Biogeochemical validation of an interannual simulation of the ROMS-PISCES coupled model in the Southeast Pacific

Validación biogeoquímica de una simulación interanual del modelo acoplado ROMS-PISCES en el Pacífico Sudeste

Dante Espinoza-Morriberon 1*, Vincent Echevin 2, Jorge Tam 1, Jesús Ledesma 1, Ricardo Oliveros-Ramos 1, Jorge Ramos 1 and Carlos Y. Romero 1,3

Abstract
Currently biogeochemical models are used to understand and quantify biogeochemical processes in the ocean. The objective of the present study was to validate predictive ability of a regional configuration of the PISCES biogeochemical model on main biogeochemical variables in Humboldt Current Large Marine Ecosystem (HCLME). The statistical indicators used to evaluate the model were the bias, root-mean-square error, correlation coefficient and, graphically, the Taylor’s diagram. The results showed that the model reproduces the dynamics of the main biogeochemical variables (chlorophyll, dissolved oxygen and nutrients); in particular, the impact of El Niño 1997-1998 in the chlorophyll (decrease) and oxygen minimum zone depth (increase). However, it is necessary to carry out sensitivity studies of the PISCES model with different parameters values to obtain a more accurate representation of the properties of the Ocean.

Keywords: PISCES model; HCLME; chlorophyll; dissolved oxygen; nutrients.

Resumen
Los modelos biogeoquímicos en la actualidad son utilizados para entender y cuantificar los principales procesos biogeoquímicos que suceden en el océano. El objetivo del presente estudio es validar estadísticamente la habilidad predictiva de una simulación del modelo biogeoquímico PISCES en reproducir la dinámica de las principales variables biogeoquímicas del Ecosistema de la Corriente de Humboldt (ECH). Para evaluar el modelo se utilizaron indicadores estadísticos: sesgo, error raiz cuadrado medio, coeficiente de correlación y gráficamente el diagrama de Taylor. Los resultados muestran que el modelo es capaz de reproducir la dinámica de las principales variables biogeoquímicas (clorofila, oxígeno disuelto y nutrientes), captando bien el impacto que tiene El Niño 1997-1998 en la clorofila (disminución) y profundidad de la zona mínima de oxígeno (incremento). Es necesario llevar a cabo estudios de sensibilidad del modelo PISCES usando diferentes valores de los parámetros principales para obtener una mejor representación de las propiedades biogeoquímicas del océano.

Palabras claves: modelo PISCES; ECH; clorofila; oxígeno disuelto; nutrientes.
Introduction

Oceanographic models are used to study several ocean processes (i.e. Kelvin waves impacts, coastal upwelling) and, in recent years, biogeochemical models have been developed to understand, quantify and predict the main biogeochemical processes and the complex dynamics between nutrients and plankton. In addition, biogeochemical models are coupled or used as forcings for High Trophic Level (HTL) models, aiming to understand ecosystem dynamics.

Intermediate complexity biogeochemical models, like NPZD type models, simulate the interaction between nitrate, phytoplankton, zooplankton and detritus (Edwards 2001, Heine & Slawig 2013). More complex biogeochemical models like PISCES model (Aumont & Bopp 2006) or BioEBUS model (Gutknecht et al. 2013) include more variables and processes, simulating a more complete representation of the ocean biogeochemistry and plankton dynamics.

Biogeochemical models have been used in previous studies in Humboldt Current Large Marine Ecosystem (HCLME) in order to understand the climatological processes affecting the Oxygen Minimum Zone (OMZ) (Montes et al. 2014) and the climatological (Echevin et al. 2008, Albert et al. 2010) and interannual (Echevin et al. 2014) variability of phytoplankton and nutrients.

In the present study, we used statistical metrics to assess the skill of the coupled hydrodynamic and biogeochemical ROMS-PISCES model using data from surveys, remote sensing and international databases.

We analyzed the climatological and interannual (1992 – 2008) biogeochemical variability, including the impact of El Niño 1997 – 1998 event on the chlorophyll concentration and depth of the OMZ.

Material and methods

Study area.- The study area comprised the HCLME (0°S – 40°S and 70°-W – 100°W). To validate the biogeochemical model, two zones were defined: a) the area off Peru within 200 km and between 06°S – 16°S; b) the area off Chile within 200 km and between 30°S – 40°S (Fig. 1). The seasons were defined as follows: summer (January – March), autumn (April – June), winter (July – September) and spring (October – December).

PISCES biogeochemical model.- The PISCES biogeochemical model (Pelagic Interaction Scheme for Carbon and Ecosystem Studies) simulates marine biological productivity and describes biogeochemical cycles like carbon and major nutrients in the ocean (Aumont & Bopp 2006). PISCES model assumes that phytoplankton growth depends on external concentration of nutrients and that the main nutrients in the medium follow the Redfield ratio (C:N:P = 106:16:1) (Redfield et al 1963). PISCES has 24 state variables, such as: nutrients (Phosphorus, Nitrogen, Silica and Iron), dissolved oxygen, two types of detritus (large and small), two classes of zooplankton (microzoooplankton and mesozoooplankton) and two classes of phytoplankton (nanophytoplankton and diatoms) (Fig. 2). Diatoms differ from nanophytoplankton in their requirements of silicates, an increased consumption of iron and higher levels of nutrient saturation due to its larger size (Echevin et al. 2008).

Figure 1. Coastal zone between 0-200 m (grey) and zones for analyzing the behavior of the modeled variables off Peru and Chile.

Model configuration.- In the present study the PISCES model is coupled to the hydrodynamical model ROMS (Regional Ocean Modeling System) (Shchepetkin & McWilliams 2005), following the coupling principles of Gruber et al. (2006), who coupled ROMS to a simpler biogeochemical model than PISCES. The physical model ROMS has been used in several previous works to simulated the dynamics of the HCLME (Penven et al. 2005, Colas et al. 2008, Montes et al. 2010, Echevin et al. 2012, Illig et al. 2014).

Our simulation covered a larger area to the north (10°N) than the HCLME to reproduce more accurately the equatorial circulation, because surface and subsurface equatorial currents can affect the dynamics (Montes et al. 2010), oxygenation and productivity (Espinoza-Morriberon 2012) of Peruvian waters. The model spatial resolution was 1/6° with 32 sigma vertical levels (which follow the topography of the ocean floor), the output simulations were recorded of every averaged 6 days from 1992 to 2008.

Atmospheric forcings were obtained from two sources: (1) merging climatological SCOW data (Risien & Chelton 2008) with NCEP anomalies (www.ncep.noaa.gov/) for wind fields and, (2) merging COADS climatology data (Da Silva et al. 1994) with NCEP anomalies for the heat fluxes and air temperatures. For boundary conditions the outputs of the global ORCA2-PISCES physical-biogeochemical coupled model (Aumont & Bopp 2006) were used. More details about model configuration are described in Echevin et al. (2010, 2012) and Cambon et al. (2013).
surface chlorophyll), the following climatologies and interannual time series were used (Table 1):

**Chemical.** Oxygen and nutrients climatologies were obtained from CSIRO Atlas of Regional Seas “CARS 2009” (Ridgway et al. 2002) at a spatial resolution of 1/2º and a vertical resolution of every 5 m depth (from 0 – 200 m) and every 100 m depth (from 200 – 1000 m). Monthly series from 1992 to 2008 of the depth of the upper limit of the OMZ were obtained by merging the database of the Instituto del Mar del Peru (IMARPE) and the international database of WOD09 (García et al. 2010), following the methodology developed by Musial et al. (2011); the reader is referred to Bertrand et al. (2011) for more details about oxygen data processing.

**Biological.** Surface chlorophyll climatological and interannual time series was obtained from the database of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) (Behrenfeld & Falkowski 1997) at a spatial resolution of 1/12º and a monthly temporal resolution from 1997 to 2008.

All observed time series were interpolated on the model temporal resolution using the Laplace transformation method implemented in the FERRET program (http://ferret.pmel.noaa.gov/Ferret) to make comparisons between model and observations.

**Data analysis.** Goodness of fit between observed and simulated data was measured using the following statistical indicators (Vichi and Masina 2009, Lehmann et al. 2009):

- **Bias (S),** which reflects the overestimation (positive) or underestimation (negative) of a simulation:
  
  $$ S = \frac{1}{n} \sum (M - O) $$  
  
  Where \( n \) represents the number of pairs between observed \((O)\) and modeled \((M)\) data by time each bin. However it is necessary to keep in mind that positive and negative deviations in the model tend to cancel, so that the bias would represent the persistence of a difference between the modeled \((M)\) and observed \((O)\) values.

- **Root mean square error (RMSE),** reflects whether the modeled values differ in magnitude from the observed values, so that a small RMSE would indicate good agreement between modeled and observed values:
  
  $$ \text{RMSE} = \sqrt{\frac{\sum (M - O)^2}{n}} $$  
  
  Where \( \text{cov} (M, O) \) is the covariance between the time series of \( M \) and \( O \), and \( \sigma_M \) and \( \sigma_O \) are the standard deviations of \( M \) and \( O \) respectively.

- **Correlation coefficient,** allows us to assess how strongly the temporal variations of the modeled and the observed values are related:
  
  $$ \rho(M, O) = \frac{\text{cov}(M, O)}{\sigma_M \sigma_O} $$  

- **Taylor diagrams** (Taylor 2001) were used to evaluate the goodness of fit of the model: 1) between areas (off Peru and Chile), and 2) between temporal variations (climatological and interannual). Taylor diagram integrates the RMSE, the correlation coefficient and the standard deviation in a single diagram, to compare the fields that are tested (coming from one or more models) and a reference field (representing the “truth” or the observed field).

**Results and discussion**

**Surface chlorophyll.** SeaWiFS interannual satellite data (1997-2008) were interpolated on the model spatial resolution

---

**Figure 2.** PISCES model architecture. The ecosystem model is shown omitting the carbonate system. The grey arrows indicate the processes that influence the concentration of oxygen (modified from Aumont & Bopp 2006).
(~18 km) and the climatologies (for satellite and simulated data) were calculated between the years 2000-2008 to reduce the effect of the strong 1997-98 “El Niño”. Figure 3 (a and b) shows that the simulation represents adequately the average spatial distribution of surface chlorophyll observed, reproducing larger chlorophyll values near the coast. In addition, the area of greatest productivity of the HCLME (Northern Central Peru) and the simulated distribution of the 1 mg Chl/m³ isoline, used as a proxy to define the upwelling area (Nixon & Thomas 2001), follows the annual pattern observed in SeaWiFS data. The high productivity observed off the Northern Center Peruvian region could be generated by the widening of the continental shelf in this area, which would provide a high amount of iron to the water column. The iron is generally a key limiting element for phytoplankton (Bruland et al. 2005).

Goodness of fit indicators validated the model within the study area: within 200 km of the coast in Peruvian area, mean values of RMSE around 3.2 mg Chl/m³ and a mean positive bias of ~+2.5 mg Chl/m³ (Fig. 3c and 3d) were observed; furthermore, in the whole modelling region RMSE values mainly fluctuated between 0 at 0.5 mg Chl/m³ and bias values between -0.1 to 0.1 mg Chl/m³. The overestimation of the simulations compared to the observed pattern could be influenced by the presence of clouds near the coast off Peru that could have masked some phytoplankton blooms. Another explanation for the model overestimation would be the lack of two-way coupling between PISCES and high trophic level model (to connect large zooplankton to top predators), which could influence phytoplankton mortality (Traver & Shin 2010). Also, a slightly positive bias was observed in the equatorial upwelling, which

<table>
<thead>
<tr>
<th>MODELED VARIABLE</th>
<th>OBSERVED DATA</th>
<th>VALIDATION AREA</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMZ depth</td>
<td>Climatology</td>
<td>Peru, Chile</td>
</tr>
<tr>
<td></td>
<td>Interannual</td>
<td>Peru</td>
</tr>
<tr>
<td>Surface chlorophyll</td>
<td>Climatology</td>
<td>Peru, Chile</td>
</tr>
<tr>
<td></td>
<td>Interannual</td>
<td>Peru</td>
</tr>
<tr>
<td>Surface nutrients</td>
<td>Climatology</td>
<td>Peru, Chile</td>
</tr>
</tbody>
</table>

Table 1. Modeled variables in PISCES and observed data bases used in the present study.

Figure 3. Comparison of the observed and modeled surface chlorophyll. Annual average of the modeled surface chlorophyll (a) and SeaWIFS (b) in mg Chl/m³ with the isoline of 1 mg Chl/m³. Map of RMSE values (c) and bias (d) calculated between PISCES model and SeaWIFS surface chlorophyll a for the period simulation.
has also been observed by Albert et al. (2010), possibly because the model overestimate the equatorial upwelling.

The annual cycle of surface chlorophyll within 200 km was acceptably reproduced by the model off the coasts of Peru and Chile compared to the SeaWiFS data (from here, satellite values will be given in parentheses). Seasonality off Peru had larger amplitude than in Chile. The mean seasonal chlorophyll concentration presented highest values off Peru during the summer and autumn months with 3.5 (3.2) mg Chl/m³ and lower values during winter with 2.5 (1.2) mg Chl/m³, while off Chile, highest values were observed during spring and summer with 1.2 (1.5) mg Chl/m³ and lower values during winter with 0.9 (0.92) mg Chl/m³ (Fig. 4). Within the coastal area off Peru, overestimation of simulations compared to SeaWiFS was also observed by Echevin et al. (2008) during all climatological months (+1 mg Chl/m³) and by Albert et al. (2010) especially during winter, who mentioned that it could be due to the reduced availability of satellite data near the coast by the presence of clouds in this period. In Chile, this overestimation is weaker than off Peru due to fewer clouds in this area (Demarcq, pers. comm.).

In figure 5 we can observe that the model reproduced the interannual pattern of SeaWiFS chlorophyll data off Perú; however with more productive values in most of the studied period (1997 to 2008) with the exception of the summer months; furthermore off Chile the simulation also followed adequately the interannual pattern and the magnitude of SeaWiFS values. The correlation between observations and simulations was 0.56 (Peru) and 0.43 (Chile) (p < 0.05 for both correlations). Echevin et al. (2014) also found an interannual overestimation compared to SeaWiFS off Peru (almost the same magnitude as our simulation) and mentioned that cloud presence makes difficult the validation process. The effects of El Niño were captured by the model showing decreased values of chlorophyll (0.6 mg Chl/m³) during the years 1997/1998 off Peru. The decreased level of chlorophyll during “El Niño” along Peruvian coast is caused by poor waters advected from Equator region and by more frequency of downwelling Kelvin waves that produce the deepening of the nutricline (Echevin et al. 2014).
Taylor diagram (Fig. 6) comparing climatological and interannual variations of chlorophyll in the areas off Peru and Chile within 200 km of the coast showed that the model has a better goodness of fit off Peru than off Chile; and in addition, goodness of fit was better for the chlorophyll seasonal cycle than for the interannual variation in both areas.

Surface nutrients.- Nutrients climatology from CARS was also interpolated on 18 km resolution grid and simulated nutrients climatology was calculated for the period 1992 – 2008 to be comparable with CARS data, taking into account about 50 years of observed data.

On annual average, the model represented adequately in the Peruvian region the spatial distribution of observed nutrients, with higher concentrations near the coast; however, for the three nutrients used in the present study (nitrates, phosphates and silicates) a positive bias of the concentrations was obtained for the most study area (not shown). The overestimation of nutrients in the PISCES model compared to observations, although with less magnitude in our simulations, has also been observed by Echevin et al. (2008) and Albert et al. (2010), who explained that it could be due to high concentrations of nutrients in the boundary conditions and that the upwelling of nutrients during winter would be stronger in the model than in reality. In the configuration of Echevin et al. (2014), nitrate and phosphate are also overestimated (+1 y +0.3 μmol/L, respectively) and silicate is underestimated (~-1.5 μmol/L) compared to IMARPE data.

The seasonal cycle showed more nutrients during winter-spring possibly due to more vertical mixing; and less concentration during summer, when there is more nutrient consumption by phytoplankton. This variability is also observed by Calienes et al. (1985).

In Figure 7 Taylor diagram indicated that modelled nutrients had a higher goodness of fit off Peru than off Chile.

Depth of the OMZ.- The depth of the OMZ in this study is defined as the depth at which the concentration of dissolved oxygen is 88 μmol/kg (2 mL/L), because it is the concentration limit for the optimal development of pelagic species (Bertrand et al. 2011). Dissolved oxygen data from CARS on surface and vertical levels were interpolated to 18 km.

Figure 6. Taylor diagram to assess the goodness of fit of the climatological and interannual series of modeled surface chlorophyll compared to SeaWIFS data off Peru (blue) and off Chile (red) region.

Figure 7. Taylor diagrams to assess the goodness of fit of the climatology of modeled nutrients (silicate, phosphate and nitrate) compared to climatological values observed in CARS off Peru and off Chile region.
The model represented adequately the spatial distribution of the depth of the OMZ, with a shallower OMZ near the ocean off Peru, as observed in CARS (Fig. 8a and 8b). The presence of a stronger and shallower OMZ off Peru could be due to: 1) a high rate of remineralization influenced by high primary productivity, 2) a lack of ventilation due to the presence of a strong pycnocline that keeps subsurface waters isolated from the atmosphere, plus deoxygenation of equatorial currents that are the main source of oxygen in this zone (Stramma et al. 2010.) and; 3) the long residence time of the waters with a cumulative reduction of oxygen (Czeschel et al. 2011).

RMSE values fluctuated between 0 to 50 µmol/kg and the bias between -20 to 40 µmol/kg. Lower values of RMSE and bias were obtained off Peru; however, an underestimation in the equatorial zone and an overestimation at the south of the HCLME (between 20°S to 40°S) of the depth of the OMZ was observed (Fig. 8c and 8d). The greatest differences observed at the equator and at the south of the HCLME could be due to a slight lack of representativeness of the intensities of the currents in the boundary conditions used in the configuration of our simulation in these areas, which could lead to less oxygen flow in the equatorial zone and greater oxygen flow off Peru (Espinoza-Morriberón 2012).

The model represented adequately the seasonal signal of the OMZ upper limit depth off Peru; however the values from CARS and IMARPE were underestimated on average throughout the year. The OMZ is shallower during the summer-autumn and deeper during winter-spring (Fig. 9b), probably related to: (a) during winter the vertical mixing is stronger due to more wind intensity and (b) during autumn after high phytoplankton biomass dying, organic matter is consumed by bacteria during the remineralization process with a consequently high consumption of dissolved oxygen (Libes 1992, Paulmier et al. 2006). Interannually, the model reproduced the oxygenation events during “El Niño” (shallower OMZ) and the observed dissolved oxygen from IMARPE; however, it presented some deficiencies in following the interannual pattern (Fig. 9a). The correlation between simulations and observations off Peru was 0.42 (p<0.05). The deepening of the OMZ during “El Niño” could be due to the impact of Kelvin waves and the intrusion of surface tropical waters rich in oxygen and poor in nutrients. Taylor diagram (Fig. 9c) showed for OMZ depth that correlation with IMARPE data was higher for the climatology than for interannual variations.

In other hand, vertically the model reproduced relatively well the vertical structure of dissolved oxygen off Peru and Chile compared to CARS data (Fig. 10); however, the modeled OMZ off Peru tend to be larger than CARS OMZ. The greater intensity of the OMZ in front of Peru is the result of high primary production reproduced in the model; another aspect to consider is the values of the remineralization of dissolved organic carbon (0.3 d−1) and nitrification (0.05 d−1) rates used in our configuration that influence oxygen consumption in the model (Espinoza-Morriberón 2012). Sensitivity studies of PISCES model at different values of these parameters need to be explored in the future. CARS climatological data base could also present deficiencies in accurately represent the values within the OMZ, because they are based on World Ocean Atlas (WOA)

---

**Figure 8.** Annual average OMZ depth in meters, defined by the 88 µmol O₂/kg isoline from ROMS-PISCES model (a) and CARS data (b). Map of the RMSE values (c) and the bias (d) calculated of the OMZ depth from PISCES model and CARS data for the period simulation.
Figure 9. OMZ depth in meters, defined by the 88 µmol O₂/kg isoline. (a) Mean monthly series (3 month moving average) from 1992 to 2008 of the modeled OMZ depth (black line) and observed data from IMARPE (dashed blue line) off Peru; (b) seasonal variation of modeled OMZ depth (black line), the observed OMZ depth in CARS (dashed black line) and IMARPE data (dashed blue line) off Peru; and (c) Taylor diagram assessing the goodness of fit of the climatological and interannual series of modeled OMZ depth compared to observed OMZ depth from CARS and IMARPE respectively off Peru. The time series were computed within 0-200 km from the coast and from 4°S to 16°S.

Figure 10. Zonal section at 10°S (top) and 30°S (bottom) of the annual average of modeled dissolved oxygen (left) and observed data in CARS (right) in µmol/kg from 0 to 1000 meters in depth.
data, which tend to overestimate the O₂ within the OMZ due to various factors such as: positive biases in old measurements, interpolation artifacts and the effects of variability in ocean circulation (Bianchi et al. 2012).

Coupled ROMS-PISCES model has been validated in other works for the Peruvian region, explaining the interactions of phytoplankton and nutrients in climatological (Echevin et al. 2008, Albert et al. 2010) and interannual (Echevin et al. 2014) mode; however in the present configuration we reproduced the interannual dynamics of biogeochemical variables on a longer period (17 years) and during the occurrence of an extreme El Niño (1997−1998). On the other hand, the dissolved oxygen dynamics was studied and we could observe El Niño impact on it.

Conclusions
The present configuration of the ROMS-PISCES model reproduced adequately the dynamics of the main biogeochemical variables (chlorophyll, nutrients and the depth of the oxycline) in the HCLME. The model showed a very productive system near the coast and the seasonal cycles of chlorophyll, nutrients and depth of the OMZ were acceptably reproduced. Regarding the interannual signal, the model captured well the decrease in chlorophyll concentration and the increase in OMZ depth produced by "El Niño"; however we have to keep in mind that the model represented the climatological variations better than the interannual variations.

It is noteworthy that this study presents one of the first attempts to reproduce the dynamics of biogeochemical variables in the HCLME (Echevin et al. 2008, Albert et al. 2010, Echevin et al. 2014, Montes et al. 2014). However, the present configuration of the model still has several problems to be solved. It is necessary to conduct sensitivity analysis of the model with different boundary conditions, forcings and different values of key parameters of the PISCES model, such as rates of remineralization and nitrification, among others.

Acknowledgements
We acknowledge Christophe Hourdin for providing model configurations and running the simulation performed within the PEPS project, Carlos Quispe from LMOECC (IMARPE) for their valuable comments on the statistical analysis. This work was published with the support of IRD and the JEAI EMACEP (Ecología Marina Cuantitativa del Ecosistema de A/fluoramiento) and the LMI DISCOH (Dinamica de Sistema de Corriente de Humboldt).

Bibliography


