ICES Journal of Marine Science

ICES Journal of Marine Science; doi:10.1093/icesjms/fsv204

Recovery of photosynthesis parameters from *in situ* profiles of phytoplankton production

Žarko Kovač[1]*, Trevor Platt[2], Shubha Sathyendranath[2], Mira Morović[1], and Thomas Jackson[2]

[1]*Physical Oceanography Laboratory, Institute of Oceanography and Fisheries, Split 21000, Croatia*
[2]*Remote Sensing Group, Plymouth Marine Laboratory, Devon PL1 3DH, UK*

**Corresponding author. e-mail: kovac@izor.hr*

Kovač, Ž., Platt, T., Sathyendranath, S., Morović, M., and Jackson, T. Recovery of photosynthesis parameters from *in situ* profiles of phytoplankton production. – ICES Journal of Marine Science, doi: 10.1093/icesjms/fsv204.

Received 2 April 2015; revised 15 October 2015; accepted 17 October 2015.

We examine a model of the rate of phytoplankton production in the ocean and its dependence on depth. The model is analysed as a function of photosynthesis parameters and it is shown that: (i) production profiles with depth are determined uniquely by the parameter values; (ii) daily water column production is not uniquely determined by the parameter values; (iii) a unique combination of parameters exists for which the model best fits a measured production profile. An inverse procedure is developed to recover photosynthesis parameters from measured profiles of primary production, and its performance tested by application to profiles of primary production collected at the Hawaii Ocean Time Series. For each profile tested, the method is successful in recovery of the photosynthesis parameters. The method can be applied to the estimation of photosynthesis parameters from data on *in situ* production profiles, which have been collected globally for more than half a century, thereby augmenting the world archive of these parameters for application in ecosystem modelling and estimation of primary production from remotely sensed data.

Keywords: model analysis, model optimization, photosynthesis parameters, phytoplankton production.

Introduction

In the history of measurements of primary production in the ocean, an important milestone was the introduction of the C14 method in 1952, which allowed precise estimates of photosynthetic rate to be made at sea (Steemann Nielsen, 1952; Barber and Hilting, 2002). At the beginning, the typical implementation of the method was to construct profiles of *in situ* production through the photic zone, yielding an estimate of primary production for the entire water column and for the period of incubation with the tracer (Peterson, 1980). Many such experiments were performed, with the result that there is a rich archive of data on water column production by the *in situ* method from stations scattered throughout the world ocean (Buitenhuis *et al.*, 2013).

An alternative implementation of the tracer method is to incubate the samples in a controlled light gradient so that the photosynthesis response curve (light saturation curve) can be established, and the photosynthesis parameters recovered (Platt and Gallegos, 1980). Although this method has more generality than the *in situ* method, and is now more usually preferred, the world archive on primary production in the ocean still has far more entries of data collected by the *in situ* method than by any other.

In several areas of biological oceanography, the photosynthesis parameters are essential data. For example in the estimation of primary production using data on visible radiometry collected by instruments in Earth orbit (ocean colour), an essential step is the assignment of the photosynthesis parameters (Platt and Sathyendranath, 1988, 1993). A similar assignment step is required for numerical models of the marine ecosystem (Platt and Sathyendranath, 1991; Franks, 2002; Gentleman, 2002). But at the world scale, there is a lamentable paucity of *in vitro* data, given the historical dominance of the *in situ* method (Buitenhuis *et al.*, 2013).

The question then arises whether the photosynthesis parameters might be recovered from data collected by the *in situ* method, using an inverse procedure? If that were possible, we could exploit the more-numerous *in situ* databases to augment the sparse database on photosynthesis parameters recovered from the photosynthesis light curve. Here, we establish such an inverse procedure. We show that it is possible to recover the photosynthesis parameters

from *in situ* profiles of primary production, and we demonstrate the utility of the method by applying it to profiles of primary production collected at the Hawaii Ocean Time Series (HOT). Using the retrieved parameters, the water column production for the HOT data could be calculated to within 10% of the measured value. Of course, the retrieved parameters could also be employed in other applications, such as the estimation of water column production from remotely-sensed data on ocean colour, as we illustrate with an example, and as input data in numerical models of the marine ecosystem.

Matrix model

It is convenient, and computationally efficient, to introduce a matrix formalism for the calculations in this paper. The formalism is essentially a discrete version of the analytical production model put forward by Platt *et al.* (1990). In the analytical model phytoplankton primary production is modelled by expressing mathematically the photosynthesis–irradiance relationship through the use of the photosynthesis–irradiance function $p^B(E)$. The $p^B(E)$ function gives the amount of carbon assimilated, per unit time and unit biomass, as a function of available irradiance. Primary production at depth z and time t is then expressed as $P(z,t) = B(z)p^B(E(z,t))$, where $B(z)$ is the time-independent biomass and $E(z,t)$ is the irradiance. This expression is integrated over time and depth to yield the production of the water column. Here we cast the established analytical model into a matrix form.

Let the model have N vertical levels at depths z_n, labelled by the index n and J time intervals labelled by the index j. The depth z is taken positive downwards and the index n increases with depth such that $z_n < z_{n+1}$. The total water column depth is calculated as follows: $Z = Z_N + (Z_N - Z_{N-1})/2$. Each time interval equals $\Delta T = D/J$, where D is the integration period. We label the light intensity at depth z_n at time $j\Delta T$ as $E_{nj} = E(z_n, j\Delta T)$ and collect all the light intensities into an irradiance matrix $\mathbf{E}$ (size $N \times J$), whose elements equal E_{nj}. A total description of the light conditions in the model is thus gathered into a single matrix. Now the photosynthesis light function $p^B(E)$ (Platt and Jassby, 1976) is used on each element of the irradiance matrix to calculate the instantaneous rates of production normalized to phytoplankton biomass (indexed as chlorophyll concentration B): $p^B_{nj} = p^B(E_{nj})$. More formally, $p^B[\cdot]$ can be regarded as an elementwise operator that acts on the irradiance matrix to give the biomass-normalized production matrix, and we write $\mathbf{P}^B = p^B[\mathbf{E}]$. Each row of this matrix equals the time evolution of biomass-normalized production at the depth z_n and each column equals the vertical profile of biomass-normalized production at time $j\Delta T$. To calculate the absolute values of production $p_{nj} = B(z_n)\, p^B(E_{nj})$, each row of the biomass-normalized production matrix has to be multiplied by the corresponding biomass value. First the biomass vector $\mathbf{b}$ (size $N \times 1$) is defined, such that each element in it equals the biomass at a certain depth $b_n = B(z_n)$, where $B(z)$ is the biomass profile. Next the biomass matrix $\mathbf{B}$ (size $N \times N$) is defined as a diagonal matrix $B_{nm} = \delta_{nm} b_n$ (where δ_{nm} is the Kronecker delta), which holds the values of the biomass profile on its main diagonal. Multiplying the biomass matrix with the biomass-normalized production matrix, we obtain the production matrix:

$$\mathbf{P} = \mathbf{B}\mathbf{P}^B. \tag{1}$$

The production matrix holds all the information that is needed to calculate the daily production profile $\mathbf{p}_T$ (size $N \times 1$), or the daily water column production $P_{Z,T}$. By defining the time vector $\boldsymbol{\tau}$ (size $J \times 1$), whose elements are all equal $\tau_j = \Delta T$, temporal integration is achieved simply by:

$$\mathbf{p}_T = \mathbf{P}\boldsymbol{\tau}, \tag{2}$$

and the daily production profile calculated. To calculate $P_{Z,T}$, vertical integration of $\mathbf{p}_T$ is required. We define the vertical increment vector as a row vector $\boldsymbol{\zeta}$ (size $1 \times N$) whose elements equal vertical increments around each depth $\zeta_n = \Delta z_n$, with $\Delta z_n = (z_{n+1} - z_{n-1})/2$ for $n = 2,3,\ldots,N-1$. The first vertical increment is $\Delta z_1 = (z_1 + z_2)/2$, and the last is $\Delta z_N = z_N - z_{N-1}$. For $\boldsymbol{\zeta}$ we have $\sum_{n=1}^{N} \zeta_n = Z$. Daily water column production is now given by:

$$P_{Z,T} = \boldsymbol{\zeta}\mathbf{P}\boldsymbol{\tau}. \tag{3}$$

Equations (1), (2) and (3) form the basis of the matrix model for calculating primary production. To calculate the irradiance matrix, a light model is required to supply the values of E_{nj}. Here we use a simple model of the form:

$$E_{nj} = E_{0j}\, \exp(-Kz_n), \tag{4}$$

where E_{0j} marks the surface PAR at time $j\Delta T$ and K is the diffuse attenuation coefficient for downwelling irradiance (Kirk, 2011).

Optimization

The model presented above is implemented in an optimization procedure in which the model results are compared with measurements to extract optimal values of model parameters for given measurements. The optimization procedure should provide an estimate of the parameters for which the model results best match the measurements under the given model assumptions. A basic treatment of optimization problems in oceanography can be found in Glover *et al.* (2011), while a more advanced treatment of inverse modelling and data assimilation is found in Wunsch (1996) and Bennett (2005).

The key element in this model is the $p^B(E)$ parameterization, which is determined uniquely by two parameters, the initial slope α^B and the assimilation number P^B_m (Platt *et al.*, 1977). Their ratio equals the photoadaptation parameter $E_k = P^B_m/\alpha^B$. The $p^B(E)$ function can be written as $p^B(E|\alpha^B, P^B_m)$ to indicate that it is a function of light intensity depending on two parameters. Let us assume that we have measurements of the biomass profile $\widetilde{\mathbf{b}}$, attenuation coefficient $\widetilde{K}$, and surface PAR $\widetilde{E}_{0j}$, where $j = 1,2,\ldots,J$. We denote a measured value of a variable x as $\widetilde{x}$. These measurements also give information on the integration length D, model depths z_n, and number of time increments J. From the biomass vector $\mathbf{b}$ we have $\mathbf{B}$. Using the optical model (4) we can calculate the irradiance matrix $\mathbf{E}$. We can then treat the model equation (2) as being dependent only on α^B and P^B_m and write:

$$\mathbf{p}_T(\alpha^B, P^B_m) = \mathbf{B}\mathbf{P}^B(\alpha^B, P^B_m)\,\boldsymbol{\tau}. \tag{5}$$

Let us also assume that we have a measured daily production profile consisting of incubations at N depths which have been carried out over the incubation period D. We gather all the measurements into a vector $\widetilde{\mathbf{p}}_T$ (size $N \times 1$) with each element given by $\widetilde{P}_T(z_n)$. The notation $\widetilde{P}_T(z_n)$ is used to indicate the measured daily production at depth z_n. The model production given by (5) can be indicated by $\mathbf{p}_T(\alpha^B, P^B_m)$. Subtraction of the measured from the modelled

profile yields the error profile:

$$\Delta\mathbf{p}_T(\alpha^B, P_m^B) = \mathbf{p}_T(\alpha^B, P_m^B) - \widetilde{\mathbf{p}}_T, \tag{6}$$

which gives the difference between the modelled and the measured production at each depth. The error depends only on α^B and P_m^B since everything else is fixed. If we wish the model profile to match as closely as possible the measured profile, we need to minimize the difference between the two. The Euclidian norm of the error profile $||\Delta\mathbf{p}_T||$ is taken as a measure of the difference and it is written as a function of the parameters:

$$\mathcal{P}(\alpha^B, P_m^B) = ||\Delta\mathbf{p}_T(\alpha^B, P_m^B)||. \tag{7}$$

Fixing $\widetilde{\mathbf{b}}, \widetilde{K}, \widetilde{E}_{0t}$, and $\widetilde{\mathbf{p}}_T$ by measurement gives the error as a function of α^B and P_m^B. To find the smallest error, the minimum of $\mathcal{P}$ has to be found in the space spanned by α^B and P_m^B (parameter space). At the point where $\mathcal{P}$ reaches its minimum value, the parameters reach their optimal values (under given assumptions of the model) and the model profile matches the measured profile as well as it can.

To give a measure of the accuracy of the estimated parameters that is not dependent on the specific application relative model errors are needed. We define the relative model error of daily water column production as:

$$R(\alpha^B, P_m^B) = \frac{P_{Z,T}(\alpha^B, P_m^B)}{\widetilde{P}_{Z,T}} - 1, \tag{8}$$

which gives the percentage error in daily water column production. We also define the relative profile error as:

$$r(\alpha^B, P_m^B) = \frac{\mathcal{P}(\alpha^B, P_m^B)}{||\widetilde{\mathbf{p}}_T||}, \tag{9}$$

which relates the error function to the norm of the measured daily production profile $||\widetilde{\mathbf{p}}_T||$. This error gives the relative distance between the modelled and the measured profile and shows how much the two profiles resemble each other. These two errors (8, 9) enable the comparison of model fitnesses resulting from the application of the inverse procedure to different sets of measurements.

In cases where the model assumptions are not strongly violated, we want to find the values of parameters that give as low a value of the error $\mathcal{P}$ as possible. To do so all that remains to be done is to choose an optimization algorithm to carry out the search for the minimum. Irrespective of the algorithm, the final result should be the optimal parameter combination $O = (\hat{\alpha}, \hat{P}_m)$ for which $\mathcal{P}$ has a minimal value. The optimization procedure indicates which one of the variety of model profiles best describes the measured profile. That the profiles are indeed determined by the photosynthesis parameters is shown in Appendix.

Parameter space analysis

In real applications the values $\widetilde{\mathbf{b}}, \widetilde{K}, \widetilde{E}_{0t}$, and $\widetilde{\mathbf{p}}_T$ will be given by measurements, and the goal is to estimate α^B and P_m^B based on these measurements. Because the search is performed on only two parameters, we can view the model results as being dependent only on α^B and P_m^B. In the plane spanned by α^B and P_m^B, a plot of daily production $P_{Z,T}(\alpha^B, P_m^B)$ yields contour lines that do not intersect (Figure 1a). The same value for $P_{Z,T}$ can be obtained through the entire range of parameter values, and there appears to be no constraint on

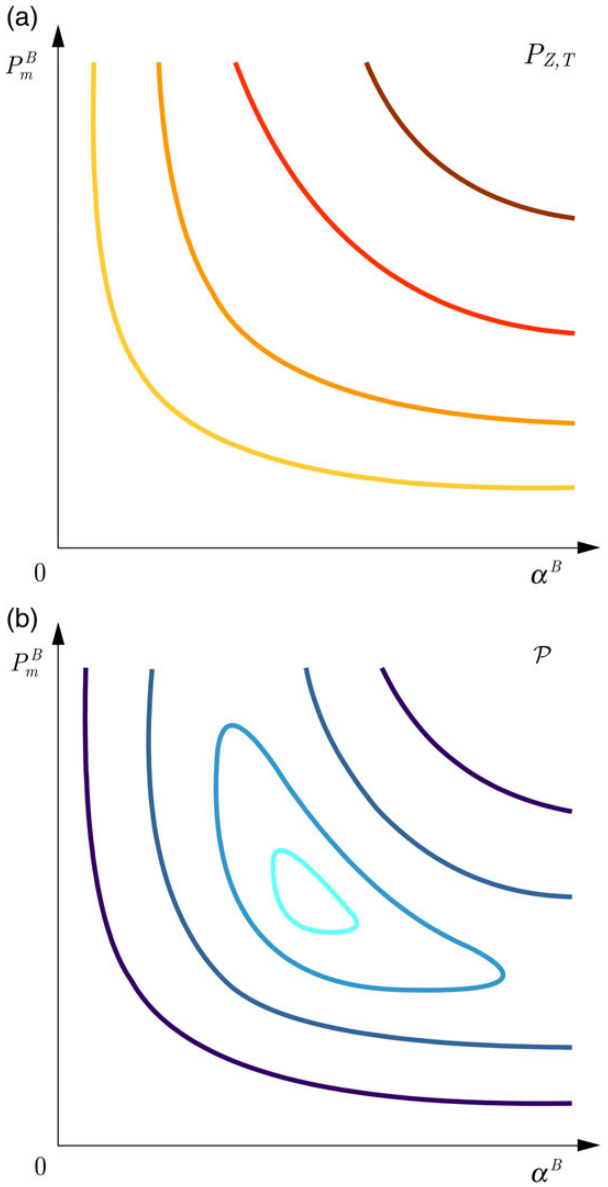

Figure 1. (a) Contour plot of daily water column production $P_{Z,T}$ in the parameter space. Contours in darker colour have higher values of $P_{Z,T}$. (b) Contour plot of the error function P^{ER} in the parameter space. Contours in darker colour have higher values of P^{ER}. An area with low values of P^{ER} surrounding the minimum can be seen.

them. If the model is to give a certain value for $P_{Z,T}$, any combination of parameters that lies on that contour line will do just as well as any other. This emphasizes that the measurement of daily water column production does not enable the determination of the parameters. Nor does it allow us even to constrain the range of possible parameter values.

For the parameters to become uniquely determined, additional information needs to be taken into account. It is provided by $\widetilde{\mathbf{p}}_T$, which holds information about the profile of daily production. The influence of the measured production profile is taken into account by the error function $\mathcal{P}(\alpha^B, P_m^B)$. When the error function $\mathcal{P}(\alpha^B, P_m^B)$ is plotted in the plane spanned by α^B and P_m^B, the plot also yields contour lines that do not intersect (Figure 1b). But unlike the contour lines of $P_{Z,T}$, some of the contour lines of $\mathcal{P}$ are

closed. The closed contour lines enclose other contour lines with lower values of $\mathcal{P}$, implying the existence of an area in the parameter space where the model performs better than outside it. All the points in the area enclosed by a certain contour are those for which the model outperforms the ones in the area outside it. Closure of contour lines also implies setting ranges on α^B and P_m^B, which could not have been done for the case of $P_{Z,T}$.

To see why this occurs, let us rewrite the error function slightly. It is related to the scalar product between the modelled and the measured profiles $\tilde{\mathbf{p}}_T \cdot \mathbf{p}_T$, under the following expression:

$$\mathcal{P} = \sqrt{||\mathbf{p}_T||^2 - 2\tilde{\mathbf{p}}_T \cdot \mathbf{p}_T + ||\tilde{\mathbf{p}}_T||^2}, \tag{10}$$

where $\mathbf{p}_T = \mathbf{p}_T(\alpha^B, P_m^B)$ and $\mathcal{P} = \mathcal{P}(\alpha^B, P_m^B)$. The higher the value of the scalar product $\tilde{\mathbf{p}}_T \cdot \mathbf{p}_T$, the smaller the value of $\mathcal{P}$. It is obvious that for a particular value of $\mathcal{P}$ there will be more than one $\mathbf{p}_T$. In fact all the profiles with the same Euclidean distance from $\tilde{\mathbf{p}}_T$ form a hypersphere around the measured profile. What is seen in the parameter space are the contours joining the values of the parameters that correspond to these model profiles on a specific hypersphere. That is why there are closed contour lines of $\mathcal{P}$ in the parameter space. As the Euclidean distance of the model profile $\mathbf{p}_T$ from the measured profile $\tilde{\mathbf{p}}_T$ gets smaller, the volume inside the hyphersphere becomes smaller and there are fewer parameter combinations remaining on the hyphersphere. That is the reason for the surfaces enclosed by contours in the parameter space having a lesser extent as $\mathcal{P}$ becomes smaller. By requiring $\mathcal{P}$ to go to zero we also require $\tilde{\mathbf{p}}_T \cdot \mathbf{p}_T$ to reach its highest value. Ideally, if the model could fit the measured profile perfectly, the error function would equal zero. In that case the scalar product $\tilde{\mathbf{p}}_T \cdot \mathbf{p}_T$ reaches a maximum and the right-hand side of (10) vanishes. The ideal case would be achieved if we took a model profile with a pre-specified set of parameters, instead of a measured profile, and then formed the error.

Taking both considerations into account, namely one contour line of $P_{Z,T}$ and one of $\mathcal{P}$, we see that they intersect at two points (Figure 2). Let these points be labelled by X and Y. Since X and Y belong to both contour lines the following holds: $P_{Z,T}(X) = P_{Z,T}(Y)$ and $\mathcal{P}(X) = \mathcal{P}(Y)$. This means that the model will give the same daily production and the same error for two different combinations of parameters. Knowing that some of the contour lines of $\mathcal{P}$ are closed, whereas none of the $P_{Z,T}$ are, when we consider smaller and smaller values of $\mathcal{P}$ we see that the two points, X and Y, converge to the point O, which is then the minimum of $\mathcal{P}$. That point is the optimal estimate of parameters for the given measured daily production profile $\tilde{\mathbf{p}}_T$. The contour line of $P_{Z,T}$ to which O belongs will not necessarily correspond to the contour line of $\tilde{P}_{Z,T}$. Setting the condition that $P_{Z,T}(O) = \tilde{P}_{Z,T}$ could be considered as a constraint on the search in the parameter space. Owing to the nature of the measurement of primary production, that constraint is not necessary, and variation of $P_{Z,T}(O)$ around $\tilde{P}_{Z,T}$ is allowed.

Limitations

The first limitation of the inverse method comes from the production model we use. Because the photosynthesis parameters are assumed to be independent of depth, $\alpha^B \neq \alpha^B(z)$ and $P_m^B \neq P_m^B(z)$, and leaving aside the possible influence of photoinhibition, the normalized daily production has to decline with depth, which is expressed as:

$$P_T^B(z_n) > P_T^B(z_{n+1}). \tag{11}$$

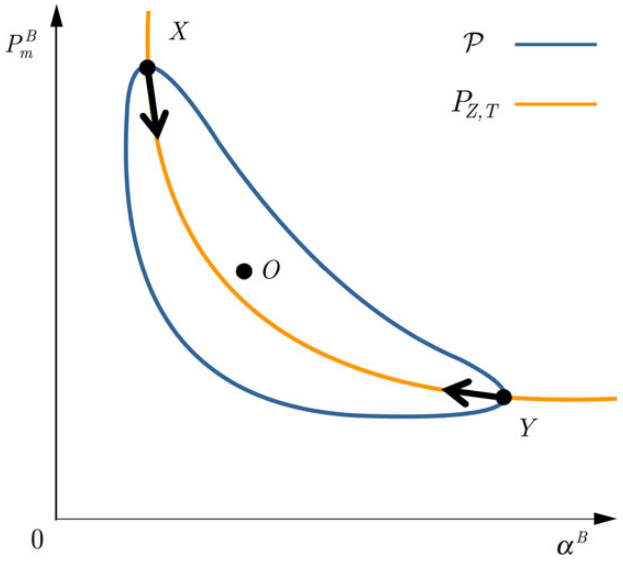

Figure 2. Intersecion of two contour lines, one belonging to $\mathcal{P}$ and the other to $P_{Z,T}$. The value of $\mathcal{P}$ is arbitrary, while the value of $P_{Z,T}$ equals the measured daily water column production $\tilde{P}_{Z,T}$. Reducing the error function value causes the points X and Y to converge onto the optimal point O. That is the point with the smallest value of $\mathcal{P}$.

The validity of this condition can be proved as follows. Let us consider the irradiances at any two vertical levels z_n and z_{n+1}, at time $j\Delta T$. The corresponding elements of the irradiance matrix are E_{nj} and $E_{(n+1)j}$. Due to the decrease of irradiance with depth, it is always valid that the irradiance at depth z_n is higher than the irradiance at depth z_{n+1}. For the light model used we have $E_{(n+1)j} = E_{nj} \exp(-K(z_n - z_{n+1}))$. Therefore, for any two elements of the irradiance matrix, it holds that $E_{nj} > E_{(n+1)j}$. The production normalized to biomass is calculated for these two irradiances as $p_{nj}^B = p^B(E_{nj})$ and $p_{(n+1)j}^B = p^B(E_{(n+1)j})$. The $p^B(E)$ is a positive increasing function for $E > 0$, has a negative curvature $\partial^2 p^B(E)/\partial E^2 < 0$, and an asymptote $\lim_{E\to\infty} p^B(E) = P_m^B$ (Platt *et al.*, 1977). For $E_{nj} > E_{(n+1)j}$ it gives $p_{nj}^B > p_{(n+1)j}^B$. Total daily production normalized to biomass $P_T^B(z_n)$ is a sum of p_{nj}^B over $j = 1, 2, \ldots, J$. Comparing $P_T^B(z_n)$ with $P_T^B(z_{n+1})$ we see that every element of the sum for the z_n level is greater than the corresponding element of the sum for the z_{n+1} level. Therefore, we conclude that the daily production normalized to biomass at level z_n will be greater than the daily production normalized to biomass at level z_{n+1}, and that is what condition (11) says. In simple terms, for the case of vertically uniform parameters the profile of production normalized to biomass always decreases with increasing depth.

Vertical uniformity of parameters can be expected for a mixed layer. For the case of vertical non-uniformity of parameters, the condition (11) can be violated and there is no reason to expect a model with vertically uniform parameters to perform well. We can justify the application of our procedure by considering whether or not the measurements were carried out in the mixed layer. We can also justify the procedure by the goodness of fit between the model and the measurements, given by the value of the relative profile error r, and by considering the relative model error of daily water column production R. If the condition (11) is satisfied by the measured profile of production normalized to biomass, and

the errors are low, we can be confident in the recovered optimal values of parameters.

In case the parameter values increase with depth, the condition (11) may be violated. If this happens we are certain that the parameters have changed in the vertical, because nothing else could cause an increase in the biomass-normalized production with depth. By simply calculating the differences for the measurements $\widetilde{P}_T^B(z_n) - \widetilde{P}_T^B(z_{n+1})$, with $n = 1,2,\ldots,N-1$, we can see whether or not the parameters have indeed changed with depth. If the difference is positive, they are either fixed of decreasing, but in case the difference is negative they are increasing. When relying on a database of productivity profiles, the condition (11) can be used as an easy test for checking the justifiability of applying the inverse procedure to obtain the values of parameters.

The other limitation of the inverse method comes from the relation between irradiance and the photoadaptation parameter $E_k = P_m^B/\alpha^B$. For low values of irradiance $E \ll E_k$, the $p^B(E)$ function is determined by the initial slope α^B, and as a first approximation we have $p^B(E) \approx \alpha^B E$. For high values of irradiance $E \gg E_k$, $p^B(E)$ is determined only by P_m^B and we have $p^B(E) \approx P_m^B$. In the intermediate range where E is around E^k, the $p^B(E)$ function is determined by both parameters. For a given irradiance matrix $\mathbf{E}$, let us label the highest value of irradiance that it contains as $E_{\max}$, and the lowest value, different from zero, as $E_{\min}$. A zero value for irradiance does not interest us because it results in no production in our model. We distinguish three possibilities (Figure 3):

(a) $E_{\max} < E_k$

In this case the biomass-normalized production matrix $\mathbf{P}^B$ is determined only by α^B and we can write $\mathbf{P}^B \approx \alpha^B\mathbf{E}$. The model daily production profile equals $\mathbf{p}_T = \alpha^B\mathbf{BE}\boldsymbol{\tau}$. As a consequence we can determine only α^B by the inverse procedure. Any value for P_m^B can be taken because it has no influence on $\mathbf{p}_T$. This happens in the area of the parameter space where the parameter combinations yield a high value for E_k and is the reason why the contour lines of $\mathcal{P}$ and $P_{Z,T}$ are parallel with the ordinate. In this region we obtain high values of P_m^B and low values of α^B.

(b) $E_{\min} < E_k < E_{\max}$

For E_k in this range of irradiances both of the parameters can be determined. Model daily production profile is given by (2) and depends on both parameters. This occurs throughout most of the parameter space and the parameter combinations yield realistic values for E_k, that are also expected to be obtained as a result of measurements of the $P-E$ curve.

(c) $E_k < E_{\min}$

In this case we are confronted with the situation in which the biomass-normalized production matrix $\mathbf{P}^B$ can be written in first approximation as $\mathbf{P}^B \approx P_m^B\mathbf{U}$, where $\mathbf{U}$ is a unit matrix of the same size as $\mathbf{P}^B$. The model daily production profile now equals $\mathbf{p}_T = P_m^B\mathbf{BU}\boldsymbol{\tau}$, and we can determine P_m^B only, because α^B has a low influence on $\mathbf{p}_T$. Unlike the first case (a), this occurs in the area of the parameter space where the parameter combinations yield a low value for E_k. That happens for combinations of low P_m^B with high α^B and explains why the contour lines of $\mathcal{P}$ and $P_{Z,T}$ are parallel with the abscissa in this region.

When these conditions will be satisfied depends on the model irradiances. The lowest value of irradiance will occur at the greatest model depth, at the time of lowest surface irradiance. This value can become higher than the photoadaptation parameter (condition (c)) for shallow clear waters and incubations carried out in the interval around noon, on a cloud-free day. Surface irradiance will be high, saturation will occur during the whole interval of incubation and production will not be light-limited at any depth. Only P_m^B can be determined in this case.

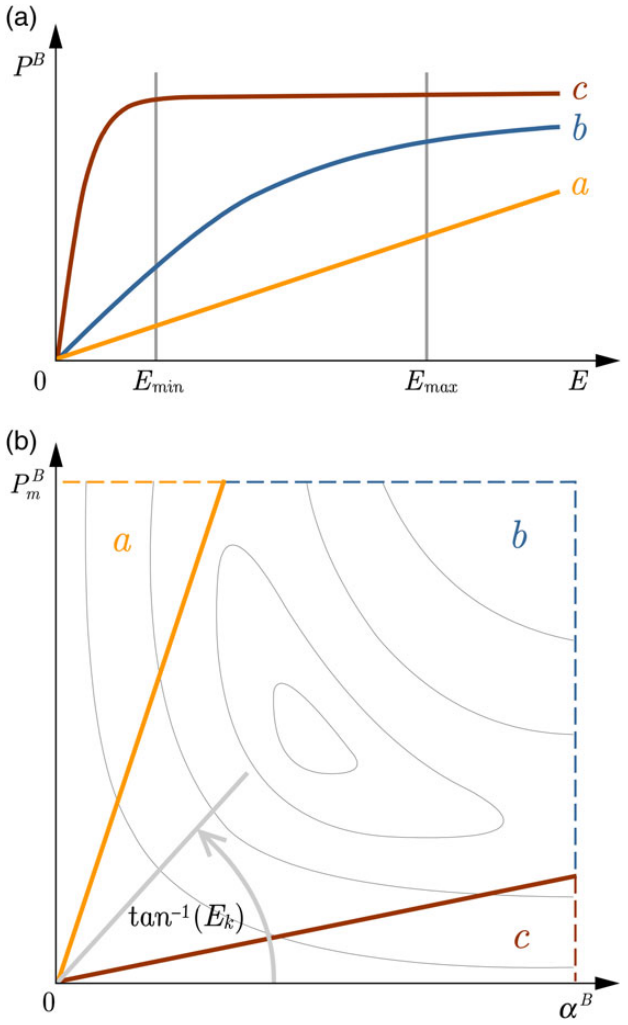

Figure 3. (a) Plots of $p^B(E)$ curves in accordance with the second limitation of the method. (b) Division of the parameter space into three areas based on the ability of the method to recover the parameters. In the a area, the method can only recover α^B, whereas in the c area, only P_m^B. Both parameters can be recovered in the b area. In this plot, E_k equals the slope of the line passing through the origin. Contours of $\mathcal{P}$ are shown in the background.

At the other extreme, the highest value of irradiance will occur at the shallowest model depth, at the time of highest surface irradiance. This value can become smaller than the photoadaptation parameter (condition a) if the incubations are carried out on an overcast day. The effect will be strengthened still further if the water is turbid. Light will be strongly attenuated and production will be light-limited at all model depths. Then, only α^B can be determined. This is likely to be the case in deep eutrophic waters.

For all the remaining cases (and these are in fact the vast majority of cases encountered), the parameters can be recovered successfully.

In practice, for incubations carried out over the entire course of the day, light conditions will provide production values that range both below saturation and above it. Light saturation will occur at shallow depths and those measured values of production will be used for estimation of P_m^B. The values of production at greater depths will be used to estimate α^B. If we are to recover both parameters well, phytoplankton should experience the whole range of light intensities, from light limiting to saturating.

Applications

We tested the inverse method on a dataset from the Hawaii Ocean Time Series (HOT). The data are publicly available at hahana.soest.hawaii.edu/hot/hot-dogs/. Details about the HOT data can be found in Karl and Lukas (1996) and Karl *et al.* (2001). Some recent publications using HOT primary production data include Luo *et al.* (2012) and Nicholson *et al.* (2012). We found that data from 169 HOT cruises were suitable for testing the inverse method. At the time we accessed the dataset it consisted of 237 cruises, meaning that we used 71% of the available data. The data are not seasonally clustered. Starting from January the number of profiles per month is 15, 16, 13, 15, 14, 11, 12, 13, 15, 19, 14, 12. All of cruise data we used had measurements of surface PAR, carried out at 10 min intervals during the entire incubation period D. Surface PAR was given in μE m^{-2} s^{-1} and the conversion to W m^{-2} was done using Smith and Morel's procedure (Morel and Smith, 1974). Most of the incubations were carried out from sunrise to sunset. Sampling depths were at 5, 25, 45, 75, 100, 125, 150 and 175 m. All the chlorophyll profiles were measured at 8 depths, 75 production profiles were measured at 8 depths, and 94 at 6 depths (lacking measurement at 150 and 175 m). One per cent light level was given for 128 cruises, from which the diffuse attenuation coefficient was calculated. For the remaining cases the average value of 0.0436 m^{-1} for K was used. As for the remainder of the HOT dataset it had no surface PAR, or optical measurements required by the method. In other words, the 29% of the profiles we did not use were incomplete, within the requirements of the model.

Before applying the inverse procedure we tested all the normalized production profiles against the condition given by (11). For 79 profiles it was violated at the first vertical level. We assume that that is due to photoinhibition, which is not accounted for in this model. The error in parameter estimation caused by this violation is small, because the effect is limited only to the first vertical level. Some profiles violate this condition at greater depths, but this occurs only rarely.

Let us denote each cruise with l, where $l = 1,2,\ldots,169$ (these numbers do not correspond to the HOT cruise numbers). The given data allowed formulation of the error function (7) for each cruise $\mathcal{P}_l(\alpha^B, P_m^B)$, and with it the recovery of optimal values of parameters $(\hat{\alpha}^B, \hat{P}_m^B)_l = O_l$ and the optimal value of the photoadaptation parameter $(\hat{E}_k)_l$, as their ratio. Measured chlorophyll profiles gave $\tilde{\mathbf{b}}_l$ and consequently $\mathbf{B}_l$, whereas production profiles gave $(\tilde{\mathbf{p}}_T)_l$. Using measured surface PAR $(\tilde{E}_{0t})_l$, together with $\tilde{K}_l$, we calculated the irradiance matrix for each cruise $\mathbf{E}_l$. It is worth mentioning that the irradiance matrix $\mathbf{E}_l$ is calculated only once for each cruise. To calculate the production matrix $\mathbf{P}_l$ we took the equation of Platt *et al.* (1980):

$$p^B(E) = P_m^B\left(1 - \exp\frac{-\alpha^B E}{P_m^B}\right). \tag{12}$$

The production matrix for each cruise:

$$\mathbf{P}_l = \mathbf{B}_l\ p^B[\mathbf{E}_l|\alpha^B, P_m^B], \tag{13}$$

is recalculated every time the optimization algorithm requires the calculation of the error function $\mathcal{P}_l(\alpha^B, P_m^B)$. Finally, the Nelder–Mead method (Nelder and Mead, 1965) was used as an optimization method for locating the minimum of the error function $\mathcal{P}_l$. A description of the mathematical properties of the method in low dimensional cases can be found in (Lagarias *et al.*, 1998). Lewis *et al.* (2000) give a survey of a number of classical direct search methods for unconstrained optimization, including the Nelder–Mead method. In our application values for the reflection, expansion, contraction, and shrink coefficient were set to 1, 2, 0.5, and 0.5, respectively. These are the standard values for the Nelder–Mead method. Optimization was performed without any additional constraints and convergence to the optimal point O_l was achieved for each cruise. Examples of optimal model profiles and measured profiles are given in Figure 4. Model profiles are calculated with optimal values of parameters for each cruise $\mathbf{p}_T(O_l)$ by using (5).

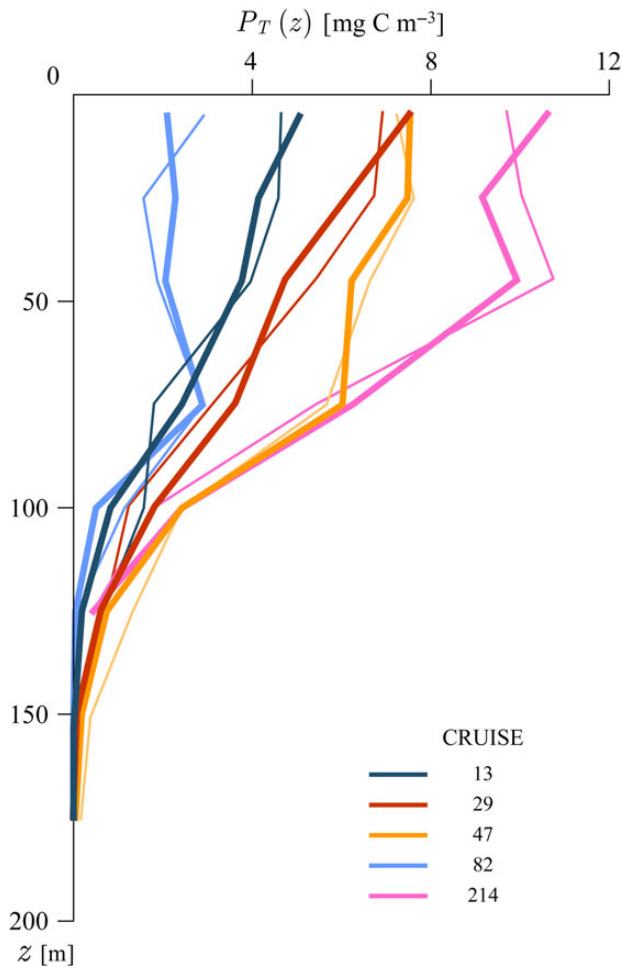

Figure 4. Examples of optimal model profiles (thick lines) and measured profiles (thin lines) for HOT cruises: 13, 29, 47, 82, and 214. Optimal values of parameters for these cruises are: $\alpha^B = 0.25$, $P_m^B = 4.33$ (cruise 13); $\alpha^B = 0.01$, $P_m^B = 2.88$ (cruise 82); $\alpha^B = 0.24$, $P_m^B = 9.57$ (cruise 29); $\alpha^B = 0.28$, $P_m^B = 7.53$ (cruise 47) and $\alpha^B = 0.33$, $P_m^B = 10.02$ (cruise 214). α^B is given in mg C m^2 (mg Chl W h)$^{-1}$ and P_m^B is given in mg C (mg Chl h)$^{-1}$.

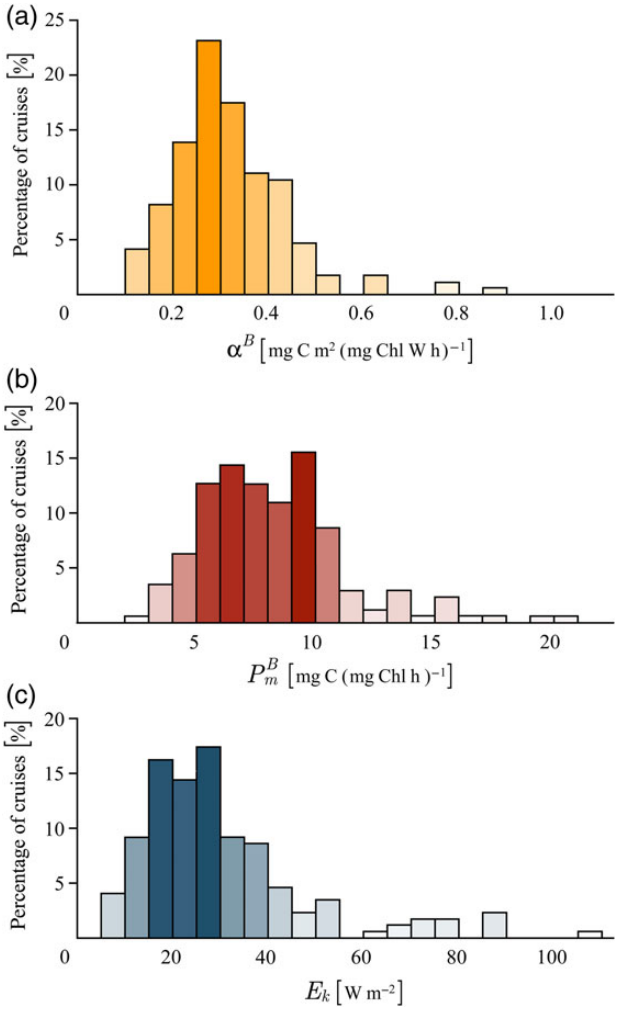

Figure 5. Histograms of recovered parameter values for all the cruises. Abscissa corresponds to parameter values and the ordinate gives percentage of cruises that fall into a certain interval of parameter values. (a) Distribution of the initial slope α^B. (b) Distribution of the assimilation number P_m^B. (c) Distribution of the photoadaptation parameter E_k.

The distribution of recovered parameters and the distribution of the photoadaptation parameter are shown in Figure 5. There are some outliers, but the distributions resemble a normal distribution. By careful inspection of the cruise data for which the outliers occur, we see that they correspond to cases for which the normalized production profiles have peaks at depth. These peaks violate condition (11), pointing to vertical non-uniformity of parameter values. Parameters attain high values as the optimization procedure attempts to fit the measured profile of production. Our interest here was not to interpret this particular dataset to provide an explanation for these values, but only to show the efficiency of the inverse method in parameter recovery. To interpret these values fully, a careful analysis of each cruise from an oceanographic perspective would be required, which is beyond the scope of this paper.

It is worth showing the relative model error of daily water column production (8), and the relative error of the model profile (9), when the model is run with optimal parameter values for each cruise. The optimal parameter combination O_l is taken in calculation of the production matrix for each cruise separately $\mathbf{P}_l = p^B[\mathbf{E}_l|O_l]$. Then, the

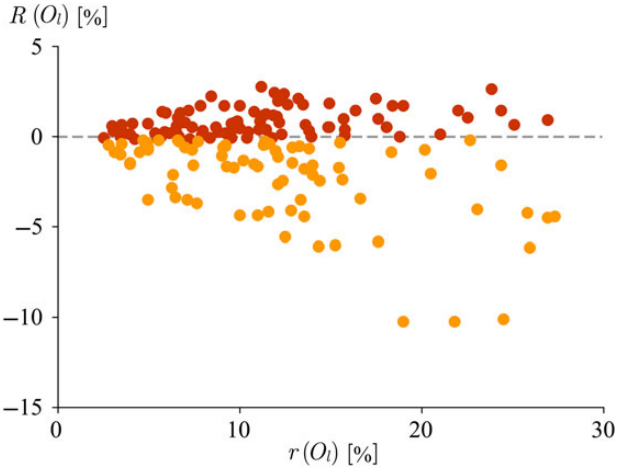

Figure 6. Plot of $R(O_l)$ vs. $r(O_l)$. Points above the dashed horizontal line represent those model results for which $\widetilde{P}_{Z,T}$ is overestimated by the model. Points below the line are those for which $\widetilde{P}_{Z,T}$ is underestimated.

errors $R(O_l)$ and $r(O_l)$ are calculated (8, 9). These are shown in Figure 6, expressed as percentages. For 85% of cruises the value of $P_{Z,T}$ does not exceed the measured daily water column production $\widetilde{P}_{Z,T}$ by more than $\pm 5\%$, with the relative profile errors $r(O_l)$ below 20% of the norm of the measured production profile. The mean value of $r(O_l)$ is 11.6%, with a standard deviation of 6.6%. The minimum value is 2.7% and the maximum value is 27.2%. The mean value of $R(O_l)$ is -0.6%, with a standard deviation of 2.2%. The minimum value is -10.2% and the maximum is 2.6%. Low values of $R(O_l)$ show that this simple model can give accurate estimates of water column production when the model is run with optimal values for parameters. The product of the relative error of water column production at the optimal point $R(O_l)$ and the measured water column production of that cruise gives the amount, predicted by the model, of excess/deficit carbon assimilated in the entire water column. As for the profiles of daily production, the error $r(O_l)$ does not have a straightforward interpretation. It can be visualized as the relative deviation of the model profile from the measured production profile. The spread of points indicates that there is no simple relation between $R(O_l)$ and $r(O_l)$. For most cruises, the model gives a low error for both, but there are also cruises when this does not occur, emphasizing that the model can give an accurate value for $P_{Z,T}$ regardless of $\mathcal{P}$.

This happens because condition (11) is violated at the first vertical level. The first vertical level could be ignored and the error calculated with information from the remaining vertical levels. For the profiles that have 8 points in the vertical, disregarding one would mean throwing away 12.5% of data and 16.6%, for the case of profiles with 6 points in the vertical. We performed this calculation and found that the results for some cruises yielded unrealistically high values for E_k, whereas the remaining distributions were only slightly altered. Higher values for E_k were obtained because the measured value of production from the first vertical level helps constrain the magnitude of P_m^B. Increasing P_m^B causes a proportional increase in E_k. A considerable amount of information on P_m^B is lost if measured production from the first vertical level is lacking.

Discussion

We may regard the inverse procedure presented here from two points of view: first it can be seen as a way to recover the

photosynthesis parameters when measurements of *in situ* production profiles exist; and second as a means of optimizing the model to perform as well as possible under the given limitations of both the model and the measurements. Both interpretations are useful, but from a practical standpoint, what is really important are the errors. Considering that errors are rather low (Figure 6), the inverse procedure can be declared successful within the range of its applicability. The extension of the range of applicability is limited heavily by the number of vertical levels at which the production profile is sampled. If there were more vertical levels, a more complex model (for example with depth-dependent parameters) for calculating production could be used. Historically, most of the profiles that were routinely measured at stations all over the world had measurements at a pre-specified number of depths, which hardly ever exceeded 10 in number. Also, surface PAR was rarely measured on these cruises. Even nowadays it is difficult to find datasets with a high vertical resolution of production and chlorophyll profiles, which also have measurements of surface PAR and of optical properties. We decided to use a simple model given the limitations imposed by the small number of sampling depths.

One way of avoiding the constraints imposed by the need for vertical uniformity of parameters would be to assume some functional form for the vertical dependence of parameters. But that would merely shift the issue from the estimation of photosynthesis parameters to the estimation of the parameters that would occur in these functions. In the case of vertical uniformity of photosynthesis parameters, a large number of vertical levels is not required, since there are only two parameters that need to be estimated. Vertical variation of parameter values in the model would require *in situ* measurements at more vertical levels, to capture this variation properly, if indeed it occurs *in situ*. Nevertheless, we see that the error in the estimation of daily water column production of our model, with vertically uniform parameters, does not exceed more than 10% for the HOT dataset. Using vertically uniform parameters, similar results have been obtained by Herman and Platt (1986). Note that, for $E > E_k$, it is essential only to capture the right value of P_m^B, and for $E < E_k$, the important parameter is α^B.

However, if we seek representative values for photosynthesis parameters of the entire population, it is natural to estimate the parameters simultaneously for the entire population, as is done here. With the usual application of the *in vitro* method, where the phytoplankton are sampled at a single depth, this cannot be done. Even if samples are taken from a number of depths, the parameters are still estimated for each depth separately. When using the *in vitro* method, it is assumed implicitly that the parameters extracted at a single depth are suitable for the entire water column. Instead, we find that it is more suitable to search for the optimal parameter combination in the way presented in this paper. One advantage of having just a single set of parameter values for the entire population is that those values represent the optimal photosynthesis parameters of the population, with respect to the measured production profile. When calculating water column primary production, it is better to have one value of parameters estimated from incubations at a number of depths (*in situ*), than one value estimated from multiple incubations at a single depth (*in vitro*). In a sense, with this approach, the water column is treated as a compound photosynthetic system (Talling, 1957). A similar approach was followed by Siegel *et al.* (2001) to estimate the parameters for their model of the production profile, and by Behrenfeld and Falkowski (1997) in their vertically generalized production model. However, these approaches do not use the fundamental parameters of the photosynthesis–irradiance function (α^B and P_m^B), but altered parameters, tailored for use in the models, and having limited scope for application outside them. Siegel *et al.* (2001, Section 5.3) and Behrenfeld and Falkowski (1997, Relative vertical distribution model section) acknowledge the discrepancies between the parameters of their model and the parameters of the photosynthesis–irradiance function.

For example, the model developed by Behrenfeld and Falkowski (1997) to estimate primary production at large scales by remote sensing, uses a parameter P_{opt}^B, based on the analysis of water primary production profiles measured *in situ*. This parameter is the maximum chlorophyll-normalized production observed in any particular profile of daily primary production. Since, in the course of a day, the available light varies from zero near dawn to a maximum, typically at local noon in the absence of clouds, and back to zero around dusk, the phytoplankton would have been light-limited during parts of the day, and possibly light-saturated at some other parts of the day. So, P_{opt}^B would lie somewhere between P_m^B and α^B: it would tend towards α^B if light remained low throughout the day, and towards P_m^B if light levels remained at saturating values for most of the day, but would always remain less than P_m^B because of low light levels at dawn and dusk. The interpretation of P_{opt}^B is therefore difficult, since it is representative only of the variable light conditions that existed at the time of the *in situ* primary production measurements. On the other hand, our model uses precisely α^B and P_m^B, and that is one characteristic setting it apart from other models of the production profile. Also, the photosynthesis parameters of our model can now be estimated from both *in situ* and *in vitro* data, as opposed to other models which can use only *in vitro* or *in situ* estimated parameters, but not both. The P_m^B and α^B retrieved in our model are true parameters of photosynthesis–irradiance models. They are designed to answer the "what-if" question: if we want to know what would the primary production be if the light level were such and such, the two parameters can be used to provide the answer.

A straightforward way of testing the precision of the inverse method would be to measure the production profile simultaneously with the P–E curve. The inverse procedure would be applied to the *in situ* profile and the parameters recovered. Once recovered, they could be compared with the measured values from the P–E curve. Platt and Sathyendranath (1988) demonstrate the agreement between water column production measured *in situ* and calculated from photosynthesis parameters measured *in vitro*.

Reliability of the recovered parameter values can also be tested by the magnitude of the errors of the optimal profile. When the model is run with the optimal parameter combination, the error is guaranteed to be the lowest possible, for a given set of light conditions. If the light conditions are erroneous, the recovered parameters may be suspect. If the model is run with measured parameters, but erroneous light conditions, the same error will occur. Calculating the error by comparing two erroneous results does not enable the detection of the error in parameters caused by incorrect light conditions. A way to avoid this problem would be to run the inverse model with perturbations in the irradiance matrix and thereafter checking the sensitivity of recovered parameters to these perturbations. But there appears no reason to run the model with different light conditions, since the light conditions experienced by the incubated phytoplankton are indeed those that are caused by the measured surface PAR. A model run with measured parameters would also be forced with the same surface PAR. To question the measurement of surface PAR is one thing, but to question the value of the diffuse attenuation

coefficient for down-welling irradiance K is somewhat different. For the HOT dataset we considered it best to estimate the value of K from the values of 1% per cent light level because it was available for a majority of cruises. This could be done because the region is oligotrophic with a low value of K. When considering the application of the inverse procedure to more complicated cases (eutrophic waters), careful treatment of light conditions is required. The HOT dataset also has a high temporal resolution of surface PAR (10 min sampling interval) for all the cruises. If such a high temporal resolution were not available, a way of proceeding would be to construct, as accurately as possible, a model for surface PAR (Iqbal, 1984). Having the surface PAR data together with *in situ* incubations enabled us to estimate the photosynthesis parameters, which otherwise would not be feasible. *In vitro* incubations would be required to do this. Therefore, cruises with *in situ* production measurements which include PAR measurements can be considered to be more informative with respect to the possibility of photosynthesis parameters estimation, via the method outlined.

Finally, we illustrate the application of our results to estimation of primary production from remotely sensed data on ocean colour (Longhurst *et al.*, 1995; Sathyedranath *et al.*, 1995). Using monthly chlorophyll data from the Ocean Colour Climate Change Initiative (www.oceancolour.org) of the European Space Agency and photosynthetically active radiation (PAR) from NASA, primary production was computed as a test product using the approach of Longhurst *et al.* (1995), with updated photosynthesis and chlorophyll profile parameters from Méin and Hoepffner (2011), as part of the Trans-Boundary Water Assessment Project (TWAP). The computations are based on a depth-resolved and wavelength-resolved model, following the original formalism of Platt and Sathyendranath (1988). In primary production model intercomparison exercises, this model compared favourably with respect to others (Friedrichs *et al.*, 2009; Saba *et al.*, 2010, 2011; see also Buitenhuis *et al.*, 2013). Primary production was computed using monthly chlorophyll fields at 9 × 9 km resolution and corresponding PAR fields. We found 65 match-ups between daily observations at HOT and monthly satellite-based primary production fields: in other words, in this preliminary demonstration, the daily in-water primary production was taken to be representative of the monthly production at that site. It must be emphasized to point out here that Longhurst *et al.* (1995) highlighted the paucity of information on photosynthesis–irradiance parameters for many parts of the world oceans, which obliged them to use parameters from similar biomes in the Atlantic Ocean to fill gaps in the Pacific Ocean. In fact, for the North Pacific Subtropical Gyre, in which the HOT station falls, the primary production computations in the TWAP project were executed using fixed parameter values for the whole year: 0.055 mg C m^2 (mg Chl W h)$^{-1}$ for α^B and 5 mg C (mg Chl h)$^{-1}$ for P_m^B.

Figure 7 shows the comparison between the production estimated using satellite data and the measured production for all the match-ups. The agreement is rather poor. To understand the sources of the discrepancy, the model presented here was re-run with the same parameter values as for the initial satellite computation, keeping all other inputs (including PAR, diffuse attenuation coefficient K and the chlorophyll profile) unchanged. We show also in Figure 7 that the new results are quite comparable with the satellite observations: this highlights the importance of assigning appropriate parameter values to get reasonable estimates from satellite-based computations. It is only when the appropriate parameters are used that the present model gives results that are consistent with *in situ* observations, as also shown in Figure 7 (see also Figure 6). Even though the satellite algorithm used (which is spectrally resolved, and allows for non-uniform structure in chlorophyll concentration) is different from the model presented here (non-spectral), and although the temporal match-ups are not ideal (daily *in situ* production and monthly satellite calculations), the differences between the two models is modest, when common photosynthesis–irradiance parameters are used, compared with the differences associated with parameter assignment. In this oligotrophic region where the ranges of variability in chlorophyll concentration and PAR are low, it is particularly important to get the parameters right.

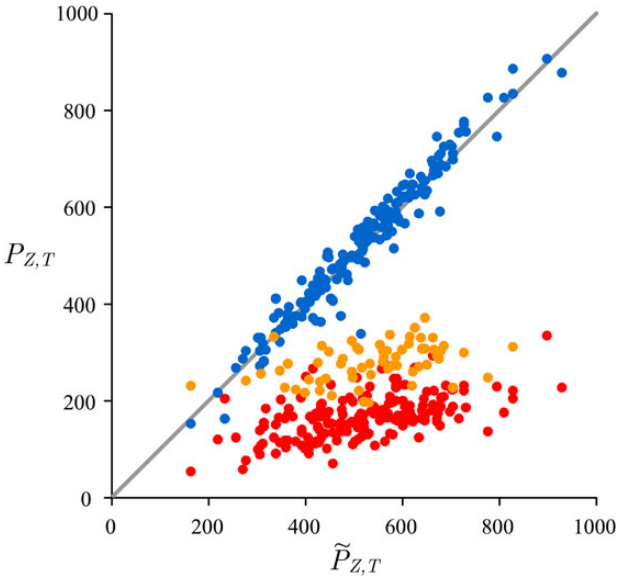

Figure 7. Plot of measured $\widetilde{P}_{Z,T}$ vs. modelled $P_{Z,T}$ estimated using satellite data with the parameter values from Longhurst *et al.* (1995) (orange points), estimated from the matrix model with the parameter values also from Longhurst *et al.* (1995) (red points) and from the matrix model with the parameter values estimated in this study (blue points).

Conclusions

The inverse procedure presented here is able to recover photosynthesis parameters from *in situ* vertical profiles of phytoplankton production. The parameters extracted using this method are not limited simply to the model used, they are general and can be used in any model which has P–E parameters, traditionally determined through the use of *in vitro* incubations. Therefore, the ecological relevance of our work lies in extracting the photosynthesis parameters from the data that were not originally intended to be used for such purposes and have not been used since the implementation of the *in situ* method in 1952 (Steemann Nielsen, 1952). The approach outlined changes our understanding in a way that two implementations of the C14 method, which were considered distinct, can indeed yield the same outcome, that outcome being the values of photosynthesis parameters. *In vitro* experiments were constructed specifically to yield this result, whereas *in situ* experiments were constructed to yield an estimate of water column production. With respect to the photosynthesis parameters the two approaches are now consolidated.

All required inputs for the inverse procedure can be measured easily, and are routinely measured on oceanographic cruises. The

iterative procedure was shown to be robust by unconstrained convergence for data from all 169 HOT cruises tested, which span a period of more than 20 years. When run with optimal values of parameters, the estimation errors of the model for daily, water column production are extremely low, although the model is quite simple compared with spectral models of primary production (Sathyendranath and Platt, 1989). The method can be applied to recover the values of photosynthesis parameters whenever the outlined limitations are not violated. The recovered parameters can be further used in a remote sensing context as inputs to algorithms for calculating primary production (Platt and Sathyendranath, 1988; Longhurst *et al.*, 1995). In a modelling context the parameters can be used in equations describing the spatial and temporal evolution of phytoplankton biomass (Fennel and Boss, 2003; Platt *et al.*, 2003).

Acknowledgements

This research was conducted under the 2013 POGO-SCOR Fellowship Programme. The activity benefited from the workshop on primary production and follow-up activities organized under the Indo-Mareclim Project of the European Union. This work is a contribution to the ESA projects "Photosynthesis Parameters from Space" and "Photosynthetically Active Radiation". This work has been supported in part by Croatian Science Foundation under the project (IP-2014-09-3606). We acknowledge the Ocean Colour Climate Change Initiative (OC-CCI) of the European Space Agency and the Trans-boundary Water Assessment Project for access to chlorophyll and primary production data, respectively, and NASA for access to PAR data. We acknowledge the Hawaii Ocean Time Series programme for the publicly available database and thank their scientists and staff.

References

Barber, R. T., and Hilting, A. K. 2002. History of the study of plankton productivity. *In* P.J. Le B. Williams, D.N. Thomas, and C.S. Reynolds (eds), Phytoplankton Productivity – Carbon Assimilation in Marine and Freshwater Ecosystems, pp. 16–43. Blackwell, Oxford.

Behrenfeld, M. J., and Falkowski, P. G. 1997. Photosynthetic rates derived from satellite-based chlorophyll concentration. Limnology and Oceanography, 42: 1–20.

Bennett, A. F. 2005. Inverse Modeling of the Ocean and Atmosphere, 1st edn. Cambridge University Press, Cambridge

Buitenhuis, E. T., Hashioka, T., and Quéré, L. 2013. Combined constraints on global ocean primary production using observations and models. Global Biogeochemical Cycles, 27: 847–858.

Fennel, K., and Boss, E. 2003. Subsurface maxima of phytoplankton and chlorophyll: steady-state solutions from a simple model. Limnology and Oceanography, 48: 1521–1534.

Franks, P. J. S. 2002. NPZ models of plankton dynamics: their construction, coupling to physics, and application. Journal of Oceanography, 58: 379–387.

Friedrichs, M., Carr, M.-E., Barber, R., Scardi, M., Antoine, D., Armstrong, R., Asanuma, I., *et al.* 2009. Assessing the uncertainties of model estimates of primary productivity in the tropical pacific ocean. Journal of Marine Systems, 76: 113–133.

Gentleman, W. 2002. A chronology of plankton dynamics in silico: how computer models have been used to study marine ecosystems. Hydrobiologia, 480: 69–85.

Glover, D. M., Jenkins, W. J., and Doney, S. C. 2011. Modeling Methods for Marine Science, 1st edn. Cambridge University Press, Cambridge.

Herman, A. W., and Platt, T. 1986. Primary production profiles in the ocean: Estimation from a chlorophyll/light model. Oceanologica Acta, 9: 31–40.

Iqbal, M. 1984. Introduction to Solar Radiation, 1st edn. Academic Press, New York.

Karl, D. M., Dore, J. E., Lukas, R., Michaels, A. F., Bates, R. N., and Knap, A. 2001. Building the long-term picture: the U. S. JGOFS time-series programs. Oceanography, 14: 6–17.

Karl, D. M., and Lukas, R. 1996. The Hawaii Ocean Time-series (HOT) program: background, rationale and field implementation. Deep-Sea Research II, 43: 129–156.

Kirk, J. T. O. 2011. Light and Photosynthesis in Aquatic Ecosystems, 3rd edn. Cambridge University Press, Cambridge.

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. SIAM Journal on Optimization, 9: 112–147.

Lewis, R. M., Torczon, V., and Trosset, M. W. 2000. Direct search methods: then and now. Journal of Computational and Applied Mathematics, 159: 245–254.

Longhurst, A., Sathyendranath, S., and Platt, T. A. 1995. An estimate of global primary production in the ocean from satellite radiometer data. Journal of Plankton Research, 17: 1245–1271.

Luo, Y., Ducklow, H. W., Friedrichs, M. A. M., Church, M. J., Karl, D. M., and Doney, S. C. 2012. Interannual variability of primary production and dissolved organic nitrogen storage in the north pacific subtropical gyre. Journal of Geophysical Research: Biogeosciences, 117: G03019.

Mélin, F., and Hoepffner, N. 2011. Monitoring phytoplankton productivity from satellite. *In* Handbook of Satellite Remote Sensing Image Interpretation: Applications for Marine Living Resources Conservation and Management, pp. 79–93. Ed. by J. Morels, V. Stuart, T. Platt and S. Sathyendranath. EU PRESPO and IOCCG, Darthouth, Canada.

Morel, A., and Smith, R. C. 1974. Relation between total quanta and total energy for aquatic photosynthesis. Limnology and Oceanography, 19: 591–600.

Nelder, J. A., and Mead, R. 1965. A simplex method for function minimization. Computer Journal, 7: 308–313.

Nicholson, D. P., Stanley, R. H. R., Barkan, E., Karl, D. M., Luz, B., Quay, P. D., and Doney, S. C. 2012. Evaluating triple oxygen isotope estimates of gross primary production at the Hawaii Ocean Time-series and Bermuda Atlantic Time-series study sites. Journal of Geophysical Research: Oceans, 117: C05012.

Peterson, B. J. 1980. Aquatic primary production and the 14C-CO_2 method: A history of the productivity problem. Annual Review of Ecology, Evolution and Systematics, 11: 359–385.

Platt, T., Broomhead, D. S., Sathyendranath, S., Edwards, A. M., and Murphy, E. J. 2003. Phytoplankton biomass and residual nitrate in the pelagic ecosystem. Proceeding of the Royal Society A, 459: 1063–1073.

Platt, T., Denman, K. L., and Jassby, A. D. 1977. Modelling the productivity of phytoplankton. *In* The sea: ideas and observations on progress in the study of the seas, pp. 807–856. Wiley, New York.

Platt, T., and Gallegos, C. L. 1980. Modelling primary production. *In* Primary Productivity in the Sea, pp. 339–362. Plenum Publishing Corporation.

Platt, T., Gallegos, C. L., and Harrison, W. G. 1980. Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. Journal of Marine Research, 38: 687–701.

Platt, T., and Jassby, A. 1976. The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. Journal of Phycology, 12: 421–430.

Platt, T., and Sathyendranath, S. 1988. Oceanic primary production: Estimation by remote sensing at local and regional scales. Science, 241: 1613–1620.

Platt, T., and Sathyendranath, S. 1991. Biological production models as elements of coupled, atmosphere-ocean models for climate research. Journal of Geophysical Research, 96: 2585–2592.

Platt, T., and Sathyendranath, S. 1993. Estimators of primary production for interpretation of remotely sensed data on ocean color. Journal of Geophysical Research, 98: 14561–14576.

Platt, T., Sathyendranath, S., and Ravindran, R. 1990. Primary production by phytoplankton: Analytic solutions for daily rates per unit area of water surface. Proceeding of the Royal Society B, 241: 101–111.

Saba, V., Friedrichs, M., Antoine, A., Armstrong, R., Asanuma, I., Behrenfeld, M., Ciotti, A., *et al.* 2011. An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. Biogeosciences, 8: 489–503.

Saba, V., Friedrichs, M., Carr, M.-E., Antoine, A., Armstrong, R., Asanuma, I., Aumont, O., *et al.* 2010. The challenges of modeling depth integrated marine primary productivity over multiple decades: A case study at bats and hot. Global Biogeochemical Cycles, 24: GB3020.

Sathyedranath, S., Longhurst, A., Caverhill, C. M., and Platt, T. 1995. Regionally and seasonally differentiated primary production in the north atlantic. Deep Sea Research 1, 42: 1773–1802.

Sathyendranath, S., and Platt, T. 1989. Computation of aquatic primary production: extended formalism to include the effect of angular and spectral distribution of light. Limnology and Oceanography, 34: 188–198.

Siegel, D. A., Westberry, M. C., O'Brien, M. C., Nelson, N. B., Michaels, A. F., Morrison, J. R., Scott, A., *et al.* 2001. Bio-optical modeling of primary production on regional scales: the Bermuda biooptics project. Deep-Sea Research II, 48: 1865–1896.

Steemann Nielsen, E. 1952. The use radioactive carbon (^{14}C) for measuring organic production in the sea. Journaldu Conseil. Conseil Permanent International pour l'Exploration de la Mer, 18: 117–140.

Talling, J. F. 1957. The phytoplankton population as a compound photosynthetic system. New Phytologist, 56: 133–149.

Wunsch, C. 1996. The Ocean Circulation Inverse Problem, 1st edn. Cambridge University Press.

Appendix

Let us take two points in the parameter space, namely $X = (\alpha^B_x, P^B_{m,x})$ and $Y = (\alpha^B_y, P^B_{m,y})$ such that $\alpha^B_x \neq \alpha^B_y$ and $P^B_{m,x} \neq P^B_{m,y}$. For a unique biomass profile $\mathbf{b}$ and irradiance matrix $\mathbf{E}$, for all such X and Y it holds that:

$$\mathbf{p}_T(X) \neq \mathbf{p}_T(Y), \tag{A1}$$

while only for some X and Y it holds that:

$$P_{Z,T}(X) = P_{Z,T}(Y). \tag{A2}$$

These conditions say that the model with two different parameter combinations can give the same value of daily water column production (A2), but with two different vertical profiles of daily production (A1), for a unique biomass profile and irradiance matrix. This means that each profile, for any point in the parameter space, is unique, whereas the daily production at that same point is not unique with respect to other points.

The assumption $X \neq Y$ leads to two different $p^B(E)$ functions, such that $p^B(E|X) \neq p^B(E|Y)$, and consequently to different $\mathbf{P}^B$ matrices for the same irradiance matrix $\mathbf{E}$. Finally by $\mathbf{p}_T = \mathbf{P}\boldsymbol{\tau}$ we get $\mathbf{p}_T(X) \neq \mathbf{p}_T(Y)$. If we assume the opposite of (A1), that is $\mathbf{p}_T(X) = \mathbf{p}_T(Y)$ and write it using $\mathbf{p}_T = \mathbf{P}\boldsymbol{\tau}$, we get $\mathbf{B}\mathbf{P}^B(X)\boldsymbol{\tau} = \mathbf{B}\mathbf{P}^B(Y)\boldsymbol{\tau}$. For this to hold it is required that the biomass-normalized production matrices are equal $\mathbf{P}^B(X) = \mathbf{P}^B(Y)$. That is possible only for the case of $p^B(E|X) = p^B(E|Y)$, which is valid only if $X = Y$, contradicting the initial assumption of $X \neq Y$. So, by assuming that $X \neq Y$ and $\mathbf{p}_T(X) = \mathbf{p}_T(Y)$, we find that $X = Y$ for it to hold. This proves the uniqueness of the vertical profile of daily production in the parameter space. Put more simply, to every point (α^B, P^B_m) there corresponds a unique model profile of daily production $\mathbf{p}_T(\alpha^B, P^B_m)$. The given proof enables us to conclude that the inverse problem is well defined. To every parameter combination there is joined a distinct profile.

We now prove that equality of daily water column production is valid for $X \neq Y$ and does not lead to the same contradiction as in the previous case. Let us assume $\boldsymbol{P}_{Z,T}(X) = \boldsymbol{P}_{Z,T}(Y)$ holds. We can write it using $\boldsymbol{P}_{Z,T} = \boldsymbol{\zeta}\mathbf{P}\boldsymbol{\tau}$ as $\boldsymbol{\zeta}\mathbf{P}(X)\boldsymbol{\tau} = \boldsymbol{\zeta}\mathbf{P}(Y)\boldsymbol{\tau}$. Considering that $\mathbf{p}_T = \mathbf{P}\boldsymbol{\tau}$, we can write $\boldsymbol{\zeta}\mathrm{p}_T(X) = \boldsymbol{\zeta}\mathbf{p}_T(Y)$, which is nothing more than an equality between two scalar products. Because the scalar product is not uniquely determined, it is not necessary for the two daily production profiles to be equal for them to have the same value of $P_{Z,T}$. Following (A1), to every different vertical profile there corresponds a unique combination of parameters and we conclude that the two profiles can indeed have different values of parameters, although they yield the same value of $P_{Z,T}$. So, the equality $P_{Z,T}(X) = P_{Z,T}(Y)$ does not require $p^B(E|X) = p^B(E|Y)$.

Handling editor: Matthew Oliver