Remote sensing of ocean color

Bruno Pelletier and Robert Frouin

Department of Mathematics, University Rennes II, France

Scripps Institution of Oceanography, University of California at San Diego

Journées STAR, Rennes, Octobre 22, 2009.

Outline

1. Introduction
2. Basics of radiative transfer
3. Remote sensing of ocean color with function fields
4. Estimation of the aerosol vertical distribution with maximum entropy regularization

1. Introduction

2. Basics of radiative transfer

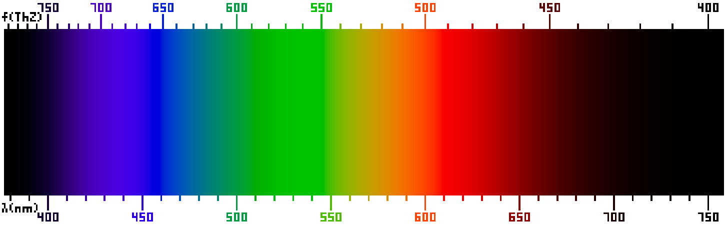
3. Remote sensing of ocean color with function fields

4. Estimation of the aerosol vertical distribution with maximum entropy regularization

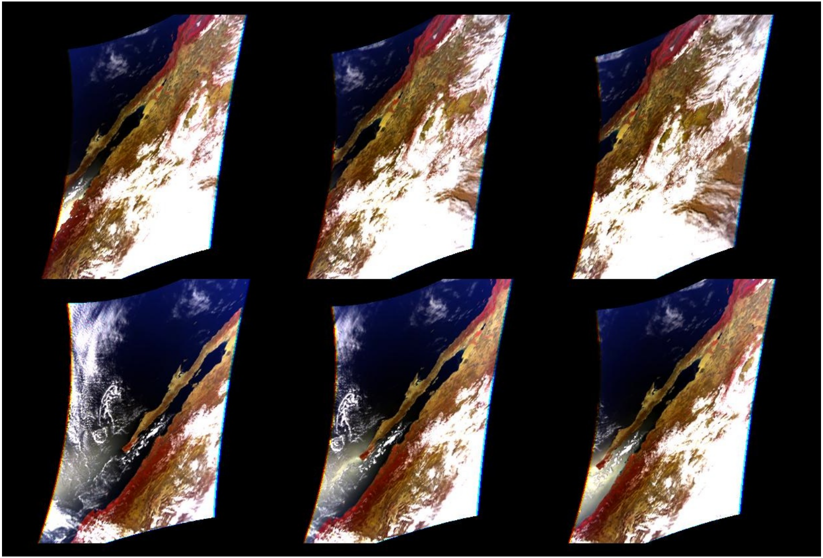
Why ocean color ?

Sunlight

- The spectral characteristics of the sunlight are altered by the medium it propagates through.
- Optical measurements in selected spectral bands (in the visible and near infra-red) give access to optically active parameters, either oceanic or atmospheric.

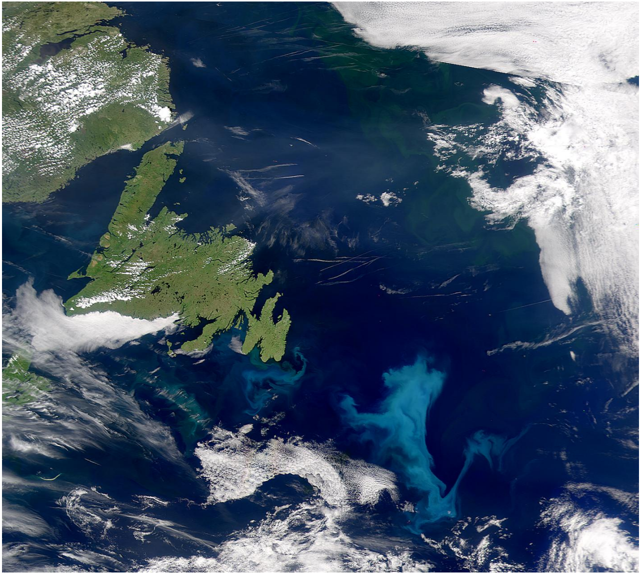


Total radiance over California, 1 orbit, Polder

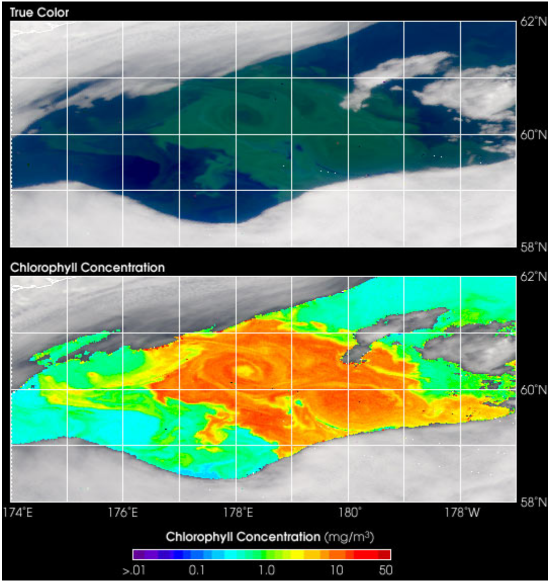


Red: 865 nm ; Green : 670 nmn ; Blue 443 nm.

Phytoplankton bloom in the Bering Sea



Phytoplankton bloom in the Bering Sea



Past, present, and future missions

The beginning

- Began in 1978 with the Coastal Zone Color Scanner (CZCS, NASA): 6 bands, 825 m resolution.
- CZCS: initially an experimental mission... continued up to 1986.

Past and actual missions

- POLDER-2(2003, CNES): 9 bands, resolution 6,000 m
- SeaWiFS (1997, NASA): 8 bands, resolution 1,100 m
- MERIS (2002, ESA): 15 bands, resolution 300/1,200 m
- MODIS-Aqua (1999, NASA): 36 bands, resolution 1,000m

Scheduled sensors

- OLCI (2012, ESA): 15 bands, resolution 300/1,200 m
- VIIRS (2010, NOAA): 22 bands, resolution 370/740 m.

Motivations for ocean color data

Ocean carbon cycle

- Produce fields of chlorophyll pigments, an index of phytoplankton biomass.
- Quantify the carbon flux for climate research and at a global scale.

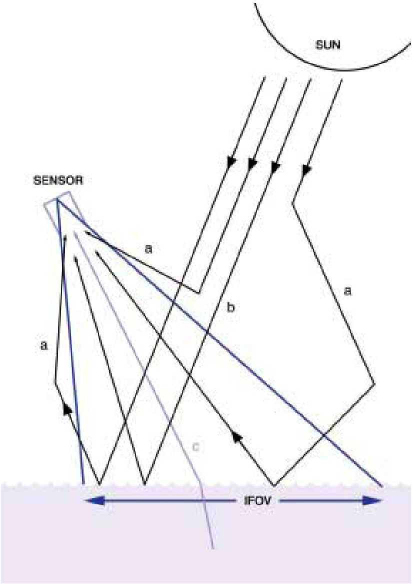
Oceanic mixed layer

- The mixed layer (from surface to $25 \sim 200$m depth) is the link between the atmosphere and the deep ocean.
- A key component in climate studies, biological productivity, and marine pollution.

Coastal zone management

- Monitoring harmful algal blooms.
- Monitoring coastal pollution.
- Fisheries management.

Pathways of light: Atmosphere



Major processes in the atmosphere

In the atmosphere, the solar radiation is subject to absorption and scattering by:

- Air molecules, e.g., oxygen, water, ozone.
- Aerosols, i.e. suspended liquid and particulate matter including smokes, water and hydrogen sulfate droplets, dust, ashes, pollen, spores,...

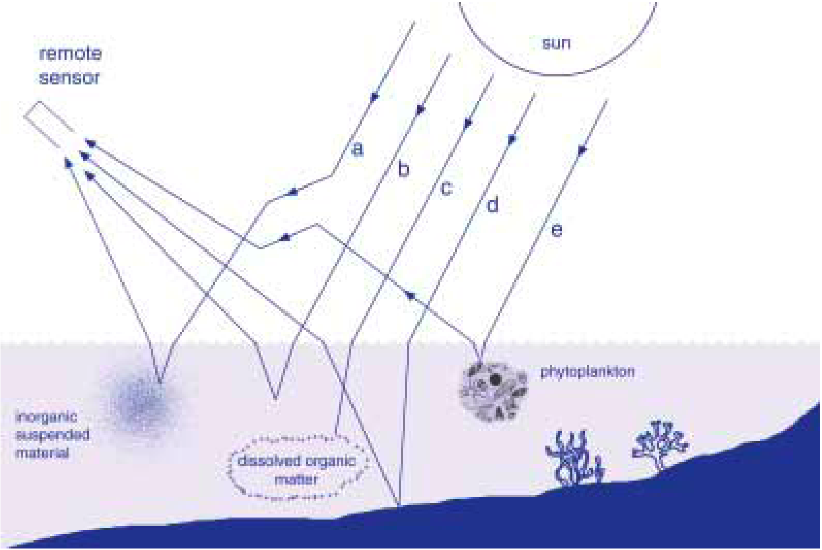
Main forms of scattering

- Particules with diameter $\ll \lambda$: Rayleigh scattering.
- Particules with diameter $\sim \lambda$: scattering varies in a complex fashion. For spherical particules, scattering may be described by Mie theory.

Rayleigh scattering

- Wavelength dependence $\sim \lambda^{-4}$
- More scattering in the blue than in the red
- The eye is more sensitive to the blue than to the violet.

Pathways of light: Ocean



Optical properties of oceans

Three main components, in addition to pure water

- Phytoplankton: microsopic algae with chlorohyll pigments.
- Suspended inorganic materials, sediments.
- Yellow substances: coloured, dissolved, organic substances.

Types of waters

- Case I waters: optically dominated by the phytoplankton. Typical of open oceans (away from coasts).
- Case II waters: optically complex waters, with several optically active components (phytoplankton, sediments, yellow substances). Typical of estuaries and shorelines.

Case I and Case II waters

Decomposition of the satellite signal

A simplified model (neglecting absorption phenomenons) of the top-of-atmosphere radiance is :

$$L_{toa} = L_{atm} + \left[L_G + L_F + L_w \right]$$

- L_{toa} : measured signal ;
- L_{atm} : atmospheric contribution ;
- L_G : glitter contribution, i.e., specular reflection by the surface;
- L_F : foam contribution, i.e., diffuse reflection by the surface;
- L_w : water contribution, i.e., photons which have penetrated the water body.

Remark: generally expressed in terms of reflectance $R(\lambda) = \frac{L_u(\lambda)}{L_d(\lambda)}$.

Reference inversion algorithms: principles

Main steps :

- Atmospheric correction : yields L_w;
- Inversion of L_w towards oceanic parameters.

Atmospheric correction

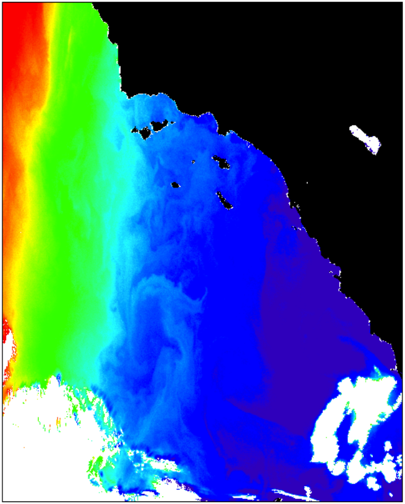
- Central assumption: black ocean in the near infra-red.
- Extrapolation of L_{atm} to the lower part of the spectrum.

Inversion of L_w

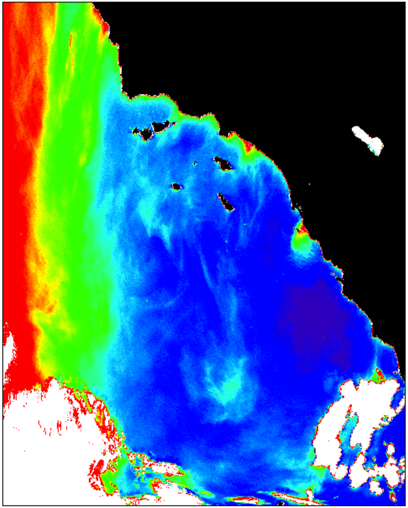
- Via bio-optical models;
  - Case I: total chlorophyll pigments content;
  - Case II: subject to intensive research.

Example: black ocean assumption

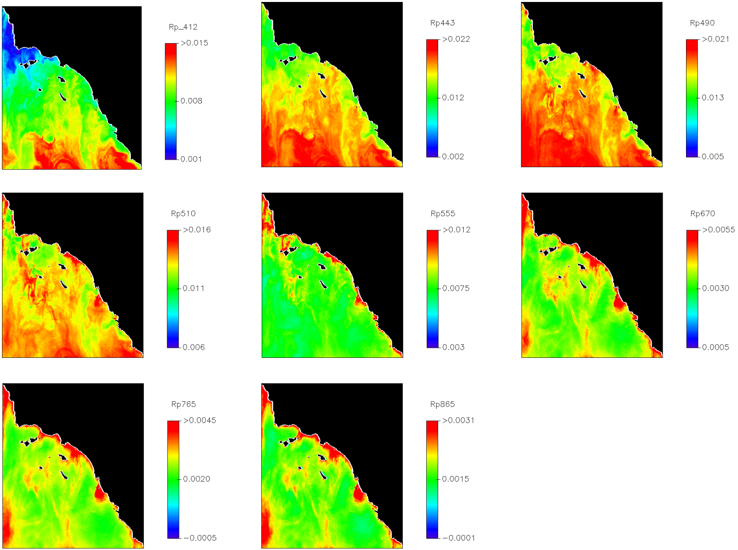
412 nm
Scale: $0.080 - 0.132$



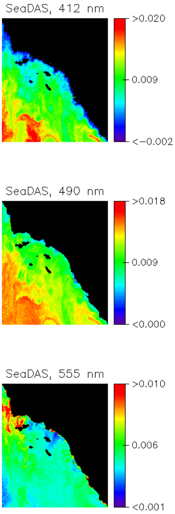
865 nm
Scale: $0.005 - 0.011$



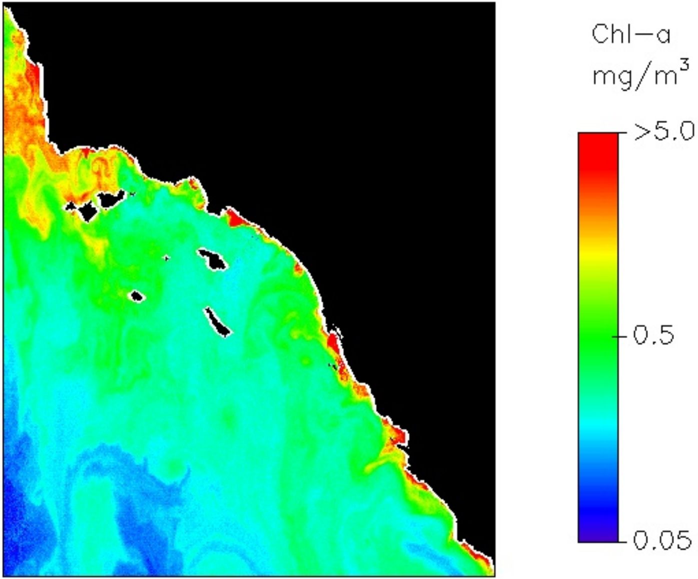
Example: Southern California, R_{toa} (1/3)



Example: Southern California, R_w (2/3)



Example: Southern California, [Chl-a] (3/3)



Issues and difficulties in remote sensing of ocean color

- The contribution of the ocean represents less than 10% of the measured signal.
- Calibration of sensors
- Measurement noise.
- Uncertainties in radiation transfer modeling:
  - aerosol properties (e.g., type, vertical distribution);
  - Case II waters modeling.
- Actual errors on chlorophyll-a concentration are often above the 35% requirement for biological applications.
- Computer time.

Context

We observe

$$y = \Phi_t(x_w, x_a) + \varepsilon,$$

where

- y: top-of-atmosphere reflectance;
- x_w: oceanic parameters, e.g., [Chl-a], to be recovered;
- x_a: atmospheric parameters, unknown;
- t: vector of 3 angular parameters, observed;
- ε: noise term.

Approaches to nonlinear inverse problems:

- Tikhonov regularization;
- Bayesian framework.

Inverse problems

$$y^{obs} = \mathcal{K}x + \varepsilon.$$

Regularization

- Minimize $\|y^{obs} - \mathcal{K}x\|^2 + \alpha\mathcal{R}(x)$.
- α: regularization parameter; $\mathcal{R}(x)$: regularization functional.
- Possible preliminary step: nonparametric reconstruction of y from y^{obs} [Bissantz *et al*, 2005].

Bayesian framework

- Find the distribution of $X|Y$ given the prior P_X and noise distribution P_ε.
- Possible solution: $\mathbb{E}\left[X|Y = y^{obs}\right]$ with quadratic loss.

Function fields and varying coefficient models

In ocean color remote sensing, we have

$$Y = \Phi_t(X) + \varepsilon,$$

with $t \in T$: vector of three angular variables, observed and deterministic.

Idea: "attach" a regression model at each $t \in T$, and demand that the attachment is continuous:

$$\zeta : T \longrightarrow \mathcal{M} \subset \mathbb{R}^{\mathcal{X}} \quad ; \quad \left(\mathbb{R}^{\mathcal{X}}\right)^T \cong \mathbb{R}^{T \times \mathcal{X}}.$$

Varying coefficients models [Hastie and Tibshirani, 1993]:

$$Y = \alpha_1(t)X_1 + \cdots + \alpha_p(t)X_p + \varepsilon.$$

Density results

Theorem

Let X locally compact Hausdorff, T compact, metric Hausdorff
Let $\mathcal{M}$ dense subset of $\mathbb{R}^X$.
Then $\mathcal{M}^T$ is dense in $\left(\mathbb{R}^X\right)^T$.

- In particular, if $\mathcal{G}$ is fundamental in $\mathbb{R}^X$, then $\left(\operatorname{span}\mathcal{G}\right)^T$ is dense in $\left(\mathbb{R}^X\right)^T$.
- If $\mathcal{M} := \operatorname{span}\mathcal{G}$ is a linear space: $\mathcal{M}^T$ identified with $\mathcal{C}(T) \otimes \mathcal{G}$, e.g., polynomials.

Approximation by ridge functions

A ridge function on $\mathbb{R}^d$ is a function of the type $h(\langle a, x\rangle)$, where $h : \mathbb{R} \to \mathbb{R}$, and where $a \in \mathbb{R}^d$.

Approximation by ridge functions:

$$\begin{aligned} \mathcal{R}_k(A) &= \left\{ \sum_{i=1}^{k} c_i h_i\left(\langle a_i, x\rangle\right), \quad c_i \in \mathbb{R}, \quad a_i \in A \subset \mathbb{R}^d, \quad h_i \in \mathcal{C}(\mathbb{R}) \right\} \\ \mathcal{R}(A) &= \bigcup_k \mathcal{R}_k(A). \end{aligned}$$

Special case, for fixed $h : \mathbb{R} \to \mathbb{R}$:

$$\begin{aligned} \mathcal{M}_k(A) &= \left\{ \sum_{i=1}^{k} c_i h\left(\langle a_i, x\rangle\right), \quad c_i \in \mathbb{R}, \quad a_i \in A \subset \mathbb{R}^d \right\} \\ \mathcal{M}(A) &= \bigcup_n \mathcal{M}_k(A). \end{aligned}$$

Remark: $\mathcal{R}$ and M are not linear spaces, and are dense in $\mathcal{C}(\mathbb{R}^d)$ (Lin and Pinkus, 1993; Barron, 1993).

Parametrization of continuous ridge function fields

$$\mathcal{M}_k = \left\{ \sum_{i=1}^{k} c_i h\left(\langle a_i, x + b_i \rangle\right), \quad b_i \in \mathbb{R}, \quad c_i \in \mathbb{R}, \quad a_i \in \mathbb{R}^d \right\}$$

$$\Theta_k = \{(c_i, a_i, b_i)\} := \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$$

$j_k : \Theta_k \to \mathcal{M}_k$: continuous surjection.

Theorem

Let $\mathcal{T}_a \subset \mathcal{C}(T, \mathbb{R}^d)$, and let $\mathcal{T}_b$ and $\mathcal{T}_c$ be subsets of $\mathcal{C}(T)$. Let $\mathcal{R}(\mathcal{T}_a, \mathcal{T}_b, \mathcal{T}_c)$ be the set of ridge function fields $\zeta : T \to \mathcal{M}$ such that

$$\zeta_*(x, t) = \sum_{i=1}^{k} c_i(t) h\left(\langle a_i(t), x \rangle + b_i(t)\right),$$

for some k, $c_i \in \mathcal{T}_c$, $a_i \in \mathcal{T}_a$, and $b_i \in \mathcal{T}_b$. For $\mathcal{R}(\mathcal{T}_a, \mathcal{T}_b, \mathcal{T}_c)$ to be dense in $\left(\mathcal{C}(T)\right)^T$, it is sufficient that

- i) *$\mathcal{T}_a$ and $\mathcal{T}_c$ contain the constant functions;*
- ii) *$\mathcal{T}_b$ contain the affine functions.*

Nonparametric regression

- $(X_1, T_1, Y_1), \ldots, (X_n, T_n, Y_n)$ and i.i.d. random sample $\sim \mu$ on $\mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}$.
- Basis expansions (e.g., splines):

$$
\begin{aligned}
\mathcal{T}(L) &= \left\{ t \mapsto \sum_{j=1}^{L} a_j \Psi_j(t) \,:\, a_1, \ldots, a_k \in \mathbb{R} \right\}, \\
\mathcal{T}_\rho(L) &= \left\{ t \mapsto \sum_{j=1}^{L} a_j \Psi_j(t) \,:\, a_1, \ldots, a_k \in \mathbb{R}\,, \sum_{j=1}^{L} |a_j| \leq \rho \right\}.
\end{aligned}
$$

- Least-squares estimate $m_n^\dagger$ and m_n:

$$
\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} \left(m_n^\dagger(X_i, T_i) - Y_i \right)^2 &= \inf_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^{n} \left(f(X_i, T_i) - Y_i \right)^2 \\
m_n &= T_{\beta_n} m_n^\dagger.
\end{aligned}
$$

Consistency

Theorem

If

$$K_n \to \infty, \quad L_n^a \to \infty, \quad L_n^b \to \infty, \quad L_n^c \to \infty, \quad \beta_n \to \infty, \quad \rho_n \to \infty,$$

as $n \to \infty$ *in such a way that*

$$\frac{K_n \beta_n^4 (L_n^a + L_n^b + L_n^c) \log(\beta_n \rho_n K_n)}{n} \to 0 \quad \textit{and} \quad \frac{\beta_n^4}{n^{1-\delta}} \to 0,$$

for some $\delta > 0$ *as* $n \to \infty$*, then*

$$\lim_{n\to\infty} \int \big(m_n(x,t) - m(x,t)\big)^2 \mu(dx, dt) = 0 \quad \textit{with probability 1},$$

for every distribution of (X, T, Y)*.*

Algorithm

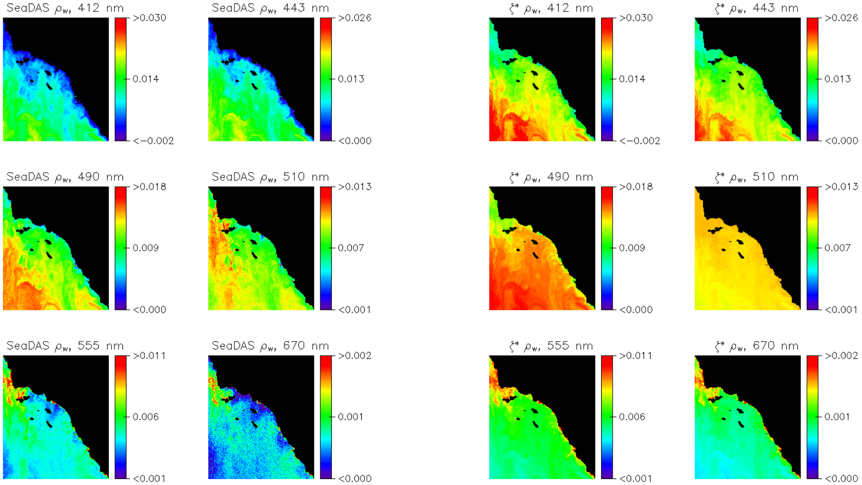
1. Simulate a data set $(x_1, y_1; t_1), \ldots, (y_n, x_n; t_n)$ using a RT model.
2. Select a noise model, i.e., a distribution for ε.
3. Consider a model of order K, parametric classes $\mathcal{T}_a$, $\mathcal{T}_b$,and $\mathcal{T}_c$: global parameter $\theta_K \in \Theta_K$:

$$\zeta_*(x,t) = \sum_{i=1}^{K} c_i(t) h\Big(\langle a_i(t), x\rangle + b_i(t)\Big)$$

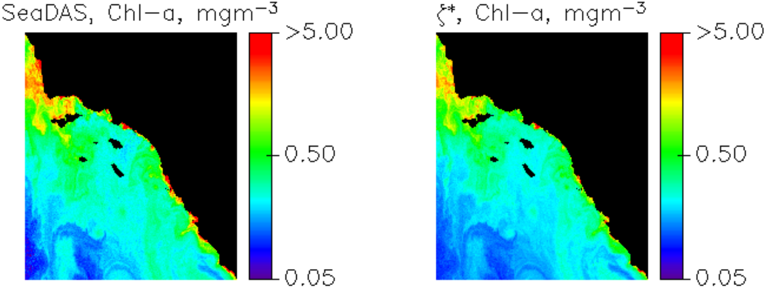
4. Fit the model by a stochastic gradient descent algorithm:
   - Initial value $\theta_K^{(0)}$;
   - Pick a data $(x_i, y_i; t_i)$;
   - Simulate ε_i and set $\tilde{y}_i = y_i + \varepsilon_i$.
   - At step j, update $\theta_K^{(j)}$ according to:

$$\theta_K^{(j+1)} = \theta_K^{(j)} + \alpha\left(\tilde{y}_i - \zeta_*(x_i, t_i)\right).$$

Application to SoCal SeaWiFS Imagery



Application to SoCal SeaWiFS Imagery



Motivations

Aerosols force climate by:

- scattering/absorbing solar radiation;
- absorbing/emitting terrestrial radiation;
- affecting cloud properties;
- reacting with greenhouse gases.

Aerosols impact:

- air quality;
- visibility.

Informations on aerosols are necessary for

- air quality forecasts;
- satellite imagery of ocean colour.

A key piece of information is the aerosol vertical profile.

Methodology

Radiance measurements in two spectral bands:

- a spectral band in the oxygen absorption A-band (e.g., 761-765 nm);
- a close spectral band outside the oxygen A-band (eg., 745-755 nm).

Assumptions:

- the atmosphere is homogeneous horizontally;
- scattering is only due to aerosols;
- the aerosol type does not vary with altitude;
- the surface is not reflecting.

Downward radiances in the single scattering approximation

$$
\begin{aligned}
L_1(\theta_0, \theta; p_s) &= C_1 \int_0^{p_s} n_a(p) T(\theta_0, \theta; p, p_s) dp \quad \text{in A-band} \\
L_2(\theta_0, \theta; p_s) &= C_2 \int_0^{p_s} n_a(p) dp := C_2 N_a \quad \text{outside A-band.}
\end{aligned}
$$

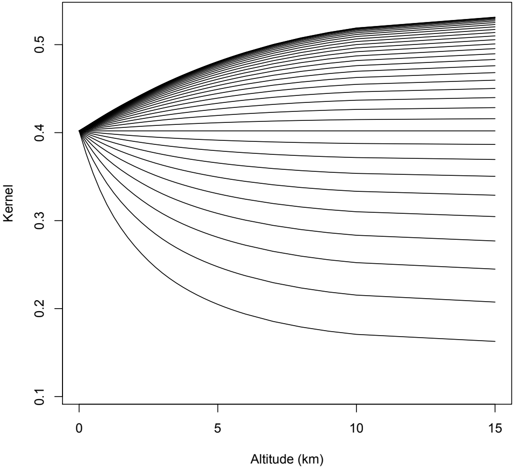
Close spectral bands: $C_1 \simeq C_2$, yielding the radiance ratio

$$
R(\theta_0, \theta; p_s) = \int_0^{p_s} T(\theta_0, \theta; p, p_s) \langle n_a(p) \rangle dp \quad \text{where} \quad \langle n_a(p) \rangle = n_a(p)/N_a.
$$

- $n_a(p)$: number of particule per m^2 per mb.
- Fredholm integral equation of the first kind $\sim$ Moment problem.

Kernel functions $k(., \theta)$

$$y(\theta) = (\mathcal{K}x)(\theta) = \int k(u, \theta)x(u)\mathrm{d}u.$$



Maximum entropy regularization

Context:

- μ_0: a reference measure;
- Set $\mu(\mathrm{d}u) = x(u)\mu_0(\mathrm{d}u)$.
- Set $\Phi : \mathbb{R} \to \mathbb{R}^n$ with $\Phi^i(u) = k(u, \theta_i)$, $i = 1, \ldots, n$.

We observe

$$y^{obs} = \int \Phi(u)\mu(\mathrm{d}u) + \varepsilon.$$

and the problem is to recover μ.

Regularization: Minimize $\mathcal{R}(\mu)$ w.r.t. μ subject to:

$$\int \Phi(u)\mu(\mathrm{d}u) = y^{obs},$$

or in a convex $K \ni y^{obs}$.

Maximum entropy regularization

Convex regularization functional:

- $\varphi : \mathbb{R} \to \mathbb{R}$ strictly convex, + additional assumptions e.g., $\varphi(x) = x\log(x)$;
- Define

$$I_\varphi(\mu) = \int \varphi\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}(u)\Big)\mu_0(\mathrm{d}u) \quad \text{if } \mu \ll \mu_0 \quad ; \quad +\infty \text{ otherwise.}$$

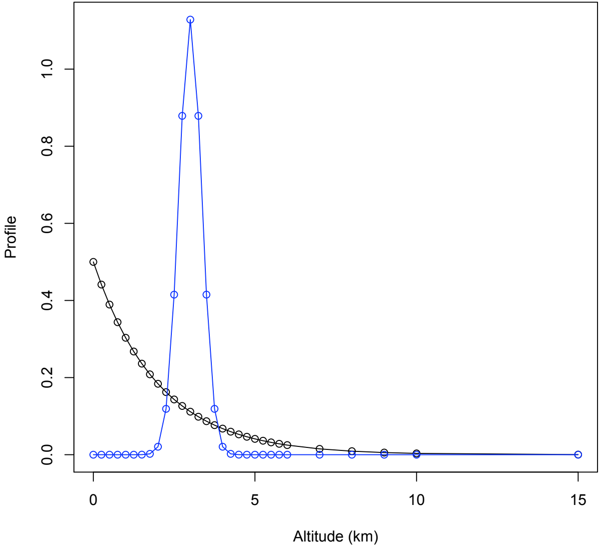
Then (Borwein and Lewis, 1993a, 1993b):

$$\hat{\mu} = \varphi^{*\prime}\big(\langle v^\star, \Phi(u)\rangle\big)\mu_0.$$

where $v^\star$ is any solution to

$$\max_{v\in\mathbb{R}^n} \quad \langle y^{obs}, v\rangle - \int_{\mathcal{X}} \varphi^*\big(\langle v, \Phi(u)\rangle\big)\mu_0(\mathrm{d}u).$$

Exponential and Gaussian profiles



Exponential

$$x(z) = \frac{1}{H_a} \exp\left(-\frac{z}{H_a}\right)$$

$H_a = 2$ km

Gaussian

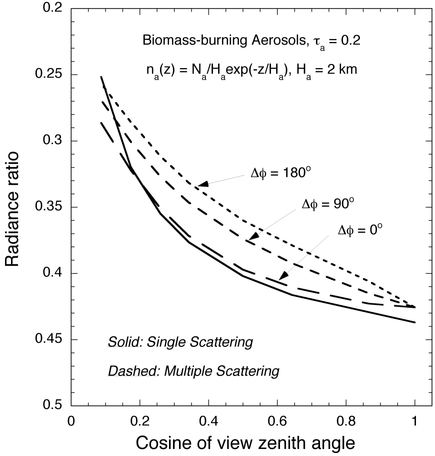
$$x(z) =$$

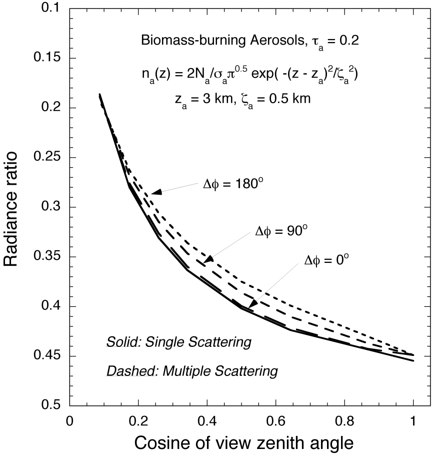
$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$\mu = 3$ km

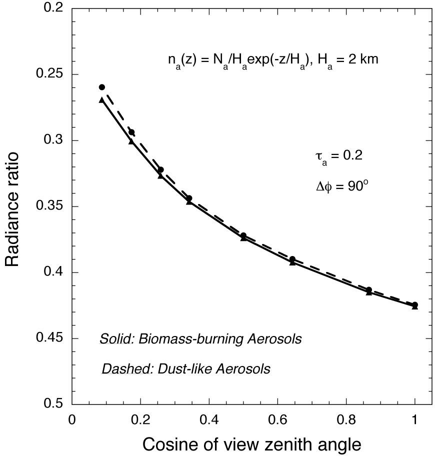
$\sigma = \frac{1}{2\sqrt{2}}$

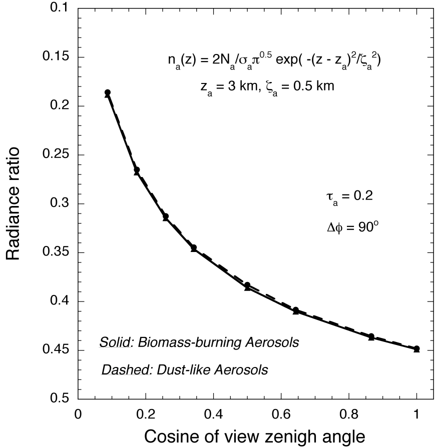
Radiance ratios



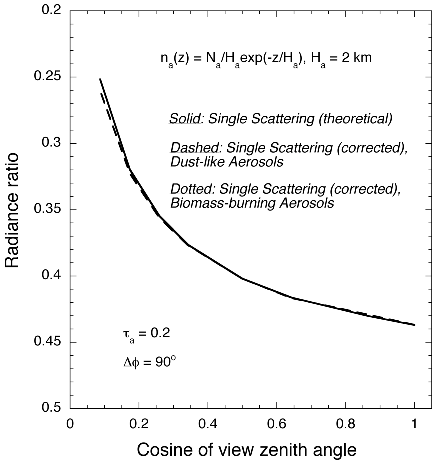


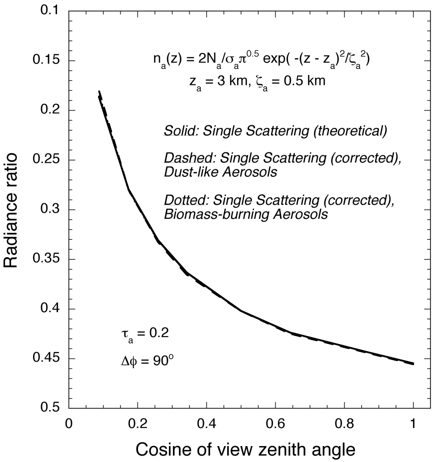
Radiance ratios: weakly dependent on aerosol type



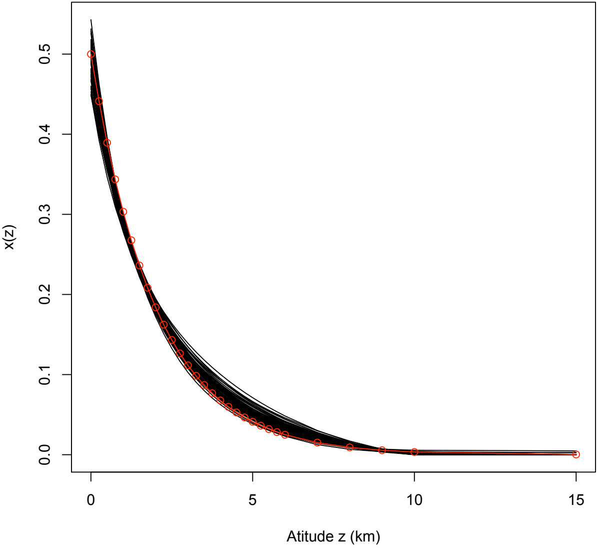


Radiance ratios: corrected for multiple-scattering



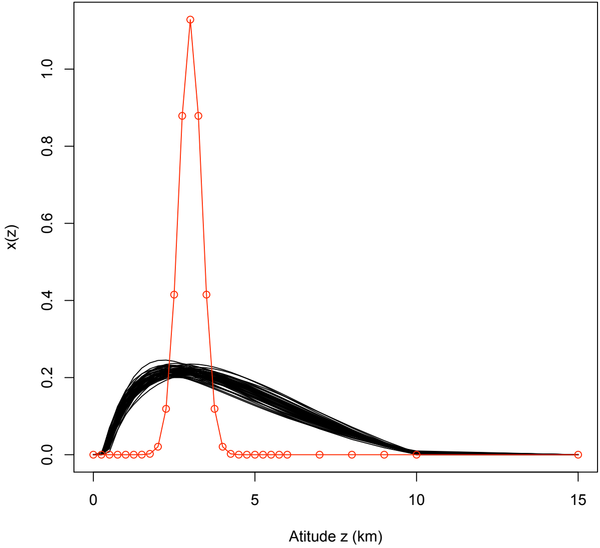


Exponential profile – Noise level 0.5%



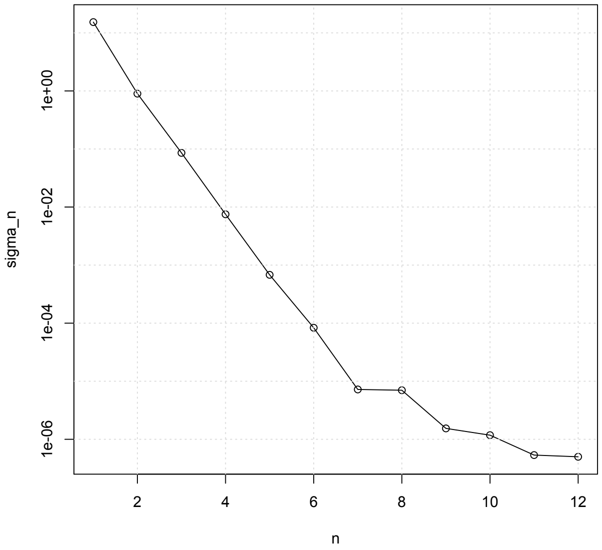
$\|\hat{x}_1 - x_1\|_2$ stat.	
Min	0.0104
Mean	0.0618
Max	0.1478
Std Dev	0.0308

Gaussian profile – Noise level 0.5%



$\|\hat{x}_1 - x_1\|_2$ stat.	
Min	1.472
Mean	1.491
Max	1.507
Std Dev	0.008

Singular values of $\mathcal{K}$



Conclusion

- $\sigma_n \sim e^{-n}$
- Exponential decay
- Severely ill-posed problem