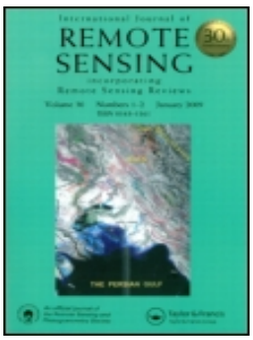

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

# An assessment of cloud masking schemes for satellite ocean colour data of marine optical extremes

Andrew Clive Banks & Frédéric Mélin

**To cite this article:** Andrew Clive Banks & Frédéric Mélin (2015) An assessment of cloud masking schemes for satellite ocean colour data of marine optical extremes, International Journal of Remote Sensing, 36:3, 797-821, DOI: 10.1080/01431161.2014.1001085

**To link to this article:** https://doi.org/10.1080/01431161.2014.1001085

© 2015 European Union. Published by Taylor & Francis.

Published online: 10 Feb 2015.

Submit your article to this journal

Article views: 1526

View related articles

View Crossmark data

Citing articles: 12 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tres20

*International Journal of Remote Sensing*, 2015
Vol. 36, No. 3, 797–821, http://dx.doi.org/10.1080/01431161.2014.1001085

Taylor & Francis
Taylor & Francis Group

# An assessment of cloud masking schemes for satellite ocean colour data of marine optical extremes

Andrew Clive Banks* and Frédéric Mélin

*European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), Water Resources Unit, Ispra (Va), Italy*

(*Received 7 August 2014; accepted 28 November 2014*)

One of the most important steps in utilizing ocean colour remote-sensing data is subtracting the contribution of the atmosphere from the signal at the satellite to obtain marine water-leaving radiance. To be carried out accurately, this requires clear-sky conditions, i.e. all clouds need to be excluded or masked from the data prior to atmospheric correction. The standard cloud mask used routinely in the processing of NASA global ocean colour data is based on a simple threshold applied to the Rayleigh-corrected top-of-atmosphere (TOA) radiance. The threshold is kept purposefully low to ensure high-quality processing at a global scale. As a consequence, the standard scheme can sometimes inadvertently mask important extreme optical events such as intense blue–green algal (cyanobacteria) blooms or the outflow of sediment-rich waters from some of the world's largest rivers. However, the importance of these extreme conditions, both for ecological and hydrological applications, requires that they should be appropriately monitored. Therefore, an assessment of existing cloud masking schemes that could provide valuable alternatives was carried out. A new hybrid cloud mask was also proposed and similarly tested. The selected schemes were systematically assessed over a full annual cycle of satellite ocean colour data on three example regions: the Baltic Sea, the Black and Azov Seas, and the Amazon River delta. The results indicate that the application of alternative cloud masking schemes produces a significant increase in clear-sky diagnostics that varies with the scheme and the region. Major occurrences of extreme optical conditions, such as cyanobacteria blooms, or river deltas formerly excluded from any processing may be recovered, but some schemes may underestimate the amount of thin clouds potentially detrimental to ocean colour atmospheric correction.

## 1. Introduction

The most fundamental quantity in ocean colour remote sensing is marine water-leaving radiance in the visible part of the electromagnetic spectrum. From this quantity, most ocean colour products are derived. In order to estimate radiance at the Earth's surface precisely, in addition to a well-calibrated sensor, it is essential to accurately subtract the contribution of the atmosphere from the signal at the satellite. This requires clear-sky conditions, i.e. all clouds or thick aerosol plumes need to be excluded or masked from the data prior to atmospheric correction. This step is usually referred to as cloud masking.

Processing of satellite imagery sampling the worlds' oceans every day is automated and usually excludes situations with clouds, deep aerosol plumes, or ice, which are filtered out with rather simple tests. The processing code Sea-viewing Wide Field-of-view Sensor

*Corresponding author. Email: andrew.banks@jrc.ec.europa.eu

© 2015 European Union. Published by Taylor & Francis.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

(SeaWiFS) Data Analysis System (SeaDAS, Fu, Baith, and McClain 1998) from NASA, by far the most commonly used by the ocean colour community, is using the signal from one satellite band in the near infrared for cloud masking: if the signal for a given pixel is brighter than a fixed threshold, then this pixel is masked out. This threshold has so far been kept purposefully low to avoid contaminating data records with unwanted cases and ensuring the quality of global data series. With respect to the open ocean, coastal regions and marginal seas are characterized by an optical complexity that includes high reflectance in the near infrared. Some algal blooms with large concentrations of cells or detrital material accumulating at the surface, or large sediment loads found close to the coasts and some river mouths may sometimes be bright enough to trigger the exclusion of the associated data even in clear atmospheric conditions, leading to a loss of precious information in downstream products.

In a context of climate change, extreme conditions may become more frequent (Easterling et al. 2000). Coupled with the direct impacts from anthropogenic nutrient inputs and fishing in coastal regions (Galloway et al. 2008; Stewart et al. 2010), they may trigger sharp responses from phytoplankton communities. For instance, ecologically disruptive or harmful algal blooms show signs of increased frequencies (Hallegraeff 2010; Fu, Tatters, and Hutchins 2012). Some of these blooms are also optical extremes, which might prompt their exclusion from standard ocean colour processing. Considering the ecological importance of these blooms on marine ecosystems and even human activities, they should be described quantitatively by remote sensing to be properly integrated into the statistical treatment of extreme events (Katz 2010).

Another condition typically leading to excessive masking is high concentrations of sediments, which is limiting the potential of ocean colour to support hydrological and coastal studies. Here again, more comprehensive satellite time series of suspended and dissolved organic matter would be an asset to study the impact of river inputs on coastal waters, such as eutrophication of coastal ecosystems and the role that inputs of terrestrial origin play in marine productivity (Nixon 1995; Smith, Tilman, and Nekola 1999; EEA 2001; Banks et al. 2012; references therein). This may also include the impact of modifications in the upstream precipitation regimes or of human activities in extending the network of river impoundments (Humborg et al. 1997; Vörösmarty et al. 2003; Gong et al. 2006).

The standard NASA cloud mask, operational in the processing of ocean colour missions, including SeaWiFS, the Moderate Resolution Imaging Spectroradiometer (MODIS), and the Medium Resolution Imaging Spectrometer (MERIS), was implemented concurrently with a change in the imaging duty cycle of SeaWiFS in 2000, which resulted in data from higher solar zenith angles being collected (up to 83°, Patt et al. 2003). Prior to this, the cloud detection algorithm for SeaWiFS was based on a mean cloud albedo threshold. This was sensitive enough to be incorrectly triggered by just the increased Rayleigh path radiance at high solar zenith angles. The threshold of the updated standard cloud mask was set at 2.7% of the Rayleigh-subtracted reflectance value at 865 nm (869 nm for MODIS Aqua) in order to match the performance of the original mean cloud albedo mask at low-to-moderate solar and viewing geometries over the open ocean, while correctly masking less of the data at higher angles.

There are many other cloud masking schemes reported in the literature for different applications. These include histogram analysis methods (Phulpin et al. 1983), other threshold-based schemes (e.g. Saunders and Kriebel 1988; Simpson and Gobat 1996; Simpson, Schmidt, and Harris 1998; Santer et al. 1997; Ackerman et al. 1998; Cervino

et al. 2000; Birks 2007; Plummer 2008), masks that use spatial variability in time series (e.g., Hagolle et al. 2010; Lyapustin, Wang, and Frey 2008), and pattern recognition and region segmentation techniques (Garant and Weinman 1986; Sedano et al. 2011). However, not all of these are applicable to cloud masking for ocean colour sensors.

For SeaWiFS, the limitation of the wavelength range, particularly the lack of thermal infrared bands, has limited more accurate cloud detection for this sensor (Okada, Mukai, and Sano 2003). An improvement for SeaWiFS has been offered by Nordkvist, Loisel, and Duforet Gaurier (2009), who used a ratio threshold in the visible and near-infrared wavelengths to add to the clear pixels already defined by the standard cloud mask.

In contrast, for the MODIS sensors (Terra and Aqua) and their increased number of spectral bands and potential applications, a dedicated MODIS atmosphere group is looking at the discrimination of cloud from clear sky and producing an operational cloud mask product for MODIS (Ackerman et al. 1998, 2010). This has taken the form of a confidence cloud mask that is based on 10 separate cloud detection tests and their confidence levels when the sensor is over the ocean during the day. Wang and Shi (2006) also formulated a cloud mask designed specifically for MODIS ocean colour data using the short wave infrared bands at 1240 and 1640 nm.

Of additional relevance to ocean colour remote sensing are the more recent cloud masks developed specifically for the Visible Infrared Imaging Radiometer Suite (VIIRS) and MERIS sensors. For VIIRS, the cloud mask was built using a confidence cloud mask framework similar to the MODIS cloud mask. However, this included some improvements and additional cloud tests created to exploit the VIIRS design (Hutchison et al. 2005, 2008). Although the VIIRS-specific cloud mask may offer some improvements, testing it systematically at this point in time was not pursued because the data products available are limited (2012–2013) and are still considered beta quality by the National Polar-orbiting Partnership science team (NASA 2013).

For MERIS, the NASA SeaDAS processing of level-1 TOA data to level-2 products also applies the standard cloud mask. The European Space Agency (ESA) has a separate processing chain for MERIS and the production of level-2 products where their operational cloud mask is based on a pixel classification algorithm (Santer et al. 1997; ESA 2011). Originally, the cloud masking part of the pixel classification was based on a combination of three threshold tests on spectral radiance and five pressure estimate tests derived from oxygen absorption calculations. For the third reprocessing (March 2011) of all MERIS data, cloud masking was based on refinements of the same general methodologies (ESA 2011).

The above introduction to the problems of cloud masking for satellite ocean colour data and the different cloud masks available begs a number of questions, i.e. what are the optical situations in the world's oceans that can be subject to a loss of data due to excessive cloud masking? How much data are lost over longer time periods and are there temporal and spatial patterns in this loss? Finally, could alternative cloud masks solve these problems while remaining conservative enough for application to the entire ocean colour archive and operationally for future ocean colour missions? This article attempts to contribute elements of answers to these issues.

Some of the alternative approaches to cloud masking listed above are not easily applicable to ocean colour sensors or were often tested on a small set of satellite images. The objectives of this article are thus to assess a limited set of cloud masking schemes that are relevant and applicable to past, present, and planned ocean colour sensors by testing them on significant datasets (typically a year of data) and to obtain conclusions on representative test cases of marine optical extremes.

## 2. Methods

### 2.1. Satellite data

MODIS Aqua and SeaWiFS level 1 data used for this research were downloaded from NASA's Ocean Color Web site (oceancolor.gsfc.nasa.gov). The MODIS Aqua cloud mask data specific to the MODIS atmosphere group were downloaded from NASA's level 1 and Atmosphere Archive and Distribution System (ladsweb.nascom.nasa.gov). MERIS level 1 and ESA processed level 2 data were obtained from ESA.

### 2.2. Cloud mask application

The cloud masks selected for testing were as follows:

(1) The standard NASA cloud mask (STD). This mask is activated for any pixels that have a TOA Rayleigh-corrected reflectance ($\rho_s$) value in the 865/869 nm-centred spectral band >0.027 (Patt et al. 2003).

(2) The Nordkvist, Loisel, and Duforet Gaurier (2009) cloud mask (N09). Initially, a $\rho_s$ value <0.027 in the 865 nm-centred spectral band indicates pixels that are clear. These are added to those further identified as clear sky using a threshold of the ratio ($\varepsilon_{max}$) in the visible and near infrared:

$$\varepsilon_{max} = \rho_{smax}/\rho_{smin}, \tag{1}$$

where $\rho_{smax}$ is the highest surface reflectance between selected SeaWiFS or MODIS Aqua bands (412, 555, 670, 865 nm-centred spectral bands for SeaWiFS and 412, 555, 667, 859/869 nm-centred spectral bands for MODIS Aqua), and $\rho_{smin}$ is the lowest surface reflectance between the same bands.

The threshold for $\varepsilon_{max}$ was set at 2.5 (as specified by Nordkvist, Loisel, and Duforet Gaurier 2009) with any values above this indicating additional pixels considered to be clear. All other pixels are masked out.

(3) The Wang and Shi (2006) cloud mask (WS06). This cloud mask uses the 1240 nm-centred short wave infrared band of MODIS Aqua. Specifically, a Rayleigh-corrected reflectance ($\rho_s$) value for this band >0.0235 indicates cloudy conditions.

(4) The MODIS Atmosphere Group (MAG) cloud mask (Ackermann et al. 2010). For satellite acquisition over the ocean during the day, which is the part of the orbit relevant to ocean colour, this cloud mask employs a combination of 10 separate cloud masks as well as non-cloud obstruction (heavy aerosol) and suspended dust tests. The cloud masks are as follows.

- Temperature threshold cloud masks at 6.7, 11.0, and 13.9 μm. Low, mid-point, and high confidence levels of clear sky in Kelvin for these wavelengths are: 215, 220, 225; 267, 270, 273; and 222, 224, 226, respectively.
- Temperature difference cloud masks using 8.6 vs. 11.0 and 11.0 vs. 3.9 μm. Low, mid-point, and high confidence levels of clear sky in Kelvin for these differences are 0.0, −0.5, −1.0, and −8.0, −6.0, −10.0, respectively.
- Temperature difference cloud mask for cirrus cloud using a 'split window' technique on 11.0 and 12.0 μm (Saunders and Kriebel 1988; Key 2002). This initially adjusts for the scan angle dependence of brightness temperature

difference for these two channels, with thresholds varying according to temperature from the 11.0 μm channel.

- Water vapour cirrus test at 1.38 μm. This is a threshold test that relies on the strong water vapour absorption at this wavelength. Thresholds in the reflectance at 1.38 μm are 0.0050, 0.0125, and 0.0350 for confidently clear, the mid-point, and confidently cloudy conditions, respectively.
- Visible or near-infrared threshold test. The TOA reflectance at 0.86 μm is used over water, making it a near-infrared test that uses thresholds at 0.030, 0.045, and 0.065 reflectance for high confidence clear, mid-point, and low confidence clear conditions, respectively.
- Near-infrared/visible ratio test. This test uses the ratio between the reflectances at 0.86 and 0.65 μm (0.86/0.65) and over the ocean uses thresholds of 0.85, 0.90, and 0.95 for confidently clear, midpoint, and confidently cloudy conditions, respectively.
- Surface temperature cloud mask. This cloud mask uses the difference between the bulk sea surface temperature and the brightness temperature at 11.0 μm. The thresholds for these differences are 3.0, 2.5, and 2.0 K for low, middle, and high confidence of clear sky, respectively.

As detailed above, each of these cloud masks has threshold ranges set to give confidence levels for clear sky/cloudiness. These are then combined to give an overall confidence level cloud mask described as: confidently clear; probably clear; probably cloudy; or confidently cloudy. For this study, probably cloudy and confidently cloudy were used to define a binary cloud mask that could be compared with the other cloud masks.

(5) A hybrid cloud mask developed for this study. Initial test scene results for cyanobacteria blooms indicated that cloud masks that employed visible and/or near-infrared spectral bands could be causing incorrect masking of this particular marine phenomenon (see Section 3). Therefore, a new cloud mask was formulated by removing the near-infrared and near-infrared/visible ratio tests from the MODIS Atmosphere Group cloud mask and replacing them with the WS06 cloud mask. The final hybrid cloud mask then followed the same method as the original MODIS Atmosphere Group mask, except that heavy aerosol and suspended dust tests were not used as these also utilized visible and near-infrared bands.

(6) The operational MERIS cloud mask (L2ESA). Initially, this cloud mask screens for bright pixels using two reflectance tests in the blue wavelengths: the first test employing a look-up table of the maximum reflectance at 442 nm over oceanic waters depending on the sun and observation geometry; the second test using a single global threshold on the Rayleigh reflectance. Further refinement is carried out using the 'apparent height of the scatterer (Pscatt)' which is derived from the near-infrared bands of MERIS, and in particular the oxygen absorption characteristics of band 11 at 760 nm (ESA 2011).

Each of the cloud masks was implemented and visualized using custom written IDL programs and the NASA SeaDAS software. The MERIS cloud mask was extracted from standard ESA level 2 products. The BEAM-Visat software (v4.11) from Brockmann Consult GmbH was also used to aid in the visualization and testing of cloud mask extraction from the MERIS data.

### 2.3. Test scenes

Potential optical extremes in the ocean that were examined here were:

- intense blue/green algal (cyanobacteria) blooms, for example, in the Baltic Sea;
- intense coccolithophore blooms, for example, in the Black Sea and the North West European Shelf;
- other extreme phytoplankton/algal blooms, for example, the intense blooms and occasional red tides associated with the strong upwelling off the west coast of Africa;
- large sediment loads due to outflow from major rivers into the ocean, for example, the Amazon and Congo rivers.

Even though these examples do not represent an exhaustive list of optical extremes, they are often associated with a strong reflectance signature and were selected as a diverse set of candidates to incorrectly trigger the standard NASA cloud mask. Initially, each optical extreme was examined using the best available imagery from SeaWiFS, MODIS Aqua, and MERIS to determine whether the standard NASA cloud mask was incorrectly masking any of the clear-sky data. Following the application of the cloud masks to these test scenes, time series analysis was performed on selected areas as described below.

### 2.4. Time series analysis

The cloud masking programs (Section 2.2) were adapted to run in batch mode over thousands of images to produce results for the time series chosen (Figure 1). For the Baltic Sea, this was applied to 2003 and 2005, for the Black Sea to 2003, and for the Amazon region to 2003, years where test scenes had shown problems with the standard cloud masking (see Section 3). Each scene was re-mapped onto a fixed geographical domain (2 km grid), excluding pixels associated with large solar zenith angles or viewing angles (larger than 70º and 60º, respectively). These limits are those enforced for the creation of level 3 products in operational processing so that statistics derived in the current work are meaningful in the context of actual time series. For each test year and grid point and for each cloud masking scheme, the number of days considered clear sky or cloudy is counted. Similarly, for each pair of cloud masking schemes, the number of days when the two schemes agree or disagree is recorded.

A given satellite pixel can be characterized by a continuum of conditions that go from clear-sky conditions to thick clouds, without mentioning the complicating effect of aerosols, so that selecting one cloud data set as ‘truth’ is somewhat arbitrary. In this work, comparisons between products were carried out by quantifying the relative change away from the STD cloud mask. In this way, the individual scene testing was designed to highlight which type of extreme optical conditions in the ocean may be affected and give some indications of the performance of cloud masking algorithms. The time series analyses were also designed to quantitatively show how much the data might be affected through temporal scales of a year or more and also whether there are any spatial patterns within the relevant geographical areas. It was for this reason that the analyses were extended to yearly time series of the Black Sea (2003), the Baltic Sea (2003, 2005), and the Amazon outflow into the Atlantic (2003). The year 2003 was chosen for two reasons: because it allows the possibility to use SeaWiFS, MODIS, and MERIS data; and because there were extensive

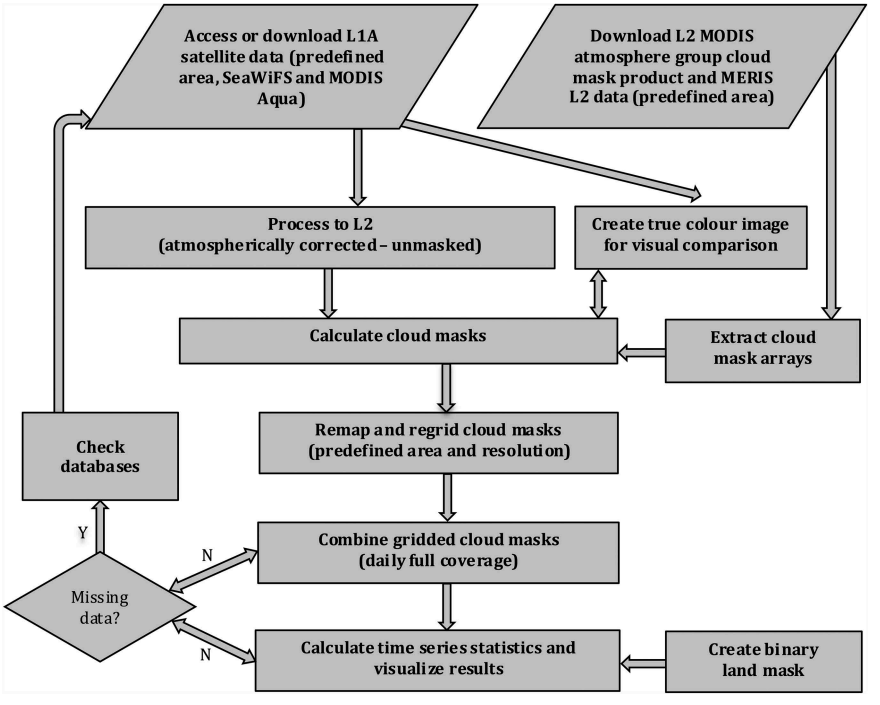


Figure 1. Summary of the automation of the cloud masking process for time series analysis.

coccolithophore blooms in the Black Sea, a large amount of suspended sediment observed in the Azov Sea, and previously observed problems with the cloud masking of single scenes of the Amazon outflow. The year 2005, on the other hand, was chosen exclusively as an addition to the analysis of the Baltic Sea because the most intense cyanobacteria bloom ever recorded by satellite occurred during that year.

## 3. Results

### 3.1. *Cyanobacteria/blue–green algal blooms*

On 8 July 2005, large and intense cyanobacteria blooms were observed in the Baltic Sea. Figure 2(*a*) offers a coloured rendition of the basin at that time and the possibility of visually assessing the extent of cloud coverage. In Figures 2(*b*) (for MODIS) and (*h*) (for MERIS), the erroneous masking of a large part of this bloom by the standard cloud mask can be seen. Figure 2 and Table 1 also give an indication of the relative performance of the various alternative cloud masks. The cloud masks that use the visible and near-infrared wavelengths, N09 (Figure 2(*c*)) and MAG (Figure 2(*e*)), had similar problems to the standard cloud mask (Figure 2(*b*)) and incorrectly masked a portion of the algal bloom. Furthermore, the N09 cloud mask missed a significant amount of the thinner cloud over Denmark and southern Sweden. The MAG cloud mask also seemed to underestimate the extent of cloud (e.g. the outcropping of cloud between southern Sweden and northern Germany). The WS06 mask (Figure 2(*d*)) is not affected by the brightness of the algae in the near infrared and thus did not mask any of the bloom area. However, it missed some thin clouds between Denmark and northern Germany.

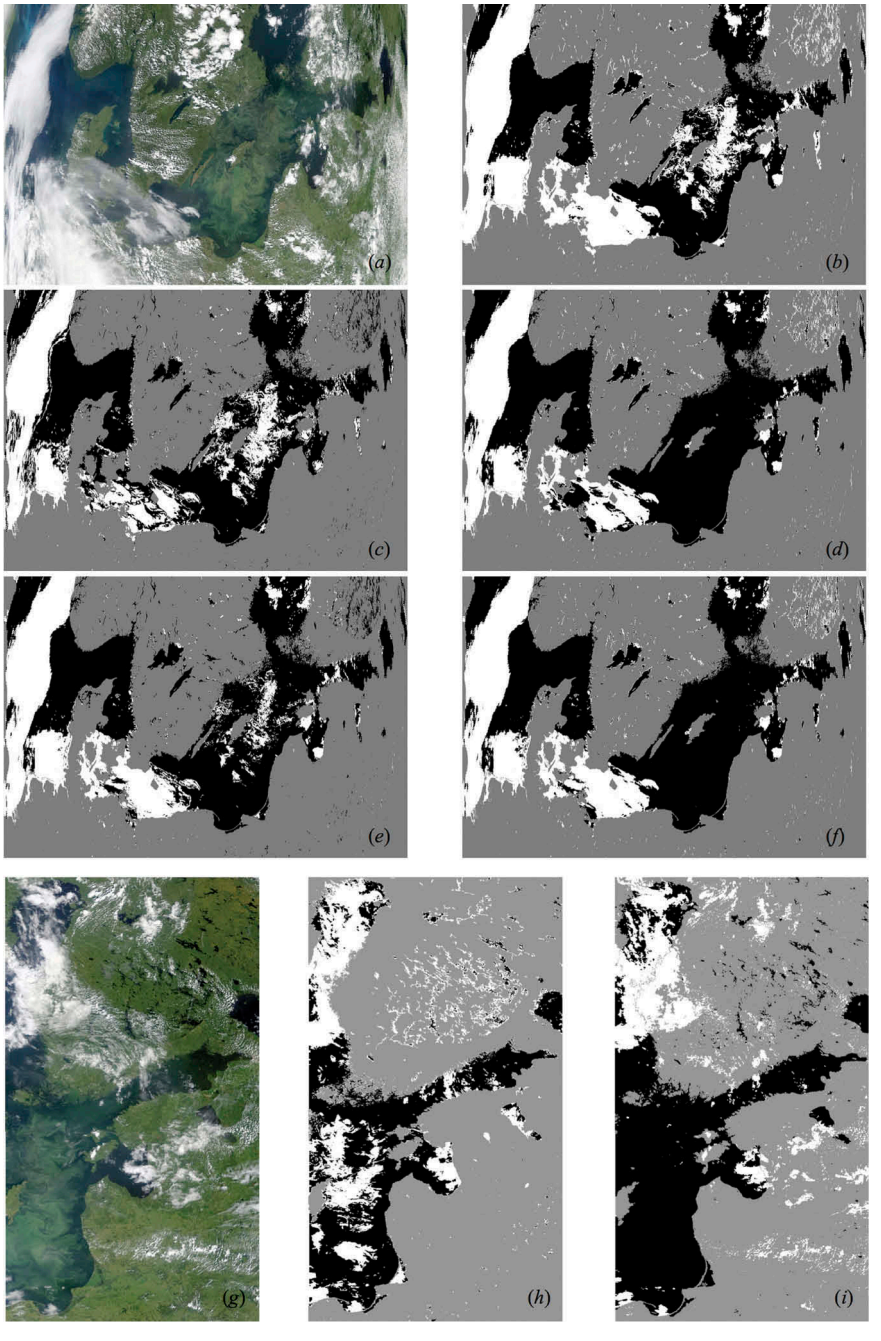


Figure 2. Cloud masks applied to MODIS Aqua and MERIS images of a cyanobacteria bloom in the Baltic Sea (8 July 2005): (*a*) Rayleigh-corrected MODIS Aqua RGB image (645, 555, 469 nm; latitude, longitude of corners: 57.04° N, 4.86° E top left; 62.97° N, 36.24° E top right; 49.35° N, 3.10° E bottom left; 54.11° N, 36.35° E bottom right); (*b*) STD (grey = overlaid land mask, black = clear atmosphere over water, white = cloud); (*c*) N09; (*d*) WS06; (*e*) MAG; (*f*) Hybrid; (*g*) Rayleigh-corrected MERIS RGB image (665, 560, 412 nm; 66.08° N, 21.67° E top left; 64.47° N, 34.17° E top right; 54.40° N, 17.25° E bottom left; 53.12° N, 26.22° E bottom right); (*h*) STD; (*i*) L2ESA.

Table 1. Statistics for the cloud masking of the 8 July 2005 MODIS Aqua image of a cyanobacteria bloom in the Baltic Sea (statistics computed excluding the North Sea area).

| Sensor | Cloud mask | Change in total clear-sky pixels from the standard cloud mask (%) |
|---|---|---|
| MODIS Aqua | STD | – |
| MODIS Aqua | N09 | +1.7 |
| MODIS Aqua | WS06 | +27.0 |
| MODIS Aqua | MAG | +19.8 |
| MODIS Aqua | Hybrid | +16.8 |
| MERIS | STD | – |
| MERIS | L2ESA | +31.9 |

The hybrid and L2ESA cloud masks (Figures 2(*f*) and (*h*)) did not erroneously mask out the cyanobacteria bloom and also visually appeared to do well masking for the rest of the clouds in the image. These visual interpretations are supported by the statistics of cloud coverage (Table 1). The N09 mask shows very little difference (+1.7%), whereas, in contrast, the WS06 mask shows the highest relative change for MODIS (+27.0%), suggesting a possible cloud underestimation problem for WS06. For the MAG cloud mask, the incorrect masking of the bloom is offset in the percentage change results (+19.8%) by its underestimation of cloud coverage in other areas. The hybrid cloud mask appears to strike a balance between preserving the bloom area and effectively masking the cloudy areas in the southwestern part of the image. From this single scene analysis, it appeared to mask real clouds correctly with perhaps an underestimation in the areas of thin cloud (which might be related to the fact that tests for heavy aerosol and suspended dust over water were not included – see Section 2.2). The L2ESA mask gave similar good results although the statistics cannot be directly compared with those of MODIS as the two images cover different parts of the Baltic Sea.

The temporal analysis of cloud masking in the Baltic Sea is restricted to Baltic waters in the window 52.75–66.00º N, 3.50–30.50º E. The results show similar overall levels of difference to the STD cloud mask as the single scene results (see Tables 1 and 2). Figure 3

Table 2. Statistics for the 2005 cloud masking time series for the Baltic Sea (52.75–66.00° N, 3.50–30.50° E – statistics exclude non-Baltic waters).

| Sensor | Cloud mask | Average number of clear-sky days, rounded to whole days (2005 time series) | Change in total clear-sky pixels from the standard cloud mask (%) |
|---|---|---|---|
| MODIS Aqua | STD | 109 | – |
| MODIS Aqua | N09 | 119 | +8.6 |
| MODIS Aqua | WS06 | 149 | +35.6 |
| MODIS Aqua | MAG | 139 | +27.2 |
| MODIS Aqua | Hybrid | 129 | +17.5 |
| MERIS | STD | 78 | – |
| MERIS | L2ESA | 108 | +39.2 |

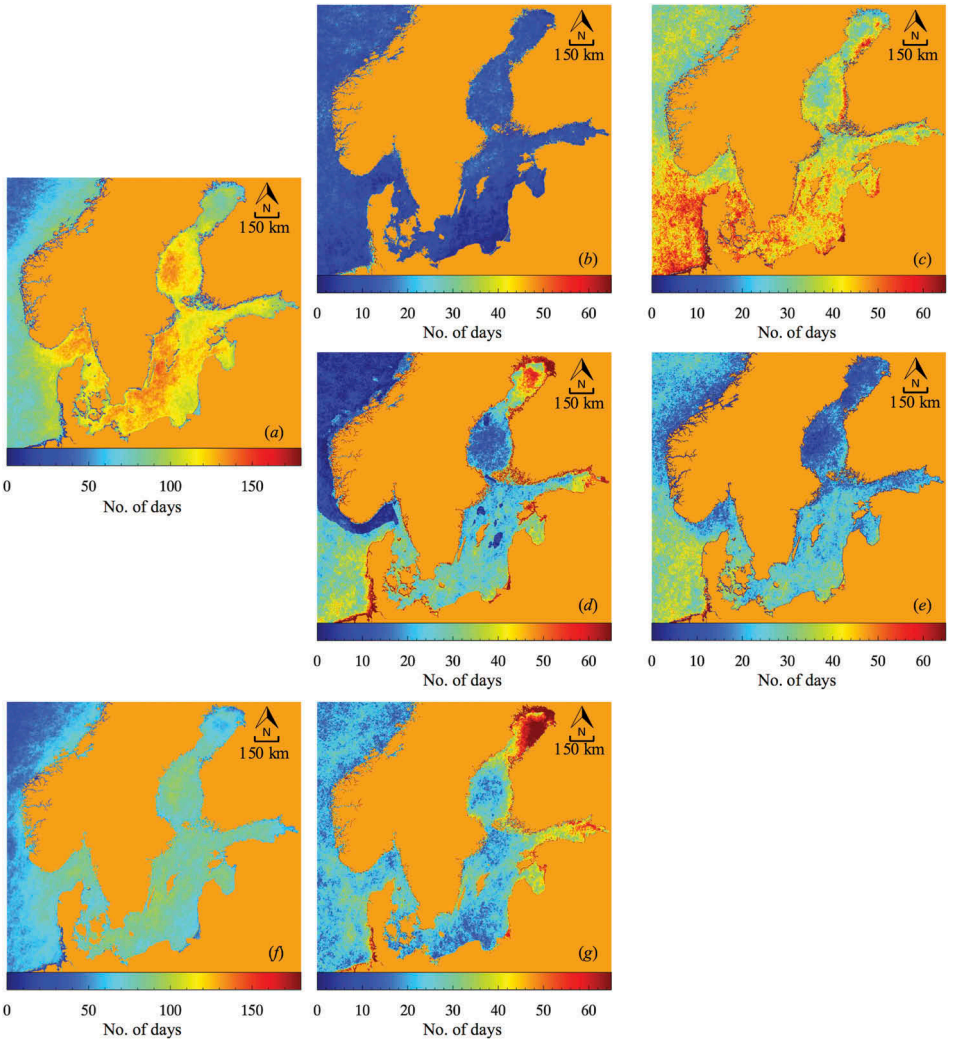


Figure 3. Results of yearly cloud masking analysis for the Baltic Sea (52.75–66.00° N, 3.50–30.50° E). Land is masked in orange and colour scale is the number of days as indicated. The example year is 2005. Results for MODIS Aqua: (*a*) number of clear days, STD; (*b*) difference between STD and N09; (*c*) difference between STD and WS06; (*d*) difference between STD and MAG; (*e*) difference between STD and Hybrid. Results for MERIS: (*f*) number of clear days, STD; (*g*) difference between STD and L2ESA.

also serves to highlight any spatial patterns there may be in the cloud masking through time. For example, apart from the N09 cloud mask, which shows very little difference to the STD cloud mask, there does appear to be a faint pattern of greater increase in clear days in the southern Baltic in the main bloom area for the alternative masks applied to MODIS. This increase in spatial coverage can be associated with the lower sensitivity to surface water conditions of cloud masks relying only on infrared bands. This may also be partly due to the meridional gradient in data coverage in the basin, with less data northwards. The increase in clear-sky days offered by the MAG cloud mask (Figure 3(*d*))

shows sharp features with values becoming very low around Norway and the central Baltic. This is explained by the use of bathymetry in some of the tests making up this cloud mask, specifically distinguishing from the EOS Land/Sea mask between shallow, moderate, and deep ocean (Ackerman et al. 2010). The largest increase in clear-sky days is associated with the WS06 cloud masking, with an average of 149 days over the basin (Table 2). MERIS results are dominated by a large increase in clear-sky days in the northern Baltic Sea (Gulf of Bothnia). This may be due to a greater ability to distinguish between cloud and bright sea ice.

Figure 4 highlights how the differences between cloud masks changed through the year for MODIS and MERIS (Figures 4(*a*) and (*b*), respectively). There are fewer clear days in winter because of a more frequent occurrence of clouds in that season as well as an absence of data collection for very high solar zenith angles. The ranking between the cloud masking schemes is in a varying extent conserved through the year: the schemes N09, hybrid, MAG, and WS06 offer an increasing number of clear days. Only in winter does the N09 cloud mask provide a few more clear days than the other schemes. For MERIS, the relative increase in clear days with the L2ESA cloud mask varies around 30%, except in February–March when it exceeds 80%.

Table 3, a matrix compiling the cases of similar or differing outcomes of the cloud masking approaches, shows that the number of cases where the STD cloud mask diagnosed clear sky and the alternative cloud masks diagnosed cloudy conditions for any given pixel is very low, not exceeding 1%. In line with Table 2, the WS06 scheme shows the largest difference with the STD cloud mask, with ~11% of cases when the two

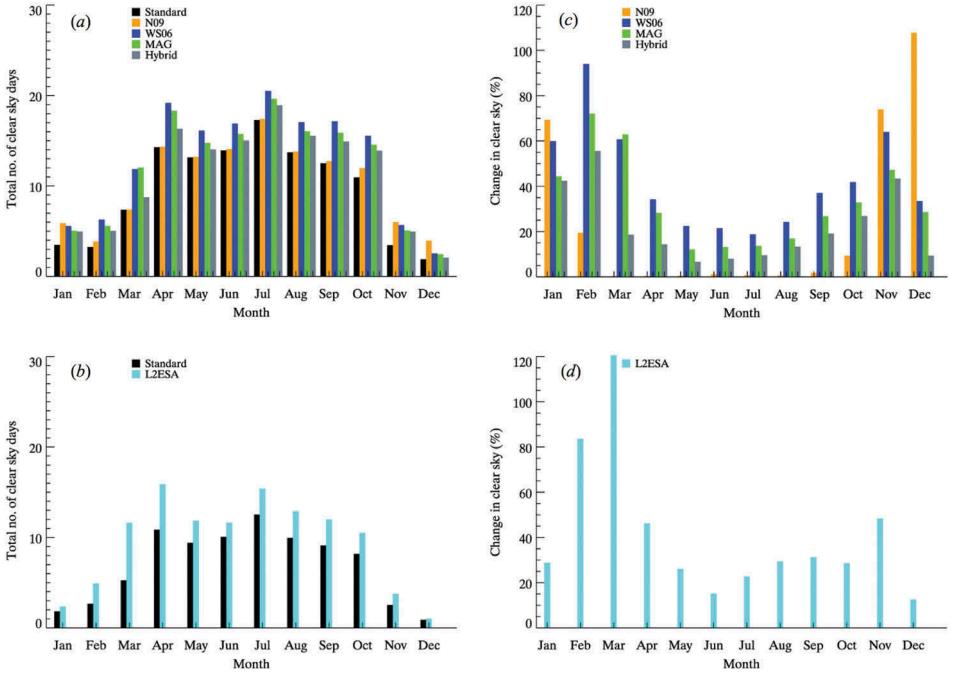


Figure 4. Total clear-sky days for (*a*) MODIS Aqua and (*b*) MERIS and the percentage change in clear-sky days per month for (*c*) MODIS Aqua and (*d*) MERIS for the Baltic Sea in 2005.

Table 3. Contingency matrix for the 2005 cloud masking time series for the Baltic Sea (52.75–66.00° N, 3.50–30.50° E – statistics exclude non-Baltic waters).

| Sensor | Cloud mask comparison | Proportion of time series where STD clear – other cloudy (%) | Proportion of time series where STD cloudy – other clear (%) | Proportion of time series where both cloudy or clear (%) |
|---|---|---|---|---|
| MODIS Aqua | STD vs. N09 | 0.0 | 2.6 | 97.4 |
| MODIS Aqua | STD vs. WS06 | 0.2 | 10.9 | 89.0 |
| MODIS Aqua | STD vs. MAG | 0.3 | 8.5 | 91.2 |
| MODIS Aqua | STD vs. Hybrid | 1.1 | 6.3 | 92.6 |
| MERIS | STD vs. L2ESA | 0.3 | 8.7 | 91.0 |

schemes disagree, the former diagnosing clear-sky conditions. It is noted that the percentage values listed in the contingency matrix (Table 3 for the Baltic Sea) are lower than those quantifying the relative increase in the number of clear-sky days (Table 2). This is explained by the inclusion in the contingency matrix of all the cases when both masking approaches classify a pixel as cloud, largely increasing the number of cases where both schemes agree.

### 3.2. *Coccolithophore blooms*

The results of the examination of single scenes affected by coccolithophore blooming in both the Black Sea (Figure 5(*b*)) and the North West European Shelf showed that these blooms are not masked incorrectly by the STD cloud mask. It should be noted that the often sediment loaded Azov Sea, which is part of the Black Sea region, did show some problems with the STD mask only in the very shallow northeast (Figure 5(*d*)).

As for the Baltic Sea, a temporal analysis was conducted for the Black Sea region with the test year 2003 and the window 40.75–47.50º N, 27.25–42.00º E. Again, the N09 cloud mask gave almost the same results as the STD mask for both SeaWiFS and MODIS data, as reflected both in the difference images and the percentage change

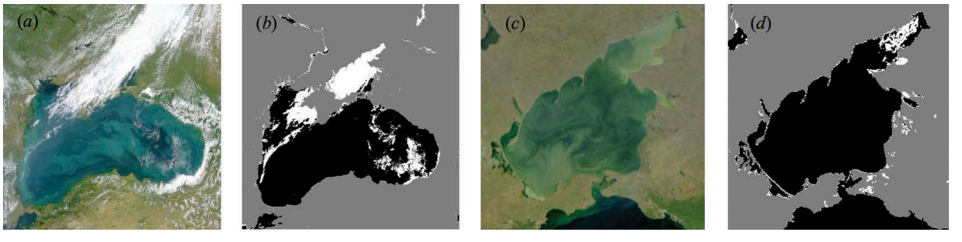


Figure 5. Standard cloud mask applied to SeaWiFS images of a coccolithophore bloom in the Black Sea (15 June 2002) and high sediment load in the Azov Sea (2 October 2000): (*a*) Rayleigh-corrected RGB image (670, 555, 412 nm; 15 June 2002; latitude, longitude of corners: 50.59° N, 27.21° E top left; 47.55° N, 47.30° E top right; 40.67° N, 25.41° E bottom left; 38.13° N, 42.51° E bottom right); (*b*) STD (grey = overlaid land mask, black = clear atmosphere over water, white = cloud); (*c*) Rayleigh-corrected RGB image (670, 555, 412 nm; 2 October 2000; 47.95° N, 35.14° E top left; 47.24° N, 39.71° E top right; 45.12° N, 34.12° E bottom left; 44.43° N, 38.23° E bottom right); (*d*) STD.

statistics (0.5 and 0.4%, respectively; Figure 6 and Table 4). The other cloud masks indicate an increase in clear-sky conditions through the year (between 26% and 41%, or 27 to 51 days), with the WS06 mask giving the highest increase and the other two remaining approximately the same. There do not appear to be any discernible spatial patterns in the results, with the exception of shallow coastal areas and river mouths such as the Danube and northeast Azov Sea where there is an increase in the number of clear days according to all the alternative cloud masks (Figure 6). In the MAG results, again there are some sharp gradients associated with bathymetry (transition between the northwest shelf and the deep basin, Figure 6(*f*)) that affects some tests of the scheme (see Section 3.1). The MAG cloud scheme also shows a relatively larger increase in clear-sky conditions in the Azov Sea. On the other hand, in the central basin, the hybrid mask shows a slightly higher increase in clear sky with respect to the MAG mask (Figure 6(*g*)). This could be due to the hybrid mask underestimating the amount of thin cloud or heavy aerosol/dust because the tests for heavy aerosol and suspended dust over water were not included (see Section 2.2). The changes through the year, as seen in Figure 7, show a clear annual cycle in the overall number of clear

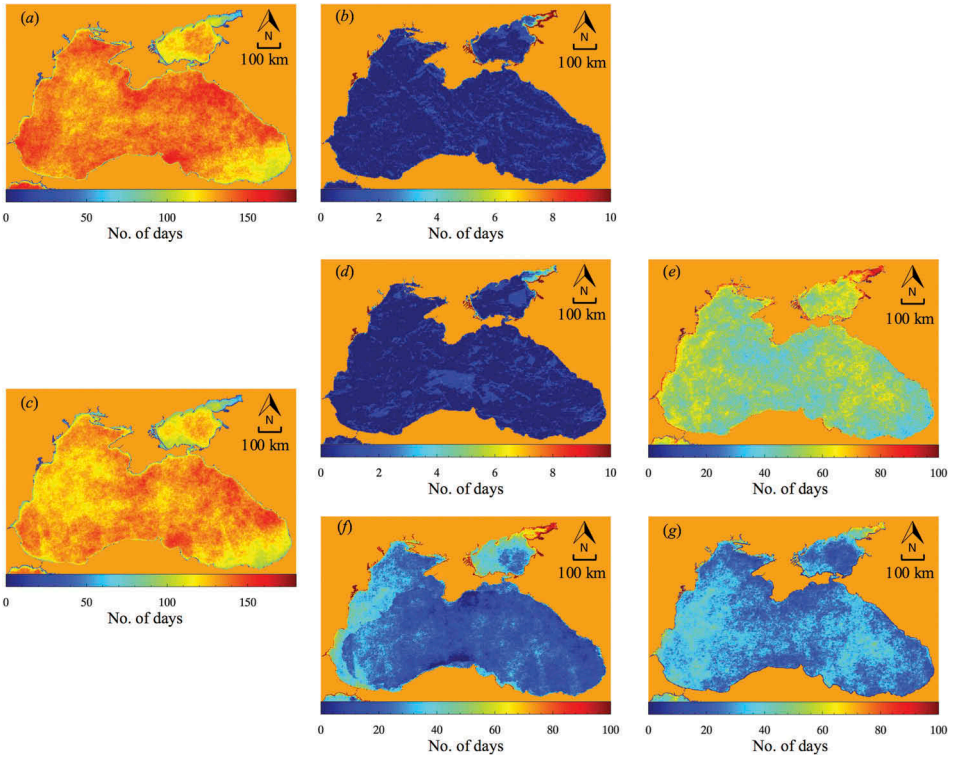


Figure 6. Results of yearly cloud masking analysis for the Black Sea (40.75–47.50° N, 27.25–42.00° E). Land is masked in orange and colour scale is the number of days as indicated. The example year is 2003. Results for SeaWiFS: (*a*) number of clear days, standard cloud mask; (*b*) difference between STD and N09. Results for MODIS Aqua: (*c*) number of clear days, standard cloud mask; (*d*) difference between STD and N09; (*e*) difference between STD and WS06; (*f*) difference between STD and MAG; (*g*) difference between STD and hybrid.

Table 4. Statistics for the 2003 cloud masking time series for the Black Sea (40.75–47.50° N, 27.25–42.00° E).

| Sensor | Cloud mask | Average number of clear-sky days, rounded to whole days (2003 time series) | Change in total clear-sky pixels from the standard cloud mask (%) |
|---|---|---|---|
| SeaWiFS | STD | 131 | – |
| SeaWiFS | N09 | 132 | +0.5 |
| MODIS Aqua | STD | 126 | – |
| MODIS Aqua | N09 | 126 | +0.4 |
| MODIS Aqua | WS06 | 177 | +41.0 |
| MODIS Aqua | MAG | 158 | +25.5 |
| MODIS Aqua | Hybrid | 161 | +28.6 |

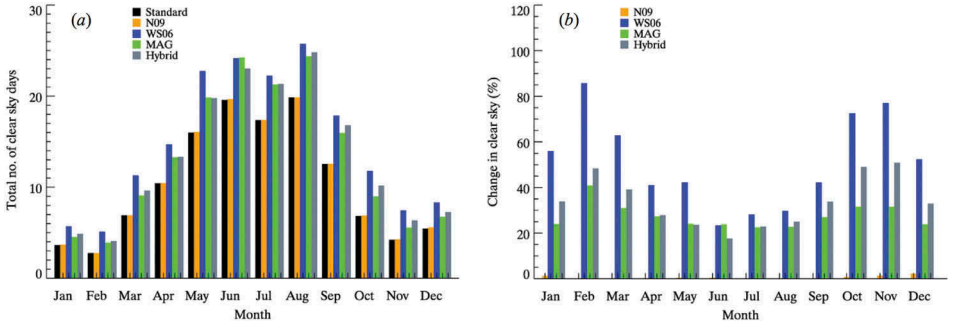


Figure 7. Total clear-sky days (*a*) and percentage change in clear-sky days per month (*b*) for the Black Sea in 2003 – MODIS Aqua.

days for all masks and the difference between masks. Apart from the N09 mask, all the masking schemes show a trend of greatest relative difference in the winter months decreasing through the spring and increasing again through the autumn (see Figure 7). The WS06 cloud masking is the scheme producing the largest amount of clear-sky conditions through the entire year.

### *3.3. Sediment laden river outflows*

Selected SeaWiFS, MODIS Aqua, and MERIS scenes of 2003 were used to show a representative example of the treatment of the Amazon River outflow region by the different cloud masking schemes, with different levels of masking of the sediment-rich area of the Amazon estuary according to the STD mask and all the alternative cloud masks (Figure 8). These results showed incorrect (excessive) masking by the STD mask in the outflow into the sea and along the river's channels. Unlike cyanobacteria and coccolithophore blooms, but similarly to the high sediment loads of the northeast Azov Sea, the N09 mask shows an improvement and a distinct difference to the STD mask in the estuarine area (+15.6 and +4.7%, respectively, for SeaWiFS and MODIS, Table 5). The other alternative cloud masks for MODIS and

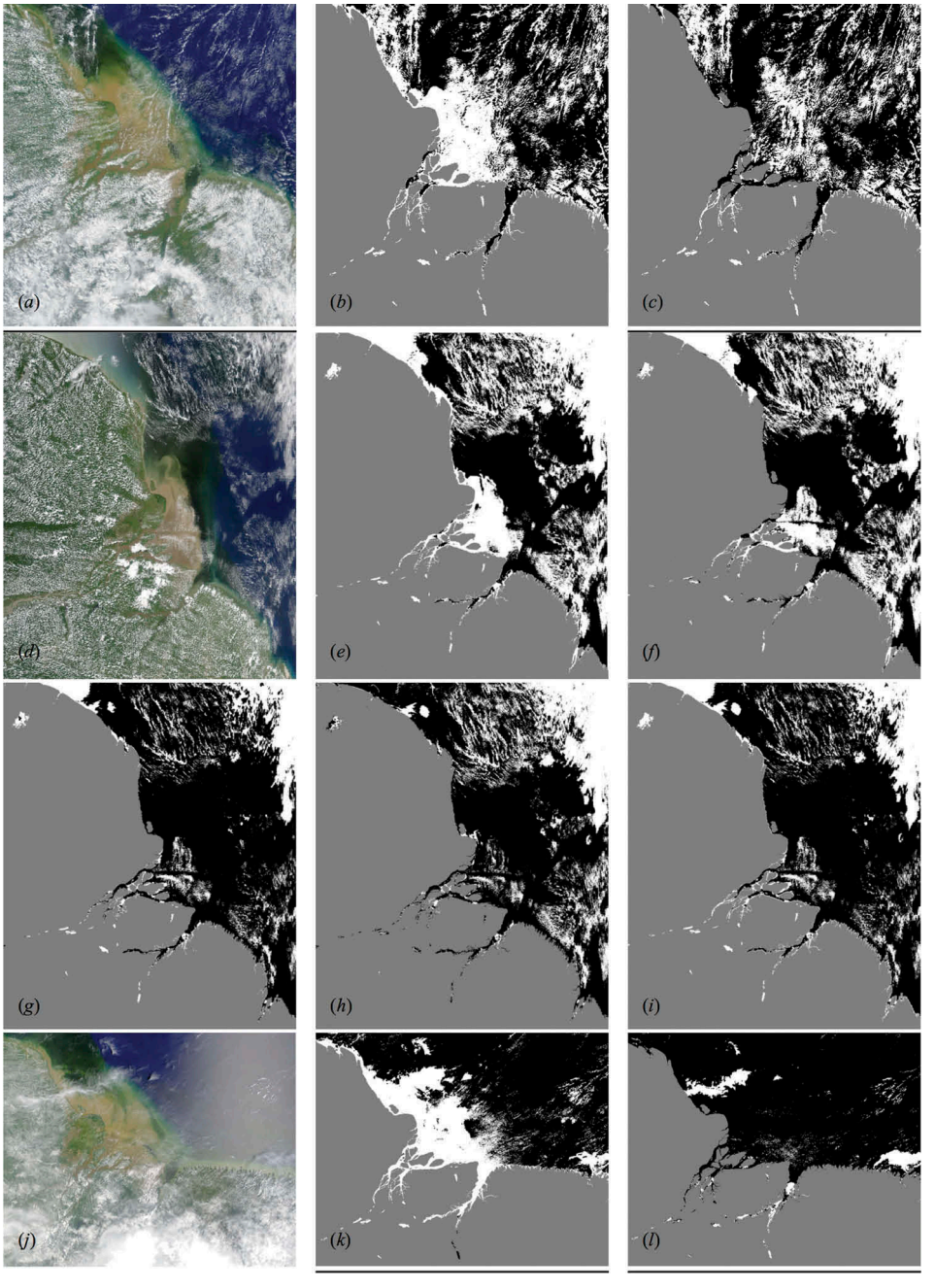


Figure 8. Cloud masks applied to SeaWiFS, MODIS Aqua, and MERIS images of the Amazon River outflow into the western Atlantic: (*a*) Rayleigh-corrected SeaWiFS RGB image (670, 555, 412 nm; 3 November 2003; latitude, longitude of corners: 4.70° N, 52.51° W top left; 3.54° N, 43.34° W top right; 3.02° S, 54.18° W bottom left; 4.13° S, 45.00° W bottom right); (*b*) STD (grey = overlaid land mask, black = clear atmosphere over water, white = cloud); (*c*) N09; (*d*) Rayleigh-corrected MODIS Aqua RGB image (645, 555, 469 nm; 21 August 2003; 5.77° N, 55.85° W top left; 7.35° N, 44.24° W top right; 4.78° S, 53.59° W bottom left; 3.07° S, 42.03° W bottom right); (*e*) STD; (*f*) N09 cloud; (*g*) WS06; (*h*) MAG; (*i*) Hybrid; (*j*) Rayleigh-corrected MERIS RGB image (665, 560, 412 nm; 4 November 2003; latitude, longitude of corners: 4.86° N, 52.17° W top left; 2.83° N, 42.96° W top right; 2.69° S, 53.83° W bottom left; 4.73° S, 44.57° W bottom right); (*k*) STD; (*l*) L2ESA.

Table 5. Statistics for the cloud masking of the SeaWiFS, MODIS Aqua, and MERIS 2003 images of the Amazon outflow region (Figure 8).

| Sensor | Cloud mask | Change in total clear-sky pixels from the standard cloud mask (%) |
|---|---|---|
| SeaWiFS | STD | – |
| SeaWiFS | N09 | 15.6 |
| MODISA | STD | – |
| MODISA | N09 | +4.7 |
| MODISA | WS06 | +48.1 |
| MODISA | MAG | +49.8 |
| MODISA | Hybrid | +43.4 |
| MERIS | STD | – |
| MERIS | L2ESA | +49.7 |

MERIS all show a much larger difference to the STD mask (43–49%), with clear-sky conditions diagnosed for almost the entire estuary, but also more clear-sky conditions far offshore.

The temporal analysis over the Amazon region is conducted over the window 3.0º S–3.5º N, 52.5º–45.0º W for the test year 2003. The results show similar trends in the differences to the STD cloud mask as the single scene results (see Tables 5 and 6), but higher differences because statistics are computed over a smaller area (compare Figures 8 and 9). From Figure 9 and Table 6, it can initially be noted that the overall levels of cloudiness are much higher and the number of clear days much lower in the selected Amazon region than in the areas considered previously (56 vs. 126 and 109 clear days for the Black and Baltic Seas, respectively, according to the STD mask applied to MODIS). There is also a clear spatial pattern in the clear-sky distribution, with the lowest values associated with the outflow region (Figures 9(*a*), (*c*), and (*h*)); as already suggested by Figure 8, this indicates that the STD cloud mask systematically excludes conditions of high sediment concentrations.

Table 6. Statistics for the 2003 cloud masking time series for the Amazon River outflow into the Atlantic Ocean (3.00° S–3.50° N, 52.50–45.00° W).

| Sensor | Cloud mask | Average number of clear-sky days, rounded to whole days (2003 time series) | Change in total clear-sky pixels from the standard cloud mask (%) |
|---|---|---|---|
| SeaWiFS | STD | 60 | – |
| SeaWiFS | N09 | 72 | 21.6 |
| MODIS Aqua | STD | 56 | – |
| MODIS Aqua | N09 | 63 | 10.8 |
| MODIS Aqua | WS06 | 128 | 126.2 |
| MODIS Aqua | MAG | 130 | 129.9 |
| MODIS Aqua | Hybrid | 111 | 96.1 |
| MERIS | STD | 54 | – |
| MERIS | L2ESA | 96 | 78.7 |

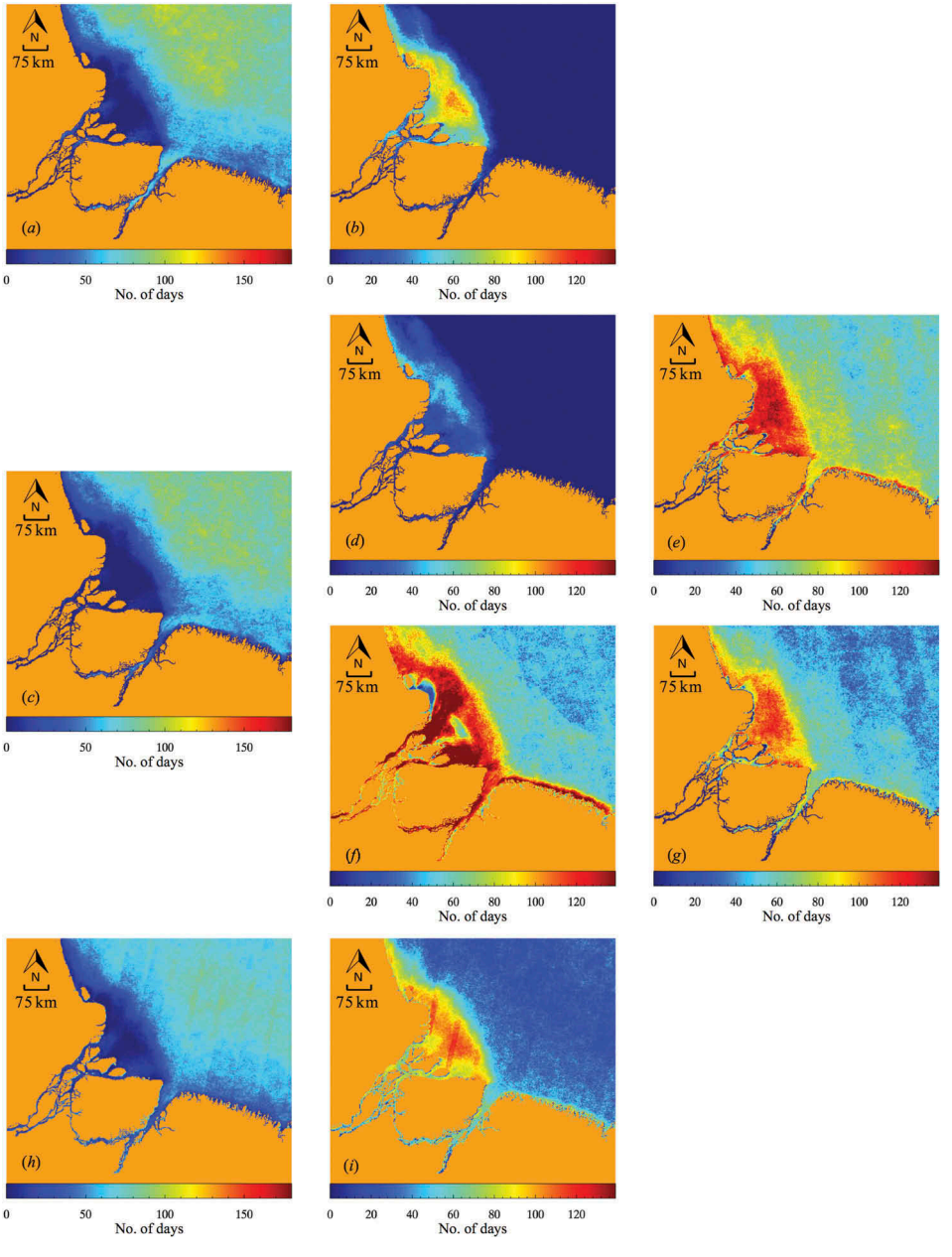


Figure 9. Results of yearly cloud masking analysis for the Amazon River outflow into the Atlantic Ocean (3.00° S–3.50° N, 52.50–45.00° W). Land is masked in orange and colour scale is the number of days as indicated. The example year is 2003. Results for SeaWiFS: (*a*) number of clear days, standard cloud mask; (*b*) difference between STD and N09. Results for MODIS Aqua: (*c*) number of clear days, standard cloud mask; (*d*) difference between STD and N09; (*e*) difference between STD and WS06; (*f*) difference between STD and MAGa; (*g*) difference between STD and hybrid. Results for MERIS: (*h*) number of clear days, standard cloud mask; (*i*) difference between STD and ESA cloud mask.

Although the N09 mask only gives 7 and 12 extra clear days per year for MODIS Aqua and SeaWiFS, respectively (Table 6), this translates to a percentage change of 10–20% given the low baseline from the STD mask. More importantly, the increase in clear-sky conditions is well associated with the outflow region and does not extend offshore (Figures 9(*b*) and (*d*)). The other cloud masks show many more clear days than the N09 mask. Again because of the low baseline, this translates into percentage increases between 96 and 129%. As for the N09 mask, a definite bias towards greater increase in clear-sky days can be seen for all the alternative cloud masks in the areas of highest estuarine sediment occurrence. Even if the amplitudes are lower, significant increases extend offshore for the entire domain (Figures 9(*e*), (*f*), (*g*), and (*i*) to a lesser extent). The increase in clear-sky conditions also extends upstream in the river channels, particularly for the MAG scheme (Figures 9(*f*) and (*i*)).

The annual cycle found for the cloud coverage (Figure 10) appears to reflect the precipitation in the Amazon basin (Zeng 1999). The number of clear-sky days is lowest in boreal winter during the rainy season, while the relative increase in clear-sky conditions diagnosed by alternative masking exceeds 100% in this period for MODIS and MERIS. The WS06 and MAG schemes compete for the highest relative increase, with the latter providing the largest score from August to October.

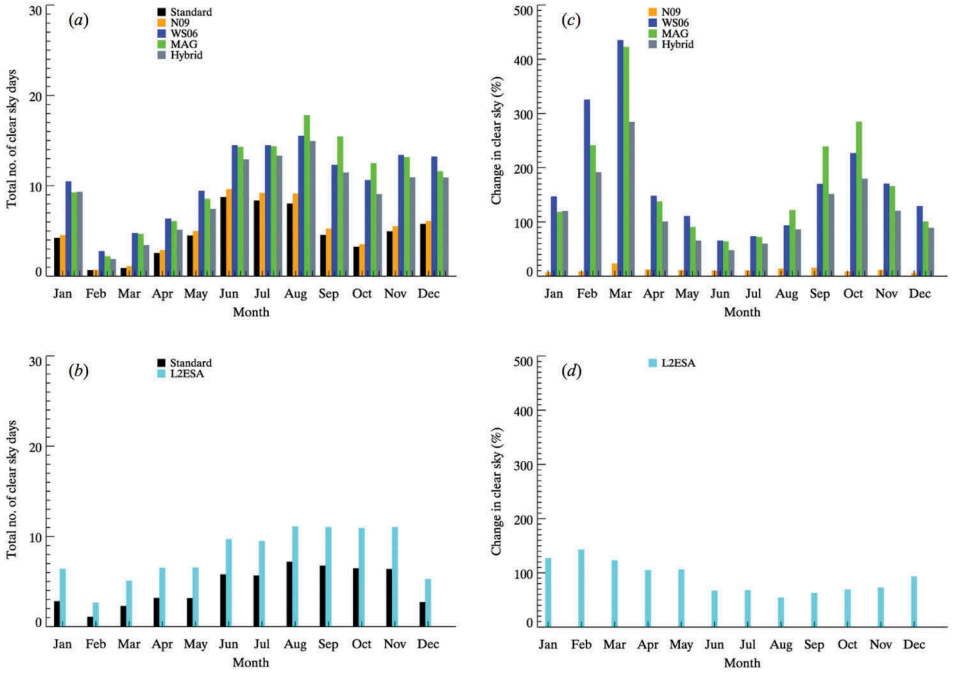


Figure 10. Total clear-sky days for (*a*) MODIS Aqua and (*b*) MERIS, and percentage change in clear-sky days per month for (*c*) MODIS Aqua and (*d*) MERIS for the Amazon River outflow into the Atlantic Ocean in 2003.

## 4. Discussion

Generally, wherever the reflectance of the ocean in the near infrared is particularly bright, this area is excluded from further processing by the so-called cloud mask (STD) of the NASA ocean colour processing chain. For marine scenes, this may happen in the presence of clouds, thick aerosol plumes, ice, or unusually bright surface water phenomena. It is noted that the regions studied here are far from sources of dense aerosol concentrations from biomass burning or desert dust, and usually not affected by ice (except for the northern Baltic and Azov waters in winter when ocean colour imagery is anyway restricted).

The initial testing of the STD mask on single scenes representative of extreme optical situations in the world's oceans indicated that there is only a limited set of circumstances where this algorithm incorrectly masks water surface phenomena as cloud. The adversely affected situations, where there is a significant loss of data, are very intense cyanobacteria blooms and particular types of sediment-laden waters, either in coastal estuaries or in the outflow from large rivers (see Figures 2 and 8). Following testing on multiple scenes of intense coccolithophore and phytoplankton blooms, including those that have led to red tides (e.g. in the Benguela upwelling region, off the southwest coast of Africa), it was concluded that these optical extremes do not exhibit the same response in the near infrared and thus are not mistaken for cloud by the STD mask. However, it cannot be totally excluded that particularly extreme cases that were not surveyed in this study could display anomalous responses.

In order to try and confirm these results through time, time series analysis was conducted for selected regions. The Azov Sea was considered together with the Black Sea, being a heavily sediment laden and shallow sea. For this region, the results can be seen in Figures 6 and 7 and Table 4. The only notable differences between the STD mask and the N09 mask were found in some near coastal estuarine areas, such as the Danube delta and the northeast Azov Sea, possibly due to high sediment loads near the surface. It is also noticeable that the rest of the Azov Sea showed little difference over the whole year even though it is the shallowest sea in the world. This may be explained if high concentrations of sediment are not close enough to the surface, except in the northeast, to influence reflectance at the near-infrared wavelengths. For the Black Sea proper, it is notable that the WS06 mask diagnosed significantly more clear-sky conditions across the basin and the annual cycle, indicating that it is systematically more tolerant with respect to cloudy conditions. There did not seem to be any strong spatial patterns in the differences between the cloud mask outputs. These results may indicate that the observed increases are simply related to a cumulative effect throughout the year, with the alternative cloud masking being systematically less conservative compared with the STD mask.

For those conditions in the ocean that are incorrectly masked by the STD mask, it was found that the effects are not universal across the globe. For example, the large sediment outflow from the Congo River into the eastern Atlantic might reasonably be expected to be incorrectly masked, but all examined imagery does not suffer the same cloud masking problems as the Amazon and other major rivers around the world, for example the Yangtze and the Rio de la Plata estuary (Nordkvist, Loisel, and Duforet Gaurier 2009). This may indicate that the type and depth of sediment plume from a river is influential, with the shallow deposition inclined environment of the Amazon estuary, and those of the other rivers mentioned, producing higher reflectance values in the near infrared. In contrast, the deeper channel of the Congo limits sediment deposition near its mouth

with its course continuing offshore for hundreds of kilometres as a deep-sea fan and underwater canyon system (Savoye et al. 2009).

The outflow from the Amazon River exits on the edge of a tropical rain forest environment with a dry season (June–September) and a rainy and thus cloudier season (rest of the year). The low baseline of clear days per year according to the STD mask reflects this climate. In contrast to the Black Sea and the Congo River, results found in the Amazon delta showed very spatially distinct misclassifications in the STD masking when single scenes were examined (see example in Figure 8). From the time series analysis, a high level of difference between the STD and other cloud masks was observed with a very clear spatial pattern associated with the delta. The results reported by Nordkvist, Loisel, and Duforet Gaurier (2009) indicated that their cloud mask was performing well for the Amazon and other river outflows, which is more generally supported by the results given here even though the amount of additional coverage in the estuarine region is not as high as for other cloud masks. It is worth noting that the L2ESA mask was also not incorrectly triggered by the sediment-laden waters and appeared to correctly identify real cloud (see Figures 8(*l*) and 9(*i*)). In contrast, the results of other cloud masks (WS06, MAG, and hybrid) appeared to greatly underestimate the amount of cloud, including offshore. This likely results from a combination of less incorrect masking of the highly sediment loaded estuarine and river channel areas than the STD and N09 masks, and underestimation of thin cloud or high levels of aerosols. Some of this clear-sky overestimation may have been due to the data quality of some of the infrared MODIS bands used and seen as striping in some of the time series results.

The individual scene testing of optical extremes highlighted intense cyanobacteria blooms in the Baltic Sea as the main algae-related condition that inadvertently triggers the STD mask, because of a large reflectance signal in the near infrared (Kutser 2004). However, in the Baltic Sea, the bloom of cyanobacteria and the accompanying detrital material needs to be very intense to do this over large areas; for example, the extreme blooms in 1999, 2003, 2008 (results not shown) and 2005 (see Figures 2 and 3).

Visible cyanobacteria occur nearly every year in the Baltic Sea (Stal et al. 2003; Kahru, Savchuk, and Elmgren 2007; Öberg 2013) but rarely do parts of the bloom trigger the STD mask over large areas of the bloom. Individual scenes from other years (e.g. 2007 and 2009–2012) where the seasonal cyanobacteria bloom was less intense (Hansson and Öberg 2012) were examined and found not to trigger the standard cloud mask to such a large degree as in 2005. Some of the particular atmospheric and environmental conditions that could account for the intense blooms may also be those that determine whether or not the cloud mask is incorrectly triggered. From the individual satellite scenes examined here, it would appear that this depends on how well the bloom is expressed at the surface and this in turn may be related to its intensity. Of further importance may be whether or not the prevailing hydrographic conditions encourage or deter the cyanobacteria and the associated detrital material from accumulating near the surface.

Therefore, in the time series analysis, the greatest increase in clear sky according to the alternative cloud masks may be expected where the blooms are the most intense, but this was not seen clearly in the yearly results. It is likely that the spatial patterns of increased clear-sky frequency are not as distinct as those of the Amazon case because the cyanobacteria blooms are seasonal events that usually occur in the summer for a matter of weeks, whereas the discharge of large quantities of sediment by the Amazon occurs almost year round. Furthermore, the WS06 mask again

appeared to underestimate the amount of cloud throughout the region during the course of the year, which has the tendency of further blurring any spatial preference due to the cyanobacteria bloom. Nevertheless, if it is assumed that the hybrid cloud mask and standard ESA operational cloud mask for MERIS are performing well for the Baltic Sea, as was the case in the individual scene analysis, then the time series results still pointed to a 18% and 39% yearly increase in clear-sky diagnostics compared with the STD approach (for MODIS and MERIS, respectively). When the potential increase of 22% in clear sky for the Amazon using the most conservative alternative mask (N09 – SeaWiFS) is also considered, then the potential loss of data from the ocean colour archive due to incorrect cloud masking can be seen to be regionally significant. As mentioned in Section 1, these particular extremes are important for marine ecosystems and coastal human activities, and this level of data loss is limiting the quantitative description of the marine environment using remote sensing in these regions. Furthermore, if extreme events are indeed becoming more frequent as climate is changing, then this suggests an urgency to complete the ocean colour climate record using more accurate cloud masking to appropriately treat these areas.

Sky conditions may vary from clear to heavily cloudy through a variety of conditions, so it is hard to unambiguously identify a pixel as cloudy and to derive a 'true' cloud coverage. Therefore, this work focused on assessing several cloud masking schemes but did not aim at identifying the best performing one. It is also reminded that the considered schemes are not applicable to all sensors. With respect to other studies that evaluated a limited number of masking schemes tested on a few selected satellite images, the current analysis has extended the assessment to a full annual cycle for each tested scheme to comprehend the consequences of applying one particular approach at the scale of a region and through time. It is only through such comprehensive testing that an informed selection can be made. In the context of ocean colour remote sensing, the definition of cloudy is operational, a condition for which the operation of the atmospheric correction would produce results of marine reflectance with degraded quality with respect to clear-sky conditions. Additional work is required to definitively answer whether any of the alternative cloud masks are conservative enough for application to the entire ocean colour archive and operationally for future ocean colour missions. Generally, contamination by cloud is minimized in the global datasets by the conservative nature of the STD mask maintaining the high quality of the global ocean colour data archive. However, for areas such as the Amazon delta and the Baltic Sea, this research has exposed a clear need for the regional application of alternative cloud masks to the full time series of ocean colour data in these areas to restore excluded data. Without contaminating the processing with extra cloud, encouraging results have been noted for the N09 mask in the Amazon region and the hybrid and L2ESA masks in the Baltic Sea. The WS06 approach appears as the least conservative, systematically producing a significant additional coverage in all cases, including the open Atlantic waters far from the Amazon delta.

Going beyond these results, it is important to bear in mind that an extension of the area diagnosed as clear sky, even if appropriate from the point of view of cloud identification, does not necessarily result in additional coverage with valid marine reflectance values. Indeed, the atmospheric correction might still have difficulty in handling intense surface optical signatures that might trigger additional excluding or warning flags or lead to the failure of the atmospheric correction algorithm. This may be due to known problems that the standard near-infrared atmospheric correction

scheme (e.g., Ruddick, Ovidio, and Rijkeboer 2000) can encounter with turbid waters, such as characterized by intense cyanobacteria blooms and high sediment loads. A preliminary analysis of selected examples in the Baltic Sea has shown that the application of the standard NASA atmospheric correction to the central part of intense cyanobacteria blooms formerly excluded by the standard cloud mask only rarely led to valid reflectance retrievals. So, improvements of cloud masking for ocean colour application need to be accompanied by parallel progress in atmospheric correction algorithms.

## 5. Conclusions

It has been shown here and in previous research that ocean colour processing does not produce valid results in some specific conditions such as some intense algal blooms or river outflows. Progress is needed in atmospheric correction and cloud detection and masking to ensure that these conditions are appropriately detected, monitored, and their inter-annual evolutions studied and included in diverse applications such as ecosystem monitoring, global primary productivity modelling, and climate change studies. This work has highlighted that distinguishing between the predominant state of 'dark' ocean and 'bright' cloud is complicated by some extreme optical situations in parts of the ocean (e.g. intense algal blooms and high sediment loads). This study has focused on assessing cloud masking schemes in relevant representative conditions.

Generally, the NASA operational standard cloud mask for ocean colour data seems to perform adequately for the global SeaWiFS, MODIS Aqua, and MERIS datasets. However, there are a limited number of circumstances, optically and geographically in the ocean, where this cloud mask incorrectly excludes important surface phenomena. This article has highlighted intense cyanobacteria (blue–green algae) blooms and sediment-laden coastal waters from the outflow of major rivers as the main extreme optical conditions where this is the case. In particular, certain conditions in the Baltic Sea and the Amazon River's outflow into the Atlantic Ocean, as case studies, have exemplified this current loss of data. Other sediment-laden waters (e.g. the outflow from the Congo River into the Atlantic), coccolithophore blooms, and other intense phytoplankton blooms do not seem to be incorrectly masked as cloud, even though particularly extreme conditions cannot be excluded. The work also indicates possible alternatives to the standard cloud mask, for example, the hybrid cloud mask proposed here for MODIS, the ESA MERIS cloud masking technique, as well as the N09 cloud mask for the case of river outflow regions. Thus, the results of this assessment recommend working towards an operational cloud mask combining the advantages of a number of different existing masks. This study indicates that such a mask may benefit from the inclusion of short wave and thermal infrared channels as used in the MAG and hybrid cloud masks as well as calculation of the apparent height of the scatterer as used by ESA for MERIS processing. This could ensure that for future missions, such as the Sentinel-3 Ocean and Land Colour Instrument, the exclusion of important data from optical extremes in the ocean and cloud contamination are both minimized.

## Acknowledgements

The authors would like to thank the Ocean Biology Processing Group of NASA and NASA itself for continuing dedication to the principles of open satellite data access for research purposes and

support in the processing of such data. ESA is likewise acknowledged for the distribution of the MERIS imagery.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley. 1998. "Discriminating Clear Sky from Clouds with MODIS." *Journal of Geophysical Research* 103 (D24): 32141–32157. doi:10.1029/1998JD200032.

Ackerman, S., R. Frey, K. Strabala, Y. Liu, L. Gumley, B. Baum, and P. Menzel. 2010. *Discriminating Clear-sky from Cloud with MODIS: Algorithm Theoretical Basis Document (MOD35). Version 6.1.* Greenbelt, MD: NASA Goddard Space Flight Center.

Banks, A. C., P. Prunet, J. Chimot, P. Pina, J. Donnadille, E. Jeansou, M. Lux, G. Petihakis, G. Korres, G. Triantafyllou, C. Fontana, C. Ulses, and L. Fernandez. 2012. "A Satellite Ocean Colour Observation Operator System for Eutrophication Assessment in Coastal Waters." *Journal of Marine Systems* 94: S2–S15. doi:10.1016/j.jmarsys.2011.11.001.

Birks, A. R. 2007. *AATSR Technical Note: Improvements to the AATSR IPF relating to Land Surface Temperature Retrieval and Cloud Clearing over Land*. Chilton: Rutherford Appleton Laboratory.

Cervino, M., V. Levizzani, C. Serafini, A. Bartoloni, M. Mochi, P. Colandrea, and B. Greco. 2000. "Cloud Fraction within GOME Footprint Using a Refined Cloud Clearing Algorithm." *Advances in Space Research* 25 (5): 993–996. doi:10.1016/S0273-1177(99)00462-7.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changon, T. R. Karl, and L. O. Mearns. 2000. "Climate Extremes: Observations, Modeling, and Impacts." *Science* 289: 2068–2074. doi:10.1126/science.289.5487.2068.

EEA. 2001. *Eutrophication in Europe's Coastal Waters*. European Environment Agency, Environmental Assessment Series No. 7. Copenhagen: EEA.

ESA. 2011. *MERIS Algorithm Theoretical Basis Document 2-17 – Pixel Classification*, *Date: 30.05.2011 Issue: 5.0.* Frascati: ESA ESRIN.

Fu, F. X., A. O. Tatters, and D. A. Hutchins. 2012. "Global Change and the Future of Harmful Algal Blooms in the Ocean." *Marine Ecology Progress Series* 470: 207–233. doi:10.3354/meps10047.

Fu, G., K. Baith, and C. McClain, 1998. The SeaWiFS Data Analysis System. In *Proceedings of the 4th Pacific Ocean Remote Sensing Conference*, edited by M.-X. He and G. Chen, Qingdao, China, July.

Galloway, J. N., A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. A. Martinelli, S. P. Seitzinger, and M. A. Sutton. 2008. "Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions." *Science* 320: 889–892. doi:10.1126/science.1136674.

Garant, L., and J. A. Weinman. 1986. "A Structural Stochastic Model for the Analysis and Synthesis of Cloud Images." *Journal of Climate and Applied Meteorology* 25: 1052–1068. doi:10.1175/1520-0450(1986)025<1052:ASSMFT>2.0.CO;2.

Gong, G.-C., J. Chang, K.-P. Chiang, T.-M. Hsiung, -C.-C. Hung, S.-W. Duan, and L. A. Codispoti. 2006. "Reduction of Primary Production and Changing of Nutrient Ratio in the East China Sea: Effect of the Three Gorges Dam?" *Geophysical Research Letters* 33: L07610. doi:10.1029/2006GL025800.

Hagolle, O., M. Huc, D. Villa Pascual, and G. Dedieu. 2010. "A Multi-Temporal Method for Cloud Detection, Applied to FORMOSAT-2, Venμs, LANDSAT and SENTINEL-2 Images." *Remote Sensing of Environment* 114: 1747–1755. doi:10.1016/j.rse.2010.03.002.

Hallegraeff, G. M. 2010. "Ocean Climate Change, Phytoplankton Community Responses, and Harmful Algal Blooms: A Formidable Predictive Challenge." *Journal of Phycology* 46: 220–235. doi:10.1111/j.1529-8817.2010.00815.x.

Hansson, M., and J. Öberg, 2012. "Cyanobacterial Blooms in the Baltic Sea in 2012." HELCOM Environmental Fact Sheet 2012. Accessed April 16, 2014. http://helcom.fi/Documents/Baltic%20sea%20trends/Environment%20fact%20sheets/BSEF_Cyanobacterial%20blooms%20in%20the%20Baltic%20Sea%20in%202012.pdf

Humborg, C., V. Ittekkot, A. Cociasu, and B. V. Bodungen. 1997. "Effect of Danube River Dam on Black Sea Biogeochemistry and Ecosystem Structure." *Nature* 386: 385–388. doi:10.1038/386385a0.

Hutchison, K. D., B. D. Iisager, T. J. Kopp, and M. Jackson. 2008. "Discriminating between Clouds and Aerosols in the VIIRS Cloud Mask Algorithms." *Journal of Atmospheric and Oceanic Technology* 25: 501–518. doi:10.1175/2007JTECHA1004.1.

Hutchison, K. D., J. K. Roskovensky, J. M. Jackson, A. K. Heidinger, T. J. Kopp, M. J. Pavolonis, and R. Frey. 2005. "Automated Cloud Detection and Classification of Data Collected by the Visible Infrared Imager Radiometer Suite (VIIRS)." *International Journal of Remote Sensing* 26: 4681–4706. doi:10.1080/01431160500196786.

Kahru, M., O. P. Savchuk, and R. Elmgren. 2007. "Satellite Measurements of Cyanobacterial Bloom Frequency in the Baltic Sea: Interannual and Spatial Variability." *Marine Ecology Progress Series* 343: 15–23. doi:10.3354/meps06943.

Katz, R. W. 2010. "Statistics of Extremes in Climate Change." *Climatic Change* 100: 71–76. doi:10.1007/s10584-010-9834-5.

Key, J. 2002. *The Cloud and Surface Parameter Retrieval (CASPR) System for Polar AVHRR, Users Guide*. Madison, WI: Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin.

Kutser, T. 2004. "Quantitative Detection of Chlorophyll in Cyanobacterial Blooms by Satellite Remote Sensing." *Limnology and Oceanography* 49: 2179–2189. doi:10.4319/lo.2004.49.6.2179.

Lyapustin, A., Y. Wang, and R. Frey. 2008. "An Automatic Cloud Mask Algorithm Based on Time Series of MODIS Measurements." *Journal of Geophysical Research* 113: D16207. doi:10.1029/2007JD009641.

NASA. 2013. "NASA Goddard Space Flight Center, NPP VIIRS QA - Product Quality Documentation." Accessed November 22, 2013. http://landweb.nascom.nasa.gov/cgi-bin/NPP_QA/NPPpage.cgi?fileName=qaFlag&subdir=forPage

Nixon, S. W. 1995. "Coastal Marine Eutrophication: A Definition, Social Causes and Future Concerns." *Ophelia* 41: 199–219. doi:10.1080/00785236.1995.10422044.

Nordkvist, K., H. Loisel, and L. Duforet Gaurier. 2009. "Cloud Masking of SeaWiFS Images over Coastal Waters Using Spectral Variability." *Optics Express* 17: 12246–12258. doi:10.1364/OE.17.012246.

Öberg, J., 2013. "Cyanobacterial Blooms in the Baltic Sea in 2013." HELCOM Environmental Fact Sheet 2013. Accessed April 16, 2014. http://helcom.fi/baltic-sea-trends/environment-fact-sheets/eutrophication/cyanobacterial-blooms-in-the-baltic-sea

Okada, Y., S. Mukai, and I. Sano. 2003. "Modified Cloud Flag for SeaWiFS Data over Turbid Water Regions." *Geophysical Research Letters* 30: 4. doi:10.1029/2002GL015714.

Patt, F. S., R. A. Barnes, R. E. Eplee Jr., B. A. Franz, W. D. Robinson, G. C. Feldman, S. W. Bailey, J. Gales, P. J. Werdell, M. Wang, R. Frouin, R. P. Stumpf, R. A. Arnone, R. W. Gould Jr., P. M. Martinolich, V. Ransibrahmanakul, J. E. O'Reilly, and J. A. Yoder. 2003. "NASA Technical Memorandum 2003-206892, Volume 22: Algorithm Updates for the Fourth SeaWiFS Data Reprocessing." In *SeaWiFS Postlaunch Technical Report Series*, edited by S. B. Hooker and E. R. Firestone. Greenbelt, MD: NASA.

Phulbin, T., M. Derrien, and A. Brard. 1983. "A Two Dimensional Histogram Procedure to Analyse Cloud Cover from NOAA Satellite High Resolution Imagery." *Journal of Climate and Applied Meteorology* 22: 1332–1345. doi:10.1175/1520-0450(1983)022<1332:ATDHPT>2.0.CO;2.

Plummer, S. E. 2008. "The GLOBCARBON Cloud Detection System for the Along-Track Scanning Radiometer (ATSR) Sensor Series." *IEEE Transactions on Geoscience and Remote Sensing* 46 (6): 1718–1727. doi:10.1109/TGRS.2008.916200.

Ruddick, K. G., F. Ovidio, and M. Rijkeboer. 2000. "Atmospheric Correction of SeaWiFS Imagery for Turbid Coastal and Inland Waters." *Applied Optics* 39 (6): 897–912. doi:10.1364/AO.39.000897.

Santer, R., V. Carrère, D. Dessailly, P. Dubuisson, and J.-C. Roger. 1997. *MERIS Algorithm Theoretical Basis Document 2.17, Pixel Classification, Date: 05.11. 2007,Issue: 4.0*. Frascati: ESA ESRIN.

Saunders, R. W., and K. T. Kriebel. 1988. "An Improved Method for Detecting Clear Sky and Cloudy Radiances from AVHRR Data." *International Journal of Remote Sensing* 9: 123–150. doi:10.1080/01431168808954841.

Savoye, B., N. Babonneau, B. Dennielou, and M. Bez. 2009. "Geological Overview of the Angola-Congo Margin, the Congo Deep-Sea Fan and Its Submarine Valleys." *Deep Sea Research Part II: Topical Studies in Oceanography* 56 (23): 2169–2182. doi:10.1016/j.dsr2.2009.04.001.

Sedano, F., P. Kempeneers, P. Strobl, J. Kucera, P. Vogt, L. Seebach, and J. San-Miguel-Ayanz. 2011. "A Cloud Mask Methodology for High Resolution Remote Sensing Data Combining Information from High and Medium Resolution Optical Sensors." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 588–596. doi:10.1016/j.isprsjprs.2011.03.005.

Simpson, J. J., and J. I. Gobat. 1996. "Improved Cloud Detection for Daytime AVHRR Scenes over Land." *Remote Sensing of Environment* 55: 21–49. doi:10.1016/0034-4257(95)00188-3.

Simpson, J. J., A. Schmidt, and A. Harris. 1998. "Improved Cloud Detection in Along Track Scanning Radiometer (ATSR) Data over the Ocean." *Remote Sensing of Environment* 65: 1–24. doi:10.1016/S0034-4257(98)00025-X.

Smith, V. H., G. D. Tilman, and J. C. Nekola. 1999. "Eutrophication: Impacts of Excess Nutrient Inputs on Freshwater, Marine, and Terrestrial Ecosystems." *Environmental Pollution* 100: 179–196. doi:10.1016/S0269-7491(99)00091-3.

Stal, L. J., P. Albertano, B. Bergman, K. Von Bröckel, J. R. Gallon, P. K. Kayes, K. Sivonen, and A. E. Walsby. 2003. "BASIC: Baltic Sea Cyanobacteria. an Investigation of the Structure and Dynamics of Water Blooms of Cyanobacteria in the Baltic Sea—Responses to a Changing Environment." *Continental Shelf Research* 23: 1696–1714. doi:10.1016/j.csr.2003.06.001.

Stewart, K. R., R. L. Lewison, D. C. Dunn, R. H. Bjorkland, S. Kelez, P. N. Halpin, and L. B. Crowder. 2010. "Characterizing Fishing Effort and Spatial Extent of Coastal Fisheries." *PLoS ONE* 5 (12): e14451. doi:10.1371/journal.pone.0014451.

Vörösmarty, C. J., M. Meybeck, B. Fekete, K. Sharma, P. Green, and J. P. M. Syvitski. 2003. "Anthropogenic Sediment Retention: Major Global Impact from Registered River Impoundments." *Global and Planetary Change* 39: 169–190. doi:10.1016/S0921-8181(03)00023-7.

Wang, M., and W. Shi. 2006. "Cloud Masking for Ocean Color Data Processing in the Coastal Regions." *IEEE Transactions on Geoscience and Remote Sensing* 44: 3196–3105. doi:10.1109/TGRS.2006.876293.

Zeng, N. 1999. "Seasonal Cycle and Interannual Variability in the Amazon Hydrologic Cycle." *Journal of Geophysical Research* 104: 9097–9106. doi:10.1029/1998JD200088.