Generalization and Validation of MACHI: Minimal Atmospheric Compensation for Hyperspectral Images

Master's thesis in Engineering Mathematics and Computational Science

EDVIN DANNÄS

DEPARTMENT OF SPACE, EARTH AND ENVIRONMENT

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Generalization and Validation of MACHI: Minimal Atmospheric Compensation for Hyperspectral Images

Edvin Dannäs

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Space, Earth and Environment
Division of Geoscience and Remote Sensing
Chalmers University of Technology
Gothenburg, Sweden 2025

Edvin Dannäs

Generalization and Validation of MACHI: Minimal Atmospheric Compensation for Hyperspectral Images

Master's thesis in Remote sensing
September 2025

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

NTNU

Generalization and Validation of MACHI:
Minimal Atmospheric Compensation for Hyperspectral Images
EDVIN DANNÄS

Supervisor:
Dr. Joseph Landon Garrett, Norwegian University of Science and Technology

Examiner:
Prof. Patrick Eriksson, Department of Space, Earth and Environment

Master's Thesis 2025

Department of Space, Earth and Environment
Division of Geoscience and Remote Sensing
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics
Norwegian University of Science and Technology
NO-7491 Trondheim
Telephone +47 73 59 50 00

Typeset in LaTeX
Gothenburg, Sweden 2025

Generalization and Validation of MACHI:
Minimal Atmospheric Compensation for Hyperspectral Images
EDVIN DANNÄS
Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

Atmospheric correction (AC) is essential for deriving accurate surface reflectance from satellite imagery for ocean color studies. Previous attempts to apply AC to satellite data from the Norwegian small-satellite mission HYPSO (HYPerspectral Smallsat for Ocean observation) have been unsuccessful, with a problem of producing negative blue-band reflectance. To address this, the Minimal Atmospheric Compensation for Hyperspectral Images (MACHI) algorithm was developed as a fast, unsupervised alternative that requires no external atmospheric data or sensor-specific tuning. MACHI estimates atmospheric parameters in a simplified atmospheric model by optimizing for spectral smoothness in the retrieved ground reflectance under physical constraints.

This thesis generalizes the theoretical framework of MACHI to support arbitrary smoothing kernels and introduces a new Python implementation capable of efficient batch processing. The algorithm was applied to a selection of scenes from HYPSO-2, the second satellite of the HYPSO mission, and evaluated through a convergence analysis and validation against ground-based measurements from the AERONET-OC program, as well as a comparison to retrievals from POLYMER, an established AC algorithm for water applications.

Convergence analysis shows fast and stable convergence of MACHI. Validation indicates that MACHI retrieves physically plausible atmospheric parameters and surface reflectance spectra that are comparable to both ground measurements and established AC methods. This while also avoiding the negative reflectance artifacts observed in previous attempts. However, the validation results also reveal a systematic overestimation, particularly in the blue region, as well as a residual absorption feature near 760 nm, indicating a need for further development and testing of MACHI.

Overall, this work demonstrates the potential of smoothness-based optimization for hyperspectral AC and provides a foundation for future enhancements, such as more detailed atmospheric modeling, refined initialization, and testing on other sensors.

Keywords: atmospheric correction, hyperspectral imagery, earth observation, remote sensing, surface reflectance retrieval, regularization, optimization, MACHI

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Joseph Landon Garrett, for proposing such an exciting project and for generously sharing his expertise and enthusiasm for the topic of remote sensing. Working on this thesis has been a rewarding experience.

I am also grateful to NTNU for hosting this project at the Small Satellite Lab and to Chalmers for enabling this collaboration. I thank Morten Omholt Alver and Vishnu Perumthuruthil Suseelan for welcoming me to NTNU, and Patrick Eriksson for overseeing the project from Chalmers and for offering valuable advice throughout the process.

Finally, I want to thank my partner, Mina Alexandra, for her unwavering support and patience, making this challenging journey far brighter. I want to extend this gratitude to everyone else, family and friends, who has encouraged this effort.

I am proud of this work and hope that readers will find both enjoyment and insight in the results of many months of hard work.

Edvin Dannäs, Trondheim, September 2025

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

AC	Atmospheric correction
BoA	Bottom-of-atmosphere
DN	Digital Number
DOS	Dark object subtraction technique
HYPSO (name)	Hyper-Spectral Small Satellite for Ocean Observation
HyVa (name)	Hyperspectral remote sensing of water quality in Norwegian lakes and coasts
MACHI (name)	Minimal Atmospheric Correction for Hyperspectral Imagers
NIR	Near-infrared
NTNU (name)	Norwegian University of Science and Technology
RTS	Radiative transfer simulation
SR	Surface reflectance
ToA	Top-of-atmosphere

Nomenclature

Below is the nomenclature of indices, sets, parameters, and variables that have been used throughout this thesis.

Indices

i	Index for pixels (spatial dimension)
n	Index for bands (spectral dimension)
j	Index for kernel responses
k	Index for convolution

Sets

$[0, N-1]$	Full wavelength index range
$[0, I-1]$	Full pixel index range
$[a, b]$	Support of the kernel function $h[k]$
$[l, u]$	Local window around band n for optimization

Parameters

$R_{\text{ToA}}^{(i)}[n]$	Top-of-atmosphere reflectance for pixel i and band n
$h[k]$	Smoothing kernel function with support $[a, b]$
$\tilde{h}[k]$	Flipped smoothing kernel function
N	Total number of wavelength bands
I	Total number of pixels
L	Length of the smoothing kernel ($L = b - a + 1$)
ϵ	Convergence tolerance
maxIter	Maximum number of iterations

Variables

$S[n]$	Estimated atmospheric scattering at band n
$T[n]$	Estimated atmospheric transmittance at band n
$\beta[n]$	Estimated normalized atmospheric transmittance at band n
$R^{(i)}_{\mathrm{BoA}}[n]$	Estimated bottom-of-atmosphere reflectance for pixel i and band n
R_{rs}	Remote-sensing reflectance
$c^{(i)}_j$	Kernel response at position j for pixel i
$p^{(i)}[n]$	Smoothness penalty at pixel i, band n
P	Total smoothness penalty across all pixels and bands
$P_{\mathrm{local}}[n]$	Localized penalty around band n

Contents

List of Figures

List of Tables

1
Introduction

This section introduces atmospheric correction for hyperspectral images, including essential theory, concepts and techniques, with a focus on ocean color studies. Additionally, it contains a brief overview of current algorithms - their status and challenges. There is also a subsection on the HYPSO satellite mission and current atmospheric correction efforts for the HYPSO satellites. Finally, previous work on MACHI and the original prototype is presented.

1.1 Atmospheric correction for hyperspectral images

Hyperspectral satellite imaging enables the identification and characterization of Earth's surface from spectral signatures by providing data with high spectral resolution over continuous spectral bands. However, the signal received by the sensor, referred to as top-of-atmosphere (ToA) radiance is contaminated by interactions in the atmosphere, primarily due to molecular (Rayleigh) scattering, aerosol (Mie) scattering, and gas absorption by water vapor, ozone, and carbon dioxide [1]. In order to study Earth's surface these atmospheric effects need to be estimated and removed to retrieve a corrected spectral profile of the surface, or bottom-of-atmosphere (BoA) reflectance. This process is known as atmospheric correction.

1.1.1 Basic theory and concepts

Some essential concepts for this thesis are introduced and described. First, the relation between radiance and reflectance, and relevant conversion formulas, are described. Then, a simple atmospheric model is presented, outlining the physical interactions that are encoded in MACHI. Finally, a smoothness assumption for ground spectra is formulated, followed by how this can be used in atmospheric correction, with regularization methods as an example.

Radiance and reflectance

Atmospheric correction (AC) is the process of retrieving surface parameters from ToA *radiance* measurements. A radiance measurement can be converted to *reflectance* by dividing with the incoming irradiance. In contrast to radiance, reflectance is a material property and is therefore suitable for surface characterization [2]. It is commonly expressed in [sr^{-1}], for reflectance in a specific direction, or unit-

less for the total reflection in all directions. For a Lambertian surface the difference between directional and total reflectance is simply a scaling factor π.

So, a total ToA reflectance (R_{ToA}) can be calculated from ToA outgoing radiance (L_{ToA}) by dividing with the ToA solar irradiance (E_{ToA}):

$$R_{\mathrm{ToA}}(\lambda) = \pi \cdot \frac{L_{\mathrm{ToA}}(\lambda)}{E_{\mathrm{ToA}}} \tag{1.1}$$

$$E_{\mathrm{ToA}}(\lambda) = \frac{E_0(\lambda) \cdot \cos(\theta_{\mathrm{s}})}{d^2} \tag{1.2}$$

where E_0 is the mean incident exoatmospheric solar irradiance, θ_{s} the solar zenith angle and d the Sun-Earth distance in astronomical units. d can be estimated according to the following formula [3]:

$$d = 1 - 0.01672 \cdot \cos(0.9856^\circ \cdot (\text{Julian Day} - 4)) \tag{1.3}$$

In a similar fashion a BoA reflectance (R_{BoA}) can be calculated as the ratio of BoA radiance to downwelling irradiance just above the surface (E_{d}). For ocean color studies the BoA reflectance is commonly divided into two components: one from the interaction on the water surface (surface reflectance, SR) and another from the upwelling radiance from the ocean (*remote-sensing reflectance*, R_{rs}), where the latter is a more informative quantity regarding the inherent optical and biogeochemical properties of the water body.

Let L_{w} be the water-leaving radiance. Then the remote-sensing reflectance can be expressed as

$$R_{\mathrm{rs}}(\lambda) = \frac{L_{\mathrm{w}}(\lambda)}{E_{\mathrm{d}}(\lambda)} \quad [\mathrm{sr}^{-1}] \tag{1.4}$$

and the (non-dimensional) *normalized water-leaving reflectance*:

$$\rho_{\mathrm{w}}(\lambda) = \pi \cdot R_{\mathrm{rs}}(\lambda) \quad [-] \tag{1.5}$$

If a ground measurement of $E_{\mathrm{d}}(\lambda)$ is not available, it can be estimated by propagating E_{ToA} through the atmosphere according to the Beer–Lambert law:

$$E_{\mathrm{d}}(\lambda) = E_{\mathrm{ToA}} \cdot T_{\mathrm{d}}(\lambda) \tag{1.6}$$

$$T_{\mathrm{d}}(\lambda) = \exp\left(-\frac{\tau(\lambda)}{\cos(\theta_0)}\right) \tag{1.7}$$

where $E_0(\lambda)$ is the extraterrestrial solar irradiance [mW/(cm$^2 \cdot \mu$m)], θ_0 the solar zenith angle, $T_{\mathrm{d}}(\lambda)$ the direct atmospheric transmittance of the atmosphere along the solar path, and $\tau(\lambda)$ the total optical depth of the atmosphere at wavelength λ.

Again assuming Lambertian reflectance for simplicity[1] we get the non-directional unit-less $\rho_{\mathrm{w}}(\lambda)$ as

$$\rho_{\mathrm{w}}(\lambda) = \frac{\pi \cdot L_{\mathrm{w}}(\lambda) \cdot d^2}{E_0(\lambda) \cdot \cos(\theta_0)} \cdot \exp\left(\frac{\tau(\lambda)}{\cos(\theta_0)}\right) \tag{1.8}$$

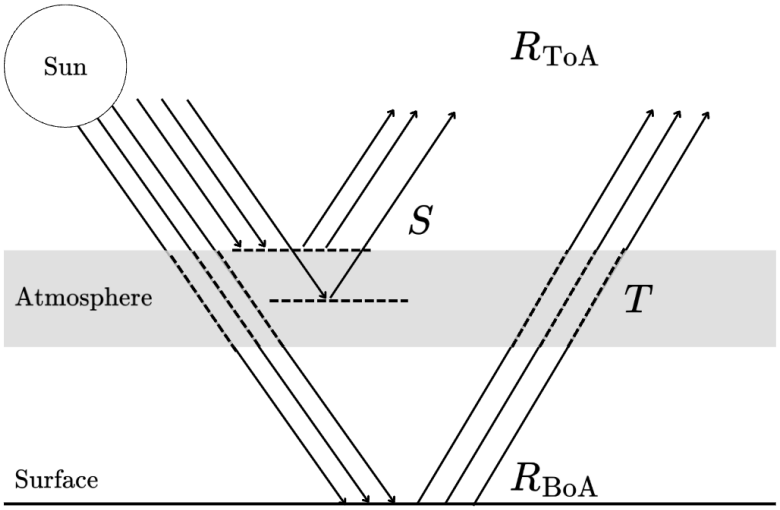


Figure 1.1: A graphical representation of the simple atmospheric model.

Simple atmospheric model

The atmospheric effects can be modeled as two components: a spectral offset resulting from back-scattered light from the atmosphere, and a spectral gain factor accounting for attenuation of the surface signal through the atmosphere. The offset is referred to as the *scattering*, or *path reflectance*, of the atmosphere, and the gain factor as the *transmittance* of the atmosphere. The ToA reflectance can then be expressed as the sum of an attenuated BoA reflectance[2] and an atmospheric scattering term [4]. In its simplest form (without any partitioning into different atmospheric contributions or surface effects) it can be written as follows:

$$R_{\mathrm{ToA}}(\lambda) = R_{\mathrm{BoA}}(\lambda) \cdot T(\lambda) + S(\lambda) \tag{1.9}$$

where R_{ToA} is the ToA reflectance, R_{BoA} the BoA reflectance, S the fraction of solar irradiance back-scattered by the atmosphere toward the satellite and T the

[1]The Lambertian assumption is an idealization; in reality, water-leaving radiance exhibits anisotropy. However, the simplification is widely adopted in satellite ocean color processing to allow consistent definitions of reflectance.

[2]Note that speaking of an attenuated reflectance does not make much physical sense, since reflectance is a surface property and does not travel through the atmosphere, as opposed to radiance, and is thereby not attenuated in itself. However, we here mean "attenuated reflectance" in the sense "reflectance based on attenuated radiance and incident irradiation". This terminology is well established and is therefore used further on in this report [2].

transmittance of the atmosphere, as a fraction of incident light; all functions of wavelength λ (see Figure 1.1). The physical boundaries are naturally incorporated as the constraints

$$R_{\mathrm{ToA}}(\lambda) - S(\lambda) \geq 0 \tag{1.10}$$

$$T(\lambda) \in (0, 1] \tag{1.11}$$

where (1.10) is interpreted as that the atmospheric scattering is bounded by the total signal, or that there is a non-negative BoA contribution, and (1.11) restricts the transmittance to its physical range. The inclusion/exclusion of zero here is a technical detail given that it is more convenient programmatically to enforce constraints on closed sets as compared to open sets.

Over the ocean around 90% of the ToA signal is atmospheric contribution, S, and in coastal waters this can be even higher, especially in the green and blue bands [5]. Since only a smaller part of the measured radiance is transmitted surface radiance, small uncertainties in the correction can give a large impact on the results. For example, if propagating a 1% uncertainty in AC (using our simple atmospheric model) for a capture over dark water assuming atmospheric contribution being 95% of the measured ToA signal, we get a ~20% uncertainty of the BoA retrieval. Now, for a capture over desert assuming an atmospheric contribution of 20% of the total signal, the same AC and uncertainty propagation gives an uncertainty of only 0.25% in the retrieved product. This simple example specifically highlights the importance of an accurate atmospheric correction for ocean color studies.

Spectral smoothness and regularization

Especially relevant to this work is the concept of spectral smoothness and its application to atmospheric correction, for example using *regularization*. While ToA reflectance often contains sharp spectral features due to gaseous absorption in specific and narrow bands in the atmosphere, the corresponding surface reflectance can be assumed to vary more smoothly with wavelength. This means that high-frequency components observed in the ToA measurement would be caused by atmospheric effects rather than an attribute of the surface reflectance, assuming an accurate calibration of the sensor. It follows that a poor atmospheric correction, in the sense of a poor estimation of gaseous and particle volumes in the atmosphere, would result in sharp features in the retrieved BoA reflectance [6]. This property is frequently exploited in AC methods by in some way optimizing for, or regularizing, smoothness in the retrieved BoA spectra. Examples of algorithms making use of this are HATCH and STEAC (see section 1.1.2 for overview of existing algorithms). Smoothness can also be applied as a spatial constraint/regularizer, given that optical properties also tend to change slowly in the spatial domain [7].

Retrieval of surface reflectance is an inverse problem where atmospheric parameters are inferred from measured ToA radiance or reflectance. The radiative transfer model cannot be solved analytically, but can instead be solved by an optimization procedure, where the atmospheric parameters are updated in an altering manner

to minimize an objective function. Due to this problem usually being ill-posed additional information on the desired solution is required for stabilization of the optimization procedure, and enforcing the uniqueness of the solution. Tikhonov regularization (TR) and the optimal estimation method (OEM) are common ways of doing this.

While OEM require prior statistical information of natural variability of the expected signal, TR enforces smoothness and/or magnitude of the solution through a regularization (penalty) term [8]. If we have the state to retrieve as $\mathbf{x} \in \mathbb{R}^m$, observed ToA reflectance as $\mathbf{b} \in \mathbb{R}^n$, and let $\mathbf{A} \in \mathbb{R}^{n \times m}$ be the coefficient matrix for the forward model of radiative transfer, we can express the optimization problem for TR as

$$\min_{\mathbf{x}} \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda^2 \|\mathbf{L}\mathbf{x}\|_2^2 \right\} \tag{1.12}$$

where $\mathbf{L}$ is a regularization matrix (e.g., derivative operator to enforce spectral smoothness) and λ the regularization parameter balancing data fidelity and penalty (e.g. smoothness).

1.1.2 Overview of existing algorithms

Having a reliable atmospheric correction algorithm is essential for remote sensing applications, such as ocean color studies and land cover classification, since it directly influences the quality of derived surface parameters. Therefore many approaches have been developed for this purpose, tailored to different platforms, sensors, and atmospheric conditions, some of which are presented in this section.

Method classes

Current methods can be divided in two main categories: empirical approaches and models based on radiative transfer simulations.
The first category of methods require a-priori information on the ground target, such as in situ measurements, or to make assumptions on the spectral profile or distribution in the image. Examples are the flat field correction approach, the empirical line approach, and the dark object subtraction technique (DOS) [9][10].
The other main category of atmospheric correction methods do not require ground truth data but instead rely on complex simulations, or similar tabular values, as well as ancillary inputs in the form of atmospheric profiles or parameters. They are based on radiative transfer simulations (RTS) using software such as MODTRAN (MODerate resolution atmospheric TRANsmission) [11] or 6S (Second Simulation of the Satellite Signal in the Solar Spectrum) [12]. Notable examples of atmospheric correction methods using RTS include ATREM (ATmosphere REMoval program) [4], based on 6S; FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) [13], based on MODTRAN, and HATCH (High-accuracy ATmospheric Correction for Hyperspectral data) [14]. ATREM was for a long time after its release the industry standard but later models such as FLAASH and HATCH have incorporated more accurate and sophisticated RTS [15]. The complexity of these

models can be a disadvantage, as well as the fact that many are designed for a specific set of conditions, such as a specific satellite system, or spectral and/or spatial resolution [16]. Many of these algorithms have historically focused primarily on land imagery, but there are also algorithms developed specifically for water applications, such as C2RCC (Case 2 Regional Coast Color) [17], POLYMER [18] or ACOLITE [19].

Status and challenges

There exists a wide range of commonly used atmospheric correction algorithms. However, many of these pose high uncertainties, or require pre-defined averaged atmospheric parameters or a certain atmospheric profile, which gives a limited adaptability for spatial and temporal variations [20]. Many have noted the potential and need of improvement in current atmospheric correction algorithms. Among these are Warren et al. (2019), that evaluated six common algorithms for retrieval of water leaving reflectance (ACOLITE, C2RCC, iCOR, l2gen, POLYMER and Sen2Cor). In this work all methods displayed extremely high uncertainties in the red and near-infrared (NIR) bands when applied to ESA Sentinel-2 Multi Spectral Imager (MSI) [21].

Several approaches have been proposed for the improvement of atmospheric correction algorithms. Yang et al. (2022) suggest establishing a relationship between satellite imagery and UAV (Unmanned Aerial Vehicle) image reflectance as a way to create a universal atmospheric correction model [20]. Others emphasizes the capabilities and potential of machine learning in atmospheric correction, one example being OC-SMART, a data analysis platform for ocean color based on machine learning algorithms [22]. There has also been interest for the development of unsupervised algorithms, that apart from many other methods would not need ground measurements or atmospheric profiles. The QUAC (QUick Atmospheric Correction) method [23] is one *weakly supervised* algorithm that finds a BoA spectra based on a predetermined dictionary of ground spectral profiles. QUAC needs to establish this reference dictionary as a prior, which can be seen as a simpler sort of "ground truth", but importantly it does not require explicit in-situ measurements for the observed scene. Building on the QUAC algorithm, Cubero-Castan et al. (2015) presented a new method, STEAC (Smoothing Technique for Empirical Atmospheric Correction), making use of a smoothing function, similar in sense to the topic of this work [24]. However, unlike STEAC, which works by simple empirical smoothing of the mean scene radiance, MACHI explicitly models and optimizes the atmospheric scattering and transmittance components under smoothness and physical constraints.

1.2 The HYPSO mission

HYPerspectral Smallsat for Ocean observation (HYPSO), is a small-satellite mission at the Norwegian University of Science and Technology (NTNU) with the purpose of delivering high-resolution remote sensing data for studying ocean color events, such as algal blooms. There are currently two satellites in (sun-synchronous) orbit,

HYPSO-1 and HYPSO-2, both 6U CubeSats at 500 km altitude hosting custom-built hyperspectral imagers. They measure in 120 binned bands in the range 387–801 nm, with a usable range from 430nm [25]. Further it has a swath width of 70 km and a 100 m spatial resolution. A third satellite, HYPSO-3, is in the concept phase, aiming to achieve significantly higher spatial resolution than its predecessors, enabling improved data quality for studies on coastal and terrestrial areas, as well as water quality monitoring in lakes and rivers [26][27].

In order to retrieve ocean color properties, such as chlorophyll-a concentration and suspended matter, there is a need for an efficient atmospheric correction algorithm for HYPSO images, something that is emphasized in [25]. Lately, some methods for atmospheric correction have been tested on HYPSO data, with varying success. Within a recent master's thesis ACOLITE was tested and compared to two other methods (6S and ELF) on HYPSO-1 images. It was shown that while ACOLITE provides the best estimate out of the three there are issues with negative reflectance around 400-500 nm, and more validation is needed [28]. This leaves room for the development and/or testing of other algorithms for HYPSO data.

1.2.1 The HyVa project and MACHI prototype

The HyVa (Hyperspectral remote sensing of water quality in Norwegian lakes and coasts) project was a collaborative initiative between NTNU, Norwegian Institute for Water Research (NIVA), and OceanObs AS during 2024. The project investigated new methods for retrieval of ocean color products from the HYPSO satellites for monitoring of Norwegian water resources and was mainly funded by the Norwegian Space Agency (NRS) [29]. A project report was released summarizing the results of this effort [30].

The report reiterated the problems with negative reflectance from ACOLITE applied to HYPSO, as well as an unexpected broad peak in the red-NIR range. It was also noted that HYPSO lacking longer wavelength NIR and SWIR bands is problematic for many mainstream atmospheric correction algorithms that rely on these bands. Possible other causes to a misalignment between retrieved products from HYPSO with those from SENTINEL-3 Ocean and Land Color Instrument (OLCI) and in-situ measurements were listed; for example bias in the HYPSO imager, sun glint effects, and low spatial resolution leading to contamination of land reflectance in coastal water captures. Due to these problems a new method, MACHI (Minimal Atmospheric Compensation for Hyperspectral Images), was outlined, tailored to the HYPSO satellite to mitigate the unwanted spectral artifacts seen in retrieved BoA spectra using conventional software. MACHI was presented as a "cyclic block coordinate descent" approach [31] optimizing smoothness in the retrieved spectra using an alternating iterative optimization procedure. A prototype of MACHI showed promising results, retrieving sufficiently smooth spectra and eliminating the issue with negative reflectance. However this version was limited to using a forward difference convolution kernel ($[1, -1]$) for the smoothness estimation, and problems regarding this selection of kernel were discovered. The prototype was made open-

source [32], but there exists no further documentation of this implementation.

1.3 Aim

The aim of this work is to, building on the MACHI prototype from the HyVa project, (1) formulate a general theoretical foundation for MACHI that can handle different smoothness measures (any convolution kernel), (2) build an implementation of this algorithm and test the performance of the implementation and (3) validate the retrieved products against in-situ data as well as compare to other methods. The broader purpose is to provide a simple atmospheric correction model based on few assumptions, and with few input parameters, that can provide reliable remote sensing products from the HYPSO satellites for ocean color studies, and that can be easily applied to other satellite data without any tuning or training of parameters (unsupervised).

2
Methods

This section contains a complete derivation of the MACHI algorithm and the optimization procedure, as well as a description of the convergence test and validation of results carried out as part of this work.

2.1 MACHI overview

The MACHI algorithm focuses on estimating the two unknowns S and T from a ToA capture. This method is based on the smooth ground spectra assumption (see section 1.1.1) in that the atmospheric scattering, S, and transmittance, T, are updated to optimize a smoothness based objective function applied to the retrieved BoA signal. Spectral derivatives of $R_{\text{BoA}}^{(i)}$, in each pixel i, are approximated by convolution with a suitable kernel, h, and incorporated into an objective function (penalty) P. This formulation of the penalty function allows generalization to arbitrary derivative orders, by selecting the appropriate kernel h. S and T are then updated in an alternating manner to minimize the smoothness penalty on R_{BoA}, in a cyclic block coordinate descent approach [31]. The summation over all pixels in P provides robustness to spatial noise and reduces variance in the estimate. The algorithm produces a single estimate of S and T for the whole image, based on the assumption that the significant atmospheric effects are global, i.e. the same for all pixels within the image.

2.2 Problem formulation

Let N be the "length" of R_{ToA}, i.e. the number of discrete wavelengths (and/or bands) in the spectrum. Then we can solve the BoA reflectance from equation (1.9) for a discrete set of wavelengths $\{\lambda_n\}_{n=0}^{N-1}$:

$$R_{\text{BoA}}[n] = \frac{R_{\text{ToA}}[n] - S[n]}{T[n]} \tag{2.1}$$

This formulation can be unstable under optimization, so we introduce $\beta[n] = (1 - T[n])/T[n] \geq 0$ to linearize the expression:

$$R_{\text{BoA}}[n] = \left\{ T[n] = \frac{1}{1 + \beta[n]} \right\} = (R_{\text{ToA}}[n] - S[n])(1 + \beta[n]) \tag{2.2}$$

For a discrete signal we can approximate derivatives using discrete convolution with a suitable kernel. Let h be a kernel that approximates the derivative operator of desired order. Then we get the discrete convolution of R_{BoA} with h as:

$$(R_{\text{BoA}} * h)[n] = \sum_{k=-\infty}^{\infty} R_{\text{BoA}}[n-k]\, h[k] = \sum_{k=a}^{b} R_{\text{BoA}}[n-k]\, h[k] \tag{2.3}$$

where a and b determine the support of the kernel and $*$ is the convolution operator.[1] This holds when $0 \leq n-k \leq N-1$. Since $a < b$ we have $b \leq n \leq N-1+a$. We use $i \in [0, I-1]$ to index the pixels in the image. We can then define the smoothness penalty at wavelength n and pixel i as:[2]

$$p^{(i)}[n] = (R^{(i)}_{\text{BoA}} * h)[n]^2 \tag{2.4}$$

and the total penalty is a summation over the spectral and spatial domain using the previously derived boundaries. We get total penalty P:

$$P = \sum_{i=0}^{I-1} \sum_{n=b}^{N-1+a} p^{(i)}[n] \tag{2.5}$$

Further, since we are aiming to minimize the total penalty, i.e. a summation over the convolution with h for all wavelengths, and assuming $0 \in [a, b]$, the support ($[a, b]$) of h is in fact arbitrary. This comes from the exclusion of the convolutions where the kernels are outside the range of n. We can instead just see h as a vector of length L, indexed from 0 by k. We now use c and j for indexing *kernel responses*, $(R_{\text{BoA}} * h)$, so that $c^{(i)}_j$ is the j:th kernel response starting from $j = 0$:

$$c^{(i)}_j = \sum_{k=0}^{L-1} R^{(i)}_{\text{BoA}}[(j + (L-1) - k]\, h[k] \tag{2.6}$$

For clarity and readability in implementation, we can rewrite this discrete convolution as a cross-correlation, applying the kernel directly, such as

$$c^{(i)}_j = \sum_{k=0}^{L-1} R^{(i)}_{\text{BoA}}[(j+k]\, \tilde{h}[k], \quad \tilde{h}[k] = h[L-1-k] \tag{2.7}$$

which corresponds to a standard convolution with flipped kernel $\tilde{h}[k]$. We simplify P:

$$P = \sum_{i=0}^{I-1} \sum_{n=b}^{N-1+a} p^{(i)}[n] = \sum_{i=0}^{I-1} \sum_{n=b}^{N-1+a} \left(\sum_{k=a}^{b} R^{(i)}_{\text{BoA}}[n-k]\, h[k] \right)^2 = \sum_{i=0}^{I-1} \sum_{j=0}^{N-L} (c^{(i)}_j)^2 \tag{2.8}$$

This can be thought of as the kernel window sliding over the signal creating the kernel responses c_n, which is visualized for one pixel in Figure 2.1. These kernel responses are no longer a function of a specific n; there is no local reference after losing

[1] For instance, a first-order derivative can be approximated by a forward difference kernel $h_1 = [1, -1]$, so that $(R_{\text{BoA}} * h_1)[n] = R_{\text{BoA}}[n] - R_{\text{BoA}}[n-1]$.

[2] The square in the penalty function ensures - besides computational convenience - that positive and negative slopes are penalized equally, if considering a derivative kernel.

information about the support of h. However, the global summation is equivalent and only this is interesting for our task. This notation provides a more intuitive indexing in further calculations.

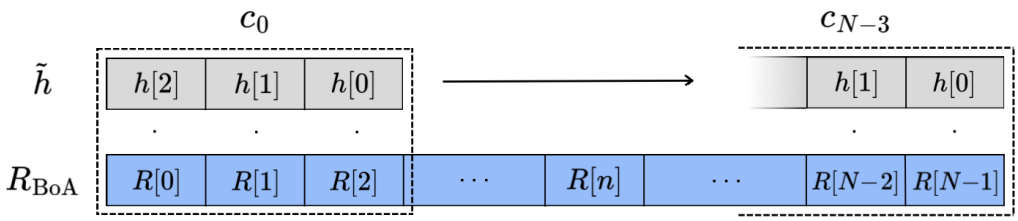


Figure 2.1: The flipped kernel window $\tilde{h}$ for a kernel (h) of length three sliding over the BoA signal R_{BoA} (denoted R in the figure for simplicity) for a single pixel i to create the kernel responses c_n as used in the total penalty function P.

To find optimal estimates for $S[n]$ and $\beta[n]$, we minimize the smoothness penalty P subject to the physical constraints. We can now formulate the minimization problem for MACHI:

$$\begin{aligned} \min_{S,\beta} \quad & \sum_{i=0}^{I-1} \sum_{j=0}^{N-L} (c_j^{(i)})^2 \\ \text{s.t.} \quad & R_{\mathrm{ToA}}^{(i)}[n] - S[n] \geq 0, \quad \forall n, i \\ & \beta[n] \geq 0, \quad \forall n \end{aligned} \tag{2.9}$$

2.2.1 Comparison to Tikhonov regularization

We can note similarities between this formulation and Tikhonov regularization with the key difference that in MACHI the data-fidelity term is replaced by a hard constraint:[3]

$$R_{\mathrm{BoA}} = (R_{\mathrm{ToA}} - S) \odot (\mathbf{1} + \beta) \tag{2.10}$$

with $\odot$ as the element-wise multiplication operator.

Let $\oslash$ denote element-wise division. Then, under the MACHI constraint, the residual between estimated and observed R_{ToA} vanishes by construction:

$$\begin{aligned} R_{\mathrm{ToA}}^{\mathrm{est}}(R_{\mathrm{BoA}}^{\mathrm{est}}, S, \beta) - R_{\mathrm{ToA}}^{\mathrm{obs}} &= \Big(R_{\mathrm{BoA}}^{\mathrm{est}}(R_{\mathrm{ToA}}^{\mathrm{obs}}, S, \beta) \oslash (\mathbf{1} + \beta) + S\Big) - R_{\mathrm{ToA}}^{\mathrm{obs}} \\ &= \Big((R_{\mathrm{ToA}}^{\mathrm{obs}} - S) \odot (\mathbf{1} + \beta) \oslash (\mathbf{1} + \beta) + S\Big) - R_{\mathrm{ToA}}^{\mathrm{obs}} \\ &= R_{\mathrm{ToA}}^{\mathrm{obs}} - R_{\mathrm{ToA}}^{\mathrm{obs}} = \mathbf{0}. \end{aligned} \tag{2.11}$$

Further on, since the estimated and observed top-of-atmosphere reflectance are identical under the MACHI constraint, as demonstrated, we can omit this distinction

[3] Again, this strictly constrained formulation enables analytical solutions in each update, providing efficiency and accuracy to the implementation.

in notation. The constraint is then substituted into the regularization term, giving the TR formulation of MACHI as something like this:

$$\begin{aligned} \min_{R_{\mathrm{BoA}}} \quad & \|\mathbf{H} R_{\mathrm{BoA}}\|_2^2 \\ \text{s.t.} \quad & R_{\mathrm{BoA}} = (R_{\mathrm{ToA}} - S) \odot (\mathbf{1} + \beta) \end{aligned} \quad \Rightarrow \quad \min_{S,\beta} \quad \|\mathbf{H}\left[(R_{\mathrm{ToA}} - S) \odot (\mathbf{1} + \beta)\right]\|_2^2 \tag{2.12}$$

Where H is a regularization matrix stacking the kernel h along the diagonal. In contrast to solving this Tikhonov-style problem, MACHI uses the formulation in equation (2.9) to construct localized, restricted minimization problems. These admit simple analytical solutions, which enable fast updates with a smaller computational load in each step, and allow easy and flexible handling of the constraints.

2.3 Optimization

The MACHI algorithm iteratively updates S and β to minimize P, as expressed in equation (2.9), using a "cyclic block coordinate descent approach" approach [31]. When updating S, β is fixed, and the entries in S are updated sequentially over the spectral bands. The same procedure holds for the updates of β. This alternating optimization is iterated until convergence or a maximum number of iterations is met.[4]

2.3.1 Localized penalty function

Let's consider an arbitrary $\hat{n} \in [0,\, N-1]$. The updates for $S[\hat{n}]$ and $\beta[\hat{n}]$ will only affect P through the kernels that overlap with the current iterate ($S[\hat{n}]$ or $\beta[\hat{n}]$). We find these as the $c_j^{(i)}$ where j is in the range $[\hat{n}-L+1,\, \hat{n}] \cap [0,\, N-L]$. The first interval ($[\hat{n}-L+1,\, \hat{n}]$) defines a local window around $\hat{n}$ while the second interval ($[0,\, N-L]$) represents the valid range of kernel positions within the data. The kernel responses outside this range stay constant during the update and can be omitted from the summation, creating a much smaller problem and computational load. We can now construct a localized penalty function $P_{\text{local}}[n]$ and express this for $\hat{n}$ as

$$P_{\text{local}}[\hat{n}] = \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} (c_j^{(i)})^2 \tag{2.13}$$

where $\hat{l} = \max(\hat{n}-L+1,\, 0)$ and $\hat{u} = \min(\hat{n},\, N-L)$. Figure 2.2 shows the kernel responses which constitutes the localized penalty function around $\hat{n}$ for one pixel. From this we can analytically derive the updates for S and β.

2.3.2 Update for $S[n]$

Let us now derive the explicit expressions for the update of $S[n]$, again initially considering a specific $n = \hat{n}$. First, we formulate the subproblem where we find the $S[\hat{n}]$ that minimizes the local penalty function $P_{\text{local}}[\hat{n}]$:

$$\begin{aligned} \min_{S[\hat{n}]} \quad & \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} (c_j^{(i)})^2 \\ \text{s.t.} \quad & R_{\text{ToA}}^{(i)}[\hat{n}] - S[\hat{n}] \geq 0, \quad \forall i \end{aligned} \tag{2.14}$$

with $\hat{l} = \max(\hat{n}-L+1,\, 0)$ and $\hat{u} = \min(\hat{n},\, N-L)$. Note that we only need to enforce the constraints that apply to the current iterate.

Due to $P_{\text{local}}[\hat{n}]$ being quadratic and convex in $S[\hat{n}]$ we can directly find the minimum by differentiating and finding the root.

[4] Due to P being quadratic and convex in both S and β this optimization approach allows for exact updates in every step, i.e. we can directly find the minimum of the local optimization problem and hence we can avoid super-parameters like step length, or gain.

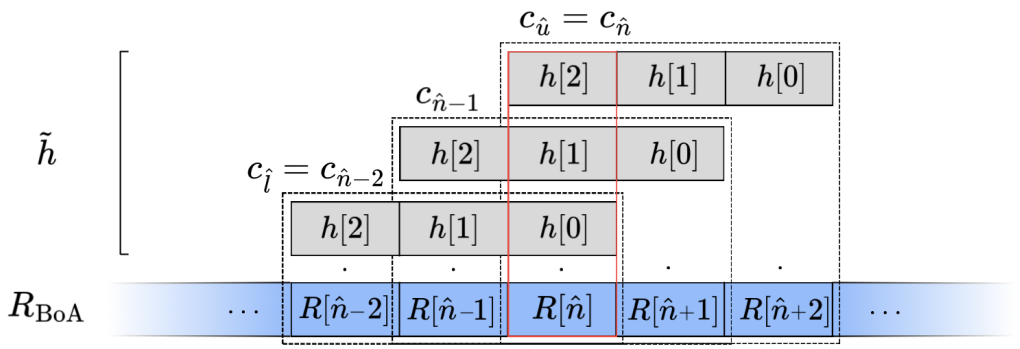


Figure 2.2: The kernel responses c_n which contribute to the localized penalty function P_{local} around $\hat{n}$ for a single pixel i, shown for a kernel (h) of length three. The flipped kernel is denoted $\tilde{h}$, and R is used to represent the BoA signal R_{BoA} in the figure for simplicity.

$$\frac{\partial P_{\mathrm{local}}}{\partial S}[\hat{n}] = \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} 2 \left(c_j^{(i)} \cdot \frac{\partial c_j^{(i)}}{\partial S[\hat{n}]} \right) \tag{2.15}$$

with

$$\begin{aligned} \frac{\partial c_j^{(i)}}{\partial S[\hat{n}]} &= \frac{\partial}{\partial S[\hat{n}]} \left(\sum_{k=0}^{L-1} R_{\mathrm{BoA}}^{(i)}[j+k]\, \tilde{h}[k] \right) \\ &= \frac{\partial}{\partial S[\hat{n}]} \left(\sum_{k=0}^{L-1} (R_{\mathrm{ToA}}^{(i)}[j+k] - S[j+k])(1+\beta[j+k])\, \tilde{h}[k] \right) \end{aligned} \tag{2.16}$$

All terms in the summation will disappear except when $S[j+k] = S[\hat{n}]$, i.e. when $k = \hat{n} - j$. We get

$$\frac{\partial c_j^{(i)}}{\partial S[\hat{n}]} = \frac{\partial}{\partial S[\hat{n}]} \left((R_{\mathrm{ToA}}^{(i)}[\hat{n}] - S[\hat{n}])(1+\beta[\hat{n}])\, \tilde{h}[\hat{n}-j] \right) = -(1+\beta[\hat{n}])\, \tilde{h}[\hat{n}-j] \tag{2.17}$$

$$\begin{aligned} \xrightarrow{\text{eq. (2.15)}} \quad \frac{\partial P_{\mathrm{local}}}{\partial S}[\hat{n}] &= \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} 2 \left(\sum_{k=0}^{L-1} R_{\mathrm{BoA}}^{(i)}[j+k]\, \tilde{h}[k] \right) \cdot \left(-(1+\beta[\hat{n}])\, \tilde{h}[\hat{n}-j] \right) \\ &= -2(1+\beta[\hat{n}]) \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n}-j] \cdot \left(\sum_{k=0}^{L-1} (R_{\mathrm{ToA}}^{(i)}[j+k] \right. \\ &\qquad \left. - S[j+k])(1+\beta[j+k])\, \tilde{h}[k] \right) \end{aligned} \tag{2.18}$$

We now find the roots by setting $(\partial P_{\text{local}}/\partial S)[\hat{n}] = 0$

$$\xrightarrow{\beta[\hat{n}]\geq 0} \sum_{i=0}^{I-1}\sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n}-j]\left(\sum_{k=0}^{L-1}(R_{\text{ToA}}^{(i)}[j+k]-S[j+k])(1+\beta[j+k])\,\tilde{h}[k]\right)=0 \tag{2.19}$$

We isolate the terms involving $S[\hat{n}]$, which we find when $j+k=\hat{n} \Rightarrow$ $k=\hat{n}-j$ (which per the definition of $[\hat{l},\hat{u}]$ is in the feasible range $[1,L]$). This gives:

$$\begin{aligned}\sum_{i=0}^{I-1}\sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n}-j]\Bigg(\sum_{\substack{k\in[0,L-1]\\k\neq\hat{n}-j}}(R_{\text{ToA}}^{(i)}[j+k]-S[j+k])(1+\beta[j+k])\,\tilde{h}[k]\Bigg)\\ +\sum_{i=0}^{I-1}\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]\Big((R_{\text{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])(1+\beta[\hat{n}])\,\tilde{h}[\hat{n}-j]\Big)=0\end{aligned} \tag{2.20}$$

and now using equation (2.2) and splitting the second summation:

$$\begin{aligned}\Rightarrow \sum_{i=0}^{I-1}\sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n}-j]\sum_{\substack{k\in[0,L-1]\\k\neq\hat{n}-j}} R_{\text{BoA}}^{(i)}[j+k]\,\tilde{h}[k]\\ +(1+\beta[\hat{n}])\sum_{i=0}^{I-1}R_{\text{ToA}}^{(i)}[\hat{n}]\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]^2\\ -S[\hat{n}](1+\beta[\hat{n}])\sum_{i=0}^{I-1}\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]^2=0\end{aligned} \tag{2.21}$$

Finally, solving for $S[\hat{n}]$ and generalizing to n we get the update as

$$\begin{aligned}S[n]=\frac{1}{(1+\beta[n])\cdot I\sum_{j=l}^{u}\tilde{h}[n-j]^2}\cdot\\ \Bigg(\sum_{i=0}^{I-1}\sum_{j=l}^{u}\tilde{h}[n-j]\sum_{\substack{k\in[0,L-1]\\k\neq n-j}}R_{\text{BoA}}^{(i)}[j+k]\,\tilde{h}[k]\\ +(1+\beta[n])\sum_{i=0}^{I-1}R_{\text{ToA}}^{(i)}[n]\sum_{j=l}^{u}\tilde{h}[n-j]^2\Bigg)\end{aligned} \tag{2.22}$$

with $l=\max(n-L+1,\,0)$ and $u=\min(n,\,N-L)$. In order to ensure alignment with constraints we project this onto the feasible set: $(S[n]\leq R_{\text{ToA}}^{(i)}[n]$ for all $i)$:

$$S[n]\leftarrow \min\Big(S[n],\,\min_i\Big(R_{\text{ToA}}^{(i)}[n]\Big)\Big) \tag{2.23}$$

2.3.3 Update for $\beta[n]$

In a similar way as for S (section 2.3.2) we formulate a local subproblem as:

$$\begin{aligned} \min_{\beta[\hat{n}]} \quad & \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} (c_j^{(i)})^2 \\ \text{s.t.} \quad & \beta[\hat{n}] \geq 0 \end{aligned} \tag{2.24}$$

$\hat{l} = \max(\hat{n} - L + 1,\, 0)$, $\hat{u} = \min(\hat{n},\, N - L)$.

$P_{\text{local}}[\hat{n}]$ is quadratic and convex also in $\beta[\hat{n}]$ so again we can find the minimum directly by differentiating and finding the root. We now differentiate w.r.t. $\beta[\hat{n}]$:

$$\frac{\partial P_{\text{local}}}{\partial \beta}[\hat{n}] = \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} 2 \left(c_j^{(i)} \cdot \frac{\partial c_j^{(i)}}{\partial \beta[\hat{n}]} \right) \tag{2.25}$$

with

$$\frac{\partial c_j^{(i)}}{\partial \beta[\hat{n}]} = \frac{\partial}{\partial \beta[\hat{n}]} \left(\sum_{k=0}^{L-1} (R_{\text{ToA}}^{(i)}[j+k] - S[j+k])(1 + \beta[j+k])\, \tilde{h}[k] \right) \tag{2.26}$$

Again all terms disappear except when $\beta[j+k] = \beta[\hat{n}]$, i.e. when $k = \hat{n} - j$. We get

$$\frac{\partial c_j^{(i)}}{\partial \beta[\hat{n}]} = (R_{\text{ToA}}^{(i)}[\hat{n}] - S[\hat{n}])\, \tilde{h}[\hat{n} - j] \tag{2.27}$$

$$\begin{aligned} \xrightarrow{\text{eq. (2.25)}} \quad \frac{\partial P_{\text{local}}}{\partial \beta[\hat{n}]} &= \sum_{i=0}^{I-1} \sum_{j=\hat{l}}^{\hat{u}} 2 \left(\sum_{k=0}^{L-1} R_{\text{BoA}}^{(i)}[j+k]\, \tilde{h}[k] \right) \cdot \left((R_{\text{ToA}}^{(i)}[\hat{n}] - S[\hat{n}])\, \tilde{h}[\hat{n} - j] \right) \\ &= 2 \cdot \sum_{i=0}^{I-1} (R_{\text{ToA}}^{(i)}[\hat{n}] - S[\hat{n}]) \sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n} - j] \cdot \Bigg(\sum_{k=0}^{L-1} (R_{\text{ToA}}^{(i)}[j+k] \\ &\qquad - S[j+k])(1 + \beta[j+k])\, \tilde{h}[k] \Bigg) \end{aligned} \tag{2.28}$$

And finding the root ($\partial P_{\text{local}} / \partial \beta[\hat{n}] = 0$):

$$\begin{aligned} \sum_{i=0}^{I-1} (R_{\text{ToA}}^{(i)}[\hat{n}] - S[\hat{n}]) \sum_{j=\hat{l}}^{\hat{u}} \tilde{h}[\hat{n} - j] \Bigg(\sum_{k=0}^{L-1} & (R_{\text{ToA}}^{(i)}[j+k] - S[j+k]) \\ & \cdot (1 + \beta[j+k])\, \tilde{h}[k] \Bigg) = 0 \end{aligned} \tag{2.29}$$

Now we isolate the terms with $\beta[\hat{n}]$ in the same manner as for $S[\hat{n}]$, starting by breaking the case $k = \hat{n} - j$ out of the summation. We get:

$$\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]\Bigg(\sum_{\substack{k\in[0,L-1]\\k\neq\hat{n}-j}}(R_{\mathrm{ToA}}^{(i)}[j+k]-S[j+k]) \cdot(1+\beta[j+k])\,\tilde{h}[k]\Bigg) + \sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]\Big((R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])(1+\beta[\hat{n}])\,\tilde{h}[\hat{n}-j]\Big)=0 \tag{2.30}$$

and now using equation (2.2) and splitting the second summation:

$$\begin{aligned}\Rightarrow \sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]\sum_{\substack{k\in[0,L-1]\\k\neq\hat{n}-j}}R_{\mathrm{BoA}}^{(i)}[j+k]\,\tilde{h}[k]\\ +\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])^2\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]^2\\ +\beta[\hat{n}]\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[\hat{n}]-S[\hat{n}])^2\sum_{j=\hat{l}}^{\hat{u}}\tilde{h}[\hat{n}-j]^2=0\end{aligned} \tag{2.31}$$

Rearranging for $\beta[n]$ and generalizing to n, we obtain:

$$\begin{aligned}\beta[n]=\frac{1}{\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[n]-S[n])^2\sum_{j=l}^{u}\tilde{h}[n-j]^2}\cdot\\ \Bigg(-\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[n]-S[n])\sum_{j=l}^{u}\tilde{h}[n-j]\sum_{\substack{k\in[0,L-1]\\k\neq n-j}}R_{\mathrm{BoA}}^{(i)}[j+k]\,\tilde{h}[k]\\ -\sum_{i=0}^{I-1}(R_{\mathrm{ToA}}^{(i)}[n]-S[n])^2\sum_{j=l}^{u}\tilde{h}[n-j]^2\Bigg)\end{aligned} \tag{2.32}$$

with $l=\max(\hat{n}-L+1,\,0)$ and $u=\min(n,\,N-L)$. To maintain alignment with constraints we project this onto the feasible set: ($\beta[n]\geq 0$):

$$\beta[n]\leftarrow\max(\beta[n],\,0) \tag{2.33}$$

2.4 Implementation

This sections contains details on kernel selection, initialization, stopping criteria and batching, for the MACHI implementation made for this thesis. There is also a pseudo-code outline of the algorithm.

2.4.1 Kernel selection

The choice of convolution kernel h critically influences the behavior of the smoothness penalty. It determines the order and locality of the regularization applied to the estimated BoA reflectance. Note that no support needs to be defined, see section 2.2. The performance of MACHI was evaluated for the following kernels:[5]

Operator	**Kernel**	**Label**
First-order forward difference	$[1, -1]$	h_1
First-order central difference	$[1, 0, -1]$	h_2
Second-order central difference	$[1, -2, 1]$	h_3
Third-order central difference	$[1, -3, 3, -1]$	h_4

Table 2.1: Selected kernels for evaluation with MACHI.

Without loss of generality (w.l.o.g.), the kernel h is normalized such that:

$$\sum_k |h[k]| = 1 \tag{2.34}$$

This prevents the scale of the regularization term from depending on the absolute magnitude of the kernel and ensures comparability across choices.

2.4.2 Parameter initialization

The initial estimate of S[n] is obtained using a simplified DOS [10] strategy. This is done by finding the pixel in the image with the lowest total reflectance over all bands, and setting $S_0[n]$ to the spectrum measured in this pixel. This method generally assumes that certain regions in the image exhibit near-zero reflectance (e.g., deep water), and that the signal for these dark objects consists entirely of the atmospheric contribution. However this assumption is not as critical when the estimate is only used for initialization.

The transmission $T[n]$ is then initialized as:

$$T_0[n] = 1 - S_0[n] \Rightarrow \beta_0[n] = \frac{1 - (1 - S_0[n])}{1 - S_0[n]} = \frac{S_0[n]}{1 - S_0[n]} \tag{2.35}$$

The physical interpretation of this initialization of $T[n]$ is simply that all incoming light that is not scattered (S) is transmitted (T).

[5] Absolute magnitude does not matter here since the kernel is later normalized. Note also that h_1 is the kernel used in the original MACHI prototype.

2.4.3 Batching

In order to efficiently process a large image, a smaller (random) selection of pixels is made at each iteration (one update for both S and T). This enables the processing of the whole image while keeping down the requirements on memory and computation power for the algorithm. The batch size indicates the number of pixels selected in each iteration.

2.4.4 Stopping criteria

The optimization is terminated when one of two stopping criteria is met:

1. The algorithm is considered to have converged when the relative change of the penalty for one update step of both S and β is below a small threshold ϵ. Note that the estimate is for a new batch each iteration and the penalty will likely not converge monotonically.

$$\frac{P_{\text{before}} - P_{\text{after}}}{P_{\text{before}}} < \epsilon$$

   Note that the exact solutions to the optimization in each step should ensure $P_{\text{before}} \geq P_{\text{after}}$.

2. A maximum number of iterations is applied as an emergency break in the case of slow or no convergence.

2.4.5 Algorithm outline

The following pseudo-code outlines the MACHI algorithm as implemented in this thesis.

Algorithm 1 MACHI Atmospheric Correction Algorithm

Input: ToA reflectance $R^{(i)}_{\text{ToA}}[n]$ for pixels $i = 0, \ldots, I-1$ and wavelengths $n = 0, \ldots, N-1$
Parameters: Kernel $h[k]$, maximum iterations maxIter, tolerance ϵ
Initialization:
    Initialize $S_0[n] \leftarrow$ DOS
    Initialize $\beta_0[n] \leftarrow S_0[n]/(1 - S_0[n])$
    Define flipped kernel $\tilde{h}[k] = h[L-1-k]$
    Normalize kernel $\tilde{h} \leftarrow \frac{\tilde{h}}{\sum_k |\tilde{h}[k]|}$
while $iter \leq maxIter$ **do**
    Create batch $R^{(i)}_{\text{ToA, batch}}[n], \quad i \in [0, I_{\text{batch}} - 1]$
    Calculate total penalty P_{before}
    Step 1: Update $S[n]$**, fixed** $\beta[n]$
    for $n = 0$ to $N-1$ **do**
        Determine valid local window:

$$l = \max(n - L + 1,\, 0), u = \min(n,\, N - L)$$

        Update S[n] (eq. (2.22))
        Project onto feasible set ($S[n] \leq R^{(i)}_{\text{ToA}}[n]$ for all i):

$$S[n] = \min\left(S[n],\, \min_i \left(R^{(i)}_{\text{ToA}}[n]\right)\right)$$

    end for
    Step 2: Update $\beta[n]$**, fixed** $S[n]$
    for $n = 1$ to N **do**
        Determine valid local window:

$$l = \max(n - L + 1,\, 0), u = \min(n,\, N - L)$$

        Update $\beta[n]$ (eq. (2.32))
        Project onto feasible set ($\beta[n] \geq 0$):

$$\beta[n] = \max(\beta[n],\, 0)$$

    end for
    Calculate total penalty P_{after}
    Step 3: Check for convergence
    if $(P_{\text{before}} - P_{\text{after}})/(P_{\text{before}} + P_{\text{after}}) < \epsilon$ **then**
        break
    end if
end while
Output: Estimated atmospheric parameters $S[n]$ and $T[n] = 1/(1 + \beta[n])$, and the retrieved R_{BoA} as $R_{\text{BoA}}[n] = (R_{\text{ToA}}[n] - S[n])(1 + \beta[n])$

2.5 Algorithm testing

The performance of the algorithm was evaluated for different selections of h, and validated on HYPSO-2 images over ground measurement sites. This section contains a description of the necessary pre-processing of the images, the convergence evaluation and the validation with ground measurements.

2.5.1 HYPSO-2 data pre-processing

The images used for testing the MACHI algorithm contain 653 016 pixels in the shape (598, 1092). The raw image contains, for each pixel, 120 bands with measured radiance intensities as Digital Numbers (DN) in the range 0 to 36 855. The data is cropped to the usable spectral range 430 to 800 nm, and further, the two highest bands (787.9 nm, 791.3 nm) are removed due to poor data. This leaves 103 bands in the range 432.6 to 784.5 nm. To avoid the distorted spectra of saturated pixels and non-linear effects when approaching the sensor limit, all pixels where any band is above 90% of saturation were masked out. The DN values were then converted to reflectances using the custom `hypso` Python package [33]. This conversion is based on equation (1.1).

2.5.2 Convergence analysis

For a convergence test MACHI was run on two captures, Oslo fjord (2025-06-20) and Chesapeake Bay (2025-02-24), for all kernels (h_1-h_4) and a selection of batch sizes (100, 1000, 10000 and 100000). The convergence of the total penalty was characterized and compared between the batch sizes for all kernels. Then, the final estimates of S and T are compared between the batch sizes, as well as the retrieved BoA reflectance spectra for two pixels, one over land and one over water. Based on these results a batch size of 1000 is selected as default for the following runs, providing both a fast convergence and reliable output. All MACHI runs were made with 16GB RAM on an Apple M2 Pro chip, with $\epsilon = 10^{-2}$ as convergence threshold.

2.5.3 Validation

The retrieved BoA reflectance spectra were validated against ground measurements from Aerosol Robotic Network for Ocean Color (AERONET-OC) [34] on four sites: "AAOT" (Venice, Italy), "Ariake Tower 2" (Ariake Sea, Japan), Martha's Vineyard Coastal Observatory - "MVCO" (Massachusetts, USA) and "Section-7 Platform" (Constanta, Romania). For each site a data pair, i.e. HYPSO capture and corresponding ground measurement, with good temporal alignment was found. The pixel from the HYPSO capture with pixel coordinates closest to the ground measurement site was extracted, i.e. the overlapping data.

Water-leaving radiance (L_w) was acquired from the AERONET-OC Level 1.5 data,[6] and remote-sensing reflectances calculated by equation (1.8). The solar spectrum

[6] Level 1.5 indicate cloud-screened and quality-controlled data.

was taken from Thuillier et al. (2003) [35], while the solar zenith angle and total optical depth were obtained from the AERONET-OC data. The total optical depth was calculated as a sum of individual components as

$$\tau = \tau_{\text{Rayleigh}} + \tau_{\text{aerosol}} + \tau_{\text{ozone}} + \tau_{\text{NO}_2} + \tau_{\text{water-vapor}} \tag{2.36}$$

The spectra were compared to retrievals from POLYMER [18] for the same captures, an established AC algorithm for water applications that had been previously applied to HYPSO satellites. These output were kindly provided by Dr. Vishnu Perumthuruthil Suseelan.

3
Results

This chapter contains two sections: first, the convergence analysis of MACHI, then the validation of the corrected reflectances over AERONET-OC sites.

3.1 Convergence analysis

Figure 3.1 shows an RGB plot of the uncorrected image over the Oslo fjord. There are 77 228 saturated pixels in the image, which is around 12 % of all pixels. The distribution of saturation across the bands is seen in Figure 3.2.

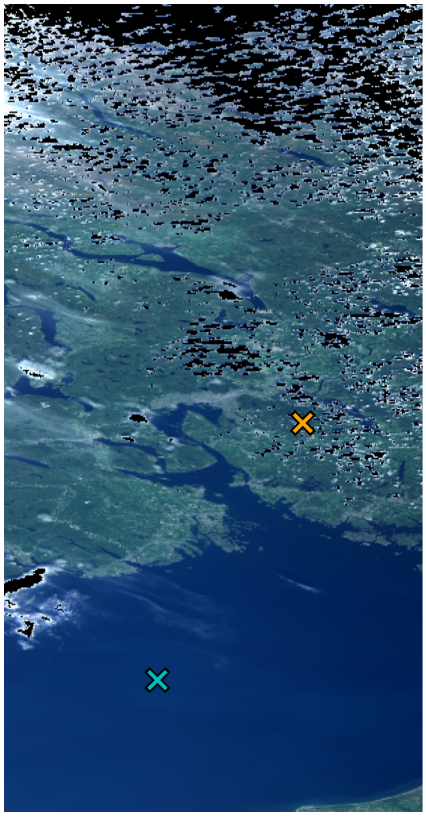


Figure 3.1: RGB plot of ToA reflectance from a HYPSO-2 capture over the Oslo fjord on 2025-06-20, with saturation mask (black), and pixel locations marked in turquoise (water) and orange (land).

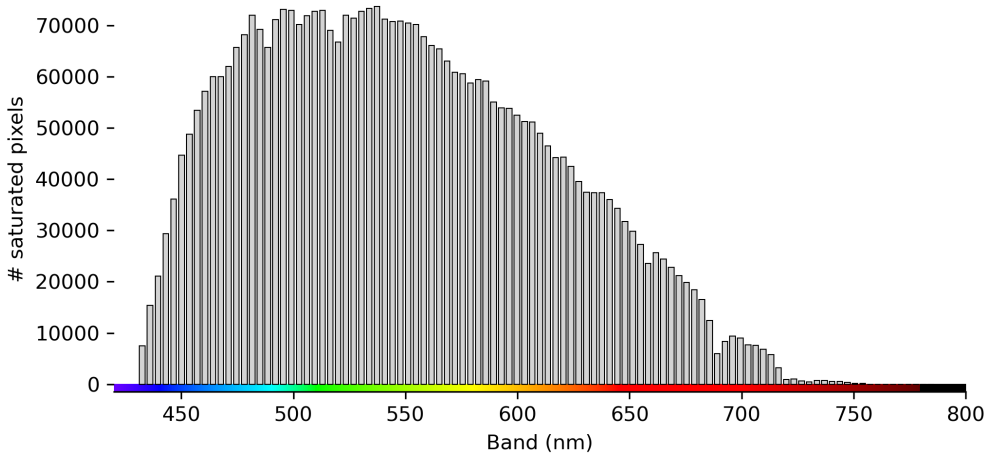


Figure 3.2: Number of saturated pixels per band for the HYPSO capture over the Oslo fjord on 2025-06-20.

The convergence of MACHI on this HYPSO capture over the Oslo fjord can be seen in Figure 3.3, for kernels h_1-h_4 and batch sizes of 100, 1000, 10000, and 100000 pixels. The measure used for analyzing convergence is the total penalty for the batch. Since this includes a summation over the pixels, this measure is proportional to the batch size, so for comparison the penalty is divided with the batch size. This is the value on the y-axis in the convergence plot. Further, we base the analysis of convergence *rate* on the product of iteration count and batch size. This can be seen as a proxy for number of computations, or similarly the total runtime, as found in Table 3.1. In the convergence plot, two values are plotted per iteration: the penalty before and the penalty after the updates of S and β. Since there is a new batching each iteration the penalty is not expected to decrease monotonically. Note also that the penalty is not comparable between kernels since they estimate different properties of the spectrum.

Batch size	h_1	h_2	h_3	h_4
100	0.52	1.64	1.79	1.35
1000	4.80	8.87	5.60	8.53
10000	41.83	95.46	3534.07	1083.89
100000	5536.92	12194.06	11517.68	4476.07

Table 3.1: Runtime (in seconds) of MACHI for kernels h_1-h_4 and all batch sizes, over the Oslo fjord.

We can observe that larger batch sizes provide slower but more stable convergence. The larger batch sizes also seem to show almost monotonic convergence, which can be a good indication that the algorithm approaches a global optimum, since the minimization seem to be robust to the batch selection.

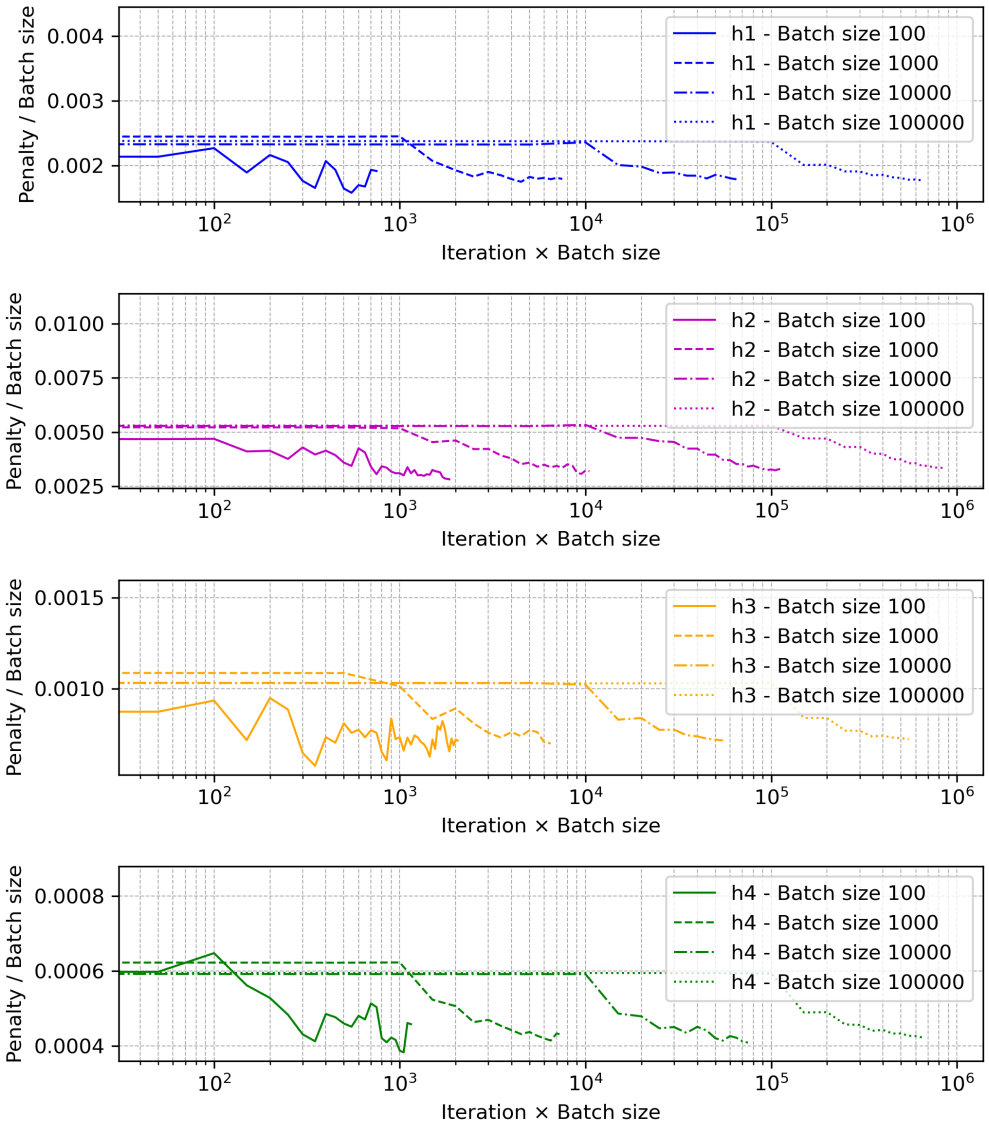


Figure 3.3: Scaled penalty (y-axis) by scaled iteration index (x-axis) for kernels h_1-h_4 and all batch sizes, over the Oslo fjord. Note log-scale on x-axis.

The final estimates also have similar normalized total penalty (y-axis) for the final batch, when comparing between the batch sizes, however, a better measure of the quality of the retrieved BoA spectrum is the total penalty over all pixels. In Table 3.2 we can see that even this is very similar between the batch sizes. Interestingly, the smallest batch size (100), while displaying a jagged convergence pattern seems to, with some consistency, converge to a smaller value than the larger batch sizes.

Batch size	h_1	h_2	h_3	h_4
100	1019	1832	403.8	236.8
1000	1011	1862	413.7	241.1
10000	1023	1863	415.9	240.2
100000	1020	1929	416.7	243.0

Table 3.2: Total penalty over all pixels after convergence, for kernels h_1-h_4 and the different batch sizes, over the Oslo fjord.

Let us now look at the final estimates of S and T and the retrieved BoA reflectance spectra for the two pixels (land and water) as marked in Figure 3.1. As we can observe in Figure 3.4, there is a very good alignment of the estimates of atmospheric parameters between the batch sizes. Apart from batch size 100 deviating some for kernel h_2, it is not possible to see any difference in performance.

When instead looking at the retrieved BoA spectra (Figure 3.5) we see that while the spectra for the land pixel are almost inseparable, there are some minor discrepancies between the spectra for the water pixel, especially in the lower bands ($<$ 500nm). The convergence analysis for the HYPSO-2 capture over Chesapeake Bay show similar results, but with significant discrepancies in the lower bands ($<$ 600nm) for the water pixel BoA spectra (Appendix A). For a qualitative analysis of retrieved spectra from MACHI, see next section (3.2).

When combining the convergence results for the selected batch sizes, a batch size of 1000 seems to provide a good trade-off between stability, performance and computational efficiency, and is therefore used in validation.

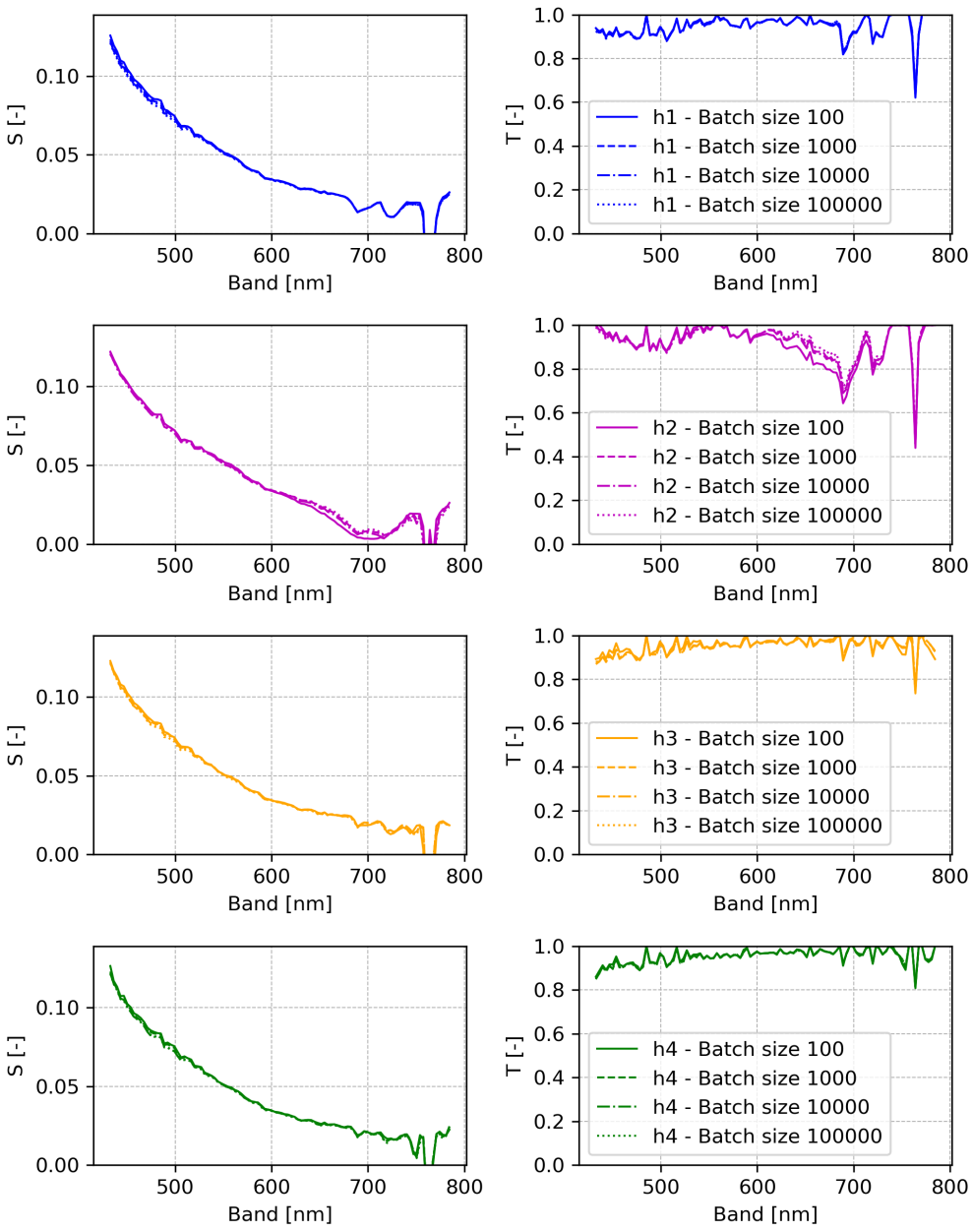


Figure 3.4: Final S and T from MACHI over the Oslo fjord.

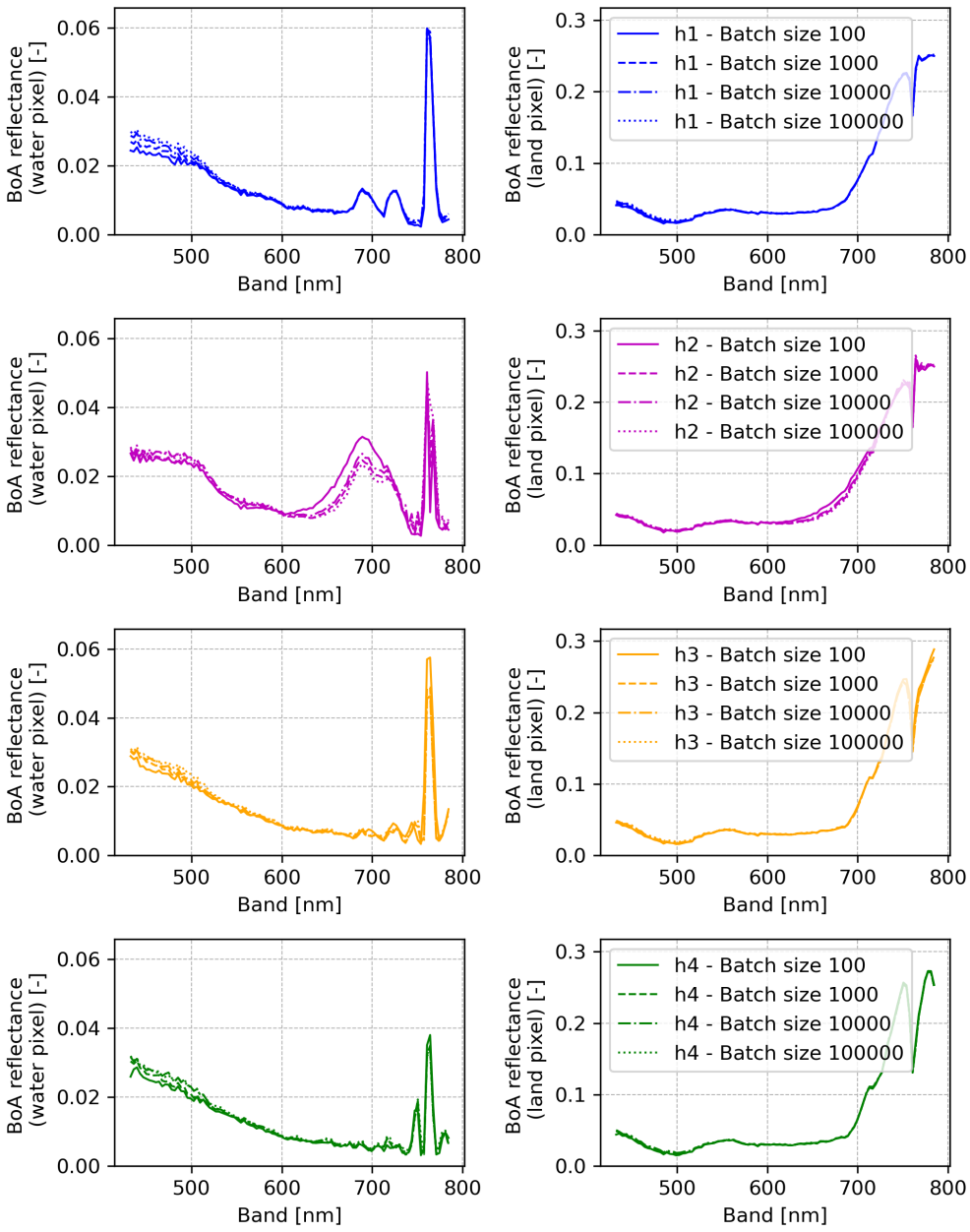


Figure 3.5: Retrieved BoA reflectance from MACHI for two pixels (land and water) over the Oslo fjord.

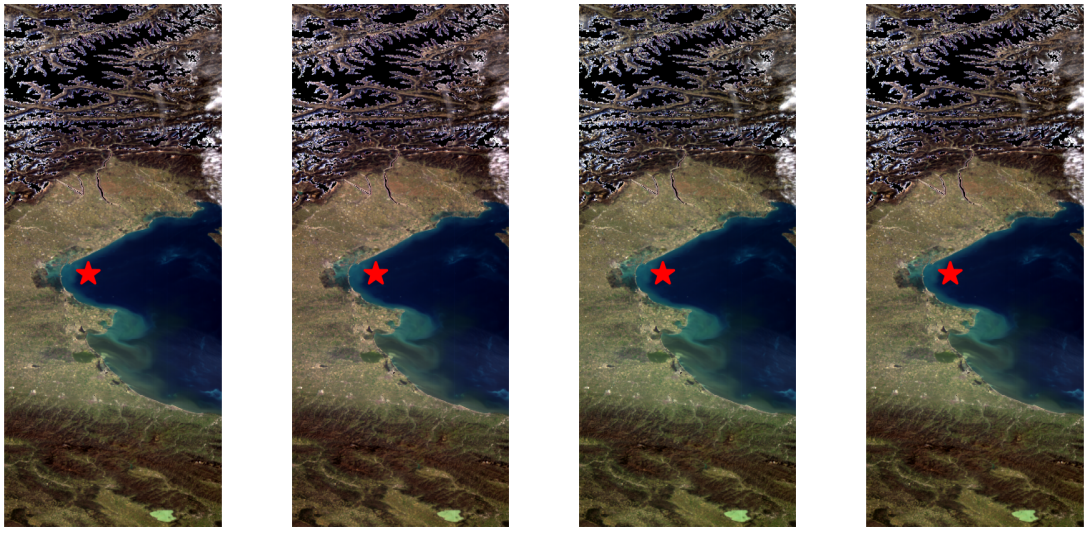
3.2 Validation

This section contains the validation of MACHI spectra over the AERONET-OC sites, with a detailed analysis of "AAOT", and a summary section for the other sites.

3.2.1 "AAOT"

The "AAOT" measurement site in Venice (Italy) is located at coordinates 45.3139°N, 12.5083°E, where a measurement was found for 4 March 2025 at 10:39:40, closely coinciding with the HYPSO-2 satellite capture at 10:38:05 on the same day. The uncorrected HYPSO-2 image with the validation site marked is shown in Figure 3.6. The corresponding corrected images (RGB) from kernels h_1-h_4 are found in Figure 3.7. We see that all kernels remove the disproportionate blue contribution (Rayleigh) from the atmosphere, with only minor differences in color nuances between the kernels.

Figure 3.6: RGB plot of (uncorrected) ToA reflectance from HYPSO-2 capture over Venice (Italy) with AERONET-OC validation site "AAOT" marked with a red star, and saturation mask in black, on 2025-03-04.



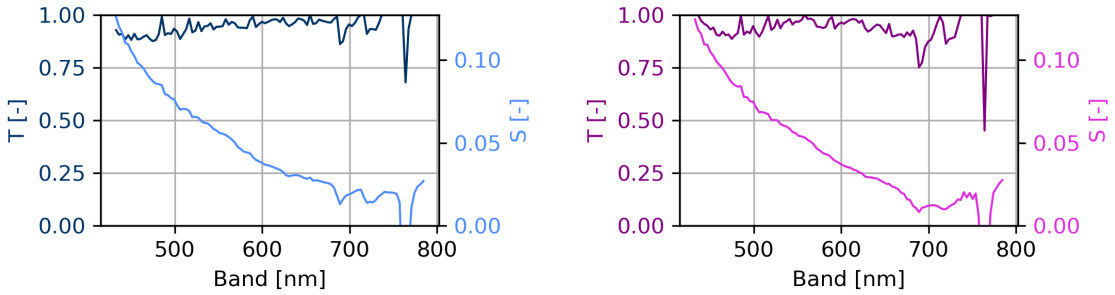
(a) Kernel h_1 **(b)** Kernel h_2 **(c)** Kernel h_3 **(d)** Kernel h_4

Figure 3.7: RGB plot of (corrected) BoA reflectance from HYPSO-2 capture over Venice (Italy) with AERONET-OC validation site "AAOT" marked with a red star, and saturation mask in black, on 2025-03-04.

Figure 3.8 shows the estimated atmospheric parameters for the same capture, with the spectra showing similar characteristics across the kernels. The most prominent feature in the transmission spectra is the oxygen A-band observed around 760 nm for all kernels. In addition, kernels h_1 and h_2 display a possible second oxygen B-band at 690 nm along with a weaker absorption feature in the water vapor region around ~720 nm.

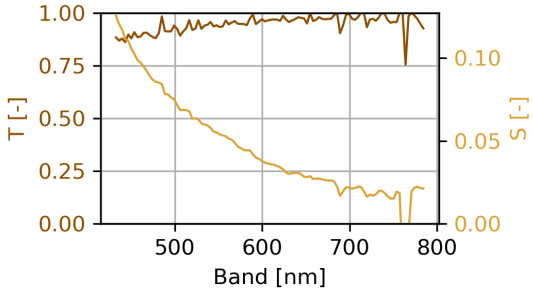
We now validate the BoA reflectance spectra from MACHI against AERONET-OC ground measurements. Figure 3.9 shows the retrieval from all kernels against the POLYMER retrieval and AERONET-OC ground measurement. We can observe that we get a lowered corrected spectrum relative to the uncorrected ToA reflectance that is also smooth in the lower bands. However, there is a broad peak appearing for h_2 (and to a smaller extent for h_1) around 700 nm, and all kernels fail to accurately correct for the absorption peak at 760 nm.

In Figure 3.10 we look closer at the spectrum of interest for water color studies, mainly the blue and green regions. We can observe that MACHI aligns relatively well with ground measurements apart from a major deviation for the lowest bands, where there is an overestimation by MACHI. There seems to be overall a small overestimation by MACHI.

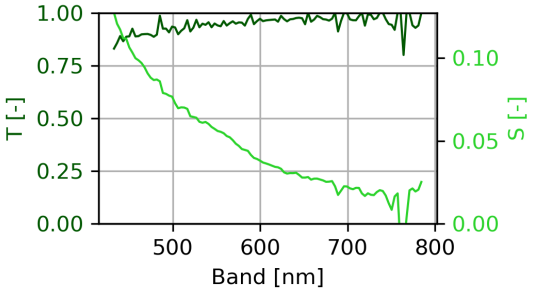


(a) Kernel h_1

(b) Kernel h_2



(c) Kernel h_3



(d) Kernel h_4

Figure 3.8: Estimated S and T from MACHI over "AAOT" validation site.

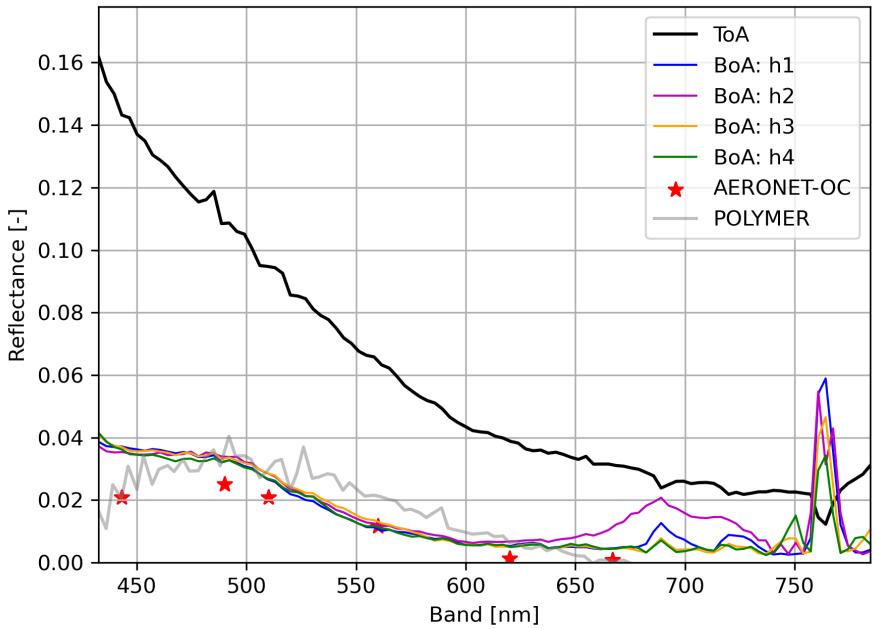


Figure 3.9: Retrieved BoA reflectance from MACHI compared to POLYMER retrieval and AERONET-OC ground measurement, at "AAOT" (full spectrum range).

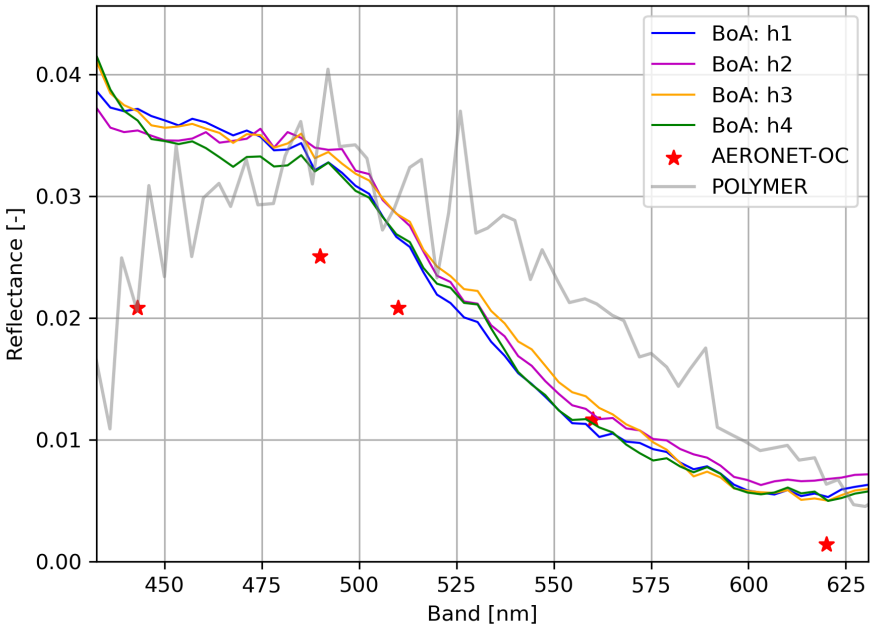


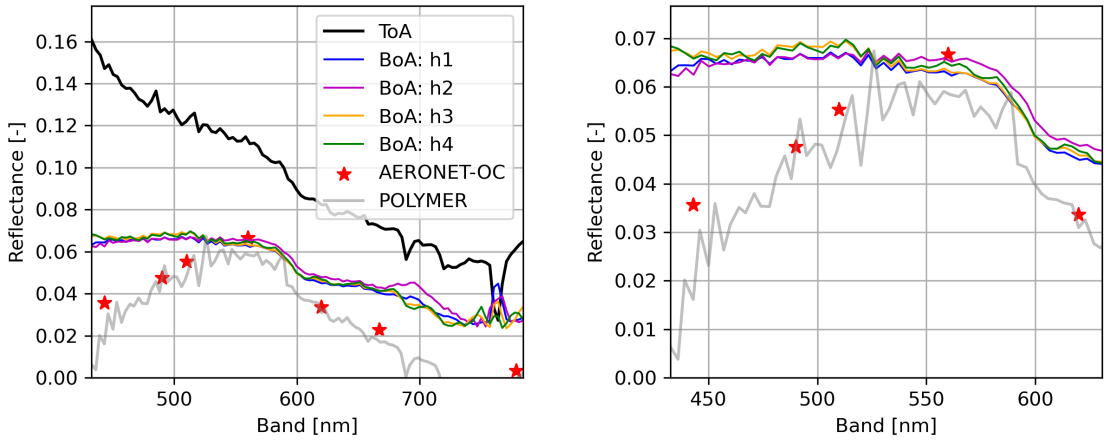
Figure 3.10: Retrieved BoA reflectance from MACHI compared to POLYMER retrieval and AERONET-OC ground measurement at "AAOT" (water spectrum range).

3.2.2 Other sites

We now take a look at the three other validation sites to assess the performance of MACHI in different settings. A summarized description of the ground measurements and related HYPSO captures is given in Table 3.3. The validation on these sites, as seen in Figures 3.11-3.13 shows similar characteristics of the MACHI retrievals, such as an overestimation in the lower bands and difficulties correcting for the absorption peak at 760 nm. The MACHI performance is further discussed in the next section. All validation material for these sites are found in Appendices B-D.

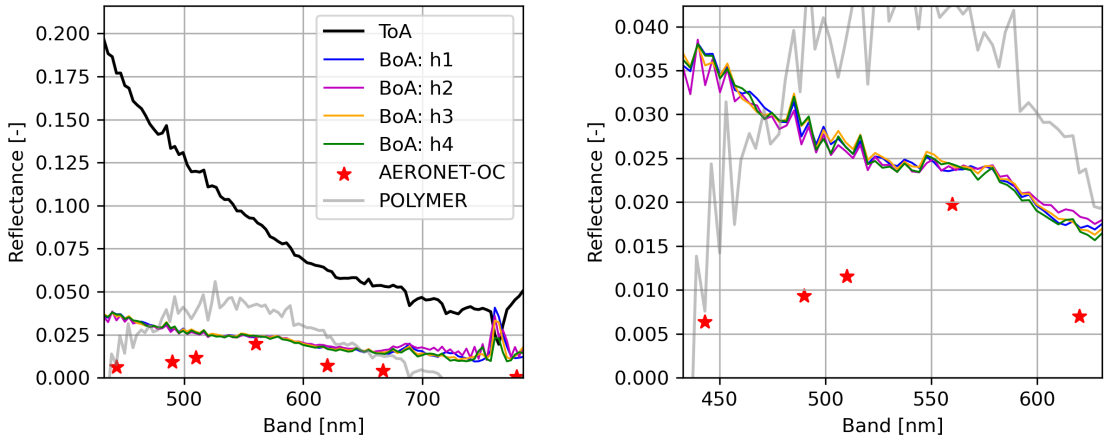
Site Observation station Geografic location	**Coordinates** Latitude Longitude	**Date || Time** AERONET-OC HYPSO
Ariake Tower 2 Ariake sea, Japan	33.114°N 130.29°E	2025-04-04 || 02:32:00 2025-04-04 || 02:27:25
Section-7 Platform Constanta, Romania	44.5458°N 29.4466°E	2025-02-02 || 07:50:48 2025-02-02 || 09:06:35
MVCO Massachusetts, USA	41.322°N -70.5667°E	2025-04-10 || 15:53:21 2025-04-10 || 15:49:22

Table 3.3: Name and location of validation sites, as well as time of ground measurement from AERONET-OC and time of HYPSO-2 capture.



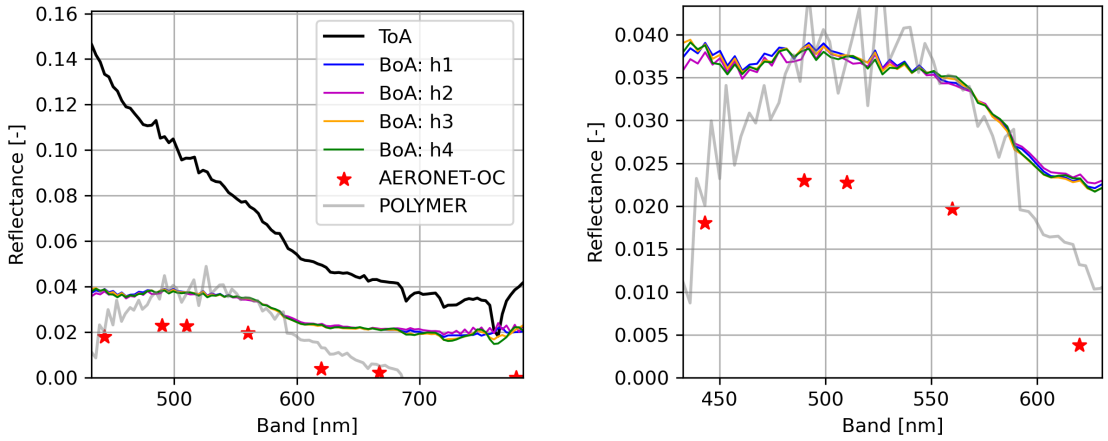
(a) **(b)**

Figure 3.11: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "Ariake Tower 2" for (a) full spectrum range and (b) water spectrum range.



(a) **(b)**

Figure 3.12: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "Section-7 Platform" for (a) full spectrum range and (b) water spectrum range.



(a) **(b)**

Figure 3.13: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "MVCO" for (a) full spectrum range and (b) water spectrum range.

4

Discussion

The results are here discussed in terms of convergence and validation as well as recommendations for future research on MACHI and conclusions.

4.1 Convergence

The convergence test shows that MACHI converges fast in terms of number of iterations. The larger batch sizes provides more stable convergence, while the smallest batch size tested (100) in many cases converged to a solution with a smaller total penalty, i.e. more optimal, while also being computationally favorable. This is not completely unexpected since a smaller batch size favors exploration at the cost of unstable and sometimes no convergence. Similarly, the near-monotonic convergence of large batch sizes is expected, since a sufficiently large batch size is likely to capture all essential features of the image.

4.2 Validation

The atmospheric parameters estimated by MACHI (e.g. Figure 3.8) align roughly with expectation, capturing both expected transmission spectra and the Rayleigh atmospheric scattering component. This even though the Rayleigh effect is spectrally smooth, and not necessarily expected to be modeled well in a smoothness-based approach such as MACHI. These results provide some confirmation of the physical modeling in MACHI. The validation with in-situ measurements shows a varying performance of MACHI across the sites, and can be discussed in terms of different characteristics:

- *Smoothness* : MACHI provides smooth spectra in the water region (< ~650 nm), where the output from POLYMER instead is very jagged. However MACHI fails to correct for the major absorption peak at ~760nm. This however is likely not a feature of MACHI but instead due to lacking correction of the smile effect in HYPSO-2, meaning that the wavelength at which this relatively sharp negative peak is occurring in the ToA spectrum is slightly shifted between the pixels, so no global optimum to the smoothing problem can be found.
- *Accuracy* : MACHI provides reflectances that are comparable to ground measurements, and in comparision to POLYMER retrievals, MACHI seems to give

good estimates for "AAOT" and "Section-7 Platform", while giving slightly worse estimates for "Ariake Tower 2" and "MVCO".

- *Shape* : There seems to be a systematic overestimation by MACHI in the lower bands which leads to a lower resemblance with the overall shape of the expected ground spectra. An overestimation, while not as significant for the water spectrum range, exists also for the upper bands, possibly as a consequence of the absorption lines here.

4.3 Future work

Future studies could further explore the convergence behavior of MACHI in order to find an optimal configuration for convergence. A potential strategy would be to apply batching in two (or more) stages: first using a small batch size to rapidly approximate a good solution, and then switching to a larger batch size for more stable convergence at a lower convergence threshold. Ideally, the convergence testing would be done with a finer resolution (more batch sizes) to find the optimal configuration. It is likely that such an improvement will give more accurate retrievals in the lower bands, since the batch selection significantly affected the output in this region (Figure A.5). Additionally, a more careful initialization would be beneficial, e.g. to do DOS per band and with a more sophisticated selection of dark pixels, e.g. using a histogram.

Another potential improvement would be to adopt a more physically detailed atmospheric model that separates the different contributors to transmission and surface reflectance. Explicitly modeling the atmospheric contributors, such as Rayleigh scattering, aerosol scattering, and gaseous absorption as independent components, and splitting surface effects from e.g. water-leaving reflectance would allow a more precise and informative analysis, with a possibility of improving performance. For example, better understanding how MACHI responds to spectrally smooth effects such as Rayleigh scattering could help improve performance in the blue region.

MACHI displays similar performance across the kernels tested in this thesis. However, MACHI could and should be further tested with other kernels, not restricted to derivative kernels, and possibly modified to allow a multi-kernel selection by incorporating multiple convolutions in the penalty function. This would allow correcting for smooth (low-frequency) effects while reducing high-frequency noise.

The validation is also affected by quality issues in the HYPSO images. For example, implementing an effective smile correction for HYPSO-2 would be expected to significantly increase the potential for accurate atmospheric correction using MACHI, especially around the absorption peaks. It is also possible that stray light in the HYPSO data contributes to the bias observed in the lower bands. Additionally, it could be worth noting a possible bias introduced by an uneven distribution of spectral profiles of excluded pixels, as indicated e.g. in Figure 3.2. For future studies, MACHI could be tested on other quality-controlled hyperspectral satellite data, e.g. from PACE OCI [36], to reduce the risk of poor data quality affecting the validation.

Furthermore, strictly enforcing non-negativity on noisy data can introduce a systematic positive bias in spectral regions where the signal is low, as negative fluctuations are truncated while positive ones remain. Therefore, the treatment of constraints in MACHI and the possible trade-off between preserving physicality of data and bias reduction should be further investigated.

MACHI displays a varying performance for different sites, which creates a need to investigate under which conditions MACHI performs better/worse compared to other methods. This will require further testing and validation. Other suggestions for future work include sun glint correction, incorporating spatial effects/smoothing, and optimizing the code (e.g. parallelization) to enable larger batch sizes and more efficient computation.

4.4 Conclusions

This thesis establishes a solid theoretical foundation for the Minimal Atmospheric Compensation for Hyperspectral Images (MACHI) algorithm and presents a new Python implementation capable of handling arbitrary kernels and efficient batch processing. Validation against AERONET-OC ground measurements and POLYMER retrievals shows that MACHI retrieves physically reasonable atmospheric parameters and surface reflectances that are generally comparable to established methods. Notably, MACHI avoids the negative blue-band reflectances that have limited the use of ACOLITE for HYPSO data.

While the method delivers smooth corrected spectra across several test sites, the performance varies to some degree, with a systematic overestimation remaining to be addressed, particularly in the lower spectral bands. The sharp oxygen-A absorption near 760 nm is not efficiently corrected, likely due to uncorrected smile effects in HYPSO-2 imagery. Nevertheless, the results demonstrate the potential of smoothness-based optimization for atmospheric correction and indicate that MACHI can provide fast, unsupervised, and parameter-free processing of hyperspectral data.

This work places the MACHI prototype on a firmer scientific footing and outlines clear paths for further development, including more detailed atmospheric modeling, refined initialization and batching strategies, and application to additional high-quality hyperspectral sensors such as PACE OCI.

4.5 Code availability

The MACHI implementation created for this thesis is written in Python and the algorithm itself depends only on `numpy`. The open-source code repository is available at `https://github.com/edannas/machi` or by correspondence with the author.

References

[1] K. Arai, "Monte Carlo Ray Tracing Based Sensitivity Analysis of the Atmospheric and the Ocean Parameters on Top of the Atmosphere Radiance," *International Journal of Advanced Computer Science and Applications*, vol. 3, no. 12, 2012.

[2] C. Mobley, J. Werdell, B. Franz, Z. Ahmad, and S. Bailey, "Atmospheric Correction for Satellite Ocean Color Radiometry," tech. rep., NASA, 06 2016.

[3] W. Pervez, S. A. Khan, and Valiuddin, "Hyperspectral Hyperion Imagery Analysis and its Application Using Spectral Analysis," in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-3/W2, (Munich, Germany), pp. 25–27, ISPRS, Mar. 2015. PIA15+HRIGI15 – Joint ISPRS Conference 2015.

[4] B.-C. Gao, M. Montes, Z. Ahmad, and C. Davis, "Atmospheric correction algorithm for hyperspectral remote sensing of ocean color from space," *Applied Optics*, vol. 39, pp. 887–896, 02 2000.

[5] P. R. Renosh, D. Doxaran, L. D. Keukelaere, and J. I. Gossn, "Evaluation of Atmospheric Correction Algorithms for Sentinel-2-MSI and Sentinel-3-OLCI in Highly Turbid Estuarine Waters," *Remote Sensing*, vol. 12, no. 8, 2020.

[6] R. Marion, R. Michel, and C. Faye, "Atmospheric correction of hyperspectral data over dark surfaces via simulated annealing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1566–1574, 2006.

[7] A. M. Filippi and T. Kubota, "Introduction of spatial smoothness constraints via linear diffusion for optimization-based hyperspectral coastal ocean remote-sensing inversion," *Journal of Geophysical Research: Oceans*, vol. 113, no. C3, 2008.

[8] J. Xu, L. Rao, F. Schreier, D. S. Efremenko, A. Doicu, and T. Trautmann, "Insight into Construction of Tikhonov-Type Regularization for Atmospheric Retrievals," *Atmosphere*, vol. 11, no. 10, 2020.

[9] B.-C. Gao, M. J. Montes, C. O. Davis, and A. F. Goetz, "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sensing of Environment*, vol. 113, pp. S17–S24, 2009. Imaging Spectroscopy Special Issue.

[10] P. S. Chavez, "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data," *Remote Sensing of Environment*, vol. 24, no. 3, pp. 459–479, 1988.

[11] A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, "MODTRAN® 6: A major upgrade of the MODTRAN® radiative transfer

code," in *2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, 2014.

[12] E. Vermote, D. Tanre, J. Deuze, M. Herman, and J.-J. Morcette, "Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 675–686, 1997.

[13] T. Cooley, G. Anderson, G. Felde, M. Hoke, A. Ratkowski, J. Chetwynd, J. Gardner, S. Adler-Golden, M. Matthew, A. Berk, L. Bernstein, P. Acharya, D. Miller, and P. Lewis, "FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, pp. 1414–1418 vol.3, 2002.

[14] Z. Qu, B. Kindel, and A. Goetz, "The High Accuracy Atmospheric Correction for Hyperspectral Data (HATCH) model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1223–1231, 2003.

[15] M. Matthew, S. Adler-Golden, A. Berk, G. Felde, G. Anderson, D. Gorodetzky, S. Paswaters, and M. Shippert, "Atmospheric correction of spectral imagery: evaluation of the FLAASH algorithm with AVIRIS data," in *Applied Imagery Pattern Recognition Workshop, 2002. Proceedings.*, pp. 157–163, 2002.

[16] L. V. Katkovsky, A. O. Martinov, V. A. Siliuk, D. A. Ivanov, and A. A. Kokhanovsky, "Fast Atmospheric Correction Method for Hyperspectral Data," *Remote Sensing*, vol. 10, no. 11, 2018.

[17] C. Brockmann, R. Doerffer, M. Peters, S. Kerstin, S. Embacher, and A. Ruescas, "Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters," in *Living Planet Symposium*, vol. 740, p. 54, 2016.

[18] F. Steinmetz, P.-Y. Deschamps, and D. Ramon, "Atmospheric correction in presence of sun glint: application to MERIS," *Opt. Express*, vol. 19, pp. 9783–9800, May 2011.

[19] Q. Vanhellemont and K. Ruddick, "Acolite for Sentinel-2: Aquatic applications of MSI imagery," in *Proceedings of the 2016 ESA Living Planet Symposium, Prague, Czech Republic*, vol. 9, 2016.

[20] H. Yang, J. Kong, H. Hu, Y. Du, M. Gao, and F. Chen, "A Review of Remote Sensing for Water Quality Retrieval: Progress and Challenges," *Remote Sensing*, vol. 14, no. 8, 2022.

[21] M. Warren, S. Simis, V. Martinez-Vicente, K. Poser, M. Bresciani, K. Alikas, E. Spyrakos, C. Giardino, and A. Ansper, "Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters," *Remote Sensing of Environment*, vol. 225, pp. 267–289, 2019.

[22] Y. Fan, W. Li, N. Chen, J.-H. Ahn, Y.-J. Park, S. Kratzer, T. Schroeder, J. Ishizaka, R. Chang, and K. Stamnes, "OC-SMART: A machine learning based data analysis platform for satellite ocean color sensors," *Remote Sensing of Environment*, vol. 253, p. 112236, 2021.

[23] L. Bernstein, S. Adler-Golden, R. Sundberg, R. Levine, T. Perkins, A. Berk, A. Ratkowski, G. Felde, and M. Hoke, "Validation of the QUick Atmospheric Correction (QUAC) algorithm for VNIR-SWIR multi- and hyperspectral im-

agery," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5806, 06 2005.

[24] M. Cubero-Castan, D. Constantin, K. Barbieux, V. Nouchi, Y. Akhtman, and B. Merminod, "A new smoothness based strategy for semi-supervised atmospheric correction: Application to the léman-Baïkal campaign," in *2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, 2015.

[25] S. Bakken, M. B. Henriksen, R. Birkeland, D. D. Langer, A. E. Oudijk, S. Berg, Y. Pursley, J. L. Garrett, F. Gran-Jansen, E. Honoré-Livermore, M. E. Grøtte, B. A. Kristiansen, M. Orlandic, P. Gader, A. J. Sørensen, F. Sigernes, G. Johnsen, and T. A. Johansen, "HYPSO-1 CubeSat: First Images and In-Orbit Characterization," *Remote Sensing*, vol. 15, no. 3, 2023.

[26] M. E. Grøtte, R. Birkeland, E. Honoré-Livermore, S. Bakken, J. L. Garrett, E. F. Prentice, F. Sigernes, M. Orlandić, J. T. Gravdahl, and T. A. Johansen, "Ocean Color Hyperspectral Remote Sensing With High Resolution and Low Latency—The HYPSO-1 CubeSat Mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–19, 2022.

[27] NTNU Small Satellite Lab, "https://www.ntnu.edu/smallsat/," 2025. Accessed: 2025-05-31.

[28] E. B. Vigre, "Atmospheric Correction for HYPSO-1 Data," master's thesis, Norwegian University of Science and Technology, Trondheim, NO, June 2023. Available at `https://hdl.handle.net/11250/3100359`.

[29] NTNU Small Satellite Lab, "https://www.ntnu.edu/smallsat/hyva," 2025. Accessed: 2025-06-14.

[30] T. A. Johansen and O. Nordbeck, "FINAL REPORT, SUMMARY: HyVa - Hyperspectral remote sensing of water quality in Norwegian lakes and coasts," tech. rep., Norwegian Space Agency, 2024.

[31] K. Lange, *Block Relaxation*, pp. 171–183. New York, NY: Springer New York, 2013.

[32] J. L. Garrett, "Machi-prototype." `https://github.com/NTNU-SmallSat-Lab/MACHI`, 2025.

[33] A. Romero, C. Penne, M. Henriksen, J. Garrett, A. Alinejad, S. Bakken, and D. Langer, "hypso-package." `https://github.com/NTNU-SmallSat-Lab/hypso-package/tree/main/hypso`, 2025.

[34] G. Zibordi, B. Holben, I. Slutsker, D. Giles, D. D'Alimonte, F. Mélin, J.-F. Berthon, D. Vandemark, H. Feng, G. Schuster, B. Fabbri, and S. Kaitala, "AERONET-OC: A network for the validation of ocean color primary products," *Journal of Atmospheric and Oceanic Technology - J ATMOS OCEAN TECHNOL*, vol. 26, 08 2009.

[35] G. Thuillier, M. Hersé, D. Labs, T. Foujols, W. Peetermans, D. Gillotay, P. C. Simon, and H. Mandel, "The Solar Spectral Irradiance from 200 to 2400 nm as Measured by the SOLSPEC Spectrometer from the Atlas and Eureca Missions," *Solar Physics*, vol. 214, no. 1, pp. 1–22, 2003.

[36] G. Meister, J. J. Knuble, L. H. Chemerys, H. Choi, N. R. Collins, R. E. Eplee, U. Gliese, E. T. Gorman, K. Jepsen, S. Kitchen-McKinley, *et al.*, "Test Results From the Prelaunch Characterization Campaign of the Engineering Test Unit

of the Ocean Color Instrument of NASA's Plankton, Aerosol, Cloud and Ocean Ecosystem (PACE) Mission," *Frontiers in Remote Sensing*, vol. 3, p. 875863, 2022.

A

Convergence analysis (Chesapeake Bay)

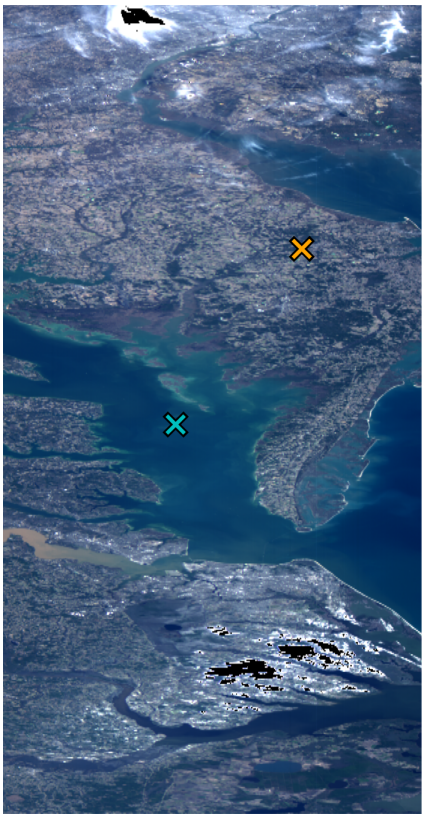


Figure A.1: RGB plot of ToA reflectance from HYPSO-2 capture over Chesapeake Bay on 2025-02-24, with saturation mask (black), and pixel locations marked in turquoise (water) and orange (land).

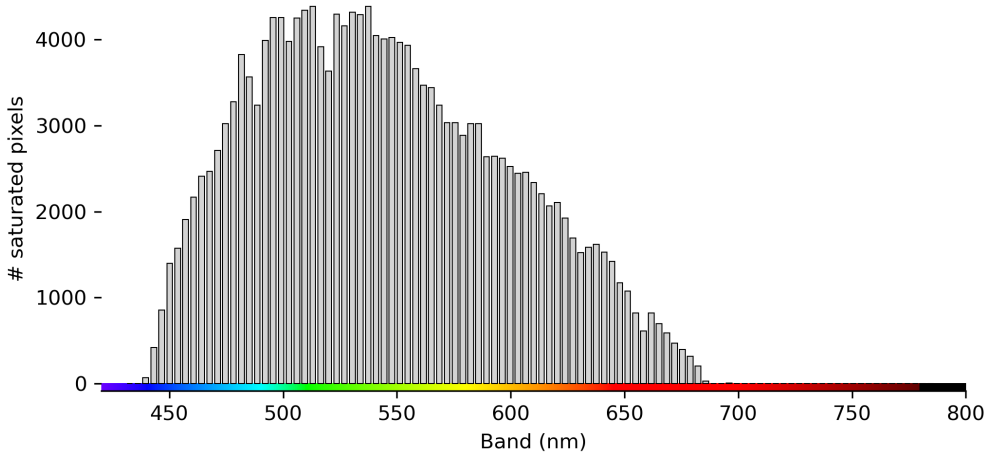


Figure A.2: Number of saturated pixels per band for the HYPSO capture over Chesapeake Bay on 2025-06-20.

Batch size	h_1	h_2	h_3	h_4
100	0.74	1.72	1.12	1.38
1000	4.83	11.84	5.84	8.80
10000	47.86	116.74	57.66	86.83
100000	500.77	671.60	1685.42	773.69

Table A.1: Runtime (in seconds) of MACHI for kernels h1-h4 and all batch sizes, over Chesapeake Bay.

Batch size	h_1	h_2	h_3	h_4
100	526.3	921.4	236.6	144.9
1000	527.4	925.4	238.5	148.2
10000	526.8	925.8	238.4	147.8
100000	526.8	1063.0	238.4	150.2

Table A.2: Total penalty over all pixels after convergence, for kernels h1-h4 and the different batch sizes, over Chesapeake Bay.

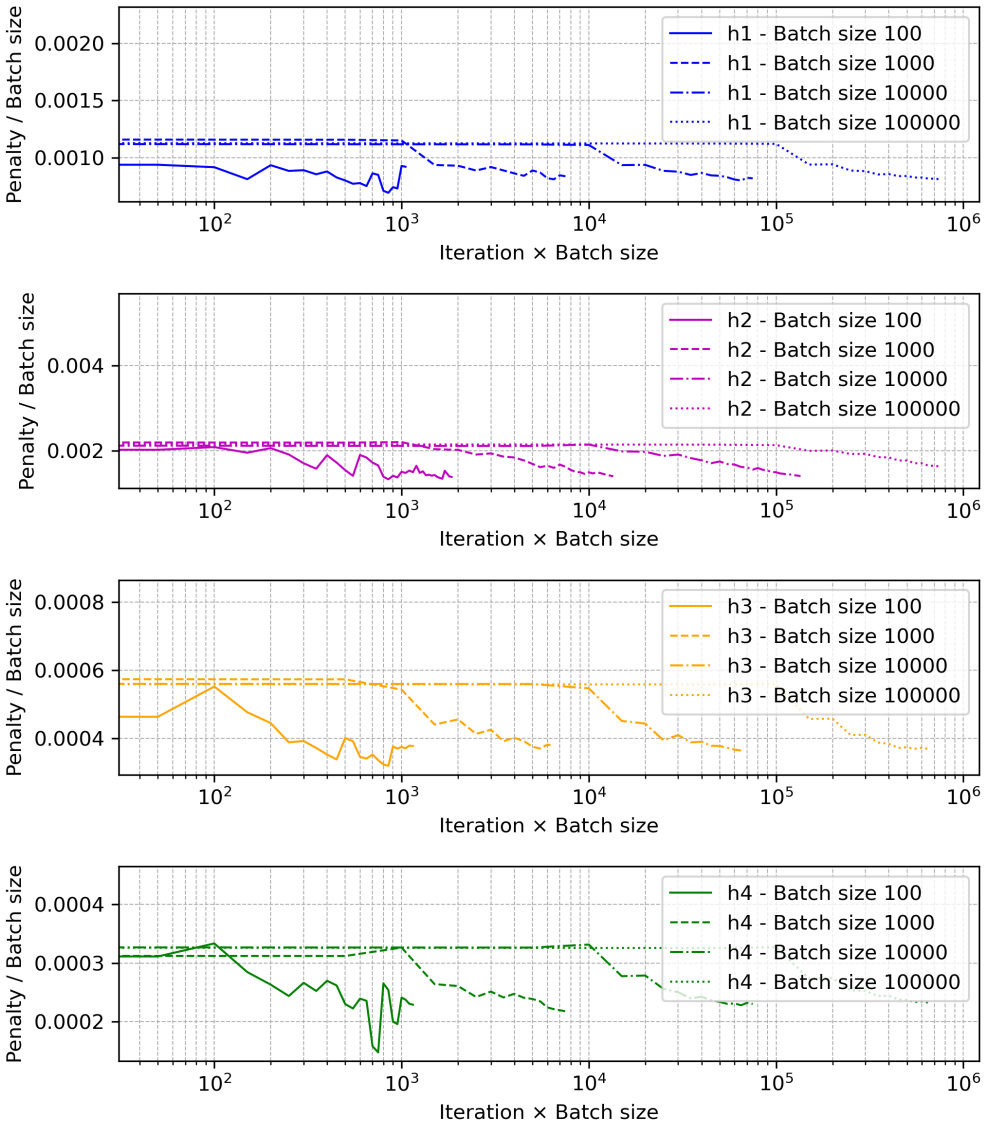


Figure A.3: Scaled penalty (y-axis) by scaled iteration index (x-axis) for kernels h1-h4 and selection of batch sizes, over Chesapeake Bay. Note log-scale on x-axis.

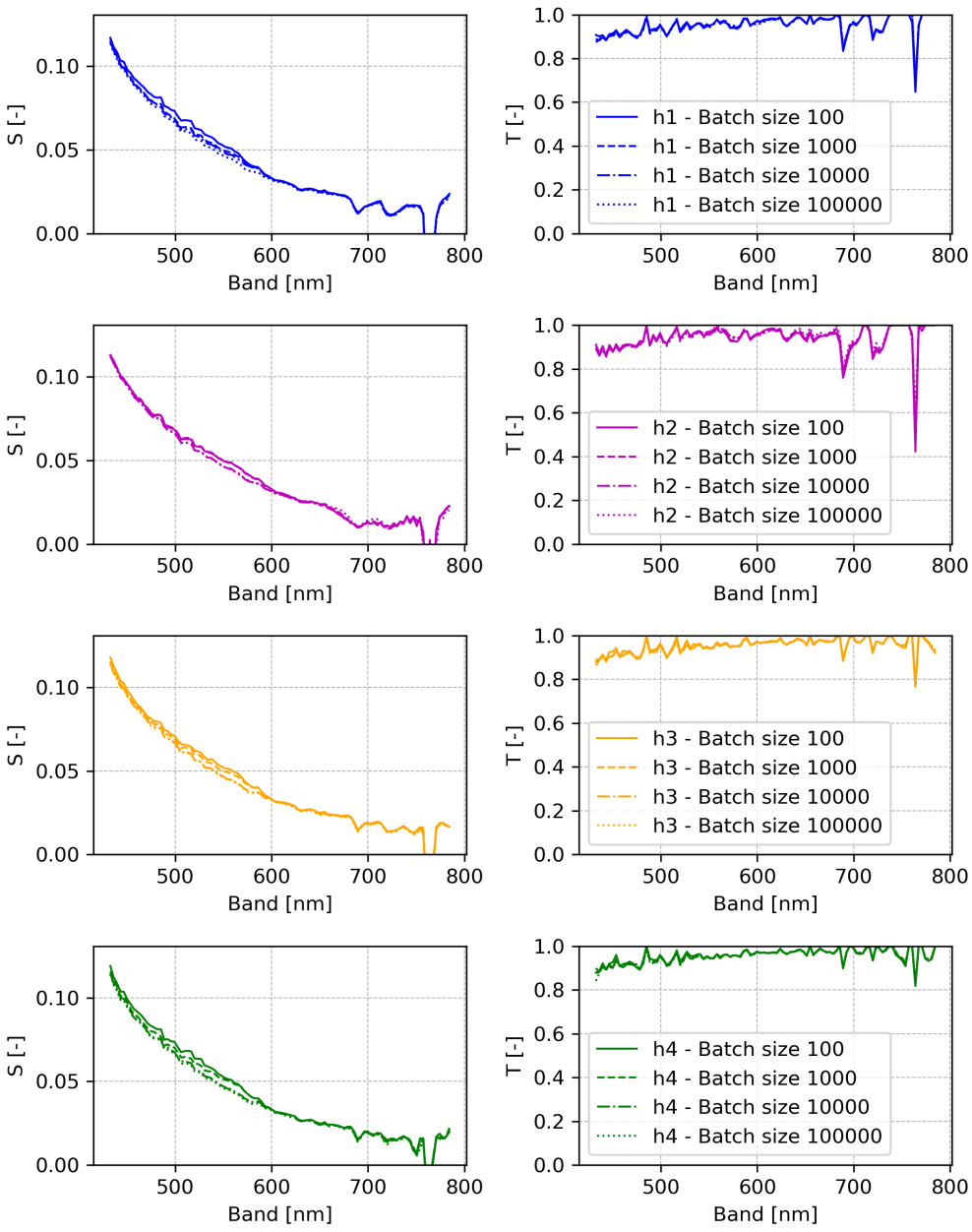


Figure A.4: Final S and T from MACHI over Chesapeake Bay.

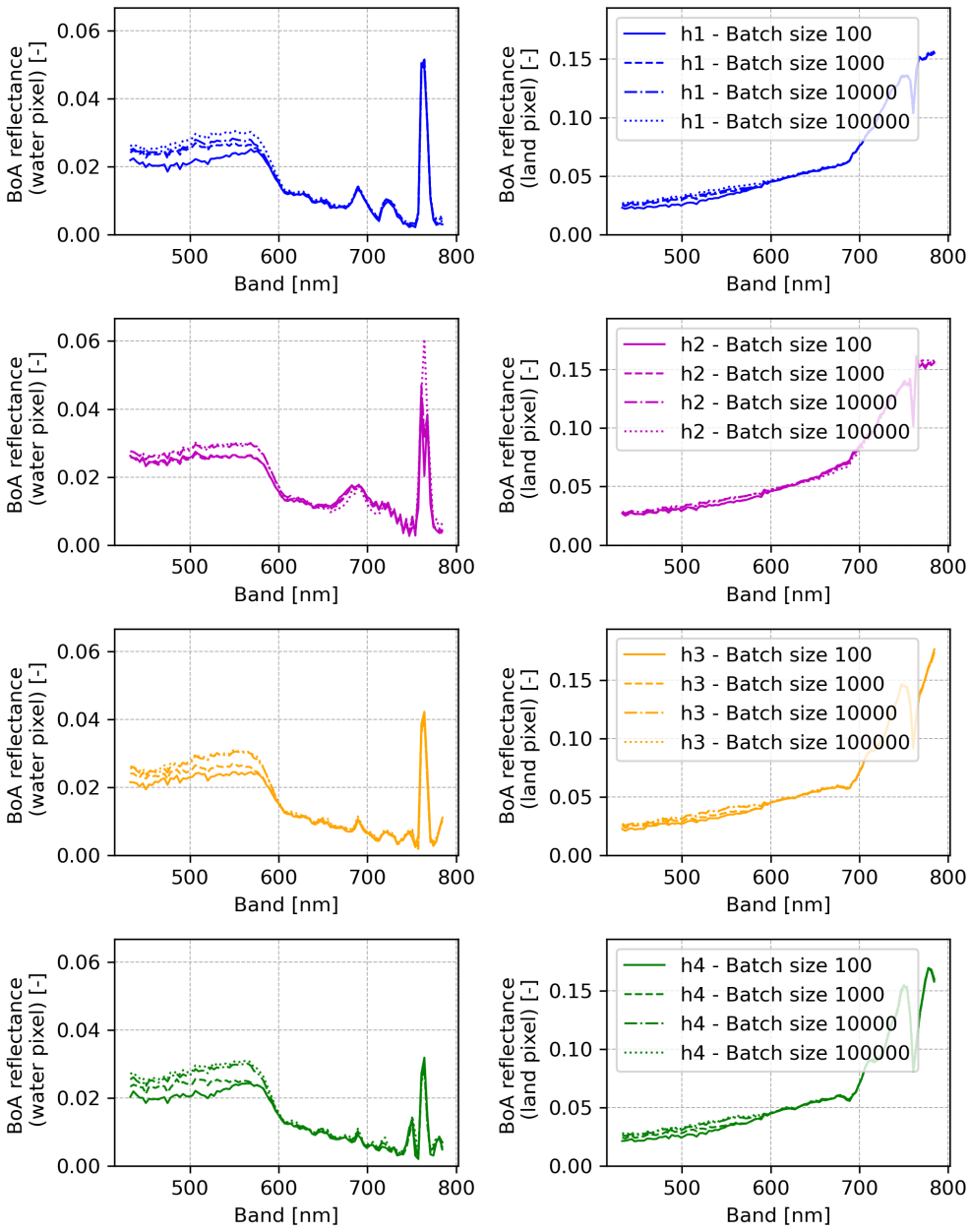
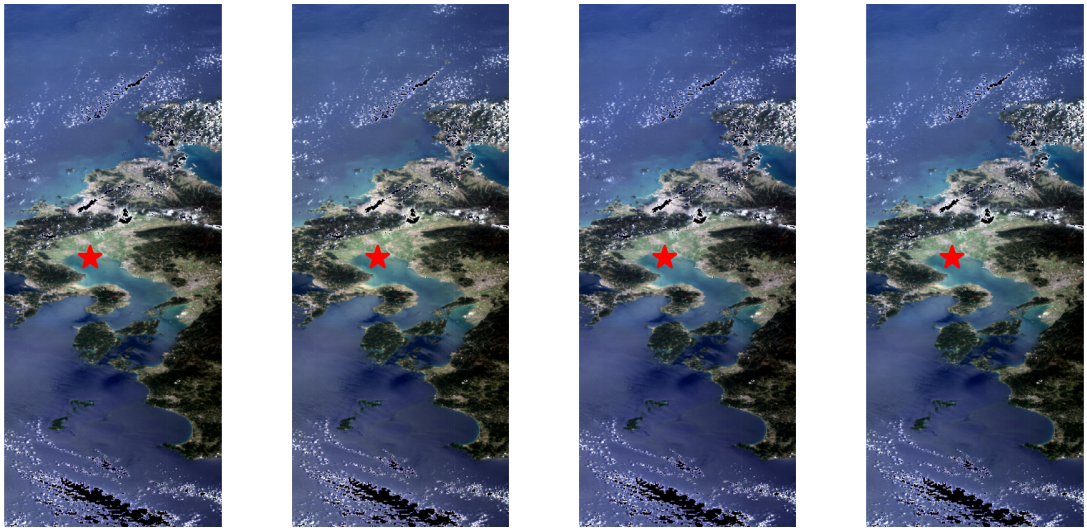


Figure A.5: Retrieved BoA reflectance from MACHI for two pixels (land and water) over Chesapeake Bay.

B

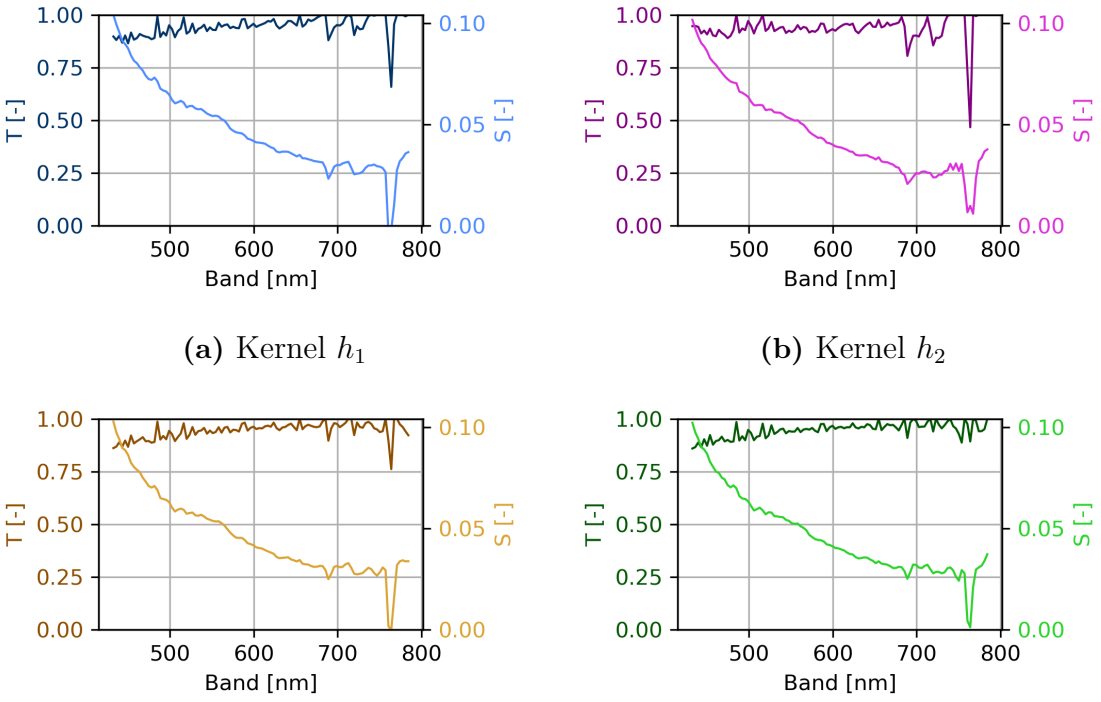
Validation (Ariake Tower 2)

Figure B.1: RGB plot of (uncorrected) ToA reflectance from HYPSO-2 capture over "Ariake Tower 2" validation site in Ariake Sea, Japan, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-03-04.



(a) Kernel h_1 **(b)** Kernel h_2 **(c)** Kernel h_3 **(d)** Kernel h_4

Figure B.2: RGB plot of (corrected) BoA reflectance from HYPSO-2 capture over "Ariake Tower 2" validation site in Ariake Sea, Japan, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-04-04.



(a) Kernel h_1 **(b)** Kernel h_2

(c) Kernel h_3 **(d)** Kernel h_4

Figure B.3: Estimated S and T from MACHI over "Ariake Tower 2" validation site.

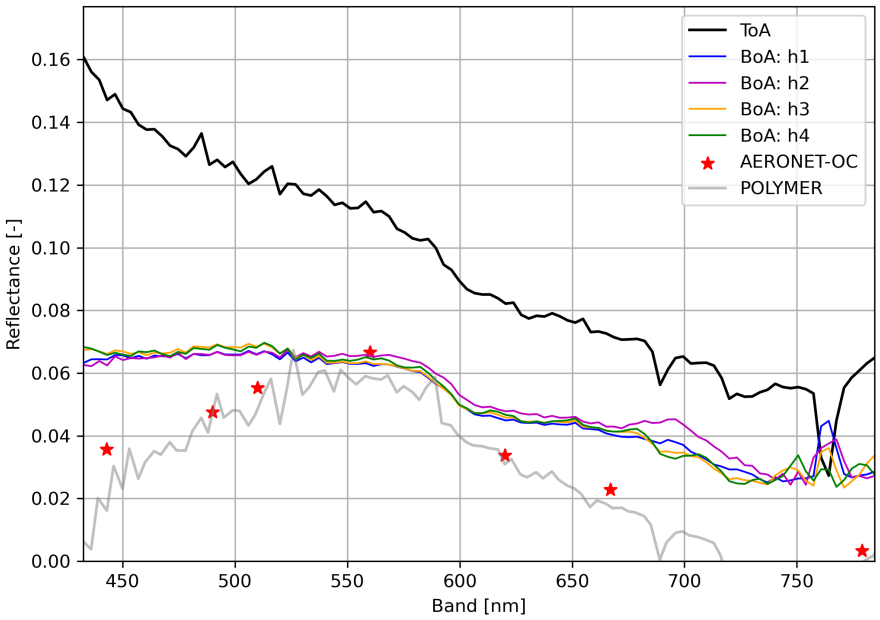


Figure B.4: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "Ariake Tower 2" validation site in Ariake sea, Japan.

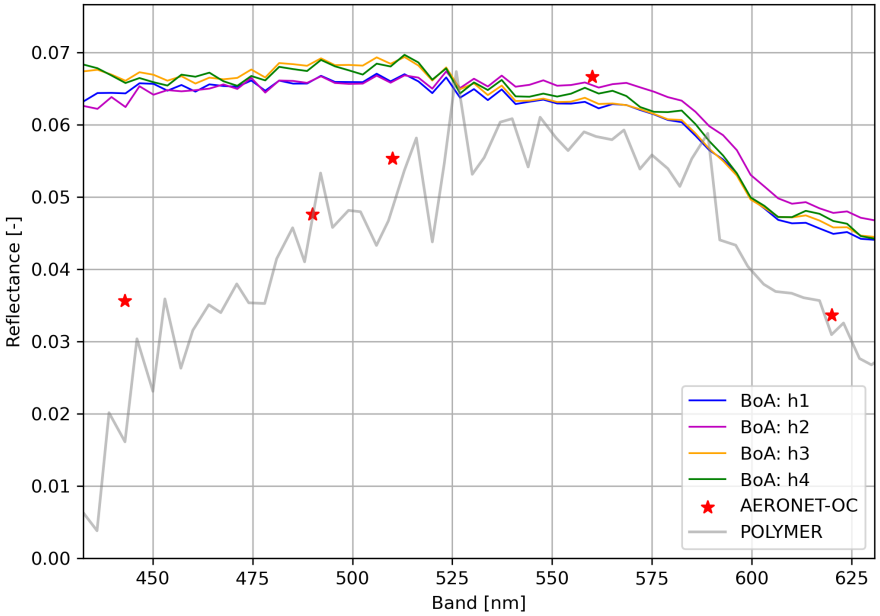


Figure B.5: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement at "Ariake Tower 2" validation site in Ariake sea, Japan (water spectrum range).

C

Validation (Section-7 Platform)

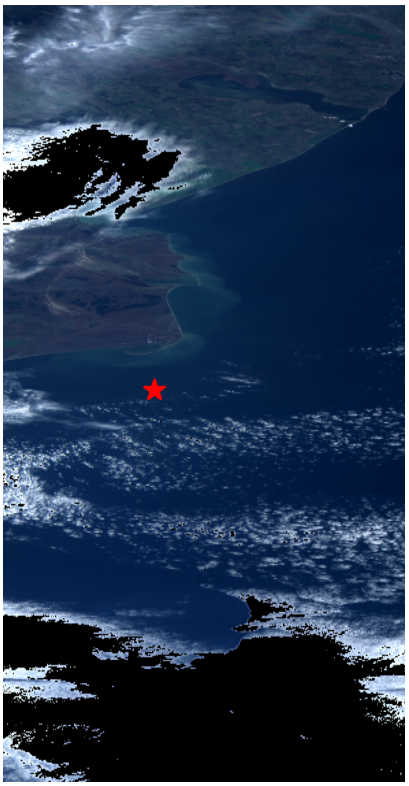
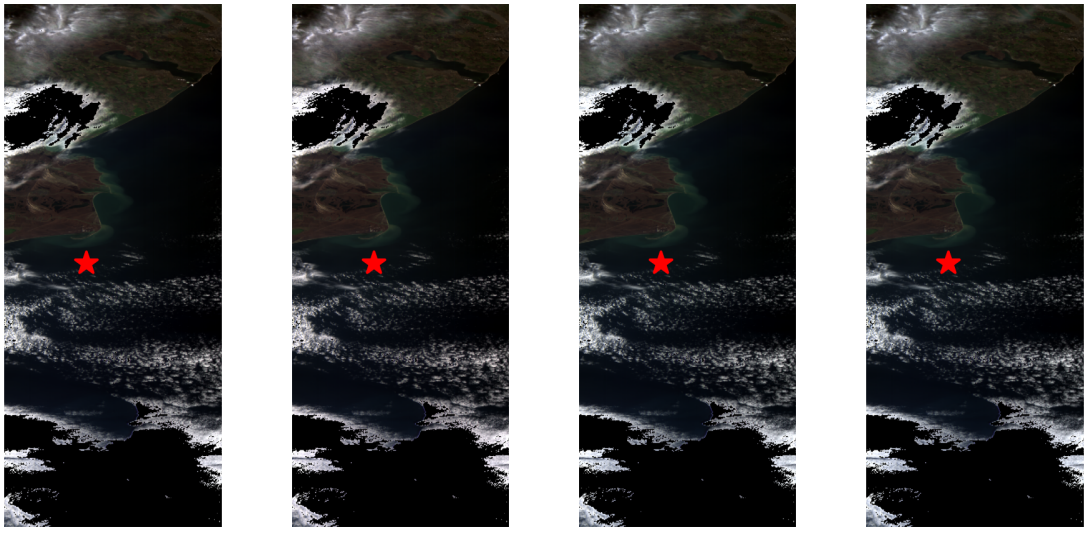
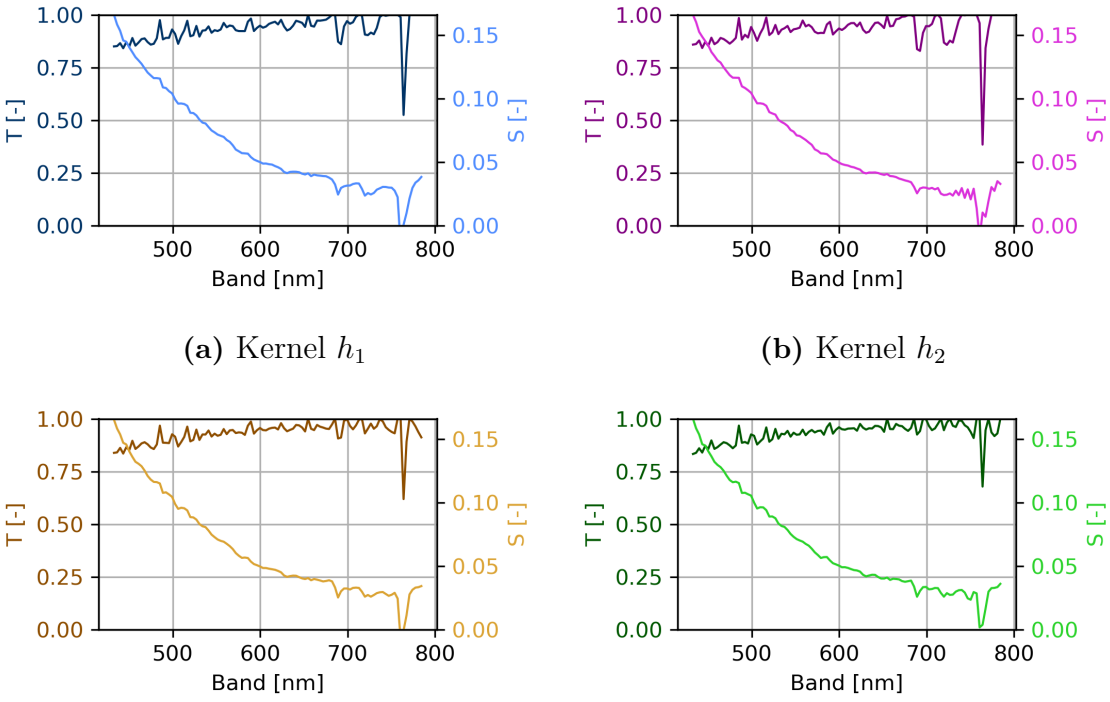


Figure C.1: RGB plot of (uncorrected) ToA reflectance from HYPSO-2 capture over "Section-7 Platform" validation site in Constanta, Romania, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-03-04.



(a) Kernel h_1 **(b)** Kernel h_2 **(c)** Kernel h_3 **(d)** Kernel h_4

Figure C.2: RGB plot of (corrected) BoA reflectance from HYPSO-2 capture over "Section-7 Platform" validation site in Constanta, Romania, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-03-04.



(a) Kernel h_1 **(b)** Kernel h_2

(c) Kernel h_3 **(d)** Kernel h_4

Figure C.3: Estimated S and T from MACHI over "Section-7 Platform" validation site.

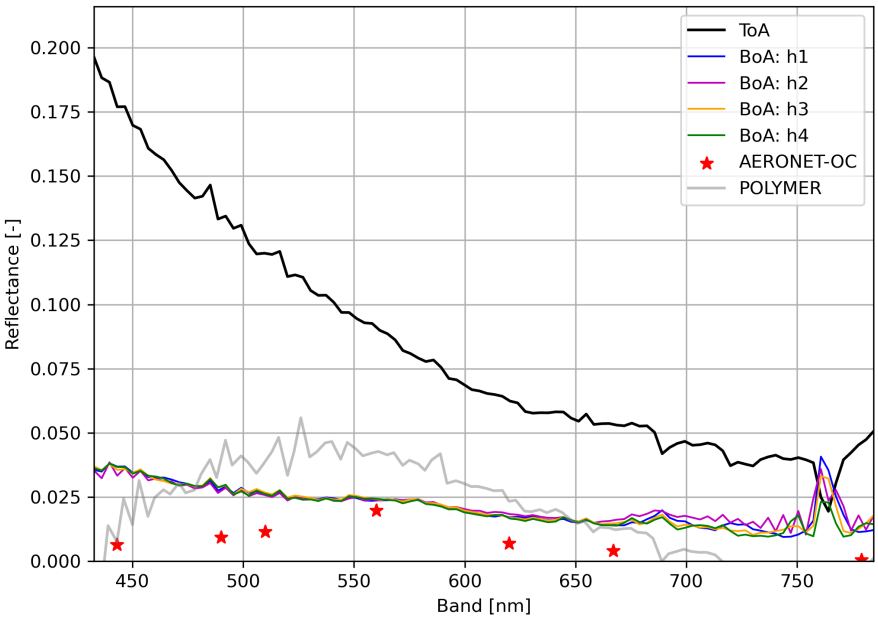


Figure C.4: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "Section-7 Platform" validation site in Constanta, Romania.

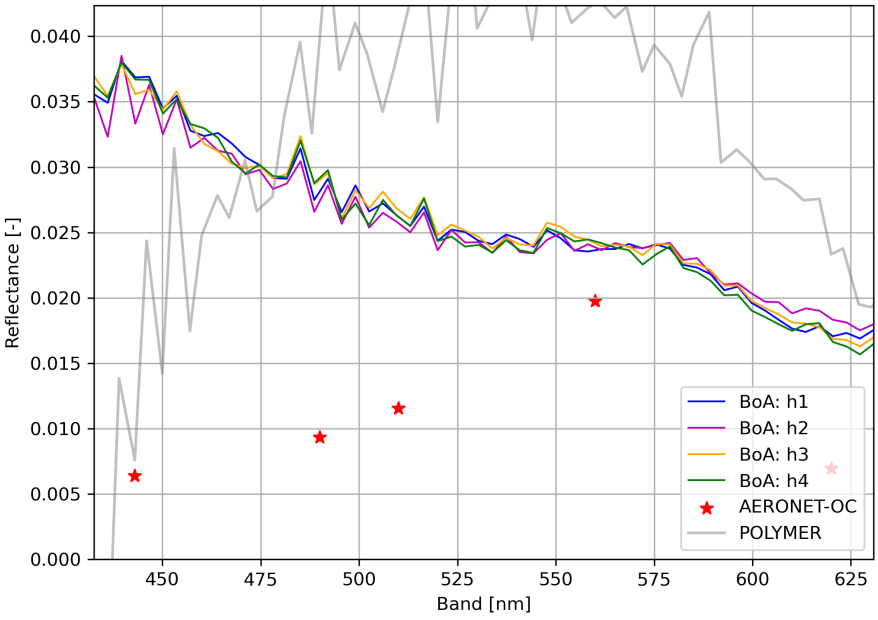


Figure C.5: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement at "Section-7 Platform" validation site in Constanta, Romania, (water spectrum range).

D

Validation (MVCO)

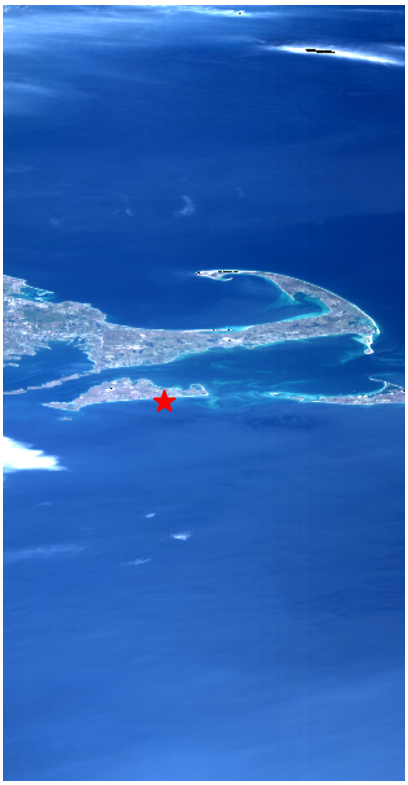
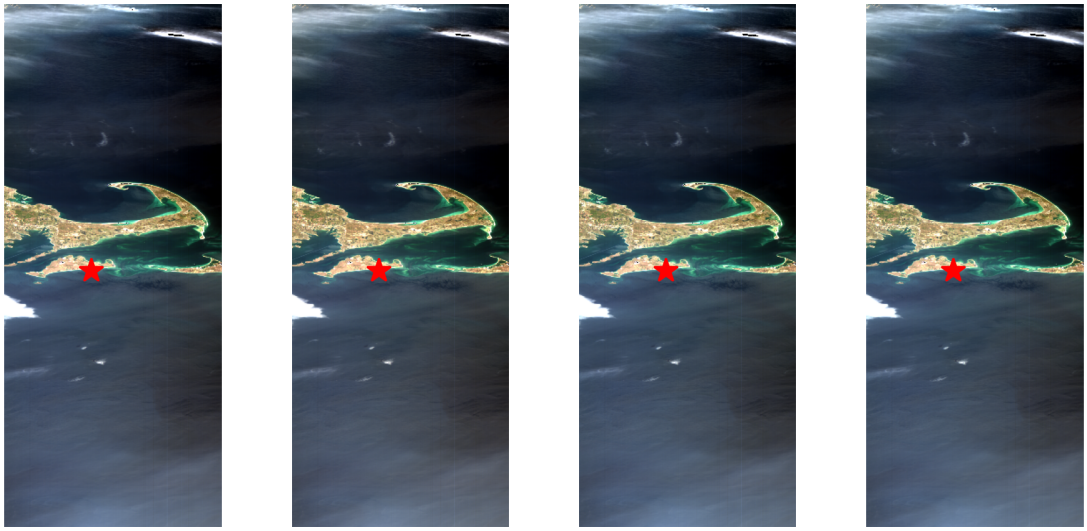
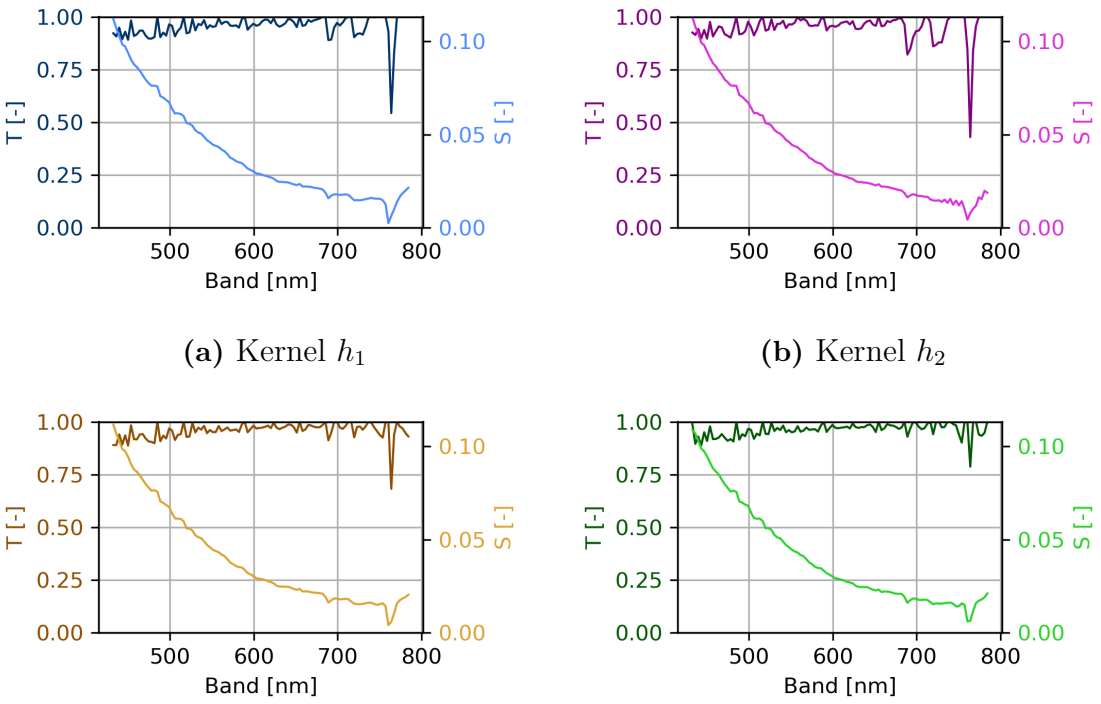


Figure D.1: RGB plot of (uncorrected) ToA reflectance from HYPSO-2 capture over ”MVCO” validation site in Massachusetts, USA, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-04-10.



(a) Kernel h_1 **(b)** Kernel h_2 **(c)** Kernel h_3 **(d)** Kernel h_4

Figure D.2: RGB plot of (corrected) BoA reflectance from HYPSO-2 capture over "MVCO" validation site in Massachusetts, USA, with AERONET-OC validation site marked with a red star, and saturation mask in black, on 2025-04-10.



(a) Kernel h_1 **(b)** Kernel h_2

(c) Kernel h_3 **(d)** Kernel h_4

Figure D.3: Estimated S and T from MACHI over "MVCO" validation site.

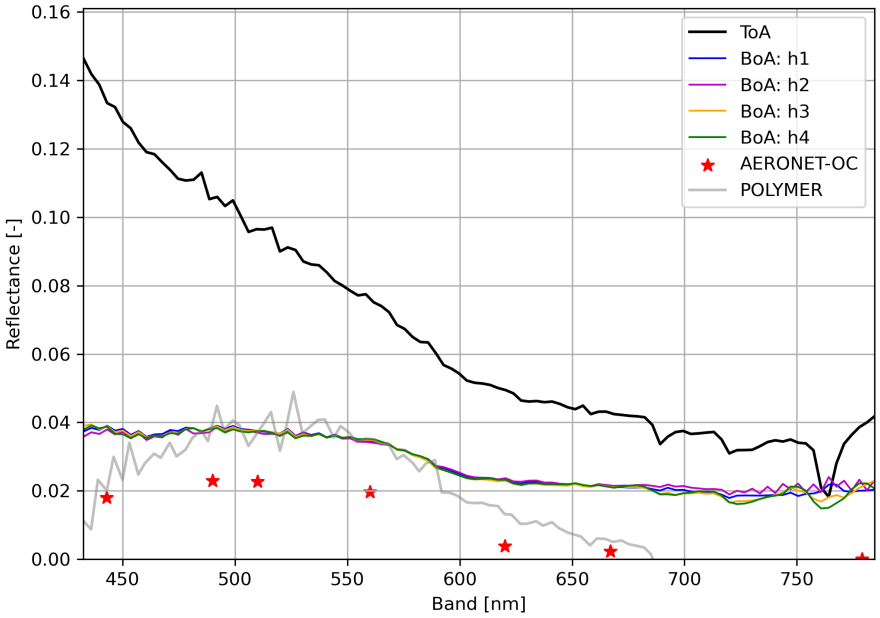


Figure D.4: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement, at "MVCO" validation site in Massachusetts, USA.

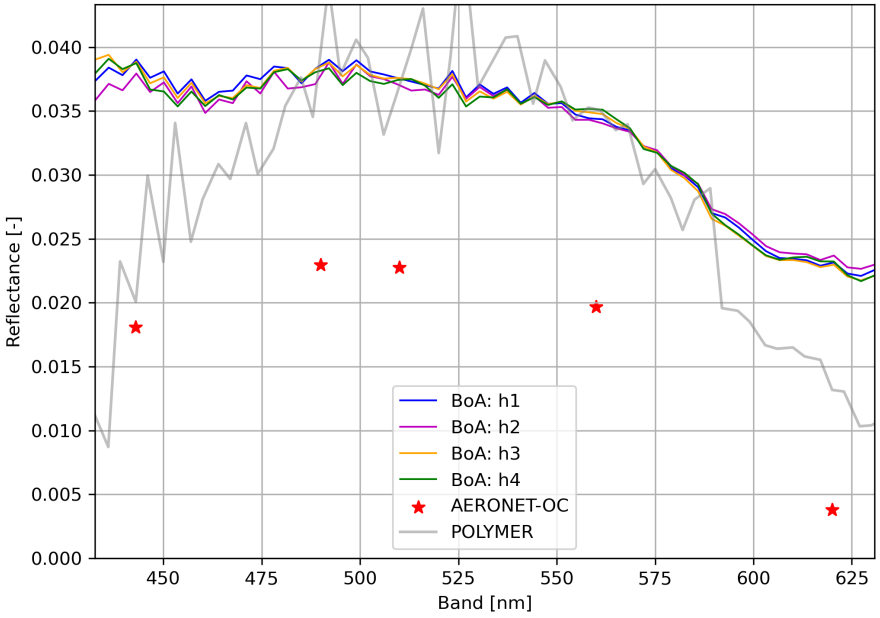


Figure D.5: Retrieved BoA reflectance from MACHI compared to AERONET-OC ground measurement at "MVCO" validation site in Massachusetts, USA (water spectrum range).

DEPARTMENT OF SPACE, EARTH AND ENVIRONMENT
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY