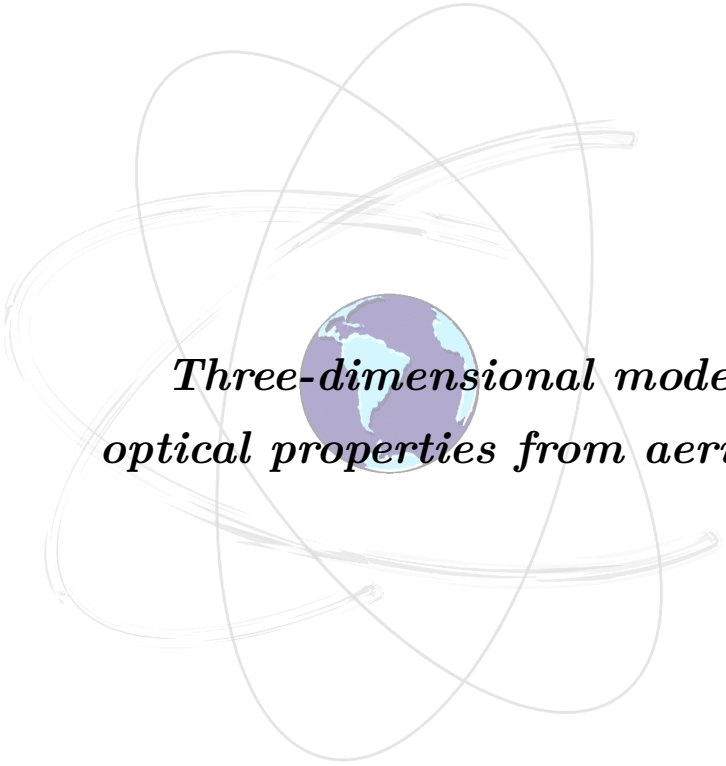
SÃO PAULO STATE UNIVERSITY
CÂMPUS PRESIDENTE PRUDENTE
SCHOOL OF SCIENCES AND TECHNOLOGY
GRADUATE PROGRAM OF CARTOGRAPHIC SCIENCES

Alisson Fernando Coelho do Carmo



Three-dimensional modeling of inland waters optical properties from aerial hyperspectral images.

Dissertation (Doctoral)

Presidente Prudente
2019

Alisson Fernando Coelho do Carmo

Three-dimensional modeling of inland waters optical properties from aerial hyperspectral images.

Doctoral dissertation presented to the Graduate Program of Cartographic Sciences of the São Paulo State University, Campus Presidente Prudente to obtain the Doctoral Degree in Cartographic Sciences.

São Paulo State University (UNESP),

School of Sciences and Technology (FCT),

Graduate Program of Cartographic Sciences (PPGCC),

Campus Presidente Prudente – Brazil

Supervisor: Prof. Dr. Nilton Nobuhiro Imai

Co-supervisor: Prof. Dr. Milton Hirokazu Shimabukuro

Presidente Prudente

2019

C287t

Carmo, Alisson Fernando Coelho do
Three-dimensional modeling of inland waters optical properties from aerial hyperspectral images / Alisson Fernando Coelho do Carmo. -- Presidente Prudente, 2019
90 f.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente
Orientador: Nilton Nobuhiro Imai
Coorientador: Milton Hirokazu Shimabukuro

1. Remote sensing. 2. Image processing. 3. Quality of water. 4. Underwater radiometric data. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

unesp

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: Three-dimensional modeling of inland waters optical properties from aerial hyperspectral images

AUTOR: ÁLISSON FERNANDO COELHO DO CARMO
ORIENTADOR: NILTON NOBUHIRO IMAI
COORIENTADOR: MILTON HIROKAZU SHIMABUKURO

Aprovado como parte das exigências para obtenção do Título de Doutor em CIÊNCIAS CARTOGRÁFICAS, área: Aquisição, Análise e Representação de Informações Espaciais pela Comissão Examinadora:

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

Prof. Dr. ARCILAN TREVENZOLI ASSIREU
Instituto de Recursos Naturais / Universidade Federal de Itajubá

Prof. Dr. MOACIR ANTONELLI PONTI
Departamento de Ciências da Computação / Universidade de São Paulo

Profa. Dra. MARIA DE LOURDES BUENO TRINDADE GALO
Departamento de Cartografia / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Dr. LUIZ HENRIQUE DA SILVA ROTTA
Pós-doutorando PNPD / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Presidente Prudente, 26 de abril de 2019

À Deus,

Aos familiares e amigos,

em especial meus pais João e Rozangela

e aos meus avós Manoel e Izabel,

à minha esposa Nariane,

e nossa linda filha Cecilia.

AGRADECIMENTOS

Tudo o que somos e tudo o que fazemos só é possível graças ao Deus de infinita bondade que nos presenteou com o dom da vida e que guia cada um de nossos passos. Obrigado Deus por se fazer tão presente em nossas vidas, em cada momento e em cada detalhe.

Gostaria de agradecer a todas as pessoas de minha família, que sempre estiveram presentes, apoiando e incentivando todo esforço, e sobretudo, capazes de compreender o tempo direcionado para realização desse trabalho. De modo especial quero agradecer aos meus pais João e Rozangela e aos meus avós Manoel e Izabel pelo apoio incondicional, todos os sacrífícios para transformar esse sonho em realidade e toda a grandeza dos meus maiores exemplos de inspiração.

Agradeço àqueles que desempenharam o importante papel de acompanhar e direcionar cada passo desse caminho, meus orientadores do doutorado Nilton Imai e Milton, que sempre confiaram paciente e corajosamente na exploração de alterantivas e tentativas ousadas para novas soluções. Agradeço também ao Enner, Luiz, Fernanda, Lourdes, Tommaselli, Carol Ambrósio, Carol Campos, Carol Pifer, Bruno Faga, Oliver, Yara, Gabi, Marcus, Adilson e todos os importantes talentos do grupo SERTIE, o qual tem desenvolvido e apresentado ao mundo importantes contribuições científicas e aplicadas.

A pós-graduação transforma a vida de quem participa, e esse doutorado, em especial, por meio das grandes bençãos de Deus foi o fio condutor inicial de uma união que permanecerá eternamente gravada direto no coração, pela qual agradeço, hoje e sempre, à minha menina, amiga, parceira, companheira e esposa Nariane e nossa filha Cecilia que está sendo gerada com todo carinho e amor.

Não poderia deixar de agradecer a todos meus amigos que compartilharam diversos momentos, sem os quais, não tenho dúvida de que essa caminhada seria mais difícil. Agradecimentos aos amigos da pós-graduação, amigos do Ceará AC, amigos do Futebol

e Tênis de mesa, aos irmãos de caminhada da Paróquia Nossa Senhora Mãe da Igreja, que muitas vezes foram os essenciais responsáveis pelos momentos de reabastecimento das energias.

Agradeço aos colegas de trabalho da Toledo Prudente Centro Universitário, na pessoa da reitora Zelly, por todo apoio e por sempre acreditarem na formação como um processo transformador que precisa ser inovador. Agradeço à toda comunidade FCT/UNESP que me acolheu e se transformou em minha casa desde 2008 e 11 anos depois, se encerra para a abertura de um novo ciclo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecimentos também à Fundação de Amparo à pesquisa do Estado de São Paulo (FAPESP), pelo financiamento dos projetos no 2012/19821-1, 2015/21586-9, 2015/18525-8 e 2019/00259-0, e ao Conselho Nacional de Tecnologia e Desenvolvimento (CNPq) pelos projetos dos processos 472131/2012-5, 400881/2013-6 e 482605/2013-8, em especial o processo 141909/2015-3 pela importante bolsa concedida no período do doutorado.

"O sol se levanta, o sol se põe e se apressa para voltar a seu lugar, onde renasce. O vento gira para o sul e dobra para o norte; passando ao redor de todas as coisas, ele prossegue e volta aos seus rodeios. Todos os rios correm para o mar, e o mar contudo não transborda; para o lugar de onde saíram voltam os rios, no seu percurso. Todas as coisas são difíceis e não se pode explicá-las com palavras. A vista não se cansa de ver, nem o ouvido se farta de ouvir."

(Eclesiastes, 1:5-8)

RESUMO

A aquisição de dados por meio da combinação de Sensoriamento Remoto e amostragens in-situ permite que várias fontes de dados sejam integradas para a análise e observação de características do alvo de interesse e pode exigir métodos computacionais para apoiar o processamento, exploração e análise de dados. A necessidade de integrar dados de diferentes fontes é destacada em estudos de ambientes dinâmicos e complexos que se alteram frequentemente, como os reservatórios hidrelétricos. Os reservatórios são ecossistemas artificiais, que influenciam diretamente nas características regionais, principalmente devido ao seu uso múltiplo uso. As interações da energia eletromagnética com os componentes opticamente ativos ocorrem ao longo de toda a coluna d'água, de modo que o comportamento do campo de luz reflete as mudanças aplicadas ao longo da zona eufótica. No entanto, as grandezas registradas nas imagens são usados de acordo com o plano e limitadas ao respectivo ponto ou área da superfície. A calibração de modelos bio-ópticos, considerando apenas os dados de amostragem da superfície, pode não fornecer resultados totalmente eficazes, porque a radiação eletromagnética interage com os componentes localizados ao longo da coluna de água e, consequentemente, a resposta capturada pelos sensores não representa apenas o valor associado à superfície. Este trabalho propõe uma investigação sobre a influência da distribuição vertical das propriedades ópticas ao longo da coluna d'água, a fim de contemplar registros sobre a interação em diferentes níveis de profundidade, além da própria disposição espacial, configurando-se como um ambiente genuinamente tridimensional. Para isso, dados coletados por sensores espectrais in-situ e por câmera hiperespectral embarcada em uma plataforma aérea foram usados para construir o modelo. Por fim, espera-se que o produto final desta investigação, composto pelo modelo tridimensional ajustado a partir de relações semi-analíticas e empíricas, possa ser uma contribuição inédita para a área de Sensoriamento Remoto e permitir identificar relações ao longo do volume de água que pode fornecer subsídios para a melhoria de modelos bio-ópticos e derivação de outros estudos relacionados à propriedades ópticas em diferentes profundidades.

Palavras-chave: Modelo bio-óptico tridimensional. Propriedades ópticas. Distribuição vertical da luz. Águas Interiores.

ABSTRACT

The acquisition of data using Remote Sensing and in-situ sampling allows several data sources to be integrated for the analysis and observations of environmental characteristics and may require computational methods to support the data processing, exploration and analysis. The need to integrate data from different sources is highlighted in studies of dynamic and complex environments that frequently change, such as hydroelectric reservoirs. Reservoirs are artificial ecosystems, which influence directly the regional characteristics, mainly because of their multipurpose use. The interactions of the electromagnetic energy with the optically active components occur along the entire water column, so that the behavior of the light field reflects the changes applied along the entire euphotic zone. However, the values taken from images are used accordingly to a plane and associated with the respective point or area of surface. The calibration of bio-optical models considering only the surface sampling data can not deliver fully effective results because the electromagnetic radiation interacts with the components located along the water column and the response captured by the sensors does not only represent the value associated with the surface. Considering this scenario, this work proposes an investigation on the influence of the vertical distribution of the optical properties along the water column, in order to contemplate records about the interaction in different levels of depth, besides the spatial disposition itself, being configured as a genuinely three-dimensional environment. For that, data collected by in situ spectral sensors and by hyperspectral camera in an aerial platform are used to construct the model. Finally, it is expected that the final product of this investigation, composed of the three-dimensional model adjusted from semi-analytical and empirical relations, may be an unprecedented contribution to the Remote Sensing area, and allow to identify relations along the water body volume which may assist as subsidies for the improvement of bio-optical models and derivation from other related studies about depth-related optical properties.

Keywords: Three-dimensional bio-optical model. Optical Properties. Vertical distribution of light. Inland Water.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

CHAPTER 1

INTRODUCTION

Brazil has a set of energy power plants producing over 151.6 million kilowatts, which more than 61% is generated from hydro-power managed by the 219 hydroelectric plants built in the country (ANEEL, 2019). The greatest energy potentials come from hydroelectric built on accumulation reservoirs, in which the river is dammed and a large area is flooded. The impacts from these artificial systems result in changes in regional hydrology, flooding of riparian forests, interference in physical and chemical processes, expansion of river geometry, an increase of material discharge, aggressive implantation of a lentic system, a change in the reproduction cycle of species, among other factors that cause the gradual eutrophication of the system.

The reservoirs changes the physical, chemical and biological characteristics of water bodies, whose magnitude of impacts is directly related to the size of the reservoir and the technology adopted. Elseways, reservoirs have positive aspects, which can produce socioeconomic benefits, such as: energy generation; water for consumption; recreational and sports activities; flood control; irrigation; economic development; jobs and navigation (KORNIJóW, 2009; TUNDISI; MATSUMURA-TUNDISI; TUNDISI, 2008). The effects and impacts are cumulative in reservoirs arranged sequentially in cascade systems on the same river (NOGUEIRA et al., 2006).

The Paranapanema River is the second river in Brazil with the largest number of cascading reservoirs, comprising eleven hydroelectric plants installed. One characteristic of the Paranapanema River is the integration with large tributaries, such as the Tibagi River. The confluence region between the Tibagi and Paranapanema rivers is part of the lake of the Capivara reservoir.

Large tributaries act as important contributors to nutrient availability, changing

the effects of eutrophication of the dammed regions (LANSAC-TÔHA et al., 2004). Despite the impacts attributed to anthropogenic intervention on the natural flow of the river, large tributaries are essential for maintaining the changed environment and mitigating the consequences of damming. The tributaries may have lotic or semi-lotic features and preserve some original features of the system (HOFFMANN; ORSI; SHIBATTA, 2005; ORSI et al., 2002; AGOSTINHO; PELICICE; GOMES, 2008; BENNEMANN et al., 2006).

Monitoring and control complex environments which can be characterized by secondary circulation, flux separation, and vertical layer segmentation (KENWORTHY; RHOADS, 1995), requires studies that consider data from sampling in transects and in vertical profile with different depths. Constant and periodic data acquisition are essential for effective management of dynamic ecosystems that frequently change. Remote sensing techniques based on images produced by orbital sensors and photographs generated by other aerial imaging devices have been consolidated as important resources to obtain data about the characteristics of targets of interest with various temporal, spectral and spatial resolution. Hyperspectral sensors can be useful to analysis of water bodies with complex mixtures, such as found in tropical inland water. The remote sensors offer an alternative to the analysis, which, integrated with local sampling methodologies (in situ), represents an important tool for environmental monitoring, with benefits related to the temporal (higher frequency) and spatial (larger area) resolution of the analyzes with no impact on the execution cost.

Remote sensing products represent the interaction of electromagnetic energy with the optically active constituents (OAC) of the target and elements present on atmosphere. The electromagnetic wave interacts with the targets from the path in the water column to the depth of the light field can react (euphotic zone), which is directly dependent on the specfic wavelength (MCCLUNEY, 1974). The images represent the characteristics of the composition of the water column, according to the associated electromagnetic spectrum. However, the values from the orbital images are usually processed under a plane and associated with the respective point or area of the surface, because they are usually adjusted from surface samplings. Different interactions along the water column can not be decomposed and segmented precisely in depth using only the response of the remote sensors.

Since the energy recorded by the sensors interacts with the components along the water column, and the vertical variation and depth of light penetration directly depends on the wavelength, the response captured by the sensors does not represent only the value associated with the surface, and the calibration of bio-optical models considering only surface sampling data may not agree fully effective results. Therefore, it is still necessary to understand the process of interaction of light within the water column.

The use of surface-based sampling method for calibration and validation of bio-

optical models highlights the main concern of this work: Is it possible to estimate the behavior of optical properties at different depths from remote sensing images?

To improve models calibrated with surface data and to bring a new contribution to the area, studies related to the volumetric bio-optic characterization of the water body are necessary to understand the processes present in the water column. For this, it is necessary to use specific tools that allows: (1) the acquisition of real data in different depth, minimizing the errors inserted in the process of estimation of these values; (2) effective methods capable of integrating and processing the collected data; and (3) genuinely three-dimensional data management and processing approaches to enable exploration and analysis in 3D space.

1.1 Goal

The main aim of this research is to develop automated approaches to integrate and process the data from multiple sensors using vertical in-situ data and hyperspectral image to calculate depth-related optical properties.

1.1.1 Specific objectives

- Model a spatial database infrastructure to enable the integration of collected datasets from various sources and allow to explore and analysis with the support of an automated system;
- Configure, acquire, correct and process the images from a hyperspectral camera transported in an aerial platform building a unique mosaic for each band;
- Investigate and develop models to transform the aerial image according to the optical properties collected in-situ, such as reflectance;
- Construct empirical and semi-analytical models to map the spatial variability of optical properties along the water column from remote sensing images and in situ data.

1.2 Document structure

This document is composed of seven chapters, including this one introduction that performs a brief context of the problem and delimits the scientific research and objectives. In Chapter 2 are presented the approaches that guide the execution of this research, composed by the delimitation of the study area, as well as the methodological aspects that define the materials and processes used to develop the thesis.

Each of the four specific objectives was discussed in specific papers submitted for review in qualified journals. The papers present specific and sequential contributions

that incorporate the chapters of this dissertation. The subjects of each chapter were defined to emphasize the sequential synergy between them and avoid overlap. The Figure 1 synthesizes the relationship and integration between each of the development chapters.

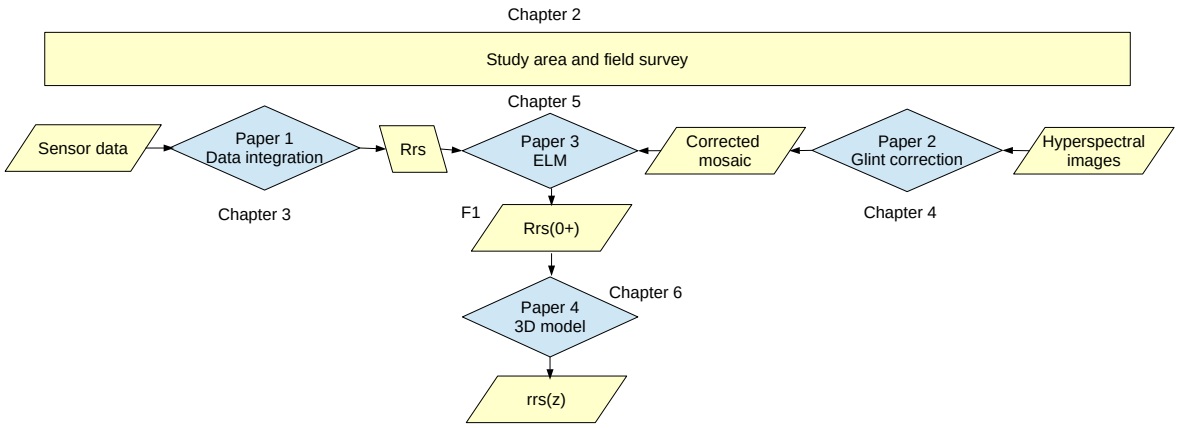


Figure 1 – Diagram of relationship between each of the development chapters

In order to improve the reading, all the information about the study area and field survey, common to all articles, has been moved to Chapter 2.

In Chapter 3 we presents an infrastructure containing a data model to integrate information from multiple sources, and a workflow to correct the data from different sensors. In Chapter 4 are discussed methods to correct specular reflection in hyperspectral aerial image. We propose a approach to correct and construct of the mosaic based on the overlap of the images and the capture of points in different views. In Chapter 5 we discuss approaches to transform the digital numbers registered in the remote sensing images into reflectance. We propose a new approach to the settings of empirical lines using artificial reference targets and water sample points. Chapter 6 performs the integration of the results produced by the previous specific papers and produces a new model for estimating depth-related reflectance from the remote sensing images. Finally, in chapter 7 we present conclusions about the work and recommendations for future work.

CHAPTER 2

STUDY AREA AND DATA ACQUISITION

The analysis of water composition and its optically active components, from the study of their apparent and inherent optical properties require the integration of different data collected by sensors. In this chapter, we describe the study area and the main characteristics of the collected data. This chapter shows the materials used for collecting and obtaining the data, and those used for the processing, exploration and analysis of the data.

2.1 Study area

The confluence region between the Tibagi and Paranapanema rivers is formed by the merging of two large rivers in a portion of the lake of the reservoir of the Capivara hydroelectric power plant. The reservoir is located on the border between the states of São Paulo and Paraná, as can be seen in Figure 2 and is the largest of the eleven hydroelectric plants installed in the Paranapanema River.

The eleven cascading reservoirs along the Paranapanema River, with only three of them configured with accumulation lakes (Jurumirim, Chavantes and Capivara), creates a strong relationship of dependence between the characteristics of these reservoirs(FELISBERTO; RODRIGUES, 2005). However, large tributaries, such as Tibagi River, act as important contributors to carry in nutrients into different regions, resulting the possible eutrophic effects of the dammed regions (LANSAC-TÔHA et al., 2004).

The Tibagi River carries from the south of the Paraná state at about 1060 meters altitude and has an extension of about 550 kilometers to the north, where it flows into the Paranapanema River at 298 meters altitude. The Paranapanema River has a greater

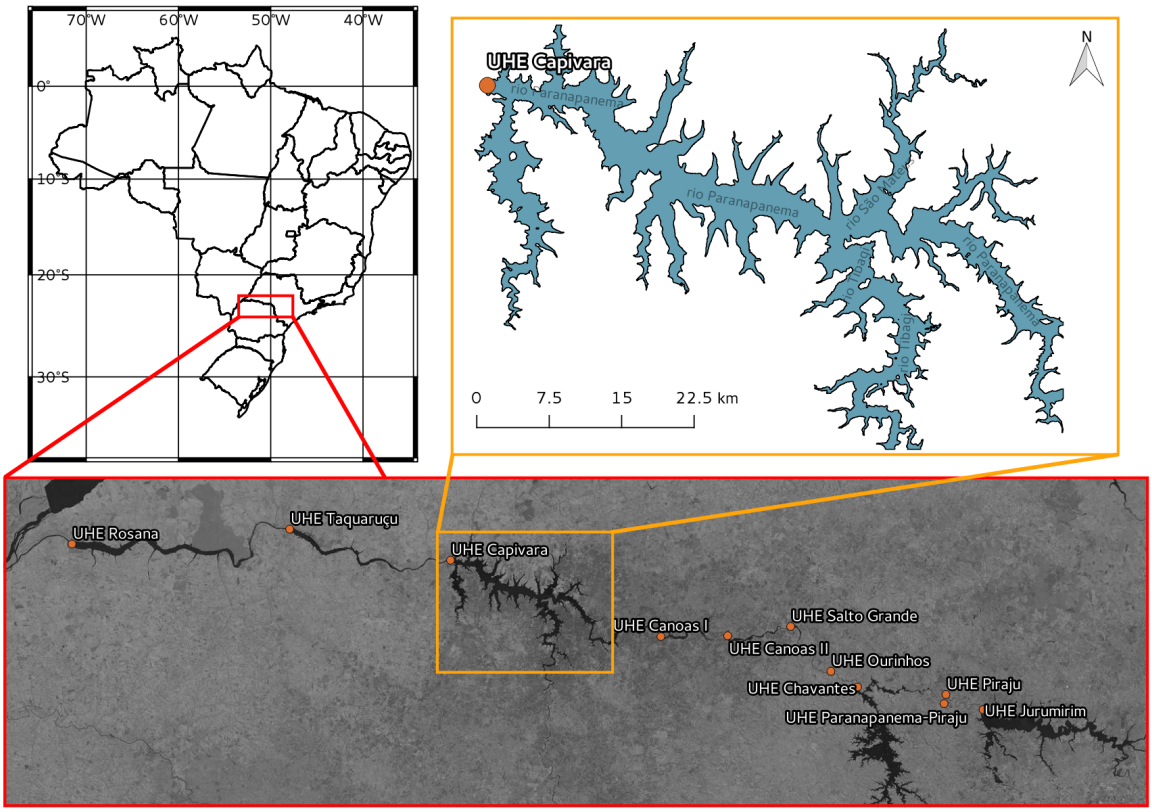


Figure 2 – Study area and location of eleven hydroelectric power plant on Paranapanema river

extension, with 929 kilometers in the East-West direction, corresponding mostly to the border between the states of Paraná and São Paulo. Its source in the southeast region of the São Paulo state, at 809 meters of altitude to drain into the Paraná River at an altitude of 239 meters. While the Tibagi River has an average slope of 138 centimeters per kilometer, the Paranapanema River has an average slope of only 61 centimeters per kilometer, transforming it into rivers with distinct flow characteristics (MEDRI, 2002).

2.2 Data acquisition

The field survey was performed between August 29 and September 2, 2016. The sampling design, as can be seen in Figure 3 was designed to contain 26 sampling stations.

From the sampling design, we performed the flight planning to cover the area of sampling. The area of interest and respective flight ranges defined in the planning step can be seen in Figure 4. The flight starts around 1 P.M. (local time) performed at 3600 meters of altitude and cover the ROI with six flight paths, resulting in a 40% lateral overlap. We achieved a longitudinal overlap of approximately 80%.

The onboard system had a GNSS receiver, a hyperspectral camera, an RGB camera and a RaspBerry with integrated circuits for image acquisition management. The

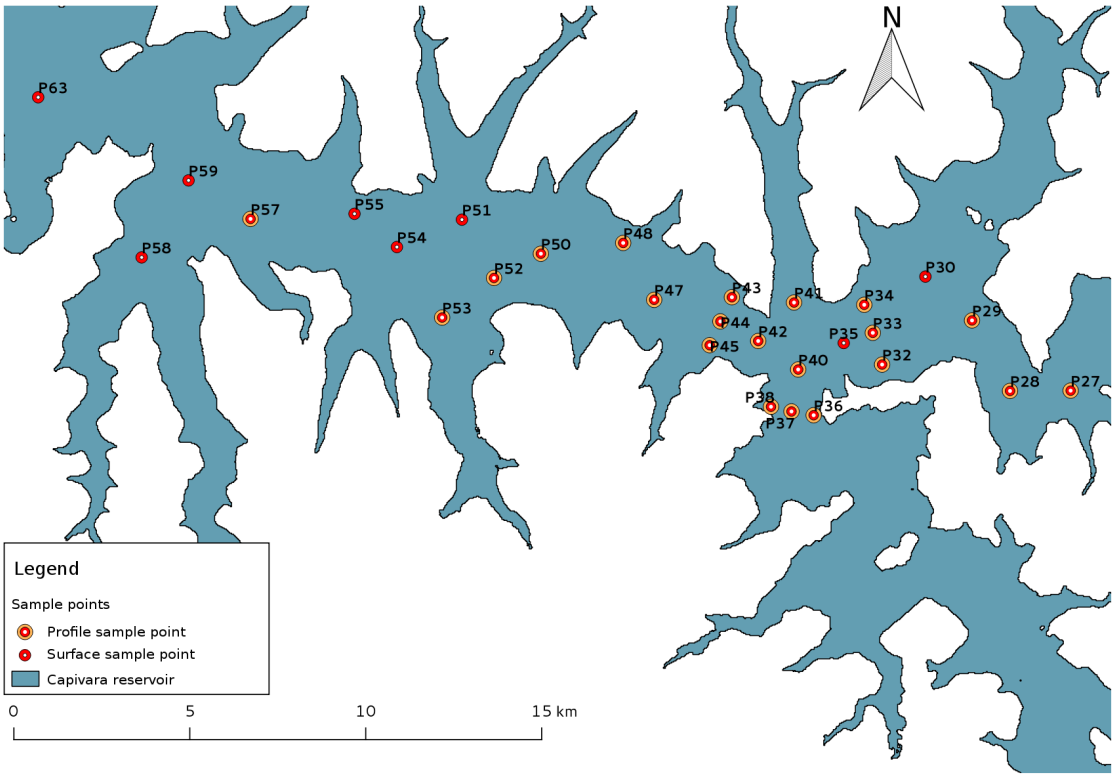


Figure 3 – Sampling design of the field campaign with the location of 26 stations

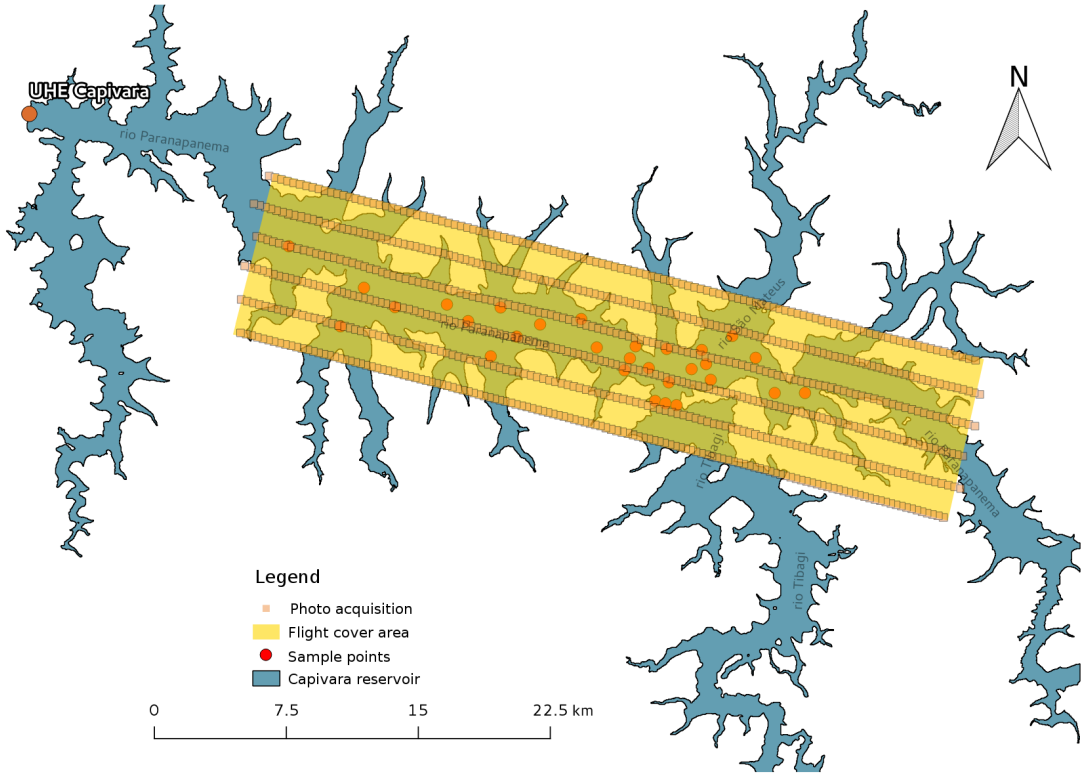


Figure 4 – Flight plan and respective paths and position of the photos

hyperspectral camera manufactured by Rikola, model FPI2014, can collect up to fifty spectral bands at customized intervals. However, because of a manufacturer software limitation, only twenty-five bands have been configured. We set the camera channels according to the main orbital sensor wavelengths used for aquatic environment studies (MODIS, Sentinel and Landsat) and for coverage of interval spectra useful for the analysis. We used the following central wavelengths and respective full widths at half maximum (FWHM), both expressed in *nanometers* (*nm*), for the 25 Rikola bands: 505.37 (9.51), 515.31 (14.05), 528.55 (14.82), 539.87 (14.03), 546.99 (14.31), 554.99 (13.26), 560.22 (12.11), 570.44 (14.31), 579.58 (13.26), 592.57 (16.57), 605.73 (14.98), 620.22 (16.26), 625.92 (15.47), 655.06 (12.55), 665.72 (15.59), 670.03 (15.74), 681.33 (15.89), 695.04 (14.96), 700.46 (15.44), 707.96 (15.32), 715.19 (15.37), 725.37 (14.72), 737.29 (14.98), 749.87 (15.08) and 780.1 (14.72). The central wavelengths in Rikola with the full widths at half maximum (FWHM) simulated from a normal distribution are showed on Figure5.

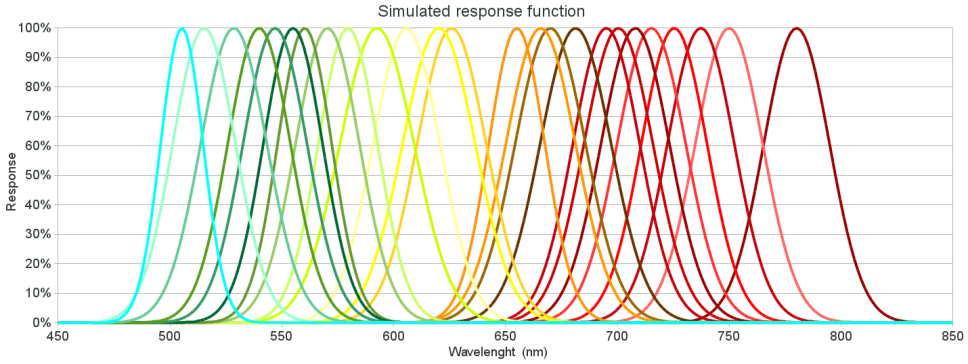


Figure 5 – Response function of the 25 bands of Rikola camera simulated with a normal distribution

The Rikola camera is a hyperspectral frame imager based on a time-variant tunable filter, which record bands sequentially with a time lag between the bands (AASEN et al., 2018). The bands are captured in a slightly distinct position and orientation and need to be coregistered to correct a gap between each band acquisition (HONKAVAARA et al., 2013; HONKAVAARA et al., 2017). We did not apply band coregistration for all works, because some tasks process each band individually, working in a preprocessing step. Coregistration can be performed after band mosaicking by determining the orientations and georeferencing the individual bands separately based on a reference band (HONKAVAARA et al., 2013; Honkavaara et al., 2012; HONKAVAARA et al., 2012).

In this survey, we did not use an inertial measurement unit (IMU) system. To perform georeferencing using bundle adjustment, we need identifiable targets distributed in the image block. However, the identification of targets in all scenes of large homogeneous areas, such as water bodies, can be a difficult task. To overcome this challenge and avoid decreasing the image resolution, we adopt a different methodology: we use a Sony RGB camera, with a higher field of view (FOV) than the Rikola camera, to cover a region of

the river bank in all scenes (captured simultaneously with both cameras) and enable the measurement of tie points and further bundle adjustment processing. Table 1 presents the specifications of each camera.

Table 1 – Hyperspectral and RGB camera specifications used in the flight survey

Specification	**Rikola**	**Sony**	**Unit**
Model	FPI-2014	A6000	
Sensor	CMOS-CMV400	CMOS	
F-number	∼2.8	∼3.5	
FOV	37	83	*degrees*
Spectral range	500 - 900	RGB	ηm
Image	1024 x 648	6000 x 4000	pixels
Pixel	5.5	3.92	μm
Focal length	∼9	∼16	mm
Weight	∼700	∼470	gr

The triggering of both cameras was synchronized by a Raspberry PI electronic controller to enable simultaneous image shots. The external orientation parameters (EOP) taken by the RGB camera and adjusted with tie points were computed by bundle adjustment and applied to the Rikola images. We perform individual georeferencing on each image.

During the field campaign, we collected optical data using spectoradiometers to measure water and reference targets responses. We acquired data from three calibration tarpaulins: white, gray and black. Each target had a dimension of 9m x 9m and was placed on the irregular grass field near the river margin. The tarpaulin presents some small irregularities due to grass surface which cause undesirable reflections.

Using two ASD Handheld spectroradiometers, we acquired simultaneous radiance measurements from artificial targets (L_t) and a reference Lambertian panel (L_g) to enable calculate the remote sensing reflectance (R_{rs}).

The in-situ radiometric data of the sample points along the reservoir were acquired with hyperspectral sensors TriOS. The TriOS data are composed of two different types of measurements: above and below the water surface. Four radiometers of the TriOS RAMSES model were used, two of which are irradiance sensors and the others are radiance sensors.

To maximize the amount of profile data, three sensors were arranged in the structure to be submerged and only one radiance sensor was operated out of the water and repositioned at each measurement, as can be seen in Figure 6, which presents two moments of different positioning of the sensors, before and during the profile.

In order to increase the frequency of data acquisition and corresponding record availability at different depths, the TriOS profile data was performed continuously during submersion in the water column. Post-processing was performed to integrate the data

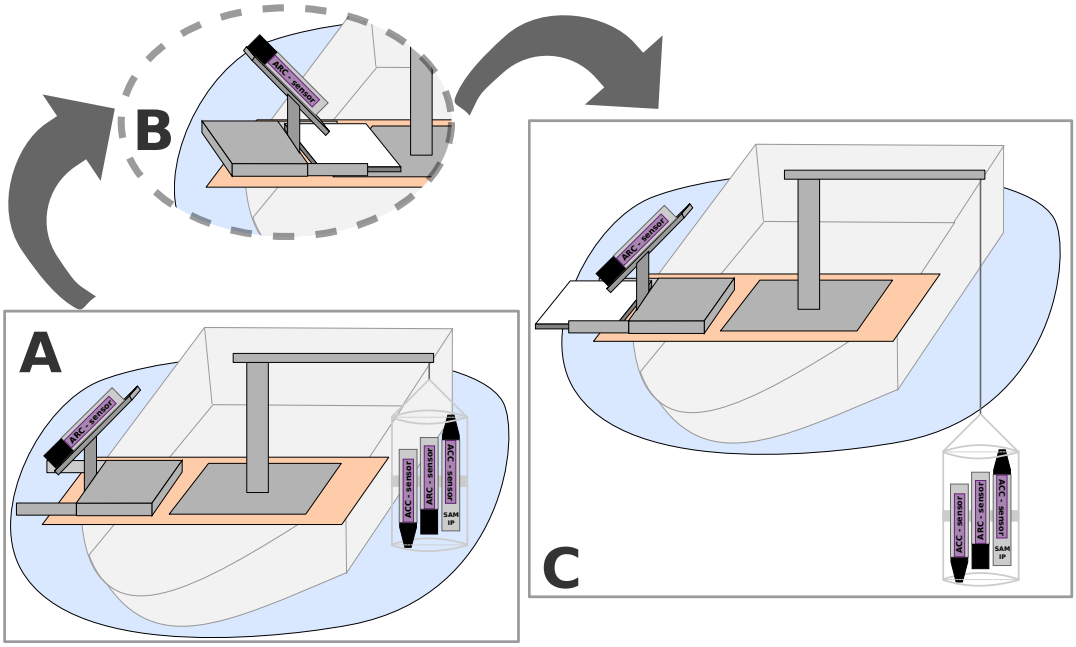


Figure 6 – The position of the sensors before and during the profile and some setting changes of position of the radiance sensor. (A) Radiance sensor pointed at the water and the other three above the surface. (B) Radiance sensor pointed at the sky and the other three remain above the surface. (C) Radiance sensor pointed to the reference plate and realization of the profile with the other three submerged sensors.

according to the depth captured by the Trios SAMIP radiometer, which has inclination and pressure sensor. In this setting, the frequency of acquisition at different depths is directly dependent on the submersion speed of the structure by the operator and the integration time required by the sensors for energy acquisition, which varies from 500 milliseconds to 8000 milliseconds, allowing usually average acquisitions every 8 centimeters of the water column.

CHAPTER 3

AQUOPTS: A MULTISOURCE PROCESSING SYSTEM FOR MULTIDIMENSIONAL BIO-OPTICAL DATA INTEGRATION AND CORRECTION

Abstract

Field surveys are important source of data in several scientific researches. Hydrological optics investigations require a large amount of data about optical properties. Sensors from distinct manufacturers built with specific technologies are used to acquire such measurements. Consequently, a fast and integrated processing is not a trivial task due to several challenges, such as: non-standard outputs, multidimensional resolutions of data (spatial, temporal and spectral), mismatched sampling, sensor-specific corrections, manufacturer proprietary software, and so on. This paper proposes a procedure to overcome the drawbacks of processing multisource and multidimensional datasets. Our main goal is provide a platform to integrate and analyze the datasets acquired from five sensors commonly used by hydrological optics community. We can synthesize the contribution of this work by three resources: a database model for management of datasets from field surveys; a processing workflow that describes all the correction steps by grouping protocols to integrate the data; and an online system, Aquopts, which all resources for storage, integration, correction,

processing and analysis are implemented based on the proposed methods. The Aquopts have provided resources that were used in several papers already published in recognized journals of geoscience supported by our system delivering a fast and integrated datasets processing.

3.1 Introduction

Several research projects, mainly on the Remote Sensing area, depend on field survey to collect data for analysis and calibration of bio-optical models. Studies about underwater light field to characterize water bodies and models for water quality monitoring require spatio-temporal data about Inherent Optical Properties (IOP) and Apparent Optical Properties (AOP) (BRACCHINI et al., 2005; DOXARAN; CHERUKURU; LAVENDER, 2006; MENON et al., 2006).

In-situ measurements can derive AOPs, in which the angular distribution of the light field and the substances on the water body have influence about it (SATHYENDRANATH et al., 2000). Otherwise, the IOPs, such as absorption and scattering coefficients, are independent of the angular distribution and geometry of the light field making possible to analyze them from water samples in a laboratory (KIRK, 1994). Both optical properties are directly dependent on the type and amount of present substances in the water (KIRK, 1994).

When light penetrates the water body it suffers attenuations by the Optical Significant Components (OSCs) present in the water. The light propagates in the water column and changes the radiance distribution (BUKATA et al., 2018; RIDDICK et al., 2015). Spectoradiometers, such as TriOS sensors (TriOS, Rastade), acquire AOP data (DOXARAN; CHERUKURU; LAVENDER, 2006; NOVOA; WERNAND; WOERD, 2015). The AC-S (WET Labs, Philomath), ECO-BB9 (WET Labs, Philomath) and Hydroscat-6p (HOBILabs, Bellevue) spectrophotometers are commonly used to collect the IOP datasets (KIM JOONG YONG PARK, 2010; LIN et al., 2018; WOźNIAK et al., 2011).

In hydrological optical studies there are relevant additional dimensions, such as spectral, temporal and spatial resolutions also related to depth (MOORE et al., 2009; LEWIS, 2008). Integrating multidimensional data obtained from multiple sources is a relevant challenge (KRÖGER et al., 2009; HU BRIAN B. BARNES, 2013). Using many sensors demands effective methods on integrating the data from different sources, stored on distinct structures with specific formats (IMAI et al., 2015). Each new sensors developed increases the combining options become more complex the data integration requiring appropriated models and development of general usage protocols (MOORE et al., 2009).

In this context, the main goal of this work is to present a framework to automate

and optimize the processing steps to assist the researcher on data management. This paper proposes the use of three resources to support the integration, process and analysis of datasets: a database model, a workflow process and an online system. We built the database model to represent all data collected from field campaigns with relationships among the data retrieved from sensors, images and laboratory analysis. The workflow process describes all steps of correction and integration the data acquired from five sensors. We built the Aquopts system to optimize and automate tasks applied on multidimensional and multisourcing datasets. The full workflow process and some relevant developed tools are publicly available (CARMO et al., 2019; CARMO, 2019). In addition, the Aquopts system which implements all features presented in this work is available [1], however, it is necessary to register and request permission to access all functions.

Several works developed by the Remote Sensing and Information Technology for Environmental Monitoring Group (SERTIE [2]) have been using the proposed platform and frameworks (ALCâNTARA et al., 2016; RODRIGUES et al., 2016; WATANABE et al., 2018). The framework is an information technology solution to enable the automation of several processes that are usually performed manually by researchers, making the process costly. The Aquopts implements a number of features and tools to reduce the costs of generation and standardization of bio-optical datasets, so the researchers can focus on analyzing and modeling solutions. We validated the system using the datasets collected by SERTIE, highlighting its performance to improve storage infrastructure, integration, correction, filtering and other important preprocessing tasks required to provide adequate data for analysis.

3.2 Critical points and challenges

There are many parameters important to describe water bodies, justifying the usage of different sensors to collect a vast set of attributes (MOORE et al., 2009). Field campaigns are essential to acquire qualified dataset and enable the parametrization of bio-optical modeling, and the data preprocessing plays a important role on remote sensing of water environment (DALL'OLMO; GITELSON, 2005; MA; TANG; DAI, 2006). In our case, we use five sensors, in which Table 2 shows a summary of the technical specifications of each sensor.

Multivariate analysis and study of the correlations between collected parameters require protocols and good practices to guarantee the compatibility between collected datasets. Usually, once the field campaign has been over, the analyst run a data exportation from each sensor and starts the processes of integrity validation, correction enforcement, data normalization for data integration. Such actions are frequently repeated for each new

[1] http://sertie.fct.unesp.br/aquopts
[2] http://sertie.fct.unesp.br

Table 2 – Summary of spectral sensors features used in this work

Model	Manufacturer	Type	Parameters	Specifications
AC-Spectra (AC-S)	WET Labs	Spectrophotometer	Absorption and Attenuation	4 nm resolution, 80 wavelengths
ECO-BB9	WET Labs	Spectrophotometer	Backscattering	9 wavelengths, acquisition angle of 117º
CTD	Sea-Bird	Multiparameter Sensor	Conductivity, Temperature, Pressure	4 Hz sampling
HydroScat-6p	HOBI Labs	Spectrophotometer and Fluorometer	Backscattering, Fluorescence, Depth	Backscattering at 6 wavelengths, acquisition angle of 140º, Fluorescence at 2 wavelengths
Ramses	TriOS	Spectroradiometers	Radiance (Ramses-ARC), Irradiance (Ramses-ACC-VIS)	Hyperspectral resolution (3.3 nm)

field campaign increasing the cost of these processes. Among the important aspects related to field campaigns, it is possible to name other issues about (IMAI et al., 2015): spatial attribute and mismatch sampling; spectral data; multisource and multidimensional data and dataset management.

3.2.1 Spatial attribute and mismatch sampling

The sampling position defines the spatial coordinates of data acquisition. The aquatic environment field surveys require a special care about the data management of the sampling locations. It does not limit the spatial attributes to a bi-dimensional set of latitude and longitude, but also have an important third coordinate representing the depth on the water column (Z-coordinates) (CURTARELLI et al., 2015).

The depth attribute is relevant because it is a parameter directly used to reference and establish the link to the dataset, since it is a real component of the 3D coordinates of the collected data. Fixing standard depth intervals is essential to the integration of data measured by different sensors, but it is not easily obtained on the field survey. As the sensor displacement in the water column, depends on the intervention of the operator, the submersion route in each execution cycle is inevitably variable. Performing the sample in an unstable platform also amplifies the sensor displacement variation.

3.2.2 Spectral data

The electromagnetic spectrum allows highlighting absorption and scattering features caused by specific OSCs, i.e, it is a proxy to identify water components (KUTSER; VAHTMÄE; METSAMAA, 2006; BABIN et al., 2003). However, the usage of different equipments composed of diverse sensors, built with many technologies, adopting distinct

storage structures and measuring at different wavelengths, highlight an essential care in the standardization of collected data aiming the integration of dataset (MOORE et al., 2009; ZIELINSKI et al., 2009).

Some equipments hold a determined quantity of fixed wavelengths, such as the ECO-BB9 and the Hydroscat-6p. Other equipments acquire hyperspectral data, such as the AC-S and the TriOS spectoradiometers.

3.2.3 Multisource and multidimensional data

To produce useful information, the analysis process integrates different datasets stored in various structures. The dataset integration requires a shareable parameter and compatible structure to synchronize the data. There are a global effort to develop specification to assure the data standardization and interoperability, which it is an important source to organize concepts and technical products to add economic, social and legal value to the system (BOULOS et al., 2011; Di; Moe; van Zyl, 2010; Moe et al., 2008; Suri et al., 2007).

The data structures vary because of specific characteristics on the collected attributes, including the dimensionality (scalar, vector, spectral, and others), and by the noncompliance output format standards (MOORE et al., 2009). Proprietary tools as commercial database files illustrate this issue, in which only specific softwares can read the data from original data structure (ZIELINSKI et al., 2009). The use of non-standard data structures increases the dependence on specific resources, affects the flexibility for data processing and the automation of processing by other solutions (GLASGOW et al., 2004).

3.2.4 Data management

Each field survey produces a significant amount of data, which are a rich source of information for different studies (TEILLET et al., 2002). The data produced by the sensors are exported, processed and stored in individual files. The data files can be grouped into repositories that meet research groups, or, at worst, be stored and used only by that user who conducted the field survey. The infrastructure that meets the organization of files has a fundamental importance to allow the reuse of previously acquired dataset (TEILLET et al., 2002).

There are significant global efforts for adopting storage and data management standards and good practices to facilitate interoperability and reusability of datasets (SCHMACHTENBERG; BIZER; PAULHEIM, 2014; TEILLET et al., 2002; Di; Moe; van Zyl, 2010; Moe et al., 2008; Suri et al., 2007). We can adopt standards outlined by the Open Geospatial Consortium (OGC) or concepts such as Earth Observation Sensor Web (EOSW), aimed at building identifiers of datasets according to a common

and shared data, enabling the creation of references between these datasets (BOULOS et al., 2011; Di; Moe; van Zyl, 2010; Moe et al., 2008; Suri et al., 2007). The concept of Web of Data (WoD) is also useful to describe a unified network formed by the interconnection of heterogeneous datasets (BIZER; HEATH; BERNERS-LEE, 2009).

3.3 Sensor data integration

The factors presented in the previous section highlight the relevance of developing resources to make large volumes of data from multiple data sources suitable for analysis. Considering all sensor characteristics, the dataset present five dimensions: X-Y horizontal coordinates, depth (Z), time and spectral. Each parameter has specific protocols to correct and transform the data. This session presents three solutions, one for each task to correct, store, and explore the dataset.

3.3.1 Dataset Model

We used a Database Management System (DBMS) to amplify the capacity of storing, exploring and processing the data. Spatial attributes in the dataset motivate the adoption of PostgreSQL, which are extended with the PostGIS geospatial module. We designed the database model to manage the collected data from field surveys and integrate with other data sources such as laboratory data and remote sensing images.

The data model provides the resources of management in a wide range of researches that collect *in-situ* data, providing a useful and flexible framework to several scenarios of data integration. The Figure 7 shows the Entity-Relationship Diagram (ERD) of the database model.

The central elements named *d_field* and *d_set* are the core of the database model. They concentrate the relationships with all the other elements and has all attributes related to the execution of the field survey. The *d_field* has references to the time and sampling location of *in-situ* measurements, the equipment and sensors, the operator user who performed the collection, and the link with the measure set.

We stored the collected data in specific entities constructed to meet the characteristics of the data structure produced by each specific sensor. To increase the flexibility of the model, all entities destined to store data are linked only to *d_set*. We describe the process using the metadata format. Metadata is an essential information for the data user who needs to know about the composition and integrity of the data.

We suppress other hidden entities on the full database model for a clearer presentation. The hidden entities are used to manage the system platform with dynamic information about the pages, users and restrictions. The full database model, implemented

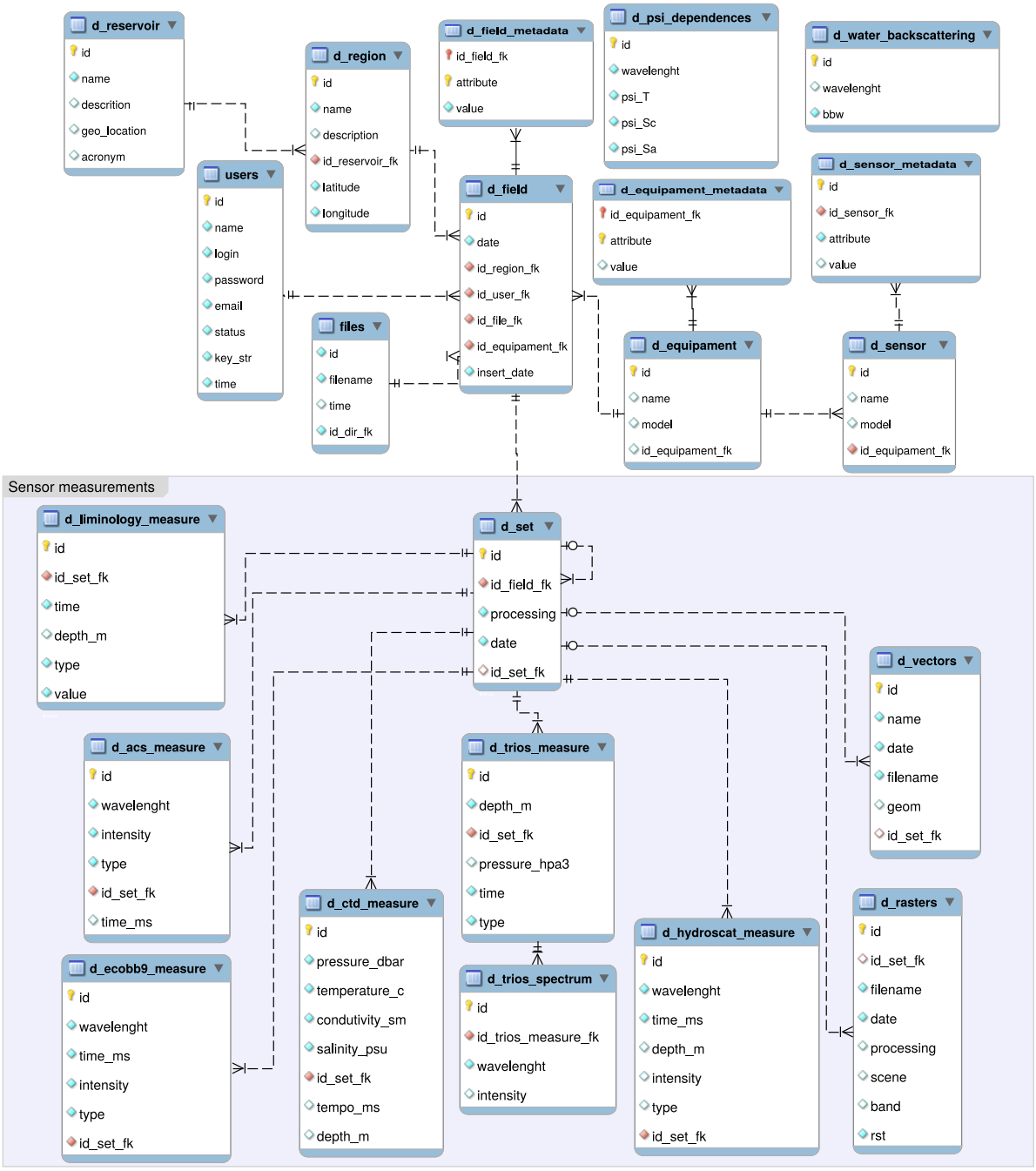


Figure 7 – ERD of the data model. The entities to store sensor measurements are highlighted

using PostgreSQL and PostGIS spatial extension, is also available (CARMO et al., 2019)

3.3.2 Processing workflow

Some corrections are required by each sensor and need previous data integration from other sensors, because the models require parameters not collected by a single sensor. The manufacturer or previous works usually proposes correction protocols, but some steps are not clear enough to ensure the entire processing, mainly due to integrating data from

different sensors. To contribute with the researcher community of water color remote sensing that frequently use these sensors, we described all processes of correction and integration about the five sensors on a single workflow, as shown at Figure 8.

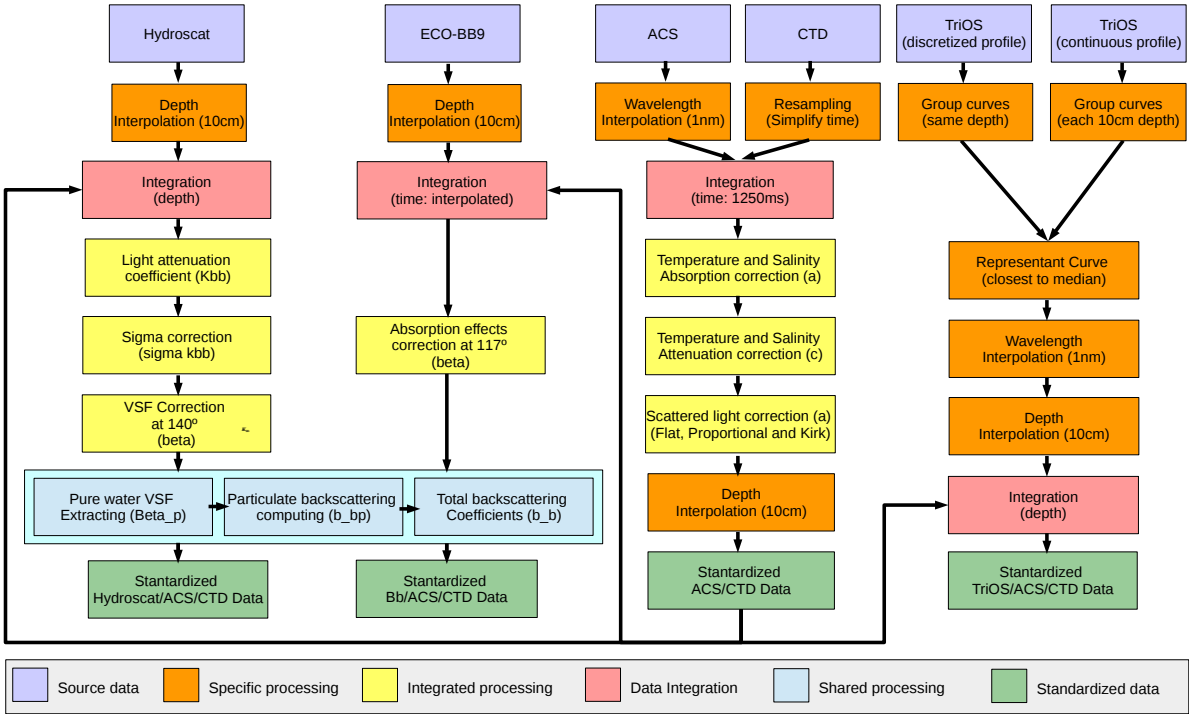


Figure 8 – Processing workflow of sensor data

The operations are grouped by color to identify the processing performed under the input data, correction functions or output structures (Figure 8). The arrows show the processing flow and the produced data by each step. We adopted some standards to integrate the data from different sensors. Depending on the sensor, we carry out interpolations to adjust the sampling rate of attributes, such as wavelength and depth.

The data produced by the CTD and AC-S sensors are the core of processing algorithms, because their data are integrated with all other sensors to perform corrections. After the spectral and temporal resampling of wavelengths for AC-S and CTD, respectively, the next step is to integrate both datasets using the time as key (sensors working synchronized). Consequently, the values of conductivity, temperature and pressure were included into the measurements of absorption (a) and attenuation (c) coefficients taken at different wavelengths at several depths.

Once integrating CTD and AC-S, a and c values were corrected. The first correction is important to remove the effects of temperature and salinity on a (eq. 3.1) and c (eq. 3.2) (LABORATORIES, 2008).

$$a_{mts} = a_m - [\Psi_t \times (t - t_r) + \Psi_{Sa} \times S] \tag{3.1}$$

$$c_{mts} = c_m - [\Psi_t \times (t - t_r) + \Psi_{Sc} \times S] \tag{3.2}$$

where, a_{mts} and c_{mts} are absorption and attenuation coefficients corrected for temperature and salinity, respectively; a_m and c_m are the measurements of a and c; Ψ_t is the temperature dependency of pure water; $\Psi_S a$ and $\Psi_S c$ are the salinity dependency of pure water absorption and attenuation, respectively ; t is in situ temperature measurement; t_r is the reference temperature; and S is a salinity measurement. We extract the temperature and salinity dependencies of pure water (Ψ_t, $\Psi_S a$ and $\Psi_S c$) from a tabular constant data (SULLIVAN et al., 2006).

The absorption measurement is overestimated due to the scattered light lost in specific angles inside the absorption tube. We adopted the three most commonly scattering errors corrections (CARVALHO et al., 2015): Flat Method (eq. 3.3), Kirk Method (eq. 3.4) and Proportional Method ((eq. 3.5)).

$$a_{Flat} = a_{mts} - a_{mts}(\lambda_r) \tag{3.3}$$

$$a_{Kirk} = a_{mts} - \varepsilon \times b_{mts} \tag{3.4}$$

$$a_{Proportional} = a_{mts} - a_{mts}(\lambda_r)\frac{b_{mts}}{b_{mts}(\lambda_r)} \tag{3.5}$$

where, a_{Flat}, a_{Kirk} and $a_{Proportional}$ are absorption coefficients corrected by the Flat, Kirk and Proportional methods, respectively; a_{mts} and b_{mts} are absorption and scattering coefficients corrected for temperature and salinity (note that $b = c - a$); $a_{mts}(\lambda_r)$ and $b_{mts}(\lambda_r)$ are absorption and scattering coefficients corrected for temperature and salinity at a given reference wavelength; and ε is a constant of scattering proportion that is not detected by the sensor. After all AC-S and CTD data processing, the dataset records are interpolated using the depth to transform into a standard resolution of ten centimeters.

ECO-BB9 and Hydroscat-6p spectrophotometers generate the Volume Scattering Function (VSF), known as β, at pre-established wavelengths. The process of obtaining and correcting the VSF from the measured signals is specific for each sensor because of acquisition geometry differences, in which ECO-BB9 and Hydroscat-6p uses a sensor view of 140° and 117°, respectively. However, the same equations can be applied on the next steps to achieve the backscattering (b_b) from corrected β. We integrated the measurements taken by the ECO-BB9 with the AC-S dataset by using the time as key. After the dataset

integration, it is possible to correct the ECO-BB9 VSF (eq. 3.6) to eliminate the absorption dependency (HYDRO-OPTICS; LABORATORIES, 2010).

$$\beta(117^{\circ}) = \beta_u(117^{\circ}) \times \mathrm{e}^{0.0391a} \tag{3.6}$$

where, a is the absorption coefficient measured by AC-S sensor; and $\beta(117^{\circ})$ and $\beta_u(117^{\circ})$ are the corrected and measured VSF, respectively.

We interpolated the Hydroscat-6p measurements at a depth of ten centimeters to match with AC-S data, because this sensors is not synchronized. After interpolating and integrating datasets, it is calculated the attenuation of light that runs into the Hydroscat-6p internal volume (K_{bb}, eq. 3.7), at each wavelength, to use in the sigma function compensation of sensor light attenuation(eq. 3.8) and correct the VSF generating the $\beta(140^{\circ})$ (eq. 3.9) (HYDRO-OPTICS; LABORATORIES, 2010).

$$K_{bb} = a + 0.4b \tag{3.7}$$

$$\sigma(K_{bb}) = k_1 e^{k_{\exp} K_{bb}} \tag{3.8}$$

$$\beta(140^{\circ}) = \beta_u(140^{\circ}) \times \sigma(K_{bb}) \tag{3.9}$$

where, K_{bb} is the attenuation coefficient of the light traveling on the sensor volume that returns to the detector; $\sigma(K_{bb})$ is an adjustment of light proportion that is lost inside the instrument volume; k_1 is a factor to satisfy $\sigma(K_{bbw}) = 1$, normally set to 1; $k_{\exp}$ is a spectral factor defined by instrument characteristics from equipment calibration; and $\beta_u(140^{\circ})$ and $\beta(140^{\circ})$ are, respectively, the measured and sigma-corrected VSF.

After the specific VSF corrections generating a corrected β, the b_b can be calculated using the same steps for both sensors. The particulate VSF (eq. 3.10) is calculated extracting pure water VSF (BOSS; PEGAU, 2001; MOREL, 1974) from total VSF. Finally, the particulate backscattering coefficient (b_{bp}) is generated from particulate VSF (eq. 3.11) to achieve the total backscattering coefficient (b_b) adding the pure water backscattering (eq. 3.12).

$$\beta_p = \beta - \beta_w \tag{3.10}$$

$$b_{bp} = 2\pi\chi_{bb}\beta_p \tag{3.11}$$

$$b_b = b_{bp} + b_{bw} \tag{3.12}$$

where, β_w, β and β_p are, respectively, the pure water (calculated from (BOSS; PEGAU, 2001; MOREL, 1974)), total and particulate VSF ; χ_{bb} is a constant of proportionality between β and b_b for particle scattering (the factory value of $2\pi\chi_{bb}$ is stored in calibration file); and b_{bp}, b_{bw} and b_b are the particulate, pure water (tabular values extracted from (SMITH; BAKER, 1981)) and total backscattering coefficients, respectively.

The TriOS sensor acquires data at different depths. There are two ways of performing the profile, in which each of them produces a dataset with different characteristics: continuous profile and discrete static profile. The continuous profile measurements are performed during the sensor movement on the water column based on its frequency rate of acquisition. On the static profile, the sensor is placed on specific depths to read redundant measurements on the same position. Depending on the profile type, we use different approaches to group spectral curves and choose a representative data. The data collected in a continuous profile may present noises related to the sensor movement in the water. To avoid noise effects, we grouped the spectral measurements in blocks of ten centimeters to choose a single representative curve. Therefore, using static profile way, we grouped the measurements in each acquisition depth, and select a single spectral curve. In some cases, the profile can be repeated to avoid errors caused by bubbles on the sensor, and the analyst is responsible to making decision about the path to use.

We used an algorithm based on median processing to compute the reference curve, in both profile case. First, we calculated a median value from each wavelength to generate a reference median spectrum. In the next step we choose the measured curve closest to the reference median spectrum, based on the lower Sum of Squares Error (SSE), Equation 5.4. It is worth noting that in both cases we do not produce or calculate any artificial value, because we selected a representative curve from measurements closest to a computed reference. To successfully apply the median, we are assuming that the amount of spurious data is inferior to the quality data or the noises are randomly distributed about the median.

$$\text{SSE}_i = \sum_{\lambda} [Y_i(\lambda) - \widetilde{Y}(\lambda)]^2 \tag{3.13}$$

where, λ is the wavelength; $\widetilde{Y}(\lambda)$ is the reference value of median curve at wavelength λ; $Y_i(\lambda)$ is the value of i-th spectral curve of group and; SSE_i is the resultant error of i-th curve.

All processing of TriOS dataset are independent of other sensors dataset. Despite that, the integration with dataset from other sensors is possible using the depth attribute as

key, after standard interpolation. Any other parameter can be generated by implementing a specific function to complement the processing workflow.

3.3.3 Data Integration, processing and exploring using Aquopts

We developed a system, named as Aquopts, to automate the process of parametrization and correction of data to make it useful for analysis. The system is the concrete implementation of all correction steps described in the workflow (Figure 8). The Aquopts provides resources to full automate the frequently repeated tasks from field surveys. We built the Aquopts using technologies based on a client-server architecture. The Figure 9 shows the system architecture with all components and logical flow of server-side processing and client-side exploration.

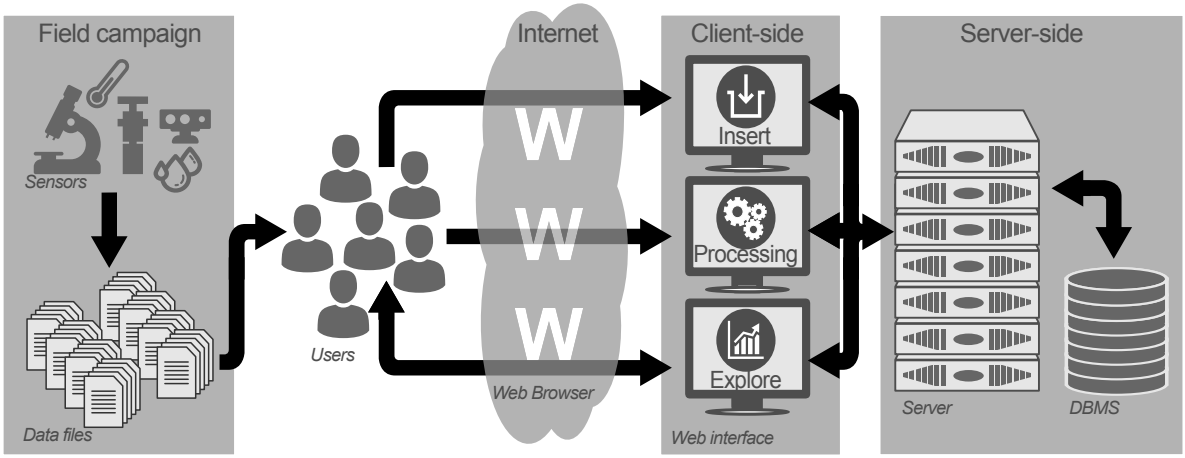


Figure 9 – Aquopts system architecture and logical flow of server-side processing and client-side navigation

We implemented the server-side with all algorithms for data integration and correction using PHP and Python programming language. We construct the web interface with HTML and Javascript. The system restricts the access of resources by registration and identification for each user. After login, the system checks the permissions and provides the access to the system.

3.4 Results and system applications

The system Aquopts has several features to support all stages of processing and analyzing data collected by the sensors. The graphical interface presents descriptive guide information about all processing steps. The Aquopts have three main interfaces: insertion of new datasets, processing stored data and visually exploring data.

The insertion interface is the unique way of input data into the system. We developed the insertion interface to read all original data file from the sensors using specific

predefined parameters (Figure 10). The user can customize the reading to match with the different data structure. In addition, all data related to the field survey are set in this step of data insertion. After the insertion, the original data file is stored and the data are registered on the database using standardized format defined on the model.

All data stored in the database can be used in the processing interface. The processing interface present all resources and tools to transform the data by normalizations, corrections and interpolations. The procedure are available accordingly to the selected dataset based on the functions expressed on the processing workflow (Figure 8), because each sensor need different processing steps.

There are some global parameters to filter the datasets to enable additional functions, but the option to export the data is not sensor-dependent which is available for all type of sensor (Figure 11). The output of exported data is a standard tabular text file which others softwares can explore it. Several works have used the export function to proceed the analysis process (ALCâNTARA et al., 2016; RODRIGUES et al., 2016; WATANABE et al., 2018).

The processing interface is able to integrate different datasets. This integration can be performed by manually selecting the dataset or automatically by the system based on the sensor type, date and location of the field survey. All historical processing keep stored on the database to track the process sequence as connected datasets.

After processing and correcting the data, users can visualize the results from the exploring interface (12). Filters may be applied to select records based on groups of constraints and conditions composed by AND/OR sentences. Finally, data can be viewed from interactive graphs that enable the analyst to obtain data graphical representation in order to interpret results, refine the search by zooming and moving the graphs, and re-evaluate the filtering conditions for reconstruction of new representations.

The generated graph is dynamic and interactive to increase the analysis capabilities. The graph parameters are chosen by the database attributes to define the x-axis, y-axis and the label to group the curves. After setting all parameters, the system generates the interactive graph.

The line graph is a quick way to analyze the data behavior and can support the decision about the data quality at the end of a field campaign day. Using the system, the analyst can run the processing steps and preliminary analysis before finishing the full field campaign.

3.4.1 Derived results using the Aquopts

The implemented workflow have been adopted by researchers of SERTIE and contributed to develop many scientific works to process sensor data acquired from many

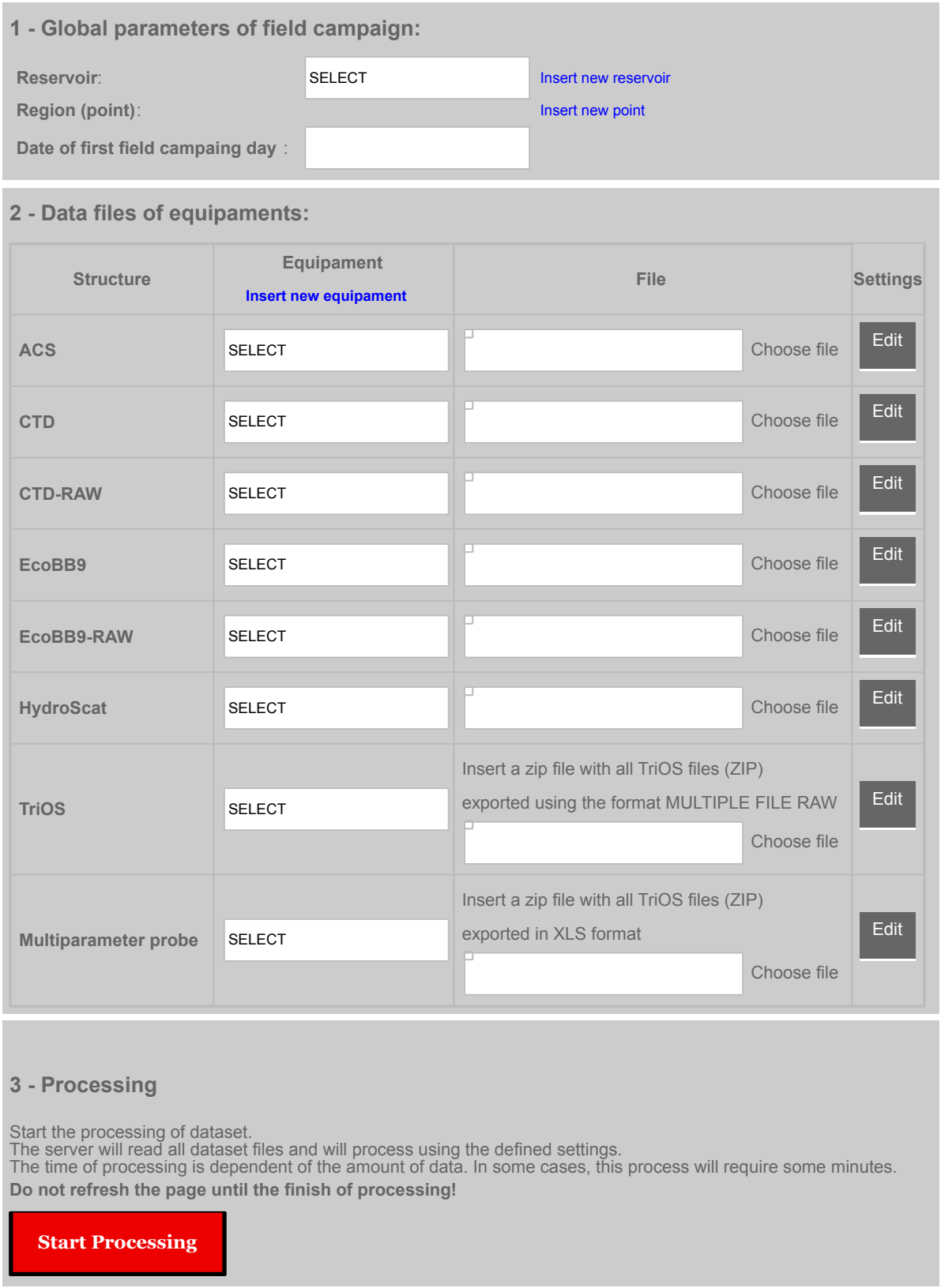


Figure 10 – Insertion interface to transform the original data file structure into database register

field surveys performed on water bodies of hydroelectric reservoirs. Based on results of previous works, we updated frequently the Aquopts to optimize the implemented protocols.

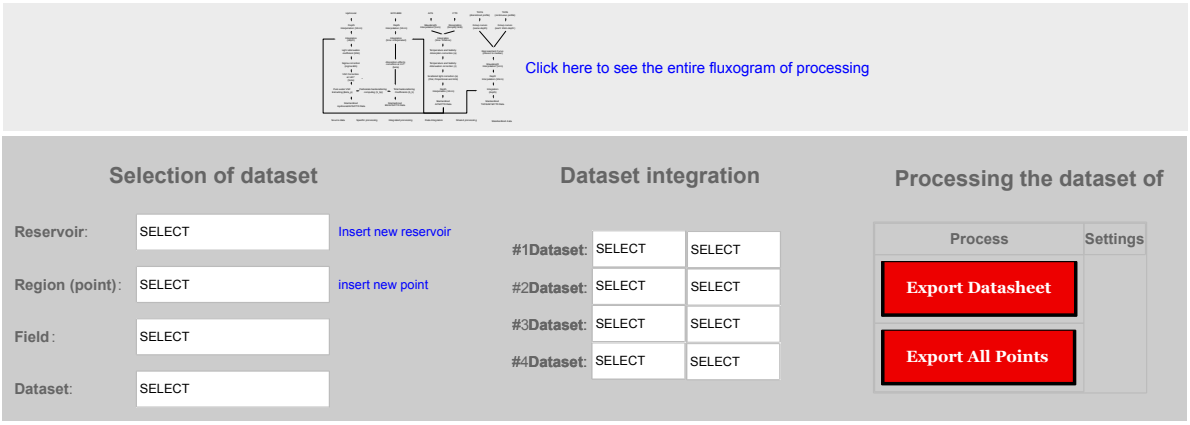


Figure 11 – Processing interface presenting some global parameters and default exportation functions. All other functions are sensor-dependent and appear accordingly with the selected dataset

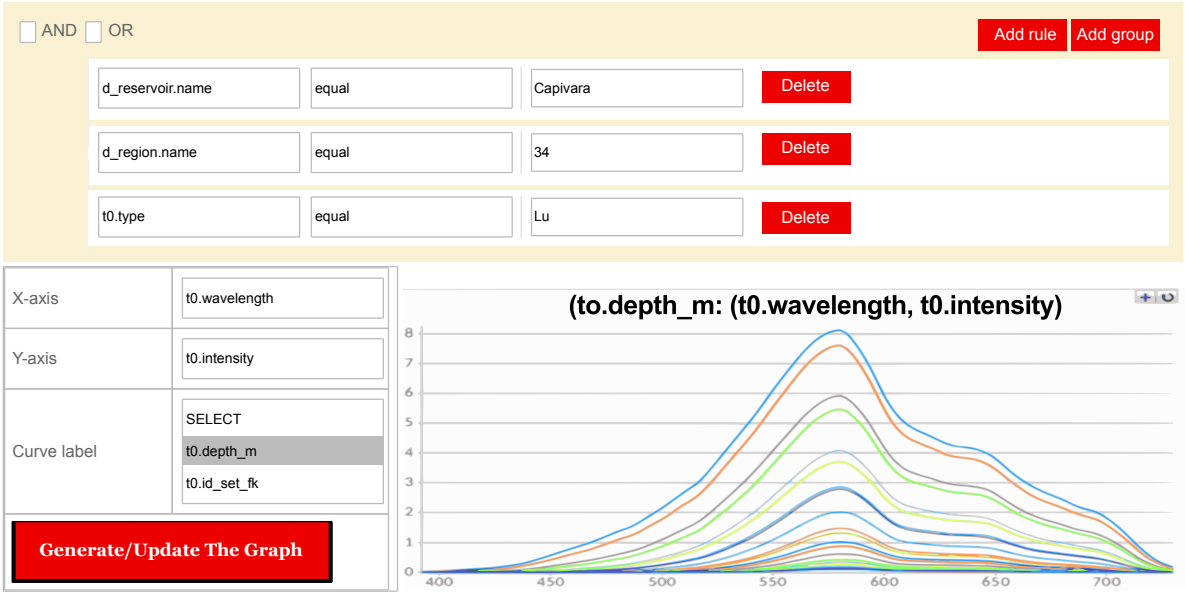


Figure 12 – Exploring interface with resources to dynamically filter the data and visualize the results with interactive graph

The Aquopts was used to process data collected by Hydroscat-6p and ECO-BB9 sensors to compute corrected backscattering coefficient from field surveys performed at Nova Avanhadava (NAV) and Barra Bonita (BB) reservoirs in Tietê River, São Paulo, Brazil (ALCâNTARA et al., 2016). After correcting the datasets, the authors choose values at 442 nm (Hydroscat-6p) and at 440 nm (ECO-BB9) to analyze the b_b, because these wavelengths are widely used and accepted as reference to study the OSCs of the water body.

The corrected backscattering coefficient profile is shown on Figure 13 with the higher and lower b_b values of two field campaigns for each reservoir. Looking the highest value graph, we can observe a high difference among the spectrum related to the wavelengths in the NAV, but BB present a closest behavior among the wavelengths. Nevertheless, the

lower values graph indicates a clear water column in NAV than BB, due to lower b_b values for NAV. In general, NAV present a high amplitude between the lowest and highest b_b spectrum, while BB present only a small difference between the high and low b_b spectrum. There is a high homogeneity in BB with high turbidity, in contrast to NAV that present a heterogeneity water, but with significantly low b_b. This behavior highlights the filter effects of the reservoir cascading system at Tietê River, because NAV stay at downstream position.

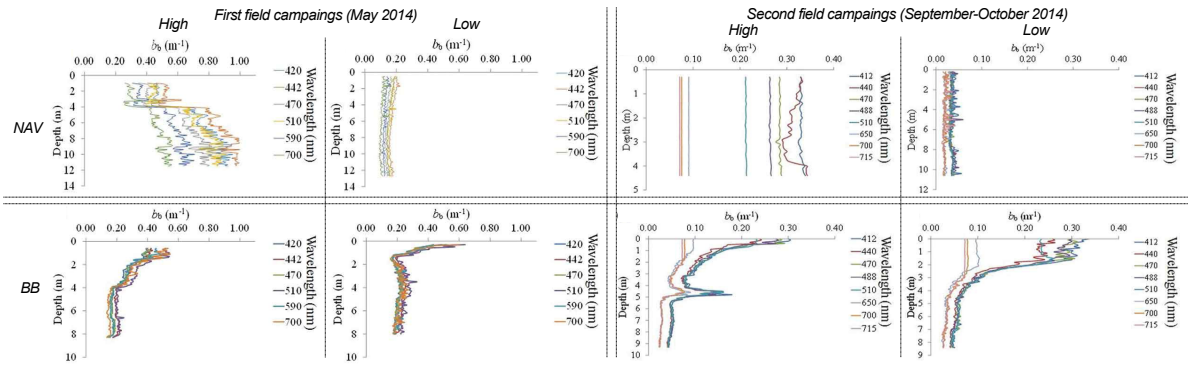


Figure 13 – Vertical profiles of b_b, corrected with the Aquopts, from the low and high value graphs of each field campaigns in the Nova Avanhadava and Barra Bonita reservoirs. Adapted from (ALCâNTARA et al., 2016)

The Aquopts system interpolated and corrected a TriOS dataset with the representative curve process, from a field campaign at Nova Avanhadava reservoir using the static profile way (RODRIGUES et al., 2016). The authors use radiometric data from TriOS to calculate remote sensing reflectance (R_{rs})(Figure 14) and discuss the low turbidity of the water body. The authors concluded that the R_{rs} curves represent a water body without prominent absorption features due to the shape of spectrum, but the magnitude of curves can indicate a turbid water.

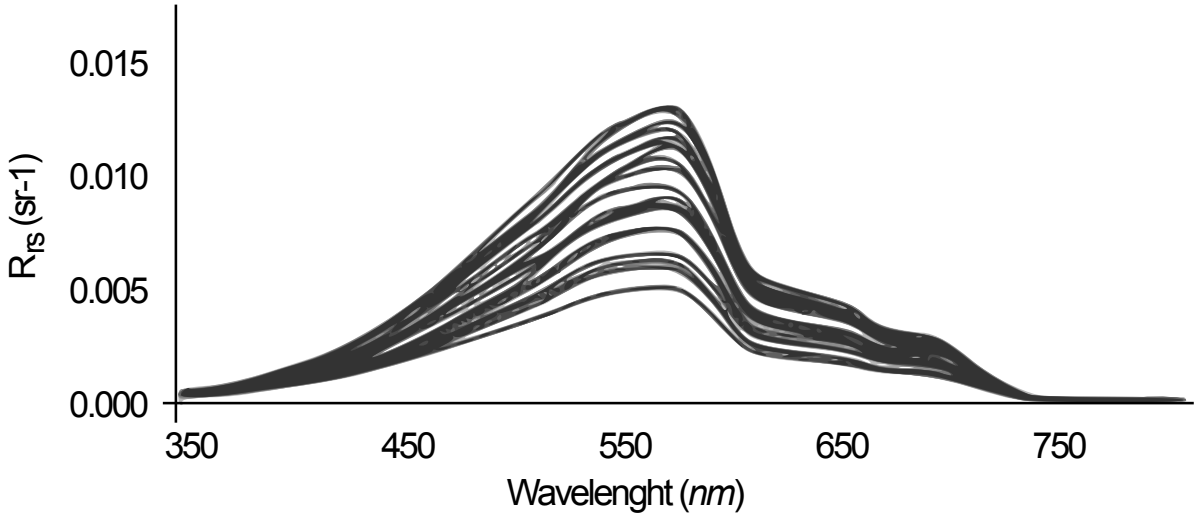


Figure 14 – Remote sensing reflectance of a field campaign performed in May 2014 at Nova Avanhandava reservoir. Adapted from (RODRIGUES et al., 2016)

The performance of the three methods of scattering error correction implemented on the Aquopts was analyzed to study the most suitable method to correct the absorption coefficient measured with AC-S sensor in a tropical eutrophic reservoir (WATANABE et al., 2018). The dataset was collected in two field campaigns and the data were corrected using the Aquopts. The results are compared to the total absorption coefficient extracted in the laboratory. The error analysis shown that among three correction methods, the flat and proportional algorithms achieved the best performances when applied in a dataset collected in a highly productive aquatic system, such as the Barra Bonita reservoir. The Figure 15 shows the comparison of absorption coefficient measured in laboratory and also the output of correction algorithms.

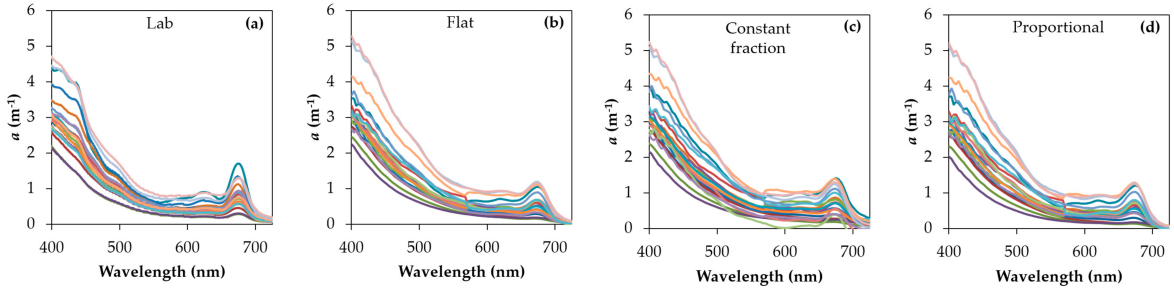


Figure 15 – Absorption measurements of a dataset collected in May 2014 from Barra Bonita reservoir presenting the laboratory sample analysis (a), compared with the *in-situ* AC-S measurements corrected and integrated with the online system using the flat (b), kirk (c) and proportional (d) algorithms. Adapted from (WATANABE et al., 2018)

3.5 Conclusions

This paper presents three important resources to be applied in geoscience works which perform field campaigns to collect datasets about water bodies: a database model for storing and managing multidimensional data obtained from multiple sources; a workflow with corrections, integration and processing suggested for five sensors commonly used in hydrological optics ; and the online system Aquopts developed with the previous features to automate the processes of storing, correcting, managing, and analyzing collected datasets.

We can apply the database model and the process workflow in other scenarios using additional approaches, since they represent an abstract logical concept. The Aquopts is a concrete implementation of these concepts that allows the expansion and addition of new resources, according to the research context. The database, workflow and system were validated with several datasets from field works performed by the SERTIE group. The system have been essential to the development of many valuable works.

The integration and normalization of data from different sensors and the automation of a controlled processing task are the main design guidelines for this work. Thus, this

approach is adequate to integrate sensors able to produce valuable information related to water components. Using the Aquopts, the preprocessing and integration tasks are easy and fast because it transform a manually high-cost job into a full automated process with controlled data quality. Consequently, researchers can focus on analysis of the dataset due to the optimization applied in the data preprocessing tasks.

Data availability

The dataset, python algorithms and additional information that support the findings of this study are openly available in IEEE DataPort™ at `<http://dx.doi.org/10.17632/f2tz548v2c.1>` (CARMO et al., 2019). Then online system can be accessed at `http://sertie.fct.unesp.br/aquopts` from a prior registration.

CHAPTER 4

A MOSAICKING TECHNIQUE TO AVOID SUN GLINT EFFECTS IN AIRBORNE HYPERSPECTRAL IMAGES

Abstract

Environments with complex mixtures, such as inland waters, present great challenges in the analysis and quantification of optically active components. Water interacts with electromagnetic energy, attenuating much of the radiation, and the resulting energy can be affected by noise. Specular reflection, also known as glint, is one of the main destructive effects of water signals in remote sensing. Flight surveys exhibit angular variations caused by aerial platform motion that can lead to the occurrence of glint in specific regions of images, but by using digital cameras, the number of acquired images can be increased to ensure respective overlap for redundancy at each image point. This work presents a new mosaicking approach that employs highly overlapped aerial images of a water body. The technique is based on a new multiscale algorithm that uses information from different views from global, regional and local neighborhoods to choose pixels that are less influenced by noise. We apply the proposed method to the image dataset of the Capivara reservoir with hyperspectral cubes for comparison with other known approaches based on statistical operators. The proposed multiscale algorithm presents the best results among all the tested algorithms, generating an output mosaic with a reduced number of outliers, decreasing inappropriate variability and producing no visual marks caused by noise.

4.1 Introduction

The remote sensing techniques applied in images produced by orbital satellites and photographs generated by aerial imaging platforms have been consolidated as important resources for obtaining data about water bodies (SU; CHOU, 2015). Water bio-optical models can be optimized using hyperspectral sensors because they offer a detailed spectral response and can be used for the recognition and quantification of specific features of optically active components (OAC) (LEE; CARDER, 2004). Water interacts with electromagnetic energy, with much of the radiation being attenuated. Noise can mask and override the low response of water captured by sensors of remote sensing platform. Specular reflection, also known as glint, is one of the main destructive forms of noise in remote sensing signals in water bodies (MARTIN et al., 2016; ZHOU et al., 2017; GROETSCH et al., 2017; ZHANG; WANG, 2010). Glint effects depend on the viewing geometry, sun elevation and azimuth, illumination conditions, and wind speed and direction (KUTSER et al., 2013).

Several works have applied glint correction to the final image mosaic (KAY; HEDLEY; LAVENDER, 2009; HOCHBERG; ANDREFOUET; TYLER, 2003; KUTSER; VAHTMäE; PRAKS, 2009). Among these methods, two types of glint corrections are mainly adopted: the first type is based on near-infrared channel spectral information (HEDLEY; HARBORNE; MUMBY, 2005; HOCHBERG; ANDREFOUET; TYLER, 2003), and the second type uses geometric modeling of a mirroring effect according to the field of view of the sensor and probabilistic models (COX; MUNK, 1954; COX; MUNK, 1956; YUROVSKAYA et al., 2018). Some methods assume conditions that might be invalid in some environments, such as a negligible water response in the infrared spectral bands, which is generally false for inland waters. Furthermore, data redundancy due to imaging of the same point from different views in overlapping areas, is important information that can be lost if glint correction is processed after acquiring the mosaic.

Glint affects portions of images in different ways and magnitudes (DOERFFER et al., 2008), but more information can be obtained by increasing the image overlap. Frame-based digital cameras allow increase of the longitudinal overlap, and the side overlap can be improved with additional flight paths. So, considering the data redundancy in overlapped area, with regions imaged from different viewpoints during a flight survey increasing the probability of at least one image present no glint effects due to the change caused by waves (COX; MUNK, 1954; COX; MUNK, 1956), in a preprocessing step, is it possible to choose pixels less affected by noise in order to build a mosaic to mitigate glint and other noise effects?

The hypothesis of this work is that highly overlapped images acquired using digital sensors in aerial platforms are sufficient to guarantee that all captured points have a best possible value representing the lowest glint-affected pixel based on the probability of waves

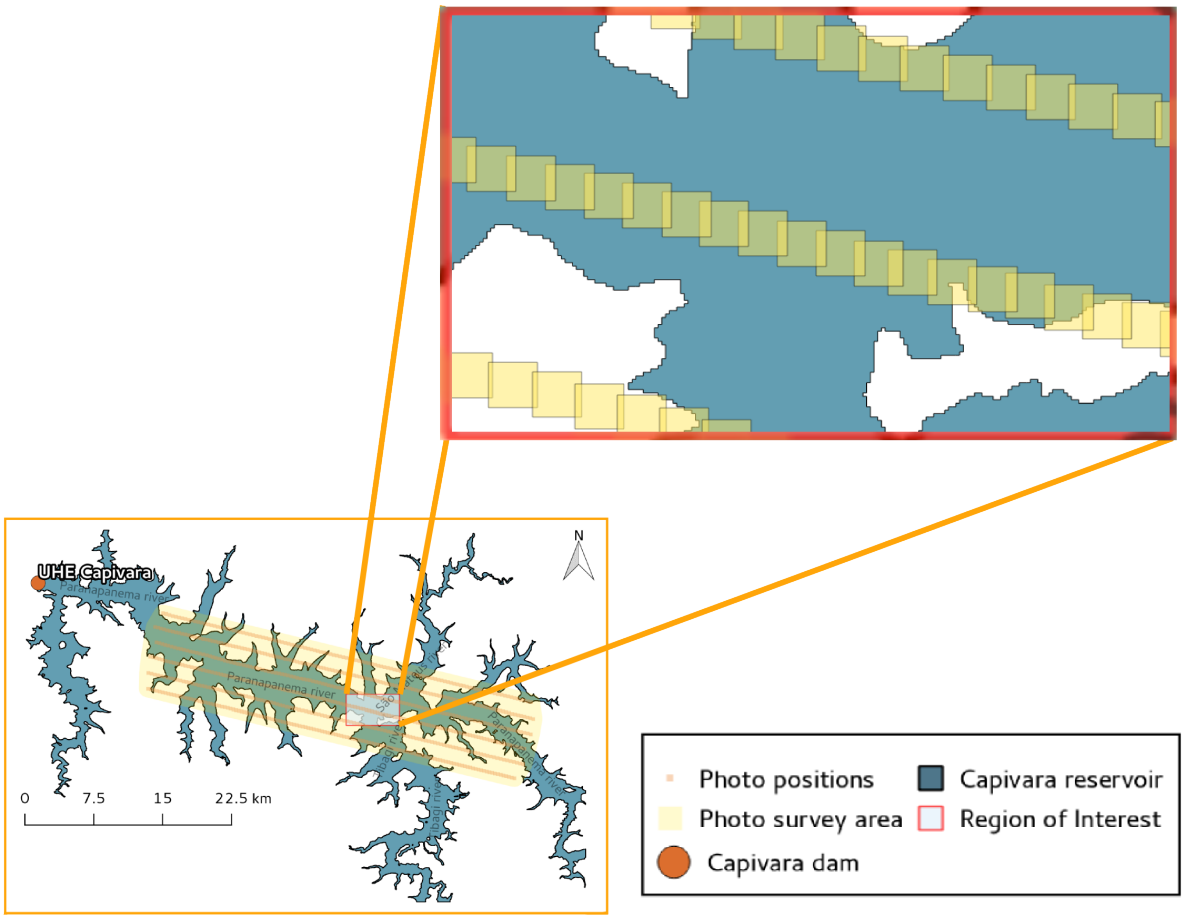


Figure 16 – Study area delimitation with flight path details and the region of interest

to change the water surface and consequently the geometry of sensor view (COX; MUNK, 1954; COX; MUNK, 1956). This work presents a new mosaicking approach with highly overlapped aerial images of water bodies, using a new algorithm based on local, regional and global scales to choose pixels less influenced by noise.

The proposed algorithm, named here as 6rb-multiscale, processes each image band separately and searches for a reference value to choose the closest candidate pixel for each point. The algorithm was developed and inspired by the concept of quadtree to define and process subblocks recursively (Strobach, 1991). Frequent and recursive block segmentation ensures that all local reference values reflect influences of the entire image.

4.2 Methods

We choose the strongest glint-affected region to delimit a region of interest (ROI) for the water body, as highlighted in Figure 16. Using a specific ROI, we can optimize the simultaneous presentation of all results to perform the visual and statistical analysis in this work.

In this work, we analyzed nine approaches to mosaicking individual georeferenced

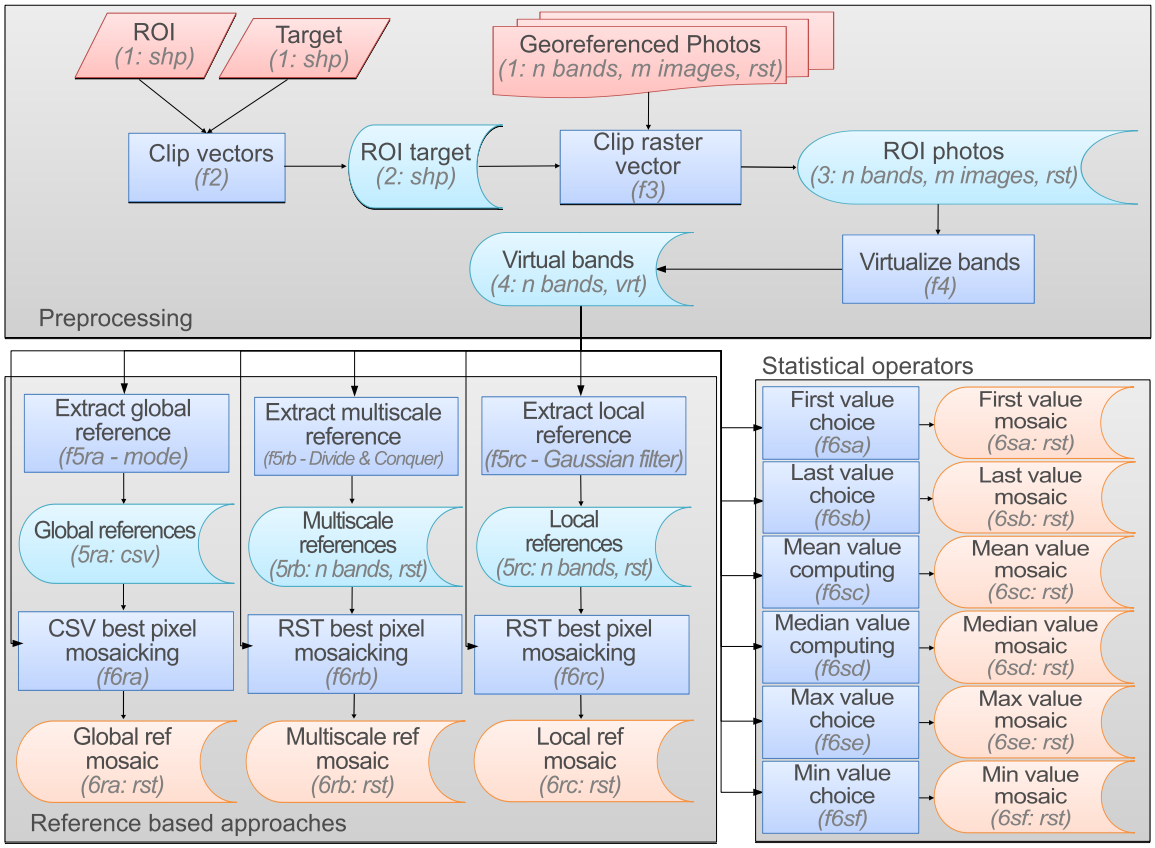


Figure 17 – Mosaicking steps of individual georeferencing photographs. The sequential identifier shows which steps can run in parallel programming, and the output file type is presented below each step.

images, which are described in Figure 17. The approaches are grouped into two larger categories: statistical and reference-based approaches. The statistical approaches process the mosaic using six known statistical operators (see Figure 17) that are commonly applied in geoprocessing software. Reference-based approaches, which are employed in this work, rely on a best pixel value choice using three new algorithms (Figure 17), each of them using a different technique to calculate a reference value and choose the nearest pixel value.

Each algorithm, presented in Figure 17, is identified with the nomenclature fx, where x represents the sequential processing order and the resulting product of the function. This nomenclature was used to demonstrate that the processes can be performed simultaneously in order to optimize parallel processing. The data types generated in each step are identified by the extension of output file: shp is a shapefile file with a vector structure; rst is a raster file with a matrix structure; vrt is a raster file with a virtual matrix structure, and csv is a file of pure text separated by semicolons.

4.2.1 Hyperspectral images preprocessing

The preprocessing task of this work has some important steps to ensure that all the redundant data in the images overlap. A ROI can be defined and cropped from the entire water body vector in step f2. Each band has a specific displacement due to a capture delay (see Section 4.2), then we apply a 100-meter buffer to the final ROI vector to keep only water pixels in the image. In step f4, all images are individually clipped to extract the intersection region. In this step, the data volume decreases, because the step discards files with no information on the ROI.

The last stage of preprocessing uses an OGR-GDAL library functionality to build a virtual band set. Each file is inserted into an individual layer of the virtual band set, so that all overlapping pixels are available with simplify access through links to the source file. The virtual band set avoids duplicate information and optimizes access to the values of each georeferencing position. The algorithms applied in this paper use the original digital number (DN) of the images, i.e., the pixel values lie in the normalized DN range of 0-255.

The virtual band set of each image band constitutes a information cube of dimensions mxnxp, where m represents the number of lines in the scene, n is the number of columns, and p is the dynamic number of overlapped pixels at each point. The mosaicking process has the objective of transforming 3d cubes of dimensions mxnxp into 2d matrices mxn representing each band.

4.2.2 Statistical approaches

Each pixel has different values in the overlapping regions, but a unique value is required for the mosaic. In this work, the following known statistical operators were used for comparison with the proposed methods: 6sa-first (value of the first image); 6sb-last (value of the last image); 6sc-mean; 6sd-median; 6se-max and 6sf-min.

Other operators using weighted averages and gradient analysis can be used to minimize the transition effects between adjacent images (BURT; ADELSON, 1983; PRADOS et al., 2012). However, the weighted mean was not used in this study because this technique should be applied to each pair of images, leading to possible noise effects in the images and making the process computationally expensive.

4.2.3 Reference-based approaches

The proposed method uses three new approaches to choose the best pixel for each point instead of obtaining an artificial value calculated from the input values of the overlapped images. Reference-based approaches can be described by two distinct processes, as shown in Figure 17: reference value computation and choosing the closest pixel to the reference value. The task of choosing the closest pixel to the reference is the same for all

proposed approaches, but the reference value computation is different due to scale view: global, local and multiscale.

The global reference, the 6ra-global algorithm, was calculated from a single band mode value for the entire scene. The mode was chosen taking into account the homogeneous behavior of a water body and assuming that there are more uncorrupted values than noise-affected values .

The local reference, the 6rc-local algorithm, was calculated from the neighborhood of each image pixel. Processing the full band cube with a moving window requires a large computational effort. To simplify the process, we calculate the median of all overlapped pixels and use a moving Gaussian filter.

The Gaussian morphological operator provides a good representation of the expected spatial variability of a water body in which the regions closest to a central point (high weight) have similar characteristics that gradually change (decreased weight) as the distance increases.

We developed the reference-based multiscale algorithm, 6rb-multi, to compute each reference value using different weights from all scene pixels. The impact of an entire scene, even a small impact, will be represented by each reference, according to a contribution factor. Our approach uses the concept of quadtree (Strobach, 1991) to break up the image and process it from a global scale that is recursively subdivided into several levels until reaching a local region.

In each multiscale algorithm interaction, the parent block is divided into four child blocks. Segmentation is performed in the plane to keep all overlapped pixel values on the third axis. For each block division, we compute the median of the three-dimensional block and calculate a weighted average of the parent block reference using a contribution factor. Other measures with a central tendency can be used in addition to the median. This process runs recursively for all child blocks until it reaches the minimum block size limit.

The contribution factor and the block size limit can be modified to match specific data characteristics. We use a contribution factor of 10% and a minimum block size limit of 10x10 pixels, i.e., in each algorithm iteration, the reference value will be computed by 90% of the child blocks median and 10% of the previewed parent block reference, and this process will run until the child block reaches the 10x10 size limit.

4.2.4 Assessment of the results

We tested nine algorithms for rendering the mosaic from individual georeferenced images. We evaluated the performance of each algorithm in three different ways: boxplot graphics, information entropy and false color compositions.

Boxplot graphs were used to analyze the behavior and variability of the data. With this representation, it is possible to verify the existence of outliers and to compare the dispersion of data and their respective concentrations in each of the significant quartiles.

Entropy is a metric of information theory that is used to quantify the average degree of uncertainty in a dataset, i.e., the entropy expresses the amount of information in the dataset. A theoretical model was originally described to quantify noise in digital message communication (SHANNON, 1948); however, the concept of information entropy is widely used in different contexts (STRAIT; DEWEY, 1996; ZENG et al., 2010; JAT; GARG; KHARE, 2008). Considering $V_1, V_2, ..., V_n$ as the pixel image values with respective occurrence probabilities of $P(V_1), P(V_2), ..., P(V_n)$, then the information entropy using b logarithm bases is defined in (4.1).

$$H(V) = -\sum_{i=1}^{n} P(V_i) \times log_b P(V_i) \tag{4.1}$$

In this work, we use base-2 log entropy to describe the variability of the algorithm outputs in binary units (bits). Binary units help to interpret how many bits are required to represent the information in the dataset, e.g., if a dataset has eight different values with the same frequency, then the base-2 log entropy will be three because this is the number of bits necessary to represent the data ($2^3 = 8$). We expect homogeneous behavior without great variations along the water body, so a high entropy value represents high variation in the pixel values, indicating the possible existence of noise. Thus, a low entropy represents greater homogeneity and identifies the best results.

The final results are plotted with a false color composition using bands 19, 3 and 1 for the RGB configuration. Unexpected variations can be identified by visual marks inside the water body.

4.3 Results

Glint effects are visible with the increase in the brightness values, so the affected points are saturated. Figure 18 depicts the response of all pixels representing the water body region of interest using the result from the f4 process (virtual band set) with all original non-corrected values.

All bands have similar data dispersion around the median. The first quartile represents 25% of the data below the median and is concentrated in a smaller range than third quartile, which represents 25% of the data above the median, i.e., there is more data variability above of the median than below it. In Figure 18, we can also note that values smaller than the median are still within the lower limit of the outlier data, while the values

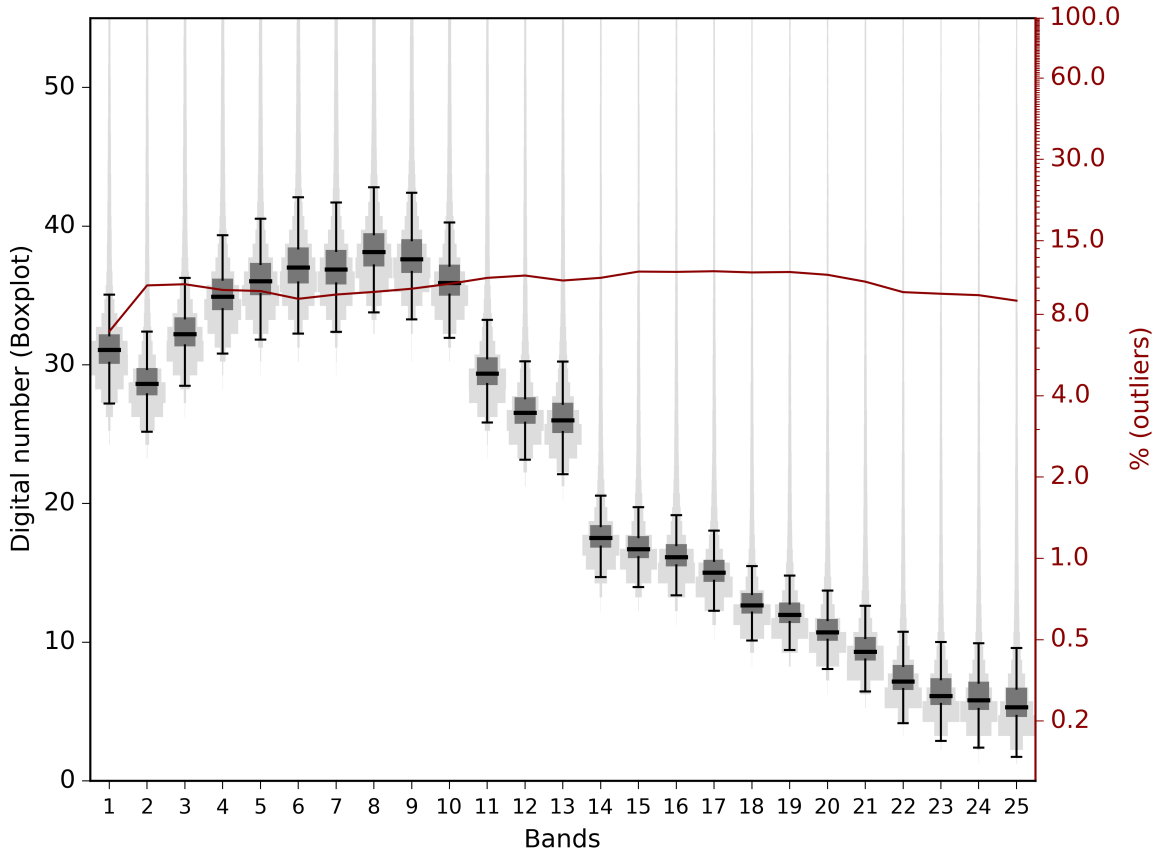


Figure 18 – Boxplot graph of central measures and variation in all water pixels for each band. The y boxplot axis (left) is limited by two standard deviations. There are outliers above the maximum boxplot limit represented by the outlier percentage in line (right axis on a logarithmic scale).

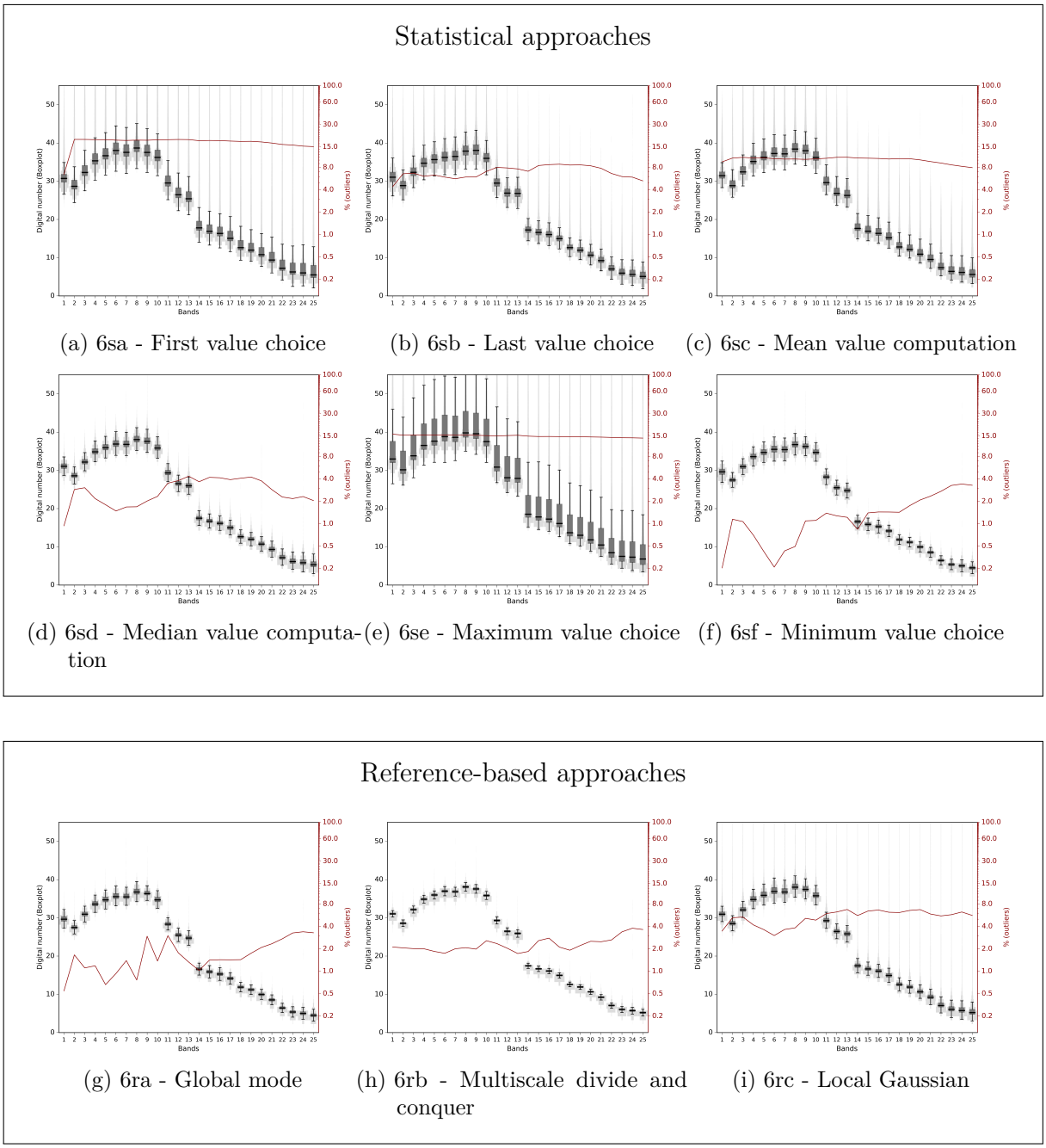
above the median stratify continuously on the upper limit, with the number of outliers represented by the line graph on the same plot.

Figure 19 presents boxplot graphs for each algorithm to enable an assessment of the process efficiency.

All the algorithm outputs maintain approximately the same average spectral signature. In general, reference-based approaches show a more significant decrease in the data variability than statistical operators. The 6se-max and 6sa-first algorithms present the highest number of outliers and interquartile range, with average digital numbers of 6.08 and 2.66, respectively. The 6rb-multiscale and 6ra-global algorithms exhibit smaller interquartile amplitudes, with average digital numbers of 0.43 and 0.90, respectively. The 6sf-min, 6ra-global and 6rb-multiscale algorithms have the smallest number of outliers.

Observing the information entropy, (Figure 20), we can quantify the variation that remains after the glint correction of each algorithm.

The entropy graphs show the different behavior of each applied algorithm. The



(a) 6sa - First value choice (b) 6sb - Last value choice (c) 6sc - Mean value computation

(d) 6sd - Median value computation (e) 6se - Maximum value choice (f) 6sf - Minimum value choice

(g) 6ra - Global mode (h) 6rb - Multiscale divide and conquer (i) 6rc - Local Gaussian

Figure 19 – Graphs of the central tendency and variability of all bands after applying each algorithm to a clipped region of interest. The left axis is for the boxplot, and the line (right axis) represents the outlier percentages.

6rb-multiscale algorithm presents the lowest values for all image bands. On the other hand, the 6se-max algorithm has the highest entropy levels in all image bands, i.e., there is greater data variability even with glint correction. The difference between the algorithms with higher and smallest entropy is about a factor of four.

The high data variability and existence of outliers due to glint noise result in saturation areas in the image. Figure 21 presents the results of algorithms applied to the ROI. To enhance the variations of the resulting images, the image values produced by

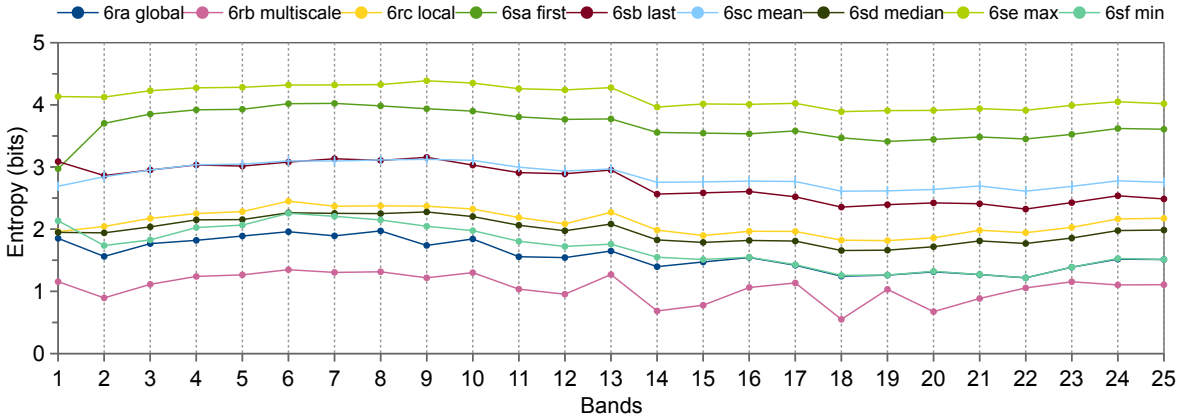


Figure 20 – Information entropy of all results from each algorithm

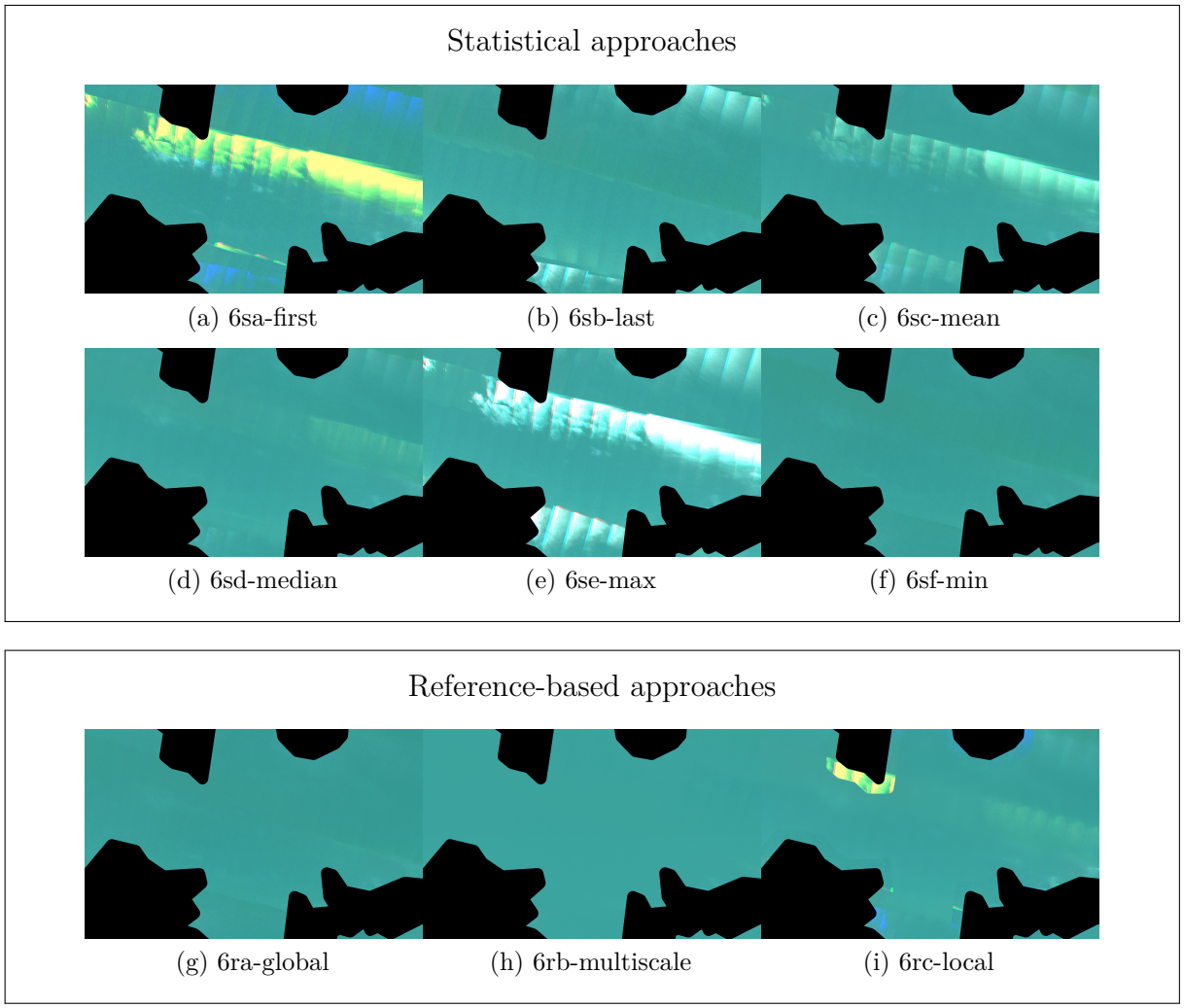


Figure 21 – Processing results of all algorithms applied to the image clipped area. The images are shown using RGB false color composition with bands 19, 3 and 1.

each approach was rescaled to a range of 0 to 50 DN. This range was adopted because it represents the interval of about 95% of the image data, i.e., two standard deviations with respect to the global mean.

We can note some noise by visually analyzing the resulting processed images in Figure 21. Some of the applied algorithms highlight the glint effects in the whole image, as seen in the results for the 6sa-first, 6sc-mean and 6se-max algorithms. The 6sb-last, 6sd-median, and 6rc-local algorithms have better performance when calculating or choosing each pixel value and exhibit glint effects only in a few regions of the image. The 6sf-min, 6ra-global and 6rb-multiscale algorithms present great homogeneity throughout the image, but some soft marks remain in the results for the 6sf-min and 6ra-global algorithms.

4.4 Discussion

Water absorbs much of the electromagnetic energy from the sun, and only a small amount of energy is backscattered and captured by the sensor. Because of the small magnitude of the backscattered energy leaving the water body, noise affects the signal critically.

All bands are affected by noise that increases the DN and deteriorates the signal. Observing the original values from the uncorrected images, expressed in the boxplot of Figure 19g, we can see the strong incidence of glint indicated by the amount of data above the median. There is a variation below the median, but the amount of data is insignificant compared to the values above the median, emphasizing the glint behavior increases and saturates the pixel values. There are about 8% outliers in every band.

Applying the correction algorithms, we expect data convergence to a central value so that sparse variations of the spatial data are decreased. This is true for a water body because of the high spatial correlation of the data. A water body has soft-variation behavior with no abrupt changes, and even in highly turbid water, a spatial dependence exists. In this way, we consider that results with lower variability are better than results with sparse variations.

From the corrected image boxplots, Figure 19, and false color compositions, Figure 21, we can see that the 6rb-multiscale algorithm presents the best results, with the lowest data variation. The 6sf-min and 6ra-global algorithms show good variation behavior in the boxplots, but some visual marks remain in the images of the false RGB compositions. Note that, due to a high interquartile range limit, the 6sf-min algorithm (Figure 19f) presents a small outlier percentage level for all bands, but the image results (Figure 21f) display identifiable visual marks.

The entropy information presented in Figure 20 highlights the good performance of the 6rb-multiscale algorithm. Note that the two other algorithms that presented a low boxplot variation, 6sf-min and 6ra-global, achieve small entropy values in all bands, but they are nonetheless larger than that of the 6rb-multiscale algorithm, confirming the 6rb-multiscale algorithm as the best algorithm considered in this case study.

4.5 Conclusions

Several works have presented different methods for glint correction. However, all the methods are applied on a processed full mosaic in order to eliminate noise caused by glint effects. To solve these drawbacks, this work proposes a new approach to eliminate or minimize glint before generating the mosaic, using the overlapping information of the captured points to choose the most representative pixel.

The approach proposed in this work does not produce new pixel values since it assumes that all redundant points captured from the scene have at least one candidate value that is not contaminated with noise. The good results obtained with the proposed approaches based on the best choice of the pixel value closest to a given reference demonstrate the power of the approaches compared to other statistical techniques. The 6rb-multiscale algorithm proposed in this work is a novel technique that allows the acquisition of references of each image point using a recursive quadtree strategy to reflect the influences of all neighbor levels, resulting in a reference value map constructed from global, regional and local characteristics.

The application of our approach to water bodies has special benefits due to the high probability of changes in water surface due to wave, where smooth variations with a high level of spatial correlation is expected. A reference-based approach can compute suitable results and be applied in different scenarios according to the following criteria: the area of interest must have high spatial correlation ; and the overlapping images obtained by a frame camera should be sufficient to support the capture of all points from different acquisition geometries and ensure that at least one pixel is not influenced by noise. The number of bands, radiometric resolution and acquisition platforms are independent of the proposed approach.

Data availability

The dataset, python algorithms and additional visual information that support the findings of this study are openly available in IEEE DataPort™ at `<http://dx.doi.org/10.21227/dxgz-3m95>` (CARMO et al., 2019b).

CHAPTER 5

IMPROVING ON EMPIRICAL LINE METHOD APPLIED TO INLAND WATER HYPERSPECTRAL IMAGES COMBINING REFERENCE TARGETS AND IN SITU WATER MEASUREMENTS

Abstract

Empirical line methods are frequently used to correct images from aerial remote sensing. This method of correction is performed in two steps: the first stage finds the calibration equation that represents the interval of the data and the second step transforms the image data into the established quantity using the equation found. Many works have been successfully applied empirical lines over land areas, but it is still a great challenge to correct images from water bodies. Remote sensing of aquatic environments captures only a small amount of energy because the water absorbs much of it. The small signal response of the water is proportionally smaller than the signal from other land surface targets. This work presents a new approach to calibrate empirical lines combining reference calibration panels with a water point. We optimize the method until reaches the best result. For this purpose, we evaluated several combinations of targets using both linear and exponential fit. The best matching was provided using an exponential fit using a single gray reference panel combined with a water point resulting in a coefficient of determination about 0.87,

a root mean squared error of 0.002 sr^{-1} and a mean absolute percentage error of 18%. This approach presented suitable results to derive reflectance data from the raw digital numbers from images.

5.1 Introduction

The use of empirical line method (ELM) to calibrate remotely sensed images to surface reflectance is the most widely used approach for airborne imagery because it is straightforward and effective (SMITH; MILTON, 1999). This method is important to generate reflectance data that are most comparable unity to other spectra measured in the field or in the laboratory (GAO et al., 2009). The ELM consist to establish a prediction equation for each waveband to find the coefficients of the curve resulted from the fit made between the digital number (DN) of calibration panel in the image and the in situ sensor measurement (reflectance) (WANG; MYINT, 2015; STOW et al., 1996). Several works implemented the ELM with variable success (SMITH; MILTON, 1999; KARPOUZLI; MALTHUS, 2003).

The ELM requires reflectance field measurements for at least one bright and one dark reference captured by the image to find an equation representing the variation of the calibration data (CONEL et al., 1987). The reference targets must cover the whole spectrum with an appropriate size related to the pixel resolution of images (KARPOUZLI; MALTHUS, 2003). These prediction equations, which attempt to remove the illumination and atmospheric effects, are then applied to the remotely sensed data to produce images in units of reflectance (SMITH; MILTON, 1999).

The ELM is usually applied on remotely sensed data of land surface, but is a challenging to apply it on water body images due to the low radiance from water targets, which in some cases it is lower than black references adopted in land surface image calibration (PöLöNEN H.-H. PUUPPONEN, 2014). Water response presents a lower contribution to at-sensor signal than other land targets (KARPOUZLI; MALTHUS, 2003). The water has a low-reflectance response because of its high absorption coefficient and is very important the removal of atmospheric effects or its normalization (HADJIMITSIS; CLAYTON; RETALIS, 2009). Deriving the water leaving reflectance from imaging data requires a very accurate modeling of ELM and different targets and calibration panels (GAO et al., 2009).

To improve the performance achieved by ELM, different amount of targets can be used, but the main fact is the contrast between each target used as the reference spectra (SMITH; MILTON, 1999; KARPOUZLI; MALTHUS, 2003; NAZEER; NICHOL; YUNG, 2014). Another approach is the combination of both artificial calibration panels with field spectral samples to calibrate the ELM. This setting archived good results over a

land surface, but the error on water image are still high (STABEN et al., 2012).

The success of ELM applications over a land surface and the inaccurate results over water highlight a need to analyze the application of ELM in aquatic environments. We evaluated several combinations of targets using both linear and exponential fit to correct water body images. The dataset has 26 water spectra acquired with a hyperspectral camera on board of a lightweight aircraft in a region of inland water. We also checked the calibration of ELM using combinations of three reference panels, white, gray and black.

5.2 Methods

The Rikola camera has a radiometric resolution of 12 bits and produces Digital Numbers (DN) on the range from 0 to 255 DN, with float precision. There is a redundant value of each point – because of images overlap – and we use a median to get a unique representative spectrum of each point (PöLöNEN H.-H. PUUPPONEN, 2014). We extract the representative spectral DN curves of all twenty-five bands from the artificial targets and the water sample points. We present the syntheses of the method steps on the Figure 22. The workflow shows all interaction to process the variables to generate the inputs of the ELM.

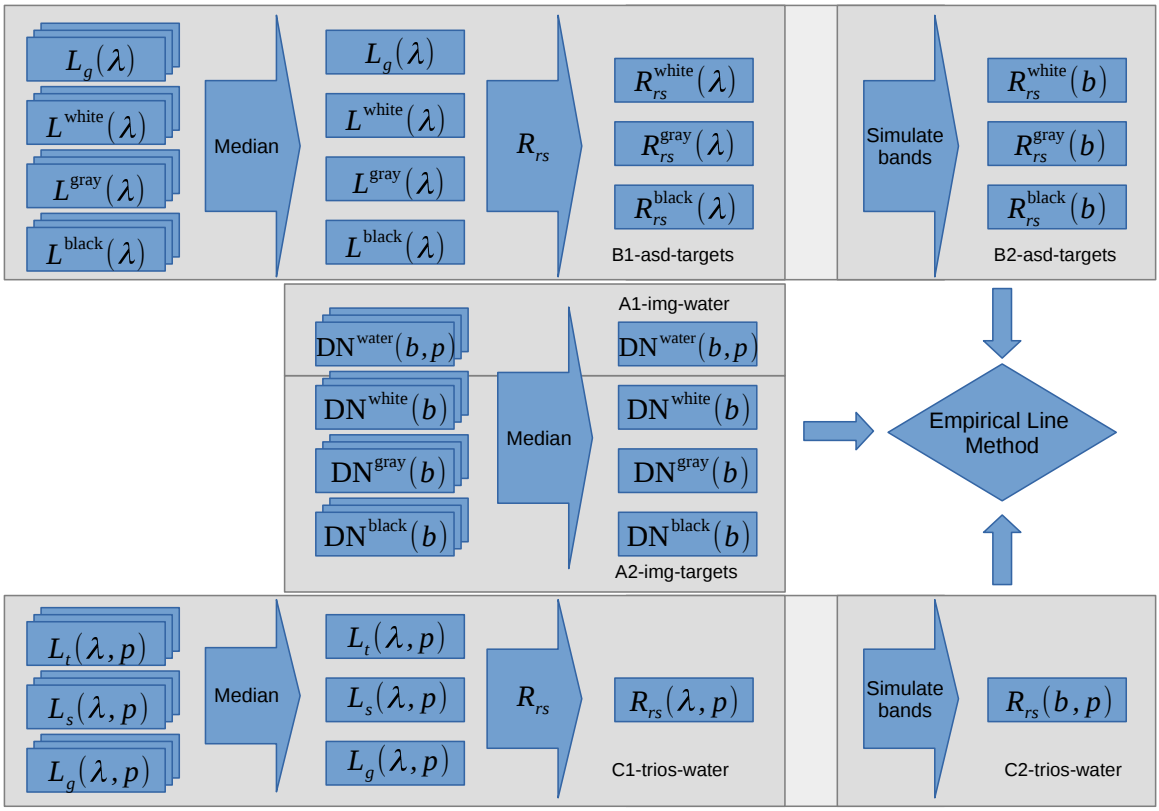


Figure 22 – Workflow to processing the measurements of water and reference targets to calibrate the Empirical Line Model. Some variables are expressed in function of points (p), wavelengths (λ) and bands (b).

We process the ELM using two different fit function to transform DN to R_{rs}: linear and exponential. Two parameters a and b describes both models, linear and exponential, as showed in the Equations (5.1) and (5.2).

$$R_{rs} = a \times (\mathrm{DN}) + b \qquad (5.1) \qquad R_{rs} = a \times e^{b \times (\mathrm{DN})} \qquad (5.2)$$

To analyze the results of each model, we use the coefficient of determination (r^2), the root-mean-square error (RMSE) and the mean absolute percentage error (MAPE), defined by the Equations (5.3), (5.4) and (5.5), respectively.

$$r^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \overline{y})^2} \qquad (5.3)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \qquad (5.4) \qquad \mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - f_i}{y_i}\right| \times 100 \qquad (5.5)$$

At the end of process, we have data from the images (DNs) and from the spectroradiometers (R_{rs}) of 26 water sample points and 3 calibration panels (white, gray and black). The spectroradiometer R_{rs} data is the reference value to transform all DN using a specific ELM for each of 25 bands. We build ELMs using different combinations using two approaches: only calibration panels and other including a water spectrum as a calibration point. We calculated a median spectrum from 26 points collected in situ and choose the closest curve as the water calibration sample. The median was adopted due to glint noise present on the water samples, in this way, the median isolates the saturated values and presented a low noise influence than average.

5.3 Results and Discussion

The results of processing, in function of the points and the bands, are the input data to the ELM: R_{rs} of calibration panels, R_{rs} of water points, DN of calibration panels and DN of water points. The Figure 23 shows the spectral curves of digital number and reflectance from water and calibration targets.

Analyzing the water spectra to compare the digital number to the reflectance, we can note that a direct linear relationship are not visible for all data. The water response is very small in comparison to the calibration panel. This issue is one of the main challenge in the remote sensing of water bodies since the reflected light from water bodies is lower than regular black tarpaulins.

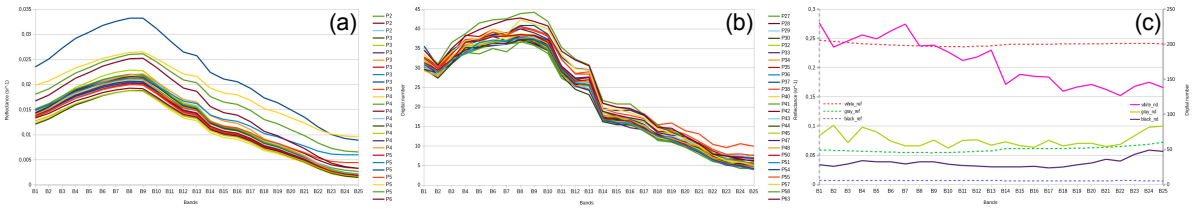


Figure 23 – Input data of ELM. Water reflectance (a) and digital number (b). The reflectance and digital number of the calibration panels (c) share the same x-axis but with specific y-axis in different unities.

To analyze the performance using the usually adopted approach, we calibrated a linear and exponential ELM using the white, gray and black panel evaluating with the twenty-six water points. The Figure 24 depict the ELM application and its associated errors using a linear and an exponential fit.

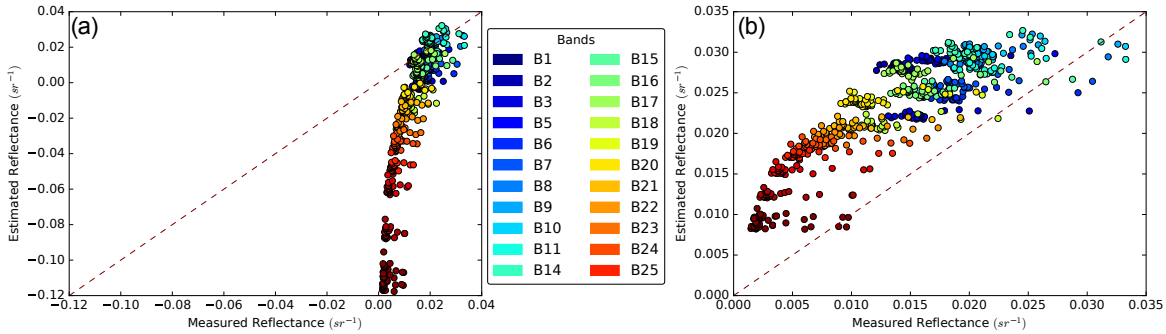


Figure 24 – Linear (a) and exponential (b) fit of ELM calibrated with white, gray and black panels evaluating using all 26 water points.

High errors are obtained using only the three panels to calibrate the ELM, as shown in Figure 24. The values predicted by the linear fit underestimate the reflectance values, generating negative values, while the exponential fit overestimate all data. The errors in the exponential fit are lower than linear fit, and the coefficient of determination stay around 75.68%, against 73.4% of the linear fit.

The linear and exponential function do not result in a good fit when calibrated only with the white, gray and black panels because the water reflectance values stay out of the range of calibration. The spectrum of water was usually lower than black panel, so, including a water spectrum point to calibrate the ELM can improve the match of the model. We add the P28 sample on the calibration process because it is the closest spectrum to a median of the 26 water reflectance. The linear and exponential fit calibrated with the three panels plus the P28 spectrum improve the results, as showed on Figure 25, and produce a coefficient of determination of 82.49% and 80.06%, respectively.

The error in shorter wavelengths decreases when including a water point on the ELM calibration, but the longer wavelengths present a high error yet. The end of visible

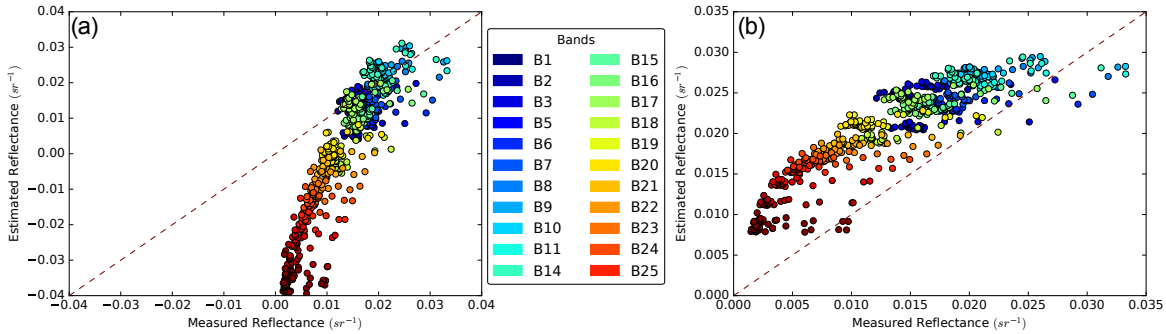


Figure 25 – Linear (a) and exponential (b) fit of ELM calibrated with white, gray and black panels plus the P28 water spectrum and evaluated using all 26 water points

spectrum has a lower reflectance of the entire spectrum. Negative values stay occurring with the linear fit.

Due to the low reflectance of water, the responses of water point is closest to the black panel (Figure 26). In this case, the black panel is redundant because of the proximity with the water spectra.

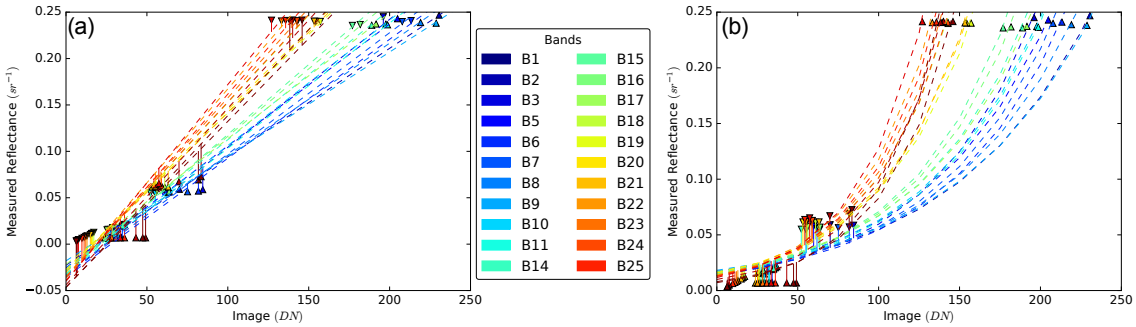


Figure 26 – Linear (a) and exponential (b) fit of ELM calibrated with white, gray and black panels plus the P28 water spectrum and evaluated with the same points. The vertical line represent the error of each measurement and estimated point.

If we remove the black panel, the white and gray panel will have a greater weight compared to the low response of water point, reducing the coefficient of determination to 79.61% and 75.98% for the linear and exponential fit, respectively. Looking at Figure 26 it is possible to note that the low values produced by the high wavelengths change the behavior of the function. Therefore, the ELM calibrated using over one reference panel – which values are considerably higher than those of water – employ a greater tendency to fit well only for the targets with higher reflectance and deteriorate the adjust of low reflectance water points.

A calibration set with a water point and any target to ensure the variability of the data - it does not apply to the black panel - can improve the results of ELM. Using P28 and the white panel to calibrate the ELM, a determination coefficient of 85.81% and

86.46% can be obtained for the linear and exponential adjustments, respectively. However, the smallest error and the best fit, showed on Figure 27, is achieved by calibrating the ELM with P28 and the gray panel, resulting in a coefficient of determination of 86.52% and 87.31% for the linear and exponential fit, respectively.

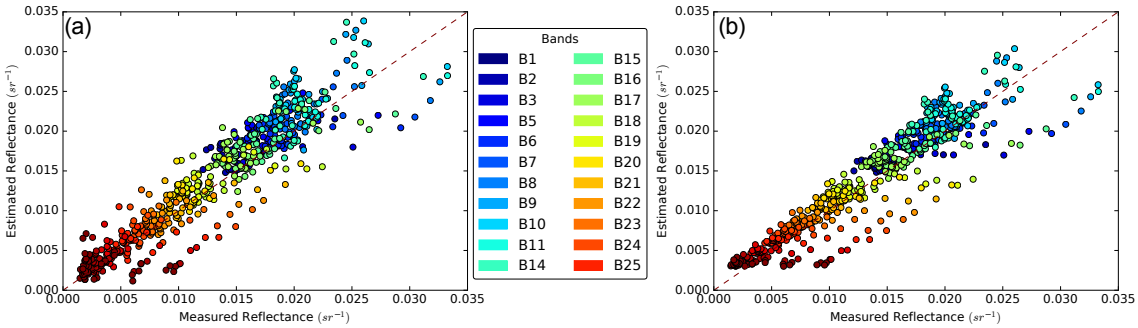


Figure 27 – Linear (a) and exponential (b) fit of ELM calibrated with gray panel and the P28 sample and evaluated with all 26 water points

The ELM calibrated with the gray panel and a water point improve the reflectance response and mitigate the negative values (Figure 28). The reflectance predicted by the ELM calibrated with the gray panel and a water point maintain the features of the measured reflectance (Figure 23(a)), mainly considering the linear fit (Figure 28(a)).

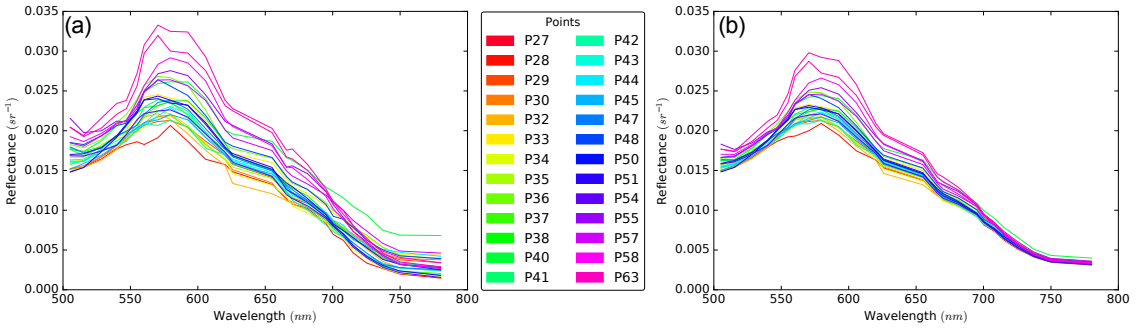


Figure 28 – Reflectance from linear (a) and exponential (b) fit of ELM calibrated with gray panel and the P28 sample applied to all 26 water points

Table 3 shows a statistical synthesis of all ELM evaluated in this work. The results of ELM calibrated with the gray panel and P28 is very similar for both linear and exponential fit. The largest difference between the two models lies on the higher wavelengths because in this range the linear fit overestimates the values due to the small amount of energy reflected. Exponential fit shows a good response over the entire spectrum.

The ELM calibrated with only one water point and a single reference panel far from the water point produced the best results, independent of the adjusted model. The best model eliminated negative values, except the calibration performed with the white panel and P28, which presented negative values for some curves at the longest wavelength.

Table 3 – Synthesis of all ELM evaluated in this work ordered by the best coefficient of determination

Panels	Water	Fit	Negative values	r^2	RMSE	MAPE
gray	P28	exponential	no	0.8757	0.0025	18.0076
white	P28	exponential	no	0.8676	0.0025	17.2695
gray	P28	linear	no	0.8675	0.0029	19.8540
white	P28	linear	yes*	0.8612	0.0030	24.3768
white, black	P28	exponential	no	0.8589	0.0049	29.1378
gray, black	P28	linear	yes	0.8549	0.0067	86.8265
gray, black	P28	exponential	no	0.8509	0.0048	40.2822
white, gray, black	P28	exponential	no	0.8011	0.0079	94.8108
white, black	P28	linear	yes	0.7978	0.0143	216.7235
white, gray, black	P28	linear	yes	0.7895	0.0179	273.6657
white, gray	P28	linear	yes	0.7877	0.0106	150.1309
white, gray, black	–	exponential	no	0.7566	0.0099	113.0873
white, gray	P28	exponential	no	0.7562	0.0136	159.0230
white, gray, black	–	linear	yes	0.5420	0.0438	691.3565
black	P28	exponential	no	0.3337	0.0747	97.1490
black	P28	linear	yes	0.2364	0.2388	250.0960

*: negative values only on the last band

Combining the water point with only the black panel generates the worst results. We expected this behavior because the values of the black panel are close to the water values. The proximity of the two spectrum does not contemplate the variability of the data and the curve using only these targets becomes biased.

The differences on the results applying different setting of reference targets can be due to specular reflection. The tarpaulin targets was placed on the irregular ground, and the small variation of the surface can caused specular reflections. The better fit achieved using the gray target combined with a water point can indicate a low degree of specular reflectance influence on the gray tarpaulin than white and black targets.

5.4 Conclusions

This work presented a new approach to improve the performance of empirical lines modeled to transform the values of digital numbers captured by hyperspectral images of water bodies in reflectance. The water attenuates the energy strongly, resulting in a small amount captured by the sensors, which makes it inviable to calibrate the ELM with just reference panels. The purpose of this work is to include a reference value of water on the calibration process.

Using artificial calibration targets, as successfully applied in the literature with other terrestrial targets, does not describe the variability of water response. The use of calibration panels is fundamental to extend the variability contemplated by the empirical

line, however, its isolated use does not give good results applied in water bodies. Integrating a single point of water with any of the targets sufficiently far from the water spectrum has verified effective to improve the performance of the empirical line applied to the water body.

These results contribute to improve the ELM performance in water bodies, but also confer benefits to the field survey in aquatic environments, as it eliminates the need to manage unnecessarily many targets. The ELM calibration proposed in this work differs from those applied in other works, and we believe this approach improve the application of ELM to correct water imagery data.

Data availability

The dataset, python algorithms and additional graph information that support the findings of this study are openly available in IEEE DataPort™ at <http://dx.doi.org/10.21227/fbks-rj40> (CARMO et al., 2019a).

CHAPTER 6

USING DIFFUSE ATTENUATION COEFFICIENTS TO DERIVE DEPTH-RELATED OPTICAL PROPERTIES FROM HYPERSPECTRAL IMAGES

Abstract

Remote sensing products applied in aquatic environments represent the interaction of electromagnetic radiation with the optically active components along the euphotic zone of the water column. Understanding the light attenuation processes along the water column can be useful to generate and improve bio-optics models. In this work we present a semi-analytical model to derive depth-related reflectance using the above water reflectance integrated with downwelling and upwelling vertical diffuse attenuation coefficients. We use the above water reflectance to estimate the optically active properties, such as reflectance, from different depths. The results of more then 83% and 0.02 for R^2 and RMSE, respectively, show a good performance of the model. The performance of the novel model, based on semi-analytical and empirical steps, can be applied to understand the electromagnetic energy behavior on underwater light field.

6.1 Introduction

The remote sensing techniques are successfully applied to study water bodies. The electromagnetic waves interact with components present on water, along the path in the water column to the depth which the light is able to react. The euphotic zone is directly dependent on the wavelength (GORDON; MCCLUNEY, 1975). Remote sensing images represent the characteristics of the water column, related to the entire euphotic zone.

The underwater light field is dependent on the environmental characteristics and the light penetration depth is lower in turbid waters than in clear water (MCCLUNEY, 1974). However, the values taken from remote sensing images are processed in accordance with a plane and associated to the respective point or area of the surface. Different interactions along the water column can not be precisely decomposed and segmented in depth because the inversion of bio-optical models are usually adjusted from surface samplings (SHANMUGAM et al., 2010).

Since the electromagnetic radiation interacts in a variable space, because the depth of light penetration depends on the specific energy of each spectral range, and dynamic, due to the variability of the optically significant components (OSC), the response captured by the sensors does not only represent the value associated with the surface, and calibration of models considering only surface sampling data may not provide effective results. Therefore, knowing that the imagery sensors acquire data not only from the surface but from the euphotic zone of the water column, is it possible to estimate the underwater light field behavior to retrieve depth-related optical properties from hyperspectral image data?

This work proposes a model combining empirical and semi-analytical approaches to use the vertical diffuse attenuation coefficients to derive depth-related reflectance. Our main objective was to estimate the reflectance along of water column using reflectance data from hyperspectral images and vertical diffuse attenuation coefficient obtained from in-situ profile data or estimated from bio-optical models. The remote sensing reflectance is dependent just on the inherent optical properties (IOP) and can be calculated at different depths (MOBLEY et al., 2005).

We use the dataset acquired in Capivara reservoir, Chapter 2. The proposed model uses three measurements to estimate the depth-related reflectance: downwelling (E_d) and upwelling (E_u) vertical diffuse attenuation coefficients and the reflectance above water surface. Using the proposed model, we are able to express the reflectance spatially and along the water column, resulting in a three-dimensional view of the optical property.

6.2 Methods

6.2.1 Deriving depth-related reflectance

There are two main reflectance data usually applied on water body remote sensing: Irradiances reflectance (R) and Remote Sensing Reflectance (R_{rs}). Both reflectances are described from a ratio of the amount of energy available on the water from the sun, but R is based on the upwelling irradiance and the R_{rs} on the upwelling radiance, as we can see on its equations:

$$R = \frac{E_u}{E_d}(\%) \qquad (6.1) \qquad R_{rs} = \frac{L_u}{E_d}(sr^{-1}) \qquad (6.2)$$

The basic difference between R and R_{rs} is the field of view used to capture the energy signal from the water (SCHAEPMAN-STRUB et al., 2006). We need a cosine collector to capture the hemispherical light used to compute R(MORROW MICHAEL S. DUHIG, 1994), but, on the other hands, a radiance sensor is able to detect the light in a single solid angle to produce R_{rs} (MOBLEY, 1999). Because of the specific angle used to calculate R_{rs}, this unity is less influenced by light conditions changes in comparison to R (MOREL; GENTILI, 1991; MOREL; GENTILI, 1996).

If we acquire data on the same environmental characteristics, the reflectances can be related using the anisotropy factor of light field distribution (Q), defined by the ratio of E_u/E_d (MOREL; GENTILI, 1996). Using Q, is possible to express the reflectance relationship from:

$$R = QR_{rs} \rightarrow R_{rs} = \frac{R}{Q} \qquad (6.3)$$

The vertical diffuse attenuation coefficient, also named as K-functions, express how the light field change along water column (KIRK, 1986). All subsurface signals can be expressed using K-functions (LEE et al., 1998). In this work, the vertical diffuse attenuation of upwelling (K_u) and downwelling (K_d) irradiances play a important role to derive depth-related data. Using K_d and K_u, we can estimate the light along water column (z) from the subsurface value (0_-):

$$E_d(z) = E_d(0_-)e^{(-zK_d)} \qquad (6.4) \qquad E_u(z) = E_u(0_-)e^{(-zK_u)} \qquad (6.5)$$

From the K_d and K_u relationships, we can rewrite the irradiance reflectance equation to express the depth-related reflectance from the subsurface signal:

$$R(z) = \frac{E_u(z)}{E_d(z)} = \frac{E_u(0_-)e^{(-zK_u)}}{E_d(0_-)e^{(-zK_d)}} = R(0_-)\frac{e^{(-zK_u)}}{e^{(-zK_d)}} = R(0_-)e^{z(K_d-K_u)} \tag{6.6}$$

Nevertheless, irradiance reflectance can be influenced by environmental light changes, and because of this, the remote sensing reflectance is usually adopted on hydrological remote sensing (MOBLEY et al., 2005). Using the relationship 6.3 we can rewrite the final model:

$$R_{rs}(z) = \frac{R(z)}{Q} = \frac{R(0_-)e^{z(K_d-K_u)}}{Q} = \frac{QR_{rs}(0_-)e^{z(K_d-K_u)}}{Q} = R_{rs}(0_-)e^{z(K_d-K_u)} \tag{6.7}$$

Using the R_{rs} also is possible to apply several analytical, semi-analytical and empirical models discussed by others works. Imagery sensors data is able to generate R_{rs} above water surface (0_+), but a just-below water surface remote sensing reflectance (0_-) can be estimated using (LEE; CARDER; ARNONE, 2002):

$$R_{rs}(0_-) = \frac{R_{rs}(0_+)}{0.52 + 1.7R_{rs}(0_+)} \tag{6.8}$$

6.2.2 Processing workflow

At each sample point, we used the TriOS sensor to measure the total radiance of water(L_t), sky radiance (L_s) and a reference Lambertian panel (L_g). Downwelling irradiance (E_u) was collected using another TriOS sensor. We obtained the remote sensing reflectance (R_rs) using the equation 6.9. We adopt a fixed rho equals to 0.028 to extract the direct reflection of the water surface from the total radiance, but we know the rho value is spectral-dependent (MOBLEY, 1999).

$$R_{rs} = \frac{L_t - \rho L_s}{\pi L_g}(sr^{-1}) \tag{6.9}$$

We use the results from the works discussed on the previous Chapter. Using the hyperspectral corrected mosaic(Chapter 4) transformed in reflectance values using ELM(Chapter 5), we process and compare with the corrected in-situ dataset(Chapter 3). We present the syntheses of the method steps on the Figure 29. The workflow shows all interaction to process the variables.

The inputs of the method are the hyperspectral images captured by the Rikola camera, the radiance values (total - L_t, sky - L_s and reference panel - L_g) for each river

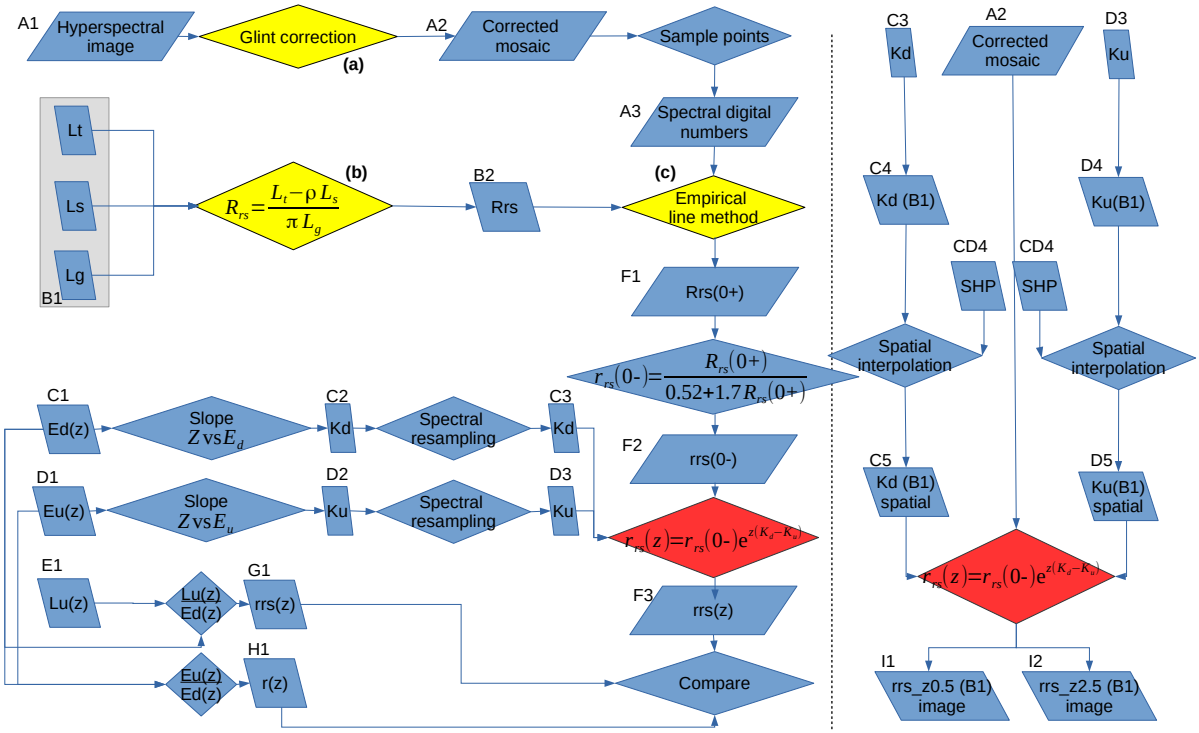


Figure 29 – Workflow to processing the measurements. Steps analyzed in previous works are highlighted with yellow color: Approach to correct glint and mosaic the hyperspectral images (a); Processing and standardization using the Aquopts (b); and the empirical line calibration (c). The proposed model is in red color.

sample point, the profile of upwelling radiance (L_u) and irradiance (E_u) also well of downwelling irradiance (E_d). All other parameters are derived from these inputs, according to the method described in Figure 29. There are two groups of steps on the method: the first one use only in-situ data and the second is about processing the images to apply the calculated parameters from in-situ data. All steps are identified with a label to easy reading the workflow.

6.3 Results and Discussion

We process all data based on the workflow presented in the Figure 29. Some preprocessing are required to calculate the input attributes of the proposed method. The first steps are used to obtain R_{rs} above the water surface, calculated using the radiance acquired by the sensor above the water surface, in step (b) of the method, and the empirical line generated with the digital numbers obtained in the image, step A3. At the end of this first stage, in step F1 is produced $Rs(0+)$.

The proposed model uses the measure of r_{rs} just below the water surface. We have applied the method proposed by (LEE; CARDER; ARNONE, 2002) to introduce into the water the measure of R_{rs} taken above the surface. Figure 30 displays the curves of r_{rs} just below the water surface, after resampling of the hyperspectral channels of the sensors

according to the bands of Rikola camera.

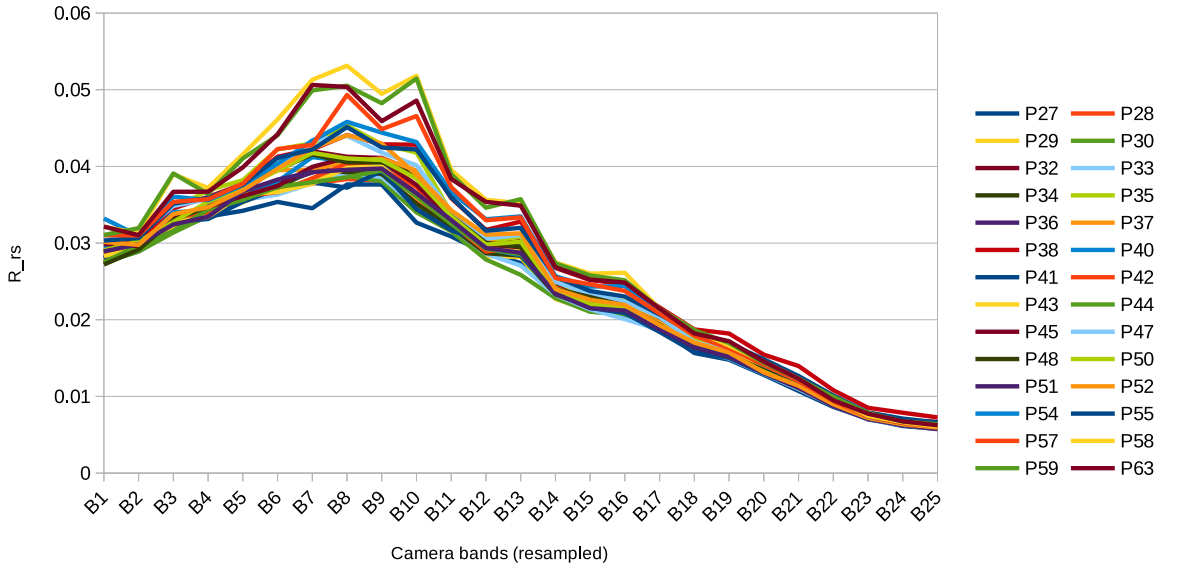


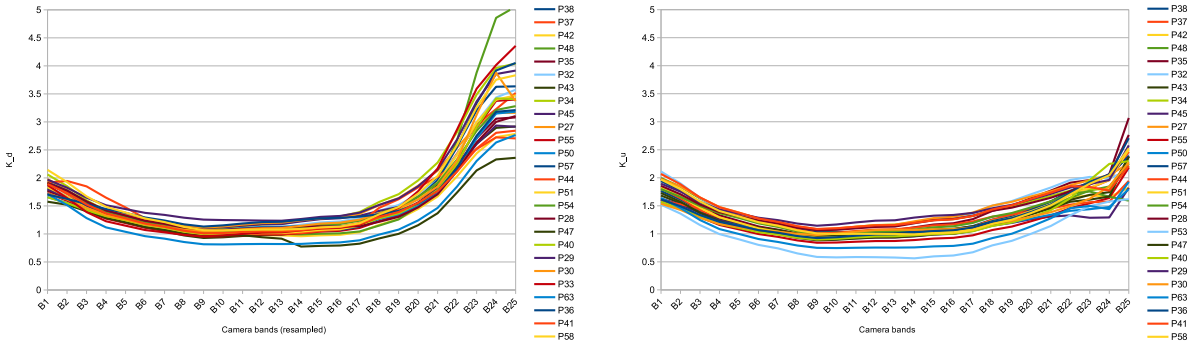
Figure 30 – Remote Sensing Reflectance just bellow water surface for each point resampled from hyperspectral profile sensor

The spectral curves of remote sensing reflectance indicate a small variation of optically significant components on the reservoir. All the curves present the same spectral behavior and anyone highlight distinct features, representing great homogeneity of the optical properties. The largest differences in spectral behavior are found with a peak close to band 8, which registers the range of 570 nm, in the green region, indicating interaction with chlorophyll.

The proposed model uses the diffuse vertical attenuation information to estimate the reflectance values along the water column. The values of K_d and K_u were calculated by fitting an exponential curve on the E_u and E_d data. The adjustment was made from sensor measurements along the water column with variable acquisition frequency rate according to the light availability and the displacement speed to submerge the sensors. With this methodology, the mean gap between each measurement was 0.28 meters with the sensors submerged until reaching 10% of the measured energy obtained on the surface.

After fitting the K_d and K_u in step C2 and D2 for each sampling point using the hyperspectral sensors data, the curves were resampled according to the 25 bands of camera. The result of K_d and K_u resampled in steps C3 and D3 can be seen in Figure 31.

The values of K_d and K_u highlight the same homogeneous behavior among the collected points, because the spectral values are closer. There is some variation in the curves in the last wavelengths, about to the red region captured by bands B18 and B19. The increase of the slope in the near infrared wavelengths, B20-B25 indicates the strong attenuation of the water in this region, consequently some curves presented a random



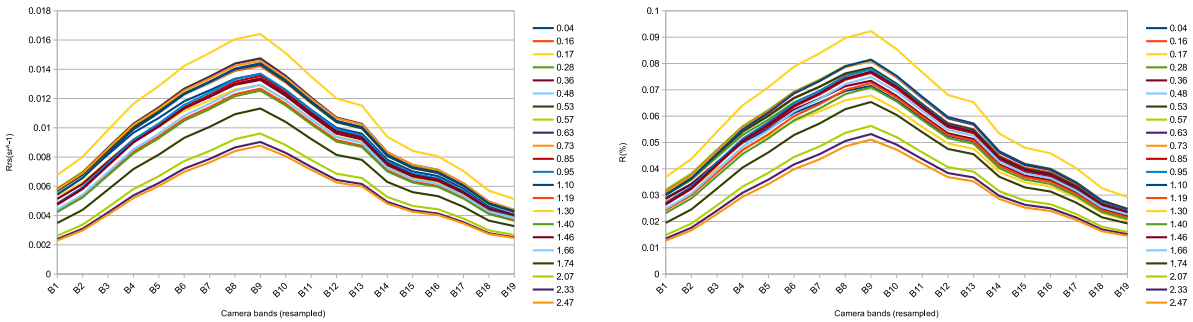
(a) Resampled K_d (b) Resampled K_u

Figure 31 – Vertical diffuse attenuation coefficient of downwelling and upwelling irradiance, K_d and K_u, respectively, resampled from the hyperspectral profile sensor

behavior in that region, that may have been influenced by noises due to the resulting small energy. The high absorption of water from the near-infrared wavelengths resulting in high noise amplified proportionally as the depth increases, makes it impossible to use these wavelengths in the proposed model. For this reason, we discard the bands B20-B25.

The same values from E_u and E_d were applied to calculate the reflectance bellow water in different depths. We used the ratio of L_u and E_u over E_d to obtain two types of reflectance. The first reflectance was calculated from the ratio between upwelling (E_u) and downwelling irradiance (E_d), resulting in the dimensionless reflectance (R). The other reflectance was obtained by using the upwelling radiance (L_u) at a specific solid angle (sr) over the respective downwelling irradiance.

In Figure 32, we can note the behavior of both reflectances along the water column at point P51. Point P51 was chosen to represent the dataset because it is the closest to the average curve of K_d, and considering the homogeneity of K_d and K_u, all points show a similar behavior of the light field within the water.



(a) Resampled R_{rs} (b) Resampled R

Figure 32 – Depth related reflectances calculated from the profile sensor measurements in point P51

By analyzing the reflectance curves shown in Figure 32 we can see that both

reflectances have the same spectral signature varying only the amplitude of the values. The relationship between the two reflectances is described by the anisotropy factor (Q). Anisotropy is defined by the optical response according to the angular variation and represented by the ratio between a hemispherical measurement, in this case E_u, over the angular radiance, in this case L_u. As can be seen in Figure 33, the Q factor approaches 2π.

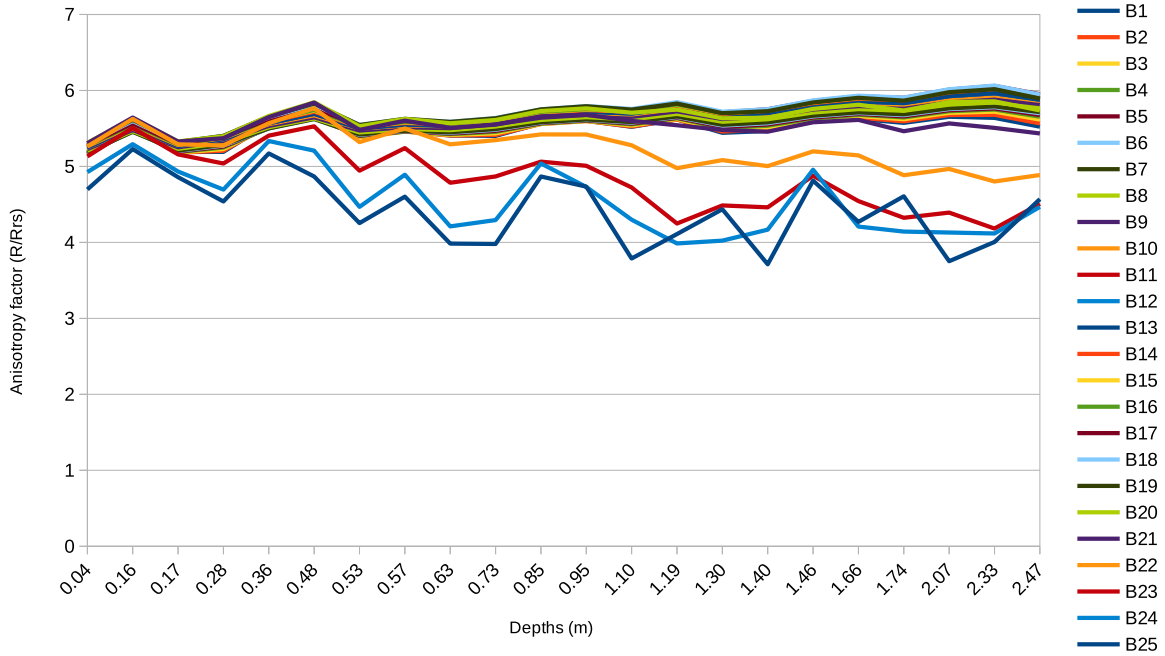


Figure 33 – Anisotropy factor obtained from the ratio of R over R_{rs} in depths of point P51 from the hyperspectral profile sensor

The behavior of the Q factor highlights the strong noise in the near-infrared bands. By analyzing the values of Q, we see that all the bands have the same spectral behavior, however, the noise in the near-infrared bands changes significantly the behavior by inserting random noise. The strong noise demonstrated by the Q factor emphasizes the need to discard the near-infrared bands.

The values of K_d, K_u and $r_{rs}(0-)$ are the inputs of the proposed model to estimate the reflectance value along the water column. For a given depth value at a specific location, the model is applied considering its input parameters. Figure 34 shows the graph with the results of model applied considering a depth variation of 0.5 to 2.5 meters at point P51.

The wavelengths close to red, B19, starts an exponential increase trend. Such behavior occurred due to the tendency of increase in the red wavelength present in the three input parameters of the model, K_d, K_u e $r_{rs}(0-)$. The combination of the increase present in the three parameters amplified exponentially the resulting value.

The shape of the curves is consistent in comparison with the reflectance calculated from the submerged sensors, Figure 32. There is a high correlation between the depth

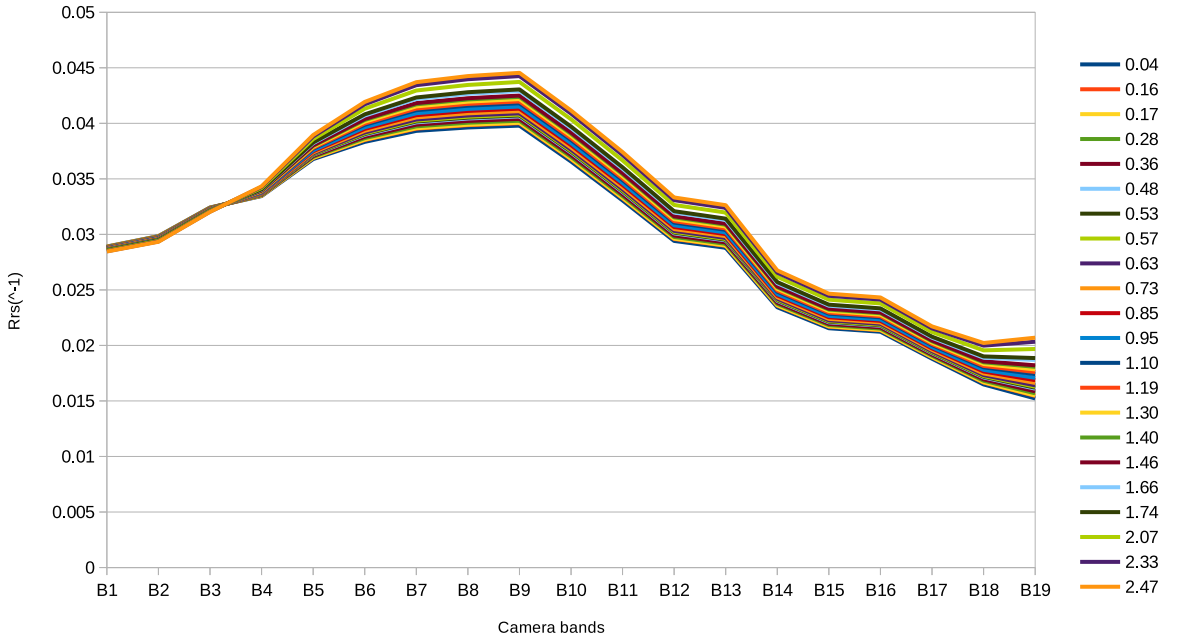


Figure 34 – Depth related reflectances calculated using the proposed model over the surface profile sensor measurements, K_d and K_u in point P51

reflectance estimated by the model against the depth reflectances calculated from the data measured in-situ, as can be seen in Figure 35.

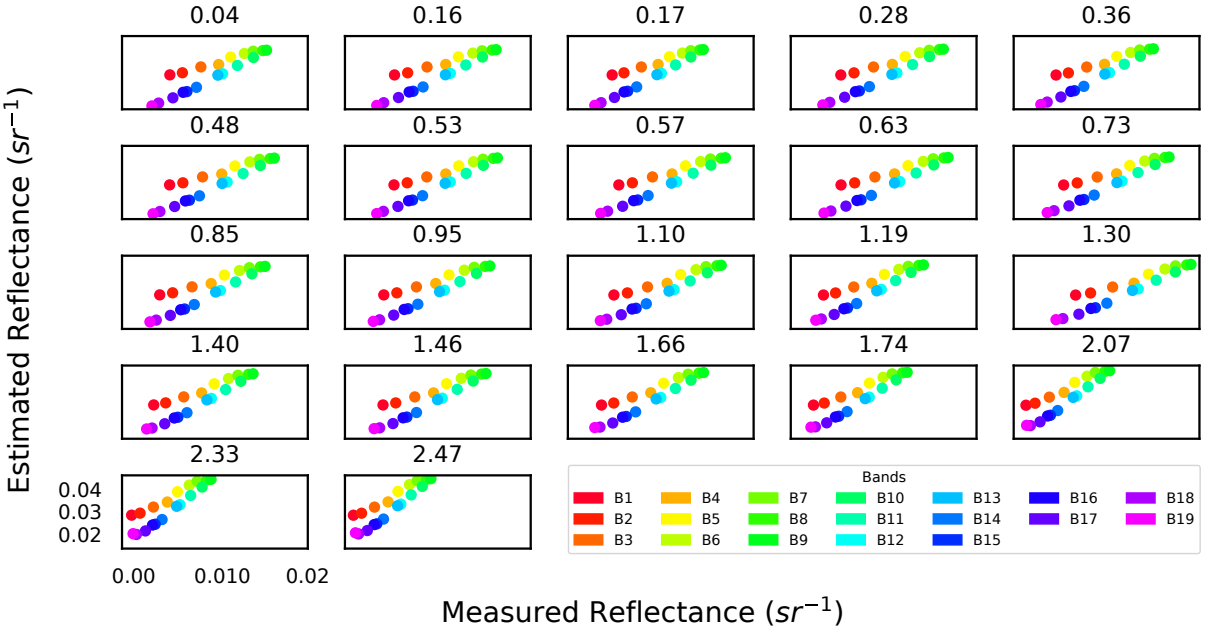


Figure 35 – Spectral correlation of $R_{rs}(z)$ with the values calculated from the model and in-situ measurements in point P51 with variant depth

Looking at the scatter-plots shown in Figure 35 it is possible to verify a good matching that occurs at all depths. However, the amplitude of the values needs to be readjusted since the reflectances calculated by the model presented a smaller interval with respect to the calculated reflectances from the in-situ measurements. The good behavior

of the model can be seen from the statistics presented in Table 4, in which the model obtained more than 83% and 0.02 of R^2 and RMSE, respectively.

Table 4 – Statistics about the error of the depth-related reflectances from model in comparison to depth-related reflectance from in-situ measurements

Statistics	R2	RMSE (sr^{-1})
Minimum	81.385%	0.0208
Maximum	89.223%	0.0280
Average	83.703%	0.0230
Standard Deviation	1.771%	0.0022

The depth reflectance curve calculated by the sensor measurements is directly dependent on the sensibility of the sensor to capture the energy required to perform the measurement. The availability of light decrease exponentially along the water column and in higher depths the sensor is not able to get the small amount of energy. The exponential decay is present both in the measurements of K_d and K_u, so the combination of the two values in the model cancel the effect of the sensor sensibility to the presence of light. No matter how small the light availability, the equation will still return some result different of zero.

To apply the proposed model in the image, the input parameters must be spatially interpolated. The spatial reflectance can be directly obtained from the image by the application of the empirical line method in the entire image. However, to obtain spatially values of K_d and K_u, the measurements need to be spatially interpolated from the in-situ measurements or obtained directly from the image. To maintain compatibility with the data collected in the field and allow comparisons, the K_d and K_u values was spatially interpolated. Due to the homogeneity of the data, we used the spatial interpolation with the algorithm inverse distance weight (IDW). In the Figure 36 we can see the result of the interpolation of K_d and K_u for the first band of the camera, which is the closest to the 490 wavelength commonly used in bio-optical models to estimate the k-functions.

Once all input parameters have been spatialized, the proposed model can be applied to each pixel of the hyperspectral mosaic. The model was applied in all bands, but for better visualization, in Figure 37, the depth-related reflectances results of the estimated depth of 0.5 meters and 2.5 meters for the first camera band are presented.

Analyzing the Figure 37, we can see the spatial variability along the water column tends to increase proportionally to depth. The amplitude of the values increased at both extremes, altering the largest and smallest values found closest to the surface. Such behavior is associated to the combination of K_d and K_u values considered in the model as constants along the water column. Since both determine the exponential decay of the available energy and its difference is corresponding to the exponent of the model, when

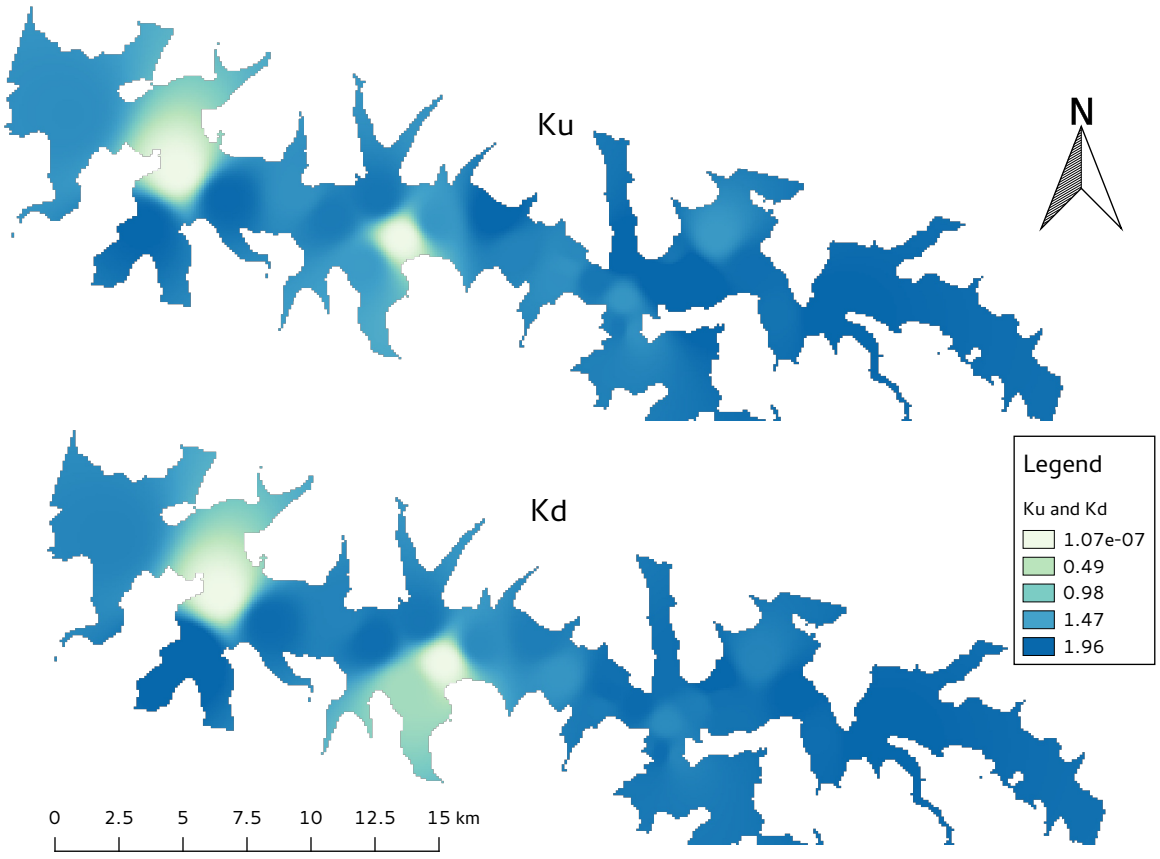


Figure 36 – Spatially interpolated K_d and K_u in the band B1

increase the difference of K_d and K_u values along the water column, consequently the exponent increases too.

The proposed model is based on an exponential equation, in which all the parameters present positive behavior. For this reason, as the depth increases, the resulting reflectance will also increase if it maintains positive behavior. However, the main influencing factor in the model performance is the precision of the K_d and K_u values, because when K_u is greater than K_d, then the model will present a decreasing exponential curve, as expected. However, in cases where K_d is greater than K_u, the result will be an increasing exponential curve, as shown in Figure 37.

6.4 Conclusions

Studies about the light field below the water surface are critical for the development and refinement of new bio-optical models. The interactions of electromagnetic energy along the water column is a rich information source that can be used for optical characterization of complex aquatic systems. This work presented a model for calculating reflectance along the water column using a remote sensing image and estimated diffuse vertical attenuation

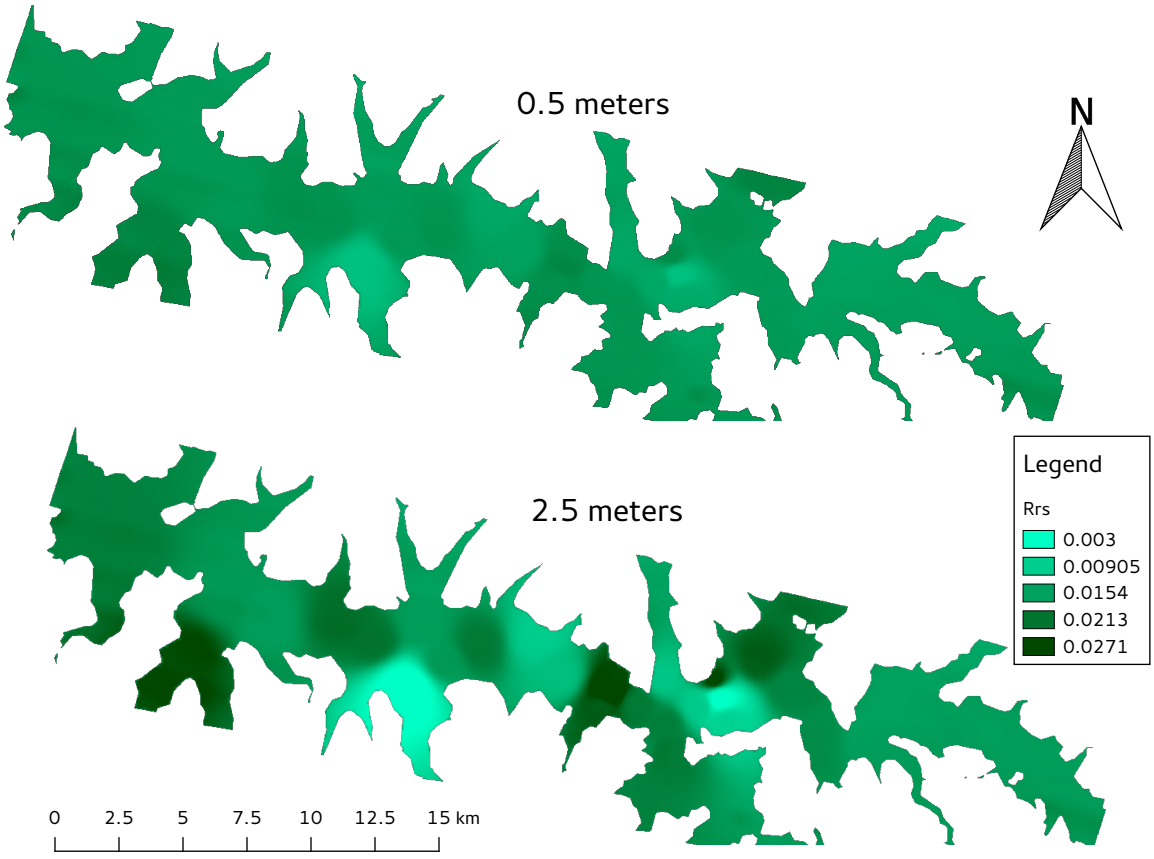


Figure 37 – Bellow water reflectance on 0.5 and 2.5 meters of depth for the band B1

coefficients.

The results demonstrate a potential of application of the proposed model since the estimated reflectance curves were shown with compatible shapes in relation to the curves obtained through the profile of the water column with sensors. Additionally, the coefficient of determination about 83% show a good matching between the measured and modeled depth-related reflectance, but there is a factor to corrected the range of values. The fragility of the model was the estimation of K_d and K_u, because the proposed model directly depends on these parameters. More accurate methods for estimating the vertical attenuation coefficients, aiming to avoid the use of interpolations, are essential to improve the final performance of the model.

The challenge of estimating depth-related optical properties using remote sensing images is justified to amplify the ability to apply analytical and empirical models along the water column, in addition to those methods currently directed to the surface.

CHAPTER 7

CONCLUSIONS AND FINAL REMARKS

The behavior of the light field below the water surface is a significant challenge in understanding the interaction of electromagnetic energy along the water column. This work presented analysis and contributions to all processes, from data acquisition, processing, integration, transformation and modeling. The topics covered by each chapter, unified in individual papers, were directed to compose the research minimizing the overlapping of the chapters, but highlighting the synergy among the subjects by the serial and gradual integration of each chapter.

The Geosciences area, and several other contexts dependent of a field campaign using different methods of data collection based on sensors. However, integrating data from different sources is not a trivial task since each manufacturer produces specific tools to manipulate the data of the respective sensor. The solution proposed in this work solves the problem of integrating the data of five sensors frequently applied to field work in aquatic environments. The system including data management and automatic information processing was efficient and fundamental for the reporting of this work, and other impact studies of the group. We automated all manual process performed after each field survey using the system for data standardization, automatic processing and quality control.

Remote sensing images of water bodies are often affected by noises and specular reflection. Using images with high spatial resolution, such as that used in this work, amplifies the occurrence of glint because increase the probability of acquired variations from the surface caused by waves. The original images achieved with the field survey had a high noise level, mainly on the border of the images. We evaluated several approaches, and the proposed model based on the multi-scale reference algorithm showed to be more efficient than conventional methods. The proposed algorithm does not produce any new

value or data because it searches for the best candidate pixel with the minor influence of noise. However, this approach considers that the overlap between images is sufficient to guarantee the existence of at least one pixel not affected by noise, limiting its applications in previously mosaic images that do not present redundancy in the values of each pixel.

After correct the images and eliminate or reduce noises, the process of radiometric transformation of the digital number values into physically meaning quantities can begin. Images acquired from aerial surveys require field work to support further processing, justifying the application of the empirical line method in this scenario for a local calibration. We evaluated several targets and samples settings for empirical line calibration. The results showed the difficulty in applying this method to recover information about the water body, due to the intense absorption through the water. However, the adaptation of the method to including a sample point measured in water along with another artificial target produced better results.

Finally, after all the processing to integrate and correct of the information sources, in the last chapter we present a proposal for a semi-analytical model to determine the measures of reflectance along the water column from the acquired images. We constructed the model from physical and mathematical relations known in the literature. The proposed model uses the data of the vertical attenuation coefficients of the light, both ascending and descending, to determine the variation of light at any depth and to estimate the optical property. The data of the vertical attenuation coefficients of the light have a great influence on the model, and its precision and accuracy is essential for the good performance of the model. Unexpected values of K_d and K_u will produce uncertain results of the model.

In all the stages of this work, we have an effort to keep the data, algorithms, scripts, and system publicly accessible to the whole community. The discovery and access to the data collected in the field is a current demand, considering the large amount of data collected, observing different attributes, following different methods, using diverse tools and stored in multiple data formats. Given this scenario, methodologies for standardization and definition of the data infrastructure are necessary to promote the interoperability of the data, especially regarding data sharing.

As future possibilities to expand the subjects addressed in this work, we can cite recommendations on:

- Test and validate the models using another set of images acquired in another aquatic system with different characteristics of composition;
- Investigate visual analytics techniques to include on the system to amplify visual and interactive exploration of data from multiple sources;
- Analyze the use of parallel processing on the glint mitigation algorithm to optimize

the execution of large scenes;

- Check the performance of the empirical line using a gradient gray scale of artificial targets to find the one that best fits in this and another dataset;
- Use simulations based on the radiative transfer function, such as Hydrolight, to validate and verify the behavior of the proposed model to estimate depth-related reflectance in an environment with variant composition;
- Verify the application of models to calculate the vertical attenuation coefficients from image to replace the use of spatial interpolation of these values;
- Investigate the relation of the Secchi disk with the optical properties related to each depth. This relationship can represent a factor to be included on the proposed model.

We hope the results obtained and showed in this work can provide subsidies, applicable both to the area of Geosciences and Computer Science, to amplify the investigations and discussions to understand the light field below the water surface mainly to the recovery depth-related information using remote sensing images as an input source.

BIBLIOGRAPHY

AASEN, H. et al. Quantitative remote sensing at ultra-high resolution with uav spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows. *Remote Sensing*, v. 10, n. 7, 2018. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/10/7/1091>. Cited on page 24.

AGOSTINHO, A. A.; PELICICE, F. M.; GOMES, L. C. Dams and the fish fauna of the neotropical region: impacts and management related to diversity and fisheries. *Brazilian Journal of Biology*, SciELO Brasil, v. 68, n. 4, p. 1119–1132, 2008. Cited on page 18.

ALCâNTARA, E. et al. Field measurements of the backscattering coefficient in a cascading reservoir system: first results from nova avanhandava and barra bonita reservoirs (são paulo, brazil). *Remote Sensing Letters*, Taylor & Francis, v. 7, n. 5, p. 417–426, 2016. Disponível em: <https://doi.org/10.1080/2150704X.2016.1145361>. Cited 5 times on pages 11, 29, 39, 41, and 42.

ANEEL, A. N. d. E. E. *Capacidade de Geração do Brasil*. 2019. Disponível em: <http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm>. Cited on page 17.

BABIN, M. et al. Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around europe. *Journal of Geophysical Research: Oceans*, v. 108, n. C7, 2003. Disponível em: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JC000882>. Cited on page 30.

BENNEMANN, S. T. et al. Dinâmica trófica de plagioscion squamosissimus (perciformes, sciaenidae) em trechos de influência da represa capivara (rios paranapanema e tibagi). *Iheringia, Série Zoologia*, SciELO Brasil, v. 96, n. 1, p. 115–119, 2006. Cited on page 18.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data - the story so far. 2009. Disponível em: <https://eprints.soton.ac.uk/271285/>. Cited on page 32.

BOSS, E.; PEGAU, W. S. Relationship of light scattering at an angle in the backward direction to the backscattering coefficient. *Appl. Opt.*, OSA, v. 40, n. 30, p. 5503–5507, Oct 2001. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-40-30-5503>. Cited 2 times on pages 36 and 37.

BOULOS, M. N. K. et al. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, ogc standards and application examples. *International Journal of Health Geographics*, v. 10, n. 1, p. 67, Dec 2011. ISSN 1476-072X. Disponível em: <https://doi.org/10.1186/1476-072X-10-67>. Cited 2 times on pages 31 and 32.

BRACCHINI, L. et al. Spatial and temporal variations of the inherent and apparent optical properties in a shallow coastal lake. *Journal of Photochemistry and Photobiology B: Biology*, v. 80, n. 3, p. 161 – 177, 2005. ISSN 1011-1344. Disponível em: <http://www.sciencedirect.com/science/article/pii/S101113440500103X>. Cited on page 28.

BUKATA, R. P. et al. *Optical properties and remote sensing of inland and coastal waters*. [S.l.]: CRC press, 2018. Cited on page 28.

BURT, P. J.; ADELSON, E. H. A multiresolution spline with application to image mosaics. *ACM Trans. Graph.*, ACM, New York, NY, USA, v. 2, n. 4, p. 217–236, out. 1983. ISSN 0730-0301. Disponível em: <http://doi.acm.org/10.1145/245.247>. Cited on page 49.

CARMO, A. et al. *Correction workflow and spatial database model of Aquopts - A Hydrological Optical Data Processing System*. 2019. Disponível em: <http://dx.doi.org/10.17632/f2tz548v2c.1>. Cited 3 times on pages 29, 33, and 44.

CARMO, A. F. C. do. *carmoafc/aquopts: First Version - eng-Interface - pt-SourceCode*. 2019. Disponível em: <https://doi.org/10.5281/zenodo.2605523>. Cited on page 29.

CARMO, A. F. C. do et al. *Python algorithms and dataset of empirical line method applied to inland water hyperspectral images combining reference targets and in situ water measurements*. IEEE Dataport, 2019. Disponível em: <http://dx.doi.org/10.21227/fbks-rj40>. Cited on page 65.

CARMO, A. F. C. do et al. *Processed airborne hyperspectral images and mosaics of a Paranapanema River region in Capivara reservoir, Brazil*. IEEE Dataport, 2019. Disponível em: <http://dx.doi.org/10.21227/dxgz-3m95>. Cited on page 56.

CARVALHO, L. A. S. de et al. Implications of scatter corrections for absorption measurements on optical closure of amazon floodplain lakes using the spectral absorption and attenuation meter (ac-s-wetlabs). *Remote Sensing of Environment*, v. 157, p. 123 – 137, 2015. ISSN 0034-4257. Special Issue: Remote Sensing of Inland Waters. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425714002387>. Cited on page 35.

CONEL, J. E. et al. Ais-2 radiometry and a comparison of methods for the recovery of ground reflectance. 1987. Cited on page 58.

COX, C.; MUNK, W. Measurement of the roughness of the sea surface from photographs of the sun's glitter. *Josa*, Optical Society of America, v. 44, n. 11, p. 838–850, 1954. Cited 2 times on pages 46 and 47.

COX, C.; MUNK, W. Slopes of the sea surface deduced from photographs of sun glitter. 1956. Cited 2 times on pages 46 and 47.

CURTARELLI, M. et al. Coupling remote sensing bio-optical and three-dimensional hydrodynamic modeling to study the phytoplankton dynamics in a tropical hydroelectric reservoir. *Remote Sensing of Environment*, v. 157, p. 185 – 198, 2015. ISSN 0034-4257. Special Issue: Remote Sensing of Inland Waters. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425714002260>. Cited on page 30.

DALL'OLMO, G.; GITELSON, A. A. Effect of bio-optical parameter variability on the remote estimation of chlorophyll-a concentration in turbid productive waters: experimental results. *Appl. Opt.*, OSA, v. 44, n. 3, p. 412–422, Jan 2005. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-44-3-412>. Cited on page 29.

Di, L.; Moe, K.; van Zyl, T. L. Earth observation sensor web: An overview. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 3, n. 4, p. 415–417, Dec 2010. ISSN 1939-1404. Cited 2 times on pages 31 and 32.

DOERFFER, R. et al. The impact of sun glint on the retrieval of water parameters and possibilities for the correction of meris scenes. In: CITESEER. *Proceedings of the 2nd MERIS/(A) ATSR User Workshop, ESA SP-666,(ESA, Frascati, Italy, 2008)*. [S.l.], 2008. Cited on page 46.

DOXARAN, D.; CHERUKURU, N.; LAVENDER, S. J. Apparent and inherent optical properties of turbid estuarine waters: measurements, empirical quantification relationships, and modeling. *Appl. Opt.*, OSA, v. 45, n. 10, p. 2310–2324, Apr 2006. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-45-10-2310>. Cited on page 28.

FELISBERTO, S. A.; RODRIGUES, L. Periphytic community of reservoirs cascade in the paranapanema river, brazil. *Acta Scientiarum*, v. 27, n. 3, p. 215–223, 2005. Cited on page 21.

GAO, B.-C. et al. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment*, v. 113, p. S17 – S24, 2009. ISSN 0034-4257. Imaging Spectroscopy Special Issue. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425709000741>. Cited on page 58.

GLASGOW, H. B. et al. Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. *Journal of Experimental Marine Biology and Ecology*, v. 300, n. 1, p. 409 – 448, 2004. ISSN 0022-0981. VOLUME 300 Special Issue. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022098104001066>. Cited on page 31.

GORDON, H. R.; MCCLUNEY, W. R. Estimation of the depth of sunlight penetration in the sea for remote sensing. *Appl. Opt.*, OSA, v. 14, n. 2, p. 413–416, Feb 1975. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-14-2-413>. Cited on page 67.

GROETSCH, P. M. M. et al. Validation of a spectral correction procedure for sun and sky reflections in above-water reflectance measurements. *Opt. Express*, OSA, v. 25, n. 16, p. A742–A761, Aug 2017. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-25-16-A742>. Cited on page 46.

HADJIMITSIS, D.; CLAYTON, C.; RETALIS, A. The use of selected pseudo-invariant targets for the application of atmospheric correction in multi-temporal studies using satellite remotely sensed imagery. *International Journal of Applied Earth Observation and Geoinformation*, v. 11, n. 3, p. 192 – 200, 2009. ISSN 0303-2434. Disponível em: <http://www.sciencedirect.com/science/article/pii/S030324340900004X>. Cited on page 58.

HEDLEY, J. D.; HARBORNE, A. R.; MUMBY, P. J. Technical note: Simple and robust removal of sun glint for mapping shallow water benthos. *International Journal of Remote Sensing*, Taylor & Francis, v. 26, n. 10, p. 2107–2112, 2005. Disponível em: <https://doi.org/10.1080/01431160500034086>. Cited on page 46.

HOCHBERG, E. J.; ANDREFOUET, S.; TYLER, M. R. Sea surface correction of high spatial resolution ikonos images to improve bottom mapping in near-shore environments. *IEEE Transactions on Geoscience and Remote Sensing*, v. 41, n. 7, p. 1724–1729, July 2003. ISSN 0196-2892. Cited on page 46.

HOFFMANN, A. C.; ORSI, M. L.; SHIBATTA, O. A. Diversidade de peixes do reservatório da uhe escola engenharia mackenzie (capivara), rio paranapanema, bacia do alto rio paraná, brasil, ea importância dos grandes tributários na sua manutenção. *Iheringia. Série Zoologia*, SciELO Brasil, v. 95, n. 3, p. 319–325, 2005. Cited on page 18.

HONKAVAARA, E. et al. A process for radiometric correction of uav image blocks. *PFG Photogrammetrie, Fernerkundung, Geoinformation*, Schweizerbart Science Publishers, Stuttgart, Germany, v. 2012, n. 2, p. 115–127, 05 2012. Disponível em: <http://dx.doi.org/10.1127/1432-8364/2012/0106>. Cited on page 24.

Honkavaara, E. et al. Hyperspectral Reflectance Signatures and Point Clouds for Precision Agriculture by Light Weight Uav Imaging System. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, p. 353–358, jul. 2012. Cited on page 24.

HONKAVAARA, E. et al. Band registration of tuneable frame format hyperspectral uav imagers in complex scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 134, p. 96 – 109, 2017. ISSN 0924-2716. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0924271616305378>. Cited on page 24.

HONKAVAARA, E. et al. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight uav spectral camera for precision agriculture. *Remote Sensing*, v. 5, n. 10, p. 5006–5039, 2013. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/5/10/5006>. Cited on page 24.

HU BRIAN B. BARNES, B. M. P. R. C. C. Satellite-based virtual buoy system to monitor coastal water quality. *Optical Engineering*, v. 53, n. 5, p. 1 – 11 – 11, 2013. Disponível em: <https://doi.org/10.1117/1.OE.53.5.051402>. Cited on page 28.

HYDRO-OPTICS, B. H.; LABORATORIES, I. *HydroScat-6P Spectral Backscattering Sensor and Fluorometer User Manual*. [S.l.], 2010. Cited on page 36.

IMAI, N. N. et al. Bio-optical data integration based on a 4d database system approach. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3, p. 635–641, 2015. Disponível em: <https:

//www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-7-W3/635/2015/>. Cited 2 times on pages 28 and 30.

JAT, M. K.; GARG, P.; KHARE, D. Monitoring and modelling of urban sprawl using remote sensing and gis techniques. *International Journal of Applied Earth Observation and Geoinformation*, v. 10, n. 1, p. 26 – 43, 2008. ISSN 0303-2434. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0303243407000335>. Cited on page 51.

KARPOUZLI, E.; MALTHUS, T. The empirical line method for the atmospheric correction of ikonos imagery. *International Journal of Remote Sensing*, Taylor & Francis, v. 24, n. 5, p. 1143–1150, 2003. Disponível em: <https://doi.org/10.1080/0143116021000026779>. Cited on page 58.

KAY, S.; HEDLEY, J. D.; LAVENDER, S. Sun glint correction of high and low spatial resolution images of aquatic scenes: a review of methods for visible and near-infrared wavelengths. *Remote Sensing*, v. 1, n. 4, p. 697–730, 2009. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/1/4/697>. Cited on page 46.

KENWORTHY, S. T.; RHOADS, B. L. Hydrologic control of spatial patterns of suspended sediment concentration at a stream confluence. *Journal of Hydrology*, Elsevier, v. 168, n. 1-4, p. 251–263, 1995. Cited on page 18.

KIM JOONG YONG PARK, G. T. M. A constrained optimization technique for estimating environmental parameters from czmil hyperspectral and lidar data. In: . [s.n.], 2010. v. 7695. Disponível em: <https://doi.org/10.1117/12.851989>. Cited on page 28.

KIRK, J. Optical limnology—a manifesto. In: *Limnology in Australia*. [S.l.]: Springer, 1986. p. 33–62. Cited on page 68.

KIRK, J. T. *Light and photosynthesis in aquatic ecosystems*. [S.l.]: Cambridge university press, 1994. Cited on page 28.

KORNIJóW, R. Controversies around dam reservoirs: benefits, costs and future. *Ecohydrology and Hydrobiology*, v. 9, n. 2–4, p. 141 – 148, 2009. ISSN 1642-3593. Disponível em: <"www.sciencedirect.com/science/article/pii/S1642359309700400">. Cited on page 17.

KRÖGER, S. et al. Sensors for observing ecosystem status. *Ocean Science*, Copernicus GmbH, v. 5, n. 4, p. 523–535, 2009. Cited on page 28.

KUTSER, T.; VAHTMÄE, E.; METSAMAA, L. Spectral library of macroalgae and benthic substrates in estonian coastal waters. *Proc. Estonian Acad. Sci. Biol. Ecol*, v. 55, n. 4, p. 329–340, 2006. Cited on page 30.

KUTSER, T. et al. Removing glint effects from field radiometry data measured in optically complex coastal and inland waters. *Remote Sensing of Environment*, v. 133, p. 85 – 89, 2013. ISSN 0034-4257. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425713000527>. Cited on page 46.

KUTSER, T.; VAHTMäE, E.; PRAKS, J. A sun glint correction method for hyperspectral imagery containing areas with non-negligible water leaving nir signal. *Remote Sensing of Environment*, v. 113, n. 10, p. 2267 – 2274, 2009. ISSN 0034-4257. Disponível em:

<http://www.sciencedirect.com/science/article/pii/S0034425709001928>. Cited on page 46.

LABORATORIES, I. W. *Ac Meter Protocol Document*. [S.l.], 2008. Cited on page 34.

LANSAC-TÔHA, F. A. et al. Abundance of planktonic ciliates in a cascading reservoirs of the paranapanema river, brazil. *Acta Scientiarum. Biological Sciences*, v. 26, p. 407–413, 2004. Cited 2 times on pages 18 and 21.

LEE, Z.; CARDER, K. L. Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. *Remote Sensing of Environment*, v. 89, n. 3, p. 361 – 368, 2004. ISSN 0034-4257. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425703002876>. Cited on page 46.

LEE, Z.; CARDER, K. L.; ARNONE, R. A. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.*, OSA, v. 41, n. 27, p. 5755–5772, Sep 2002. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-41-27-5755>. Cited 2 times on pages 69 and 70.

LEE, Z. et al. Hyperspectral remote sensing for shallow waters. i. a semianalytical model. *Appl. Opt.*, OSA, v. 37, n. 27, p. 6329–6338, Sep 1998. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-37-27-6329>. Cited on page 68.

LEWIS, M. Measurement of apparent optical properties for diagnosis of harmful algal blooms. *Real-time Coastal Observing System Dynamics and Harmful Algal Blooms*, p. 207–236, 2008. Cited on page 28.

LIN, J. et al. Hyperspectral absorption and backscattering coefficients of bulk water retrieved from a combination of remote-sensing reflectance and attenuation coefficient. *Opt. Express*, OSA, v. 26, n. 2, p. A157–A177, Jan 2018. Disponível em: <http://www.opticsexpress.org/abstract.cfm?URI=oe-26-2-A157>. Cited on page 28.

MA, R.; TANG, J.; DAI, J. Bio-optical model with optimal parameter suitable for taihu lake in water colour remote sensing. *International Journal of Remote Sensing*, Taylor & Francis, v. 27, n. 19, p. 4305–4328, 2006. Disponível em: <https://doi.org/10.1080/01431160600857428>. Cited on page 29.

MARTIN, J. et al. Automatic sun glint removal of multispectral high-resolution worldview-2 imagery for retrieving coastal shallow water parameters. *Remote Sensing*, v. 8, n. 1, 2016. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/8/1/37>. Cited on page 46.

MCCLUNEY, W. Estimation of sunlight penetration in the sea for remote sensing. 1974. Cited 2 times on pages 18 and 67.

MEDRI, M. *A bacia do rio Tibagi*. Universidade Estadual de Londrina, 2002. ISBN 9788590239017. Disponível em: <https://books.google.com.br/books?id=UQ_eAAAACAAJ>. Cited on page 22.

MENON, H. B. et al. Variability of remote sensing reflectance and implications for optical remote sensing—a study along the eastern and northeastern waters of arabian sea. *Geophysical Research Letters*, v. 33, n. 15, 2006. Disponível em: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026026>. Cited on page 28.

MOBLEY, C. D. Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Opt.*, OSA, v. 38, n. 36, p. 7442–7455, Dec 1999. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-38-36-7442>. Cited 2 times on pages 68 and 69.

MOBLEY, C. D. et al. Interpretation of hyperspectral remote-sensing imagery by spectrum matching and look-up tables. *Appl. Opt.*, OSA, v. 44, n. 17, p. 3576–3592, Jun 2005. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-44-17-3576>. Cited 2 times on pages 67 and 69.

Moe, K. et al. Sensor web technologies for nasa earth science. In: *2008 IEEE Aerospace Conference*. [S.l.: s.n.], 2008. p. 1–7. ISSN 1095-323X. Cited 2 times on pages 31 and 32.

MOORE, C. et al. Optical tools for ocean monitoring and research. Copernicus Publications on behalf of the European Geosciences Union, 2009. Cited 3 times on pages 28, 29, and 31.

MOREL, A. Optical properties of pure water and pure sea water. *Optical aspects of oceanography*, San Diego, Calif., v. 1, n. 1, p. 1–24, 1974. Cited 2 times on pages 36 and 37.

MOREL, A.; GENTILI, B. Diffuse reflectance of oceanic waters: its dependence on sun angle as influenced by the molecular scattering contribution. *Appl. Opt.*, OSA, v. 30, n. 30, p. 4427–4438, Oct 1991. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-30-30-4427>. Cited on page 68.

MOREL, A.; GENTILI, B. Diffuse reflectance of oceanic waters. iii. implication of bidirectionality for the remote-sensing problem. *Appl. Opt.*, OSA, v. 35, n. 24, p. 4850–4862, Aug 1996. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-35-24-4850>. Cited on page 68.

MORROW MICHAEL S. DUHIG, C. R. B. J. H. Design and evaluation of a cosine collector for a seawifs-compatible marine reflectance radiometer. In: . [s.n.], 1994. v. 2258. Disponível em: <https://doi.org/10.1117/12.190041>. Cited on page 68.

NAZEER, M.; NICHOL, J. E.; YUNG, Y.-K. Evaluation of atmospheric correction models and landsat surface reflectance product in an urban coastal environment. *International Journal of Remote Sensing*, Taylor & Francis, v. 35, n. 16, p. 6271–6291, 2014. Disponível em: <https://doi.org/10.1080/01431161.2014.951742>. Cited on page 58.

NOGUEIRA, M. et al. Reservatórios em cascata e os efeitos na limnologia e organização das comunidades bióticas (fitoplâncton, zooplâncton e zoobentos): Um estudo de caso no rio paranapanema (sp/pr). *Ecologias de reservatórios: impactos potenciais, ações de manejo e sistema em cascata. São Carlos: Rima*, p. 83–125, 2006. Cited on page 17.

NOVOA, S.; WERNAND, M.; WOERD, H. J. van der. Wacodi: A generic algorithm to derive the intrinsic color of natural waters from digital images. *Limnology and Oceanography: Methods*, v. 13, n. 12, p. 697–711, 2015. Disponível em: <https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.1002/lom3.10059>. Cited on page 28.

ORSI, M. L. et al. Caracterização biológica de populações de peixes do rio tibagi, localidade de sertanópolis. *A bacia do rio Tibagi. Londrina: UEL*, p. 425–432, 2002. Cited on page 18.

PRADOS, R. et al. A novel blending technique for underwater gigamosaicing. *IEEE Journal of Oceanic Engineering*, v. 37, n. 4, p. 626–644, Oct 2012. ISSN 0364-9059. Cited on page 49.

PöLöNEN H.-H. PUUPPONEN, E. H. A. L. H. S. L. M. T. H. K. N. I. Uav-based hyperspectral monitoring of small freshwater area. In: . [s.n.], 2014. v. 9239, p. 9239 – 9239 – 9. Disponível em: <https://doi.org/10.1117/12.2067422>. Cited 2 times on pages 58 and 59.

RIDDICK, C. A. et al. Spatial variability of absorption coefficients over a biogeochemical gradient in a large and optically complex shallow lake. *Journal of Geophysical Research: Oceans*, Wiley Online Library, v. 120, n. 10, p. 7040–7066, 2015. Cited on page 28.

RODRIGUES, T. W. P. et al. Sampling design in reservoirs based on Landsat-8/OLI images: a case study in Nova Avanhandava reservoir (SÃO PAULO State, BRAZIL). *Boletim de Ciẽncias Geodésicas*, scielo, v. 22, p. 303 – 323, 06 2016. ISSN 1982-2170. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1982-21702016000200303&nrm=iso>. Cited 4 times on pages 11, 29, 39, and 42.

SATHYENDRANATH, S. et al. Remote sensing of ocean colour in coastal, and other optically-complex, waters. International Ocean Colour Coordinating Group (IOCCG), 2000. Cited on page 28.

SCHAEPMAN-STRUB, G. et al. Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sensing of Environment*, v. 103, n. 1, p. 27 – 42, 2006. ISSN 0034-4257. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425706001167>. Cited on page 68.

SCHMACHTENBERG, M.; BIZER, C.; PAULHEIM, H. Adoption of the linked data best practices in different topical domains. In: MIKA, P. et al. (Ed.). *The Semantic Web – ISWC 2014*. Cham: Springer International Publishing, 2014. p. 245–260. ISBN 978-3-319-11964-9. Cited on page 31.

SHANMUGAM, P. et al. An evaluation of inversion models for retrieval of inherent optical properties from ocean color in coastal and open sea waters around korea. *Journal of Oceanography*, v. 66, n. 6, p. 815–830, Dec 2010. ISSN 1573-868X. Disponível em: <https://doi.org/10.1007/s10872-010-0066-0>. Cited on page 67.

SHANNON, C. E. A mathematical theory of communication. *Bell system technical journal*, Wiley Online Library, v. 27, n. 3, p. 379–423, 1948. Cited on page 51.

SMITH, G. M.; MILTON, E. J. The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing*, Taylor & Francis, v. 20, n. 13, p. 2653–2662, 1999. Disponível em: <https://doi.org/10.1080/014311699211994>. Cited on page 58.

SMITH, R. C.; BAKER, K. S. Optical properties of the clearest natural waters (200–800 nm). *Appl. Opt.*, OSA, v. 20, n. 2, p. 177–184, Jan 1981. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-20-2-177>. Cited on page 37.

STABEN, G. W. et al. Empirical line calibration of worldview-2 satellite imagery to reflectance data: using quadratic prediction equations. *Remote Sensing Letters*, Taylor & Francis, v. 3, n. 6, p. 521–530, 2012. Disponível em: <https://doi.org/10.1080/01431161.2011.609187>. Cited on page 59.

STOW, D. et al. Monitoring detailed land surface changes using an airborne multispectral digital camera system. *IEEE Transactions on Geoscience and Remote Sensing*, v. 34, n. 5, p. 1191–1203, Sep. 1996. ISSN 0196-2892. Cited on page 58.

STRAIT, B. J.; DEWEY, T. G. The shannon information entropy of protein sequences. *Biophysical journal*, The Biophysical Society, v. 71, n. 1, p. 148, 1996. Cited on page 51.

Strobach, P. Quadtree-structured recursive plane decomposition coding of images. *IEEE Transactions on Signal Processing*, v. 39, n. 6, p. 1380–1397, June 1991. ISSN 1053-587X. Cited 2 times on pages 47 and 50.

SU, T.-C.; CHOU, H.-T. Application of multispectral sensors carried on unmanned aerial vehicle (uav) to trophic state mapping of small reservoirs: A case study of tain-pu reservoir in kinmen, taiwan. *Remote Sensing*, v. 7, n. 8, p. 10078–10097, 2015. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/7/8/10078>. Cited on page 46.

SULLIVAN, J. M. et al. Hyperspectral temperature and salt dependencies of absorption by water and heavy water in the 400-750 nm spectral range. *Appl. Opt.*, OSA, v. 45, n. 21, p. 5294–5309, Jul 2006. Disponível em: <http://ao.osa.org/abstract.cfm?URI=ao-45-21-5294>. Cited on page 35.

Suri, D. et al. A multi-agent architecture provides smart sensing for the nasa sensor web. In: *2007 IEEE Aerospace Conference*. [S.l.: s.n.], 2007. p. 1–9. ISSN 1095-323X. Cited 2 times on pages 31 and 32.

TEILLET, P. M. et al. Towards integrated earth sensing: Advanced technologies for in situ sensing in the context of earth observation. *Canadian Journal of Remote Sensing*, Taylor & Francis, v. 28, n. 6, p. 713–718, 2002. Disponível em: <https://doi.org/10.5589/m02-068>. Cited on page 31.

TUNDISI, J.; MATSUMURA-TUNDISI, T.; TUNDISI, J. Reservoirs and human well being: new challenges for evaluating impacts and benefits in the neotropics. *Brazilian Journal of Biology*, scielo, v. 68, p. 1133 – 1135, 11 2008. ISSN 1519-6984. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-69842008000500020&nrm=iso>. Cited on page 17.

WANG, C.; MYINT, S. W. A simplified empirical line method of radiometric calibration for small unmanned aircraft systems-based remote sensing. *IEEE Journal of Selected*

Topics in Applied Earth Observations and Remote Sensing, v. 8, n. 5, p. 1876–1885, May 2015. ISSN 1939-1404. Cited on page 58.

WATANABE, F. et al. Assessment of scattering error correction techniques for ac-s meter in a tropical eutrophic reservoir. *Remote Sensing*, v. 10, n. 5, 2018. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/10/5/740>. Cited 4 times on pages 11, 29, 39, and 43.

WOźNIAK, S. B. et al. Inherent optical properties of suspended particulate matter in the southern baltic sea**financial support for this study was provided by research project grant no. n306 2838 33 awarded to s.b. woźniak by the polish ministry of science and higher education and by statutory research programme no. i.1 at the institute of oceanology, polish academy of sciences, sopot, poland. *Oceanologia*, v. 53, n. 3, p. 691 – 729, 2011. ISSN 0078-3234. Disponível em: <http://www.sciencedirect.com/science/article/pii/S007832341150020X>. Cited on page 28.

YUROVSKAYA, M. et al. Wave spectrum retrieval from airborne sunglitter images. *Remote Sensing of Environment*, v. 217, p. 61 – 71, 2018. ISSN 0034-4257. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425718303596>. Cited on page 46.

ZENG, Y. et al. Fusion of satellite images in urban area: Assessing the quality of resulting images. In: *2010 18th International Conference on Geoinformatics*. [S.l.: s.n.], 2010. p. 1–4. ISSN 2161-024X. Cited on page 51.

ZHANG, H.; WANG, M. Evaluation of sun glint models using modis measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, v. 111, n. 3, p. 492 – 506, 2010. ISSN 0022-4073. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022407309003021>. Cited on page 46.

ZHOU, G. et al. Versatile time-dependent spatial distribution model of sun glint for satellite-based ocean imaging. *Journal of Applied Remote Sensing*, v. 11, p. 11 – 11 – 18, 2017. Disponível em: <https://doi.org/10.1117/1.JRS.11.016020>. Cited on page 46.

ZIELINSKI, O. et al. Detecting marine hazardous substances and organisms: sensors for pollutants, toxins, and pathogens. 2009. Cited on page 31.