



**HAL**  
open science

## Automated SWIR based empirical sun glint correction of Landsat 8-OLI data over coastal turbid water

Noelia Abascal Zorrilla, Vincent Vantrepotte, Dat Dinh Ngoc, Nicolas  
Huybrechts, Antoine Gardel

► **To cite this version:**

Noelia Abascal Zorrilla, Vincent Vantrepotte, Dat Dinh Ngoc, Nicolas Huybrechts, Antoine Gardel.  
Automated SWIR based empirical sun glint correction of Landsat 8-OLI data over coastal turbid  
water. *Optics Express*, 2019, 27 (8), pp.A294. 10.1364/OE.27.00A294 . hal-02347926

**HAL Id: hal-02347926**

**<https://hal.science/hal-02347926>**

Submitted on 5 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Automated SWIR based empirical sun glint correction of Landsat 8-OLI data over coastal turbid water

NOELIA ABASCAL ZORRILLA,<sup>1,\*</sup> VINCENT VANTREPOTTE,<sup>1,2</sup> DAT DINH NGOC,<sup>3</sup> NICOLAS HUYBRECHTS,<sup>4,5</sup> AND ANTOINE GARDEL<sup>1</sup>

<sup>1</sup>Centre National de La Recherche Scientifique – CNRS, USR3456, 97334 Cayenne, French Guiana, France

<sup>2</sup>Laboratoire d'Océanologie et de Géosciences, Université du Littoral-Côte-d'Opale, CNRS, Université de Lille, 32 avenue Foch, Wimereux, France

<sup>3</sup>Space Technology Institute, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam

<sup>4</sup>Cerema Eau Mer Fleuve, 134 rue de Beauvais, 60280 Margny Les Compiègne, France

<sup>5</sup>Sorbonne Université, Laboratoire Roberval –FRE UTC – CNRS 2012, Université de Technologie de Compiègne, Centre de Recherche de Royallieu CS 60319 – 60203 Compiègne, France

\*noelia.abascal-zorrilla@cnsr.fr

**Abstract:** The optimal exploitation of the oceanic information provided by recent high spatial resolution sensors such as Landsat 8-OLI is strongly conditioned by the quality of the water reflectance signal retrieval. One main issue stands in the ability to correct water pixels for the contamination of the sun glint, which might induce a seasonal or permanent loss of data according to the latitude. The SWIR information now provided for the most recent high spatial resolution sensors was used for evaluating the sun glint level and correcting the radiative signal for its effect. This has been performed transposing historical empirical formalisms based on the NIR signal. An automated SWIR-based sun glint correction procedure was then developed using a 4-year OLI archive gathered over very turbid waters of French Guiana (227 scenes). This procedure allows the practical limitations associated with past similar empirical methods (sensitivity to water turbidity and manual image per image correction) to be overcome. While a satisfactory preservation of the information over sun glint free pixels was observed, comparison exercises based on in situ  $R_{rs}$  data gathered in sun glint affected areas emphasize the relevance of the proposed methodology (correction by a factor of 14 of the averaged bias in the  $R_{rs}$  values after removing sun glint effects). Current limitations in the applicability of this SWIR-based empirical automated method are mainly associated with the presence of high cloud coverage, thin clouds in the OLI scene or highly spatially variable marine or atmospheric signal (around 47%, 42% and 11%, respectively, of the total of 12% of failure over French Guiana OLI archive). The potential large applicability of the procedure developed in this work was eventually demonstrated over contrasted coastal environments.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

The spectral and radiometric improvement of the recent high spatial resolution satellite sensors (e.g., Landsat8-OLI, Sentinel2-MSI) now offers innovative possibilities for investigating the biogeochemical variability of inland, estuarine or coastal environments at fine spatial scale. Numerous recent works have documented the potential of such remote sensing observation for obtaining new insights into the distribution of a variety of key biogeochemical parameters including suspended particulate matter [1–4], phytoplankton concentration [5,6], or dissolved organic matter [7,8]. Besides the development of adapted bio-optical inversion algorithms, the full exploitation of high spatial resolution data for

quantitatively describing the distribution of the latter biogeochemical descriptors is strongly conditioned by the quality of the marine reflectance assessed from these sensors. While numerous efforts have been performed for adapting existing atmospheric corrections schemes to high spatial resolution data through different processing platforms (e.g. ACOLITE [1,2], SeaDAS [1,2,5]), issues related to other specific patterns still have to be specifically addressed [9]. In particular, the specular reflection of the sun on the ocean surface can represent a strong limitation in the use of ocean color satellite data [10]. While at medium latitudes sun glint contamination affects the marine reflectance only seasonally [5,11], permanent sun glint issues prevail over tropical areas [12] being therefore responsible for a huge loss of information. The different methods currently available for correcting the water reflectance for the sun glint contamination are based on different assumptions according to sensors spatial resolution [10]. For medium resolution sensors, historical standard models rely on geometric based approaches [13,14] requiring information on the sea surface state assessed from wind speed based model [15]. More recent approaches based on neural networks [16] or polynomial formulations (POLYMER [17]) have demonstrated their potential, which however still needs to be evaluated for high spatial resolution sensors.

Diverse alternative methods specifically designed for correcting high spatial resolution data for sun glint contamination have been proposed. These methods rely generally on near-infrared (NIR) based relationships used for assessing the sun glint contamination level and then correcting the marine reflectance in the visible for its effects [18–21]. In practice, these approaches assume that 1) a stable NIR-visible dependency exists over areas presenting increasing levels of sun glint and 2) the NIR information can be used as a proxy for estimating the sun glint level due to a negligible water-leaving radiance in this spectral domain. A comparison exercise by Kay et al. [10] has documented the potential of the different NIR-based correction approaches illustrating their relevance but also the limits in their practical applicability. One major limitation is related to the assumption of a negligible marine contribution in the NIR, being invalid for highly turbid environments [22]. Further, the current NIR-based procedures are based on user-defined relationships assessed manually for each individual scene to be corrected. Until now, no automated procedure has been proposed for routinely applying such empirical sun glint correction methods over series of high resolution data, thus restricting the scope of application of their latter methods to punctual or local studies. The possibility to overcome the latter issues have been recently documented by Harmel et al. [11] who proposed a semi-analytical procedure for correcting Sentinel2-MSI data for sun glint effects taking advantage of the short-wavelength infrared (SWIR) information now available for the most recent high spatial resolution sensors. The main limitation of this method however stands in the need of a priori information on the atmospheric properties.

This study aims first at evaluating the performance of simple SWIR-based empirical approaches transposing the historical methods developed based on the NIR information [19,20,23] for correcting large data set of Operational Land Imager (OLI) scenes in coastal waters. An automated empirical SWIR-based sun glint correction procedure is further proposed and applied to the OLI 4-year archive gathered over the turbid coastal waters of French Guiana while its potential is illustrated considering other contrasted coastal regions (Vietnam and Mexico).

## **2. Material and methods**

### **2.1 Study area**

The coastal waters of French Guiana (Fig. 1), permanently influenced by the Amazon river outputs, are globally characterized by high turbidity levels, with suspended particulate matter concentration reaching up to  $325 \text{ g m}^{-3}$  in the shallowest waters [24–26].

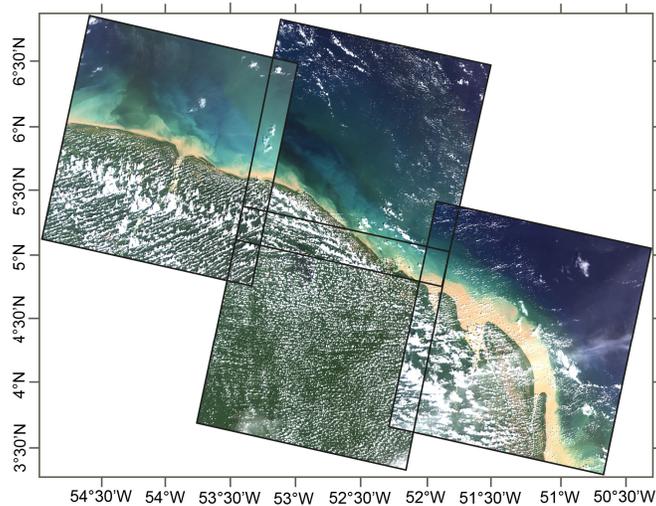


Fig. 1. Landsat-8 OLI true color images illustrating the coverage over the coastal waters of French Guiana (4 OLI scenes).

Four different OLI scenes cover this coastal area (Fig. 1). The low latitude of French Guiana (3°30'-6°30'N) combined with the solar and viewing path geometry of OLI, induce the presence of a quasi-permanent sun glint contamination towards the eastern part of each OLI scene. Standard OLI  $R_{rs}$  (Remote sensing reflectance) processing (e.g. using ACOLITE software) allow a recovery of less than 10% of the information over the 4-year OLI archive for the corresponding pixels (Fig. 2), using the standard bright pixels (e.g. cloud, land, sun glint pixels) masking threshold in ACOLITE (0.021 for the band 1609nm).

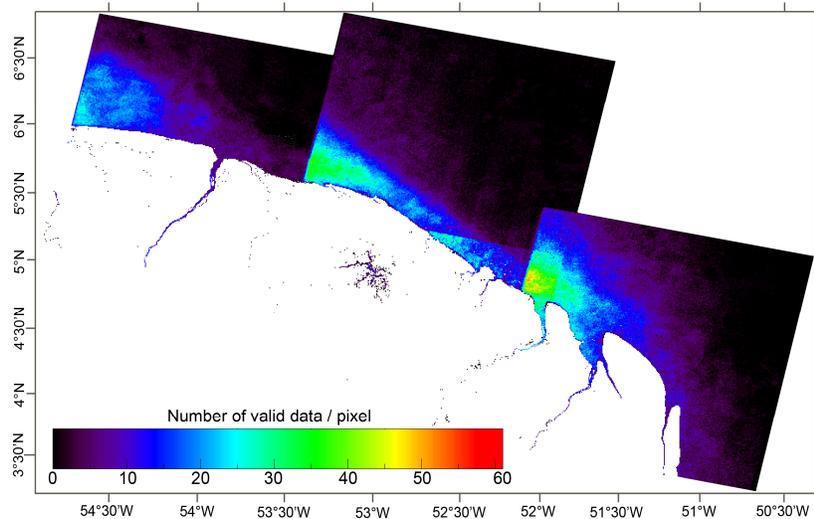


Fig. 2. Number of valid pixels gathered from the OLI archive over the coastal waters of French Guiana.  $R_{rs}$  values have been processed using ACOLITE pixel per pixel SWIR processing from April 2013 to April 2017 (265 scenes).

## 2.2 *In situ* measurements

Marine reflectance measurements have been performed over estuarine and coastal waters of French Guiana quasi permanently affected by sun glint contamination during concomitant OLI acquisition in order to assess the performance of the sun glint correction procedure

developed in the frame of this study. In practice, two in situ cruises (19/10/2016 and 04/11/2016) were performed for this matchup exercise. Above-water hyperspectral radiometric measurements (3 nm resolution) were performed in the 350-900 nm spectral domain using TRiOS radiometers.  $R_{rs}(\lambda)$  spectra were computed using the standard protocol fully described in Ruddick et al. [22]. It is worth mentioning that the  $R_{rs}$  in situ data are here assumed to be free of sun glint effects thanks to the specific geometry considered when performing above water radiometric measurements for minimizing these effects (135° azimuth angle to the sun and 45° from nadir [27]). 21 stations were kept after data quality control, taking into account the temporal interval with OLI overpass (< 90 min) and the potential impact of cloud shadow. Hyperspectral  $R_{rs}$  data were spectrally computed to the nominal OLI bands (B1 to B7) using the spectral response function reported for this sensor [27].

### 2.3 Satellite data

#### 2.3.1. Landsat 8-OLI data set

OLI level 1 data (L1T) were downloaded from the EarthExplorer platform (<http://earthexplorer.usgs.gov/>) for the four scenes covering the study area between April 2013 and April 2017 leading to a total of 265 L8/OLI images. A selection of the OLI scenes presenting at least 5% of cloud-free pixels led to the consideration of a regional data set gathering 227 OLI images.

Top-of-atmosphere (TOA) radiances,  $L_{TOA}$ , were computed from digital numbers, DN, as follows:

$$L_{TOA} = ML \cdot DN + AL \quad (1)$$

with ML (multiplicative factor, gain) and AL (additive factor, offset) values provided in the OLI metadata file. TOA reflectance ( $R_{TOA}$ ) was computed by normalizing  $L_{TOA}$  to the band averaged solar irradiance:

$$R_{TOA} = \frac{\pi L_{TOA} d^2}{F_0 \cos \theta_0} \quad (2)$$

where  $d$  is the earth-sun distance in Astronomical Units,  $\theta_0$  the sun zenith angle and  $F_0$  the band averaged extraterrestrial solar irradiance [2].

#### 2.3.2. OLI $R_{rs}(\lambda)$ processing

Landsat 8 L1T data have been processed using the ACOLITE software (acolite\_win 20170718.0) developed at RBINS (<https://odnature.naturalsciences.be/remsem/software-and-data/acolite>). The ACOLITE atmospheric correction chain allows simple and fast processing (no external measurements required) of OLI images for marine and inland water applications. Considering the high turbidity level of French Guiana coastal waters, marine reflectance ( $\rho_w(\lambda)$ ) were computed using the SWIR bands approach [22] which rely on OLI bands 6 (1609 nm) and 7 (2201 nm) for estimating the aerosol reflectance [2]. Note that ACOLITE allows users to define the aerosol properties ( $\epsilon$  parameter) on a pixel per pixel basis or using a single value for the whole scene. The current processing of OLI database by ACOLITE was performed regionally setting the threshold value on band 6 (1609 nm) to 0.032 for masking pixels contaminated by sun glint, which was established empirically based on a OLI regional sub-data set.

### 2.4 Principle of the SWIR adapted sun glint correction methods

Diverse methods for correcting the marine signal provided by historical high-resolution sensors from sun glint contamination have been documented [18–20,23,28]. All the latter

methods are based on the use of the NIR information. They estimate the level of sun glint assuming that the water-leaving radiance in this spectral domain is negligible. Therefore, any signal remaining in the NIR after atmospheric correction can be related to sun glint effects. Considering that the Fresnel reflectance affects the whole spectrum with weak variability, sun glint should affect both NIR and visible spectrum in a predictable manner (i.e. linear relationship) [10]. Such NIR-based methods can be used alternatively after or before (under the condition of a uniform atmosphere across the image) atmospherically correcting the top-of-atmosphere signal.

Globally, two kinds of NIR-based sun glint correction approaches can be considered. The first one aims at correcting the whole image for NIR-visible relationships derived from a selection of pixels presenting gradual levels of sun glint and a stable marine (and atmospheric) signal. The various approaches proposed in this frame [18–20,23] mainly differ regarding the way used to define the NIR-visible spectral dependency. Other methods, such as the one by Goodman et al. [21], perform a sun glint correction on a pixel-by-pixel basis. The latter approach, however suffers from several issues limiting its practical applicability [10]. Goodman et al. [21] further assume a unique NIR-visible relationship for correcting all the visible bands, potentially biasing the sun glint corrected signal. This method has therefore not been considered in the frame of this study.

A common limitation of all the latter historical NIR-based procedures stands in their assumption of a null marine signal in the NIR, which prevent their practical use over very turbid environments [22]. Conversely, the latter assumption of null marine signal over very turbid waters remains true when considering the SWIR domain [29], which information is now available for recent high spatial resolution sensors (Landsat 8-OLI, Sentinel 2-MSI). As a matter of fact, the advantage of the presence of SWIR bands for correcting sun glint effects have been recently documented by Harmel et al. [11], based on a semi-analytical scheme applied to Sentinel-2 data. The main principle of this method is to derive the BRDF values of the air-water interface in the SWIR bands for which the atmosphere is weakly impacted by scattering of air molecules and aerosols. Once the BRDF value is retrieved in the SWIR, the sun glint radiance in the visible-NIR bands is computed considering for a theoretical representation of the BRDF spectral dependence [11]. The main limitation of the latter automated method however stands in the need of a priori knowledge on the aerosol optical properties, which implicitly requires the use of ground-based (e.g. AERONET) or external data (e.g. CAMS global reanalysis and forecasts data sets) or relies on a coupling with an atmospheric correction procedure able to spectrally describe the aerosol optical thickness. Besides a correction of the visible reflectance, the presence of multiple information in the SWIR (e.g. 2 SWIR bands for OLI) also potentially allows the correction of NIR and SWIR signals for the sun glint effect being of interest for further applying atmospheric correction [1,2] or some bio-optical algorithms [30].

In the framework of this study, a simple transposition of three NIR and image based empirical methods: Hedley et al., Lyzenga et al., and Joyce [19,20,23] has been performed in order to correct OLI data using the SWIR information for deriving the sun glint level. Considering that a seawater signal can potentially still be found in the OLI SWIR 1 (1609 nm) channel over ultra turbid waters [29], sun glint correction was based on the information hold by the SWIR 2 channel (2201 nm). The latter three methods have fundamental similarities. As defined in the original approaches, a selection of pixels showing a variable range of sun glint has first to be performed by the user in order to derive the relationship between the SWIR 2 and the band to be corrected for sun glint effects. Considering a sun glint correction performed before atmospheric correction ( $R_{TOA}$  or digital number), the whole scene can then be corrected, using a Hedley-like approach, as follows:

$$R_{TOAi}^i = R_{TOAi} - b_i \cdot (R_{SWIR2} - offset_{SWIR2}) \quad (3)$$

where  $R_{TOA_i}^c$  is the corrected TOA reflectance of the band  $i$ ,  $R_{TOA_i}$  is the TOA reflectance of the non-corrected band  $i$ ,  $b_i$  is the spectral dependency between the band  $i$  and the SWIR 2 band,  $R_{TOASWIR2}$  is the TOA reflectance of the SWIR2 and ( $offset_{SWIR2}$ ) is the SWIR2 level for the pixels not affected by sun glint. The spectral dependence of sun glint is primarily function of the spectral shape of the direct sun light. As a matter of fact, the spectral slope of the SWIR-visible based relationships is expected to increase from the blue to the red part of the spectrum. Besides this general feature, the spectral dependence of sun glint is modulated according to atmospheric conditions and observation geometry. The definition of the offset and slope parameter in Eq. (3) has therefore to be defined on a scene specific basis. Note that the use of the latter parameter implicitly assumes the presence of a homogeneous atmosphere over a defined OLI scene, which is quite reasonable according to the relatively small size of typical OLI scenes (i.e.  $\sim 200 \times 200$  km over French Guiana coastal waters).

The definition of the spectral dependency for the method proposed by Lyzenga et al. and Joyce [20,23] is very similar to Hedley et al. [19] specifically regarding the slope term ( $b_i$  in Eq. (3)). Since it is based on the covariance information between SWIR 2 and each of the bands to be corrected, it is equivalent to the assessment of the slope parameter from a least squares linear regression [10]. The latter three methods therefore mainly differ on the definition of the offset value, which is assumed to be the minimum SWIR 2 value over the whole image (or within a sub-sample) by Hedley et al. [19] approach, while it is considered as the average value for Lyzenga et al. [20] or the mode for Joyce [23].

The method proposed in the frame of this study has been applied to the TOA information. This enables further processing of the  $R_{rs}$  signal (including areas initially affected by sun glint contamination) using the various processing chains specifically developed to atmospherically correct the OLI images (e.g. ACOLITE, SeaDas) [1,2]. Note that the Eq. (3) can also be potentially considered for correcting the NIR (using SWIR2 vs NIR relationship) and SWIR data (using SWIR1-SWIR2 reverse relationships). This potential has been also investigated in the frame of this work.

## 2.5 Statistics

Four statistical indicators have been used to characterize the impact of sun glint correction procedure on reflectance data restitution. The root mean squared difference (RMSD), the average relative percentage error (BIAS), the mean absolute error (MAE), the mean absolute relative percentage error (MARE) have been calculated based on original values as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (y_i - x_i)^2}{N}} \quad (4)$$

$$BIAS = \frac{1}{N} \sum_{i=1}^N \frac{y_i - x_i}{x_i} \cdot 100 \quad (5)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (6)$$

$$MARE = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - y_i|}{x_i} \cdot 100 \quad (7)$$

where  $N$  is the number of samples in the data set,  $x_i$  the radiometric measured (or initial) value,  $y_i$  the sun glint corrected data.

Robust linear regression procedure [31], adapted for deriving a linear regression between variables potentially affected by the presence of outliers, has been used for computing the

spectral relationships considered in the frame of this study (i.e. SWIR 2 band vs visible and NIR, SWIR 1 vs SWIR 2).

### 3. Results and discussion

#### 3.1 Illustration of the performance of the SWIR based algorithms

An illustration of the performance of the three SWIR based methods considered in the frame of this study (i.e. Hedley-SWIR, Lyzenga-SWIR and Joyce-SWIR) was first performed on a single test image over French Guiana coastal waters (LC82270562015258LGN00). Figure 3(b) shows the relationships between the SWIR 2 band and the other OLI bands over a transect [Fig. 3(a)] gathering clear waters pixels with increasing levels of sun glint. As expected, strong linear relationships between SWIR 2 and each of the visible bands can be drawn [Fig. 3(b)], confirming the spectral co-variation of the sun glint contamination over clear waters, in relation with the low spectral variability of refraction index from the visible to the SWIR spectral domain [32]. While previous studies were focusing on the definition of NIR vs visible bands relationships, it is worth to notice that strong relationships also exist between SWIR 2 and NIR and SWIR 1 data, suggesting the possibility to correct the sun glint effect also for the latter spectral bands using such empirical methods.

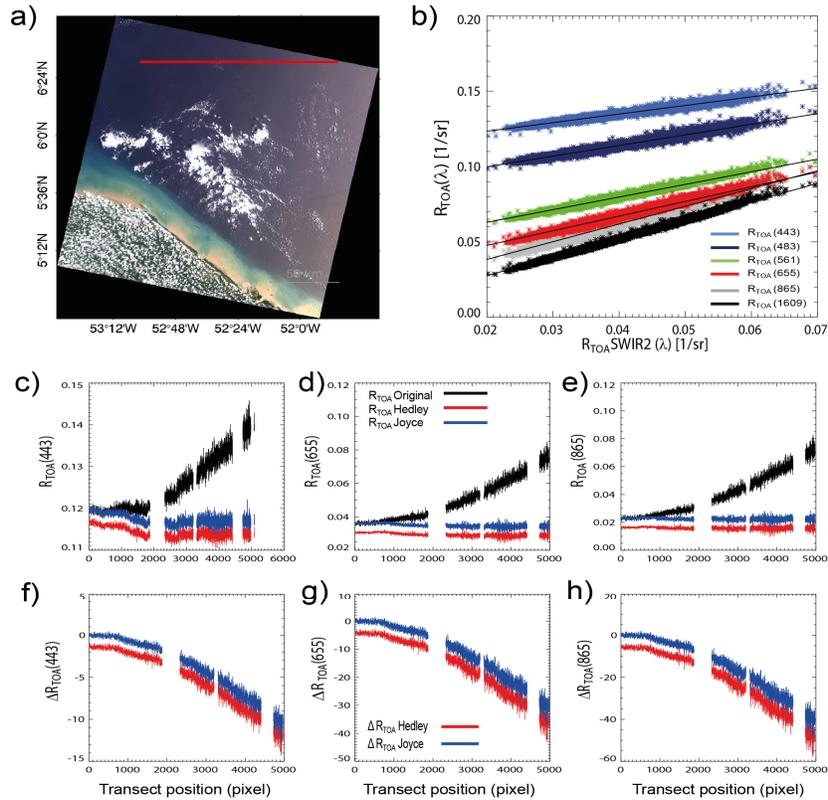


Fig. 3. Performance of the three SWIR based methods of this study over the scene LC82270562015258LGN00 (15 September 2015). Transect (a) used (true color image) to estimate the linear relationship (b) between the band SWIR 2 and the other OLI bands B1 (443 nm), B5 (655 nm) and B6 (1609 nm) and to illustrate the sun glint correction provided by the Hedley-SWIR and Joyce-SWIR. (c to e) Performance of the latter two different correction methods over  $R_{TOA}$  values at 443, 655 and 865 nm, respectively. (f to h) Relative difference (in %) between  $R_{TOA}$  values before and after the sun glint correction at 443, 655 and 865 nm, respectively. Note that the results for the Lyzenga-SWIR approach are coincident with those provided by Joyce-SWIR and are thus here omitted.

The spectral relationship between the SWIR2 band and the visible bands is not constant with slopes increasing from 0.56 at 443 nm to 1.02 at 665 nm (Table 1). The latter spectral variation can be associated with the spectral shape of direct sun glint reflectance. As mentioned previously, the terms  $b_i$  for Hedley-SWIR or  $r_{ij}$  for Lyzenga-SWIR or Joyce-SWIR show very similar values due to their statistical equivalence as already pointed out by Kay et al. [10]. Variations according to the method are more significant when considering the offset value defined by each approach (Table 1, Hedley-SWIR SWIR2 minimum over the scene, Joyce-SWIR and Lyzenga-SWIR, mode and average of SWIR2 over the selected transect respectively).

**Table 1. Comparison of the coefficients of the SWIR based relationships considered in the Hedley-SWIR, Lyzenga-SWIR and Joyce-SWIR method used for correcting the French Guiana sample OLI image in Fig. 3 from the sun glint effects.**

Band	Hedley		Lyzenga		Joyce	
	$b_i$	Min offset	$r_{ij}$	Mean offset	$r_{ij}$	Mode Offset
443	0.565		0.565		0.565	
483	0.705		0.705		0.705	
561	0.876		0.876		0.876	
665	1.028		1.028		1.028	
865	1.233		1.233		1.233	
1609	1.252	0.0055	1.252	0.012	1.252	0.0119
2201	0.794	0.0027	0.794	0.0087	0.794	0.0080

The  $R_{TOA}$  values corrected for sun glint contamination using the three methods selected are represented in Fig. 3(c), 3(d) and 3(e). Globally, the shape of the corrected transect is similar from one method to another, due to the similarity in the slope parameter used for correcting the TOA values using the relative level of sun glint detected for the SWIR2 signal. The observed variation in the amplitude of the returned  $R_{TOA}$  is more likely related to the differences in the way considered for defining the offset parameter in each method. The level of the  $R_{TOA}$  corrected using Joyce-SWIR and Lyzenga-SWIR methods is similar providing a better preservation of the initial values for the glint-free pixels when compared to the  $R_{TOA}$  returned by the Hedley-SWIR approach which tends to underestimate  $R_{TOA}$  (by up to 5% in this example) for these pixels. This underlines the potential bias induced by the use of the scene SWIR 2 actual minimum value as an offset parameter.

The three latter methods of sun glint correction tends to smooth the  $R_{TOA}$  signal as already reported by Kay et al. [10] for the similar NIR based approaches. Results in Fig. 3(f), 3(g) and 3(h) illustrate the relative importance of sun glint contamination [Eq. (9)] which maximal values increase for the illustration transect in Fig. 3(a) from 15% at 443 nm to 55% at 865 nm the latter spectral pattern being in agreement with previous observations [10].

The previous results on a single test image illustrate the potential for these simple empirical SWIR based methods to correct OLI data from sun glint effects, also underlining the importance of the way chosen to define the region of interest (ROI) considered for deriving the SWIR based parameterization.

### 3.2 Automated parameterization of the SWIR2 based relationships

Current applications of the previous NIR based empirical methods were mainly aiming at correcting a restricted amount of historical high resolution sensors data [33], based on user defined and image per image parameterizations.

Despite the overall simplicity of the general principle of the latter empirical methods, a procedure defining their practical application for routinely processing an archive of high spatial resolution satellite data has not been achieved yet. More generally, few automated procedures have been currently proposed for correcting high spatial resolution data for sun glint effects, the existing ones being mainly based on semi-analytical procedures (SWIR based: Harmel et al. [11], NIR based: Martin et al. [34], which application however requires a priori knowledge specifically on the atmospheric conditions.

The possible definition of an automated procedure was, therefore here assessed from an archive gathering 227 OLI images over French Guiana coastal waters. The application of SWIR 2 based automatic sun glint correction requires three key steps which have been specifically investigated on the basis of the latter regional data set: 1) identification of homogeneous water pixels impacted by different levels of sun glint, 2) determination of the slope of the SWIR2 based spectral relationships and 3) SWIR2 offset selection.

In practice, the automated approach proposed in the frame of this work was developed following a two-steps procedure. In a first development step, the empirical SWIR based sun glint correction was performed without any a priori knowledge of the data to be corrected in order to assess the practical issues related to the spectral relationships definition and the sensitivity of the TOA signal correction to regional variation in atmospheric and oceanic conditions. Based on these first developments, an automated procedure was then proposed to empirically correct the OLI marine reflectance signal for sun glint effects.

### 3.2.1 Region of interest (ROI) definition

The first step for applying a blind correction of sun glint effects using the considered empirical methods consists of the detection of clear water areas affected and not affected by sun glint for a defined OLI scene. It is worth to notice that the “clear waters” definition is considered here as the pixels having a constant background signal over marine pixels showing different sun glint levels. Different ways can be considered for defining the corresponding ROI. In the absence of sun glint, empirical methods using a threshold value based on the TOA Rayleigh corrected signal have been for instance proposed by Vanhellemont and Ruddick [2]. The direct use of such a threshold is here avoided considering that clear water pixels have to be selected over different levels of sun glint. Further, the presence of pixels impacted by cloud shadow, darker than surrounding waters, might also represent another issue in the clear waters detection process. In practice, the ROI used for parameterizing the SWIR-based relationships was performed considering the following successive masks:

- 1) *Water mask*: A land and cloud detection was first performed on the whole image. The method used here is based on the recent works by Ngoc et al., (in review [35]), which detects cloud and cloud shadow over glint-free pixels and which is based on the statistical analysis of the Rayleigh corrected TOA reflectance (i.e. spectral shape analysis, Hue-Saturation-Value analysis). Considering the current issues related to the retrieval of cloud shadow in the presence of sun glint, water pixels surrounding cloud regions were here arbitrarily discarded from our analysis.
- 2) *“Clear water” mask*: The detection of clear water pixels needs to be performed over the whole OLI scene (including sun glint affected pixels). As a first approach, considering the absence of a priori knowledge of sun glint correction parameters to be applied over French Guiana area, the detection of clear water pixels was here based on the ratio B4/B1. These two bands have been corrected for the SWIR 2 signal in order to roughly remove the sun glint information (see section 3.1):

$$R_{clear} = \frac{R_{TOA}^4 - C1 \cdot R_{TOA}^7}{R_{TOA}^1 - C2 \cdot R_{TOA}^7} \quad (8)$$

where factors C1 and C2, defined to balance the heterogeneity of the sun glint impact at 443 and 655 nm, were assessed empirically on a local sub data set and set to 1 and 0.6 respectively.

A relative threshold was used for assessing the “clearest waters” present in each OLI scene defined as the pixels for which the  $R_{clear}$  ratio was lower than a defined percentile value computed over the whole water pixel of the corresponding map (10% percentile as default value).

- 3) *Sun glint mask*: The definition of the sun glinted mask was based on the relative increase of the value in the SWIR 2 signal from the map average SWIR2 background signal as follows:

$$Glint\ area = where \left( \left( \frac{SWIR2 - \overline{SWIR2}_{min}}{\overline{SWIR2}_{min}} \right) \cdot 100 \right) > 15 \quad (9)$$

where  $\overline{SWIR2}_{min}$  corresponds to the averaged minimum of the SWIR 2 signal computed over the pixels showing the map lowest SWIR 2 data (10% percentile as the default threshold value).

Conversely the pixels not (or slightly) affected by sun glint were defined as:

$$No\ glint\ area = where \left( \left( \frac{SWIR2 - \overline{SWIR2}_{min}}{\overline{SWIR2}_{min}} \right) \cdot 100 \right) < 5 \quad (10)$$

The previous water, clear waters, sun glinted and non-glinton masks were then used for developing the automated sun glint procedure proposed in the frame of this work, so the detection of the glint and no glint area should be regionally adapted.

### 3.2.2 Slope estimation

The definition of the spectral slope parameter aims at relating the sun glint level observed in the SWIR2 signal to the one of the bands to be corrected. In the context of a correction applied before atmospheric correction, the slope has to be assessed considering pixels showing different levels of sun glint and spatially stable marine (usually assumed to be the case in clear waters region) and atmospheric signals. Ideally, the considered pixels should not be affected by cloud shadow pixels or thin cloud pixels, which might induce scatter in the relationships and bias the SWIR based spectral linear relationships.

In this development step, spectral slopes were computed according to the following process. A first estimation of the slope parameter of the SWIR2 based linear relationships was performed for all the bands on the clear water pixels defined from the  $R_{clear}$  ratio [Eq. (8)]. In order to check the spatial homogeneity of the marine pixels selected for deriving the slope of the SWIR2 based linear relationships for a defined OLI map, three successive ROI based on decreasing percentile value thresholds applied to  $R_{clear}$  were considered (i.e. 10, 5, and 1%). A variation lower than 5% in the computed slopes was considered as an indicator of the presence of a homogeneous marine/atmospheric spatial pattern, validating the representativeness of the slope defined.

Conversely, a variation higher than 5% between the slope values indicates a higher spatial heterogeneity in the map clearest pixels water and/or atmospheric conditions. This feature can be potentially related to an imperfect selection of clear pixels from the  $R_{clear}$  criterion. An approximate band specific correction was in this case performed for selecting the “clear waters” data on which the slope was newly assessed as follows:

$$R_{clear2} = R_{TOA}^i - C3 \cdot R_{TOA}^7 \quad (11)$$

where, C3 corresponds to the slope firstly defined considering the 10% percentile, in order to minimize the weight of potential outliers on the linear fit procedure.

A new test on the slope stability was performed as previously on considering this refined clearest water ROI selection using the same three successive percentile values (10, 5 and 1%).

The analysis of the stability of the selected clear waters has been performed on the 227 images gathered over French Guiana. The vast majority of the OLI scenes considered (e.g. from 74 to 85% from the Band 1 to the Band 7, respectively) were presenting spatially homogeneous clearest waters based on the use of the  $R_{\text{clear}}$  threshold [Eq. (8)]. The percentage of the remaining OLI scenes showing spatially stable patterns over clear waters using the approximate per band correction based on the  $R_{\text{clear}2}$  criterion was  $61 \pm 13\%$ . Consequently, only 10% of the OLI scenes gathered over French Guiana (22 images over 227) were showing a highly variable signal over clearest waters indicating that the corresponding sun glint corrected map should be considered with caution.

**Table 2. Slope values (mean, minimum, maximum, standard deviation and variation coefficient) computed from the whole data set gathered over French Guiana coastal waters (227 images) for each OLI band.**

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
Mean	0.743	0.825	0.901	0.985	1.079	1.089	0.846
Std	0.166	0.154	0.139	0.067	0.126	0.108	0.169
Min	0.039	0.085	0.220	0.675	0.563	0.802	0.175
Max	1.283	1.354	1.422	1.393	1.497	1.484	1.213
CV	22.485	18.875	15.628	6.813	10.599	9.576	19.999

Statistics associated with the slope values obtained for each bands considering the whole OLI data set gathered over French Guiana coastal waters are presented in Table 2. Slope values increase from the blue ( $0.74 \pm 0.17$ ) to the red band ( $0.99 \pm 0.067$ ), in agreement with previous NIR based empirical [36] or SWIR based theoretical works [11]. Strong variations in  $b_i$  are further observed at regional scale (227 OLI images) with a higher variability in the blue ( $CV = 22.5\%$  at 443 nm) than in the red part of the spectral ( $CV = 6.81\%$  at 655nm). Such variation in the slope of the SWIR to visible relationships can be related to a variation in the spectral shape of sun glint according to atmospheric properties and data geometry [11]. The latter results further indicate spectrally invariant methods such as Goodman et al. [21] can lead to significant errors in the sun glint correction procedure. It is worth noticing that the latter statement is in agreement with a recent study aiming to correct in situ above-water reflectance measurements for sun reflection at the sea surface [37]. As a matter of fact, the comparison of the correction performed using the average  $b_i$  value found from our data set (Table 2) or the average  $b_i$  modulated by the corresponding spectral standard deviation ( $b_i \pm \text{std}$ ) indicates that the use of regional average  $b_i$  value would induce a relative error in the returned  $R_{\text{TOA}}$  of about 20% in the blue and 100% in the red band over glint free area. This indicates the need to use image specific slope values for correcting the radiative signal for sun glint effects.

### 3.2.3 Offset definition

The offset value corresponding to the background SWIR 2 signal was assessed over areas not affected by sun glint [Eq. (10)]. The consideration of the actual SWIR2 minimum for a defined image is sensitive to noise in the SWIR2 signal potentially generating important errors in the corrected  $R_{\text{TOA}}$  [Fig. 3(c), 3(d) and 3(e)]. The minimum offset value was therefore estimated as the average value of the pixels showing a SWIR2 signal lower than the 10% percentile of a defined image in agreement with Harmel et al. [11].

Using the latter ROI, the offset definition still has to consider the actual distribution of the SWIR data that might be significantly impacted by the presence of thin clouds, not detectable using current existing cloud masks over sun glinted areas. Thin cloud free scenes globally

show a normal distribution of the lowest SWIR2 values [Fig. 4(a) and 4(c)] inducing a general convergence of the correction performed alternatively with the average or modal offset definition (i.e. Lyzenga and Joyce-SWIR2 approaches). In the presence of thin clouds a bimodal distribution of the lowest SWIR2 is instead observed inducing a potential misestimation of the SWIR2 background signal. In such cases, the first peak corresponds to the signal to be considered to set the SWIR2 background of the image while the second peak is associated with pixels impacted by thin clouds. In the presence of a bimodal distribution in the SWIR2 values, a specific selection of the pixels corresponding to the first peak was thus performed using local minima and maxima information.

The impact of the offset definition on the  $R_{TOA}$  restitution was assessed over sun glint free area. In the presence of a SWIR2 unimodal distribution, the  $R_{TOA}$  signal returned considering the mean or the mode as the SWIR2 offset value are as expected consistent, but strong differences are provided when using the actual map averaged minimum (with a maximum of 26% for the SWIR bands for the example in Fig. 4(e)).

In the presence of the bimodal distribution, the impact of a misdefinition of the SWIR2 offset on the performance of the correction is more relevant [Fig. 4(f)]. As observed previously, the use of the minimum SWIR2 value as offset parameter induces a strong modification in the signal returned over sun glint free pixels (with relative errors reaching 11% and 55% in the red and SWIR2 bands respectively). The best preservation of the signal is provided when defining the offset as the SWIR2 modal value. If the average offset value is used, results will be biased due to the partial representation of the SWIR2 background signal when selecting the pixels corresponding to the first SWIR2 peak [Fig. 4(f)].

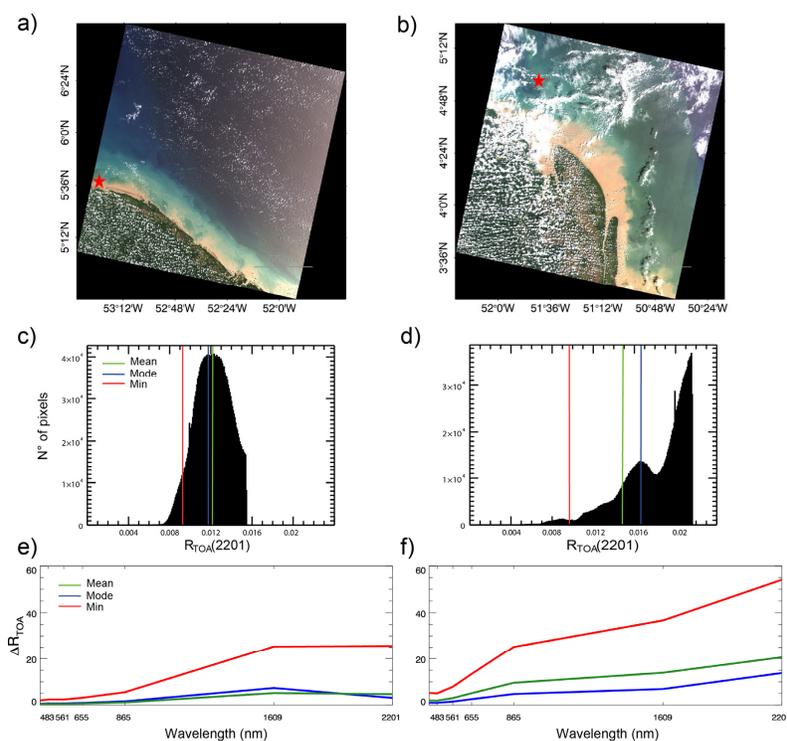


Fig. 4. Illustration of the sensitivity of the method to the offset value definition for two OLI scenes LC82270562016261LGN00 (17 September 2016) (a) and LC82260572014072 (13 March 2017) (b). Panels (c) and (d) show the two kind of distributions of SWIR2 in the non glinted area, while the relative errors (e and f) in the  $R_{TOA}$  restitution related to the use of the three different offsets: mean, mode and averaged minimum (lower than 10%) are illustrated in panels e and f for the pixels identified by a red star in the panel a and b, respectively.

The processing of the whole OLI data set over French Guiana coastal waters indicates that the offset SWIR2  $R_{TOA}$  values vary strongly at regional scale from 0.00094 to 0.0516 with an average of 0.0118. This indicates the need to consider image specific offset definition in contrast with other studies which were relying on the use of a unique regional offset parameter [36]. For the vast majority of the images (87%) the definition of the offset as the average of the glint free SWIR2 pixels was providing the best preservation of the  $R_{TOA}$  signal over sun glint free pixels. The mode value was instead found to be more appropriated for the remaining maps which showed a bimodal distribution of SWIR signal.

### 3.3 Automated sun glint correction procedure

Based on the results documented in the previous sections, an automated empirical procedure proposed here for correcting OLI data for sun glint effects (using DN as input variables) is described in Fig. 5 and consists in the application of the following main steps:

#### 1) Mask definition:

- Water pixel definition: the approach chosen for masking land and cloud pixels was here based on Ngoc et al., (in review) [35]. Note that a main issue still remains regarding the identification of thin cloud affected pixels over sun glinted areas.
- Clear waters mask: the procedure is based on the identification of the clearest water pixels (considering the relative distribution for each band) including non glinted and glinted areas which are defined from the Eq. (11) using spectral average parameters in Table 2.
- Sun glint/no glint masks: pixels affected and non-affected by sun glint are defined considering relative thresholds as defined in Eqs. (9) and (10). These thresholds might be refined regionally according to the average relative level of sun glint affecting a defined coastal region.

#### 2) Definition of the offset parameter:

The definition of the offset SWIR2 signal has to be performed for each scene individually and has to consider the actual distribution of the SWIR2 background signal over glint free pixels (uni/bimodal). In the case of a unimodal SWIR2 distribution, the average SWIR2 value has to be preferentially considered. In the presence of a bimodal distribution, generally related to the presence of thin clouds in the corresponding scene, the offset should be defined as the mode of the first peak.

#### 3) Definition of the slope parameter:

The slope parameter has to be computed for each image individually and for each band to be corrected. The clearest water pixels have first to be identified based on a relative threshold for assessing the slope of the SWIR based relationships (using robust fit procedure). At this step the stability of the TOA signal (and therefore of the atmospheric and/or marine conditions) should be assessed in order to identify the representativeness of the slope definition and thus potential uncertainties in the signal corrected for the sun glint effects (based on the use of successive relative thresholds when defining the clearest waters mask).

#### 4) Validity of the sun glint correction

The validity of sun glint correction for a defined OLI scene has to be checked on each individual band considering spectral variation in the performance of the sun glint correction procedure. In the absence of concomitant sea truth information, this can be performed by checking the preservation of the signal over glint free areas

(user defined threshold for discarding scenes for which the sun glint correction has induced a strong modification of the initial information).

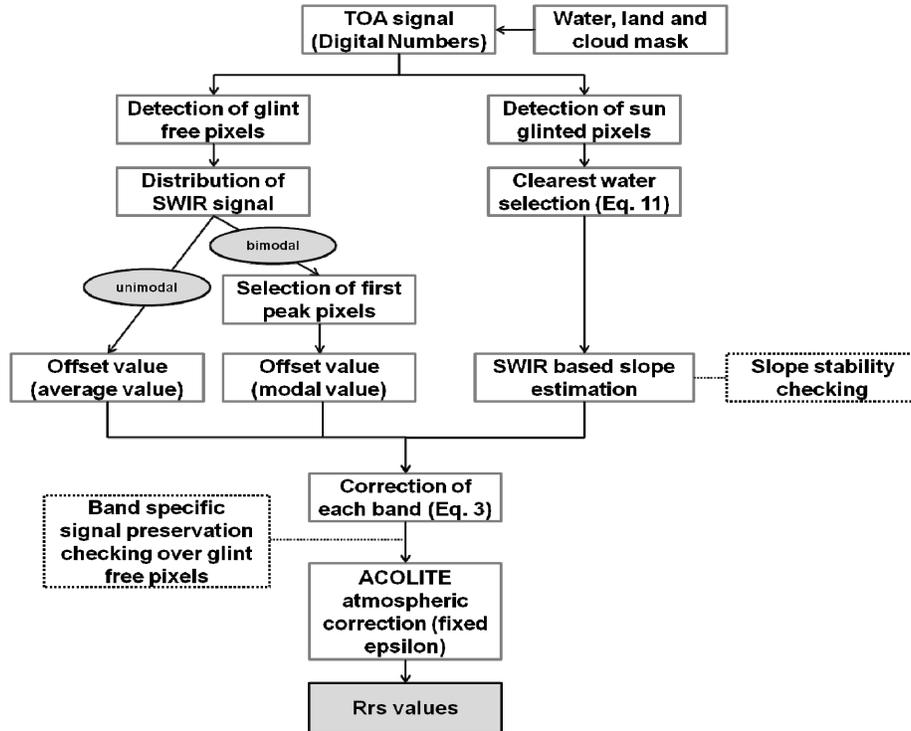


Fig. 5. Synthetic flowchart describing the different steps corresponding to the SWIR2 based automated empirical correction procedure proposed in the frame of this study.

### 3.3.1 Performance of TOA signal correction

The previous automated procedure has been applied to Digital Numbers corresponding to the 227 OLI scenes over French Guiana coastal waters. A first visual examination of the  $R_{TOA}$  distribution after sun glint correction indicates a good restitution of the spatial patterns without any apparent discontinuity between non-glintoned and sun glint corrected areas [Fig. 7(a), 7(b), 7(e) and 7(f)].

The validity of the sun glint correction procedure was statistically assessed checking the preservation of the radiometric signal after the application of the automated procedure (Fig. 5) over glint free areas. Results indicate a global good preservation of the information over sun glint free pixels with relative differences in the DN before and after correction reaching a maximum of 1.7% in the NIR, while the relative error in the visible domain show an average of 1.05% with a slight increasing pattern from the blue to the red band (Table 3).

Table 3. Statistics values illustrating the preservation of the TOA signal (DN) over sun glint free pixels after the correction the whole OLI data set gathered over French Guiana using the automated procedure defined in Fig. 5.

	Band 443	Band 483	Band 565	Band 655	Band 865
Mean (DN)	10837	9997	8725	7507	6449
RMSD	105	11	120	131	143
BIAS	0.22	0.97	1.72	1.30	1.11
MARE	0.78	0.9	1.13	1.40	1.74
MAE	86.4	91	99.1	107.8	117.6

### 3.3.2 Sun glint corrected $R_{rs}$ restitution

The SWIR based atmospheric correction approach available for atmospherically correcting OLI data over highly turbid waters [2], was here chosen for computing the marine reflectance signal over French Guiana coastal waters (section 2.3.2.). The use of this atmospheric correction scheme implicitly induces the need to correct the SWIR channels from sun glint effects, the SWIR information being indeed specifically required for assessing the aerosol properties described through the  $\epsilon$  parameter (defined here as the regression slope between the two OLI SWIR bands). Applying the sun glint correction procedure to the SWIR channel tends to smooth the SWIR signal over the whole OLI scene generating a homogeneous SWIR spatial distribution (Fig. 6). The average  $\epsilon$  values over sun glint free pixels remain unchanged when correcting OLI data from sun glint effects [Fig. 6(c)], suggesting the possible definition of  $\epsilon$  alternatively using the SWIR data before or after sun glint correction. The current results however emphasize that the consideration of sun glint corrected data tend to generate a higher noise level in the SWIR1 to SWIR2 ratio [Fig. 6(c)]. In the context of this study the initial  $\epsilon$  values were therefore considered for processing the OLI marine reflectance data.

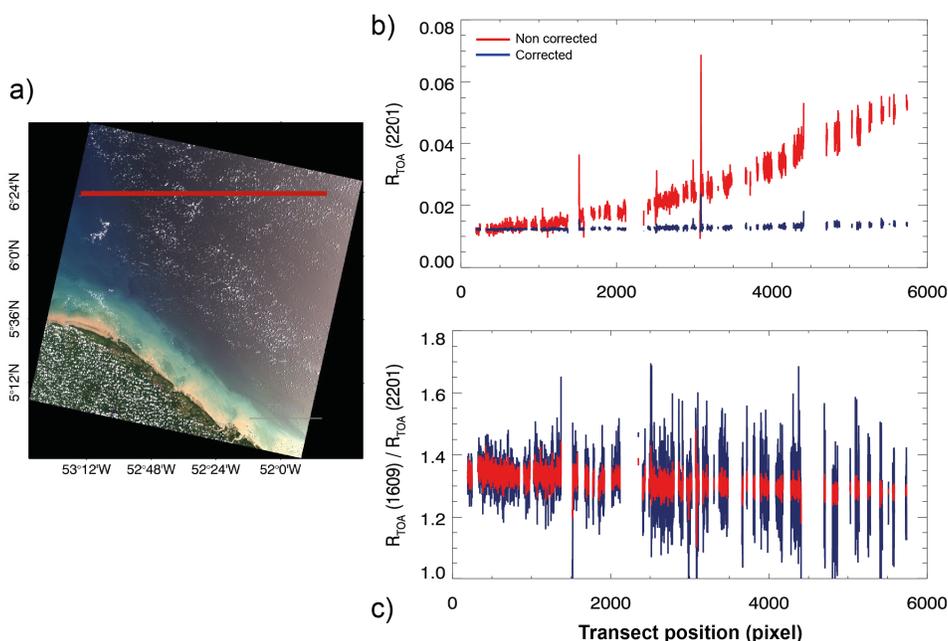


Fig. 6. Illustration of the impact of the sun glint correction process on the SWIR signal (LC82270562015338 scene – 4 December 2015) along a transect (a) in French Guiana water. Panels (b) and (c) show the distribution of the SWIR 2 and of the ratio SWIR 1 and SWIR 2 before and after correction, respectively.

Similarly to the results obtained for  $R_{TOA}$  data, a visual examination of the OLI maps before and after correcting for sun glint effects seems to indicate a satisfactory recovery of all the information associated with the pixels initially flagged as sun glinted pixels with an apparent preservation of the  $R_{rs}$  spatial patterns [Fig. 7(c), 7(d), 7(g) and 7(h)].

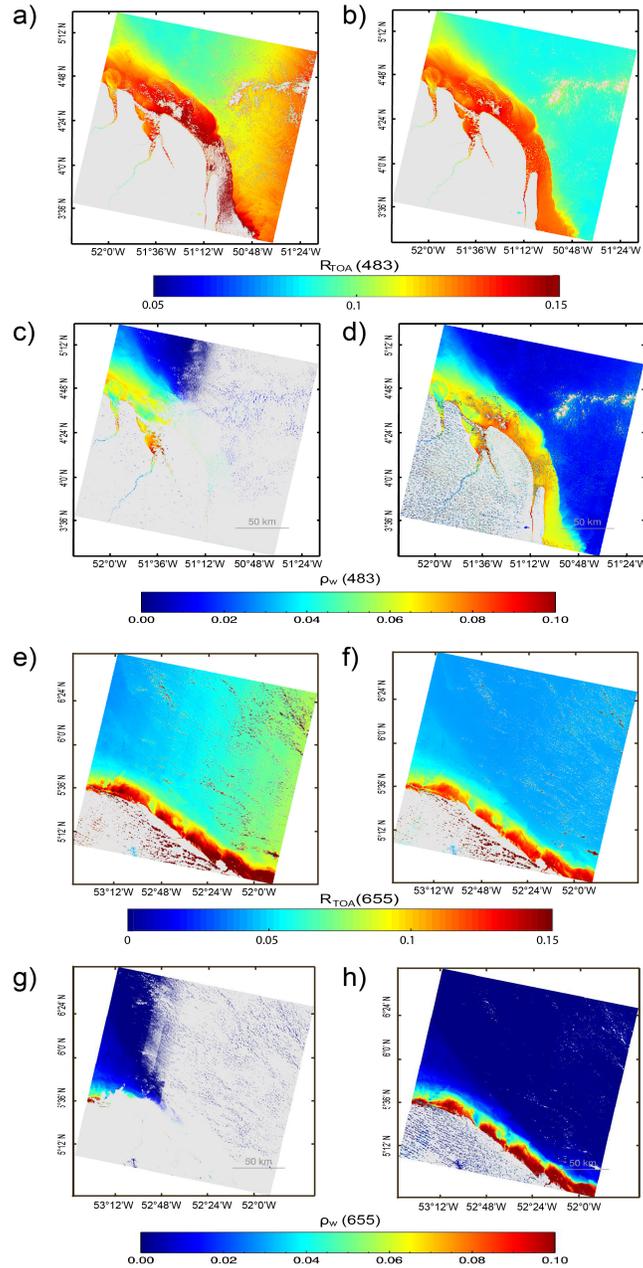


Fig. 7. Illustration over two coastal contrasted OLI scenes (coastline orientation, glint free water masses distribution) over French Guiana of the performance of the SWIR-based sun glint empirical automated procedure.  $R_{TOA}$  values before (a) and after correction (b) and  $R_{rs}$  values before (c) and after correction (d) at 483 nm for the scene LC82260572014264 (21 September 2014).  $R_{TOA}$  values before (e) and after correction (f) and  $R_{rs}$  values before (g) and after correction (h) at 655 nm for the scene LC82270562015242 (30 August 2015).

As performed previously for the  $R_{TOA}$  values, a quality control of the sun glint corrected  $R_{rs}$  was performed checking the preservation of the data before and after correction over sun glint free coastal areas (Fig. 8). Results show an overall good preservation of the marine reflectance signal whatever the spectral band considered. Similarly, spectral ratios usually

considered in various bio-optical inversion models (e.g.  $R_{rs}(443)/R_{rs}(565)$  [37,38]) are also preserved when correcting OLI images for the sun glint effects [Fig. 8(f)]. Note however that a general higher scatter is found for the  $R_{rs}$  values than for the DN with a reversal trend with wavelength, relative errors in the  $R_{rs}$  restitution increasing with decreasing wavelength, except for band 865 nm, with higher scatter than the visible bands (Table 3). As a matter of fact, when considering the whole data set over French Guiana, a better preservation of the marine signal is globally observed in the red channel (variation in coastal waters  $R_{rs}$  lower than 5% for 80% of the scenes) than in the blue one (75% of the OLI scenes at 483 nm). Regarding the  $R_{rs}$  blue/green ratio, the variation remains lower than 5% for the 84% of the scenes. The latter features indicate that modifications of the original data induced by the sun glint correction process are amplified when correcting the TOA information for the effects of the atmosphere. This is particularly true for the shortest wavelengths (i.e. 443, 483 nm) for which the marine signal is the lowest in relation to the atmospheric one, especially for turbid waters [39–41].

In some restricted cases (from 7% to 12% for band 443 and NIR band, respectively, over the 227 OLI  $R_{rs}$  images considered) a failure in the sun glint correction occurred with a modification of the non-glinted  $R_{rs}$  data higher than 20% in coastal sun glint free waters. Such peculiar cases are related to OLI scenes showing high levels of clouds and clouds shadow coverage as well as to the presence of thin clouds, which have led to a wrong parameterization of the SWIR2 based relationships (misestimation of the  $b_i$  and offset parameters). Figure 9 illustrates through two examples the possible failure in the sun glint correction procedure emphasizing the limits of the approach proposed in the frame of this study. A visual examination of the OLI scenes showing a failure in the sun glint correction process emphasizes that at regional scale 47% of these failure cases are associated with high cloud-cloud shadow coverage, 42% with the presence of thin cloud and 11% with the presence of highly heterogeneous TOA signal. In these specific OLI scenes [Fig. 9(a) and 9(b)], the presence of thin clouds is responsible for the poor performance of the sun glint correction leading to a modification of the  $R_{rs}$  for sun glint free pixels reaching about 30% in the blue and 100% in the NIR bands, respectively. While a priori detection of invalid OLI scene is difficult to achieve (e.g. 53.13% and 6.25%, of the invalid French Guiana OLI scenes being associated with a low significance level of the corresponding SWIR2-visible relationships for the bands 483 and 655 nm, respectively), a proper way to flag the occurrence of failure in the sun glint correction process should consist in checking the preservation of the signal over glint free areas. This should be performed on a spectral basis (considering the potential spectral variation in the performance of the sun glint correction) taking into account the user requirements (in terms of bands of interest and associated level of precision), which might vary according to the aimed application of sun glint corrected reflectance data. Figure 9(c) and 9(d) show situations when the marine signal was too heterogeneous for deriving a relevant linear SWIR to visible dependency due to the absence of clear (or spatially stable) waters for the corresponding map.

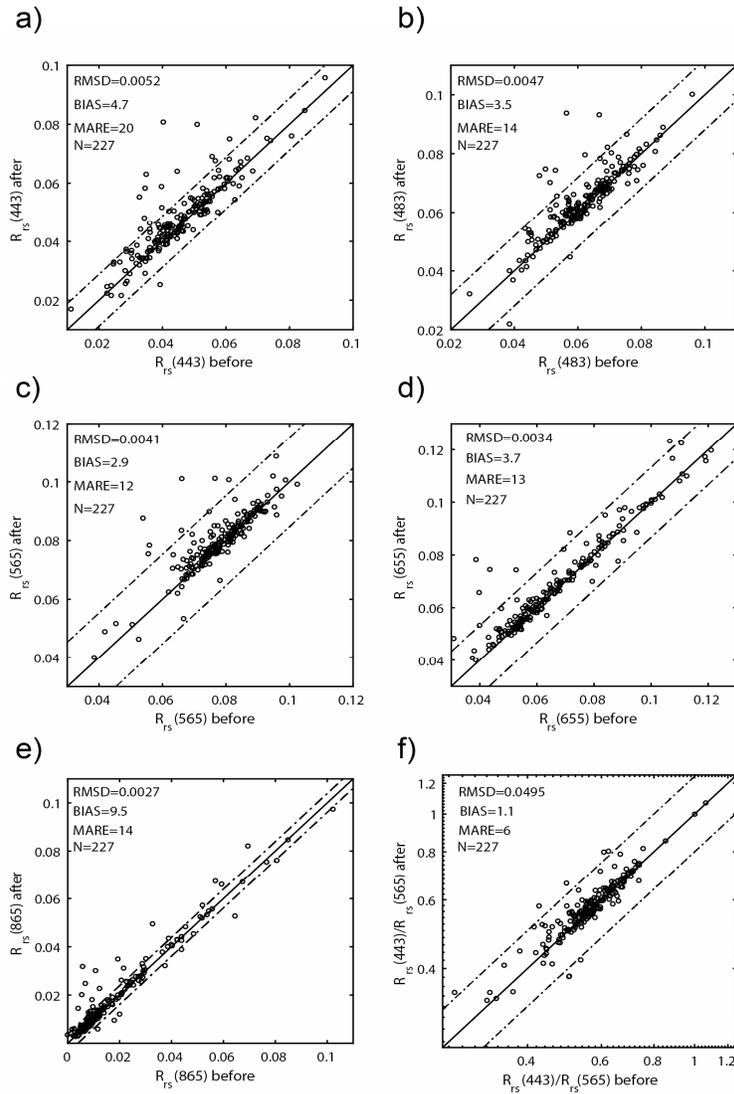


Fig. 8. Illustration of the  $R_{rs}$  preservation after the correction of the sun glint effect using average values over French Guiana coastal sun glint free areas for the 227 OLI scenes considered in the frame of this study. The solid and dashed black lines represent the 1:1 and the  $\pm 20\%$  error lines, respectively. The same representation is provided in panel f) for the  $R_{rs}(443)/R_{rs}(565)$  ratio.

The application of image-based sun glint correction is expected to impact the original signal even in areas not affected by sun glint. Considering the potential spectral variation in the precision of the sun glint correction process and in the context of further biogeochemical applications based on sun glint corrected OLI data, an examination of the  $R_{rs}$  preservation over glint free areas should be performed on a band specific basis considering the user input bands of interest associated with a defined bio-optical inversion model.

Note that the OLI data set considered in the frame of this study over French Guiana was voluntarily not restrictive regarding the level of cloud coverage (selection of the OLI scenes showing at least 5% of cloud-free pixels). The average performances here reported therefore include extreme cases, presumably not considered by the user in the context of an exploitation of OLI data for biogeochemical applications.

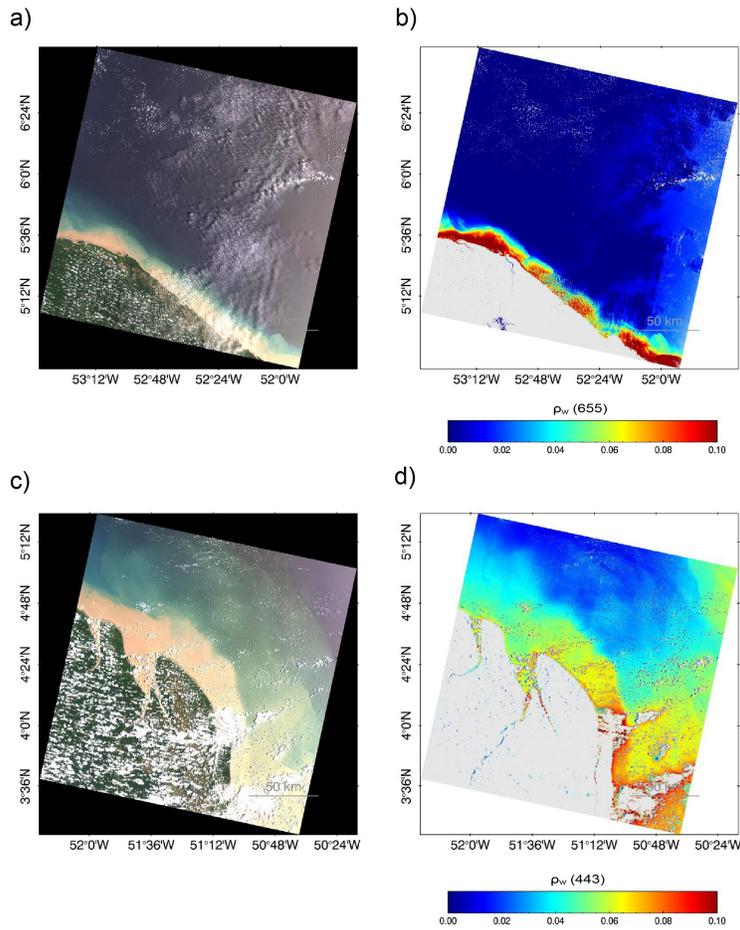


Fig. 9. OLI scenes over French Guiana coastal waters illustrating the failure in the sun glint correction procedure. True color image (a) and  $R_{rs}(655)$  (b) values after sun glint correction obtained in the case of failure due to the presence of thin clouds (LC82270562015274, 1 October 2015). True color image (c) and  $R_{rs}(443)$  values (d) after correction illustrating highly heterogeneous marine patterns related issues (LC82260572013261, 19 September 2013)

### 3.3.3 Validation

A comparison of the in situ and satellite  $R_{rs}$  values gathered simultaneously during OLI overpass was performed. Results of these  $R_{rs}$  matchup exercises performed in coastal waters of French Guiana affected by sun glint are presented in the Fig. 10 and Table 4. The different in situ stations presented moderated levels of sun glint contamination as estimated from the SWIR 2 signal using Eq. (9) (10 to 25% of increase in the  $R_{TOA}(2201)$  signal [Fig. 10(b)].

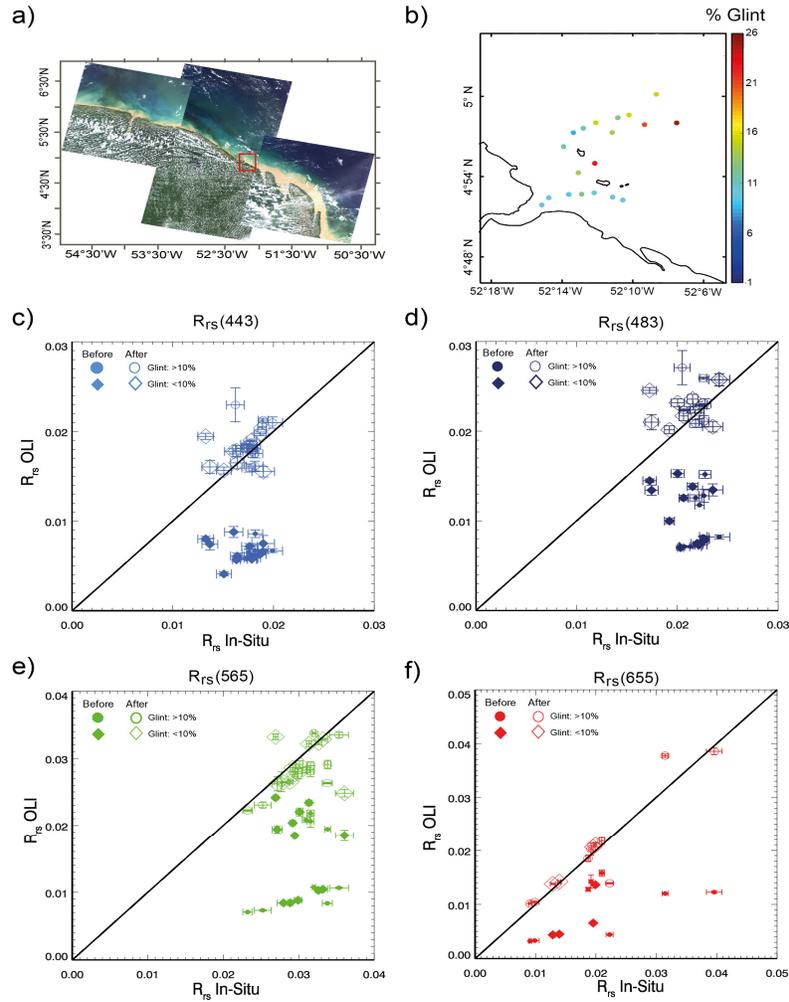


Fig. 10. Validation of the sun glint correction method through match-up data gathered in sun glint affected coastal areas of French Guiana (a). The panel b) illustrates the percentage of sun glint present in the sampled stations, estimated using Eq. (9), while the comparison between in situ and OLI  $R_{rs}$  data (before and after correction) is provided for the OLI visible bands in panels c) to f), making a distinction between the points with sun glint percentage in the SWIR2 lower or greater than 10%.

The statistics derived from the matchup data show a sharp underestimation of the OLI  $R_{rs}$  in sun glint conditions (Fig. 10, Table 4) related to an overestimation of the atmospheric contribution using the SWIR information impacted by sun glint contamination. A good agreement between in situ and ACOLITE  $R_{rs}$  is, however, obtained after applying the automated empirical SWIR based sun glint correction procedure developed in the frame of this study as illustrated by the strong decrease in the bias (globally divided by a factor of 14) and MARE (globally divided by a factor of 6) values when comparing the statistics derived for all bands before and after the correcting for the  $R_{rs}$  data for sun glint effect. Note that in spite of a relative low bias, the agreement between in situ and ACOLITE  $R_{rs}$  after sun glint correction was globally lower in the blue spectral domain ( $R^2 = 0.191$  and  $0.145$  at 443 and 483 nm respectively). This higher scattering in the blue part of the spectrum can be related to the low signal of these wavelengths in very turbid environments as those of French Guiana (SPM reaching up to 1200 mg/l for the current matchups data with an average of  $57.73 \pm$

283.43 mg/l). Conversely, a higher agreement with the  $R_{rs}$  measured in situ was found in the green and red part of the spectrum where the marine signal is stronger ( $R^2 = 0.493$  and  $0.937$  at 561 and 655 nm, respectively). Further, data presenting a relatively restricted amount of sun glint contamination are globally well preserved (no systematic over correction [Fig. 10]). It is worth mentioning that  $R_{rs}$  matchup results are rarely close to perfect due to limitation inherent to the radiometric in situ measurements [37], as well as to atmospheric correction procedures [1,2]. Taking into account these limitations, these results although derived from a quite restricted matchup data set tend to confirm the overall good performance of the sun glint correction procedure proposed in this work as well as of the ACOLITE atmospheric correction over such turbid waters.

**Table 4. Statistics of the matchup comparison between in situ and OLI-ACOLITE  $R_{rs}$  data (before and after sun glint correction).**

	Band 443		Band 483		Band 561		Band 655	
	Before	After	Before	After	Before	After	Before	After
RMSD	0.011	0.0025	0.011	0.0028	0.016	0.0037	0.012	0.0031
Correction factor RMSD		4.40		3.93		4.32		3.87
BIAS(%)	-61.6	4.9	-50.1	7.1	-44.9	-3.9	-54.6	2.1
Correction factor BIAS		12.68		7.06		11.63		26.52
MARE(%)	61.6	10.4	50.1	10.5	49.9	8.5	54.6	8.9
Correction factor MARE		5.92		4.77		5.88		6.11
$R^2$	-0.075	0.191	0.090	0.145	0.134	0.493	0.325	0.937
slope	-0.059	0.3082	-0.535	0.208	0.266	0.546	0.321	0.949

### 3.3.4 Illustration for the performance of the SWIR2 automated sun glint correction over contrasted coastal waters

The potential of the automated empirical procedure summarized in Fig. 5 for application over contrasted coastal sites (in terms of turbidity and sun glint coverage) is here illustrated considering two different OLI scenes.

Specifically, a first image gathered in coastal waters of Mexico ( $17^{\circ}50'-19^{\circ}48'N$  and  $93^{\circ}24'-95^{\circ}-32'W$ , LC08\_L1TP\_023047\_20170713\_20170726\_01\_T1) provides an example of sun glint correction over low-moderate turbid waters atmospherically corrected using the NIR information (ACOLITE NIR formalism) to the difference of the development data set (ACOLITE SWIR formalism). The second illustration scene depicts Vietnamese turbid coastal waters ( $9^{\circ}06'-11^{\circ}12'N$  and  $106^{\circ}28'-108^{\circ}30'E$ , LC08\_L1TP\_124053\_20170615\_20170628\_01\_T1, ACOLITE SWIR processing) and differs from the coastal waters of French Guiana regarding the part of clear water affected by the sun glint due to different shoreline orientation (east–west vs north–south, respectively).

Note that the ROI definition and SWIR based correction were performed without regionally adapting the procedure proposed in the frame of this work (e.g. the clear waters pixels definition using the Eq. (11) and coefficients in Table 2).

Results of the sun glint correction applied to these two sample maps are provided in Fig. 11 and tend to indicate the potential large applicability of the method here proposed with an apparent good restitution of the  $R_{TOA}$  and  $R_{rs}$ , signal in non glinted areas and a satisfying recovery of the information over sun glinted pixels.

The presence for both scenes in Fig. 11 of other issues in the  $R_{rs}$  spatial patterns depicted from OLI data, such as band striping, are clearly visible. The correction of these spatial artifacts requires special pre-processing steps that are beyond the scope of this paper.

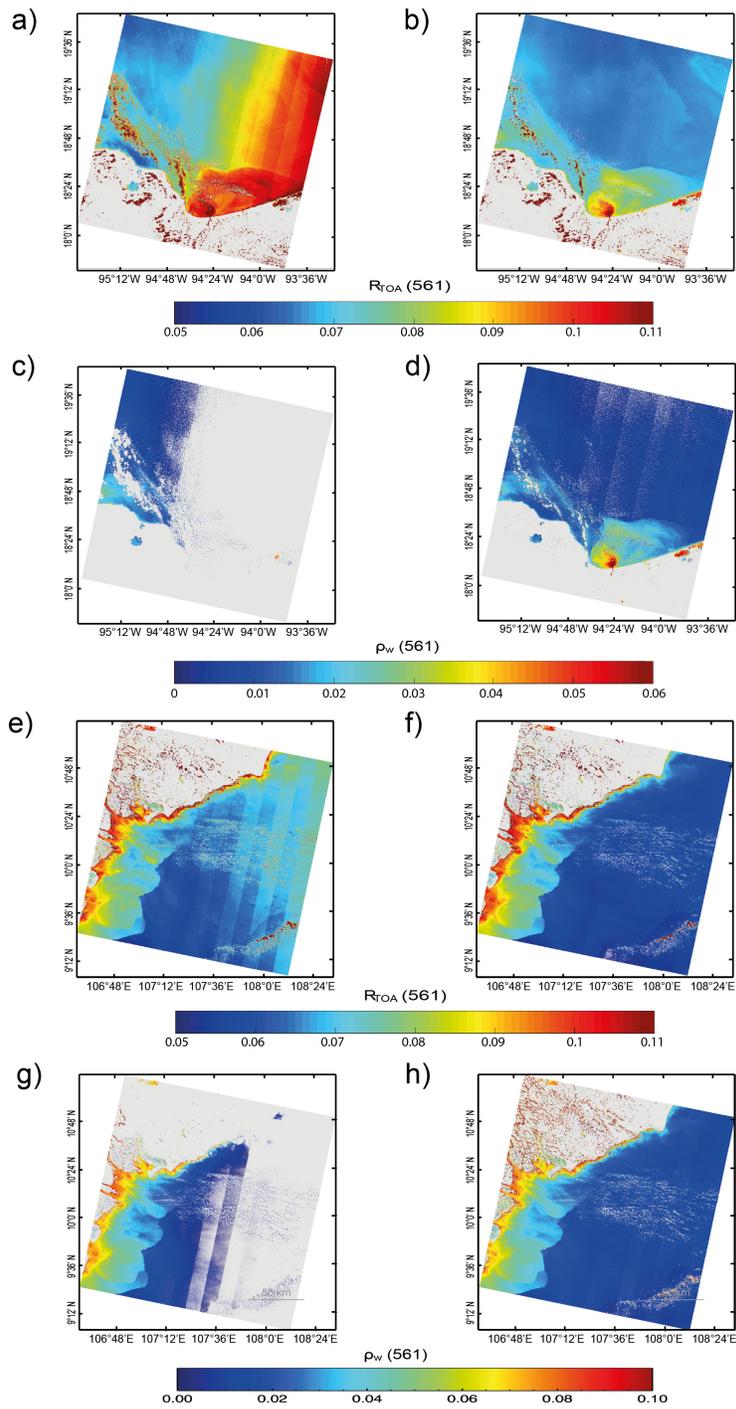


Fig. 11. Illustration of the potential applicability of the automated procedure developed over French Guiana coastal waters over contrasted coastal sites in Mexico (a to d, LC08\_L1TP\_023047\_20170713\_20170726\_01\_T1) and in Vietnam (e to f, LC08\_L1TP\_124053\_20170615\_20170628\_01\_T1). Values of  $R_{TOA}$  and  $R_{rs}$  for band 561nm before (a,c,e,g) and after (b,d,f,h) sun glint correction. Land and cloud masks are represented in grey.

## 4. Conclusions

Water pixel contamination from sun glint effects represents a main issue to overcome in order to fully exploit the potential now offered by recent high spatial resolution sensors for gathering detailed information on biogeochemical properties of coastal water masses. The performance of simple empirical SWIR based methods was evaluated in the frame of this work considering a 4-year OLI archive (227 OLI scenes) gathered over the low latitude coastal waters of French Guiana, which are permanently affected by sun glint contamination.

This study first provides evidences of the interest of such empirical approaches which are not relying, to the difference of other proposed methodologies [11,34], on any theoretical step nor on a priori information on the atmospheric conditions.

This work has also emphasized the possibility to overcome the two main limitations in the applicability of historical empirical sun glint corrections schemes initially based on the use of the NIR signal. First, the results have illustrated the advantage of the availability for the most recent sensors, such as OLI, of information in the SWIR spectral domain, where the marine information is negligible, for now correcting the marine signal over coastal waters including the most turbid ones. Further, the developments performed based on the considered regional OLI data set have demonstrated the possibility to routinely apply SWIR based empirical sun glint correction approaches. The proposed automated sun glint correction procedure, which aims at correcting the water signal before atmospheric correction, mainly relies on the definition of adapted masks (clearest waters, glinted and non-glinted pixels) based on scene specific relative statistical descriptors.

The overall performance of the proposed methodology is satisfactory as illustrated by the global good recovery of the signal over the whole sun glint contaminated area, as also confirmed from the match ups exercises performed in French Guiana coastal waters. Our results have further demonstrated that the current image based correction procedure allows the marine signal over glint free pixels to be preserved. The main limitation of the current automated empirical approach is related to the presence of thin clouds and cloud shadows (which are not detected over glinted areas) biasing the estimation of the SWIR based relationships parameters. Further developments should therefore aim at coupling the current sun glint correction approach with cloud masking procedures in order to optimize the identification of the water pixels over the whole map and the recovery of the Rrs information for biogeochemical applications.

Another main limitation of the proposed procedure is the presence over a defined OLI scene of a high heterogeneity in the TOA signal associated with the highly spatially variable atmospheric and/or water patterns. While an a priori definition of the level of homogeneity suitable for performing SWIR based empirical sun glint corrections is difficult to achieve, a quality control of the sun glint correction can be easily performed on an image basis checking the preservation of the Rrs signal over glint free areas. This should be performed spectrally considering the user's bands of interest and requirement regarding the level of precision in the marine signal restitution.

## Funding

Centre National d'études spatiales (CNES) (HARDECOT TOSCA); Direction de l'environnement, de l'aménagement et du logement (DEAL) ; Grand Port Maritime de Guyane.

## Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

## References

1. Q. Vanhellemont and K. Ruddick, "Turbid wakes associated with offshore wind turbines observed with Landsat 8," *Remote Sens. Environ.* **145**, 105–115 (2014).

2. Q. Vanhellemont and K. Ruddick, "Advantages of high quality SWIR bands for ocean colour processing: examples from Landsat-8," *Remote Sens. Environ.* **161**, 89–106 (2015).
3. B. Han, H. Loisel, V. Vantrepotte, X. Mériaux, P. Bryère, S. Ouillon, D. Dessailly, Q. Xing, and J. Zhu, "Development of a semi-analytical algorithm for the retrieval of suspended particulate matter from remote sensing over clear to very turbid waters," *Remote Sens.* **8**(3), 211 (2016).
4. Z. Qiu, C. Xiao, W. Perrie, D. Sun, S. Wang, H. Shen, D. Yang, and Y. He, "Using Landsat 8 data to estimate suspended particulate matter in the Yellow River estuary," *J. Geophys. Res. Oceans* **122**(1), 276–290 (2017).
5. B. A. Franz, S. W. Bailey, N. Kuring, and P. J. Werdell, "Ocean color measurements with the Operational Land Imager on Landsat-8: implementation and evaluation in SeaDAS," *J. Appl. Remote Sens.* **9**(1), 096070 (2015).
6. E. Watanabe, J. Onodera, N. Harada, M. N. Aita, A. Ishida, and M. J. Kishi, "Wind-driven interannual variability of sea ice algal production in the western Arctic Chukchi Borderland," *Biogeosciences* **12**(20), 6147–6168 (2015).
7. E. T. Slonecker, D. K. Jones, and B. A. Pellerin, "The new Landsat 8 potential for remote sensing of colored dissolved organic matter (CDOM)," *Mar. Pollut. Bull.* **107**(2), 518–527 (2016).
8. T. Kutser, C. Verpoorter, B. Paavel, and L. J. Tranvik, "Estimating lake carbon fractions from remote sensing data," *Remote Sens. Environ.* **157**, 138–146 (2015).
9. N. Pahlevan, J. R. Schott, B. A. Franz, G. Zibordi, B. Markham, S. Bailey, C. B. Schaaf, M. Ondrusek, S. Greb, and C. M. Strait, "Landsat 8 remote sensing reflectance (Rrs) products: Evaluations, intercomparisons, and enhancements," *Remote Sens. Environ.* **190**, 289–301 (2017).
10. S. Kay, J. D. Hedley, and S. Lavender, "Sun glint correction of high and low spatial resolution images of aquatic scenes: A review of methods for visible and near-infrared wavelengths," *Remote Sens.* **1**(4), 697–730 (2009).
11. T. Harmel, M. Chami, T. Tormos, N. Reynaud, and P. A. Danis, "Sun glint correction of the Multi-Spectral Instrument (MSI)-SENTINEL-2 imagery over inland and sea waters from SWIR bands," *Remote Sens. Environ.* **204**, 308–321 (2018).
12. P. Jagalingam, B. J. Akshaya, and A. V. Hegde, "Bathymetry mapping using landsat 8 satellite imagery," *Procedia Eng.* **116**, 560–566 (2015).
13. M. Wang and S. W. Bailey, "Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products," *Appl. Opt.* **40**(27), 4790–4798 (2001).
14. L. Bourg, F. Montagner, V. Billat, and S. Belanger, "Sun glint flag algorithm," MERIS ATBD 2.13, version 4.3 (7 July 2011). <https://earth.esa.int/instruments/meris/atbd/>
15. C. S. Cox and W. Munk, "Slopes of the sea surface deduced from photographs of sun glitter," *Bull. Scripps Inst. Oceanogr.* **6**, 401–488 (1956).
16. B. Heim, E. Abramova, R. Doerffer, F. Günther, J. Hölemann, A. Kraberg, H. Lantuit, A. Loginova, F. Martynov, P. P. Overduin, and C. Wegner, "Ocean colour remote sensing in the southern Laptev sea: evaluation and applications," *Biogeosciences* **11**(15), 4191–4210 (2014).
17. F. Steinmetz, P.-Y. Deschamps, and D. Ramon, "Atmospheric correction in presence of sun glint: application to MERIS," *Opt. Express* **19**(10), 9783–9800 (2011).
18. E. J. Hochberg, M. J. Atkinson, A. Apprill, and S. Andréfouët, "Spectral reflectance of coral," *Coral Reefs* **23**, 84–95 (2004).
19. J. D. Hedley, A. R. Harborne, and P. J. Mumby, "Technical note: simple and robust removal of sun glint for mapping shallow-water benthos," *Int. J. Remote Sens.* **26**(10), 2107–2112 (2005).
20. D. R. Lyzenga, N. P. Malinas, and F. J. Tanis, "Multispectral bathymetry using a simple physically based algorithm," *IEEE Trans. Geosci. Remote Sens.* **44**(8), 2251–2259 (2006).
21. J. A. Goodman, Z. Lee, and S. L. Ustin, "Influence of atmospheric and sea-surface corrections on retrieval of bottom depth and reflectance using a semi-analytical model: a case study in Kaneohe Bay, Hawaii," *Appl. Opt.* **47**(28), F1–F11 (2008).
22. K. G. Ruddick, V. De Cauwer, Y. Park, and G. Moore, "Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters," *Limnol. Oceanogr.* **51**(2), 1167–1179 (2006).
23. K. E. A. Joyce, "Method for mapping live coral cover using remote sensing". Ph.D. thesis, University of Queensland: Brisbane, Australia (2004).
24. H. Loisel, X. Mériaux, A. Poteau, L. F. Artigas, B. Lubac, A. Gardel, J. Caillaud, and S. Lesourd, "Analyze of the inherent optical properties of French Guiana coastal waters for remote sensing applications," *J. Coast. Res. SI* **56**, 1532–1536 (2009).
25. V. Vantrepotte, H. Loisel, X. Meriaux, G. Neukermans, D. Dessailly, C. Jamet, E. Gensac, and A. Gardel, "Seasonal and inter-annual (2002-2010) variability of the suspended particulate matter as retrieved from satellite ocean color sensor over the French Guiana coastal waters," *J. Coast. Res. SI* **64**, 1750–1754 (2011).
26. V. Vantrepotte, H. Loisel, D. Dessailly, and X. Mériaux, "Optical classification of contrasted coastal waters," *Remote Sens. Environ.* **123**, 306–323 (2012).
27. C. D. Mobley, *Light and water: radiative transfer in natural waters* (Academic Press, 1994).
28. J. A. Barsi, K. Lee, G. Kvaran, B. L. Markham, and J. A. Pedely, "The spectral response of the Landsat-8 operational land imager," *Remote Sens.* **6**(10), 10232–10251 (2014).
29. T. Kutser, E. Vahtmäe, and J. Praks, "A sun glint correction method for hyperspectral imagery containing areas with non-negligible water leaving NIR signal," *Remote Sens. Environ.* **113**(10), 2267–2274 (2009).
30. W. Shi and M. Wang, "An assessment of the black ocean pixel assumption for MODIS SWIR bands," *Remote Sens. Environ.* **113**(8), 1587–1597 (2009).

31. E. Knaeps, K. G. Ruddick, D. Doxaran, A. I. Dogliotti, B. Nechad, D. Raymaekers, and S. Sterckx, "A SWIR based algorithm to retrieve total suspended matter in extremely turbid waters," *Remote Sens. Environ.* **168**, 66–79 (2015).
32. P. Legendre and L. Legendre, "Numerical Ecology, Volume 24," *Developments Environ. Model.* **24**, 870 (1988).
33. G. Doxani, M. Papadopoulou, P. Lafazani, M. Tsakiri-Strati, and E. Mavridou, "Sun glint correction of very high spatial resolution images," *Thales, Honor Prof. Emeritus Michael E. Contadakis* 329–340 (2013).
34. J. Martin, F. Eugenio, J. Marcello, and A. Medina, "Automatic sun glint removal of multispectral high-resolution WorldView-2 imagery for retrieving coastal shallow water parameters," *Remote Sens.* **8**(1), 1–16 (2016).
35. D. Dinh Ngoc, H. Loisel, C. Jamet, V. Vantrepotte, L. Duforêt-Gaurier, C. Doan Minh, A. Mangin "Coastal and inland water pixels extraction algorithm (WiPE) from high spatial resolution optical sensors OLI on Landsat 8 and MSI on Sentinel-2" (in review).
36. C. Hu, "An empirical approach to derive MODIS ocean color patterns under severe sun glint," *Geophys. Res. Lett.* **38**(1), 1–5 (2011).
37. P. M. M. Groetsch, P. Gege, S. G. H. Simis, M. A. Eleveld, and S. W. M. Peters, "Validation of a spectral correction procedure for sun and sky reflections in above-water reflectance measurements," *Opt. Express* **25**(16), A742–A761 (2017).
38. J. E. O'Reilly, S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. R. McClain, "Ocean color chlorophyll algorithms for SeaWiFS," *J. Geophys. Res.* **103**(C11), 24937–24953 (1998).
39. K. L. Carder, F. R. Chen, Z. P. Lee, S. K. Hawes, and D. Kamykowski, "Algorithms for chlorophyll a and absorption with bio-optical cases  $R_{rs}(0^+) = Q(\bullet) b$ ," **104**, 5403–5421 (1999).
40. E. F. Vermote and S. Kotchenova, "Atmospheric correction for the monitoring of land surfaces," *J. Geophys. Res. Atmos.* **113**(D23), 1–12 (2008).
41. J. Ju, D. P. Roy, E. Vermote, J. Masek, and V. Kovalskyy, "Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods," *Remote Sens. Environ.* **122**, 175–184 (2012).