UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Programa de Pós-Graduação em Ciências Cartográficas

SARAH CRISTINA ARAÚJO MARTINS

THE a_{CDOM} SPATIAL AND TEMPORAL DISTRIBUTION ANALYSIS IN FUNIL RESERVOIR

Presidente Prudente
2017

UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Programa de Pós-Graduação em Ciências Cartográficas

SARAH CRISTINA ARAÚJO MARTINS

THE a_{CDOM} SPATIAL AND TEMPORAL DISTRIBUTION ANALYSIS IN FUNIL RESERVOIR

A dissertation submitted to the Faculty of Science and Technology of Sao Paulo State University in partial fulfillment of the requirements for the degree of Master of Cartographic Sciences.

Advisor: Prof. Dr. Enner Alcântara
Co-advisor: Prof. Dr. Karem Chokmani

Presidente Prudente
2017

FICHA CATALOGRÁFICA

Martins, Sarah Cristina Araújo.
M346a The a_{CDOM} spatial and temporal distribution analysis in Funil reservoir / Sarah Cristina Araújo Martins. - Presidente Prudente : [s.n.], 2017
xxviii, 102 f. : il.

Orientador: Enner Herenio de Alcântara
Coorientador: Karem Chokmani
Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia
Inclui bibliografia

1. Colored dissolved organic matter (CDOM). 2. Bio-optical models. 3. CDOM absorption coefficient. I. Alcântara, Enner Herenio de. II. Chokmani, Karem. III. Universidade Estadual Paulista. Faculdade de Ciências e Tecnologia. IV. Título.

unesp

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: THE aCDOM SPATIAL AND TEMPORAL DISTRIBUTION ANALYSIS IN FUNIL RESERVOIR

AUTORA: SARAH CRISTINA ARAUJO MARTINS
ORIENTADOR: ENNER HERENIO DE ALCÂNTARA

Aprovada como parte das exigências para obtenção do Título de Mestra em CIÊNCIAS CARTOGRÁFICAS, área: AQUISIÇÃO, ANÁLISE E REPRESENTAÇÃO DE INFORMAÇOES ESPACIAIS pela Comissão Examinadora:

Prof. Dr. ENNER HERENIO DE ALCÂNTARA
Departamento de Engenharia Ambiental / Instituto de Ciência e Tecnologia de São José dos Campos

Profa. Dra. THANAN WALESZA PEQUENO RODRIGUES GUIMARÃES
IFPA / Instituto Federal do Pará

Prof. Dr. NILTON NOBUHIRO IMAI
Departamento de Cartografia / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Presidente Prudente, 03 de agosto de 2017

Aos professores que tive durante toda a vida, que me deram as armas com as quais lutar. À minha família, sem a qual não teria aprendido pelo que lutar. E aos amigos, sem os quais eu jamais teria a força e o apoio necessários para enfrentar as batalhas diárias da vida.... a todos vocês eu dedico esta conquista e um amor imenso.

AGRADECIMENTO

Agradeço a todos que participaram de alguma forma deste trabalho. Tenho sorte por tê-los comigo.

Agradeço especialmente:

À minha família, por todo apoio e força, capazes de superar distâncias tão grandes.

Ao meu marido, Victor, pelo amor e por ter tanta paciência comigo, quando eu mesma não tinha.

Ao meu orientador na UNESP, Enner Alcântara, que não me negou esta oportunidade, mesmo sem me conhecer. Agradeço por sempre exigir mais, por ter um olhar diferente sobre todas as coisas, por sempre contribuir.

Ao meu diretor de pesquisa no INRS, Karem Chokmani, que igualmente aceitou um desafio de me receber. Agradeço por todo apoio e suporte durante o estágio de pesquisa, e às oportunidades de compartilhar conhecimento.

Aos professores do Programa de Ciências Cartográficas da UNESP, por todo o conhecimento adquirido.

Aos amigos e colegas da UNESP e do INRS, que tanto compartilharam conhecimento comigo, contribuindo não só com esta pesquisa, mas com a minha formação pessoal e profissional. Meu obrigada especial a: Carol Piffer pelas infinitas correções de inglês; Carol Campos por sempre me receber sua casa nas minhas idas à Presidente Prudente; Fer Watanabe, Luiz Rotta, Carol Ambrósio, Bruno Faga e Nariane Bernardo pela disponibilidade em ajudar a encontrar soluções; e Anas El Alem por me ajudar a automatizar algumas etapas deste trabalho… *vous êtes super*!

À Cinthia da Seção de Pós-graduação, sempre disposta a ajudar e a resolver problemas.

Aos membros da banca, por aceitarem avaliar e contribuir com este trabalho.

À Universidade Estadual Paulista (UNESP – São Paulo, Br) e ao Institut National de la Recherche Scientifique (INRS – Québec, Ca) pela infraestrutura para realização desta pesquisa.

Ao Instituto Nacional de Pesquisa Espacial (INPE) e ao pesquisador Igor Ogashawara, hoje na Indiana University-Purdue University Indianapolis (IUPUI – Indianapolis, USA), por ceder os dados de campo utilizados neste trabalho.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – Ministério da Ciência e Tecnologia do Brasil) e ao *Programme des Futurs Leaders dans les Amériques* (PFLA – Gouvernement du Canada) pelo apoio financeiro na forma de bolsas de estudos.

"Quem não sonha já morreu.

O exercício do sonho, de sonhar o que nos parece impossível, equivale a vencer o desafio de decifrar o futuro de um ponto de vista próprio.

Quem não for insensato nos seus sonhos estará condenado a viver feito cego de espírito.

O sonho é o olho aberto para o futuro e o quase delírio de imaginá-lo melhor.

Mas o sonho não se basta.

Viver?

Viver é perseguir o sonho.

Viver é viver o sonho ou não é viver.

Viver é acordar o sonho."

(César FNM.)

RESUMO

A matéria orgânica dissolvida (DOM) é a componente da água que pode ser usada como indicativo de sua qualidade, pois possui duas fontes: uma alóctone, relacionada com descargas de material terrestre, estando vinculada aos ácidos húmicos, e outra autóctone, associada às descargas fluviais ou produção própria do corpo hídrico estudado, estando relacionada aos ácidos fúlvicos. A matéria orgânica dissolvida colorida (CDOM) é a fração colorida da DOM, que pode ser usada como *proxy* para a observação desta última em águas interiores. O reservatório hidrelétrico de Funil (FHR) foi o corpo hídrico escolhido como área de estudo deste trabalho. Neste contexto, o objetivo geral desta pesquisa foi identificar e avaliar as mudanças no coeficiente de absorção da CDOM (a_{CDOM}) na superfície da água ao longo do tempo (1995 – 2010), bem como entender a sua relação com mudanças no uso e cobertura da terra (LULC) na bacia de contribuição do FHR. Para alcançar tal objetivo foram realizados: (i) o mapeamento histórico de LULC (1995 – 2010, com 5 anos de intervalo) para detecção de mudanças; (ii) o estudo de um conjunto de modelos bio-ópticos baseados na literatura, bem como de um novo modelo empírico desenvolvido para estimar a_{CDOM} via reflectância simulada ($R_{rs_{simulated}}$) para o sensor *Thematic Mapper* (TM); (iii) a distribuição espaço-temporal do a_{CDOM} por meio da aplicação de um modelo bio-óptico em imagens TM/Landsat-5 de 1995 a 2010, e (iv) a análise das fontes possíveis de CDOM/DOM , assim como do comportamento/distribuição do a_{CDOM} no FHR ao longo do tempo. Assim, o primeiro estudo desenvolvido nesta pesquisa foi o da parametrização do algoritmo maquinas de vetores de suporte (SVM) de acordo com as características da área de estudo para classificação supervisionada de LULC na bacia de contribuição do FHR. A detecção de mudança da classificação obtida para LULC demonstrou que a parametrização proposta para o SVM tornou o algoritmo capaz de diferenciar classes grandes e contínuas, classes estreitas e alongadas, além de áreas não contínuas e pequenas localizadas dentro de outra classe maior. A classificação obtida para o SVM apresentou boa avaliação estatística, com acurácia geral entre 86% e 96% para toda a série temporal, acurácia do produtor de 90%, acurácia do usuário maior do que 86% e índice Kappa entre 86% e 91%. Ainda, foi observado que o LULC desenvolvido na área de estudo se manteve relativamente estável ao longo da série histórica analisada. O segundo estudo realizado proporcionou o desenvolvimento de um modelo empírico em um comprimento de onda (485 nm) e uma razão de bandas (B4/B1) alternativos para estimativa de a_{CDOM} via $R_{rs_{simulated}}$ para o TM/Landsat-5 (RMSE = 7%, Nash = 0.91). Este modelo também pôde identificar mesmo pequenas variações nos valores de reflectância via dados orbitais, assim como pode diferenciar variações sutis no

a_{CDOM}. Ainda, foram identificados dois padrões de comportamento da CDOM para o FHR: um associado ao LULC e à ocorrência de chuva/lixiviação, bem como outro relacionado à Clorofila-a (Chl-a) em situações de floração de algas. Os referidos estudos que compõe esta pesquisa foram padronizados como artigos científicos para a confecção deste documento. O primeiro estudo, sobre a parametrização do SVM, foi publicado na revista Modelling Earth Systems Environment – Springer (DOI 10.1007/s40808-016-0190-y). O segundo estudo, sobre a distribuição histórica do a_{CDOM} está na etapa de revisão para futura submissão.

Palavras-chave: Matéria orgânica dissolvida colorida (CDOM); coeficiente de absorção da CDOM (a_{CDOM}); modelos bio-ópticos; TM/Landsat-5; uso e cobertura da terra; algoritmo máquina de vetores de suporte (SVM).

ABSTRACT

The dissolved organic matter (DOM) is a water compound related to water quality, since it has two sources: one allochthonous, related to terrestrial discharges that can be linked to humic acids, and another autochthonous, associated with river input and itself production, so related to or fulvic acids. The colored dissolved organic carbon (CDOM) is the colored fraction of DOM that could be used as a proxy for its occurrence in inland waters. The Funil hydroelectric reservoir (FHR) was chosen as the study site for this work. In this context, the general aim of this research was to identify and to evaluate the changes in CDOM absorption coefficient (a_{CDOM}) at the water surface over time (1995 – 2010), and to understand its relationship with land cover land use (LULC) changes in FHR watershed. For match this goal, (i) a LULC historical mapping (1995 – 2010, with 5 years of interval) was made to change detection, (ii) a bio-optical model set and a new model were studied in order to estimate a_{CDOM} from simulated reflectance ($R_{rs\,simulated}$) for Thematic Mapper (TM) sensor, (iii) a a_{CDOM} spatial and temporal distribution was obtained by applying a bio-optical model in TM/Landsat-5 imagery from 1995 to 2010, and (iv) the possible CDOM/DOM sources in FHR were found, as well a_{CDOM} historical behavior/distribution over time was analysed. Thus, the first study was the support vector machine algorithm (SVM) parameterization according to study area characteristics to LULC supervised classification in FHR watershed. The obtained LULC change detection analysis demonstrates that the proposed SVM parameterization made the algorithm able to differentiate large and continuous classes, lengthy and thin areas, and non-continuous small areas located inside wide classes. The obtained classification had great statistics with overall accuracy among 86% and 98% over the time series, the producer accuracy of 90%, the user accuracy higher than 86%, and the Kappa statistics ranged from 86% to 91%. In addition, no significant changes in LULC were identified in the study site over all time series. The second study provides a bio-optical model at alternatives wavelength (485 nm) and a band ratio (B4/B1) for a_{CDOM} estimation using simulated R_{rs} for TM/Landsat-5 (RMSE = 7%, Nash = 0.91). This model could identify even small variations in reflectance values from orbital data, as well as differentiate even slight alterations in a_{CDOM}. Two significantly different a_{CDOM} behaviors were also identified for FHR: one associated with LULC and rainfall/runoff occurrence, and other correlated to Chlorophyll-*a* high concentrations (Chl-*a*) in algal blooms situations. The referred studies that compose this research ware standardized as academic articles in this document. The first study, about SVM parameterization, was published yet in

Modeling Earth Systems Environment – Springer (DOI 10.1007/s40808-016-0190-y). The second study, about a_{CDOM} historical distribution is in the revision step to future submission.

Keywords: Colored dissolved organic matter (CDOM); CDOM absorption coefficient (a_{CDOM}); bio-optical models; TM/Landsat-5; land cover land use (LULC); support vector machines algorithm (SVM).

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AOP – Apparent optical properties

PPA – Permanent protection area

C – Carbon

CDOM – Colored dissolved organic matter

Chl-*a* – Chlorophyll-*a*

CSI – Consortium for Spatial Information

DEM – Digital elevation model

DIC –Dissolved inorganic carbon

DN – Digital number

DOC – Dissolved organic carbon

DOM – Dissolved organic matter

FAO – Food and Agriculture Organization

FLAASH – Fast line-of-sight atmospheric analysis of hypercubes

INEA – Environmental State Institute of Rio de Janeiro

INPE – National Institute for Space Research

IOP – Inherent optical properties

IR-MAD – Iteratively reweighted multivariate alteration detection

ISS – Inorganic suspended solids

LOOCV – Leave one out cross validation

LULC – Land cover land use

MAD – Multivariate Alteration Detection

MLC – Maximum likelihood classifier

MODTRAN4 – Moderate spectral transmittance – version 4

Nash – Sutcliffe model efficiency coefficient

NIR – Near Infrared

NRMSE – Normalized root mean square error

OD – Optical density

OSC – Optically significant constituents

OSS – Organic suspended solids

POC – Particulate organic carbon

QAA –Quasi-Analytical Algorithm

QAA-BBHR – QAA developed for eutrophic inland waters

QAA-v5 – QAA version 5

RBF –Radial Base Function

RHF – Funil hydroelectric reservoir

RMSE – Root mean square error

SPM – Suspended particles matter

SRM – Structural risk minimization

SRTM – Shuttle Radar Topography Mission

SVM – Support Vector Machines

TM –Thematic Mapper

USGS – United States Geological Survey

UV – Ultraviolet

LIST OF SYMBOLS

CH_4 – Methane

CO_2 – Carbon dioxide

N_2O – Nitrous oxide

$a(\lambda)$ – Absorption coefficient of the total, $a_w + a_\phi + a_{CDOM} + a_{dg}$

a_{CDOM} – Absorption coefficient of colored dissolved organic matter

a_{dg} – Absorption coefficient of colored dissolved organic matter and detritus

a_p – Absorption coefficient of suspended particles

a_{ph} – Absorption coefficient of phytoplankton pigments

a_w – Absorption coefficient of pure water

b_b – Backscattering coefficient of the total, $b_{bw} + b_{bp}$

b_{bp} – Backscattering coefficient of suspended particles

b_{bw} – Backscattering coefficient of pure water

E_s – Spectral downwelling irradiance (Sun)

L_t – Total water-leaving spectral radiance (water)

L_{TOA} – Radiance top of atmosphere

R_{rs} – Remote sensing reflectance above surface

$R_{rs_{simulated}}$ – Simulated remote sensing reflectance above surface

r_{rs} – Remote sensing reflectance below surface

z – Depth

α – Significance level (used $\alpha = 5\%$)

θ – Zenith angle

ϕ – Azimuth angle

λ – Wavelength

λ_0 – Reference/base wavelength

ρ_{TOA} – Reflectance top of atmosphere

$\rho_{surface}$ – Surface reflectance

CONTENTS

CHAPTER 1: Introduction and objectives

1.1 INTRODUCTION

The global carbon balance computes the Earth surface as an important carbon sink (Denman and Brasseur 2007). The global sink is estimated in approximately $2.6 \pm 1{,}7$ carbon Pg per year^{-1}, excluding the deforestation emissions (Luyssaert et al. 2008). Lakes, reservoirs, and rivers are part of the earth surface, but water bodies were not considered in gases global emissions balance. However, surveys indicate a tendency to carbon (C) emissions in eutrophic waters by respiration (primary production) and dissolved organic carbon (DOC) photodegradation (Pacheco, Roland, and Downing 2014). Thus, inland waters can be a substantial source of carbon dioxide (CO_2), and methane (NH_4), contributing to global warming and climate changes (Bastviken et al. 2011).

Climate changes are considered as the hot topic to worldwide ecosystems (Rosenzweig et al. 2007). Monitoring and understanding its effects are difficult activities mainly because of the different response patterns associated with each ecosystem diversity and to spatial variation in landscapes worldwide. Surveys indicate that inland waters are sensitive to weather/climate, and environment variations since they respond quickly to physical, chemical and biological alterations (Curtarelli et al. 2014). Still, studies demonstrated that inland waters can be considered as great indicators of global climate change because of its capacity to integrate the results related to alterations in surroundings and atmosphere (Carpenter et al. 2007).

In summary, inland waters can be considered as climate change indicator since: (i) aquatic systems are well delimited, and sustainable studies are possible in this kind of target, (ii) they respond directly to environment/atmospheric changes, incorporating its effects in watershed area, (iii) they integrate changes response over time, which can be filtered to remove aleatory errors, and (iv) water bodies are a type of target distributed worldwide, which is useful to compare climate change effects in different longitudes (Adrian et al. 2009).

Besides that, environmental changes that occur in the flux and discharge volume, water temperature, water compounds structure (as phytoplankton concentration, dissolved organic matter –DOM- concentration, suspended particle matter – SPM – quantities), and dissolved oxygen values, etc., impact the reservoir structure and its water quality condition. Then, alterations in a single water body may lead to changes in all aquatic system downstream.

Reservoirs are important for transportation, transformation and storage of C quantities, especially due to sediment occurrence (Tranvik et al. 2009). The dissolved inorganic carbon (DIC) is generally associated with terrestrial sources, and it is particularly important to gases

exchange between atmosphere and water surface, playing an important role in the global carbon cycling either. In addition, DIC is important as an indicator of biological activity because of its control of the oxygen exchange.

DOM is a substantial factor to soil and water biogeochemical process and is originated in water bodies by leaching, when it is associated to soil microbial activities, and by itself production, when it is associated with the water compounds structure (distribution and concentration). The related DOC is determined by factors as substrate composition, microbial community structure, temperature, water flows, and surrounding characteristics, which produce chemical variants in water body composition. Thus, the soil management practices as well as river discharges characteristics may affect DOM, modifying DOC point of balance in a reservoir (Cole et al. 2007; Matsuoka et al. 2015; F. S. Pacheco et al. 2015; Rochelle-Newall and Fisher 2002).

Indeed, inland waters are capable of affecting the weather/climate in local and regional scales through the heat transfer between water and atmosphere, presenting an important function in carbon global cycling as a C emitter or sink. It is true since consumption and/or production of CO_2, CH_4, and nitrous oxide (N_2O) by microorganisms influence the concentration of these greenhouse gases in the atmosphere, and its heat balance (Krinner 2003; Pacheco, Roland, and Downing 2014).

As a rule, DIC and DOC are the main sources of C in lakes and reservoirs, followed by organic suspended sediments (OSS) and inorganic suspended solids (ISS). The relative importance of these sources varies according to water body location, and to its hydrologic dynamics, generating longitudinal differences in reservoirs C fractions worldwide. They may also reflect the different characteristics of weather, soil, geochemical structure (soil and water), river discharges, nutrient availability, and land cover land use (LULC) possibilities, etc. (Effler et al. 2015).

Colored carbon fractions present in water bodies can be estimated by bio-optical models, which use remote sensing reflectance (R_{rs}, sr^{-1}) to estimate water optically significant constituents (OSC) or to retrieve water inherent optical properties (IOP), as total absorption ($a(\lambda)$) and scattering ($b(\lambda)$, m^{-1}) coefficients.

IOPs are the water properties which can be directly measured, are independent of the light field, and are linked to the interaction between the OSCs and solar irradiation. There are also the apparent optical properties (AOP), which are not directly measured, but derived from other quantities, and are highly dependent on the light field, the depth and the quantity of incident energy (Kirk 2011; Mobley 2001). As an example of AOPs we have R_{rs}.

Both, IOPs and AOPs are related themselves, allowing to retrieve $a(\lambda)$ and backscattering $b_b(\lambda)$ coefficients from R_{rs} using models or algorithms specially developed for this purpose. Based on this possibility, we studied a bio-optical models set performance for retrieving the colored dissolved organic matter absorption coefficient in a given wavelength ($a_{CDOM}(\lambda)$) for a tropical eutrophic reservoir. This set is composed by Quasi-analytical algorithm (QAA- Lee, Carder, and Arnone 2002) and empirical models developed by Del Castillo and Miller (2008); D'Sa and Miller (2003); Ficek, Zapadka, and Dera (2011); Griffin et al. (2011); Kutser et al. (2005); Mannino, Russ, and Hooker (2008), and studied by (Zhu et al. 2014), all of them specifically developed to retrieve $a(\lambda)$ from R_{rs} in complex inland waters.

QAA is an algorithm based on the radiative transfer equations and its main objective is to accurately retrieve a and b_b coefficients from R_{rs} through analytical, semi-analytical and empirical sequenced steps, which ensure QAA application in almost all water cases with just few modifications (Lee, Carder, and Arnone 2002). Empirical steps use statistics regressions, being geographic and temporal dependent on the data; semi-analytical ones are originated from numeric optimization guaranteeing geographical, and temporal flexibility, and finally, analytical steps are linked to radiative transfer equations having independence from local and time. QAA can also provide the specific absorption coefficient of CDOM plus detritus (a_{dg}), or just CDOM (a_{CDOM}) and phytoplankton (a_{ph}) (Cheng Feng Le et al. 2009; Lee, Carder, and Arnone 2002; Watanabe et al. 2016; Yang et al. 2013; Zhu et al. 2011; Zhu and Yu 2013).

In contrast with QAA spatial flexibility, empirical models do not use the radiative transfer equations in estimations, but the direct mathematical relation among IOPs and AOPs by numeric regressions. Thus, a rigorous parameterization procedure is needed to construct an efficient and reliable model to IOPs retrieving by using field or simulated data to calibrate and validate the model. Therefore, empirical models are dependent on its baseline data, and are geographically restricted as a consequence (Zhu et al. 2014).

CDOM is the colored fraction of the dissolved organic matter (DOM). Dissolved organic carbon (DOC) is a great part of DOM for inland waters. According to this relationship, CDOM or a_{CDOM} may be used as a proxy to retrieve the total DOC concentration, since a high correlation between them is frequently found (Esteves 1998; Zhu et al. 2014). However, recent surveys already demonstrated that CDOM can also be correlated to phytoplankton and Chlorophyll-a (Chl-a) in specific situations (Danhiez et al. 2017; Kirchman et al. 1991; Sasaki et al. 2005; Xue, Zhang, and Duan 2016; Zhang et al. 2009).

Depending on CDOM main source, DOM content in inland waters can be associated with: i) external dynamics and structure, which means LULC and weather/climate characteristics that provide organic matter (OM) to reservoirs by leaching (Allan 2004; Cole et al. 2007; Lee, Carder, and Arnone 2002; Zhu et al. 2014), and ii) internal dynamics that involves the reservoir contents, river discharges and aquatic environment behavior, which define the water compounds distribution (Cole et al. 2007; Xue, Zhang, and Duan 2016; Zhu et al. 2014). Despite the CDOM source and its correlation with DOC or Chl-a, the relationship between CDOM and DOM is still the same, both increase or decrease proportionally, and linearly (Brezonik et al. 2015; Danhiez et al. 2017).

a_{CDOM} estimates using orbital data is possible by applying a bio-optical model in satellite imagery; it was already made and presented successful results for Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+), Operational Long Imager (OLI) (Kutser et al. 2005, 2016; Zhu et al. 2014), Moderate Resolution Imaging Spectroradiometer (MODIS) (Chen and Zhang 2015) and Medium Resolution Imaging Spectrometer (MERIS) (Kutser et al. 2015; Yang et al. 2013), for example. In this research, the TM/Landsat-5 legacy was used to obtain the spatial and temporal distribution of a_{CDOM} for a tropical eutrophic reservoir between 1995 and 2010, as well as to verify the historical LULC in the same period of time.

To ensure the bio-optical model applicability in orbital data, the field hyperspectral measurements were simulated according to the related signal registration at the TM/Landsat-5 bands central wavelength. By this procedure, the hyperspectral field data became simplified/general, as it would be registered by the TM multispectral sensor. The resultant data are useful to calibrate and validate bio-optical models to obtain reliable estimations from orbital imagery.

Land cover is referent to the natural cover present in a given area, while land use includes the natural or artificial surfaces made/managed by humans (Briassoulis 2000; FAO 1994). This classification segments the surface in clusters with similar features, making easier to associate them to OM terrestrial sources to the Funil hydroelectric reservoir.

Certainly, historic LULC study and mapping, allied with the a_{CDOM} spatial and temporal distribution in reservoirs can be useful to understand the main causes/sources of C in inland waters, providing data to water quality, water balance, air quality, and even global C cycling researches (Briassoulis 2000; Cole et al. 2007). In this context, LULC changes mapping and a_{CDOM} estimates can be made by remote sensing, and the relationship among them for inland waters can be established and analysed.

1.2 HYPOTHESIS

The DOM is originated from internal (phytoplankton, periphyton, and macrophytes), and external sources (dissolved humic substances from the watershed), but this last one is the main source for tropical inland water bodies (Esteves 1998). Thus, changes associated with LULC in the reservoir watershed may affect DOM and a_{CDOM} in the water surface. Based on this, a bio-optical model parameterization allied to its application in satellite imagery may indicate spatial and temporal changes in a_{CDOM} in reservoirs, which should be mainly dependent on LULC changes.

1.3 OBJECTIVES

1.3.1 MAIN OBJECTIVE

The main aim of this research is to identify and to evaluate the changes in a_{CDOM} at the water surface over time, and to understand its relationship with LULC changes in FHR watershed.

1.3.2 SPECIFIC OBJECTIVES

To map the historical LULC in the study area from 1995-2010, to verify its changes.

To calibrate and validate a bio-optical model (empirical model or QAA) in order to estimate $a_{CDOM}(\lambda)$ from R_{rs}.

To retrieve a_{CDOM} spatial and temporal distribution by applying a bio-optical model in TM/Landsat-5 imagery from 1995 to 2010.

To identify the possible a_{CDOM} sources in FHR, as well as to analyse its distribution.

1.4 DISSERTATION STRUCTURE

This dissertation is divided in five chapters plus references which concerns the themes described below. Generally, we have three chapters containing the research structure, as theme contextualization, objectives, hypotheses, study area characterization, fieldwork description, and general conclusions (Chapters 1, 2, and 5, respectively). Furthermore, Chapters 3 and 4 are related to two scientific papers, which contain the results found during the two main steps of this research development. They are structured according to all topics needed in a scientific article, as abstract, introduction, study area, methods and materials for data acquisition, results, discussion, and conclusions. Just the references of them all are not inside each chapter, but are unified in a specific topic (References) to avoid repetitions.

CHAPTER 1: Introduction and objectives

In Chapter 1 the general introduction of this work is presented, to contextualise and to justify the themes covered in Chapters three and four. This chapter also treats the hypothesis that originates/motivates this study, objectives that should be matched, and its document structure.

CHAPTER 2: Study area and fieldwork

Chapter 2 presents the reservoir where this research was developed according to its watershed geomorphology, weather/climate structure, and historical human occupation. In addition, the fieldwork procedures, and the methods used to obtain the aims are presented.

CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)

Chapter 3 mainly reports the study of a supervised classification algorithm, called Support Vector Machines (SVM) used to obtain the FHR watershed LULC historical classification. In this section, there is an overview of FHR watershed delimitation procedures; the detailed satellite images preprocessing made for all orbital imagery; a complete analysis of SVM parameters; the evaluation of LULC classification and change detection; and the SVM performance review. As a result, an optimal parameterization of SVM was obtained, which allowed an accurate LULC historical classification and its related change detection analysis over a 15-years time series (1995-2010).

This paper was published in Model Earth Systems Environment (Springer), under the reference: Model. Earth Syst. Environ. (2016) 2:138 - DOI 10.1007/s40808-016-0190-y

CHAPTER 4: Spatiotemporal distribution of $\boldsymbol{a_{CDOM}}$ in a tropical eutrophic reservoir

Chapter 4 contains a brief discussion about bio-optical models to estimate IOPs from R_{rs}; a complete study of a bio-optical models set performance to retrieve a_{CDOM} using simulated reflectance to TM sensor; and the study of possible sources of CDOM variations in FHR. As a result, an empirical model was proposed to accurately estimate a_{CDOM} using TM/Landsat-5 data, and the possible sources of a_{CDOM} were linked to LULC and Chl-a, variating over time. This paper was not published yet.

CHAPTER 5: Final Conclusions

Chapter 5 is the last one, and concerns to general conclusions, based on the obtained results in the two scientific papers presented above.

CHAPTER 2: Study area and fieldwork

2.1 STUDY AREA

The study area is descried in the papers that compose this document (*CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)* and *CHAPTER 4: Spatiotemporal distribution of* a_{CDOM} *in a tropical eutrophic reservoir*), according to the characteristics that are important to the subject of each one. To avoid topic repetition, just the FHR geographic region and watershed historic occupation is described in this chapter.

The FHR is part of the Funil Hydropower system, located in the middle reaches of Paraíba do Sul River. This reservoir in the frontier between São Paulo and Rio de Janeiro states, Brazil (Figure 1), and has 6.2 billion m^3, and 40 km^2 of flooded areas in its maximum volume, employed for electric power production.

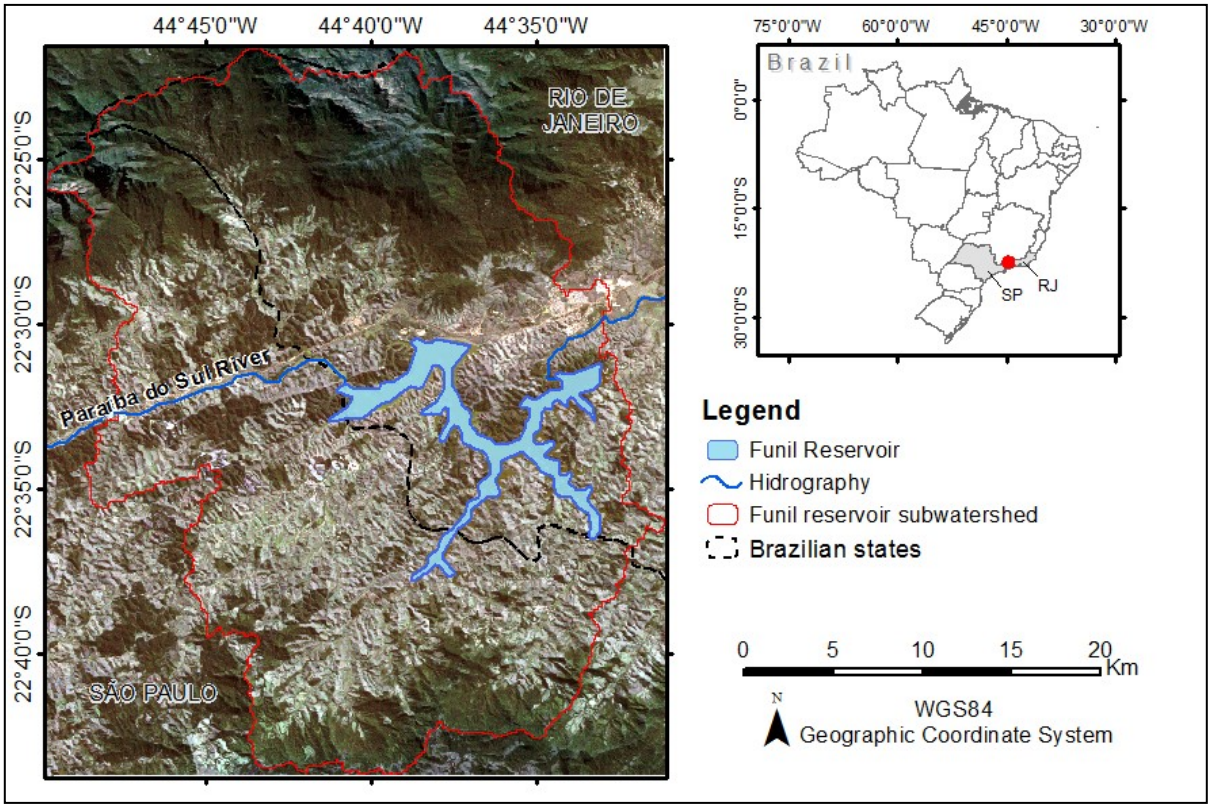

Figure 1. Funil reservoir sub-watershed location. TM/Landsat-5 surface reflectance images, R3G2B1 color composite. Date: 08/01/2010. Path/row: 218/76.

Historically, the human occupation in Paraiba do Sul River surrounding caused great changes in its water characteristics (flow, temperature, oxygen, etc.) and compounds (SPM, OM, phytoplankton concentration, etc.). These changes have been occurring since the 17th century, when the deforestation starts as a consequence of sugarcane plantations in São Paulo state and then followed by coffee plantations (18th and 19th centuries), pasture for farming activities (20th), and industry (21st). These last two human activities coexist until actual days and still affect the Vale do Paraiba do Sul water bodies in terms of sediments and nutrients availability (INEA 2012, 2016).

Actually, the original characteristics are almost completely modified by deforestation, vegetation substitution, river transpositions, damming, unstructured urban growth, and industrial pollution. Because of these alterations, the Paraiba do Sul watershed was segmented in zones according to the main environmental configurations. The FHR is a part of the Funil-Santa Cecília Zone, where an increase of water contamination by domestic effluents related to the absence of sewage treatment, landslides and inundation produced by the rising deforestation and irregular occupation in Permanent Protection Areas (PPA – established by law), and industrial effluent discharges in water is observed (INEA 2012).

Despite the referred substantial water degradation sources, the water quality is not completely committed and is considered as "acceptable level" according to the current legislation. However, Paraiba do Sul River water ecosystems were regarded as "not adequate" by the Environmental State Institute of Rio de Janeiro (INEA 2012, 2016).

Specifically in FHR surroundings, we can find the presence of livestock activity and urban areas, and the direct influence of the Vale do Paraiba industrial complex located upstream, next to São Paulo city (Branco et al. 2002). As a consequence of these contributions, modifications in FHR water compounds have been identified, promoting condition for cyanobacteria presence, eutrophication, and water quality degradation. This conjuncture preclude or restrict water utilization in general for this reservoir (INEA 2016).

It is important to highlight that the Vale do Paraiba industrial complex generates potentially toxic discharges in FHR, which can be accumulated there. This is possible since reservoirs act as barriers allowing depuration and sedimentation for these pollutants, avoiding its transportation downstream (INEA 2016). Additionally, the tendency of accumulation is also applicable to sediments and nutrients, which are drained by water confluence or leached from surroundings, making both potential sources of C for FHR.

Therefore, FHR is strongly affected by effluent discharges (domestic, industrial and agricultural) and leaching, which are maximized by the weather characteristics, and the natural

tendency for erosion. This environmental structure makes possible C transportation in all forms: dissolved or particulates, inorganic or organic, which are substantial for water quality conditions, depth reduction, generating social, ecological and economic damages.

2.2 MATERIAL AND METHODS

2.2.1 DATA COLLECTION

Both radiometric and absorption data were collected in 16 sample stations at FHR (Brazil) during April 2013. The orbital data was taken for the period between 1995 and 2010 at the same reservoir, that which is 218/76 (path/row) images.

2.2.1.1 IN SITU *RADIOMETRIC DATA*

In situ radiometric data were taken in FHR for all stations at the water surface (0+) in a spectral range between 320 and 950 nm, and spectral resolution of 3.3 nm. Two inter calibrated RAMSES spectrometers (TriOS GmbH, Germany) were used to measure the total water-leaving spectral radiance (L_t - RAMSES-ARC) and spectral downwelling irradiance above the surface (E_s - RAMSES-ACC) according to its acquisition geometry (Figure 2).

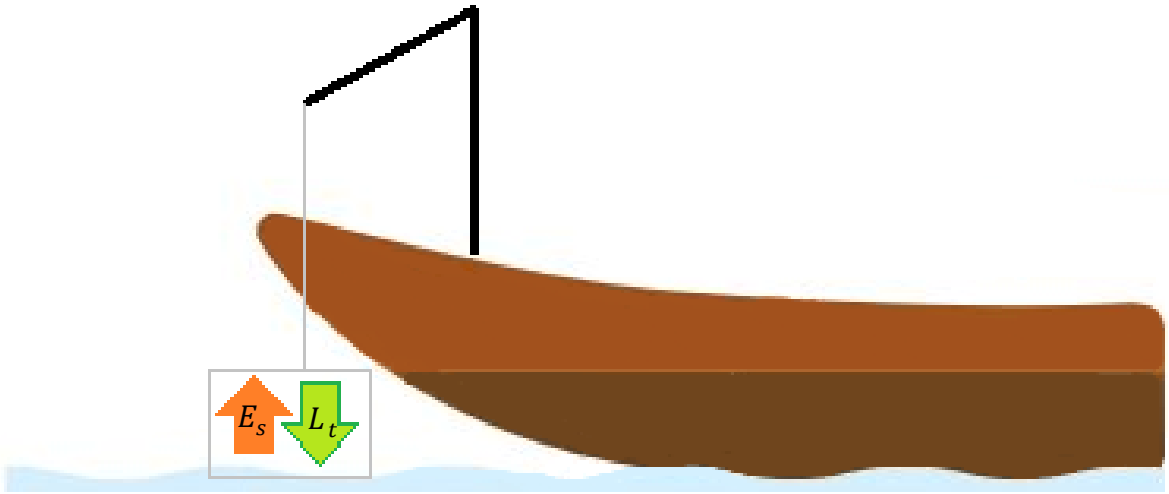

Figure 2. Acquisition geometry: RAMSES-ARC pointed to target intending to measure the total water-leaving spectral radiance ($L_t(\lambda)$), and RAMSES-ACC pointed to sky to measure the spectral downwelling irradiance ($E_s(\lambda)$).

The equipment receives the light that is conducted by an optical fiber to a polychromator, where there is color decomposition by a holographic grating, to detect light by photodiodes. In RAMSES-ARC, the optical fiber set is disposed next to the lenses with a 7° as field of view, and spray protection over them. In RAMSE-ACC, the light acquired by the optical fiber set passes through a plan diffuser, and the optical configuration was defined as cosine function for detection angle (cosine collector) (Figure 3). This procedure was made as far as possible from the boat to avoid the interference of its shadow or light beams. Additionally, at each sample

station, several measures of R_{rs} were taken and their average was used as the representative spectrum for each point of collection (Ogashawara et al. 2016).

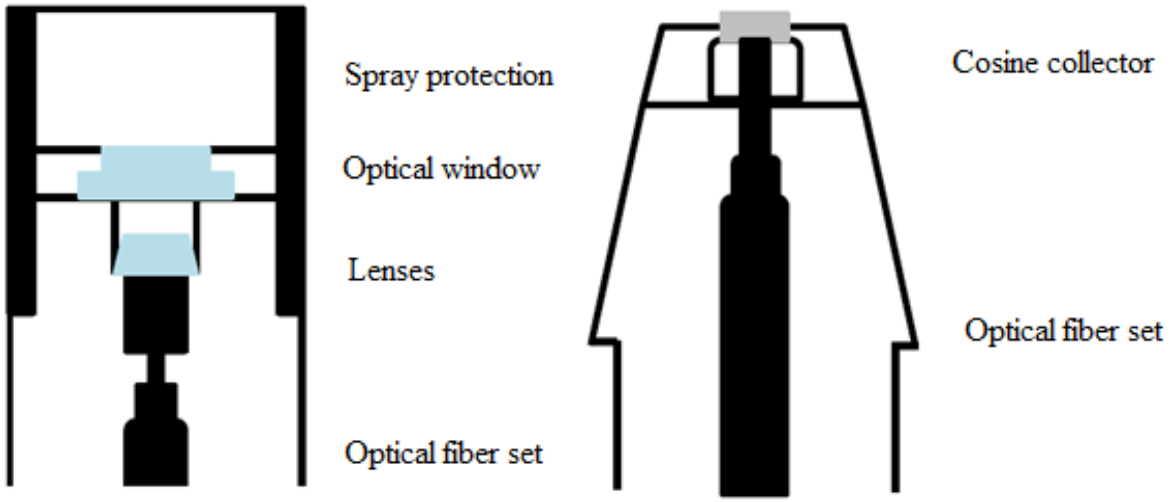

Figure 3. TriOS sensors schematic illustration: RAMSES-ARC (at left) and RAMSES-ACC (at right). Source: TriOS GmbH, Germany

TriOS operation is intermediated by a computer with MSDA_XE software, where parameters configuration and equipment operations are possible. This software allows us to define collection choice (above or below the water surface) and to monitor the acquired data in actual time.

Finally, the remote sensing reflectance data (R_{rs}, sr^{-1}) was obtained as Eq. 1 (Kirk 2011):

$$R_{rs}(0+) = \frac{L_t(\lambda,\theta,\phi)}{E_s(\lambda,\theta,\phi)} \quad (1)$$

where, L_t is the total water-leaving spectral radiance [W/m^2sr], E_s is the spectral downwelling irradiance above the water surface [W/m^2], λ is wavelength, θ is the zenithal angle, and ϕ is the azimuthal angle.

2.2.1.2 REFLECTANCE DATA SIMULATION

The *in situ* hyperspectral remote sensing reflectance measurements were used to simulate the related signal registration at the central wavelengths of TM/Landsat-5 bands, according to the specific spectral response function of each one (Figure 4, Table 1). Then, the resulting simulated remote sensing reflectance ($R_{rs_{simulated}}$) was calculated by the weighted sum of the hyperspectral field reflectance data into each bandwidth by a convolving procedure between the signal registered and the bands spectral function, as Eq. 2 (Kidder and Vonder Haar 1995; Schowengerdt 2006):

$$R_{rs_{simulated}} = \frac{\int_{x\,min}^{x\,max} R_{rs\,field}.Fr.dx}{\int_{x\,min}^{x\,max} Fr.dx} \quad (2)$$

where, $R_{rs_{field}}$ is the hyperspectral R_{rs}, Fr is the sensor function, dx is the wavelength range which defines the sensor spectral resolution.

Table 1. TM/Landsat-5 band configuration.

Spectral band	Wavelength (µm)	Central wavelength (µm)	Spatial resolution (m)
B1 – blue	0,45 – 0,52	0,485	30
B2 – green	0,52 – 0,60	0,560	30
B3 – red	0,63 – 0,69	0,660	30
B4 – near infrared	0,76 – 0,90	0,830	30
B5 – infrared short wave	1,55 – 1,75	1,650	30
B6 – infrared thermal	10,40 – 12,50	11,450	120
B7 – infrared short wave	2,08 – 2,35	2,215	30

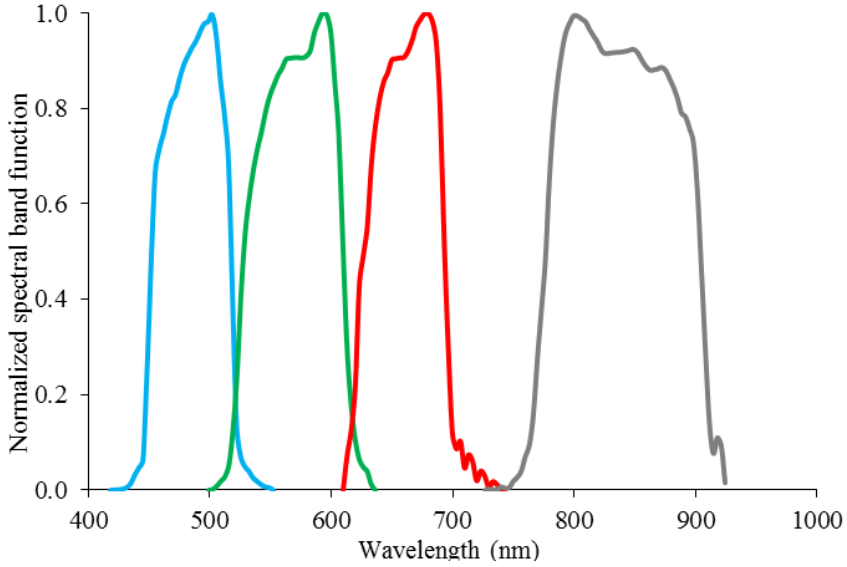

Figure 4. The TM/Landsat-5 spectral band function distribution: B1 (blue), B2 (green), B3 (red), and B4 (gray).

The radiometric data simulation is useful since it aligns the hyperspectral field data with the multispectral TM/Landsat-5 band configuration allowing the calibration of the bio-optical models according to the wavelength where the registered signal is maximized at the sensor. This procedure makes the model able to be applied in orbital data to obtain a_{CDOM}

2.2.1.3 $a_{CDOM}(\lambda)$ *DATA*

The water samples were taken in FHR for all stations, and filtered right after collection using a 0.2 µm nucleopore membrane. The $a_{CDOM}(\lambda)$ analysis was made using a Perkin Elmer lambda 35 UV/Vis system (UV-2600 - Perkin Elmer Inc, USA) with a transparent cuvette with 10 cm of linear length, and the Millipore Milli-Q water (Millipore Corporation, USA) was the reference of pure water. The optical density (OD) was measured in a range between 190 and 1100 nm, with 1 nm as spectral resolution. The absorbance correction was made according to

the average of the Milli-Q water absorbance baseline fluctuations at each wavelength (James L. Mueller 2000). The $a_{CDOM}(\lambda)$ values were obtained for a spectral range between 390 and 730 nm, with intervals of 1 nm, using Eq. 3 (Bricaud et al., 1981):

$$a_{CDOM} = \frac{2{,}3\ OD(\lambda)}{l} \qquad (3)$$

where, λ corresponds to wavelength, l is the cuvette linear length in meters.

It is important to highlight that the field data was collected in an algal bloom situation, which implies more a_{CDOM} correlation with Chl-a than with DOC (Danhiez et al. 2017; Kirchman et al. 1991; Le and Hu 2013; Rochelle-Newall and Fisher 2002; Sasaki et al. 2005). Despite that, the relationship between CDOM and DOM is still the same, both increase or decrease proportionally, and linearly (Brezonik et al. 2015; Danhiez et al. 2017; Rochelle-Newall and Fisher 2002). The data collection in bloom situation also allows calibrating a flexible model, which is capable to retrieve a_{CDOM} in bloom and no-bloom situation, since this collection take into account a wide range of CDOM samples.

2.2.1.4 ORBITAL DATA

The TM/Landsat-5 imagery was acquired from the United States Geological Survey (USGS - http://earthexplorer.usgs.gov/). This sensor was chosen due to its temporal, radiometric, and spatial resolutions that were considered sufficient for this study. In addition, FHR is limited to just one scene of this sensor coverage, which eliminates the impact of using different scenes to extract information from the same water target. Besides that, TM/Landsat-5 data acquisition costs were also taken into account.

As orbital dataset, all available images (Path/row: 218/76) without cloud cover above FHR since 1995 to 2010 were obtained, totalizing 64 images. To LULC analysis, we used the images taken in August with 5-years interval (1995, 2000, 2005, 2010), totalizing 4 images. For a_{CDOM} retrieving, no date restriction was imposed, and all images were used. This differentiation was needed since LULC presented a relatively stable behavior over time, and comparing images taken in different months can add errors linked to vegetation phenotypic stage and to sun relative position. Contrarily, water characteristics are dynamic, and can quickly change, thus it must be analysed using as much as possible of the available data (John R. 2011).

Intending to ensure the high-quality of surface information analyses, image preprocessing was needed, aiming to compare orbital data taken in different atmospheric and illumination conditions (Hadjimitsis and Clayton 2009).

2.2.2 IMAGE PREPROCESSING

Image preprocessing was needed to minimize the effects of the signal by atmospheric scattering and absorption as well as to reduce the differences attributed to the acquisition geometry and illumination conditions (Chander, Markham, and Helder 2009). Thus, to differentiate these interferences from the real target, radiometric signal is essential to have robust information from orbital data. This procedure is composed by radiometric calibration, atmospheric correction, and radiometric normalization, as detailed in *CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)*.

The image preprocessing procedure was conducted for all orbital data, varying just the image base according to the target characteristics. For LULC mapping, the oldest image in the data set was chosen (August/1995), since it conserves the original sensor configuration parameters defined before the launching, in order to guarantee high quality for land signal correction. For a_{CDOM} retrieving, we used the image acquired in the driest period (September/2001) to ensure low atmospheric influence as a way to avoid negative reflectance values at the water surface after atmospheric correction.

2.2.3 LAND COVER CLASSIFICATION EVALUATION AND CHANGE DETECTION ANALYSIS

The methods and materials used for LULC classification are completely described and discussed in *CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)* which contains the first article of this research.

2.2.4 BIO-OPTICAL MODELS FOR a_{CDOM} RETRIEVING

The materials and methods to study the bio-optical models set performance are completely described and discussed in *CHAPTER 4: Spatiotemporal distribution of* a_{CDOM} *in a tropical eutrophic reservoir* which contains the second article derived from this work.

2.2.5 MATERIALS AND METHODS SUMMARY

The Figure 5 presents the general flowchart of this research, in order to familiarize the reader with this work and steps dependence.

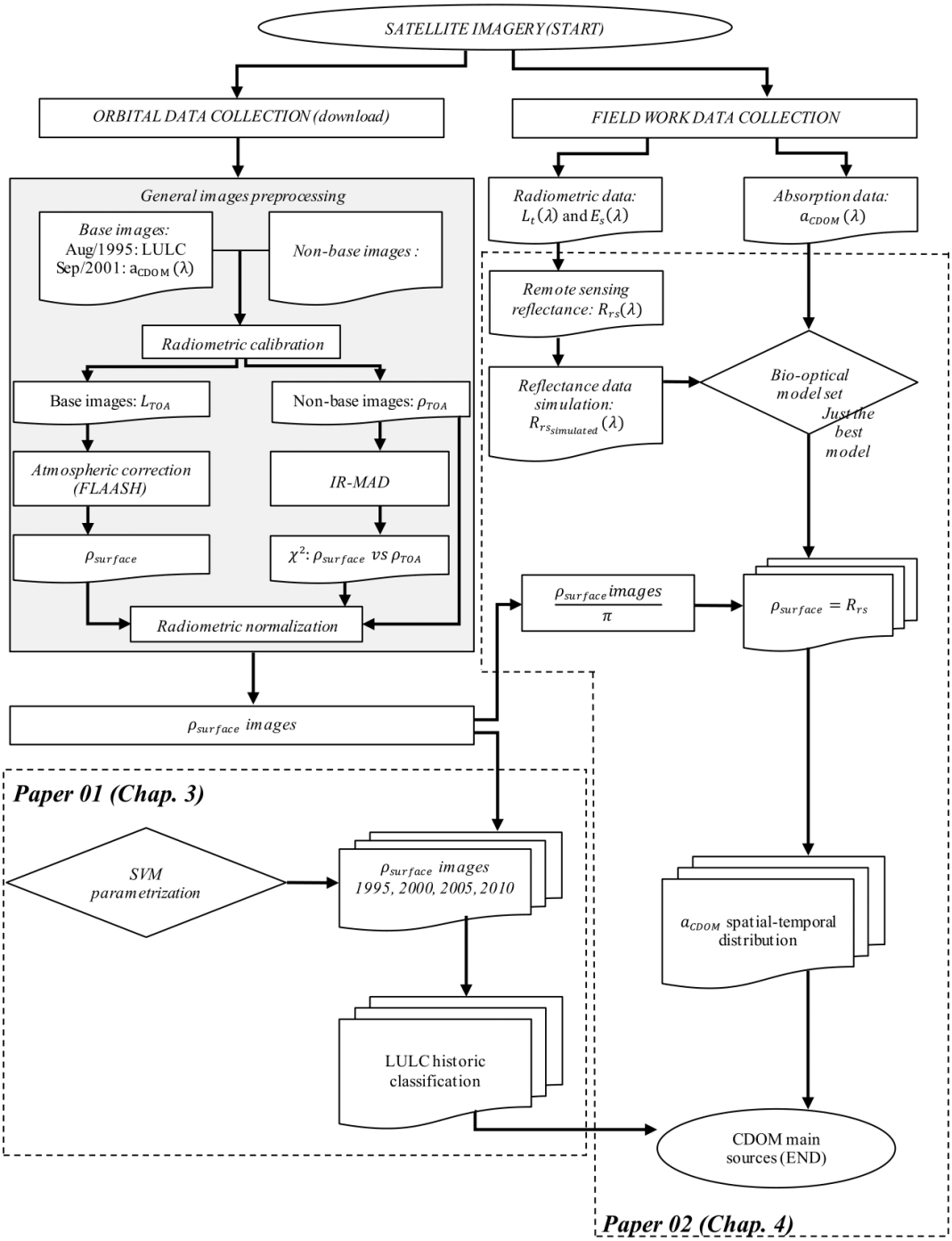

Figure 5. Materials and methods flowchart.

Where, a_{CDOM} is the absorption coefficient of colored dissolved organic matter, E_s is the spectral downwelling irradiance, L_t is total water-leaving spectral radiance, R_{rs} is the remote sensing reflectance above the surface, $R_{rs_{simulated}}$ is the simulated remote sensing reflectance above the surface, ρ_{TOA} is the reflectance top of atmosphere, L_{TOA} is the radiance top of atmosphere, $\rho_{surface}$ is the surface reflectance.

CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)

*Sarah MARTINS[1], Nariane BERNARDO[1], Igor OGASHAWARA[2], Enner ALCÂNTARA[1]

Reference: Model. Earth Syst. Environ. (2016) 2:138 - DOI 10.1007/s40808-016-0190-y

Abstract: Change detection in Land Use and Land Cover (LULC) using Support Vector Machines (SVM) to mapping a geographic area is a way that has been studded and improved because of its advantages as low costs in terms of computing, field research and staff team. To use SVM, it is needed firstly to define the most efficient function to be used (linear, polynomial, and radial base function - RBF) and secondly to establish the most appropriate input parameters of them, based on the study area, which was the main challenge of using SVM algorithm. The main goal of this work was to test the performance of polynomial function and RBF, and to identify which input parameters combination are the best to use SVM algorithm for Funil Hydroelectric Reservoir (FHR) sub-watershed LULC mapping, using TM/Landsat-5 time-series images. After several tests and based on the obtained results, the RBF was identified as the best SVM's function, which was used to classify the time-series images. The referred SVM function has the following parameters to be defined: the error tolerance (ξ or C), the pyramid depths (P), the radial basis function parameter (γ), and the threshold. The most effective combination of input parameters to RBF was C = 100; P=2, γ =0.1, threshold = 0.05. LULC change detection analyses demonstrate that the obtained SVM parameterization made the algorithm able to differentiate large and continuous classes, lengthy and thin areas, as borders, and not continuous small areas located inside wide classes, through the usage of effective, but small, training sample. The parameterization proposed for this work to FHR sub-watershed area resulted in great statistics classification with the overall's accuracy among 86% and 98% over the time-series, the producer's accuracy of 90%, the user's accuracy higher than 86%, and the Kappa statistics ranged from 86% to 91%.

Keywords: SVM parameters, change detection, TM/Landsat-5, automatic classification.

[1] Department of Cartography, São Paulo State University, Presidente Prudente, SP, Brazil
[2] Department of Earth Sciences, Indiana University-Purdue University Indianapolis, Indianapolis, IN, USA

3.1 INTRODUCTION

To planning and management the natural resources in a sustainable way, detailed knowledge about Land Use and Land Cover (LULC) and its changes detection is considered as a key parameter to evaluate the environment progress over time (Usman et al. 2015; Verburg, Neumann, and Nol 2011). In this context, LULC encompass two separate terminologies which are often used interchangeably: land cover means the observed biophysical cover of the Earth's surface, and land use is the human use of land resources (Briassoulis 2000; Dimyati et al. 1996). The LULC thematic mapping from remote sensing data is commonly based on image classification techniques and the choice of an appropriate method is considered crucial for reliable mapping (Lu and Weng 2007).

Classifying remote sensing images into thematic maps still a challenge mainly due to selected image data, complexity of the landscape, image processing and classification approaches (Lu and Weng 2007). Numerous classification algorithms have been developed since the first TM/Landsat-1 image was acquired in 1972 (Lu and Weng 2007; Tewkesbury et al. 2015) and among the most popular algorithms available we can highlight the maximum likelihood classifier (MLC) (Huang, Davis, and Townshend 2002). This classifier, MLC, is a parametric method that assumes a normal distribution for the data set, what can be invalid in some cases, as to classes consisting of several subclasses or to classes with different spectral features (Taskin and Selcuk 2008). To overcome the referred kind of limitation, non-parametric classification techniques have been introduced, such as the Support Vector Machines (SVM). SVMs are based on the statistical learning theory proposed by (Vapnik 1995).

SVMs are particularly suitable for remote sensing field because this method has the ability to handle successfully with small training datasets, producing higher classification accuracy than the traditional methods. SVMs minimize classification error on unseen data without prior assumptions made on the probability distribution of the data (Mantero, Moser, and Serpico 2005). The main challenge in using the SVM method to classify satellite images is to select the most appropriate kernel function type (Radial Basis Function - RBF, Linear, Polynomial or Sigmoid) and its parameters, which influences SVM performance. The RBF and the polynomial kernel are commonly used in the literature for the remotely sensed images classification (Pal and Mather 2005). The SVMs need user-defined parameters and each parameter has different impact on kernels function performance. Hence, the classification accuracy of SVMs is based upon the choice of the parameters and kernels (Ustuner, Sanli, and Dixon 2015).

In order to produce a better LULC mapping, the main goal of this work was to assess which combination of SVM's function and parameterization produce the most efficient classification, and consequently, represents the best combination to map LULC in the Funil Hydroelectric Reservoir (FHR) sub-watershed. The analysis ranged from 1995 to 2010 allowing assessing the change detection to provide useful information to explain the water quality in the reservoir.

3.2 MATERIALS AND METHODS

3.2.1 STUDY AREA

The FHR is part of the Funil Hydropower system, located in the middle reaches of Paraíba do Sul River, at the frontier between São Paulo and Rio de Janeiro states, Brazil (Figure 1). The FHR has 6.2 billion m^3, and 40 km^2 of flooded areas in its maximum volume, employed for electric power production, generating average power about 216 MW, since 1969). Moreover, at upstream there is a large industrial complex from the region called "Vale do Paraíba", at east of São Paulo Metropolitan Region (Branco et al. 2002).

3.2.2 FUNIL SUB WATERSHED DELIMITATION

The FHR sub-watershed delimitation was conducted using a 90m digital elevation model (DEM) produced by Shuttle Radar Topography Mission (SRTM) data (Rabus et al. 2003), available from Consortium for Spatial Information (CGIAR-CSI - http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp). To process this data, the ArcHydro Tool available on ArcGIS-10 was used (ESRI, 2015).

Sub-watershed delimitation was based on filling DEM's sinks to rectify the original data, and also used to define flow directions and accumulation, which were used to estimate the drainage. The estimated drainage localization (streams) was defined based on slope and altitude characteristics. In addition to these data, relief characteristics (water divisor) and dam spot (bath point) were also computed to define two sub-watersheds: one for Funil reservoir and another at the reservoir immediately upstream (Walsh et al. 2015). Therefore, the FHR' sub-watershed was clipped from the images (Figure 6).

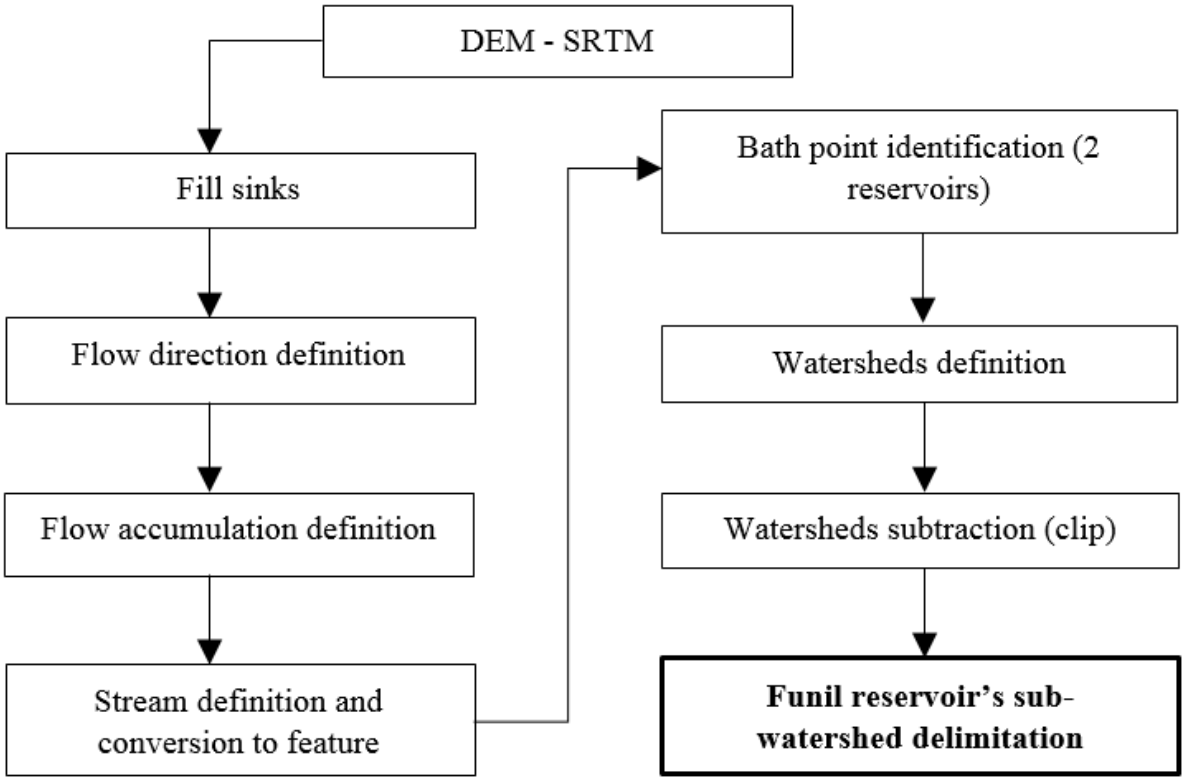

Figure 6. FHR sub-watershed delimitation flowchart.

3.2.3 SATELLITE IMAGES PREPROCESSING

The TM/Landsat-5 images used in this work were acquired from the United States Geological Survey (USGS - http://earthexplorer.usgs.gov/). Images from 1995 to 2010, set in regular intervals of five years, were selected to apply SVM algorithm for LULC classification (see Table 2 for LULC class definition). TM/Landsat-5 was chosen because of its long and continuous time series of data with high quality and adequate spectral and spatial resolutions for the purposes of this work.

Table 2. TM/Landsat-5 images characteristics.

Image date	Path/ row	Solar zenith angle	Solar azimuth angle
08/26/1995	218/076	56.019	57.195
08/21/2000	218/076	50.241	49.712
08/03/2005	218/076	53.031	43.287
08/01/2010	218/076	53.072	42.244

To ensure the high-quality of the surface information analyses it is necessary to apply atmospheric correction and normalization procedures to the images (Hadjimitsis and Clayton 2009). These procedures were made to minimize the effects on the signal by atmospheric scattering and absorption as well as differences attributed to the geometry acquisition and illumination conditions. Thus, it is essential to differentiate these interferences from the real radiometric signal from the targets. Therefore, radiometric calibration based on sensor

parameters and atmospheric correction of the images can be used to obtain the surface reflectance (Schroeder et al. 2006). The atmospheric correction was conducted using Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) application; while the radiometric normalization process was made by the Iteratively Reweighted Multivariate Alteration Detection - IR-MAD (Canty, Nielsen, and Schimidt 2004).

FLAASH is an algorithm that uses the atmospheric compounds to establish regional models in proposing to compensate atmospheric effects in satellite spectral images by using MODerate spectral resolution atmospheric TRANsmittance version 4 (MODTRAN4) (Adler-Golden et al. 1999). This algorithm was applied to convert the 1995 image to surface reflectance. This image was chosen as reference because it is the oldest in the image data set; therefore it has the closest characteristics from those originally defined for the sensor before its launching. In addition to sensor configuration parameters FLAASH set was: rural area model to estimate aerosol contribution, tropical as atmospheric model, and Kaufman over water as multispectral settings.

IR-MAD is an algorithm used to compute the radiometric normalization procedure, which has as main objective to find the pseudo invariant object in a bi-temporal images comparison. This procedure uses canonical components through sequences of MAD (Multivariate Alteration Detection) transformations until it reaches the stop criterion (Canty 2009; Canty and Nielsen 2008; Canty, Nielsen, and Schimidt 2004). IR-MAD was performed with surface reflectance image as a base data (1995 image), and relates it with an at-sensor reflectance data (2000, 2005 and 2020 images). The stop criteria were set as 30 interactions or no significant change in canonical correlations.

To conduct IR-MAD application, it is necessary to have a previous registration of all images to ensure that all of them have a correspondence in their pixels localization, which will be essential for the change detection analyses. The maximum mean square error premised to ensure spatial correspondence among the image pixels for this registration process was 5%. Finally, the time-series images radiometric normalization was processed based on 1995' surface reflectance image (FLAASH output), and statistical parameters (qui-square – IR-MAD output), generating surface reflectance images for all years as a result. Figure 7 presents a flowchart of the steps of the preprocessing used in this work.

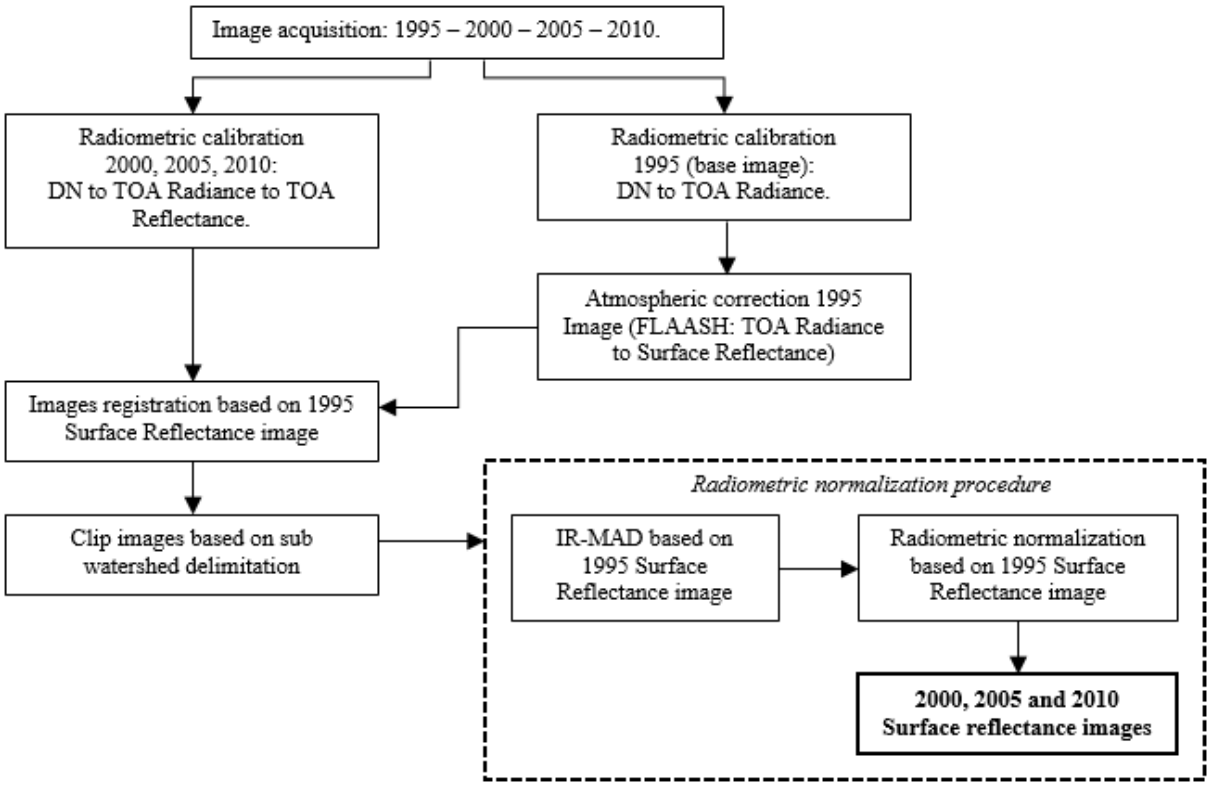

Figure 7. Image preprocessing flowchart.
where, DN is digital number, and TOA is top of atmosphere.

3.2.4 SUPERVISED CLASSIFICATION

The SVM main framework is the structural risk minimization (SRM) which can be achieved by two pathways: i) by fixing the empirical risk to obtain the smallest confidence interval to reduce the errors; ii) by finding an optimal hyperplane which maximizes the separation among the closest data of different classes. The data which must to be considered the class boundary corresponds to the support vector, and could incorporate some error tolerance (ξ or C), like a soft margin for the information class (Canty 2009; Huang, Davis, and Townshend 2002; Kavzoglu and Colkesen 2009; Mountrakis, Im, and Ogole 2011).

SVM works considering each pixel as vector in a multidimensional system, where each class have a pixel as support vector to establish the class boundary. To perform the classification, it is necessary a training stage to allow the machine to learn the process by using examples which indicate the correspondence between a training sample and a class. The class has an unknown probability distribution which is compared to a training sample and, based on this comparison, the support vectors of each classes are defined. After the algorithm training stage, the machine is able to classify a non-training sample using the support vectors parameters. The training data are useful to define the support vector for each classes' boundaries, making statistical data

unnecessary in this supervised classification (Canty 2009; Huang, Davis, and Townshend 2002; Kavzoglu and Colkesen 2009; Mountrakis, Im, and Ogole 2011).

SVM was originally projected to perform separation between two linear classes and was adapted to work for multi-classes with nonlinear separation by the use of the kernelization method (Canty 2009; Mountrakis, Im, and Ogole 2011). For non-linear classes, there are four SVM's functions available: linear, quadratic, polynomial and radial basis function. According to (Kavzoglu and Colkesen 2009) the last two functions are more efficient, presenting better performances than the others. However, it is still necessary to select one of them and the most suitable parameters for the chosen function to ensure a good performance of the algorithm. Because of that, we evaluated the best function and parameterization to be used on FHR sub-watershed.

Using the sample configuration described above, two tests were performed: i) the first to decide what function must be used; and ii) the second to define the parameters to be used in the previously selected function. The 2010 surface reflectance image was chosen to perform the tests, because this is the image where LULC areas are discontinued, wich means classes great fragmented. This fragmentation is ideal to evaluate the capability of the algorithm by observing if the algorithm could identify all of fragments. The information classes were based on the land cover classification system proposed by Food and Agriculture Organization (FAO - FAO 1994), see Table 3.

To perform the tests, the information classes present in the study area were identified and one representative sample (area) of each one was taken in order to contain "pure" pixels. The training samples quantity, size, and localization were based on the premise stated by Mountrakis, Im, and Ogole (2011), who suggest that SVM algorithm is less sensitive to sample size and could be able to obtain successful results using few samples.

The first test was based on Huang, Davis, and Townshend (2002) findings which affirmed that the best results for SVM could be obtained by using the polynomial and radial basis kernel functions. These types of functions improve the results by defining complex decision boundaries among the information classes. It occurs because of the data transformation from non-linear to linear boundaries, which is useful to multi-dimensional spatial analysis (Huang, Davis, and Townshend 2002). Additionally, these authors also presented the accuracy of the processing according to polynomial order (p) and Radial Basis Function (RBF) parameter (γ); and affirmed that the selection of parameters affects the final results.

To select the best parameters to be used in the RBF, a second evaluation was performed. For this type of kernel function, besides γ value, defined in the first test, the available parameters

are: the soft margin penalization (*C*), corresponding to the tolerance to classification errors; the pyramid depths (*P*), that speed up the classification; and threshold value, which limits the perceptual of data that will be inserted in each information classes. The results of this second evaluation defined the parameters to be applied in all surface reflectance images in the multi-temporal series to obtain the LULC classification of study area using the SVM.

Table 3. Information classes used in the SVM classification, corresponding to the land cover identified in study area with their respective descriptions (adapted from (FAO 1994)).

Information classes	Description
Natural and Semi-Natural Vegetation (NSNV)	Natural vegetated areas are defined as areas where the vegetative cover is in balance with the abiotic and biotic forces of its biotope. Semi-natural vegetation is defined as vegetation not planted by humans but influenced by human actions. These anthropogenic actions may result from grazing; possibly overgrazing the natural phytocenoses, or else from practices such as selective logging in a natural forest whereby the floristic composition has been changed. Previously cultivated areas which have been abandoned and where vegetation is regenerating are also included in this class.
Terrestrial Primarily Non-Vegetated Areas (TPNVA)	The cover is influenced by the edaphic substratum(including bare soils).
Artificial Waterbodies (AW)	This class applies to areas those are covered by water due to the construction of artefacts such as reservoirs, canals, artificial lakes, etc. Thus, without these artefacts, the area would not be covered by water.
Cultivated and Managed Terrestrial Areas (CMTA)	This class refers to areas where the natural vegetation has been removed or modified and replaced by other types of vegetative cover of anthropogenic origin. This vegetation is artificial and requires human activities to maintain it in the long term. Between human activities, or before starting cropping, the surface can be temporarily without vegetative cover.
Artificial Surfaces and Associated Areas (ASAA)	This class describes areas that have an artificially covered by human activities such as construction (cities, towns, transportation), extraction (open mines and quarries) or waste disposal.

2.3.5 LAND COVER CLASSIFICATION EVALUATION AND CHANGE DETECTION ANALYSIS

Accuracy and Kappa statistic (Congalton 1991) were used to evaluate the SVM performance to land cover classification by a comparison between the classified data and regions of interest, obtained by visual analysis for each image in time series. In the first test, one area was selected for each individual class. For the second one, it was selected a data set for each class, aiming to generate a robust output data. For both tests, two regions of interest sets were constructed, one to calibrate and other to evaluate the algorithm performance, both obtained by visual analsis and taken before the classification step.

Accuracy indicates the classification quality by relating the reference and the classified data in terms of the global value or in terms of its categories using the confusion matrix. Therefore, the classification result was evaluated by comparing the representativeness of the training pixels to the ones in each category (Lillesand, Kiefer, and Chipman 1999). While the global error is estimated by the overall accuracy; the categories errors are estimated by the producer's and the user's accuracies (Congalton 1991).

Producer's accuracy measures omissions errors, relating actual correctly classified pixels in a category with the amount of pixels classified in the same category on training step. User's accuracy indicates the probability of some classified pixels really correspond to the same category in the reference data (regions of interest), relating the actual correct classification pixels in a category with the same class amount of pixels classified on the actual classification (Congalton 1991).

3.3 RESULTS AND DISCUSSION

The results of the first test to find the best parameters combination function to classify the study area with the lowest error as possible are show in Figure 8. The use of RBF function with γ values varying from 0.1, 0.3 and 0.5 showed no changes for accuracy (Figure 8a). The use of polynomial kernel function with orders of 3, 6 and 9 showed that the best order was 6 (Figure 8b).

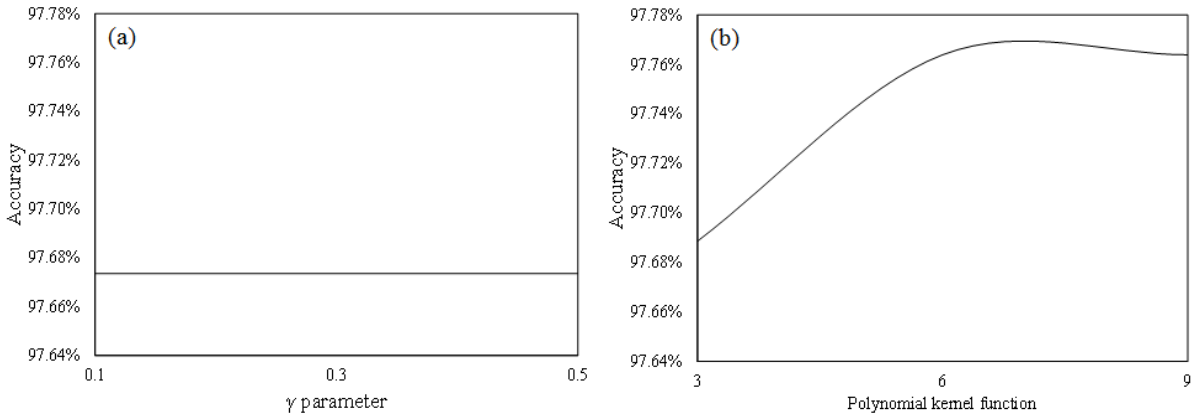

Figure 8. Overall accuracy (%) resultant of tests between RBF function with γ values 0.1, 0.3 and 0.5 (a), and polynomial kernel function with orders 3, 6 and 9 (b).

Since there is no significant difference between these two functions, the lowest computational effort was chosen as the decision criteria to select the function to be used. Thus, the RBF kernel was selected to be performed in this work with $\gamma = 0.1$. The Table 4 shows the overall accuracy for the tested functions and its parameters.

Table 4. Kappa and overall accuracy values obtained for both kernel functions according to their tests values.

Radial basis kernel function		
γ	**Kappa**	**Overall accuracy %**
0.1	0.9589	97.6735%
0.3	0.9589	97.6735%
0.5	0.9589	97.6735%
Polynomial kernel function		
p	**Kappa**	**Overall accuracy %**
3	0.9592	97.6886%
6	0.9605	97.7639%
9	0.9605	97.7639%

The best parameters values for the chosen function (RBF) were defined in the second evaluation. Besides the γ value, the parameters available to change for this type of function were: margin penalization (*C*), pyramid depths (*P*), and threshold value. Combinations among these parameters in pairs were conducted and the best results, in terms of overall accuracy and Kappa statistic, were chosen to perform SVM classifications (Table 5).

Table 5. Kappa and accuracy values obtained in RBF parameters tests to define the values that will be used to apply SVM algorithm.

γ	***C***	***P***	**Threshold**	**Kappa**	**Overal Accuracy (%)**
0.1	120	0	0.05	0.9727	98.0874
	120	0	0.00	0.9727	98.0874
	100	0	0.05	0.9727	98.0838
	100	0	0.00	0.9727	98.0838
	120	2	0.05	0.9844	98.9071
	100	2	0.00	0.9873	99.1075
	100	2	0.05	0.9873	99.1075
	120	2	0.00	0.9878	99.1403

The obtained results show that Kappa and overall accuracy are weakly influenced by variables *C* and the threshold. The *C* variable's influence is probably related to a less restrictive classification with the increase of its value, allowing a larger distance between the optimal hyperplane and the misclassified data. The threshold's influence is presumably related to the largest percent of acceptance for the classes when its value decreases, resulting in a more restrictive classification for this value incensement. Based on these analyses, the variables values selected were the second higher Kappa and overall accuracy (highlighted line in table 4). The selection for the second higher values was due to the fact that it is more restrictive than the highest one and the results were similar.

The results obtained indicated that the best BRF γ value was 0.1, because of the lack of significant improvement in the accuracy with changes in its value for the same function.

Therefore, there was no relevant improvement in the Kappa and overall accuracy in a comparison between this RBF configuration and the best polynomial function parameter (Kappa difference = 0.0016 and accuracy difference = 0.0904%). Therefore, the RBF with a γ value = 0.1 was chosen due to the lowest computational cost. The other parameters defined for the RBF were: C = 100, because it was the most restrictive value with highest accuracy; Threshold = 0.05 (acceptance = 95%), because it improved Kappa and overall accuracy values, since it is more restrictive; and P = 2, to improve the computational performance.

Once defined the SVM parameters, the classification was performed and a confusion matrix was generated for each image classification. The overall accuracy (Table 6) showed a good performance of SVM classification, which was equal to 98% for 1995, 2005 and 2010 images and equal to 86% for 2000 image. These results indicate that the training sample had premised low errors in all classification categories relative to the image's pixels amount.

Table 6. Kappa and overall accuracy values obtained for SVM classification per year.

Year	Overall accuracy	Kappa
1995	0.98	0.96
2000	0.86	0.81
2005	0.98	0.97
2010	0.98	0.96

Producer's accuracy demonstrates that the SVM had a good performance to NSNV, CMTA and ASAA categories for all images, maintaining more than 90% of accuracy for all images classification in the referred categories (Table 7).

Table 7. Product and user's accuracy of SVM classification based on confusion matrix.

Year	Producer's accuracy (%)					User's accuracy (%)				
	NSNV	TPNVA	AW	CMTA	ASAA	NSNV	TPNVA	AW	CMTA	ASAA
1995	99.40	91.32	98.99	93.15	96.30	99.64	91.92	98.43	96.06	96.18
2000	91.93	95.47	34.48	99.44	96.84	93.26	76.85	99.49	85.96	96.81
2005	96.62	99.2	82.96	97.05	99.79	97.81	98.34	92.42	97.02	99.87
2010	99.98	76.96	99.08	93.67	95.89	99.93	86.66	97.77	96.89	97.51

The poorest performance was observed in the AW category for the 2000 image with a 34.48% of accuracy, and for TPNVA category in the 2010 image with a 76.96% of accuracy. However, for the other images the same categories had satisfactory performances, with accuracy values higher than 80% for AW and 90% for TPNVA.

The poorest performance for AW category in the 2000 image classification is probably related to the low water level period in the entire system, river and reservoir, which made it difficult to collect samples from this category without any spectral interference of river-margin vegetation, margin soil, and areas with riverbed exposition, especially within the river's channel. The

classifications errors for TPNVA category in the 2010 image could be attributed to the presence of underbrush vegetation in contact to areas between TPNVA and NSNV, making it difficult to collect samples without spectral influences of the vegetation. For user's accuracy, all results were satisfactory, showing for most of the classes more than 90% of accuracy, the same was observed for the overall classification (Table 7).

The two, local and by category accuracies, corroborate to the overall accuracy results indicating the high consistence of SVM parameterization proposed. The complementary analysis of these classification results is the Kappa statistics that can variates from 1, representing total agreement, to 0 or negative numbers, meaning no agreement (Lillesand, Kiefer, and Chipman 1999). Kappa statistics results confirmed the high agreement between the training and actual classifications, indicating that some other random classification may obtain close results with this work classification. Kappa values were 0.81 for the 2000 image, 0.96 for the 1995 and 2010 images, and 0.97 for the 2005 image (Table 6).

For LULC mapping analysis, globally, it could be observed that the SVM parameterization made the algorithm sensible for variations not only large periods, as 1995 to 2010 (Table 8), also for small change, as year-by-year analysis (Table 9, Table 10).

Table 8. LULC change from first to last year (1995 – 2010).

Class		1995 (%)				
		NSNTV	**TPNVA**	**AW**	**CMTA**	**ASAA**
2010 (%)	**NSNTV**	88.07	21.48	0.08	9.08	0.83
	TPNVA	2.03	58.46	0.01	0.71	0.38
	AW	0.02	0.01	93.79	0.55	12.29
	CMTA	9.89	20.02	5.42	89.31	57.11
	ASAA	0.01	0.03	0.70	0.36	29.39
Image Difference		**1.42**	**-25.50**	**11.89**	**3.70**	**-54.99**

As a SVM application result, the LULC changes between the time-series images can be accessed in Table 8. Through these results, it is possible to observe small changes in LULC, around 3% or lower (1995-2000: CMTA class = 2%; 1995–2010: NSNTV = 1.42 and CMTA = 3.70%, for example).

Table 9. LULC classification areas for each class.

Year	LULC (km²)				
	NSNV	**TPNVA**	**AW**	**CMTA**	**ASAA**
1995	328.68	49.44	28.20	433.16	12.80
2000	312.26	60.13	26.55	441.74	11.60
2005	291.52	52.95	26.32	470.20	11.28
2010	338.80	39.24	32.09	436.64	5.51

Table 10. LULC change year by year.

Comparison period	LULC change (%)				
	NSNV	TPNVA	AW	CMTA	ASAA
1995 - 2000	-5%	22%	-6%	2%	-9%
2000 - 2005	-7%	-12%	-1%	6%	-3%
2005 - 2010	16%	-26%	22%	-7%	-51%

Observing the Figure 9, is it possible to see that the obtained parameterization enables the SVM algorithm to differentiate the large and continuous classes, as well as lengthy and thin areas, as borders, and not continuous small areas located inside wide classes.

The conserved TPNVA area was around 58%, 21.48 % of this area was converted into NSNTV, probably because of the regeneration in the borders. Around 20% of this area was converted in CMTA. Water cover (AW), was the class with the lowest change, keeping 93.79% of its original area; 5.42% were changed to CMTA, probably because of the water level reduction in the river channel, exposing fertile areas which could be used for agricultural and livestock activities.

Agricultural and livestock class (CMTA) had maintained 89.31% of its original area, and was mostly changed to natural areas (NSNTV - 9.08%). The anthropic areas (ASAA), kept only 29.39% of its original area, becoming the lowest maintenance rate among all classes. The major change in ASAA class occurred in the conversion to CMTA, 57.11%, indicating the reduction of exposed soil in rural areas, and 12.29% of ASAA class were converted in water area (AW), probably because of the reservoir water level growth, recovering the margin area which was characterized by exposed soil. In general, all percentage lower than 1% probably is associated with classification error real but unrepresentative land cover changes.

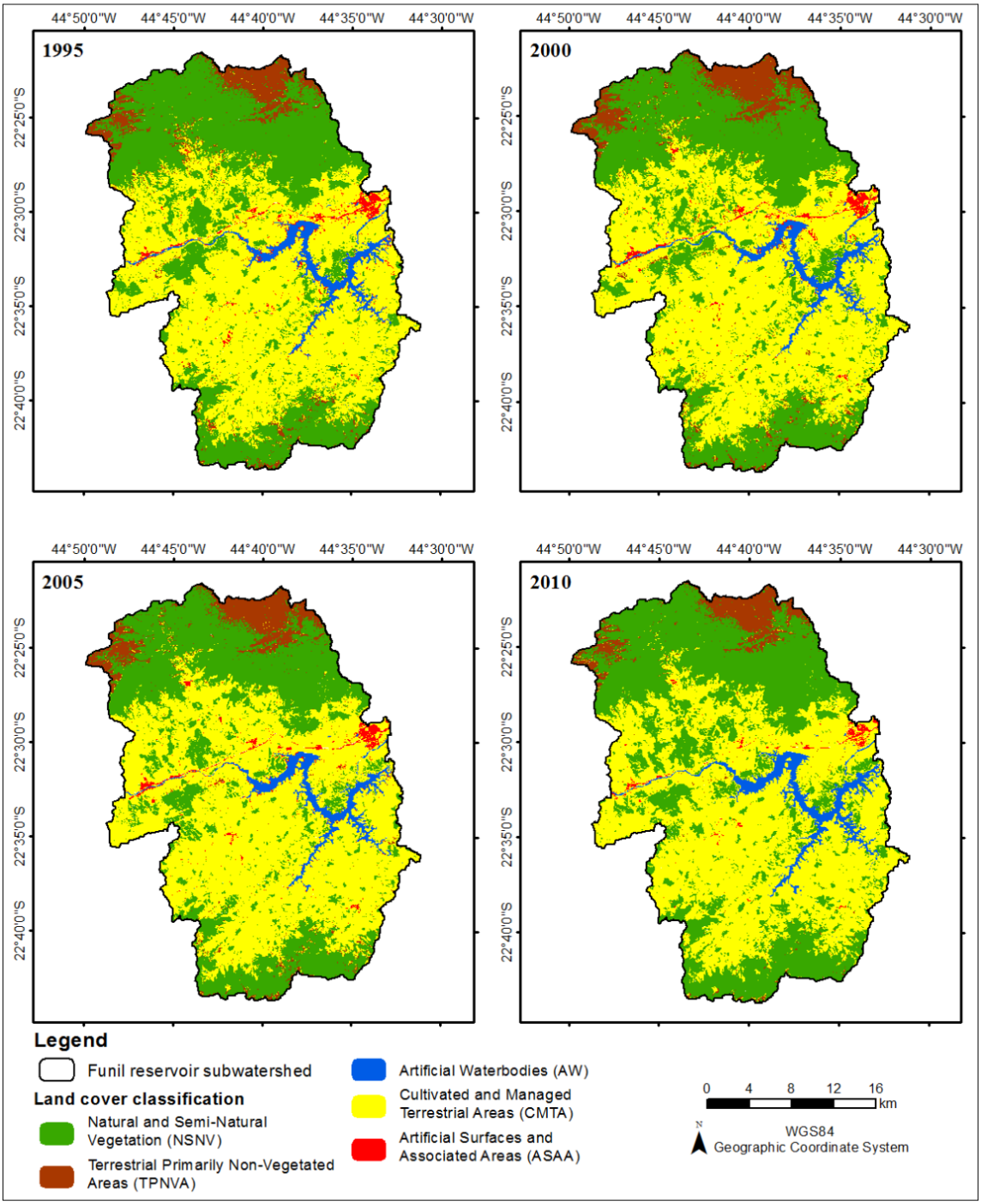

Figure 9. LULC classification maps per year.

3.4 CONCLUSION

The tested SVM's parameters ensured the optimal parameterization of algorithm and allowed mapping the main changes in FHR, including the very small ones. Results demonstrated that the degree of polynomial function modifies the accuracy, and RBF showed best results when compared to polynomial function. Input parameters of RBF were also tested and shown that the highest regularization parameter (C = 120) did not retrieve the higher accuracy. For this study area, the higher accuracy was obtained using a lower C-value (=100), lower basis function parameter ($\gamma = 0.1$), to reduce the computational cost, and the pyramid depths (=2) to improve the performance of algorithm.

The change detection analyses demonstrate that the obtained SVM parameterization allows a great differentiation between the trained classes, supporting the premise of low sensitivity to sample size and high sensitivity to representative training sample selection. In addition to differentiation between the large and continuous classes, the parameterization obtained in this work made the algorithm able to differentiate lengthy and thin areas, as borders, and not continuous small areas located inside wide classes.

The LULC mapping performed using the obtained SVM function and parameterization, for this multi-temporal series was efficient and provided trustful and real results, ensuring high values of accuracy, both overall (minimum 90%), producer's (90%) and user's (86%), and Kappa statistics (86% and 91% over time), allowing high level of confidence to LULC mapping obtained.

CHAPTER 4: Spatiotemporal distribution of a_{CDOM} in a tropical eutrophic reservoir

*Sarah MARTINS[1,2], Karem CHOKMANI[4], Enner ALCÂNTARA[3], Igor OGASHAWARA[4], Anas EL ALEM[4]

Reference: not published yet

Abstract: The dissolved organic matter (DOM) is a water compound that can be associated with humic (terrestrial discharges – allochthonous origin) or fulvic (river input - autochthonous origin) acids. Thus, its correlation with dissolved organic carbon (DOC) or chlorophyll-*a* (Chl-*a*), depends on the environment dynamics and its related changes. Based on this, DOM can be an indicative of water quality. Since colored dissolved organic matter (CDOM) is the colored fraction of DOM, it can be used as a proxy for DOM historical behavior. It is possible by applying a bio-optical model in satellite imagery as a way to remotely obtain data from water surfaces. The aims of this research are: i) to evaluate the performance of a set of bio-optical models composed by empirical models and quasi-analytical algorithm (QAA) using simulated reflectance data for the thematic mapper sensor (TM), and ii) to analyse the possible sources of CDOM/DOM in a tropical eutrophic reservoir for a 15-years long TM/Landsat-5 time series. Results showed that the literature-based models failed in retrieving a_{CDOM} from simulated reflectance data. However, an alternative wavelength (485 nm) and a band ratio (B4/B1) obtained the best relation for a_{CDOM} estimation after correlation and collinearity analyses (2D color correlation plot). The proposed model presented satisfactory performance ($R^2 = 0.91$, p-value < 0.0001, RMSE = 7%, Nash = 0.91), and it could also identify even small variations in reflectance values from orbital data, as well as differentiate even slight alterations in a_{CDOM}. Besides that, two significantly different a_{CDOM} behaviors were identified for the Funil hydroelectric reservoir (FHR): one associated to land cover land use and rainfall/runoff occurrence, and other correlated to Chl-*a* in algal blooms situations.

Keywords: Bio-optical model; QAA; empirical model; a_{CDOM}; Chlorophyll-a (Chl-a); land cover land use (LULC); TM/Landsat-5 imagery.

[1] São Paulo State University (UNESP), Department of Cartography –Presidente Prudente (SP), Brazil.

[2] National Institute of Scientific Research (INRS), Eau Terre Environment Research Centre (Research Centre on Water, Earth, and the Environment) – Québec (QC), Canada.

[3] São Paulo State University (UNESP), Department of Environmental Engineering –São José dos Campos (SP), Brazil.

[4] Indiana University - Purdue University Indianapolis (IUPUI), Department of Earth Sciences – Indianapolis (IN), USA

4.1 INTRODUCTION

Reservoirs are useful to study environmental changes because of their sensitivity and fast response to physical, chemical and biological alterations (Carpenter et al. 2007; Curtarelli et al. 2014). Changes in the reservoirs natural structure caused by internal or external sources may affect the phytoplankton mass and chlorophyll-a (Chl-a) concentration, the balance among suspended and dissolved organic/inorganic carbon, the organic matter distribution, the dissolved oxygen, the nutrient availability, etc.

The colored dissolved organic matter (CDOM) is the colored fraction of the dissolved organic matter (DOM). Likewise, dissolved organic carbon (DOC) is a great part of DOM in inland waters. Thus, CDOM should be useful as an indicator of the total DOM or DOC (Esteves 1998; Zhu et al. 2014). Moreover, the DOM/DOC content of inland waters can be associated with: i) external dynamics and structure, which means land use land cover (LULC) and climate/weather characteristics, that provide terrestrial organic matter (OM) to reservoirs by leaching, and ii) internal dynamics that involve the reservoir contents, river discharges and aquatic environment behavior (Cole et al. 2007; Esteves 1998; Zhu et al. 2014). Because of these associations, monitoring changes in DOM/DOC content may indicate reservoir water quality evolution over time.

Since colored carbon fractions interact with light at ultraviolet (UV), visible and near-infrared (NIR) spectral regions, CDOM estimations are useful to determine the total carbon content, and have been computed via satellite data. This is possible through the application of a bio-optical model in satellite imagery. This kind of model uses remote sensing reflectance (R_{rs}) from determined spectral regions which are adequate to estimate water optically significant constituents (OSC) or to retrieve water inherent optical properties (IOP), as absorption. Bio-optical modeling can be divided in different types such: empirical, semi-empirical, semi-analytical, quasi-analytical, and analytical models (Gordon and Morel 1983; Kirk 2011; A. Morel 1988; A. Morel and Loisel 1998; A. Y. Morel and Gordon 1979).

In this research, we analysed a bio-optical models set composed by quasi-analytical algorithm (QAA-Lee, Carder, and Arnone 2002), developed to retrieve the absorption coefficient of CDOM plus detritus (a_{dg}), and empirical models (Del Castillo and Miller 2008; D'Sa and Miller 2003; Ficek, Zapadka, and Dera 2011; Griffin et al. 2011; Kutser et al. 2005; Mannino, Russ, and Hooker 2008), studied by Zhu et al. (2014), all of them developed to direct retrieve the CDOM absorption coefficient ($a_{CDOM}(\lambda)$) in a given wavelength. QAA is an algorithm which the main objective is to accurately retrieve water absorption ($a(\lambda)$) and backscattering

($b_b(\lambda)$) coefficients from R_{rs}, based on radiative transfer equations, through empirical, semi-analytical, and analytical steps (Lee and Carder 2004; Lee, Carder, and Arnone 2002). This algorithm was developed to oceanic and coastal waters, but currently it has been studied and re-parametrized to be used in inland turbid waters with great concentration of organic and inorganic carbon, as shown by (Cheng Feng Le et al. 2009; Watanabe et al. 2016; Yang et al. 2013; Zhu et al. 2011; Zhu and Yu 2013).

In this research QAA-BBHR (Watanabe et al. 2016) was used since this is a reparametrization of QAA version 5 (Lee, Carder, and Arnone 2002), made for using this algorithm in tropical eutrophic inland waters, developed using a reservoir (Barra Bonita Hydroelectric Reservoir – São Paulo, Brazil) in the same geographic region where the Funil Hydroelectric Reservoir (FHR) is located. QAA-BBHR was used in its original version to retrieve $a_{dg}(\lambda)$, and also with the addition of some steps based on Zhu et al. (2011); Zhu and Yu (2013) to directly derive $a_{CDOM}(\lambda)$.

Empirical models use a direct mathematical relation between R_{rs} and measured data in order to construct a regression capable to accurately estimate a given variable, which guarantees spatial and temporal dependence. The empirical models set used in this research was already tested by Zhu et al. (2014), who have analysed 15 models to retrieve $a_{CDOM}(\lambda)$ and $a_{dg}(\lambda)$ based on neural networks, matrix inversion, semi-analytical algorithm, and empirical model methods. From this set, all empirical models developed to estimate a_{CDOM} were used in this study.

CDOM/DOM and DOC estimation via orbital data is possible by applying a bio-optical model in satellite imagery; it was already conducted with successful results for Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+), Operational Land Imager (OLI) (Kutser et al. 2005, 2016; Zhu et al. 2014), Moderate Resolution Imaging Spectroradiometer (MODIS) (Chen and Zhang 2015) and Medium Resolution Imaging Spectrometer (MERIS) (Kutser et al. 2015; Yang et al. 2013), for example. In this research, the TM/Landsat-5 legacy was used to obtain the spatial and temporal distribution of CDOM for a tropical eutrophic reservoir between 1995 and 2010.

The aims of this research are: a) to evaluate the performance of bio-optical models set composed by empirical models (Zhu et al. 2014) and QAA (Watanabe et al. 2016; Zhu et al. 2013) using simulated reflectance data for TM/Landsat-5 in order to obtain a model capable to accurately retrieve a_{CDOM} from orbital data for a tropical eutrophic reservoir, and b) to analyse the relationships: a_{CDOM} vs Chl-a and a_{CDOM} vs LULC, in order to identify the possible sources of a_{CDOM} in a tropical eutrophic reservoir for a 15-years TM/Landsat-5 time series (1995 to 2010).

4.2 MATERIALS AND METHODS

4.2.1 STUDY AREA

FHR is located at the border of São Paulo and Rio de Janeiro states in Brazil (Figure 10). This reservoir serves to hydroelectric generation, and has as its maximum capacity 40 km² of inundated area, and 6.2 billion m³ of volume. Additionally, Funil is a part of Paraiba do Sul river drainage system, which contributes to FHR water characteristics together with other small streams.

The FHR watershed geomorphology is characterized by a great declivity and strong tendency to erosion, which provide a tendency for a strong surface runoff (Guerra and Botelho 2006; Marengo and Alves Lincoln M. 2005). The climate is influenced by atmospheric systems that promote temperature and rainfall alternations throughout the year, generating two great-defined periods: rainy/wet (from October to April), and dry (from May to September) (Marengo and Alves Lincoln M. 2005; Mendonça and Danni-Oliveira 2007).

Besides that, FHR surroundings are characterized by extensive agriculture and pasture activities, a great industrial park presence, and urban areas with low levels of effluent treatment (INEA 2012, 2016). This LULC dynamic makes nutrients and organic/inorganic matter available to be drained to the reservoir by leaching or water confluence. Thus, carbon and nutrients availability, relief configuration and climate characteristics, are crucial environmental characteristics for this reservoir be considered as susceptible to organic matter and Chl-*a* variation, making CDOM alterations reliable. However, processes such as algal blooms occurrence in FHR are not strongly correlated with surrounding surface cover or climatic variability, but to non-treated domestic effluent discharges that change biotic and abiotic characteristics of water environment (Guedes et al. 2014; INEA 2012, 2016).

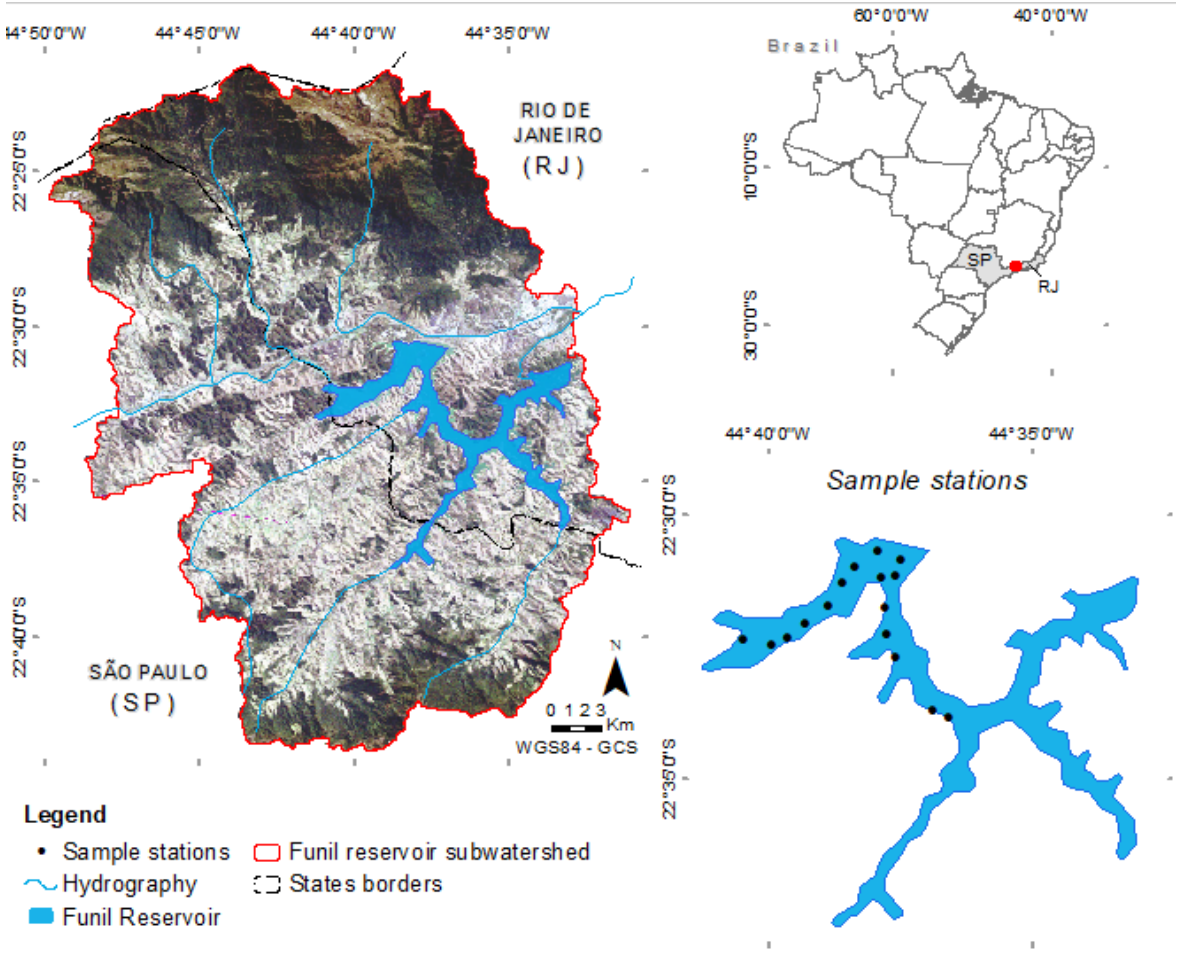

Figure 10. FHR localization map with sample stations distribution. TM/Landsat-5 surface reflectance image, R3G2B1 color composite. Date: 01/08/2010. Path/row: 218/76.

4.2.2 RADIOMETRIC DATA

In situ radiometric data were taken in FHR in April/2013 for 16 sample stations (Figure 10) at above the water surface (0+), in a spectral range between 320 and 950 nm, and spectral resolution of 3.3 nm. Two inter calibrated RAMSES spectrometers (TriOS GmbH, Germany) were used to measure the total water-leaving spectral radiance (L_t - RAMSES-ARC) and the spectral irradiance (E_s - RAMSES-ACC).

The equipment was connected to an optical fiber in order to yield the radiometric measurements using 7° as the field of view and a spray protection to RAMSES-ARC, and a cosine collector to RAMSES-ACC. This procedure was conducted as far as possible from the boat to avoid the interference of its shadow or light beams. Additionally, at each sample station, several measures were taken and their average was used as the representative spectrum for each collection point (Ogashawara et al. 2016).

Finally, the remote sensing reflectance data (R_{rs}, sr^{-1}) was obtained as Eq. 3 (Kirk 2011):

$$R_{rs}(0^+) = \frac{L_t\,(\lambda,\theta,\phi)}{E_s(\lambda,\theta,\phi)} \quad (3)$$

where, L_t is the total water-leaving spectral radiance [W/m^2sr], E_s is the spectral downwelling irradiance above the water surface [W/m^2], λ is wavelength, θ is the zenithal angle, and ϕ is the azimuthal angle.

4.2.3 REFLECTANCE SIMULATION

In situ hyperspectral measurements were used to simulate the related signal registration at the central wavelengths of TM/Landsat-5 bands, according to the specific spectral response function of each one (Figure 11, Table 11). Then, the resulting simulated remote sensing reflectance ($R_{rs_{simulated}}$) was calculated by the weighted sum of the hyperspectral field remote sensing reflectance data into each bandwidth by a convolving procedure between the signal registered and the bands spectral function as Eq. 4 (Kidder and Vonder Haar 1995; Schowengerdt 2006):

$$R_{rs_{simulated}} = \frac{\int_{x\,min}^{x\,max} R_{rs\,field}.Fr.dx}{\int_{x\,min}^{x\,max} Fr.dx} \quad (4)$$

where, $R_{rs_{field}}$ is the hyperspectral R_{rs}, Fr is the sensor function, dx is the wavelength range which defines the sensor spectral resolution.

Table 11. TM/Landsat-5 band configuration.

Spectral band	Wavelength (µm)	Central wavelength (µm)	Spatial resolution (m)
B1 – blue	0,45 – 0,52	0,485	30
B2 – green	0,52 – 0,60	0,560	30
B3 – red	0,63 – 0,69	0,660	30
B4 – near infrared	0,76 – 0,90	0,830	30
B5 – infrared short wave	1,55 – 1,75	1,650	30
B6 – infrared thermal	10,40 – 12,50	11,450	120
B7 – infrared short wave	2,08 – 2,35	2,215	30

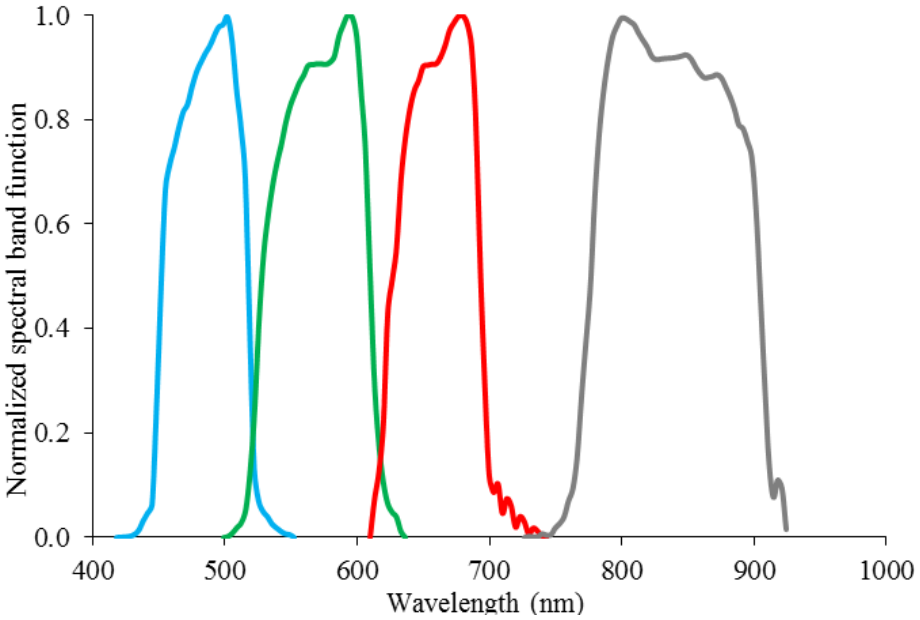

Figure 11. The TM/Landsat-5 spectral band function distribution: B1 (blue), B2 (green), B3 (red), and B4 (gray).

The radiometric data simulation is useful since it allows the adequacy of hyperspectral field data to the multispectral TM/Landsat-5 band configuration, in order to calibrate the bio-optical models according to the wavelength where the signal is potentialized (Figure 12). For this reason, the models capable of being applied to TM orbital data to accurately estimate a_{CDOM}.

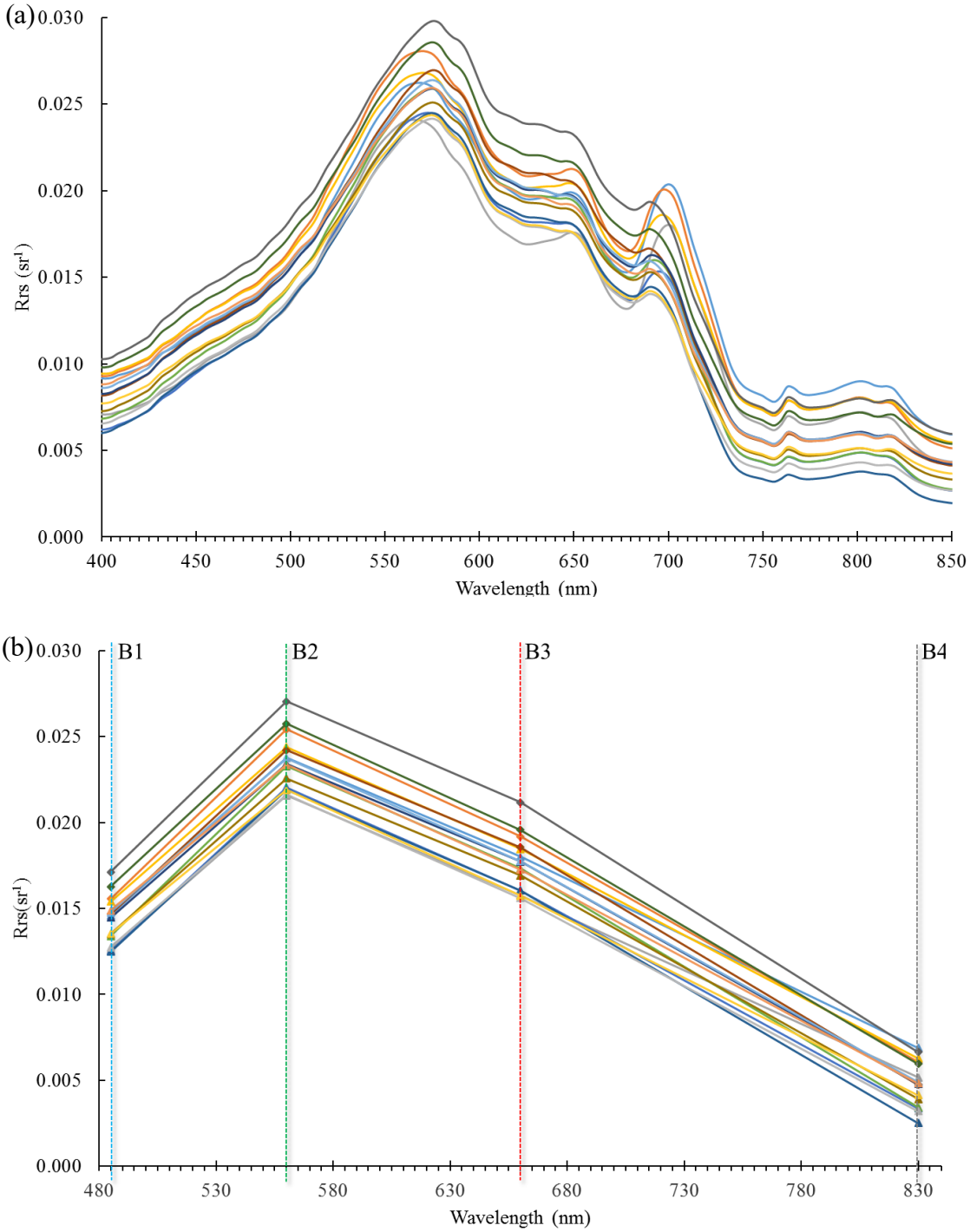

Figure 12. The hyperspectral field R_{rs} (a) and its related simulated multispectral R_{rs} (b) with TM/Landsat-5 band indication.

2.4.5 $a_{CDOM}(\lambda)$ DATA

The samples were taken in FHR in the same fieldwork, and at the same stations used for *in situ* reflectance data collection. Water samples were filtered right after collection using a 0.2 µm nucleopore membrane. A Perkin Elmer lambda 35 UV/Vis system (UV-2600 - Perkin Elmer Inc, USA) with a transparent quartz cuvette with 10 cm of linear length was used to determine

$a_{CDOM}(\lambda)$. Millipore Milli-Q water (Millipore Corporation, USA) was the reference for pure water, and the absorbance correction was made according to the average of the Milli-Q water absorbance baseline fluctuations at each wavelength (James L. Mueller 2000). The optical density (OD) was measured in a range between 190 and 1100 nm, with 1 nm as spectral resolution. The $a_{CDOM}(\lambda)$ values were obtained for a spectral range between 390 and 730 nm, with intervals of 1 nm (Figure 13), using Bricaud et al., 1981 (Eq. 5):

$$a_{CDOM} = \frac{2{,}3\ OD(\lambda)}{l} \qquad (5)$$

where, λ corresponds to wavelength, l is the cuvette linear length in meters.

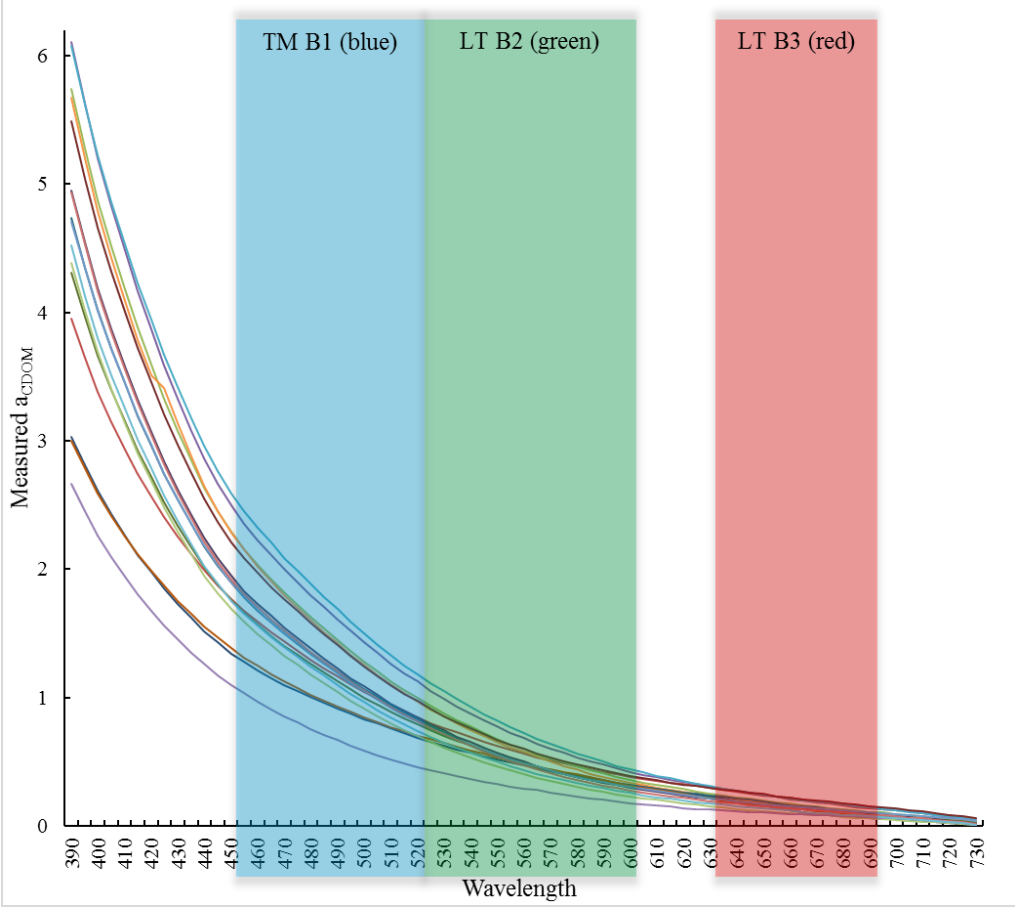

Figure 13. The $a_{CDOM}(\lambda)$ measured spectra with TM/Landsat-5 bands of the visible spectrum.

4.2.5 BIO-OPTICAL MODELS FOR $a_{CDOM}(\lambda)$ ESTIMATION

In order to estimate $a_{CDOM}(\lambda)$ a bio-optical models set composed by empirical models, and QAA-BBHR configurations was tested using the $R_{rs\,simulated}$ as input data (Table 12). Empirical models use a direct mathematical relation between R_{rs} and the $a_{CDOM}(\lambda)$ field data to make it possible to retrieve this variable from other R_{rs} data sets. This direct relationship guarantees spatial and temporal dependence, which implies in recalibration procedures to make the model capable of being applied in a different water body from the other one that was the base for its parameterization (Le and Hu 2013). The empirical models set used in this research is composed by all empirical models (Del Castillo and Miller 2008; D'Sa and Miller 2003; Ficek, Zapadka, and Dera 2011; Griffin et al. 2011; Kutser et al. 2005; Mannino, Russ, and Hooker 2008) tested by (Zhu et al. 2014) to derive a_{CDOM} (Appendix 1).

QAA is a kind of algorithm based on the radiative transfer equations to retrieve a and b_b coefficients from R_{rs} by analytical, semi-analytical and empirical sequenced steps that can ensure the application of this kind of algorithm in almost all water cases with just few modifications (Lee, Carder, and Arnone 2002). The QAA-BBHR (Watanabe et al. 2016) is a QAA-v5 (Lee, Carder, and Arnone 2002) re-parameterization developed specifically for tropical eutrophic waters (Watanabe et al. 2016) to estimate IOPs. It is substantial to evidence the importance of an accurate definition of the reference wavelength (λ_0) at the wavelength/spectral band where water absorption is the strongest.

Since the goal was to estimate a_{dg} and a_{CDOM}, four configurations of QAA-BBHR were tested, two of them using λ_0 at B3 (660 nm) and B4 (830 nm) to obtain a_{dg}, and two at the same referred λ_0 with addiction of some steps in order to obtain directly $a_{CDOM}(\lambda)$, based on Zhu et al. (2011) and Zhu and Yu (2013), here called as QAA-BBHR+CDOM (Appendix 1). In this research, we are assuming that both absorption quantities, a_{dg} and a_{CDOM} are similarly correlated to measurements of CDOM absorption (Cheng Feng Le et al. 2009; Lee and Carder 2004; Lee, Carder, and Arnone 2002; Yang et al. 2013; Zhu and Yu 2013).

It is important to highlight that the main objective of $a_{CDOM}(\lambda)$ historical distribution retrieving is to accurate capture the CDOM behavior/alterations over the time series, but not to obtain its accurate historic values. Because of that, the sample, orbital data and models sets, as well as the reflectance calculation were considered sufficient for this work.

Table 12. The $a_{CDOM}(\lambda)$ bio-optical models set.

Algorithm name	Type	Input R_{rs}[a]	Output	References
Griffin	EM	B1, B2, B3	$a_{CDOM}(400)$	Griffin et al. (2011); Zhu et al. (2014)
D'sa	EM	B1, B2	$a_{CDOM}(412)$	D'Sa and Miller (2003); Zhu et al. (2014)
Castilo	EM	B2, B3	$a_{CDOM}(412)$	Del Castillo and Miller (2008); Zhu et al. (2014)
Kutser	EM	B2, B3	$a_{CDOM}(420)$	Kutser et al. (2005); Zhu et al. (2014)
Ficek	EM	B2, B3	$a_{CDOM}(440)$	Ficek, Zapadka, and Dera (2011); Zhu et al. (2014)
Mannino	EM	B1, B2	$a_{CDOM}(443)$	Mannino, Russ, and Hooker (2008); Zhu et al. (2014)
QAA-BBHR(660)	QA	B1, B2, B3, B4	$a_{dg}(485)$	Watanabe et al. (2016)
QAA-BBHR(830)	QA	B1, B2, B3, B4	$a_{dg}(485)$	Watanabe et al. (2016)
QAA-BBHR+CDOM(660)	QA	B1, B2, B3, B4	$a_{CDOM}(485)$	Watanabe et al. (2016); Zhu and Yu (2013)
QAA-BBHR+CDOM(830)	QA	B1, B2, B3, B4	$a_{CDOM}(485)$	Watanabe et al. (2016); Zhu and Yu (2013)

[a] Simulated R_{rs} values were used as input data. / EM: empirical model. / QA: Quasi-analytical models.

4.2.6 ALTERNATIVE EMPIRICAL MODEL PARAMETRIZATION

Likewise literature-based bio-optical models were tested; a specific empirical model for a_{CDOM} retrieving from simulated R_{rs} in FHR was also parametrized. This is justified by the difficult to successfully apply an empirical model to an environment different from the one which it was parameterized (Le and Hu 2013).

To propose an alternative empirical model capable of retrieving a_{CDOM} for FHR, a detailed study was necessary to verify the relationship between the measured $a_{CDOM}(\lambda)$ and the $R_{rs_{simulated}}$ data. To do that, we used a correlation analysis, and also a 2 dimensional (2D) color correlation plot (ICE - Ogashawara et al. 2014) to obtain the model reference wavelength (λ) and its ideal index (band ratio or single band). Furthermore, the ternary plot among the absorption coefficients of non-algal particles (NAP), CDOM and phytoplankton presented by Ogashawara et al. (2016) was also used to ensure that at the chosen wavelength (λ) the absorption predominance is related to CDOM.

We have also tested the a_{CDOM} estimation at different wavelengths choices using the found index, just varying the regression type among linear, power (*pwr*), exponential (*exp*), and logarithmic (*ln*) for the simple band ratio (*sp*) and normalized difference (*ND*). Only the most significative model will be applied in the satellite imagery.

Independent from the bio-optical model choice, a mask was constructed to restrict the bio-optical model application to water surface in all satellite images despite the water level variations, to ensure that the $a_{CDOM}(\lambda)$ retrieval is related only to the water surface reflectance values and to derive the statistics assessment of mathematical comparisons among the same pixel amount. This mask was based on the image taken in the driest period (September/2001) as well as it was used to imagery preprocessing described below (4.2.9 SATELLITE IMAGERY DATA AND PREPROCESSING).

4.2.7 MODEL CALIBRATION AND VALIDATION

The validation procedure was conducted just for empirical models, since the QAA-BBHR different configurations were directly applied in its original parameterization. Empirical models are based on the relationship between the measured absorption data and $R_{rs_{simulated}}$, so two data sets (validation and calibration) are needed, or a leave one out cross validation (LOOCV) methodology can be used to verify the model predictive power (Montanher et al. 2014).

For empirical models, a previous adjustment analysis was conducted to select which literature-based models could be adequate for the $a_{CDOM}(\lambda)$ retrieval. This analysis was based on a linear tendency line between the index (band ratio) of each model and the measured a_{CDOM} to obtain R^2 and p-value. Then, the LOOCV technique was applied just for the models that presented a good performance in adjustment analysis. After validation procedures, the error analysis was made just for the validated models.

For QAA-BBHR(660), QAA-BBHR(830), QAA-BHR+CDOM (660) and QAA-BBHR+CDOM(830), the adjustment analysis and error analysis were conducted together at the same stage for all tested configurations by comparing the estimated $a_{CDOM}(\lambda)$ with the measured $a_{CDOM}(\lambda)$.

4.2.8 MODELS PERFORMANCE ASSESSMENT

For both types of bio-optical models, the error analyses were evaluated using the normalized root mean squared error (NRMSE) and the root mean squared error percent (%RMSE) to verify the residual, bias to analyse the tendency, and Sutcliffe model efficiency coefficient (Nash) to evaluate the model performance. They are defined as presented in Eq. 6, Eq. 7, Eq. 8, Eq.9:

$$NRMSE = \frac{RMSE}{(x_{max}^{M} - x_{min}^{M})} \quad (6)$$

$$\%RMSE = \left(\frac{RMSE * 100 * n}{\sum x_i^M}\right) \quad (7)$$

$$Bias = \frac{\sum(x_i^E - x_i^M)}{n} \quad (8)$$

$$Nash = 1 - \left(\frac{\sum(x_i^E - x_i^M)^2}{\sum(x_i^M - \overline{x^M})^2}\right) \quad (9)$$

where, x_i^E is the estimated value by the model, x_i^M is the measured value, n is the sample quantity, x_{max}^M is the maximum measured value, x_{min}^M is the minimum measured value, $\overline{x^M}$ is the average of measured values, and $RMSE = \sqrt{\frac{\sum(x_i^E - x_i^M)^2}{n}}$.

4.2.9 SATELLITE IMAGERY DATA AND PREPROCESSING

The TM/Landsat-5 images were acquired from the United States Geological Survey (USGS - http://earthexplorer.usgs.gov/). This sensor was chosen due to its temporal, radiometric, and spatial resolutions that were considered sufficient to this study. Besides that, this sensor provides orbital data for the whole period intended in this research, from 1995 to 2010. In addition, we also considered the fact that the studied water body is limited just to one scene of TM coverage, which eliminates the impact of using different scenes to extract information from the same water target.

As orbital data sets, all available images (Path/row: 218/76) without cloud cover above FHR since 1995 to 2010 were used, totalizing 64 images. No date restriction was imposed, and when there were two images per month, the surface reflectance average was taken to use as much as possible of orbital data. The imagery preprocessing was conducted as previously described by (Martins et al. 2016), intending to ensure the analysis quality, since each image that compose the time series was taken in different dates, and under different atmospheric, geometric and illumination conditions.

An image taken during the driest period for the study area was chosen as the base image for the imagery preprocessing described below in order to ensure low influence of atmospheric compounds in the signal registered by the sensor, and to avoid negative values of reflectance after atmospheric correction. According to this choice criterion, the image from September 2001 was used as base image for the pre-processing procedure.

Generally, radiometric calibration was first made to compute the Top Of Atmosphere (TOA) reflectance from digital numbers following Chander, Markham, and Helder (2009). Then, the

atmospheric correction was conducted using fast line-of-sight analysis of hypercubes (FLAASH - Adler-Golden et al. 1999) in the base image, in order to derive its surface reflectance image. After, the iteratively reweighted multivariate alteration detection (IR-MAD - Canty, Nielsen, and Schimidt 2004) among the surface reflectance base image and the TOA reflectance images was conducted to derive the statistical data to the following radiometric normalization step. In this last step, atmospheric corrected base image (FLAASH), TOA reflectance images (radiometric calibration), and the χ^2 statistics (IR-MAD) were used to convert all time series images to surface reflectance ($R_{surface}$).

After the described image preprocessing, the images were divided by the mathematical constant π to make the $R_{surface}$ comparable to R_{rs}, in order to apply the bio-optical model chosen in all images to obtain $a_{CDOM}(\lambda)$ historical distribution (Ahn and Shanmugam 2006; Moses et al. 2012).

4.3 RESULTS

4.3.1 BIO-OPTICAL MODELS

The literature-based empirical models failed in the adjustment analyses stage (model index vs measured $a_{CDOM}(\lambda)$), with $R^2 < 0.0058$, and p-value > 0.7789 (Table 13), consequently they were not recalibrated and validated. These models probably did not work well because of its band ratios definition and a_{CDOM} reference wavelength are not defined where the best correlation with $a_{CDOM}(\lambda)$ field data and $R_{rs\,simulated}$ values were found. Additionally, an unusual $a_{CDOM}(\lambda)$ behavior was found in FHR during the field work, that does not follow the typical shape with exponential decrease for wavelength increase, but presenting high spectral response even at high wavelengths, probably as a consequence of CDOM significant presence (Ogashawara et al. 2016).

Table 13. Adjustment analysis among the tested empirical model index and the measured a_{CDOM} at the model reference wavelength.

Bio-optical model	index	Wavelength (nm)	R²*	p-value*
Mannino	B1/B2	443	0.0002	0.9635
D'sa	B1/B2	412	0.0001	0.9642
Griffin	B3+B2/B1	400	0.0000	0.9911
Kutser	B2/B3	420	0.0046	0.8029
Castilo	B2/B3	412	0.0038	0.8206
Ficek	B2/B3	440	0.0058	0.7789
$a_{CDOM}(485)$ simple ratio	B4/B1	485	0.91	<0.0001
$a_{CDOM}(485)$ normalized difference	B4-B1/B4+B1	485	0.76	<0.0001
$a_{CDOM}(547)$ simple ratio	B4/B1	547	0.66	0.00
$a_{CDOM}(547)$ normalized difference	B4-B1/B4+B1	547	0.69	0.00

* For linear tendency line and $\alpha = 5\%$.

In the same way, the literature-based empirical models use B1, B2, and B3 to estimate a_{CDOM} at 400, 412, 420, and 443 nm, but the best correlations were found for $a_{CDOM}(485)$ and $a_{CDOM}(547)$ using B4/B1 TM/Landsat-5 band ratio, after correlation and 2D color correlation plot (Figure 14). The a_{CDOM} models at alternative wavelengths 485 and 547 nm using B1 and B4 had better adjustment than all other bio-optical models analysed (Table 13, Table 14).

After adjustment analysis verification, the proposed empirical models at 485 and 547 nm were the only ones that fitted the criteria to be validated by LOOCV. During this validation procedure, the regression coefficients for all regressions types tested did not have great variations over LOOCV interactions, with no significant changes since it was observed only from the fifth decimal place, characterizing a stable model.

The obtained results are corroborated by other researches in complex inland waters, which found that the water compounds variation may affect the IOPs, indicating that $a_{CDOM}(\lambda)$ can be measured/estimated at longer wavelengths for inland eutrophic waters, since CDOM behavior is directly dependent on the suspended particle matter (SPM) and the Chl-a concentrations (Brezonik et al. 2015; Matsuoka et al. 2015; Rochelle-Newall and Fisher 2002; Zhang et al. 2009; Zhu and Yu 2013). Additionally, despite the strongest CDOM absorption at shorter wavelengths, it can be more useful to retrieve a_{CDOM} by R_{rs} at higher ones, as a way to overcome the influence of the other OSCs, such as phytoplankton and SPM, in its spectral features (Brezonik et al. 2015; J. Chen et al. 2017; Ogashawara et al. 2016). The best adjusted model was the one proposed in this research for the estimation of $a_{CDOM}(485)$, which presented a $R^2 = 0.91$, and p-value $< \alpha$.

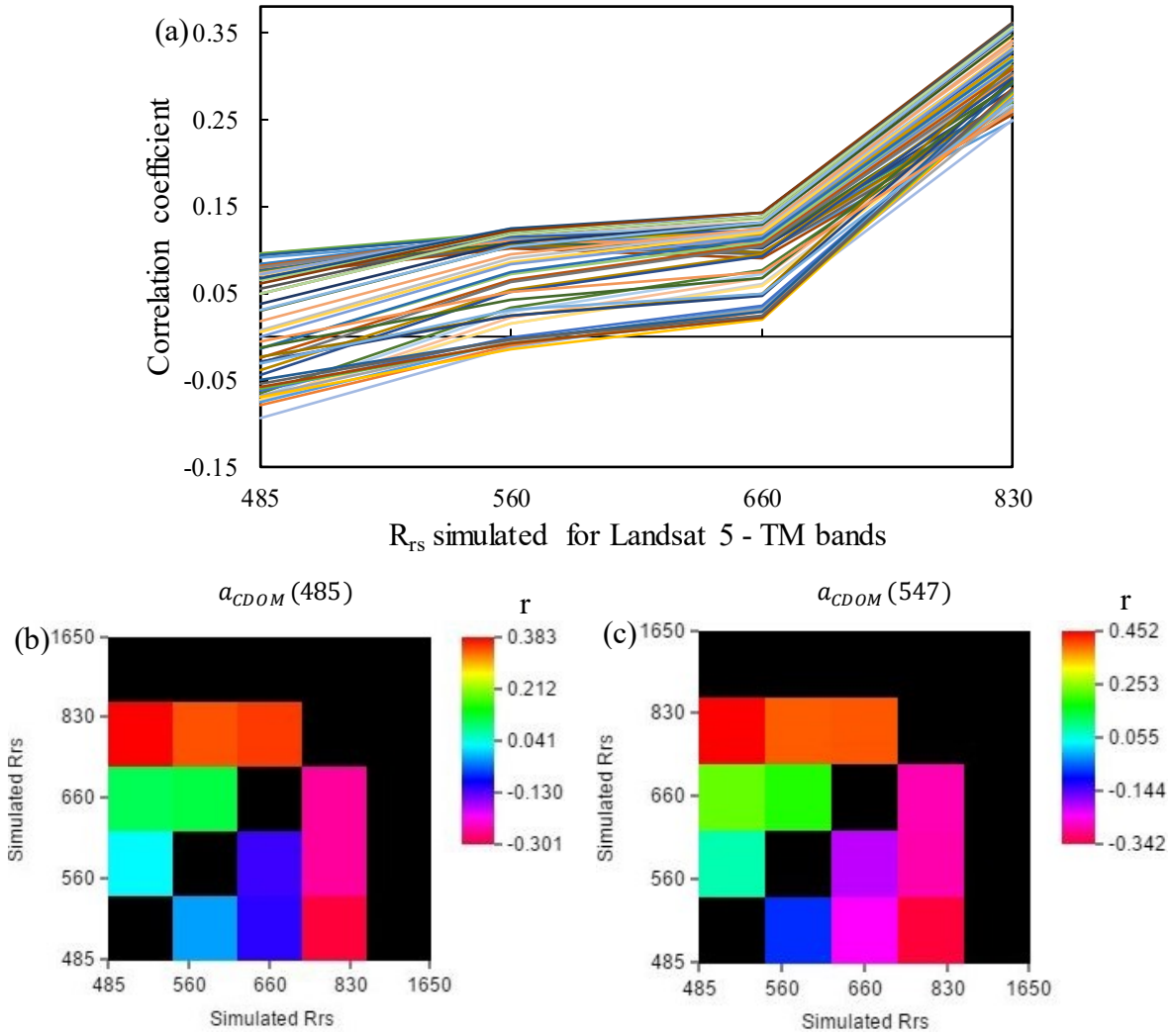

Figure 14. Correlation analysis among all measured a_{CDOM} and the simulated R_{rs} values (a), and the best results found by 2D color correlation plot for correlation analysis – r – among band ratios using simulated R_{rs} and a_{CDOM} measured at 485 (b) and 547 nm (c).

Even QAA-BBHR configurations had low adjustment between measured and estimated data, demonstrating that they cannot be used to retrieve $a_{CDOM}(\lambda)$ and $a_{dg}(\lambda)$ by $R_{rs\,simulated}$ in FHR ($R^2 < 0.1461$, p-value between 0.14 and 0.76 - Table 14). The error analyses also demonstrated a QAA-BBHR or QAA-BBHR+CDOM poor performance, with NRMSE between 0.6 and 1.17, %RMSE variating from 41 to 98, and negative values for Nash for all configurations (Table 14). It probably occurred because of the used QAA parameterization was constructed using non-simulated R_{rs} data, and was based on different water environment, so the empirical steps and regression constants are not adequate to FHR characteristics. Furthermore, CDOM huge concentration strongly impacts QAA performance for FHR, because its values push the a_{CDOM} slope for the high wavelengths, as already shown by Ogashawara et al. (2016). This performance can also be related to the TM/Landsat-5 band width that is not the optimal spectral configuration for QAA application.

Table 14. QAA-BBRH adjustment analysis and error evaluation.

Bio-optical models	R^2[a]	p-value[a]	NRMSE[b]	% RMSE[b]	Bias[b]	Nash[b]
QAA-BBHR(830)	0.14	0.16	1.17	98.22	-1.23	-20.53
QAA-BBHR(660)	0.01	0.63	0.50	41.86	-0.46	-2.91
QAA-BBHR +CDOM(830)	0.15	0.14	1.17	98.34	-1.23	-20.59
QAA-BBHR +CDOM(660)	0.01	0.76	0.60	50.00	-0.58	-4.58

[a] Related to the adjustment between the model index and the measured a_{CDOM} at the model wavelength, for a linear tendency line, and $\alpha = 5\%$.
[b] Related to the adjustment between measured and estimated a_{CDOM}, for a linear tendency line, and $\alpha = 5\%$.

Finally, the error analysis among all validated empirical models varying just the regression types demonstrated that the best model was the linear $a_{CDOM}(485)$ for a simple band ratio (Eq. 10), since there was no significant difference in error evaluation (RMSE=7.27%), bias (<0.001) and Nash (0.91) among all possible regressions for a_{CDOM} estimations at the referred λ (Table 15), and due to linear regressions have less sensitivity related to its coefficients, and more dependence on the regression index (band ratio), making the model more reliable to environment variations.

$$a_{CDOM}(485) = -0.5986 + 5.5510 * \left(\frac{B4}{B1}\right) \quad (10)$$

The proposed empirical model could identify small variations in reflectance values and it was possible to differentiate even slight alterations in a_{CDOM} after its application in the satellite imagery (Figure 20, Figure 21).

Table 15. Proposed empirical models error evaluation after comparison between measured and estimated $a_{CDOM}(\lambda)$ values.

Bio-optical models	NRMSE	% RMSE	Bias	Nash
$a_{CDOM}(485)_{sp\ pwr}$	0.08	6.88	0.00	0.92
$a_{CDOM}(485)_{sp\ exp}$	0.08	6.83	0.00	0.92
$a_{CDOM}(485)_{sp\ linear}$	0.09	7.27	0.00	0.91
$a_{CDOM}(485)_{ND\ ln}$	0.12	10.26	0.00	0.80
$a_{CDOM}(485)_{ND\ exp}$	0.12	10.47	0.00	0.79
$a_{CDOM}(485)_{ND\ linear}$	7.50	644.65	-7.92	-794.44
$a_{CDOM}(547)_{sp\ pwr}$	0.15	12.90	0.01	0.55
$a_{CDOM}(547)_{sp\ exp}$	0.07	6.03	0.12	-0.83
$a_{CDOM}(547)_{sp\ linear}$	0.16	13.48	0.00	0.66
$a_{CDOM}(547)_{ND\ ln}$	2.00	175.09	1.02	-63.26
$a_{CDOM}(547)_{ND\ linear}$	3.60	314.76	1.84	-206.67

4.3.2 a_{CDOM} HISTORICAL DISTRIBUTION

According to rainfall behavior in the study area, the year can be divided in two periods: the rainy/wet (from October to April), and the dry (from May to September). Relating these periods to seasons, we can identify austral summer (January to March) as a wet period, austral fall (April to June) as a transition to wet from dry, austral winter (July to September) as dry, and austral spring as a transition between dry and wet (Figure 15). When we consider the a_{CDOM} average according to seasons, a specific behavior could be identified and related to rainfall occurrence either.

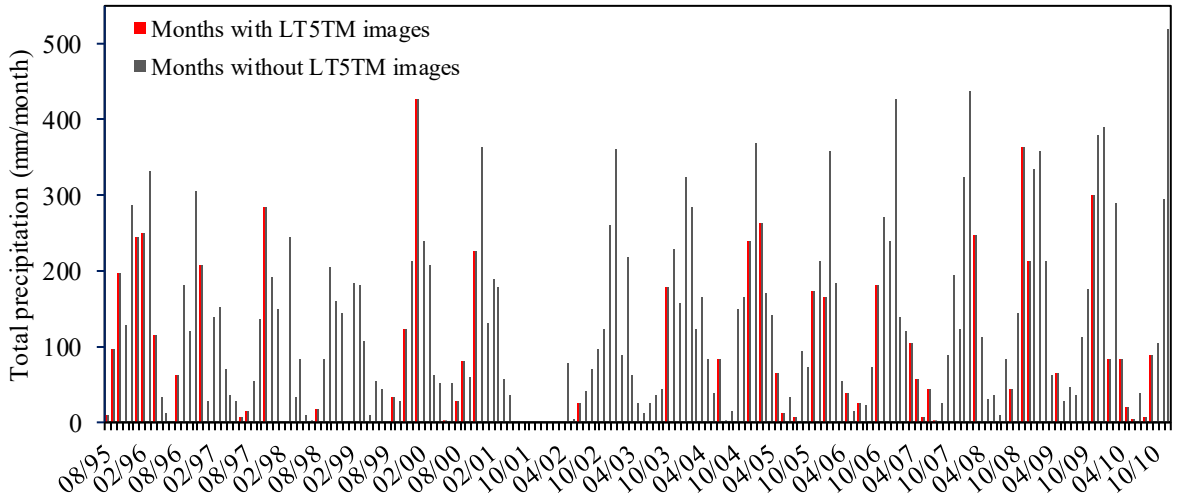

Figure 15. Monthly precipitation with indication of imagery occurrence (orange bars).
Source: BDMEP – INMET for rainfall data station: Resende/RJ - 83738
(http://www.inmet.gov.br/projetos/rede/pesquisa/)

Using a_{CDOM} average according to seasons (Figure 16, Figure 17), generally we can identify three different patterns to CDOM behavior: one associated with austral summer, when an almost homogeneous a_{CDOM} behavior can be found, producing the lowest values for average (<7 m^{-1}) and variance with the highest frequencies (>140) occurrence, which is associated with the lowest a_{CDOM} average (3 m^{-1}) values. Contrary, the other found pattern is associated with austral winter season presenting the highest average (>11 m^{-1}) and variance (maximum variance = 14) values wich are linked to the lowest frenquency (<90) for the lowest a_{CDOM} average values, indicating a great heterogeneous distribution for CDOM values.

Finally, the third pattern found is associated with austral spring and austral fall, when a moderate scale for average and variance were found, demonstrating that these seasons have a behavior not homogeneous as austral summer season, but without the great variation found in austral winter, with average most concentrated between 3 and 8.5 m^{-1}, low variance values in

general, and moderate frequencies (> 50) associated with the lowest a_{CDOM} average values (< 5 m^{-1}).

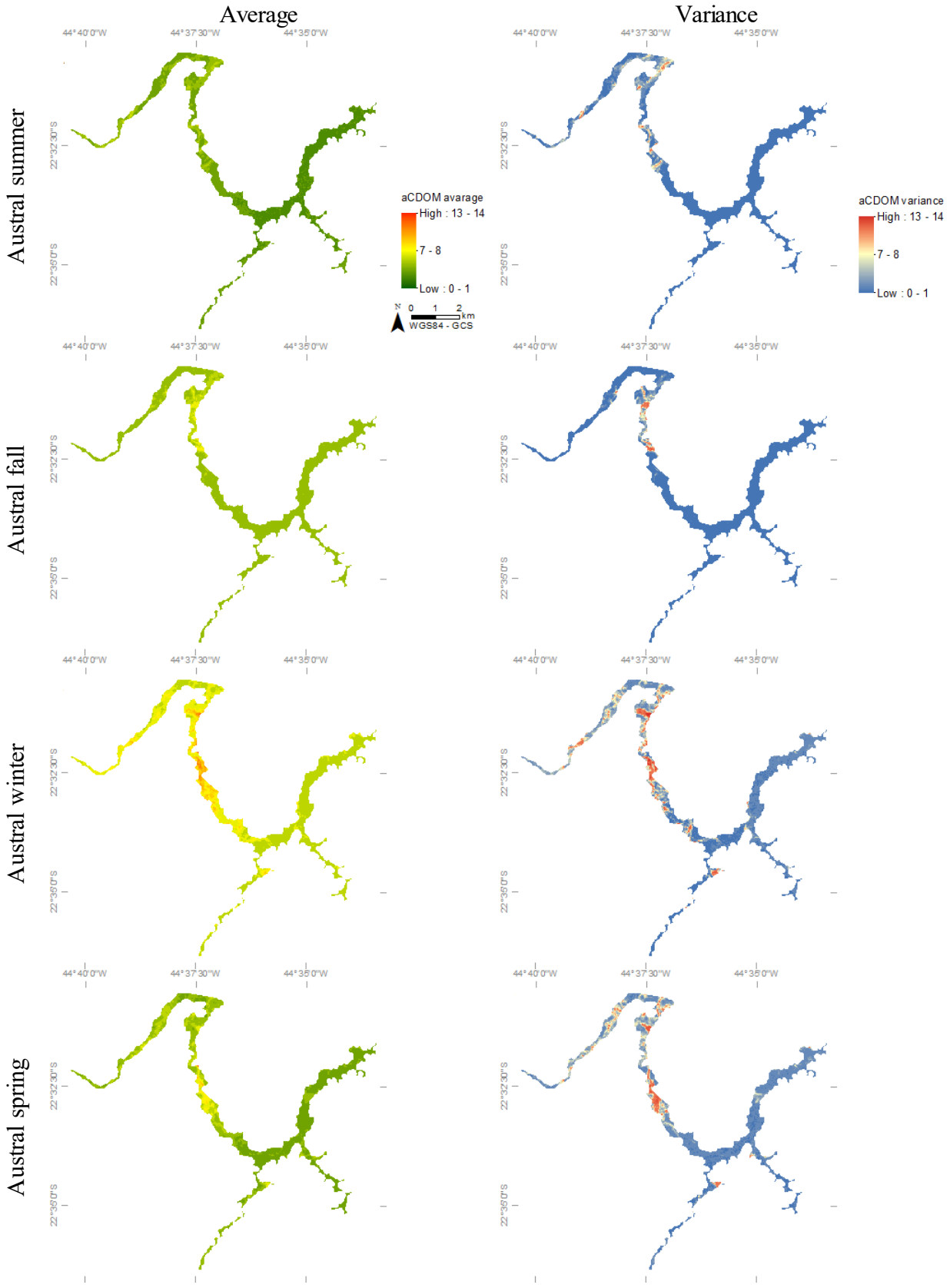

Figure 16. $a_{CDOM}(485)$ historical average and variance distribution by season.

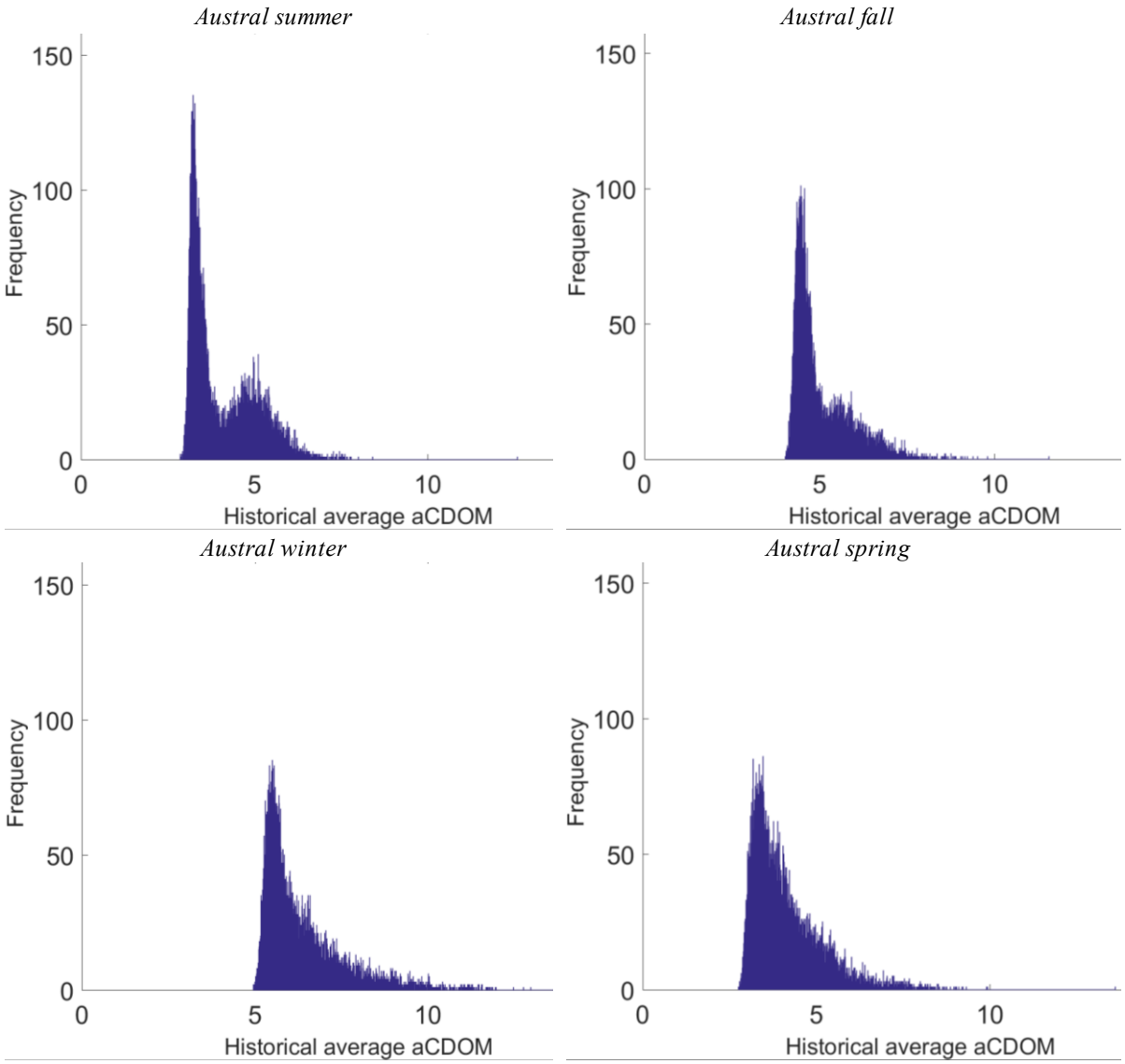

Figure 17. $a_{CDOM}(485)$ histograms presenting historical average values by season and its related frequencies.

In summary, the general season average values indicate that the austral summer and winter can be defined as extreme periods, one almost homogeneous and other great heterogeneous, respectively. The pattern austral spring/fall can also be differentiated from the other two, and these seasons can be defined as transition periods.

The described CDOM behavior can be well observed when we regard the average of each image in the boxplot presented in (Figure 18), which presents the average for all $a_{CDOM}(485)$ images separated by seasons. In austral summer we can observe a little outlier occurrence and a low dispersion of them between 7 and 9 m^{-1}. Contrary, in austral winter a great presence and dispersion of outliers can be detected (from 9 to 14 m^{-1}), indicating a homogeneous a_{CDOM} behavior in this season, while for austral spring and austral fall the outliers are distributed according to approximately the same range (7 m^{-1}), but ranging from 5 to 11 m^{-1} in austral fall, and 7 to 13 m^{-1} in austral spring.

The a_{CDOM} average distribution obtained in the image analysis agrees with the study of a_{CDOM} average values by seasons presented in histograms, both indicating a cyclical tendency for CDOM behavior over year. The data dispersion observed in box plot also contributes with this conclusion, since the median position and variance indicate a most concentrated data in austral summer that became most scattered in austral fall, getting the peak in austral winter, but starts to be less dispersed until spring, indicating a tendency for summer pattern occurrence.

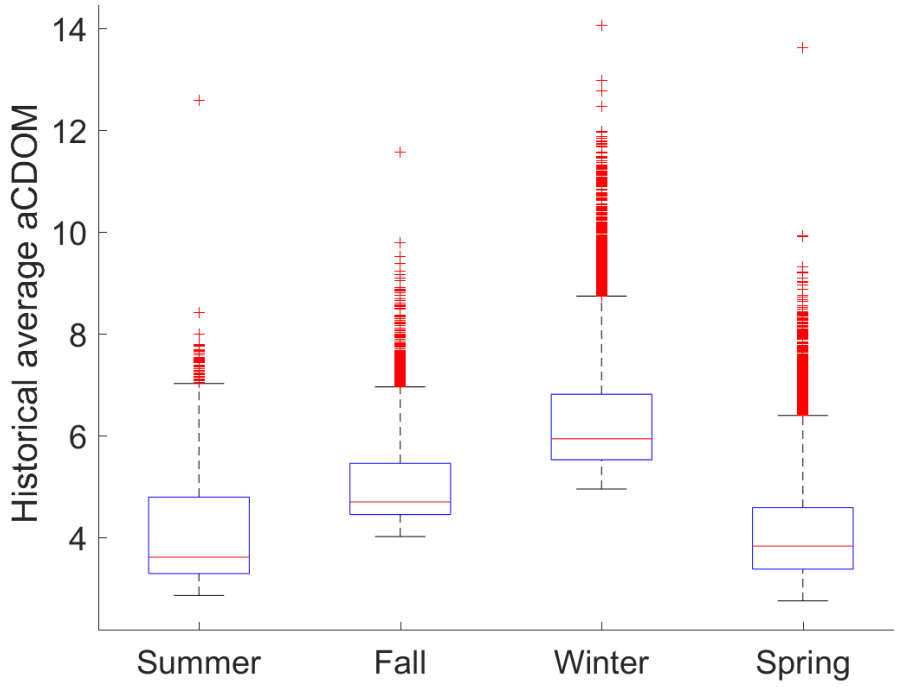

Figure 18. Boxplot presenting the average obtained for all a_{CDOM} (485) images grouped according to seasons.

These results suggest a CDOM sources mix over time, which can be related to LULC and/or to Chl-a and bacterial activity, providing a relationship among them and CDOM that is not constant in space and time, as indicated by Allan (2004) and Matsuoka et al. (2015). The observed cyclical tendency over seasons and year corroborate this conclusion, since the a_{CDOM} averages (by season and by image) can be correlated to rainfall and drought periods, as well as the transition between them.

It is important to highlight that the field data was collected in an algal bloom situation, which allows to construct a flexible model able to retrieve a_{CDOM} in bloom and no-bloom situation, since the samples points take in account a wide range of CDOM collection. On the other hand, taking samples in this situation also implies in more correlation for CDOM with Chl-a than with DOC (Danhiez et al. 2017; Kirchman et al. 1991; Le and Hu 2013; Rochelle-Newall and Fisher 2002; Sasaki et al. 2005). Despite that, the relationship between CDOM and DOM still

the same, both increase or decrease proportionally, and linearly (Brezonik et al. 2015; Danhiez et al. 2017; Rochelle-Newall and Fisher 2002).

4.3 DISCUSSION

The CDOM variation can be associated with humic acids, which means terrestrial matter source and correlation with DOC, or can be linked to fluvial acids, that is an indication of rivers input as a source of organic matter and Chl-a significant presence (Xue, Zhang, and Duan 2016). Surveys have shown that the Chl-a high concentrations allied to phytoplankton degradation after bloom situation, and watershed LULC constitution are two important sources of DOM/CDOM for inland waters (Allan 2004; Danhiez et al. 2017; Kirchman et al. 1991; Matsuoka et al. 2015; Sasaki et al. 2005; Zhang et al. 2009).

Based on this, two different patterns of CDOM behavior were found: one associated with leaching from FHR watershed, and other related to algal blooms occurrence. Regarding rainfall seasonality, phytoplankton blooms occurrence in satellite images and $a_{CDOM}(485)$ behavior found over year together, it is possible to observe that the highest a_{CDOM} average values and variances occurred in images collected in dry months. This situation is completely opposite to that found in wet/rainy months.

Observing Figure 19, we can see this great difference between dry and wet CDOM averages, which can be explained by: (a) algal bloom occurrence, that is more frequent in the dry period and when the highest variances were found despite average values rage among images and (b) rainy occurrence that promotes a higher dissolution of river discharges effluents avoiding algal blooms occurrence as well as allowing an almost homogeneous distribution of terrestrial DOM leashed to water surface, resulting in small variances for a_{CDOM} values.

It is important to observe that a_{CDOM} averages values without variance analysis are not sufficient to explain the CDOM behavior since high values can occur in just small areas associated with small blooms, being masked by the prevailing low values occurred in the most part of the reservoir. Then, we must conclude that a pattern based on algal bloom situations and others linked to rainfall occurrence are clearly identified.

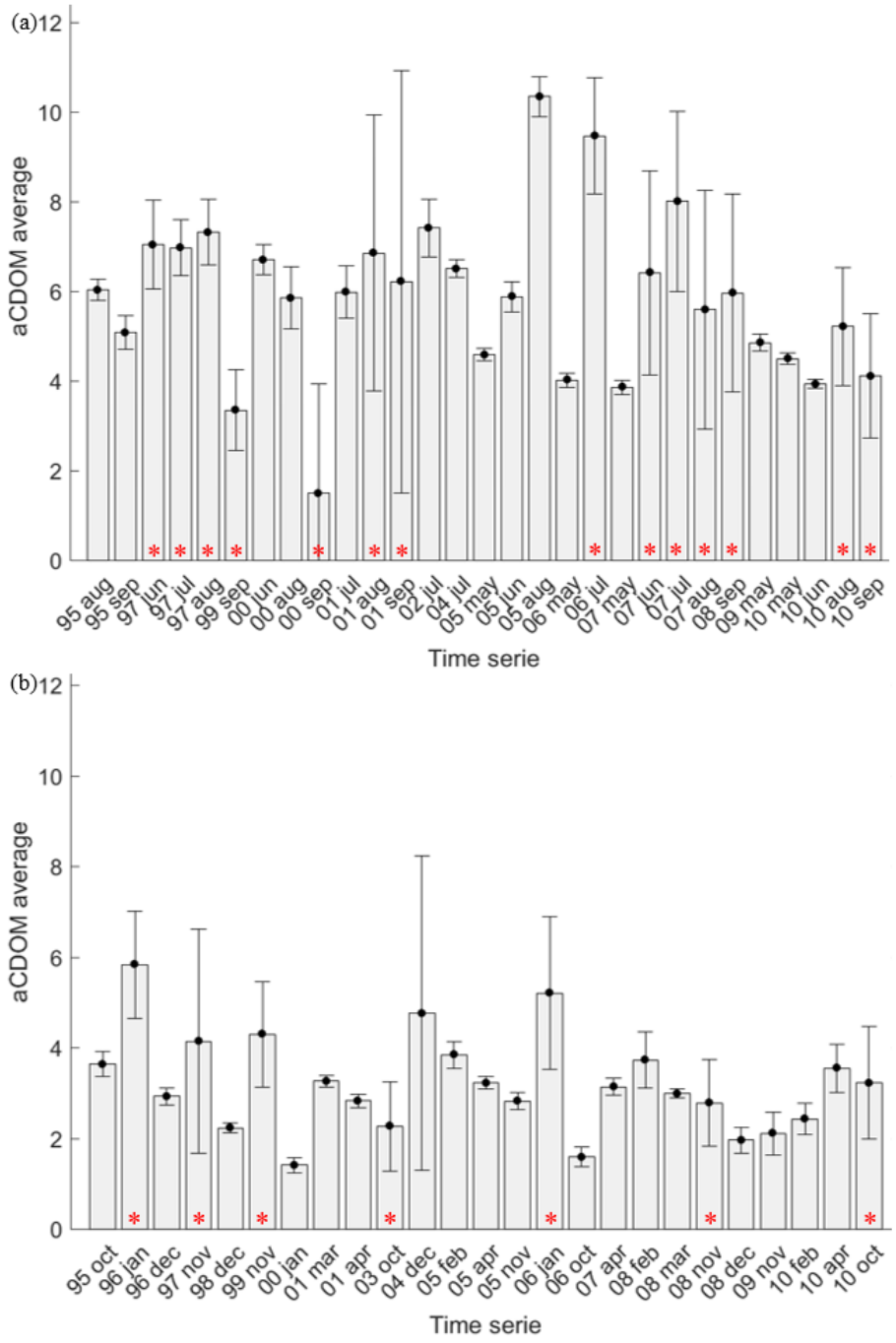

Figure 19. Plot of $a_{CDOM}(485)$ average by image over (a) dry and (b) wet periods. The images with bloom presence are identified with a red mark (*).

In the non-bloom images, $a_{CDOM}(485)$ distribution suggests a spatial-temporal homogeneity, with higher frequencies distributed among the low-to-medium scale (< 18 m^{-1}). This configuration was more common during the wet period images and for those taken after rainfall occurrence for both periods (Figure 15, Figure 20). The $a_{CDOM}(485)$ variations for these images probably are related to leaching that conduces the OM provided from LULC to the water body surface. Thus, in this situation the CDOM behavior is associated with terrestrial humic acids and environment dynamics.

It could be affirmed since elevation in $a_{CDOM}(485)$ values were observed in post rainfall images; it is associated with almost homogeneous CDOM distribution, and the LULC classification demonstrated a predominance of cultivated and managed terrestrial areas (Martins et al. 2016). This human activity strongly affects water bodies condition by nutrients, pesticides and OM inputs, but it is possible just after leaching (Allan 2004), that is generalized for the reservoir surrounding because of watershed geomorphology.

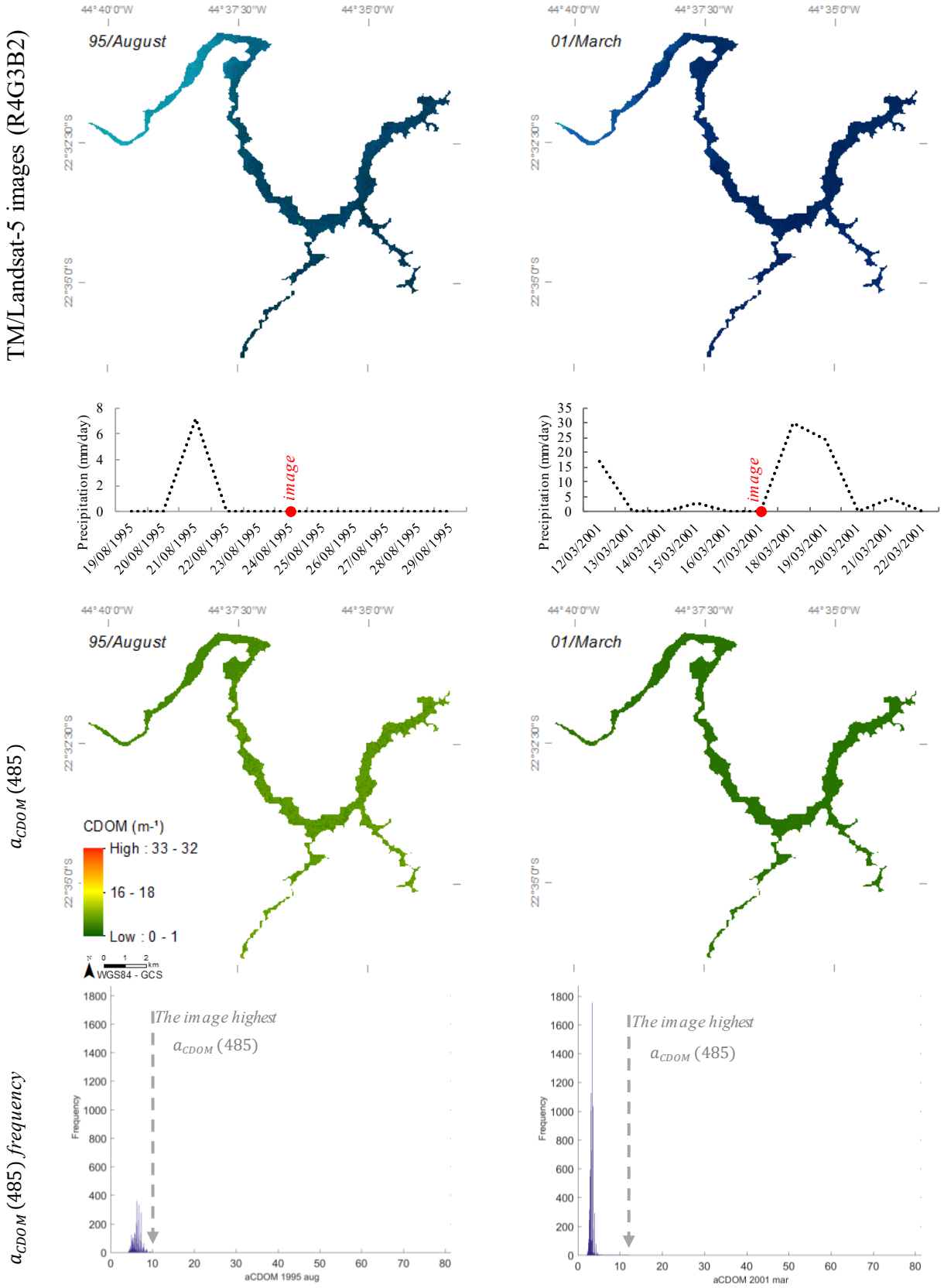

Figure 20. Examples of non-bloom situation in FHR using satellite images in R4G3B2 color composition, precipitation graphs, CDOM distribution (map) and frequencies (histograms). The precipitation graph indicates pluviometric condition for five days before and after the image capture (indicated by a red point).

In bloom images, the CDOM spatial-temporal distribution is related to the occurrence of $a_{CDOM}(485)$ peaks, that are associated with high CDOM levels (a_{CDOM}>20 m^{-1}). This patter was not found in non-bloom images, and it is also possible to visually identify a strong adequacy between $a_{CDOM}(485)$ high values and algal bloom patches in R4G3B2 satellite images composite color (Figure 21). This composition of bands makes possible to observe well Chl-*a* behavior, since it has a great reflectance in the near-infrared (NIR) and blue regions (TM/Landsat-5 B4 and B2, respectively) and also absorbs well in red region (B3) (Ahn and Shanmugam 2006; G. Wetzel 2001; Gurlin, Gitelson, and Moses 2011; John R. 2011; Kirk 2011), resulting in a pink color for bloom areas and almost blue for non-bloom areas, as it can be seen in Figure 21.

The referred correlation between Chl-*a* and a_{CDOM} high concentrations were more common for dry period images, but still did occur in the wet period. It could be associated with the phytoplankton occurrence, Chl-*a* elevation, and bacterial activity. Besides that, the water reduction or scarcity may increase the influence of effluent constitution in the water body, since it becomes less dissolved.

It is well known that CDOM is not directly produced from Chl-*a*, but both can be correlated in algal bloom situations. CDOM is a by-product of phytoplankton degradation by cyanobacteria, with DOM generation as consequence after DOC consumption. In this situation, CDOM starts to increase at approximately the half part of bloom period, and starts to reduce few days after the end of the bloom (Danhiez et al. 2017; Kirchman et al. 1991; Matsuoka et al. 2015; Rochelle-Newall and Fisher 2002; Sasaki et al. 2005; Zhang et al. 2009). Due to that a weak correlation (R^2 = 0.001) between measured $a_{CDOM}(\lambda)$ and field DOC was observed, which is common for complex inland waters, because the relationship between them is variable in space and time, and due to its strong dependence of the others water compounds concentration (Brezonik et al. 2015; Zhu and Yu 2013). It could also be justified because water sample collection was taken in an algal bloom situation, which indicates that a_{CDOM} should be more related to fulvic acids (Chl-*a*) than humic ones (DOC) (Xue, Zhang, and Duan 2016).

Additionally, surveys have demonstrated that a weak correlation between DOC and a_{CDOM} can be associated with a significant presence of Chl-*a*, suggesting a predominant production of non-chromophoric DOC in Chl-*a* maximum production areas even for water bodies where CDOM and DOC are generally high correlated (Rochelle-Newall and Fisher 2002). Thus, bloom images analyses suggest a significant correlation between Chl-*a* and a_{CDOM}, which is related to

phytoplankton degradation with huge bacterial activity; and the DOC constitution associated with the presence of non-chromophoric compounds.

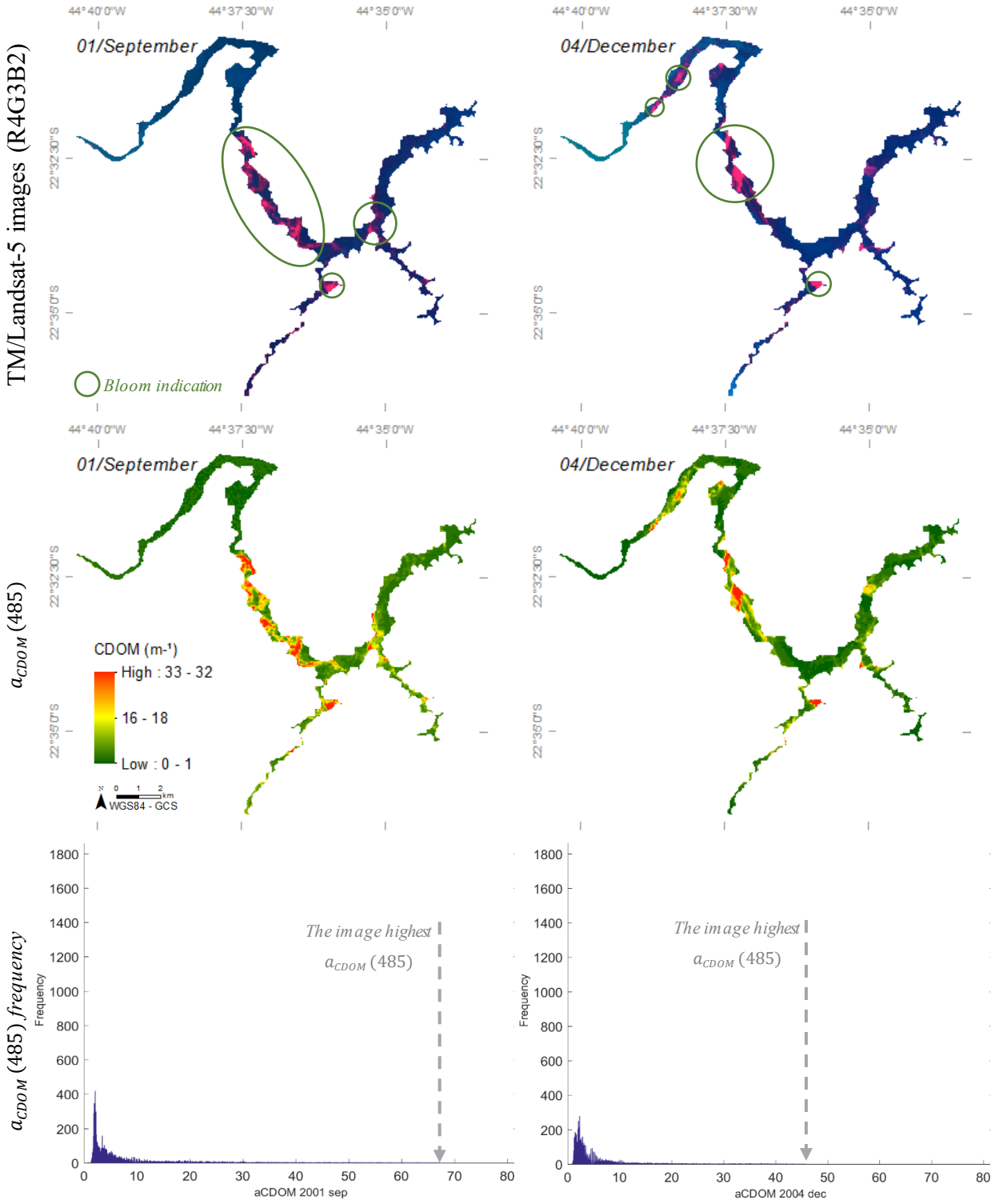

Figure 21. Examples of bloom situation images in FHR using satellite images in R4G3B2 color composition, CDOM distribution (map) and its frequencies (histograms). In satellite images composite color R4B3G2, the bloom patches appear as pink features which are evidenced by green circles.

4. CONCLUSION

This research proposed to evaluate a set of bio-optical models to retrieve $a_{\mathrm{CDOM}}(\lambda)$ spatial-temporal distribution in FHR using simulated reflectance data to TM/Landsat-5 imagery. It was also part of the goals to analyse the a_{CDOM} behavior to infer about its possible sources for tropical eutrophic water bodies. For achieving these goals, the performance of four QAA-BBHR configurations, six literature-based empirical models, and two new models were tested. Even though the literature-based empirical models and QAA-BBHR were developed to complex inland waters application, they failed to retrieve a_{CDOM} by simulated reflectance for the studied reservoir. Therefore, an alternative empirical model at 485 nm using B4/B1 ratio based on a linear regression was constructed, and demonstrated high sensibility to the water body behavior, and to environment variations.

Atmospheric correction, radiometric normalization, and reflectance data simulation played a huge role in making the developed model able to retrieve $a_{CDOM}(485)$ accurately from orbital data. Using atmospherically corrected and radiometrically normalized images, the noise registered at the sensor, the difference in acquisition geometry, and the variations in images illumination could be minimized. Besides that, reflectance data simulation made the model adequate to the TM sensor configuration, using the wavelengths where the signal is potentialized.

After application in orbital data, the proposed model was capable of identifying small variations in reflectance values, and to differentiate even slight alterations in CDOM quantities from TM/Landsat-5 data. Two different patterns of CDOM behavior could be observed: one associated with LULC and climate dynamics, relating a_{CDOM} $\mathrm{a_{CDOM}}$ distribution to leaching, and other linked to phytoplankton blooms and alterations in river discharges, which implies in correlation with Chl-*a* high concentration.

When CDOM had a homogeneous distribution, a relationship among rainfall, LULC, and CDOM smooth variance along all FHR was observed. This configuration was more common in wet period images and for those taken after rainfall occurrence for both periods. For this situation, CDOM concentration was more related to leaching, which conduces the organic matter provided from LULC to the FHR water surface. On this occasion, we had just a small elevation in CDOM values, and its concentration did not pass the medium scale.

A direct relationship between CDOM and Chl-*a* high concentration was also identified. In this situation, a_{CDOM} presented a heterogeneous distribution with occurrence of high and localized concentration, which could be associated with algal bloom shapes in satellite images, strongly

relating CDOM and Chl-*a* peaks just in bloom situations. This behavior was more frequent for dry period, but it was also identified in some wet period images.

It is well known that CDOM is not directly produced from Chl-*a*, but can be associated with the phytoplankton degradation by bacterial activity, with DOM generation as consequence. This situation can be associated with alterations in river discharges constitution that promotes the conditions to bloom occurrence. Besides that, the water reduction or scarcity related to dry periods may increase the influence of effluent in the water body, since it becomes less dissolved.

In conclusion, a model at a high wavelength (485 nm), using B1 and B4 TM bands can be precisely used to retrieve CDOM quantities for a tropical eutrophic reservoir using TM/Landsat-5 legacy. Furthermore, it could be verified that the sources of CDOM for the FHR can be related to LULC and Chl-*a*.

These results can support future water quality studies in FHR, providing its historical distribution of CDOM. They are also useful to support further evaluations in other reservoirs with the same environmental characteristics, in order to verify if the proposed model can provide CDOM distribution away from FHR.

4.5 ACKNOWLEDGMENTS

The authors would like to thank UNESP (São Paulo State University) and INRS (National Institute of Scientific Research) for the research infrastructure, INPE (Brazilian National Institute of Spatial Research) for yielding this research data. For financial support, we would like to thank CNPq (National Council of Scientific and Technological Development - Brazil) and ELAP (Emerging Leaders in the Americas Program - Canada). We would also like to thank the São Paulo Research Foundation (FAPESP) under grant number 2011/19523-8.

APPENDIX 1

QAA-BBHR formulation:

QAA-BBHR original formulation ($\lambda_0 = 709\ nm$) (Watanabe et al. 2016)			
Step	**Property**	**Formulation**	**Type**
0.	r_{rs}	$= \frac{R_{rs}}{(0{,}52 + 1{,}7R_{rs})} = g_0 u(\lambda) + g_1[u(\lambda)]^2$	SA
1.	$u(\lambda)$	$= \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} = \frac{-g_0+\sqrt{g_0^2+4g_1 r_{rs}(\lambda)}}{2g_1}$, $g_0 = 0.089$ $g_1 = 0{,}125$	SA
2.	$a(\lambda_0)$	$= a_w(\lambda_0) + 10^{-0{,}7702\ -0{,}0999\chi+0{,}0566\ \chi^2}$, $\chi = \log\left(\frac{r_{rs}(443) + r_{rs}(665)}{r_{rs}(\lambda_0) + 5\left(\frac{r_{rs}(620)}{r_{rs}(443)}\right) r_{rs}(620)}\right)$	EP
3.	$b_{bp}(\lambda_0)$	$= \frac{u(\lambda_0)a(\lambda_0)}{1-u(\lambda_0)} - b_{bw}(\lambda_0)$	AN
4.	Y	$= a\left(b - c * e^{\left(-d\frac{r_{rs}(\lambda_1)}{r_{rs}(\lambda_0)}\right)}\right)$	EP
5.	$b_{bp}(\lambda)$	$= b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^Y$	SA
6.	$a(\lambda)$	$= \frac{1-u(\lambda)(b_{bw}(\lambda)+b_{bp}(\lambda))}{u(\lambda)}$	AN
7.	ζ	$= \frac{a_{ph}(665)}{a_{ph}(\lambda_0)} = 0{,}3 + \frac{0{,}2}{0{,}8+\frac{r_{rs}(665)}{r_{rs}(\lambda_0)}}$	AP
8.	ξ	$= \frac{a_{dg}(411)}{a_{dg}(443)} = e^{S(443-411)}$, $S = 0{,}014 + \left(\frac{0{,}002}{0{,}6+\left(\frac{r_{rs}(443)}{r_{rs}(\lambda_0)}\right)}\right)$	SA
9.	$a_{dg}(443)$	$= \frac{a(411)-\zeta a(443)}{\xi-\zeta} - \frac{a_w(411)-\zeta a_w(443)}{\xi-\zeta} = \left(\frac{a(411)-\zeta a(443)}{\xi-\zeta}\right) - \left(\frac{a_w(411)-\zeta a_w(443)}{\xi-\zeta}\right)$, $a_{dg}(\lambda) = a_{dg}(443) * e^{-S(\lambda-443)}$	AN

EP= Empirical / SA = Semi-analical / AN = Analitical

where, λ corresponds to a non-specific wavelength, λ_0 indicates the algorithm reference wavelength, where it should be the highest water absorption value, R_{rs} is the remote sensing reflectance at the surface obtained by Kirk (2011); r_{rs} is the reflectance under surface obtained from R_{rs}; u corresponds to the relation between absorption and backscattering that can be obtained by r_{rs} and the constants g_0 and g_1; χ indicates the relationship among r_{rs} and the spectral slope of particles and pure water backscattering (b_{bp} and b_{bw}); Y is related to decay coefficient of particles; a, b, c are adjustment parameters; ξ are the variable used to obtain $a_{dg}(\lambda)$ and ζ is the variable used to calculate $a_{ph}(\lambda)$.

+Steps used to obtain **QAA-BBHR+CDOM** (Watanabe et al. 2016; Zhu and Yu 2013)		
9.	$a_p(440)$	$= j_1 * b_{bp}(\lambda_0)^{j_2}$
10.	$a_{CDOM}(440) =$	$=a(440) - a_w(440) - a_p(440)$
Where, j_1 and j_2 are constant with values defined as 63 e 0,88 (Zhu et al. 2011, 2013).		

Empirical models formulation (Zhu et al. 2014):

Model	**Formulation**
Mannino	$a_{CDOM}(443) = -0{,}0736 \ln\left(\frac{0{,}408 R_{rs}(B1)}{R_{rs}(B2)} - 0{,}173\right)$
D'sa	$a_{CDOM}(412) = 0{,}134\left(\frac{R_{rs}(B1)}{R_{rs}(B2)}\right)^{-2{,}025}$
Griffin	$a_{CDOM}(400) = \exp\left(-1{,}145 + 26{,}529 * B3 + 0{,}603\left(\frac{B2}{B1}\right)\right)$
Kutser	$a_{CDOM}(420) = 5{,}13\left(\frac{B2}{B3}\right)^{-2{,}67}$
Castillo	$a_{CDOM}(412) = -0{,}90\left(\frac{R_{rs}(B2)}{R_{rs}(B3)}\right) + 2{,}34$
Ficek	$a_{CDOM}(440) = 3.65\left(\frac{R_{rs}(B2)}{R_{rs}(B3)}\right)^{-1.93}$
where, Bx are the TM/Landsat-5 bands.	

CHAPTER 5: Final Conclusions

This research proposed to identify and to study variations in a_{CDOM} at the water surface from 1995 to 2010 using TM/Landsat-5 data. Besides that, it was verified the LULC contribution for a_{CDOM} historic distribution changes. According to the presented results, we can affirm that all objectives of this research were accomplished.

In regard to the referred aims, two main steps of research development could be differentiated: one to accurately obtain LULC in FHR watershed (*CHAPTER 3: Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)*), and other to study a set of bio-optical models to retrieve $a_{CDOM}(\lambda)$ spatial-temporal distribution in Funil hydroelectric reservoir water surface by simulated reflectance data to TM sensor (CHAPTER 4: Spatiotemporal distribution of $\boldsymbol{a_{CDOM}}$ in a tropical eutrophic reservoir). In this last step we could also analyse the *v* historical behavior to infer about its possible sources for a tropical eutrophic inland water body.

In an effort to obtain LULC historical evolution, the SVM parameters were studied and tested to ensure the optimal configuration for this algorithm, in order to allow accurately mapping the main changes in FHR watershed LULC. As results for SVM performance in this study area, the highest accuracy was obtained using C-value =100, the lowest basis function parameter (γ = 0.1), and pyramid depths =2. The change detection analyses demonstrate that the used SVM parameterization allows a great differentiation among LULC classes, supporting this algorithm premise of low sensitivity to sample size and high sensitivity to representative training sample selection. In addition, the parameterization obtained in this work made the algorithm able to differentiate lengthy and thin areas, as borders, and not continuous small areas located inside wide classes.

In conclusion, the LULC mapping performed using the obtained SVM parameterization was efficient and provided trustful and real results, ensuring high values of Kappa statistics (> 96%, except for 2000 image = 81%), and accuracy (overall > 98%; producer > 76%; user > 85%, except for 2000 image: 86%, > 34%, > 76%, respectively), allowing high level of confidence to the obtained LULC mapping and change detection analysis. The LULC change detection demonstrated that no strong changes were observed over the time series used by this research, and this classification also demonstrated the predominance of Cultivated and Managed Terrestrial Areas (CMTA), that means agricultural and livestock class major occurrence.

With regard to a_{CDOM} retrieving and its possible sources study, a bio-optical model set composed by QAA-BBHR and empirical models was tested and analysed. Even though the

literature-based empirical models and QAA-BBHR were developed to complex inland waters application, they failed in retrieving$a_{\mathrm{CDOM}}(\lambda)$by simulated R_{rs} for FHR. Therefore, an alternative empirical model at 485 nm using B4/B1 ratio based on a linear regression was constructed, and demonstrated high sensibility to the water body behavior, and to environment variations (R^2 = 0.91, p-value < α, NRMSE=0.09, %RMSE=7.27, Bias=0.0008, Nash=0.91).

The proposed model was capable of identifying small variations in reflectance values, making possible to differentiate even slight alterations in a_{CDOM} from TM/Landsat-5 historical imagery. Two great different patterns were observed in historical a_{CDOM} behavior: one associated with LULC and climate dynamics, relatinga_{CDOM} distribution to leaching; and other linked to phytoplankton bloom shapes and alterations in river discharges, implying in correlation with Chl-a high concentration.

When a_{CDOM} could be related to LULC, a homogeneous distribution and smooth variance along all FHR was observed, relating organic matter discharges in water to leaching situations occurred after rainfall, which were more common in wet period images. On this occasion, we had just a small elevation in a_{CDOM} values that did not pass the medium scale (a_{CDOM} < 15 m^{-1}).

When a direct relationship between a_{CDOM} and Chl-a high concentration was identified, a_{CDOM} presented a heterogeneous distribution with occurrence of high (a_{CDOM}> 20 m^{-1}) and punctual concentration, which could be associated with algal bloom shapes in satellite images. This behavior was more frequent for dry period, but it also occurred in wet period images, but never after rainfall situations. It is well known that a_{CDOM} is not directly produced from Chl-a, but can be associated with phytoplankton degradation by bacterial activity, with DOM generation as consequence. This situation can be associated with alterations in the river discharges constitution, and to increase the effluent influence in water during dry periods or in no-rainfall images, since it becomes less dissolved.

In conclusion, a model at high wavelength (485 nm), using B1 and B4 TM/Landsat-5 bands can be used to retrieve a_{CDOM} using TM/Landsat-5 legacy. Furthermore, it could be verified that the sources of CDOM for the FHR can be related to LULC and Chl-a.

It is important to highlight that atmospheric correction, radiometric normalization, and reflectance data simulation played a huge role in making the images appropriated to algorithms application in orbital data for both research steps (mapping LULC and retrieving CDOM). Using atmospherically corrected and radiometrically normalized images, the noise registered at

the sensor, the difference in acquisition geometry, and the variations in images illumination could be minimized.

All things considered, we must refuse this work hypothesis that associated CDOM alterations mainly to LULC changes since a great correlation with Chl-a was also found. Thus, we can affirm that, both LULC and Chl-a can be related as CDOM source for FHR according equal proportion depending on weather conditions, LULC predominant class, effluent discharges, algal bloom occurrence, and cyanobacterial activity.

In summary, these results can support future water quality studies in FHR, providing its a_{CDOM} historical distribution and main sources. They can also be useful to support further works in other reservoirs with the same environmental characteristics, in order to indicate an alternative wavelength/spectral region to retrieve a_{CDOM}; an unusual band ratio to derive it; the possible sources of CDOM depending on the weather behavior; and also providing a way to accurately map LULC to obtain terrestrial C contribution for a given water body.

REFERENCES

Adler-Golden, Steven M. et al. 1999. "Atmospheric Correction for Shortwave Spectral Imagery Based on MODTRAN4." *Imaging Spectrometry V* 3753(July): 61–69. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=994756.

Adrian, Rita et al. 2009. "Lakes as Sentinels of Climate Change." *Limnology and Oceanography* 54(6 part 2): 2283–97.

Ahn, Yu Hwan, and Palanisamy Shanmugam. 2006. "Detecting the Red Tide Algal Blooms from Satellite Ocean Color Observations in Optically Complex Northeast-Asia Coastal Waters." *Remote Sensing of Environment* 103(4): 419–37.

Allan, J David. 2004. "Landscapes and Riverscapes : The Influence of Land Use on Stream Ecosystems." *Annual Review of Ecology, Evolution, and Systematics* 35(2004): 257–84. http://www.jstor.org/stable/30034117.

Bastviken, D. et al. 2011. "Freshwater Methane Emissions Offset the Continental Carbon Sink." *Science* 331(6013): 50–50. http://www.sciencemag.org/cgi/doi/10.1126/science.1196808.

Branco, Christina W C et al. 2002. "Limnological Features of Funil Reservoir (R.J., Brazil) and Indicator Properties of Rotifers and Cladocerans of the Zooplankton Community." *Lakes and Reservoirs: Research and Management* 7(2): 87–92. http://doi.wiley.com/10.1046/j.1440-169X.2002.00177.x.

Brezonik, Patrick L, Leif G Olmanson, Jacques C Finlay, and Marvin E Bauer. 2015. "Factors Affecting the Measurement of CDOM by Remote Sensing of Optically Complex Inland Waters." *Remote Sensing of Environment* 157: 199–215. http://dx.doi.org/10.1016/j.rse.2014.04.033.

Briassoulis, Helen. 2000. "Analyses of Land Use Change: Theoretical and Modeling Approaches." In *The Web Book of Regional Science.*, Regional Research Institute, Wes Virginia University. http://www.rri.wvu.edu/webbook/briassoulis/contents.htm.

Canty, Morton J. 2009. *Image Analysis, Classification, and Change Detection in Remote Sensing: With Algorithms for ENVI/IDL*. 2nd ed. Boca Raton: CRC Press.

Canty, Morton J, and Allan A Nielsen. 2008. "Automatic Radiometric Normalization of Multitemporal Satellite Imagery with the Iteratively Re-Weighted MAD Transformation." *Remote Sensing of Environment* 112(3): 1025–36.

Canty, Morton J, Allan A Nielsen, and Michael Schimidt. 2004. "Automatic Radiometric Normalizatio of Multitemporal Satellite Imagery." *Remote Sensing of Environment* 91: 441–51.

Carpenter, Stephen R. et al. 2007. "Understanding Regional Change: A Comparison of Two Lake Districts." *BioScience* 57(4): 323. https://academic.oup.com/bioscience/article-lookup/doi/10.1641/B570407.

Del Castillo, Carlos E., and Richard L. Miller. 2008. "On the Use of Ocean Color Remote Sensing to Measure the Transport of Dissolved Organic Carbon by the Mississippi River Plume." *Remote Sensing of Environment* 112(3): 836–44. http://linkinghub.elsevier.com/retrieve/pii/S0034425707002933.

Chander, Gyanesh, Brian L Markham, and Dennis L Helder. 2009. "Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors." *Remote Sensing of Environment* 113(5): 893–903. http://dx.doi.org/10.1016/j.rse.2009.01.007.

Chen, Jiang, Wei-ning Zhu, Yong Q Tian, and Qian Yu. 2017. "Estimation of Colored Dissolved Organic Matter From Landsat-8 Imagery for Complex Inland Water: Case Study of Lake Huron." *IEEE Transactions on Geoscience and Remote Sensing* (January): 1–12. http://ieeexplore.ieee.org/document/7815300/.

Chen, Shuguo, and Tinglu Zhang. 2015. "Evaluation of a QAA-Based Algorithm Using

MODIS Land Bands Data for Retrieval of IOPs in the Eastern China Seas." *Optics Express* 23(11): 13953. https://www.osapublishing.org/abstract.cfm?URI=oe-23-11-13953.

Cheng Feng Le et al. 2009. "Validation of a Quasi-Analytical Algorithm for Highly Turbid Eutrophic Water of Meiliang Bay in Taihu Lake, China." *IEEE Transactions on Geoscience and Remote Sensing* 47(8): 2492–2500. http://ieeexplore.ieee.org/document/4815462/.

Cole, J J et al. 2007. "Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget." *Ecosystems* 10(1): 172–85. http://link.springer.com/10.1007/s10021-006-9013-8.

Congalton, Russell G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37(1): 35–46.

Curtarelli, M. P., E. H. Alcântara, C. D. Rennó, and J. L. Stech. 2014. "Physical Changes within a Large Tropical Hydroelectric Reservoir Induced by Wintertime Cold Front Activity." *Hydrology and Earth System Sciences* 18(8): 3079–93. http://www.hydrol-earth-syst-sci.net/18/3079/2014/.

D'Sa, Eurico J, and Richard L Miller. 2003. "Bio-Optical Properties in Waters Influenced by the Mississippi River during Low Flow Conditions." *Remote Sensing of Environment* 84(4): 538–49. http://linkinghub.elsevier.com/retrieve/pii/S0034425702001633.

Danhiez, F P et al. 2017. "Optical Properties of Chromophoric Dissolved Organic Matter during a Phytoplankton Bloom. Implication for DOC Estimates from CDOM Absorption." *Limnology and Oceanography* (2015): 1–17. http://doi.wiley.com/10.1002/lno.10507.

Denman, K.L., and G. Brasseur. 2007. "Climate Change 2007: The Physical Science Basis." In *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.*, eds. S. SOLOMON et al. New York: Cambridge Univesitya Press. https://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm.

Dimyati, Muh, Kei Mizuno, Shintaro Kobayashi, and Teitaro Kitamura. 1996. "An Analysis of Land Use/cover Change in Indonesia." *International Journal of Remote Sensing* 17(5): 931–44. http://www.tandfonline.com/doi/abs/10.1080/01431169608949056.

Effler, Adam J P et al. 2015. "Linking CDOM Patterns in Cayuga Lake, New York, USA, to Terrigenous Inputs." *Inland Waters* 5(4): 355–70.

Esteves, Francisco de Assis. 1998. *Fundamentos de Limnologia*. 2nd ed. Rio de Janeiro: Interciência.

FAO. 1994. "Land Cover Classification System."

Ficek, Dariusz, Tomasz Zapadka, and Jerzy Dera. 2011. "Remote Sensing Reflectance of Pomeranian Lakes and the Baltic**The Study Was Partially Financed by MNiSW (Ministry of Science and Higher Education) as a Research Project N N306 066434 in the Years 2008–2011. The Partial Support for This Study Was Also pro." *Oceanologia* 53(4): 959–70. http://linkinghub.elsevier.com/retrieve/pii/S0078323411500338.

G. Wetzel, Robert. 2001. Academic Press Elsevier *Limnology Lake and River Ecosystems*.

Gordon, Howard R., and André Y. Morel. 1983. 4 *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*. Washington, D. C.: American Geophysical Union. http://doi.wiley.com/10.1029/LN004.

Griffin, Claire G, Karen E Frey, John Rogan, and Robert M Holmes. 2011. "Spatial and Interannual Variability of Dissolved Organic Matter in the Kolyma River, East Siberia, Observed Using Satellite Imagery." *Journal of Geophysical Research: Biogeosciences* 116(3): 1–12.

Guedes, Iamê Alves et al. 2014. "Fluctuations in Microcystin Concentrations, Potentially Toxic Microcystis and Genotype Diversity in a Cyanobacterial Community from a Tropical

Reservoir." *Harmful Algae* 39: 303–9.

Guerra, Antônio José T., and Rosangela G.M. Botelho. 2006. "Erosão Dos Solos." In *Geomorfologia Do Brasil.*, eds. Sandra B. Cunha and Antônio José T. Guerra. Rio de Janeiro: Bertrand Brasil.

Gurlin, Daniela, Anatoly A. Gitelson, and Wesley J. Moses. 2011. "Remote Estimation of Chl-a Concentration in Turbid Productive Waters — Return to a Simple Two-Band NIR-Red Model?" *Remote Sensing of Environment* 115(12): 3479–90. http://www.sciencedirect.com/science/article/pii/S0034425711003130 (June 1, 2017).

Hadjimitsis, Diofantos G, and Chris Clayton. 2009. "Assessment of Temporal Variations of Water Quality in Inland Water Bodies Using Atmospheric Corrected Satellite Remotely Sensed Image Data." *Environmental Monitoring and Assessment* 159(1–4): 281–92.

Huang, C., L. S. Davis, and J. R. G. Townshend. 2002. "An Assessment of Support Vector Machines for Land Cover Classification." *International Journal of Remote Sensing* 23(4): 725–49. http://www.tandfonline.com/doi/abs/10.1080/01431160110040323.

INEA. 2012. "Avaliação Ambiental Do Rio Paraíba Do Sul: Trecho Funil Santa Cecília." : 109.

———. 2016. "Reservatório de Funil. Gestão de Qualidade Da Água."

John R., Jensen. 2011. *Remote Sensing of the Environment: An Earth Resource Perspective*. 2nd ed. Minneapolis: Pearson Prentice Hall.

Kavzoglu, T, and I Colkesen. 2009. "A Kernel Functions Analysis for Support Vector Machines for Land Cover Classification." *International Journal of Applied Earth Observation and Geoinformation* 11(5): 352–59.

Kidder, Stanley Q., and Thomas H. Vonder Haar. 1995. *Satellite Meteorology an Introduction*. San Diego: Academic Prees.

Kirchman, David, Yoshimi Suzuki, Christopher Garside, and W. Hugh Ducklow. 1991. "High Turnover Rates of Dissolved Organic Carbon during a Spring Phytoplankton Bloom." *Letters of Nature*: 612–14.

Kirk, John T. O. 2011. *Light and Photosynthesis in Aquatic Ecosystems*. 3rd ed. Cmbridge: United Kingdom at the University Press.

Krinner, Gerhard. 2003. "Impact of Lakes and Wetlands on Boreal Climate." *Journal of Geophysical Research* 108(D16): 4520. http://doi.wiley.com/10.1029/2002JD002597.

Kutser, Tiit et al. 2005. "Mapping Lake CDOM by Satellite Remote Sensing." *Remote Sensing of Environment* 94(4): 535–40.

———. 2016. "Mapping Inland Water Carbon Content with Landsat 8 Data." *International Journal of Remote Sensing* 37(13): 2950–61.

Kutser, Tiit, Charles Verpoorter, Birgot Paavel, and Lars J Tranvik. 2015. "Estimating Lake Carbon Fractions from Remote Sensing Data." *Remote Sensing of Environment* 157: 138–46. http://dx.doi.org/10.1016/j.rse.2014.05.020.

Le, Chengfeng, and Chuanmin Hu. 2013. "A Hybrid Approach to Estimate Chromophoric Dissolved Organic Matter in Turbid Estuaries from Satellite Measurements: A Case Study for Tampa Bay." *Optics express* 21(16): 18849–71. http://www.ncbi.nlm.nih.gov/pubmed/23938799.

Lee, Zhongping, and Kendall L Carder. 2004. "Absorption Spectrum of Phytoplankton Pigments Derived from Hyperspectral Remote-Sensing Reflectance." *Remote Sensing of Environment* 89(3): 361–68. http://linkinghub.elsevier.com/retrieve/pii/S0034425703002876.

Lee, Zhongping, Kendall L Carder, and Robert A Arnone. 2002. "Deriving Inherent Optical Properties from Water Color: A Multiband Quasi-Analytical Algorithm for Optically Deep Waters." *Applyed optics* 41(27): 5755–72.

Lillesand, Thomas, Ralph W. Kiefer, and Jonathan Chipman. 1999. *Remote Sensing and Image Interpretation*. 4th ed. Hoboken: John Wiley & Sons.

Lu, D., and Q. Weng. 2007. "A Survey of Image Classification Methods and Techniques for Improving Classification Performance." *International Journal of Remote Sensing* 28(5): 823–70. http://www.tandfonline.com/doi/abs/10.1080/01431160600746456.

Luyssaert, Sebastiaan et al. 2008. "Old-Growth Forests as Global Carbon Sinks." *Nature* 455(7210): 213–15. http://www.nature.com/doifinder/10.1038/nature07276.

Mannino, Antonio, Mary E Russ, and Stanford B Hooker. 2008. "Algorithm Development and Validation for Satellite-Derived Distributions of DOC and CDOM in the U.S. Middle Atlantic Bight." *Journal of Geophysical Research: Oceans* 113(7): 1–19.

Mantero, P., G. Moser, and S.B. Serpico. 2005. "Partially Supervised Classification of Remote Sensing Images through SVM-Based Probability Density Estimation." *IEEE Transactions on Geoscience and Remote Sensing* 43(3): 559–70. http://ieeexplore.ieee.org/document/1396328/.

Marengo, José A., and Alves Lincoln M. 2005. "Tnedências Hidrológicas Da Bacia Do Rio Paraíba Do Sul." *Revista Barsileira de Meteorologia* 20(2): 215–26. http://www.inea.rj.gov.br/cs/groups/public/documents/document/zwew/mtez/~edisp/inea0113198.pdf.

Martins, Sarah, Nariane Bernardo, Igor Ogashawara, and Enner Alcantara. 2016. "Support Vector Machine Algorithm Optimal Parameterization for Change Detection Mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)." *Modeling Earth Systems and Environment* 2(3): 138. http://link.springer.com/10.1007/s40808-016-0190-y.

Matsuoka, Atsushi et al. 2015. "Characteristics of Colored Dissolved Organic Matter (CDOM) in the Western Arctic Ocean: Relationships with Microbial Activities." *Deep-Sea Research Part II: Topical Studies in Oceanography* 118: 44–52. http://dx.doi.org/10.1016/j.dsr2.2015.02.012.

Mendonça, Francisco, and Inês M. Danni-Oliveira. 2007. *Climatologia: Noções Básicas E Climas Do Brasil*. 1st ed. Rio de Janeiro: Oficina de Textos.

Mobley, C.D. 2001. "Radiative Transfer in the Ocean." In *Encyclopedia of Ocean Sciences*, Elsevier, 2321–30. http://linkinghub.elsevier.com/retrieve/pii/B012227430X004694.

Montanher, Otávio C et al. 2014. "Empirical Models for Estimating the Suspended Sediment Concentration in Amazonian White Water Rivers Using Landsat 5/TM." *International Journal of Applied Earth Observation and Geoinformation* 29(1): 67–77. http://dx.doi.org/10.1016/j.jag.2014.01.001.

Morel, André. 1988. "Optical Modeling of the Upper Ocean in Relation to Its Biogenous Matter Content (Case I Waters)." *Journal of Geophysical Research* 93(C9): 10749. http://doi.wiley.com/10.1029/JC093iC09p10749.

Morel, André, and Hubert Loisel. 1998. "Apparent Optical Properties of Oceanic Water: Dependence on the Molecular Scattering Contribution." *Applied Optics* 37(21): 4765. https://www.osapublishing.org/abstract.cfm?URI=ao-37-21-4765.

Morel, Andrey Y., and Howard R. Gordon. 1979. "Report of the Working Group on Water Color."

Moses, Wesley J et al. 2012. "Estimation of Chlorophyll-a Concentration in Turbid Productive Waters Using Airborne Hyperspectral Data." *Water Research* 46(4): 993–1004. http://dx.doi.org/10.1016/j.watres.2011.11.068.

Mountrakis, Giorgos, Jungho Im, and Caesar Ogole. 2011. "Support Vector Machines in Remote Sensing: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 66(3): 247–59.

Mueller, James L. 2000. "In-Water Radiometric Profile Measurements and Data Analysis Protocols." In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*, eds. J. L. Mueller et al. Greenbelt: NASA/TM - National Aeronautical and Space administration, 13–20.

Ogashawara, Igor et al. 2014. "Interactive Correlation Environment (ICE) — A Statistical Web Tool for Data Collinearity Analysis." *Remote Sensing* 6(4): 3059–74. http://www.mdpi.com/2072-4292/6/4/3059/.

———. 2016. "Re-Parameterization of a Quasi-Analytical Algorithm for Colored Dissolved Organic Matter Dominant Inland Waters." *International Journal of Applied Earth Observations and Geoinformation* 53: 128–45. http://dx.doi.org/10.1016/j.jag.2016.09.001.

Pacheco, F. S. et al. 2015. "The Effects of River Inflow and Retention Time on the Spatial Heterogeneity of Chlorophyll and Water-Air CO2 Fluxes in a Tropical Hydropower Reservoir." *Biogeosciences* 12(1): 147–62.

Pacheco, Felipe S., Fabio Roland, and John A. Downing. 2014. "Eutrophication Reverses Whole-Lake Carbon Budgets." *Inland Waters* 4(1): 41–48.

Pal, M., and P. M. Mather. 2005. "Support Vector Machines for Classification in Remote Sensing." *International Journal of Remote Sensing* 26(5): 1007–11. http://www.tandfonline.com/doi/abs/10.1080/01431160512331314083.

Rabus, Bernhard, Michael Eineder, Achim Roth, and Richard Bamler. 2003. "The Shuttle Radar Topography Mission—a New Class of Digital Elevation Models Acquired by Spaceborne Radar." *ISPRS Journal of Photogrammetry and Remote Sensing* 57(4): 241–62. http://linkinghub.elsevier.com/retrieve/pii/S0924271602001247.

Rochelle-Newall, E.J J., and T.R R. Fisher. 2002. "Chromophoric Dissolved Organic Matter and Dissolved Organic Carbon in Chesapeake Bay." *Marine Chemistry* 77(1): 23–41.

Rosenzweig, Cynthia et al. 2007. "Assessment of Observed Changes and Responses in Natural and Managed Systems." In *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* eds. M.L. Parry et al. Cmabridge: Cambridge University Press, 79–131.

Sasaki, Hiroaki, Tsuyoshi Miyamura, Sei Ichi Saitoh, and Joji Ishizaka. 2005. "Seasonal Variation of Absorption by Particles and Colored Dissolved Organic Matter (CDOM) in Funka Bay, Southwestern Hokkaido, Japan." *Estuarine, Coastal and Shelf Science* 64(2–3): 447–58.

Schowengerdt, Robert A. 2006. *Remote Sensing: Models and Methods for Image Processing*. Academic Press.

Schroeder, Todd A et al. 2006. "Radiometric Correction of Multi-Temporal Landsat Data for Characterization of Early Successional Forest Patterns in Western Oregon." *Remote Sensing of Environment* 103(1): 16–26.

Taskin, Kavzoglu, and Reis Selcuk. 2008. "Performance Analysis of Maximum Likelihood and Artificial Neural Network Classifiers for Training Sets with Mixed Pixels Preview." *GIScience* 45(3): 330–42.

Tewkesbury, Andrew P. et al. 2015. "A Critical Synthesis of Remotely Sensed Optical Image Change Detection Techniques." *Remote Sensing of Environment* 160: 1–14. http://linkinghub.elsevier.com/retrieve/pii/S0034425715000152.

Tranvik, Lars J et al. 2009. "Lakes and Reservoirs as Regulators of Carbon Cycling and Climate." *Limnology and Oceanography* 54(6part2): 2298–2314. http://doi.wiley.com/10.4319/lo.2009.54.6_part_2.2298.

Usman, M., R. Liedl, M. A. Shahid, and A. Abbas. 2015. "Land Use/land Cover Classification and Its Change Detection Using Multi-Temporal MODIS NDVI Data." *Journal of Geographical Sciences* 25(12): 1479–1506. http://link.springer.com/10.1007/s11442-015-1247-y.

Ustuner, Mustafa, Fusun Balik Sanli, and Barnali Dixon. 2015. "Application of Support Vector Machines for Landuse Classification Using High-Resolution RapidEye Images: A

Sensitivity Analysis." *European Journal of Remote Sensing* 48(1): 403–22. https://www.tandfonline.com/doi/full/10.5721/EuJRS20154823.

Vapnik, Vladimir. 1995. *The Nature of Statistical Learning Theory*. New York: Springer.

Verburg, Peter H., Kathleen Neumann, and Linda Nol. 2011. "Challenges in Using Land Use and Land Cover Data for Global Change Studies." *Global Change Biology* 17(2): 974–89. http://doi.wiley.com/10.1111/j.1365-2486.2010.02307.x.

Walsh, Stephen J. et al. 2015. "A Reservoir Siting Tool for North Carolina: System Design & Operations for Screening and Evaluation." *Applied Geography* 60: 139–49. http://linkinghub.elsevier.com/retrieve/pii/S0143622815000818.

Watanabe, Fernanda et al. 2016. "Parametrization and Calibration of a Quasi-Analytical Algorithm for Tropical Eutrophic Waters." *ISPRS Journal of Photogrammetry and Remote Sensing* 121: 28–47. http://linkinghub.elsevier.com/retrieve/pii/S0924271616303410.

Xue, Kun, Yuchao Zhang, and Hongtao Duan. 2016. "Variability of Light Absorption Properties in Optically Complex Inland Waters of Lake Chaohu, China." *Journal of Great Lakes Research*.

Yang, Wei et al. 2013. "Retrieval of Inherent Optical Properties for Turbid Inland Waters From Remote-Sensing Reflectance." *IEEE Transactions on Geoscience and Remote Sensing* 51(6): 3761–73.

Zhang, Yunlin et al. 2009. "The Contribution of Phytoplankton Degradation to Chromophoric Dissolved Organic Matter (CDOM) in Eutrophic Shallow Lakes: Field and Experimental Evidence." *Water Research* 43(18): 4685–97. http://dx.doi.org/10.1016/j.watres.2009.07.024.

Zhu, Weining et al. 2011. "Estimation of Chromophoric Dissolved Organic Matter in the Mississippi and Atchafalaya River Plume Regions Using above-Surface Hyperspectral Remote Sensing." *Journal of Geophysical Research* 116(C2): C02011. http://doi.wiley.com/10.1029/2010JC006523.

———. 2014. "An Assessment of Remote Sensing Algorithms for Colored Dissolved Organic Matter in Complex Freshwater Environments." *Remote Sensing of Environment* 140: 766–78. http://dx.doi.org/10.1016/j.rse.2013.10.015.

Zhu, Weining, Yong Q Tian, Qian Yu, and Brian L Becker. 2013. "Using Hyperion Imagery to Monitor the Spatial and Temporal Distribution of Colored Dissolved Organic Matter in Estuarine and Coastal Regions." *Remote Sensing of Environment* 134: 342–54. http://dx.doi.org/10.1016/j.rse.2013.03.009.

Zhu, Weining, and Qian Yu. 2013. "Inversion of Chromophoric Dissolved Organic Matter From EO-1 Hyperion Imagery for Turbid Estuarine and Coastal Waters." *IEEE Transactions on Geoscience and Remote Sensing* 51(6): 3286–98.