

OPTIMIZACION DE LA EJECUCIÓN DE LOS MODELOS REGIONAL OCEAN MODELING SYSTEM (ROMS) Y COASTAL AND REGIONAL OCEAN COMMUNITY (CROCO) EN EL CLÚSTER COMPUTACIONAL DE ALTO RENDIMIENTO DEL IMARPE

OPTIMIZING THE EXECUTION OF THE REGIONAL OCEAN MODELING SYSTEM (ROMS) AND COASTAL AND REGIONAL OCEAN COMMUNITY (CROCO) MODELS ON THE HIGH-PERFORMANCE COMPUTATIONAL CLUSTER OF IMARPE

Jorge Ramos-Flores¹ Carlos Romero¹ Jorge Tam¹
 Adolfo Chamorro-Gómez¹

RESUMEN

RAMOS-FLORES, J., ROMERO, C., TAM, J. & CHAMORRO-GÓMEZ, A. (2023). Optimización de la ejecución de los modelos Regional Ocean Modeling System (ROMS) y Coastal and Regional Ocean Community (CROCO) en el clúster computacional de alto rendimiento del IMARPE. *Inf Inst Mar Perú*, 50(4), 460-465.- Este trabajo tiene como propósito determinar el número de núcleos en los cuales el modelo *Regional Ocean Modeling System* (ROMS) versión AGRIF y el modelo *Coastal and Regional Ocean Community* (CROCO), consiguen su mejor tiempo de ejecución, para un dominio y resolución de grilla determinados, cuando es ejecutado en el Clúster computacional de alto rendimiento del Instituto del Mar del Perú (IMARPE).

PALABRAS CLAVE: modelado, clúster, ROMS, CROCO, Perú

ABSTRACT

RAMOS-FLORES, J., ROMERO, C., TAM, J. & CHAMORRO-GÓMEZ, A. (2023). Optimizing the execution of the Regional Ocean Modeling System (ROMS) and Coastal and Regional Ocean Community (CROCO) models on the high-performance computational cluster of IMARPE. *Inf Inst Mar Perú*, 50(4), 460-465.- This study aims to determine the optimal number of cores for achieving the best execution time for the Regional Ocean Modeling System (ROMS) AGRIF version and the Coastal and Regional Ocean Community (CROCO) model, given a specific domain and grid resolution, when executed on the high-performance computational cluster of IMARPE.

KEYWORDS: modeling, cluster, ROMS, CROCO, Peru

1. INTRODUCCIÓN

En el 2017 se instaló un Clúster Computacional de Alto Rendimiento (CCAR) en el Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC) del Instituto del Mar del Perú (IMARPE), el cual fue adquirido dentro del Proyecto "Adaptación al cambio climático del Sector Pesquero y el Ecosistema Marino-Costero del Perú" financiado por el Banco Interamericano de Desarrollo (BID) (GUTIÉRREZ *et al.*, 2019). El clúster denominado "Norte Humboldt" está dedicado a ejecutar varios modelos de simulación para el proceso de escenarios de cambio climático en el mar peruano.

En el Perú, se han realizado diversas investigaciones con el modelo *Regional Ocean Modeling System* (ROMS) para estudiar la variación climatológica (PENVEN, ECHEVIN, PASAPERA, COLAS & TAM, 2005; MONTES, COLAS,

CAPET & SCHNEIDER, 2010), los efectos de El Niño 1997-1998 (COLAS, CAPET, McWILLIAMS & SHCHEPETKIN, 2008), y la variación intraestacional de la circulación (ECHEVIN *et al.*, 2014). También para estudiar los mecanismos que fuerzan la variabilidad de la temperatura superficial del mar (ILLIG *et al.*, 2014), la sensibilidad de la zona de mínimo oxígeno (ZMO) a la circulación tropical (MONTES *et al.* 2014), así como la variabilidad de la ZMO asociada a El Niño Oscilación del Sur (ESPINOZA- MORRIBERON *et al.*, 2019) y para estudiar los impactos físicos y biogeoquímicos en la región bajo escenarios de cambio climático (ECHEVIN *et al.*, 2020).

Por su parte, CROCO se ha utilizado para estudiar, entre otros, el forzamiento y evolución de El Niño Costero 2017 (ECHEVIN *et al.*, 2018), las fluctuaciones de la ZMO acoplándolo al modelo BioEBUS (JOSE, STRAMMA, SCHMIDTKO & OSCHLIES, 2019), filamentos y frentes de submesoescala

¹ IMARPE, DGIOCC, Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC).
jramos@imarpe.gob.pe

(HAUSCHILDT *et al.*, 2021) y la variabilidad de la clorofila-a en el sistema de afloramiento acoplándolo al modelo PISCES (ECHEVIN, HAUSCHILDT, COLAS, THOMSEN & AUMONT, 2021). También se ha usado para estudiar la forma en que la profundidad de la capa de mezcla influye en el funcionamiento del ecosistema de afloramiento costero peruano (XUE, FRENGER, PROWE, JOSE & OSCHLIES, 2022) y en el modelado de extremo a extremo acoplando el modelo CROCO-BioEBUS a un modelo ecosistémico (OSMOSE) (HILL *et al.*, 2022).

En el LMOECC se utilizan de manera rutinaria los modelos ROMS y CROCO, cada uno con su propia configuración de dominio y resolución. Por tal motivo, el objetivo del presente trabajo fue optimizar la ejecución de dichas configuraciones de los dos modelos, determinando el número de núcleos con los que se obtiene la mayor velocidad de procesamiento en el CCAR “Norte Humboldt” del LMOECC-IMARPE. Los resultados obtenidos permitirán administrar los recursos computacionales del CCAR, de manera que se puedan realizar simulaciones óptimas de manera simultánea.

2. MATERIALES Y MÉTODOS

En el LMOECC ambos modelos ROMS y CROCO son utilizados para simular las variables físicas del mar a alta resolución espacial reproduciendo las condiciones del océano frente a la costa del Perú, contribuyendo de esta manera con los diversos estudios que se realizan en el laboratorio.

Modelo ROMS

El modelo *Regional Ocean Modeling System* (ROMS) fue diseñado para realizar simulaciones precisas de sistemas oceánico regionales y ha sido aplicado para la simulación regional de una amplia variedad de regiones diferentes de los océanos del mundo. ROMS resuelve las Ecuaciones Primitivas en un entorno rotatorio centrado en la Tierra, basado en la aproximación Boussinesq y en el balance de momento vertical hidrostático (SHCHEPETKIN & McWILLIAMS, 2005).

La ventaja del modelo ROMS sobre otros modelos es doble; primero desde un punto de vista físico

reproduce procesos alterados por el supuesto de tapa rígida (*rigid-lid*), como movimientos de marea, relaciones de dispersión de las ondas de Rossby, etc., y segundo logra una economía computacional, mediante algoritmos rápidos que incluyen el núcleo no lineal computacional de ROMS, descritos en SHCHEPETKIN y McWILLIAMS (2005).

Modelo CROCO

El modelo *Coastal and Regional Ocean Community* (CROCO) está construido alrededor del núcleo de ROMS (versión AGRIF), con la incorporación de un solucionador no hidrostático y con capacidades de acoplamiento para la atmósfera, las ondas superficiales, la dinámica de sedimentos marinos, la biogeoquímica oceánica y los ecosistemas, es decir es la componente oceánica de un sistema acoplado complejo. CROCO tiene como un objetivo importante el resolver escalas muy finas (especialmente en la zona costera), y sus interacciones con escalas más grandes (CROCO, 2022).

Ejecución de los Modelos

ROMS y CROCO son modelos oceánicos utilizados tanto para la investigación como para el pronóstico operativo. Tanto ROMS como CROCO (CROCO, 2022) necesitan un compilador Fortran95 y pueden ser compilados en modo OMP (*Open Message Passing*), en modo MPI (*Message Passing Interface*) y en modo openMPI, indicando la opción adecuada en el proceso de configuración.

Clúster computacional de alto rendimiento (CCAR)

El CCAR instalado en el LMOECC cuenta con 12 nodos computacionales (servidores Blade Dell). Cada nodo contiene 2 procesadores Intel Xeon E5-2680 de 2,50 GHz, con 12 núcleos por cada procesador, lo que resulta en un total de 288 núcleos. El CCAR tiene una memoria RAM de 128 GB por servidor, lo que hace 1,536 TB de memoria principal, un dispositivo de almacenamiento de 400 TB y sus nodos computacionales están interconectados mediante conectividad *infiniband* FDR de 56 Gbps y alcanza una capacidad teórica pico de ~12 TFLOPS (Fig. 1).



Figura 1.- Clúster computacional de alto rendimiento implementado en el Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC)

Tabla 1.- Número de núcleos y la configuración de los dominios MPI utilizados.

Nro. Núcleos	Parámetros XI y ETA
6	NP_XI=2 NP_ETA=3
12	NP_XI=3 NP_ETA=4
24	NP_XI=3 NP_ETA=8
36	NP_XI=3 NP_ETA=12
48	NP_XI=4 NP_ETA=12
60	NP_XI=4 NP_ETA=15
72	NP_XI=4 NP_ETA=18
84	NP_XI=6 NP_ETA=14
96	NP_XI=6 NP_ETA=16
108	NP_XI=6 NP_ETA=18
120	NP_XI=6 NP_ETA=20

Simulaciones

Se realizaron una serie de simulaciones con ambos modelos para así determinar el tiempo que demora la ejecución de un mes de simulación para cada modelo y lo que varía es el número de núcleos asignados para ejecutar el modelo.

Se procedió a elaborar un *script* en *bash* (intento), para medir con exactitud cuánto tiempo demoraba cada simulación que se iba lanzando. El número de núcleos varió entre 6 y 120, con incremento de 12

núcleos para la simulación siguiente, salvo cuando se ejecutó con 6 cores.

El número de núcleos utilizados, así como la configuración del número de subdominios MPI en las direcciones XI y ETA se aprecian en la Tabla 1.

En ambos modelos se asignó el mismo número de subdominios MPI en las direcciones XI y ETA de acuerdo con el número total de núcleos a utilizar en cada simulación.

Configuración de modelos

El dominio del modelo ROMS se extendió de 82°W a 70°W y de 8°S a 18°S, a una resolución horizontal de 0,037°, con 325 x 279 puntos de grilla y 32 niveles verticales (Fig. 2).

El dominio en el caso del modelo CROCO abarcó entre 96°W - 70°W y entre 5°N - 22°S, con resolución horizontal de 0,1°, con 261 x 277 puntos de grilla y 32 niveles verticales. Se usó la configuración del modelo que utiliza el LMOECC para ejecutar ROMS de modo operativo (Fig. 3).

Los archivos ejecutables de ambos modelos fueron generados utilizando los compiladores *icc* e *ifort* de Intel versión 17.04. La librería *netCDF* con la que el modelo fue enlazado fue una compilación clásica realizada también con los compiladores de Intel *icc*, *icpc*, *ifort* (UCP, 2019) indicando el flag *disable-netcdf-4*.

3. RESULTADOS

Determinación del número óptimo de núcleos

Los resultados obtenidos luego de ejecutar los modelos en el CCAR del LMOECC - IMARPE se pueden apreciar en la Figura 4, en la que se presenta el tiempo de ejecución de 1 mes de simulación en función al número de núcleos activos en el CCAR.

Para el modelo ROMS, se observa que el tiempo de ejecución disminuye inicialmente hasta alcanzar los 24 núcleos, después de este punto el tiempo de ejecución comienza a aumentar, a pesar de que el número de núcleos utilizados en la simulación se esté incrementando. Por lo tanto, el número óptimo de núcleos del modelo ROMS en el CCAR "Norte Humboldt" es de 24 núcleos. La utilización de un mayor número de núcleos solo ralentizaría dicho tiempo de ejecución y se estaría desperdiciando los recursos computacionales del clúster.

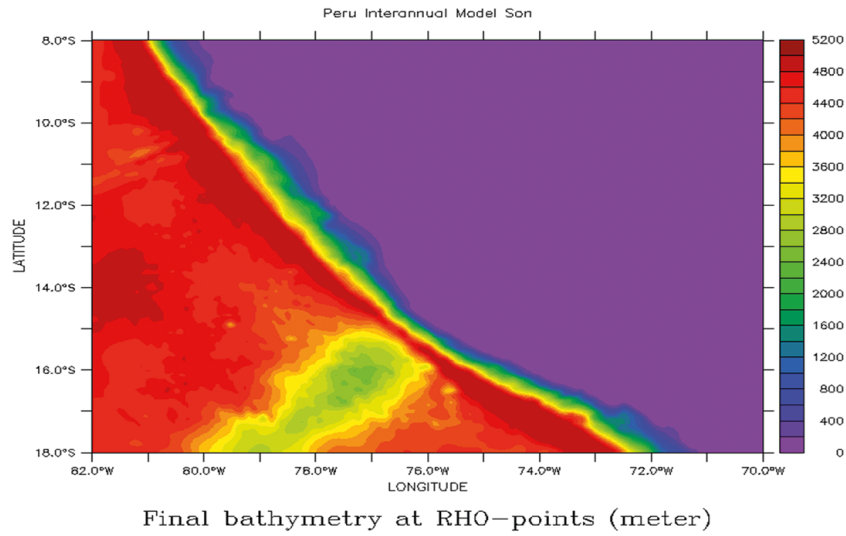


Figura 2.- Dominio del modelo ROMS usado para evaluar la eficiencia del CCAR

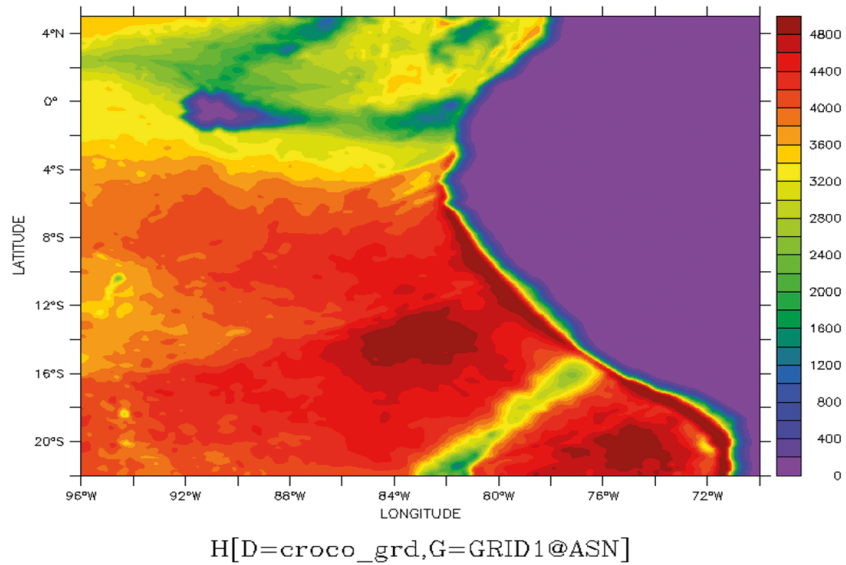


Figura 3.- Dominio del modelo CROCO usado para evaluar la eficiencia del CCAR

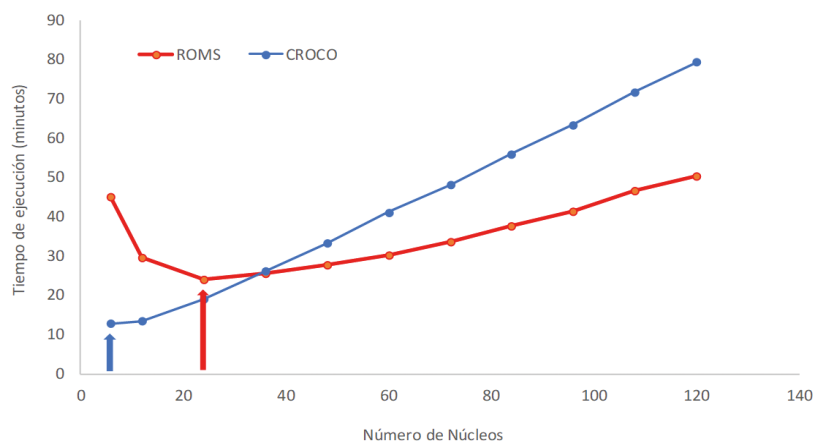


Figura 4.- Tiempo de ejecución por mes de simulación de los modelos ROMS y CROCO en función del número de núcleos utilizado, indicando el número óptimo de núcleos (flechas roja y azul, respectivamente)

Para el modelo CROCO, en la Figura 4, se observa desde el inicio que a mayor cantidad de cores empleados en la simulación, el tiempo también se incrementa. Por lo tanto, el número óptimo de núcleos del modelo CROCO en el CCAR "Norte Humboldt" es 6.

Como la simulación realizada con ROMS tuvo mayor cantidad de puntos de grilla en longitud y latitud ha necesitado mayor cantidad de núcleos para alcanzar su mejor tiempo de ejecución, por lo que también se debería ejecutar esta configuración con el modelo CROCO, para ver si es posible bajar los tiempos de ejecución obtenidos con ROMS.

Existen algunos estudios recientes a diferentes escalas espacio/temporal realizados con el modelo ROMS/CROCO. Mogollón *et al.* (sometido), estudió la estructura vertical de las olas de calor marinas frente al Perú e informó que para la simulación utilizó 48 núcleos, con una grilla de 285 latitudes (LLm0) y 303 longitudes (MMm0), periodo de simulación 27 años, para lo cual necesitó 648 horas, es decir con eficiencia de 120 minutos por mes de simulación. Para estudiar la estratificación de la Bahía de Paracas (Perú), ARELLANO *et al.* (2023), usó 64 núcleos con una grilla de 600 latitudes (LLm0) y 635 longitudes (MMm0), con eficiencia de 102 minutos por cada mes de simulación (datos técnicos por información personal de los autores).

Debido a que en la ejecución de los modelos numéricos ROMS y CROCO no se utilizan todos los núcleos que tiene el CCAR del LMOECC - IMARPE, posibilita que se puedan ejecutar, de manera simultánea, varias simulaciones de di-

chos modelos. Por ejemplo, podrían correrse 4 simulaciones como las que realizó ARELLANO *et al.* (2023) para su estudio, lo que permitiría acortar tiempos de investigación, y varios investigadores pueden ejecutar su simulación sin tener que esperar a que termine la simulación de otro investigador.

4. CONCLUSIONES Y RECOMENDACIONES

El número óptimo de núcleos para realizar simulaciones con el modelo ROMS fue de 24 núcleos usando el Clúster Computacional de Alto Rendimiento "Norte Humboldt" del LMOECC - IMARPE, para un dominio entre 82°W - 70°W y entre 8°S - 18°S y una resolución espacial de 0,037°.

Para el caso del modelo CROCO, el número óptimo de núcleos para realizar las simulaciones fue de 6 núcleos usando el CCAR "Norte Humboldt" del LMOECC - IMARPE, con un dominio 96°W - 70°W y entre 5°N - 22°S y una resolución horizontal de 0,1°.

Se recomienda realizar pruebas que permitan comparar el desempeño de ejecución de los modelos cuando son configurados, para utilizar librerías netcdf compiladas tanto con GNU como con Intel.

Por último, recomendamos que antes de realizar simulaciones con el modelo CROCO-PISCES, se realicen pruebas para determinar el número óptimo de núcleos con el que debería ejecutarse la simulación.

5. REFERENCIAS

- ARELLANO, C., ECHEVIN, V., MERMA-MORA, L., CHAMORRO, A., GUTIERREZ, D., AGUIRRE-VELARDE, A., TAM, J. & COLAS, F. (2023). Circulation and stratifications drivers during the summer season in the upwelling bay of Paracas (Peru): A modelling study. *Continental Shelf Research*, 254, 104923. <https://doi.org/10.1016/j.csr.2022.104923>.
- COLAS, F., CAPET, X., MCWILLIAMS, J.C. & SHCHEPETKIN, A. (2008). 1997-1998 El Niño off Peru: A numerical study. *Progress in Oceanogr.*, 79(2-4), 138-155. <https://doi.org/10.1016/j.pocean.2008.10.015>
- CROCO. (2022). Coastal and Regional Ocean Community model. In line: <https://www.croco-ocean.org/>
- ECHEVIN, V., HAUSCHILDT, J., COLAS, F., THOMSEN, S., & AUMONT, O. (2021). Impact of chlorophyll shading on the Peruvian upwelling system. *Geophysical Research Letters*, 48(19), e2021GL094429. <https://doi.org/10.1029/2021GL094429>
- ECHEVIN, V., COLAS, F., ESPINOZA-MORRIBERON, D., VASQUEZ, L., ANCULLE, T. & GUTIERREZ, D. (2018). Forcings and Evolution of the 2017 Coastal El Niño Off Northern Peru and Ecuador. *Front. Mar. Sci.* 5, 367. doi: 10.3389/fmars.2018.00367
- ECHEVIN, V., AURÉLIE, A., LÉVY, M., GRACO, M., AUMONT, O., PIÉTRI, A. & GARRIC, G. (2014). Intraseasonal variability of nearshore productivity in the Northern Humboldt Current System: The role of coastal trapped waves. *Continental Shelf Research*, (73), 14-30. <http://dx.doi.org/10.1016/j.csr.2013.11.015>
- ECHEVIN, V., GÉVAUDAN, M., ESPINOZA-MORRIBERÓN, D., TAM, J., AUMONT, O., GUTIERREZ, D. & COLAS, F. (2020). Physical and biogeochemical impacts of RCP8.5 scenarios in the Peru upwelling System. *Biogeosciences*, 17(12), 3317-3341. <https://doi.org/10.5194/bg-17-3317-2020>

- ESPIÑOZA-MORRIBERÓN, D., ECHEVIN, V., COLAS, F., TAM, J., GUTIERREZ, D., GRACO, M., LEDESMA, J. & QUISPE-CCALLUARI, C. (2019). Oxygen Variability During ENSO in the Tropical South Eastern Pacific. *Front. Mar. Sci.* 5, 526. doi:10.3389/fmars.2018.00526
- GUTIÉRREZ, D., TAM, J., REGUERO, B. G., RAMOS CASTILLEJOS J., OLIVEROS, R., CHAMORRO, A., GÉVAUDAN, M., ESPINOZA, D., COLAS, F., ECHEVIN, V., CORREA, D., DOMÍNGUEZ, N., ZAVALA, R., GONZALES, N., RAMOS, J., GRADOS, D. & ROMERO, C. Y. (2019). Fortalecimiento del conocimiento actual sobre los impactos del cambio climático en la pesquería peruana. Pp. 38 -71. En: R. Zavala, D. Gutiérrez, R. Morales, A. Grunwaldt, N. Gonzales, J. Tam, C. Rodriguez, S. Bucaram (Eds.). *Avances del Perú en la adaptación al cambio climático del sector pesquero y del ecosistema marino-costero*. Monografía del BID, Serie IDB-MG-679, Lima, 125 p. DOI:10.18235/0001647
- HAUSCHILDT, J., THOMSEN, S., ECHEVIN, V., OSCHLIES, A., JOSE, Y. S., KRAHMANN, G., BRISTOW, L. A. & LAVIK, G. (2021). The fate of upwelled nitrate off Peru shaped by submesoscale filaments and fronts. *Biogeosciences*, 18, 3605–3629. <https://doi.org/10.5194/bg-18-3605-2021>
- HILL-CRUZ, M., FRENGER, I., GETZLAFF, J., KRIEST, I., XUE, T. & YUNNE-JAI, S. (2022) Understanding the drivers of fish variability in an end-to-end model of the Northern Humboldt Current System. *Ecological Modelling*, 472(110097), 16. <https://doi.org/10.1016/j.ecolmodel.2022.110097>
- ILLIG, S., DEWITTE, B., GOUBANOVA, K., CAMBON, G., BOUCHAREL, J., MONETTI, F., ROMERO, C., PURCA, S. & FLORES, R. (2014). Forcing mechanisms of intraseasonal SST variability off central Peru in 2000–2008. *J. Geophys. Res. Oceans*, 119(6), 3548–3573. <https://doi.org/10.1002/2013JC009779>
- JOSE, Y. S., STRAMMA, L., SCHMIDTKO, S. & OSCHLIES, A. (2019). ENSO-driven fluctuations in oxygen supply and vertical extent of oxygen-poor waters in the oxygen minimum zone of the Eastern Tropical South Pacific, *Biogeosciences Discuss.* [preprint]. <https://doi.org/10.5194/bg-2019-155>, in review
- MONTES, I., COLAS, F., CAPET, X. & SCHNEIDER, W. (2010), On the pathways of the equatorial subsurface currents in the eastern equatorial Pacific and their contributions to the Peru-Chile Undercurrent, *J. Geophys. Res.*, 115(C09003). Doi: 10.1029/2009JC005710
- MONTES, I., DEWITTE, B., GUTKNECHT, E., PAULMIER, A., DADOU, I., OSCHLIES, A. & GARÇON, V. (2014). High-resolution modeling of the Eastern Tropical Pacific oxygen minimum zone: Sensitivity to the tropical oceanic circulation. *J. Geophys. Res. Oceans*, 119, 5515–5532. doi:10.1002/2014JC009858
- PENVEN, P., ECHEVIN, V., PASAPERA, J., COLAS, F. & TAM, J. (2005). Average circulation, seasonal cycle, and mesoscale dynamics of the Peru Current System: A modeling approach. *J. Geophys. Res.*, 110(C10021). doi:10.1029/2005JC002945
- SHCHEPETKIN, A. F. & McWILLIAMS, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean modelling*, 9(4), 347–404. <https://doi.org/10.1016/j.ocemod.2004.08.002>
- UCP. (2019). Network Common Data Form (NetCDF). En línea: <https://www.unidata.ucar.edu/software/netcdf/>
- XUE, T., FRENGER, I., PROWE, A. E. F., JOSE, Y. S. & OSCHLIES, A. (2022). Mixed layer depth dominates over upwelling in regulating the seasonality of ecosystem functioning in the Peruvian upwelling system. *Biogeosciences*, 19, 455–475, <https://doi.org/10.5194/bg-19-455-2022>