar polarized light are available. Not all measurements are done at full night-time and these need to be filtered from the regression.

The measurements are filtered on time – between 23h at night and 2h in the morning and outliers are removed (i.e. cloud contaminated measurement). Measurements with negative and positive phase angles are split to be able to produce a separate regression on both sides. About 70000 measurements per phase sign are used to perform the model regression. The spectral bands are treated separately.

Definition of the angle α

Finding the angle between both the instrument internal reference coordinate system to the actual coordinate, we must define the angle between the instrument 'perfect installation' position and the plane sun – moon – observer (the CIMEL instrument).

The orientation of observer is therefore important with respect to the orientation of the polarizing element (the moon). This orientation is defined by the rotation between the moon's 'upright' position and the orientation of the instrument lever arm. It is also defined as the zenith angle of the bright limb of the moon with respect to the observer.

This angle α can be calculated with the next formula. The difference between the position angle of the moon bright limb (τ) and parallactic angle of the moon at the time and position of the observer CIMEL instrument (q).

$$\alpha = \tau - q$$

With the solution for q :

$$\tan(q) = \frac{\sin(H)}{\tan(\varphi) * \cos(\delta) - \sin(\delta) * \cos(H)}$$

H : hour angle CIMEL at time of observation

φ : geographical latitude of CIMEL

δ : declination of the moon

and for τ :

$$\tan(\tau) = \frac{\cos(\delta_0) * \sin(\alpha_0 - \theta)}{\cos(\alpha_0) * \cos(\delta) - \cos(\delta_0) * \sin(\delta)\cos(\alpha_0 - \theta)}$$

δ_0 : sun declination

α_0 : sun right ascension

δ : moon declination

θ : moon right ascension

A detailed explanation of the calculations can be found in [RD11] in chapters 14 and 48.

2.8.3 DoLP model derivation

Publications have shown a negative polarization for phase angles smaller than around 22 degrees (inversion angle) [RD5]. The method and definition currently applied with the CE318-TP9 measurements show comparable results for the different instrument wavelengths.

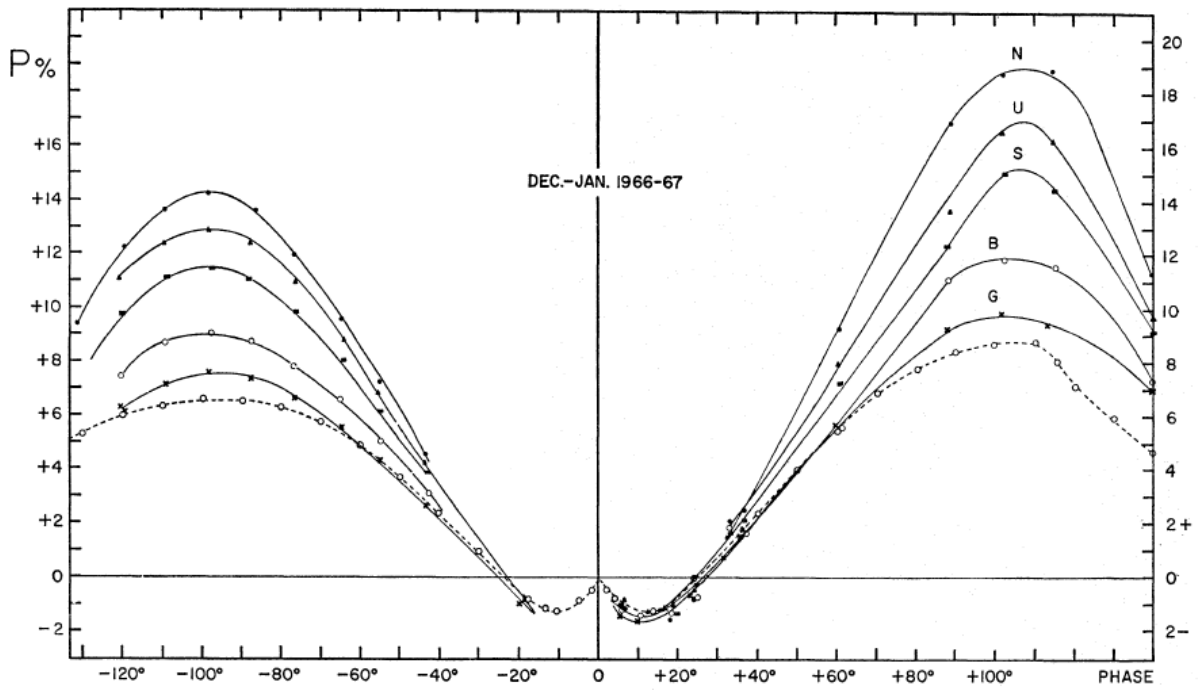

Figure 16: Model for lunar DOLP curve with phase angle at wavelength 430 nm [RD14].

As observed in , the raw DoLP measurements are, as output by the CIMEL, have quite some scatter and clear subgroups appearing in the results. The table below was taken from [RD14].

TABLE III. Maximum polarization of the moon, and phase angle at which this occurs.

λ_{eff} (μ)	Waxing moon P_{max}	Waxing moon phase	Waning moon P_{max}	Waning moon phase
Sept–Oct 1966				
0.336	12.0%	−96°	...	...
0.367	11.4	−96	...	...
Dec–Jan 1966–67				
0.336	14.2	−99	19.1%	+109°
0.367	12.8	−99	16.9	+106
0.383	11.4	−99	15.2	+105
0.449	9.0	−99	11.9	+104
0.519	7.6	−98	9.8	+103
May–June 1967				
0.336	15.4	...	20.4	+104
0.363	13.9	...	18.8	+105
0.383	12.2	...	16.8	+104
0.450	10.0	...	13.2	+103
0.534	8.1	...	10.9	+102
Lyot average curve, 1922–1925				
Visual	6.6	−95	8.8	+108

Figure 17: Coyne G. V. and Pellicori S.F., Wavelength dependence of polarization

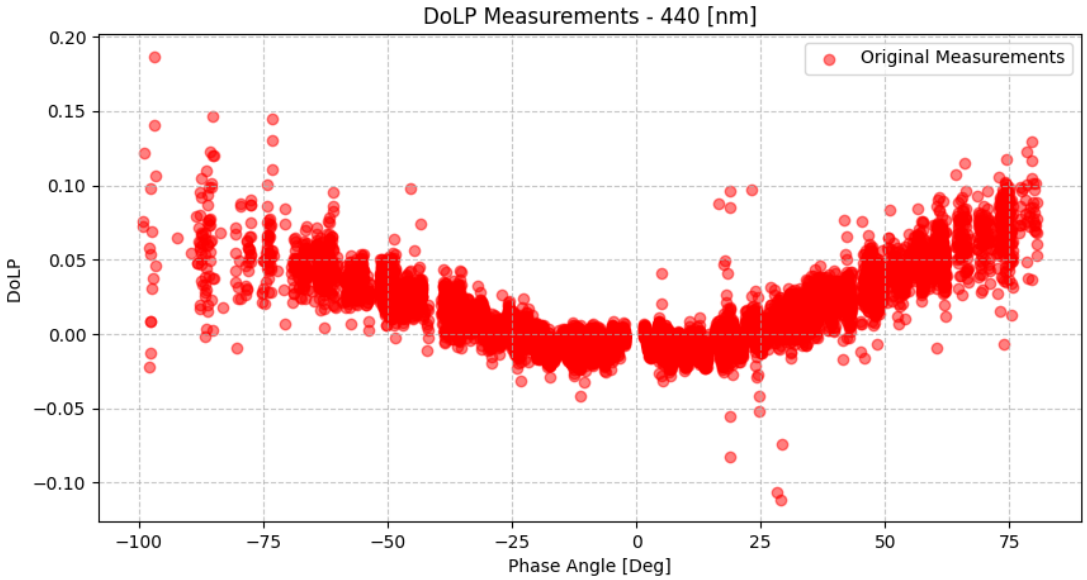

Figure 18: DoLP raw measurements.

In the next plot, all the measurements are binned per 2 degrees, for clarity of the image. It shows the DoLP levels for all wavelengths over the phase angle.

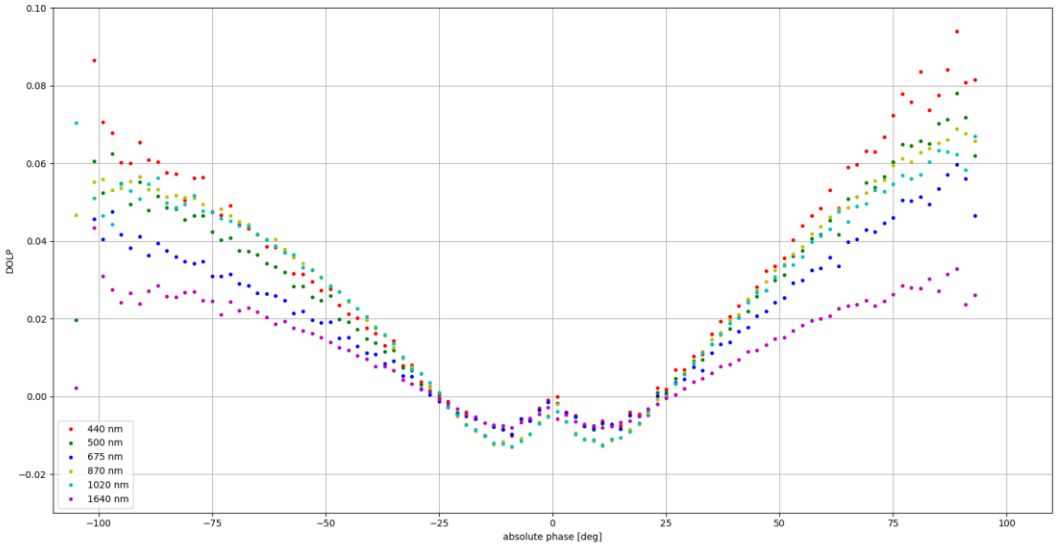

Figure 19: Binned DoLP measurement curves (2nm bins) agains the phase angle

From one can see that the DoLP is zero with zero phase angles. The DoLP stays negative until around 25 degrees absolute phase.

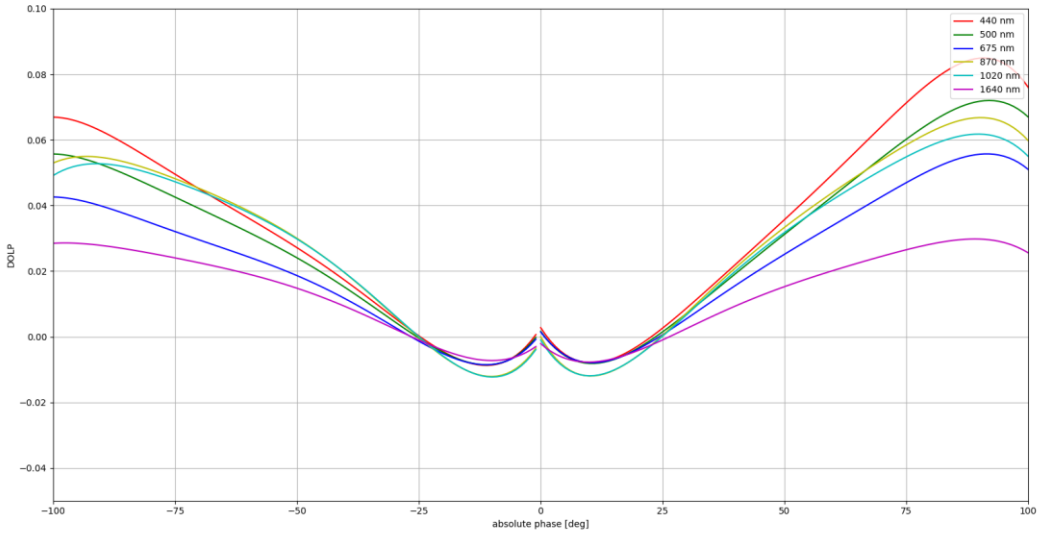

Figure 20: DoLP models for all wavelength and phase signs

The model is limited to be between 0° and 90° absolute phase angle. All polarization measurements outside these angles are removed from the regression of DoLP curves.

As can be observed in **Error! Reference source not found.**, the DoLP can be modelled using a fifth order polynomial with the intercept set to zero. From this polynomial the DoLP value is calculated directly for a certain phase angle.

$$DoLP_{wl} = \sum_{n=0}^{N} a_n * g^n$$

g is the signed phase angle;

N is the model degree

a_n are the model coefficients as tabulated in (negative phase) and (positive phase).

Table 4: Positive phase polynomial coefficients for DoLP model.

wl [nm]	a0	a1	a2	a3	a4	a5
440	9.6270E-11	2.7641E-08	3.0044E-06	1.5056E-04	2.3852E-03	2.8779E-03
500	7.7942E-11	2.3029E-08	2.5842E-06	1.3265E-04	2.1184E-03	1.7867E-03
675	7.5651E-11	2.2397E-08	2.4868E-06	1.2419E-04	1.9682E-03	1.0637E-03
870	-4.1063E-11	-1.1805E-08	-1.2911E-06	-6.3855E-05	-9.7848E-04	8.9767E-05
1020	-3.9467E-11	-1.1828E-08	-1.3381E-06	-6.6992E-05	-1.0151E-03	3.4751E-04
1640	4.5753E-11	1.3918E-08	1.6020E-06	8.1938E-05	1.2126E-03	-1.8716E-03

Table 5: Negative phase polynomial coefficients for DoLP model.

wl [nm]	a0	a1	a2	a3	a4	a5
440	-1.3734E-10	3.4041E-08	-3.2872E-06	1.5636E-04	-2.3378E-03	2.7407E-03
500	-1.0868E-10	2.8131E-08	-2.8475E-06	1.3984E-04	-2.1188E-03	1.6256E-03
675	-1.0426E-10	2.7432E-08	-2.7840E-06	1.3348E-04	-2.0176E-03	1.4354E-03
870	5.2922E-11	-1.4322E-08	1.4963E-06	-7.2403E-05	1.1001E-03	-1.9945E-04
1020	6.1859E-11	-1.6285E-08	1.6499E-06	-7.7351E-05	1.1583E-03	-4.7921E-04
1640	-6.9375E-11	1.8715E-08	-1.9372E-06	9.1733E-05	-1.3018E-03	-2.1319E-03

The polarization model will allow for a degree of polarization provided by the lunar model with respect to the input phase angle. It is provided as an optional extra output.

2.8.4 Angle of Linear Polarization model

The definition of the angle of linear polarization is drawn in . The AoLP is a measure of the magnitude of the polarization in the direction of the reference plane axis (x/y). The red arrow is the major axis of the polarization ellipse.

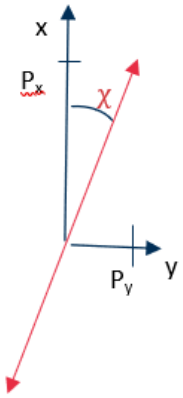

Figure 21: Representation of the angle of linear polarisation

The angle of linear polarization (AoLP) can be calculated directly with the derived Stokes parameters using the next formula:

$$\chi = \frac{1}{2} * \text{atan}\left(\frac{U}{Q}\right)$$

The angle is for all measurements at 440nm is plotted in (grey dots). As can be observed, the results for this calculation is very scattered at low phase angles. The lower signal levels and DoLP might prevent a stable measurement of the AoLP. When the polarization is zero, the value of AoLP is expected to be zero as well. We do not observe such a value in the current measurements and consequently in the model regression. Further investigation might be required. Reference document [RD11] indicates that the absolute level of AoLP for the moon is around 0° Degrees with a magnitude of a few degrees.

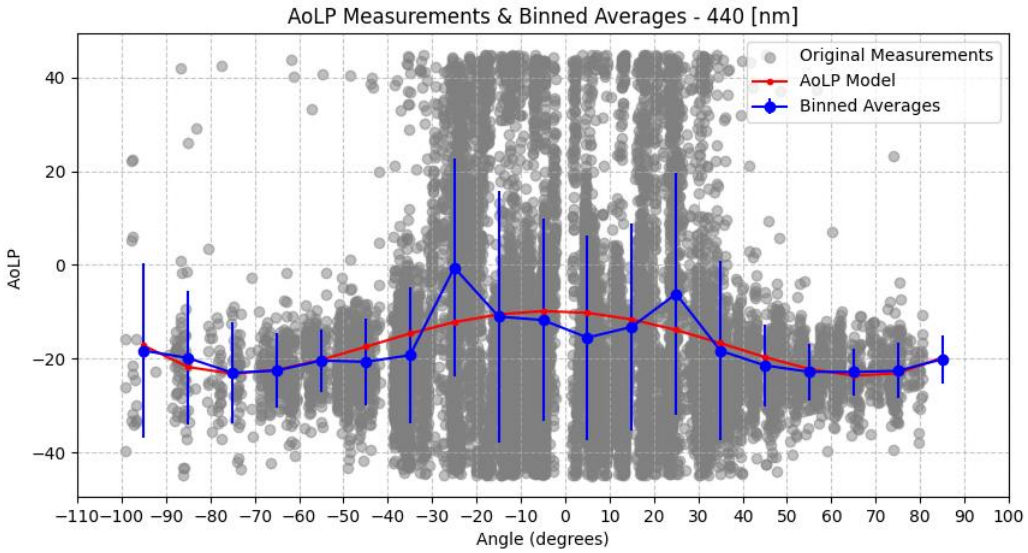

Figure 22: Measurements and model for the 440nm wavelength vs. lunar phase angle for AoLP. Blue line is the binned average with 1σ stdev as used to create the AoLP model.

The polynomial regression function for $AoLP_{wl}$ is calculated per wavelength (wl) :

$$AoLP_{wl} = \sum_{n=0}^{N} b_n * g^n$$

g is the signed phase angle;

N is the model degree

b_n are the model coefficients as tabulated in Table 6.

Table 6 : AoLP model coefficients

wl [nm]	a0	a1	a2	a3	a4	a5
440	-9.8097E+00	-7.3223E-03	-5.5030E-03	-1.2104E-05	5.5256E-07	2.1221E-09
500	-1.4225E+01	-6.3487E-02	-3.2590E-03	1.5817E-05	2.9748E-07	-6.2084E-10
675	-1.5240E+01	-4.2586E-02	-2.7201E-03	1.5767E-06	2.4438E-07	1.0091E-09
870	-1.0256E+01	-1.9410E-02	-7.0429E-05	1.1101E-06	-2.4675E-08	-6.2007E-11
1020	-8.6459E+00	-1.2387E-02	-1.1641E-03	-1.2822E-06	1.2430E-07	1.2293E-10
1640	-1.3332E+01	-6.8507E-03	-2.5133E-03	8.4554E-06	2.1761E-07	-9.7784E-10

3 Lunar irradiance model uncertainties

To enable the propagation of uncertainties up to the final model simulated hyperspectral irradiances, a Monte Carlo approach is implemented throughout the processing chain using the CoMet toolkit (www.comet-toolkit.org). The process is broken down into multiple steps. First, uncertainties on the TOA (Langley-corrected) lunar reflectance from the CIMEL measurements are determined. Next, these data are used in fitting the LIME model. Finally, uncertainties in the spectral adjustment process are detailed and combined with a solar irradiance model to calculate the lunar hyperspectral irradiance uncertainties. In this section, we will first detail some of the uncertainty concepts, and then discuss each of the aforementioned steps of the uncertainty propagation in detail.

3.1 Uncertainty framework

3.1.1 Initial concepts

The Guide to the Expression of Uncertainty in Measurement (GUM (BIPM, 2008)), and its supplements provide guidance on how to express, determine, combine and propagate uncertainty. The GUM and its supplements are maintained by the JCGM (Joint Committee for Guides in Metrology), a joint committee of all the relevant standards organisations and the International Bureau of Weights and Measures, the BIPM. The documents describe both the "Law of Propagation of Uncertainty" (hereafter, LPU) and Monte Carlo methods of uncertainty propagation.

The LPU propagates standard uncertainties for the input quantities through a locally-linear first-order Taylor series expansion of the measurement function to obtain the standard uncertainty associated with the measurand. Higher order approximations can be applied if necessary.

Monte Carlo (MC) methods approximate the input probability distributions by finite sets of random draws from those distributions and propagate the sets of input values through the measurement function to obtain a set of output values regarded as random draws from the probability distribution of the measurand. The output values are then analysed statistically, for example to obtain expectation values, standard deviations, and error covariances. The measurement function in this case need not be linear nor written algebraically. Steps such as inverse retrievals and iterative processes can be addressed in this way. The input probability distributions can be as complex as needed, and can include distributions for digitised quantities, which are very common in EO, where signals are digitised for on-board recording and transmission to ground.

Monte Carlo methods can provide information about the shape of the output probability distribution for the measurand, deal better with highly non-linear measurement functions and with more complex probability distributions, and can be the only option for models that cannot be written algebraically (including for iteration). However, they are computationally more expensive, which is an important consideration with the very high data volumes of EO.

Both LPU and MC methods can inherently take into account error-correlation matrices (or equivalently covariance matrices). Here (as in the GUM) error refers to the difference between the measurement and true value (which is not known, and can be positive or negative), whereas uncertainty is a statistical property associated with the width of the probability distribution function of the errors. Even though individual errors cannot be known, often information is available whether the errors for different measurements are correlated to each other (e.g. any error on the sensor gain, will affect all the measurements by that sensor in the same way, and thus the errors for these measurements will be correlated).

Considering these error-correlations is important when combining the uncertainties of various measurements. E.g. when integrating over the spectral response function of a satellite sensor, the integrated uncertainties will depend on whether the input uncertainties are random or systematic. Throughout the lunar processing stream, error-correlations with respect to wavelength are calculated, propagated and stored in each step. For some of the steps, additional error-correlations (e.g. between the various model parameters in Section Lunar model definition2.2) are also considered.

3.1.2 Fitting

In this project there are several fitting processes. First for each night's data, a straight line is fitted to the raw data for the Langley plot method; second the lunar model (separated into linear and nonlinear components) is fitted to the TOA lunar irradiance values obtained for each night; and third the spectral data are fitted to a spectral model. The process is further complicated by iteration both at the Langley plot stage to consider the lunar phase change during the Langley plot, and with the lunar model fitting outlier removal process.

We can consider uncertainties at several levels of complexity in fitting processes:

- At the simplest level, the fitting is done with an unweighted least squares approach (or non-linear, iterative, equivalent, such as the Levenberg-Marquardt method), taking no account of uncertainties in the fit itself.
- At the next simplest level, we may use uncertainties to weight the data going into the fitting process. This could be a weighted-least-squares fit (where the residual is divided by the uncertainty before being squared and summed), or, more simply, by introducing a cut-off (outlier removal)
- In a more robust method, the covariance of the input quantities and uncertainties in the X_i and Y_i would be considered in the fitting process. Here the full input quantity covariance would be included in the fit process. (e.g. methods such as generalised least squares – if the uncertainties are in one quantity only, or orthogonal distance regression – if the uncertainties are in both axes).
- In the most robust method, a measurement model is developed that includes error quantities for the input measurements and fits the values of those errors as well as the desired model parameters as part of the fit process. This is known as "errors in variables" fitting.

Note that each of these methods would not only provide a different uncertainty associated with the fitting but would also give a different value for the fit. When uncertainties and error covariance are taken into account in the fit process, then the fit will be different. To understand this, consider the simplest example of the difference between a simple mean and a weighted mean. If some measured values have much larger uncertainties than others, then in a simple mean the fit will be closer to these, in a weighted mean, it would be closer to those points with smaller uncertainties.

With the more robust approaches, the covariance of the fit parameters (the uncertainty associated with each parameter and the error-correlation between the errors on any pairs of parameters) would be determined "automatically" – in the sense that it can be easily calculated from the available information and would normally be calculated as part of the analysis. With the simpler approaches there is a need to calculate the uncertainty separately.

For a simple linear regression, the uncertainty can be calculated analytically. Alternatively, and more easily for complex multi-stage or iterative regressions, Monte Carlo methods can provide an uncertainty. Whether analytical or Monte Carlo approaches are used, care must be taken in

interpreting the uncertainty determined – the determined uncertainty will be based on the assumptions. It assumes that the model being fitted is consistent with the data.

3.2 Uncertainties in the TOA CIMEL data

3.2.1 Uncertainty components and their error-correlation

There are multiple components of uncertainty affecting the TOA CIMEL data. These come from two sources. First there are the uncertainties in the Langley Plot intercept. These uncertainties come from the straight line fit to the logarithm of signal as a function of airmass, and are dominated by noise in the measurement. The errors associated with these uncertainties are not correlated either with wavelength or with other times. The uncertainties will be further discussed in Section 3.2.2.

Next there are the uncertainties on the gains of the CIMEL spectral channels. All CIMEL measurements are directly proportional to these gains, and any error in the gains for a given wavelength will thus affect all CIMEL measurements in the same way (i.e. they are fully systematic with respect to time). These uncertainties are due to the calibration of the CIMEL photometer in the lab. As all CIMEL points are affected in the same way, any error present in the gains, will also directly affect the Langley Plot intercept and thus the TOA CIMEL data. Based on how the photometers were calibrated in the lab, some part of the uncertainty on the gains will be in common (systematic) between different wavelengths and some part of that will be independent (random) between different wavelengths. These uncertainty components will be discussed in more detail in Section 3.2.3.

In summary, there are three components on the TOA CIMEL measurements:

- Noise on the Langley plot intercept (random with respect to wavelength, random with respect to time)
- Random uncertainties on gain (random with respect to wavelength, systematic with respect to time)
- Systematic uncertainties on the gain (systematic with respect to wavelength, systematic with respect to time)

3.2.2 Uncertainties in the Langley Plot intercept

The Langley Plot fits a straight line to the logarithm of signal as a function of airmass. The algorithm has remained mostly unchanged from phase 1 of the LIME project, but has now followed the recommendations from Phase 1, and uncertainties on the Langley plot intercept are now calculated for each fit individually.

We summarise here the main uncertainties as quantified in Phase 1 (see also Appendix A for details on how the logarithms inherent in the Langley method are dealt with). The uncertainty associated with airmass is considered negligible. The uncertainty associated with the signal (corrected for instrument temperature effects and for lunar phase changes during the Langley[1], as well as for sun-moon and moon-Earth distances) is dominated by the noise in the measurement. This measurement noise was estimated from the statistics of the triplet (each observation being three observations made very close together in time). shows some of the uncertainties quantified as part of Phase 1.

[1] The correction for lunar phase changes during the Langley is performed by iterating the Langley plot and lunar model fitting several times.

Table 7: Table from Phase 1 of the LIME project showing the contributions to the Langley plot intercept uncertainty from various sources.

Term	Uncertainty [%]					
	1640nm	1020nm	870nm	675nm	500nm	440nm
D(λ, t)	0.07	0.05	0.02	0.01	0.03	0.04
FT(λ)	0.0027	0.13	0.18	0.17	0.15	0.053
FT(λ)	0.002	0.037	0.001	0.002	0.003	0.003
Kdist	0	0	0	0	0	0
A(tref,λ)/A(t,λ)	0.006	0.006	0.006	0.006	0.006	0.006
+0 (aerosol's diurnal cycle)	0	0	0	0	0	0
Combined standard uncertainty	0.070	0.144	0.181	0.170	0.153	0.067

The residuals observed in Phase 1 indicated these uncertainties did not account for the full uncertainties on the intercept (e.g. because there was variation in the atmosphere during the Langley or the triplet variation, which showed instrument stability over a very short period of time, underestimated the instrument variation during the full Langley period). To account for this, in cases where the residuals indicated the uncertainties were underestimated, the uncertainties on the signal were incrementally increased until the expected residuals were obtained.

Once the uncertainties on the signal are defined, the Langley plots can be fitted using a straight-line fit algorithm that takes into account uncertainty information. The algorithm used is detailed in Appendix B. Two examples are shown in and gives the range and average of the intercept uncertainties of all the individual Langley plots. These individual uncertainties were used when determining the uncertainties in the derivation of the LIME model (i.e. more uncertain Langley plot intercepts carry less weight in the fit).

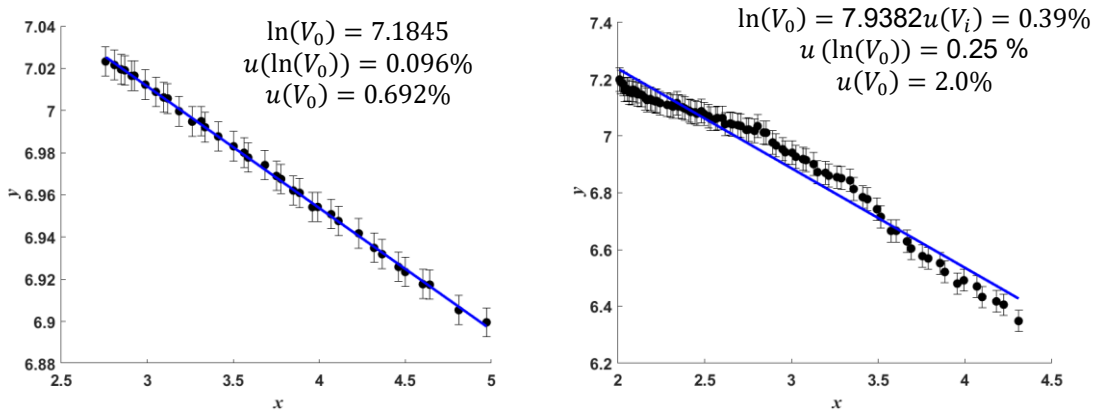

Figure 23: Langley plots and statistics for two examples. Left: example where the uncertainties did not need to be increased to pass the residual check. Right: example where the uncertainties were increased until the check was passed. The latter results in higher final uncertainties on the Langley intercept V_0.

Table 8 : uncertainty (k=1) statistics on the Langley plot intercept.

	440 nm	500 nm	675 nm	870 nm	1020 nm	1640 nm
Min	0.0379%	0.0396%	0.0381%	0.0353%	0.0336%	0.0426%
Max	0.8352%	0.6185%	0.6819%	0.6569%	0.8846%	0.8876%
Average	0.1510%	0.1207%	0.1032%	0.1009%	0.1064%	0.1029%

3.2.3 Uncertainties on the CIMEL calibration

The uncertainties on the gains originate from the calibration of the CIMEL photometer at NPL and are calculated from the uncertainty analysis performed as part of Phase 1 of the LIME project. The different contributions were reanalysed and combined into 2 separate components. One that is considered completely independent (random) with respect to wavelength and one that is considered entirely in common (systematic) with respect to wavelength. The combined uncertainties for these two components are given in .

Table 9 : Uncertainties (k=1) on the CIMEL gains from the NPL calibration.

	440 nm	500 nm	675 nm	870 nm	1020 nm	1640 nm
$u_{rand}(gain)$	0.39%	0.36%	0.42%	0.25%	0.30%	0.30%
$u_{syst}(gain)$	0.91%	0.87%	0.83%	0.90%	1.01%	1.01%

3.3 Uncertainties in the derivation of the LIME model

3.3.1 Fitting the lunar model

The lunar model fit is described in section 2. This is a multistep process where the linear part of the model is fit for each band, then outliers are removed, then the non-linear part is fit (all bands simultaneously), there is further outlier removal and finally the linear part is fit again. The whole multistep process is itself iterated.

To understand the uncertainties associated with the method, we use Monte Carlo (MC) Uncertainty Analysis using the CoMet toolkit. The MC method consists generally of three stages:

- Formulation: Defining the measurand, the input quantities X, and the measurement function (as a model relating Y and X). These definitions are provided in Section 2. One also needs to assign Probability Density Functions (PDF) of each of the input quantities, as well as define the correlation between them (through joint PDF). In the previous section, we have discussed the uncertainties in the TOA signal and how the errors are correlated. The probability density functions are assumed to be Gaussian with a standard deviation matching the discussed uncertainties. Uncertainties on the geometric angles are considered negligible. Finally, samples of the joint PDF are drawn for each of the input quantities. More details on the generation of the MC samples is provided in Section 3.3.2.
- Propagation: Propagate the PDFs for the input quantities through the model to obtain the PDF for the measurand Y. Here each MC draw is run through the model regression process described in Section 2.5, and the resulting model parameters form the output sample (PDF) of Y.

- Summarizing: Use the PDF for Y to obtain the expectation of Y, the standard uncertainty u(Y) associated with Y (from the standard deviation), and the covariance between the different values in Y. Some results are shown in Section.

3.3.2 Generating MC samples of input quantities

Generating MC samples for uncorrelated gaussian uncertainties is fairly trivial with most random number generators. However, here we have measurements for which the errors are correlated among multiple dimensions (wavelength & time). The comet_maths package in the CoMet toolkit allows to generate samples of errors which are correlated to each other. In short, this is done by first generating uncorrelated samples, and then correlating these samples using the Cholesky decomposition method. More details on this method are provided in the comet_maths ATBD[2].

To generate the full sample of input quantities for the TOA lunar reflectances, we generate samples of errors for each of the three uncertainty components listed in Section 3.2. First a fully uncorrelated sample with 1000 iterations of relative random errors $\epsilon_{rand,k,i}$ on the gains on each band k for each iteration i is generated. Next a fully correlated sample of 1000 iterations i of relative systematic errors $\epsilon_{syst,\lambda,i}$ on the gains on each band k is generated. Then, an uncorrelated sample of 1000 iterations i of relative random errors $\delta_{rand,k,j,i}$ on the individual Langley-corrected signals $\rho_{j,k}$.

The final MC sample of TOA lunar reflectances is then calculated as:

$$\rho_{i,j,k} = \rho_{j,k}(1 + \epsilon_{rand,i,k})(1 + \epsilon_{syst,\lambda,i})(1 + \delta_{rand,i,j,k})$$

for iteration i, TOA signal j (for different dates/times), and band k.

3.3.3 Propagating each iterations through the measurement function

The LIME model is then fitted to each iteration i, and each of the fitted parameters for each iteration are stored in an output MC sample. The model fit to each iteration includes the full multistep process with the linear fit, outlier removal and non-linear fit, as described in Toledano et al (in prep.). shows an example of the output sample of the fitted values of the a_0 parameter for the 440 nm band.

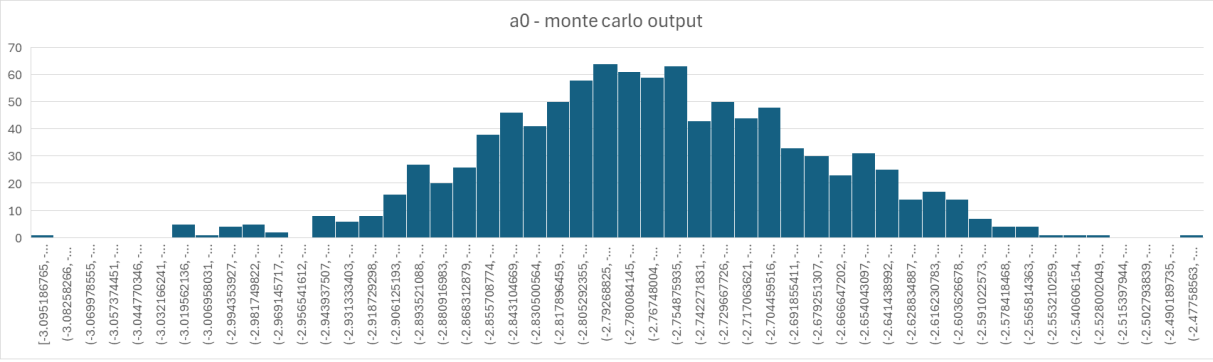

Figure 24 : Monte Carlo output for coefficient a0.

[2] https://comet-maths.readthedocs.io/en/latest/content/random_generator_atbd.html

3.3.4 Calculating the uncertainty and covariance of the fit parameters

At the end of the MC process, we have 1000 versions of the model that differ from one another in a way that is consistent with the uncertainties and covariances of the input quantities. We can use these to estimate the uncertainty and error-correlation associated with the model. This is done using the punpy module from the CoMet toolkit. Punpy calculated the uncertainties on each parameter from the standard deviation between the different iterations. The uncertainties in percent are shown in . Some parameters are much better constrained than others. Even though some of these parameters are poorly constrained, they don't necessarily have a big effect on the final uncertainty on the reflectance (see next section).

Table 10: uncertainties in percent for each of the LIME model parameters.

	A0	**a1**	**a2**	**a3**	**b1**	**b2**	**b3**	**c1**	**c2**	**c3**	**c4**	**d1**	**d2**	**d3**	**p1**	**p2**	**p3**	**p4**
440	0.9	5.1	320	18.4	1.9	26.5	17.0	125.4	4.5	5.9	9.3	18.5	7.4	36.8	0.5	0.4	0.4	0.2
500	0.7	3.3	47.5	9.5	1.8	13.8	11.1	183.0	3.5	4.5	5.3	5.6	6.0	13.7	0.5	0.4	0.4	0.2
675	0.8	2.8	20.1	6.1	1.6	10.3	10.3	26.1	3.3	3.3	6.5	10.1	7.5	11.1	0.5	0.4	0.4	0.2
870	0.8	2.8	26.5	6.4	1.5	7.8	7.5	22.4	2.9	3.0	5.1	7.3	7.2	11.3	0.5	0.4	0.4	0.2
1020	0.8	2.8	21.6	6.4	1.5	10.3	10.3	22.3	2.8	2.6	6.4	5.9	8.2	13.5	0.5	0.4	0.4	0.2
1640	1.0	2.8	30.9	5.9	1.2	13.7	12.6	10.9	3.0	2.7	5.8	3.8	10.8	13.0	0.5	0.4	0.4	0.2

The error-correlation matrix is calculated by punpy using the Pearson correlation coefficients between the each pairs of variables (i.e. the different coefficients at different wavelengths). There are 18 coefficients and 6 wavelengths in the LIME model. The error correlation matrix will thus span $18 \times 6 = 108$ columns and rows. The resulting error correlation matrix is shown in . As expected, all values on the main diagonal are 1. There are also some combinations of parameters for which the errors are clearly anti-correlated (~-1, e.g. for combination of a_0 and a_1 see the dark line below and above the main diagonal). Also interesting to note is that for the parameters p (the ones in the bottom right of the plot), all the error-correlation values for each parameter are the same, irrespective of wavelength. This is expected as the p parameters in the model do not vary with wavelength.

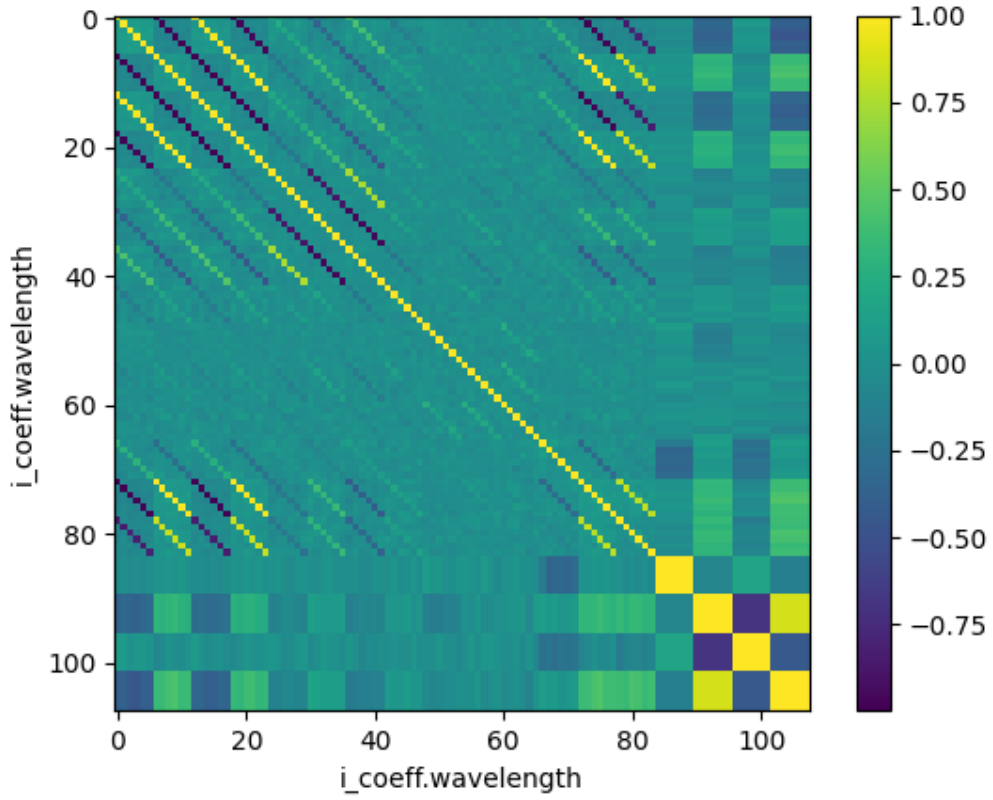

Figure 25: Error-correlation matrix for all model coefficients.

3.4 Calculating the uncertainty on reflectance

Now that the uncertainties and error-correlations on each of the LIME model coefficients are determined, these can easily be propagated to the lunar reflectance for a given observation by again following a MC approach. Within the LIME-TBX, this uncertainty propagation is done using punpy once the appropriate angles (lunar phase angle, selenographic latitude and longitude of the observer and selenographic longitude of the Sun) are determined from the user input. The uncertainties in the lime coefficients can then be propagated through the measurement model which now is simply the LIME model (as given in Section 2.2). This allows to calculate the uncertainties for each of the included bands, while taking into account the error-correlation between the different coefficients. Using the same approach it is also possible to calculate the error-correlation between the lunar reflectances for the different bands.

3.5 Calculating uncertainties on the spectral adjustment

The spectral interpolation methodology described in Section 2.6.2 is implemented within the comet_maths tool. The same tool also propagates uncertainties through this process. The following uncertainty contributions are included:

- Uncertainties on the LIME model reflectances for the CIMEL wavelengths, as described in previous section.
- Uncertainties on the spectral reference data. For the ASD case, there is a random uncertainty contribution which comes from the standard deviation between the various observations within a single phase angle bin (see Section 2.6.1). There is also a contribution from the uncertainty on the ASD calibration, which is assumed to be 3% (but does not have a big

influence on the final uncertainty as this contribution is systematic and largely cancels out through the interpolation approach).

- Model uncertainty from comparing linear, quadratic and cubic interpolation methods (standard deviation between results for these three methods is used).

3.6 Calculating the uncertainty on irradiance

Uncertainties on the hyperspectral reflectances $\rho_k(\lambda)$ and (TSIS-1) solar irradiances $E_k(\lambda)$ are propagated to uncertainties on the lunar irradiances using the measurement function defined in Section 2.7.1 and the punpy uncertainty propagation tool. No uncertainties on the lunar solid angle, or the different distances used are available, and are thus not included (though these are expected to be negligible).

In a subsequent step, these uncertainties are then propagated to the band-integrated instrument lunar irradiances using punpy.

4 Conclusions

This report has described the process involved in fitting a lunar-reflectance model to the measured Langley plot intercepts. The lunar-reflectance has been calculated using the 1088 CIMEL instrument specified, purchased, calibrated, and installed as part of this project. A degree of linear polarization model was derived from the raw polarized CIMEL instrument observations.

With the current uncertainty budget analysis, a full system uncertainty characterization has been performed. Based on the lab calibrations, uncertainties for the different measurement stages are identified:

- Systematic uncertainties (common to all),
- Uncertainties linked to instrument spectral band,
- Uncertainties linked to every measurement separately.

A method has been derived to estimate the uncertainty for every measurement, based on the Langley fitting process. This algorithm has been plugged into the measurement facilities at the Izaña institute and will provide in the future a measurement specific uncertainty. In the next iteration, these individual uncertainties will also be plugged into the Monte-Carlo analysis, allowing for a more accurate characterization of the model uncertainties.

With the current setup, using the estimated 'average' Langley uncertainties for every measurement the outcome of the Monte-Carlo analysis shows a reasonably flat uncertainty value for all model spectral bands. The 95 % confidence interval shows an uncertainty level of 2 % or less for all bands. The 99 % gave uncertainties between 2.5% to 3% depending on the spectral band. We believe that, with increasing number of measurements from the 1088 instrument and more accurate uncertainty estimates, the levels of uncertainty will slightly decrease and flatten out. This is however to be confirmed in the next iteration.

In summary, future work to improve the model will include:

- Obtaining additional measurements (6 years of data are required to cover the full range of different lunar cycles), including measurements at a new CIMEL channel around 2 µm which will constraint the spectral interpolation in the SWIR.
- Improved understanding of negative degree of polarization.
- Improved quality checks and uncertainty propagation of both CIMEL and ASD data.

5 Acknowledgements

Special thanks go to Tom Stone for having several fruitful discussions on the existing ROLO model and its possible adaptations. Without his help and legacy, this project would not have reached its current level of success.

Also, thanks to EUMETSAT for providing the GIRO model, to be used in the comparison with the new developed model. A subset of the measurements from this project will be contributed to the GLOD, the GIRO Lunar Observations Database, to be used in future inter-comparisons.

APPENDIX A – Dealing with logs in the uncertainty analysis

It is not possible to apply the Law of Propagation of Uncertainties to a logarithm directly. The logarithm and exponential functions are highly nonlinear and in this case we do not have dimensionless quantities. If we just take a logarithm of the uncertainty, we get the wrong uncertainty for the logarithm of the value. Instead we calculate the uncertainty associated with the logarithm numerically.

To understand this, we consider a simple quantity, y, which is the natural logarithm of the measured signal V, thus

$$y = \ln(V)$$

Therefore, we can write

$$V = \exp(y).$$

To evaluate an uncertainty, we perturb the measured signal, V by a small perturbation, δV.

$$V + \delta V = \exp(y + \delta y)$$

Taking a logarithm of both sides we get

$$\ln(V + \delta V) = (y + \delta y)$$

And rearranging:

$$\delta y = \ln(V + \delta V) - y.$$

We make δV equal to the uncertainty associated with the signal, and use the result δy as the uncertainty associated with the y-axis model process.

To check for symmetry, try $\delta V = +u(V)$ and $\delta V = -u(V)$. For highly non-linear functions this may not be symmetrical.

APPENDIX B – Fitting a straight line with uncertainty information

In this case we consider fitting a straight line to measured data points with some associated uncertainties which may differ from point to point. The method described here calculates both a slope and intercept for the straight line and their associated uncertainties and covariance using the uncertainties associated with the y-values. There is assumed to be no uncertainty associated with the x-values.

The calculation of the slope and offset is as follows:

The weights[3] are defined as

$$w_i = \frac{1}{u(y_i)} \tag{0.1}$$

where $u(y_i)$ is the uncertainty associated with the measured value y_i at the set value x_i.

The reference values are given by

$$x_0 = \frac{\sum_{i=1}^{N} w_i^2 x_i}{\sum_{i=1}^{N} w_i^2}, \qquad y_0 = \frac{\sum_{i=1}^{N} w_i^2 y_i}{\sum_{i=1}^{N} w_i^2} \ . \tag{0.2}$$

The slope is then calculated as

$$b = \frac{\sum_{i=1}^{N} w_i^2 (x_i - x_0)(y_i - y_0)}{\sum_{i=1}^{N} w_i^2 (x_i - x_0)^2} \tag{0.3}$$

and the intercept as

$$a = y_0 - b x_0 \ . \tag{0.4}$$

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

$$\begin{aligned} u^2(b) &= \frac{1}{\sum_{i=1}^{N} w_i^2 (x_i - x_0)^2}, \\ u^2(a) &= \frac{1}{\sum_{i=1}^{N} w_i^2} + \frac{x_0^2}{\sum_{i=1}^{N} w_i^2 (x_i - x_0)^2}, \\ u(a,b) &= \frac{-x_0}{\sum_{i=1}^{N} w_i^2 (x_i - x_0)^2}. \end{aligned} \tag{0.5}$$

[3] Because this term is squared in the subsequent equation, the actual weight is inversely proportional to the square of the uncertainty.

APPENDIX C – Producing a covariance matrix for the input observations

To do a full uncertainty analysis and a fit that fully takes into account the covariance at the lunar model fitting, we would need a covariance matrix for the input observations. We have an input observation model of

$$E_{i,\lambda} = E_{i,\lambda}^{\mathrm{True}} \times \left(1 + R_{i,\lambda}\right)(1 + S_{\lambda})(1 + C)$$

A covariance matrix for the full set of observations (all bands, all actual measurements) would have as its diagonal the uncertainty associated with a single observation, squared, i.e.

$$u^2\left(E_{i,\lambda}\right) = \left(E_{i,\lambda}^{\mathrm{meas}}\right)^2 \times \left[u_{\mathrm{rel}}^2\left(R_{i,\lambda}\right) + u_{\mathrm{rel}}^2(S_{\lambda}) + u_{\mathrm{rel}}^2(C)\right]$$

The covariance for two measurements on the same band would be

$$u\left(E_{i,\lambda j}, E_{i,\lambda k}\right) = \left(E_{i,\lambda j}^{\mathrm{meas}}\right)\left(E_{i,\lambda k}^{\mathrm{meas}}\right) \times \left[u_{\mathrm{rel}}^2(S_{\lambda}) + u_{\mathrm{rel}}^2(C)\right]$$

And the covariance for two measurements in different bands would be

$$u\left(E_{i,\lambda j}, E_{i,\lambda k}\right) = \left(E_{ij,\lambda j}^{\mathrm{meas}}\right)\left(E_{ik,\lambda k}^{\mathrm{meas}}\right) \times \left[u_{\mathrm{rel}}^2(C)\right]$$

In practice because we have to do our model in terms of log reflectance, this model would first have to be propagated to reflectance (straightforward, analytic) and then to log reflectance (numerically).

APPENDIX D – Uncertainties in the TOA CIMEL data

This appendix contains the original text from the first version of Lunar irradiance Model Algorithm and Theoretical Basis Document that was written in 2019. In the current document many of the uncertainty related sections were improved and rewritten, nevertheless we keep the original text for the reference here.

Uncertainties in the Langley Plot intercept

The Langley Plot fits a straight line to the logarithm of signal as a function of airmass. As discussed in our report [AD2], the uncertainty associated with airmass is considered negligible. The uncertainty associated with the signal (corrected for instrument temperature effects and for lunar phase changes during the Langley[4], as well as for sun-moon and moon-Earth distances) is dominated by the noise in the measurement.

This measurement noise was estimated in the D4 report [AD2] from the statistics of the triplet (each observation being three observations made very close together in time).

Table 11 Table given in D4 (there Table 28) with additional line for combined standard uncertainty.

Term	Uncertainty [%]					
	1640nm	1020nm	870nm	675nm	500nm	440nm
$D(\lambda, t)$	0.07	0.05	0.02	0.01	0.03	0.04
$F_T(\lambda)$	0.0027	0.13	0.18	0.17	0.15	0.053
$F_T(\lambda)$	0.002	0.037	0.001	0.002	0.003	0.003
K_{dist}	0	0	0	0	0	0
$A(t_{ref},\lambda)/A(t,\lambda)$	0.006	0.006	0.006	0.006	0.006	0.006
+0 (aerosol's diurnal cycle)	0	0	0	0	0	0
Combined standard uncertainty	0.070	0.144	0.181	0.170	0.153	0.067

The residuals observed in the initial model fit however indicate that the uncertainty associated with individual Langley plots using these values is an underestimation.

In D4 we discussed that this could be in part because of changes in aerosol (and other atmospheric) properties during the Langley. We described there that at least some of such a variation may not be visible in the Langley analysis itself as the curve would "still look linear", however, since we do see variations in the Langleys that have a clear "shape" to them, we also allow for the possibility that these atmospheric variations may also cause random or semi-random effects in the Langley fitting.

Table 12: Uncertainty associated with atmospheric effects during the Langley [AD2].

	Uncertainty in Vo [%]							
	1640nm	1020nm	870nm	675nm	500nm	440nm	380nm	340nm
Aerosol	0.2	0.2	0.3	0.3	0.5	0.5	0.5	0.7
Other	0.17	0.25	0.01	0.12	0.17	0.19	0.31	0.5
Total	0.37	0.45	0.31	0.42	0.67	0.69	0.81	1.2

[4] The correction for lunar phase changes during the Langley is performed by iterating the Langley plot and lunar model fitting several times.

It is also reasonable to assume that using a standard deviation of the triplets, which shows instrument stability over a very short period of time, underestimate the uncertainty associated with the stability of the instrument for the duration of the Langley.

Fitting the Langley

The original fitting for this project used a simple least squares analysis fit that does not take into account the uncertainty associated with each data point, nor provides an uncertainty associated with the intercept.

Based on the residuals observed indicating a potential underestimation of the uncertainty associated with the Langleys, we performed a more rigorous analysis where each data point was given the same relative uncertainty taken from the standard deviation of the triplets. (See APPENDIX A for more information about how this fit was done and APPENDIX C for how logs were dealt with). The fit routine we used then also calculated the uncertainty associated with the intercept and the χ^2 value of the fit. Where the observed χ^2 was smaller than the expected χ^2, the intercept uncertainty was accepted as given. Where the observed χ^2 was larger than the expected χ^2, relative uncertainty on each data point was increased by small increments until the χ^2 test was passed.

Example Langley statistics

When performing the fit and applying the χ^2 test, very few of the Langley plots passed using the original uncertainty on the input parameters. This is as expected from the observed residuals in the model fit. For those that failed, uncertainty on input parameters was increased incrementally until the test was passed and uncertainty associated with the y-intercept, $\ln(V_0)$ was determined. A small selection required a small increase in uncertainty for each data point, and a few had high uncertainty indicating that the fit should be considered for removal from the dataset or have very low weighting in the final model.

Example Langley plots are shown below:

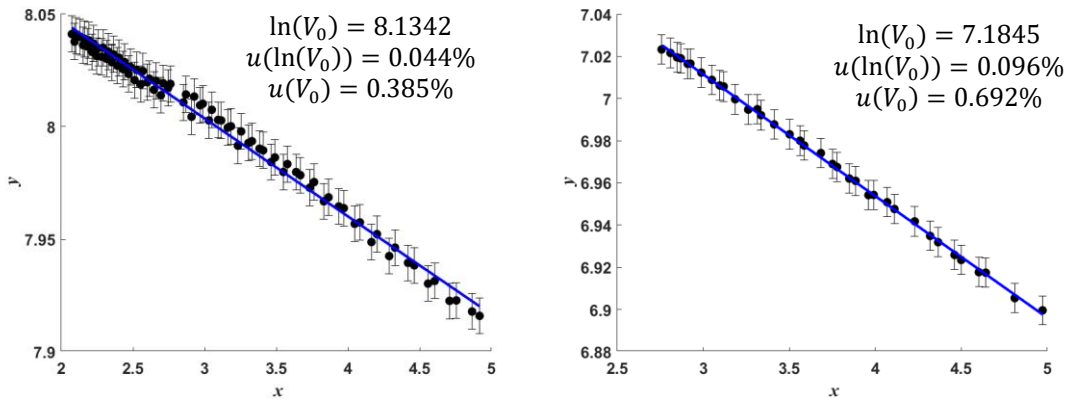

Figure 26: Langley plots which pass χ^2.

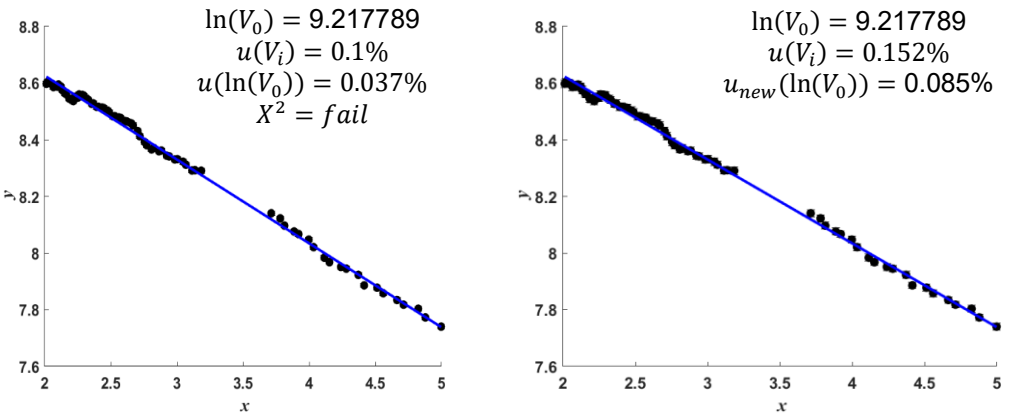

Figure 27: Langley before increase uncertainty (left) and after (right).

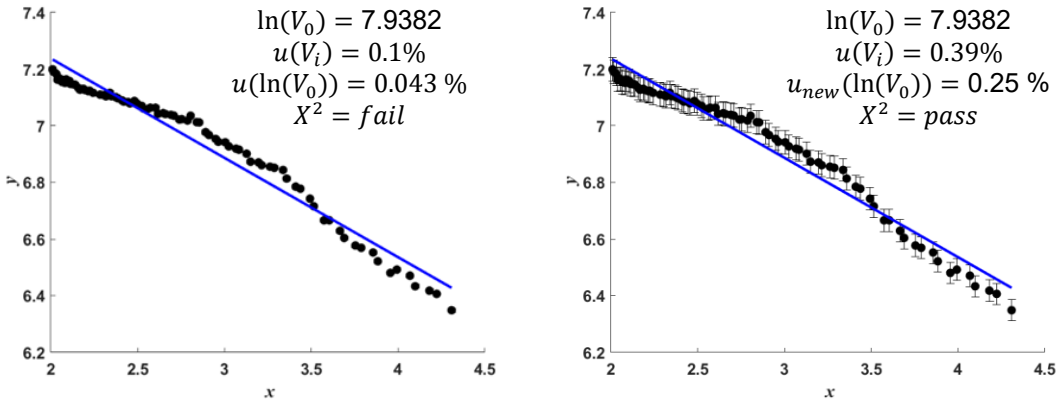

Figure 28: Langley before increase uncertainty (left) and after (right).

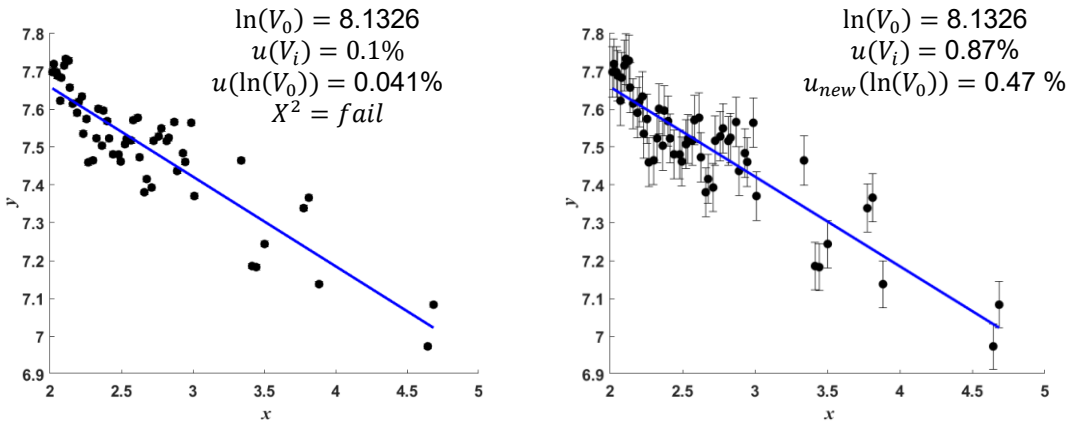

Figure 29: Langley before increase uncertainty (left) and after (right).

Using the values of $u(\ln[V_0])$ an and the $u_{new}(\ln[V_0])$ where appropriate we plotted a histogram of the uncertainty associated with the y-intercept of the Langley plots for each wavelength, and from this determined a 'typical' uncertainty that would be used in the Monte Carlo uncertainty analysis input parameters. Ideally, we should carry forward the individual uncertainty determined for each Langley and weight each value of V_0 determined from individual Langley plots in the model fit, and this is something to be considered in future improvements of the model.

Table 13: Estimated uncertainty in y-intercept of Langley plots.

	440 nm	500 nm	675 nm	870 nm	1020 nm	1640 nm
$\boldsymbol{u(ln[V_0])}$	0.35%	0.25%	0.16%	0.18%	0.47%	0.61%

The initial estimates for uncertainty in $ln\ (V_0\ )$ presented in Table 6 are the 'typical' uncertainties determined by statistical analysis of the range of intercept uncertainties for the set of Langley plots in this analysis. These estimates are obtained before considering the later outlier removal process implemented by UVa which iteratively removes any 3-sigma outliers during an initial model fit process.

A large proportion of Langley plots with high uncertainties in the fit are removed during the model fit outlier removal, however some are not, and this highlights the need to consider this fit uncertainty in the next iteration of the model.

After removal of these outliers identified in the model fit process we find that the typical uncertainty associated with the linear fit of the Langleys reduces to those values presented in . As a first approximation these values are used in this iteration of the model, although more ideally we would consider the uncertainty in each individual Langley plot and weight each data point in model fit accordingly. We also set an upper limit whereby any Langley plots with uncertainty 5 times higher than the typical uncertainty are currently removed from the dataset.

Table 14: Estimated uncertainty in the y-intercept of the Langley plots, after removal of those data points filtered by the model fit outlier removal process.

	440 nm	500 nm	675 nm	870 nm	1020 nm	1640 nm
$\boldsymbol{u(ln[V_0])}$	0.21%	0.16%	0.13%	0.12%	0.12%	0.21%

The uncertainties associated with systematic errors

The uncertainties associated with the systematic errors S_λ and C originate from the calibration of the CIMEL photometer at NPL and calculated from the uncertainty analysis outlined in deliverable D4 [AD4], and are described in the previous section.

The uncertainties associated with the calibration were separated into 4 categories:

a. Fully independent effects (e.g. noise) where the error varies statistically from observation to observation (e.g. is different at different distances and for the different methods),
b. Fully common effects (e.g. instrument alignment) where the error is (almost) identical for all measurements with all sources at all distances,
c. FEL399 effects (e.g. its calibration) where the error is common to all methods that use FEL399,
d. Method common effects where the error is common to the measurements at different distances with this method but which is different for other methods.

Here we identify that categories a and d vary from spectral band to spectral bad, and so will make up the S_λ input parameters for the MCUA and, categories b and d will be largely systematic for all channels and so make up C. These values are presented in .

Note that in D4 [AD4], we considered one spectral band at a time, and were combining results from multiple calibration measurements. Therefore, these categories do not transfer perfectly from that process to here (where band-to-band error correlation is more important). However, there is sufficient overlap to use those categories as a starting point. Future activity could propagate band-correlation separately.

Table 15 : Systematic uncertainties per band S_λ and to all measurements C.

	440 nm	500 nm	675 nm	870 nm	1020 nm	1640 nm
S_λ	0.77%	0.73%	0.55%	0.63%	0.31%	0.31%
C	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%

The uncertainties associated with random errors

The uncertainty associated with random errors is given by the uncertainty in the Langley Plot intercepts. We showed in that the uncertainty in the intercept could range from 0.16 % to 0.61 % however this is reduced once we consider the outlier removal in the model fit.

Because the uncertainty associated with systematic effects is, by definition, causing a common error to all the observation values being fitted by the lunar model, the variation between the measured data points and the model must be explained either by inaccuracies in the model form, or by the uncertainties associated with random effects.

We can use an initial estimate of the model parameters, calculated without taking uncertainty into account, to compute the residuals – the difference between the model and the measured value. These differences should fall within the uncertainties. To test this we plotted a histogram of these residuals divided by their associated uncertainties. i.e. we have

$$\Delta_{residual} = \frac{(V_{meas} - V_{model})}{u(V_{meas})}$$

Where, V_{meas} and V_{model} are the irradiance values for the measurement and model respectively and $u(V_{\mathrm{meas}})$ is the absolute uncertainty (i.e. the relative uncertainty in percent multiplied by the value $u_{\mathrm{rel}}(V_{\mathrm{meas}}) \times V_{\mathrm{meas}}$) associated with that measurement.

When the model is well fitted by the measurements, and when the measurement uncertainty is correctly estimated, this [either form!] will be a Gaussian[5] distribution with a mean of zero and standard deviation of 1. If the standard deviation is very much less than 1, then uncertainties are overestimated, if it is very much greater than 1, then uncertainties are underestimated and if the shape is not Gaussian, then the model may not be a good fit to the data, or there are significant outliers.

When we do this calculation with the initial uncertainties we get the following plots:

[5] For very small numbers of data points it will be a T-distribution.

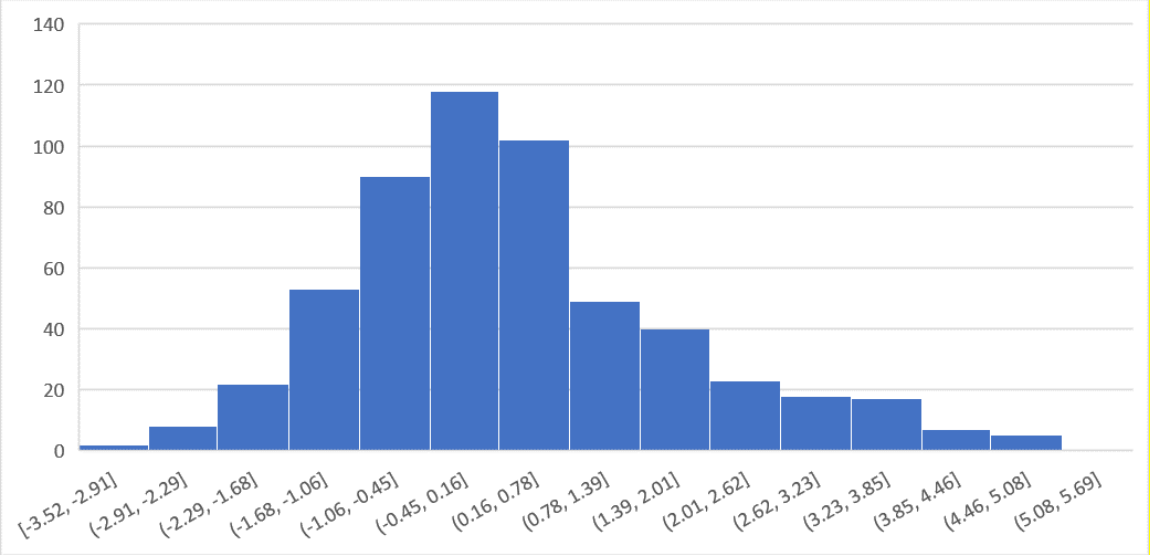

Figure 30: Relative residual histogram 1088+933 model before filtering.

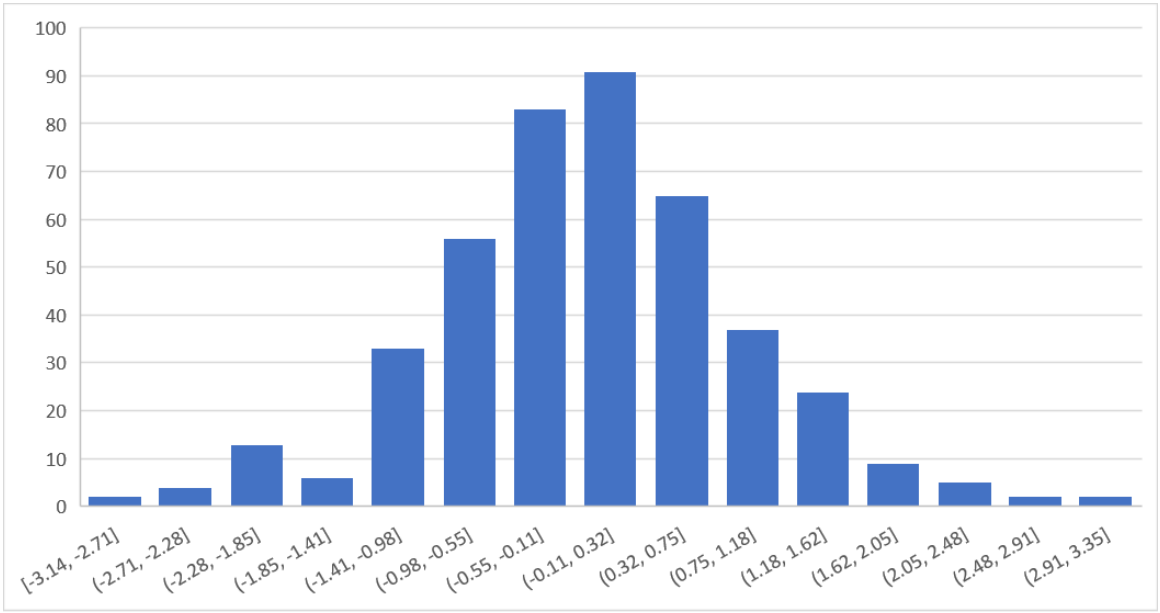

Figure 31: Relative residual histogram 1088+933 model after filtering.

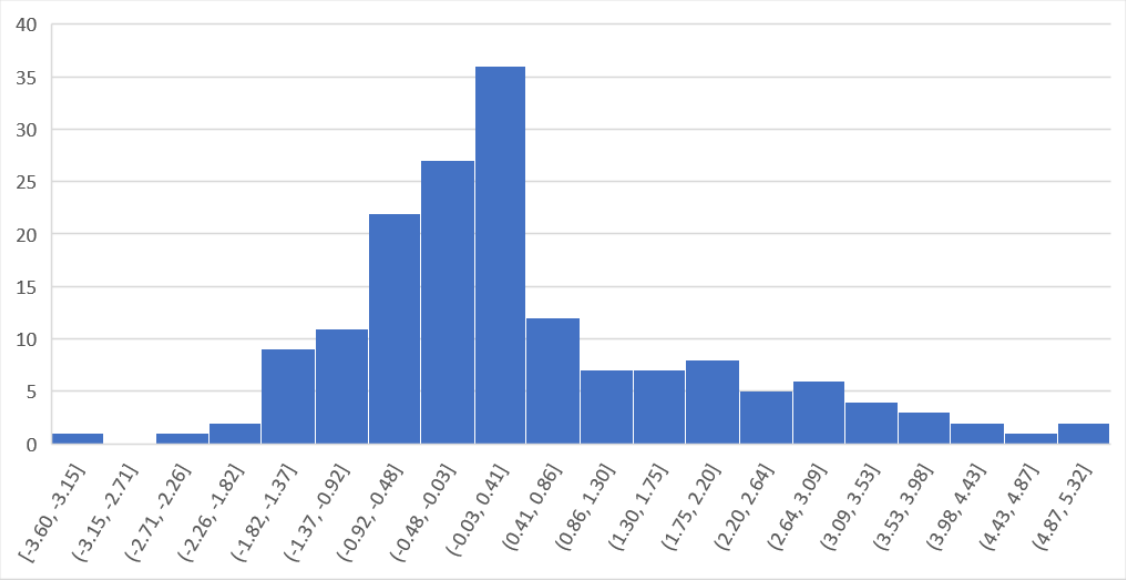

Figure 32: Relative residual histogram 1088 model before filtering.

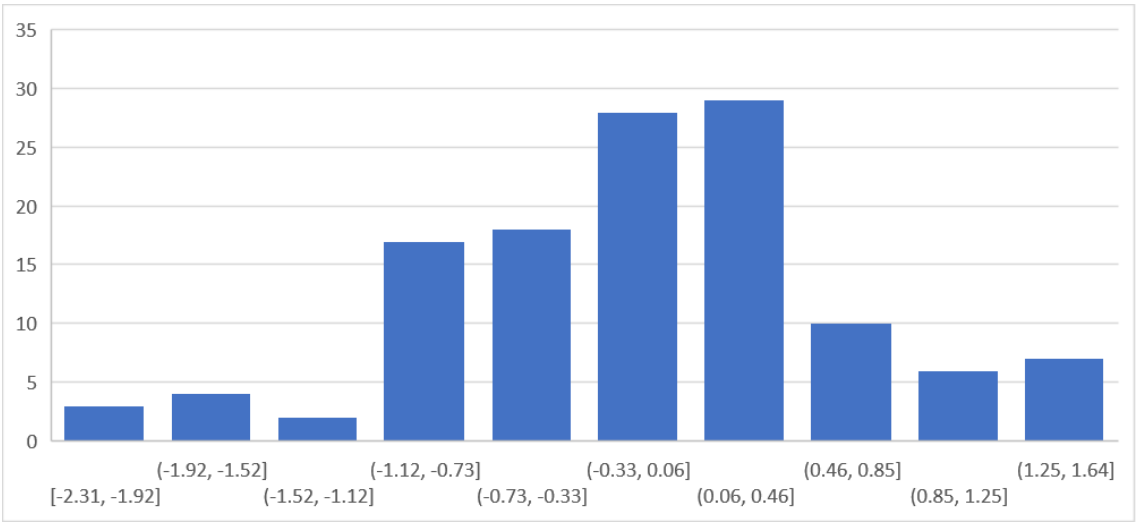

Figure 33: Relative residual histogram 1088 model after filtering.

The plots show that these residuals are more gaussian shaped after the filtering process. It is also quite clear that more samples (555 for 1088+933 as opposed to 166 for 1088) results in a more standard gaussian distribution.

Statistics in the next table show that after filtering, the averages are already close to zero for both models and the standard deviation is close to one, after the filtering process. We hope to be able to improve the number, when we updated the model to include individual uncertainties for every measurement.

Table 16 : Statistics associated with model relative residuals.

model	average unfiltered	stdev unfiltered	average filtered	stdev filtered
1088+933	0.298	1.475	-0.026	0.951
1088	0.386	1.535	0.011	0.927

Uncertainties in the derivation of the model

Fitting the lunar model

The lunar model fit is described in section 2. This is a multistep process where the linear part of the model is fit for each band, then outliers are removed, then the non-linear part is fit (all bands simultaneously), there is further outlier removal and finally the linear part is fit again. The whole multistep process is itself iterated.

To understand the uncertainties associated with the method, we use Monte Carlo Uncertainty Analysis (MCUA). The MCUA process is based on a measurement model. In this case we treat the input irradiance values (the TOA irradiance values for each night obtained by the Langley Plot process) as

$$E_{i,\lambda} = E_{i,\lambda}^{\text{True}} \times \left(1 + R_{i,\lambda}\right)(1 + S_\lambda)(1 + C)$$

Here:

$E_{i,\lambda}^{\text{True}}$ is the nominal "true" value for the TOA irradiance in spectral band λ for the ith observation.

$R_{i,\lambda}$ is the error in the observation in spectral band λ for the ith observation due to random effects, expressed in relative terms, e.g. as a percentage of the true value.

S_λ is the error in the observation that is common for all measurements in this band, expressed in relative terms, e.g. as a percentage of the true value.

C is the error in the observation that is common for all measurements in all bands, expressed in relative terms, e.g. as a percentage of the true value.

The error values are unknown; but are draws from a probability distribution with a standard deviation given by the relative uncertainty associated with this effect and with an expectation value (central value) of zero.

$R_{i,\lambda}$ takes a different value for every observation. This comes from random processes relating to the measurement of the TOA irradiance for a particular night. These include instrument noise, instrument temperature changes and atmospheric changes – and relates to the relative uncertainty in the Langley Plot intercept.

S_λ takes the same value for every observation for a single spectral band. This comes from effects that are common for that band – and mostly that is from the NPL calibration of the instrument. Any uncertainty associated with the NPL calibration is "fixed" into that calibration and applied to all measurements.

C takes the same value for every single observation in all spectral bands. This comes from effects in the NPL calibration that are wavelength independent, e.g. from a distance offset on an instrument alignment.

Performing the Monte Carlo Uncertainty Analysis

The MCUA is performed only for the final iterative step. Here, we run the fit routine 1000 times. For each iteration we create a single value of the error C drawn randomly from a Gaussian distribution with a central value 0 and a standard deviation equal to the uncertainty associated with C. We draw 6 values S_λ, each corresponding to a different spectral band, and we draw as many values of $R_{i,\lambda}$ as are needed – the number of spectral bands multiplied by the number of observations.

Conceptually, we alter the input values by these errors, perform the fit and then obtain a model based on those errors. We then repeat this 1000 times to give 1000 different models. Thus, in pseudo code we have:

```
For k = 1 to 1000
    Choose C from a random Gaussian of width u(C)
    Choose S_lambda for each band from the appropriate random Gaussians
    For i = 1 to the number of observations, N
        Choose a set of R_i for each band from Gaussians
        Calculate irradiance from Equation (above)
        Fit the model, get fit parameters set i
    Loop observations
Loop Monte Carlo run
```

Practically, we need to modify this somewhat to account for the fact that the fit occurs on the logarithm of reflectance rather than the irradiance, and is for the final iteration of the final, post outlier removal, linear fit.

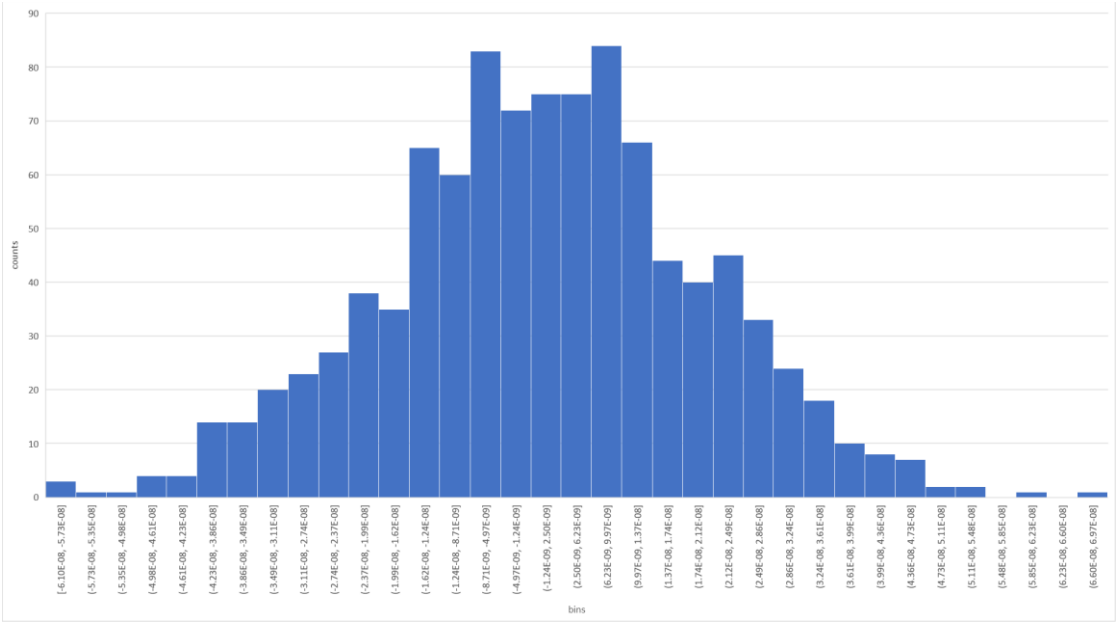

Figure 34: example of perturbated input irradiance for one measurement.

For one iteration, all measurements of all bands are perturbated with a factor (%) calculated from the above scheme - 1000 times a new model is fit.

Calculating the uncertainty and covariance of the fit parameters

At the end of the MCUA process we have 1000 versions of the model that differ from one another in a way that is consistent with the uncertainties and covariances of the input quantities. We can use these to estimate the uncertainty associated with the model. The first step is to estimate the uncertainty associated with the fit parameters and the covariance between pairs of fit parameters.

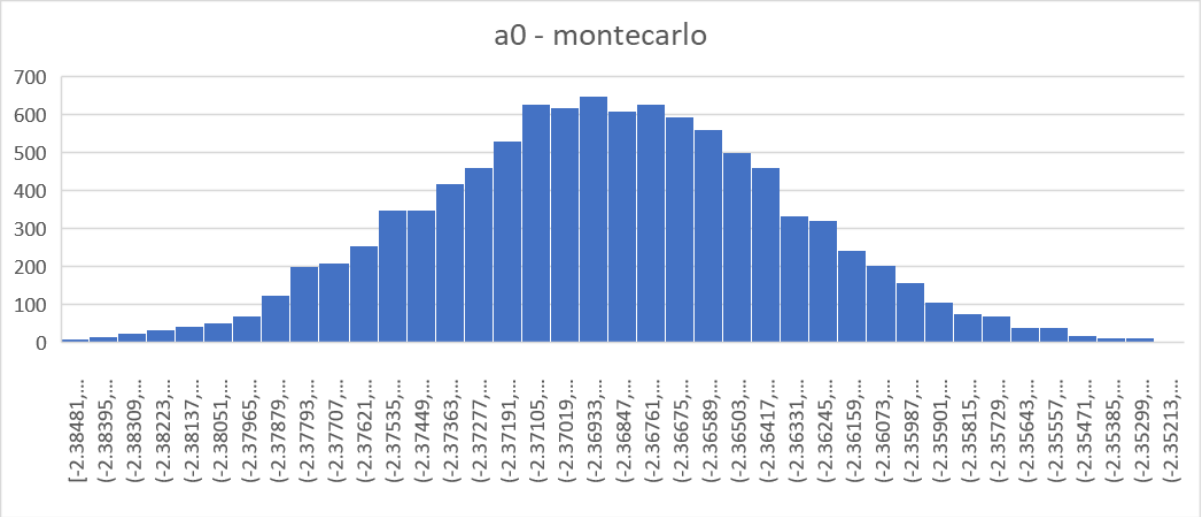

Figure 35 : Monte Carlo output for coefficient a0.

This is done statistically. The uncertainty is determined by taking the sample standard deviation of the 1000 instances of each fit parameter. We get for the first band (440 nm):

Table 17: Statistics for model 440 band (mean fit parameter from MCUA, standard deviation of the 1000 estimates, and the standard deviation expressed as a percentage of the mean value).

wl[nm]	**a0**	**a1**	**a2**	**a3**	**b1**	**b2**	**b3**
440	-2.76443	-0.77946	-0.28446	-0.02657	0.051998	0.011877	-0.00584
wl[nm]	**c1**	**c2**	**c3**	**c4**	**d1**	**d2**	**d3**
440	0.00144	-8.4E-05	0.001911	0.001031	1.111408	2E+132	0.003075
	p1	**p2**	**p3**	**p4**			
all	1.448495	18.99534	10.77744	9.002714			

It is also valuable to estimate the correlation coefficient of the different fit parameters for a single spectral band, or for the different spectral band for a given fit parameter. This is calculated using the standard formula for the sample Pearson correlation coefficient.

	a0	a1	a2	a3	b1	b2	b3	c1	c2	c3	c4	d1	d2	d3	p1	p2	p3	p4
a0	1.000000	-0.539161	0.530155	-0.518675	0.070539	-0.123045	0.149257	0.031141	-0.007576	-0.029292	-0.057731	0.282807	-0.283089	0.248621	-0.404896	-0.404829	-0.319940	0.319177
a1	-	1.000000	-0.995087	0.982249	-0.129463	0.226003	-0.275749	-0.058253	0.034145	0.046046	0.113490	-0.515960	0.516468	-0.454136	0.731448	0.731092	0.579855	-0.578564
a2	-	-	1.000000	-0.995748	0.137411	-0.242317	0.298290	0.060667	-0.039522	-0.046223	-0.127477	0.500355	-0.500858	0.436304	-0.714102	-0.714017	-0.562116	0.560846
a3	-	-	-	1.000000	-0.145271	0.259627	-0.323479	-0.064996	0.042775	0.048682	0.145481	-0.484153	0.484647	-0.419807	0.694384	0.694501	0.543619	-0.542380
b1	-	-	-	-	1.000000	-0.886272	0.745499	0.009393	0.079067	0.029071	-0.060304	0.067467	-0.067538	0.063356	-0.093377	-0.093412	-0.071492	0.071398
b2	-	-	-	-	-	1.000000	-0.959775	0.002928	0.003774	0.066227	0.147046	-0.121846	0.121954	-0.104609	0.164763	0.164198	0.136932	-0.136656
b3	-	-	-	-	-	-	1.000000	0.007164	-0.047092	-0.112758	-0.210064	0.145869	-0.145997	0.118802	-0.199157	-0.198383	-0.167309	0.166918
c1	-	-	-	-	-	-	-	1.000000	-0.402845	-0.053920	-0.019270	0.058748	-0.058770	0.055017	-0.051686	-0.050823	-0.050337	0.050590
c2	-	-	-	-	-	-	-	-	1.000000	0.037962	-0.062008	-0.029820	0.029820	0.001789	0.019704	0.019199	0.023283	-0.023477
c3	-	-	-	-	-	-	-	-	-	1.000000	-0.388492	-0.038111	0.038130	-0.036052	0.040598	0.039759	0.039443	-0.039466
c4	-	-	-	-	-	-	-	-	-	-	1.000000	-0.004544	0.004616	-0.009300	0.047189	0.049170	0.011108	-0.010816
d1	-	-	-	-	-	-	-	-	-	-	-	1.000000	-0.999999	0.627920	-0.717032	-0.690887	-0.894020	0.898366
d2	-	-	-	-	-	-	-	-	-	-	-	-	1.000000	-0.628000	0.717719	0.691612	0.894047	-0.898389
d3	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000000	-0.584447	-0.567397	-0.689432	0.689495
p1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000000	0.998949	0.796552	-0.794968
p2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000000	0.769137	-0.767422
p3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000000	-0.999948
p4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000000

Figure 36: Correlation matrix 440 nm coefficients.

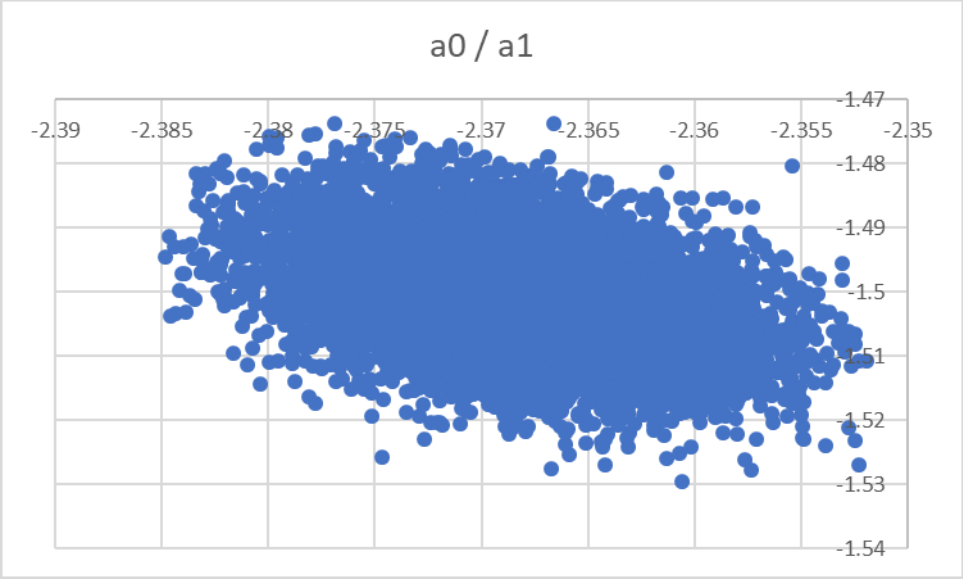

Figure 37: Low correlation between a0 and a1.

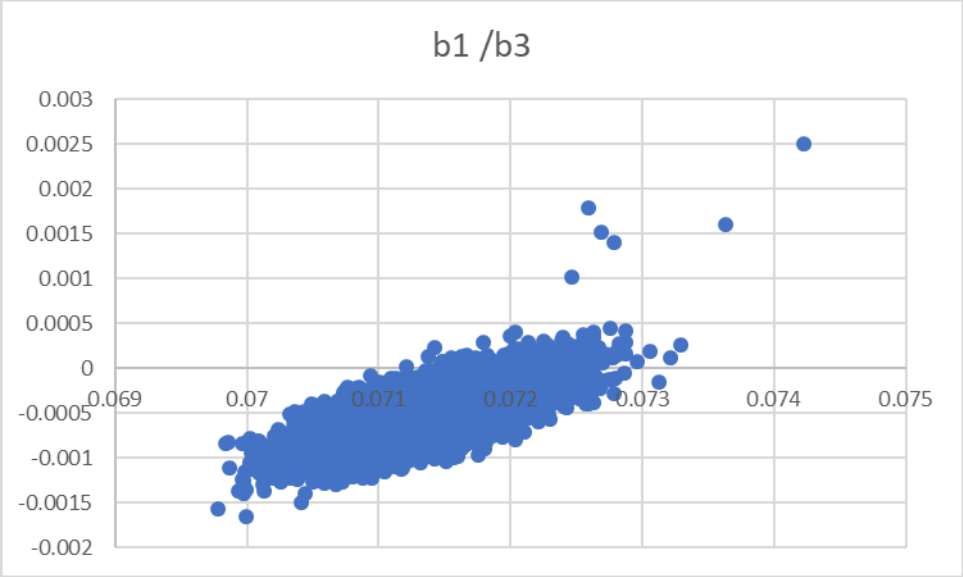

Figure 38: High correlation of 0.74 between b1 and b3 band.

Any high correlation (>0.5) suggests that the model is not well defined – i.e. that you could add something to one and compensate by removing it from the other. Such models can be hard to fit, so this may indicate the value of removing one of those parameters. Permanent evaluation of the MCUA results is part of the future work when measurements become available. These results might ultimately result in an updated formulation of the model. Currently we don't draw any conclusions on these results, as individual measurement uncertainties are not yet included, as well as the number of 1088 measurements is limited to 2 years.

A covariance matrix is calculated for the fit parameters by having on-diagonal terms equal to the square of the absolute uncertainty associated with that fit parameter and the off-diagonal terms equal to $u(a_i, a_j) = u(a_i)(a_j)r(a_i, a_j)$.

	a0	a1	a2	a3	b1	b2	b3	c1	c2	c3	c4	d1	d2	d3	p1	p2	p3	p4
a0	3.430E-05	-3.642E-05	3.874E-05	-1.279E-05	1.947E-07	-5.325E-07	2.351E-07	4.567E-09	-1.126E-09	-5.044E-09	-9.766E-09	1.475E-01	-1.476E-01	3.070E-07	-1.432E-04	-1.279E-04	-1.524E-05	2.011E-08
a1	-	1.330E-04	-1.432E-04	4.770E-05	-7.039E-07	1.926E-06	-8.552E-07	-1.683E-08	9.996E-09	1.562E-08	3.781E-08	-5.299E-01	5.305E-01	-1.104E-06	5.094E-04	4.549E-04	5.440E-05	-7.179E-08
a2	-	-	1.557E-04	-5.231E-05	8.082E-07	-2.234E-06	1.001E-06	1.896E-08	-1.252E-08	-1.696E-08	-4.594E-08	5.559E-01	-5.565E-01	1.148E-06	-5.380E-04	-4.806E-04	-5.705E-05	7.528E-08
a3	-	-	-	1.773E-05	-2.884E-07	8.079E-07	-3.663E-07	-6.854E-09	4.572E-09	6.028E-09	1.770E-08	-1.815E-01	1.817E-01	-3.727E-07	1.766E-04	1.578E-04	1.862E-05	-2.457E-08
b1	-	-	-	-	2.222E-07	-3.087E-07	9.450E-08	1.109E-10	9.461E-10	4.030E-10	-8.211E-10	2.832E-03	-2.835E-03	6.297E-09	-2.658E-06	-2.375E-06	-2.741E-07	3.621E-10
b2	-	-	-	-	-	5.461E-07	-1.907E-07	5.419E-11	7.078E-11	1.439E-09	3.139E-09	-8.018E-03	8.026E-03	-1.630E-08	7.352E-06	6.546E-06	8.231E-07	-1.086E-09
b3	-	-	-	-	-	-	7.230E-08	4.824E-11	-3.214E-10	-8.915E-10	-1.632E-09	3.493E-03	-3.496E-03	6.735E-09	-3.234E-06	-2.878E-06	-3.660E-07	4.829E-10
c1	-	-	-	-	-	-	-	6.272E-10	-2.561E-10	-3.970E-11	-1.394E-11	1.310E-04	-1.311E-04	2.905E-10	-7.816E-08	-6.866E-08	-1.025E-08	1.363E-11
c2	-	-	-	-	-	-	-	-	6.443E-10	2.833E-11	-4.546E-11	-6.740E-05	6.741E-05	9.575E-12	3.020E-08	2.629E-08	4.807E-09	-6.411E-12
c3	-	-	-	-	-	-	-	-	-	8.646E-10	-3.300E-10	-9.979E-05	9.984E-05	-2.235E-10	7.208E-08	6.307E-08	9.434E-09	-1.248E-11
c4	-	-	-	-	-	-	-	-	-	-	8.343E-10	-1.169E-05	1.187E-05	-5.664E-11	8.231E-08	7.662E-08	2.610E-09	-3.361E-12
d1	-	-	-	-	-	-	-	-	-	-	-	7.930E+03	-7.930E+03	1.179E-02	-3.856E+00	-3.319E+00	-6.476E-01	8.606E-04
d2	-	-	-	-	-	-	-	-	-	-	-	-	7.931E+03	-1.179E-02	3.859E+00	3.323E+00	6.476E-01	-8.607E-04
d3	-	-	-	-	-	-	-	-	-	-	-	-	-	4.445E-08	-7.440E-06	-6.453E-06	-1.182E-06	1.564E-09
p1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.646E-03	3.254E-03	3.913E-04	-5.164E-07
p2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.910E-03	3.375E-04	-4.454E-07
p3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.617E-05	-8.751E-08
p4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.157E-10

Figure 39: covariance matrix for all model coefficients of 440nm.

Reflectance model total covariance matrix

In further processing, the total covariance matrix is used as input to the calculation of the individual uncertainties.

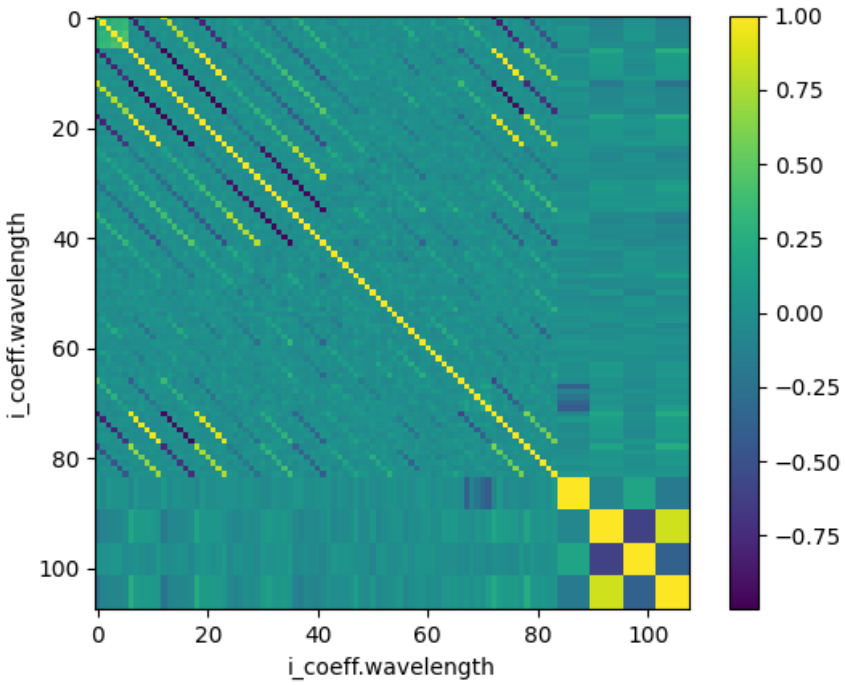

Figure 40: covariance matrix for all model coefficients

Calculating the uncertainty associated with the model

The full model combines these spectral band observations with a hyperspectral lunar reflectance profile, based on either lunar rock lab measurements [RD2] or a phase depended ASD ground measurements, performed during the most recent developments in the project developments.

In the current iteration of the model we assume (erroneously, but in the absence of other information) that there is no uncertainty associated with the observations used in the spectral interpolation, nor with the interpolation process itself. Instead, we will simply propagate our uncertainties by continuing the MCUA, creating 1000 hyperspectral moon models and considering the variability. Uncertainty associated with the spectral observations and interpolation will be included in later phases of the project.

The irradiance is calculated for every generated input measurement used in the coefficient regression. The results per input measurement is the 1000x applied model. The irradiance is calculated using the instrument response curves.

An example of this output is plotted, showing a normal like distribution. Such plot could be produced for every measurement and hence the mean and standard deviation is calculated.

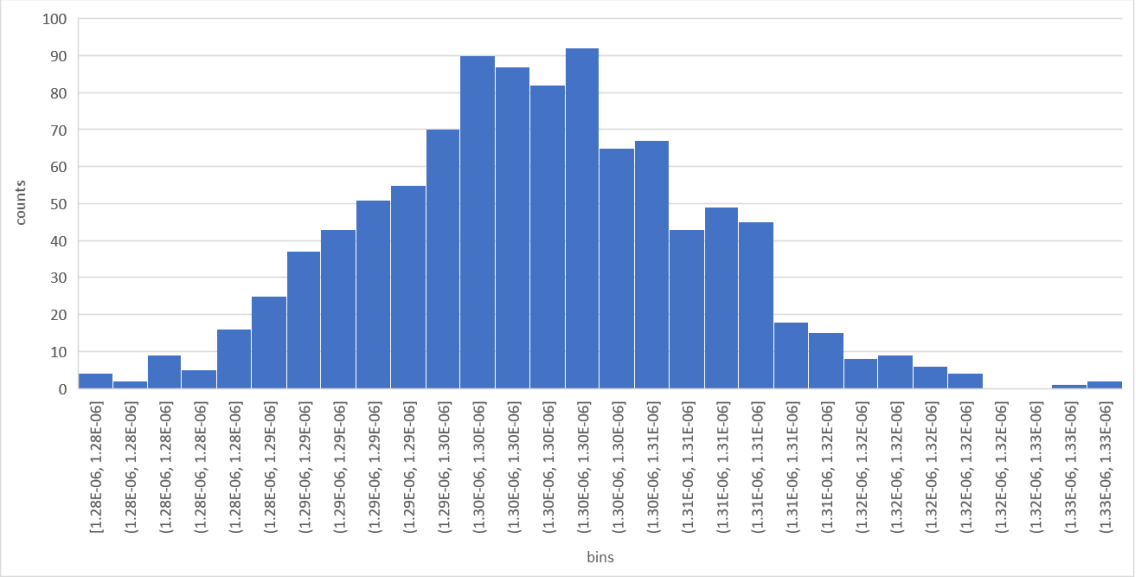

Figure 41: Distribution of the results for one input measurement of band 440nm.

The extended relative uncertainty is calculated from the standard deviation of the output distributions.

$$u(E_{i,\lambda}) = k * \frac{s(E_{i,\lambda})}{E_{i,\lambda}}$$

$E_{i,\lambda}$ is the measured irradiance value, before any perturbation is applied, $u(E_{i,\lambda})$ is the relative uncertainty, *k* is the factor applied to obtain the required confidence interval for a normal distribution, $s(E_{i,\lambda})$ is the standard deviation of all models results for this measurement.

From the results, as shown in the uncertainty level is quite stable over phase angle, except for angles close to outside the model phase angle limits [-90.0:-2.0-2.0:90.0]. In the summarizing plots, the uncertainty values have been averaged per 5° phase angle.

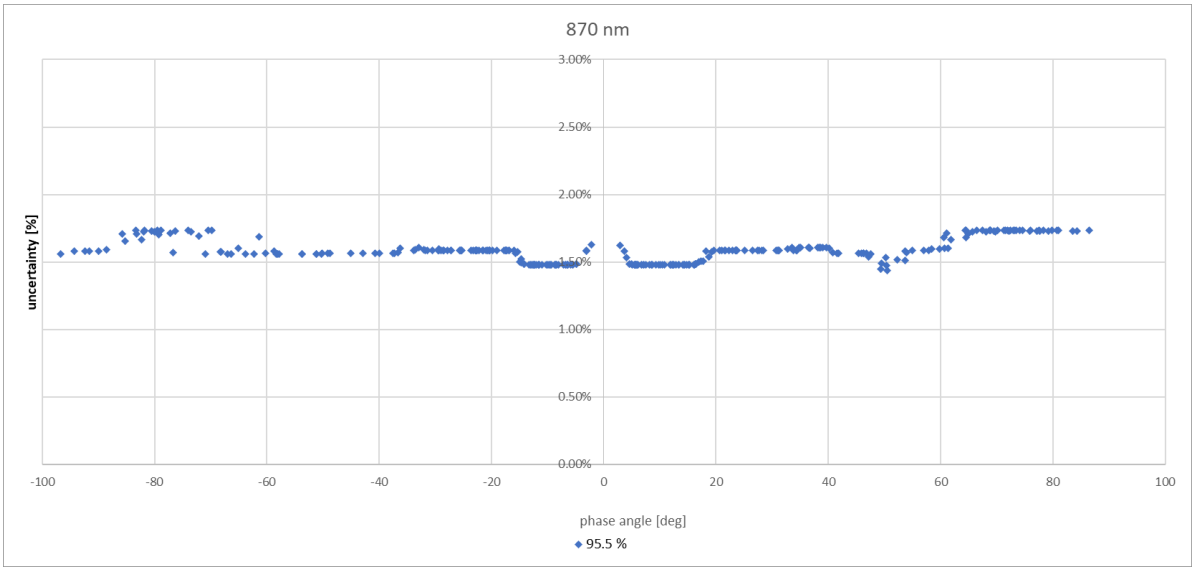

Figure 42: 95.5% uncertainty band 870 nm.

up to show the uncertainty levels for all bands, averaged per 5 degrees. Uncertainty levels at 95.5 % (k = 2.0) and 99.7 % (k = 3.0) confidence level are shown, as well as the mean $E_{i,\lambda}$ obtained averaged over 1000 results. This is not the irradiance from the "root" model output.

When looking at the plots, one can see that for the 95.5% confidence interval, all bands perform well below 2 % except for band 440 nm.

For the 99.7 % confidence level, all bands perform approximately at 2.5 % uncertainty, except for the 400 nm and 500 nm bands, which are slightly above.

The model and its uncertainties here are provided for the 1088 instrument. It must be noted that the Langley plots with the high uncertainty in the intercept are currently excluded in the model fitting process. The measurements with higher uncertainties introduce a much higher scatter in the residuals for the non-linear part.

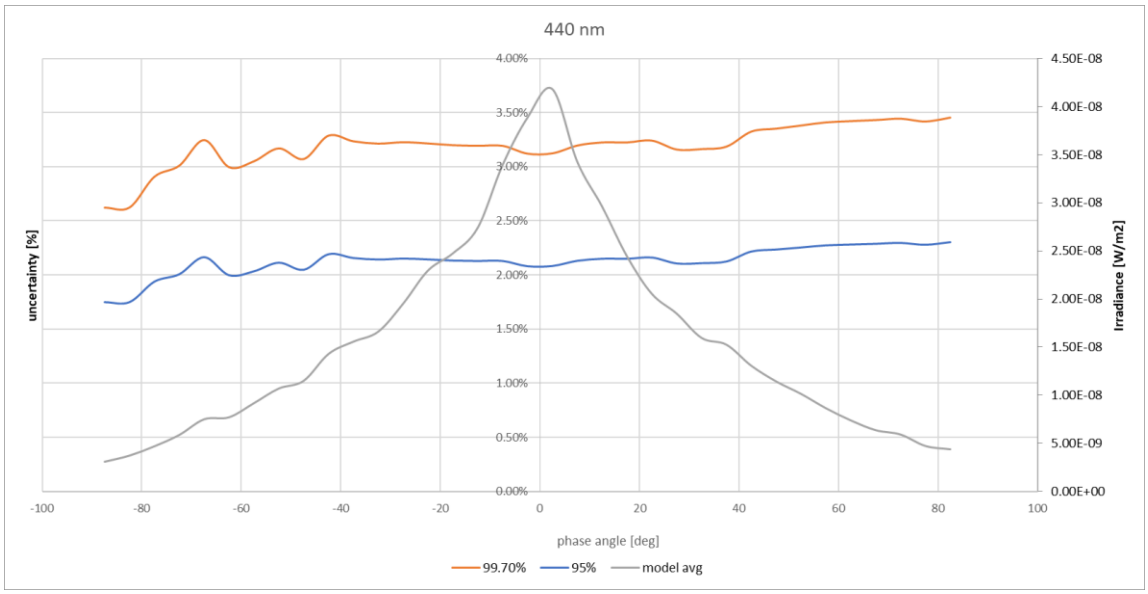

Figure 43: Uncertainty levels for the 440 nm band.

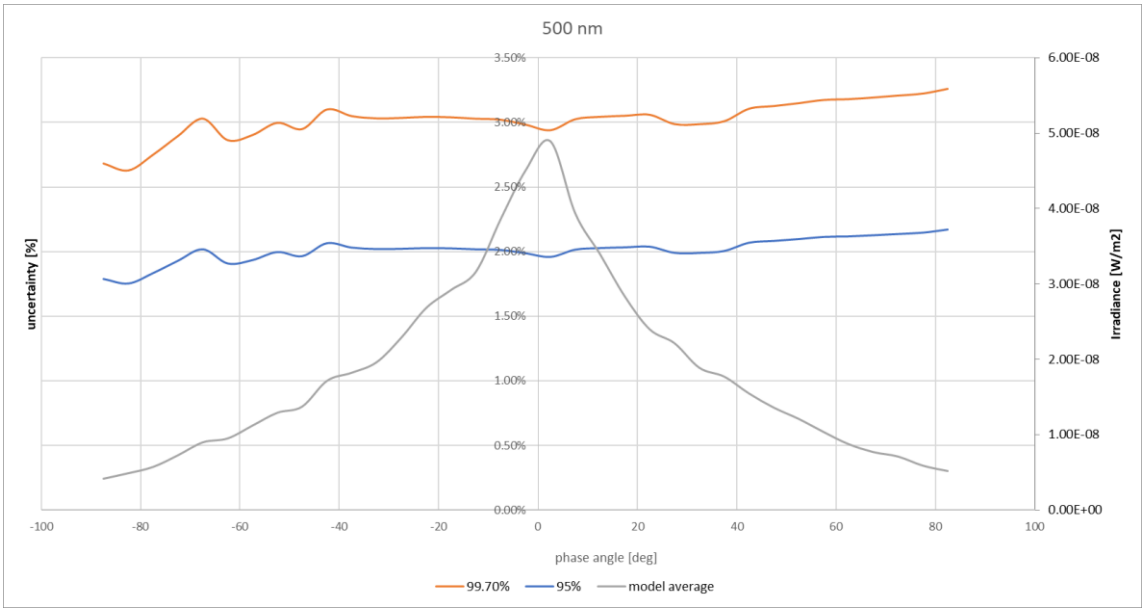

Figure 44: Uncertainty levels for the 500 nm band.

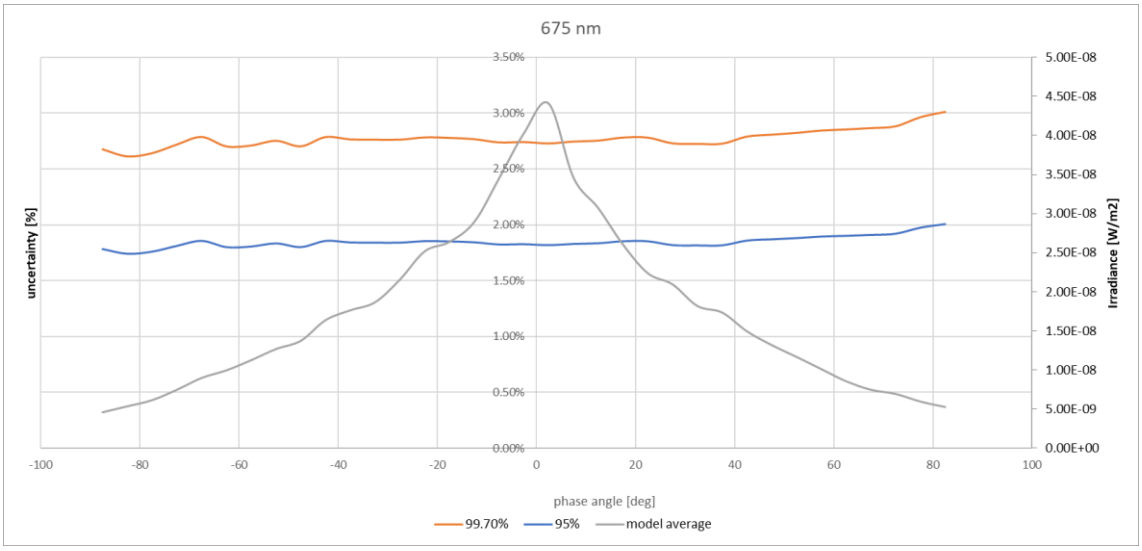

Figure 45: Uncertainty levels for the 675 nm band.

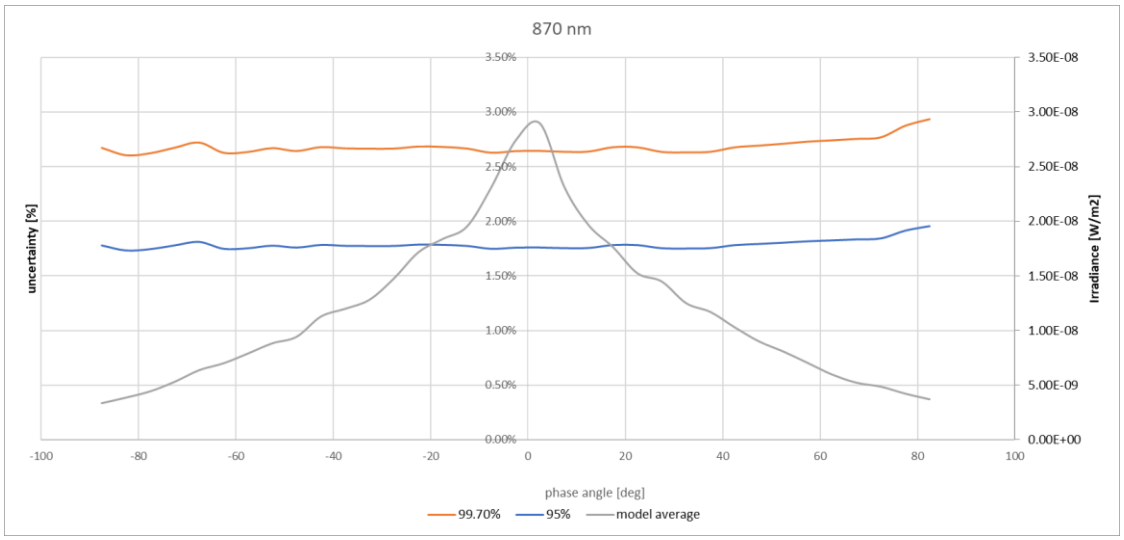

Figure 46: Uncertainty levels for the 870 nm band.

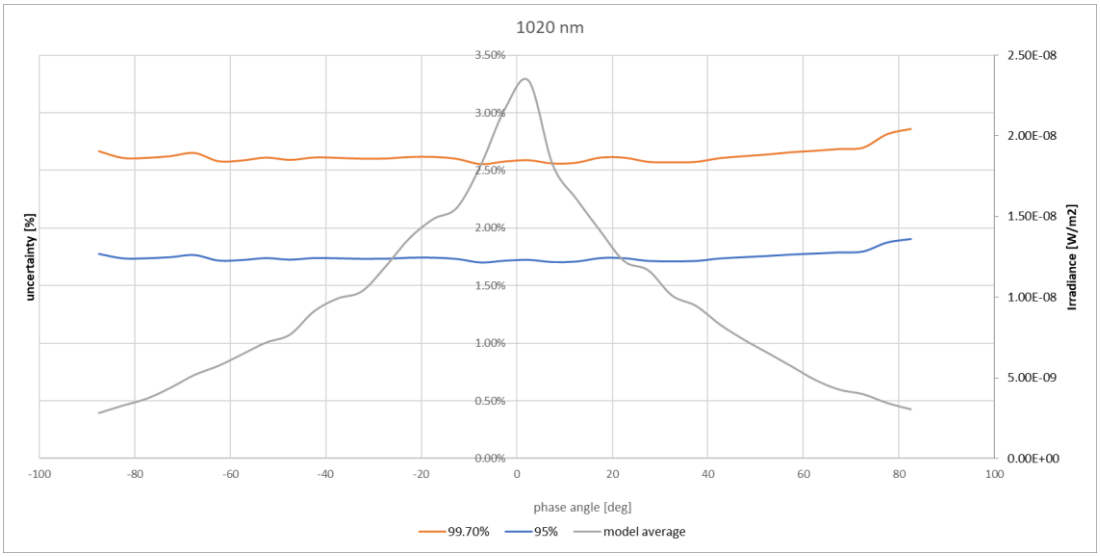

Figure 47: Uncertainty levels for the 1020 nm band.

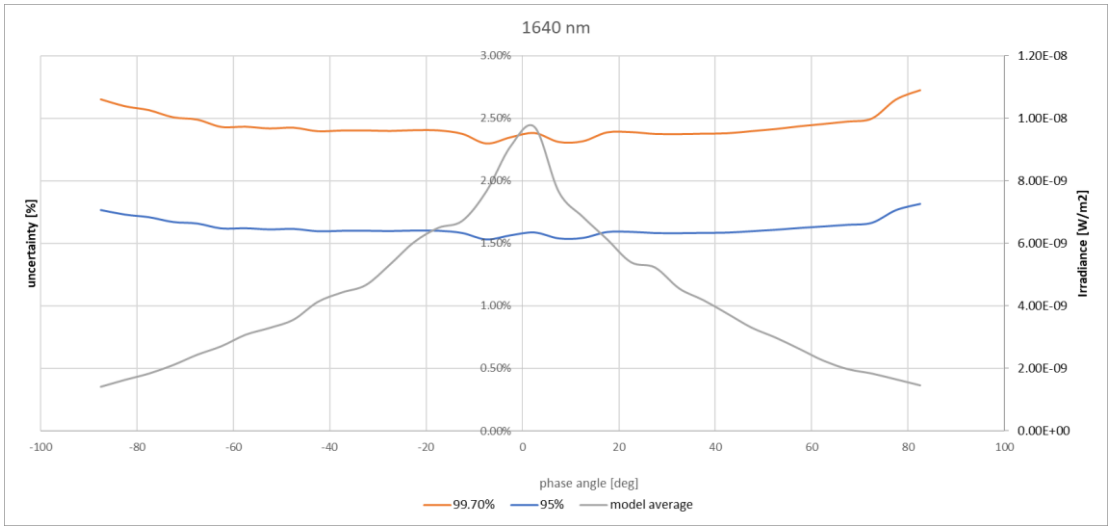

Figure 48: Uncertainty levels for the 1640 nm band.

Evolution of model uncertainties with number of measurements

To get a clear understanding of the evolution of the uncertainty levels with respect to the amount of used measurements, the MCA is performed on both the 1088 measurements only and the 1088+933 measurements. In the beginning of the project measurements, there was little data and therefore the regression needed to be done on datasets from both instruments to get decent regression.

It appears that in general the uncertainty level stays the same for both regressions. There is a slight difference of around 0.2% with large positive phase angles and about 0.1% for specific low phase angles. This can be interpreted as very low differences. This must be re-evaluated during the process of model iterations, when more 1088 measurements become available. Applying the uncertainty budgets from the 1088 instrument on the 933 instrument is not good practice and should be avoided.

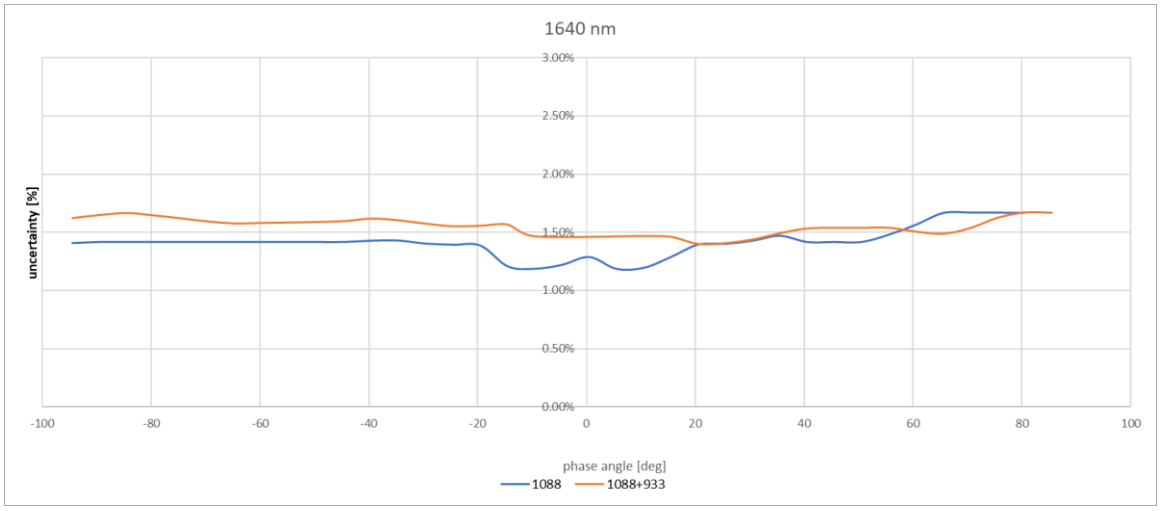

Figure 49: Uncertainty level for both 1088 and 1088+933 model regression