Lab 11: R_{RS} Semi-analytic Inversion   29 June 2023

written and designed by Collin Roesler – updated by Patrick Gray

1.1 Introduction and Lab Goals

The purpose of satellite remote sensing is to derive properties of the earth, ocean, or atmosphere from observations of electromagnetic radiation either emanating from the target of interest (e.g., NIR emission), reflecting from the solar source (e.g., ocean color), or reflecting from a satellite-based source (e.g., RADAR or LIDAR). Approaches range from purely empirical models (e.g., chlorophyll estimation from ocean color) to nearly purely analytic (e.g., altimetry), to somewhere in between semi-analytic (e.g., most derived products). Today we will explore the semi-analytic approach to estimating oceanic IOPs from ocean color remote sensing. The goals are to (1) provide students with improved intuition about how variations in IOPs drive variations in the magnitude and spectral dependence of remote sensing reflectance, and (2) provide a simple Matlab code for semi-analytic inversion to investigate how results vary with range of data sets, models and parameterization.

Recall that the relationship between ocean IOPs and the remote sensing reflectance can be modeled in the forward sense using radiative transfer codes such as Hydrolight®. The inverse approach requires an approximation to the radiative transfer equation; the most common results from the quasi-single scattering approximation (QSSA) which yields the equation:

$$R_{rs}(\lambda) = \sum_i g_i \times \left(\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^i$$

Where i is typically 1 or 2 and sometimes the b_b term in the denominator is removed when $a \gg b_b$. The form of the equation suggests that because the spectral variations in the IOPs determine the spectral variation in the reflectance, an optimization approach can be used to statistically determine the relative magnitudes for each IOP that best determine reflectance, which is a typical regression approach. Recall that the IOPs are additive and thus the spectral variations in the IOPs can be further decomposed into components that are determined by operational considerations or because of similar spectra:

$$a(\lambda) = \sum_i^N a_i(\lambda) = a_w(\lambda) + a_{phyt}(\lambda) + a_{NAP}(\lambda) + a_{CDOM}(\lambda)$$

$$b_b(\lambda) = \sum_i^N b_{b_i}(\lambda) = b_{b_w}(\lambda) + b_{b_p}(\lambda)$$

Beer's Law states that the IOPs are proportional to the concentration of the material doing the absorption or backscattering. Thus, each component IOP can be separated into a scalar (i.e., concentration, C_i) and a vector (e.g., concentration-specific spectrum, $a_i^*(\lambda)$ or $b_{b_i}^*(\lambda)$):

$$\boldsymbol{a(\lambda) = a_w(\lambda) + \sum_i^N C_i \times a_i^*(\lambda)}$$

$$b_b(\lambda) = b_{b_w}(\lambda) + \sum_i^N C_i \times b_{bp_i}^*(\lambda)$$

Lab 11: R_{RS} semi-analytic inversion

The selection of the concentration-specific IOP spectra will depend upon the specific component and the operational consideration for how concentration or magnitude is measured. For example, the chlorophyll concentration is often used to define the concentration-specific phytoplankton absorption spectrum:

$$a_{phyt}(\lambda)(m^{-1}) = C_{phyt} \times a_{phyt}^*(\lambda) = [Chl](\frac{mg}{m^3}) \times a_{phyt}^*(\lambda)(\frac{m^2}{mg})$$

where the chlorophyll-specific phytoplankton absorption, $a_{phyt}^*(\lambda)$, is estimated from measurements (and therefore is one of the sources of "semi" in semi-analytic inversion). Alternatively, the absorption at a reference wavelength, λ_{ref}, is used to define the CDOM concentration-specific absorption spectrum:

$$a_{CDOM}(\lambda)(m^{-1}) = C_{CDOM} \times a_{CDOM}^*(\lambda) = a_{CDOM}(\lambda_{ref})(m^{-1}) \times \exp\left(-S_{CDOM} \times (\lambda - \lambda_{ref})\right)$$

and the spectral dependence is approximated by the analytic exponential function parameterized by the single coefficient for the exponential slope, S_{CDOM}. Comparable expressions are used for non-algal absorption, or a single component that is the sum of CDOM and NAP (i.e., referred to as colored detrital matter, CDM, is that a good descriptor?). Particle backscattering is generally described by a power function, the spectral dependence of which is parameterized by the exponent eta, η:

$$b_{bp}(\lambda)(m^{-1}) = C_{bp} \times b_{bp}^*(\lambda) = b_{bp}(\lambda_{ref})(m^{-1}) \times \left(\frac{\lambda_{ref}}{\lambda}\right)^\eta$$

Substituting the expressions for each component IOP into the QSSA for the reflectance leads to the canonical form of a non-linear regression where the concentration-specific IOP spectra are known eigenvectors (blue terms) and the concentration terms are eigenvalues (red terms) to be estimated by least squares minimization. In the simplest form:

$$R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{bw}(\lambda) + C_p \times b_{b_p}^*(\lambda)}{a_w(\lambda) + C_{phyt} \times a_{phyt}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda) + C_{CDOM} \times a_{CDOM}^*(\lambda)}\right)$$

In this lab, students will explore how well this model estimates the magnitude and spectral dependence of a measured remote sensing reflectance spectrum, and how the robustness of the fit varies with the number and parameterization of eigenvectors. The code is explicitly provided to allow students to easily see how definitions and parameterizations are incorporated into the inversion so that they can easily modify it to replicate other forms of the reflectance equation, other published algorithms, other parametrizations of the eigenvectors, and new forms that they may develop themselves.

1.2 Activities

1.2.1 Semi-analytic inversion by eye using Excel (*Ha, ha! Wait, what? Really? Yes, really.*)

Open up the excel file *Semi_analytic_inversion_Roesler.xlsx*

As a class we will go through the inversion by guessing the eigenvector values. Yes, it is possible to set up a macro to have Excel do this for you but that is not the point of this

exercise. The point is to have students develop intuition about reasonable values for the eigenvalues in the ocean and about how various combinations of the eigenvectors reconstruct reflectance spectra. This approach will be used for three different measured reflectance spectra that represent the blue waters of the oligotrophic open ocean, green waters of the eutrophic coastal ocean, and the yellowish waters of an estuarine system. As a group we will compare eigenvalues between students and consider the reasonableness of the answers.

1.1.2 Code to perform semi-analytic inversions.

Data sets

Four data sets of field-measured irradiance reflectance spectra are provided for the four oceanic regimes of Roesler and Perry (1995). These were measured with a LiCOR hyperspectral irradiance sensor that was deployed to measure the upwelling irradiance spectrum, $E_u(\lambda)$, just beneath the sea surface, and the downwelling irradiance spectrum, $E_d(\lambda)$, just above the sea surface by flipping the sensor. A fifth data set is provided that consists of a set of Hydrolight®-simulated (*thank you Curt*) radiance reflectance spectra, where $L_u(\lambda)$ is modeled for a nadir-viewing 10° solid angle sensor like the HTSRB just beneath the sea surface, and $E_d(\lambda)$ is modeled just beneath the sea surface. The Excel file *Rrs_L_HL_simulation.xls* provides the details of Hydrolight® inputs for each of the 24 spectra (and thus the data for validating the inversion results).

Once the data set is selected, students will need to provide an appropriate f/Q value for the data set. For example, if the data set is irradiance reflectance, f/Q is approximately 0.33; if the data set is radiance reflectance, f/Q is approximately 0.0825.

Python code

The code and data to run this lab are available here: https://github.com/patrickcgray/inverse_modeling. A direct link to a colab running the notebook is here: https://colab.research.google.com/github/patrickcgray/inverse_modeling/blob/main/inverse modelRC.ipynb. All the datasets are available as .csv files which need to be placed in a directory called data/ and are pulled into the notebook from there.

Matlab code

The main program to run is *Rrs_inversion_RP1995.m*. There are two programs that provide phytoplankton absorption spectra; *phyto_avg_abs.m* is a generic chlorophyll-specific absorption spectrum generated from the mean of a large number of spectrophotometrically-measured phytoplankton absorption spectra and *phyto_species_abs.m* provides four species-specific chlorophyll-specific phytoplankton absorption. The function *denan.m* simply removes rows of data for which a NaNs appear; it is used to clean up spectrally-interpolated reflectance arrays for which there are no data.

Eigenvectors

The eigenvectors provided include the measured phytoplankton absorption spectra, as well as

analytic functions for CDOM and NAP absorption and power functions for particle backscattering. Parameters defining the exponential slopes, e.g., S_{CDOM} and S_{NAP}, and power exponent, η, are defined in the code, but students are encouraged to modify these values to explore the sensitivity of the retrievals, or to put in values that better represent their data sets and parameter ranges.

Models

The inversion code provides 8 different inversion models for students to explore. This is by no means exhaustive, but by providing these 8 examples, it should be clear how to modify the code to add other models. For example, the code is provided for the simplest form of the QSSA reflectance equation, the ratio of backscattering to absorption. Think about what lines of code to add to model the form of the equation that is the ratio of backscattering to the sum of absorption plus backscattering; then how to formulate the two-term QSSA. The model numbers and expressions are:

$$(1)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_p \times b_{b_p}^*(\lambda)}{a_w(\lambda) + C_{phyt} \times a_{phyt}^*(\lambda) + C_{CDM} \times a_{CDM}^*(\lambda)} \right)$$

$$(2)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_p \times b_{b_p}^*(\lambda)}{a_w(\lambda) + C_{diat} \times a_{diat}^*(\lambda) + C_{dino} \times a_{dino}^*(\lambda) + C_{CDM} \times a_{CDM}^*(\lambda)} \right)$$

$$(3)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_p \times b_{b_p}^*(\lambda)}{a_w(\lambda) + C_{phyt} \times a_{phyt}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda) + C_{CDOM} \times a_{CDOM}^*(\lambda)} \right)$$

$$(4)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_p \times b_{b_p}^*(\lambda)}{a_w(\lambda) + C_{diat} \times a_{diat}^*(\lambda) + C_{dino} \times a_{dino}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda) + C_{CDOM} \times a_{CDOM}^*(\lambda)} \right)$$

$$(5)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_{psmall} \times b_{b_{psmall}}^*(\lambda) + C_{plarge} \times b_{b_{plarge}}^*(\lambda)}{a_w(\lambda) + C_{phyt} \times a_{phyt}^*(\lambda) + C_{CDM} \times a_{CDM}^*(\lambda)} \right)$$

$$(6)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_{psmall} \times b_{b_{psmall}}^*(\lambda) + C_{plarge} \times b_{b_{plarge}}^*(\lambda)}{a_w(\lambda) + C_{diat} \times a_{diat}^*(\lambda) + C_{dino} \times a_{dino}^*(\lambda) + C_{CDM} \times a_{CDM}^*(\lambda)} \right)$$

$$(7)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_{psmall} \times b_{b_{psmall}}^*(\lambda) + C_{plarge} \times b_{b_{plarge}}^*(\lambda)}{a_w(\lambda) + C_{phyt} \times a_{phyt}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda) + C_{CDOM} \times a_{CDOM}^*(\lambda)} \right)$$

$$(8)\ R_{rs}(\lambda) = \frac{f}{Q} \times \left(\frac{b_{b_w}(\lambda) + C_{psmall} \times b_{b_{psmall}}^*(\lambda) + C_{plarge} \times b_{b_{plarge}}^*(\lambda)}{a_w(\lambda) + C_{diat} \times a_{diat}^*(\lambda) + C_{dino} \times a_{dino}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda) + C_{CDOM} \times a_{CDOM}^*(\lambda)} \right)$$

$$(9)\ a_p(\lambda) = C_{diat} \times a_{diat}^*(\lambda) + C_{dino} \times a_{dino}^*(\lambda) + C_{NAP} \times a_{NAP}^*(\lambda)$$

The specific eigenvectors and parameters that are used to define them are summarized in Table 1.

Table 1. Details of the 8 models in the inversion code *Rrs_inversion_RP1995.m* and the values of the parameters defining the eigenvector equations.

Model	b_{bp}	b_{bp_s}, b_{bp_l}	a_{cdm}	a_{cdom}, a_{nap}	a_{phyt}	a_{diatom}, a_{dino}
eigenvector function	1 power function	2 power functions	1 exponential	2 exponentials	1 spectrum	2 spectra

Lab 11: R$_{RS}$ semi-analytic inversion

parameters	$\eta=0.75$	$\eta_{small}=1$ $\eta_{small}=0$	$S_{CDM}=-0.0145$	$S_{NAP}=-0,01$ $S_{CDOM}=-0.02$	generic	diatom, dino
1	X		X		X	
2	X		X			X
3	X			X	X	
4	X			X		X
5		X	X		X	
6		X	X			X
7		X		X	X	
8		X		X		X
9				X (nap only)		X

Inversion exercises

1. Spectral resolution: Select one of the data sets. Select one model. Run in both full spectral resolution and at the MODIS wavelengths. How does spectral resolution impact the retrieved coefficients and the fit between modeled and measured R?

2. Model selection: Select one of the data sets. Select full spectral resolution. How much do the inversion results vary between the 8 models? Does the fit between measured and modeled R change as you add more eigenvectors? Which eigenvectors induce the most change? Can you ascribe any pattern to the inclusion of one or more eigenvectors?

3. Eigenvector parameterization: Select one data set, full spectral resolution, one model. Keeping all else constant, modify the eigenvector parameters one by one (exponential slopes of CDOM and NAP, power exponent(s) for particle backscattering, temperature and salinity for water IOPs). How much variation in retrieved eigenvalues do you find as you modify the eigenvector parameterization? How does the fit between measured and modeled R change? To which eigenvector parameterization is the model most sensitive? Are the results the same for all data sets? Why or why not?

4. There is a lot of interest in using this approach to solve for multiple pigment-based phytoplankton groups. How does the fit between the measured and modeled reflectance spectra change as the number of phytoplankton eigenvectors increases? Look at the differences between measured and modeled reflectance; which wavebands exhibit the most difference? Are they associated with known pigment absorption peaks? How can this be further exploited?

5. How does inverting multiple phytoplankton groups directly from ap spectra compare to inverting from Rrs? You can explore this with your own data to see how different the output is with

6. There is a lot of interest in using this approach to solve for spectral slope of bbp. Some solve for eta, the power function exponent, while models 5-8 provide two endmember power functions meant to represent communities of small particles (backscattering slope of -1) compared to large particles (backscattering slope of 0). The retrieved eigenvalues

Lab 11: R$_{RS}$ semi-analytic inversion

for these two endmembers can be combined to allow for infinite solutions of the particle backscattering slope between -1 and 0. Think about the stability of a least-squares minimization will respond to these two approaches.

7. Validation is the most important part of algorithm development.

    a. For each model, what measurements do you need to make, along with the radiometric ones to derive the measured $R(\lambda)$, to provide a full validation data set? What considerations are you going to have to make to ensure the most accurate validation schema? What approximations will you have to make?

    b. As shown in the lecture, there was validation data associated with each of the spectra in Roesler and Perry (1995). For your validation exercise, compare the retrieved ($eigenvalue \times eigenvector(\lambda)$) (component IOP spectra) with those input to the Hydrolight® run. Divide and conquer to determine the conditions for which the inversion results yield strong validation, and for which are they weak.

Lab 11: R$_{RS}$ semi-analytic inversion