Strategy for 4-dimensional Variational Assimilation of Remotely Sensed Ocean Colour SeaWiFS Data in a Coupled Biological-Physical Model of the North Atlantic Ecosystem

Pierre-Philippe Mathieu
Strategy for 4-dimensional variational assimilation of remotely sensed ocean colour SeaWiFS data in a coupled biological-physical model of the North Atlantic ecosystem.

Pierre-Philippe MATHIEU

Joint Research Center of the European Commission
Space Application Institute - Marine Environment Unit - TP 442
I-21020 Ispra - ITALY
pp.mathieu@altavista.net
http://www.gotm.net
LEGAL NOTICE

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of the following information.
# TABLE OF CONTENT

1. INTRODUCTION ................................................................................................................................. 3
   1.1. BACKGROUND AND RATIONALE ......................................................................................... 3
   1.2. SCIENTIFIC GOALS AND APPROACH ................................................................................ 3

2. PHYSICAL MODEL "ISPRAMIX" ........................................................................................................ 4
   2.1. MATHEMATICAL MODEL ......................................................................................................... 4
   2.2. NUMERICAL MODEL ................................................................................................................ 4
   2.3. MODEL SET-UP FOR THE NORTH ATLANTIC ....................................................................... 5

3. ECOLOGICAL MODEL "EMA" ............................................................................................................. 6
   3.1. CHOICE OF THE MODEL .......................................................................................................... 6
         3.1.1 What is expected from the model? .................................................................................... 6
         3.1.2 What level of complexity is needed? .................................................................................. 7
   3.2. MODEL COMPARTMENTS ......................................................................................................... 7
         3.2.1 State variables .................................................................................................................... 7
         3.2.2 Currency ............................................................................................................................ 7
   3.3. MODEL EQUATIONS ................................................................................................................... 8
         3.3.1 Parameterisation of biological processes .......................................................................... 8
         3.3.2 Source/sink terms for biological variables ...................................................................... 10
   3.4. NEXT STEPS TO BE DONE ..................................................................................................... 11
         3.4.1 Coupling with ISPRAMIX ................................................................................................. 11
         3.4.2 Validation with in-situ data ............................................................................................... 11
         3.4.3 Oriented sensitivity study .................................................................................................. 11

4. DATA ASSIMILATION TECHNIQUE "4D-VAR" ................................................................................ 12
   4.1. PHILOSOPHY OF THE METHOD ............................................................................................. 12
         4.1.1 The control variables ........................................................................................................ 13
         4.1.2 The cost function ............................................................................................................... 13
         4.1.3 The gradient and the adjoint model ................................................................................... 14
         4.1.4 The optimisation technique .............................................................................................. 15
   4.2. CRUCIAL BENEFITS OF 4D-VAR ......................................................................................... 15
   4.3. NEXT STEPS TO BE DONE .................................................................................................... 16
         4.3.1 Implementation of the data assimilation system ................................................................. 16
         4.3.2 Validation of the adjoint and gradient .............................................................................. 17
         4.3.3 Evaluation of the performance of the method ................................................................. 17

5. CONCLUSIONS .................................................................................................................................... 17

6. ACKNOWLEDGEMENTS .................................................................................................................... 17

7. REFERENCES ....................................................................................................................................... 17

8. FIGURES
1. Introduction

1.1. Background and rationale

As concern of the impact of human activities on global environmental change and in particular climate change [Houghton et al., IPCC report, 1995], the “Space Applications Institute” has recently initiated a “Global Environmental Information System (GEIS)” project which aims at better understanding the global carbon cycle [Fig 1] in the terrestrial and marine ecosystem (with special emphasis on the requirements for the implementation and verification of environmental treaties/conventions such as the Kyoto Protocol or the UN Convention for Biodiversity).

Within the GEIS project, the specific task “Global Ocean Monitoring Resource (GOMOR)” of the “Marine Environment Unit (ME)” aims at identifying the oceanic sinks/sources of CO₂ [Fig 2] which is currently one of the major challenge in global change studies [Gruber, 1998; Takashi et al., 1995]. This goal can only be achieved accurately by use of numerical biogeochemical models coupled with fully 3D hydrodynamical models and carefully validated and calibrated versus in-situ and remote measurements [Orr, 1997 – report of the Ocean Carbon-Cycle Intercomparison IGBP project]. In particular, as the main uncertainties about the ocean uptake of CO₂ come from the important biotic component [Volk & Hoffert, 1985], it is essential that the distribution of the oceanic plankton should be simulated correctly. In this to validate and to better understand the ecosystem model, ocean colour data remotely sensed by the space-borne Sea-Viewing Wide Field [SeaWiFS, Hooker & Esaias, 1993] play a role of paramount importance since, they are the only global data set for the near-surface phytoplankton pigment biomass [Fig 3].

1.2. Scientific goals and approach

In this context and as a preliminary step to the ME component of the GEIS project, we propose:

- To build an adequate biological model of the pelagic ecosystem and couple it to the hydrodynamical model ISPRAMIX [Demirov et al., 1998] running on the North Atlantic basin (domain between latitude 30°S and 60°N).
- To use 4-dimensional¹ variational assimilation of SeaWiFs data to calibrate poorly known ecological parameters in an “optimal” way (i.e. by minimising the mismatches between model results and remotely sensed data of surface chlorophyll).

The present note aims to present the strategy that we recommend to build such an integrated framework. In particular, we will discuss the rationale for the choices of the ecological model and the data assimilation technique and detail the various step of their implementation. This note is

¹ 3D in space + 1D in time (since it takes into account evolving data).
therefore intended as a roadmap document (we explain what we will do and how we will proceed) while modelling results will be presented in a future report. The text is structured as follows: In section (2), we describe the physical model ISPRAMIX (equations, underlying assumptions and numerical schemes). In section (3), we discuss the choice of the biological model (complexity needed, parameterisations of biological processes). In section (4), we explain the data assimilation technique and how we will design the assimilation system. In section (5), we draw some conclusions.

2. Physical model “ISPRAMIX”

The model used in this study is the 3D hydrodynamical model ISPRAMIX [Demirov et al., 1998] running on the Atlantic basin with the 4D-Var facility developed by Ouberdous et al. [1998]. In this section, we briefly recall the main mathematical and numerical features of the ISPRAMIX model. The reader is invited to consult the comprehensive reports of Demirov et al. [1998, 1999], Eiffer & Schrimpf [1992] and Ouberdous et al. [1998] for a complete description of the model equations and the 4D-Var assimilation system of Sea-Surface-Temperature data. ISPRAMIX reproduces well the mixed layer dynamics and circulation pattern in the North Atlantic [Demirov et al., 1999]. This is important for the coupled ecological perspective since the performances of the biological model are only as good as the physical model in which it is embedded.

2.1. Mathematical model

ISPRAMIX is based on the three-dimensional z-coordinate primitive hydrostatic equations, simplified by use of the Boussinesq and thin shell and approximations [Bryan, 1969]. The horizontal mixing by sub-grid lateral eddies is parameterised in a simple way by use of a Laplacian operator with a constant horizontal eddy coefficient. By contrast, the vertical eddy viscosity (for momentum) and diffusivity (for tracers) are calculated by a sophisticated turbulence closure following the model of Nihoul & Djenidi [1986] adapted by Eiffer & Schrimpf [1992]. Namely, the turbulence closure calculates eddy coefficients as a function of (i) a local stability function that depends on the flux-Richardson number [Nihoul & Djenidi, 1986], (ii) the Turbulent Kinetic Energy (TKE) computed by use of a differential equation and (iii) an algebraic macro eddy length scale “L” [Eiffer & Schrimpf, 1992]. It is worth stressing that a good parameterisation of turbulence is crucial for an adequate representation of biological processes since the mixed layer depth is known to control the primary production [Archer, 1995].

2.2. Numerical model

- Spatial discretisation by finite volume:
The model equations are solved numerically by using a finite volume method on a staggered C-grid of Arakawa & Lamb [1977]. This numerical scheme has the key advantage that it enables the conservation of the globally averaged variables.

♦ **Non-stationary vertical coordinate:**
The free-surface in ISPRAMIX is handled by use of a variable-layer method adapted from Dukowicz & Smith [1994]. Namely, the grid moves with the sea-surface, which means that the thickness of each numerical cell changes in time according to the oscillations of the sea-surface elevation.

♦ **Advection scheme:**
The QUICK "Quadratic Upstream Interpolation for Convective Kinematics" scheme of Leonard et al. [1995] (third order, monotonic, high-frequency filtering) is used to transport tracers (such as temperature and salinity) and biological variables, while the less expensive "Centered Difference" scheme (second order, dispersive) is used to transport momentum. It is worth stressing that the use of a highly performant advection scheme is especially critical for a good representation of biological variables (which should not overshoot/undershoot).

♦ **Temporal discretisation by mode splitting:**
In order to increase the efficiency of the numerical model algorithm, the mode-splitting method of Killworth et al. [1991] is used in ISPRAMIX. This method is based on the splitting the velocity field in two components: the rapid barotropic mode which corresponds to the depth-averaged velocity (barotropic transport) and the slow baroclinic mode which corresponds to the fluctuations around the mean transport. The method consists in resolving the rapid barotropic mode with a cheap 2D model and a small time step and the slow baroclinic mode with a large time step but with an expansive 3D model.

♦ **Implicit resolution of rotation and vertical mixing terms:**
In order to relax the numerical stability condition and therefore increase the time-step, two terms in the model equations are resolved implicitly. The Coriolis term is resolved via the pseudo-analytical rotation method of Backhaus [1985] and the vertical mixing term is resolved fully implicitly through the resolution of a tri-diagonal system.

### 2.3. Model set-up for the North Atlantic

The Atlantic version of ISPRAMIX that will be used has been fully tested and calibrated by Demirov et al. [1999]. The model covers the North Atlantic (between 30°S and 60°N) with a horizontal resolution of 0.1° × 0.1° (119 x 82 grid points) and 20 levels on the vertical (ranging from a thickness of 5 m at surface to 400 m in the abyssal ocean). The model bathymetry has been
obtained from the ETOPO 11 database and interpolated on the model grid [Fig 4]. At the ocean surface, the seasonal meteorological wind stress of Hellerman & Rosenstein [1983] and the heat flux of Esbensen & Kushnir [1981] are prescribed. A weak restoring of surface temperature and salinity toward the climatological Levitus [1994] data is also used to diminish the model’s drift. The Northern boundary (at 60°N) is assumed closed (zero normal velocity) and the Greenland-Iceland Norway sea is mimicked by use of a T/S relaxation toward Levitus climatology. The southern boundary (at 30°S) is forced by geostrophic velocities reconstructed from the observed density field and the T/S fields relaxed toward their observed climatological values. A comprehensive description of model results and validation procedure is given in Demirov et al. [1999].

3. Ecological model “EMA”

In this section, we discuss the choice of the trophic model “EMA” (Ecosystem Model for the Atlantic) that we have developed and that will be coupled with ISPRAMIX in the near future. The key advantage of EMA is that it provides a family of ecological parameterisations so that the user can test various biological hypotheses. The choice of biological parameterisations is done via a set of functions that serve as building blocks of the model.

3.1. Choice of the model

3.1.1. What is expected from the model?

Regarding the main objectives of GEIS, the biological model EMA is expected at least:

- To reproduce properly the annual cycle of the phytoplankton and its spatial distribution. This implies a capability to reproduce the major biological processes of the pelagic ecosystem in the Atlantic (e.g. basic food chain, primary production and its limitation by nutrient, re-mineralisation and exportation of inorganic matter in the deep ocean, replenishment of nutrient in the mixed layer via upwelling).

- To capture adequately the cycle of carbon in order to be coupled with an inorganic carbon chemistry model\(^2\) that computes the CO\(_2\) flux at the air-sea interface [Antoine & Morel, 1995; Bacastow, 1981; Peng et al., 1987].

- To be consistent with the framework in the context of inverse modelling. For example, Prunet et al. [1996] demonstrated that there is no need to use a highly complex ecosystem model to

---

\(^2\) In the momentum equations, advection plays a less important role since the Coriolis and pressure gradient are the dominant terms.

\(^3\) By contrast to biological modelling, chemistry modelling of the carbonic acid system causes no great controversy.
assimilate surface data since only a small amount of the biological parameters are independent!

3.1.2. What level of complexity is needed?

In order to fulfil the above criteria of success, one should question how many compartments are needed in EMA? Based on our review of available literature (Fasham et al. [1990]; Evans & Garçon [1997]; Kuhn & Radach [1997], Oschlies & Garçon [1999]; Oschlies & Garçon [1998]), we have concluded that a complex model (e.g. with numerous variables, bacterial loop, size-dependent classes...) is not suitable for our goal. Indeed, although it would permit a more accurate representation of the ecosystem, it provides a huge number of little known parameters and furthermore its results are much less interpretable. We have therefore opted for a simple pelagic 4-compartments NPZD model (Nitrogen-Phytoplankton-Zooplankton-Detritus) with a few adjustable parameters (Fig [5]). This model is conceptually and computationally simple but retains enough complexity to reproduce the essential processes of the Atlantic ecosystem that we want to represent.

3.2. Model compartments

3.2.1. State variables

The simple model we have designed comprises the following compartments:

- "P": The "Phytoplankton" variable encompasses all group-size of marine plants.
- "Z": The "Zooplankton" variable includes the biomass of herbivorous of micro and macro sizes.
- "N": The "Nutrient" variable lumps\(^4\) Nitrates, Phosphates and Ammonium.
- "D": The "Detritus" variable corresponds to all non-living organic matter (phytodetritus, fecal pellets, exuded mucus, and dead organisms).

3.2.2. Currency

All biological state variables and parameters are expressed in a common nitrogen unit "\(\text{mmolN}\)" since nutrients are regarded as the major limit of biological production in most parts of the ocean and in particular in the Atlantic [Fasham et al., 1990]. This has been preferred to mixed units of Carbon/Nitrogen/Chlorophyll since it enables a rapid check of equation dimension and global balance. When it is necessary to compare results with carbon or chlorophyll, we assume constant Redfield's ratios (\(1gC = 20mg\ Chlorophyll = 10\ mmolN\) from Henderson & Steele, [1995]).

\(^4\) Note that as the model does not partition \(\text{NO}_3\) and \(\text{NH}_4\), it cannot distinguish between the new and regenerated production [Dugdale & Goering, 1997].
3.3. Model equations

Each biological variable "Y" is governed by the following equation:

\[
\frac{\partial Y}{\partial t} + v \cdot \nabla Y = \nabla \cdot (K \cdot \nabla Y) + Q_r
\]  

(1)

where, \(Y, t, v, \nabla, K, Q_r\) are respectively a generic biological state variable, the time, the velocity vector, the gradient operator, the horizontal/vertical mixing tensor (which parameterises sub-grid scale turbulence) and the ecological source/sink term (which parameterises the exchanges between the different compartments). The successful choice of approximation for "\(Q_r\)" is the key issue in biogeochemical modelling and is discussed below.

3.3.1. Parameterisation of biological processes

In EMA, we have taken into account the important processes relevant to represent adequately the Atlantic ecosystem. One of the key originality of the EMA code is that it provides a portfolio of biological parameterisations by appealing to a functional formalism. Namely, each biological process for a variable "X" is represented by a generic function "\(G^N_m(X, a, b)\)" of Class "\(M^r\)" (referring to the nature of parameters "\(a\)" and "\(b\)") and Type "\(N\)" (referring to the asymptotic behaviour of the function). The set of available functions is described in Table [2] and their behaviour illustrated in Fig [6]. The definition of function parameters and an associated default value are given in Table [1].

It is worth stressing that the functional formalism has several powerful advantages: (i) it enhances lisibility of the Fortran code reducing thereby the risk of bug (no need to copy 10 times the same expression), (ii) it easily permits the use and inter-comparison of various biological processes parameterisations, (iii) it is extremely useful to develop the adjoint code since a change the biological parameterisation is automatically taken into account in the adjoint.

* PAR attenuation and self-shading

The Photosynthetically Available Radiation "I" (called PAR) corresponds to the short-wave spectral band of the penetrating radiation. The PAR intensity at depth "z" depends on the PAR at surface via the Beert exponential attenuation law,

\[
I(z) = I(0) e^{-kz}
\]  

(2.a)

with an extinction coefficient,

\[
k = k_p + k_w \frac{1}{z} \int_0^z P(z', t) dz'
\]  

(2.b)

taking into account the influence of water turbidity "\(k_w\)" (e.g. inorganic suspended matter including yellow substances) and self-shading by biotic matter "\(k_p\)" (such as the Chlorophyll of phytoplankton) (see Table [1]):
Photosynthesis primary production

The Gross\(^5\) primary production by photosynthesis (i.e. the total uptake of $CO_2$) has a maximal rate of $\mu_p$ that is influenced by several factors:

$$G_{pp} = \sigma_I \times \sigma_N \times \frac{Q_{10}^{(T-10)/T}}{\mu_p} \times P$$

(3.a)

- **Limitation by light**

The maximal primary production is limited by the PAR. In EMA, the following classes of $P-I$ curves (functional dependence of $\sigma_I$ on "PAR") can be considered according to the process that is intended to be represented.

(a) **Photo-saturation** is parameterised by use of various functions that have a "Linear-Constant" asymptotic behaviour and are defined by their initial slope value (Class "B"):

$$\sigma_I = G_0^I(I, \alpha_I, 1)$$

(3.b)

(b) **Photo-inhibition** is parameterised by use of various functions with a "Linear-Hyperbolic" asymptotic behaviour and are defined by their optimal value (Class "C"):

$$\sigma_I = G_0^{1,3}(I, I_o, 1)$$

(3.c)

- **Limitation by nutrients**

To parameterise the limitation by nutrient, we classically use function of Michaelis-Menten type with a "Linear-Constant" saturation behaviour and defined by their half-saturation value (Class "A"):

$$\sigma_N = G_A^I(N, K_N, 1)$$

(3.c)

- **Dependence on temperature**

The temperature dependence of growth (and also of grazing) is taken into account via "$Q_{10}$" coefficient where the exponent measures the deviation compared to a reference temperature of 10 degrees [Eppley, 1972]. Note that it is important to consider the effect of temperature on biology of the Atlantic ecosystem because of the rather large range of temperature in the North Atlantic basin.

- **Grazing by herbivorous**

The grazing of Phytoplankton by Zooplankton is assumed to be a linear function of $Z$:

$$Graz_Z = \sigma_i \times \frac{Q_{10}^{(T-10)/T}}{\mu_z} \times Z$$

(4.a)

- **Limitation by phytoplankton**

The maximal uptake is limited by the presence of phytoplankton. This can be represented by a function that saturates and has a linear or quadratic behaviour for small values depending on the

---

\(^5\) Note that it is sometimes convenient to use the Net Primary Production which is the Gross Primary Production minus the linear respiration losses.
underlying biological assumption (see Fasham [1995] for a discussion of the choice of grazing function):

\[ \sigma_p = G_p^{1.1}(P, K_p, 1) \]  

(4.b)

**Losses to non-living matter**

The process of respiration is assumed to be linear in the variable and to create a loss toward the nutrient compartment,

\[ \text{Resp}_Y = \alpha_{Y,N} Y \quad \text{where} \quad Y = Z, P \]  

(5)

The mortality term represents the effects of natural mortality but also parameterises the effects of higher trophic level (e.g. consumption of zooplankton by higher predators). The term is assumed to be quadratic in zooplankton and to create a loss toward the detritus compartment,

\[ \text{Mort}_Y = \beta_{Y,D} Y^2 \quad \text{where} \quad Y = Z, P \]  

(6)

Note that the quadratic dependence (which means that the mortality rate per capita is proportional to the zooplankton biomass) assumes that higher predators have a biomass proportional to their prey.

**Remineralisation of organic detritus in the aphotic deep ocean**

Detritus are assumed to sink with a prescribed vertical velocity "\( w_o \)" and to be mineralised to nutrients with a constant rate,

\[ \text{Mine}_D = \alpha_{D,N} D \]  

(7)

This parameterisation of the mineralisation process enables the replenishment of surface water with nutrient upwelled from the deep ocean.

### 3.3.2. Source/sink terms for biological variables

By identifying the direction of the flows between the various compartments as in Fig [5], the source/sink term of the various biological variables read:

\[ Q_p = +G_{pp} - \text{Resp}_p - \text{Mort}_p \quad - \text{Graz}_z \]  

\[ Q_z = \quad - \text{Resp}_z - \text{Mort}_z \quad + \text{Graz}_z \quad \sigma_{z,D} \]  

(7)

\[ Q_N = -G_{pp} + \text{Resp}_p \quad + \text{Resp}_z \quad + \text{Mort}_z \quad + \text{Mine}_N \]  

\[ Q_D = \quad + \text{Mort}_p \quad + \text{Mort}_z \quad - \text{Mine}_N \quad + \text{Graz}_z (1 - \sigma_{z,D}) \]

The above system of equations simply expresses that:

- The phytoplankton grows via photosynthesis and decreases due to zooplankton grazing and natural losses.
• The zooplankton growths by uptake of phytoplankton and decreases by respiration, and mortality losses. Note that one fraction of the grazing is not assimilated but excreted directly to the detritus.
• The nutrients are regenerated by mineralisation of detritus and respiration losses and are consumed by phytoplankton for the primary production.
• The detritus increase by mortality and excretion of biotic compartments and decrease by the bacterial activity that remineralise them.

3.4. Next steps to be done

3.4.1. Coupling with ISPRAMIX

The biological model EMA will be coupled with ISPRAMIX. This will be pursued by taking the ISPRAMIX routine for an advected/diffused tracer (such as temperature, salinity) and include in it the above source/sink terms of interactions between biological variables. The effective vertical velocity of transport of biological variables will be the sum of the migration/sinking velocity plus the the Eulerian velocity computed by ISPRAMIX.

3.4.2. Validation with in-situ data

The integrated model ISPRAMIX-EMA should be validated by the complementary use of “in-situ data” (available only at some location in space but for the water column depth and for various biological compartments) and “ocean colour remote sensing data” (available with a global coverage in time and space but only at the top surface for the phytoplankton compartment). We propose first to grossly tune EMA versus in-situ data via the “hit & miss” method and then to calibrate optimally EMA versus SeaWifs data via the “4D-Var” method (see section [4]). The best available data sets for biology are measured at Ocean Weather Stations. We therefore recommend to pre-calibrate EMA by use of the long time series of phytoplankton, nutrients and detritus available at the U.S. JGOFS Bermuda Atlantic Time-Series Study (BATS) [Doney et al., 1986].

3.4.3. Oriented sensitivity study

A sensitivity study of EMA is required not only to better understand the properties of the model and verify its skill but also in view of applying properly the 4D-Var method. Namely, parameters that are susceptible to create a change in the surface chlorophyll should be identified to provide candidates of potentially adjustable parameter for the data assimilation system. It is indeed obvious that if the surface phytoplankton is not modified by the value assigned to a parameter, there is no point to try to adjust this parameter to fit SeaWifs data. We propose to test more precisely the model sensitivity by perturbing the different parameters within their physical error bar and establishing sensitivity indexes based on the ratio of surface phytoplankton change and parameter perturbation.
In order to perform the sensitivity analysis and the pre-calibration of EMA, we recommend to use a simple 1D framework which is highly beneficial for an easier understanding of results and in term of its low CPU cost that enables extensive sensitivity studies. Furthermore, in order to take into account three-dimensional information, the 1D integrated model could be forced by the velocity field obtained by SPRAMIX (a procedure similar to Oschlies & Garçon [1999]). We propose to couple EMA with the model GOTM (General Ocean Turbulence Model) developed at the JRC by Burchard et al. [1999]. The coupled model can then be tested at the BATS site to provide the first guess value of biological parameters.

4. Data assimilation technique "4D-Var"


In this study, we will use and examine in details the very powerful four-dimensional variational method (also known as 4D-Var method). Namely, 4D-Var is a mathematical technique that enables to combine simultaneously all available information on the system: (i) measurements (e.g. in situ or remote observations), (ii) their error characteristics and (iii) the dynamical constraints (e.g. primitive equations, chemical reactions, biological model) in view of providing the "best" estimate of the state of the system. As illustrated in Fig [8], the crucial benefit of 4D-Var method is that it fully exploits the information contained in the model evolutive equations!

4.1. Philosophy of the method

The basic philosophy of 4D-Var illustrated in Fig [7] and Fig [8] is to adjust a set of control variables ("C") by minimising globally the misfit between model and observations [Courtier & Talagrand, 1987]. The various steps of the 4D-Var method are illustrated in Fig [7]: (1) the numerical model is integrated forward, (2) the cost function measuring distance between model and data is evaluated, (3) the gradient of the cost function with respect to the control parameter is evaluated by running the adjoint model backward, (4) the optimisation package uses the gradient as a search direction to adjust the control variables in a "best" way. We discuss further below the various basic ingredients of the 4D-Var method.
4.1.1. The control variables

The so-called control variables “C” are the variables that must be recovered optimally by the 4D-Var method [Lions, 1968]. They can be of various nature such as: (i) the model initial conditions (see meteorological forecasting problem [LeDimet & Tallagrand, 1986; Bennet & Miller, 1990], (b) poorly-known coefficients in the model or of the forcing fields (see parameter estimation problems in 1D biogeochemical models [Ishizaba, 1993; Matear, 1995; Prunet et al., 1996]), (c) little estimated or missing data (see how to reconstruct ozone field with dynamical consistency [Fisher & Larry, 1995]).

In the scope of this study, the problem is to optimally calibrate poorly known ecological parameters by use of remotely sensed surface phytoplankton data. Selecting the biological parameter as the only control variables implies that the physical model is assumed “perfect” in the sense that the discrepancy between observations and simulations are considered to be caused by the biological model solely and not by the parameterisations or resolution of the physical model.

After a careful review of various biogeochemical models (Steele & Frost [1977]; Evans & Parslow [1985]; Forst [1987]; Hofmann & Ambler [1988]; Wroblewski [1989]; Fasham et al. [1990]; Steele & Henderson [1992,1993]; Fasham [1993]; Armstrong [1994]), we have selected as control the following parameters “Kn”, “Kp”, “μϕ”, “μz”, “αi” that have large error bars. The range of uncertainty of these biological parameters simply reflects (i) the difficulty of measuring (e.g. the growth and the mortality rates describe the evolution of a mixture of species and cannot therefore be measured in a laboratory or on site) and (ii) the weak spatio-temporal coverage of observations (most of the observations for nutrient and zooplankton come from weather station and transects) and (iii) the inherent variability of the coefficient due to different bio-physical regime.

Among the potential candidates for control parameters, we have to ensure that they have an impact on the surface phytoplankton. A targeted sensitivity analysis is therefore a crucial prerequisite to the choice of control parameters. Furthermore, the initial value for parameter should be chosen realistic enough since it entirely conditions the success or the failure of the assimilation procedure (e.g. for example, if the first guess value is too far from the optimum, the minimisation procedure could be stuck in a local minimum).

4.1.2. The cost function

The cost function is the sum of 2 components:

- A “goodness of fit” functional that measures the distance on a spatio-temporal domain between the observations and the model-equivalent of measurements. By using a least square estimator for the norm, this term reads:

\[
J_{GOF} = \frac{1}{2} \int \int \delta(x) (H(\mathbf{y}) - \mathbf{y}_0)^T \mathbf{R}^{-1} (H(\mathbf{y}) - \mathbf{y}_0) dV dt
\]  

- 13 -
where "Y_o" is the observation (e.g. SeaWifs chlorophyll pigment value), "\( H(Y) \)" is the model-derived quantities equivalent to the observation (e.g. the phytoplankton concentration at surface), "\( R \)" is the observation error co-variance matrix (used to provide more weight to the more precise) data, "\( ^T \)" is the transpose operator, "\( \delta_m \)" is the Heaviside function (value 1 when and where observations are available and zero otherwise).

- A "background" or "penalty" functional term that ensures that the biological parameter does not deviate too much from a target value remaining thereby in an acceptable physical range (the background term becomes infinite if the parameters exceeds the bounds):

\[
J_{\text{BACK}} = \frac{(C_o - C)^2}{(C_u - C_L)^2}
\]  

where "\( C_o, C_u, C_L \)" are respectively the target, upper and lower values of the control parameters.

### 4.1.3. The gradient and the adjoint model

The gradient of "\( J \)" with respect to "\( C \)" is needed to provide a search direction for the optimisation routine that aims to minimise "\( J \)". However as "\( J \)" is not explicitly expressed in term of "\( C \)" the explicit evaluation of the gradient is a cumbersome task requiring an extensive use of the chain rule derivative. A more elegant way is to appeal to an intermediate mathematical tool called the "adjoint" model that maps the cost function and model equations into the gradient. The adjoint model has no physical interpretation but emerges naturally from the Euler-Lagrange equations. These equations introduce an additional variable called the Lagrangian multiplier to transform a constrained minimisation problem in an unconstrained one [Betsekas, 1982]. Let us build the following Lagrangian functional,

\[
L = \frac{1}{2} \iint_{V_t} \delta_m (Y - Y_o)^2 \, dt \, dV + \iint_{V_t} \lambda \left( \frac{\partial Y}{\partial t} + \nu \cdot \nabla Y - \nabla \cdot (K \cdot \nabla Y) - Q_Y \right) \, dt \, dV
\]

where the first term is the observationnal cost \( J_{\text{COF}} \) function (observations have been assumed uncorrelated so that the covariance matrix simplifies to a diagonal array and \( J \) becomes a weighted sum of square differences) and the second term is the model equations pondered by the Lagrangian multiplier "\( \lambda \)" and integrated over the spatio-temporal domain.

The variational calculus theory tells us that the optimal control variable that minimises the cost function and verifies model equations at the same time is obtained when the Lagrangian is stationary at a saddle point [Betsekas, 1982]. This means that all its partial derivatives vanish simultaneously:

\[
\frac{\partial L}{\partial \lambda} = 0 \Rightarrow \frac{\partial Y}{\partial t} + \nu \cdot \nabla Y - \nabla \cdot (K \cdot \nabla Y) - Q_Y = 0
\]
The variation of the Lagrangian with respect to the Lagrangian variable simply returns the "forward" or "direct" model equations which must be satisfied exactly (strong constraint problem).

\[ \frac{\partial L}{\partial Y} = 0 \Rightarrow \frac{\partial \lambda}{\partial t} + \nabla \cdot (v \lambda) + \nabla \cdot (K \cdot \nabla \lambda) + \delta_m (Y_m - Y) + \lambda \frac{\partial Q}{\partial Y} = 0 \]  
(12)

After integration by part of all gradient terms in "L" (in order to transfer the derivative on "\lambda" rather than "Y"), the derivation of the Lagrangian with respect to the dependent state variable "Y" yields the "adjoint" or "inverse" equation which governs the evolution of the Lagrangian multiplier "\lambda(x,y,z,t)". The adjoint differential equation looks dynamically similar to the forward model but has however a conservative advection form, an anti-mixing operator and a modified source/sink term including the derivative of the biological exchange and the data misfit. Furthermore, the adjoint model must be run backward in time since it has its boundary condition fixed at the final time by the integration by part.

\[ \frac{\partial L}{\partial C} = 0 \Rightarrow \frac{\partial J}{\partial C} = \int_{t_0}^{t_f} \lambda(x,y,z,t) dt \]  
(13)

The variation of the Lagrangian with respect to the control variable yields the gradient of the cost function with respect to control parameters.

Eq [13] demonstrates that the gradient of the cost function can be simply obtained by solving first the adjoint model for "\lambda" and then integrating the Lagrangian multiplicator in time. The Lagrange multiplier appears therefore as an intermediate purely mathematical variable that eases the computation of the gradient.

4.1.4. The optimisation technique

The minimisation routine is based on a descent algorithm that searches the optimal value of the parameter (the one that minimises J) in a direction computed from the gradient. More precisely, the algorithm we will use in the model is the highly performant routine “N1QN3” from MODUPLLOT library in INRIA [Gilbert & Lemarechal, 1989] based on the variable storage quasi-Newton method and the BFGS formula [Liu & Nocedal, 1989].

4.2. Crucial benefits of 4D-Var

In this section, we describe the powerful advantages of the 4D-Var technique over classic methods like sequential assimilation via optimal interpolation which simply consists of inserting data information while integrating the model (see Fig [8] for illustration of the method):
• 4D-Var provides a global analysis in time that takes into account within a single analysis all the observations in time (past and future) and space (from surface to depth) while data insertion is only local in time.
• 4D-Var has the maximum accuracy of the ensemble of observations while optimal interpolation cannot be more accurate than the last measurement.

• 4D-Var includes all the available information in a smooth way while data insertion is likely to create an unbalance in the system due to successive corrections in an evolving system.

• 4D-Var is a dynamical method in the sense that it adds information contained in the dynamical equations of the system (e.g. primitive equations, biological/chemical relations...). In our problem of optimal calibration, the variational method is susceptible to improve the overall ecosystem model (all biological parameters) with information solely on phytoplankton situated at surface. Indeed, since 4D-Var takes into account the equation of the biological model which express the transfer of matter between state variables, the method should be able to transfer information from surface to depth and between the different biological compartments.

4.3. Next steps to be done

4.3.1. Implementation of the data assimilation system

The first step will be to code the various ingredients of the 4D-Var method and integrate them in a common computer framework called the data assimilation system and schematically described in Fig [7]. This will require:

• The interpolation of SeaWifs data on the spherical 0.1°x0.1° grid of ISPRAMIX (in time and in space) and their conversion in biomass of phytoplankton at surface in order to build a discrete form of the cost function.

• The design of the adjoint code of the biological model either: (i) directly from the direct model code by using the recipes of Glering [1992], or (ii) indirectly from the linear tangent code (linearised version of the computer code) [Tallongrand, 1991]. Note that the adjoint of the 3D advection/diffusion operator in ISPRAMIX has already been designed by Ouberdous et al. [1999].

• The calibration of the input parameters of the optimisation technique of the INRIA library.

The design of the assimilation system will strongly benefit from the work Lellouche et al. [2000] on the 4D-Var biogeochemical assimilation in the Adriatic Sea and the work of Ouberdous et al. [1999] on the assimilation of SST. These authors have indeed developed an assimilation system that runs off-line (by saving the ISPRAMIX model trajectory into an intermediate file) and could be easily adapted to another application domain.
4.3.2. Validation of the adjoint and gradient

We will perform the Taylor test to validate the adjoint. The test consists in perturbing the cost function with an increment \( \delta C \) in the space of parameter and to check if the verification of the Taylor expansion formula in the limit of \( \gamma \to 0 \):

\[
J(C + \gamma \delta C) = J(C) + \gamma \nabla_C J \delta C + O(\gamma^2)
\]  

(14)

If the above formula is verified, it means that the adjoint code has computed a good gradient \( \nabla J \) (note that this is a necessary but not sufficient condition for the validity of the adjoint!).

4.3.3. Evaluation of the performance of the method

In order to evaluate the performances of the data assimilation system, we will perform "twin" experiments. This test consists in (i) building a set of self-consistent "ideal" parameters by running the model with self-constructed observations, (ii) perturbing the parameters (e.g. by Gaussian noise) and (iii) trying to recover their true initial value via the data assimilation system. Once we have validated the method on self-constructed data, we will test it with real SeaWifs data. If in this case, we get unrealistic parameters, it would mean that the biological model is not adapted to fit the data and should be questioned!

5. Conclusions

One of the key opened question in global climate change study is how the ocean will absorb the anthropogenic carbon dioxide. This research requires a better understanding of the marine ecosystem that plays a role of paramount importance in the global carbon cycle. This challenging task can only be achieved by using 3-dimensional biogeochemical models thoroughly validated by local in-situ measurements and global remote sensing data (e.g. near-surface phytoplankton data derived from ocean colour sensors like SeaWIFS).

As a first step in this direction, we have discussed in the present note the strategy to build a numerical tool integrating ecological-physical models and optimal assimilation of remote-sensing data by the 4D-Var method. An 4-compartment ecological model of the North Atlantic ecosystem "EMA" has been designed and will be coupled to the circulation model ISPRAMIX. We have also discussed how to implement the 4D-variational assimilation technique which provides an ideal, elegant and automated approach to squeeze the juice of satellite data and model dynamical information.
6. Acknowledgement

Marc Bouvet, Mohamed Ouerdous, Jean-Michel Lellouche, Karsten Bolding, Fabrice Veerse are kindly acknowledged for providing helpful comments.

7. References


Clais, P. & Le Quéré, C. [1998]. "ERCA, Urban Air Pollution to Extra Planets". S. Boutron, C.:

Courtier, P. & Tallagrand, O., [1990]. "Variational assimilation of meteorological observations with the direct and adjoint shallow water equations." Tellus. Ser. A. 42. 531-549.


Evans, G. & Garçon, V. [1997]. "One-dimensional models of the water column biogeochemistry." JGOFs report. Toulouse (France).


- 19 -
Ouberdous, M., EIFLER, W., Demirov, E., Villacastin, C., & Nykjaer, L. [1999]. "Data assimilation of remote sensed sea surface temperature on the Atlantic basin and a nested window via a hydrodynamic model." Copenhagen, Denmark.


Figure 1: Reservoirs and exchanges in the natural global carbon cycle. (the word natural means before the human perturbation). The reader is referred to Ciais & LeQuéré [1998] for a review of the flux exchange and the associated process.
Figure 2: Oceanic carbon cycle. The ocean which covers 70% of the Earth surface is one of the largest acting carbon pool. The carbon element is present in the ocean in various forms. We distinguish the inorganic carbon in dissolved form (DIC which includes (a) CO₂ [0.1%], (b) carbonates CO₃⁻ [10%] and (c) bicarbonates HCO₃⁻) and the organic carbon present in dissolved form (DOC), particulate form (POC) or biotic form (marine organisms). The ocean plays a major role in the global carbon cycle and it controls the atmospheric CO₂ via 3 main effects: (1) the chemical solubility pump (buffer effect of carbonate chemistry into the water), (2) the biological pump (marine organisms like phytoplankton absorb dissolved CO₂ via photosynthesis and export particulate organic carbon from the surface layer toward the deep ocean where they are remineralised) and (3) the carbonate counter-pump (the formation of carbonate shell by phytoplankton consumes carbonate ions increasing thereby the pCO₂). The net effect of ocean biota (combined effect 2 & 3) is to reduce the ocean pCO₂ at surface. Model studies [Bacastow & Maier-Reimer, 1990] and GEOSECS analysis [Volk & Hoffert, 1985] suggest that the biological pump contributes for more than 60% of the total pump of DIC in the world ocean.
Figure 3: Satellite image of chlorophyll. Colour representation of phytoplankton pigment concentration derived from the radiances measured by the Sea-Viewing Wide Field [Hooker & Esaias, 1993]. The SeaWiFS sensor is operational since 1997 on the synchronous polar orbit, it has 8 spectral bands in the visible and near-infrared and provides a near global coverage every two days with a nominal resolution of about 1.14 km at nadir (when the sensor is directly overhead). The ocean colour data combined with measurements of sea-surface temperature (obtained from a satellite) and irradiance can be used to provide an estimate of the primary production through empirical formula [Bouvet & Dowell, 1999].
Figure 4: Bathymetry of the North Atlantic used in ISPRAMIX. The Bathymetry has been reconstructed from the ETOPO data and has not been smoothed. Note that the Gibrallar Strait is closed and the effect of the Mediterranean outflow is parameterized via restoring of T/S towards observed values.
Figure 5: N-P-Z-D pelagic model. This schematic representation illustrates the interactions between the various biotic (P & Z) and abiotic (N & Z) reservoirs. The arrows indicate flows of matter in the system via various biological processes (yellow boxes). Arrows not ending in a biological compartment (like sinking) are net losses to the system.
Figure 6: Functions behaviour. The type of function represents the property of the biological process (either growing, saturated or inhibited). Note that within each class, there is a set of functions that have equivalent asymptotic behaviour (Table [2]).
Figure 7: Sketch of the 4D-Var assimilation system (from Lawson et al. [1995]).
Figure 8: Data assimilation methods: Philosophy and benefit. Schematic representation of the [a] the sequential and [b] 4D-Var data assimilation technique (as taken from Eskes et al. [1997]). The inverse problem treated in this example is how to improve the initial condition of a model run to get a simulation trajectory that is closer to observations (this is a typical forecasting problem in meteorology called re-analyses). In the sequential method (top sketch), the model is integrated forward and when an observation (dot) is encountered, a new field is computed. This method creates thereby jumps in the model trajectory [Eskes et al., 1997]. By contrast, in the 4D-Var method (bottom sketch) all data are assimilation in one single analysis. This is archived via the following iterative method: (1) the direct model is integrated forward until final time to provide a first guess model trajectory (dashed line), (2) the misfits (dotted arrows) between the first guess forecast and the observations (circled) are computed, (3) the "adjoint model" (lower solid curve) forced by observation mismatches (vertical steps in the curve) is integrated backward to provide the gradient, (4) an optimisation technique gives a new starting point that provides an updated smooth forecast (solid curve) including all observations. The difference between both methods is clearly apparent at the first measurement where data insertion would suggest a positive correction while 4D-Var provides a negative correction.
<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>UNIT</th>
<th>VAL</th>
<th>DESCRIPTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_w$</td>
<td>$m^{-1}$</td>
<td>0.04</td>
<td>Extinction coefficient of water</td>
<td>Fasham [1993]</td>
</tr>
<tr>
<td>$k_p$</td>
<td>$m^{-1}/(mmolN.m^{-3})$</td>
<td>0.03</td>
<td>Self-shading extinction coefficient (Phy chlorophyll)</td>
<td>Fasham et al. [1993]</td>
</tr>
<tr>
<td>$I_o$</td>
<td>$W/m^2$</td>
<td>100</td>
<td>Optimal PAR intensity for the P-I curve</td>
<td></td>
</tr>
<tr>
<td>$\alpha_I$</td>
<td>$day^{-1}/W/m^2$</td>
<td>0.065</td>
<td>Initial slope of the P-I curve</td>
<td>Kuhn &amp; Radach [1997]</td>
</tr>
<tr>
<td>$\mu_p$</td>
<td>$day^{-1}$</td>
<td>1.5</td>
<td>Maximum uptake rate for primary production (10°C)</td>
<td>Fasham et al. [1993]</td>
</tr>
<tr>
<td>$\alpha_{P,N}$</td>
<td>$day^{-1}$</td>
<td>0.1</td>
<td>Respiration linear loss rate from $P\rightarrow N$</td>
<td>Franks et al. [1986]</td>
</tr>
<tr>
<td>$\beta_{P,D}$</td>
<td>$day^{-1}/mmolN/m^3$</td>
<td>0.1</td>
<td>Mortality quadratic loss rate from $P\rightarrow D$</td>
<td></td>
</tr>
<tr>
<td>$K_p$</td>
<td>$mmolN/m^3$</td>
<td>0.05</td>
<td>Half-saturation for $P$ in $Z$ mortality</td>
<td></td>
</tr>
<tr>
<td>$K_N$</td>
<td>$mmolN/m^3$</td>
<td>1.0</td>
<td>Half-saturation for nutrient uptake</td>
<td>Franks et al. [1986]</td>
</tr>
<tr>
<td>$Q_{10}$</td>
<td>/</td>
<td>2.08</td>
<td>Temperature dependence coefficient for growth</td>
<td>Fransz &amp; Verhagen [1985]</td>
</tr>
</tbody>
</table>

**Herbivoros compartment**

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>UNIT</th>
<th>VAL</th>
<th>DESCRIPTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_z$</td>
<td>$day^{-1}$</td>
<td>1.0</td>
<td>Maximum grazing rate of $Z$ at 10°C</td>
<td>Fasham et al. [1990]</td>
</tr>
<tr>
<td>$\alpha_{Z,N}$</td>
<td>$day^{-1}$</td>
<td>0.2</td>
<td>Respiration linear loss rate from $Z\rightarrow N$</td>
<td>Franks et al. [1986]</td>
</tr>
<tr>
<td>$\beta_{Z,N}$</td>
<td>$day^{-1}/mmolN/m^3$</td>
<td>0.2</td>
<td>Mortality quadratic loss rate from $Z\rightarrow N$</td>
<td>Fasham et al. [1990]</td>
</tr>
<tr>
<td>$\sigma_{Z,D}$</td>
<td>/</td>
<td>0.75</td>
<td>Assimilation efficiency of $Z$ uptake (excretion)</td>
<td>Oechsli &amp; Gnacon [1999]</td>
</tr>
<tr>
<td>$K_Z$</td>
<td>$mmolN/m^3$</td>
<td>0.5</td>
<td>Half-saturation concentration for $Z$ grazing</td>
<td></td>
</tr>
<tr>
<td>$Q_{10}$</td>
<td>/</td>
<td>3.1</td>
<td>Temperature dependence coefficient for grazing</td>
<td>Fransz &amp; Verhagen [1985]</td>
</tr>
</tbody>
</table>

**Detrital compartment**

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>UNIT</th>
<th>VAL</th>
<th>DESCRIPTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{D,N}$</td>
<td>$day^{-1}$</td>
<td>0.05</td>
<td>Detrital remineralisation rate from $D\rightarrow N$</td>
<td>McCreary et al. [1996]</td>
</tr>
<tr>
<td>$w_p$</td>
<td>$m/day$</td>
<td>1.</td>
<td>Sinking velocity of $D$</td>
<td>Fasham et al. [1990]</td>
</tr>
</tbody>
</table>

**Table 1: Estimation of pelagic parameters.** For seek of clarity, we adopt the concise and informative convention similar to Evans & Garçon [1997]. For the biological parameters, we use the symbols: "μ" for a maximal value, "K" for a half-saturation constant, "α" for a fraction (between 0 and 1), "w" for a vertical velocity, "α" for linear and "β" for quadratic coefficients of the Taylor expansion of a generic function around zero. Furthermore, the parameter subscript is used to refer to the state variable associated with the parameter. A single subscript implies that the parameter appears in only one equation while a double subscript means that the parameter is associated with a process of exchange and therefore will appear in two equations. This automatic cross-reference is very powerful to provide a direct and clear view of the model. For example, when we see the linear loss parameter "α_{P,D}" in the phytoplankton equation [Evans & Garçon, 1997], we already know that it will also appear in the detritus equation (but with opposite sign). The biological parameter are measured in units of "mmolN" but they can be converted in the appropriate unit by using the conversion equivalence: 1 $gC = 20 mg$ Chlorophyll = 10 $mmolN$ [Henderson & Steele, 1995].
<table>
<thead>
<tr>
<th>FUNCTION ( a ) = ( G(X, a, b) = G_X(X, a, b) = G(X, a, b) )</th>
<th>TYPE</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A: ( a ) = Half-saturation abscissa (X value for which ( G = k/2 )). ( b ) = Max value.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G^1_a = b \frac{X}{X + a} )</td>
<td>L-C</td>
<td>Limitation Michaelis-Menten</td>
</tr>
<tr>
<td>( G^2_a = h \frac{X^2}{X^2 + a^2} )</td>
<td>Q-C</td>
<td>Grazing [Denmann, 1974]</td>
</tr>
<tr>
<td>( G^3_a = b \left(1 - 0.5X^{1/4}\right) )</td>
<td>L-C</td>
<td>Grazing [McGillicudy et al., 1996]</td>
</tr>
<tr>
<td>Class B: ( a ) = Initial slope (first derivative of ( G ) at ( X=0 )). ( b ) = Max value.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G^1_b = b \frac{a X}{a X + b} )</td>
<td>L-C</td>
<td>P-I curve [Denman, 1974]</td>
</tr>
<tr>
<td>( G^2_b = b \frac{a X}{\sqrt{a^2 X^2 + b^2}} )</td>
<td>L-C</td>
<td>P-I curve [Fasham et al., 1990]</td>
</tr>
<tr>
<td>( G^3_b = b \left(1 - \exp\left(-\frac{a X}{b}\right)\right) )</td>
<td>L-C</td>
<td>P-I curve [Webb et al., 1974]</td>
</tr>
<tr>
<td>( G^4_b = b \tanh\left(\frac{a X}{b}\right) )</td>
<td>L-C</td>
<td>P-I curve [Platt-Jasby, 1996]</td>
</tr>
<tr>
<td>Class C: ( a ) = Optimal value (abscissa ( X ) for which ( G ) is maximal) ( b ) = Maximal value.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G^1_c = b \frac{2(X/a)}{(X/a)^2 + 1} )</td>
<td>L-H</td>
<td>P-I curve [Eilers &amp; Peeters, 1988]</td>
</tr>
<tr>
<td>( G^2_c = b \frac{4(X/a)}{(X/a)^2 + 2(X/a) + 1} )</td>
<td>L-H</td>
<td>P-I curve [Eilers &amp; Peeters, 1988]</td>
</tr>
<tr>
<td>( G^3_c = b X \frac{a}{a} \exp\left(1 - \frac{X}{a}\right) )</td>
<td>L-H</td>
<td>P-I curve [Steele, 1962]</td>
</tr>
<tr>
<td>Class D: ( a ) = Asymptotic linear coefficient (first derivative of ( G )). ( b ) = Asymptotic quadratic coefficient (second derivative of ( G )).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G^1_d = a X + b X^2 )</td>
<td>L-Q</td>
<td>Mortality [Doney et al., 1996]</td>
</tr>
<tr>
<td>( G^2_d = a b X^2 \frac{X}{X + a} )</td>
<td>Q-L</td>
<td>Mortality [Orange, 1994]</td>
</tr>
<tr>
<td>( G^3_d = a X \left{1 - e^{-\frac{b X}{a}}\right} )</td>
<td>Q-L</td>
<td>Grazing [Doney et al., 1996]</td>
</tr>
</tbody>
</table>

**Table 2:** Set of functions of Class "M" and Type "N". The function "\( G \)" is described by (a) its Class "M" that refers to the nature of parameters "\( a \)" and "\( b \)" and by (b) its Type "N" that refers to the asymptotic behaviour of the function "\( G \)" for small/high values of "\( X \)" (it can be either Linear "L", Quadratic "Q", Constant "C" or Hyperbolic "H").