Global Lakes Sentinel Services

Grant number 313256

WP3 Algorithm development for S2 and S3

(pre-processing and data reduction)

D3.3 Optical pre-classification method

Marieke A. Eleveld, Ana Ruescas,

Annelies Hommersom (VU/VUmc, BC, WI)

2014-10

Global Lakes Sentinel Services
Grant number 313256

WP3 Algorithm development for S2 and S3 (pre-processing and data reduction)

D 3.3 Optical pre-classification method

STICHTING VU-VUMC, BROCKMANN CONSULT, WATER INSIGHT BV

Due date: 2014-05 (extended to 2014-09)

Submitted: 2014-10

Change records

Version	Date	Description	Contributors
0.1	2014-05	First draft that includes a literature review and test results from a first implementation in BEAM	M.A. Eleveld (VU/VUmc), Incl. figs of A. Ruescas, D. Odermatt (BC)
0.2	2014-08	Data for a new cluster analysis: SYKE, TO, WI, VU/VU-mc, CNR, Yunlin Zhang (extra), Added to: New Hampshire, Lake Erie, Spanish lake datasets (Moore et al., 2014) (extra) Cluster analysis: T.S. Moore (UNH, extra) Software development: Marco Peters (BC) Testing: WI & BC, VU (Dutch lakes) Coordinated by A. Hommersom, A. Ruescas & M.A. Eleveld	
0.3	2014-09	Comparison of the Diversity Inland and GLaSS-5c OWT classification; Comparison of GLaSS-5c, 6c, and 6c-normalised for Dutch Lake IJsselmeer and Markermeer; A proposed processing chain for nearby lakes	A. Ruescas & D. Odermatt (Diversity, BC, EXTRA) M.A. Eleveld (VU/VUmc) A. Hommersom (WI)
0.9	2014-09	Draft to GLaSS core team, advisory board and Tim Moore	M.A. Eleveld (VU/VUmc)
1.0	2014-10	Implementing comments Comments Adapting and adding to Appendices	M.A. Eleveld (Vu/VU-mc), A. Ruescas (BC), A. Hommersom (WI) K. Kangro (TO), K. Schenk & T. Heege (EOMAP), S. Koponen (SYKE) M.A. Eleveld (Vu/VU-mc), A. Ruescas and D. Odermatt (BC), C. Giardino, M. Bresciani (CNR)

Consortium

No	Name	Short Name
1	WATER INSIGHT BV	WI
2	SUOMEN YMPARISTOKESKUS	SYKE
3	EOMAP GmbH & Co.KG	EOMAP
4	STICHTING VU-VUMC	VU/VUmc
5	BROCKMANN CONSULT GMBH	BC
6	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR
7	TARTU OBSERVATORY	TO
8	BROCKMANN GEOMATICS SWEDEN AB	BG

Reference
Please refer to this report as: GLaSS Deliverable 3.3, 2014. Global Lakes Sentinel Services, D3.3: Optical pre-classification method, VU/VUmc, BC, WI. Available via: www.glass-project.eu/downloads

Scope

The overall aim of the GLaSS project is the setting up of a system that is able to handle large quantities of Sentinel data. This report is about the development of an optical pre-classification algorithm for main water-types. This was deemed necessary because inversion algorithms often have a limited range of values – of basically total inherent optical properties (IOPs), a and b – for which they perform optimally. Because of the actual, extreme variability in range of spectral absorption (a) and scattering (b) between and within global lakes, satellite borne optical sensors capture the distinctly different colours. We decided to classify these spectra with a spectral classification method (Moore et al., 2014), because satellite sensors collect these apparent optical properties (AOPs) for all (global) lakes.

To distinguish the main water-types, we first collected atmospherically corrected satellite reflectances and radiances (also from Task 3.2). In addition, clustering of the collected in situ spectra resulted in three classification schemes that generate different lake water classes. Thus we could identify and map water types with distinctly different reflectances and total Inherent Optical Properties (IOPs). Given the limited information about the contribution of individual constituents, we refrain from labelling of these water types further. This is where the actual classification with retrieval algorithms (Task 3.4) comes in.

The purpose of Task 3.3 is to develop a method that facilitates to pre-select which atmospheric correction and water quality retrieval algorithm could be used for which lake. The OWT pre-classification tool, which was implemented in BEAM, maps the water type of the class spectrum that matches the remotely sensed spectrum best. It has been tested for GLaSS nearby lakes, which are those lakes that have been studied by individual GLaSS partners for long time, and for which we have additional in situ data. This resulted in a protocol, or processing line that can also be applied or adapted for ENVISAT heritage Sentinel-3 OLCI data sets.

Environmental conditions can cause optics to be variable, and this variability could be used as indicator for validation of biogeochemical and hydrological models, such as the HYPE (GEOLAND-2) model (Task 4.3). The spectral classification method is also being taken up by Task 3.6, data mining, and is being considered for the development of case studies WP5. BEAM now features several GLaSS OWT processors for further exploration by GLaSS partners, and we also provide handles (description of method, processing line) to these Tasks and WP.
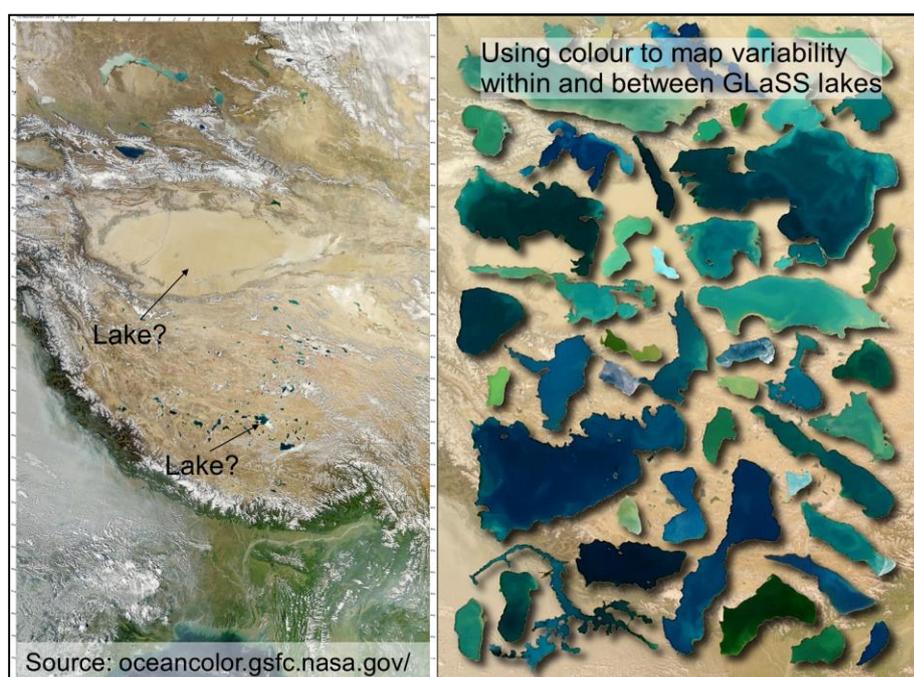
Abstract

Shallow and deep lakes receive and recycle organic and inorganic substances from within the confines of the lakes proper, the watershed and beyond. Hence, the huge ranges in absorption and scattering, and the extreme optical variability between and within global lakes. Inversion algorithms use the observed spectral variability for retrieving IOPs and concentrations of individual optically active substances. However, the large ranges and extreme variability are a challenge for (parameterization of) the currently available atmospheric correction and/or in-water retrieval algorithms. We aimed to develop a method and protocols that help optical water quality professionals selecting (from a set of existing) atmospheric correction algorithms and in-water retrieval algorithms by means of a pre-classification.

We have been adapting a spectrally-based optical pre-classification scheme (Moore et al., 2014, RS Env. 113) for global lakes, by calculating clusters from in situ spectra collected by partners and advisors of the EU-GLaSS project. We also have access to a large set of MERIS data, atmospheric correction tools and dedicated lake water quality retrieval algorithms. The tool maps the water type of the class spectrum that matches the remotely sensed spectrum best, and has been integrated in BEAM. It has been tested for GLaSS nearby lakes, which are those lakes that have been studied by individual GLaSS partners for long time, and for which we have additional in situ data. This resulted in a protocol, or processing line that can also be applied or adapted for ENVISAT heritage Sentinel-3 OLCI data sets.

We also anticipate that these maps and protocols for nearby lakes will suggest similar settings for lakes of comparable biophysical types (the GLaSS use cases). Once these settings are verified and robust we can apply them to the huge majority lakes for which we only have the satellite spectra. We are aware that local specific IOPs (sIOPs) can be different, and that this should only be used as a first indication of a potentially good atmospheric correction (AC) and water quality (WQ) retrieval approach, or as a best guess in case these sIOPs are not known.

Graphical abstract

Source: oceancolor.gsfc.nasa.gov/

Table of contents

1 Introduction

The upcoming ESA satellites Sentinel-2 and Sentinel-3 will provide unprecedented monitoring capabilities of Earth surfaces thanks to the high overpass frequency of S-3 and the high spatial resolution of S-2. However, the large data volumes also require new methods to turn data into useful information for end-users. Within the GLaSS project, these issues are tackled, and a prototype infrastructure targeted specifically at lakes is set up.

The Algorithm and methods development work package, WP3 (Fig. 1) contains Task 3.3, which aims at the design of an optical pre-classification algorithm that can distinguish the main water-types and facilitate selection of water quality (WQ) algorithms. After exploring the topic of optical pre-classification, the fuzzy Optical Water Type processor (Moore et al., 2014) is introduced and applied. One of the additional innovative uses of this processor is testing the impact of different atmospheric corrections (GLaSS WP 3.2, the 'feedback loop' in Fig. 1).
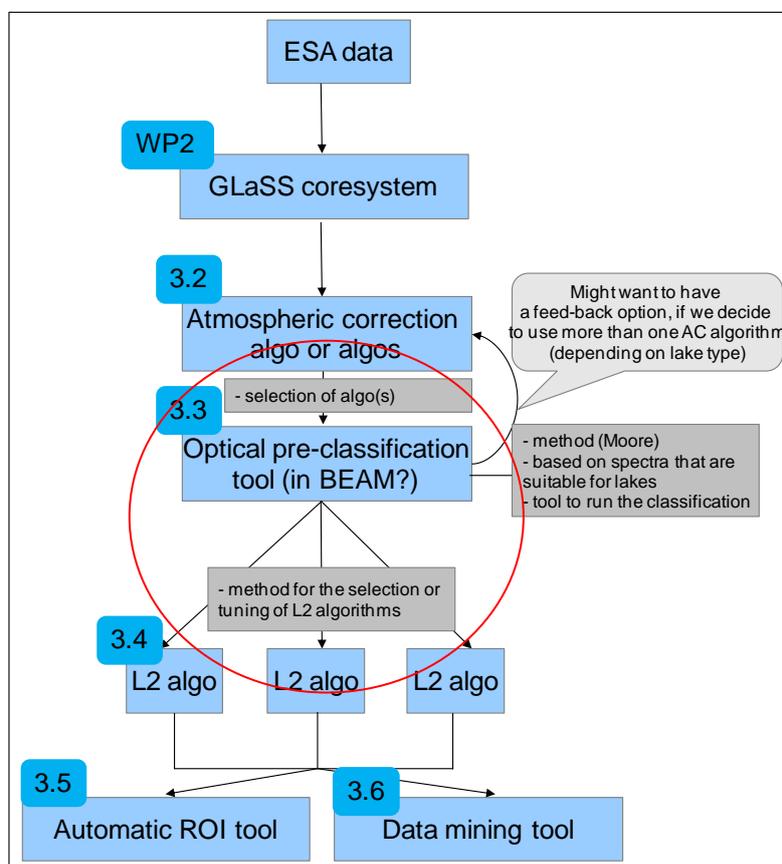

Figure 1. Task 3.3 Optical pre-classification, within the WP3 Algorithm and methods development workflow.

1.1 Why do we need an optical pre-classification of lakes?

AOPs and IOPs

Our common theoretical framework regarding the behaviour of in-water light is that it is defined by the main inherent optical properties (IOPs, absorption, scattering, and the scattering phase function) and apparent optical properties (AOPs, such as irradiance, radiance, and reflectance). Total IOPs of surface waters depend on absorption and scattering of the pure water and by major optically active constituents that reside dissolved and suspended in the water column. Apparent Optical Properties (AOPs) are those optical properties that are influenced by the angular distribution of the light field, as well as by the IOPs. For known angular distributions of the incoming light, radiative transfer models can link AOPs such as remote sensing reflectance to total IOPs. Decomposition into concentrations of optically active substances such as chlorophyll-a (CHL), total suspended matter (TSM), coloured dissolved organic matter (CDOM), and occasionally phycocyanin (PC), requires additional information about concentration-specific inherent optical properties (sIOPs). Retrieval algorithms unravel non-linear, spectrally-varying interactions among remotely-sensed signals, sIOPs and concentrations.

Challenges for retrieval algorithms

The retrieval is challenging (Fig.2), and there might be several reasons why we need optical pre-classification. First of all, the changes in signal with changes in the sIOPs or concentrations of the individual constituents are often very small compared to changes brought about by atmospheric optically active substances, making it difficult to guarantee the quality and reliability of the extracted in-water concentrations. Also, at many wavelengths, two or more substances may influence the optical signal in a similar manner, making it difficult to distinguish one material from another. Then, remotely-sensed observations are often inverted for widely varying conditions (Blondeau-Patissier et al., 2009), and the assumption that requisite transfer function and sIOPs are constant in time and for all pixels might not hold. Finally, each individual satellite sensor band has its own sensitivity and saturation characteristics, while together they encompass a large dynamic range so that they cover dark as well as bright signals. All these effects cause algorithms to have a range of values (viz., total IOPs) for which they perform optimally (D' Alimonte et al., 2003; Odermatt et al., 2012). To resolve this, these algorithms have switches incorporated, such as Case-2 adaptations of the standard OC3 algorithm (Gohin et al., 2002 and O'Reilly et al, 1998 respectively), or use many wavelengths (Doerffer and Schiller, 2007; Schroeder et al., 2007; Van der Woerd and Pasterkamp, 2008). Nonetheless, application of a semi-analytical, or spectral matching algorithm parameterised with adaptable sIOPs implies a physically sound procedure and has been giving good results (Brando et al., 2012; Tilstone et al., 2012, Heege et al. 2014). In fact, these spectral matching algorithms – e.g. Van der Woerd and Pasterkamp, 2007, and c-distort by Stamnes (Stamnes et al., 1988; Jin and Stamnes, 1994; Stamnes et al., 2000; Gjerstad et al., 2003) – work if the bio-optical model is well-constrained (sIOPs are set) and the necessary bands are included. For inland waters, appropriate adjacency correction procedures are usually necessary (Kiselev et al. 2014). For the detection of phytoplankton in case-2 waters, MERIS bands 7 and 9 are required to capture chlorophyll-*a* absorption at 676 nm with respect to a minimum in the combined absorption by phytoplankton pigments and water at 705 nm (Gons et al, 2005; Gurlin et al., 2011). However, full sIOP sets are rare, particularly if accessory pigments such as phycoerythrin (PE) and phycocyanin (PC) are also of interest, and therefore such algorithms are not yet commonly applied, and might need further validation (Brando et al., 2012; Odermatt et al., 2012).
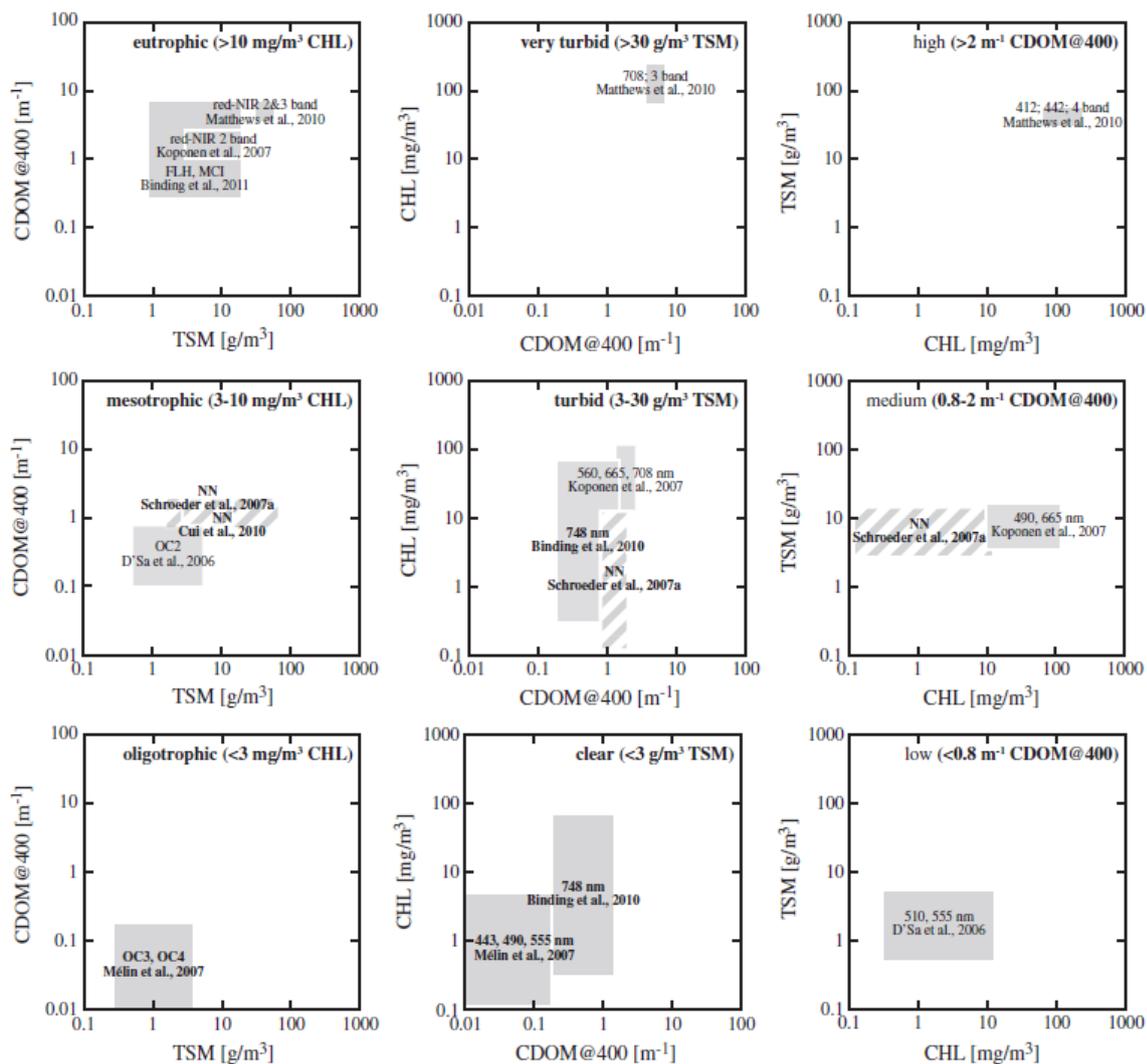
Figure 2. Algorithm validation ranges for different optical conditions (Odermatt et al., 2012, Fig. 4)

Pre-classification for lake waters

Hence, there might be several reasons why an ***optical pre-classification*** might be of interest, for lakes in particular. An optical pre-classification enables identifying areas with common bio-optical properties.

Biomass and pigment concentrations can be high in lakes, and there are also extremely humic (CDOM) and mineral (SPM dominated, glacial) lakes. Thus concentrations of several independently varying optically active substances and their IOPs can vary significantly within lakes and between lakes. The concentration-normalised sIOPs have a more narrow dynamic range, but can still contain extremes, particularly for "special conditions", such as dominance by cyanobacteria, and mine tailing ponds. Retrieval of concentrations with empirical algorithms without switches will simply fail (or give erroneous results) as soon as one of the concentrations, or sIOPs from the optically active constituents changes.

Dekker et al. (1995), Durand et al. (1999), and Lindell et al. (1999) have shown that the empirical approach to remote sensing of (TSM) in inland waters cannot produce multi-temporal quantitative maps, nor does it allow retrospective analysis. They recommend forward and inverse modelling between the remotely sensed signal and the water constituents, via the specific inherent optical properties (Dekker et al., 2002.). The previously mentioned state-of-the-art spectral matching methods that incorporate this approach are currently mainly retrieving the three main constituents

(CHL, TSM, CDOM) at once from remote sensing reflecance. Only limited research has been done on incorporation of additional absorption of PC (or PE) and scattering of cyanobacteria.

In our nearby and long list of global lakes (Eleveld et al., 2014), we will use a measure that is available for all lakes, remote sensing reflectances of global lakes (at all optical and the first NIR bands of respective sensor) for *spectral* pre-classification of inland waters (Moore et al., 2014). In this case, the spectral matching is aimed primarily at indicating the most appropriate atmospheric correction and retrieval algorithm.

1.2 Lake classification systems

Classifiers addressing the heterogeneous nature of lakes

We use classification systems to assess the state of the environment at any point in time. In general, classification systems address the heterogeneous nature of lakes. Some of the most important characteristics include the lake's morphometry (size, shape and depth), the activities that occur in the lands that drain into it (watershed or basin), the location or ecoregion in which the lake is located, and when and how the lake basin was formed. Lakes and rivers within ecoregions often have similar physical characteristics, water chemistry, and biological communities, because they occur in an area of similar land type. The number, appearance, and condition of lakes vary among ecoregions because of glacial history, geology, soil type, land use, and climate. A lake is also a reflection of its watershed: what happens on the land and the basic characteristics of the land (soil, geology, vegetation, drainage, etc.) affects the quality and health of a lake. These factors, acting in various combinations, have created the multitude of lake types, with different trophic state and use (Eleveld et al., 2014). Trophic state is related to CHL (Odermatt et al., 2012) and has a link to transparency (e.g. Carlson, 1977). Thus, the classification can be based on descriptors of lake ecosystems (morphometry, hydrodynamics, and trophic state), its environment (hydrology and land use in the watershed, ecoregions, geology), use, and optics. Trophic thresholds may vary with ecosystem-specific limitations to primary productivity, while the validity of remote sensing algorithms is determined by the spatio-temporal variability in optical properties.

Similarly, the ocean colour community has been perceiving the oceans as biogeochemical provinces (Longhurst, 1995, 1998, 2006; IOCCG, 2009) based on physical circulation patterns (forcings), availability of light and nutrients (distance from continental masses), complexity of the marine food web, and common bio-optical properties. Recently, Lee and Hu (2006), Mélin et al. (2011) realized that in practice, the only feasible means to map case-1 versus case-2 waters is to devise an inclusive remote-sensing criterion using remote sensing reflectances from ocean color satellite sensors and bio-optical models from extensive measurements.

Historic 'optical' water type classification systems
If the emphasis of the classification is on identifying areas with common bio-optical properties, the results may be called optical water types. If remotely sensed spectra (an AOPs) are the main optical information source, we actually deal with spectral water types. However, this distinction is not always made. There certainly has been a long tradition of classification into 'optical' water types. Jerlov (1976) identified optically different classes for open ocean water I, IA, IB, II and III, and for coastal water 1 to 9, based on the curves of percent transmittance of downward irradiance against wavelength, which he derived through the use of broad band colour filters (which absorb part of the spectrum) in the surface layer. The Jerlov water types are in essence a classification based on water clarity as quantified by vertical diffuse attenuation (Kd) at a certain depth just below the sea surface, and wavelength (Mobley, 1994). This classification scheme can be contrasted with the case 1 and case 2 classification described by Morel and Prieur (1977), who classified ocean water into two classes (case 1 and case 2) based on IOPs, the type of absorbent particle suspended in the water column (Moore et al., 2001). Kirk (1980) extended this classification explicitly with soluble fractions, which is of interest for inland waters. In type G waters, gilvin (CDOM) absorbs light more strongly at all wavelengths in the photosynthetic range than the

particulate fraction. Prieur and Sathyendranath (1981) further expanded and elaborated the case 2 classification. The idea is to take an optical property of interest, say absorption at 440 nm or reflectance at a certain wavelength, for a sample taken at a particular location and time, and compute the fractional contributions due to phytoplankton, yellow substances, and suspended material. The sample can then be characterised optically by a single point on a triangular plot, in which the axes are the fractional contributions due to each of the three components. Waters in which only one component dominates belonging to these cases would fall into one of the three smaller triangles demarcating areas close to the three apices. Cases where all the components play important roles end up in the central, inner triangle. In principle, this would be the most complex water type, from an optical point of view (Sathyendranath/IOCCG, 2000).

For our project we foresee optical lake types that are optically dominated by chlorophyll (CHL), total suspended matter (TSM), and yellow substance (CDOM) mixtures, as well as lakes optically dominated by cyanobacteria, or mine tailing ponds. Even then, lakes such as the Australian salt lake Kati Thanda–Lake Eyre can have exceptional colours or spectra that will fall even out of these categories. Boundaries between water types can vary in time (e.g. with season) and are dynamic (Van der Woerd et al., 2004; Hommersom et al., 2010, 2011).

1.3 Discussion on cluster approaches

An overview of possibilities for optical pre-classification was extracted from the peer-reviewed literature. This ranges from pre-classification on an IOP level to clustering on reflectance (Moore et al., 2001, 2009) or sIOPs (Tilstone et al., 2012), or both. Spectra are non-linearly related to total IOPs. Successive decomposition of an IOP into component IOPs and subsequent conversion to concentrations through sIOPs adds variability and ambiguity. Should the pre-classification discriminate highly absorbing (high a), high scattering (high b), intermediate/low a & b (see e.g. Reinart et al. 2002)?

Optical clustering based on information from nearby lakes
For some (nearby) lakes we have quality controlled spectra, and sometimes concentrations and IOPs. These datasets enable constructing a classification framework for grouping waters into various types. The resulting classes could subsequently be used for pre-classification of all global lakes if we assume that unclassified pixels belong to a 'novelty' class, which comprises those spectra that cannot be assigned to any of the known classes (Schiller et al., 2007; Moore et al., 2014). A good example of clustering on the concentration-specific absorption characteristics in bands that can be assigned to individual substances is given by Tilstone et al. (2012), scattering phase functions were not considered, because there are simply not enough measurements available (Tilstone et al., 2012). This illustrates that, if we want to cluster on optical properties, ideally we have to cluster on sIOPs, or full CSI sets, from which we derive also sIOPs. Such a quality controlled database for inland waters is currently being set-up under the name Limnades (http://www.globolakes.ac.uk/limnades/), and will prove to be extremely beneficial towards the understanding of limnological ecology on a global scale, and will also provide critical information for global climate change effects on lakes models (Mooij et al., 2005, 2007; Nõges et al., 2008).

Spectral clustering
If we focus on spectra as primary data source for all lakes, we can use a simple approach for the identification of the three basic water types. This consists in the comparison of the values from MERIS bands 7, 8 and 9 (central wavelengths 664.6, 680.8, and 708.3 nm, respectively), which allows for a simple classification of every pixel into one of the three types: highly absorbing, highly scattering or both (Shi et al., 2013a and b). We could also add another class using an additional blue band to pre-classify for high absorption dominated by high CDOM. In Shi et al. (2013b) a cluster analysis was applied to classify the entire set of Rrs(λ) spectra into homogeneous types. To preserve spectral shape information of Rrs(λ) in the classification, each Rrs(λ) was previously

normalized by its integral, computed over the entire PAR spectrum (Lubac and Loisel, 2007; Le et al., 2011). This amplifies the absorption characteristics.

Moore et al. (2001) applied a fuzzy logic classification scheme applied to the satellite-derived water leaving radiance data to select and blend class-specific algorithms. Local in situ bio-optical data were used to characterize optically-distinct water classes *a priori*. A membership function expresses the likelihood that the satellite remote sensing reflectance belongs to a class with a known reflectance distribution. The χ^2 implementation has similarities with other spectral matching algorithms (Hommersom et al., 2010, Brando et al., 2012, Tilstone et al., 2012). Class memberships can also be used to weight the class-specific retrievals and obtain a final blended retrieval for each pixel. Moore et al. (2009) used this technique to assess the increase in uncertainty of CHL estimates from ocean to coast with the standard OC3M algorithm (based on O'Reilly et al., 1998). Moore et al. (2014) aim at blending retrievals between such a blue/green (OC3M) and a red/NIR based algorithm for lakes and coastal water. This water type-specific approach is intrinsically independent of location and the time and therefore presumably more appropriate for geographically extensive application than empirical regional classification (Vantrepotte et al., 2012; Shi et al., 2013b).

We will test the case-2 spectral classification method of Moore et al. (2014). The main reason is that we have satellite spectra for all lakes. Also, an approach focusing on sIOPs, the relationships between individual spectral inherent optical properties and optically active substances in lakes (Brando et al., 2012; Tilstone et al., 2012) tends towards retrieval instead of pre-classification.

2 Implementation of the OWT approach

2.1 Clustering of in situ spectra

Moore et al. (2001, 2014) designed a fuzzy logic spectral classification scheme that we adapted for GLaSS lakes. In situ hyperspectral data were used to characterize optically-distinct water classes *a priori.* The aggregate data come from multiple sources. The GLaSS dataset covers a wide range in concentrations, also for CDOM and SPM (Table 1).

Area	Spectrometers	Spectral range (nm)	Spectral resolution (in the VIS-NIR)*	Spectral resolution of interp. output	# of spectra	Range CDOM (at 443) (m^{-1})	Range SPM (g m^{-3})	Range Chl (mg m^{-3})
Finland	ASD FieldSpec FR Pro Jr.,	350 – 2500	3 nm	1 nm	16	0.5-10	0.8-3.4	1.7-11
Estonia	TriOS RAMSES	400-800	~7 nm	2.5 nm	34	1.7-4.2	1.8-18.7	2.7-45.3
The Netherlands WI	WISP-3	400 – 800	3.9 nm for E_d sensor; 4.9 nm for L_w and L_{sky} sensors	1 nm	177	Estimate 0.5-1.5	1.3-30	10-50
The Netherlands IVM	Photo Research PR650	380-748		4 nm	5 (IJM)		13.0-26.0	33.4-87.3 (40.7-104.6TCHL)
					3 (MM)		29.7-39.2	36.6-42.6 (46.4- 53.3TCHL)
Italy	ASD FieldSpec Full Range Pro	350 – 2500	3 nm	1 nm	90	0.04-1.25	0.1-15	0.1-10
	SpectraScan Colorimeter PR-650	380 -780	8 nm	4 nm	3			
	WISP-3	400 – 800	3.9 nm for E_d sensor; 4.9 nm for L_w and L_{sky} sensors	1 nm	13			
China	ASD	350 – 1000	3 nm	1 nm	243	0.3-2.4	10-286	5-940
TOTAL					584	0.1-10	0.1-290	1.7-940

Table 1. In situ data from various GLaSS partners and advisory board member Yunlin Zhang.

The GLaSS in situ spectra were measured with spectrometers listed in Table 1. The measurements were performed according to standard protocols (e.g. Mueller et al., 2003). and included the measurement of: 1) Light emerging from water (L_w) with about 42 degree elevation angle from nadir and about 135 degree azimuth angle from the Sun; 2) Light from sky (L_{sky}) with about 42 degree angle from zenith and about 135 degree azimuth angle from the Sun; 3) Downwelling irradiance (E_d).

The remote sensing reflectance, Rrs (in sr^{-1}) is then computed with:

$$R_{rs} = \frac{L_W - \rho L_{Sky}}{E_d},$$

(Eq. 1)

where the air-sea interface reflectance factor ρ is approx. 0.028 at a zenith angle of 42 degrees (Mobley, 1999). The number of measured spectra is 550. The conversion from above water Rrs (Eq. 1) to the below water RRs, for which the pre-classification works, is (Lee et al., 2002), Eq. 2.

Rrs (0,-)= Rrs(0,+)/(0.52+1.7*Rrs(0,+)) (in sr^{-1})   (Eq. 2)

The GLaSS data were Rrs filtered and band averaged at 3 nm resolution. Subsequently, Rrs and Chl data were merged with Moore et al. (2014) 'lake only' data (N=320), that consist of assorted lake data from New England and the Great Salt Lake (Shane Bradt), Lake Erie (Tim Moore), Spanish lake datasets (Antonio Ruiz-Verdu). Rrs data collected on Lake Erie in 2013 by Moore were added (N=16). The hyperspectral resolution captures spectral features in the visible spectrum and near infrared ranges (VIS-NIR) and provides flexibility for adapting the derived OWT spectral reflectance characteristics to many sensors, such as S3 OLCI. This combined dataset (Fig. 3) ideally should include a complete representation of all possible water types.

The merged data (with N about 926) were quality controlled and reduced to N=871. Several chlorophyll algorithms were run for consistency checking, and all data sets followed similar behaviour.
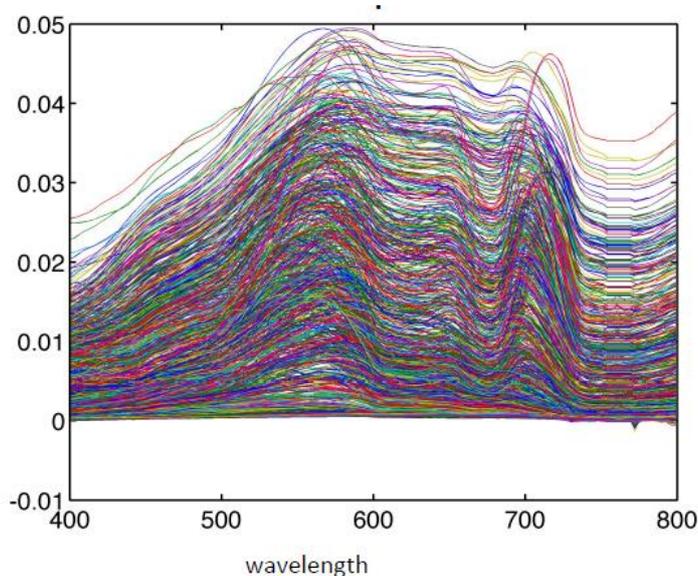
Figure 3. Total Rrs input data, after quality control (courtesy Tim Moore).

Merged, quality controlled data were subset to the MERIS bands (2-10, Table 1) before clustering, to suppress the influence of correlation between bands that did occur in using the full hyperspectral settings. They were used in the fuzzy c-mean (FCM) algorithm Bezdek (1981, see Moore et al., 2014) for normal and PAR-normalised (Vantrepotte et al., 2012; Craig et al., 2012) reflectances. The latter to stress spectral shape (and thus absorption) characteristics. The FCM algorithm produces a fuzzy clustering of data into a specified number of clusters. The basic function of this algorithm is to choose clusters that minimise the distance between the data points and the prototype cluster centres (means). It is impacted by both the shape and magnitude of Rrs. Cluster centres are iteratively adjusted until optimization criteria are met: the relation of data points and cluster centres (mean vectors) are collectively in a better configuration in terms of compactness and separation aspects, than for other cluster choices. Thus the clusters were identified by Rrs with band subsets, but then sorted and re-created with all the bands. (assuming the subset does a good job of representing the important features). This way, the classification can be adapted to any satellite or band configuration.

Hence, the clustering routine returns the optimal mean reflectance vector and co-variance matrix for each of the cluster classes, and thereby the "optical:" water types and memberships functions. These mean reflectances are the visual representations of the total in-water optical properties.

For non-normalized clustering, the optimal number of classes (c) is 5 or 6. For PAR-normalized clustering, optimal c is 6 clusters. A trapezoidal numerical integration over PAR was used for the normalisation. Wavelength choices for integration match the nine MERIS bands (Table 1). Also, class means and one common inverted covariance matrix for all classes has been computed. (The matrix needed to be pre-inverted because of the normalisation and use of a common covariance

matrix for each class.)

These three classifications were analysed, in order to understand which one represents better the different global lakes. In Appendix 2, results from all three methods were compared for the Dutch Lakes case. In Appendix 3 results of the GLaSS 5c and from a complementary ESA project Diversity II (http://www.diversity2.info/), a 7c inland classification are given. Appendix 3 (courtesy CNR) summarises a test on Italian alpine Lakes.

Table 1. MERIS band number, centre wavelength and bandwidth(nm)

2	442.55902	9.946
3	489.88202	9.957
4	509.81903	9.961
5	559.69403	9.970
6	619.601	9.979
7	664.57306	9.985
8	680.82104	7.488
9	708.32904	9.992
10	753.37103	7.495

2.2 Implementation in BEAM

The clustering process results in clusters that form the classes, which are defined by the means and co-variances of the clusters. These are the data that are needed for the membership function that BEAM can use. In order to update the processing the updated tables need to be provided. For normalized classes, the pixel spectra need to be normalized as well.

Classification method by Moore et al. (2014) had been implemented in BEAM for first tests. However, for the global lakes of GLaSS, lakes with extreme CDOM concentrations, and lakes mainly influenced by SPM were missing from the training set. Therefore, three new GLaSS classifications were set up. A class with five clusters (5C), a class with 6 clusters (6C) and a class in which the spectra were normalised to their total intensity (the area below the graph) before clustering: the 6C_normalised clustering. Tim Moore provided these three data tables to BC, who implemented the classification method with these new values in the OWT tool in BEAM 5, so that we have a processor tool for GLaSS partners to use embedded into BEAM.

The OWT can be found under VISAT's *Processing* tab, on the *Thematic Water Processing* sub-label. It is called *MERIS OWT classification*. The interface with the users consist on a window with the input and output parameters (source and target products), selection of directory, format of the output product and if the derived product should be open in VISAT or not. The *Processing Parameters* interface lets the user select the type of OWT and indicates the prefix used to identify the reflectance bands on the input product. In the example case the prefix is "reflec". Reflectances can be somewhat different when produced with different atmospheric correction schemes (e.g. standard MEGS, CoastColour, C2R, etc.) and others. The user can opt for OWT input as RADIANCE_REFLECTANCES or IRRADIANCE_REFLECTANCES (ESA compatible).
Remote sensing, or water leaving radiance reflectance (Eq. 3) is expressed in sr-1 units (theta_s is solar zenith angle).

$$Rrs(+) = Lw/(E0*\cos(theta\_s)) \hspace{4cm} (Eq. 3)$$

Water leaving (irradiance) reflectance is implemented in BEAM according to Eq. 4 and is dimensionless.

$$rho\_w(+) = RLw = Lw*PI//(E0*\cos(theta\_s)) \hspace{3cm} (Eq.4)$$

Based on this input, the OWT Tool will convert RLw into Rrs, or not. Internal calculations will also convert the satellite signal from Rrs(+) to Rrs(-) (Eq.2), and perform normalisation, so that remote sensing data sets, and in situ classes were treated equally, and there's spectral matching on equal grounds.

The OWT processor expects to find the reflectances following a naming convention which is <prefix>_i, where <prefix> is a user defined text and "i" is the MERIS band number, starting with 1. For instance "reflec_2_*".The <prefix> can be specified here in this text field.
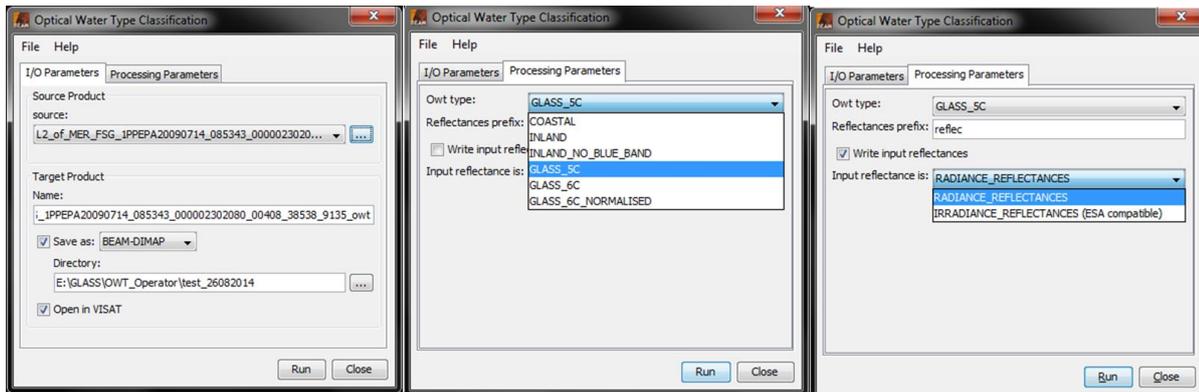
Figure 4. Optical Water Type classification tool in VISAT.

The output file consists on one band per each cluster, without and with normalisation, one band indicating the dominant class, the class sum and the normalised class sum band (Figure 5). All pixels with reflectance data are evaluated and a probability of pertaining to the different clusters is generated. These are the membership maps. In some cases, it is possible that some pixels do not have membership in any of the OWTs, in that case they are assigned to NaN and depict in black. The dominant class band shows which of the OWT classes is dominant on each pixel, that is, which one is the maximum membership. The class sum band is the the sum of all class memberships. The normalization used is the simple division of the class by the class sum; therefore the norm_class_sum should be always 1 when the pixel has been classified in any of the clusters.
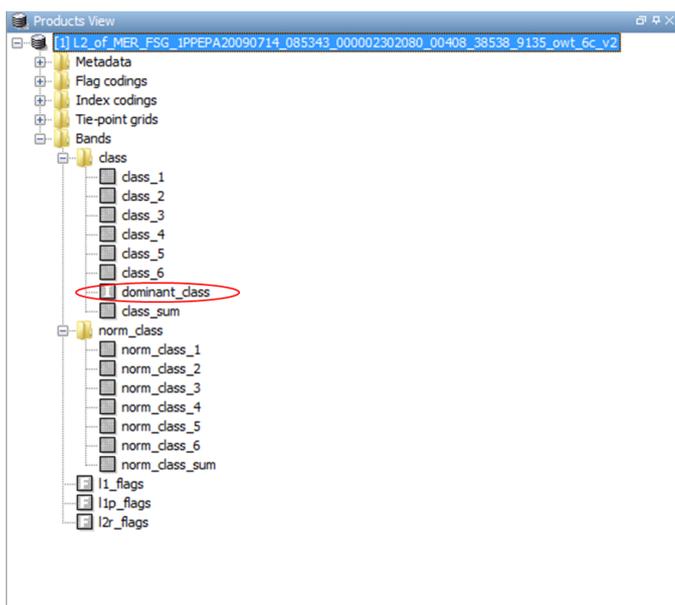
Figure 5 Output bands after OWT classification

15

The number of classes depends highly on the input in situ data used; and it is possible and even likely, that the data collected do not represent all conditions

The objectives of the first testing of the approach were:
(1) to investigate how the MERIS spectra of the GLaSS lakes fit to the class-spectra;
(2) to test Tim Moore's Optical Water Types in situ spectra from GLaSS lakes.

3 Results

3.1 The GLaSS clusters

For each of the three GLaSS classifications, in situ data ended up in different classes. Tables 3a, b and c show for every lake in which classes the in situ data ended up. The normalised classification resulted in a larger range of classes for many lakes, and notably the Finish ones (Table 3c)

a)

Non-normalised: 5 Classes

Source	Type 1	2	3	4	5	Total
Finnish	0	16	0	0	0	16
Taihu	0	0	14	129	91	234
Peipsi	0	0	24	10	0	34
Markermeer99	0	0	2	0	0	2
IJsselmeer99	0	0	5	0	0	5
IJsselmeer	2	6	45	0	0	53
Markermeer	0	0	71	0	0	71
ITA ASD	85	0	2	3	0	90
ITA WSD	12	0	0	0	0	12
ITA PR 650	0	3	0	0	0	3
Betuwe	0	12	4	0	0	16
NH Lakes	38	42	97	2	0	179
Spanish lakes	28	86	25	1	0	140
Lake Erie	4	0	9	3	0	16
Totals	169	165	298	148	91	871

b)

Non-normalised: 6 Classes

Source	Type 1	2	3	4	5	6	Total
Finnish	0	15	1	0	0	0	16
Taihu	0	0	1	41	108	84	234
Peipsi	0	0	10	21	3	0	34
Markermeer99	0	0	0	2	0	0	2
IJsselmeer99	0	0	1	4	0	0	5
IJsselmeer	0	0	49	4	0	0	53
Markermeer	0	0	16	55	0	0	71
ITA ASD	81	0	4	3	2	0	90
ITA WSD	12	0	0	0	0	0	12
ITA PR 650	0	3	0	0	0	0	3
Betuwe	0	8	8	0	0	0	16
NH Lakes	32	29	77	39	2	0	179
Spanish lakes	28	72	19	20	1	0	140
Lake Erie	2	0	4	10	0	0	16
Totals	155	127	190	199	116	84	871

c)

Normalised Method: 6 classes

Source	Type 1	2	3	4	5	6	Total
Finnish	0	11	0	1	4	0	16
Taihu	0	0	8	79	22	125	234
Peipsi	0	0	0	9	25	0	34
Markermeer99	0	0	0	1	1	0	2
IJsselmeer99	0	0	0	4	1	0	5
IJsselmeer	0	6	0	8	39	0	53
Markermeer	0	8	0	29	34	0	71
ITA ASD	79	9	0	1	1	0	90
ITA WSD	11	1	0	0	0	0	12
ITA PR 650	1	2	0	0	0	0	3
Betuwe	0	8	0	6	2	0	16
NH Lakes	27	60	26	32	33	1	179
Spanish lakes	25	52	1	37	25	0	140
Lake Erie	2	7	0	3	4	0	16
Totals	145	164	35	210	191	126	871

Tables 3a, b and c. In situ classes per lake.

The difference between the clusters can be easily seen when their reflectance means are plotted (Figs 6a, b and c). They are sorted purely on Rrs distribution, but are representations of optical conditions governed by the total absorption and scattering properties (IOPs) of the in-water constituents. The wavelength dependencies give us clues for the interpretation. This was done by the optical specialists for their nearby lakes in section 3.2. We refrain from a more generic attribution of absorption and scattering characteristic to the impact of the individual optically active components, or a more formal classification over here, also because this is beyond the scope of Task 3.2.

6a)

6b)
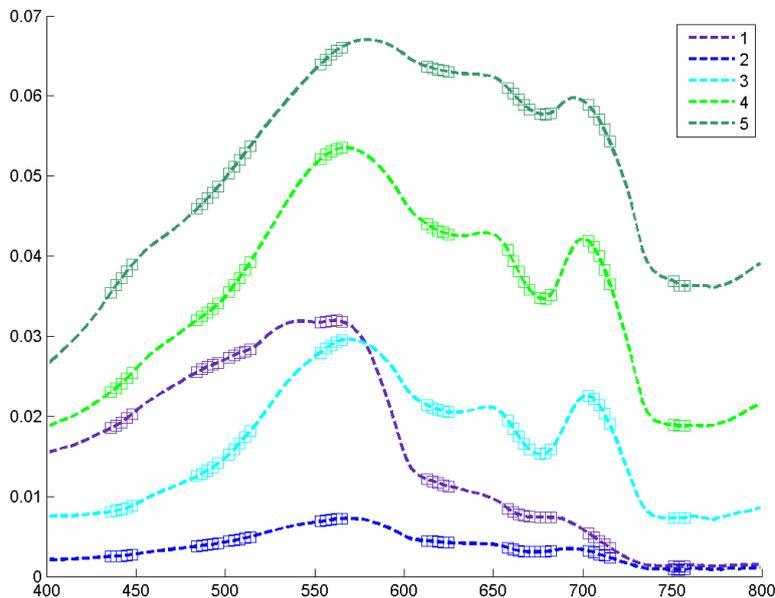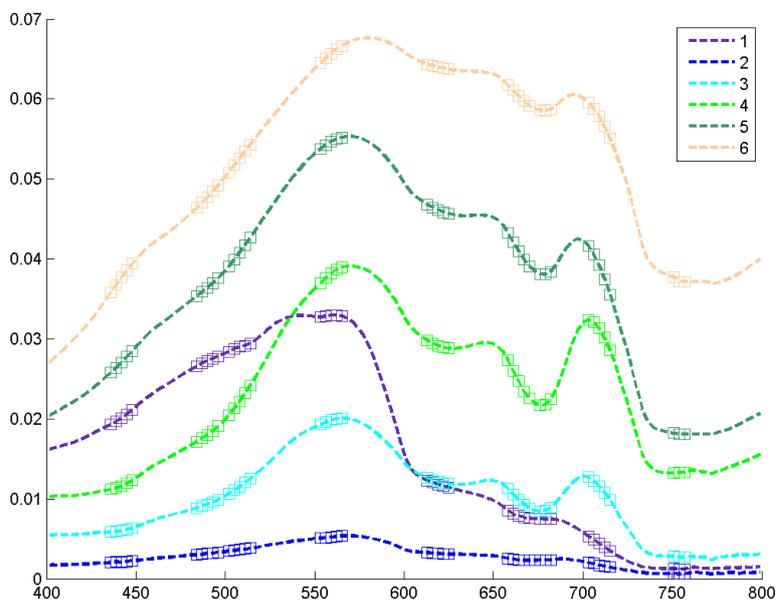
6c)
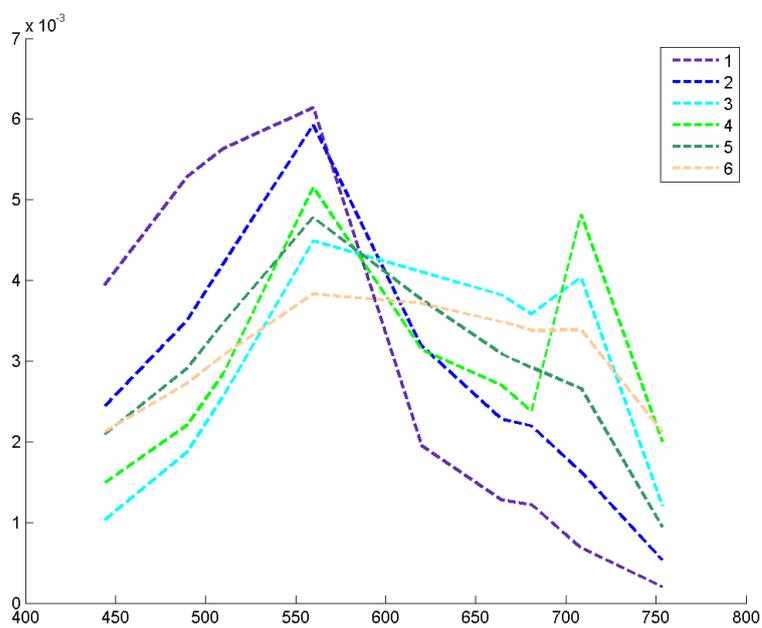
Figures 6a, b and c. Plots of the mean class spectra.

3.2 Methods for classification of the GLaSS nearby lakes

The three classification methods (5C, 6C and 6Cnormalised) were applied to the GLaSS nearby lakes in Estonia, Finland, Italy, the Netherlands and Sweden. Before the classification can be applied, pre-processing and atmospheric correction are required. MERIS 3rd reprocessing data were used as input.

Next, a suitable atmospheric correction method was chosen based on the results (Chi-square, Spectral Angle) as presented in the GLaSS D3.2 report on atmospheric correction. This atmospheric correction method differed per area. For some areas more than one atmospheric correction method was applied because the scores in D3.2 were comparable, or because one method had led to better results for Chi-square, while the other method apparently led to a better spectral shape and therefore a lower Spectral Angle.
So for the classification of the nearby lakes, not a full matrix with all atmospheric correction processors combined with all three classification methods was used. Just the atmospheric correction processors that had shown the best results either for the Chi-square or for the Spectral angle in D3.2 for specific regions were applied here for those regions.

However, not all atmospheric correction processors produce an output that is suitable as input for the OWT tool. The OWT tool requests all MERIS spectral bands 1-10, while for example the FUB AC-processor does not produce all of these as output. SCAPE-M was not included in the classification analysis, because of known problems with MERIS band 2, which would have a large influence on the produced classes. The MEGS processor was not included because of the extremely low valid pixels it produced in the atmospheric correction tests (0-14%, depending on the lake, D3.2). The results of 6S and ATCOR compared to the other atmospheric correction processors were not the best for any of the selected lakes, and have therefore been left out of this exercise. The remaining atmospheric correction processors that were used to generate the input for the OWT tool were therefore C2R, CC2R and MIP. The resulting classes are presented in sections 3.3-3.5. In a next stage, the images were pre-processed with ICOL to reduce the adjacency effects of nearby land in coastal pixels. These results are presented in 3.6. The classified maps of the nearby lakes are presented in section 3.7.

The classification tool is vulnerable for differences is spectral shape and therefore also for errors in the atmospheric correction. From D3.2 and literature it is known that atmospheric correction over highly absorbing lakes is extremely difficult and most processors do not produce suitable results for these lakes. The absence of FUB in this exercise is therefore expected to led to less good classification results for the highly absorbing lake types e.g. in Finland and Sweden, because FUB produced the best atmospheric corrected products for these lakes (D3.2).

3.3 Results 5C classification

The 5C classification results (Table 4) could not easily be explained. Within one image, lakes that are seen as different types (GLaSS D. 5.1, Eleveld et al., 2014) were classified in one class, such as the Dutch lakes Markermeer (dominated by SPM) and IJsselmeer (dominated by Chl). The same is seen for clear and CDOM dominated Swedish Vänern and its very turbid bay Dättern. Between images, the same occurs, e.g. the Dutch and Finnish lakes being classified as 3. Also unexpected was that parts of the blue Italian lakes were classified as 3. Part of this is probably due to adjacency effects in these clear lakes, which leds to higher reflectance in the NIR and therefore a too high class.

Table 4. 5C classification results

Area + date	AC-corr	Class 1	Class 2	Class 3	Class 4	Class 5	Notes and/or class sum

		"Blue"[1]	"Low"[1]	"Chl"[1]	"Chl + TSM"[1]	"TSM + Chl"[1]	note 1
Estonia 20050718	CC		North Peipsi	Mid+South Peipsi, Võrtsjärv. *All flagged as 'L2R suspect'.*			North Peipsi 0.5-1.5, Mid+South Peipsi, Võrtsjärv <0.01-0.5.
Estonia 20110727	MIP	North of Peipsi and Võrtsjärv, some parts of South of Peipsi		Most of South of Peipsi			0.04 - > 1 (lowest values is South of Peipsi)
Finland 20040805	C2R	Borders of almost all lakes		Pääjärvi, Päijänne			0.3-0.4
Finland 20060509	C2R	Borders of almost all lakes		Pääjärvi, Päijänne			0.1-1.3
Finland 20040805	CC		Part of Päijänne	Most of Päijänne, Pääjärvi			Pääjärvi < 0.01. The pixels of Päijänne with class 2 have sums around 1, the other pixels < 0.5
Finland 20060509	CC		Pääjärvi, Päijänne				0.1-1.5, large parts > 1
Finland 20070601	CC		Open parts of Päijänne	Parts of Päijänne, Pääjärvi. *Part of these areas in Päijänne and complete Pääjärvi are flagged as 'L2R suspect' or 'L2R invalid'*			Pääjärvi < 0.2. The pixels of Päijänne with class 2 have sums around 1, the other pixels < 0.5
Finland 20070823	MIP	Part of Pääjärvi and some pixels of Päijänne		Most of Päijänne and part of Pääjärvi			~0.3-1
Italy 20090911	CC	South of Garda		North Garda, Di Como, Maggiore. *All flagged as 'L2R suspect, and part of Di Como and Maggiore as 'L2R invalid'*			Garda 0.2-1, Di Como and Maggiore <0.5,
Italy 20080506	MIP			South of Garda			North of Garda not processed. Sums 0.5 - > 1
Netherlands 20110415	C2R	All small inland lakes	Most of IJsselmeer, part of Markermeer	Most of Markermeer, parts of IJsselmeer.			0.4-1.5
Netherlands 20110423	CC		Parts of IJsselmeer. *Some coastal pixels flagged as 'L2R suspect'*	Markermeer, most of IJsselmeer. *Some coastal pixels flagged as 'L2R suspect'*			0.5-1.5
Netherlands 20110928	MIP			Ijsselmeer and Markermeer			0.3 - > 2
Sweden 20030829	CC		Largest part of Vänern	Southern parts of Vänern, also Dättern. *Dättern flagged as 'L2R invalid'*			Vänern 0.5-1.5, Dättern <0.5
Sweden	MIP			All parts of Vänern			Large areas not

20090626				that were processed			processed. Sums < 0.5

Note 1: the descriptions of these classes (within quotes) are just to easier read the table and interpret the results. It is important to realise that the classes do not represent concentration ranges.
Note 2: CC2R flagged large parts of many of the lakes 'TOSA out of scope'. For readability this is not indicated in the table.

3.4 Results 6C classification

The method with 6 classes seems to better distinguish lakes within one image that the 5C method (Table 5). Markermeer and IJsselmeer are distinguished after atmospheric correction with CC or C2R. Also the Swedish lakes are classified according to expectations. For the clear lakes (Italy) and the more narrow areas (Finnish lakes) still unexpected classes (class 4) are found.

Clearly, the results are influenced by (errors in the) atmospheric correction. For example, the Finnish lakes Pääjärvi and Päijänne can be classified as 1, 2, or 3 depending on the atmospheric correction that is used.

After MIP atmospheric correction too many lakes were classified in class 1 (Pääjärvi, Päijänne, North of Peipsi, Võrtsjärv, North of Vänern). The main reason for this result seems to be the spectral shape: MIP had produced low Chi-square values in the atmospheric correction tests, however, its spectral angles were not the lowest compared to other processors (D3.2). The spectral examples in D3.2 (Figures 20-24) indicate that also this processor has a spectral tendency on increased reflectance in the blue bands, which may result from a non-optimal aerosol spectral slope applied here. This spectral tendency can explain why lakes are classified in class 1. However, also other reasons may impact the OWT classification, e.g. which SIOP's are used in the processor configuration for the coupled atmospheric and in-water influence. It may be noted, that MIP is the only processor that allows direct configuration of the aerosol slope and the SIOP's, and the only sensor independent AC program that also comes with an adjacency processor. However an optimization for MERIS was not performed here.

From the remaining two processors (CC2R and C2R), CC2R had shown the best results in the atmospheric correction test (D3.2) for MERIS. This is reflected in these classification results: the classifications with CC2R as input where more as expected than those based on C2R input.

Flagging
CC2R flagged large parts of many of the lakes 'TOSA out of scope', so this flag was ignored to obtain data. Also, the pixels flagged with this flag did not show bad results. Flags that should not be ignored are CC2R flag 'L2R suspect' and CC2R flag 'L2R invalid'. In many cases these flags could explain unexpected classes. This occurred for the shorelines of the Finnish lakes Päijänne and Pääjärvi which were classified as 4 while class 2 was expected, parts of Italian lakes Di Como and Maggiore that were classified as 4 while class 1 was expected, and for the shorelines of the Dutch lakes IJsselmeer and Markermeer that were classified as 1, while classes 3 respectively 4 were expected. The Italian and Finnish lakes and shorelines were probably influenced by adjacency effect, which elevates the NIR part of the spectrum and therefore leds to a higher class. It makes sense that especially these lakes with a low reflectance are vulnerable for this effect.
For the Dutch lakes, the explanation for raising the 'L2R suspect' flag and the following misclassification is not known.

However, just flagging out all suspect and invalid flags does not always led to the best results. In cases of very turbid lakes, that contain high concentrations of both SPM and CDOM (e.g. Estonian lake Võrtsjärv and Swedish lake Dättern), this flag flags the complete lakes so no information would be obtained, while in these cases the obtained class 4 seems suitable. The main differences between flagging these complete lakes and application of the same flags in the cases of adjacency, is that the adjacency does not occur in the centre of a lake.

Therefore, for further analysis of the 6C results, ICOL pre-processing was applied to correct for the

adjacency effect (see section 3.6).

Table 5. 6C classification results

Area + date	AC-corr	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Notes and/or class sum
		"Blue"[1]	*"Low"* [1]	*"Chl"* [1]	*"Chl+TSM"* [1]	*"TSM+Chl"* [1]	*"Extreme TSM"* [1]	note 1
Estonia 20050718	CC		Part of North Peipsi	Part of North Peipsi	South Peipsi, Võrtsjärv. *All flagged as 'L2R suspect'.*			South Peipsi, Võrtsjärv 0.001-0.5. North Peipsi 0.2- > 1.
Estonia 20110727	MIP	North Peipsi and Võrtsjärv		Few pixels in South of Peipsi	South Peipsi			0.002-0.5 (lowest values in South of Peipsi)
Finland 20040805	C2R	Pääjärvi, Päijänne						0.01-0.1
Finland 20060509	C2R	Coasts of Pääjärvi and Päijänne		Pääjärvi, Päijänne				0.01-0.9
Finland 20040805	CC		Open part of Päijänne	More narrow parts of Päijänne	Shorelines of Päijänne, Pääjärvi. *All fagged as 'L2R suspect' or 'L2R invalid'*			Pääjärvi <0.1. Päijänne pixels with class 2 sum > 0.5, Päijänne pixels class 3 or 4 sum < 0.5
Finland 20060509	CC	Some pixels Pääjärvi	Päijänne, some pixels Pääjärvi	Most of Pääjärvi	Some pixels Pääjärvi, *these are mostly flagged as 'L2R suspect' or 'L2R invalid'*			Päijänne 0.5-1, Pääjärvi 0.01-0.9
Finland 20070601	CC	Some parts of Päijänne	Open parts of Päijänne	Some parts of Päijänne *Part of these areas are flagged as 'L2R suspect' or 'L2R invalid'*	Pääjärvi. *Flagged as 'L2R suspect' or 'L2R invalid'*			Pääjärvi < 0.2. The pixels of Päijänne with class 2 have sums around 1, the other pixels < 0.5
Finland 20070823	MIP	Päijänne and Pääjärvi						0.07-0.3
Italy 20090911	CC	South Garda		North Garda, North Di Como, North Maggiore. *All fagged as 'L2R suspect'*	South Di Como and parts of Maggiore. *All flagged as 'L2R invalid'*			0.01-0.7
Italy 20080506	MIP	South Garda		North Garda				~0.15-~0.3
Netherlands 20110415	C2R	All inland lakes	Markermeer, most of IJsselmeer	Parts of IJsselmeer				0.1-1
Netherlands 20110423	CC	Some shorelines of IJsselmeer and		IJsselmeer	Most of Markermeer, some inland lakes. *The*			<0.1 – 1.5 most between 0.5-

		Markermeer. *Half of these flagged 'L2R suspect'*			*inland lakes are flagged 'L2R suspect'*			0.9
Netherlands 20110928	MIP				Ijsselmeer and Markermeer			0.1-1.2
Sweden 20030829	CC		Part of Vänern and Vättern	Part of Vänern and Vättern	Dättern, *flagged 'L2R invalid'*			<0.1 – 0.9
Sweden 20090626	MIP	Northern part of Vänern		Most of Vänern, including Dättern				0.1-0.9

Note 1: the descriptions of these classes (within quotes) are just to easier read the table and interpret the results. It is important to realise that the classes do not represent concentration ranges.
Note 2: CC2R flagged large parts of many of the lakes 'TOSA out of scope'. For readability this is not indicated in the table.

3.5 Results 6Cnormalised classification

The normalised classification was expected to have two advantages:
1) Be invulnerable to 'white errors' in atmospheric correction. Therefore, the expectation was that this method would led to good results for those atmospheric correction processors that scored well for the spectral angle (D3.2 report on atmospheric correction, because the spectral angle is a measure to compare the spectral shapes while ignoring the intensity).
2) Be able to distinguish Finnish lakes based on their spectral shape. There are large differences between the Finnish lakes with regard to Chl, TSM and CDOM. However, because most lakes have relatively higher CDOM concentrations that other 'global' lakes, the Finnish lakes tend to be all grouped in one class with lakes with low reflectance (class 2).

Indeed, the 6Cnormalised classification was the only method that could make a distinction between the low-reflecting Swedish lakes Vänern (CDOM rich) and Vättern ('blue'). However, the Finnish lakes Päijänne and Pääjärvi were still classified similarly, independent of the atmospheric correction method that was used.
For the Estonian lakes more different classes are found based on the 6Cnormalised classification (after CC2R correction), which can be expected based on the optical variability of these lakes. The Dutch lakes IJsselmeer and Markermeer are however classified as 1 and 2, which is incorrect. These lakes are of different types and belong to much higher classes.

For further analysis of the 6Cnormalised results, ICOL pre-processing was applied to correct for the adjacency effect (see section 3.6).

Table 6. 6Cnormalised classification results

Area	AC-corr	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Notes and/or class sum
		"Blue"[1]	*"Low"*[1]	*"Chl"*[1]	*"Chl+TSM"*[1]	*"TSM+Chl"*[1]	*"Extreme TSM"*[1]	note 1
Estonia 20050718	CC2R	Some lake shores	Part of North Peipsi	Part of Võrtsjärv		Parts of North Peipsi and Võrtsjärv	South of Peipsi	0.1-1.6

Estonia 20110727	MIP				Small area in south of Peipsi		Almost complete Peipsi and Võrtsjärv	Extremely low sums (<<0.001)
Finland 20040805	C2R							Not many inland lakes processed. The ones that are processed are classified as class 1. Extremely low sums (<<0.001)
Finland 20060509	C2R	Päijänne, Pääjärvi	Vesijärvi					Extremely low sums (<<0.001). Vesijärvi 0.01-0.3
Finland 20040805	CC2R	Part of Päijänne, Vesijärvi	Part of Päijänne, Vesijärvi				Most of Päijänne, Pääjärvi. Shores of Vesijärvi	Pääjärvi <<0.001 – 0.3, Päijänne <<0.001 – 1.8, Vesijärvi <<0.001
Finland 20060509	CC2R	Part of Päijänne,	Part of Päijänne, Vesijärvi				Shores of Päijänne, Pääjärvi	0.003 - >2
Finland 20070601	CC2R	Päijänne	Vesijärvi				Shores of Päijänne, Pääjärvi	Most <<0.001, however, there are pixels with sum >1
Finland 20070823	MIP	Parts of Vesijärvi and Päijänne	Part of Vesijärvi				Parts of Päijänne, complete Pääjärvi	Extremely low sums (<<0.001).
Italy 20090911	CC2R	Most of Garda					Di Como and Maggiore	Extremely low sums (<<0.001)
Italy 20080506	MIP	Garda						Extremely low sums (<<0.001).
Netherlands 20110415	C2R	Part of Ijsselmeer	Parts of Ijsselmeer and Markermeer		Part of Markermeer			Ijsselmeer <<0.001-0.3, Markermeer 0.01-0.9
Netherlands 20110423	CC2R	Part of Ijsselmeer	Most of IJsselmeer, Markermeer				Inland lakes	<<0.001-1.2
Netherlands 20110928	MIP		Markermeer and most of IJsselmeer		Part of IJsselmeer	Small area in IJsselmeer		Extremely small (<<0.001) -> 1
Sweden 20030829	CC2R	Vättern	Vänern	Dättern		Small part of Vänern that connects with Dättern		Vänern <<0.001-0.6, the parts classified as 3 or 5 sum 0.5 - > 1.
Sweden 20090626	MIP	Large parts of Vänern	Southern parts of Vänern	Part of Dättern		Small part of Vänern that connects with Dättern and some bays in the North	Part of Dättern	<<0.001 - > 2, the parts classified as 3, 5 and 6 have the largest sums.

Note 1: the descriptions of these classes (within quotes) are just to easier read the table and interpret the results. It is important to realise that the classes do not represent concentration ranges.
Note 2: CC2R-NN flagged large parts of many of the lakes 'TOSA out of scope'. For readability this is not indicated in the table.

3.6 Results after adjacency correction

Because some results of the previous section were clearly influenced by adjacency effects (too high classes were assigned in narrow parts of the lakes and close to the shores), the ICOL processor was used for pre-processing to remove the adjacency effect, before processing with the CC2R-NN atmospheric correction processor. Next, the 6C and the 6C normalised classification methods were applied. The results are presented in Tables 7 and 8.

As expected, the incorrectly assigned high classes in coastal areas and narrow lakes were not found any more after ICOL processing. This effect was mainly seen in the narrow Italian and Finnish lakes. Especially after application of the 'L2R invalid' flag, the remaining classes of these lakes were much closer to the expected class (1) than before ICOL processing. For the Italian lakes ICOL+CC2R 6Cnormalised leds to the best results (all lakes in class 1).
As expected, for the lakes with larger areas in Estonia, the Netherlands and Sweden, ICOL-preprocessing does not have an effect except of along the shorelines. There is no positive, but there also does not seem to be a negative effect. Therefore, if it has to be decided if ICOL should be applied before classification for an unknown lake, it is probably best to apply ICOL. Therefore, to produce the maps (section 3.7) ICOL has been applied.

Flagging after pre-processing with ICOL
After ICOL processing, most misclassifications related to adjacency effects seems to have disappeared. The 'L2R suspect' and 'L2R invalid' flags still remove some misclassified pixels along the lake shores (e.g. in the Netherlands, Finland), which could be caused by remaining adjacency effect, but also, especially for Finland where the issue is only seen for the first pixel, by mixed land-water pixels. However, these flags tend to flag out large areas of valid pixels (for Dättern, Võrtsjärv, most of the Italian lakes). Therefore, to determine to which class the main part of the lake belongs, it is advised not to use additional flagging (besides land and cloud related flags). The not-normalised results changed more after ICOL processing than the normalised results, indicating that the absolute classification results are more vulnerable to adjacency effects than the normalised results.

Table 7. Results of the 6C classification after adjacency correction

Area + date	AC-corr	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Note on flags
Estonia 20050718	ICOL + CC2R		Part of North Peipsi	Part of North Peipsi	South Peispi and Võrtsjär v			'L2R invalid' would flag out the south of Peipsi and Lake Võrtsjärv
Finland 20060509	ICOL + CC2R	Some pixels of Vesijärvi	Päijänne	Most of Pääjärvi and Vesijärvi				'L2R invalid' can be used to flags out some shoreline pixels that were classified as 4.
Italy 20090911	ICOL + CC2R	Main part of Garda. *A part of this area is*	Lugano *Flagged as 'L2R invalid*	The northern part of Garda and Maggiore				Other lakes, e.g. Maggiore and Di Como,

		flagged as 'L2R suspect'		*Partly flagged as 'L2R suspect' or L2R invalid'*				are > 90% flagged 'invalid'.
Netherlands 20110423	ICOL + CC2R	Small shorelines of IJsselmeer and Markermeer. *These + some areas around are flagged 'L2R suspect'*	Parts of IJsselmeer	Most of Ijsselmeer and half of Markermeer	Half of Marker meer			'L2R invalid' – this flags out some shoreline pixels that were classified as 1.
Sweden 20090626	ICOL + CC2R		Part of Vänern and Vättern	Part of Vänern and most Vättern	Dättern			Clouds and cloud shadow. 'L2R invalid' would flag out Dättern

Table 12. Results of the 6Cnormalised classification after adjacency correction

Area + date	AC-corr	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Note on flags
Estonia 20050718	ICOL + CC2R		Parts of North of Peipsi			Most of North of Peipsi, Võrtsjär v	South of Peipsi	'*L2R invalid*' would flag out the south of Peipsi and Lake Võrtsjärv
Finland 20060509	ICOL + CC2R	Some pixels of Vesijärvi	Päijänne, about 1/3 of Pääjärvi	Most of Pääjärvi and Vesijärvi				'*L2R invalid*' can be used to flags out some shoreline pixels that were classified as 4.
Italy 20090911	ICOL + CC2R	Garda, Maggiore, Di Como					Few shoreline pixels	'*L2R invalid*' flags out almost all of the lakes
Netherlands 20110423	ICOL + CC2R	Parts of IJsselmeer and some shores of Markermeer.	Most of Ijsselmeer and Markermeer					
Sweden 20090626	ICOL + CC2R	Vättern	Vänern			Small part of Vänern that connects with Dättern and Dättern		Clouds and cloud shadow. '*L2R invalid*' would flag out Dättern

3.7 Classified maps

To produce the maps, all lakes were treated the same with regard to pre-processing (3rd re-processing, followed by ICOL) and atmospheric correction (CC2R) because this combination led in the previous sections to the overall best results. Next, both the 6C and 6Cnormalised classification

methods were applied. Both maps are presented and the results discussed.

Estonian lakes

The three classes that were found in Lake Peipsi have spectra that are similar to field measurements and the spatial distribution of the classes looks suitable. The southern part of Lake Peipsi has more sediments than northern part. The southern part of Lake Peipsi (L. Pihkva) is very similar to Lake Võrtsjärv, which is also seen in the classification. The northern part of Lake Peipsi is classified with classes 2 and 3. Although class 2 could occur in this part of the lake, at the time of image acquisition (July 18 2005) there was a large phytoplankton bloom in the northern part of the lake. In the beginning of July 2005, the measured Chl-a varied between 14-74 mg m^3, being lower close to the shore and higher in the centre. In August the bloom was even more intense. Therefore, a mixture of classes 3 and 4 would be expected. The classification of 2 and 3 can probably explained by the fact the Peipsi also has a relatively high CDOM concentration of around 1 m^{-1} (at 440 nm), which is not unusual. In situ measurements from 2008-2011 gave an average value of CDOM in Peipsi s.s of 3.1 m^{-1} (Elar Asuküll's Master thesis (TO)). ICOL processing did not change the results much. Also, the southern part of Peipsi and Võrtsjärv were still flagged as 'L2R invalid' after ICOL processing.

Label	Colour	
class_1		98, 46, 179
class_2		0, 0, 255
class_3		0, 255, 255
class_4		0, 255, 0
class_5		51, 153, 102
class_6		255, 204, 153
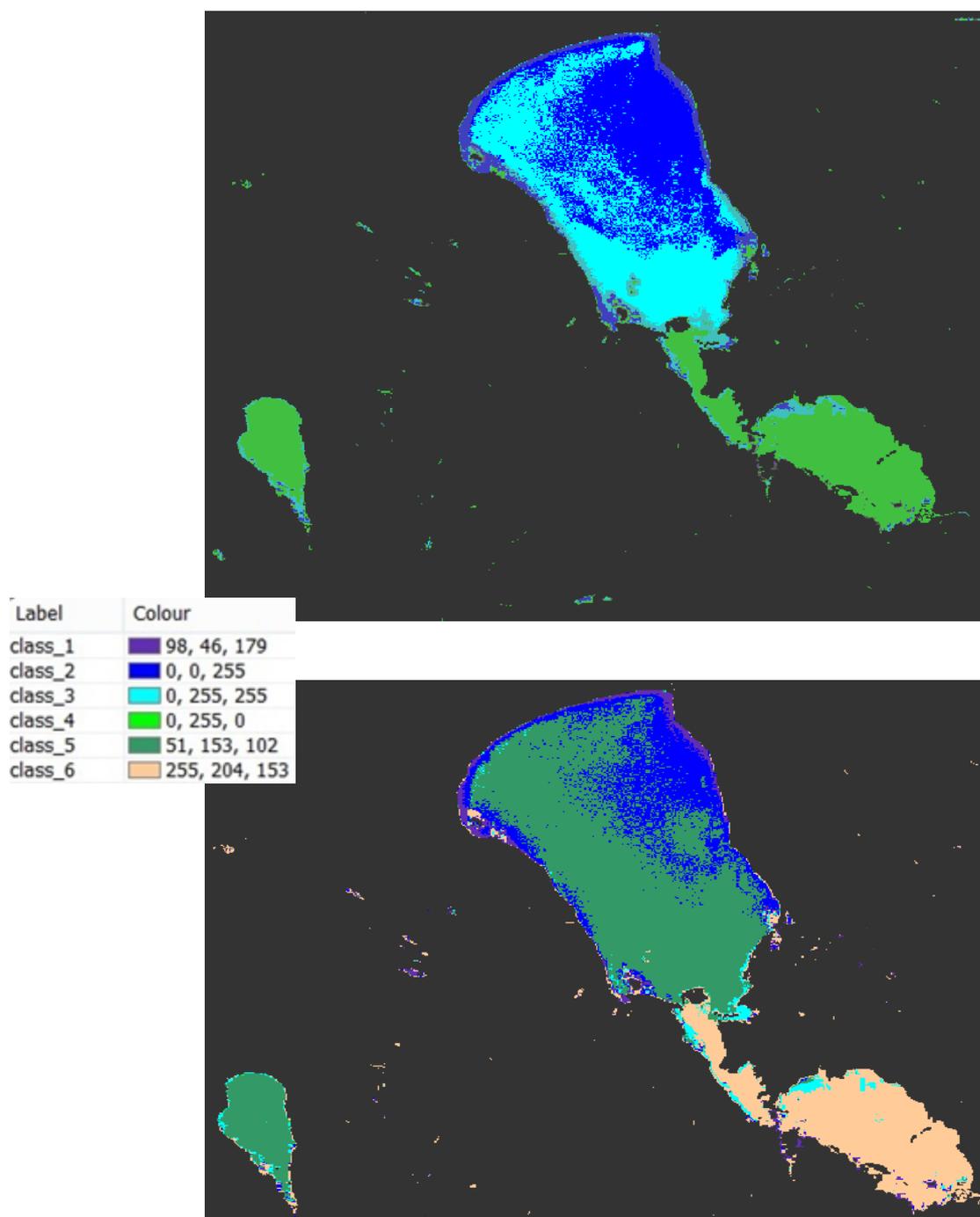
Figure 7. Estonian Lakes Peipsi (right) and Võrtsjärv (left). MERIS 20050718, ICOL+CC2R. Flagged (grey) with 'L2R suspect'. 'L2R invalid' was ignored because this would flag the south of Peipsi and complete Võrtsjärv. Top: OWT_6C, bottom: OWT_6Cnormalised

Finnish lakes

With the 6C classification, most of Päijänne and the central parts of Pääjärvi fall into class 2 (the class with the lowest Rrs).This is correct since these lakes have low reflectances (Pääjärvi due to CDOM absorption and Päijänne due to low TSM). Lake Vesijärvi (south of Päijänne and east of Pääjärvi) is in class 3. This is correct since that lake has low CDOM and typically higher TSM and Chl-a and thus higher Rrs (especially later in the summer). However, before adjacency correction, Lake Pääjärvi appeared to have a lot of variability (classes 1, 2, 3 and 4). This is not correct. After pre-processing with ICOL, Päijänne remains in class 2 and Vesijärvi in class 3 (plus a few pixels in class 1). The results for Pääjärvi improved. Although a classification with just class 2 was expected, after ICOL processing 2/3 of the lake was assigned to class 3 and 1/3 to class 2, but the higher classes in the centre of the lake disappeared. The 'L2R suspect' allows to flag out the too high class 4 pixels at the lake shores of Päijänne. With the 6Cnormalised tool, Päijänne and Pääjärvi fall in class 1 and Vesijärvi in class 2. It was indeed hoped for that the normalisation would be able to distinguish as a relatively blue lake, but Pääjärvi is not blue and should therefore have been classified as 2 instead. Because also other surrounding lakes are classified as 1, it seems that the normalization is not an improvement for Finland. At least not when applied after CC2R atmospheric correction, FUB atmospheric corrected data as input could have improved the results.
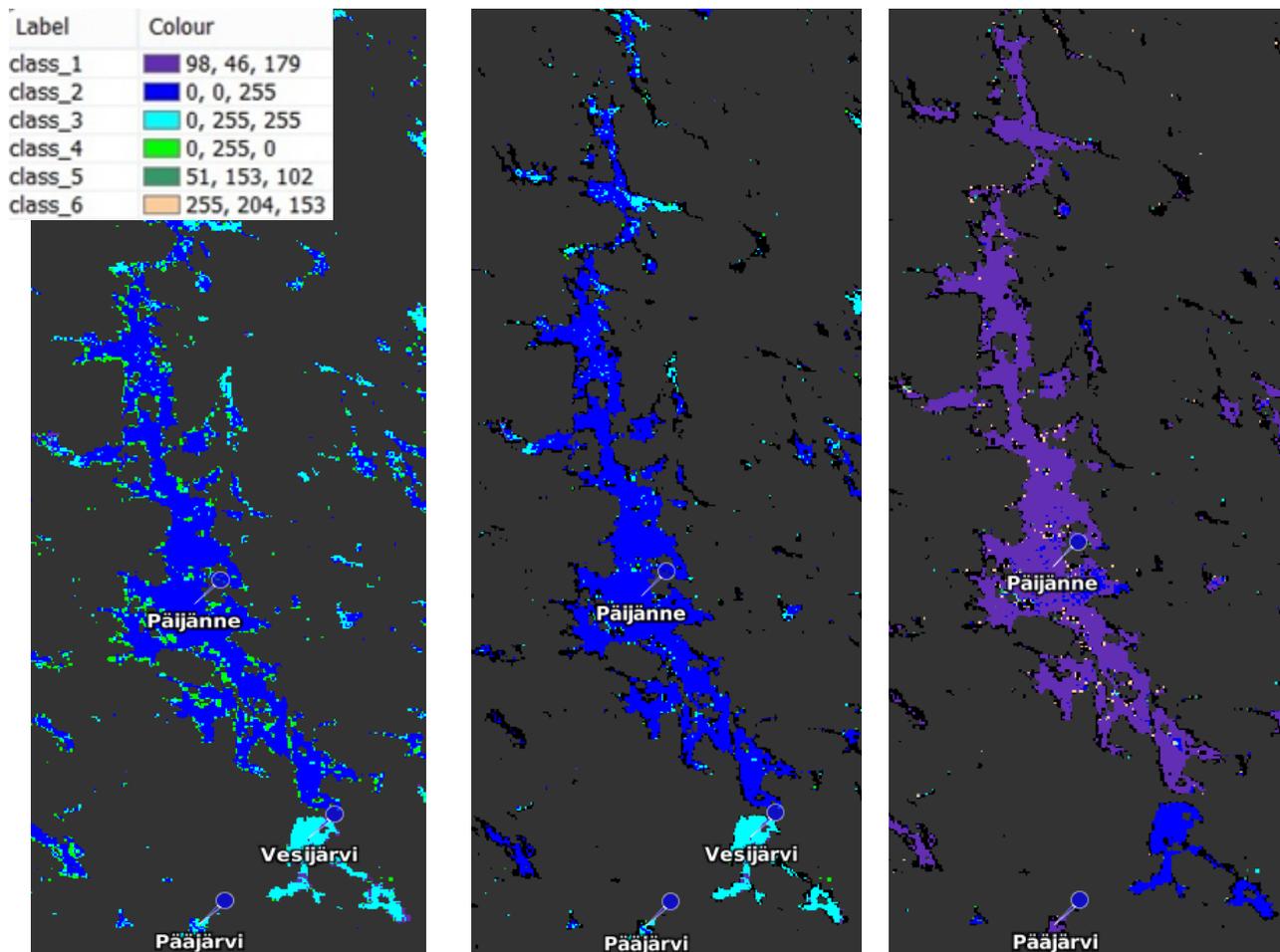
Figure 8. Finnish Lakes Päijänne, Pääjärvi and Vesijärvi. MERIS 20060509, ICOL+CC2R. Left: OWT_6C no flagging. Middle: OWT_6C 'L2R invalid' flagged out. Right: OWT_6Cnormalised with flagging

Italian lakes

Overall, class 1 should be suitable for all Italian lakes, but depending on the season and day also class 3 makes sense. The south of Garda should generally be classified as 1, so this is correct. However, without ICOL processing for adjacency effect correction, a large part of Garda and all the other lakes in the area were flagged as 'L2R invalid' or 'L2R suspect'. The classes found behind the flags were too high (e.g. classes 3 and 4 were found). This can be explained by the adjacency effect, which lifts the NIR region of the spectrum. After ICOL + CC2R processing, a very large area of the lake was flagged as 'L2R suspect' but when this flag was ignored, a larger area of Lake Garda was found to be classified as class 1 and a small part as class 3. The other Italian lakes were > 90% flagged as 'L2R invalid'. However, the classes found in these lakes (1, 2 and 3) are close to correct so the flagging seems to be too tight. Only in Lake Lugano and Idro, some parts are classified as 4, which seem incorrect.

Label	Colour	
class_1		98, 46, 179
class_2		0, 0, 255
class_3		0, 255, 255
class_4		0, 255, 0
class_5		51, 153, 102
class_6		255, 204, 153
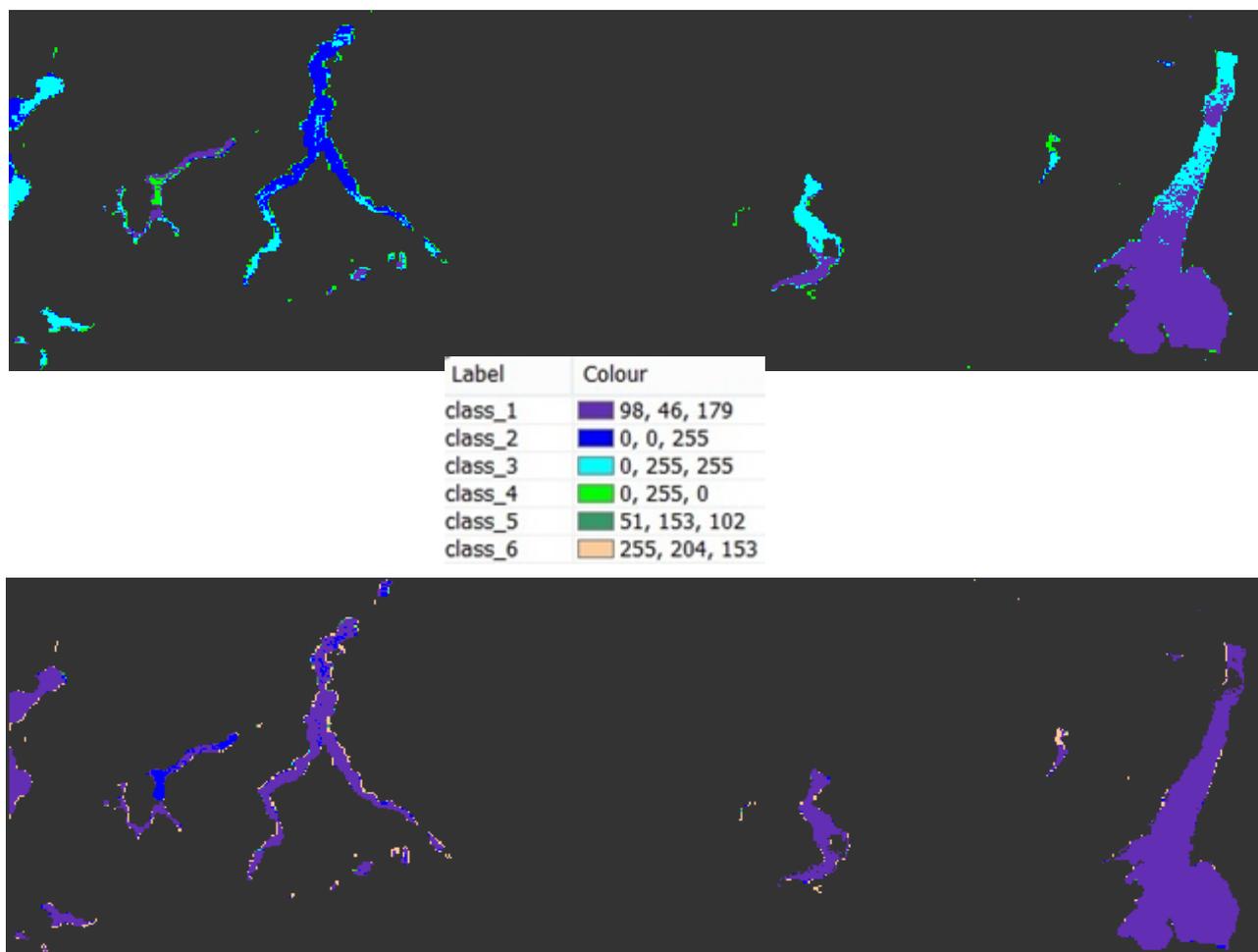
Figure 9. Italian Lakes, from left to right: Lago Maggiore (partly), Lugano (mostly a Swiss lake), Como, Iseo, Idro and Garda. MERIS 20090911, ICOL+CC2R. No flagging applied. Top: WT_6C, bottom: OWT_6Cnormalised

Lakes in the Netherlands

The Dutch Lake IJsselmeer and its splitoff Lake Markermeer have quite distinct optical properties. Markermeer is shallower and is characterized by fine, easily resuspendable sediments, and has therefore frequently higher surface TSM concentrations. The River IJssel discharges in lake IJsselmeer, which has therefore had higher nutrient inflows in the past (these have been reduced since the late 80's) and is still optically dominated by phytoplankton and cyanobacteria blooms. As expected, Markermeer is (partly) classified as class 4, while the main part of IJsselmeer is in class 3. Markermeer also contains class 3, which is correct. IJsselmeer also contains some class 2 pixels, which could be caused by high chlorophyll absorption. In both lakes some class 1 pixels were found along the shorelines. This is not correct and could not be explained. It was not solved by applying ICOL (adjacency effect would indeed not lead to a lower class). However, all these pixels were correctly flagged as 'L2R invalid' or 'L2R suspect'.

Label	Colour	
class_1		98, 46, 179
class_2		0, 0, 255
class_3		0, 255, 255
class_4		0, 255, 0
class_5		51, 153, 102
class_6		255, 204, 153
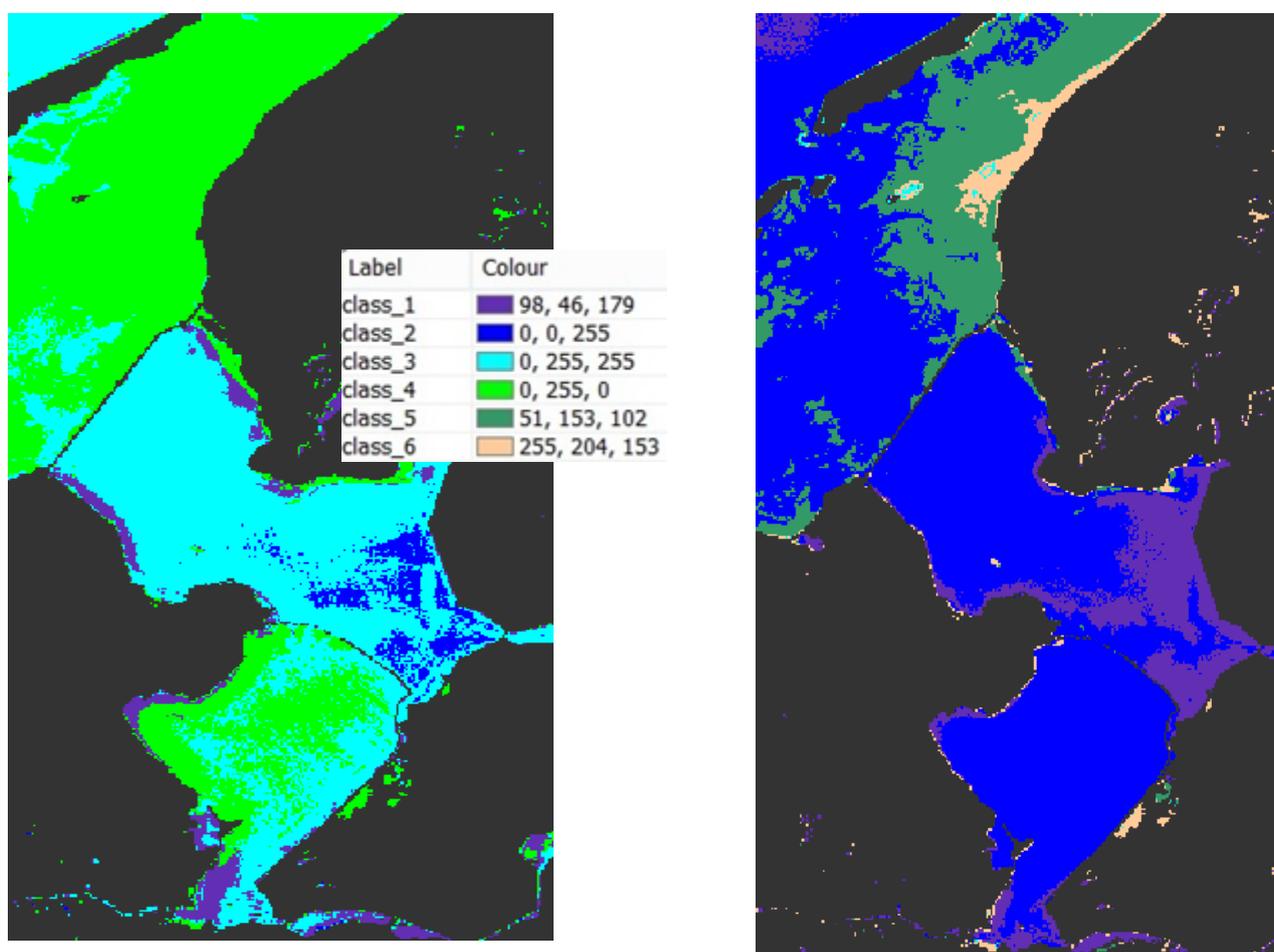
Figure 10. Dutch Lakes IJsselmeer (north of the dike) and Markermeer (south of the dike). MERIS 20110423, ICOL+CC2R. 'L2R invalid' flagged out, flagged (grey) with 'L2R suspect'. Left OWT_6C, right OWT_6Cnormalised

Swedish lakes

The two largest Swedish lakes are Vänern and Vättern. Vättern is very clear (very low concentrations of Chl, SPM and CDOM). Vänern has (low concentrated) Chl blooms and a relatively high CDOM concentration of around 1 m^{-1} (at 440 nm). Both lakes were largely classified as class 2, Vänern also partly as class 3. The small bay Dättern, on the south of the eastern basin of Vänern, is very turbid, with high concentrations of SPM (> 30 g m^{-3}), Chl (> 30 mg m^{-3} in summer) and extreme concentrations of CDOM (3-10 m^{-1}). Dättern was classified as 4, which makes sense, because of the spectral shape. Due to the high CDOM absorption it did not fall in classes 5 or 6. After ICOL processing the results for Vänern did not change much, while Vättern was now mainly classified as 3 (while 1 had been more suitable). Also, the cloud shadow flag was clearly influenced by ICOL. ICOL processes a certain number of pixels from the shore in one direction. The same shape was seen as cloud flag on one side of each cloud. Therefore, large areas were unnecessary flagged. Dättern continued to be flagged as 'L2R invalid' after ICOL processing. The result without ICOL processing is therefore presented.
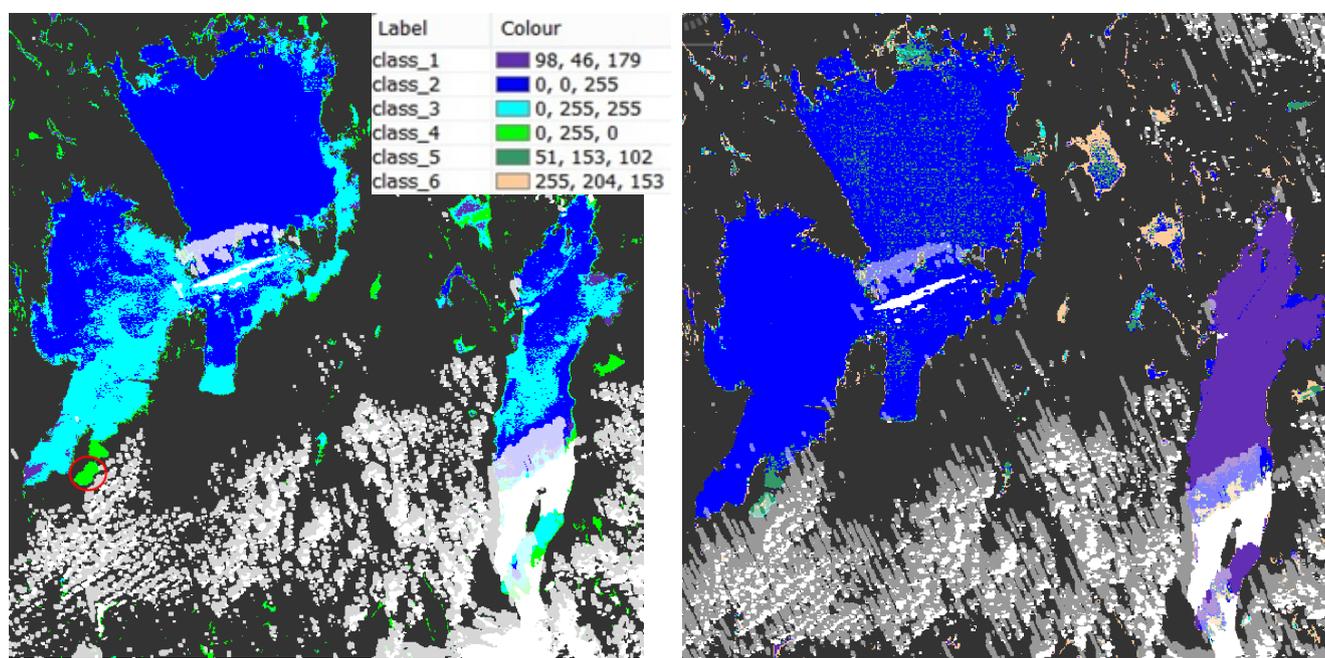
Figure 11. Swedish Lakes Vänern (left) and Vättern (right). The bay Dättern is indicated with the (red circle). MERIS 20060509, ICOL+CC2R. Only cloud (white) and cloud shadow (partly transparent white) flagging applied. Left: OWT_6C, right: OWT_6Cnormalised

3.8 Proposed processing chain

The aim of the classification tool is to produce a classified map for an unknown lake, which can assist the user in choosing the best suitable atmospheric correction schema and the best inland water constituent retrieval algorithm, or to choose the best suitable tuning of an constituent retrieval algorithm. In these cases, the classification tool will be used without much prior knowledge about the lake. Also, in these cases it will not be possible to pre-select the most suitable atmospheric correction schema and the most suitable classification method.

Therefore, based on the results of the previous sections, a processing schema is preoposed to apply to an unknown lake to obtain a proper classification of the lake.

Based on the results presented before, the proposed processing chain for unknown lakes is as follows (Figure 12. For an unknown lake, adjacency effect correction by ICOL is advised, followed by atmospheric correction with CC2R, because this combination led in the previous sections to the overall best results. Note: CC2R led to the best classes for most of the lakes. For analysis of what the best atmospheric correction schema is with regard to obtained reflectance and options to tune or apply to other sensors, the reader is referred to D3.2.

For an unknown lake it is hard to decide if the 6C or 6Cnormalised classification is best to use. For the Dutch lakes, the 6C classification is best, while for the Italian and Swedish lakes the 6Cnormalised classification provided the best results in this test. For the Estonian lakes, the 6C classification leads to somewhat lower classes than expected, while the 6Cnormalised classification leads to probably just too high classes. The assumption based on this small set of test lakes is that the 6Cnormalised classification is an improvement for the lakes or areas that were classified as class 3 with the 6C classification (part of the Italian and Swedish lakes, the Finnish lakes), while the 6Cnormalised classification looses the information for the lakes that were classified as 4 in the 6C classification (e.g. the Dutch lakes and the most turbid Estonian lakes). Therefore, both the 6C and 6Cnormalised classification method can be applied.

Although the 'L2R suspect' and 'L2R invalid' flags remove some misclassified pixels along the lake shores, which could be caused by remaining adjacency effect or by mixed land-water pixels, these flags also tend to flag out large areas of valid pixels. Therefore, to determine to which class the main part of the lake belongs, it is advised not to use additional flagging (besides land and cloud related flags). It is advised to ignore the often much higher classes that are found in the 1-2 pixels along the shores of a lake.
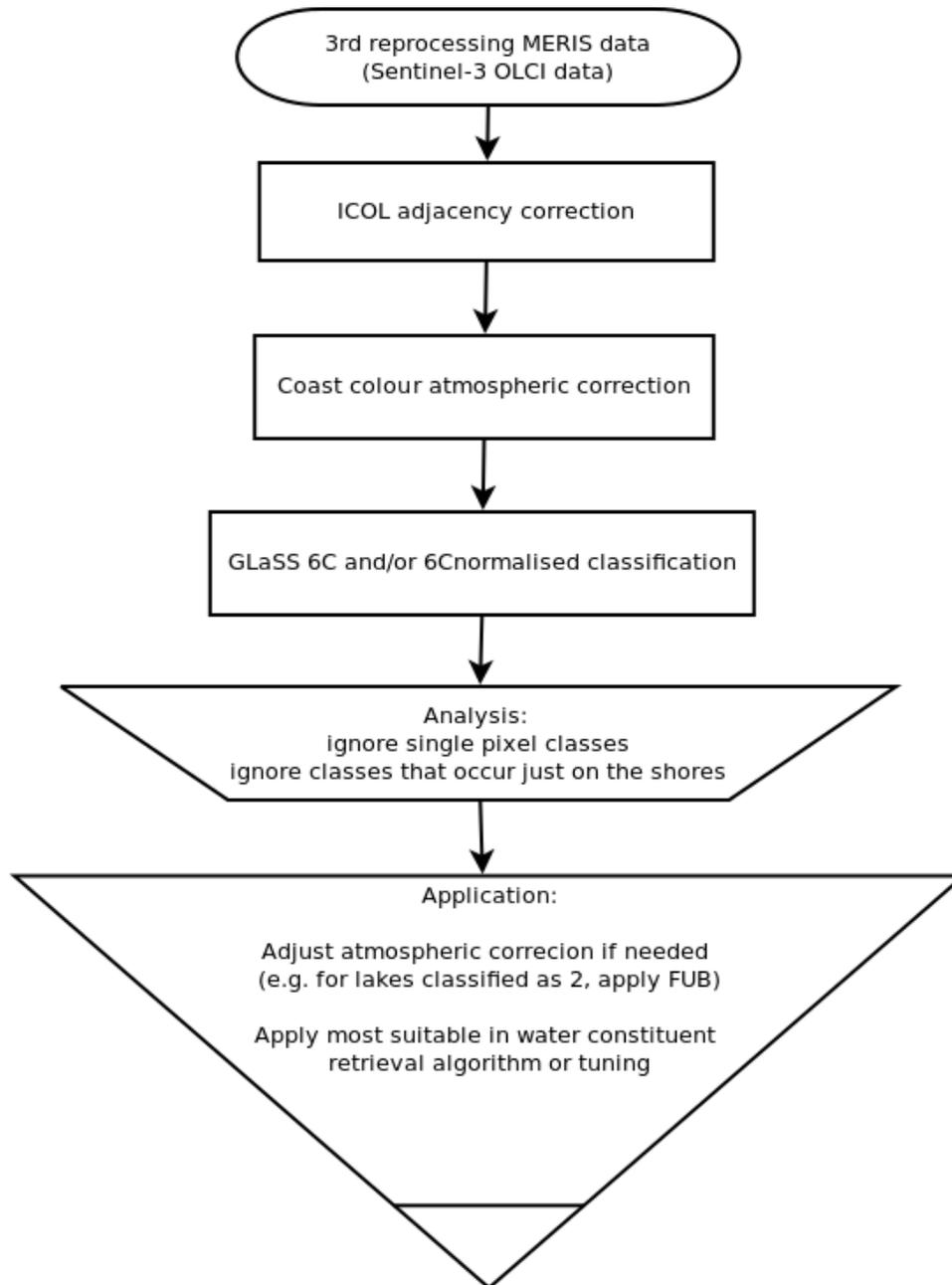
Figure 12 Proposed processing chain including the OWT tool (GLaSS 6C classification)

4 Conclusions and outlook

The differences in colour within or between lakes are a function of the optically active substances in these waters (under similar or normalised viewing and illumination geometries, and weather conditions). We classified these colours of the lakes by clustering a compiled set of in situ measured remote sensing reflectance spectra into three classification schemes (5C, 6C and 6Cnorm). These were then used to map the dominant spectral water types in several MERIS images. First results seem to indicate that processing MERIS L1b (MER_FSG_1PP) products with the CoastColour atmospheric correction (which generated CC2R2L products), and an OWT classification with 6 classes (absolute or normalised) gives credible results for many lakes. These made GLaSS partners enthusiastic about the possibilities that this tool might offer for testing different atmospheric correction schemes, and they are interested to investigate this further when developing their GLaSS case studies. In such a way, the colour of lake waters, captured as the remote sensing reflectance spectra in satellite images can be an extremely valuable indicator for a first inventory of global lakes.

We could also aim for selection of algorithms optimised for local conditions at a certain time, by using Moore et al.'s (2014) fuzzy approach to direct the algorithm selection and weigh the merging. Spectral matching could be used to assign the "best performing" algorithm to a particular water type, and class memberships can then be used to weigh retrievals into a blended product. However, this may direct us away from a more physically based approach in which we vary sIOPs directly for optimizing retrievals, and therefore this may be one step too far for GLaSS.

Nonetheless, within GLaSS, we will produce time series of satellite products and demonstrate the complementarily with field sampling for water quality and trophic status reporting. Algorithms with limited range, or hard classifiers, with boundaries that are fixed in time and space might not provide good quality output for extreme cases. In these situations, the optical preclassification can indicate where and when major changes in spectral water type occur, and show hotspots for additional collection of spectra, concentrations, and IOPs (SCI-datasets).

Now that the processor seems to work for MERIS, a logical next step would be to test on simulated OLCI and Sentinel 2 Rrs images. After all, the GLaSS core system will be built using existing components that will allow to access, pre-process and distribute very large quantities of data coming from S-2 and S-3.

5. Case studies as appendices

After the reference list, three appendices are included in the document to further investigate the use of OWT applied to MERIS images of different lakes. They also makes use of match-ups with in situ data to verify the results.

The first case study (Appendix 1) analyses the impact of different OWT classifications for different atmospheric corrections. The study is focused on shallow and turbid lakes in the Netherlands.

The second case (Appendix 2) shows how OWTs can be used to compare spatial variability (within and between lakes), as well as temporal spectral variability (in time-series). The work shows results for different European lakes, some of which are boreal, and others are trans-boundary (e.g., Lake Constance)

The third case (Appendix 3) shows the results for clear, deep lakes in the southern alpine eco-region, which are very different from the turbid shallow lake type discussed in Appendix 1.

References

Belzile, C.,. Vincent, W. F C. Howard-Williams, C. I. Hawes, I.M. R. James, M. R. Kumagai, M and Roesler C. S. (2004). Relationships between spectral optical properties and optically active substances in a clear oligotrophic lake, *Water Resour. Res.*, 40, W12512, doi:10.1029/2004WR003090.

Blondeau-Patissier, D., Brando, V. E., Oubelkheir, K. , Dekker, A. G., Clementson, L. A., Daniel, P. (2009). Bio-optical variability of the absorption and scattering properties of the Queensland inshore and reef waters, Australia. *J. Geophys. Res.* 114, C05003, doi:10.1029/2008JC005039.

Bradt, S. R. (2012). Development of bio-optical algorithms to estimate chlorophyll in the Great Salt Lake and New England lakes using in situ hyperspectral measurements. (PhD Thesis). The University of New Hampshire.

Brando, V.E, Dekker, A.G., Park. Y. J., Schroeder, T. (2012). Adaptive semianalytical inversion of ocean color radiometry in optically complex waters. *Appl. Optics* 51(15), 2808-2833.

Bresciani M., Stroppiana D., Odermatt D., Morabito G., Giardino C. (2011). Assessing remotely sensed chlorophyll-a for the implementation of the European Water Framework Directive in perialpine lakes, *Science of the total environment*, doi:10.1016/j.scitotenv.2011.05.001.

Craig, S.E., Jones, C.T., Li, W.K.W., Lazin, G., Horne, E., Caverhill, C., Cullen, J.J. (2012). Deriving optical metrics of coastal phytoplankton biomass from ocean colour. Remote Sensing of Environment 119, 72–83, doi:10.1016/j.rse.2011.12.007

Carlson, R.E. (1977) A trophic state index for lakes. *Limnology and Oceanography* 22(2): 361—369.

D'Alimonte, D., Mélin, F., Zibordi, G., Berthon, J.-F. (2003). Use of the Novelty Detection Technique to Identify the Range of Applicability of Empirical Ocean Color Algorithms *IEEE Transactions on Geoscience and Remote Sensing* 41(12), 2833-2843.

Dekker, A.G., Vos, R.J., Peters, S.W.M. (2001). Comparison of remote sensing data, model results and in situ data for total suspended matter TSM in the southern Frisian lakes. *The Science of the Total Environment* 268: 197-214.

Dekker, A.G., Vos, R.J., Peters, S.W.M. (2002). Analytical algorithms for lake water TSM estimation for retrospective analyses of TM and SPOT sensor data. *International Journal of Remote Sensing* 23(1):15-35, doi:10.1080/01431160010006917

Defoin-Platel, M., Chami, M. (2007). How ambiguous is the inverse problem of ocean color in coastal waters? *J. Geophys. Res.* 112, C03004, doi:10.1029/2006JC003847

Doerffer, R., & Schiller, H. (2007). The MERIS case 2 water algorithm. *International Journal of Remote Sensing*, 28, 517-535.

Eleveld, M. A. (2012). Wind-induced resuspension in a shallow lake from Medium Resolution Imaging Spectrometer (MERIS) full-resolution reflectances. *Water Resources Research* 48(4), W04508, doi:10.1029/2011WR011121

Eleveld M.A, Giardino, C., Bresciani, M., Gomarsca, M.A., A. Hommersom, 2014. Report socio-economic context & selected case studies. GLaSS Deliverable 5.1.

Eleveld, M.A., Peters, S.W.M., Lebreton, C., Stelzer, K., Attila, J., Simis, S.G.H., Boye-Hansen, L., Desmit, X., Vanhellemond, Q. (2013) Methods and protocols for satellite data harmonisation and error statistics. CoBiOS Deliverable 3.2.

Gallegos, C.L., Davies-Colley, R. J., Gall, M., 2008. Optical closure in lakes with contrasting extremes of reflectance. *Limnol. Oceanogr.* 53(5), 2021–2034.

Gjerstad, K., Stamnes, J.J, Hamre, B., Lotsberg, J.K., Yan, B., Stamnes, K. (2003). Monte Carlo and discrete-ordinate simulations of irradiances in the coupled atmosphere-ocean system. *Appl. Opt.* 42: 2609-2622.

GLaSS consortium, 2013. GlaSS DoW.

Gohin, F., Druon, J.N., Lampert, L. (2002). Afive channel chlorophyll algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. *International Journal of Remote Sensing* 23: 1639−1661.

Gons H., Rijkeboer M. & Ruddick K. (2005). Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. *Journal of Plankton Research* 27(1): 125–127.

Gurlin, D.; Gitelson, A.A.; Moses, W.J., 2011. Remote estimation of Chl-a concentration in turbid productive waters—Return to a simple two-band NIR-red model? *Remote Sens. Environ.* 115: 3479–3490.Heege, T., Kiselev, V., Wettle, M., Hung N.N. (2014): Operational multi-sensor monitoring of turbidity for the entire Mekong Delta . Int. J. Remote Sensing, Special Issues Remote Sensing of the Mekong, Vol. 35 (8), pp. 2910-2926

Hommersom, A., Peters, S., Wernand, M., Eleveld, M.A., van der Woerd, H.J., de Boer, J. (2011). Spectra of

a shallow sea – unmixing for coastal water class detection and monitoring. *Ocean Dynamics* 61(4): 463-480. doi:10.1007/s10236-010-0373-4

Hommersom A., Peters S., van der Woerd H.J., Eleveld M.A., de Boer J. (2010). Chi-square spectral fitting for concentration retrieval, automatic local calibration, quality control, and water type detection, *Canadian Journal of Remote Sensing* 36(6): 650-670. doi:10.5589/m11-004

IOCCG (2000). Remote sensing of ocean colour in coastal, and other optically-complex, waters. Sathyendranath, S. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 3, IOCCG, Dartmouth, Canada.

IOCCG (2009). Partition of the ocean into ecological provinces: Role of ocean-colour radiometry. Dowell, M. and Platt, T. (eds.), Reports of the International Ocean-Colour Coordinating Group, No. 9, IOCCG, Dartmouth, Canada.

IOCCG (2010). Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. Wang, M. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 10, IOCCG, Dartmouth, Canada.

Jin, Z., Stamnes, K. (1994). Radiative Transfer in nonuniformly refracting layered media: atmosphere-ocean system, *Appl. Opt.* 33, 431-442.

Kalff J, (2002). Limnology: Inland Water System. 3rd ed. Prentice Hall, Upper Saddle River (NJ).

Kirk, J.T.O. (1994). Light and Photosynthesis in Aquatic Ecosystems. 2nd ed. Cambridge University Press, Cambridge.

Kiselev, V., Bulgarelli, B., Heege, T. (2014): Sensor independent adjacency correction algorithm for coastal and inland water systemsm. Remote Sensing of Environment, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2014.07.025.Lindell, L.T., Pierson, D., Premazzi, G., Zillioli, E. (eds., 1999). Manual for Monitoring European Lakes using Remote Sensing Techniques (Luxembourg: European Communities).

Lee, Z.P., Carder, K.L., Arnone, R.A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics* 41, 5755–5772

Lee Z.P., Hu, C. (2006). Global distribution of Case-1 waters: An analysis from SeaWiFS measurements. *Remote Sensing of Environment* 101: 270 – 276.

Liu, J., Sun, D., Zhang, Y., Li, Y. (2013). Pre-classification improves relationships between water clarity, light attenuation, and suspended particulates in turbid inland waters. *Hydrobiologia* 711(1), 71-86

Lubac, B., Loisel, H. (2007). Variability and classification of remote sensing reflectance spectra in the eastern English Channel and southern North Sea. *Remote Sensing of Environment* 110, 45–58

Matthews, M.W., Bernard, S., Winter, K (2010). Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment* 114: 2070–2087.

Matthews, M.W., Bernard, S., Robertson, L. (2012) An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment* 124, 637–652.

Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J.-F., Canuti, E. (2011) Multi-sensor satellite time series of optical properties and chlorophyll-a concentration in the Adriatic Sea. *Progress in Oceanography* 91 (2011) 229–244.

Mobley, C.D. (1999) Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Opt.* 38(36), 7442–7455.

Mooij, W.M., S. Hülsmann, L.N. de Senerpont Domis, B.A. Nolet, P.L.E. Bodelier, P.C.M. Boers, L.M. Dionisio Pires, H. J. Gons, B. W. Ibelings, R. Noordhuis, R. Portielje, K. Wolfstein and E. H. R. R. Lammens (2005), The impact of climate change on lakes in the Netherlands: a review. *Aquat. Ecol.* 39(4), 381–400, doi:10.1007/s10452-005-9008-0.

Mooij, W. M., Janse, J. H., De Senerpont Domis, L.N., Hülsmann, S., Ibelings, B.W. (2007). Predicting the effect of climate change on temperate shallow lakes with the ecosystem model PCLake. *Hydrobiologia* 584(1):443-454. doi:10.1007/s10750-007-0600-2

Moore, T.S., Campbell, J.W., Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing* 39(8), 1767-1776.

Moore, T.S., Campbell, J.W., Dowell, M.D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sensing of Environment* 113(11): 2424–2430

Moore, T.S., Dowell, M.D., Bradt, S., Ruiz Verdu, A. (2014). An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. *Remote Sensing of Environment* 113(11), 2424–2430.

Nõges, P., Kangur, K., Nõges, T., Reinart., A. Simola, H., Viljanen, M. (2008). Highlights of large lake research and management in Europe. *Hydrobiologia* 599: 259-276. doi:10.1007/s10750-007-9233-8

Odermatt, D., Gitelson, A., Brando, V.E., Schaepman, M. (2012). Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of Environment* 118,116–126.

Odermatt, D; Giardino, C; Heege, T (2010). Chlorophyll retrieval with MERIS Case-2-regional in perialpine lakes. Remote Sensing of Environment 114(3):607 - 617.

O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research* 103, 24937−24953.

Palmer, SCJ, Hunter, PD, Lankester, T, Hubbard, S, Spyrakos, E, Tyler, A, Présing, M, Horváth, H, Lamb, A, Balzter, H and Tóth, VR, 2014, Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. *Remote Sensing of Environment*, in press.

Peters, S.W.M., Eleveld, M., Pasterkamp, R., van der Woerd, H., Devolder, M, Jans, S., Park, Y., Ruddick, K, Block, T., Brockmann, C., Doerffer, R, Krasemann, H., Röttgers, R., Schönfeld, W., Jorgensen, P.V., Tilstone, G., Martinez-Vicente, V., Moore, G, Sørensen, K., Høkedal, J., Johnsen, T.M, Lømsland, E.R., Aas, E. (2005).Atlas of chlorophyll-a concentration for the North Sea based on MERIS imagery of 2003.Amsterdam: Vrije Universiteit, 121 pp.

Reinart, A., Herlevi, A., Arst, H., Sipelgas, L. (2003) Preliminary optical classification of lakes and coastal waters in Estonia and south Finland. *Journal of Sea Research* 49, 357-366, doi:10.1016/S1385-1101(03)00019-4.

Ruiz-Verdu, A., Simis, S. G. H., de Hoyos, C., Gons, H. J., & Pena-Martinez, R. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. RemoteSensing of Environment, 112, 3996–4008.

Schenk et al. (GLaSS consortium, 2014). Harmonized atmospheric correction method. GLaSS Deliverable 3.2.

Schiller H, Schönfeld W, Krasemann HL, Schiller K (2007) Novelty detection—recognition and evaluation of exceptional water reflectance spectra. *Environ. Monit. Assess.* 132:339–350

Schroeder, T., Schaale, M., & Fisher, J. (2007). Retrieval of atmospheric and oceanic properties from MERIS measurements: A new Case-2 water processor for BEAM. *International Journal of Remote Sensing* 28, 5627–5632.

Shi, K., Li, Y., Li, L., Lu, H. (2013a) Absorption characteristics of optically complex inland waters: implications for water optical classification. *J. Geophys. Res. Biogeosci.* doi:10.1002/jgrg.20071

Shi, K., Li, Y., L,i L., Lu, H., Song, K., Liu, Z., Xu, Y., Li, Z. (2013b) Remote chlorophyll-a estimates for inland waters based on a cluster-based classification. *Sci Total Environ.* 444:1-15. doi: 10.1016/j.scitotenv.2012.11.058.

Stamnes, K., Tsay, S.C., Wiscome, W.J., Jayaweera, K (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.* 27, 2502-2509.

Stamnes, K., Tsay, S.C., Wiscome, W.J., Laszlo, I. (2000). DISORT, a general-purpose Fortran program for discrete-ordinate-method radiative transfer in scattering and emitting layered media: documentation of methodology. ftp://climate.gsfc.nasa.gov/pub/wiscome/MultipleScatt/.

Tilstone, G.H., Peters, S.W.M., van der Woerd H.J., Eleveld, M.A., Ruddick, K., Schönfeld, W., Krasemann, H., Martinez-Vicente, V., Blondeau-Patissier, D., Röttgers, R., Sørensen, K., Jørgensen, P. V., & Shutler, J. D. (2012). Variability in specific-absorption properties and their use in a semi-analytical Ocean Colour algorithm for MERIS in North Sea and Western English Channel coastal waters. *RS Env.* 118: 320-338. doi:10.1016/j.rse.2011.11.019

Nõges, P., Kangur, K., Nõges, T., Reinart., A. Simola, H., Viljanen, M. (2008). Highlights of large lake research and management in Europe. *Hydrobiologia* 599: 259-276. doi:10.1007/s10750-007-9233-8

Van der Woerd, H.J., Pasterkamp, R., Peters, S.W.M., Eleveld, M.A. (2004). How to deal with spatial variability in bio-optical properties in coastal waters: A case study of CHL-retrieval for the North Sea. Proceedings Ocean Optics XVII, 24–29 Oct.2004, Fremantle, Australia. (paper 184.pdf, 11 pp.)

Vantrepotte, V., Loisel, H., Dessailly, D., Mériaux, X. (2012) Optical classification of contrasted coastal waters. *Remote Sensing of Environment* 123, 306–323

Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., McClain, C.R. (2003). Unique data repository facilitates ocean color satellite validation. EOS Trans AGU, 84, 38,377.

Wetzel, R.G. (1983). Limnology. 2nd ed. Saunders College Publishing, Philadelphia.

Appendix 1: Mapping the impact of different OWT classifications for different atmospheric corrections of Dutch lakes

Marieke Eleveld

This case study was started with the following questions in mind:

- Can the GlaSS 5, 6, or 6n classes map variability between and within Dutch lakes?
- Do the 5 and 6 classes map impact of scattering because these classifiers are influenced by the height of the spectrum, while 6n maps absorption from the shape of the spectrum?
- What is the impact of different atmospheric corrections on the water type maps?

First the results from the clustering of in situ data spectra (courtesy Tim Moore) were examined. Going through the clustering of in situ classes from 5c to 6c and 6c-normalised (see the following three tables), shows that Marker- and IJsselmeer spectra end up in higher classes with each classification. Except for 2 spectra that were classified (perhaps misclassified) to class1, all spectra have been assigned to classes that show characteristics typically associated with absorption of in-water substances in case-2 waters.

Five classes

Non-normalised: 5 Classes

Source	Type 1	2	3	4	5	Total
Finnish	0	16	0	0	0	16
Taihu	0	0	14	129	91	234
Peipsi	0	0	24	10	0	34
Markermeer99	0	0	2	0	0	2
IJsselmeer99	0	0	5	0	0	5
IJsselmeer	2	6	45	0	0	53
Markermeer	0	0	71	0	0	71
ITA ASD	85	0	2	3	0	90
ITA WSD	12	0	0	0	0	12
ITA PR 650	0	3	0	0	0	3
Betuwe	0	12	4	0	0	16
NH Lakes	38	42	97	2	0	179
Spanish lakes	28	86	25	1	0	140
Lake Erie	4	0	9	3	0	16
Totals	169	165	298	148	91	871

Six classes

Non-normalised: 6 Classes

Source	Type 1	2	3	4	5	6	Total
Finnish	0	15	1	0	0	0	16
Taihu	0	0	1	41	108	84	234
Peipsi	0	0	10	21	3	0	34
Markermeer99	0	0	0	2	0	0	2
IJsselmeer99	0	0	1	4	0	0	5
IJsselmeer	0	0	49	4	0	0	53
Markermeer	0	0	16	55	0	0	71
ITA ASD	81	0	4	3	2	0	90
ITA WSD	12	0	0	0	0	0	12
ITA PR 650	0	3	0	0	0	0	3
Betuwe	0	8	8	0	0	0	16
NH Lakes	32	29	77	39	2	0	179
Spanish lakes	28	72	19	20	1	0	140
Lake Erie	2	0	4	10	0	0	16
Totals	155	127	190	199	116	84	871

For 28/09/2011 comparison of matchup in situ spectra and means from these cluster two analyses

field campaign sites

5 classes

Relative RMSE of 47.0 % between the in situ matchup and Type 2

6 classes

Relative RMSE of 8.6 % between the in situ matchup and Type 3
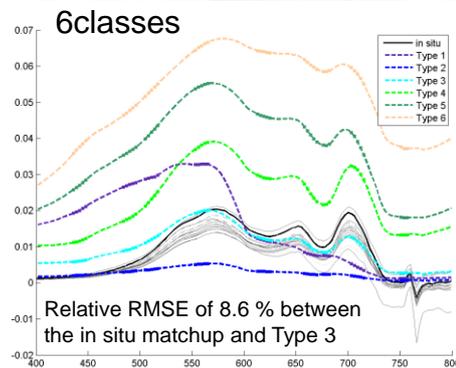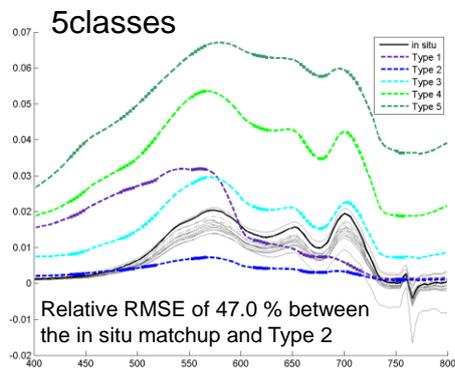
The relative RMSE between the drawn in situ line (st. 47, 1 min matchup) and all Types was calculated following Odermatt et al. (2010).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\overline{X}_i - \hat{X})^2}{N-1}}$$

$$Rel.\ RMSE = \sqrt{\frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N}\overline{X}_i}} \cdot 100$$

where N are the 134, 3 nm band intervals, and $\dot{X}_i$ and $\hat{X}_i$ are the hyperspectral in situ Rrs-, and the RRs- of the different water types, respectively. The class with the lowest relative RMSE values is given.

For 6classes the match between in situ data and Type 3 is best, X^2 and (relative) RMSE both give lowest values for this class.

For 5classes Type 2 instead of Type 3 had the lowest RMSE (see previous page Five Classes). Type 1 had a relative RMSE of (55.1%) and 3 (63.1%). Type 4 and 5 were far off.

Normalised 6 classes at MERIS bands
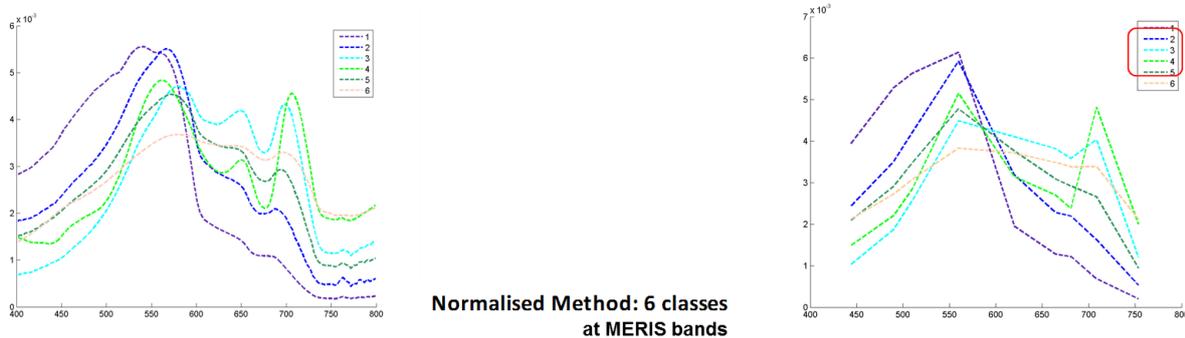
Normalised Method: 6 classes
at MERIS bands

Source	1	2	3	4	5	6	Total
Finnish	0	11	0	1	4	0	16
Taihu	0	0	8	79	22	125	234
Peipsi	0	0	0	9	25	0	34
Markermeer99	0	0	0	1	1	0	2
IJsselmeer99	0	0	0	4	1	0	5
IJsselmeer	0	6	0	8	39	0	53
Markermeer	0	8	0	29	34	0	71
ITA ASD	79	9	0	1	1	0	90
ITA WSD	11	1	0	0	0	0	12
ITA PR 650	1	2	0	0	0	0	3
Betuwe	0	8	0	6	2	0	16
NH Lakes	27	60	26	32	33	1	179
Spanish lakes	25	52	1	37	25	0	140
Lake Erie	2	7	0	3	4	0	16
Totals	145	164	35	210	191	126	871

To check if scattering and absorption have a clear impact, visual comparisons with verified TSM results from a previous study (next Figure), as a CHL indicator were enabled.

Observations of wind and TSM (HYDROPT on MEGS7.4)   Modelling resuspension at bed

Eleveld et al., 2014. WRR 48.
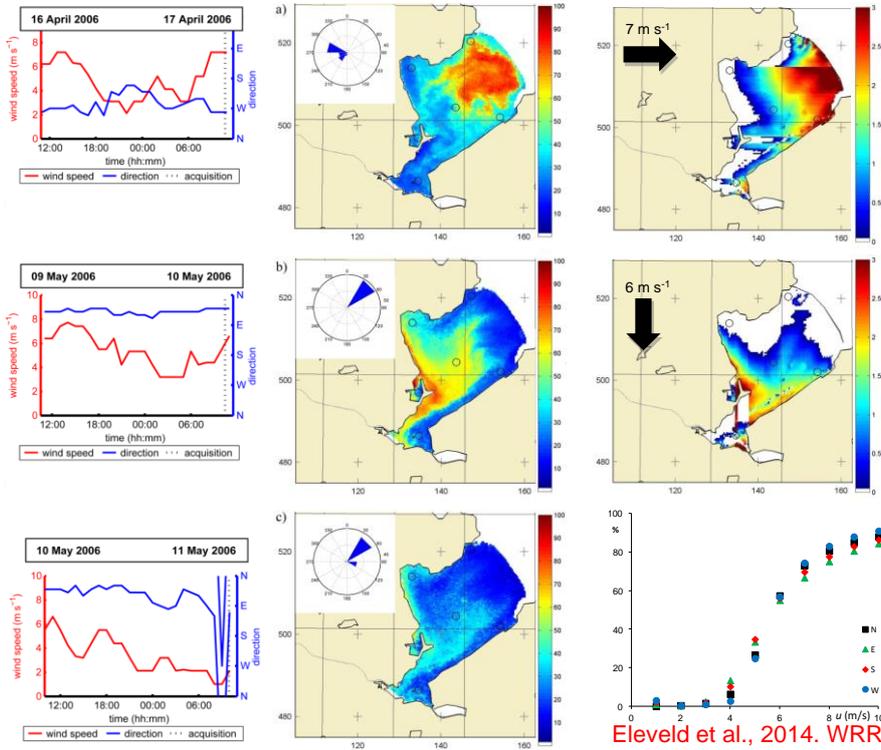
17 April 2006

MEGS8.1
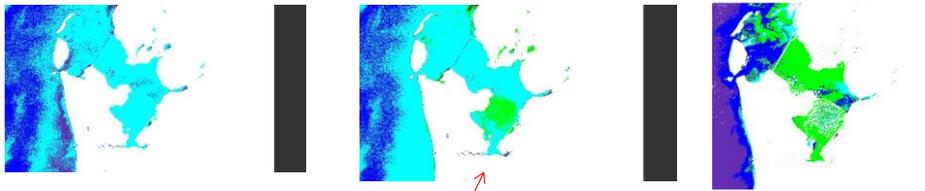
CCL2R
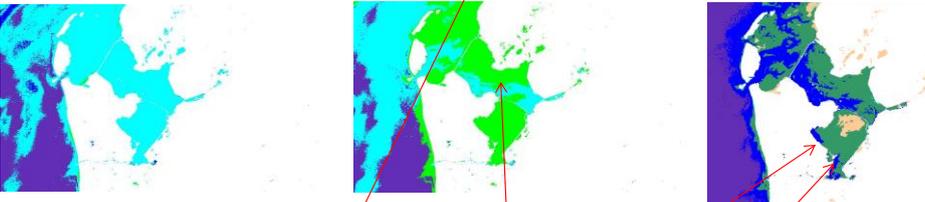
C2R Standard settings
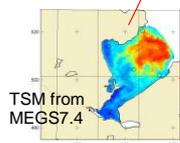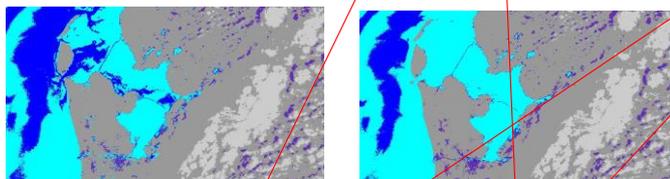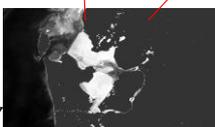
TSM from MEGS7.4

CHL index CCL2R band9/band7

Label	Colour
class_1	98, 46, 179
class_2	0, 0, 255
class_3	0, 255, 255
class_4	0, 255, 0
class_5	51, 153, 102
class_6	255, 204, 153

10 May 2006

11 May 2006

28 Sept. 2011

Huge difference between MM and IJM
Clear patterns in IJM

CHL index
CCL2R
band9/band7

Label	Colour
class_1	98, 46, 179
class_2	0, 0, 255
class_3	0, 255, 255
class_4	0, 255, 0
class_5	51, 153, 102
class_6	255, 204, 153
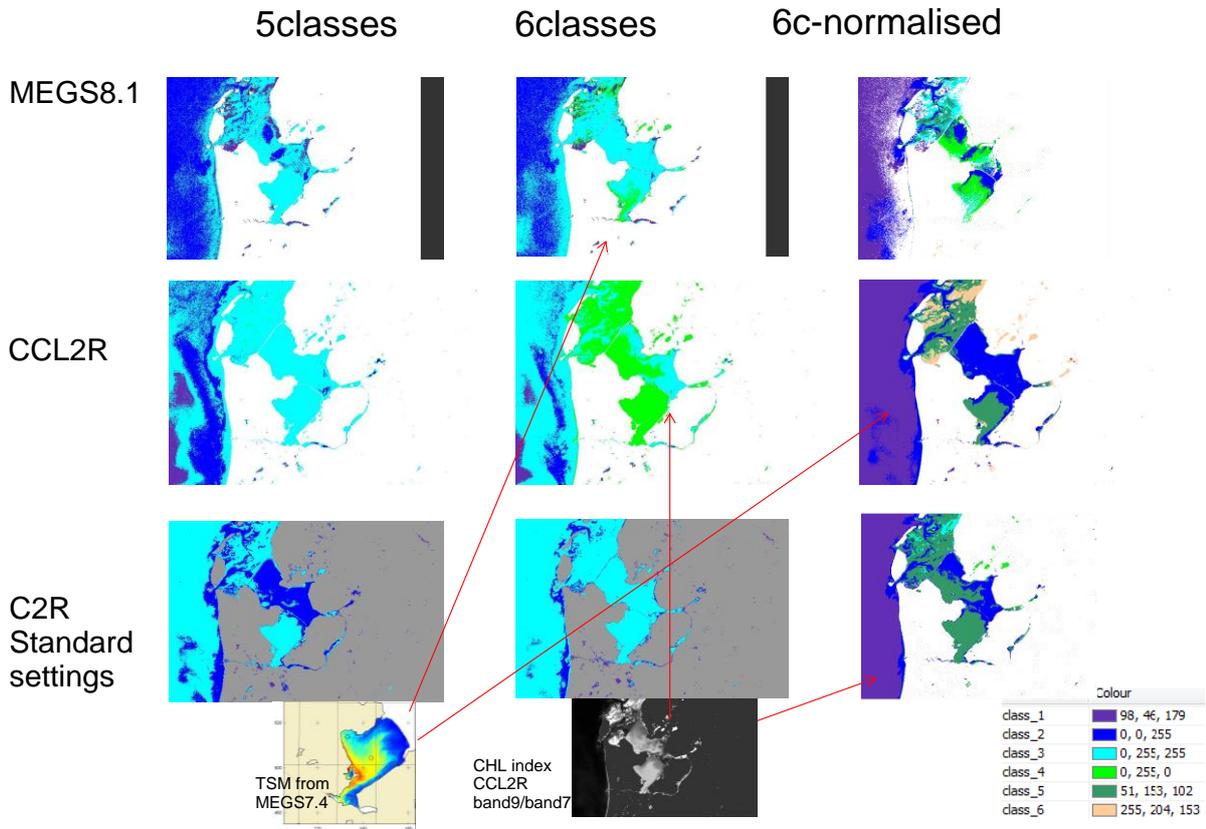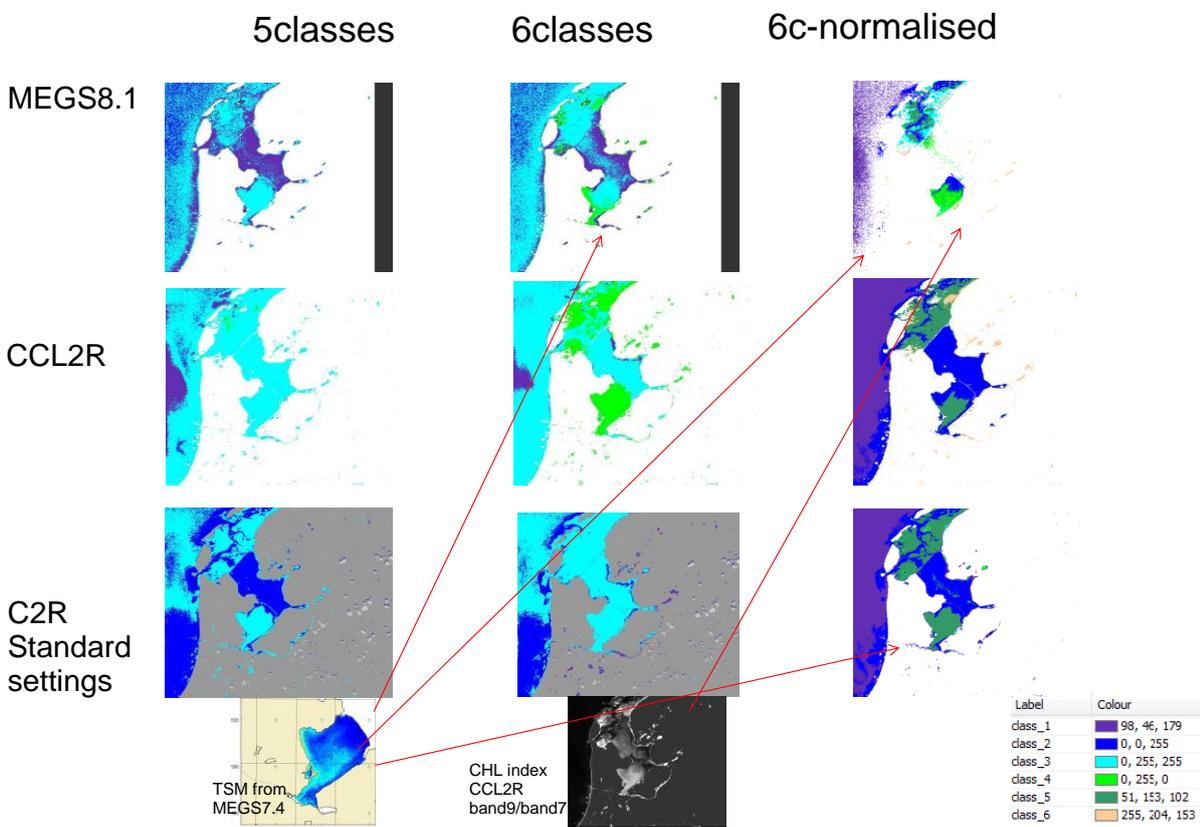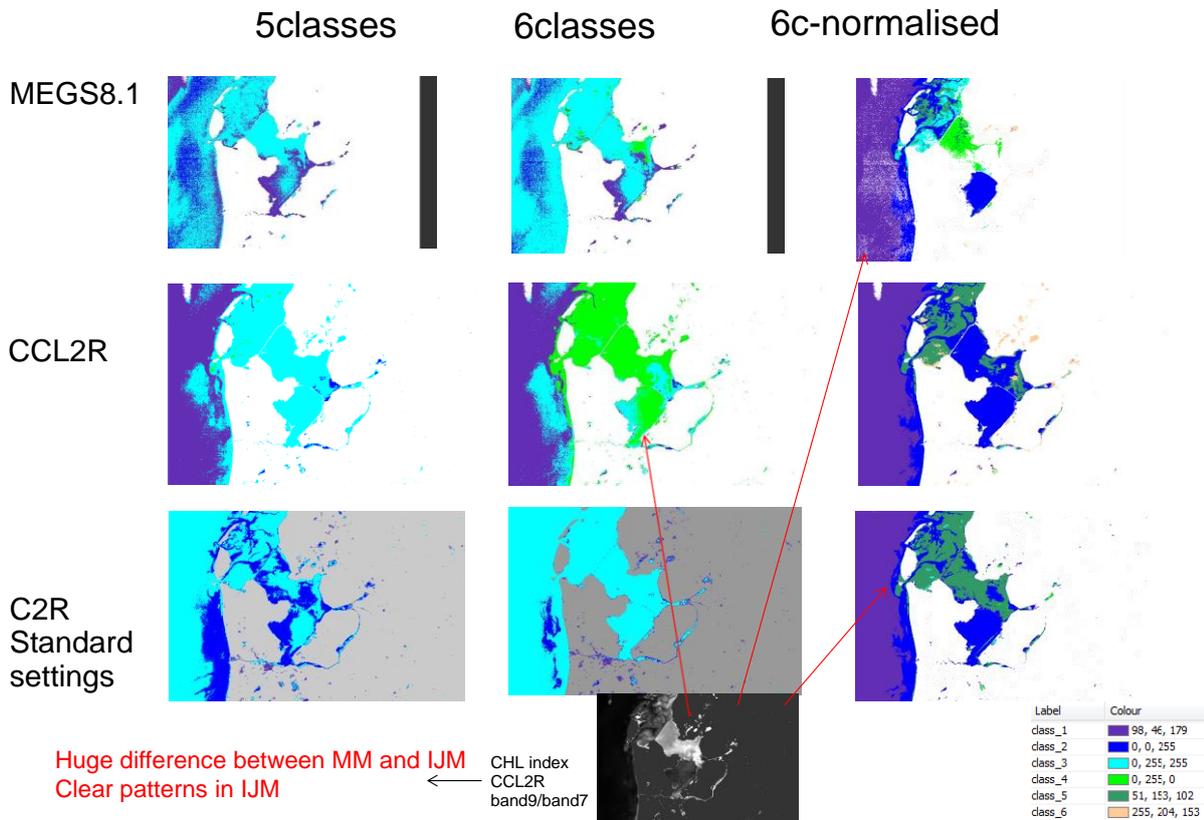
Maps resulting from the 5-classes scheme show predominance of classes 2 (highly absorptive) and 3 for different degrees of turbidity and trophic conditions. For the 6-classes scheme this shift to classes 3 and 4 (which feature spectral characteristics that can be associated with both absorption and scattering). This seems more reasonable for these environments. Normalising for spectral intensity causes Markermeer and IJseelmeer spectra mainly to end up in classes 2 to 5 for Class 6-normalised classification.

Concerning the atmospheric corrections, the Standard 3rd reprocessing (MEGS8.1 L2 products) seems to perform well, but this is not the case when negative reflectances would be masked. This results in many invalid pixels. Therefore it seems that processing L1b with the CoastColour atmospheric correction (which generated CC2R2L products), and an OWT classification with 6 classes give best results. However, there are of many more atmospheric processors available (GlaSS Task 3.2 report). As an example we include preliminary result from MIP (image from the same day, 06'31" time difference in file name). Note that all these processors were configured with standard atmospheric (and in-water) sIOPs which impact the result. They can produce better results under optimal configuration.

28 Sept. 2011

For the MERIS bands (N=9, band 2 ..10), RRs- extracted from remote sensing data $\dot{X}_i$ was compared with RRs- of the different water types $\hat{X}_i$. The class with the lowest relative RMSE values (Odermatt et al., 2010) is not always the same as the dominant class assigned by the OWT processors.
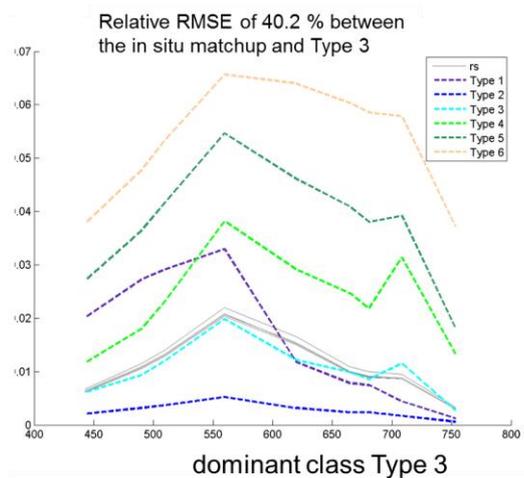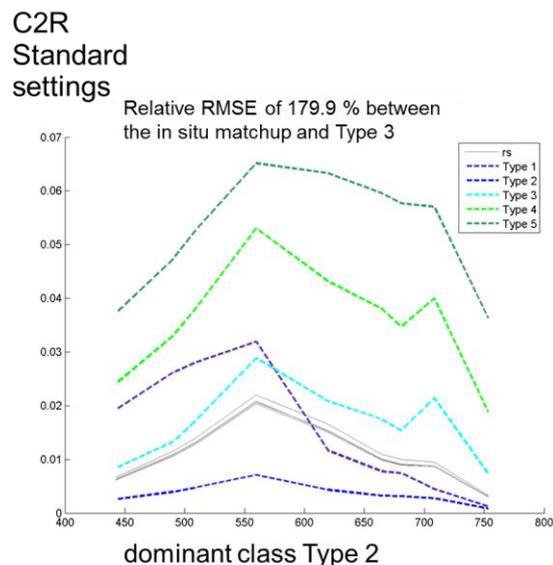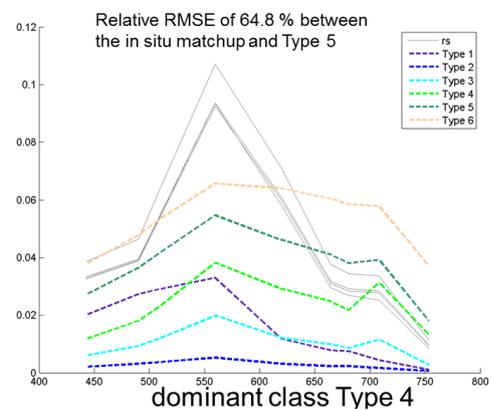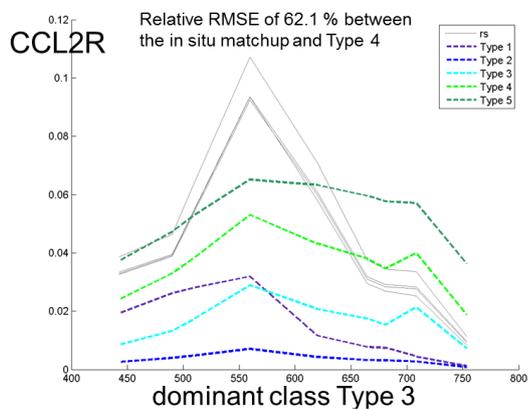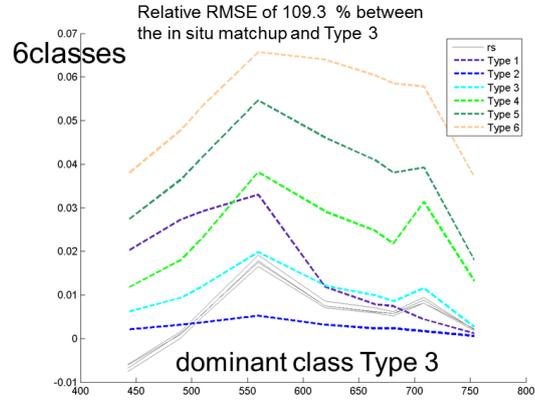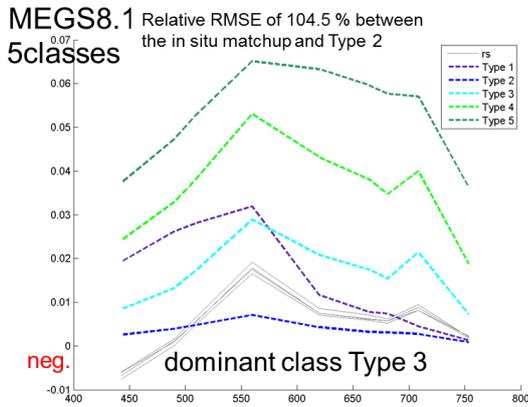
These results offer possibilities for further methodological studies, which go beyond what was anticipated for Task 3.2:

For all images, the normalised approach gives different results for the standard atmospheric correction (MEGS 8.1) than for the other neural nets (Coast Colour, C2R). Is it perhaps related to the many negative reflectances encountered in MEGS 8.1?

The X^2 GoF and the (relative) RMSE give different results: based on a (rel). RMSE different classes could be assigned to be dominant. However, the equations for both statistical measures

are quite different (see statistics handbooks).

Also, for the rel. RMSE, MERIS reflectances (Rrs-) extracted from the remote sensing data were compared to class means at the matching centre 3 nm wavelength. Hence, both the spectral response function, and the spread between the classes were not taken into account in the latter simplified approach.

However, all plots clearly do show that there can be substantial differences between the measured in situ or EO remote sensing reflectances, and the (dominant) OWT reflectances. These are wavelength dependant.

In conclusion:
- The results support the 6 and 6n classifications, five classes seem less suitable.
- Linking the 6 classes classifier to TSM scattering and 6n to mapping CHL absorption did not lead to conclusive results
- The atmospheric corrections have a huge impact remote sensing reflectances, and consequently they also impact water type maps.
- Substantial wavelength dependent differences between the measured in situ or EO remote sensing reflectances, and the (dominant) OWT reflectances were found.

Protocols used to create the maps

For testing the following steps were followed:
1) Make a subset (area covering the lakes) using Processing Geometric operations Spatial Subset from View, save Product as dimapfile just to make processing faster
2) Use several atmospheric corrections such as MEGS 8.1, CCL2R (CC2R in D 3.2), C2R (standard settings) and MIP, and run the OWT tools, 5, 6, 6-normalised.
3) In displaying the dominant class flag out land and clouds with the Manage Binary data Masks button

Detailed protocol for developing the Dutch test case

1) Open .N1 file either:
1a) MER_FSG_2PPEPA*, 1b) MER_FSG_1PPEPA*

2) click Band "some band"
Processing > Geometric Operations > Subset from view > Spatial Subset > Geo Coordinates
North latitude bound 53.50, West longitude bound 04.50, South latitude bound 51.80, East longitude bound 06.50 (this is somewhat larger than the region defined in D5.1, just to ensure that we capture lakes IJsselmeer and Markermeer.) Save product as subset_0_of_MER_FSG*

3) For different atmospheric corrections, either:
      3a) MEGS8.1 based on subset_0_of_MER_FSG_2PPEPA* Processing > Thematic Water
      Processing > MERIS OWT Classification > I/O Parameters
      Source subset_0_of_MER_FSG_2PPEPA Target
      subset_0_of_MER_FSG_2PPEPA_*_owt5, 6, 6n
      Processing Parameters GLASS_5c Reflectances prefix: reflec Write input reflectances.
      Input reflectance is Irradiance_Relectances (ESA Compatible)
      4) For diferent OWTs:
      4a) 5 classes
      4b) 6 classes
      4c) 6 classes-normalised

5) Click bands > class > dominant class, Colour manipulation > Import colourscale 0wt_5, 6_classes.cpd. Automatically distribute between max/min No. Manage binary datamasks land and clouds white. (NB note that there are Invalid reflectances in this product, but visualizing them by masking 80% gray impacts too many pixels. This was previously also the case with the previous, 2nd reprocessing). One might navigate and zoom, but the view is normalised by Export view as an image Full scene (as .jpg and tiff)

3b) CCL2R based on subset_0_of_MER_FSG_2PPEPA* Processing > Thematic Water Processing > MERIS CoastColour Processing > CoastColour Atmospheric Correction > Source subset_0_of_MER_FSG_1PPEPA Target subset_0_of_MER_FSG_1PPEPA_*_L2R On Processing Parameters Perform Smile-effect correction (others have been done in previous processing steps) Average salinity 0, Average temp a reasonable values Use NN, land mask & cloud/ice detection expression (use default) Write all reflectances as Irradiance_Relectances (ESA Compatible)

Subsequently as under 3a

(Invalid reflectances l2r_cc_input_invalid &L1b_invalid can be set to gray 80%, however to normalise between processing chains standard settings were retained)

3c) C2R

Source   subset_0_of_MER_FSG_1PPEPA   Target subset_0_of_MER_FSG_1PPEPA_*_C2IOP

Perform atmospheric correction, Smile correction Output water leaving reflectance as Irradiance reflectances (ESA compatible)

Use NN, land mask & cloud/ice detection expression (use default). Standard settings also comprise TSM conversion exponent 1, factor 1.73, CHL conversion exponent 1.04, factor 21.0 Spectrum out of scope threshold 4.0 Invalid pixel acg_flags.INVALID?

The atmospheric and in water NN are coupled. Which values should we really take. Link to Task 3.1.

Manage binary datamasks land clouds white, reduce Transparency to zero. Invalid (agc_invalid & invalid can be set to gray 80%, but standard settings were ratianed).

3d) MIP

Open in BEAM and save it as BEAM-DIMAP product. For the first ten bands, us band ratios to calculate Rrs+ (MIP produces Rrs-), using:

$$Rrs\,(0,+) = 0.52*Rrs(0,-)/(1-\ 1.7*Rrs(0,-)\,)$$

which is the reworked standard NASA algorithm for conversions between Rrs+ and Rrs-

$Rrs\,(0,-)= Rrs(0,+)/(0.52+1.7*Rrs(0,+))$ (in sr^{-1})   (Lee et al.,2002)

Next, open de .dim file in a text editor and replace the <Spectral_Band_Info> of the newly created bands with:

    <BAND_INDEX>15</BAND_INDEX>
    <BAND_DESCRIPTION>Water   leaving   radiance   reflectance   at   412.691 nm</BAND_DESCRIPTION>
    <BAND_NAME>reflec_1</BAND_NAME>
    <DATA_TYPE>float32</DATA_TYPE>
    <PHYSICAL_UNIT>sr^-1</PHYSICAL_UNIT>
    <SOLAR_FLUX>1668.2919</SOLAR_FLUX>
    <SPECTRAL_BAND_INDEX>0</SPECTRAL_BAND_INDEX>
    <BAND_WAVELEN>412.691</BAND_WAVELEN>
    <BANDWIDTH>9.937</BANDWIDTH>
    <SCALING_FACTOR>1.0</SCALING_FACTOR>
    <SCALING_OFFSET>0.0</SCALING_OFFSET>
    <LOG10_SCALED>false</LOG10_SCALED>

```
<NO_DATA_VALUE_USED>false</NO_DATA_VALUE_USED>
<NO_DATA_VALUE>0.0</NO_DATA_VALUE>
```
Do this for the ten newly createdbands. Do not forget to update the band index number.
Remove the lines `<VALID_MASK_TERM>!agc_flags.INVALID</VALID_MASK_TERM>` for each band

Appendix2: Comparison of the INLAND 7C OWT classification (Diversity) with the GLASS 5C classification

Ana Ruescas and Daniel Odermatt

Introduction

A comparison of two different classification schemas was done as a contribution of the Diversity project in Glass. The comparison consisted on using two different classification schemas and observe how the water in the lakes are assigned in each one: the seven classes distribution of the INLAND classifier used in Diversity, and the GLASS classifier with 5 classes.

The INLAND OWT classifier used in Diversity consists of seven classes whose mean and covariance where calculated using a set of in situ bio-optical data that combines coastal and inland waters. The three main sources are a dataset collected by the University of New Hampshire (UNH) in various north-eastern US lakes as well as the Great Salt lake in Utah (Bradt, 2012); a dataset from Spain covering many lakes and trophic conditions (Ruiz –Verdú et al., 2008); and a dataset obtained from NASA's SeaBASS archive primarily from US coastal marine sites (Werdell et al., 2003). All reflectance data were collected with hyperspectral instruments, which were binned at 3 nm intervals from 400 to 800 nm. The hyperspectral resolution capture spectral features throughout the visible spectrum and the NIR and provides flexibility in adapting the derive OWT spectral reflectance characteristics to many sensors. The total number of reflectance data with co-measured chlorophyll-a data was 488 points. The dataset ideally should include a complete representation of all possible water types. The statistical properties for each class or cluster found in the dataset, becomes the basis for defining membership to each class. In Figure 1 the histogram of the combine in situ chlorophyll-a values is shown.
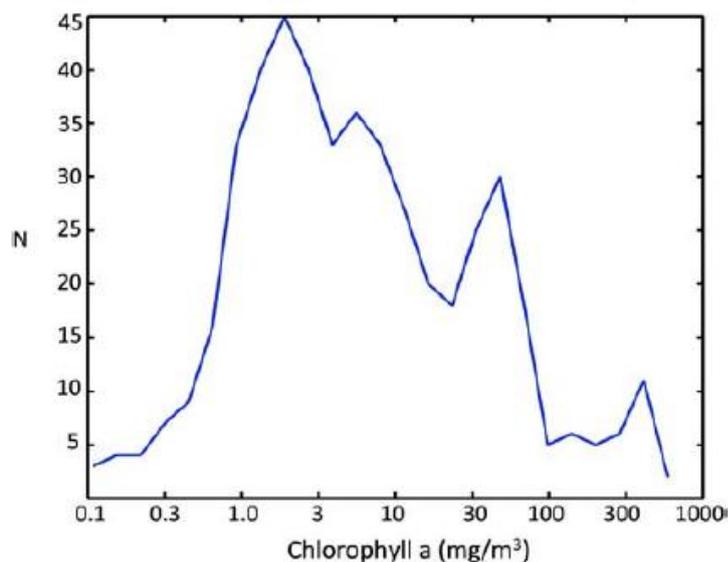
Figure 3 Histogram of the combined in situ chlorophyll-a dataset

The data cluster analysis is the fuzzy c-mean (FCM) algorithm (Bezdek, 1981) applied to the in situ reflectance data. The FCM algorithm produces a fuzzy clustering of the data into a specified number of clusters. The basic function of this algorithm is to choose clusters that minimize the distance between the data points and the prototype cluster centers (means). Cluster centres are iteratively adjusted until optimization criteria are met. The clustering routine returns the mean reflectance vectors for the cluster classes and a matrix containing the memberships of each point to each class. The best number of clusters for the datasets used here is seven. Once clusters have been identified, the mean and covariance matrix is calculated for each one. These statistics define the optical water types, and are subsequently used in the membership function. The cluster

analysis separated and differentiated sunsets based on both the shape and magnitude of the Rrs. The results are shown in Figure 2.
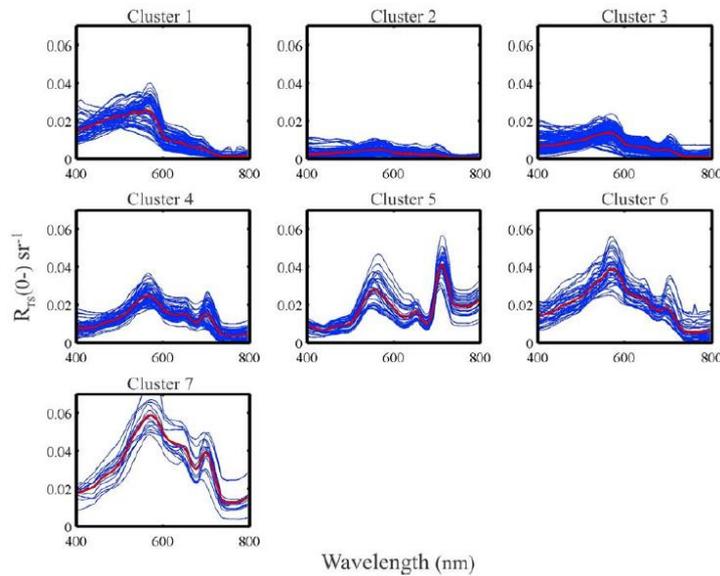
Figure 4 Reflectance data sorted into seven clusters from the fuzzy c-mean cluster analysis; blue lines: individual station reflectance data; red lines: mean reflectance (from Moore et al. 2014)

The difference between the clusters can be easily seen when their reflectance means are plotted together -see Figure 3-. They are representations of averaged conditions governed by the optical properties of the water column and depend on the absorption and scattering properties (IOPs) of the in-water constituents.
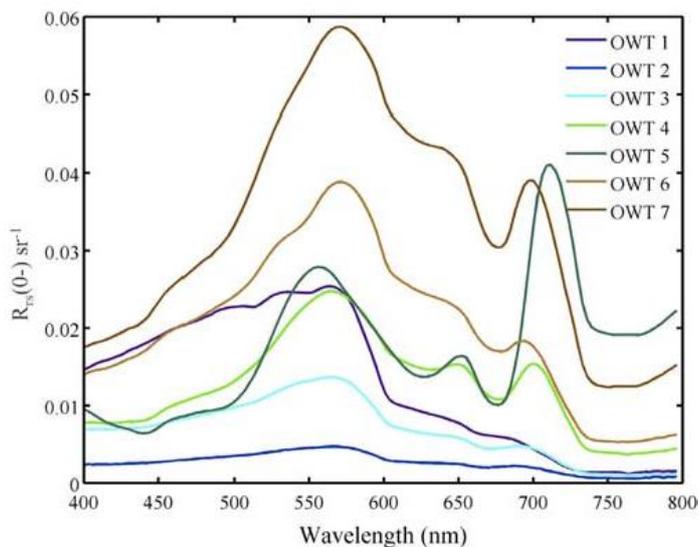
Figure 5 The reflectance means of the seven OWT in the INLAND classifier

The OWTs are organized based on spectral features and ascending chlorophyll-a concentration. The OWTs show a pattern of increasing absorption in the blue/green for low red/NIR features (OWTs 1 through 3, see Figure 2), followed by increasing peak magnitude at 555 nm (types 4 to 7). OWTs 1 through 5 show increasing chlorophyll-a concentrations, while 6 and 7 have lower chlorophyll-a values. OWTS 1-3 have low overall spectral magnitude and show relatively flat

features from 600 nm onward compared to the others OWTs. From 4 to 7 show higher overall magnitudes and more features especially in the red/NIR region. These OWTs show peaks around 700 nm, but different from each other in order of magnitude. This peak is characteristic of strong particle backscattering and has been associated with high algal particle concentration. All OWTs show a reflectance peak to some degree at or near 555 nm, which can be attributed to enhanced particle scattering from living and non-living sources. Other secondary peaks are seeing at or near 650 nm in these OWTs. Values from the mean and median of the chlorophyll-a concentration increased from OWT1 to OWT 5. OWT 6 and 7 have lower chlorophyll-a values than OWT 5, but comparable to OWTs 3 and 4.

The GLASS 5c classification has been profusely explained in the deliverable (D3.3 Optical pre-classification method). The in situ reflectance used to train the fuzzy classifier come exclusively from lake waters, marking in this way the difference with the INLAND classifier that included coastal waters reflectance too. The equivalent plot to the INLAND classification of Figure 3 is shown in Figure 4 for the GLASS 5c classification.
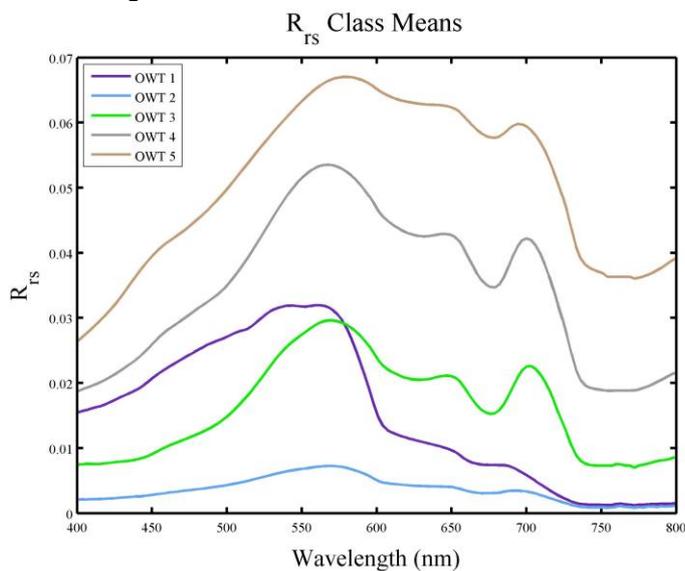
Figure 6 The reflectance means of the five OWT in the GLASS classifier

One of the clearer water classes in the INLAND classification has disappeared in the GLASS classification (INLAND class 3), and class 5 in INLAND does not exist in GLASS. The classes that represent the turbid waters (INLAND class 6 and 7) have a slightly different shape in the GLASS class 4 and 5, with a less accentuated peak in the 550-600 nm area and a smoother slope until the second peak at around 700 nm. This probably indicates a higher mixed in the water composition of the lakes compare with the more distinguishable chlorophyll/sediment waters on coastal areas. The in situ data set also counted with some high CDOM samples from Finnish lakes, but probably due to its small number compared with the other water types, the distinction of these high CDOM lakes from the clearer waters was not possible. For a closer look to the 5 classes distinguished.

In the following sections, the two classifiers were applied to the same image dataset for four lakes: Lake Balaton (Hungary), Lake Constance (Germany), Lake Peipsi (Estonia) and Lake Paijanne (Finland) have been classified with the INLAND and the GLASS classifiers. Data was atmospherically corrected with the CoastColour algorithm, and the results were classified by types of water with the two different schemas. Daily, monthly and yearly compositions (2008) are then analysed.

Lake Balaton

Ten different days (images) of the lake Balaton have been processed with the OWT tool, they are daily composites for different months in the year 2008 (see Figure 6).

Lake Balaton most common optical water types are 6 and 7 in the INLAND classification (turbid waters). The Easter basin is often classified as slightly clearer than the center of the lake. The west part is occasionally assigned to class 4. All the spatial pattern agree with the results from other investigations (Palmer et al, 2014). In the GLASS 5C classification, the eastern part remains more or less stable (always on lower classes), but the west and center parts are moved to central classes (not so turbid), and in general there is less variability.

Monthly averages confirm the turbidity of the lake in the INLAND classification (Figure 7) and the assignment of less turbid classes in the GLASS 5C classification (Figure 8). Yearly averages follow the same trend (Figure 9).
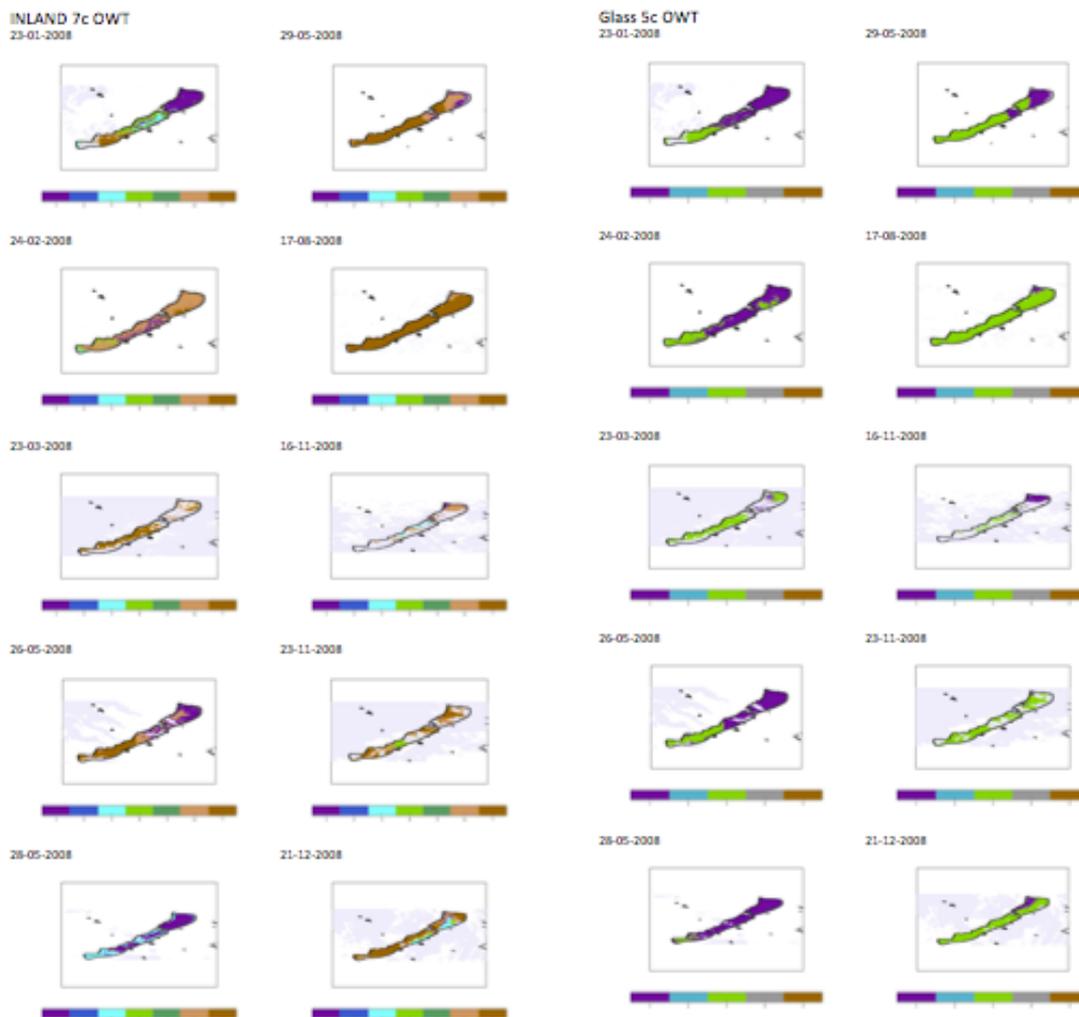
Figure 7. Lake Balaton daily examples of the INLAND (left) and GLASS (right) OWT classifications
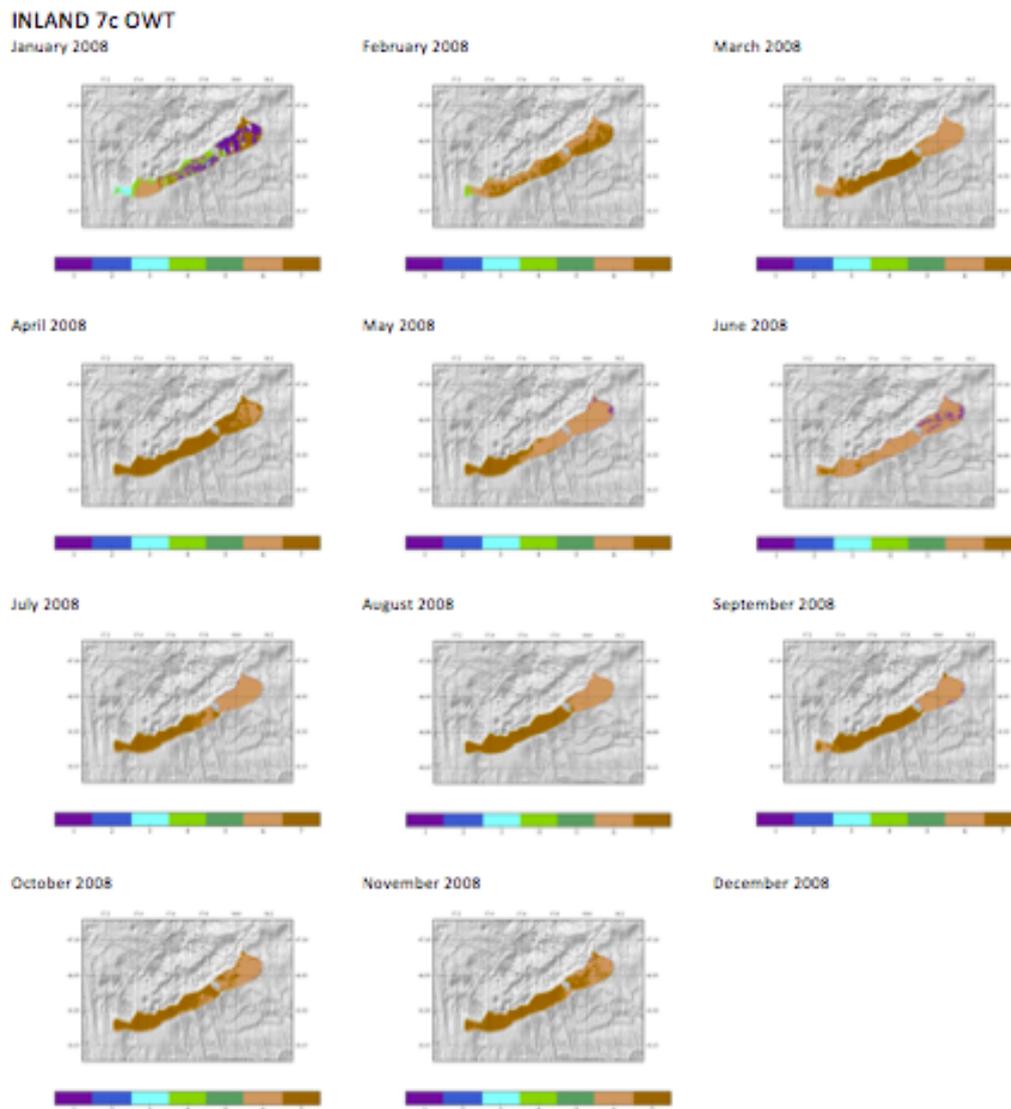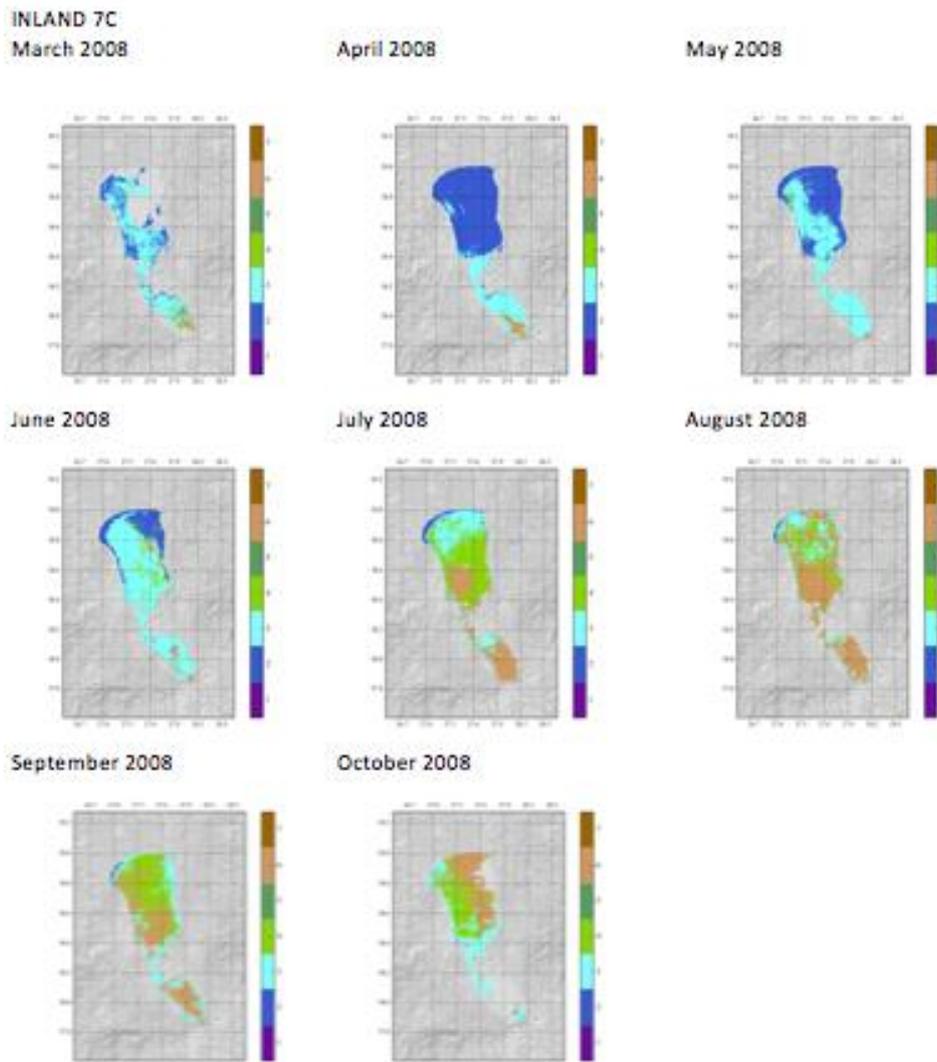
INLAND 7c OWT

Figure 8. Monthly averages of the dominant class, INLAND 7C
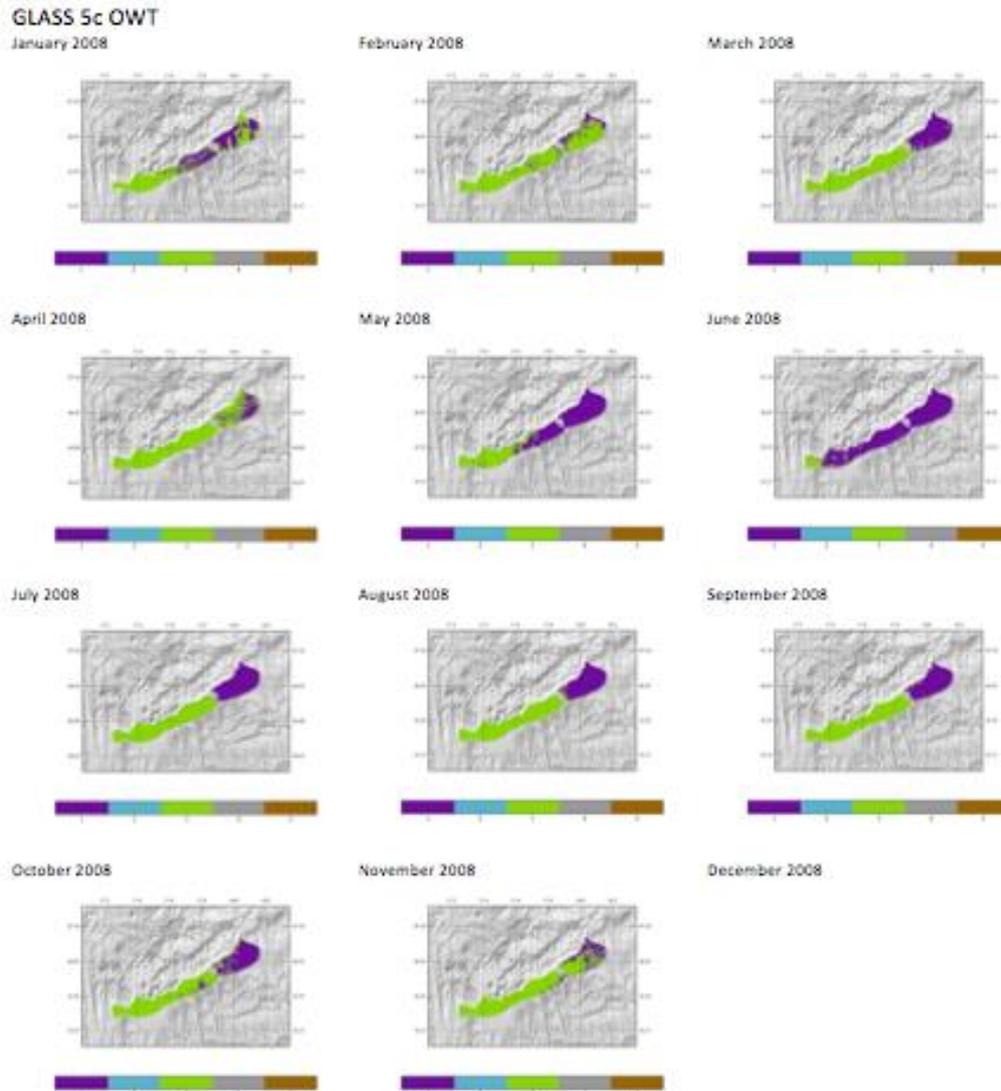
Figure 9. Monthly averages of the dominant class, GLASS 5C
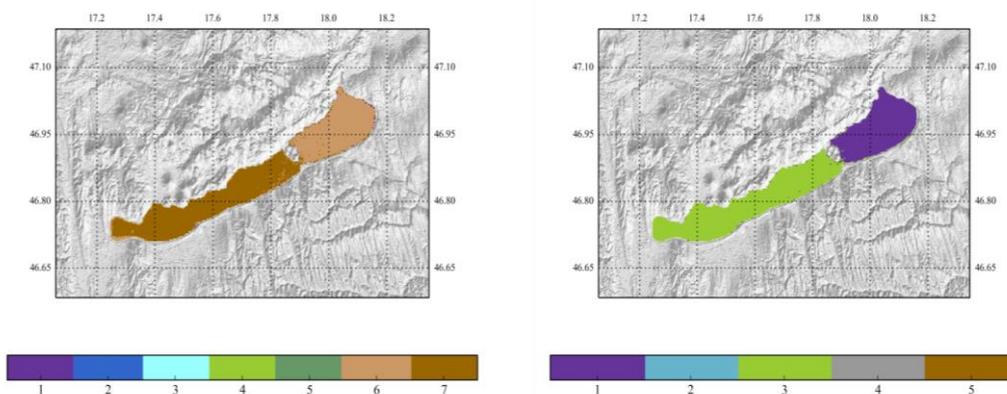
Figure 10. Yearly average (2008) of INLAND 7C and GLASS 5C dominant class

Lake Constance

For Lake Constace 15 daily images were processed and results for the two classifications are shown in Figure 10 and Figure 11. In the INLAND 7C, the lake is assigned to classes 2 or 3 most of the year, excluding the summer months, when the dominant class appears to be 1. That would mean that the status of the lake in in general of clear waters, with some chlorophyll concentrations during spring and late summer. This theory confirms quite well in the GLASS 5C classification, with a nice correspondence with the previous classification showing minimum classes during the summer months, increasing classes during spring and early autumn, but stressing the deep, clear water character of the lake throughout the year.

These observations are similar to the ones shown in Figure 12 and Figure 13. In the monthly INLAND classification the lake is assigned to classes 1 to maximum 3 (in the Autumn months), the maximum variability is given during the Spring. Within the GLASS 5c the classes drop at least one step, but the highest variability keeps showing during the Spring months.

The yearly averages show a bit more of patchiness on the INLAND 7C, assigning most of the lake to classes 2 and 3, while GLASS 5C assigns the lake to class 1. Both cases remark the clearness of the lake and the low variability in the quality of the water throughout the year

INLAND 7c OWT
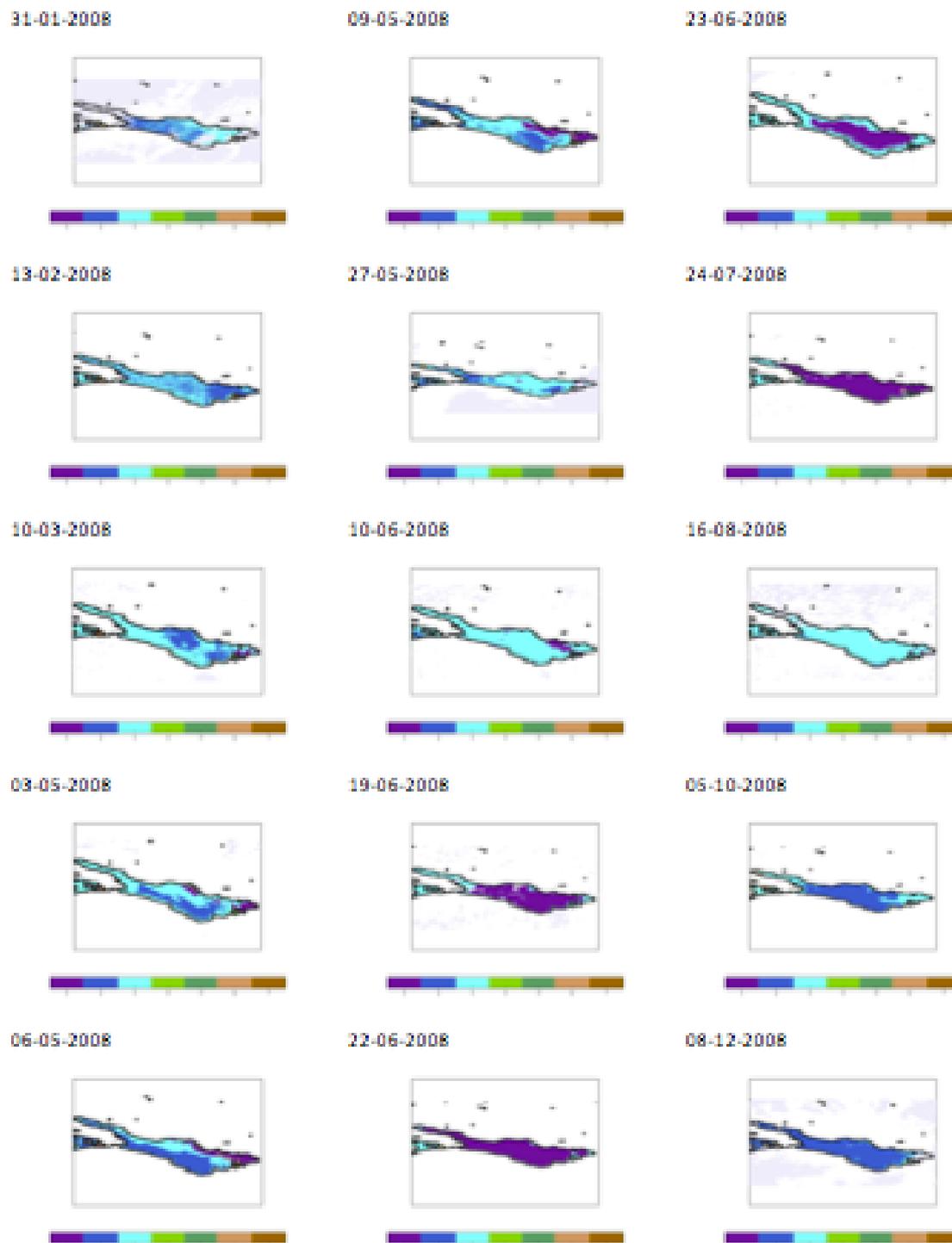
Figure 11 Lake Constance daily examples of the INLAND OWT classification
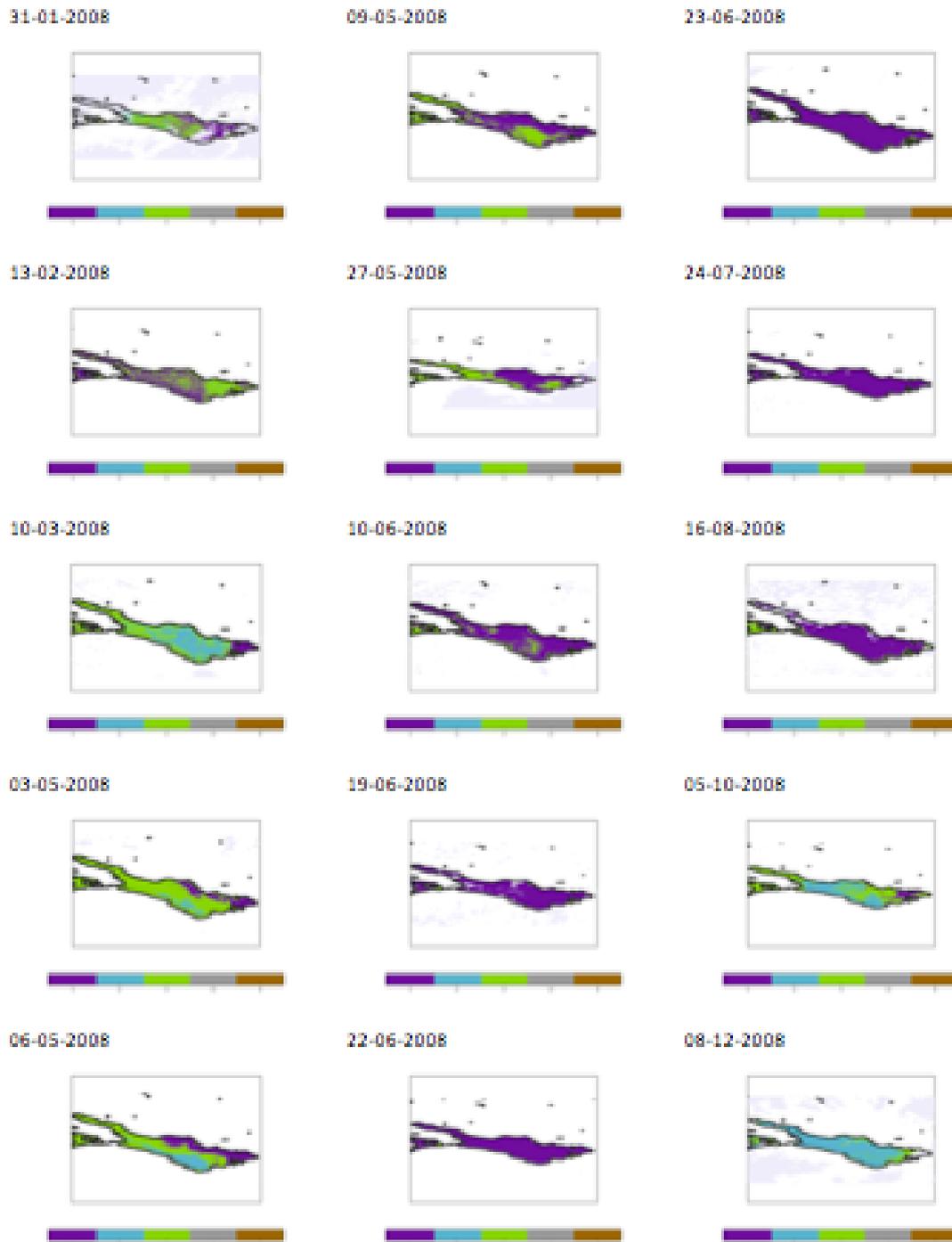
Figure 12 Lake Constance daily examples of the GLASS OWT classification

INLAND 7c OWT

January 2008
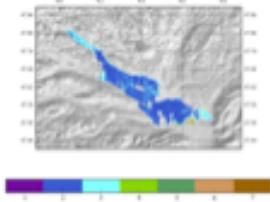
February 2008
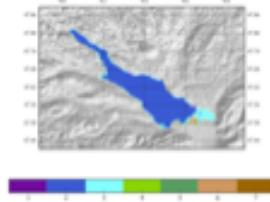
March 2008
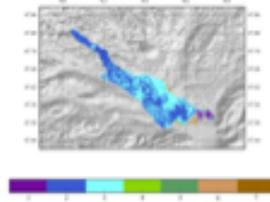
April 2008
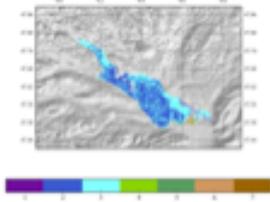
May 2008
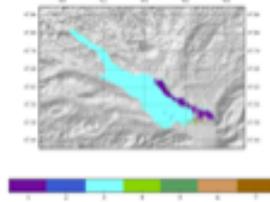
June 2008
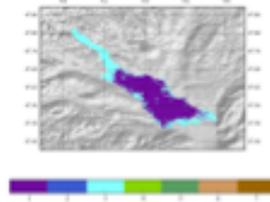
July 2008
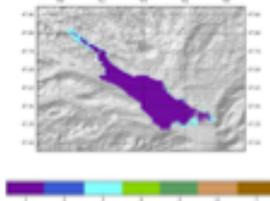
August 2008
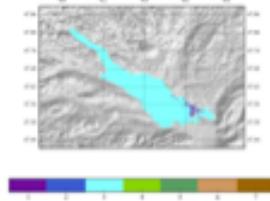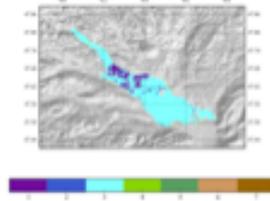
September 2008

October 2008
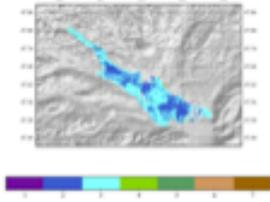
November 2008
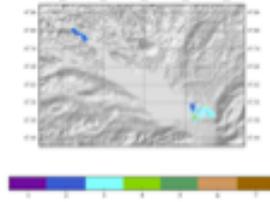
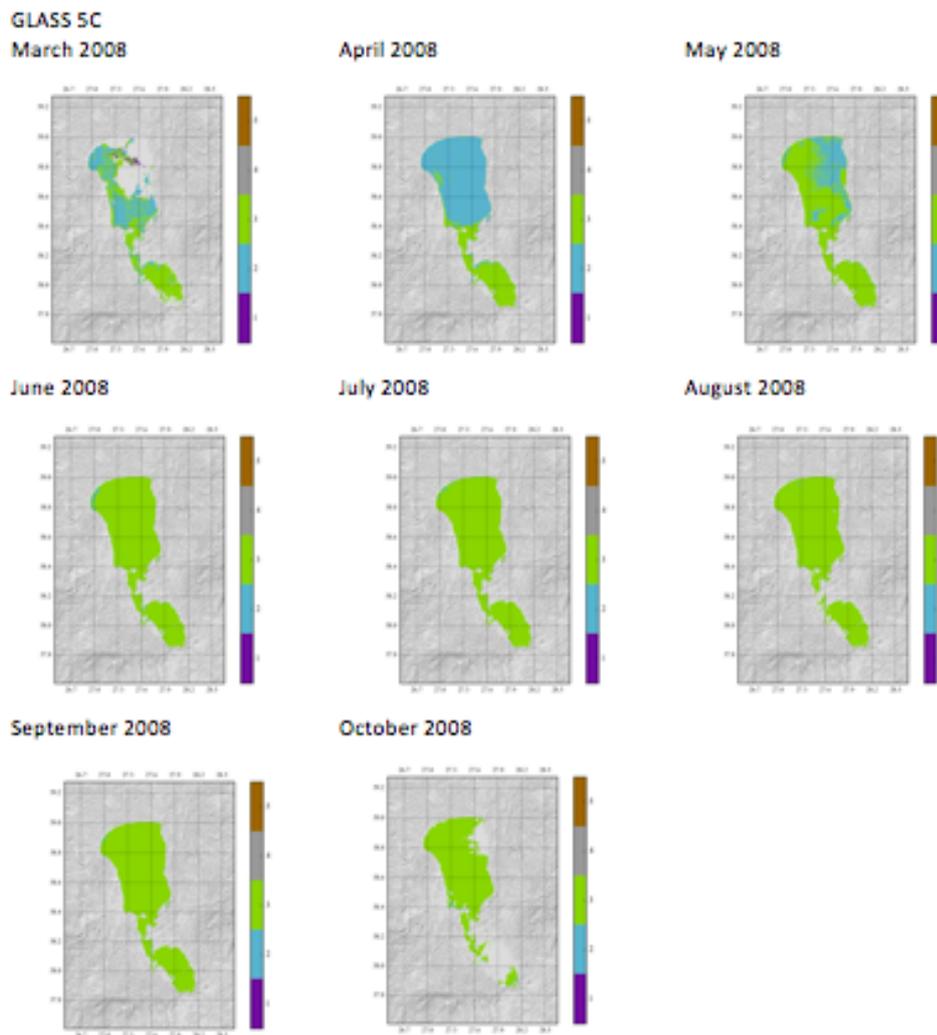*Figure 13 Monthly averages of the dominant class, INLAND 7C*

GLASS 5c OWT
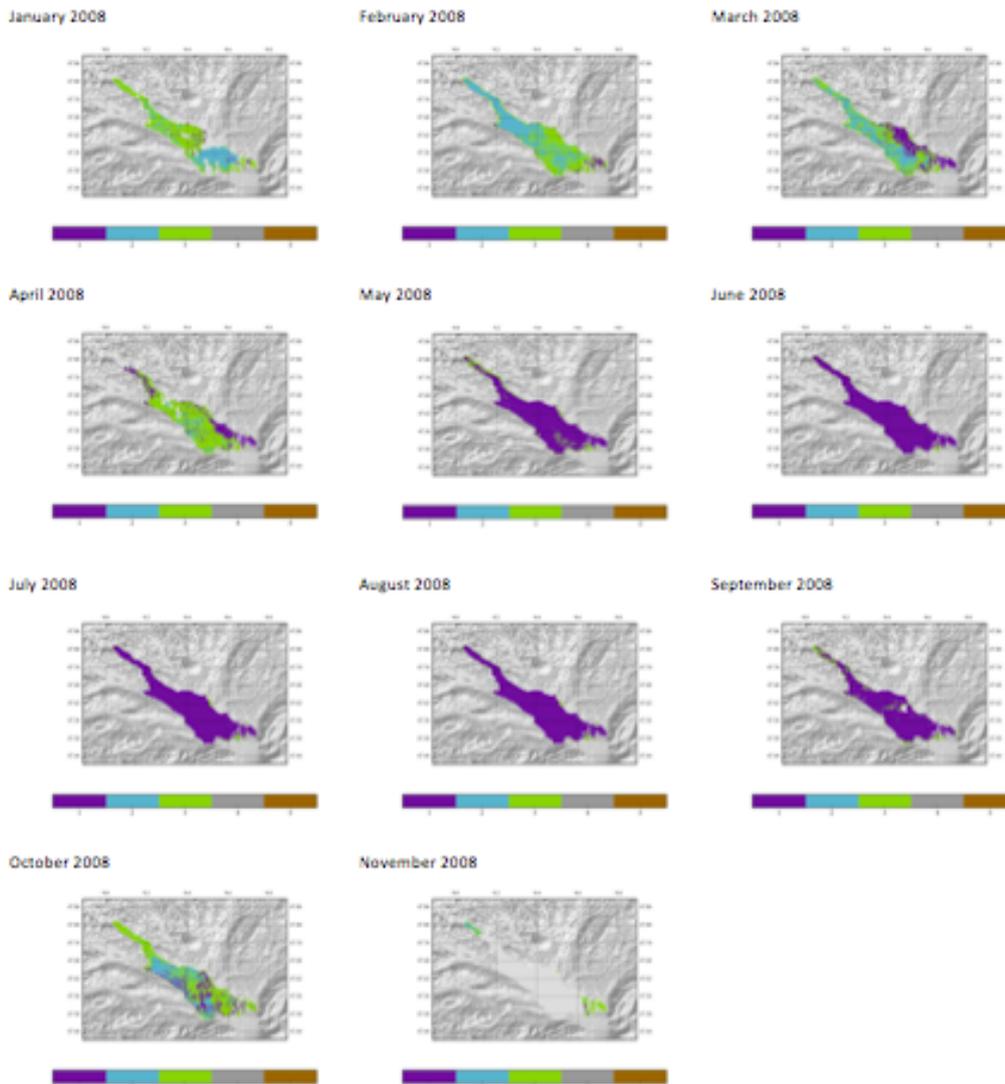
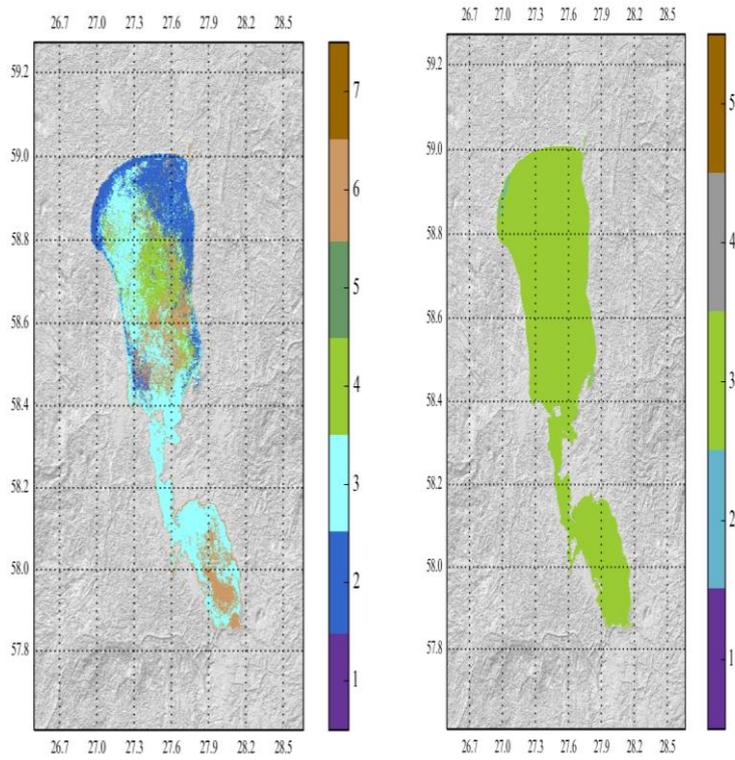Figure 14 Monthly averages of the dominant class, GLASS 5C
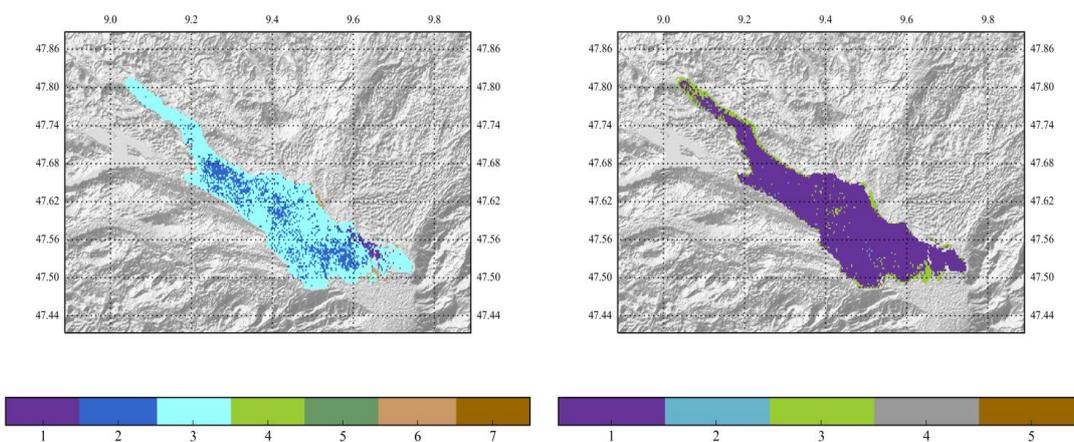

Figure 15 Yearly average (2008) of INLAND 7C and GLASS 5C dominant class

Comparison with in situ data

Two days of April 2007, the 13th and the 20th, had spectrum in situ data available, and it was used to compare with the shape spectra of the INLAND 7C and GLASS 5C classifications. In some cases, there is also the Coast Color AC derived spectra available and it was used to compare with the in situ data too.

The number beside the in situ or CC lines in the legend makes reference to the class_sum value. This should be expected to be high when the in situ/CC class is very similar to one of the spectrum classes. Unfortunately, this is not always true. Generally, a class_sum less than 1 or close to 0 is an indication of a poor match to the OWTs, but this could be for several reasons. One reason is that the type to which it belongs is not represented. In this case, there is a 'black hole' in the OWT system for this type. Another reason could be that there is enough noise in the satellite Rrs to impair membership. This is a significant problem especially for the red/NIR bands as they are the noisiest (along with bands at 412nm). This means that band choice could be a factor if it is known that the scene(s) have problems in one or more bands.

The INLAND 7C plots show that the in situ data is assigned to classes 2 or 3, as shown in the figures above (Figure 11-15) The CC results are closer to class 1. The GLASS 5C in situ data is more focused on class 2 and eventually 3; with CC results classified as class 1 or 2.

INLAND 7C	GLASS 5C
2007-04-13_08-10-00	
2007-04-13_09-05-00	
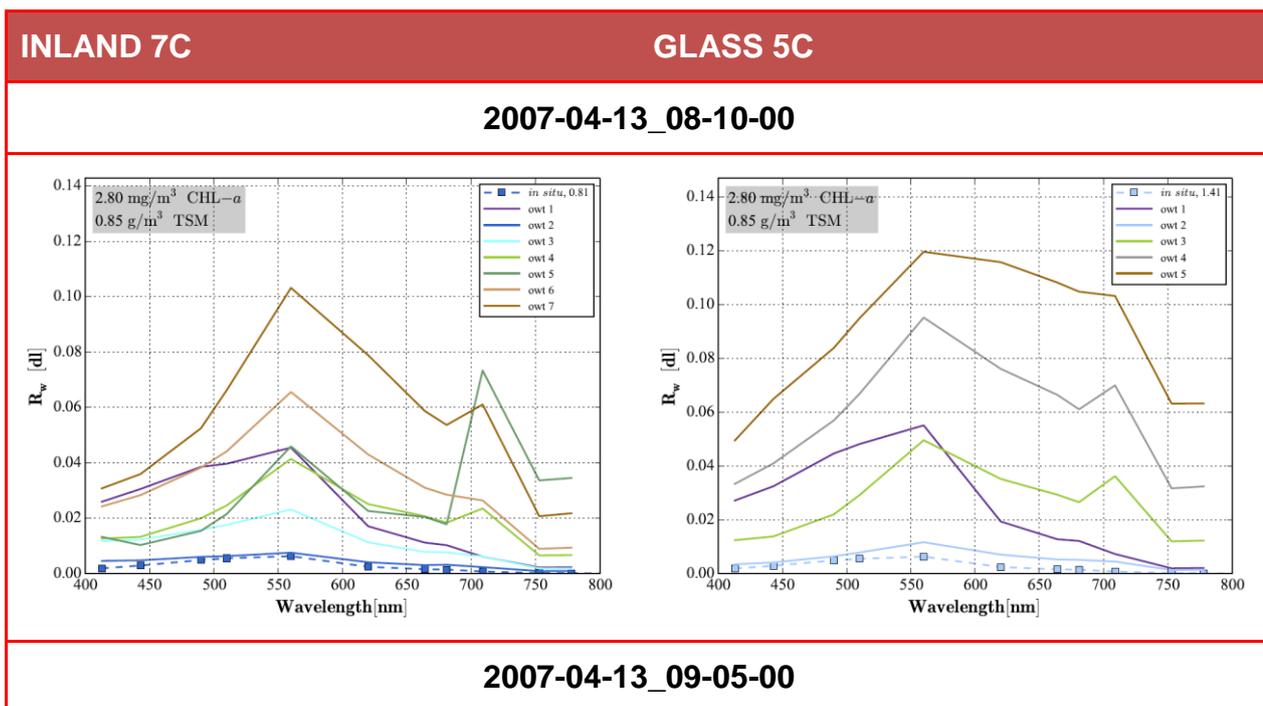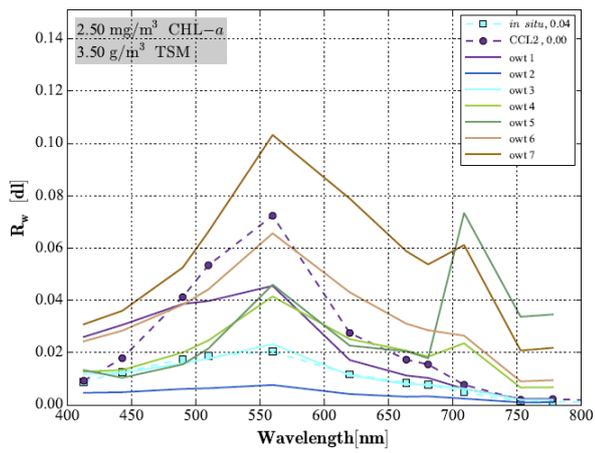
2007-04-13_09-45-00
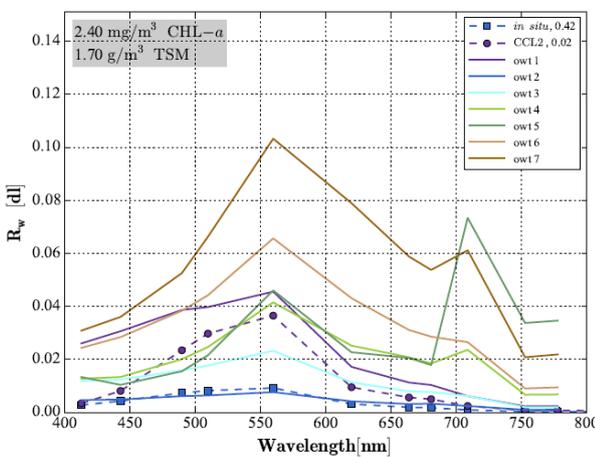
2007-04-13_10-40-00
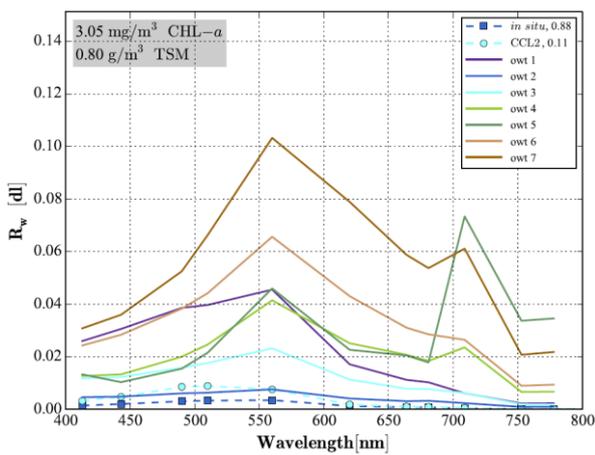
2007-04-20_08-20-00
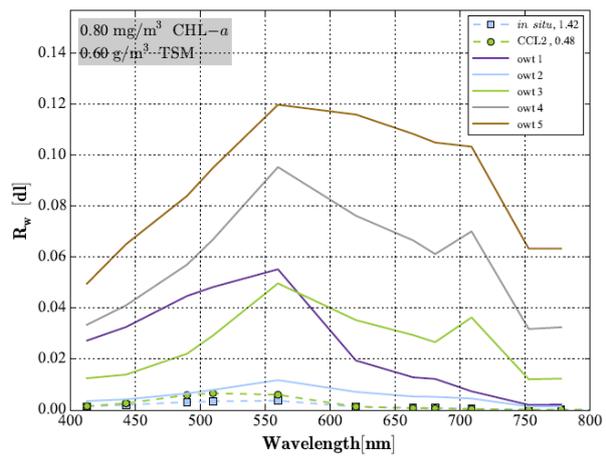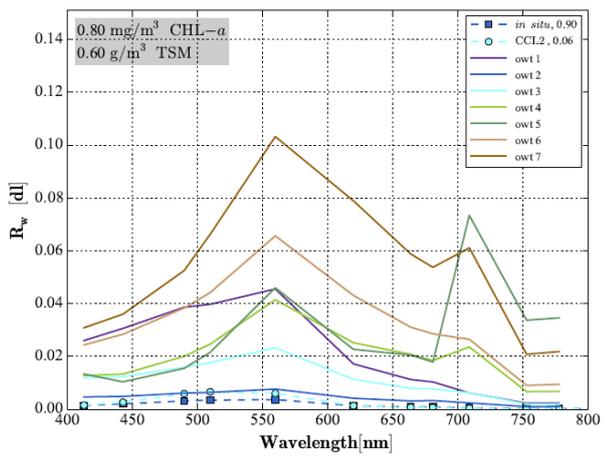
2007-04-20_09-25-00
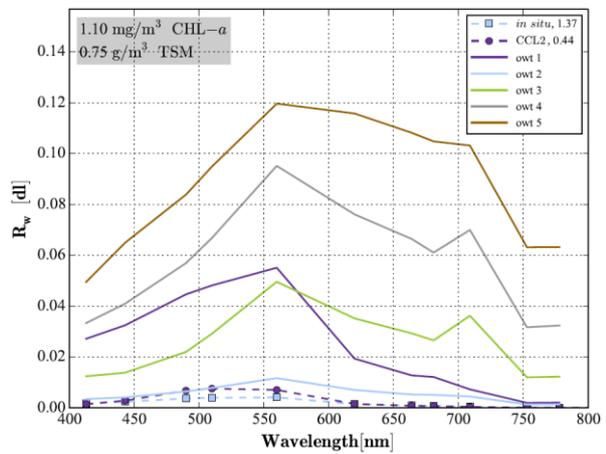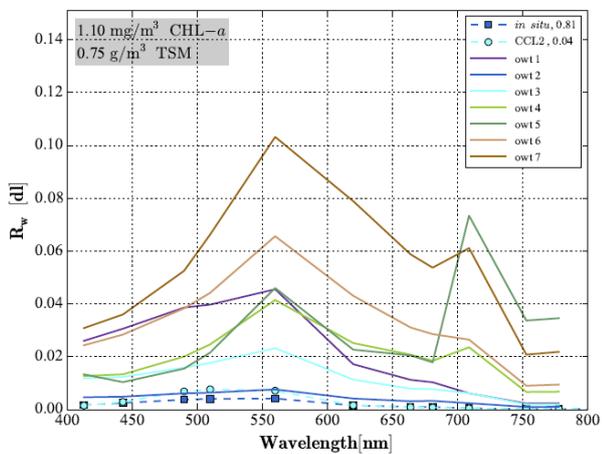
2007-04-20_10-20-00
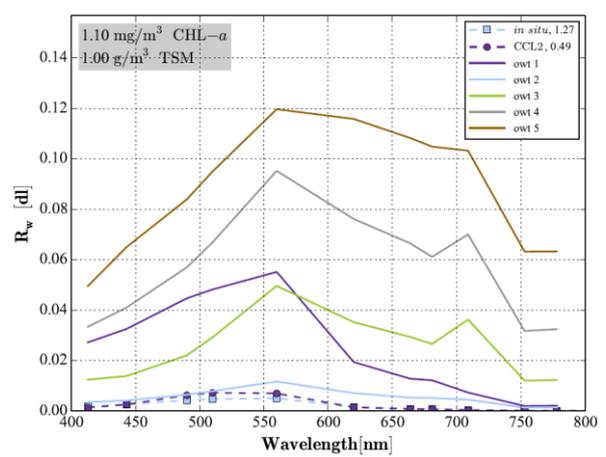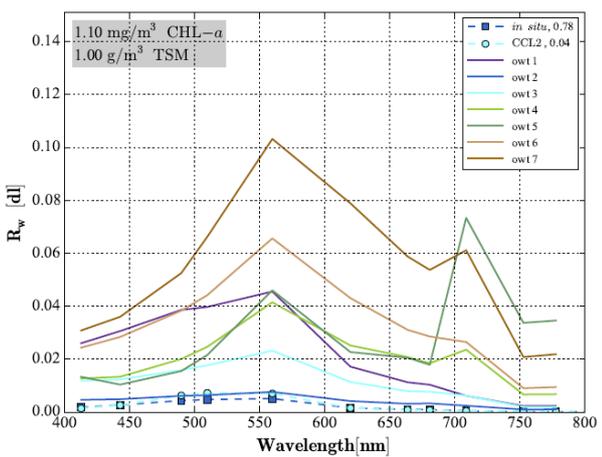
2007-04-20_11-05-00
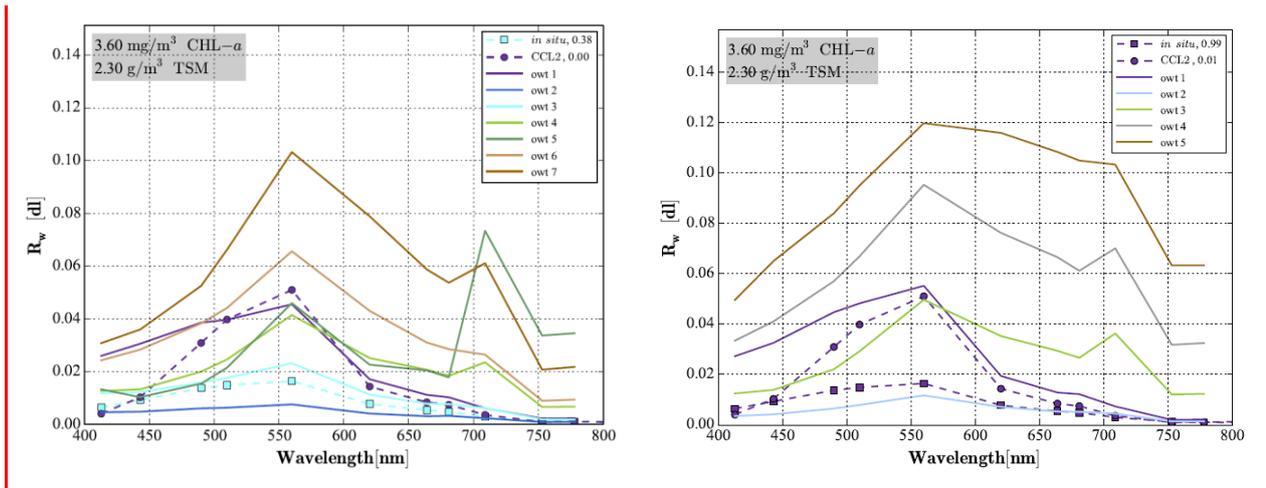
Lake Peipsi

Lake Peipsi is a shallow lake very influence by the population around it. That makes this lake quite vulnerable to eutrophication, and chlorophyll concentration values in the lake are generally high. This fact gives the lake a greenish colour.

Six daily OWT images are shown in Figure 14. All cases of the INLAND 7C classification show a lake quite complex, with several classes that separate it in two clear areas, north and south (especially visible in the April image), but also with some cases where most of the pixels could not be assigned to any class probably due to cloud covering or even ice (like the image of March). The GLASS 5C classifications seem to be simpler and able to distinguish the north and south basins in the image from April.

Some in situ data measured in the lake indicates that the chlorophyll concentrations ranked from 11-16 mg m^3 in the northern part of the lake at the end of April and in mid-June; while values up to 26-35 mg m^3 were found in Lämmijärv. The MERIS image from 21.06 is quite cloudy. The GLASS 5C seems to work better showing relatively high chl. The INLAND 7C shows additionally classes 6 & 7 in the northern part of the lake, which should mean both high chlorophyll and total suspended matter although in situ measured values does not indicate that. In the middle of July chlorophyll range goes from 22-29 mg m^3, SD ~1m in the northern and middle part of the lake. Based on the field data, northern and middle part of the lake seems quite homogeneous, and in better agreement with GLASS 5C than with INLAND 7C. The 2nd August images are relatively cloudy. Measured chlorophyll in northern part of the lake was around 25 mg m^3, being higher in the SE, as also seen from the INALND 7C classification. In the southern part of the lake chlorophyll was measured between 57-81 mg m^3. In the middle of Sept, in situ measured of chlorophyll ranges from 20-30 mg m^3 in the northern part of the lake, SD 1.3-2m, therefore the GLASS 5C seems more realistic.

Better understanding of lake can be observed in the monthly aggregations (Figure 15 and Figure 16). Again the north and south division is visible in both classifications, but it is more consistent in the INLAND 7C. In the GLASS 5C, only the image from April shows this distinction and it is too homogeneous the rest of the months, assigning almost all pixels to class 3. As seen from the in situ data, there was not high cyanobacteria bloom in the northern part of the lake in 2008. Chlorophyll was lower in March-June and higher during July to September, this is better captured from the INLAND 7C classification. This classification as well shows better the N-S gradient in the lake, where northern part is „clearer" and the southern has higher chlorophyll & TSM load. Additionally we might expect a different water type in the SW part of Peipsi, which is influenced by river inflow which is captured on the monthly averages from July-September in the INLAND 7C.

The yearly aggregate shows a large variety in Lake Peipsi with the INLAND 7C. The northern basin includes a mixture of four OWTs. In the GLASS 5c the mixture disappears and all the lake is classify within class 3, with no distinction between north and south (Figure 17). NS gradient (northern part clearer and southern more turbid) is not visible in any of the classifications. We might expect classes 4-7 in the southern part of the lake, although it shows only 3 to 6.
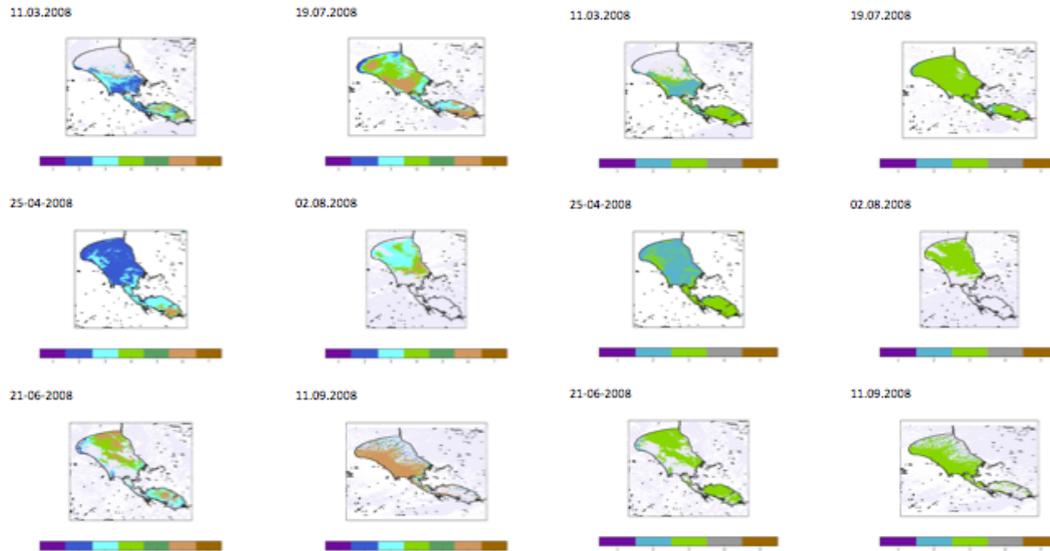
Figure 16. Lake Peipus, daily classes in INLAND 7C (left) and GLASS 5C (right) OWT classifications

Figure 17 Monthly averages of the dominant class, INLAND 7C

GLASS 5C

Figure 18 Monthly averages of the dominant class, GLASS 5C

Figure 19 Yearly average (2008) of INLAND 7C and GLASS 5C dominant class

Comparison with in situ data

Two days 2011, 12 of May and 2 of September had spectrum in situ data available, and it was used to compare with the shape spectra of the INLAND 7C and GLASS 5C classifications. In some cases, there is also the Coast Color AC derived spectra available and it was used to compare with the in situ data too.

The INLAND 7C in situ data are between classes 2 and 3; while the GLASS 5C classifies mainly on class 3. Concerning the CC AC spectra, is mainly classified as class 2 in the INLAND 7C and GLASS 5C. This confirms the observations made in the maps (Figure 16-19). The class_sum values for the GLASS 5C are close to 1, while the CC class_sum are even higher. Class_sum values are not so high in the INLAND 7C for the in situ values, but look quite good for the CC spectra.

INLAND 7C	GLASS 5C
2011-05-12_10-15-00	

| **2011-05-12_10-45-00** | |

| **2011-05-12_11-10-00** | |
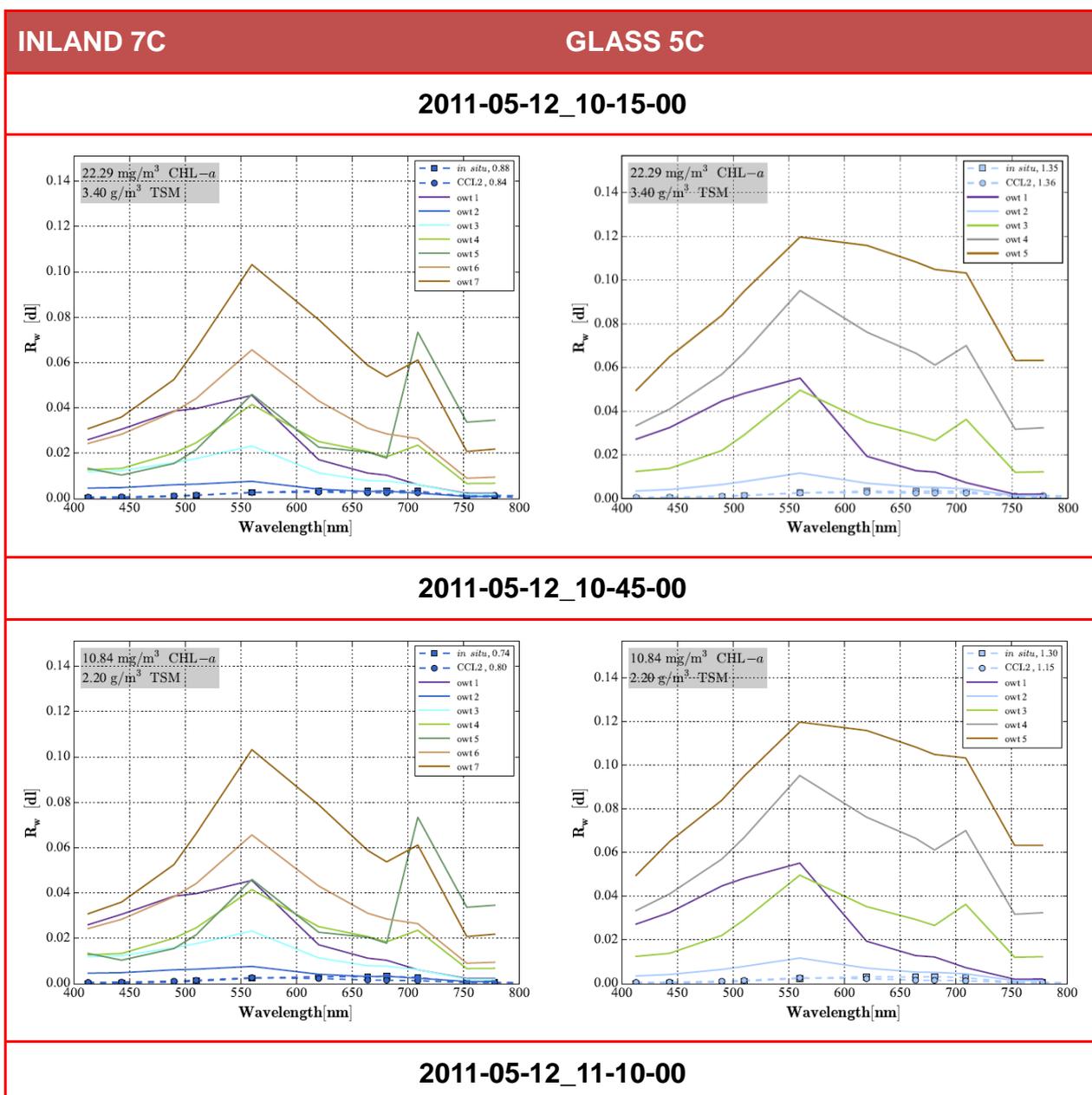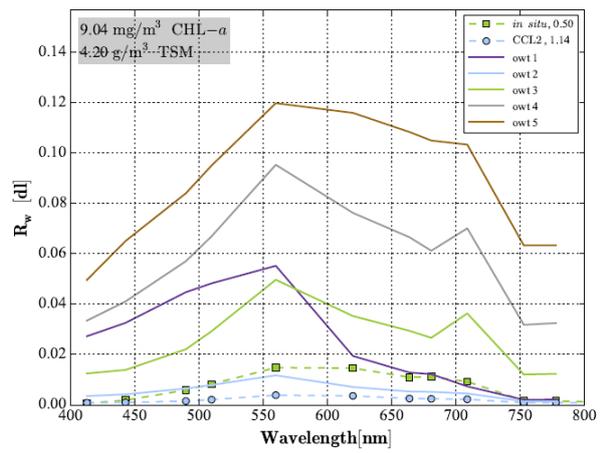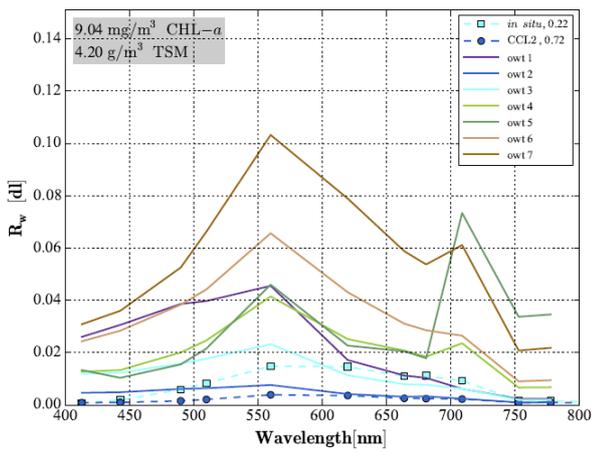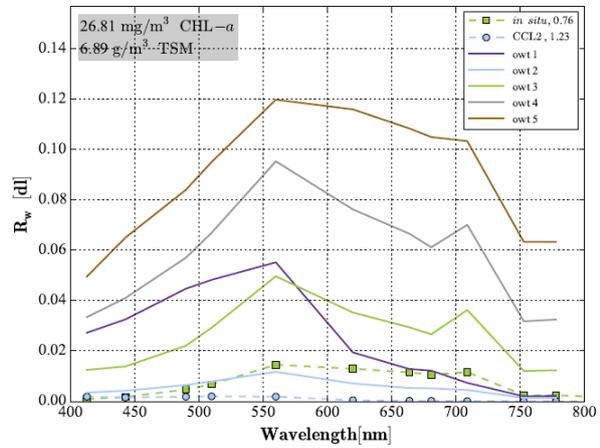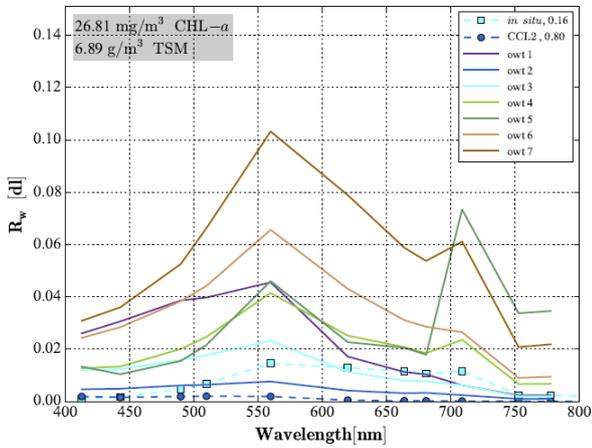
2011-09-02_10-15-00

2011-09-02_10-55-00
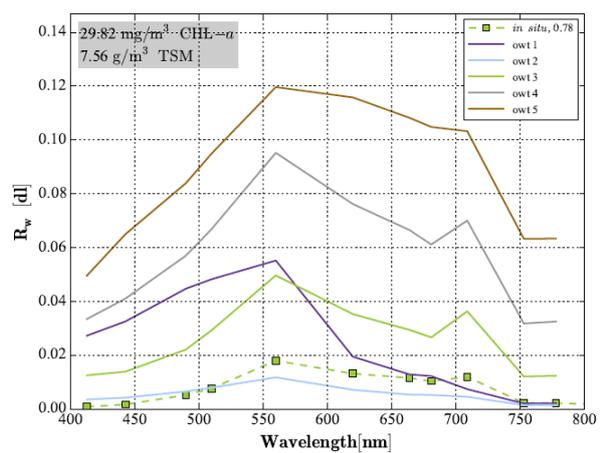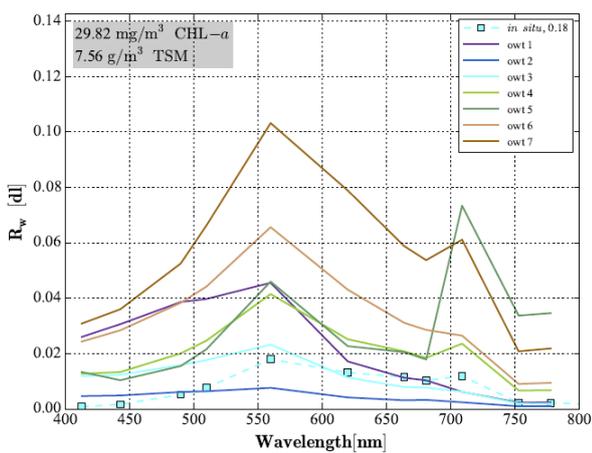
2011-09-02_11-25-00
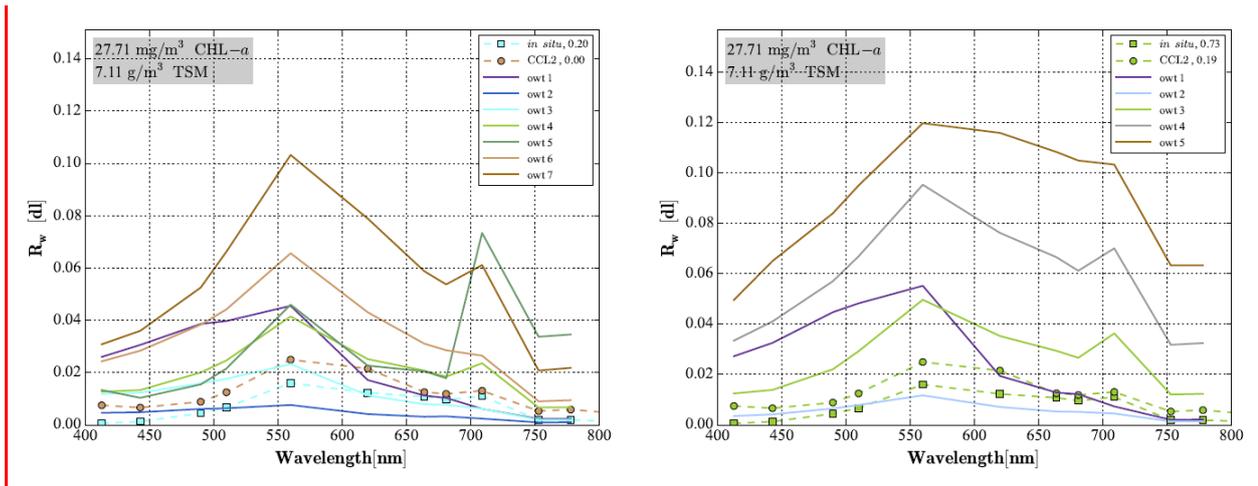
Lake Päijänne

Lake Päijänne is a boreal lake, which often have high CDOM absorption. , In Lake Päijänne TSM and Chl-a concentrations are low and CDOM is the dominant optical substance (although it is not very high when compared to other boreal lakes). Due to this, the reflectance of the lake is quite low and the lake appears dark blue. The lake is relatively deep (maximum depth almost 100 m, mean depth bout 16 m). However, the retention time is just over 2 years in Päijänne, so the potential degradation time for CDOM is quite short compared with other cases.

In Figure 18 we observe the INLAND 7C classification result for four daily averages in the central part of the year, where this blue characteristic is shown quite clearly. Most of the pixels are classified as class 2 or class 1. The image from June is not very good, probably the cloud cover was extended. In the Image from September some pixels, in the shores of the lake, appear classified as turbid. The main reason for incorrect classifications probably is the errors in the atmospheric correction (the elevated CDOM absorption causes problems for the AC).

In the GLASS 5C classification of Figure 19, most of the pixels are classified as class 3, which indicates some chlorophyll concentrations, clearly opposed to the results of the INLAND 7C processor. We see a lot of class 2, which is good. However, other classes are also visible (especially 3), which is likely to be incorrect. The reason for the failure is probably AC.

The monthly scenes shown go from May to October, in order to avoid error introduced by ice on the surface of the lake. In Figure 20, monthly averages confirm the assignation of the lake to clear water classes. During the last months of the summer the variability of the lake seems to be higher, with some pixels on the shore assigned to class 5. Figure 21, the monthly averages of the GLASS 5C classification, increase the variability in all months, with pixels assigned mainly to classes 2 and 3, but with the disappearance of the turbid water class shown in the previous classification near the shores.

23.05.2008
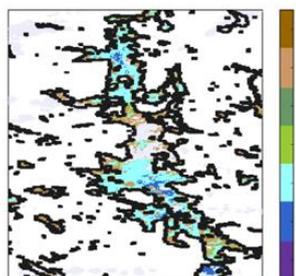
30.07.2008
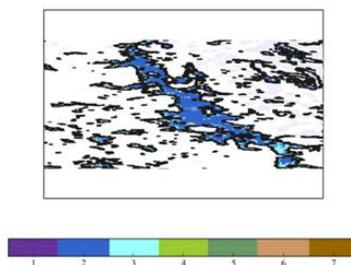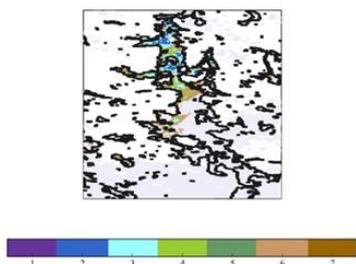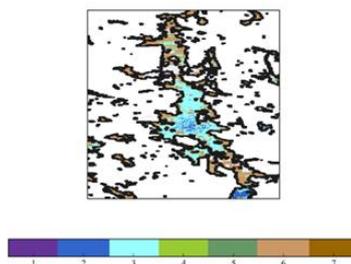
14.06.2008

08.09.2008

Figure 20 Lake Paijanne, daily classes in INLAND 7C OWT classification

23.05.2008
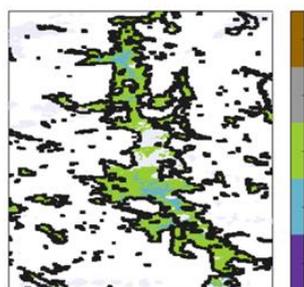
30.07.2008
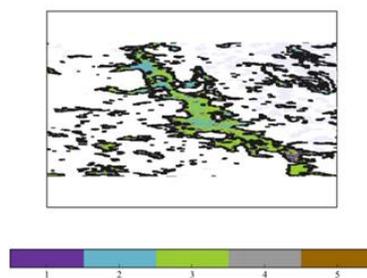
14.06.2008
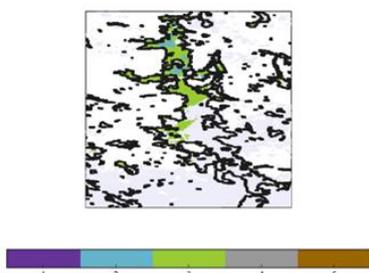
08.09.2008
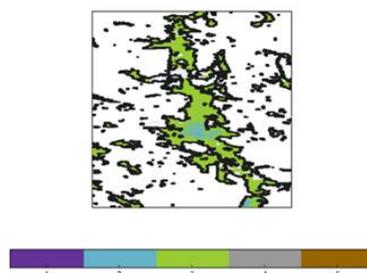
Figure 21 Lake Paijanne, daily classes in GLASS 5C OWT classification

May 2008
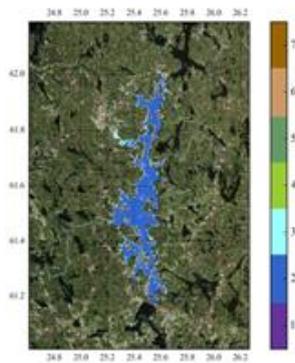
June 2008
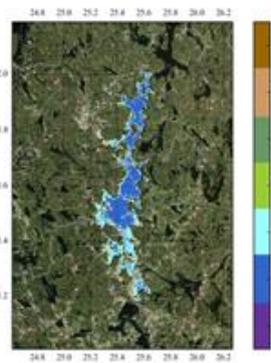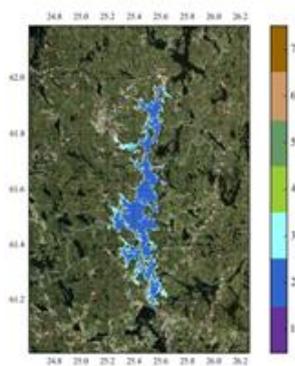
July 2008
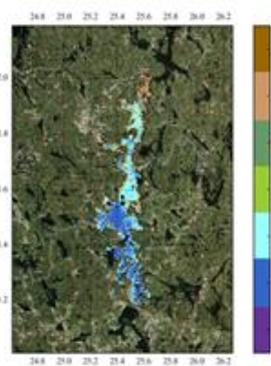
August 2008
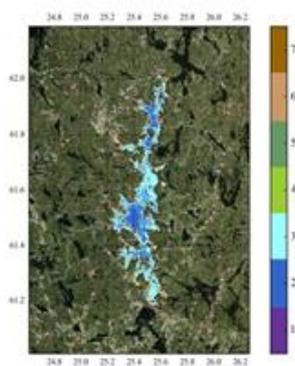
September 2008
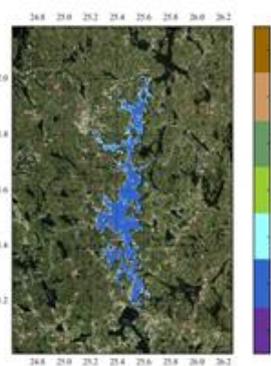
October 2008
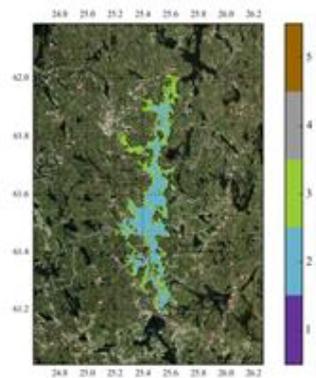

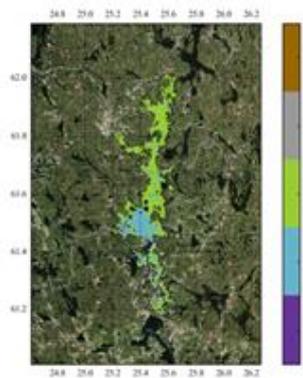*Figure 22 Monthly averages of the dominant class, INLAND 7C*
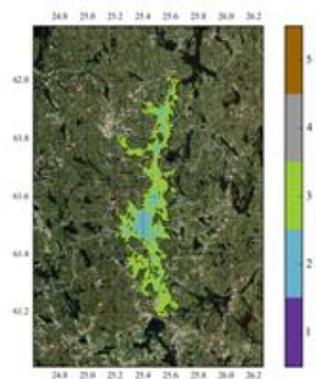
May 2008
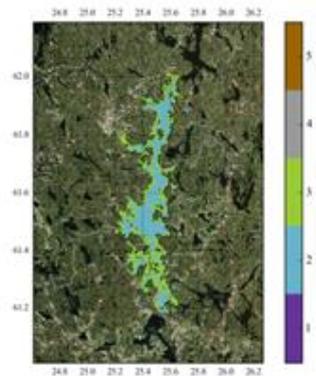
June 2008
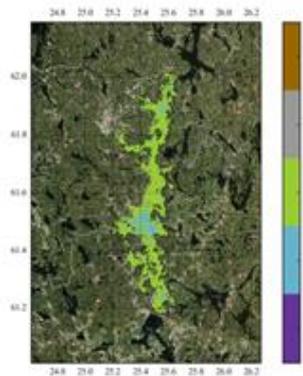
July 2008

August 2008

September 2008

October 2008
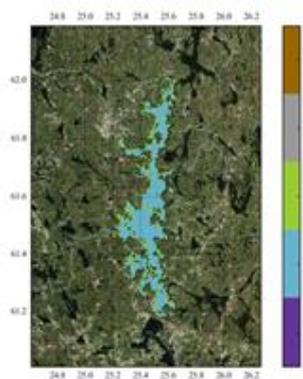
Figure 23 Monthly averages of the dominant class, GLASS 5C

Comparison with in situ data

Three days: 4 of June 2007, 7 of July 2007 and 29 of June 2011 had spectrum in situ data available, and it was used to compare with the shape spectra of the INLAND 7C and GLASS 5C classifications. In all cases, there is also the Coast Color AC derived spectra available and it was used to compare with the in situ data too.

INLAND 7C classifies the in situ spectra in classes 2 or 3; the same for GLASS 5C. The CC spectra follows quite well the shape of the in situ and is also classified as class 2 or 3. To remark, the class_sum values are higher in GLASS 5C than it INLAND 7C, indicating a better representation of these type of lakes in the GLASS 5C classification.

INLAND 7C	GLASS 5C
2007-06-04_11-30-00	

2007-08-07_10-53-00	

| 2011-06-29_12-45-00 | |
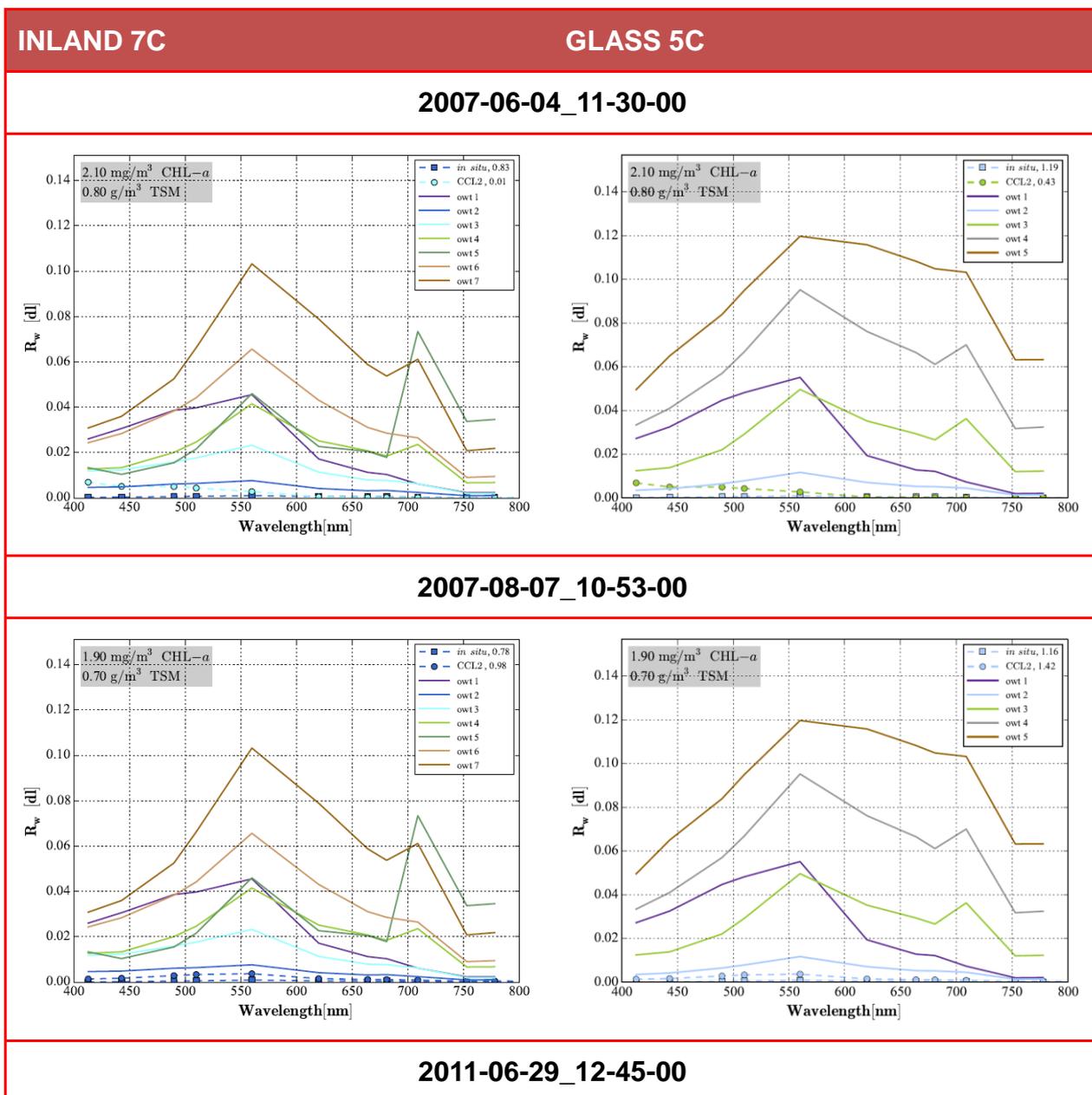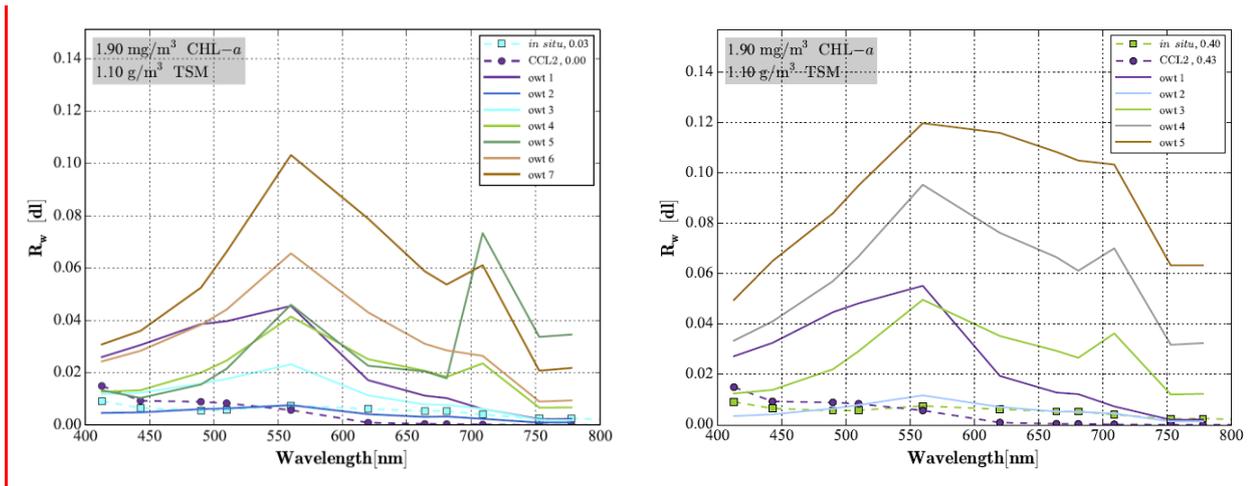
Appendix 3: Spectral water types of Alpine Italian lakes for two atmospheric corrections

Claudia Giardino and Mariano Bresciani

The following questions were addressed in this case study:

The OWT was applied for a subset of images where the knowledge about the performances of ICOL and C2R processors has been previously investigated (Odermatt et al., 2010; Bresciani et al., 2011).

For the same images also the CC processor has been used, according to the workflow selected within the task 5.3. The OWT 6_GLaSS processor has been finally applied to both set of images and a qualitative evaluation of the results has been reported (Tab. 1).

Qualitatively, for the largest lakes (from East to West; Maggiore, Como and Garda) we expect to find mostly class-1 (but also class 3 might occur in case of phytoplankton blooms, usually occurring in spring and summer). Then, we expect to find the smaller lakes classified as Class-3. However, since only Maggiore and Garda are "GLaSS Nearby Lakes", we limit the evaluation to those 2 lakes only.

Table 1. Comparison of the OWT results starting from images atmospherically corrected by using two different processors: On left the CC, on right ICOL+C2R
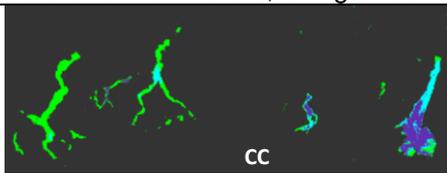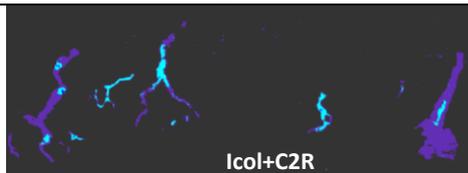
20050723	CC	Icol+C2R	Better C2R
20100726	CC	Icol+C2R	Comparable results
20100826	CC	Icol+C2R	Better CC
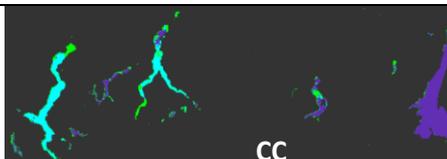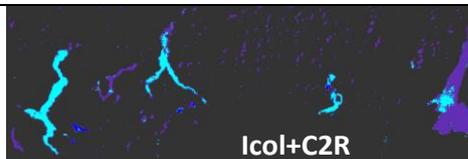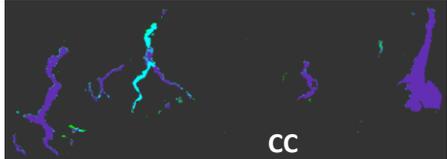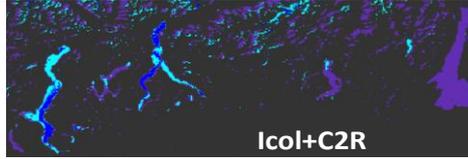
The MERIS images L1_FSG stored with the BC ftp site have been used to further increase the testing of the OWT. In this case the atmospheric correction was archived with the CC processor, according to the workflow selected within the task 5.3.

Overall (Tab. 2) the results are pretty fine for Lake Garda even the northern part is suspects as it should be also classified as Class-1. It seems that the in narrowest part of Garda the classification depend on adjacency effects. For Lake Maggiore the results are less satisfactory as many pixels are classified as Class-4, which is unrealistic. Similar as for the northern part of Lake Garda, the narrow and elongated shape of Lake Maggiore, in combination with the low signal coming from these clear lakes, makes tricky the correction for the atmospheric and adjacency effects. The use of ICOL prior to CC might improve the classification in the narrowest parts.

In conclusion, the tool indicates that CC might also work for Italian Alpine lakes, but (for Maggiore) only with ICOL. Such a conclusion is also supported by the comparison between Rrs spectra in situ measured and the corresponding MERIS data of Lake Garda. The following plot show the match-ups with in situ spectra for image data acquired in 2005 (26 July and 14 September) and on 6 May 2008. Overall, the optical closure is good, the absolute values are in the same range (despite in some cases the peak of reflectance computed by CC is anticipated with respect to in situ

measurements) and the shape and the magnitude are both those typical of class 1. Moreover, in these dates, the average concentration of Chl-a was 2.31 mgm-3 (± 0.9) so, again, it seems that for Lake Garda the class 1 is generally suitable.
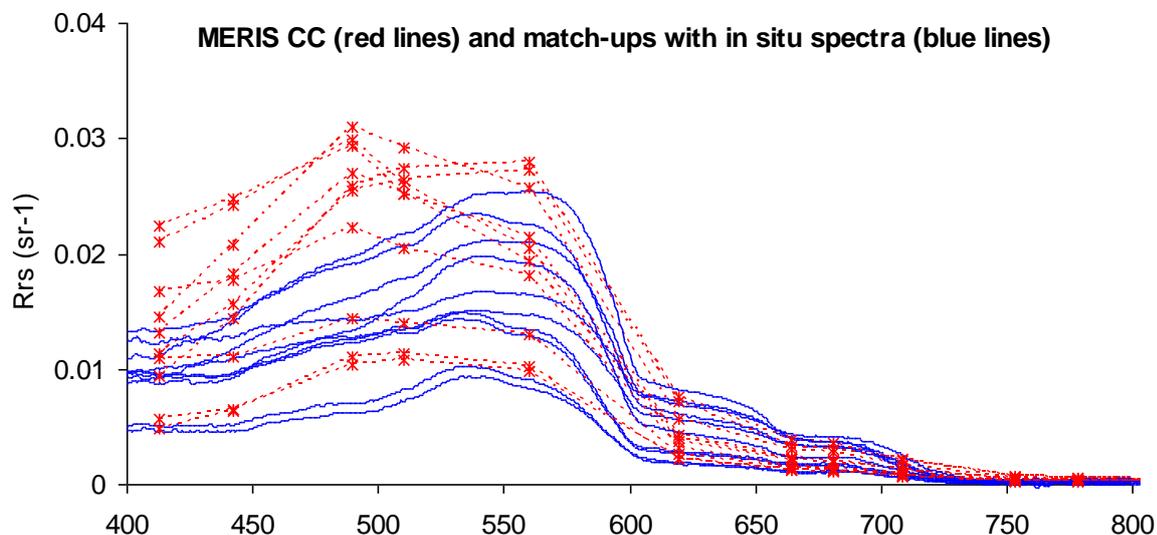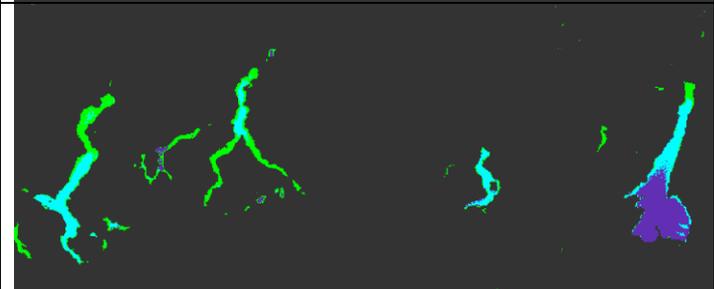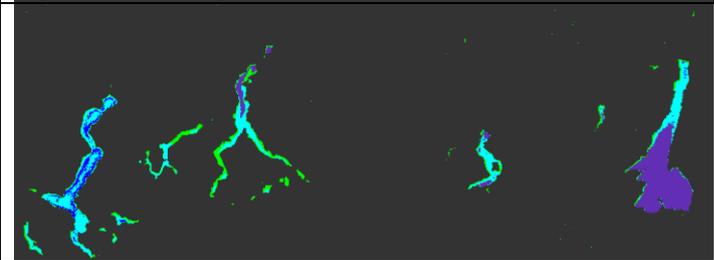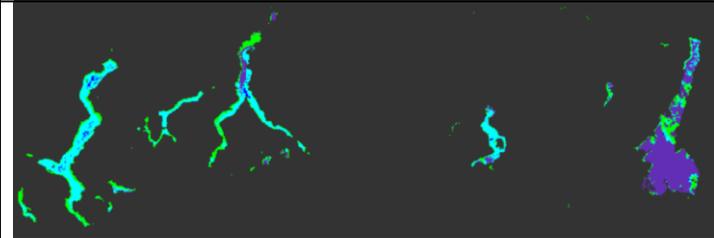

Figure 1. Comparison of the of Rrs spectra derived from in situ measurements (blue lines) and MERIS images (red lines with symbols) corrected for the atmospheric effect with the CC algorithm The data correspond to 10 stations in the southern part of Lake Garda observed on 20050914, 20050726 and 20080506.

Table 2. OWT results for a multi-temporal set of images processed with CC
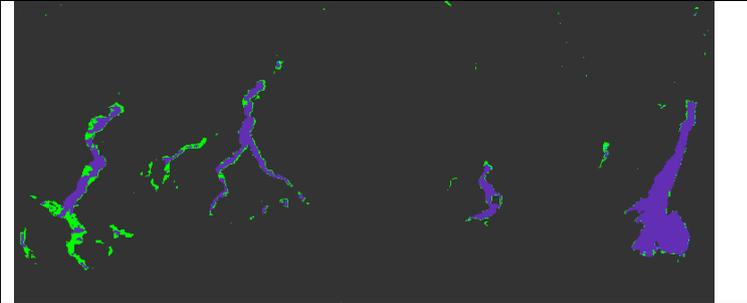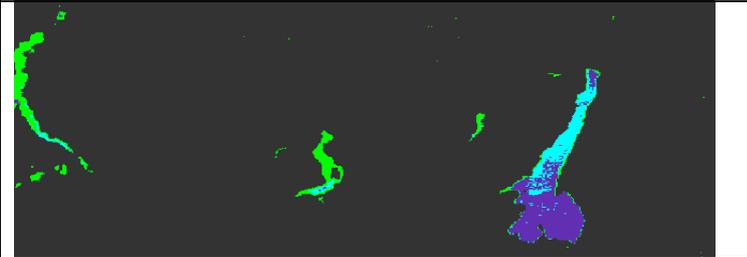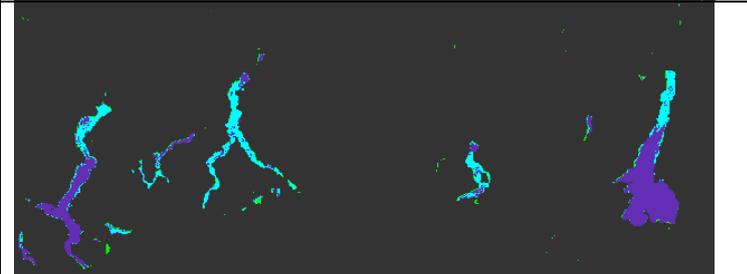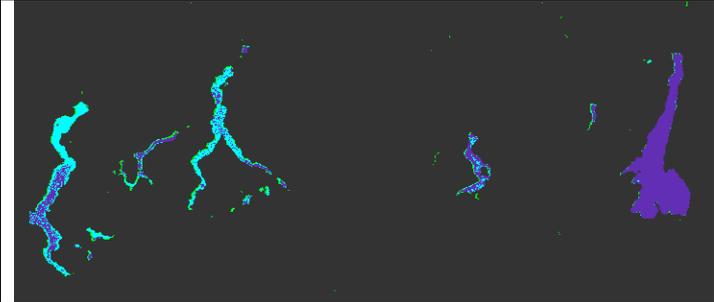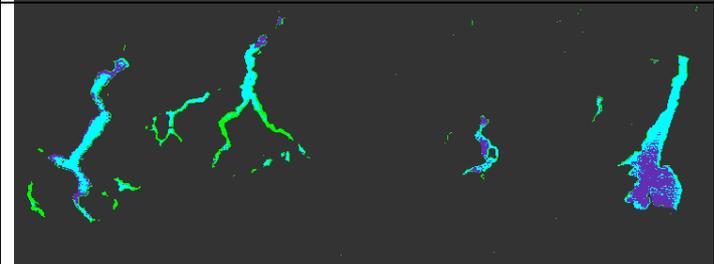
20050914	
20080506	
20090911	
20110410	
20110413	
20110626	