Sun glint correction with an inherent optical properties data processing system

Jun Chen, Xianqiang He, Zhongli Liu, Nan Lin, Qianguo Xing & Delu Pan

To cite this article: Jun Chen, Xianqiang He, Zhongli Liu, Nan Lin, Qianguo Xing & Delu Pan (2021) Sun glint correction with an inherent optical properties data processing system, International Journal of Remote Sensing, 42:2, 617-638, DOI: 10.1080/01431161.2020.1811916

To link to this article: https://doi.org/10.1080/01431161.2020.1811916

Published online: 18 Nov 2020.

Article views: 186

View related articles

View Crossmark data
Sun glint correction with an inherent optical properties data processing system

Jun Chen\textsuperscript{a,b}, Xianqiang He\textsuperscript{b}, Zhongli Liu\textsuperscript{a}, Nan Lin\textsuperscript{a}, Qianguo Xing\textsuperscript{c} and Delu Pan\textsuperscript{b}

\textsuperscript{a}School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an, China; \textsuperscript{b}State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources of the People's Republic of China, Hangzhou, China; \textsuperscript{c}Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, China

\textbf{ABSTRACT}

Frequent and extensive sun glint is a serious obstacle to real-time monitoring of ocean colour anomalies. We semi-analytically adjust an inherent optical properties (IOPs) data processing system for the open ocean to correct for sun glint (called ‘IDAS-SGC’) and for remote sensing reflectance ($R_s$) and IOPs retrievals. Tests with synthetic data validated the effectiveness of our algorithm in deriving ocean colour data from severely glint-contaminated images to produce high quality images. Evaluating results from single mission images suggested that our approach provides spatially smooth and consistent ocean colour products from both severe glint and glint-free regions for Visible Infrared Imaging Radiometer Suite and for Medium Resolution Spectral Imager II instruments. Specifically, complete coverage of circulation-caused ocean colour anomalies can be recovered from a single severe sun glint image. Comparing multi-mission images found that the inter-mission consistency for IDAS-SGC $R_s$ in sun glint regions is comparable with the inter-mission consistency in glint-free regions. Furthermore, we evaluate the performance of the Cox-Munk algorithm for sun glint estimation, and we find that our IDAS-SGC algorithm is more effective than the Cox-Munk algorithm in deriving $R_s$ products from severe sun glint regions due to the absence of accurate real-time wind data. Our results suggest that the IDAS-SGC algorithm obtains meaningful ocean colour products from sun glint-contaminated images of the open oceans.

\textbf{1. Introduction}

A series of ocean colour satellites launched in the last 50 years singular create an essential continuous climate dataset (Bojinski et al. 2014; Clark 2007; Zibordi et al. 2015). These satellites support disaster and climate research applications and related policy development. For biological oceanographic research, the satellites provide timely and up-to-date surface biomass data and communication (Lee et al. 2018; Sathyendranath et al. 2017; Jeremy et al. 2018; Kramer, Roesler, and Sosik 2018; Platt, Fuentes-Yaco, and Frank 2003). To fulfill the potential of these satellites, efforts to improve the overall quality and accuracy of ocean colour products are made continually (Hooker 2003). Since the launch
of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), much work has focused on further improving the quality of its ocean colour products (Mélin et al. 2017; Morel and Bernard 2008). In general, all improvements have two objectives: (1) to remove the low-quality pixels from ocean colour images by identifying overly restrictive mask conditions or logic errors that exclude high quality pixels (Hooker 2003; Gerace and Montanaro 2017; Zhang and Wang 2010); and (2) to increase the number of retrievals by correcting the residual error or uncertainty in low-quality pixels without destroying overall data quality (Chen et al. 2016; Hu, Feng, and Lee 2013). The former might be termed ‘image surgery’, while the latter is more like ‘image cosmetology’. Much attention has focused on ‘image surgery’ based on an increased understanding of ocean colour data over the past several decades (Kutser, Vahtmäe, and Praks 2009; Gerace and Montanaro 2017). Specifically, the original Level-2 Global Area Coverage (GAC) remote sensing reflectance ($R_{rs}$) data are operationally separated into high-quality and low-quality pixels using a 32 bit value (called ‘12_flags’) defined by National Aeronautics and Space Administration (NASA) (Hooker 2003). Using 12_flags can discard about 51% of the total nonzero Level-2 data for the global Level-3 ocean colour product composites of clear water (Chen et al. 2016). The loss of these data is harmful to spatial and temporal analysis in ecological and climate research. Thus, some ‘image cosmetology’ approaches are desired because they can provide additional valid pixels.

Sun light that is specularly reflected from wind-roughened water forms a much larger spatial pattern and, thus, larger signal than the signal from a perfectly flat surface (Zhang and Wang 2010). The larger signal can contaminate parts of the image which leads to a huge loss of information (Doerffer et al. 2008). Generally, the sun glint signal in a glint-contaminated satellite image is so strong that ocean colour data from satellites, such as $R_{rs}$ spectral products, might have huge uncertainties that render them useless. SeaWiFS can tilt the instrument from nadir by about 20° to reduce the effects of sun glint. However, instruments such as the Moderate Resolution Imaging Spectroradiometer Aqua (MODISA), Visible Infrared Imaging Radiometer Suite (VIIRS), and Medium Resolution Spectral Imager II (MERSI-II) do not have glint-tilting functionality, in backwards or forwards directions, to minimize sun glint effects. Note that the MERSI-II is the second generation of China’s polar orbiting meteorological and oceanographic instruments. The satellite was launched on 15 November 2017. This lack of glint reduction functionality might mean that more than half the ocean colour data could be contaminated by sun glint (Doerffer et al. 2008). For meaningful retrievals, the severe glint-contaminated pixels need to be flagged in NASA’s default data processing procedures (McClain et al. 2000; 1995) unless an alternative accurate corrective approach is available.

Various sun glint algorithms have been developed for sun glint retrieval (Cox and Munk 1954; Shifrin 2001; Ebuchi and Kizu 2006), but Zhang and Wang (2010) found that the model that Cox and Munk (1954) developed approach had the best correlation with MODIS measurements. In the ideal case, given the solar-satellite geometric relationship and the histogram of the wind-caused wave slopes, the distribution of glint can be accurately estimated using Cox and Munk’s model. However, because the wind at the time of overflight is hard to accurately predict, pixel by pixel, at basin scales, the Cox and Munk approach can also result in large uncertainty in the generated ocean colour products, especially in severe sun glint regions. Moreover, as pointed out by (Ottaviani et al. 2008) the Cox-Munk approach implemented in SeaWiFS Data Analysis System
SeaDAS) is based on a two-way direct transmittance of light through the atmosphere. This approach ignores the impact of multiple scattering, which leads to an error between 10% and 90% (Ottaviani et al. 2008). As a result, the severe glint-contaminated pixels still need to be flagged and are normally not used for further ocean colour application.

Data quality is a crucial part of analytical work. Time-series analysis, for example, requires the most robust products to understand temporal characteristics of a data set. However, there are several cases, such as visually estimating the surface swell wave transformation, oceanic internal waves, oil spills, and so on, where real-time spatial patterns are more important than data quality (Hu 2011; Doerffer et al. 2008). To increase the number of valid retrievals and geographic coverage, we need to closely examine how sun glint is corrected. Hu (2011) developed an empirical method that retrieves chlorophyll-a patterns under severe glint. However, for oceanographic research other optical and biogeochemical products are needed that include inherent optical properties (IOPs) (Jeremy et al. 2018). Chen et al. (2016); Chen et al. (2019) developed an IOP data processing system (IDAS) that removes residual error in satellite $R_{rs}$ ($\Delta R_{rs}$) and derives IOPs from ocean colour data. Using this scheme, the data quality and spatial coverage of the IOPs from the multi-band ocean colour satellites were improved. The core assumption of an IDAS algorithm is that the residual errors have strong spectral dependence characteristics (Chen et al. 2016; Hu, Feng, and Lee 2013) that are determined algebraically for each pixel. As a matter of fact, sun glint has a good linear spectral relationship (Hedley, Harborne, and Mumby 2005; Kay, Hedley, and Lavender 2009; Martin et al. 2016; Hu 2011) that meets the underlying assumption of an IDAS algorithm for undesired signal correction (Chen et al. 2019; 2016). This is because the sun glint reflectance depends on the change in refractive index between air and water at the air-water interface, but the change in refractive index has a weak dependence on wavelength that is not expected to lead to a significant spectral dependence of the sun glint. Therefore, if handled properly, with some adjustment, the IDAS algorithm can correct for sun glint.

The objectives of this study are to develop and evaluate the IDAS sun glint correction algorithm for the open ocean (referred to as the IDAS-SGC algorithm). Note that the IDAS algorithm is originally designed for remove residual error in satellite $R_{rs}$ data after atmospheric correction (Chen et al. 2019; 2016), which is different from sun glint correction in this study. With VIIRS and MERSI-II open ocean data as examples, the specific goals of the study are as follows: (1) to construct an open ocean IDAS-SGC algorithm for sun glint correction; (2) to validate the stability of the IDAS-SGC algorithm for sun glint correction and then for IOPs retrievals using a synthetic data set; (3) to evaluate the effectiveness of the IDAS-SGC algorithm in retrieving ocean colour data in severe glint regions using synchronized MODIS observations; and (4) to highlight the performance of the IDAS-SGC algorithm in sun glint correction by comparing IDAS-SGC with the Cox-Munk algorithm (Cox and Munk 1954).

2. Data and methods

2.1. Determining the spectral relationship of sun glint reflectance

The remote sensing reflectance associated with signals received by an instrument at the top of atmosphere ($R_i$), after correction for gaseous absorption, is expressed as the follows:
\[ R_t(\lambda) = R_r(\lambda) + R_a(\lambda) + T(\lambda)R_g(\lambda) + t(\lambda)R_{rs}(\lambda) \] (1)

\( R_t \) and \( R_a \) are the ratio of reflectance to \( \pi \) resulting from Rayleigh and aerosol scattering, respectively. \( R_g \) is the ratio of reflectance to \( \pi \) resulting from sun glint and sky glint. \( T \) and \( t \) are the direct and diffuse transmittance of the atmosphere, respectively.

Given the surface atmospheric pressure and wind speed, \( R_t \) can be determined accurately (Gordon and Voss 1999). Because the spatial pattern of aerosol optical depth is spatially continuous and smooth for the open ocean (Lee, Kim, and Lee 2014; Shikwambana and Sivakumar 2018), we propose the inverse distance weighted method for estimating \( R_a \) in glint regions from \( R_a \) in glint-free regions:

\[
R_a(p, q) = \frac{\sum_{y=1}^{N} \sum_{x=1}^{M} R_a(x, y) \left[ (p - x)^2 + (q - y)^2 \right]^{-0.5}}{\sum_{y=1}^{N} \sum_{x=1}^{M} \left[ (p - x)^2 + (q - y)^2 \right]^{-0.5}} \tag{2}
\]

\( x \) and \( y \) are the row and column numbers of the glint-free pixels. \( p \) and \( q \) are the row and column numbers of the glint-contaminated pixels.

While \( R_a \) in glint-free regions can be determined using a lookup table proposed by Gordon and Voss (1999), practically, because the relative azimuth angle varies in satellite nadir viewing regions, the inverse distance weighted approach could produce uncertainties when estimating \( R_a \). These uncertainties would be absorbed as part of \( R_g \) and be included when constructing spectral relationships. With extensive tests and evaluation, this \( R_a \) guess approach could meet the uncertainty requirements for constructing the spectral relationship when correcting for sun glint.

When the Rayleigh and aerosol reflectance are known, Equation (1) in glint regions has only two terms:

\[ R_{gw}(\lambda) = R_{tg}(\lambda) + R_{rs}(\lambda) \] (3)

\( R_{gw} = (R_t - R_r - R_g)/t \). \( R_{tg} \) refers to the ratio of \( T \times R_g \) to \( t \). \( R_{gw} \) is usually several orders larger than \( R_{rs} \), so, for the open ocean, \( R_{rs} \) can be ignored compared with \( R_{tg} \). However, to further enhance the first guess accuracy of \( R_{tg} \), as much \( R_{rs} \) as possible should be removed from \( R_{tg} \). To account for this, we suggest using Equation (2) for approximating \( R_{rs} \) in glint images from \( R_{rs} \) in glint-free images, so the first guess of \( R_{tg} \) is obtained. Note that the I2_flags technique would be used to exclude the lower quality pixels from a Level-2 image in a glint-free region before interpolating \( R_a \) and \( R_{rs} \) in sun glint regions.

Hedley, Harborne, and Mumby (2005), Kay, Hedley, and Lavender (2009), Hu (2011), and Martin et al. (2016) showed that \( R_{tg} \) is approximately linearly spectrally dependent. When the spectral \( R_{tg} \) are determined, the linear model suggested for describing the spectral relationship of \( R_{tg} \) is:

\[ R_{tg}(\lambda) = m_\lambda R_{tg}(671) + n_\lambda \] (4)

\( m_\lambda \) and \( n_\lambda \) are the scale and bias of the linear relationship for \( R_{tg} \), respectively. Values of \( m_\lambda \) and \( n_\lambda \) are fitted from a first guess of \( R_{tg} \) for glint regions. The spectral relationship of \( R_{tg} \) might change with the images due to different illumination-observation and meteorological conditions (Hu 2011), but those impacts would be very small. As a result, \( m_\lambda \) and \( n_\lambda \) are commonly considered to be fixed constants when correcting for sun glint (Hu 2011;
Hedley, Harborne, and Mumby 2005; Kutser, Vahtmäe, and Praks 2009; Martin et al. 2016. After extensive analysis, we used $m_\lambda$ and $n_\lambda$ for adjusting images, image-by-image (one spectral relationship per image), in this study.

### 2.2. Constructing the IDAS-SGC algorithm

McClain (1995) suggested ignoring further data processing for ocean colour images when sun glint radiance ($L_{GN}$) exceeds 0.005, which, in turn, leads to a loss of a large proportion of the images. $L_{GN}$ represents the normalized sun glint radiance to solar irradiance calculated using the Cox and Munk model. To obtain images with complete spatial coverage, we need an algorithm to correct for sun glint contamination, especially in a severe sun glint region ($L_{GN} > 0.005$). The IDAS algorithm was originally developed for residual error correction in satellite $R_{rs}$ and IOPs retrievals (Chen et al. 2019; 2016). The underlying assumption is that the residual error in $R_{rs}$ has spectral dependent characteristics. The $R_{tg}$ has strong spectral dependence characteristics, so it is possible to calculate $R_{tg}$ algebraically from $R_{gw}$ using the IDAS algorithm. For this study, we reconstructed the early version of the IDAS algorithm to be an algorithm to correct for sun glint (SGC).

Chen et al. (2019); (2016) showed that the IOPs-$R_{rs}$ relationship is a typical closed equation, and the uncertainty in estimating $R_{tg}$ would be an added hindrance in solving the equations for the backscattering coefficient at 555 nm ($b_b(555)$). $R_{tg}$ is usually several orders larger than $b_b(555)$ for the open ocean so a small uncertainty in estimating $R_{tg}$ could lead to a huge uncertainty in estimating $b_b(555)$ if the estimation were constrained by the IOPs-$R_{rs}$ relationship. To minimize the impacts of the $R_{tg}$ estimation uncertainty on determining $b_b(555)$ in the IOPs-$R_{rs}$ relationship, we need to decrease the difference in amplitude between the two unknowns ($b_b(555)$ and $R_{tg}$) in the IOPs-$R_{rs}$ relationship. To account for this, we use $R_{gw}(671)$ as the first guess of $R_{tg}(671)$, and then we extrapolate to other short wavelengths using Equation (4). With the first guess of $R_{tg}$ and the spectral relationship defined by Equation (4), $R_{gw}$ in Equation (3) can be corrected as $R_{fc}$:

$$R_{fc}(\lambda) = R_{tg}(\lambda) + R_{rs}(\lambda) - m_\lambda R_{gw}(671) - n_\lambda \approx R_{rs}(\lambda) - m_\lambda R_{rs}(671)$$

(5)

If we substitute Equation (5) into the IOPs-$R_{rs}$ relationship (Chen et al. 2019; 2016), we rewrite the SGC algorithm as:

$$g_0 \left[ \frac{b_b(555)}{a(555) + b_b(555)} \right]^2 + g_1 \left[ \frac{b_b(555)}{a(555) + b_b(555)} \right]^2 = \frac{R_{fc}(555) + m_{551} R_{rs}(671)}{0.52\pi + 1.7|R_{fc}(555) + m_{551} R_{rs}(671)|}$$

(6)

$$g_0 \left[ \frac{b_b(671)}{a_w(671) + b_b(671)} \right]^2 + g_1 \left[ \frac{b_b(671)}{a_w(671) + b_b(671)} \right]^2 = \frac{R_{rs}(671)}{0.52\pi + 1.7 R_{rs}(671)}$$

(7)

$b_b$ is the backscattering coefficient. $a_w$ refers to the absorption of pure water. $g_0$ and $g_1$ are empirical coefficients provided by Lee et al. (Lee, Werdell, and Amon 2009) as 0.089 sr$^{-1}$ and 0.01245 sr$^{-1}$, respectively. $a(555)$ and the spectral relationship of backscattering coefficient in Equations (6) and (7) are determined using a neural network algorithm developed by Chen et al. (2016). This algorithm semi-analytically derives IOPs from $R_{rs}$ while taking into account
the residual error in satellite $R_{rs}$ (Chen et al. 2019; 2016). Thus, we can calculate two unknowns ($b_b(555)$ and $R_{rs}(671)$) for two equations for each $R_{rs}$ spectrum (see the flow chart in Figure 1).

McClain et al. (1995) found that sun glint contamination has little effect on $R_{rs}$ retrievals when $L_{GN}$ varies from 0.0001 to 0.005. Thus, we suggest using the SGC algorithm for correcting sun glint when $L_{GN} > 0.005$. The satellite $R_{rs}$, like many other ocean colour products, is subject to residual errors due to imperfect data processing procedures such as those used for atmospheric correction. Thus, we suggest using the IDAS algorithm to improve data quality of satellite $R_{rs}$ in glint-free regions ($L_{GN} < 0.0001$). Two $R_{rs}$ results from severe glint and glint-free images need to be combined to produce smooth $R_{rs}$ data. To do this, we suggest using a weighting function to bridge those $R_{rs}$ data in slight glint regions ($0.0001 \leq L_{GN} \leq 0.005$). We express the combined new $R_{rs}$ as function of the weight:

$$R_{rs} = W R_{rs,SGC} + (1.0 - W) R_{rs,IDAS}$$

where,

$$W = \frac{\ln(L_{GN}) - \ln(L_{GN,min})}{\ln(L_{GN,max}) - \ln(L_{GN,min})} = \log_{50}(10000L_{GN})\text{, for } 0.0001 \leq L_{GN} \leq 0.005$$

where $R_{rs,SGC}$ and $R_{rs,IDAS}$ represent $R_{rs}$ obtained from slight glint regions after the sun glint correction and residual error correction are applied. The subscripts, ‘min’ and ‘max’, represent the lower ($L_{GN} = 0.0001$) and upper ($L_{GN} = 0.005$) boundaries of slight glint that we use to run the weighting function. Our data processing practice shows that the weighting function with the natural logarithm of $L_{GN}$ works well to produce smooth $R_{rs}$ data, because the sun glint radiance meets the exponential distribution in spatial coverage (Cox and Munk 1954).

### 2.3. Data used

#### 2.3.1. Synthetic data set

Optical data from satellite observations or field measurements usually include some uncertainties, but it is hard to separate those uncertainties from true values in $R_{rs}$.

![Figure 1. Flow chart of the sun glint correction method.](image-url)
Moreover, those optical data always lack desired parameter range at any desired interval and at the same time error-polluted making it difficult to assess the performance of sun glint correction models at broader scales. To overcome this problem, we used the radiative transfer code provided by Hydrolight 5.2 to generate a synthetic data set for different wavelengths of VIIRS bands. This synthetic data set included 5,934 samples. Details of the optical properties used for the Hydrolight simulation are in Chen et al. (2016). We generated a random $R_{tg}(671)$ value using the probability distribution function shown in Figure 2. Note that we obtained the $R_{tg}(671)$ from four random VIIRS images of the open ocean from 2017 using Equations (1)-(3). We extrapolated $R_{tg}(671)$ to $R_{tg}$ at other short VIIRS bands using the spectral relationships in Figure 3. We added these $R_{tg}$ to the synthetic $R_{rs}$ at all VIIRS visible bands according to Equation (3), and we used all the $R_{tg}$ and $R_{rs}$ to run the IDAS-SGC algorithm for sun glint correction and then for IOPs retrievals.

2.3.2. Satellite imagery
To test how stable and accurate our IDAS-SGC algorithm is at obtaining meaningful ocean colour products from satellite images under severe sun glint, we collected two typical VIIRS images of the Equatorial Pacific Ocean and the North Pacific Ocean that were contaminated with severe sun glint on 16 February 2017 and on 3 February 2018, respectively. We used one synchronous MODIS image collected over the North Pacific Ocean on 3 February 2018 for a matching analysis of the performance of the IDAS-SGC algorithm in retrieving $R_{rs}$ data from the VIIRS images with severe sun glint contamination.

To test the stability of the IDAS-SGC algorithm for different instruments, we randomly obtained two typical MERSI-II images with severe sun glint contamination for the Equatorial Pacific Oceans on 1 February 2018 and for the South Pacific Oceans on

![Figure 2. Frequency of $R_{tg}(671)$ determined from four VIIRS images of the open ocean in 2013 after removing atmospheric and water-leaving contributions using Equations (2)-(4).](image-url)
15 February 2019. The cross-calibration approach developed by Chen et al. (2020) was used to remove radiance calibration uncertainty in MERSI-II reflectance at the top-of-atmosphere. We collected the corresponding two synchronous MODIS images for evaluating how effective the IDAS-SGC algorithm is in correcting the sun glint for severely contaminated regions. In addition, we used four random VIIRS images of the open oceans for spectral analysis of $R_{tg}$.

### 2.4. Assessment criteria

The criteria we used for assessing the stability of the IDAS-SGC in retrieving ocean colour patterns under severe sun glint included the root-mean-square of the log transformed error ($\text{RMSE}_{\text{log}}$), the mean absolute percent difference (MAPD), the unbiased mean percent difference (UMPD), and the coefficient of variation (CV). These statistics are defined in the following equations:

$$\text{MAPD} = \frac{1}{N} \sum_{k=1}^{N} \frac{|x_{e,k} - x_{r,k}|}{x_{r,k}}$$

$$\text{UMPD} = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{x_{r,k} - x_{e,k}}{0.5x_{r,k} + 0.5x_{e,k}} \right|$$

Figure 3. Spectral relationship of sun glint-contributed reflectance estimated from four VIIRS images in 2013 after removing atmospheric and water-leaving contributions using Equations (2)-(4).
\[ CV = \sqrt{\frac{\sum_{i=1}^{3} \sum_{j=1}^{3} (x_{e,i,j} - x_{r,i,j})^2}{x_{r,m}}} \]  

(12)

\[ \text{RMSE}_{\text{log}} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left[ \log(x_{e,k}) - \log(x_{r,k}) \right]^2} \]  

(13)

where \( x_{e,k} \) is the algorithm-derived ocean colour product of the \( k^{th} \) element from satellite \( R_{rs} \), \( x_{r,k} \) is the equivalent product of the \( k^{th} \) element provided by Hydrolight or derived from the synchronous satellite \( R_{rs} \). \( N \) is the number of elements, and \( x_{r,m} \) is the mean value of \( x_{r,k} \) for a \( 3 \times 3 \) box.

3. Results and discussion

3.1. Spectral relationship of sun glint signals

The linear spectral relationship of \( R_{tg} \) is the key parameter for accurate sun glint retrieval and for correction in the IDAS-SGC algorithm. However, the spectral slope could slightly vary with different data sources due to the varying illumination-observation and atmospheric conditions of the open ocean (Hu 2011; Doerffer et al. 2008). When the \( R_{tg} \) in sun glint regions is known, the accurate slope and bias of the linear fits of Equation (4) are easily determined according to the actual \( R_{tg} \) data from a given satellite image.

Using the VIIRS image over the North Pacific Gyre on 16 February 2017 as an example, we used Equations (1)-(4) to remove the atmospheric and water-leaving contributions from \( R_{r} \). Figure 3 presents a case of the spectral relationships of \( R_{tg} \) for the VIIRS visible bands from about 0.4 million pixels. We found that the \( R_{tg} \) are monotonically spectrally related from the short to long bands, and the coefficients of determination, \( R^2 \), of the linear relationship between the satellite-observed and model-predicted values were not lower than 0.99. We formed similar conclusions for four VIIRS images over the South Pacific Gyre from May 2018. The spectral relations are the result of a weak dependence on wavelength, but strong dependence on the spectral characteristics of solar radiance, wind speed, and illumination-observation conditions (Cox and Munk 1954; Mobley 1994), which have been recognized and confirmed in historical reports (Hu 2011; Kay, Hedley, and Lavender 2009; Martin et al. 2016; Kutser, Vahtmäe, and Praks 2009). Those results suggest that the underlying assumptions of Equation (4) are robust for providing a stable spectral relationship of \( R_{tg} \) for the open ocean for glint correction.

3.2. Evaluation using synthetic data

Using the spectral relationship of \( R_{tg} \) to run the IDAS-SGC algorithm, Figure 4 presents scatterplots of the IDAS-SGC algorithm-derived \( R_{rs} \) from the \( R_{gw} \) of the synthetic data set against the known data. Based on 5,934 simulated samples, we found that the points were distributed around the 1:1 line even though random sun glint data were added into \( R_{rs} \). The \( R^2 \) of the linear relationship between the IDAS-SGC-derived and known \( R_{rs} \) were >
while the corresponding MAPD values did not exceed 1.25%. Our findings imply that the IDAS-SGC algorithm can separate \( R_{tg} \) from \( R_{gw} \) for the open ocean if the spectral relationships of \( R_{tg} \) are known.

We further evaluated the performance of the IDAS-SGC algorithm at eliminating the effects of sun glint on IOPs retrievals by using the IDAS-SGC algorithm-corrected \( R_{rs} \) data for estimating absorption and backscattering coefficients. Figure 5 compares the IOPs estimated from error-free and sun glint-corrected \( R_{rs} \). Because the IDAS-SGC algorithm can correct for sun glint, the total absorption and backscattering coefficients can be accurately determined from sun glint-corrected \( R_{rs} \) (see Figure 5(a,b)). Specifically, the \( R^2 \) for \( a(551) \) and \( b_g(551) \) are > 0.99, and the corresponding MAPD values do not exceed 1.23%. When the \( a(\lambda) \) at blue wavelengths are determined with high quality, the individual component absorption can be accurately separated from \( a(\lambda) \) for clear water (Chen et al. 2019; Lee, Werdell, and Arnone 2009). Figure 5(c,d) show the performance of the IDAS-SGC algorithm in estimating the absorption coefficient of phytoplankton \( (a_{ph}) \) and the sum of the absorption coefficients of detritus and gelbstoff \( (a_{dg}) \) at 443 nm using sun glint-corrected \( R_{rs} \) data. We found that the algorithm-derived \( a_{ph}(443) \) and \( a_{dg}(443) \) using the sun glint-corrected \( R_{rs} \) were consistent with those derived from the error-free \( R_{rs} \) data whose \( R^2 \) and MAPD values were > 0.99 and < 1.32%, respectively. Our results indicate that the IDAS-SGC algorithm can effectively remove sun glint reflectance in \( R_{gw} \) and can subsequently
obtain meaningful retrievals of IOPs from glint-contaminated satellite data, at least for our synthetic data set.

3.3. Evaluating the algorithm with VIIRS images

Figure 6 shows an example of a VIIRS image collected on 16 February 2017 over the Equatorial Pacific Ocean. We found that there were > 30% pixels in the central images that were severely contaminated by sun glint (see the pixels within the two black lines in Figure 6(a-c), where the black lines depict the boundaries of the severe sun glint region \( L_{GN} = 0.005 \)). The severe glint recognition approach suggested by McClain et al. (1995) is used to define this severely contaminated region.

To obtain the meaningful ocean colour data, NASA’s default data processing techniques are used to flag the severe glint pixels found between the two black lines in Figure 6. However, this ‘image surgery’ approach weakens the ability of satellite imagery to depict the overall spatial characteristics of some circulation patterns in near real-time. After image processing with our IDAS-SGC algorithm, however, our result showed a spatially consistent \( R_{rs} \) pattern between the severe glint and glint-free regions (Figure 6(d-f)).

Due to the significantly improved coverage, the glint-corrected images are more useful than the glint-flagged images for observing small-scale ocean circulation in the open ocean. For example, there is a small-scale eddy across the sun glint regions (Figure 6(l)). It is challenging to recognize this small-scale eddy in the glint-flagged images due to the
incomplete spatial coverage. However, the small-scale eddy is very clear in the glint-corrected images.

The spatial patterns of ocean colour retrieved from near-daily VIIRS data are meaningful for qualitatively validating numerical circulation models, and the data are useful for visually estimating surface swell wave transformation, fronts and filaments, oceanic internal waves, and surface slicks and oil spills superimposed on the ocean colour images (Hu 2011). The spatial pattern is very important in these cases. When the determined $R_{rs}$ are of high quality, the IOPs, including absorption and backscattering coefficients, can be obtained from $R_{rs}$ with high quality for the open ocean (Chen et al. 2016; Lee, Werdell, and Arnone 2009). Like the $R_{rs}$ image, Figure 6(g–k) present spatially consistent and smooth IOPs patterns between the severe glint and glint-free images. Meanwhile, the boundary of the small-scale eddy seen in the IOPs images is more clear than in the $R_{rs}$ image, because the IOPs are more directly related than $R_{rs}$ is with ocean colour anomalies caused by

Figure 6. The ocean colour data derived from a VIIRS image over the Equatorial Pacific Oceans on 16 February 2017. (a) to (c) are TOA reflectance at 443, 486, and 551 nm; (d) to (f) are remote sensing reflectance at 443, 486, and 551 nm; (g) to (i) are absorption coefficients at 443, 486, and 551 nm; (j) to (k) are absorption coefficients of phytoplankton and gelbstoff at 443 nm; and (l) is the sea surface temperature overlaid with the current direction. The two black lines are the boundaries of the region with severe sun glint provided by McClain et al. (1995). The white areas are clouds. The sea surface temperature is provided by VIIRS infrared instrument, while the current data is obtained from the Global Ocean Gridded L4 sea surface heights and derived variables reprocessed dataset.
oceanic dynamic change-dependent anomalies in the small-scale eddy. These results indicate that the IDAS-SGC algorithm improves the data quality and spatial coverage of ocean colour images under severe sun glint for the open ocean.

Furthermore, Figure 7 compares the matching VIIRS and MODIS $R_{rs}$ data in severe glint and glint-free regions over the North Pacific Ocean on 3 February 2018. The results show that sun glint can render one-third of the images useless for the VIIRS or MODIS instruments. Generally, VIIRS and MODIS have coincidental alignment of satellite orbit tracks relative to their orientation, so there is a severe glint stripe in the MODIS images that partially overlap with the VIIRS severe glint regions (see Figure 7).

When we used the default NASA data processing techniques for obtaining ocean colour products from the VIIRS images, a mask was generated for the regions under severe sun glint, removing any useless images like those MODIS images with the white stripe. As a result, the ocean colour patterns were incomplete and even distorted (see Figure 7(a,d,g)). After correcting for sun glint using the IDAS-SGC algorithm, we obtained a spatially consistent and smooth image between the severe glint and glint-free regions in ocean colour patterns (see Figure 7(b,e,h)). Specifically, some circulation patterns of a loop current and eddies are clearly observed from the images that were corrected for sun glint (deep blue region in Figure 7(b)). When we compared the corrected images, pixel-by-pixel, with the equivalent synchronized observations from MODIS (Figure 7(c,f,i)), we found that the $R_{rs}$ from the two different instruments exhibited consistent spatial patterns.

**Figure 7.** Comparison between matching VIIRS and MODIS $R_{rs}$ data over the North Pacific Ocean on 3 February 2018. The left panel is VIIRS $R_t$; the middle panel is VIIRS $R_{rs}$; and the right panel is MODIS $R_{rs}$. The region between the two black lines is the region with severe sun glint provided by McClain et al. (1995). The white areas in the VIIRS images represent clouds. The white areas in the MODIS images represent clouds and sun glint.
in the severe glint or glint-free regions. Specifically, a meso-scale ocean circulation pattern is evident around the equatorial regions from the VIIRS and MODIS images.

Figure 8 shows scatterplots of matching VIIRS and MODIS $R_{rs}$ data, like what is shown in Figure 7. We found that the IDAS-SGC algorithm accurately derived $R_{rs}$ data from the sun glint-contaminated images. For example, after correcting the contaminated images using the IDAS-SGC algorithm, the inter-mission differences between VIIRS and MODIS $R_{rs}$ are 16.56%, 7.80%, and 28.45% at 443, 486, and 555 nm (see Figure 8(a-c)), respectively, while the corresponding differences for glint-free pixels in the same images are 14.99%, 8.39%, and 38.56% (see Figure 8(d-f)), respectively. These results consistently confirm that the IDAS-SGC algorithm is effective in correcting for sun glint to provide accurate ocean colour data.

3.4. Evaluating the algorithm with MERSI-II images

Figure 9 shows an example of MERSI-II images collected on 1 February 2018 over the Equatorial Pacific Ocean after correcting for sun glint. The results show that there is an eastward cold current in the images around the equatorial region (see Figure 9(l)). The cold circulation always has high nutrients and salt that promote the growth of phytoplankton. As a result, the blue band $R_t$ and $R_{rs}$ in the cold current are significantly lower than in the neighbouring regions, but the absorption coefficients are clearly higher than in the neighbouring regions. This is because the phytoplankton and its covarying components are strongly absorbent at blue wavelengths (Mobley 1994). Generally, the pixels in the centre of the image are useful for ocean colour applications due to small geometric

![Figure 8](image_url). Scatterplots of matching VIIRS $R_{rs}$ versus MODIS $R_{rs}$ presented in Figure 7. (a) to (c) represent the results for the sun glint region after being corrected, and (d) to (f) are results for the glint-free regions.
distortion and good illumination-observation conditions. However, to obtain meaningful retrievals using NASA's default data processing techniques (McClain et al. 2000; 1995), we had to discard those pixels in the images contaminated with sun glint which could account for > 30% of total nonzero data. Consequently, we retained only the pixels around the scan edge for the ocean colour products, and the data quality of those pixels is usually poor due to their large viewing angle (> 56°) (Gordon and Voss 1999; Barnes and Hu 2016). Moreover, this 'image surgery' could lead to difficulties in presenting a complete spatial structure of the circulation from a single satellite image. Fortunately, in our case, the complete structure of the cold eddy is clearly delineated from the glint-corrected \( R_{\text{rs}} \) and IOPs images because the IDAS-SGC algorithm recovered the image coverage with high quality. More specifically, the absorption coefficient anomaly matches the sea surface temperature anomaly in the cold circulation (see Figure 9(g-l)). The ocean

Figure 9. Ocean colour data derived from a MERSI-II image over the Equatorial Pacific Ocean on 1 February 2018. (a) to (c) are reflectance at top of atmosphere at 443, 486, and 551 nm; (d) to (f) are remote sensing reflectance at 443, 486, and 551 nm; (g) to (i) are absorption coefficients at 443, 486, and 551 nm; (j) to (k) are absorption coefficients of phytoplankton and gelbstoff at 443 nm; and (l) is the sea surface temperature overlaid with the current direction. The region between the two black lines is the region with severe sun glint provided by McClain et al. (1995). The white patches represent clouds. The sea surface temperature is provided by VIIRS infrared instrument, while the current data is obtained from the Global Ocean Gridded L4 sea surface heights and derived variables reprocessed dataset.
colour products exhibit spatially smooth and consistent characteristics between the severe glint and glint-free regions. We confirm this result (see Figure 10) with the random MERSI-II images collected over the South Pacific Ocean on 15 February 2019. For these images the IDAS-SGC algorithm provided meaningful ocean colour products for severe glint or glint-free regions. These further results confirm the potential of the IDAS-SGC algorithm to correct for sun glint in satellite images from multi-missions.

Figure 11 shows the synchronous comparison between MODIS and MERSI-II $R_{rs}$ data for glint and glint-free regions, where the MERSI-II images are originally presented in Figures 9-10. The results show that the inter-mission differences are 25.48%, 15.41%, and 54.26% at 443, 490, and 555 nm in the glint regions, respectively, and the corresponding differences are 29.45%, 14.68%, and 54.53% in the glint-free regions, respectively. Such a result suggests that the usually discarded glint-contaminated pixels, which account for > 30% of the images, could be converted to meaningful ocean colour products using the IDAS-SGC algorithm. It is

![Figure 10](image_url)

**Figure 10.** The ocean colour data derived from a MERSI-II image over the South Pacific Ocean on 15 February 2019. (a) to (c) are reflectance at top of atmosphere at 443, 486, and 551 nm; (d) to (f) are remote sensing reflectance at 443, 486, and 551 nm; (g) to (i) are absorption coefficients at 443, 486, and 551 nm; (j) to (k) are absorption coefficients of phytoplankton and gelbstoff at 443 nm; and (l) is the sea surface temperature overlaid with the current direction. The region between the two black lines is the region with severe sun glint provided by McClain et al. (1995). The white patches represent clouds. The sea surface temperature is provided by VIIRS infrared instrument, while the current data is obtained from the Global Ocean Gridded L4 sea surface heights and derived variables reprocessed dataset.
It is noteworthy that the determination coefficient between inter-mission images are not very high in this study (see Figures 8 and 11). Actually, the factors influencing the inter-mission correlation between ocean colour products are very complicated, including geometric mismatch, overpass time difference, meteorological considerations, current velocity, etc. Fortunately, the correlation for sun glint regions (corrected data) are comparable with sun glint free regions (uncorrected data). This is to say that the low determination coefficients are mainly caused by the data, not by the algorithm.

### 3.5. Comparing IDAS-SGC with the Cox-Munk model

We suggest using NASA’s default data processing techniques (McClain et al. 2000), an iterative procedure combined with Cox-Munk model, for removing the sun glint and aerosol scattering contribution in the glint regions ($L_{GN} > 0.0001$) while we propose the near-infrared atmospheric correction approach for retrieving ocean colour data in the glint-free regions ($L_{GN} < 0.0001$). The images in severe glint regions should be discarded using the NASA default data processing approach, but we kept them to highlight the performance of the IDAS-SGC algorithm for sun glint correction.

**Figure 11.** Scatterplots of matching MERSI-II versus MODIS $R_{rs}$ data where the MERSI-II images have been presented in Figures 9 and 10. (a) to (c) represent the results for the sun glint region after correction, and (d) to (f) are results for glint-free regions.
With the VIIRS and MERSI-II images in Figures 6, 7, 9, and 10 as examples, Figure 12 shows the Cox-Munk $R_{rs}$ data in glint or glint-free regions. As expected, the patchiness and speckles are common in VIIRS or MERSI-II $R_{rs}$ images after correcting the images with the Cox-Munk algorithm. This is because the Cox-Munk model fails to accurately correct the sun glint from the VIIRS or MERSI-II images in severe glint regions due to the absence of accurate real-time wind speed and direction data. According to Cox and Munk (1954), overestimating (underestimating) real-time wind would lead to overestimating (underestimating) the glint contribution and would result in underestimating (overestimating) $R_{rs}$. However, the wind data contained some uncertainties that are very common in the open ocean. Generally, the sun glint reflectance is several orders higher than the water-leaving signals (Zhang and Wang 2010), so a small uncertainty in estimating the sun glint reflectance can result in a huge uncertainty in $R_{rs}$ retrievals. Figure 12 indicates that the Cox-Munk model could significantly underestimate or overestimate sun glint in the severely affected regions and could give $R_{rs}$ data patchiness and speckle noise. Specifically, there could be negative $R_{rs}$ values in Figure 12. Consequently, it is hard to

Figure 12. Ocean colour data derived from VIIRS and MERSI-II images over the open ocean using NASA default data processing techniques. (a) to (c) and (d) to (f) are VIIRS $R_{rs}$ results over the equatorial Atlantic Ocean on 16 February 2017 and over the North Pacific Gyre on 3 February 2018; and (g) to (i) and (j) to (l) are MERSI-II $R_{rs}$ results over the South Pacific Ocean on 1 February 2018 and North Pacific Ocean on 15 February 2019. The region between the two black lines is the region with severe sun glint provided by McClain et al. (1995). The white patches represent clouds.
define complete boundaries of the meso-scale ocean circulation in the Cox-Munk $R_{rs}$ images, but these boundaries are very clear in the IDAS-SGC $R_{rs}$ image.

**Figure 13** shows the synchronous comparison of MODIS $R_{rs}$ data with Cox-Munk model-derived VIIRS and MERSI-II $R_{rs}$ data for severe glints regions. We see that the UMPD values of the Cox-Munk model are about twice as large as the UMPD of the IDAS-SGC algorithm (Figures 8, 11, and 13). Specifically, a great deal of underestimated $R_{rs}$ data exist at the lower end ($R_{rs}(443) < 0.001 \text{ sr}^{-1}$), while many overestimated $R_{rs}$ appear at the higher end ($R_{rs}(443) > 0.012 \text{ sr}^{-1}$). Those $R_{rs}$ data are mainly from the extremely severe regions with $L_{GN} > 0.1$ due to the Cox-Munk model overestimating or underestimating sun glint. However, such inconsistencies significantly improved in the IDAS-SGC results (Figures 8 and 11). Those results highlight that the IDAS-SGC algorithm more effectively obtains meaningful $R_{rs}$ data from sun glint regions than the Cox-Munk algorithm.

### 4. Conclusions

Because many ocean colour instruments cannot tilt their sensors away from nadir, the ocean colour images acquired by those instruments are likely to be severely contaminated by sun glint. Sun glint can cause $> 30\%$ of total nonzero images to be discarded, which is a disadvantage to visually estimating real-time physical and chemical processes in the open ocean such as small-scale eddies and circulation (Zhang et al. 2019). The original IDAS algorithm corrects residual error in satellite $R_{rs}$ associated with imperfect data processing procedures. The sun glint radiance has good spectral dependence (see $R^2 > 0.99$) characteristics which meet the underlying assumption of the IDAS algorithm for
isolating and correcting undesired radiance. Therefore, we adjusted the IDAS algorithm to correct for sun glint. As expected, the IDAS-SGC algorithm can simultaneously correct the sun glint and derive ocean colour products from satellite images that are severely glint-contaminated.

Tests with synthetic data and satellite images indicate that the IDAS-SGC algorithm effectively corrects for sun glint and accurately estimates ocean colour products. Specifically, the $R^2$ values between the IDAS-SGC-derived and known sun glint and ocean colour data are not lower than 0.99 for the synthetic data. After processing satellite images with the IDAS-SGC algorithm, the spatial coverage of the ocean colour products improves significantly for visually estimating eddies and circulation, and the ocean colour products exhibit spatial consistency and smoothness in glint-contaminated or glint-free regions. Furthermore, the VIIRS and MERSI-II $R_{rs}$ data in severe glint regions are consistent with synchronous MODIS $R_{rs}$ data in glint-free regions. In comparison, the complete boundaries of meso-scale ocean circulation and eddies are hard to depict from Cox-Munk $R_{rs}$ images, but those boundaries are very clear in the IDAS-SGC $R_{rs}$ images.

Although the algorithm needs further testing with more data, our preliminary results indicate that the IDAS algorithm is an encouraging method for isolating and correcting sun glint in satellite top of atmosphere radiances and to subsequently convert $>30\%$ glint-contaminated ocean data into meaningful retrievals.

**Acknowledgements**

Financial support for this study was provided by the National Key R&D Program of China2018YFB0504800 (2018YFB0504802; 2018YFB0504901), Shaanxi Key Research and Development Program (2018ZDXM-GY-023), Fundamental Research Funds for the Central Universities (60020302), and National Natural Science Foundation of China (41676171). We also thank NASA, NOAA, and the National Satellite Meteorological Centre of China for providing the ocean color and meteorological data.

**Disclosure statement**

The authors declare no conflict of interest.

**References**


