

*Biogeosciences Discussions* is the access reviewed discussion forum of *Biogeosciences*

# Skill assessment of the PELAGOS global ocean biogeochemistry model over the period 1980–2000

**M. Vichi<sup>1,2</sup> and S. Masina<sup>1,2</sup>**

<sup>1</sup>Centro Euro-Mediterraneo per i Cambiamenti Climatici, Bologna, Italy

<sup>2</sup>Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

Received: 24 March 2009 – Accepted: 27 March 2009 – Published: 1 April 2009

Correspondence to: M. Vichi (vichi@bo.ingv.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

Global Ocean Biogeochemistry General Circulation models are useful tools to study biogeochemical processes at global and large scales under current climate and future scenario conditions. The accuracy of the future estimate is however dependent on the adequate representation of the current ocean biogeochemical features. To this purpose, the results of an interannual simulation of the global ocean biogeochemical model PELAGOS have been objectively compared with multi-variate observations from the last 20 years of the XX century. The model was assessed in terms of spatial and temporal variability of chlorophyll and primary production derived from satellite sensors, with a specific focus on the simulation of carbon production/consumption rates observed in the equatorial Pacific ocean and at the long-term JGOFS stations. The predicted chlorophyll is acceptable in the northern mid-latitude regions and equatorial Pacific, but is underestimated in the upwelling regions of the Atlantic and Indian Oceans and markedly overestimated in the Southern Ocean. This latter bias is linked to the inadequate representation of the mixed layer seasonal cycle in the region, which favours primary production during austral spring. Simulated primary production is comparable with satellite estimates both at the global scale and when compared with an independent data-set in the equatorial Pacific. A comparison with other models showed that PELAGOS results are as good as the estimates from state-of-the-art diagnostic models based on satellite data. The skill in reproducing the interannual variability was assessed in the equatorial Pacific and against the decadal JGOFS timeseries BATS and HOT. In the tropical Pacific our analysis suggests that interannual variability of primary production is related to the climate variability both in the observations and in the model. At the JGOFS stations PELAGOS has skill to simulate the observed bacterial biomass and shows realistic means of primary and bacterial production at BATS. These results have been further strengthened with an analysis of spatial variability of microbial carbon production/consumption and comparison with observations along a transect in the Atlantic ocean. Within the limits of the model assumption and known

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



biases, PELAGOS predicts that the system is net heterotrophic if the boreal winter period only is considered and especially in the more oligotrophic regions. However, at the annual time scale and over the global ocean, the model suggests that surface ocean is close to a slightly positive autotrophic balance in accordance with recent experimental findings and geochemical evidences.

## 1 Introduction

Ocean Biogeochemistry General Circulation Models (OBGCM) derive from the coupling of GCMs that solve the hydrodynamics of the ocean and biomass-based mathematical representations of the lower trophic levels of marine ecosystems. The state variables defining the biogeochemical model spans from dissolved inorganic nutrients at the molecular level, to unicellular autotrophic and heterotrophic organisms, and up to complex living multicellular organisms such as planktonic metazoans. Given the limitation of the biomass-based mathematical definitions, OBGCM are a rough approximation of the complexity observed in the global ocean ecosystem. The focus of these models is not the study of interactions at the wider community and ecosystem level but the quantitative representation of the major biogeochemical constituent fluxes among the first trophic levels. The planktonic components are, despite their little standing stock, the basis of the biological pump in the ocean with a global estimated primary production ranging from less than 40 to more than 60 Pg C yr<sup>-1</sup> (Carr et al., 2006). Even if the net role of the ocean biological pump in the global carbon cycle is still to be clarified (Sarmiento et al., 1998), there are concerns that alteration in ocean features due to climate changes may affect the ocean biogeochemistry with feedbacks into the climate system itself (see Denman et al., 2007, for a compendium on the current literature).

OBGCMs are promising tools to study the functioning of global biogeochemical processes and to produce future projections based on climate change scenarios. The quality of the projection is however dependent on the adequate representation of the

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ocean biogeochemical features under current climate conditions. If models have to be used in the context of climate change studies, they need to have skill in capturing the observed interannual variability of firstly the mean bulk biogeochemical properties, but secondly, and most importantly, the rates at which organic matter is processed within the food web. This task requires an objective assessment of model skill against available data, taking into account that data availability is practically limited by the inherent complexity of measuring biogeochemical parameters at the global scales. A combination of satellite and in situ data is therefore needed to obtain a plausible picture of ocean biogeochemistry under current climate conditions (Ducklow, 2003).

A set of multi-variate, multi-instrument observations collected over the last 20 years of the XX century was used in this work to evaluate the results of an interannual simulation of a global ocean comprehensive biological model (Sect. 2.1). The major substantial assumption here is that these data describe the mean state and variability of the current climate conditions from a biogeochemical point of view. This is a rather strong assumption given the little amount of data concerning carbon rates and the lack of repeated measurements that allows the definition of variability in a climatological meaning. The assessment exercise implies the usage of adequate skill scores (Sect. 2.2), an objective quantitative method that only recently has penetrated the community of biogeochemical modellers (Lynch et al., 2008). We specifically used a selected set of skill indices focusing on the evaluation of mean bulk variables at the global scale (Sect. 3) and interannual process variables at selected locations (Sects. 4 and 5). After an overall assessment of the model skill, we have further investigated the results to derive insights on the dynamics of ocean biogeochemical processes as described by the model. These tasks specifically focused on the study of the response of primary production to climate variability in the equatorial Pacific and a contribution to the discussion on the metabolic state of the global ocean comparing with data collected in the Atlantic. Our focus is mainly on the assessment of organic carbon production/consumption mechanisms and their relationship with ocean dynamics, considering also the global ocean balance of net biogenic carbon production. This final issue, presented in Sect. 6, is

---

**Skill assessment of  
the PELAGOS  
OBGCM**M. Vichi and S. Masina

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



a modelling contribution to the discussion on the metabolic state of the surface ocean (e.g. Del Giorgio and Duarte, 2002; Riser and Johnson, 2008), which has important implications for the role of microbial biogeochemistry in the ocean carbon sequestration. Section 7 finally offers a brief discussion and summarizes the conclusions.

## 2 Methods

### 2.1 Model description and setup

PELAGOS (PELAGic biogeochemistry for Global Ocean Simulations, Vichi et al., 2007b,a) is a coupling between the OPA (Océan PARallelise) general circulation model (Madec et al., 1999) and the global ocean version of the Biogeochemical Flux Model (BFM, <http://bfm.cmcc.it>) derived and modified from the ERSEM model (Baretta et al., 1995). The model grid is the irregular ORCA2 configuration (Madec and Imbard, 1996) with a nominal 2×2 degrees size and a refined latitudinal mesh of 0.5 degree in the equatorial regions.

The biogeochemical model implements a set of biomass-based differential equations that solves the fluxes of carbon, nitrogen, phosphorus, silica and iron among selected biological functional groups representing the major components of the lower trophic levels. The functional groups in the pelagic environment are represented by unicellular planktonic autotrophs (pico-, nano-phytoplankton and diatoms), zooplankton (nano-, micro- and meso-) and heterotrophic bacterioplankton. The model also simulates the dynamics of nitrate, ammonium, phosphate, biogenic silicate, iron, oxygen and has an explicit parameterization of the biochemical cycling of dissolved/particulate non-living organic matter.

The results analysed here are extracted from a multi-annual simulation over the period 1958–2001 forced with daily mean heat and momentum fluxes from the European Centre for Medium Range Weather Forecasting (ECMWF) 40-year re-analysis. The forcing functions and the results of a similar physical simulation are described

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



in Bellucci et al. (2007). The ocean physics parameterizations are as in Vichi et al. (2007b) with sea surface temperature (SST) relaxed to the daily-interpolated value of the Reynolds data set (Reynolds et al., 2002) with a coefficient of  $40 \text{ W m}^{-2}$ . The biogeochemical model is initialized as in Vichi et al. (2008) from the World Ocean Atlas nutrient data (Conkright et al., 2002) and with homogeneous low values for all the other biogeochemical variables. The model was not calibrated against the data used in this paper and the current set of parameter values was derived following the experience with the model forced by climatological atmospheric data, by means of a manual one-at-a-time modification starting from the values presented in Vichi et al. (2007b). The major changes involved an increase of the affinity constants for nutrients to the original values in Baretta-Bekker et al. (1995) and a generalized reduction of the iron limitation for all phytoplankton groups. A new list of parameter values is available as a supplemental table (<http://www.biogeosciences-discuss.net/6/3511/2009/bgd-6-3511-2009-supplement.pdf>) and in the PELAGOS page of the BFM website (<http://bfm.cmcc.it>) with the differences with respect to Vichi et al. (2007b, also available on the same page).

## 2.2 Data sets and skill indicators

The biogeochemical data sets used in this assessment encompass the last 20 years of the 20th century with a focus on the data that offer multivariate information and especially biological rates for at least a decade (e.g. primary and bacterial production). Global coverage data are however related to derived parameters of phytoplankton biomass only, available mostly through satellites and ocean color products such as the ones from the Sea Wide Field-of-view Sensor (SeaWiFS). Empirical data-models are required to translate sensor information into relevant properties such as chlorophyll-*a* concentration, and the quality of the reconstructed data is to be considered when comparing with deterministic models. For instance, Gregg and Casey (2004) report an average root mean square log error of 31% and a coefficient of determination of 0.76 for chl satellite estimates against in situ data over the whole global ocean. These

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



scores meet the objectives of the SeaWiFS mission at the global scale, but on the other hand, a thorough validation analysis against concurrent in situ chl data shows regional discrepancies with overestimation in the equatorial Atlantic and underestimation in the Southern Ocean. Satellite data and inherent optical properties can be further

combined to estimate other important biogeochemical properties such as primary production, carbon content and plankton functional groups distributions (e.g. Behrenfeld and Falkowski, 1997; Alvain et al., 2005; Aiken et al., 2007). In this work we used the estimates of global primary production derived with the Vertically Generalized Production Model (VGPM) proposed by Behrenfeld and Falkowski (1997).

The other datasets used in this assessment are in situ observations that have been selected because of their temporal and spatial coverage of primary production and other relevant biogeochemical rates or biomass data. We focused on three publicly available datasets: the ClimPP dataset (Friedrichs et al., 2008) and the Joint Global Ocean Study (JGOFS) timeseries HOT (Hawaii Ocean Timeseries at Station ALOHA, Lukas and Karl, 1999) and BATS (Bermuda Atlantic TimeSeries, Steinberg et al., 2001). These data are further described in their specific Sects. 4.1 and 5.

The choice of the performance indicators or scores is done according to recent works that focused on skill assessment (Allen et al., 2007; Lynch et al., 2008; Friedrichs et al., 2008; Stow et al., 2008). The suggested univariate indices comprises the measure of bias (B), average absolute error (AAE) and variability of the misfit measured as Root Mean Square Differences (RMSD, see Appendix). Two additional performance indicators have been applied as suggested by Allen et al. (2007) and Stow et al. (2008): the Modelling Efficiency (MEF, Nash and Sutcliffe, 1970) and the Reliability Index (RI, Leggett and Williams, 1981), which are further described in the Appendix. Regression analysis was also performed to evaluate the goodness-of-fit of prediction vs. observations taking into account the linear methods described in Smith and Rose (1995) and Pineiro et al. (2008).

To visualize the combination of the different indicators and compare the PELAGOS results with the other biogeochemical models presented in Friedrichs et al. (2008), we

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



use the Multi-Dimensional Scaling (MDS) technique. MDS is an iterative technique used to visualize proximities in a low-dimensional space first introduced in psychometrics (Borg and Groenen, 2005). This analysis carefully preserves the distance between items from multivariate datasets and allows the combined visualization of multiple information in one single plot (see Appendix).

Score values and confidence intervals were also evaluated by means of empirical  $p$ -values estimates and bootstrap techniques. The probability of obtaining a score value better than the one achieved is generally termed  $p$ -value (Mason, 2008). The empirical distribution of score values was constructed with 10 000 random re-samplings of the observation (or simulation) timeseries and computing the verification index for each new set of model-data pairs. As pointed out by Mason (2008),  $p$ -values do not answer the question whether the score value is good, but rather they provide a degree of significance with respect to random combinations. The confidence interval is instead computed by means of the bootstrap technique, in which the choice of randomly permuted model-data pairs is done by replacing the extracted pairs in the original timeseries. This procedure ensures that the quality of the new randomly-generated timeseries is as high as the original model-data pairs to be evaluated. The 95% confidence limits are then empirically computed from the distribution of the score values.

### 3 Global assessment

#### 3.1 Chlorophyll concentration

Figure 1 presents a visual comparison of simulated chlorophyll (chl) with satellite-derived estimates from SeaWiFS. Remotely-sensed chl represents the average chl concentration within one optical penetration depth, which in turns depends on the radiative extinction operated by the chl amount itself. When comparing model results with satellite data we need to consider both a data uncertainty, the 30% RMS log error described in Sect. 2.2, and the model representation of satellite chl, which may range between

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





the simulated surface values and the vertically integrated chl concentration averaged over the euphotic depth (the latter shown in Fig. 1). Simulated surface values generally underestimates the center of the gyres and the coastal upwelling areas more than vertically integrated values (not shown). The comparison is shown as annual means and variability. The variability was estimated with the anomaly coefficient of variation, which is the standard deviation of the anomaly fields computed with respect to and normalized by the monthly climatological averages. The variability allows us to assess the behavior of the model in all seasons, because means cannot be computed throughout the year in high-latitude regions.

The spatial distribution of the annual mean is generally captured in terms of high and lows. The north Atlantic and Pacific have improved with respect to the climatological simulations in Vichi et al. (2007a), and also the oligotrophic regions. Coastal maxima are instead unchanged, as they are difficult to be simulated with the coarse resolution of the ocean model. The largest discrepancies are also found in the upwelling regions of the Indian Ocean and equatorial Atlantic. The model is capable of simulating the maxima at the higher latitudes and the typical minima of variation in the subtropical regions, with distinct intervals of high variability marking the borders of the Pacific cold tongue influence. The Northern Hemisphere higher-than-observed variability, particularly evident in the North Atlantic, is caused by too low summer concentrations, a feature that has not improved from the climatological model results (Vichi et al., 2007a). The Southern Ocean is markedly overestimated, particularly in terms of variability. The cause is the large boreal spring bloom that is driven by the sudden starting of the seasonal stratification.

A more objective way of looking at this feature is through the MEF index (Fig. 2). This index is extremely strict when applied to spatial fields, because it computes a point-to-point comparison on a reference grid. With this index, the model skill is generally unsatisfactory because the value of the index indicates that the mean of the data is almost always a best predictor than the model. This analysis however gives more information on the variability of the misfit and its spatial and seasonal dependence.

**Skill assessment of  
the PELAGOS  
OBGCM**

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The problems in the Northern Hemisphere and tropical ocean are stationary and linked to the seasonal cycle. The worst skill is found during the boreal winter while during summer the model is as good a predictor as the mean of the data (see Appendix for the threshold limits of MEF). The lowest performances occur in spring in the Southern Ocean, due to the large simulated bloom in the frontal regions of the Southern Ocean.

The simulation of mixed layer depth (MLD) spatial and temporal evolution can partly explain the mismatch in the Southern Ocean. Figure 3 shows the visual comparison with annual mean data from de Boyer Montégut et al. (2004) data. The MLD was evaluated with a temperature difference criterium of  $0.2^{\circ}\text{C}$  both in the data and in the model results. The mean annual mismatch in the Southern Ocean is extremely large, particularly southern of  $40^{\circ}\text{S}$  as also evidenced in the zonal mean distribution (Fig. 3c). This discrepancy is mostly found during the onset of the stratification in October–November in the sub-Antarctic province (Fig. 3d). Even if the MLD simulation has improved with respect to Vichi et al. (2007a), mostly because of the use of interannual atmospheric forcing functions, the early stratification in the sub-Antarctic still favours the bloom of diatoms, which can maximise production through photoacclimation in the illuminated shallow MLD utilizing the abundance of nutrients typical of this region. Diatom specific production can be tuned by decreasing the iron availability or the cell membrane affinity (as previously done in Vichi et al. (2007a)) but this has consequences on their behaviour in the whole global ocean and particularly in the North Atlantic and equatorial Pacific, which resulted greatly underestimated in the previous climatological simulations.

### 3.2 Primary production

The VGPM (Behrenfeld and Falkowski, 1997) is an empirical model that estimates net primary production (NPP) from satellite-derived chlorophyll using a temperature-dependent description of chlorophyll-specific photosynthetic efficiency. The NPP is thus computed using the observed chlorophyll concentration, SST and surface available light as inputs. The comparison with PELAGOS was done with a derived prod-

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



uct of the VGPM (available at <http://www.science.oregonstate.edu/ocean.productivity/>) that implements the exponential dependence of production on temperature according to Eppley (1972). This exponential relationship is more similar to the one applied in PELAGOS (Vichi et al., 2007b).

This exercise is a model-to-model comparison (see also section below) therefore we focus here on maps of the annual means (Fig. 4) and zonal averages (Fig. 5) because they are expected to be better captured by the VGPM and satellite models in general (Campbell et al., 2002). The spatial variability is in fact very similar to the SeaWiFS chl variability (Fig. 1), since the major input of VGPM are the satellite-derived chlorophyll data. There is a good agreement in the spatial distribution of maxima, especially in the location of the frontal maximum in the Antarctic Circumpolar Current in spite of the failure in reproducing the annual chl value observed from satellite (Fig. 1). The coastal zone production is generally underestimated because of the low phytoplankton biomass. Nutrient availability plays a key role in these areas and the model is not able to fuel the surface coastal ocean with sufficient nutrient inputs. This is partly due to the resolution, as also evidenced by the mismatch in the subtropical and equatorial Atlantic and Indian Oceans, which are the smaller basins. The Mauritanian upwelling is in fact absent and the equatorial maximum in the Atlantic is found closer to the south America continent.

An interesting feature is the filament of production in the center of the south Pacific subtropical gyre, which resembles the one in the south Atlantic and Indian oceans. This marks the location of the polar front and the subduction of the sub-Antarctic waters, implying that the model tends to shift the AACC more to the north than estimated from satellite data. This limit of northward extension is also visible in the variability of chl shown in Fig. 1d. This signal is instead not visible in the SST (not shown), most likely due to the relaxation to observed surface temperature data (Sect. 2.1).

In summary, we can conclude that PELAGOS provides an estimate that lies always in the range of the VGPM estimate as evidenced by a comparison of zonally averaged primary production (Fig. 5). PELAGOS predicts higher NPP in the regions of the

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



boundary currents at 45 N and 45 S, in the sub-Antarctic region and at the equator (because of the Pacific maximum). Conversely, NPP in the inter-tropical northern region is generally underestimated, mostly due to the lack of the Indian Ocean maxima.

## 4 Primary production in the equatorial Pacific

### 4.1 The ClimPP dataset

The comparison with satellite-derived primary production shown in the previous section can be considered an assessment provided that satellite estimates are validated against in situ observations of NPP. This kind of assessment was undertaken by the series of intercomparison studies called Primary Production Assessment Round-Robin (PPARR, Campbell et al., 2002; Carr et al., 2006). In the latest published round, PPARR3, Friedrichs et al. (2008) put together a set of observations from the equatorial Pacific used as benchmarks for the reality check of satellite-based production models (SatPPM) and OBGCMs. One of their conclusions is that current state-of-the-art SatPPM are only slightly more skillfull then prognostic OBGCMs and that the actual dominance depends on the choice of the assessment score. Some models are better than other in terms of bias, while others are better in terms of variability. Incidentally, one of the major outcomes of this project is the public availability of this quality-checked dataset, which can thus be used to test model performances. The dataset consists of vertically-integrated euphotic zone measurements of primary production, combined with SST, chlorophyll and ancillary model data for MLD and surface irradiance, covering the whole tropical Pacific over the period 1983–1995. The NPP data are originally available in units of daily carbon production per unit area. A point-to-point comparison was done according to the same protocol described in Friedrichs et al. (2008) using monthly mean model data, and the assessment was done by considering the same suggested set of performance indicators.

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 4.2 PELAGOS results

Net primary production in PELAGOS is parameterized as a function of light, chlorophyll and iron cell content and dissolved silicate concentration (Vichi et al., 2007b). The cell availability of N and P do not directly control photosynthesis, but the subsequent transformation of carbohydrates into proteins and cell material in general. A portion of photosynthesized carbon is therefore released as Dissolved Organic Carbon (DOC) exudates according to the internal nutrient quota as originally described in Baretta-Bekker et al. (1995).

When comparing with in situ incubation data, especially if derived with radiocarbon techniques, the contribution of this DOC part should be neglected because of the filtering procedure that removes particles smaller than (usually)  $2\text{ }\mu\text{m}$ . The colloidal High Molecular Weight (HMW) portion is likely to remain attached to cells, while the remainder Low Molecular Weight (LMW) is released in the water. The percentual of colloidal HMW DOC during a bloom is estimated around 20% (Kepkay et al., 1993) while an overall bulk figure of LMW DOC varies from 65 to 80% (Ogawa and Tanoue, 2003), the highest loss likely to occur in oligotrophic waters. From a modelling point of view this distinction is not necessary, because all of the produced carbon will take one of the possible trophic pathways and the relevant inflow is the NPP computed as photosynthetic production minus the metabolic and/or activity respiration losses. Nevertheless, an unknown portion of the NPP that is lost to DOC should be removed when comparing the model output with data. To account for this effect, two different estimates of NPP have been used in the comparison: NPP1 is the total amount of organic carbon produced by autotrophs, while NPP2 considers an estimated conservative loss of 50% of the time-varying DOC exudates. The overall estimate of ClimPP parameters is presented in Table 1 and the graphical misfit analysis (Stow et al., 2008) of NPP is shown in Fig. 7.

The SST scores are very good, as it was expected because of the relaxation to observed data described in Sect. 2.1. Also the estimate of MLD is in accordance with the

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



model data provided in the ClimPP datasets. The modelled MLD is on average deeper but with similar standard deviation. Friedrichs et al. (2008) report that the ClimPP MLD data are in the  $\pm 20$  m range with respect to the JGOFS equatorial Pacific Process Study cruises, therefore our results fall within the same ranges. The low MEF however indicates that the two models give different time evolution and thus further independent data are needed to assess the quality of the vertical structure of the model with respect to the equatorial Pacific conditions. MLD in the equatorial Pacific is probably not the best indicator for production as the satPP models using MLD are not as skillful as the others. OBGCMs like PELAGOS use more info than the MLD and are capable to obtain better scores because they do not rely on this variable only.

Chlorophyll skill is good only regarding the indicators of average concentration. Bias and average errors are small, but the simulated standard deviation is much higher than observed and also the MEF indicates a poor if not bad predictive performance. NPP is instead much better than chlorophyll and in line with the results of the other PPARR3 models. The improvement of the NPP2 estimate of PP are clear in all the indices. The bias is much reduced and thus the total root mean square difference. It is interesting to note that also some indices of variability improve, such as the standard deviation and the correlation coefficient. This suggests that a dynamical parameterization of exudate production may contribute to a more proper estimation of the observed production. It is also important to remember that the NPP2 estimates of NPP does not change the other variable results; it only implies a different way of comparing data with a model like PELAGOS that implements a more sophisticated parameterization of primary production.

A possible way to show the combined skill of PELAGOS in the framework of the other PPARR3 models is presented in Fig. 6 with the aid of the MDS ordination (Sect. 2.2). The two-dimensional distances between the multivariate set of indices is well represented (stress is close to 0). The ClimPP data point is included in the ordination by assuming a set of indices with the highest score values (e.g.  $r=1$ ,  $B=0$ , etc.). Two additional artificial data points that represent the worst cases have been added. They

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



are obtained by combining the worst scores from all the models and by assuming a standard deviation that is either the highest or lowest than the observations found in the model parameter space. This picture clearly shows that the NPP estimates from satPP models and OBGCM are indeed comparable as pointed out by Friedrichs et al.

5 (2008). All the models lie at approximately the same distance from the data, the closest being a set of satPPMs. PELAGOS NPP2 is much better than NPP1 because it is located in the cluster of the best OBGCMs and satPPMs.

The direct visual comparison of ClimPP and PELAGOS data and misfits (Fig. 7, NPP2 results) provides additional information on the goodness-of-fit of model results. 10 The data show a small positive trend as reported by Friedrichs et al. (2008), though this is not statistically significant. The model-data misfit (Fig. 7b) apparently decreases with time though the model does not have a trend (Fig. 7a). This tendency suggests that observed primary production in the '80s was relatively lower than the one in the '90s. A similar evolution can be found in the model results if the cluster of high NPP after 1988 is not considered. The model in fact overestimates the production during and 15 after the 1988 ENSO event as it will be further discussed in Sect. 5.3. The goodness-of-fit between model and data values can be objectively assessed by means of linear regression on the scatter plot (Fig. 7c). The slope of the regression line is significantly different from 1 and the coefficient of determination is  $r^2=0.33$ . The lack of fit is however not caused by the bias, but mostly by unexplained variance and partly by the different slope. This confirms the results of Table 1 performed on the log-transformed data. The high dispersion might be due to natural variability in the observations, although it is also a possible indication of misspecification in the model. The misfit in fact increases with increasing model values (Fig. 7d) with a majority of overestimation. 20 The hypothesis  $H_0$  that the regression slope with angular coefficient 0.42 is equal 0 can be significantly rejected with  $p<0.01$ . 25

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## 5 Interannual assessment at selected locations

The number of available long term stations is rather limited and the choice of the two major JGOFS stations in the Atlantic (Bermuda) and Pacific (station ALOHA) is almost mandatory as locations for model calibration and hypothesis testing (e.g. Hurtt and Armstrong, 1996; Spitz et al., 2001; Huisman et al., 2006; Brix et al., 2006). In this section we have focused on the set of data provided at the end of JGOFS by the modelling and synthesis group because they reflect a coherent dataset based on uniform unit and method conversions for both stations.

### 5.1 The JGOFS Station BATS

BATS lies at the eastern boundary of the north Atlantic subtropical gyre and, though being stratified for most of the time, it experiences winter mixing events as deep as 200 m. This seasonal signal is found in NPP data and well captured by the model (Fig. 8). The linear regression indicates goodness-of-fit because the value of the slope is statistically equal to 1 (Fig. 8b, the equation is given in the caption). The major discrepancy with the data is due to the bias, which is mostly found during summer when our model simulates higher than observed NPP, even if most of the previous modelling studies reported underestimation during this period (Brix et al., 2006). There is however a distinct difference between the surface and the vertically-integrated NPP values: the model overestimates NPP in the euphotic zone and underestimates the surface value (not shown). This implies that PELAGOS is able to simulate a productive deep community during summertime. These results are obtained with the NPP2 estimation (cfr. Sect. 4.2), but since during summertime more colloidal DOC is exudated because of the oligotrophic conditions, it is likely that considering a constant HMW proportion in DOC when comparing with data is only partly sufficient. Direct comparison with oxygen production fluxes should help to clarify further this issue.

Primary production peaks are linked to the maxima in MLD (Fig. 9) both in the data and in the model (linear correlation coefficient  $r=0.80$  and  $r=0.65$ , respectively). MLD

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





is also well predicted by the model although there is a bias during wintertime that makes the linear test in the scatter plot not significant. The model simulates the NPP interannual variability quite well, particularly when linked to distinct physical features. The winter 1994 low-production event is in fact well predicted by the model, as it is likely that the underestimation of the other winter maxima is due to an underestimation of MLD. It is interesting to notice that 1994 was characterized by the highest export of particulate organic matter during summertime due to a bloom of diatoms (Brix et al., 2006). The model is capable to partly recovery the export after the winter negative production anomaly (not shown), though it is not as high as reported for summer 1994.

The values of the performance indicators at BATS indicate good skill of the model (Table 2). Means are correctly predicted and within the standard error for all the variables except nutrients, which are close to the detection limits in the observations and very close to 0 in the model. The linear correlation coefficient for chl and NPP are high and significant ( $p < 0.01$ ) with confidence intervals of 0.50–0.73 and 0.53–0.78, respectively (both computed with the bootstrap method, Sect. 2.2). The MEF index is larger than 0 for NPP only ( $p < 0.01$ , confidence interval 0.27–0.58), which is classified as a “good” score (see Appendix). The RI is instead rather high due to the summer overestimation, implying that on average, the predicted production can get more than twice higher than observed.

Bacterial biomass and production (BP) were compared with the observed surface values due to the larger data availability at this level. Biomass values were converted from cell counts using the cellular carbon content suggested by Gundersen et al. (2002, 10 fg/cell). BP was converted from thymidine incorporation hourly rates into daily carbon production by means of the conversion factor suggested by Fuhrman and Azam (1982) and by considering a multiplicative factor of 12, which is in the range of the ratios between leucyne and thymidine incorporations as measured by Ducklow et al. (2001) in the Arabian Sea. The mean bacterial biomass are both well simulated by the model, as confirmed by the very low bias and RI value close to 1 (Table 2). The model is however not capable to simulate the variability (not shown), which is partly linked

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to the seasonal production. This is shown here by the high values of  $\text{RMSD}_{cp}$  and the absolute average error. The small linear correlation (confidence interval 0.11–0.46,  $p < 0.01$ ) is likely caused by the presence of a weak seasonal signal both in data and model, but is not sufficient to bring MEF higher than 0. Similar considerations can be done for BP, which has a no bias at all but is characterized by a higher RI due to the low variability predicted by the model. The choice of a constant scaling factor used to convert thymidine incorporation into daily production may also play a role here, because in some seasons it can be 4 times as high as considered here (Ducklow et al., 2001). The linear correlation is slightly higher than for the biomass (confidence interval is 0.21–0.46,  $p < 0.01$ ) and the MEF is also positive (c.i. is 0.05–0.18,  $p < 0.01$ ) although still poor according to the indicative thresholds in Appendix.

## 5.2 The JGOFS station ALOHA (HOT)

Station ALOHA is located in the subtropical Pacific Ocean north of Hawaii. It presents a permanently stratified water column with a characteristic deep chlorophyll maximum (DCM) below 100 m, mostly composed of *Prochlorococcus* with temporary outbursts of diatoms and dinoflagellates (Karl et al., 2003a). The surface ocean is depleted in nutrients and the mixed layer depths occasionally reaches the location of the DCM during wintertime.

PELAGOS results at this station captures the typical low production conditions but miss the observed higher frequency variability (Fig. 10). The tendency of models to predict low NPP at HOT have been reported by other authors (e.g. Ondrusek et al., 2001). The observations are bounded by the NPP1 and NPP2 estimates, which suggests that the choice of the fraction of colloidal DOC at Sta. ALOHA may play a relevant role. If the amount of colloidal DOC at Sta. ALOHA is higher than at BATS due to the more oligotrophic conditions, it is likely that a fraction higher than 50% (as estimated with the ClimPP dataset and used with BATS, cfr. Sect. 4.2) be retained by the filter. leading to higher measured NPP. Both NPP1 and NPP2 however shows a clear lack of fit as shown by Fig. 10b.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The MLD evolution is well-reproduced by the model (Fig. 11) , but there is no clear relationship with NPP as seen for instance at BATS. It is likely that the NPP variability here is driven by small scale features, episodic but seasonally-recurrent nitrogen fixation (Dore et al., 2008) or possibly by chaotic fluctuations (Huisman et al., 2006). None of these features can be reproduced by the current model design, either due to the coarse horizontal resolution or because of the absence of diazotrophs that are still not included in PELAGOS.

The values of the performance indicators of biological variables are less good than at BATS (Table 3) although SST and MLD are very well predicted both in terms of magnitude and variability. Chl is on average one third of the observations but standard deviations are comparable and the linear correlation is statistically significant (c.i. 0.17–0.59,  $p < 0.05$ ). The same considerations apply for bacterial biomass, whose bulk value is well predicted. The number of bacterial data at HOT is however smaller than at BATS, which makes the linear correlation and RI values less significant ( $p < 0.1$ , c.i.  $-0.13 \sim 0.57$ ,  $1.3 \sim 1.4$ , respectively). The indicators of the standing stocks are thus generally acceptable but the NPP is underestimated. If we consider the NPP1 estimate of autotrophic production, PELAGOS overestimates the measured NPP with a mean value of 337 and s.d.=77. In this case the total amount of autotrophic carbon production is more similar to the values suggested by Ondrusek et al. (2001).

The model however suggests that all NPP is utilised in the surface euphotic layer by heterotrophic respiration, as it can be diagnosed with the ratio of net community production (NCP) over NPP (e- or ef-ratio), the ratio of particle production over NCP (p-ratio) and the ratio of particle production over NPP (pe-ratio, see Brix et al., 2006, for further descriptions). The long-term e-ratio is 0.04, p-ratio is 0.43 and pe-ratio is 0.015, while at BATS their values are e-ratio=0.23 p-ratio=0.23 pe-ratio=0.05. The e-ratio at HOT is much lower than estimated by Brix et al. (2006), while it is interesting that PELAGOS predicts the higher p-ratio at BATS.

Further comparison with bacterial production and respiration data are needed to assess whether the rates of bacterial carbon transformation predicted by the model are

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



realistic. At the current stage these results imply that areas like HOT are metabolically neutral from the point of view of the effects on the global carbon cycle and therefore the underestimation of NPP will not affect the average long-term fluxes of carbon with the atmosphere.

### 5.3 Interannual variability in the equatorial Pacific

The climate variability in the equatorial Pacific is dominated by the El Niño Southern Oscillation (ENSO). It is well known that phytoplankton responds to the ENSO signal as evidenced by changes in satellite-derived chlorophyll during ENSO events (e.g. Ryan et al., 2006). The high phytoplankton biomass is related to the increase in the surface availability of macro and micro nutrients from the upwelling waters in the eastern Pacific during the cold phase of ENSO. During El Niño events the upwelling is reduced and also the amount of surface chlorophyll derived from ocean color observations. During strong La Niña events, as for instance in 1998, there is a large anomalous increase in phytoplankton (Ryan et al., 2002). Besides the ocean color evidence, there have been direct observed changes in primary production associated to ENSO (see references in Pennington et al., 2006), but to our knowledge there are no direct studies linking large scale datasets as the ClimPP database and ENSO.

Friedrichs et al. (2008) described a positive trend in the ClimPP dataset that was not captured by any of the models participating to the intercomparison experiment, and also PELAGOS overestimates the production in the '80s with respect to the '90s (cfr. Sect. 4.2). To analyse the variability of the ClimPP dataset and PELAGOS results in relation to the climate conditions, we used the Multivariate ENSO Index (MEI, Wolter and Timlin, 1998) that is an indicator of the ENSO state over the whole equatorial Pacific. The NPP anomaly was computed with respect to two different periods, 1983–1989 and 1990–1995, for both ClimPP data and the PELAGOS results. Between the MEI and the NPP anomaly there is a clear visual correlation especially in the peak phases of both cold and warm ENSO states (Fig. 12). Particularly in the period 1986–1990, the NPP anomaly is negative during the positive MEI and viceversa. During the

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



long and weak ENSO phase in the early '90s (Fedorov and Philander, 2000) it is difficult to derive a clear correlation between the MEI index and the NPP anomalies due to the limited amount of observations. The only two data periods in 1992 show negative and positive NPP anomalies in correspondence with a maximum and minimum of the MEI.

5 It is also striking to note the large positive anomalies in the early '80s, apparently during a weak La Niña condition, which are found both in the data and in the model.

These results indicate that the model is indeed able to respond to the interannual variability of the equatorial Pacific and produce changes in primary production that are in accordance with the available observations.

10 **6 The metabolic balance of the global ocean**

Microbial production and consumption rates are central to the understanding of the organic carbon flux through the food webs and the implications for the biological pump. The issue of interest here is on the amount of organic carbon produced by autotrophs and consumed by the heterotrophic biota in the surface and deeper layers. The major  
15 question is whether the global ocean is net heterotrophic or how large the net heterotrophic regions are (e.g. Del Giorgio and Duarte, 2002; Riser and Johnson, 2008) and what will be the changes in future climate conditions.

OBGCMs and Earth System Models are frequently used to provide this information, by scaling up parameterizations of production and consumption processes of the lower  
20 trophic levels to the global ocean. It is therefore essential to assess the quality of the simulation of local microbial carbon fluxes and physical conditions, in order to verify the model skill and increase the confidence on the future projections when doing the dynamical extrapolation to the world ocean. Unfortunately, only sparse measures of these rates exist at the ocean basin scales. For instance, one of the largest available  
25 dataset was presented by Hoppe et al. (2002) for a meridional transect collected during the period November 1991 – January 1992 in the Atlantic. These data revealed that the tropical areas of the Atlantic with high SST values have the Bacterial Carbon

**Skill assessment of  
the PELAGOS  
OBGCM**

M. Vichi and S. Masina

Title Page

AbstractIntroduction

ConclusionsReferences

TablesFigures

◀▶

◀▶

BackClose

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Demand (BCD) that exceeds NPP at the surface of the ocean. This is an indicator of net heterotrophy at the microbial level, that is likely to reduce the importance of the biological pump as a carbon sequestration process.

Some methodological issues arise here as in Sect. 4.2 because of the direct comparison of bottle incubations with model data. The model is parameterized to produce daily values while incubation experiments have shorter time scales (usually 8–12 h). The strategy we adopted to reduce uncertainties in the extrapolation was to convert model results into data units by taking into account the experimental protocols applied during the campaign. The estimate of NPP by Hoppe et al. (2002) was done by neglecting the DOC exudation, which was estimated as 5–30%. In the model data we present both the NPP1 and NPP2 estimates as in Sect. 4.2 when comparing with the observations, but used the NPP1 only for deriving integrated carbon production at the basin and global scales because this measure is comparable with the net autotrophic oxygen production.

The model computes bacterial production (BP) from a constant bacterial growth efficiency (BGE) of 30% (Vichi et al., 2007b) and BCD is a function of the nutrient content of the available organic substrate. The comparison with observed BP values is thus the central assessment of model skill because this is the variable that is directly measured. The surface ratio of BP/NPP (Fig. 13a) computed along the meridional transect and during the same period of observations (Fig. 13c) shows a similar range of spatial variability as reported in Hoppe et al. (2002), with minima found at the higher latitudes and at the equator, and maxima in the tropics and at the southern boundary of the South-Pacific gyre. The model reproduces also the BCD/BP distribution derived from the data, with values above 100% in correspondence of the same maxima of BP/NPP from 20°N to 40°S. The values in the northern part, which are equal to the southern maximum and higher than observed, are likely due to spatial biases in the location of PP with respect to the estimates of satellite-based PP models (Fig. 4). Simulated NPP is in fact very low in the western part of the tropical Atlantic close to the Amazon river. Since the data from Hoppe et al. (2002) have been collected at the surface (although

---

**Skill assessment of  
the PELAGOS  
OBGCM**M. Vichi and S. Masina

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



with some special treatments to account for low-light production), it is interesting to analyse the integrated model results over the euphotic zone depth (Fig. 13b). In this case the model predicts that the whole Atlantic is net heterotrophic during the boreal winter period, implying that BCD is sustained by additional non-local sources of organic matter.

In the upper part of Table 4 we present a direct comparison of model results with averages from the northern and southern parts of the transect. The ranges of observations and model data overlap for all parameters and the spatial difference between north and south Atlantic is well captured. The southern part of the transects shows higher values of production mostly because of the sampling along the eastern shelf of South America and the crossing with the polar front. Surface chl is underestimated in the Northern Hemisphere but not in the southern part of the transect because the model simulates high biological activity in the frontal region that increases the mean value. The ratio BP/PP is also well simulated. The NPP2 estimates lead to higher ratios but also to larger standard deviations that still overlaps the observed values.

Given the good model skills on the transect data, it is interesting to derive basin scale estimates of the metabolic ratios first in the Atlantic and then in the global ocean. The results are shown in the lower part of Table 4 using the values integrated over the euphotic zone to provide a figure for the surface ocean. To check the relevance of seasonality, we have first extrapolated the boreal winter (NDJ) values computed over the whole simulation period 1980–2001 to the year length and compared it with the actual annual means. According to model results, if the boreal winter conditions are taken as representative of the mean microbial activity, the Atlantic ocean is net heterotrophic and the global ocean is slightly above 100%. The annual climatological means reveal that autotrophic carbon production is instead as important as heterotrophic processes, leading to values that are close to a neutral metabolic balance, if not slightly autotrophic. The overall figure of carbon production is in accordance with the satellite derived estimates (Falkowski et al., 2000). It is however likely that the model compensates for the low production in the gyres with the overestimation in the Southern

---

**Skill assessment of  
the PELAGOS  
OBGCM**

---

M. Vichi and S. Masina

---

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)



Ocean, thus leading to a value that is comparable with the satellite estimates.

The net autotrophic balance suggested by the model is in accordance with geochemical evidences of oxygen production and recent direct observations (Najjar and Keeling, 2000; Riser and Johnson, 2008). The coarse resolution of the model cannot distinguish

whether the autotrophic production is due to pulses of production as suggested by Karl et al. (2003b) or to a continuous contribution as recently suggested (Riser and Johnson, 2008). It is however clear from model results that the extrapolation of process rate variables from the local scale to the annual and basin scales may lead to misestimation of the metabolic state, with a tendency to show net heterotrophic conditions.

This may even be more pronounced when the extrapolation is done using real data that are affected by mesoscale local processes more than our coarse model simulation (Maixandeu et al., 2005a) and it is thus important to consider an entire seasonal cycle to properly estimate the trophic state (Maixandeu et al., 2005b).

## 7 Discussion and conclusions

A set of objective assessment tools have been used to test the skill of the PELAGOS model over the last 20 years of the XX century. The aims were to evaluate the performance of the model under current climate conditions in view of its usage in climate change scenario simulations in the context of Earth System Models (ESM). ESMs solve the carbon cycle in the atmospheric, terrestrial and ocean components, and the global role of ocean biology in the transformation of organic carbon and the resulting sequestration. PELAGOS was first assessed in climatological simulations focusing on bulk variables (Vichi et al., 2007a). In this work we revised the time-dependent results focusing on the production of organic carbon and the implications for the biological pump.

The predicted chlorophyll field is acceptable in the northern mid-latitude regions and equatorial Pacific, but is underestimated in the Atlantic Ocean and markedly overestimated in the Southern Ocean. The use of interannual atmospheric forcing functions

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





instead of climatological have improved the model results especially in the Northern Hemisphere. However, the improvements in the North Atlantic and equatorial Pacific with respect to the simulations in Vichi et al. (2007a, especially in terms of diatoms, not shown) led to an increase of the chl bias in the Southern Ocean. This latter bias needs to be further investigated because it is likely linked to the inadequate resolution of the mixed layer seasonal cycle in the region. The bias is further enhanced in PELAGOS due to the presence of adaptive chl:C ratio. It is therefore important to avoid overconstraints of parameters such as iron-limitation coefficients to force the model towards the observed concentrations. We consider unlikely that any parameterization of phytoplankton growth would be able to capture the observed low chl concentrations in presence of permanent stratified conditions and given the abundance of nutrients. The work on the parameter values is also a corollary aspect of modelling that needs to be supported by in situ observations.

Simulated primary production is comparable with satellite estimates both at the global scale and when compared with an independent data set in the equatorial Pacific. The interannual variability of primary production in this important climatic region is related to the climate variability both in the observations and in the model. Additionally, it is worth considering that using satellite-derived products for model assessment is still a model-to-model comparison. As demonstrated in Sect. 4.2 and also found in Friedrichs et al. (2008), the predictability skill of satellite PP models and OBGCMs is comparable. Independent data sets are thus needed to make sure that satellite-based products can be used to fill the observational gaps and robustly validate OBGCMs. It is thus extremely important that long-term data series are maintained to allow assessments as the one shown in Sect. 5.

The independent test with JGOFS station increase the confidence in the model parameterizations. The model was not calibrated to the observations at BATS, therefore the skill in capturing the observed means is remarkable. It is however likely that PELAGOS underestimates net production in more permanently stratified regions such as HOT, although further discussion on the methods to compare with primary production

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



data are needed to completely understand the discrepancies and improve the parameterizations. The background production in the surface euphotic layer is adequate, but the absence of episodic production peaks reduces the annual bulk value. The underestimation in subtropical gyres is compensated by the Southern Ocean, thus leading to an overall carbon production that is comparable with other independent estimates.

Our results underline the importance of the adequate comparison between observations and model simulations. Experimental protocols and model variable (and process) definitions need to be properly considered to maximize the information that can be extracted from data and model results. In our specific case, the quality and quantity of DOC exudated from phytoplankton under oligotrophic conditions was found to be a key variable to improve the goodness-of-fit of the model against in situ primary production observations.

The comparison with completely independent data of the carbon fluxes through bacteria increase our degree of confidence in the model results, making it suitable for studying the degradation processes of organic matter under different ocean conditions. At the current status of availability of long-term observations, the results from the model indicate that the quality and to some extent the quantity of the organic matter transfer is in line with the available long-term observations.

These results have been further strengthened with an analysis of spatial variability of microbial carbon production/consumption and comparison with observations along a transect in the Atlantic ocean. Within the limits of the model assumption and known biases, we have used PELAGOS results to estimate the metabolic balance of the global ocean in the euphotic zone. The model predicts that in boreal winter conditions and in certain oligotrophic regions there is a tendency towards net heterotrophy as observed in the field. However, in the annual mean and over basin and global scales, the surface ocean is close to a slightly positive autotrophic balance. It is therefore interesting to study the behaviour of the model in case of climate change scenarios and assess whether the induced changes in the general circulation and water-mass properties might affect this state.

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Appendix A

### Univariate skill scores

The most simple measures of distance between a set of observations  $O_n$  and model predictions  $P_n$ ,  $n=1, 2, \dots, N$  are the bias

$$B = \frac{1}{N} \sum_{n=1}^N O_n - \frac{1}{N} \sum_{n=1}^N P_n$$

and the absolute average error

$$AAE = \frac{1}{N} \sum_{n=1}^N |O_n - P_n|.$$

The total Root Mean Square Difference (RMSD) is defined as

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{n=1}^N (O_n - P_n)^2}$$

which can be further separated into a component due to the bias and a centered-pattern (unbiased) difference according to the relationship

$$\text{RMSD}^2 = B^2 + \text{RMSD}_{CP}^2.$$

The Nash-Sutcliffe Model Efficiency (Nash and Sutcliffe, 1970) is a measure of the ratio of the model error to the variability of the observations:

$$\text{MEF} = 1 - \frac{\sum_{n=1}^N (O_n - P_n)^2}{\sum_{n=1}^N (O_n - \bar{O})^2}$$

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where  $O_n$  and  $P_n$  are the  $N$  pairs of observational data and predictions, respectively. Performance levels are usually categorised as follows:  $>0.65$  excellent,  $0.65-0.5$  very good,  $0.5-0.2$  good,  $<0.2$  poor. If the index is lower than 0, it means that the model is a worse predictor than the mean of the observations. If, on the other hand is close to 0, the model is as good as the data mean.

The Reliability Index (RI, Leggett and Williams, 1981) measures the order of magnitude of model predictions with respect to data:

$$RI = \exp \sqrt{\frac{1}{N} \sum_{n=1}^N \left( \log \frac{O_n}{P_n} \right)^2}.$$

It was originally proposed as a statistical interpretation of log-normal distributed data, which is a typical distribution for most of the properties in ecology. The measure is interpreted as the value such that 68% of the model predictions fall within  $1/RI$  and  $RI$  (Smith and Rose, 1995). This index thus does not distinguish whether the multiplicative factor is related to over- or underestimation therefore it requires the concurrent analysis of the bias.

## Non-metric Multi-Dimensional Scaling (MDS)

Starting from a matrix of similarities/dissimilarities (resemblance matrix) between  $n$  items, this algorithm constructs a new set of data points in a low dimension space (usually 2-D) whose proximities are obtained through a minimization procedure that maintains the original distances in the resemblance matrix. The stress function that is minimized through iteration is the measure of the fit between proximities in the new low-dimensional space ( $d_{ij}$ ) and the distances in the original data space ( $\delta_{ij}$ ) as, for instance:

$$\text{stress} = \left[ \frac{\sum_{i=1}^n \sum_{j>i}^n (d_{ij} - f(\delta_{ij}))^2}{\sum_{i=1}^n \sum_{j>i}^n d_{ij}^2} \right]^{\frac{1}{2}}$$

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where  $f$  is a nonmetric, monotone transformation of the input data (distances). The minimization will lead to a final choice of  $f$  that reproduces the general rank-ordering of distances between the objects. A stress value lower than 0.2 is considered a sufficient description of the original proximities in the lower dimensional space. The distances between the coordinates of item  $i$  and  $j$  in the new space are Euclidean and depends on the chosen number of dimensions  $m$ , such as

$$d_{ij} = \left[ \sum_{a=1}^m (x_{ia} - x_{ja})^2 \right]^{\frac{1}{2}}$$

where  $x_a$ ,  $a=1, \dots, m$ , is the new coordinate system. MDS is a standard tool in many statistical software packages and in this particular case we used the MATLAB implementation. The multivariate distance metrics containing the combination of the skill score values can be computed in different ways. Given the non-linear nature of the skill scores, the resemblance matrix was built using the Manhattan (cityblock) distance:

$$\delta_{ij} = \sum_{k=1}^K |s_{ik} - s_{jk}|$$

where  $K$  is the number of normalized skill scores  $s$  defining each model object.

**Acknowledgements.** This work was supported by the Italian grant FISIR given to the Centro Euro-Mediterraneo per i Cambiamenti Climatici (CMCC). SeaWiFS data used in this publication were produced by the SeaWiFS project at Goddard Space Flight Centre. SeaWiFS and VGPM data have been downloaded from <http://www.science.oregonstate.edu/ocean.productivity>. Use of this data is in accord with the SeaWiFS Research Data Use Terms and Agreements. JGOFS data have been downloaded from the US JGOFS Synthesis & Modeling Project web page. MV thanks M. Azzaro and R. LaFerla for the help with the bacterioplankton production rate data.

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## References

- Aiken, J., Fishwick, J. R., Lavender, S., Barlow, R., Moore, G. F., Sessions, H., Bernard, S., Ras, J., and Hardman-Mountford, N. J.: Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October 2002: preliminary validation of new demonstration products for phytoplankton functional types and photosynthetic parameters, *Int. J. Remote Sens.*, 28, 497–516, 2007. 3517
- Allen, J., Somerfield, P., and Gilbert, F.: Quantifying uncertainty in high-resolution coupled hydrodynamic-ecosystem models, *J. Marine Syst.*, 64, 3–14, 2007. 3517
- Alvain, S., Moulin, C., Dandonneau, Y., and Breon, F. M.: Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery, *Deep-Sea Res. Pt. I*, 52, 1989–2004, 2005. 3517
- Baretta, J., Ebenhöf, W., and Ruardij, P.: The European Regional Seas Ecosystem Model, a complex marine ecosystem model, *J. Sea Res.*, 33, 233–246, 1995. 3515
- Baretta-Bekker, J., Baretta, J., and Rasmussen, E.: The microbial food web in the European Regional Seas Ecosystem Model, *J. Sea Res.*, 33, 363–379, 1995. 3516, 3523
- Behrenfeld, M. and Falkowski, P.: Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42, 1–20, 1997. 3517, 3520
- Bellucci, A., Masina, S., Dipietro, P., and Navarra, A.: Using temperature-salinity relations in a global ocean implementation of a Multivariate data assimilation scheme, *Mon. Weather Rev.*, 135, 3785–3807, 2007. 3516
- Borg, I. and Groenen, P.: *Modern Multidimensional Scaling: theory and applications*, Springer-Verlag, New York, 2nd edn., 619 pp., 2005. 3518
- Brix, H., Gruber, N., Karl, D. M., and Bates, N. R.: On the relationships between primary, net community, and export production in subtropical gyres, *Deep-Sea Res. Pt. II*, 53, 698–717, 2006. 3526, 3527, 3529
- Campbell, J., Antoine, D., Armstrong, R., Balch, W., Barber, R., Behrenfeld, M., Bidigare, R., Bishop, J., Carr, M.-E., Esaias, W., Falkowski, P., Hoepffner, N., Iverson, R., Kiefer, D., Lohrenz, S., Marr, J., Morel, A., Ryan, J., Vedernikov, V., Waters, K., Yentsch, C., and Yoder, J.: Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance, *Global Biogeochem. Cy.*, 16(3), 1035, doi:10.1029/2001GB001444, 2002. 3521, 3522
- Carr, M. E., Friedrichs, M. A. M., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., Asanuma,

**BGD**

6, 3511–3562, 2009

### Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quere, C., Lohrenz, S., Marra, J., Melin, F., Moore, K., Morel, A., Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A comparison of global estimates of marine primary production from ocean color, *Deep-Sea Res. Pt. II*, 53, 741–770, 2006. 3513, 3522

Conkright, M., Garcia, H., O'Brien, T., Locarnini, R., Boyer, T., Stephens, C., and Antonov, J.: World Ocean Atlas 2001, Volume 4: Nutrients, vol. NOAA Atlas NESDIS 52, U.S. Government Printing Office, Washington D.C., <http://www.nodc.noaa.gov/OC5/WOA01/woa01v4d.pdf>, cD-ROMs, 2002. 3516

de Boyer Montégut, C., Madec, G., Fischer, A., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology, *J. Geophys. Res.*, 109, C12003, doi:10.1029/2004JC002378, 2004. 3520, 3552

Del Giorgio, P. A. and Duarte, C. M.: Respiration in the open ocean, *Nature*, 420, 379–384, 2002. 3515, 3531

Denman, K., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P., Dickinson, R., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P., Wofsy, S., and Zhang, X.: Couplings Between Changes in the Climate System and Biogeochemistry, in: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007. 3513

Dore, J. E., Letelier, R. M., Church, M. J., Lukas, R., and Karl, D. M.: Summer phytoplankton blooms in the oligotrophic North Pacific Subtropical Gyre: Historical perspective and recent observations, *Prog. Oceanogr.*, 76, 2–38, 2008. 3529

Ducklow, H. W.: Biogeochemical provinces: towards a JGOFS synthesis, in: *Ocean biogeochemistry*, edited by: Fasham, M., Springer, 3–17, 2003. 3514

Ducklow, H. W., Smith, D. C., Campbell, L., Landry, M. R., Quinby, H. L., Steward, G. F., and Azam, F.: Heterotrophic bacterioplankton in the Arabian Sea: Basinwide response to year-round high primary productivity, *Deep-Sea Res. Pt. II*, 48, 1303–1323, 2001. 3527, 3528

Eppley, R.: Temperature and phytoplankton growth in the sea, *Fishery Bulletin*, 70, 1063–1085, 1972. 3521

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Falkowski, P., Scholes, R., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., Mackenzie, F., III, B. M., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., and Steffen, W.: The global carbon cycle: a test of our knowledge of Earth as a system, *Science*, 290, 291–296, 2000. 3533

5 Fedorov, A. and Philander, S.: Is El Nino changing?, *Science*, 288, 1997–2002, 2000. 3531

Friedrichs, M. A., Carr, M.-E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Buitenhuis, E. T., Chai, F., Christian, J. R., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Gentili, B., Gregg, W., Hoepffner, N., Ishizaka, J., Kameda, T., Lima, I., Marra, J., Mélin, F., Moore, J. K., Morel, A., O'Malley, R. T., O'Reilly, J., Saba, V. S., Schmeltz, 10 M., Smyth, T. J., Tjiputra, J., Waters, K., Westberry, T. K., and Winguth, A.: Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, *J. Marine Syst.*, doi:10.1016/j.jmarsys.2008.05.010, in press, 2008. 3517, 3522, 3524, 3525, 3530, 3535, 3555

Fuhrman, J. A. and Azam, F.: Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters – evaluation and field results, *Mar. Biol.*, 66, 109–120, 1982. 3527

Gregg, W. W. and Casey, N. W.: Global and regional evaluation of the SeaWiFS chlorophyll data set, *Remote Sens. Environ.*, 93, 463–479, 2004. 3516

Gundersen, K., Heldal, M., Norland, S., Purdie, D. A., and Knap, A. H.: Elemental C, N, and P Cell Content of Individual Bacteria Collected at the Bermuda Atlantic Time-Series Study (BATS) Site, *Limnol. Oceanogr.*, 47, 1525–1530, 2002. 3527

Hoppe, H.-G., Gocke, K., Koppe, R., and Begler, C.: Bacterial growth and primary production along a north-south transect of the Atlantic Ocean, *Nature*, 416, 168–171, doi: 10.1038/416168a, 2002. 3531, 3532, 3549, 3562

25 Huisman, J., Pham Thi, N. N., Karl, D. M., and Sommeijer, B.: Reduced mixing generates oscillations and chaos in the oceanic deep chlorophyll maximum, *Nature*, 439, 322–325, doi:10.1038/nature04245, 2006. 3526, 3529

Hurtt, G. and Armstrong, R.: A pelagic ecosystem model calibrated with BATS data, *Deep-Sea Res. Pt. II*, 43, 653–683, 1996. 3526

30 Karl, D., Bates, N., Emerson, S., Harrison, P., Jeandel, C., O., L., Liu, K.-K., Marty, J.-C., Michaels, A., Miquel, J., Neuer, S., Nojiri, Y., and Wong, C.: Temporal studies of biogeochemical processes determined from ocean time-series observations during the JGOFS era, in: *Ocean Biogeochemistry*, edited by: Fasham, M., chap. 10, 239–267, Springer, Berlin, 2003a.

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





- Karl, D. M., Laws, E. A., Morris, P., Williams, P. J. L., and Emerson, S.: Global carbon cycle: metabolic balance of the open sea., *Nature*, 426, 32, doi:10.1038/426032a, 2003b. 3534
- Kepkay, P. E., Niven, S. E. H., and Milligan, T. G.: Low molecular weight and colloidal DOC production during a phytoplankton bloom, *Mar. Ecol. Prog. Ser.*, 100, 33–244, 1993. 3523
- Leggett, R. and Williams, L.: A reliability index for models, *Ecol. Model.*, 13, 303–312, 1981. 3517, 3538
- Lukas, R. and Karl, D.: Hawaii Ocean Time-series (HOT), 1988–1998: A decade of interdisciplinary oceanography., CD-ROM 99-05, School of Ocean and Earth Science and Technology, University of Hawaii, <http://hahana.soest.hawaii.edu/hot/hot.>, 1999. 3517
- Lynch, D. R., McGillicuddy, Jr, D. J., and Werner, F.: Skill assessment for coupled biological/physical models of marine systems, *J. Marine Syst.*, doi:10.1016/j.jmarsys.2008.05.002, 2008. 3514, 3517
- Madec, G. and Imbard, M.: A global ocean mesh to overcome the North Pole singularity, *Clim. Dynam.*, 12, 381–388, 1996. 3515
- Madec, G., Delecluse, P., Imbard, M., and Levy, C.: OPA8.1 ocean general circulation model reference manual, Notes du pole de modelisation, IPSL, France, <http://www.lodyc.jussieu.fr/opa>, 1999. 3515
- Maixandeau, A., Lefevre, D., Fernandez, I. C., Sempere, R., Sohrin, R., Ras, J., Van Wambeke, F., Caniaux, G., and Queguiner, B.: Mesoscale and seasonal variability of community production and respiration in the surface waters of the NE Atlantic Ocean, *Deep-Sea Res. Pt. I*, 52, 1663–1676, 2005a. 3534
- Maixandeau, A., Lefevre, D., Karayanni, H., Christaki, U., Van Wambeke, F., Thyssen, M., Denis, M., Fernandez, C. I., Uitz, J., Leblanc, K., and Queguiner, B.: Microbial community production, respiration, and structure of the microbial food web of an ecosystem in the northeastern Atlantic Ocean, *J. Geophys. Res.*, 110, C07S17, doi:10.1029/2004JC002694, 2005b. 3534
- Mason, S. J.: Understanding forecast verification statistics, *Meteorol. Appl.*, 15, 31–40, doi:10.1002/met.51, 2008. 3518
- Najjar, R. and Keeling, R.: Mean annual cycle of the air-sea oxygen flux: A global view, *Global Biogeochem. Cy.*, 14, 573–584, 2000. 3534
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models, part 1 – a discussion of principles, *J. Hydrol.*, 10, 282–290, 1970. 3517, 3537

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Ogawa, H. and Tanoue, E.: Dissolved organic matter in oceanic waters, *J. Oceanogr.*, 59, 129–147, 2003. 3523
- Ondrusek, M. E., Bidigare, R. R., Waters, K., and Karl, D. M.: A predictive model for estimating rates of primary production in the subtropical North Pacific Ocean, *Deep-Sea Res. Pt. II*, 48, 1837–1863, 2001. 3528, 3529
- Pennington, J., Mahoney, K., Kuwahara, V., Kolber, D., Calienes, R., and Chavez, F.: Primary production in the eastern tropical Pacific: a review, *Prog. Oceanogr.*, 69, 285–317, 2006. 3530
- Pineiro, G., Perelman, S., Guerschman, J. P., and Paruelo, J. M.: How to evaluate models: Observed vs. predicted or predicted vs. observed?, *Ecol. Model.*, 216, 316–322, 2008. 3517
- Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., and Wang, W.: An Improved In Situ and Satellite SST Analysis for Climate, *J. Climate*, 15, 1609–1625, 2002. 3516
- Riser, S. C. and Johnson, K. S.: Net production of oxygen in the subtropical ocean, *Nature*, 451, 323–325, <http://dx.doi.org/10.1038/nature06441>, 2008. 3515, 3531, 3534
- Ryan, J. P., Polito, P. S., Strutton, P. G., and Chavez, F. P.: Unusual large-scale phytoplankton blooms in the equatorial Pacific, *Prog. Oceanogr.*, 55, 263–285, 2002. 3530
- Ryan, J. P., Ueki, I., Chao, Y., Zhang, H., Polito, P. S., and Chavez, F. P.: Western Pacific modulation of large phytoplankton blooms in the central and eastern equatorial Pacific phytoplankton blooms in the central and eastern equatorial Pacific, *J. Geophys. Res.*, 111, G02013, doi:10.1029/2005JG000084, 2006. 3530
- Sarmiento, J., Hughes, T., Stouffer, R., and Manabe, S.: Simulated response of the ocean carbon cycle to anthropogenic climate warming, *Nature*, 393, 245–249, 1998. 3513
- Smith, E. P. and Rose, K. A.: Model goodness-of-fit analysis using regression and related techniques, *Ecol. Model.*, 77, 49–64, 1995. 3517, 3538
- Spitz, Y., Moisan, J., and Abbott, M.: Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS), *Deep-Sea Res. Pt. II*, 48, 1733–1768, 2001. 3526
- Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., and Knap, A. H.: Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry, *Deep-Sea Res. Pt. II*, 48, 1405–1447, 2001. 3517
- Stow, C. A., Jolliff, J., Jr., D. J. M., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rosef, K. A., and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, *J. Marine Syst.*, doi:10.1016/j.jmarsys.2008.03.011, 2008. 3517, 3523

**BGD**

6, 3511–3562, 2009

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Vichi, M., Masina, S., and Navarra, A.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part II: numerical simulations, J. Marine Syst., 64, 110–134, 2007a. 3515, 3519, 3520, 3534, 3535
- 5 Vichi, M., Pinardi, N., and Masina, S.: A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: theory, J. Marine Syst., 64, 89–109, 2007b. 3515, 3516, 3521, 3523, 3532
- Vichi, M., Masina, S., and Nencioli, F.: A process-oriented model study of equatorial Pacific phytoplankton: The role of iron supply and tropical instability waves, Prog. Oceanogr., 78, 147–162, 2008. 3516
- 10 Wolter, K. and Timlin, M. S.: Measuring the strength of ENSO events – how does 1997/98 rank?, Weather, 53, 315–324, 1998. 3530

**BGD**

6, 3511–3562, 2009

---

**Skill assessment of  
the PELAGOS  
OBGCM**

M. Vichi and S. Masina

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

**Table 1.** Skill assessment indices for the ClimPP dataset. The units apply to all indices except the correlation coefficient and MEF that are non-dimensional. Skill measures are defined in the Appendix.

	SST [deg C]	MLD [m]	Chlorophyll	$\log_{10}$ PP (NPP1)	$\log_{10}$ PP (NPP2)
Pearson $r$	0.90	0.68	0.58	0.50	0.58
RMSD <sub>tot</sub>	0.89	19	0.11	0.38	0.27
$B$	−0.22	−13	−0.02	0.29	0.11
AAE	0.64	15	0.08	0.31	0.21
RMSD <sub>cp</sub>	0.85	14	0.10	0.24	0.24
s.d. (PELAGOS)	1.96	15	0.12	0.17	0.25
s.d. (ClimPP)	2.03	19	0.01	0.28	0.28
MEF	0.81	0.02	−0.10	−0.82	0.10

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

**Table 2.** Skill assessment indices for PELAGOS with the BATS JGOFS data. The units apply to all indices except the correlation coefficient, MEF and RI. Variables refer to the surface except MLD and NPP, which is integrated over the MLD in the data and to the euphotic zone depth in the model (~90 m). n.s.=non significant according to an empirical  $p$ -value estimate. Skill measures are defined in the Appendix.

	SST	MLD	Chlorophyll	NPP	Phosphate	Nitrate	Bacteria	BP
	[deg C]	[m]	[mg m <sup>-3</sup> ]	[mg C m <sup>-2</sup> d <sup>-1</sup> ]	[mmol m <sup>-3</sup> ]	[mmol m <sup>-3</sup> ]	[mg C m <sup>-3</sup> ]	[mg C m <sup>-3</sup> d <sup>-1</sup> ]
mean	23.0	43	0.14	275	$8 \cdot 10^{-4}$	0.19	4.96	2.33
mean BATS	23.5	55	0.10	251	$8 \cdot 10^{-3}$	0.04	4.58	2.33
s.d.	2.8	36	0.12	136	0.002	0.009	0.38	0.47
s.d. BATS	2.9	49	0.08	227	0.02	0.11	1.15	1.84
Pearson $r$	0.96	0.81	0.64	0.68	n.s.	n.s.	0.29	0.35
RMSD <sub>tot</sub>	0.94	31	0.10	169	0.02	0.18	1.16	1.73
$B$	-0.56	-12	0.04	24	-0.007	0.15	0.38	0.00
AAE	0.74	19	0.07	132	0.009	0.17	0.90	1.27
RMSD <sub>cp</sub>	0.75	29	0.09	167	0.02	0.11	1.10	1.73
MEF	0.90	0.59	-0.80	0.44	-0.14	-1.6	-0.03	0.11
RI	1.0	1.7	2.2	2.5	n.s.	n.s.	1.4	2.2

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

**Table 3.** Skill assessment indices for PELAGOS with the HOT JGOFS data. The units apply to all indices except the correlation coefficient, MEF and RI that are non-dimensional. n.s.=non significant according to an empirical p-value estimate. Skill measures are defined in the Appendix.

	SST [deg C]	MLD [m]	Chlorophyll [mg m <sup>-3</sup> ]	NPP [mg C m <sup>-2</sup> d <sup>-1</sup> ]	Phosphate [mmol m <sup>-3</sup> ]	Nitrate [mmol m <sup>-3</sup> ]	Bacteria [mg C m <sup>-3</sup> ]
mean	24.5	43	0.04	176	$2 \cdot 10^{-4}$	$3 \cdot 10^{-4}$	4.58
mean HOT	24.8	47	0.09	259	0.08	$5 \cdot 10^{-3}$	4.17
s.d.	1.38	20	0.02	39	$5 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	0.07
s.d. HOT	1.27	19	0.04	103	0.03	0.01	0.90
Pearson <i>r</i>	0.95	0.62	0.41	n.s.	n.s.	n.s.	0.24
RMSD <sub>tot</sub>	0.53	17	0.06	136	0.08	0.01	0.96
<i>B</i>	-0.33	-3	-0.05	-83	-0.07	-0.004	0.41
AAE	0.42	13	0.05	104	0.08	0.005	0.79
RMSD <sub>cp</sub>	0.42	17	0.03	109	0.03	0.01	0.87
MEF	0.83	0.16	-1.9	n.s.	n.s.	n.s.	-0.18
RI	1.0	1.4	3.2	n.s.	n.s.	n.s.	1.2

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

**Table 4.** Comparison of model results with data from Hoppe et al. (2002, Table 1) for the surface Atlantic in the period November 1991 – January 1992. Mean values with standard deviation between brackets. The published bacterial production (BP) data estimated via leucine incorporation were converted to carbon using the factor of  $1.5 \times 10^{-3} \mu\text{g C pmol}^{-1}$ . n.a.=not available. The lower part of the table shows the spatially-integrated annual values of metabolic properties integrated over the euphotic zone (EZ) for the Atlantic and global ocean. Mean November-December-January (NDJ) values have been extrapolated to the annual value.

Data and model	SST [deg C]	Chl [ $\mu\text{g l}^{-1}$ ]	PP [ $\mu\text{g C l}^{-1} \text{ h}^{-1}$ ]	BP [ $\mu\text{g C l}^{-1} \text{ h}^{-1}$ ]	BP/PP [%]	BCD/PP [%]
North Atlantic	20.8 (5.5)	0.35 (0.43)	0.72 (0.65)	0.05 (0.03)	18 (20)	n.a.
PELAGOS NPP1	19.9 (5.2)	0.14 (0.06)	0.77 (0.41)	0.07 (0.04)	21 (9)	53 (22)
NPP2			0.49 (0.22)		33 (21)	85 (49)
South Atlantic	14.6 (10.1)	0.86 (0.95)	1.31 (2.10)	0.09 (0.09)	23 (28)	n.a.
PELAGOS NPP1	15.0 (8.7)	0.72 (0.71)	1.91 (1.39)	0.15 (0.08)	21 (10)	54 (25)
NPP2			1.49 (1.24)		35 (22)	89 (56)
Model (EZ)			PP [ $\text{Pg C y}^{-1}$ ]	BP [ $\text{Pg C y}^{-1}$ ]	BP/PP [%]	BCD/PP [%]
Atlantic (NDJ, extrapolated)			11.31	4.93	44	112
Global (NDJ, extrapolated)			56.40	22.70	40	103
Atlantic (Annual)			11.48	4.33	38	97
Global (Annual)			53.94	20.80	39	99

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

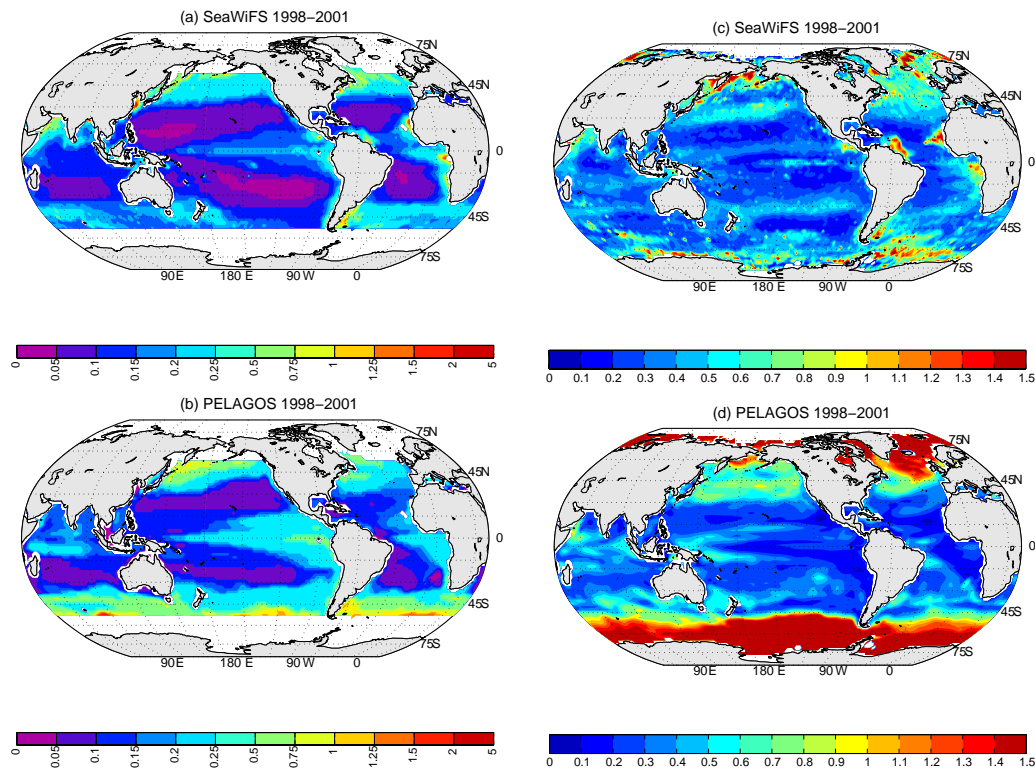
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 1.** Comparison of observed and simulated chlorophyll ( $\text{mg chl m}^{-3}$ ) annual means and coefficient of variation (non-dimensional) for the period 1998–2001: **(a, c)** SeaWiFS; **(b, d)** PELAGOS (average over the euphotic zone depth).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

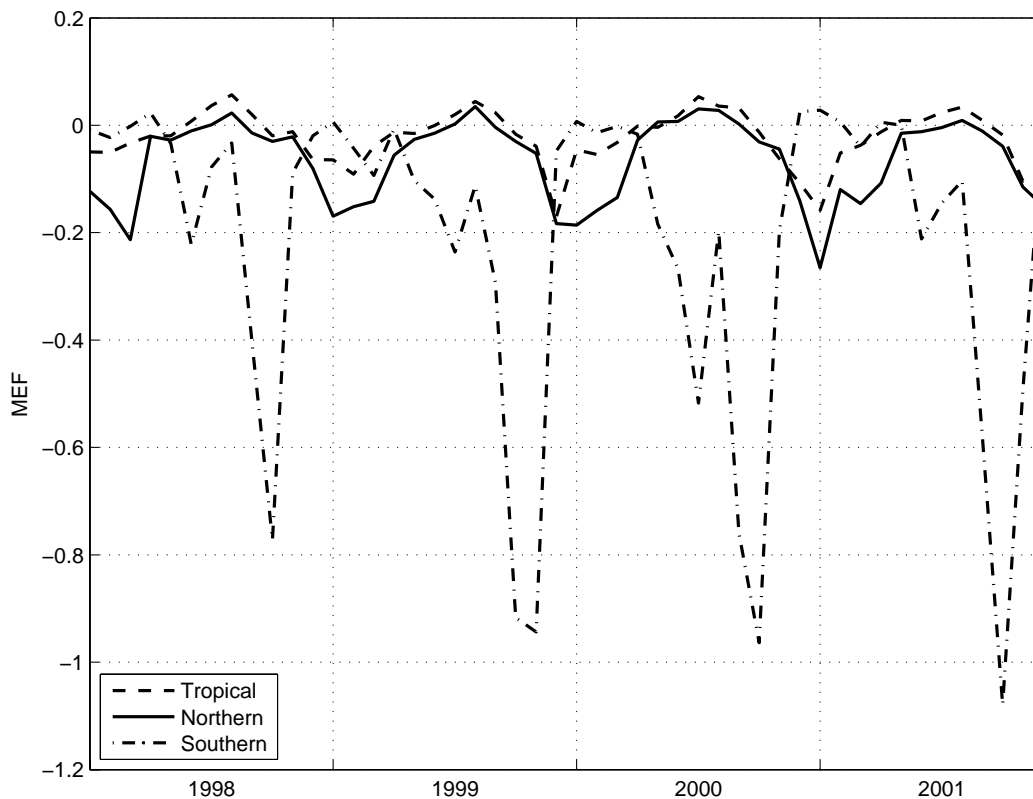
Full Screen / Esc

Printer-friendly Version

Interactive Discussion







**Fig. 2.** MEF index for PELAGOS and SeaWiFS chlorophyll data over the period 1998–2001. The regions are defined as Tropical: 20 S–20 N; Northern Hemisphere: 0–45 N; Southern Hemisphere: 0–45 S.

## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

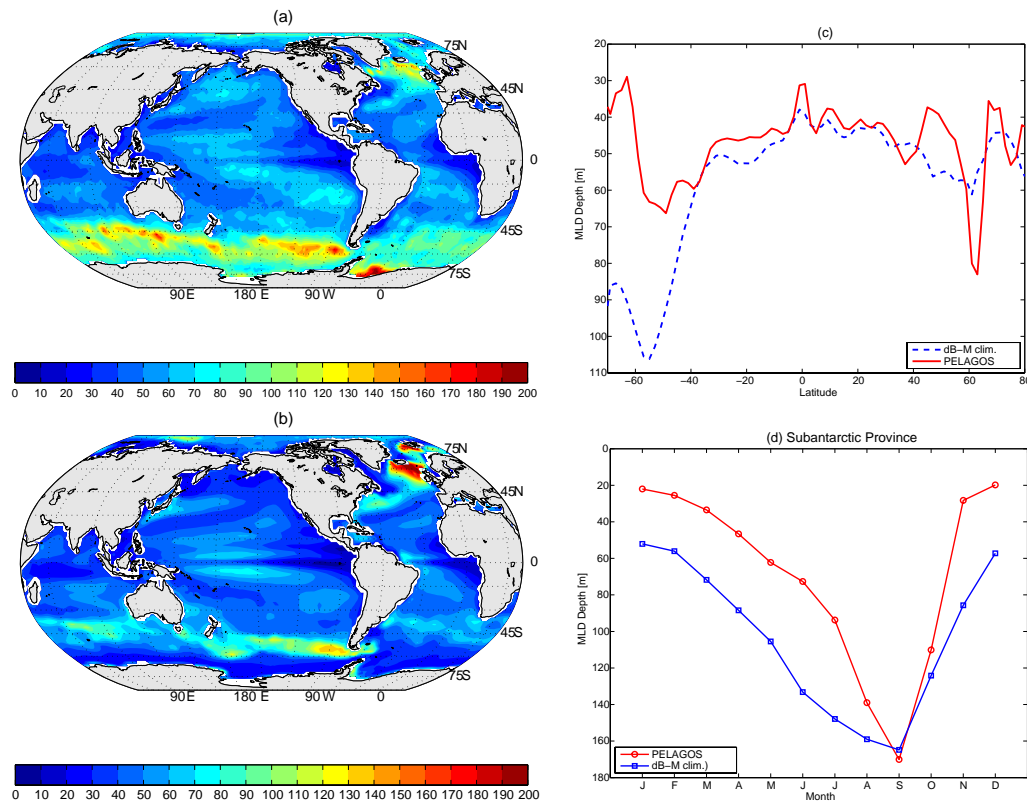
Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Fig. 3.** Comparison of observed and simulated mixed layer depths (MLD in m, temperature criterion  $\Delta T = 0.2^\circ\text{C}$ ). **(a)** de Boyer Montégut et al. (2004) **(b)** PELAGOS means over the whole simulation period; **(c)** comparison of zonal averages **(d)** mean seasonal cycle in the sub-antarctic province.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

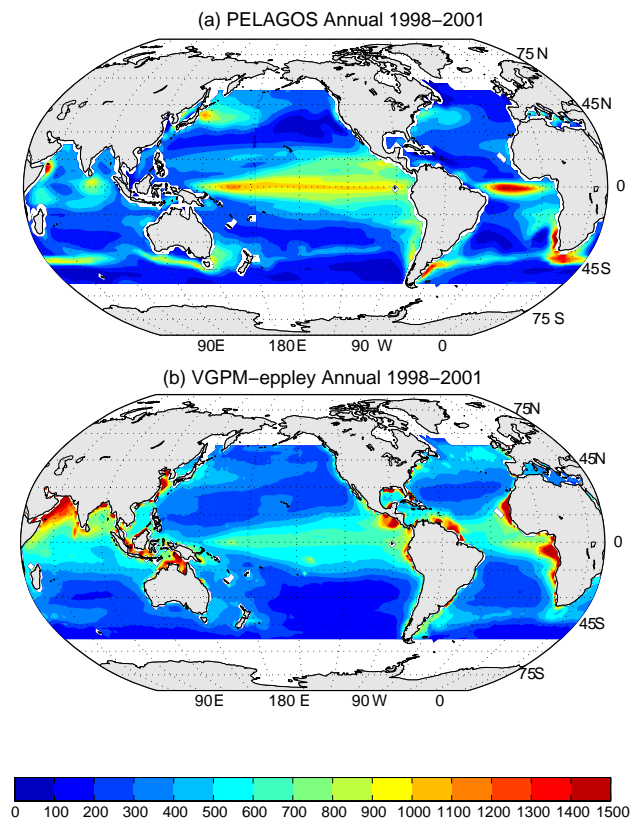
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 4.** Mean annual NPP over the euphotic zone in  $\text{mg C m}^{-2} \text{d}^{-1}$  for the period 1998–2001: **(a)** PELAGOS, **(b)** VGPM with Eppley temperature dependence.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

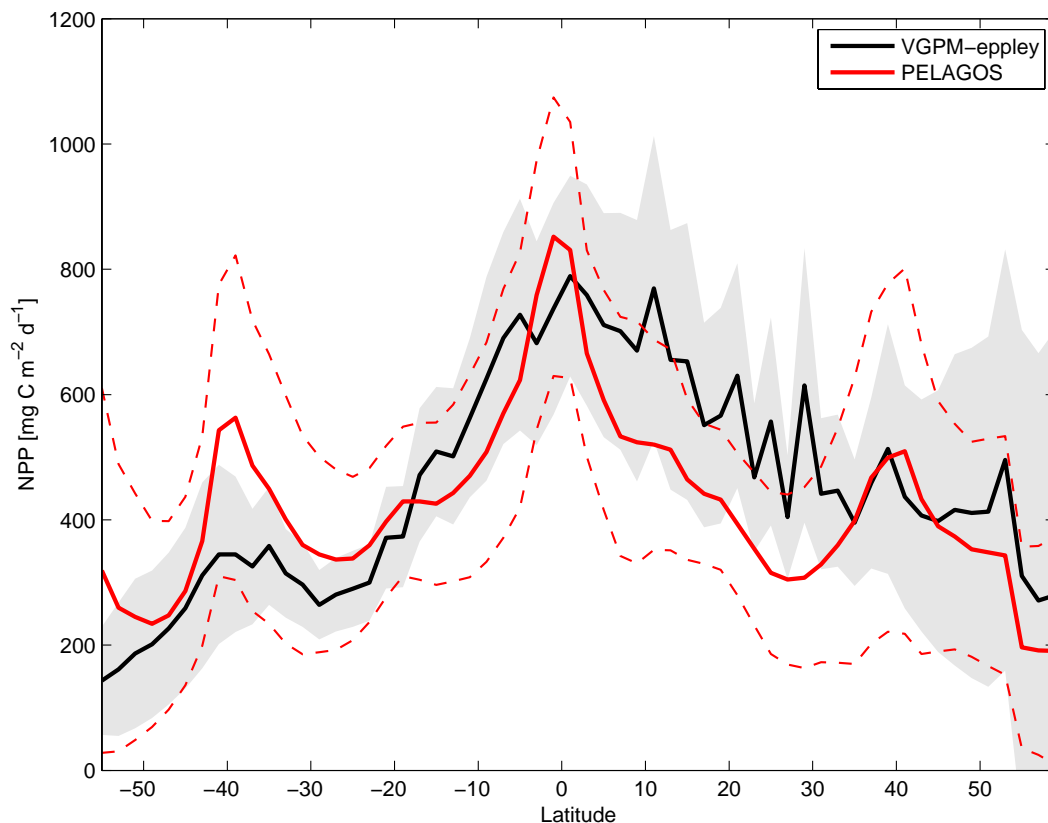
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 5.** Comparison of the zonal annual means and standard deviations of NPP predicted by the VGPM (shaded) and by PELAGOS (dashed lines) over the period 1998–2001.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

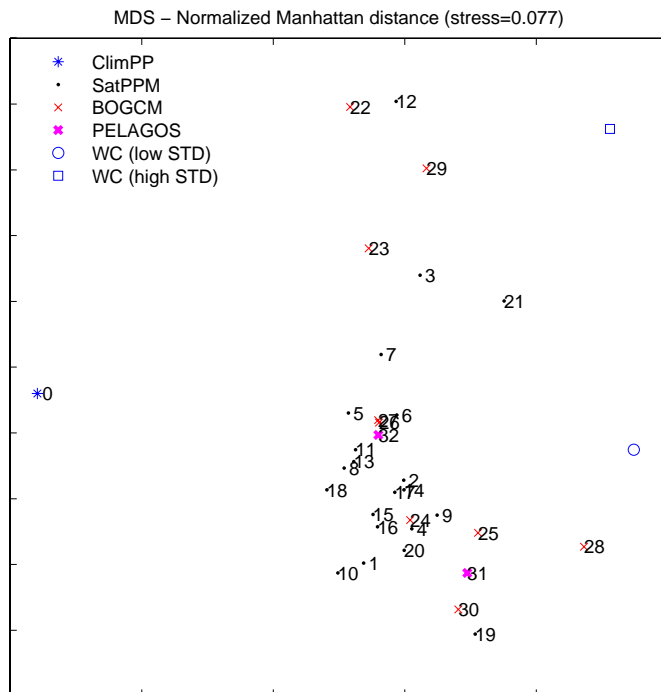
Printer-friendly Version

Interactive Discussion



## Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 6.** MDS representation of the combined performance indices for the models participating to the PPARR3 and for PELAGOS. Numbering and classification (SatPPM and BOGCM) is done according to Friedrichs et al. (2008) with the addition of PELAGOS results (31–32) and the data (0). The points marked with WC represent artificial worst cases derived from the model results space and by assuming the lowest and highest standard deviations with respect to the data.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

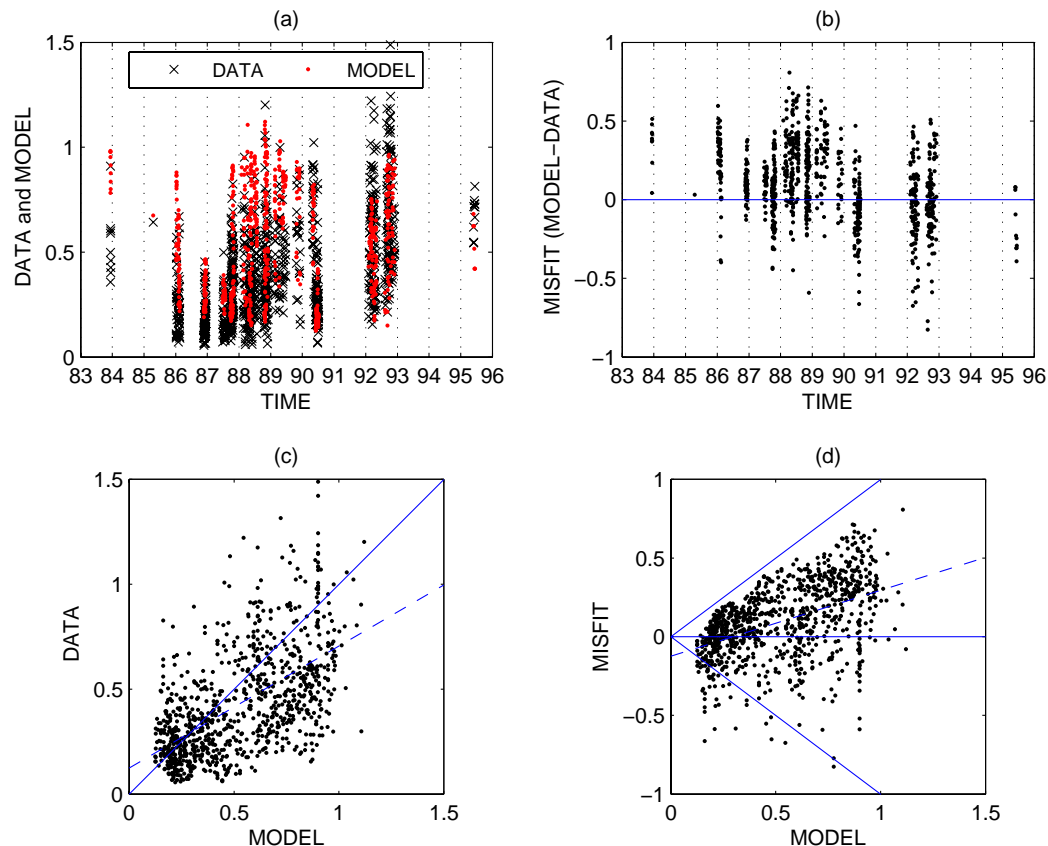
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 7.** Comparison of observed and simulated NPP values in  $\text{g C m}^{-2} \text{d}^{-1}$  for the data collected in the ClimPP dataset. **(a)** data and model timeseries; **(b)** model misfit versus time; **(c)** data and model scatter plot with the 1:1 line and regression line  $\hat{y}_D = 0.58y_M + 0.12$ ,  $r^2 = 0.33$ ; **(d)** model misfit versus model (the continuous lines represent the 0 misfit value and the 1:1 lines where the misfit is equal to the model value; the regression line is dashed).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

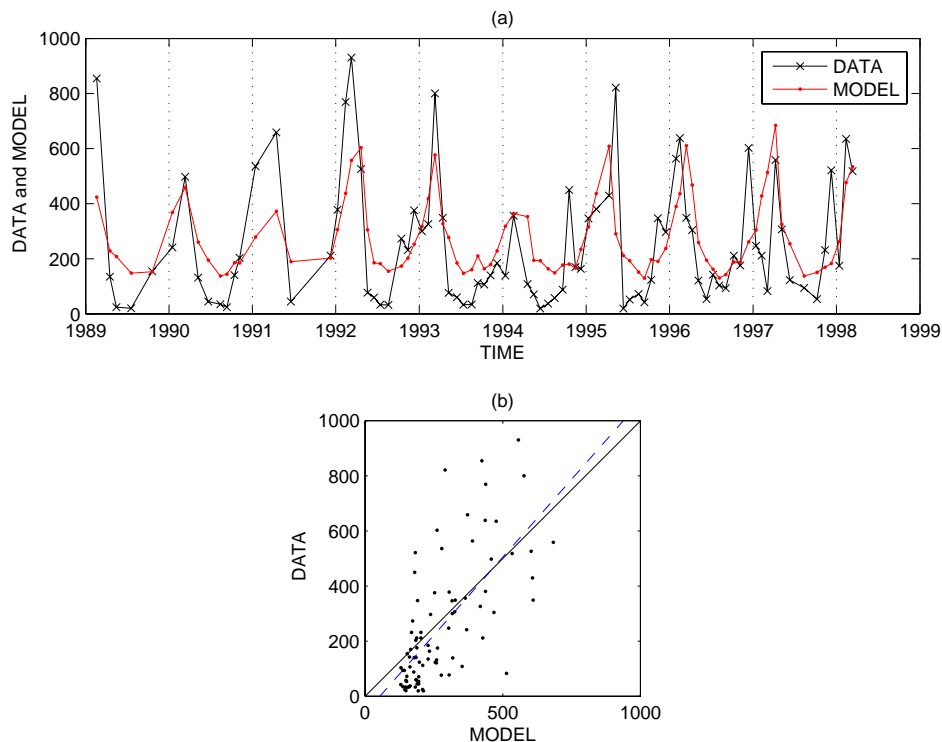
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 8.** Comparison of observed and simulated integrated primary production (in  $\text{mg C m}^{-2} \text{d}^{-1}$ ) at BATS. **(a)** JGOFS BATS time series; **(b)** scatter plot with regression line  $\hat{y}_D = 1.1y_M - 60$ ,  $r^2 = 0.46$  ( $H_0$ : slope=1 cannot be significantly rejected,  $p=0.32$ ).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

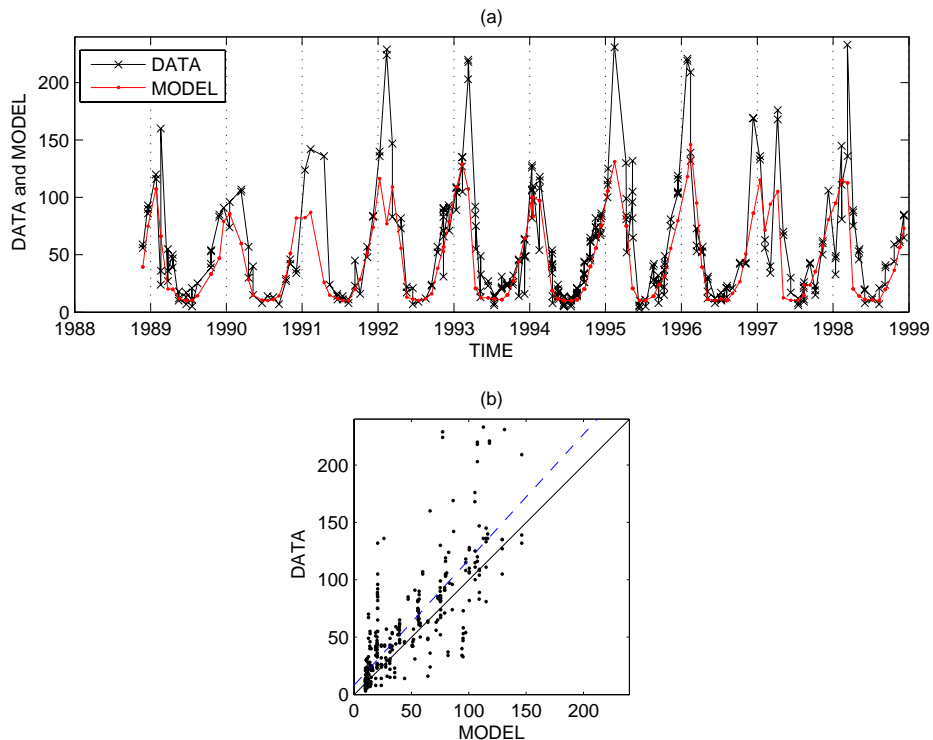
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 9.** Comparison of observed and simulated mixed layer depths (in m) at BATS. **(a)** JGOFS BATS time series; **(b)** scatter plot with regression line  $\hat{y}_D = 1.1y_M + 8$ ,  $r^2 = 0.66$  ( $H_0$ : slope=1 can be rejected according to the choice of the reference  $p$ -value since  $p=0.02$ ).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

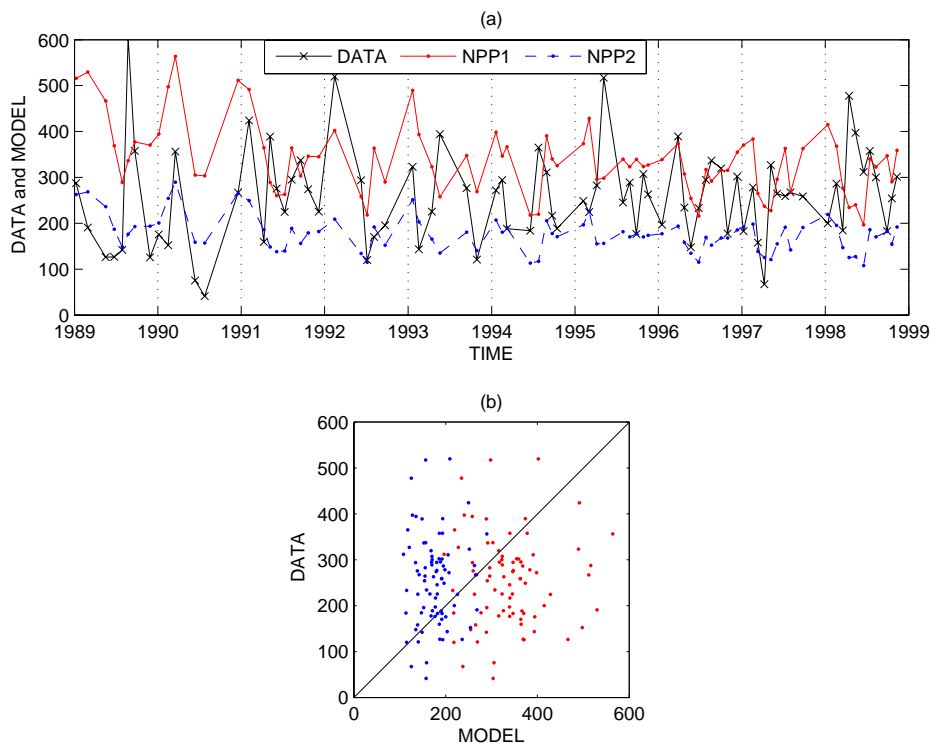
Interactive Discussion





**Skill assessment of  
the PELAGOS  
OBGCM**

M. Vichi and S. Masina

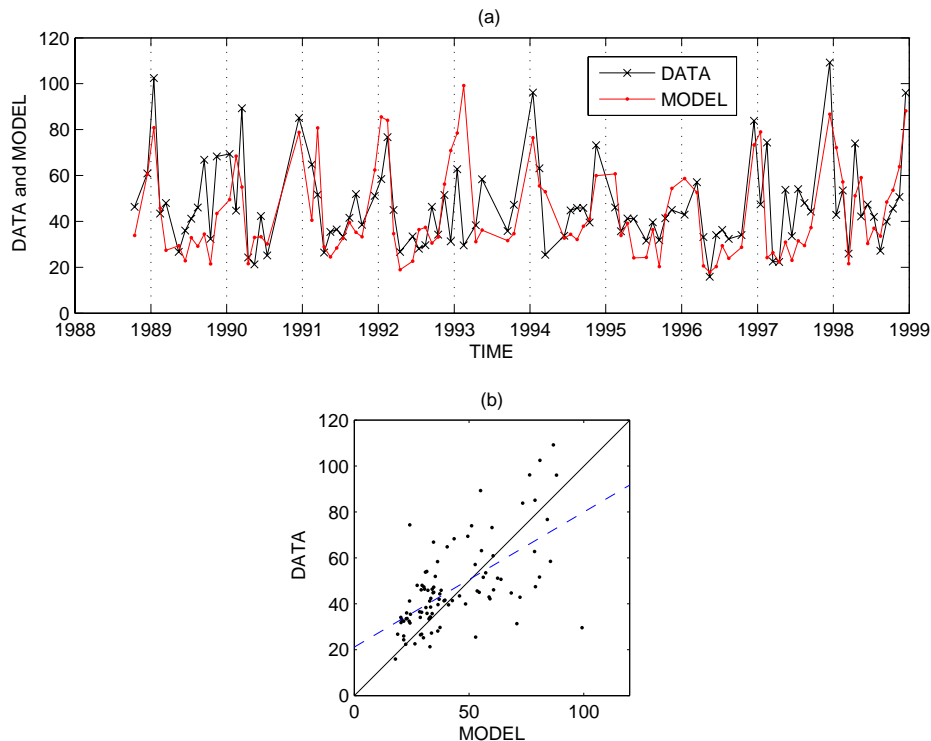


**Fig. 10.** Comparison of observed and simulated integrated primary production (mg C m<sup>-2</sup> d<sup>-1</sup>) at Sta. ALOHA. **(a)** JGOFS HOT time series for the NPP1 and NPP2 estimates (cfr. Sect. 4.2); **(b)** scatter plot of model vs. data for both estimates.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[I◀](#)[▶I](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 11.** Comparison of observed and simulated mixed layer depths (in m) at Sta. ALOHA. **(a)** JGOFS HOT time series; **(b)** scatter plot with regression line  $\hat{y}_D = 0.59y_M + 0.21$ ,  $r^2 = 0.38$  ( $H_0$ : slope=1 can be significantly rejected).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

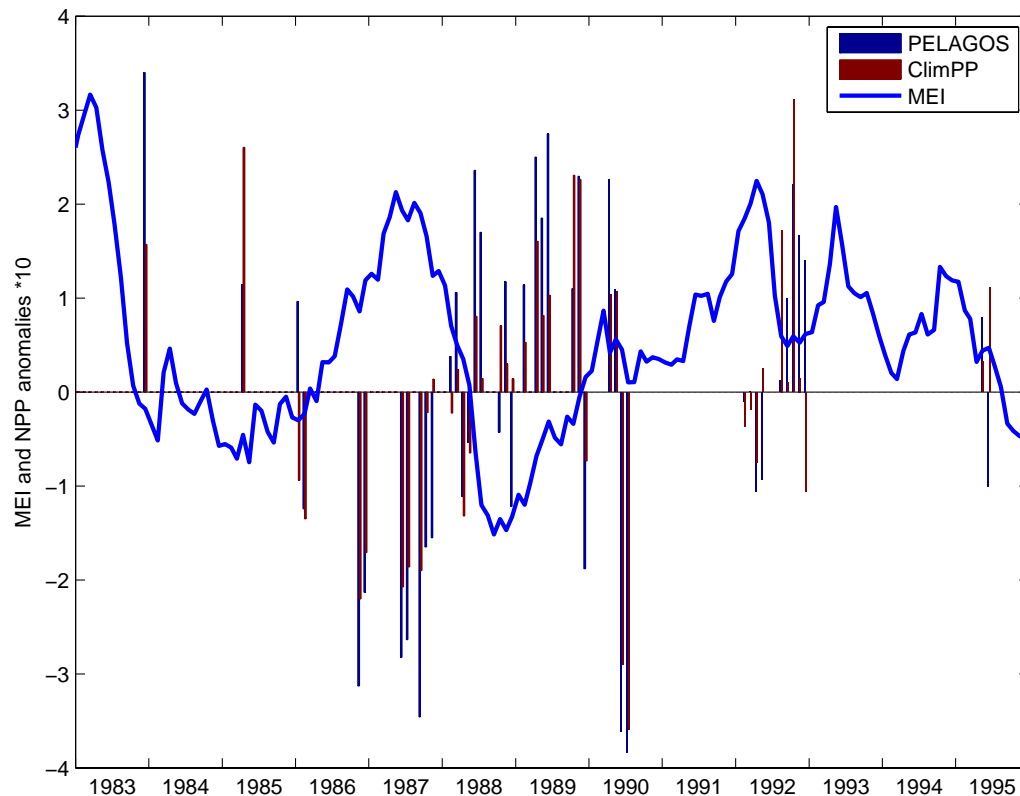
Printer-friendly Version

Interactive Discussion



# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina



**Fig. 12.** Time series of simulated and observed NPP anomalies in  $\text{g C m}^{-2} \text{ d}^{-1}$  (multiplied by 10) from the ClimPP region and time period, overlaid to the Multivariate ENSO Index (MEI) encompassing the whole tropical Pacific. The NPP anomalies are computed using the means of the period 1983–1989 for pre-1990 data and 1990–1995 for post-1990 data.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

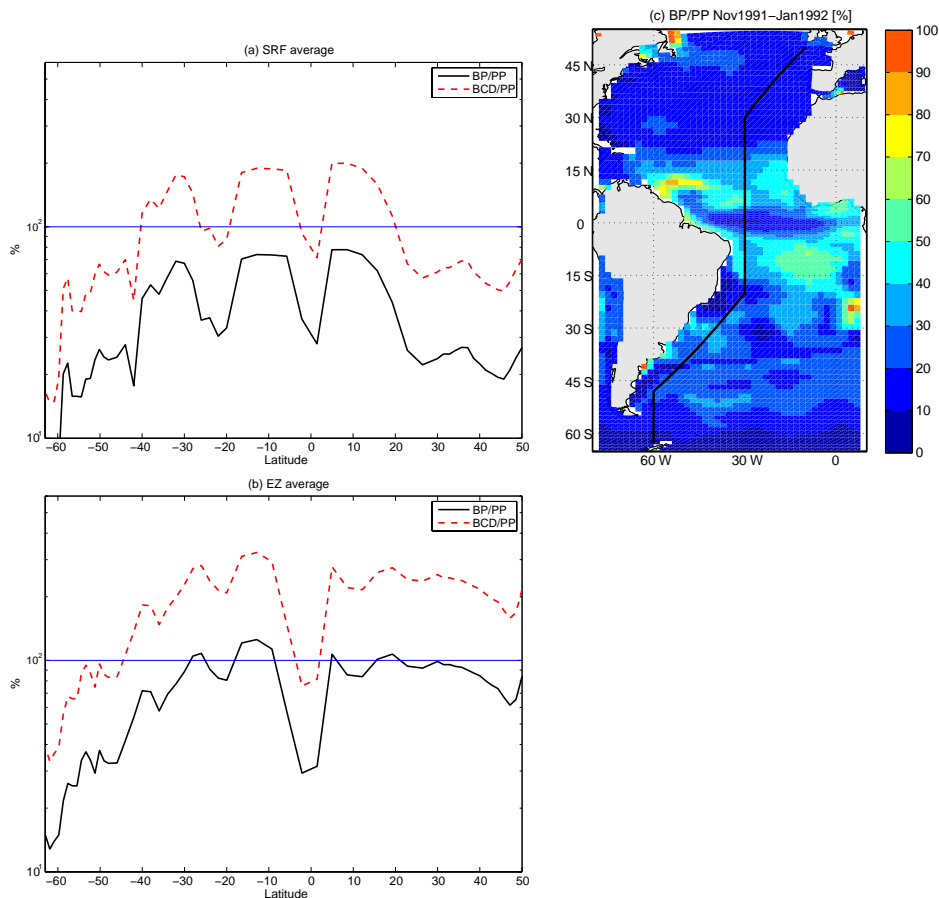
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. 13.** Ratios of Bacterial Production (BP) and Bacterial Carbon Demand (BCD) to net primary production (NPP2) along the Atlantic track of Hoppe et al. (2002): **(a)** surface values; **(b)** values integrated over the euphotic zone depth; **(c)** map of the simulated annual mean value of BP/PP with the location of the transect.

# Skill assessment of the PELAGOS OBGCM

M. Vichi and S. Masina

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion