Rayleigh, Deep Convective Clouds, and Cross-Sensor Desert Vicarious Calibration Validation for the PROBA-V Mission

Sindy Sterckx, Stefan Livens, and Stefan Adriaensen, *Member, IEEE*

***Abstract*—PROBA-V is a remote sensing satellite mission for global monitoring of vegetation. It is designed to offer almost daily coverage of all land masses and to provide data continuity with the VGT2 instrument aboard SPOT-5. Accurate radiometric calibration is key to the success of the mission; therefore, a comprehensive system for in-flight radiometric calibration has been developed. Without no onboard calibration devices, this in-flight calibration will rely fully on vicarious methods. In total nine techniques for vicarious calibration have been implemented and tested in order to meet the radiometric mission requirements. Three key methods that contribute largely to the calibration performance are presented in this paper: Rayleigh, deep convective clouds, and cross-sensor calibration over stable desert sites. As the PROBA-V sensor has still to be launched, calibration algorithm verification is performed using data from the spectrally very similar SPOT-VGT1 and SPOT-VGT2 sensors.**

***Index Terms*—Cross-sensor desert, deep convective clouds (DCC), rayleigh, vicarious calibration.**

I. INTRODUCTION

THE VEGETATION instruments on board of the SPOT-4 and SPOT-5 satellites have provided users with global low-resolution remote sensing images with almost complete daily coverage of all land masses for more than 10 years. The main objective of the missions is vegetation monitoring [1]: to provide accurate measurements of basic characteristics of vegetation canopies on an operational basis. This can be employed either for scientific studies involving both regional and global scales experiments over long time periods (e.g., development of models of the biosphere dynamics interacting with climate models), or for systems designed to monitor important vegetation resources, like crops, pastures, and forests.

As the SPOT satellites are close to or beyond the end of their foreseen lifespan, the European Space Agency has started development of a new remote sensing micro-satellite mission to ensure data continuity. The PROBA-V mission (V for Vegetation) falls within the Project for On-Board Autonomy (PROBA) program, dedicated to satellite missions that demonstrate innovative technologies. In this case, novel technologies include a very compact wide field of view telescope and new multispectral detectors.

The PROBA-V satellite will fly at an altitude of 825 km and make 16 near-polar orbits per day (similar to SPOT) to achieve daily global coverage for latitudes above 35° and two-daily coverage for lower latitudes.

To offer high quality data continuity, the radiometric accuracy of PROBA-V should match or exceed that of SPOT-VEGETATION (VGT) and provide excellent correspondence between both. The radiometric calibration requirements for PROBA-V specify 5% absolute accuracy and 3% relative accuracy (inter-band, multi-temporal, and camera relative).

A comprehensive preflight laboratory characterization and calibration of the PROBA-V instrument will be performed at CSL (Liège, Belgium) prior to integration on the platform at QinetiQ Space (Kruibeke, Belgium). Radiometric and spectral performance characteristics to be verified are: signal to noise, dark currents, linearity, stray light, pixel non-uniformity, polarization sensitivity, spectral response, and spectral misregistration. However, variations in the characteristics of the instrument are likely to occur in orbit due to outgassing phenomena during launch, aging of the optical parts, and cosmic ray damage. This makes it necessary to perform in-orbit stability monitoring and calibration. Many satellite missions are equipped with on onboard calibration devices as inner lamps, LEDs and solar diffuser plates. These are omitted in the PROBA-V instrument due to constraints on size, weight and power consumption. Hence, the PROBA-V in-flight calibration relies solely on vicarious methods.

In this paper, we present the pre-launch vicarious validation activities including 1) the verification of radiometric calibration algorithms on the basis of existing spectrally similar EO sensors and 2) the characterization of uncertainty. The paper is organized as follows. In Section II, the PROBA-V platform and instrument are introduced. In Section III, the implementation of Rayleigh, deep convective clouds (DCC), and desert cross-sensor calibration approaches is elaborated, and the expected sources of uncertainty are reviewed. Section IV lists the data used for testing and validating the algorithms. Section V describes the statistical approach for image-based uncertainty analysis. In Section VI, the validation results for SPOT-VGT1 and SPOT-VGT2 data are presented and discussed.

Manuscript received January 31, 2012; revised July 27, 2012 and November 9, 2012; accepted November 20, 2012. Date of publication February 5, 2013; date of current version February 21, 2013. This work has been supported by Belspo and ESA (PROBA-V User Segment contract 091669).

The authors are with the Flemish Institute for Technological Research, Remote Sensing and Earth Observation Processes, 2400 Mol, Belgium (e-mail: sindy.sterckx@vito.be; stefan.livens@vito.be; stefan.adriaensen@vito.be).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2012.2236682

II. PROBA-V

The main design challenge for PROBA-V was to offer the functionality and data quality of the SPOT satellites in a much smaller platform. Compared to SPOT-5, the PROBA-V platform has to achieve reductions in volume (from 22.4 m^3 to 0.38 m^3), weight (from 3000 kg to 160 kg) and power consumption (from 2100 W to 153 W). Consequently, the PROBA-V imaging instrument needs similar reductions compared to the VEGETATION instruments (from 0.85 m^3 to 0.0567 m^3 and from 152 kg to 33.3 kg). The reduction of the instrument could only be achieved by a radically different instrument design. Since it still has to cover the same very wide swath of 2250 km (corresponding field of view of 101°), an inventive solution is used.

As a single camera system consisting of a telescope and imaging sensors would be prohibitively large, three identical smaller camera systems are used together. Each camera has an equal field of view, the center camera pointing down covers a swath of about 500 km, while the side cameras pointing left and right each cover about 875 km. The telescopes themselves are three-mirror-anastigmats, a compact optical design with strong light collection. The complete telescopes are made of aluminum, which ensures that temperature variations do not affect the adjustment of the optics. Per camera, image acquisition is achieved using a set of push-broom (line) detectors that image four spectral bands. The spectral characteristics are chosen to closely match those of SPOT-VGT. The BLUE, RED, and near infrared (NIR) sensors are hosted on a common chip; they each have 5200 pixel wide arrays in the across-track direction. The short-wave infrared (SWIR) detectors are specifically developed for this system; it consists of an array of three staggered sensors, each 1024 pixels wide.

For the central camera, this results in imaging with a ground sampling distance (GSD) of 100 m in the across-track direction, with matching imaging in the along-track direction. The side cameras view the earth under larger angles, their across-track GSD varies from 100 m to more than 300 m at the far sides. The SWIR imaging pixels are twice as large but result in a similar geometry: 200 m × 200 m in the center varying to 200 m × 600 m at the edge of the side cameras. The GSD is very non-uniform, but in any case smaller than the GSD of SPOT-VGT. This allows to generate data products at higher spatial resolution. In fact, products at two different resolutions are generated: a 1 km product for maximal direct compatibility and a 300 m product which offers an important increase in resolution [2].

III. METHODS

A. Vicarious Calibration Concept

The relation between the raw digital output as registered by the sensors and the effective spectral radiance assumed to be present at the sensors is defined by the sensor radiometric model. Taking into account possible nonlinearities in the sensor response, the relationship between digital output and effective radiance can be written as

$$DN_j^k / NL\left(DN_j^k\right) = A^k \cdot g_{jm}^k \cdot G_m^k \cdot L_j^k + dc_{jm}^k \tag{1}$$

where the superscript k and subscript j identify, respectively, the spectral band and the pixel. DN is the raw digital output A the absolute calibration coefficient, L the effective radiance, G the gain and m the gain number, and dc the dark signal. g is pixel relative sensitivity or equalization coefficient. NL is the nonlinearity function. The nonlinearity is expected to be small and common for all pixels j.

The goal of the calibration activities is to quantify the parameters of the model. Initial values are determined by the pre-launch calibration measurements. After launch, these parameters are monitored, validated and, if needed, updated by the vicarious calibration activities. For every parameter, specific image acquisitions and methods will be employed [3]. For example, the dark current will be estimated from night-time observations over dark oceans, while the multi-angular equalization coefficients will be derived from images taken from homogeneous snow-covered areas over Antarctica or Greenland. The calibration methods treat the three PROBA-V cameras separately, which may introduce biases between them. However, at the edge between two cameras, a small overlap zone is imaged simultaneously by both cameras. Data of these zones will be used to verify possible inter-camera bias.

The largest effort is spent on on the accurate determination and monitoring of the absolute calibration coefficients (A^k) for the four spectral bands of the three cameras. Multiple independent approaches are used for this. The advantages of this are that systematic errors inherent to one or more techniques can be dealt with, random errors can be reduced by statistical averaging and that results can be validated independently [4].

An overview of the different methods used for the determination of A^k for the PROBA-V mission is given in Table I. Note that most methods cannot yield results for all spectral bands. The Rayleigh calibration approach allows to determine A^{BLUE} and A^{RED}. The results can be transferred to other spectral bands using sun glint and DCC calibrations which are "relative" inter-band approaches. Deserts sites with high spatial homogeneity and stability and low cloud coverage are ideal targets for absolute calibration and stability monitoring. The use of lunar observations for multi-temporal monitoring is under investigation. A statistical methodology for combining results in a hierarchical scheme based on handling and propagating of accuracies in accordance with the ISO GUM will be used to generate suitable operational calibration coefficient on the basis of the results obtained with the different methods. The approach is described in [5]. A reflectance-based method using simultaneously in-situ measured ground spectra of homogenous reference sites (e.g., Tuz Gölü in Turkey) will be also used at an ad hoc basis to validate the obtained calibration coefficients. From the methods listed, three key methods that contribute largely to the calibration performance are chosen and elaborated in this paper: Rayleigh, DCC, and cross-sensor desert calibration.

The selected methods allow to check several of the radiometric calibration requirements: with the Rayleigh method, the absolute calibration of blue and red band can be investigated, inter-band calibration accuracy can be evaluated using the DCC method, and radiometric consistency of PROBA-V versus SPOT-VGT can be assessed through the use of the cross-sensor desert method. While the use of the near simultaneous nadir

TABLE I
RADIOMETRIC CALIBRATION METHODS AND THEIR APPLICABILITY

Methods/ Targets	Absolute BLUE	Absolute RED	Absolute NIR	Absolute SWIR	Relative: Interband	Temporal	Inter camera bias	Cross Sensor	Equalization	Dark current
Rayleigh	X	X	(ref)							
DCC	X	(ref)	X		X					
Sun Glint	X	(ref)	X	X	X					
Deserts	X	X	X	X		X		X		
Moon						X		X		
Reflectance based	X	X	X	X						
Inter Camera Overlap							X			
Dark Ocean										X
Antarctica/Greenland									X	

observation method [6] requires simultaneous observations made by the reference sensor, the cross-sensor desert method can be used even after the lifetime of SPOT-VGT. In contrast to some other methods, like calibration over the Dome-C site in Antarctica [7] or on the basis of the Tuz Gölü salt lake in Turkey which can only be used during the local summers, the chosen methods can be applied through all seasons. Furthermore, no coincident labor-intensive ground-based measurements are required.

B. Rayleigh

Method: The principle behind Rayleigh calibration approach is that when the atmosphere is observed over stable, homogenous oceans, most of the signal originates from Rayleigh scattering [8]–[11]. The TOA reflectance can be predicted well using a radiative transfer model and then compared to the values measured by the sensors.

The primary assumption is that the ocean itself does not contribute to the top-of-atmosphere (TOA) signal in the NIR [12]. This holds for Case 1 waters with low chlorophyll concentration (CHL) and where phytoplankton is the only optically significant water column contributor. Suitable oligotrophic Rayleigh calibration zones have been identified by [13]. In a recent study by [14], the monthly variation of marine reflectance for the defined zones is analyzed on the basis of 9 years of SEAWIFS data. In a similar study, [15] the monthly averaged CHL concentrations for six selected oligotrophic areas (similar to or located within the zones given by [13]) are reported. Difference in the seasonal variation of CHL occurs. The seasonal variation is less pronounced and featureless in the North-West of Pacific and in the North Pacific gyre. For this latter, the highest annual mean CHL (0.059 mg m-3) are observed. The seasonal CHL pattern for the other four oligotrophic zones is comparable and rather reproducible from year to year with peaks occurring around July–August.

In these oligotrophic zones, the contribution of aerosol scattering (i.e., the aerosol optical thickness) can be derived from the NIR reference band where molecular scattering is very low assuming a fixed aerosol type. Following [9], the most suitable aerosol model to adopt for the Rayleigh calibration is the Shettle and Fenn Maritime aerosol model with 98% humidity (denoted as M98) [16]. Once the AOT is determined on the basis of the NIR band, the TOA reflectance in the other bands $(\rho_{\mathrm{TOA}}^{k,\mathrm{model}})$ (i.e., $k =$ blue and red bands) can be modeled with a radiative transfer code using the appropriate marine reflectance and/or CHL. The modeled TOA reflectance can then be compared with the sensor retrieved values $(\rho_{\mathrm{TOA}}^{k,\mathrm{model}})$ to determine the change in absolute calibration coefficients (ΔA^k) as

$$\Delta A_k = A^{k,\mathrm{new}} / A^{k,\mathrm{old}} = \rho_{\mathrm{TOA}}^{k,\mathrm{sensor}} / \rho_{\mathrm{TOA}}^{k,\mathrm{model}}. \qquad (2)$$

Implementation: As it is computationally prohibitive to run a radiative transfer model for every pixel of all the Rayleigh calibration images, a look-up-table (LUT) approach is used. The LUTs are created using 6SV. Just like its scalar predecessors, the vector version of 6S is based on the method of successive orders of scattering. 6SV takes into account the effects of radiation polarization, which is extremely important for radiative transfer calculations over dark targets such as ocean surfaces. Ignoring polarization can lead to large errors in calculated TOA reflectance: more than 10% for a molecular atmosphere and up to 5% for an aerosol atmosphere [17]. The marine reflectance in 6SV is calculated using the semi-analytical reflectance model described in [18] using CHL content as input.

In Fig. 1, the flowchart of the Rayleigh calibration workflow is given. To reduce the perturbing part of the signal due to ocean reflectance and presence of foam, strict pixel selection procedures are used. Only pixels at a distance of at least 30 km from clouds are retained. To minimize foam radiance, meteorological data are used to select zones with low wind speed (less than 5 m/s). Finally, only pixels outside sun glint direction with $\theta_n > 20^\circ$ are selected where

$$\theta_n = a\cos\left((\cos\theta_s + \cos\theta_\nu) / \left(2\cos\left(\frac{\theta p}{2}\right)\right)\right) \qquad (3)$$

with

$$\theta_p = a\cos(\cos\theta_s \cos\theta_\nu + \sin\theta_s \sin\theta_\nu \cos\phi)$$

with θ_s the solar zenith angle, θ_ν the view zenith angle, and ϕ the relative azimuth angle.

The TOA reflectance $(\rho_{\mathrm{TOA}}^{k,\mathrm{sensor}})$ is corrected for gaseous absorption of O_3 and O_2

$$\rho_{\mathrm{TOA}}^{c,k,\mathrm{sensor}} = \rho_{\mathrm{TOA}}^{k,\mathrm{sensor}} / T_r^k(\mathrm{gas}, M)$$

with

$$T_r^k(\mathrm{gas}, M) = \exp\left(a_{\mathrm{gas}}^{\mathrm{k}}.(\mathrm{airmass}.U_{\mathrm{gas}}){}^\wedge n_{\mathrm{gas}}^{\mathrm{k}}\right) \qquad (4)$$

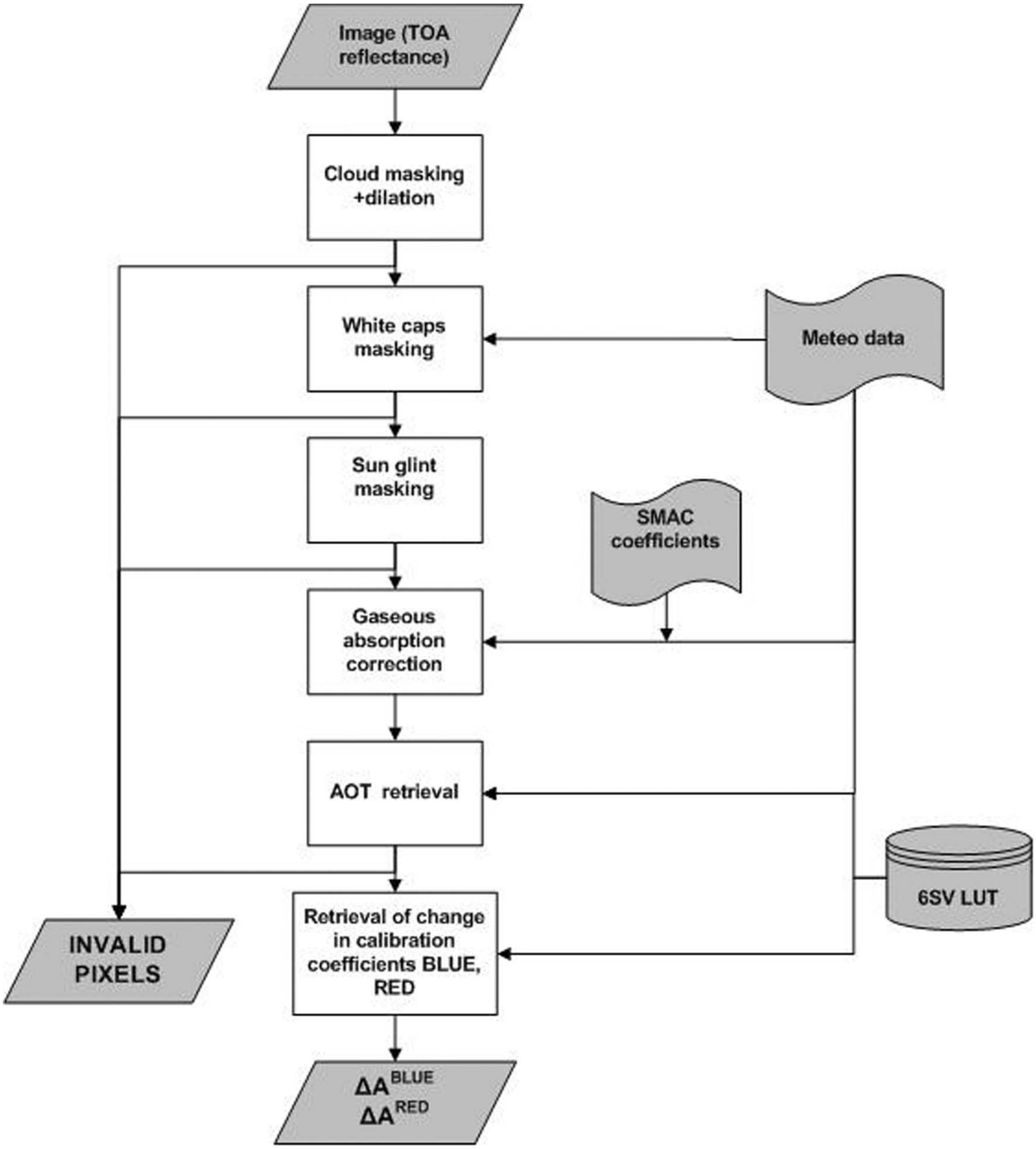


Fig. 1. Rayleigh calibration workflow.

where U_{gas} is the gaseous concentration (e.g., total amount ozone in units of cm/atm), and $a_{\text{gas}}^{\text{k}}, n_{\text{gas}}^{\text{k}}$ coefficients which depend on the response of the given spectral band. Details can be found in [19] and the SMAC model [20]. Next, the AOT per pixel is retrieved from the NIR band based on a LUT. This LUT contains the modeled $\rho_{\text{TOA}}^{c,k,\text{model}}$ for an atmosphere bounded by Fresnel reflecting wind roughened oceanic surface in function of sun zenith angle, view zenith angle, relative azimuth angle, wind speed, pressure, water vapor, CHL, and AOT. The AOT value for which the $\rho_{\text{TOA}}^{c,k,\text{model}}$ fits best with the modeled $\rho_{\text{TOA}}^{c,k,\text{model}}$ is searched for. Pixels with a retrieved AOT value above a threshold AOT value (i.e., 0.05 for NIR band) are masked. Finally, the LUT is interpolated (on the angles, water vapor, windspeed, pressure, and AOT) for each pixel to obtain $\rho_{\text{TOA}}^{c,k,\text{model}}$ for blue and red bands. Following (2) this results in an estimate of the change in absolute calibration coefficients for both red and blue bands per pixel.

Uncertainty Analysis: We perform an uncertainty analysis for the method presented and start by listing estimated uncertainties for a number of input parameters (e.g., aerosol model). Next, the error budget is determined by running the Rayleigh calibration algorithms on synthetic data, taking into account the uncertainties. The results are analyzed statistically and yield the error budget associated with respect to the uncertainty assumed for the input parameters. The considered sources of uncertainty are:

- Surface pressure: a maximal error of 10 hPa is used in the error analysis.
- Wind speed: The wind speed impacts the contribution of photons scattered by the atmosphere after their reflection over the sea surface. A typical accuracy of 2 m/s in wind speed is used in making up the error budget [21].
- NIR calibration: an error of 3% is used as input to the error analysis.

TABLE II
ERROR BUDGET RAYLEIGH CALIBRATION

	2 sigma error (%)	
Error source	BLUE	RED
Pressure	0.693	0.425
Wind speed	2.501	3.335
NIR calibration	0.720	2.065
Aerosol model	1.331	2.580
Ozone	0.036	0.333
Water vapour	0.157	0.181
Chlorophyll	1.261	0.388
Total absolute	**3.262**	**4.745**

- Aerosol model: The aerosol model is assumed fixed (M98); Uncertainties in the calibration method related to the aerosol type are analyzed by calibrating a scene generated with a coastal aerosol (C70) with a LUT generated with M98.
- Gaseous transmittance: An error in the ozone amount and water vapor of, respectively, 5% and 20% is used in the error budget [21]. An uncertainty on water vapor will have a small direct effect on the red band, and an indirect impact on the blue band due to the aerosol retrieval in the NIR band.
- Chlorophyll content: an error of 50% on the chlorophyll concentration has been used for making up the error budget. This will have a direct effect on the blue and red band; however, as the water surface can be assumed totally black in the NIR band, an error in the chlorophyll content will have no effect on the AOT retrieval.

The error analysis is performed by using a LUT generated according to the reference specifications. This LUT is then used to perform a Rayleigh calibration on an artificial image generated according to the same specifications except for the parameter of interest (e.g., C70 model instead of M98 model in LUT) or by introducing an error in the meteo-data (e.g., 20% error in water vapor). Table II summarizes the error budget (at 2-sigma) for the Rayleigh method for both blue and red band. In the blue band, the performance is determined by the accuracy on the wind speed, marine reflectance, and aerosol model. In the red band, the error in the NIR calibration, wind speed, and aerosol model are the limiting factors to accuracy.

C. DCC

Method: DCC are bright targets which have a predictable reflectance. This is exploited as the basis for the DCC calibration method: predicted TOA reflectance values are compared against values obtained by the sensors.

DCC clouds develop over subtropical warm oceans in intertropical latitudes between 30°N and 30°S. They are at the tropopause level, and hence effects of water vapor and tropospheric aerosol absorption are minimized. After correction of the ozone absorption DCC have almost a perfect white spectral behavior in visible and near-infrared (VNIR) bands. The cloud reflectance in the non-absorbing VNIR bands is mainly sensitive to the cloud optical thickness (COT), while in the absorbing SWIR bands, it is more sensitive to the droplet effective radius [22]. Due to the almost flat spectral signature, DCC are adequate for inter-band calibration [23], [24] using a well-calibrated band as reference band to derive the COT from the image pixel itself by comparing image $\rho_{\mathrm{TOA}}^{k,\mathrm{sensor}}$ data with modeled $\rho_{\mathrm{TOA}}^{k,\mathrm{model}}$ data. The COT estimated from the reference band is then used to model the cloud TOA reflectance in the other VNIR bands.

Implementation: For the generation of the DCC TOA reflectance LUTs, the LibRadtran radiative transfer code is used. LibRadtran takes into account cloud phase, microphysical properties as well as a complete description of the background atmosphere and surface. As DCC are usually thicker than 10 km, the cloud top and cloudbase heights are assumed to be 15 km and 1 km, respectively [25]. The ice particle model from [26] is used to create the LUTs. The Baum model is also employed in the MODIS operational ice cloud optical depth retrieval [27] and by [25] for calibration over DCC. The Baum model is based on the use of in-situ observations of ice particle sizes and habits to compute optical properties for a realistic ensemble of theoretical particles. A fixed effective cloud radius of 20 μm is used as cloud reflectance for the VNIR bands is rather insensitive to the effective radius. According to [25], MODIS effective particle radii for DCCs show a narrow distribution with maximum frequency at 20 μm. To calculate the Rayleigh scattering, the pressure profiles from the tropical profile are used. The ocean reflectance is assumed lambertian with a chlorophyll content of 0.1 mg/m^3.

The DCC calibration workflow is shown in Fig. 2. First, $\rho_{\mathrm{TOA}}^{k,\mathrm{sensor}}$ is corrected for the gaseous transmittance of O_3 based on predefined exponential variation with airmass and ozone content (4). As for convective clouds observations, water vapor and oxygen gases lie below the cloud level, their absorption can be neglected. The ozone layer is, however, located mainly above the clouds, and therefore ozone absorption cannot be neglected. Image pixels not suitable for DCC calibration are eliminated in a pixel selection step. These are pixels close to the specular direction or pixels not situated in large bright homogeneous DCC clouds. To reject observations corresponding to viewing geometries near the specular direction, only pixels for which the relative azimuth is between 30 and 150° are kept. To ensure that only pixels within large homogeneous clouds are kept, $\rho_{\mathrm{TOA}}^{\mathrm{NIR},\mathrm{sensor}}$ should be larger than 0.8 for the selected pixel (NIR reflectance test) and for all neighboring pixels within a Np by Np pixel window where Np depends on the spatial resolution of the sensor (Cloudy neighborhood test). Furthermore, the standard deviation of $\rho_{\mathrm{TOA}}^{\mathrm{NIR},\mathrm{sensor}}(NpXNp)/\rho_{\mathrm{TOA}}^{\mathrm{NIR},\mathrm{sensor}}$ (centerpixel) should be less than 0.03 (Homogeneity test). Next, for each pixel, the COT is retrieved from the reference red band based on a LUT approach. The LUT of simulated $(\mathrm{LUT\_}\rho_{\mathrm{TOA}}^{c,k,\mathrm{model}})$ over DCC are interpolated for the red band on the sun and viewing angles to obtain a $(\mathrm{LUT\_}\rho_{\mathrm{TOA}}^{c,k,\mathrm{model}})$ in function of only COT. Then, $\rho_{\mathrm{TOA}}^{c,\mathrm{red},\mathrm{sensor}}(\mathrm{COT}_m)$ and $\rho_{\mathrm{TOA}}^{c,\mathrm{red},\mathrm{model}}(\mathrm{COT}_{m+1})$ that surround $\rho_{\mathrm{TOA}}^{c,\mathrm{red},\mathrm{sensor}}$ are searched for. Next, COT_m and COT_{m+1} are

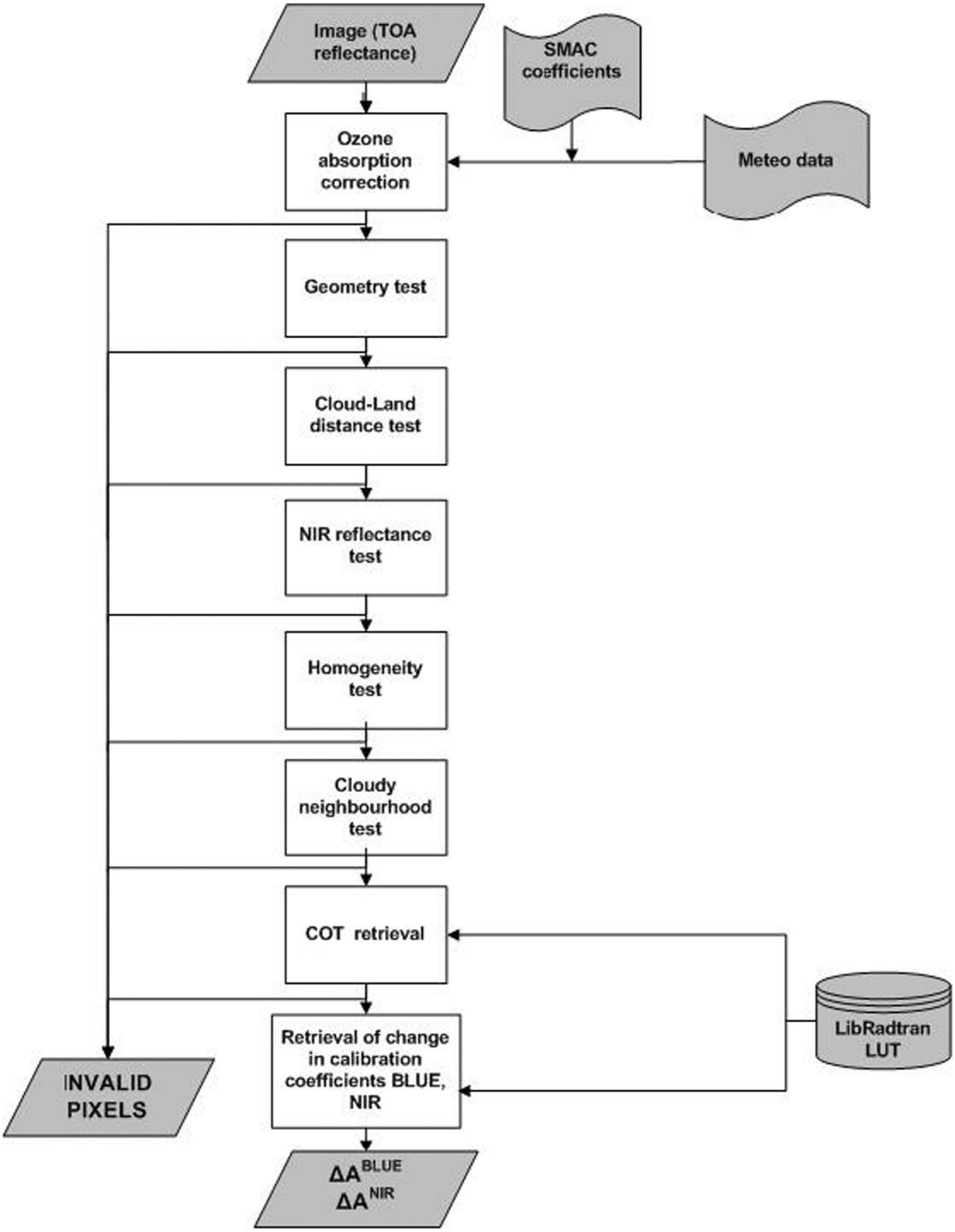


Fig. 2. DCC calibration workflow.

interpolated to retrieve the per pixel. Finally, LUT of simulated LUT_$\rho_{\mathrm{TOA}}^{c,\mathrm{blue},\mathrm{NIR},\mathrm{sensor}}$ are interpolated (on the angles and the COT) to obtain for each pixel. The estimated change in calibration coefficient can then be calculated following (2).

Uncertainty: The uncertainty analysis for the DCC method is conceptually similar to that for Rayleigh calibration. It is performed by using a LUT generated according to above listed specifications. This LUT is then used to perform a DCC calibration on an artificial image generated according to the same specification except for the parameter of interest. In the error analysis, we considered four different view zenith angles (0°, 10°, 20°, 30°), six COT values (50, 60, 70, 80, 100, 120), and four different sun zenith angles (0°, 10°, 20°, 30°). The following sources of error were considered.

- Cloud effective radius: an error of 10 μm is considered in the error analysis.
- Microphysical properties model: Baum mixture of particles compared to pure crystal shape clouds such as plates,

TABLE III
ERROR BUDGET DCC CALIBRATION

	2 sigma error (%)	
Error source	BLUE	NIR
Cloud effective radius (10 μm)	0.761	1.363
Microphysical model	0.499	1.210
Cloud top height	0.172	0.024
Cloud geometrical depth	0.046	0.005
Ozone (5%)	0.100	0.115
Atmospheric profile	0.006	0.003
Red calibration error (3%)	*2.945*	*3.061*
Total interband	**0.932**	**1.826**
Total absolute	*3.089*	*3.564*

hexagonal column, and hexagonal crystals. Calibration is performed with the Baum mixture of particles which may differ from the real microphysical model. The degree to which the calibration coefficient is dependent on the model was estimated by carrying out the calibration process with a LUT calculated with the Baum mixture on an artificial image generated using the plates model.

- Cloud top height: the uncertainty in the cloud altitude leads to an uncertainty of the Rayleigh scattering contribution above the clouds. In the error analysis, a 3 km error in cloud top altitude is considered.
- Cloud geometrical depth: an uncertainty of 4 km in the cloud geometrical depth is used for the error analysis.
- Ozone: an error of 5% in the ozone concentration is considered in the sensitivity analysis [21].
- Atmospheric profile: uncertainties are assessed by using a mid-latitude summer profile instead of the tropical profile.

Table III gives both the relative (inter-band) error budget and the absolute error for the DCC method for both the blue and NIR band. The table contains the 2-sigma error. The error sources are assumed to be non-correlated, and thus the total error is the quadratic sum of all errors. For the absolute error, the uncertainty in the red band calibration is also considered. The method inaccuracy for inter-band calibration is mostly due to uncertainty in the cloud effective radius, the microphysical properties model, and ozone content. The other uncertain parameters have only a very small effect on the accuracy of the DCC calibration method. Inter-band calibration can be performed over DCC with an accuracy of 1% (2-sigma) for the blue band and better than 2% for the NIR band. When considering the absolute calibration uncertainty, the uncertainty in the absolute calibration of the reference red band is mainly affecting the achievable accuracy, adding around 3% uncertainty to both bands.

D. Cross-Sensor Desert

Method: Cross-sensor calibration is a calibration methodology which employs data from sensors from other satellite systems, so called "exogenous data." It is a valuable addition to methods acting on the data of the satellite itself. Also, it is of great interest to obtain calibrations that are in line with those of other systems, particularly with well-established systems. For PROBA-V, this is even more essential as the prime goal of the mission is to provide data continuity and consistency with VGT.

The most straightforward way to perform cross-sensor calibration would be to use almost simultaneous observations of the same location using different sensors. This is, however, difficult to realize and will become impossible after SPOT-VGT has ceased operations. The necessity of simultaneous observations can be overcome by the use of stable sites. Deserts are well suited for this because they are usually very stable over time, spatially homogenous, and are seldom covered by clouds. Following [28], a set of suitable desert sites for vicarious calibration have been selected in North Africa and Saudi Arabia from the 20 deserts identified by [29].

Observations at different times implies different conditions such as atmospheric variations, viewing and sun geometry, and slight instability of the site. Therefore, a methodology to map observations to a common standard is needed such as described in [10]. To minimize uncertainties due to atmospheric variation between observation and reference, the comparison is performed on bottom of atmosphere (BOA) reflectance, meaning that all data (from the reference and new sensor) are atmospherically corrected considering total column ozone, water vapor, atmospheric pressure, and AOT.

Differences caused by geometry are largely reduced by comparing only observations which are acquired under approximately identical or corresponding geometrical conditions. The geometries of the reference and new data are defined by their angles: SZA (solar zenith angle), VZA (viewing zenith angle), and RAA (relative azimuth angle). The TOA reflectance is a function of the angles $(\rho_{\mathrm{TOA}}^{k,\mathrm{sensor}}(SZA, VZA, RAA))$. The same result is obtained if SZA and VZA are interchanged, due to the reciprocity principle. The reflectance is symmetrical with respect to the principal plane, so a switch from $+RAA$ to $-RAA$ also leaves the result unchanged. The two invariances define the set of corresponding angles. In practice, geometry pairs will not match exactly. Therefore, a tolerance is allowed: pairs are considered approximately matching if the quadratic sum of the angle differences is smaller than a predefined tolerance (i.e., 100).

The spectral properties of the PROBA-V sensor are approximately equal to the those of SPOT-VGT. Still, even small spectral shifts could have a appreciable influence on the results. We investigated this and implemented a method for spectral adjustment, which was used on data from both VGT1 and VGT2 sensors, which have spectral differences of comparable magnitude. The results showed that the errors of the spectral mismatch were difficult to remediate in a reliable way, mainly because they were small compared to other error sources. Because of this relative unimportance, we omit a detailed discussion.

Implementation: The implementation is shown in Fig. 3. The scene data is extracted for the predefined desert ROI. The cloud masking is based on a threshold test, if one pixel in the ROI is cloudy, the whole ROI is discarded. We only proceed with desert sites that are completely cloud free. This is less complex and is justified as the desert sites are expected to be cloud free for a sufficient large portion of the time. The average TOA reflectance, solar, and viewing angles are then calculated per

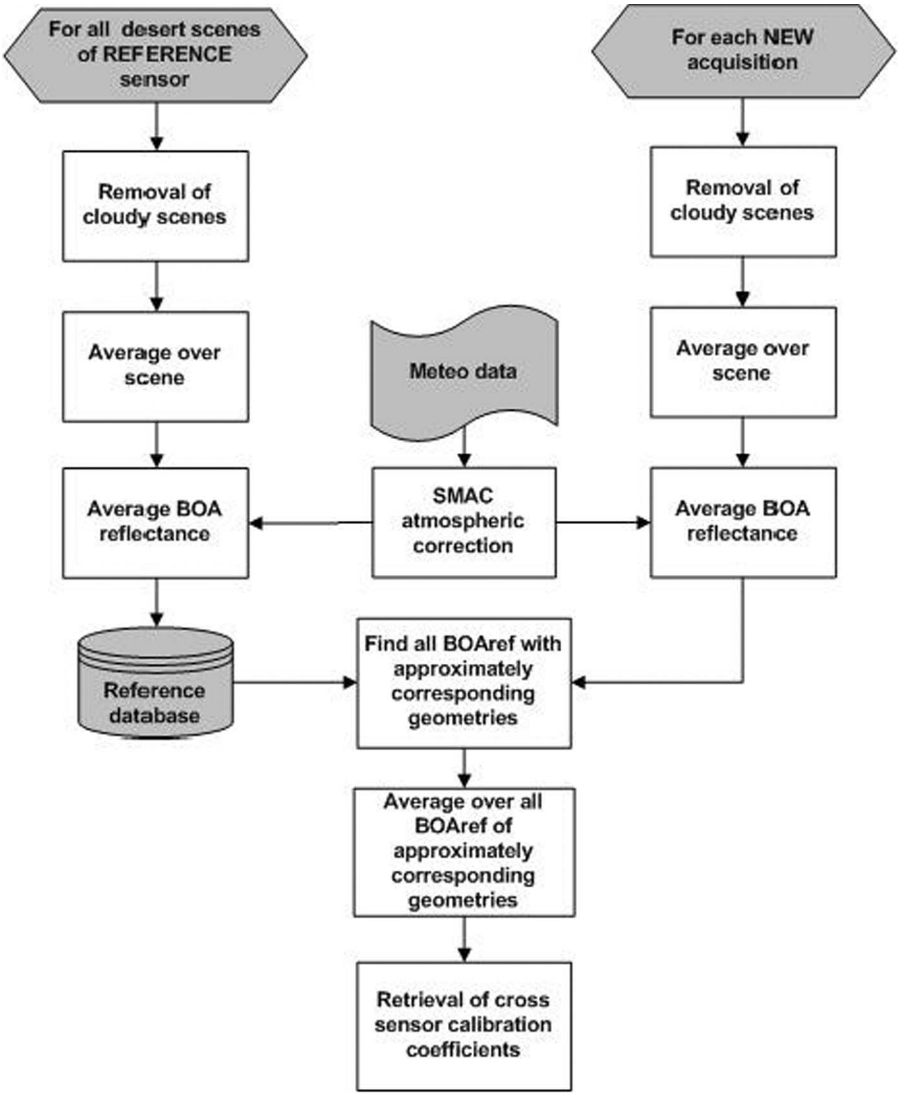


Fig. 3. Desert cross calibration workflow.

scene. The SMAC atmospheric correction scheme is applied converting TOA to BOA reflectance. This results in a single BOA reflectance value per band for an input desert scene.

For the reference sensor, the BOA reflectance of the scenes collected over a sufficient large time period is stored in the cross-sensor reference database, together with its illumination and viewing geometry and the atmospheric conditions. Observations from the new sensor processed to BOA reflectance ($\bar{\rho}_{\mathrm{BOA}}^{\mathrm{sensor},k}$) are compared against the average reference sensor BOA reflectance ($\bar{\rho}_{\mathrm{BOA}}^{ref,k}$) obtained under approximately identical or corresponding viewing and illumination geometries. Finally, an estimate of the cross-sensor calibration coefficient is obtained as

$$A_{\mathrm{Cross}}^{k} = \frac{A^{ref,k}}{A^{\mathrm{sensor},k}} = \frac{\bar{\rho}_{\mathrm{BOA}}^{\mathrm{sensor},k}}{\bar{\rho}_{\mathrm{BOA}}^{ref,k}}. \tag{5}$$

Uncertainty: The determination of the cross-sensor calibration coefficients for individual sites is affected by a number of variables, each of which is prone to random and/or systematic errors. The contributions to the errors come mainly from the following sources, which we will describe briefly.

- Properties of the target: The BRDF surface properties of the desert sites are assumed to be constant over time. Even when maximally stable targets are selected, this assumption is only approximately true (e.g., BRDF properties of sand dunes may be influenced by wind direction changing the orientation of the dunes).
- Aerosol optical thickness: Aerosol optical measurements in these deserts sites are rare, and therefore assumptions have been made on the used AOT which may deviate from the actual aerosol optical thickness. It is expected that after averaging a lot of measurements over sites and over time, that the aerosol affect will be more a random error (noise) than a fixed bias/systematic error.
- Uncertainties in ECMWF meteo information: According to [21], about 5–10 mbars is a reasonable accuracy for the atmospheric pressure. An accuracy of about 10–20 mAtm-cm is a typical accuracy for the total ozone amount of the atmosphere. An uncertainty of 20% is often used for the water vapor content.

TABLE IV
OVERVIEW OF SPOT-VGT DATA USED IN EXPERIMENT

Calibration method	Zone	Period	Sensor
Rayleigh	North Atlantic Ocean	Oct-09	SPOT-VGT1
Rayleigh	South Indien Ocean	Oct-09	SPOT-VGT1
Rayleigh	North Atlantic Ocean	Oct-09	SPOT-VGT2
Rayleigh	South Indien Ocean	Sept-09 - Oct-09	SPOT-VGT2
DCC	Maldives	June-09-July-09	SPOT-VGT2
Deserts	18 desert sites North Africa and Saudi Arabia	Jan-06-Dec-09	SPOT-VGT2 (ref)
		Jan-09-Dec-09	SPOT-VGT1

- The differences of the spectral responses for the compared similar bands or if corrected for, the uncertainty on the spectral adjustment factor
- The inaccuracy of the geolocation of the viewed pixels for each sensors

For the shorter wavelengths (particularly blue), variation in aerosol properties will have the largest contribution to the total uncertainty. For the longer wavelengths, the uncertainty by variation in target reflectance will have the largest impact.

IV. DATA

As the PROBA-V satellite is still to be launched, no real PROBA-V image data is available at present. In order to validate the methods and their implementation, satellite data from spectrally similar sensors is used as the algorithm cores can be considered as "generic" for spectrally similar sensors. The PROBA-V imaging instrument is designed to offer data continuity with the images from the VGT instruments on board the SPOT4 and SPOT5 satellites, respectively, referred here as VGT1 and VGT2. Therefore, SPOT-VGT images are a very good choice to validate the methods. Table IV gives an overview of the VGT1 or VGT2 data used for algorithmic validation.

As SPOT-VGT only acquires acquisitions over oceans for calibration purposes during dedicated periods in the year over certain region of interest, the archive of data over oceans is rather limited. For SPOT-VGT1, no data exists for 2009 over suitable regions for DCC calibration.

The cross-sensor calibration is tested and validated on SPOT-VGT1 and SPOT-VGT2 data where SPOT-VGT2 is considered as the reference sensor. Atmospherically corrected SPOT-VGT2 data over the 18 deserts sites are stored in the reference database. Atmospherically corrected SPOT-VGT1 data are then compared to comparable data of the same desert site in the SPOT-VGT2 reference database following the methodology described in the previous section.

Atmospheric pressure, wind speed at the sea level, total column ozone, and total column water vapor are accessed through the ECMWF ERA Interim database.

V. IMAGE-BASED UNCERTAINTY ANALYSIS

Rayleigh and DCC: In the previous section, the accuracy achievable with the Rayleigh and DDC methods was analyzed by a sensitivity analysis in which the effect on the calibration result of variations in every uncertain input parameter was determined. An alternative approach is to use the retrieved results to make up an image-based uncertainty budget.

For the Rayleigh method where results from two sites are obtained, the relative expanded uncertainty $u(\Delta\bar{A}^k)$ (in percentage at 2-sigma level) for the average $\Delta\bar{A}^k$ is given by

$$u(\Delta\bar{A}^k)(\%) = \frac{1.96 s(\Delta\bar{A}^k)100}{\Delta\bar{A}^k} \tag{6}$$

with

$$\Delta\bar{A}^k = \left[w_{\text{site1}}\Delta\bar{A}^k_{w,\text{site1}} + w_{\text{site2}}\Delta\bar{A}^k_{w,\text{site2}}\right]$$
$$s(\Delta\bar{A}^k) = \sqrt{\left(\Delta\bar{A}^k_{w,\text{site1}} + \Delta\bar{A}\right)^2 + \left(\Delta\bar{A}^k_{w,\text{site2}} + \Delta\bar{A}\right)^2}$$

where $\Delta\bar{A}^k_{w,\text{site1}}$ and $\Delta\bar{A}^k_{w,\text{site2}}$ are the weighted averages over all image results obtained for, respectively, site 1 and site 2. The average image results are weighted according to the number of used pixels in the image.

In the Rayleigh calibration approach, the NIR band is used as a reference band to determine AOT. An uncertainty in the absolute calibration of NIR band will therefore affect systematically the retrieved results for the blue and red bands. This source of uncertainty cannot be reduced by spatial or temporal averaging. Therefore, u^k_{calNIR}, the uncertainty on the retrieved calibration results caused by the uncertainty in the NIR calibration coefficient, needs to be added to the final uncertainty

$$u'(\Delta\bar{A}^k) = \sqrt{\left(u(\Delta\bar{A}^k)\right)^2 + \left(u^k_{\text{calNIR}}\right)^2} \tag{7}$$

with $k = $ blue or red band and where u^k_{calNIR} is listed in Table II with $u^{\text{blue}}_{\text{calNIR}}$ is 0.72% and $u^{\text{red}}_{\text{calNIR}}$ is 2.06%.

For the DCC method, only acquisitions over 1 site are available; the relative expanded uncertainty for the average calibration result obtained with the DCC is given by

$$u'(\Delta\bar{A}^k)(\%) = 1.96 s(\Delta\bar{A}^k)100/\Delta\bar{A}^k_{w,\text{site1}}\sqrt{N-1} \tag{8}$$

with

$$s(\Delta\bar{A}^k) = \sqrt{(1/(N-1))\sum_{p=1}^{N}\left(\Delta\bar{A}^k_{w,\text{site1}} - \Delta\bar{A}^k_{p,site1}\right)^2}$$

and N the number of used images.

The uncertainty u^k_{calRED} caused by the uncertainty in the red calibration coefficient needs to be added to the final uncertainty as the red band is this time used as the reference band to retrieve COT

$$u'(\Delta\bar{A}^k) = \sqrt{\left(u(\Delta\bar{A}^k)\right)^2 + \left(u^k_{\text{calRED}}\right)^2} \tag{9}$$

with $k = $ blue or NIR band and where u^k_{calRED} is listed in Table III with $u^{\text{blue}}_{\text{calNIR}}$ is 2.9% and $u^{\text{blue}}_{\text{calRED}}$ is 3%.

Deserts: For desert calibration, significant seasonal variations have repeatedly been reported in the literature, e.g., in [24]. As the cause of the variations is not fully understood, we expect to find some seasonal variations as result of our experiments as well. We estimate the significance of the variations as follows. Given a sufficient number of results over a substantial time period (e.g., daily over one year), we assume that the observed result consists of three terms: $\Delta\bar{A}^k(t) = T + S(t) + R(t)$. T is the true value in absence of variations. The other terms average to zero over a whole year and describe temporal variations with respect to T:

- $S(t)$: a seasonal term: including all effects that vary very slowly over the year, in a yearly repeat cycle.
- $R(t)$: a random term: including all other effects.

We estimate the errors induced by the both terms, using the notation N for the number of days in a month, and M for the number of months in a year. The total variance is

$$s_{\text{total}^2}\left(\Delta\bar{A}(t)\right) = \frac{1}{(NM-1)} \sum_{t=1}^{NM} \left(S(t) + R(t)\right)^2 \qquad (10)$$

Assuming that the seasonal term is constant within 1 month, the monthly average is

$$\begin{aligned}\Delta\bar{A}^k_{\text{month}}(m) &= \frac{1}{N}\sum_{t=1}^{N} \Delta\bar{A}^k(t,m)\\ &= T + S(m) + \frac{1}{N}\sum_{t=1}^{N} R(t,m)'. \qquad (11)\end{aligned}$$

The averaged intra-month variance is

$$\begin{aligned}s^2_{\text{intra}}(\Delta\bar{A}^k) &= \frac{1}{(M-1)(N-1)} \sum_{t=1}^{N}\sum_{m=1}^{M}\\ &\quad \times \left(\Delta\bar{A}^k - \Delta\bar{A}^k_{\text{month}}(m)\right)^2. \qquad (12)\end{aligned}$$

The seasonal variation is

$$s^2_{\text{seasonal}}(\Delta\bar{A}^k) = \frac{1}{(M-1)} \sum_{m=1}^{M} \left(S(m) - T\right)^2. \qquad (13)$$

It can be estimated from the total variation observed between months

$$\begin{aligned}s^2_{inter}(\Delta\bar{A}^k) &= \frac{1}{(M-1)} \sum_{m=1}^{M} \left(\Delta\bar{A}^k_{\text{month}}(m) - T\right)^2\\ &\approx s^2_{\text{seasonal}}(\Delta\bar{A}^k) + \frac{s^2_{\text{intra}}(\Delta\bar{A}^k)}{N}. \qquad (14)\end{aligned}$$

The last term is the remaining effect from random variations after averaging over a month. Its contribution to the variance is reduced because N samples are averaged within the month.

The estimates are useful because they show whether the seasonal effect is significant. If so, the best value estimate for a particular month is the monthly average $(\Delta\bar{A}^k_{\text{month}}(m))$ and its accuracy is: $\sqrt{s^2_{\text{intra}}(\Delta\bar{A}^k/N)}$. Otherwise, the best value estimate is the yearly average, with its accuracy:

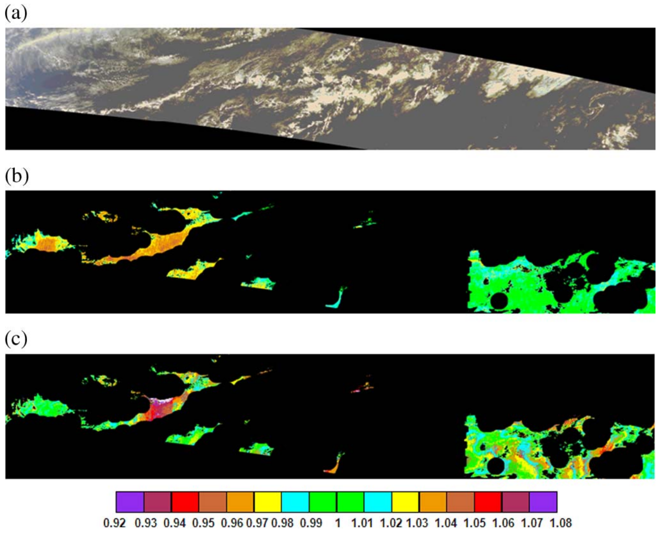


Fig. 4. Rayleigh calibration results per pixel of the input image (a) for, respectively, the BLUE band (b) and the RED band (c).

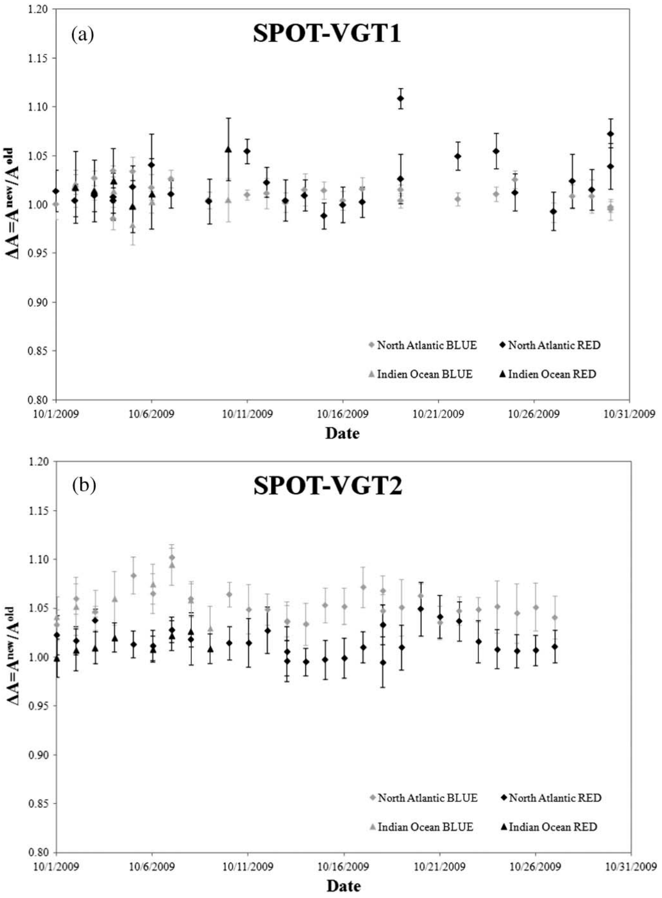


Fig. 5. Average Rayleigh calibration result per image for, respectively, (a) SPOT-VGT1 and (b) SPOT-VGT2.

TABLE V
RAYLEIGH CALIBRATION RESULTS

Site	VGT1 BLUE	VGT1 RED	VGT2 BLUE	VGT2 RED
Atlan North	1.010	1.012	1.056	1.010
Indian South	1.005	1.015	1.051	1.015

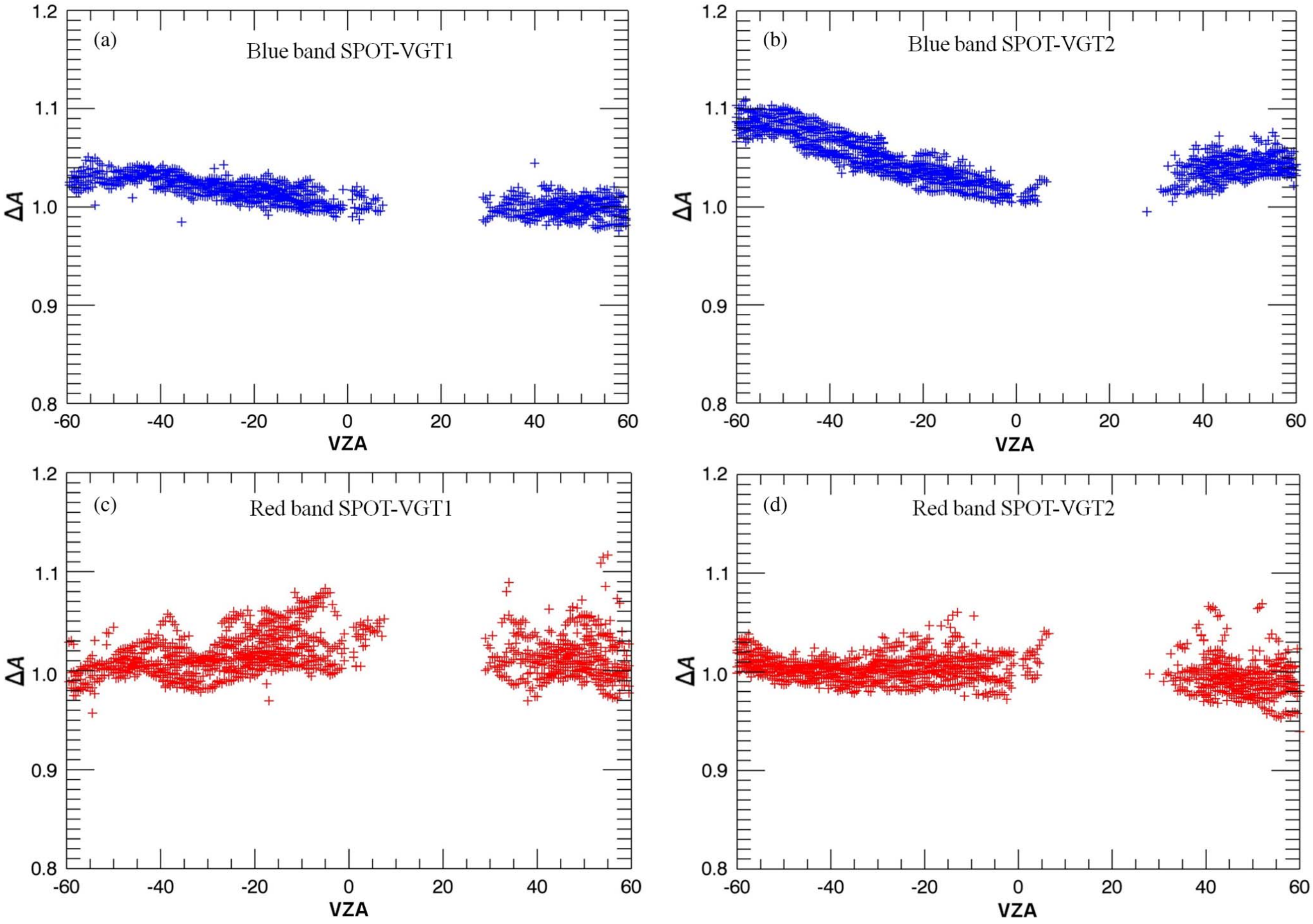


Fig. 6. Rayleigh calibration results for the North Atlantic Ocean site in function of view zenith angle for the blue and red bands of SPOT-VGT1 (a) and (d) and SPOT-VGT2 (b) and (d).

$\sqrt{s^2_{\mathrm{seasonal}}(\Delta\bar{A}^k) + (s^2_{\mathrm{intra}}(\Delta\bar{A}^k)/NM)}$. Comparing both accuracies reveals if the seasonal effect is significant.

VI. RESULTS AND DISCUSSION

A. Rayleigh

Fig. 4 illustrates the application of the Rayleigh calibration workflow to an individual image. The extension of the cloud mask in the preprocessing step is clearly visible as circular masked regions near small spots of clouds. The sun glint probability masking step masks pixels in the middle right of the image. For each non-masked pixel, the reported value gives the change of calibration coefficient with respect to the official calibration. Assuming that 1) the calibration files (absolute and equalization) are still valid for the day of the acquisition, 2) the uncertainty in the absolute and equalization coefficients is 0%, and 3) there is no uncertainty in the Rayleigh calibration itself, the change in calibration coefficients, the values (i.e., ΔA^{blue} and A^{red}) should be equal to 1 for each pixel. For this image, the observed changes ΔA^k are average less than 2%, with some larger deviation near the edges of the masked regions and close to the nadir viewing direction. However, on the basis of one single image, no well-founded conclusions can be made.

Fig. 5 plots the average results obtained per image over the two sites for both SPOT-VGT1 and SPOT-VGT2. The average and its standard deviation are given after outlier removal which removes all pixel values outside the 3-sigma confidence interval from the median. The overall weighted average per site, band, and sensor is given in Table V. To calculate the weighted average, the results of the different dates are weighted according to the number of non-masked pixels in the image. For the red band, this weighted average gives a difference about 1.0 to 1.5% for both sensors and sites. For the blue band, a significant difference between sensors is found for VGT2 with an increase of 5.1 to 5.6% in calibration coefficient detected depending on the site, while for VGT1, an increase of of 0.5 to 1% is detected. Meaning, for the blue band, a difference of 4.5% is observed between VGT 1 and VGT2. If we plot the ΔA^k values in function of view zenith angle (Fig. 6), an explanation of this difference can be given. For VGT1, plots in function of VZA are relatively flat for the blue band. For VGT2, a clear VZA dependency is detected with strongly increasing values near the edges explaining the higher average values for VGT2. Clearly, observed differences in blue would be less if included observations would be limited to VZA $< 30°$. This smile pattern is probably due some problems with the multi-angular calibration of SPOT-VGT2. For both VGT1 and VGT2 the scattering of results seems to be higher for the red band

TABLE VI
ACCURACY ESTIMATES OF THE RAYLEIGH CALIBRATION RESULTS

parameter	VGT1 BLUE	VGT1 RED	VGT2 BLUE	VGT2 RED
ΔA	1.007	1.015	1.054	1.012
$u(\Delta A)$ (%)	0.621	0.315	0.581	0.624
$u'(\Delta A)$ (%)	0.899	1.968	0.872	2.040

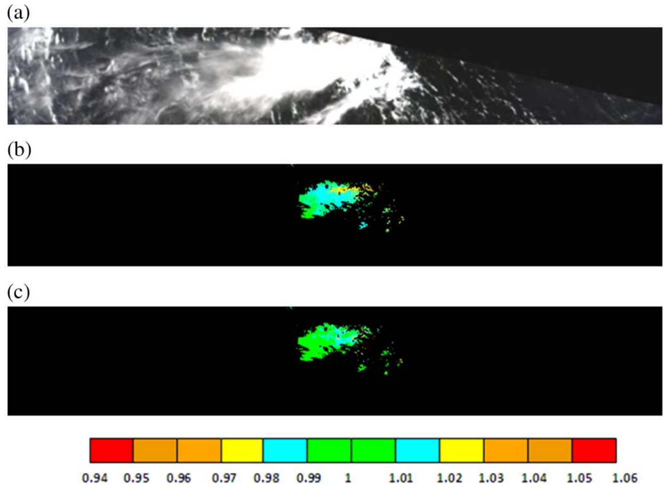


Fig. 7. Deep convective clouds calibration results per pixel of the input image (a) for, respectively, the BLUE band (b) and the NIR band (c).

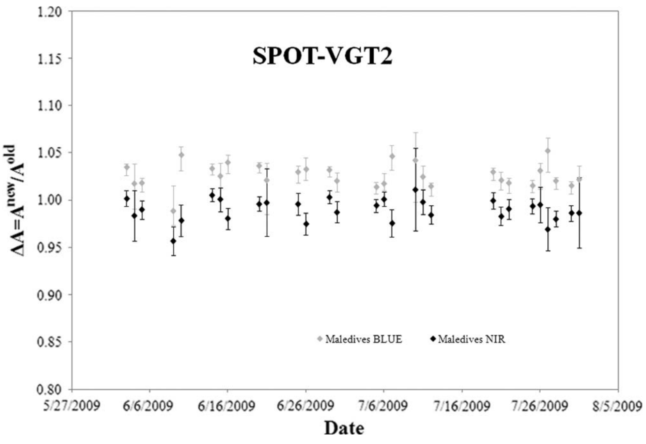


Fig. 8. Average SPOT-VGT2 DCC calibration result per image.

than for the blue band, but without major variation over the field-of-view.

Assuming that both VGT1 and VGT2 are stable for this period, the image-based uncertainty can be defined according to (7). We arrive at the results listed in Table VI. The uncertainties are well within the accuracy estimates given by the sensitivity analysis and for the red band even a factor two lower. This maybe due to too pessimistic assumptions about some of the uncertain parameters when making up the error budget.

B. DCC

The result of the DCC workflow applied to an individual image is shown in Fig. 7. Only the center part of large, bright homogenous clouds are considered. Small scattered clouds or edges of clouds are removed in the DCC specific preprocessing steps.

TABLE VII
ACCURACY ESTIMATES OF THE DEEP CONVECTIVE CLOUDS CALIBRATION RESULTS

	VGT2 BLUE	VGT2 NIR
ΔA	1.027	0.989
$u(\Delta A)$ (%)	0.466	0.448
$u'(\Delta A)$ (%)	3.841	3.871

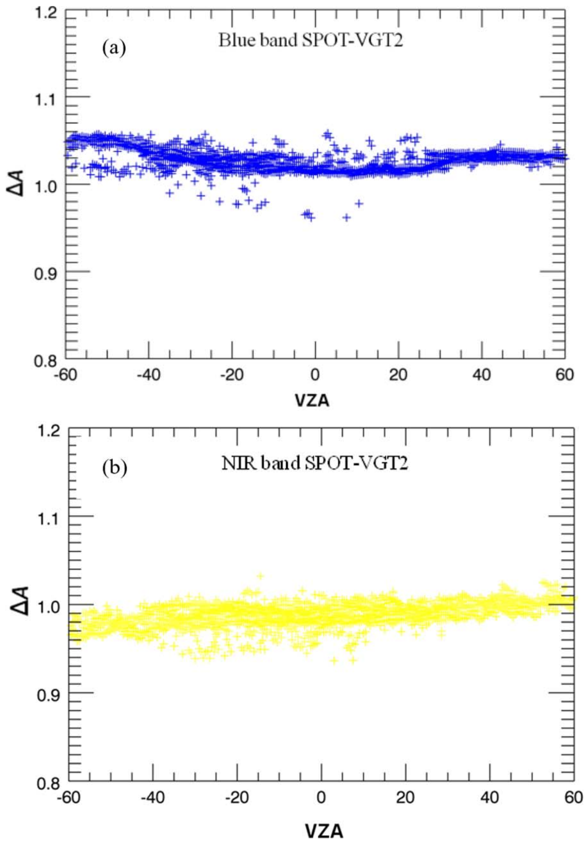


Fig. 9. Deep convective clouds calibration in function of view zenith angle for SPOT-VGT2 for, respectively, the BLUE band (a) and the NIR band (b).

Fig. 8 plots the average results obtained for each image and its standard deviation. Compared to the Rayleigh calibration results, the DCC results show relatively small variation within one scene and between different scenes. The overall weighted average over the different scenes is given in Table VII. For the blue band, an increase of around 2.7% in absolute calibration coefficient is detected, while for the NIR band, the DCC result indicates a decrease in the calibration coefficient. Similar to the Rayleigh results, this increase is mainly due to the higher values observed for larger VZA (Fig. 9). This again may indicate an issue with the VGT-2 sensor in the blue for larger VZAs. For the NIR band, no significant variation over the field-of-view is observed. Using (9), the uncertainty in the retrieved results is calculated and given in Table VII. Due to the very small inter-image variations in the retrieved estimates, the inter-band uncertainty (i.e., neglecting red band calibration uncertainty) is very small which gives high confidence in the

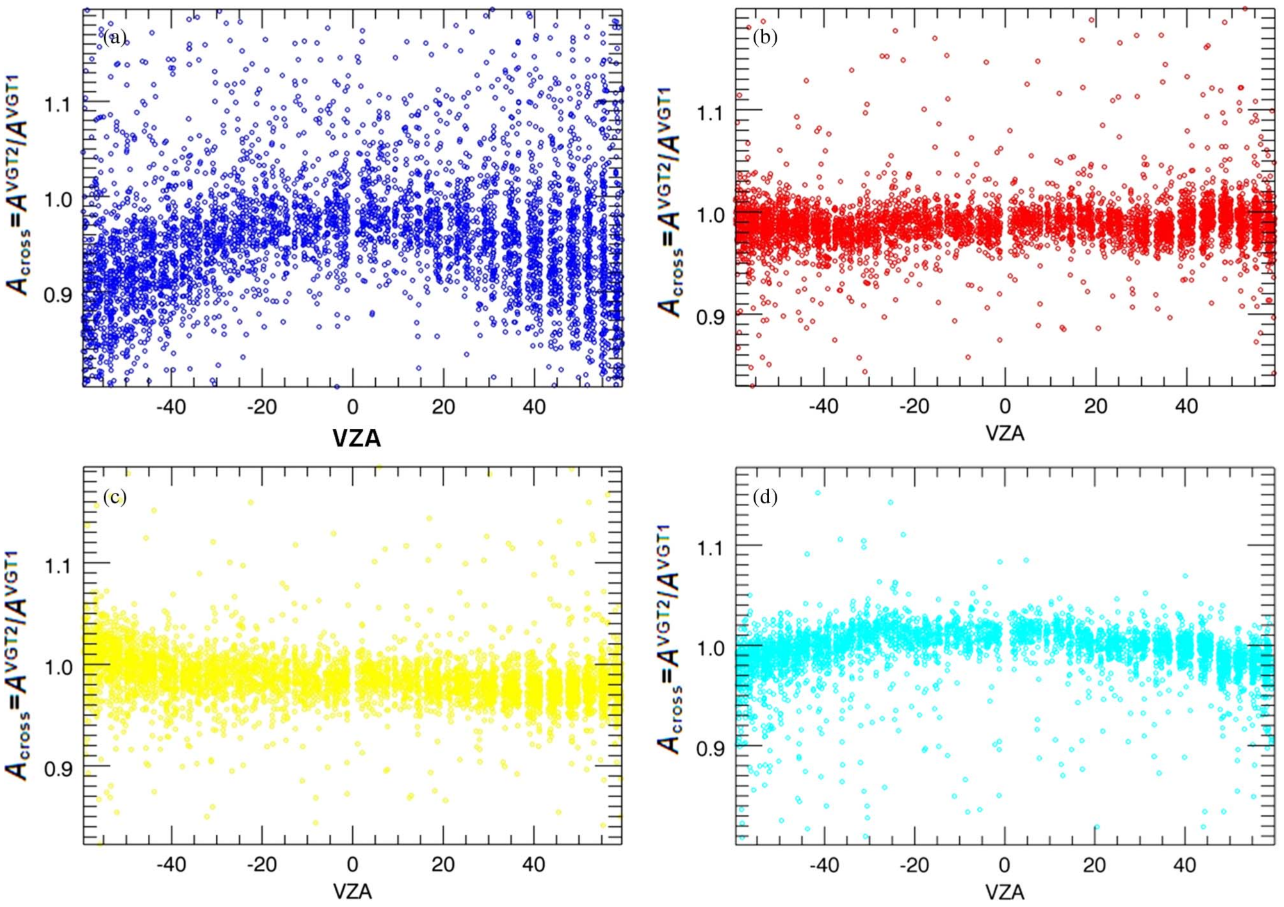


Fig. 10. Desert cross-sensor calibration results in function of view zenith angle for, respectively, (a) blue, (b) red (b), (c) NIR, and (d) SWIR bands.

retrieved inter-band results. The absolute calibration uncertainty for the DCC results is therefore almost fully dominated by the assumed calibration uncertainty of 3% for the reference red band.

C. Cross-Sensor Desert

Fig. 10 plots the cross-sensor calibration results as given by (5) obtained for all the deserts for almost one year of data in function of the VZA. The strong VZA-dependent pattern for the blue band is in line with the Rayleigh and DCC (for VGT2) results, where, on the one hand, the blue TOA reflectance values of VGT2 are in average higher than the modeled values, with the difference increasing with VZA and, on the other hand, the TOA reflectance values for VGT1 are lower than modeled with no significant VZA dependency. Therefore, in the cross-sensor calibration plots in function of VZA, these differences are even more pronounced. In addition to this strong VZA dependency, the scattering of the results is much more severe for the blue band than for the other three bands. This is probably due to instability of the aerosol optical properties between the observations which has the largest influence on the blue band. Similar to, respectively, the Rayleigh and DCC results, no significant VZA dependency is observed for red and NIR bands. A slight VZA dependency is observed for the SWIR band with VGT2 having slightly higher TOA reflectance values compared to VGT1 at large VZAs. As both Rayleigh and DCC methods are not applicable for the SWIR band, no conclusions can be made on which of the sensors is causing this effect in the SWIR.

In Fig. 11, the cross-sensor calibration results for the different desert sites are given in function of the month of the year, neglecting observations performed at VZAs larger than 30°. A small seasonal variation is visible, which seems to be consistent for red, NIR, and SWIR bands. For the blue band, the scattering of the results is a bit masking this seasonal effect. The seasonal variation has been reported by others [28]. Possible causes are instrumental behavior related to sun-earth distance, surface anisotropy variation, and aerosol optical thickness variation over the seasons. During winter months (i.e., November till February), aerosol optical values from AERONET sites are significantly lower than in summer.

The significance of the seasonal effect is quantified using the statistical considerations made in Section III-D. Results are shown in Table VIII. For all bands, the seasonal variations s_{seasonal} over the year are smaller than the random variations s_{inter}. The "Accuracy Year" shows the accuracy estimate if the seasonal variation would be ignored, while "Accuracy month" gives an estimate for the case when averaging is only performed within one month. The first clearly yields larger errors which indicates that the seasonal effect is clearly significant. For the blue band, with its larger random variations, the significance is less clear.

Omitting the observations performed with VZAs larger than 30°, observed differences averaged over a one year period

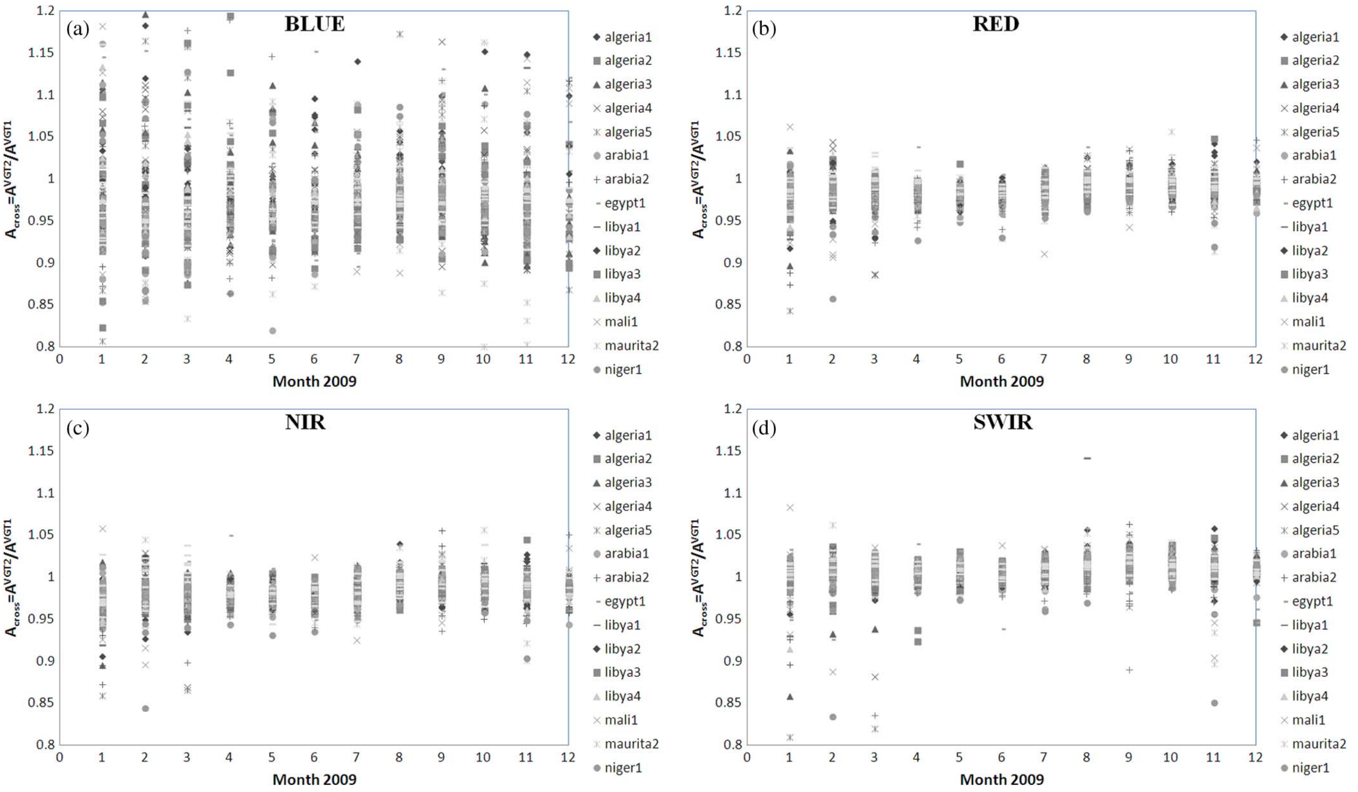


Fig. 11. Cross-sensor desert results in function of month of the year.

TABLE VIII
INTER-BAND $A^{\mathrm{VGT2}}/A^{\mathrm{VGT1}}$ DESERT CALIBRATION: VARIANCE AND ACCURACY ESTIMATES

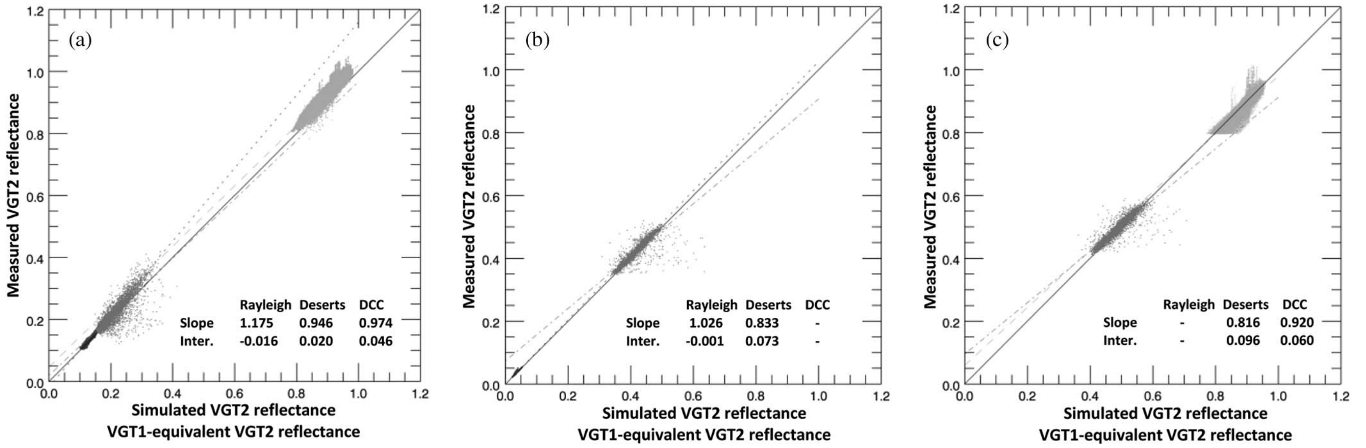
	BLUE	RED	NIR	SWIR
A^{VGT2}/A^{VGT1} *(yearly average)*	0.98	0.987	0.983	1.005
s_{total} *(%)*	3.62	1.27	1.41	1.48
s_{intra} *(%)*	3.43	1.09	1.18	1.25
s_{inter} *(%)*	0.88	0.61	0.73	0.64
$s_{seasonal}$ *(%)*	0.62	0.57	0.70	0.60
Accuracy month (%)	0.63	0.20	0.22	0.23
Accuracy year (%)	0.65	0.58	0.70	0.60

between VGT1 and VGT2 are 2% for blue, less than 1.3% and 1.7% for, respectively, red and NIR bands (Table VIII). For SWIR bands, an insignificant difference of 0.5% in the absolute calibration is detected.

D. Comparison of Methods in Scatterplots

In Fig. 12, results of the three different methods are compared for VGT2 in a scatterplot. Due to the lack of DCC data for VGT1, no similar plot is given for VGT1. For the Rayleigh method, the measured reflectance $\rho_{\mathrm{TOA}}^{c,k,\mathrm{VGT2}}$ is plotted against the modeled (simulated) reflectance $\rho_{\mathrm{TOA}}^{c,k,\mathrm{model}}$ for the blue and red band for all selected pixels of both the North Atlantic and South Indian oceanic site. Their associated regression lines are plotted as dark gray dotted lines. For the DCC method, the measured reflectance $\rho_{\mathrm{TOA}}^{c,k,\mathrm{VGT2}}$ is plotted against the modeled (simulated) reflectance $\rho_{\mathrm{TOA}}^{c,k,\mathrm{model}}$ for the blue and NIR band for all selected pixels. The associated regression lines are given in light gray dashed lines. Due to the reflectance threshold test, all plotted DCC results have a $\rho_{\mathrm{TOA}}^{c,\mathrm{NIR},\mathrm{VGT2}}$ larger than 0.8. For the DCC method, some larger spreading of the results is visible near this threshold value [Fig. 12(c)]. In Figs. 10, 11 and Table VIII, the cross-sensor desert results A_{Cross}^{k} are given following (4) with VGT2 as the reference. To plot in Fig. 12, similar to the DCC and Rayleigh method, the measured VGT2 reflectance against a reference the VGT 1 equivalent reflectance is used which is calculated as $\bar{\rho}\mathrm{VGT1}, k/A_{\mathrm{Cross}}^{k}$. The associated regression lines are given in medium gray dash dotted lines. For all the regression lines, slopes and intercepts are given in the plots.

Fig. 12 clearly shows the difference in reflectance range covered by the three methods with the Rayleigh method covering the lower range (< 0.2), the desert method, the mid-range (approximately from 0.2 in the blue to 0.6 in the NIR), and the DCC method, the higher range (> 0.8). For all methods and bands, the scattered patterns are overlapping with the $1:1$ line. For the blue bands, slopes and intercepts are relatively similar for desert and DCC methods, with slopes of, respectively, 0.946 and 0.974 and intercepts of, respectively, 0.02 and 0.046. For the Rayleigh method, an increased slope is observed (1.175) which is a bit counterbalanced by the negative intercept (-0.016). Due to very small reflectance range covered by the Rayleigh method the uncertainty in the Rayleigh slope calculations is relatively high making extrapolations to higher reflectance ranges using the Rayleigh calculated intercept and slope proud to inaccuracies. On the other hand, the uncertainty on the predicted intercepts by desert and particularly the DCC methods are high due to the higher reflectances used in the methods. Remarkably is that for the red band



(a)

	Rayleigh	Deserts	DCC
Slope	1.175	0.946	0.974
Inter.	-0.016	0.020	0.046

(b)

	Rayleigh	Deserts	DCC
Slope	1.026	0.833	-
Inter.	-0.001	0.073	-

(c)

	Rayleigh	Deserts	DCC
Slope	-	0.816	0.920
Inter.	-	0.096	0.060

Fig. 12. Comparison of the cross-sensor desert method (medium gray) against the DCC method (light gray) and Rayleigh method (dark gray) for VGT2 (a) blue, (b) red, and (c) NIR band. For the cross-sensor desert method, the measured VGT2 reflectance is plotted against the VGT1 reflectance for observations acquired under approximately identical or corresponding geometry. Regression lines for cross-sensor desert are displayed as medium gray dash dotted lines, for the DCC method as light gray dashed lines and for Rayleigh method as dark gray dotted lines. Solid black line is the 1 : 1 line presenting perfect matches.

the Rayleigh regression line matches quite closely with the 1 : 1 line.

VII. CONCLUSION

The in-flight radiometric calibration of the PROBA-V satellite mission is implemented using a set of independent vicarious methods. In this paper, three key methods were presented in detail, and an error analysis was conducted to estimate the accuracy of their results. The methods and their implementation were tested by using data from the SPOT-VGT1 and 2 missions.

A sensor artifact for the blue band of SPOT-VGT2 has been detected by all three methods. The evidence of this artifact is made clear by combining different vicarious calibration methods which cover different parts of the dynamic range of the sensor (dark oceans versus bright clouds), and which are using different references (i.e., 6SV and LibRadtran RTF calculation or observation by other sensors), therefore minimizing biases or systematic errors which may occur if only one method is used.

Both the theoretical error budgets and the image-based uncertainty analysis on the basis of obtained SPOT-VGT calibration results showed that the mission requirement specifications for the radiometric calibration, i.e., 5% absolute and 3% relative (inter-band, cross sensor) are realistic achievable goals for the PROBA-V missions although no on-board calibration devices are available.

With the methods described in this paper, no absolute calibration for the SWIR band is feasible. Therefore, other approaches like moon calibration [30], desert calibration on the basis of simulated references [28] and reflectance based methods are being integrated in the vicarious calibration chain for the PROBA-V mission.

ACKNOWLEDGMENT

ECMWF data used in this study have been obtained from the ECMWF data server. The authors wish to thank Patrice Henry from CNES, Riccardo Duca and Joe Zender from ESA for fruitful discussions, Svetlana Kotchenova (BIRA) and Alessandro IPE (RMI) for their advice on respectively 6SV and LibRadtran radiative transfer calculations. We thank the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

[1] K. Mellab, “Small satellite technology to monitor the global earth: The PROBA-V mission,” in *Proc. Int. Symp. Remote Sens. Environ.*, Stresa, Italy, 2009.

[2] J. Dries, “The SPOT VEGETATION and PROBA-V user segments,” in *Proc. Int. Symp. Remote Sens. Environ.*, Stresa, Italy, 2010.

[3] S. Adriaensen, I. Benhadj, G. Duhoux, W. Dierckx, J. Dries, W. Heyns, R. Kleihorst, S. Livens, K. Nackaerts, I. Reusen, S. Sterckx, T. Van Achteren, and J. Everaerts, “Building a calibration and validation system for the PROBA-V satellite mission,” in *Proc. EuroCOW Int. Calib. Orient. Workshop*, Castelldefels, Spain, Feb. 10, 2010.

[4] S. Livens, S. Sterckx, W. Dierckx, S. Adriaensen, and I. Reusen, “Vicarious radiometric calibration based on combined accuracy estimation,” in *Proc. CALCON Tech. Conf.*, Logan, UT, USA, 2010, vol. 30.

[5] S. Livens, S. Sterckx, W. Dierckx, S. Adriaensen, and I. Reusen, “Multiple vicarious calibration using combined accuracy estimation,” in *Proc. SPIE Remote Sens.*, Berlin, Germany, 2010, pp. 782623-1–782623-8.

[6] C. Cao, H. Xu, H. Xu, J. Sullivan, L. McMillin, P. Ciren, and Y.-T. Hou, “Intersatellite radiance biases for the high-resolution infrared radiation sounders (HIRS) on board NOAA-15, -16, -17 from simultaneous nadir observations,” *J. Atmos. Ocean. Technol.*, vol. 22, no. 4, pp. 381–395, Apr. 2005.

[7] D. Six, M. Fily, S. Alvain, P. Henry, and J.-P. Benoist, “Surface characterization of the dome concordia area (Antarctica) as a potential satellite calibration site, using spot 4/vegetation instrument,” *Remote Sens. Environ.*, vol. 89, no. 1, pp. 83–94, Jan. 2004.

[8] E. Vermote, R. Santer, P. Y. Deschamps, and M. Herman, “In-flight calibration of large field of view sensors at short wavelengths using Rayleigh scattering,” *Int. J. Remote Sens.*, vol. 13, no. 18, pp. 3409–3429, Dec. 1992.

[9] O. Hagolle and P. Goloub, “Results of POLDER in-flight calibration,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1550–1566, May 1999.

[10] B. Fougnie, G. Bracco, B. Lafrance, C. Ruffel, O. Hagolle, and C. Tinel, “PARASOL in-flight calibration and performance,” *Appl. Opt.*, vol. 46, no. 22, pp. 5435–5451, Aug. 2007.

[11] N. Martiny, R. Santer, and I. Smolskaia, “Vicarious calibration of MERIS over dark waters in the near infrared,” *Remote Sens. Environ.*, vol. 94, no. 4, pp. 475–490, Feb. 2005.

[12] H. R. Gordon, O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, “A semianalytic radiance model of ocean color,” *J. Geophys. Res.*, vol. 93, no. D9, pp. 10 909–10 924, Jan. 1988.

[13] B. Fougnie, P. Henry, A. Morel, and F. Montagner, “Identification and characterization of stable homogeneous oceanic zones: Climatology and

impact on in-flight calibration of space sensor over Rayleigh scattering," in *Proc. 16th Ocean Opt.*, Santa Fe, NM, 2002.

[14] B. Fougnie, J. Llido, L. Gross-Colzy, H. Patrice, and D. Blumstein, "Climatology of oceanic zones suitable for in-flight calibration of space sensors," in *Proc. SPIE*, 2010, vol. 7807, pp. 78070S-1–78070S-11.

[15] A. Morel, H. Claustre, and B. Gentili, "The most oligotrophic subtropical zones of the global ocean: Similarities and differences in terms of chlorophyll and yellow substance," *Biogeosciences*, vol. 7, no. 10, pp. 3139–3151, Oct. 2010.

[16] E. Shettle and R. W. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties," Air Force Geophys. Lab., Hanscom AFB, MA, Environmental Res. Paper AFGL-TR-79-0214, 1979.

[17] S. Kotchenova and E. Vermote, "Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part II. Homogeneous Lambertian and anisotropic surfaces," *Appl. Opt.*, vol. 46, no. 20, pp. 4455–4464, Jul. 2007.

[18] A. Morel, "Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters)," *J. Geophys. Res.*, vol. 93, no. C9, pp. 10 749–10 768, Jan. 1988.

[19] D. Tanré, C. Deroo, P. Duhaut, M. Herman, and J. J. Morcrette, "Description of a computer code to simulate the satellite signal in the solar spectrum: 5S code," *Int. J. Remote Sens.*, vol. 11, no. 4, pp. 659–668, Apr. 1990.

[20] H. Rahman and G. Dedieu, "SMAC: A simplified method for the atmospheric correction of satellite measurements in the solar spectrum," *Int. J. Remote Sens.*, vol. 15, no. 1, pp. 123–143, Jan. 1994.

[21] D. Antoine and A. Morrel, "MERIS ATBD 2.7: Atmospheric correction of the MERIS observations over ocean case 1 waters," Labo. Oceanograph. Villefranche, Observ. Oceanol., Villefranche-sur-Mer, France, 2005.

[22] M. D. King, S.-C. Tsay, S. E. Platnick, M. Wang, and K.-N. Liou, "Cloud retrieval algorithms for MODIS: Optical thickness,effective particle radius, thermodynamic phase," NASA Earth Observ. Syst. Proj. Sci. Off., NASA Goddard Space Flight Center, Greenbelt, MD, 1997, Algorithm Theoretical Basis Document ATBDMOD-05.

[23] B. Lafrance, O. Hagolle, B. Bonnel, Y. Fouquart, and G. Brogniez, "Interband calibration over clouds for POLDER space sensor," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 1, pp. 131–142, Jan. 2002.

[24] P. Henry and A. Meygret, "Calibration of HRVIR and VEGETATION cameras on SPOT4," *Adv. Space Res.*, vol. 28, no. 1, pp. 49–58, 2001.

[25] B. Sohn, S.-H Ham, and P. Yang, "Possibility of the visible-channel calibration using deep convective clouds overshooting the TTL," *J. Appl. Meteorol. Climatol.*, vol. 48, no. 11, pp. 2271–2283, Nov. 2009.

[26] B. A. Baum, P. Yang, A. J. Heymsfield, S. Platnick, M. D. King, Y.-X. Hu, and S. T. Bedka, "Bulk scattering models for the remote sensing of ice clouds: Part 2: Narrowband models," *J. Appl. Meteorol.*, vol. 44, no. 12, pp. 1896–1911, Dec. 2005.

[27] Z. Zhang, P. Yang, G. Kattawar, J. Riedi, L. C. -Labonnote, B. A. Baum, S. Platnick, and H.-L. Huang, "Influence of ice particle model on satellite ice cloud retrieval: Lessons learned from MODIS and POLDER cloud product comparison," *Atmos. Chem. Phys.*, vol. 9, no. 18, pp. 7115–7129, Sep. 2009.

[28] Y. M. Govaerts and M. Clerici, "Evaluation of radiative transfer simulations over bright desert calibration sites," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 1, pp. 176–187, Jan. 2004.

[29] H. Cosnefroy, M. Leroy, and X. Briottet, "Selection and characterization of Saharan and Arabian desert sites for the calibration of optical satellite sensors," *Remote Sens. Environ.*, vol. 58, no. 1, pp. 101–114, Oct. 1996.

[30] T. C. Stone, "Radiometric calibration stability and inter-calibration of the solar-band instruments in orbit using the moon," in *Proc. SPIE*, 2008, vol. 7081, pp. 1–8.

Sindy Sterckx received the M.Sc. degree in bioengineering from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1999.

After working as a Research Associate at the same university for two years, she joined the Research Unit Remote Sensing and Earth Observation Processes, Flemish Institute for Technological Research (VITO), Mol, Belgium, in 2001. Currently, she is involved in the PROBA-V PV02 User Segment Project as technical responsible for the Image Quality Center where she defines the in-flight radiometric calibration strategy for the PROBA-V mission. Furthermore, she works as a Data Processing Scientist for the Airborne Prism Experiment project where she follows up the image quality, starting from the on-ground calibration to the atmospheric correction.

Stefan Livens received the M.Sc. and Ph.D. degrees in physics from the University of Antwerp, in Antwerp, Belgium, in 1992 and 1998, respectively.

His academic research focused on the application of image analysis and pattern recognition to microscopy imaging. It was complemented with postdoctoral research on error propagation in computer vision at Surrey University, SUrrey, U.K. For 9 years, he worked at Agfa Graphics, Mortsel, Belgium, as an expert on color image quality and imaging quality control. There, he performed research and software development for large format inkjet printing. Since 2008, he has been working at the Remote Sensing research unit at VITO. He has contributed to the development of the PROBA-V satellite mission throughout its various stages, focusing on imaging systems performance, radiometric, and geometric calibration.

Stefan Adriaensen (M'08) received the M.Sc. degree in industrial engineering with specialization in digital signal analysis from Katholieke Hogeschool Kempen, Geel, Belgium.

He was an RF Measurement Engineer and partly a Software Engineer with the Catholic University of Leuven, Leuven, Belgium. Currently, he is with the Flemish Institute for Technological Research (VITO), Mol, Belgium, where he focuses on the prototyping, design, and implementation of vicarious calibration algorithms and the automated workflows.