Remote Sensing of Environment 140 (2014) 141–154

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Radiometric comparison of multispectral imagers over a pseudo-invariant calibration site using a reference radiometric model

Marc Bouvet *

European Space Agency, Noordwijk, The Netherlands

ARTICLE INFO

Article history:
Received 4 April 2013
Received in revised form 26 August 2013
Accepted 27 August 2013
Available online 22 September 2013

Keywords:
Radiometric intercomparison
Radiative transfer
Pseudo-invariant calibration sites

ABSTRACT

A model is proposed to simulate top-of-atmosphere (TOA) observations in the visible to near-infrared (NIR) spectral range, over a pseudo-invariant calibration site, the so-called Libya-4 site. The model is based on a fully physical radiative transfer model simulating the coupling between a realistic atmosphere and a spectral surface Bidirectional Reflectance Distribution Function (BRDF) model parameterised by 4 free parameters. At first, the model is 'calibrated' on 4 years of MERIS observations by inverting the 4 free parameters of the surface BRDF model that provide the best fit to the MERIS observations. The model mimics the MERIS TOA observations with a precision of approximately 1% RMSE outside water vapour and O_2 absorption features. The inverted BRDF model parameters obtained at MERIS spectral bands are then spectrally interpolated and used as input to the radiative transfer model to simulate observations from ATSR-2, AATSR, A-MODIS, MERIS, POLDER-3 and VEGETATION-2 over the 2002 to 2012 period. Depending on the spectral band considered, AATSR radiometry appears 2% to 3% above the model 'calibrated' on MERIS radiometry, A-MODIS is 0% to 3% below, POLDER-3 is 2% to 4% below and VEGETATION-2 about 4% below. ATSR-2 data during the 2002 to early 2003 period are almost 10% below their simulations. Temporal trends between simulations and observations are also measured for all sensors. The smallest linear trends are observed for the MERIS 3rd reprocessing data (below 1%/decade). The temporal trends obtained from all sensors against the coupled surface-atmosphere model are in line with expected residual errors of instrument degradation model used in temporal extrapolation: larger in blue than in the NIR. The combined temporal trends from all sensors tend to demonstrate that the Libia-4 site is radiometrically stable in the visible to the NIR to better than 1%/decade for the 2002–2012 period, thus quantitatively confirming that it is a terrestrial target particularly adequate for the assessment of the temporal stability of Earth Observation sensors.

1. Introduction

In the last decades, many space borne optical sensors have been globally imaging the Earth radiation field at the kilometric spatial resolution scale to serve both scientific and operational applications. Both types of application generally require that these space measurements be traceable to the radiometric standards of the Système International (SI). Such traceability cannot be ensured only by means of exhaustive on-ground instrument characterisation and calibration. The harsh environmental constraints imposed by the launch and in orbit environments inevitably lead to sensor performance degradation and consequently to loss of traceability as the instrument radiometry cannot be related anymore to the SI sources available on ground. Even when onboard calibration devices, such as solar diffuser plates or reference lamps, are present on the same platform to ensure radiometric traceability to on ground SI standards, these devices can suffer from similar performance degradation issues. Each sensor performance is generally affected differently in the course of the its lifetime by space environment constraints. Such performance issues are at best partially accounted for and sometimes not identified in the instrument degradation models put in place to compensate them. This not only leads to a loss of radiometric traceability at individual sensor level but also to an overall loss of radiometric consistency between sensors and by extension of the global Earth Observation system.

Space agencies and private operators have increasingly relied on vicarious calibration methodologies and radiometric intercomparison methodologies to identify and to recover from such loss of radiometric traceability in the sensor lifetime. To support these methodologies, a set of six desert sites (referred to as Algeria 3, Algeria 5, Libya 1, Libya 4, Mauritania 1 and Mauritania 2) called the Pseudo Invariant Calibration Sites (PICS). They were selected by the Committee on Earth Observation Satellites (CEOS) among 20 sites originally identified by Cosnefroy, Leroy, and Briottet (1996) as desert sites suitable for "*multitemporal*, *multiband*, or *multiangular calibration of optical satellite sensors*". All sites are located in the Sahara and were chosen chiefly for their potential radiometric stability.

* European Space Agency, Keplerlaan 1, PB 299, NL-2200 AG, Noordwijk, The Netherlands. Tel.: +31 715656571.
E-mail address: marc.bouvet@esa.int.

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.08.039

The Sahara region radiometric stability should however not be taken for granted. Its climate changed rapidly from an arid climate to a humid climate in the early Holocene (~8500 BC) and then back to today's dry climate (~5000 BC) in response to climate forcing induced by solar insolation changes due to the Earth precession and possibly a non-linear response to it (deMenocal et al., 2000; Kuper & Kröpelin, 2006). This region of the globe is subject to the recurrent so-called African Humid Periods, following a 20,000 year precession cycle, leading to significant change in particular in vegetation cover and atmospheric regime (see deMenocal & Tierney, 2012). The growing evidence for contemporary climate change and the past record of climate variability of this region invite to caution when assuming its radiometric stability. Some of the sites originally identified by Cosnefroy et al. (1996) have also recently been subject to human induced changes such as the development of oil extraction fields. A quantitative estimate of the PICS temporal stability is necessary if they are to be used to diagnostic and correct for instrument radiometric degradation.

The present paper focuses on one of the 6 PICS, the Libya-4 site. This site has been extensively used for radiometric intercomparisons of Earth Observation (EO) sensors and their vicarious calibration (e.g., Lachérade, Fougnie, Henry, & Gamet, 2013; Smith & Cox, 2013; Sterckx, Livens, & Adriaensen, 2013). The present study proposes a model of the spectral and angular variability of the TOA reflectance over the site at a 1 nm spectral resolution in the visible to Near Infrared (NIR). This model is used as a radiometric reference against which observations from various space sensors are compared. The originality of the proposed modelling lies in that it is based on a fully physical description of the radiative transfer scattering and absorption processes on the 4 St parameters of the radiation, by a Monte Carlo model, at high spectral resolution (1 nm), with a realistic atmosphere and a full description of the surface-atmosphere coupling via a hyperspectral Bidirectional Reflectance Distribution Function (BRDF) model. Such spectral resolution allows to simulate observations from sensors with similar spectral responses and compare them whilst implicitly taking into account the slight differences between their spectral responses and thus constitutes an alternative to the use of spectral bands adjustment factors derived from hyperspectral sensor data (e.g. Chander et al., 2013).

The first step of the approach is to invert the 4 free parameters of a surface spectral BRDF model from 4 years of the ENVISAT MEdium Resolution Imaging Spectrometer (MERIS) observations over the site. This process can be described as the 'calibration' of the model output onto the MERIS measurements. The BRDF model parameters inverted in MERIS spectral bands are then spectrally interpolated at a 1 nm spectral sampling and used as input to the radiative transfer model in order to simulate full time series of observations from the EO sensors: ATSR-2 (on board ERS-2), AATSR (on board ENVISAT), MODIS (on board AQUA), MERIS, POLDER-3 (on board PARASOL) and VEGETATION-2 (on board SPOT-5). Finally, the question the radiometric temporal stability of the Libya-4 site is addressed and for the first time a quantitative upper estimate of its decadal stability is estimated from the analysis of all sensors radiometric temporal trends.

2. Definition of the reference dataset of TOA reflectance over Libya-4

MERIS L1 data from the 3rd reprocessing covering a 4 year period from 01/01/2006 to 31/12/2009 were extracted from the freely available Database for Imaging Multi-spectral Instruments and Tools for Radiometric Intercomparison (DIMITRI) (http://www.argans.co.uk/dimitri/). These data consist of TOA reflectances averaged over the so-called Libya-4 region of interest (ROI). The Libya-4 site is defined as per Cosnefroy et al. (1996), i.e., a latitude/longitude box between 28.05 N–29.05 N and 22.89 E–23.89 E. A description of the site and its climate can be found in Cosnefroy et al. (1996). The radiometric spatial variability of the Libyan desert is such that the choice of the ROI centre and ROI size has a significant impact on the absolute TOA radiometric level of the time series of sensor observations. Previous studies aiming at intercomparing sensors over the so-called Libya-4 site have made use of a ROI centred at about the same location than Cosnefroy et al. (1996) than the present study but with different ROI sizes (e.g., Lachérade et al., 2013; Smith & Cox, 2013; Sterckx et al., 2013). One can distinguish two typical ROI sizes being used in such studies. Firstly, the ROIs having size of about 1° latitude and longitude; they are best suited for intercomparison of sensors with spatial resolution of the order of 1 km as they provide a large number of TOA measurements (about 1000 pixels) with a low associated standard deviation around 2%. Secondly, smaller ROIs of size about 0.2° latitude and longitude; they are more likely to be fully covered by the narrow swath of high spatial resolution sensors. In the present study, the choice of the ROI first defined by Cosnefroy et al. (1996) was driven by the fact that the sensors involved have kilometric spatial resolutions.

The L1 data were corrected for the instrument smile effect (irradiance correction only following Bourg, D'Alba, & Collagrande, 2008). They were automatically cloud screened following the MERIS-GlobCarbon scheme as per Plummer (2008). Data were further visually screened using quicklooks to exclude acquisition with residual cloud contamination. 416 MERIS acquisitions are available after this data screening (see Fig. 1). Cloud contamination over the site occurs more frequently during the winter months resulting into more acquisitions being available for low sun zenith angles (SZAs) than for high SZAs. To avoid over-constraining the inversion of the surface BRDF model (described in the next section) at low SZAs at the expense of high SZAs, a sub-selection of 200 of these cloud free MERIS TOA observations is done at times randomly selected and uniformly spread across the 4-year period.

For each acquisition, the following data are automatically extracted from the DIMITRI database: the mean TOA reflectance over ROI, the standard deviation of the TOA reflectance within the ROI, the sun and viewing direction zenith and azimuth angles (SZA, VZA, SAA, VAA), the total columnar ozone (TCO) and the total column water vapour (WV)). These TCO and WV available in the DIMITRI database correspond to the meteorological data available in the MERIS L1 products. They are data from the European Center for Medium-Range Weather Forecasts (ECMWF) operational Numerical Prediction Weather (NWP) model. They were substituted by the corresponding ECMWF ERA-Interim reanalysis data (see description in Dee et al., 2011). These ERA reanalysis meteorological data present several advantages:

1. They are generated with a single NPW model throughout the entire reanalysis period (1979 to present). In contrast the ECMWF NWP operation model has gone through 23 evolutions during the ENVISAT ERA.
2. They are generated with consolidated assimilated data.
3. They are generated on a Gaussian grid with an approximately uniform 79 km spacing compared to the 125 km spacing of the data originally available with the ENVISAT data.

3. Retrieval of the surface BRDF model parameters

The retrieval of a surface BRDF model over Libya-4 is achieved by minimizing the residual error between the previously described reference dataset of MERIS TOA reflectances and their simulations. The TOA reflectance simulations are carried out with the MYSTIC radiative transfer code.

3.1. The radiative transfer modelling: MYSTIC setup

MYSTIC was first described by Mayer (1999, 2000) as a forward Monte Carlo for plane parallel atmospheres. It took part to the International Intercomparison of 3D Radiation Codes (see Cahalan et al., 2005). The addition of a backward simulation mode and simulations in spherical atmosphere are described in Emde and Mayer (2007). The model has been further extended to simulate polarised radiation. These extensions were validated against exact solutions, benchmark results and measurements by Emde, Buras, Mayer, and Blumthaler (2010). MYSTIC has

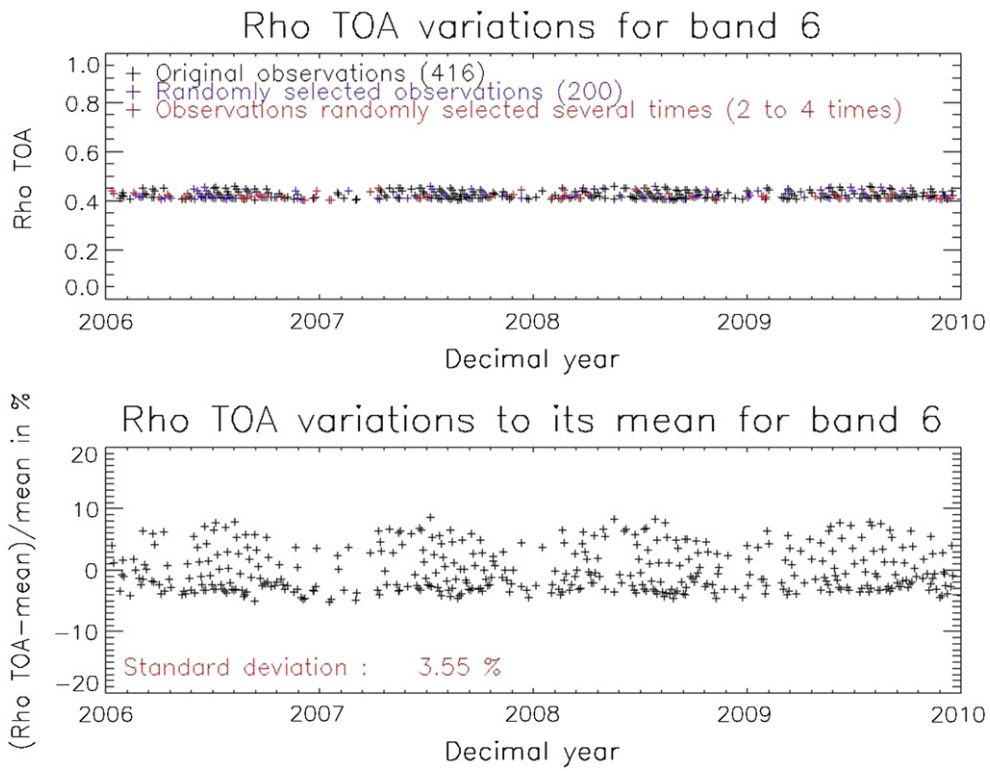

Fig. 1. In the top panel, the MERIS TOA reflectance in band 6 (620 nm) (in black). In blue, the acquisitions selected for the RPV parameter inversion and in red acquisitions selected several times. In the bottom panel, the relative difference between the TOA reflectance and its mean value over the 4 years.

further evolved as described in Buras and Mayer (2011) and Emde, Buras, and Mayer (2011) to improve its efficiency in cloudy atmospheres and for simulations at high spectral resolution. MYSTIC is freely available in its 1-D version from the libRadtran package (http://www.libradtran.org).

In this study, MYSTIC was operated in backward mode and in vector mode to provide the 4 St parameters of the radiation. In the following sections, the choices made for its set up are presented and discussed.

3.1.1. Spectral parameterisation

MYSTIC was run with the pseudo spectral gaseous absorption parameterisation of LOWTRAN 7 (Pierluissi & Maragoudakis, 1986) adopted from the SBDART code (Ricchiazzi, Yang, Gautier, & Sowle, 1998). The spectral resolution of LOWTRAN 7 is 20 cm^{-1} FWHM (Full Width at Half-Maximum) in averaged steps of 5 cm^{-1} in the spectral range of 0 to 50,000 cm^{-1} or 0.2 μm to infinity. This corresponds to a resolution of 0.32 nm at 400 nm and 2 nm at 1 μm. A single parameter band model (pressure) is used for molecular line absorption in LOWTRAN 7. Separate band models and parameters for the absorbing molecules are available for: H2O, O3, N2O, CH4, CO, O2, CO2, NO, NO2, NH3 and SO2. These band model parameters were developed with and based on degraded line-by-line spectra and validated against laboratory measurements (see Abreu & Anderson, 1996; Pierluissi & Maragoudakis, 1986). These band models allow for an accurate description of absorption processes in radiative transfer calculations performed at comparable or lower spectral resolution. The actual TOA simulations of all space sensors in this study are done with a spectral sampling of 1 nm that does not justify using a gaseous absorption parameterisation at higher spectral resolution than LOWTRAN 7 or using line-by-line calculations.

These hyperspectral simulations are convolved with the sensor spectral responses available in DIMITRI (and all originally obtained from the respective space agencies).

1000 photons are assigned to each monochromatic simulations resulting into typical photon noise of the order of 0.5% for the MERIS 412 nm band (10 nm band) down to about 0.2% in the MERIS NIR bands (see Fig. 5).

3.1.2. Atmosphere parameterisation

The atmosphere is assumed plane parallel with a mid-latitude summer atmosphere following Anderson, Clough, Kneizys, Chetwynd, and Shettle (1986). This atmosphere profile description includes trace gases profiles and default standard values for total column content. For each sensor TOA simulation, the ozone and water vapour profiles are scaled to match the TOC and the WV content of the ERA-Interim reanalysis fields at the sensor observation time. It is important to note that this effectively implies that the atmosphere used to simulate time series of sensor observations is temporally variable (while the BRDF model described later is assumed time invariant).

The aerosol profile and optical properties are chosen to follow the so-called continental average model described in Hess, Koepke, and Schult (1998). The composition of the continental average model consists of a mixture of soot and insoluble and water-soluble components.

The vertical distribution follows a 3-layer distribution. First, a 2-km thick continental aerosols layers following an exponential profile (scale height of 8 km, with a total aerosol optical thickness (AOT) of 0.151. Above this layer a 10 km thick layer of continental aerosol (with slightly difference optical properties), with an optical thickness of 0.013 is added. The last layer is a stratospheric aerosol layer extending from 12 km to 35 km altitude with an optical thickness of 0.005. Furthermore, the total aerosol optical thickness at all wavelengths is rescale by a fixed factor in such way that the total optical thickness of this three-layer aerosol vertical distribution amounts to 0.2 at 550 nm.

3.1.3. The BRDF model

The RPV model proposed by Rahman, Pinty, and Verstraete (1993) is a parametric BRDF model representing BRDFs. It is represented by the

product an amplitude ρ_0 of three separate functions accounting for both the illumination and viewing directions:

$$\rho\left(\theta_0,\theta_V,\Delta\phi,\rho_0,k,\Theta,\rho_c\right)=\rho_0 M_1\left(\theta_0,\theta_V,k\right)F_{HG}(g,\Theta)H(\rho_c,G) \quad (1)$$

Where

$$M_1\left(\theta_0,\theta_V,k\right)=\frac{\cos^{k-1}\theta_0\cos^{k-1}\theta}{(\cos\theta_0+\cos\theta)^{1-k}} \quad (2)$$

$$F_{HG}(g,\Theta)=\frac{1-\Theta^2}{\left(1+2\Theta\cos g+\Theta^2\right)^{3/2}} \quad (3)$$

$$H(\rho_c,G)=1+\frac{1-\rho_c}{1+G} \quad (4)$$

$$\cos g=\cos\theta_0\cos\theta+\sin\theta_0\sin\theta\cos\Delta\phi \quad (5)$$

$$G=\left(\tan^2\theta_0+\tan^2\theta-2\tan\theta\tan\theta_0\cos\Delta\phi\right)^{1/2} \quad (6)$$

where θ and θ_0 are the VZA and SZA, respectively. In the above formulation the relative azimuth angle $\Delta\phi$ zero when the source of illumination is behind the sensor. The model is parameterised by 4 parameters: ρ_0, k, Θ and ρ_c.

The function M_1 defines the bell vs. bowl shape of the angular field through the parameter k. The function F_{HG} defines the importance of the forward vs. backward scattering. The function H describes the hotspot with its parameter ρ_c. The BRDF model follows the principle of reciprocity and is symmetrical with respect to the principal plane.

Boucher, Cosnefroy, Petit, Serrot, and Briottet (1999) compared the performance of several BRDF model at fitting goniometric measurements. In the specific case of sand, the RPV model outperformed the other models tested. Maignan, Breon, and Lacaze (2004) also compared several BRDF models in their ability to fit atmospherically corrected POLDER measurements. They conclude that the 3-parameter RPV model and the Li-Ross model perform better than other models.

Maignan et al. (2004) also observed that for the specific case of the hotspot, all models fail to accurately reproduce the sharp reflectance increase close to the backscattering direction. As satellite remote sensing data are generally acquired at scattering angles ranging from 100 to 180°, the hot spot signature is generally captured and it must be well described by the BRDF used for the present study. Hapke, DiMucci, Nelson, and Smythe (1996) discuss the nature of physical process behind the hot spot phenomenon. They conclude the hot spots of vegetation and dry fine grained soils are different in that they are dominated by respectively shadow hiding and coherent backscatter. In their original description, Rahman et al. (1993) have set ρ_c to ρ_0. They inverted this BRDF model from AVHRR data in channels 1 and 2 (respectively the red and NIR channels) over two desert sites located in Libya and Mauritania characterised by dune fields comparable to the Libya-4 site. They found the best fit with the 3 parameters $\rho_0=\rho_c=0.538$, $k=0.826$, and $\Theta=0.0269$ in channel 1 and respectively 0.699, 0.805, 0.0097 in channel 2. They also suggest a 4-parameter RPV model with a modified hot spot function is better suited for pronounced hot spot signature. The choice of a version of the RPV BRDF model with 4 parameters, as described in Eqs. (1) to (6), with a specific parameter dedicated to the hot spot modelling, is motivated by the need to maximise the fitting performance of the MERIS observation for which backscatter geometries are common over this site.

3.2. Inversion of the Libya-4 surface BRDF from the MERIS reference dataset TOA observations

The inversion of the 4 RPV model parameters from the MERIS TOA reflectance measurements is based on the minimisation of a cost function: the root mean square deviation between the TOA reflectance from the observations and their simulations. It is defined as:

$$\text{RMSE } in\ \%(\lambda)=\chi=\sum_{i=1}^{200}\sqrt{1/200\left[\frac{\left(\rho_{sim}^{TOA}(i,\lambda)-\rho_{obs}^{TOA}(i,\lambda)\right)}{\rho_{obs}^{TOA}(i,\lambda)}\times 100\right]^2} \quad (7)$$

Where $\rho_{sim}^{TOA}(i,\lambda)$ and $\rho_{obs}^{TOA}(i,\lambda)$ are the simulated and observed TOA reflectance respectively and i is the observation index ranging from 1 to 200 and λ is the spectral band ranging from 1 to 15.

The inversion is done following the downhill simplex method implemented in IDL, also called AMOEBA (see Numerical recipes in C, 2nd Edition, Cambridge University press) and originally described in Nelder and Mead (1965). It is an iterative method, making use of repeated runs to infer a trajectory through the 4-dimensional RPV parameter space leading to a minimum of the cost function. At each run, for each spectral band, the 200 TOA MERIS observations are simulated and χ is computed. The downhill simplex method is stopped when two successive computations of the cost function agree within 5%. It generally takes few tens of iterations to reach such level of stability of the cost function.

This minimisation method does not converge necessarily to the global minimum of a cost function if local minima are also present in the predefined parameter space. Moreover, in the specific case of this study, the convergence of the cost function is hampered by the stochastic nature of the Monte Carlo simulations, in particular in regions where the cost function 4-dimensional landscape is characterised by small derivative with respect to the RPV parameters.

The downhill simplex method must be initialised with a priori values of the 4 RPV parameters and estimates of the associated uncertainties. To maximise the chances of finding the global minimum of the cost function, the minimisation process is repeated 5 times, each time with a different a priori value for the ρ_0 parameter of the RPV model, starting from the mean TOA reflectance value from the 200 MERIS acquisitions and then sampling a ± 0.2 domain around this value. The a priori values for the k, Θ and ρ_c are kept identical for the 5 runs. They are set to 0.8, 0.0 and ρ_0 respectively, with their associated uncertainties set to ± 0.3, ± 0.3 and $\pm\rho_0$ respectively. After these 5 runs of the downhill simplex method, the 4 RPV parameters giving the lowest χ are selected. More runs of the downhill simplex with a larger range of initial values for the RPV parameterisation proved unnecessary.

3.3. Results of the RPV BRDF model parameters inversion

Fig. 1 shows the temporal variations of the TOA reflectance in the MERIS band centred at 620 nm. This spectral band is chosen in this section to illustrate the inversion process and its outputs. As previously detailed, out of the 412 acquisitions, only 200 are picked at randomly selected times following a uniformly distribution over the 4 year period. This results in acquisitions actually being selected up to 4 times in rare cases (in winter in particular). On the same figure, the performance of the simplest TOA reflectance model one could think of, i.e., the mean TOA reflectance over the 4 year period, is reported.

Fig. 2 shows the relative difference between the MERIS TOA reflectance observations in band 6 and their simulation with the radiative

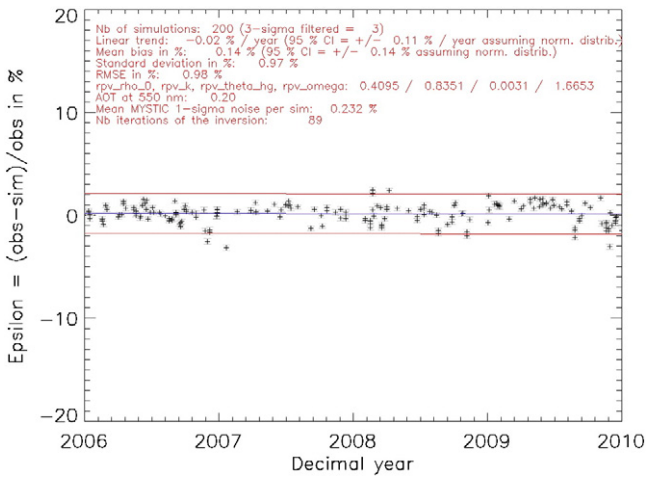

Fig. 2. The relative difference between the MERIS TOA reflectance in band 6 (620 nm) and its simulation using the inverted RPV parameters. The error bar associated to each acquisition is the Monte Carlo noise, the blue line is the linear fit and the red lines correspond to 2 standard deviations.

transfer model at the last iteration of the minimisation process. This relative difference is hereafter referred to as:

$$\varepsilon(i,\lambda) = \frac{\left(\rho_{obs}^{TOA}(i,\lambda) - \rho_{sim}^{TOA}(i,\lambda)\right)}{\rho_{obs}^{TOA}(i,\lambda)} \times 100 \quad (8)$$

In Fig. 3, the spectral variations of the inverted 4 parameters of the RPV BRDF model are plotted. The parameters obtained for band 11 (760 nm) are clear outliers (the model spectral resolution is not sufficient to accurately model this spectral band).

The values found for the parameters k and Θ controlling the overall anisotropy of the BRDF are comparable to the values found by Rahman et al. (1993) over desert sites from the inversion of AVHRR TOA observation (respectively 0.826 and 0.0269 in channel 1 and 0.805 and 0.0097 in channel 2). The values the ratio $\omega = {}^{\rho_c}/_{\rho_0}$ are generally larger than 1. Such value indicates that a hot spot of lower amplitude than what could be provided by a 3-parameter RPV BRDF model was found during the inversion. In Fig. 4, a polar representation of the surface BRDF model retrieved in band 6 is shown. The BRDF is characterised by a hot spot in the backscattering direction and a minimum in the forward scattering region for viewing angles below 30°. The yearly

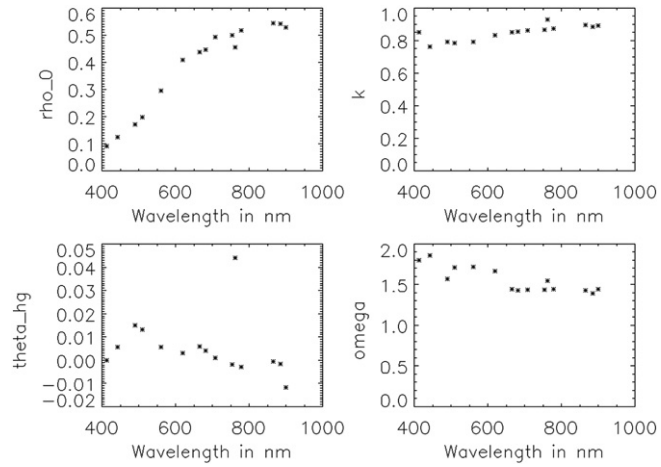

Fig. 3. The spectral variations of the RPV BRDF model parameters inverted from the MERIS TOA reflectance observations. From left to right and from top to bottom, ρ_0, k, Θ and the ratio $\omega = {}^{\rho_c}/_{\rho_0}$.

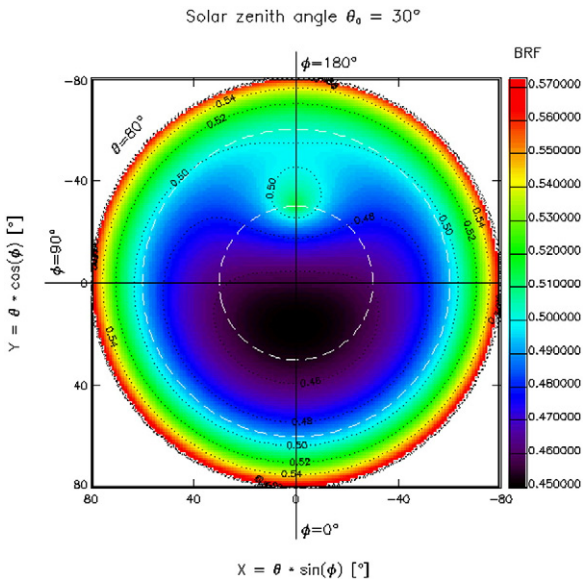

Fig. 4. RPV BRDF model representation for the parameters retrieved from MERIS band 6 TOA reflectance. Rho_0 = 0.413, k = 0.853 and theta_hg = 0.009 and rho_c = 0.664. Plotted with Anisview: software developed by P. Vogt and M. Verstraete.

variations with the SZA of this minimum in the BRDF can be observed in MERIS TOA observations for instance in the bottom panel of Fig. 1, as the lowest limit of the TOA reflectance relative to the mean.

Fig. 5 gives the χ value obtained for each MERIS spectral band at the final step of the minimisation. The spectral variations of standard deviation of the MERIS TOA reflectance over the 4 years are also shown. This figure demonstrate the ability of the model at capturing the geometrical and temporal variability of the TOA reflectance by comparing its performance to the simplest model of the TOA reflectance, namely, the mean TOA reflectance value over the 4 years (which RMSE is the standard deviation of the MERIS TOA reflectance over the 4 years).

In Fig. 5, the spectral variations of the mean values of the standard deviation of the Monte Carlo photon noise for the 200 MERIS TOA simulations is shown as well. It is generally below 0.5% and at least a factor 3 smaller than χ. Assuming χ^2 is the quadratic sum of the Monte Carlo noise and other residual errors (e.g.: MERIS instrumental noise, aerosol natural variability in measured TOA signal, etc.…), the Monte Carlo noise only contribute about 1/10th of the RMSE between MERIS

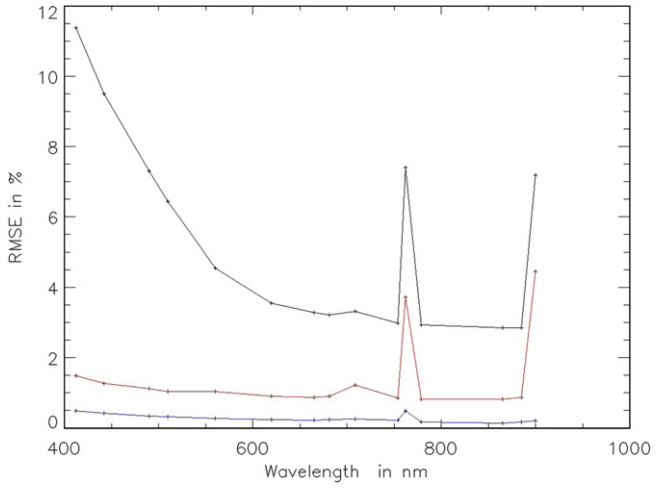

Fig. 5. The spectral variations of the RMSE between the MERIS TOA reflectance simulations and the observations is shown in red. In black, the MERIS TOA reflectance standard deviation for the 4 year period. In blue, the mean photon noise associated to a single simulation of a MERIS TOA reflectance observation.

observations and simulations. There are few outliers on this figure: band 11 (761 nm) and 15 (900 nm). Absorption in these band is the dominant process. The spectral sampling of the simulations (1 nm) is not sufficient to capture the spectral variability of the signal. Moreover, for band 15 (and to some extend for band 9 as well) the uncertainties on the water vapour from the ERA-Interim reanalysis might introduce an additional residual error to the simulations.

3.4. Analysis of the residual TOA difference between observations and simulations

For each time series of $\varepsilon(i,\lambda)$, i.e., the residual error between the MERIS TOA observations and their simulation (see for instance Fig. 2), the following statistical indicator are computed: the mean value of $\varepsilon(i,\lambda)$ over the 200 observations of the MERIS reference dataset (shown in Fig. 6), the linear trend of $\varepsilon(i,\lambda)$ with time (shown in Fig. 7), the standard deviation of the 200 $\varepsilon(i,\lambda)$ equal to χ and shown in Fig. 5). Moreover, a multivariate analysis of $\varepsilon(i,\lambda)$ was carried out against: the VZA, the SZA, the difference (SAA − VAA), the SAA, the scattering angle and the airmass. Finally, $\varepsilon(i,\lambda)$ frequency histograms were compared to Gaussian distributions with identical mean value and standard deviations. These analyses reveal that:

- The histograms of $\varepsilon(i,\lambda)$ in each band appear to approximately follow a normal distribution. This is an important result which implies that we can compute easily 95% confidence interval (CI) on the mean value of $\varepsilon(i,\lambda)$ over the 200 observations and also on the temporal linear trends. These 95% CIs are shown on the Figs. 6 and 7 respectively.
- The mean value of $\varepsilon(i,\lambda)$ over the 200 observations is not null within the 95% CI for several spectral bands (see Fig. 6). Such residual difference can be expected. In fact, the inversion of 4 parameters of the BRDF is based on the minimisation of the cost function χ, i.e., the root mean square error and not directly on the mean value of $\varepsilon(i,\lambda)$. The convergence criterion on χ was set to 5% of variation between two iterations of the inversion. Setting the conversion criterion to lower values results into 'hitting' the Monte Carlo noise of the simulation, i.e., the variations χ between two iterations starts to be dominated by Monte Carlo noise. Lowering the 5% convergence criterion in order to completely cancel out the residual mean value of $\varepsilon(i,\lambda)$ over the 200 observations would simply require to increase the number of photons used per simulation. This was not possible within the processing time available for this study.
- There is no significant correlation between $\varepsilon(i,\lambda)$ and the VZA, the SZA, the difference (SAA − VAA), the SAA, the scattering angle and the airmass ($r^2 < 0.09$ for all angles and for all bands). This is an important result indicating that the 4-parameter RPV BRDF model and the parameterisation of the atmosphere used to run MYSTIC radiative transfer code allows to capture the geometrical and temporal variability of the MERIS TOA signal.
- There appear to be statistically significant temporal linear trends in $\varepsilon(i,\lambda)$ in few spectral bands over the 4-year duration of the MERIS reference dataset (see Fig. 7). The trends are of the order of 0.1%/year. Larger trends of about 0.2%/year are observed in the 3 first bands. As discussed previously, band 11 modelling (and bands 9 and 15) requires a high spectral resolution model which was not the case in this study and a trend in the band should not be over-interpreted.
- Camera 4 for band 11 is an outlier that appears to provide measurements 5–10% brighter that the neighbouring cameras (not illustrated here).

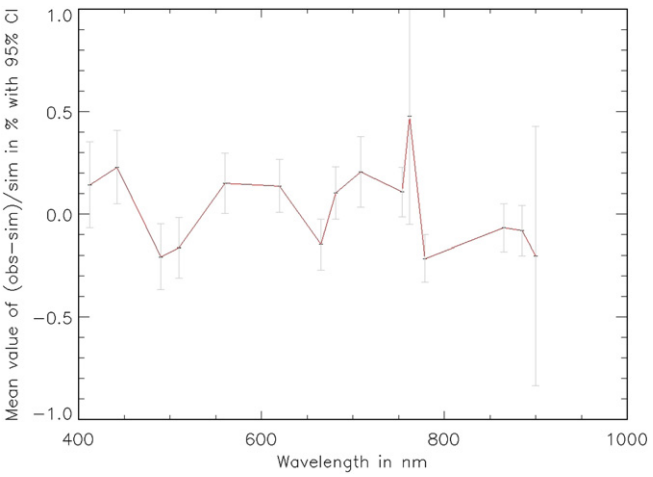

Fig. 6. The mean relative difference between the MERIS observation and their simulation and the associated uncertainty of the mean (95% CI).

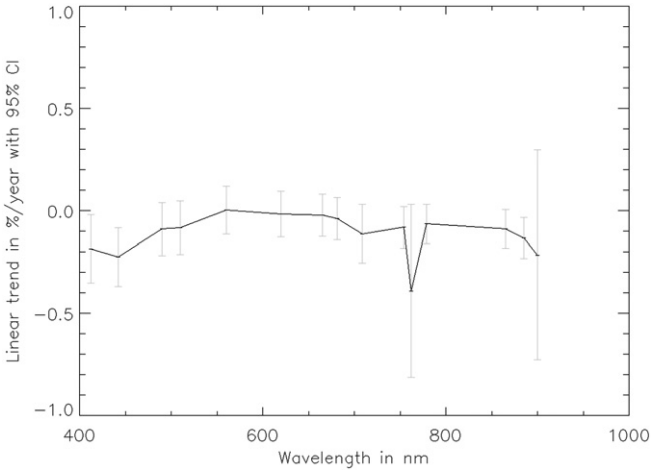

Fig. 7. The linear trends (in %/year) obtained for the time series of the relative difference between the MERIS TOA observations and their simulation. The error bar is the 95% CI.

3.5. Where does the residual error $\varepsilon(i,\lambda)$ come from?

In Fig. 5, the standard deviation associated to $\varepsilon(i,\lambda)$ varies from approximately 1.5% at 412 nm to 0.8% in the NIR (apart from the spectral bands with high gaseous absorption).

The residual error between the MERIS observations and their simulations can be split into two components: the bias and the apparently random error.

As discussed in Section 3.4, the residual bias shown in Fig. 6 is purely an artefact of the inversion process. An increased number of photons per simulation further reduce this residual bias.

The remaining random error (following a normal distribution) could not be fitted by the TOA reflectance modelling proposed in this study. What is the dominant contributor to this apparently random error? Several possible contributors and their magnitude are hereafter described.

3.5.1. MERIS instrumental noise

The MERIS signal-to-noise (SNR) was not characterised in flight. It should be comparable to the currently designed S-3/OLCI instrument which is designed to have SNR values of the order of 1000 to 2000 at a reference radiance level lower than the radiance measured over Libya-4. The contribution of this SNR to the mean TOA reflectance in the ROI is weighted by the square root of the number of pixels in the ROI (~800 pixels). Assuming 0.1% noise from the SNR we have a contribution below 0.01% to the mean in the ROI. Assuming this noise contributes quadratically to the total RMSE observed in Fig. 5, it should be completely negligible.

Another potential contributor to the observed RMSE is the smile correction applied to the data described in Bourg et al. (2008). The correction performed by the DIMITRI tool is only done at irradiance level and

does not take into account the spectral variations of the reflectance. A full correction, included irradiance and reflectance correction over the site is of the order 0.1% on average across the field of view and can reach up to 0.5% (test carried with the BEAM software available from https://earth.esa.int/web/guest/software-tools). Only correcting for the effect of smile on the irradiance results in errors in as much as 0.5% where the spectral slope of the reflectance is largest over this site.

The residual error of the correction carried out in the DIMITRI tool could partially explain the observed RMSE observed in Fig. 5 but it is unlikely to be the major contributor.

Other sources of instrumental noise, such as the straylight correction residual errors, the instrument dark current fluctuations or its non linearities might also contribute to the residual random noise observed between observations and simulations. Their individual magnitude is unknown but the RMSE shown in Fig. 5 provides an upper estimate to the combined random contribution to the TOA signal.

3.5.2. The site spatial heterogeneity

The standard deviation of the TOA reflectance in the Libya-4 site ROI is about 3% in all spectral bands. The ROI is mapped by about 800 MERIS measurements at 1 km spatial resolution for each MERIS acquisition. The value of the standard deviation of the TOA reflectance in the ROI remains approximately the same for each acquisition from MERIS, regardless of the acquisition geometry and the distribution of reflectance in the ROI follows approximately a normal distribution (not illustrate here).

In order to estimate the uncertainty on the mean TOA reflectance in the ROI induced by the pixel-to-pixel TOA reflectance in the ROI, let's assume that each one of these single pixel can be independently modelled by a random TOA reflectance generator having a normal distribution with a standard deviation of 3% and a mean value equal to the mean TOA reflectance over the ROI. For each acquisition, for each pixel in the ROI, the dies are rolled and the uncertainty of the mean TOA reflectance in the ROI follows: it is the standard deviation of the normal distribution divided by the square root of the number of pixels in the ROI, i.e., about 0.1% (1-sigma). This might be an upper estimate of the random variations induced by the spatial heterogeneity on the mean TOA reflectance over the ROI as there is a radiometric spatial correlation between pixels in the ROI that does not dramatically change between observation geometries. The site spatial heterogeneity cannot explain the RMSE illustrated in Fig. 5.

3.5.3. Residual error of the RPV model

Is the RPV BRDF model able to reproduce the geometrical variability of the actual Libya-4 site surface BRDF? Boucher et al. (1999) fitted the RPV model on a sand sample BRF measured with a goniometre. They obtained fitting RMSE of about 2% at 600 and 800 nm. They claimed precision of their BRDF measurements of the same order (between 2% and 5%). In the present study, in the NIR, where the atmosphere is optically thin and the TOA reflectance is a good first order estimate of the surface BRF, the standard deviation of $\varepsilon(i,\lambda)$ is about 1% (see Fig. 5). This is a strong indication that the RPV model is better at fitting the site BRDF than observed by Boucher et al. (1999) on sand samples and a demonstration that it can describe the site surface BRDF to at least approximately 1% RMSE.

3.5.4. MYSTIC photon noise

It was shown in Section 3.3 that the MYSTIC photon noise produces a negligible contribution to the RMSE in Fig. 5.

3.5.5. Atmospheric variability

The atmospheric surface pressure and the temperature profile are fixed input to MYSTIC. In reality these vary both in time and space. They are likely contributors to the variability observed in bands 9, 11 and 15. Water vapour is also a likely contributor to the variability of the TOA signal in bands 9 and 15 and ERA-Interim data might not be perfectly reflecting reality. The modelling of these bands with significant atmospheric absorption was done at 1 nm spectral resolution. This might not be sufficient for such bands where the spectral variation of the absorption cross sections of the gas requires better spectral resolution.

Both the atmospheric variability and the modelling limitations are likely to explain most of the residual errors observed in bands 9, 11 and 15.

At other spectral bands, gaseous absorption is much lower or well described at a lower spectral resolution (e.g.: for ozone). However, another factor of atmospheric variability might result into residual errors between observations and simulations: aerosol.

The issue of the most suitable aerosol model to use in combination with the RPV model to simulate the MERIS observations was addressed by looking at the final RMSEs obtained between TOA MERIS observations and their simulations for two aerosol models described by Hess et al. (1998): the 'average continental' aerosol model and the 'desert' aerosol model. One might expect the 'desert' aerosol model to be more representative of the Libya-4 aerosol optical properties. However, the final RMSEs obtained with the 'desert' model (not shown here) are higher by up to 0.5%, in the 400–700 nm region, than those obtained with the 'average continental' model (for which the resulting RMSEs are shown in Fig. 5). Further analysis of the $\varepsilon(i,\lambda)$ obtained with the 'desert' model reveals a residual variation of $\varepsilon(i,\lambda)$ with the scattering angle. This does not imply that the 'average continental' aerosol model is more representative of the real aerosol present over the Libya-4 site than the 'desert' model. It is rather the combination of this aerosol model optical properties (in particular the scattering phase function) and the RPV surface BRDF model that that allow the best fit of the MERIS TOA observations. In other words, the 'desert' aerosol model described by Hess et al. (1998) might be more representative of the real aerosol present over the Libya-4 site but when combined in the radiative transfer model with a RPV BRDF model it is not as efficient at mimicking MERIS TOA observations as the 'average continental' model. The 'average continental' model was thus selected for the purpose of this study.

To further assess the sensitivity of the TOA signal to aerosol two cases were simulated. In the first case, the inverted BRDF model parameters (shown in Fig. 3) and the same atmospheric parameterisation (i.e., including the 'average continental' aerosol model) used for the inversion are used to simulate the 200 MERIS observations. The total AOT at 550 nm is however set to 0.3 rather than the value of 0.2 used for the inversion for the BRDF model. This increase in AOT leads to an increase by about 1% of the TOA signal in the blue spectral bands and a decrease of about 2% towards the red and NIR.

In the second case, the same AOT value of 0.2 as used for the inversion of the RPV parameters is re-used to simulate the MERIS TOA observations but the 'average continental' aerosol model is substituted by the so-called 'desert' aerosol model (see Hess et al., 1998) to simulate the 200 MERIS observations. These simulations are meant to mimic the change in aerosol optical properties that might when desert dust is lifted off locally by wind driven process and substitutes the aerosol transported over long distances. These changes in aerosol optical properties lead to a decrease of nearly 6% of the TOA signal in the blue and an increase of 2% towards the red and NIR.

Both cases demonstrate the sensitivity of the simulated (and observed) MERIS TOA reflectances to changes in aerosol load and optical properties. The low RMSE obtained at the end of the BRDF model parameter inversion (as shown in Fig. 5) is mostly due to variations in $\varepsilon(i,\lambda)$ at the daily scale. These variations could be explained by day-to-day changes in aerosol load and optical properties over the site. No variation of $\varepsilon(i,\lambda)$ is observed at longer time scales (as illustrated in Fig. 2 for band 620 nm). In particular, no seasonal variability in $\varepsilon(i,\lambda)$ is found. This indicates that the seasonal variations of the aerosol optical properties and AOT do not need to be modelled to reproduce MERIS observations at TOA level.

3.6. What is the impact of the number of MERIS observations used to 'calibrate' the MERIS TOA reflectance model?

The objective of the surface BRDF retrieval scheme described in previous sections is to provide the surface BRDF model parameters that best allow simulating the 200 MERIS TOA observations. What is the performance of the MERIS TOA reflectance model in relation to this number of observations?

The performance of the MERIS TOA reflectance model can be directly observed at single observation level on plots of $\varepsilon(i,\lambda)$ against time as show in Fig. 2. More synoptically, the performance can be assessed in terms of simple statistical indicators: the spectral bias (plotted in Fig. 6) and the spectral standard deviation (not plotted here but well approximated by the spectral RMSE in Fig. 5). These two statistical indicators, in turn, have associated uncertainties that directly depend on the number of observations used to derive them.

Fig. 6 shows the uncertainty associated to the spectral bias as a 95% CI. The CI is computed as twice the spectral standard deviation divided by the square root of the number of observations (200). The 95% CI on the spectral bias is of the order of 0.15% for all bands marginally affected by gaseous absorption. Such CI is acceptable if this model is to be used to as reference to identify radiometric differences between MERIS and sensors of the order of few percent. For instance, if only 50 MERIS observations (about half a year of observations) had been used to retrieve the surface BRDF model parameters, the 95% CI on the TOA model spectral bias would be a factor 2 higher, i.e., about 0.3%.

The uncertainty on the spectral standard deviation can be assessed in the framework of normal distributions. The standard deviation associated to the population of $\varepsilon(i,\lambda)$ is about 1% of all bands not significantly affected by gaseous absorption. The 95% CI on a standard deviation of 1% for a population of 200 samples normally distributed can be calculated and is about $\pm 0.1\%$ (it is actually slightly asymmetrical). The 95% CI on this standard deviation becomes about $\pm 0.2\%$ if only 50 samples are used.

As a conclusion, the number of MERIS observations (200) used to 'calibrate' the TOA radiative transfer model is sufficient to ensure that its performance in terms of bias and standard deviation of its error are known with 95% CI suitable for its further use as a reference TOA model for the radiometric intercomparison of sensors (i.e., about $\pm 0.15\%$ on the bias and about $\pm 0.1\%$ on the standard deviation of the error). A lower number of MERIS observations (50) used to 'calibrate' the model would roughly double these uncertainties on the TOA model bias and the standard deviation of its error.

4. Discussion on the extension of the geometrical and spectral domain of validity of the inverted BRDF model

In the previous sections, a 4-parameter BRDF model was inverted using a 4-year reference dataset of MERIS observations. The residual errors of the resulting coupled surface-atmosphere model were assessed at MERIS spectral bands and for the MERIS observational geometries. In this section the extension of the geometrical and spectral domain of validity of the model are discussed and various options are assessed.

4.1. Geometrical domain of validity of the TOA reflectance simulations

The 200 MERIS observations used to invert the 4-parameter BRDF models have observation geometries characterised by VZAs approximately ranging from 0° to 35°, SZAs from 18° to 60° and RAAs from −180° to 50°. These 200 MERIS observation and illumination geometries however do not cover continuously nor regularly these ranges of VZA, SZA and RAA values. In fact, these angles are strongly correlated for MERIS observations due to the ENVISAT repeat cycle and the instrument geometry of observation.

It is worth noting that there is no significant correlation between $\varepsilon(i,\lambda)$ and all viewing and illumination angles: correlation coefficients (r^2) are below 0.09 for all bands. This demonstrates the geometrical fitting skills of the proposed couple surface-atmosphere modelling. However, such skills cannot be assessed outside the MERIS observational geometries and consequently simulations for other sensors will be done only for geometries matching one of the 200 geometries of the reference dataset.

4.2. Extension of the spectral domain of validity of the TOA reflectance simulation

In order to simulate the observations from other sensors than MERIS the inverted spectral BRDF model must be interpolated in spectral regions in between the MERIS spectral bands. In other words, the inverted multispectral BRDF model must be turned into a hyperspectral BRDF model. Two interpolation schemes are assessed and discussed in this section. Moreover, it is investigated if the spectral sampling of MERIS and POLDER-3 are sufficient to accurately reconstruct the spectral variations of the surface bidirectional reflectance factor.

First, the importance of the spectral sampling for the construction of an hyperspectral BRDF model can be assessed by the following experiment: only the RPV parameters inverted from MERIS TOA observations in band 443 nm, 560 nm, 665, 754 and 865 nm are used to derive the RPV parameters at a resolution of 1 nm, through the 390 nm to 960 nm spectral interval, using a spline interpolation. This is quite representative of the spectral sampling of sensors like POLDER or MISR. The result is shown in Fig. 9. Using the interpolated RPV values obtained from such interpolation approach at all MERIS spectral bands, we can simulate the MERIS TOA reflectance using the radiative transfer model and the pre-defined atmosphere and compare to the actual MERIS TOA observations. The resulting biases for the 200 MERIS acquisitions are shown in Fig. 10 and should be compared to those obtained in Fig. 6. The obtained bias reaches 9% at 510 nm. With only a subset of the MERIS spectral bands used for the spline interpolation of the RPV model parameters, their spectral variations are not sufficiently well captured as clearly indicated by the poor comparison between simulated and observed MERIS TOA reflectances.

The MERIS spectral sampling is superior to the spectral bands used for this experiment and the biases obtained in Fig. 10 should be regarded as relevant to POLDER or MISR if these sensors were to be used to invert the RPV parameters in their spectral bands and then spectrally interpolate them to simulate other sensor observations. In the particular case of using MERIS observations, it is however a warning that the interpolation of the RPV between the MERIS spectral bands should

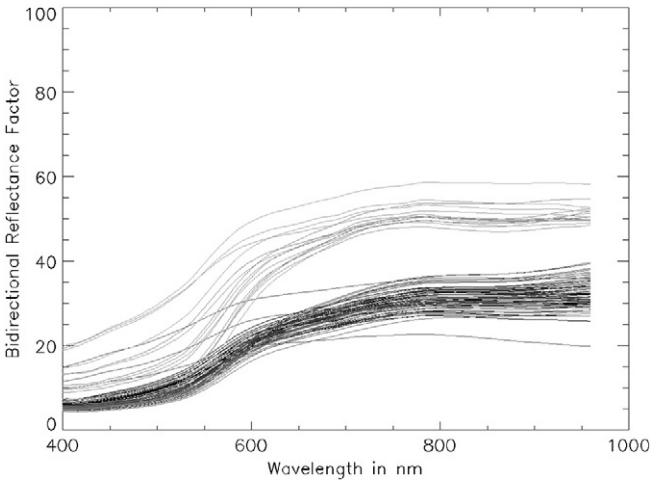

Fig. 8. All BRF sand spectra measured on samples collected at the Simpson desert (Australia), Muleshoe (Texas, USA) and in the Namibian desert. Illumination is at 45° and measurement at nadir.

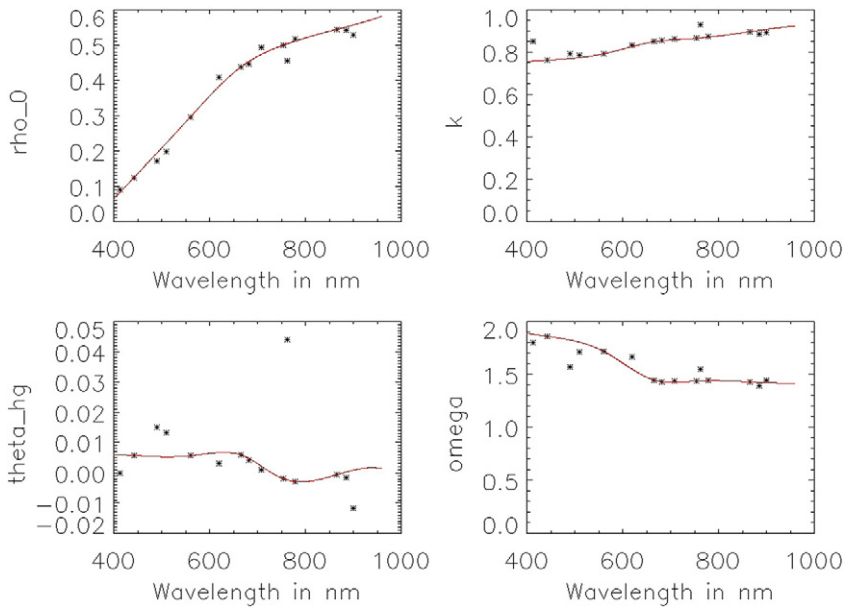

Fig. 9. The spectral variations of the RPV BRDF model parameters inverted from the MERIS TOA reflectance observations. From left to right and from top to bottom, ρ_0, k, Θ and the ratio $\omega = {}^{\rho_c}/_{\rho_0}$. The red line for each parameter corresponds to the spline interpolation based on spectral band at 443 nm, 560 nm, 665, 754 and 865 nm.

be exercised cautiously, in particularly in the spectral region 500 nm to 600 nm.

The spectral properties of sand in the visible to NIR spectral region are indeed quite variable (see Fig. 8). They depend on grain size, shape and composition. Part of this variability is explained by the presence of iron oxide coatings on sand grain surface giving colour ranging from brown and to red. The relationship between increasing redness and increasing concentrations of iron oxide in the sediments is strong and positive (Bullard & White, 2002). In visible to NIR spectral range, the absorption of these coatings is strongest in the spectral region extending between 500 nm and 600 nm. This absorption feature is visible as a steep increase of the amplitude parameter ρ_0 of the RPV BRDF model around 560 nm in Fig. 3. The spectral variations of the RPV BRDF model are mainly governed by the spectral variations of the amplitude parameter ρ_0. The spectral variations of the three other parameters of the model affect to a lesser extend the spectral variations of the output BRF. Two interpolation methodologies are assessed to interpolate the ρ_0 values obtained in the 15 spectral bands of MERIS: 1) a spline interpolation between the MERIS spectral bands and 2) a spectral reconstruction based on a prior Principal Component Analysis of a hyperspectral sand reflectance database (see spectra in Fig. 8). The hyperspectral database used for the second interpolation approach consist of 83 BRF spectra collected and measured by J. E. Bullard and K. H. White. They are spectroscopic laboratory measurements carried out on samples from the Simpson desert (Australia) and described in Bullard and White (2002)), from the Muleshoe Dunes (USA) and described in White and Bullard (2009), and unpublished similar data from the Namibian desert (Namibia). Measurements were taken in a purpose built spectroscopy lab (matte black walls and ceiling). Sand samples (flattened surface) were measured with a GER3700 spectrometer at 55 cm distance. Illumination was provided by a 1KW daylight video lamp at 45° angle. A spectralon panel was used as reference reflectance. The spectra were acquired at a spectral resolution of 3 nm between 400 nm 800 nm and 4 nm beyond. The spectra were further denoised by convolving them with a 10 nm wide square kernel.

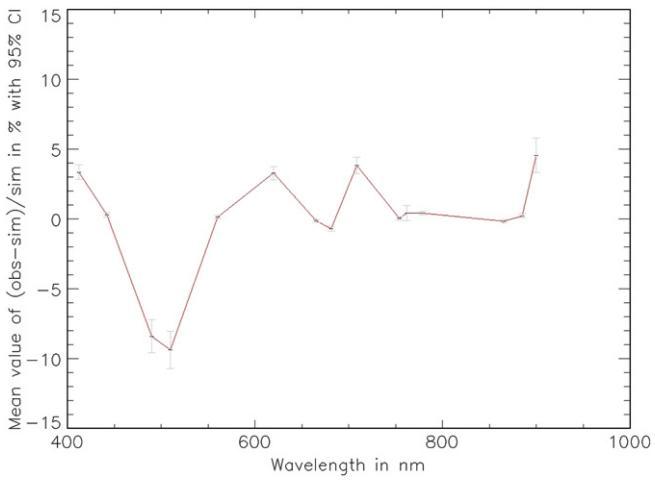

Fig. 10. The mean relative difference between the MERIS observation and their simulation and the associated uncertainty of the mean (95% CI) when using RPV parameters at 412 nm, 490 nm, 510 nm, 620 nm, 681 nm, 709 nm, 778 nm, 885 nm resulting from the interpolation of those obtained from the direct inversion of MERIS observations at 443 nm, 560 nm, 665, 754 and 865 nm.

These BRF spectra were convolved with MERIS and POLDER-3 spectral responses to generate pseudo MERIS/POLDER-3 measurements. These pseudo measurements were subsequently interpolated either by 1) a spline function or 2) using a linear combination of the first 8 hyperspectral eigen vectors derived from the hyperspectral sand reflectance database and fitted the pseudo measurements by a least square minimisation. In both cases, the interpolated spectra obtained are then compared to the original GER3700 measurements. The systematic error and the 1-sigma random error are shown in Fig. 11. Analysis of these comparisons reveals that: 1) the POLDER-3 spectral sampling combined with a spline interpolation leads to systematic BRF errors exceeding 5% in the 500 nm–600 nm spectral range and exceeding 10% at 400 nm; 2) the MERIS spectral sampling combined with a spline interpolation leads to systematic errors of few percents for all sand samples and reaching about 5% for the reddest sand samples; 3) the MERIS spectral sampling combined with an interpolation based on a linear combination of 8 eigen vectors of the hyperspectral sand reflectance database leads to systematic errors generally negligible. The latest interpolation methodology is thus selected for the generation of an hyperspectral

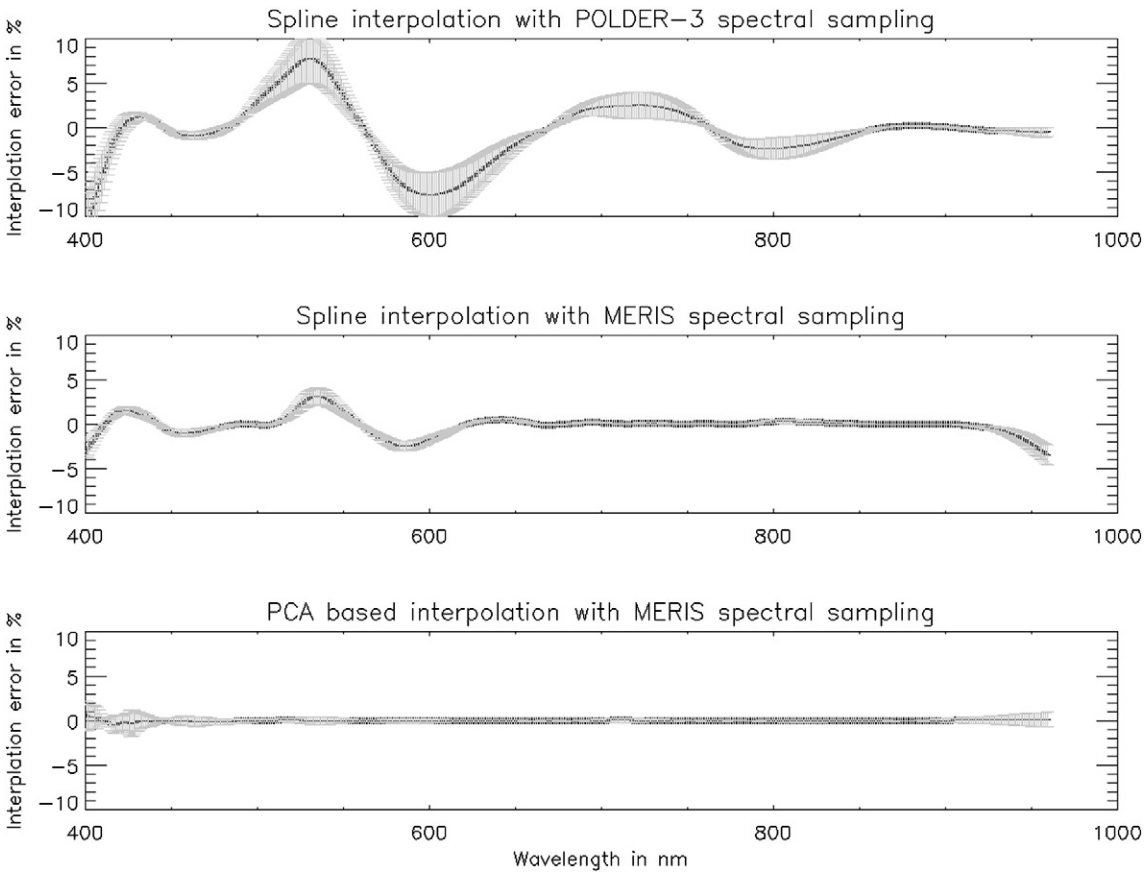

Fig. 11. The systematic error (in black) and the 1-sigma random error (grey error bar) of 3 interpolation approaches derived from the comparison of 83 sand BRF spectra and their interpolation using 1) the POLDER-3 spectral sampling with a spline interpolation (top panel), 2) the MERIS spectral sampling with a spline interpolation (centre panel) and 3) the MERIS spectral sampling with an interpolation based on the linear combination of eigen vector from the sand database.

surface BRDF model of Libya-4 with a view to simulate observations from space sensors operating in the visible to NIR spectral region.

5. Radiometric comparisons between the Libya-4 simulator and various sensors

In the following section, the previously described Libya-4 simulator is used to simulate time series of TOA observations from AATSR (2nd reprocessing), ATSR-2 (2008 reprocessing), MERIS (2nd and 3rd reprocessing), MODIS Aqua (Collection 5), POLDER-3 (Calibration 1) and VEGETATION-2 (Calibration 1). Simulations are compared to the actual satellite measurements.

5.1. Pre-processing of the data

The L1 data from these sensors were extracted from the DIMITRI database for the period 2002 to 2012. These data were cloud screened and quality checked (e.g.: no saturation).

The L1 data are then further selected on the basis of their geometry of acquisition which must be matched by one of the MERIS acquisitions used to invert the RPV BRDF model (see discussion in Section 4.1). This matching is achieved by computing the angular difference between the sensor illumination direction and each one of the 200 MERIS illumination direction and adding it to the angular difference between the sensor viewing direction and the MERIS viewing direction. Only acquisitions for which this sum is less 5° are selected. In the matching process, it is assumed that there is symmetry with respect to the principal plane and insensitivity to the solar azimuth angle (see discussion in Section 4.1).

The meteor data (WV and O3) accompanying L1 data and extracted in DIMITRI are substituted by the corresponding ERA-Interim data.

This results into about 300 AATSR acquisitions, 30 ATSR-2 acquisitions, 660 MERIS 2nd reprocessing acquisitions, 830 MERIS 3rd reprocessing acquisitions, 270 MODIS-A acquisitions, 250 PARASOL acquisitions and 250 VEGETATION-2 acquisitions.

5.2. Parameterisation of the simulations of the sensor observations

These acquisitions are simulated using MYSTIC. The parameterisation of the atmosphere is the same used for the inversion of the RPV model parameters from MERIS data from 2006 to 2009 (detailed in Section 3.1). The RPV BRDF model result from the interpolation at 1 nm spectral resolution the parameters obtained from each spectral band of the MERIS reference dataset (shown in Fig. 3). The ρ_0 parameter spectral variations are derived from the interpolation scheme using the 8 eigen vectors of the hyperspectral sand reflectance database. The spectral variations of the 3 other BRDF model parameters are obtained by simple spline interpolation. The BRDF model parameter interpolation is done excluding spectral band for which the water vapour and O2 absorption is significant (bands 9, 11 and 15).

Simulations are done using the ERA-Interim O3 and WV as input for each acquisition. Simulations of ATSR-2, AATSR, A-MODIS, MERIS, POLDER-3 and VEGETATION-2 in spectral bands for which the water vapour and O2 absorption is significant were not done as they cannot be confidently computed at the model spectral resolution.

5.3. The results: comparison between simulations and observations

The biases between sensor observations and the corresponding simulations are shown in Fig. 12. They are computed using a linear fit of the relative difference between observations and simulations, and computing

the bias on the 2008/01/01 resulting from this linear relationship. The corresponding linear trends are shown in Fig. 13.

It was observed in the analysis of the residual error between MERIS observations and simulations for the inversion of the RPV parameters (see Section 3.4), that the residual errors approximately follow a normal distribution. This is confirmed when simulating longer time series (~10 years). 95% CI on both the biases and the linear trends can be derived from the statistical formalism applicable to normal distributions. These 95% CI obtained on both the biases and the linear trends are directly proportional to the standard deviations of the residual difference between observations and simulations. The better the coupled surface-atmosphere model can capture the geometrical and temporal variations of the TOA signal, the smaller the 95% CI obtained on both the biases and the linear temporal trends. This explains that the 95% CI are obtained for instance on the linear trends of the MERIS data are remarkably low: below 0.05%/year for all spectral bands (see Fig. 14).

The simulated MERIS 3rd reprocessing time series extend from 2002 to 2012, significantly longer than the 4 years used for the inversion of the RPV parameters. Biases between the data and their simulation on 2008/01/01 ranging from 0 to 1% are found. These are larger than the biases found at the end of the RPV parameters inversion process on the MERIS reference dataset (see Fig. 6).

MERIS 2nd reprocessing data appear few tens of % lower than the MERIS 3rd reprocessing, with the highest changes seen in the blue. This is in line with what is expected from the change of calibration coefficients between the two reprocessings.

The linear trends in Fig. 13 obtained for MERIS have been reduced by the 3rd reprocessing change in calibration coefficients. There are small but statistically significant downwards trends for several bands in the 3rd reprocessing data with the maximum being for the 443 nm band with −0.08%/year (95% CI = ±0.03%/year).

Biases on the 2008/01/01 for AATSR range between 2% and 3% for its 3 spectral bands. Only the 560 nm band has a significant temporal trend of −0.15%/year (95% CI = ±0.05%/year).

Only data during 2002 and early 2003 are available from ATSR-2 sensors in the DIMITRI database. Biases extrapolated to the 2008/01/01 range between 13% and 16% for its 3 spectral bands (out of scale on Fig. 12). Biases are around 10% during the ATSR-2 acquisition period but they become larger due to downwards trends (see Fig. 13) in each band. These trends are however statistically not significant due to the low number of ATSR-2 acquisitions (about 30).

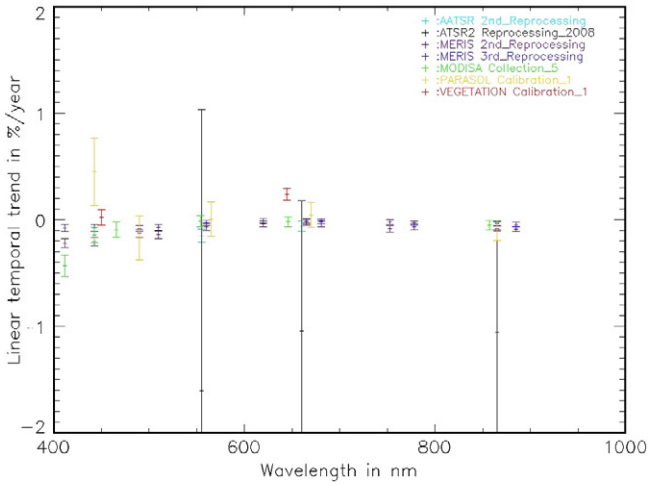

Fig. 13. The temporal trend resulting from the linear fit of the difference between observations and simulations of the AATSR, ATSR-2, MERIS, MODIS-A, POLDER-3 and VEGETATION-2 data over the period 2002–2012. Error bars are the 95% CI.

The largest bias between A-MODIS data and their simulation is found at 412 nm, −2.91% (95% CI = ±0.30%) as well as the largest downwards trend of −0.44%/year (95% CI = ±0.10%/year). All other bands have bias between 0% and −2%. The 443 nm band also has a downwards trend of −0.14%/year (95% CI = ±0.07%/year). All other bands are stable within the 95% CI.

POLDER-3 data appear about 2 to 4% below their simulations for all spectral bands. The 443 nm band shows the largest linear trend +0.45%/year (95% CI = ±0.32%/year).

VEGETATION-2 data in the 450 nm and the 550 nm bands are below their simulation by about 4% (95% CI < ±0.2%). The 550 nm band shows a positive linear trend of about +0.24%/year (95% CI about ±0.05%/year). A multivariate analysis of the 450 nm band residual error reveals a significant correlation ($r^2 = 0.27$) with the viewing zenith angle. Such values of the correlation with the VZA in this spectral region is found for no other sensor but VEGETATION-2 which rules out a surface BRDF model error. This could be explained by an across track dependency of the calibration coefficients not fully taken into account of about 2% between nadir and 40° VZA (confirmed by CNES, personal communication P. Henry).

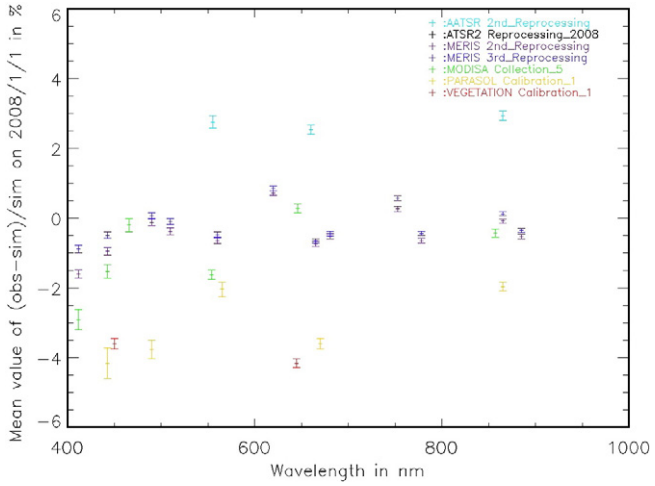

Fig. 12. The bias on the 2008/01/01 as resulting obtained from the linear fit of the time series of the relative difference between observations and simulations of the AATSR, MERIS, MODIS-A, POLDER-3 and VEGETATION-2 data over the period 2002–2012. Error bars are the 95% CI.

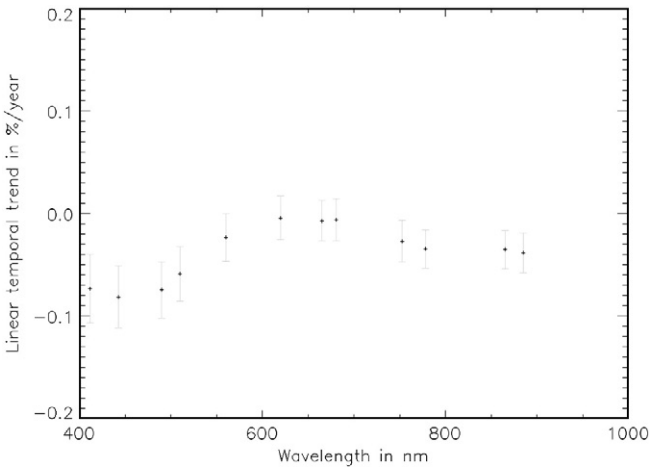

Fig. 14. The temporal trend resulting from the linear fit of the difference between observations and simulations of the MERIS 3rd reprocessing data over the period 2002–2012. Error bars are the 95% CI.

5.4. Extension of the geometrical domain of validity of the reference TOA model

The geometrical domain of validity of the surface BRDF model was cautiously restricted to the observation and illumination geometries of the 200 MERIS observations used to retrieve it. The TOA comparisons between the various sensors observations and their simulations were restricted to geometries matching any of the 200 MERIS observations, only allowing only for a 5-degree total angular difference in the geometrical matching process (see Section 5.1). To investigate the possible extension of this domain, the geometrical matching process was abandoned and simply all cloud screened and quality checked sensor observations were simulated, regardless of their geometry of acquisition. This results into approximately selecting 420 AATSR acquisitions (previously 300), 50 ATSR-2 matching acquisitions (previously 30), 770 MERIS 2nd reprocessing matching acquisitions (previously 660), 980 MERIS 3rd reprocessing matching acquisitions (previously 830), 880 MODIS-A matching acquisitions (previously 270), 1780 PARASOL matching acquisitions (previously 250) and 720 VEGETATION-2 matching acquisitions (previously 250).

The results in terms of bias between observations and simulations (on the 2008/01/01) presented in Fig. 12 are only marginally affected when all possible geometries of acquisition are considered for TOA simulations. The changes for all sensors and all spectral bands are of the order of few tenths of percent compared to the results presented in Fig. 12.

Similarly, the temporal trends presented in Fig. 13 are only marginally affected when all possible geometries of acquisition are considered for TOA simulations. The changes are all below 0.1%/year (except for PARASOL 443 nm band) compared to the results given in Fig. 13.

This demonstrates that the TOA model can be used outside the geometrical domain defined by the 200 MERIS observations that were used to retrieve the BRDF model parameters without compromising its performance.

5.5. What do all these space sensor observations tell us about the Libya-4 site?

It has been speculated that the Libya-4 BRDF might suffer from a SAA dependency due to a large scale alignment of the dunes with the wind direction. The coupled surface-atmosphere model used in the present study assumes invariance of the TOA and surface signal with respect to the SAA (but not with respect to the difference (SAA − VAA)). The multivariate analysis for the relative errors between all sensors observations and their simulation does not reveal any dependency with respect to the SAA. This is an indication that the model assumption is valid.

An implicit assumption made when analysis the temporal trends of the various bands from each sensor is that the Libya-4 site radiometry is stable. This assumption cannot be verified by comparing observations from a single sensor to their simulations. In fact, the temporal evolution of the difference between a given space sensor observations and its simulations is the result of 1) the potential radiometric temporal variability of the site BRDF and atmosphere and 2) the residual error of the degradation model of the sensor. These two contributions cannot be disentangled using the trends obtained from a single sensor. While it is clear that observations from a single instrument cannot provide alone estimates of the radiometric stability of the site, the analysis of the trends obtained from several sensors can provide indications on this stability. It is particularly interesting to note two general tendencies in the linear temporal trends obtained for all sensors in Fig. 13 (if ignoring the results obtained for ATSR-2 which bare a large uncertainty):

- Most sensor measurements are stable (within 95% CI) in the green to NIR spectral region
- The largest statistically significant trends (i.e., within 95% CI) are observed in the blue spectral region

These two facts are consistent with 1) the expected 'browning' of space instrument optics exposed to UV radiation, i.e., the faster degradation optical elements in the shorter wavelengths than in the longer wavelength, and consequently with larger residual calibration trending errors in the blue than the NIR and 2) with the assumption of radiometric stability of the Libya-4 site.

MERIS 3rd reprocessing measurements have the lowest trends across the spectrum (see Fig. 14), all trends are within 0.1%/year. Such stability can be explained by the fact that the MERIS instrument degradation model should be less prone to residual errors as 1) it has been shown that the instrument optics has degraded by at most about 5% (in the blue) over its lifetime and 2) the instrument was equipped with a first solar diffuser exposed every second week to monitor this instrument degradation and 3) a second solar diffuser exposed every 3 months was used to monitor the first diffuser degradation. Assuming the trends observed in MERIS 3rd reprocessing data (see Fig. 14) are due to the site radiometric variability only leads to an upper estimate for the Libya-4 site radiometric stability of 1%/decade over the 2002–2012 period in the spectral range covered by MERIS. It is actually likely that the residual trends in the blue are residual instrumental degradation model errors as the MERIS calibration coefficients used for the 3rd reprocessing results from the processing of the solar diffuser measurements covering only the 2002 to 2009 period and not the full mission.

5.6. Are these findings consistent with the CEOS-IVOS-WG4 findings?

In the frame of the CEOS/IVOS WG4 (working group on PICS), an intercomparison was carried out between various vicarious calibration and radiometric intercomparison methodologies. This work involved the Centre National Etudes Spatiales (CNES), ESA, the Rutherford Appelton Laboratory (RAL) and the Vlaamse instelling voor technologisch onderzoek (VITO). Intercomparisons over the sites Dome-C, Niger-2 and Libya-4 were carried on a common reference dataset composed of AATSR, A-MODIS, MERIS and POLDER-3 data spanning the 2006–2009 period (see Adriaensen et al., 2012). MERIS data were used as reference for these intercomparisons. The resulting radiometric difference between MERIS and AATSR, A-MODIS and POLDER-3 for these methodologies for the spectral band centred around 560 nm, 670 nm and 870 nm are shown in Fig. 15. These values are comparable to the statistical indicator chosen in the present study, i.e., the bias computed the 2008/01/01 following a linear fit of the time series of relative difference between observations and simulations.

6. Conclusion

A Monte Carlo vector radiative transfer model interfacing a fixed plane parallel atmosphere and a modified RPV BRDF model with 4 free parameters was used to simulate top-of-atmosphere reflectance time series over the so-called Libya-4 site. First, the model is 'calibrated' against MERIS 3rd reprocessing observations, in the MERIS 15 spectral bands, by adjusting the 4 free parameters of its surface BRDF model. These parameters were then spectrally interpolated in order to serve as input to the radiative transfer simulations of time series of observations from ATSR-2, AATSR, MODIS-A, MERIS (2nd/3rd reprocessing) and VEGETATION-2.

Depending on the spectral band considered, AATSR radiometry appears 2% to 3% above the model 'calibrated' on MERIS radiometry, MODIS is 0% to 3% below, POLDER-3 is 2% to 4% below and VEGETATION-2 about 4% below. ATSR-2 data during the 2002 to early 2003 period are almost 10% below their simulations. The presented analysis could be used to reduce the radiometric inconsistencies between these sensors. The ability of the coupled surface-atmosphere model to mimic the sensor observations allows computing accurate estimates of the temporal trends in sensor observation time series. The temporal trends obtained for all sensors against the coupled surface-atmosphere model are consistent with the expected residual errors of instrument

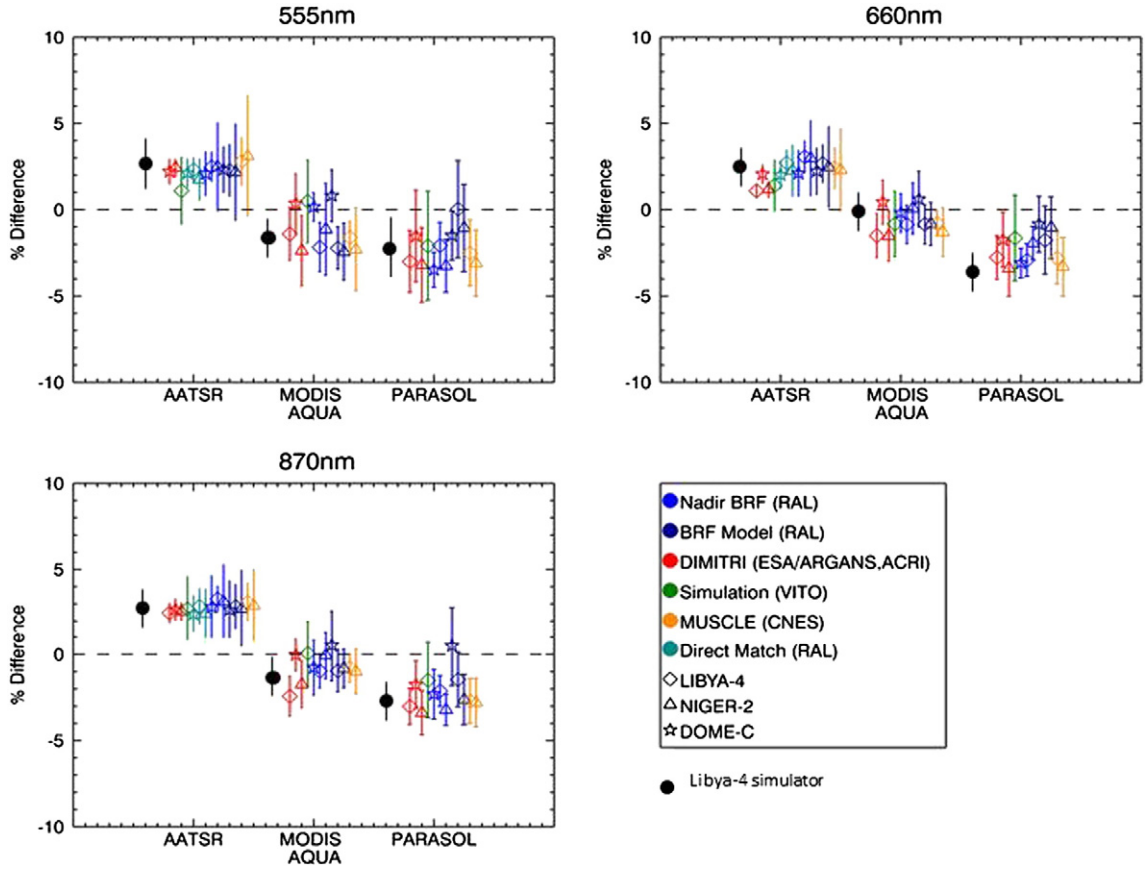

Fig. 15. Summary of the intercomparison of all methodologies participating to the CEOS/IVOS/WG4 on pseudo-invariant calibration sites showing the mean relative difference between a given sensor and MERIS 2nd reprocessing data expressed in percent. The error bar is the associated standard deviation expressed in percent. The black dots are the results obtained from the methodology presented in this paper.

degradation model used in extrapolation: larger in the blue part of the spectrum than in the NIR. This observation seems to indicate that the observed trends are linked to the sensor degradation model residual errors rather than to decadal radiometric changes of the site. The results of the present study tend to indicate that the Libia-4 site is radiometrically stable in the visible to NIR to better than 1%/decade during the 2002–2012 period, thus quantitatively confirming it is a terrestrial target particularly adequate for the assessment of the temporal stability of EO sensors. The results of the radiometric intercomparisons performed in this study compare well to the findings of the CEOS/IVOS/WG4 on PICS. It further gives confidence in these results.

Further work should be done to increase the spectral resolution of the simulations, in particular in the spectral regions where gaseous absorption from O2 and H2O dominates. The hyperspectral sand database used for the interpolation of the BRDF model should be further extended, ultimately, it should contain spectra directly obtained at the Libya-4 site. This work should also be extended to other PICS sites and also snow sites (over Antarctica and Greenland) that are also considered as pseudo-invariant.

The presented coupled surface-atmosphere model of the Libya-4 site cannot be considered as an absolute radiometric reference. The model is 'calibrated' onto the MERIS observations over the 2006–2009 period. MERIS radiometric response was characterised pre-flight and it was flown with two solar diffusers routinely exposed at different temporal frequencies that should have ensured its radiometric stability. However, it cannot be demonstrated that these devices did not suffer from any process modifying their reflective properties in the ground to space transfer or during their lifetime. MERIS, like all similar space borne instruments, suffered from a loss of radiometric traceability between its on ground characterisation and in flight operation. The model presented in this study is thus not aligned to an absolute traceable radiometric standard. Only observations over the site from a space missions like TRUTHS (Fox et al., 2011) or CLARREO (http://clarreo.larc.nasa.gov) providing measurements directly compared to an onboard traceable standard could allow such model to be 'calibrated' on a traceable and absolute radiometric standard. The model presented here could be a key element of the TRUTHS/CLARREO-like mission to transfer its radiometric standard to other space borne instruments by modelling the radiometric differences between TRUTHS/CLARREO-like observations and any other sensor observations due to geometrical and spectral response differences. This radiometric transfer would not be only possible to missions flying simultaneously to a TRUTHS/CLARREO-like mission. The radiometric stability of the pseudo-invariant calibration sites would also allow to transfer this radiometric standard to EO instrument preceding and following a TRUTHS/CLARREO-like mission. The combination of space borne traceable absolute radiometric measurements, coupled surface-atmosphere modelling and PICS will open the way to an era of increased radiometric traceability for EO measurements.

Acknowledgement

The author would like to thank J. E. Bullard and K. H. White for kindly providing the BRF measurements on sand samples from the Simpson desert (Australia), Muleshoe Dunes (USA) and the Namibian desert (Namibia).

The author would also like to thank ESA for providing the MERIS and (A)ATSR data, NASA for providing the MODIS-A data, VITO (Mol, Belgium) for providing the VEGETATION data and the CPP (CNES) for providing the PARASOL data.

Finally the author would like to thank the anonymous reviewers for their comments, questions and suggestions. They significantly contributed to improve this paper.

References

Abreu, L. W., & Anderson, G. P. (1996). The MODTRAN 2/3 report and LOWTRAN 7 model. *Contract 19628.91-C: 0132*.

Adriaensen, S., Barker, K., Bourg, L., Bouvet, M., Fougnie, B., Govaerts, Y., et al. (2012). CEOS IVOS Working Group 4: Intercomparison of vicarious calibration and radiometric intercomparison over pseudo-invariant calibration sites. available at http://calvalportal.ceos.org/cvp/web/guest/ivos/wg4

Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., & Shettle, E. (1986). AFGL atmospheric constituent profiles (0–120 km). *Tech. Rep. AFGL-TR-86-0110*. Hanscom Air Force Base, Bedford, Mass.: Air Force Geophys. Lab.

Boucher, Y., Cosnefroy, H., Petit, D., Serrot, G., & Briottet, X. (1999). Comparison of measured and modeled BRDF of natural targets. *13rd Annual International Symposium Aerosense SPIE 1999 Orlando (USA), April 05–09*.

Bourg, L., D'Alba, L., & Collagrande, P. (2008). MERIS smile effect characterisation and correction. *Tech. Note, V2.0*.

Bullard, J. E., & White, K. (2002). Quantifying iron oxide coatings on dune sands using spectrometric measurements: An example from the Simpson-Strzelecki Desert, Australia. *Journal of Geophysical Research*, *107*(B6), 2125.

Buras, R., & Mayer, B. (2011). Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: The solution. *Journal of Quantitative Spectroscopy and Radiation Transfer*, *112*, 434–447.

Cahalan, R., Oreopoulos, L., Marshak, A., Evans, K., Davis, A., Pincus, R., et al. (2005). The International Intercomparison of 3D Radiation Codes (I3RC): Bringing together the most advanced radiative transfer tools for cloudy atmospheres. *Bulletin of the American Meteorological Society*, *86*, 1275–1293.

Chander, G., Mishra, N., Helder, D. L., Aaron, D. B., Angal, A., Choi, T., et al. (2013). Applications of spectral band adjustment factors (SBAF) for cross-calibration. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(3), 1267.

Cosnefroy, H., Leroy, M., & Briottet, X. (1996). Selection of Sahara and Arabia desert sites for the calibration of optical satellite sensors. *Remote Sensing of Environment*, *58*, 101–114.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597.

deMenocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., et al. (2000). Abrupt onset and termination of the African Humid Period: Rapid climate responses to gradual insolation forcing. *Quaternary Science Reviews*, *19*, 347–361.

deMenocal, P. B., & Tierney, J. E. (2012). Green Sahara: African humid periods paced by earth's orbital changes. *Nature Education Knowledge*, *3*(10), 12.

Emde, C., Buras, R., & Mayer, B. (2011). ALIS: An efficient method to compute high spectral resolution polarized solar radiances using the Monte Carlo approach. *Journal of Quantitative Spectroscopy and Radiation Transfer*, *112*, 1622–1631.

Emde, C., Buras, R., Mayer, B., & Blumthaler, M. (2010). The impact of aerosols on polarized sky radiance: Model development, validation, and applications. *Atmospheric Chemistry and Physics*, *10*, 383–396.

Emde, C., & Mayer, B. (2007). Simulation of solar radiation during a total eclipse: A challenge for radiative transfer. *Atmospheric Chemistry and Physics*, *7*, 2259–2270.

Fox, N., Kaiser-Weiss, A., Schmutz, W., Thome, K., Young, D., Wielicki, B., et al. (2011). Accurate radiometry from space: An essential tool for climate studies. *Philosophical Transactions of the Royal Society A*, *2011*(369), 4028–4063.

Hapke, B., DiMucci, D., Nelson, R., & Smythe, W. (1996). The cause of hot spot in vegetation canopies and soils: Shadow hiding versus coherent backscatter. *Remote Sensing of Environment*, *58*, 63–68.

Hess, M., Koepke, P., & Schult, I. (1998). Optical properties of aerosols and clouds: The software package OPAC. *Bulletin of the American Meteorological Society*, *79*, 831–844.

Kuper, R., & Kröpelin, S. (2006). Climate-controlled Holocene occupation in the Sahara: Motor of Africa's evolution. *Science*, *313*, 803–807.

Lachérade, S., Fougnie, B., Henry, P., & Gamet, P. (2013). Cross calibration over desert sites: Description, methodology, and operational implementation. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(3).

Maignan, F., Breon, F. -M., & Lacaze, R. (2004). Bidirectional reflectance of Earth targets: Evaluation of analytical models using a large set of spaceborne measurements with emphasis on the Hot Spot. *Remote Sensing of Environment*, *90*, 210–220.

Mayer, B. (1999). *I3RC phase 1 results from the MYSTIC Monte Carlo model. Extended abstract for the I3RC (Intercomparison of 3D radiation codes) workshop, Tucson, Arizona, November 17–19, 1999*.

Mayer, B. (2000). *I3RC phase 2 results from the MYSTIC Monte Carlo model. Abstract for the I3RC (Intercomparison of 3D radiation codes) workshop, Tucson, Arizona, November 15–17, 2000*.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*, 308–313.

Pierluissi, J. H., & Maragoudakis, C. E. (1986). *Molecular transmission band models for LOWTRAN, AFGL-TR-86-0272, AD A180655*.

Plummer, S. E. (2008). The GLOBCARBON cloud detection system for the along-track scanning radiometer (ATSR) sensor series. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(6), 1718–1727.

Rahman, H., Pinty, B., & Verstraete, M. M. (November 20). Coupled surface-atmosphere reflectance (CSAR) model, 2, Semi empirical surface model usable with NOAA advanced very high resolution radiometer data. *Journal of Geophysical Research*, *98*(D11), 20,781–20,801.

Ricchiazzi, P., Yang, S., Gautier, C., & Sowle, D. (1998). SBDART: A research and teaching software tool for plane-parallel radiative transfer in the Earth's atmosphere. *Bulletin of the American Meteorological Society*, *79*, 2101–2114.

Smith, D. L., & Cox, C. V. (2013). (A)ATSR solar channel on-orbit radiometric calibration. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(3).

Sterckx, S., Livens, S., & Adriaensen, S. (2013). Rayleigh, deep convective clouds, and cross-sensor desert vicarious calibration for the validation for the PROBA-V. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(3).

White, K., & Bullard, J. E. (2009). Abrasion control on dune colour: Muleshoe Dunes, SW USA. *Geomorphology*, *105*(1–2), 59–66.