UNIVERSIDAD POLITECNICA DE VALENCIA

ESCUELA POLITECNICA SUPERIOR DE GANDIA

Master en Eval. y Seg. amb. de Ecos. marin. y cost





"Controles de calidad y valores de referencia para datos de oceanografía química. El caso del Radial Profundo de Canarias."

Tipología: Trabajo de investigación

TRABAJO FINAL DE MASTER

Autor/a: Irene Chamarro León

Tutor/a: Dra. Elena Tel Inst. Español de Oceanografía

Cotutor/a: Dra. Silvia Falco

GANDÍA, 2015





Resumen

Las series temporales de datos de oceanografía química recogidas a partir de muestras analizadas durante las sucesivas campañas que el IEO realiza en el Radial Profundo de Canarias, permiten llevar a cabo un estudio estadístico de la información recogida. La unificación de los criterios de validación y control de calidad y la homogeneización de los formatos de archivo permanente en formatos internacionalmente aceptados posibilita el acceso y la reutilización de los datos a través de las infraestructuras internacionales al uso (SeaDataNet). Este estudio ha permitido la generación de perfiles promedio y valores estadísticos que pueden servir como valores de referencia para la validación de futuros muestreos.

Palabras clave: series temporales, oceanografía química, control calidad, perfiles promedio, valores de referencia.

Abstract

The time series of chemical oceanographic data gathered from samples analyzed during the successive ocean surveys carried out by the Spanish Institute of Oceanography (IEO) in the Deep Section of the Canary Islands have made it possible to conduct a statistical study of the resulting information. The data quality control and permanent archiving format homogenization in internationally adopted ones have made it possible to access and reuse the data across internationally accepted infrastructures (SeaDataNet). This study has made possible to generate average vertical profiles and statistical values that may serve as reference values for future sampling.

Keywords: time series, chemical oceanographic, quality control, average vertical profiles, reference values





Agradecimientos

Deseo expresar mi gratitud al Dr. Pedro J. Vélez-Belchí, investigador principal del programa del Radial Profundo de Canarias por autorizar el uso de la información disponible. También deseo expresar mi gratitud a D. José Escánez, responsable de la toma y análisis de muestras, junto a todo el equipo que forma parte de las campañas RAPROCAN del IEO (Instituto Español de Oceanografía), cuyo trabajo ha permitido la elaboración de este estudio. Por último, deseo expresar mi gratitud a la Dra. Elena Tel, investigadora titular del IEO y tutora externa en este trabajo, por su dedicación y paciencia a largo del desarrollo de este proyecto.





Contenido

Re	sun	nen	ii
Ab	stra	act	ii
Ag	rad	ecimientos	iii
1.		Introducción	1
1	1.1.	Ámbito de estudio	2
2.		Objetivos	4
3.		Material	5
3	3.1.	Plataformas e instrumentación	5
3	3.2.	Datos y cobertura espacio temporal	6
3	3.3.	Software empleado	7
4.		Metodología	9
4	l.1.	Etapa de Muestreo	9
4	l.2.	Homogeneización de los formatos de datos	9
4	l.3.	Control de calidad de los datos	9
	Co	ontrol de duplicados	10
	Co	ontrol de la posición y fecha de las estaciones	10
	Co	ontrol de los datos registrados	12
5.		Resultados y Discusión	14
5	5.1.	Formateado	14
5	5.2.	Control de calidad	14
5	5.3.	Generación de perfiles promedio	16
5	5.4.	Perfiles promedio: zona continental, zona del archipiélago y zona oceánica.	21
5	5.5.	Generación de secciones promedio	24
6.		Conclusiones	28
7		Bibliografía	30





Αl	NEXOS	1
	I. Metodología	1
	Método Winkler	1
	Métodos colorimétricos	1
	Homogeneización de los datos	2
	Perfiles y secciones por campañas sin correcciones	4
	Campaña 022006	4
	Campaña 092006	5
	Campaña 022008	6
	Campaña 022009	8
	Campaña 122011	10
	Campaña 122012	11
	Campaña 102013	12
	III. Secciones promedio sin correcciones RAPROCAN (2006-2013)	13
	IV. Perfiles y secciones por campañas corregidos	15
	Campaña 022006	15
	Campaña 092006	16
	Campaña 022008	17
	Campaña 022009	19
	Campaña 122011	21
	Campaña 122012	. 22
	Campaña 102013	23
	V. Perfiles promedio y secciones corregidas RAPROCAN (2006-2013)	. 24
	VI. Distribución promedio de las variables RAPROCAN (2006-2013) en función d	e la
	zonificación	27





1. Introducción

En 1964, se estableció el Centro Español de Datos Oceanográficos dependiente del Instituto Español de Oceanográfia (IEO), en el marco del Sistema de Centros Nacionales de Datos Oceanográficos (NODCs) favorecidos por la Comisión Oceanográfica Intergubernamental (COI). El Centro de Datos del IEO está integrado en diferentes redes de datos internacionales y participa en distintos proyectos de investigación a nivel nacional y europeo. Un objetivo común de todos ellos es el desarrollo de estándares de formatos, y protocolos de control que faciliten el intercambio e integración de datos, y así se posibilite la reutilización de los mismos y la elaboración de productos de ámbito regional y global.

Actualmente el Centro de Datos del IEO atiende, entre otros, los compromisos europeos relativos a la red europea de observación y datos marinos EMODNET¹ a propuesta de la Comisión Europea y en el marco de la Política Marítima para la Unión Europea, que desea disponer de información que permita una mejor planificación de su política medioambiental. El IEO, en el marco de los objetivos de este proyecto (lote 4: química), está llevando a cabo el control de calidad de los datos oceanográficos de las variables químicas recogidas en aguas nacionales en el marco de sus propios programas de muestreo (RADIALES, VACLAN/RADPROF, RAPROCAN, RADMED) con el objetivo de obtener datos homogéneos y armonizados de variables químico-físicas que permitan la elaboración de mapas de distribución espacio-temporal y curvas de evolución temporal. Estos datos de calidad controlada que presentan un valor añadido al dato original, se incorporan a las infraestructuras europeas existentes (SEADATANET² y EMODNET) que favorecen su acceso controlado y la posibilidad de reutilización de los mismos.

¹ The European Marine Observation and Data Network DGMARE/2010

² En adelante SDN





1.1. Ámbito de estudio

La zona delimitada para el presente estudio forma parte del archipiélago de Canarias, que se ubica en el borde oriental del gran Giro Subtropical. Este se extiende a todo lo ancho de la cuenca atlántica y desde los 45°N hasta las zonas casi ecuatoriales. Toda el área se encuentra bajo la influencia de la corriente de Canarias, que tiene sus orígenes en la rama más oriental de la Corriente de las Azores, la cual circula con rumbo este al sur de estas islas, y en la Corriente de Portugal que circula con rumbo sur a lo largo de la costa portuguesa. La importancia de entender los mecanismos que rigen el comportamiento del giro subtropical radica en el papel clave que este juega como principal mecanismo de transporte de calor y nutrientes en el Atlántico Norte (Pelegrí, et al 2001), así como en la interacción con la atmósfera, a través de su principal modo de variabilidad en la zona: la Oscilación del Atlántico Norte (NAO).

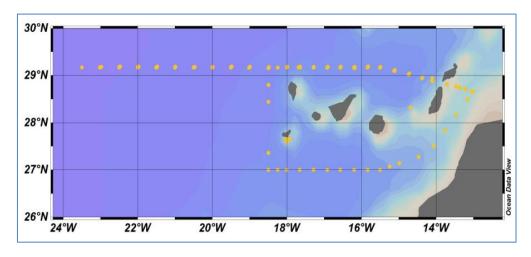


Figura 1.1. Radial Profundo de Canarias.

Actualmente existen muy pocas series oceánicas de larga duración en el Atlántico Norte, que permitan un estudio de los cambios en las masas de agua y de los importantes fenómenos asociados al giro subtropical. Partiendo de esta información, en 2006 el IEO inicia el programa RAPROCAN -Radial Profundo de Canarias- (Figura 1.1). RAPROCAN es parte del sistema de observación del océano del IEO (IEOOS³), y tiene como objetivo entender la variabilidad oceánica en las aguas circundantes a las

-

³ IEO Observing System





Islas Canarias. La componente experimental de este proyecto está compuesta originariamente por dos campañas oceanográficas anuales, una en Febrero y otra en Septiembre que permiten realizar un radial a lo largo del paralelo 29°N, desde el canal de Lanzarote (13°W) hasta el oeste de La Palma (25°W). En ellas se muestrean variables fisicoquímicas a lo largo de toda la columna de agua, desde la superficie hasta el fondo oceánico. En 2013, y a la vista de los resultados de los muestreos anteriores, se modifica esta red de muestreo acortando el radial y cerrando una caja al sur del archipiélago.





2. Objetivos

Las variables que pueden afectar a la fiabilidad de los datos registrados en campañas como la del Radial Profundo de Canarias, pueden verse incrementadas por la dificultad que presenta el medio donde se trabaja, la climatología, los errores e indeterminaciones inherentes a los propios métodos analíticos, la componente informática o incluso el propio factor humano. La escasez de información oceanográfica sobre esta zona incrementa la importancia de comprobar la validez de los datos registrados mediante un control de calidad de los mismos.

Partiendo de la ya mencionada escasez de datos existentes y de la necesidad de cumplir los estándares de calidad que se precisan para incorporar estos datos a las redes europeas, se determina llevar a cabo una evaluación de la calidad de los mismos, aportando así un valor añadido a la serie histórica y definiendo perfiles promedio e intervalos de validez de los datos propios para esta zona que puedan ser útiles en el futuro.

El estudio de los datos de oceanografía química pertenecientes a las campañas RAPROCAN del IEO se ha realizado con los siguientes objetivos:

- Unificar criterios históricos de validación y control de calidad de datos de oceanografía química realizados en el Radial Profundo de Canarias.
- 2. Unificar criterios de nomenclatura y almacenamiento de datos históricos de oceanografía química, de modo que el resultado sean archivos locales cuya información se adapte a los estándares internacionales⁴.
- 3. Favorecer el acceso y la reutilización de los mismos a través de las infraestructuras internacionales al uso (SDN).
- 4. Generar secciones promedio descriptivas de la zona de estudio mediante diversos métodos de interpolación.
- 5. Determinar unos valores estadísticos que simplifiquen la validación de datos futuros, basándose en las características locales de RAPROCAN.

_

⁴ Formatos Medar/Medatlas y ODV





3. Material

3.1. Plataformas e instrumentación

El proceso de muestreo a lo largo del Radial Profundo de Canarias se ha realizado en los buques oceanográficos operados por el IEO, B/O. Cornide de Saavedra y B/O. Ángeles Alvariño (Figuras 3.1 y 3.2). La información sobre las características de los buques puede consultarse en www.ieo.es.





Figura 3.1. B/O Cornide Saavedra.

Figura 3.2. B/O Ángeles Alvariño.

El análisis de la columna de agua se determina midiendo las variables químicas objeto de estudio mediante la utilización de una "roseta oceanográfica", formada por un conjunto de botellas Niskin (24 cilindros estancos) que recogen muestras de agua a distintas profundidades por medio del cierre controlado de las mismas desde el laboratorio de a bordo.

Para determinar analíticamente la salinidad, se emplea un salinómetro de alta precisión AutoSal Guideline (Figura 3.4), que permite obtener la conductividad de la muestra de agua respecto a la conductividad de una muestra conocida (agua estándar).

La cantidad de oxígeno disuelto en las muestras de las botellas esta relacionada con factores como la temperatura y la presión pero también con el origen profundo o superficial y por tanto, con la concentración salina del agua. La determinación de este parámetro se estima en laboratorio a partir del método Winkler (Figura 3.5). El cálculo de los parámetros relativos a los nutrientes (nitratos totales, nitritos, silicatos y fosfatos) se realiza en laboratorio aplicando protocolos específicos de química analítica basados en métodos colorimétricos (cfr. Anexo I).







Figura 3.4. Salinómetro Autosal.



Figura 3.5. Instrumentación para la determinación Oxígeno mediante el Método Winkler.

3.2. Datos y cobertura espacio temporal

La serie de datos disponibles del RAPROCAN está formada por campañas realizadas entre los años 2006 y 2013 (Tabla 3.1) correspondientes en su mayoría a los muestreos de invierno. Consta de seis campañas realizadas con el B/O Cornide de Saavedra y una campaña final con el B/O Ángeles Alvariño, en las que el número de estaciones de muestreo varía, pero no así su posición.

Tabla 3.1. Campañas realizadas y variables estudiadas.

Campaña	Buque Oceanográfico	Nº Estaciones Muestreo	Variables m	edidas
02/2006	Cornide de Saavedra	20	Oxígeno disuelto	
09/2006 Cornide de Saavedra 15		Nutrient	es	
02/2008	Cornide de Saavedra	28	Oxígeno disuelto Nutrientes	
02/2009	Cornide de Saavedra	27	Oxígeno disuelto Nutrientes	
12/2011	Cornide de Saavedra	24	Oxígeno disuelto	
12/2012	Cornide de Saavedra	40	Oxígeno disuelto Salinidad	
12/2013	Ángeles Alvariño	32	Oxígeno disuelto Salinidad	

Para el desarrollo de este trabajo se han acotado las estaciones de muestreo a las incluidas dentro del área seleccionada en la Figura 3.6.





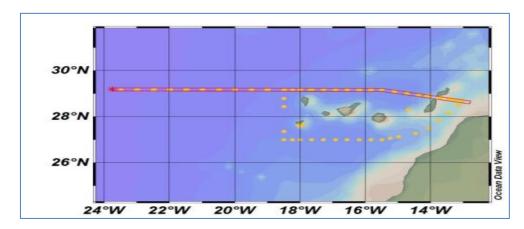


Figura 3.6. Estaciones de muestreo objeto de estudio.

La tabla 3.2 refleja los parámetros medidos para cada variable en las campañas correspondientes.

Tabla 3.2. Variables químicas medidas.

Vari	able	Código	Unidades
Oxígeno	disuelto	DOX1	ml/l
	NO ₂	NTRI	μg.mol/kg
Nutrientes	NO ₂ +NO ₃	NTRZ	μg.mol/kg
Nutrientes	PO ₄ 3-	PHOW	μg.mol/kg
	SiO ₄ ⁴⁻	SLCA	μg.mol/kg
Salinidad		PSAL	P.S.U.

3.3. Software empleado

En 1994, el IEO desarrolla un software, el QCMEDAR, siguiendo los protocolos establecidos en el marco de los proyectos MAST⁵ y MEDAR⁶ en los que los centros de datos oceanográficos del Mediterráneo determinan el almacenamiento de la información en formato Medar/Medatlas, y se establecen protocolos comunes de almacenamiento y control de la calidad basados en recomendaciones de ICES7 y COI. Esto programas se mejoraron mediante el desarrollo de una nueva versión denominada QCDAMAR (García and Nikouline, 2010) donde se incluyeron nuevas herramientas que permiten un mejor análisis y visualización de los parámetros oceanográficos de temperatura, salinidad y bioquímica. Este software es de libre disposición У puede descargarse gratuitamente desde la web

⁷ International Council for the Exploration of the Sea.

7

⁵ European Marine Science & Technology Programme.

⁶ Mediterranean Data Archaeology and Rescue





<u>http://indamar.ieo.es/productos/qcdamar</u>. Para formateado de los datos a los que se refiere este trabajo y el control de calidad primario de los mismos se emplea este programa.

En el marco de los proyectos europeos SDN y SDN-2 que comenzaron en 2004, se desarrolla OceanDataView⁸ (ODV), un programa gratuito que permite un análisis gráfico exhaustivo y más intuitivo, y que permite manejar grandes conjuntos de datos, así como ofrecer diversas salidas gráficas (Schlitzer, 2010). En este momento es una herramienta informática de amplio uso por la comunidad científica internacional. La inspección visual de los datos y el análisis manual de los mismos se ha realizado con esta herramienta.

⁸ Este software ha sido desarrollado en el Instituto Alfred Wegener (Alemania) y se encuentra disponible de forma gratuita en http://odv.awi.de/en/software/download.





4. Metodología

4.1. Etapa de Muestreo

Como se ha mencionado anteriormente, la metodología empleada en la recogida de datos se basa en la utilización de botellas Niskin acopladas en una roseta oceanográfica y un CTD que descienden desde la superficie del mar hasta el fondo oceánico. Las botellas se cierran tomando muestras a profundidades estándar para los posteriores análisis de la salinidad, nutrientes y oxígeno disuelto (Figura 4.1). Los valores registrados con el CTD permiten la elaboración de diagramas TS, que, junto con las secciones verticales de oxígeno, temperatura, salinidad, silicatos, nitratos, nitritos y fosfatos, permiten caracterizar las distintas masas de agua a lo largo de la columna, así como evaluar la calidad del muestreo y su eficiencia.



Figura 4.1. Momento de muestreo por medio de la roseta oceanográfica.

4.2. Homogeneización de los formatos de datos

Los datos registrados en las campañas RAPROCAN, se encuentran almacenados en formato *.btl (profundidades de cierre de botellas) y ficheros Excel (resultados analíticos en laboratorio). Mediante rutinas de MatLab (Matlab & Simulink, 1995) desarrolladas *ad hoc* se obtiene un fichero *.txt cuyo formato cumple con los requisitos previos para su formateo definitivo a Medar/Medatlas utilizando el QCDAMAR (cfr. Anexo I.)

4.3. Control de calidad de los datos

Se sigue un procedimiento basado en los test establecidos en el marco de los proyectos MAST y MEDAR. Se aplica una serie de test al conjunto de los datos para





asignar a cada observación un código, "flag" o marcador de calidad, que nos informa de la validez de esa observación. Los "flags" son imprescindibles para la difusión de los datos porque informan de su fiabilidad, facilitando así la reutilización de los datos por terceros y aportándoles un valor añadido. Además se utilizan los criterios establecidos por SeaDataNet para el procesado de datos en modo diferido (Tabla 4.1), puesto que son coherentes con los criterios establecidos anteriormente y son ampliamente empleados en oceanografía. Esto corresponde con los criterios seguidos en el centro de datos de IEO para la validación sistemática de datos.

Tabla 4.1. Escala de los "flag" o marcadores de calidad.

Table 111. Escale de 105 mag o marcadores de canal					
Código	Significado				
0	Sin control de calidad				
1 Dato bueno					
2	Dato probablemente bueno				
3	Dato erróneo, pero probablemente corregible				
4	Dato erróneo				
5	Dato cambiado				
8	Valor interpolado				
9	Dato perdido				

Inicialmente los datos se someten a un control (QCDAMAR) preliminar básico totalmente automático y posteriormente se aplican una serie de test más específicos de la región de estudio, donde se requiere de la inspección visual para determinar la calidad de las observaciones (cfr. Anexo II y Anexo III). A continuación se describen brevemente los test realizados:

Control de duplicados

• Test de control de duplicados: localiza registros con la misma hora y posición y elimina todos los registros excepto el primero.

Validación manual: De forma manual se guarda un solo registro.

Control de la posición y fecha de las estaciones

 Test regional de la posición: Comprueba que la posición de las observaciones está dentro de los límites del área de trabajo del barco.





- Test de fecha: Verifica que todas las observaciones llevan asociados los datos de hora y día del muestreo y que la hora de muestreo va aumentando entre observaciones consecutivas. Si falla hay que eliminar todo el registro ya que los datos carecerían de una referencia temporal válida.
- Test de la velocidad del barco: La distancia recorrida entre las dos observaciones no puede ser mayor a la permitida por la máxima velocidad que puede alcanzar el barco. Mediante este test se detectan errores de fecha/hora y posición mayoritariamente los registros en los que se sobrepasa la distancia correspondiente a una velocidad de 15kn.
- Test de profundidad de la roseta: Permite el control de la profundidad de descenso de la roseta, de manera que no se supere la profundidad que proporciona la sonda.
- Test de control de la profundidad: A partir de nueve puntos cercanos se interpola la profundidad máxima registrada para las coordenadas de la estación tomando como referencia la información batimétrica de ETOPO5 (Edwards, 1989).
- Test de la trayectoria: mediante este test se verifica que la trayectoria del barco es posible, es decir, que la posición de las observaciones está en el mar y que coincide con la ruta del barco, por tanto requiere inspección visual. Para determinar si las posiciones están en tierra o mar se utiliza la línea de costa de alta resolución GSHHS⁹ (Wessel and Smith, 1996).
- Test de control de la cabecera del perfil: Verificación de la información necesaria en la cabecera del perfil: datos de posicionamiento, fecha y hora del muestreo y la profundidad máxima del perfil.

Validación manual: Corrección del fichero y cambio de los flags que sean necesarios.

٠

⁹ Global Self-consistent Hierarchical High-resolution Shoreline.





Control de los datos registrados

- Test de datos aceptados: Cada valor de las variables medidas debe tener asociado un valor de presión. (PRES)
- Test de valores regionales imposibles: Determina que las observaciones tienen unos valores físicamente posibles. Se eliminan los registros que tienen valores de las variables fuera de los rangos establecidos para la zona de muestreo:

Tabla 4.2. Rangos aceptados para las variables regionales.

Latit	50ºN/30ºN					
Longi	0º/-12ºW					
Salini	38-33 p.s.u.					
	NO ₂	0-10 μg.mol/kg				
Nutriontos	NO ₂ +NO ₃	0-100 μg.mol/kg				
Nutrientes	PO ₄ 3-	0-10 μg.mol/kg				
	SiO ₄ ⁴⁻	0-200 μg.mol/kg				
Oxígeno		10-0 ml/l				

- Test de incremento de la profundidad: Los valores de profundidad de cada estación deben presentar valores en aumento a lo largo del perfil que corresponden de forma aproximada con los valores estándar establecidos para la toma de datos.
- Test control de la temperatura y salinidad: Evalúa la validez de los valores de las distintas variables para la zona. Se realiza mediante la inspección visual determinando los valores máximos y mínimos a partir de la posición y época de las observaciones. Se realiza a partir de valores interpolados medidos en perfiles de referencia a dichas presiones. (Perfiles de referencia MEDATLAS)
- Test de control de perfiles constantes: control de los valores para cada parámetro medido. Si el perfil presenta valores constantes de la variable respecto de la profundidad, se clasifican como erróneos asignándose un valor de flag 4.
- Test de control de las distintas presiones: Cada profundidad debe tener unos valores de presión no constantes a lo largo de la columna de agua. Su análisis se lleva a cabo mediante la comparación con la presión de fondo o con la referencia de ETOPO5.
- Test de control de spikes: Identifica valores muy distintos de las variables medidas en comparación con el valor anterior y el posterior y





los elimina porque no es posible que la diferencia sea real. Este test se define por la siguiente expresión:

$$\frac{(v2-v1)}{(P2-P1)} > borne, dato \ spike$$

 V_1 : valor anterior de la variable V_2 : valor posterior de la variable P_2 : valor posterior de la presión. P_1 : valor anterior de la variable presión. Borne: margen de diferencia aceptado.

- Test de control de inversión de densidad: El algoritmo utilizado para ello es, (dens(n)-dens(n-1), donde dens(n) es la desviación de la densidad potencial al nivel "n" calculada a partir de ecuaciones del estado del agua.
- Test de gradiente: Identifica diferencias demasiado grandes entre observaciones consecutivas que no pueden explicarse por la existencia de un gradiente en la zona. Este test se define mediante la siguiente expresión:

$$vi - \frac{(V_{i+1} + V_{i-1})}{2}$$

Donde:

vi: Valor de la variable para una profundidad "i" V_{i+1} : Valor de la variable para una profundidad "i+1" V_{i-1} : Valor de la variable para una profundidad "i-1"

- Test de comparación de la media: este test compara el valor de cada observación con el de la media de las observaciones de ±3 intervalos de muestreo. Se considera que falla si la diferencia es mayor que dos veces la desviación estándar del mismo período.
- Test de la calidad global del perfil: en la cabecera del perfil aparece un flag de calidad para cada parámetro medido en el perfil. Cada valor del flag es calculado a partir de los correspondientes flags asignados para cada nivel del perfil.

Validación manual: cambio de los flags que no sean correctos





5. Resultados y Discusión

5.1. Formateado

Se emplean rutinas MatLab de elaboración propia para generar un fichero intermedio que sirva de entrada a QCDAMAR. La salida de QCDAMAR proporciona los datos en formato MEDAR/MEDATLAS establecido como formato estándar de las bases de datos del IEO, inicialmente con flags=0.

5.2. Control de calidad

Para establecer la validez de los datos registrados en las campañas resulta imprescindible la aplicación de unos controles de calidad. Los controles de calidad aplicados en este trabajo se realizan siguiendo criterios de validación internacionalmente aceptados. Se obtienen errores de diversa tipología:

Valores de Georreferenciación.

Cada estación de muestreo está debidamente georreferenciada. Estos valores de latitud y longitud se registran *in situ* y manualmente en el documento elaborado durante el proceso de muestreo (estadillos de campaña). En el supuesto de que las coordenadas no queden correctamente registradas el análisis de calidad nos informa del error pero sin modificar o codificar el registro. Por ello, es necesaria una corrección manual que se realiza tomando como referencia el posicionamiento de la estación del muestreo anterior y posterior. En caso necesario, se pueden consultar los estadillos de campañas anteriores.

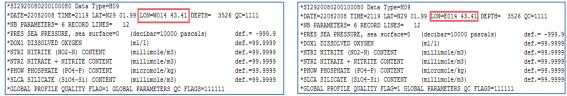


Figura 5.1. Ejemplo error y corrección de valores de georreferenciación.

Con el análisis de calidad de los datos del RAPROCAN se obtiene un volumen muy reducido de errores en los valores de georreferenciación que se asumen como casos excepcionales consecuencia de un error humano y se han corregido en el marco de este trabajo. La representación gráfica de la trayectoria del barco a lo largo del radial ha sido suficiente para comprobar que





en todos los casos detectados la longitud se ha registrado con el valor de orientación contrario (Figura 5.1). Se ha modificado el valor manualmente y se ha repetido la representación gráfica, comprobando así que coincide con la posición de la estación de muestreo. Por último, se realiza de nuevo el control de la calidad de los datos desde el principio de la campaña verificando que se ha corregido el error de posición.

Codificación de los datos.

Se obtienen algunos errores de codificación relacionados con los criterios seguidos en el registro de los valores. En los supuestos en los que la concentración está por debajo de los límites de detección o cuando no existe dato para una profundidad dada del perfil, el registro aparece como *NaN*, *0.01* o *0.02*. Es necesario sustituir estos registros de manera que en el proceso de análisis se asigne un flag "9" (Figura 5.2).

Valor del Registro	Corrección	Codificación		
NaN				
0.01	99.9999	Flag 9		
0.02				

	*SURFACE SAMPLES=						
499							
500	*PRES	NTRI	NTRZ	PHOW	SLCA		
501	15.0	0.1400	0.3300	0.2100	0.5900	11111	
502	25.0	0.0500	0.1600	0.2200	0.6600	11111	
503	50.0	0.1200	0.2200	0.2100	0.7000	11111	
504	75.0	0.0900	0.3100	0.2200	0.7900	11111	
505	100.0	0.1500	1.3400	0.2600	1.6000	11111	
506	125.0	0.0600	2.6800	0.3100	1.1400	11111	
507	150.0	0.1900	2.9100	0.3000	1.0900	11111	
508	200.0	0.0700	5.9400	0.4600	1.8000	11111	
509	250.0	0.0600	7.2000	0.4900	2.1400	11111	
510	500.0	0.1200	14.9400	0.9000	5.1800	11111	
511	600.0	0.6000	19.5300	0.9800	7.0300	11111	
512	700.0	0.0600	20.3300	1.1800	8.7500	11111	
513	800.0	0.0600	21.6700	1.2600	10.0800	11111	
514	1000.0	0.0700	22.0700	1.2600	12.0700	11111	
515	1100.0	0.5700	23.1300	2.4200	11.3600	11141	
516	1200.0	0.0800	15.9200	1.0100	9.1800	11111	
517	1500.0	0.5500	18.8800	1.0400	11.2100	11111	
518	1700.0	99.9999	99.9999	99.9999	99.9999	19999	
519	1750.0	0.0500	19.7800	1.1500	14.9300	11111	
520	2000.0	0.0400	18.1900	1.1200	16.8600	11111	
521	2500.0	0.0400	18.3300	1.2400	22.3800	11111	
522	3000.0	0.5100	22.0400	1.2000	25.7100	11111	
523	3500.0	0.0400	21.4800	1.2800	30.7000	11111	
524	3689.0	0.0400	22.4200	1.2700	31.5800	11111	
525	-999.9	99.9999	99.9999	99.9999	99.9999	99999	

Figura 5.2 Corrección Formato (Campaña 022006).

Detección de spikes.

El test de identificación de *spikes* detecta valores fuera del margen de diferencia aceptado para el rango establecido a cada variable (Figura 5.3), por lo que es necesario verificar cada uno de ellos. Para determinar si se trata de un error o de una medición correcta, se realiza una verificación teniendo en cuenta la zona de ubicación de la estación de muestreo (zona costera, zona del archipiélago o zona oceánica), la profundidad del punto de muestreo (zona superficial o profunda) y las posibles circunstancias dadas para el momento del muestreo (temporales, etc.). En base al resultado, permanece la codificación





del flag como "3" (Figuras 5.4 y 5.5), si se asume que el dato es dudoso o posiblemente erróneo o se codifica asignándole un flag "1" si corresponde a un valor real.

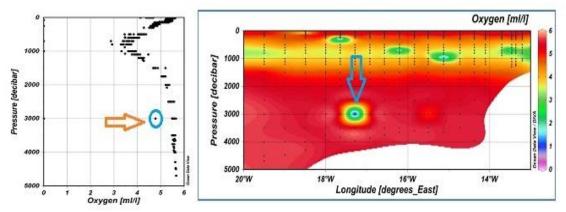


Figura 5.3 Campaña 022006 spike detectado en el oxígeno disuelto a 3000dbar.

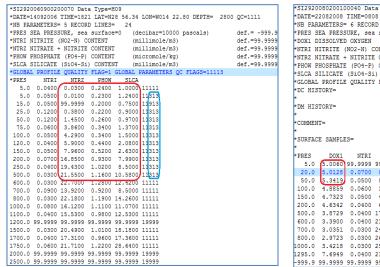


Figura 5.4 Campaña 092006 spikes.

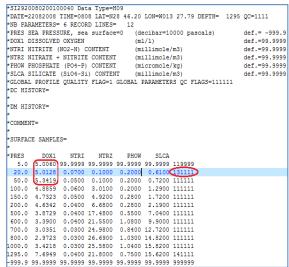


Figura 5.5 Campaña 022008 spikes.

5.3. Generación de perfiles promedio

Una vez aplicado el control de calidad y mediante el ODV, es posible obtener perfiles promedio de cada variable por campaña (cfr. Anexo IV) y para el conjunto de las campañas RAPROCAN (cfr. Anexo V) cuya representación muestra una distribución acorde con las características de la columna de agua en la región de estudio. Cabe destacar que se observa una mayor dispersión en la distribución de los datos de los nitratos totales (Figura 5.6).





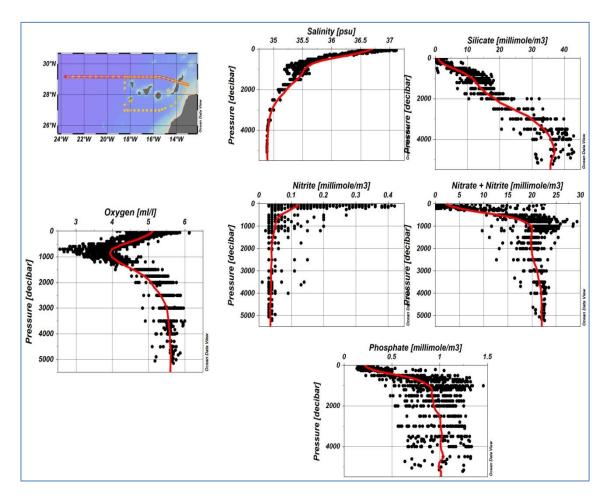


Figura 5.6 Perfiles promedio de las variables y curva promedio RAPROCAN 2006-2013.

Para comprobar la homogeneidad de la distribución, se calculan ciertos parámetros estadísticos calculados para cada perfil mostrando la media, la desviación estándar, el número total de puntos muestreados y se representan los histogramas de distribución para las variables de RAPROCAN (Tabla 5.1).

La distribución de las variables en los perfiles promedio varía en función de la profundidad, no ajustándose a una distribución normal, por lo que los valores estadísticos obtenidos para la media y la desviación estándar, si consideramos todos los perfiles desde superficie hasta fondo, no podríamos asegurar estadísticamente que la relación es representativa. Esto ocurre porque las características de las masas de agua van cambiando a lo largo del perfil (Frank, 2005). Con el objetivo de verificar la distribución normal de las concentraciones de las variables de estudio, se analizan los datos de concentración a una profundidad dada del perfil promedio para cada variable, representando los valores estadísticos (Tabla 5.2) y los histogramas de frecuencias obtenidos.





Tabla 5.1. Parámetros estadísticos de los perfiles promedio variables RAPROCAN 2006-2013.

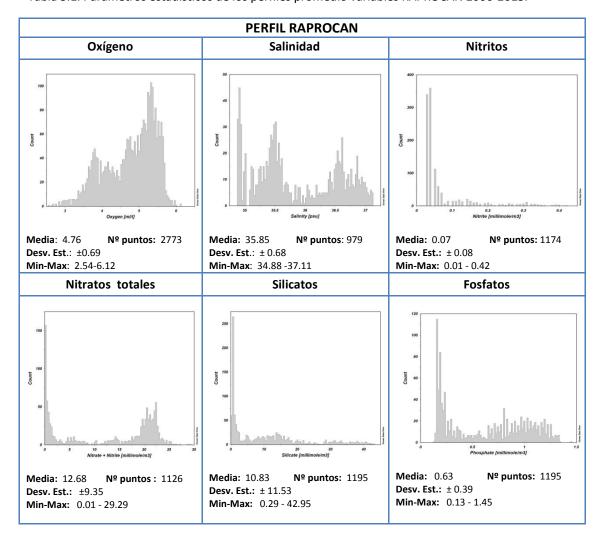


Tabla 5.2. Valores estadísticos de las variables RAPROCAN a Pres=1000dbar

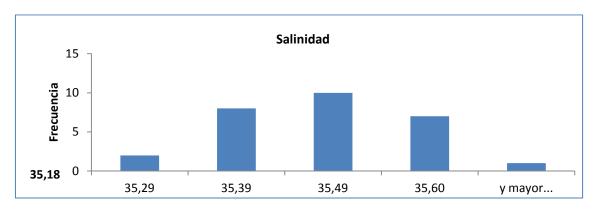
	Oxígeno. Disuelto	Salinidad	Nitritos	Nitratos	Silicatos	Fosfatos
	(ml/l)	(p.s.u.)	(µg.mol/kg)	(µg.mol/kg)	(µg.mol/kg)	(µg.mol/kg)
Promedio	3,70	35,42	0,08	21,77	13,28	0,92
Mediana	3,75	35,43	0,07	22,10	13,50	0,91
Moda	3,81	35,49	0,09	22,49	13,26	0,87
Desviación Est.	0,15	0,10	0,05	2,85	1,74	0,24

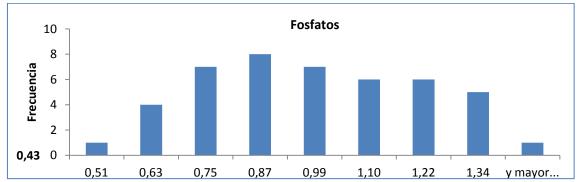
Obtenemos valores pequeños de desviación estándar en relación al valor del promedio de las variables, lo que indica una baja dispersión de los datos. Además, existen relaciones de similitud entre los valores de media, mediana y moda que nos informa del grado de simetría que presenta la distribución de los datos de dicha variable respecto del valor central o mediana del conjunto de datos.





Los histogramas de frecuencias de las variables salinidad, fosfatos y silicatos muestran una distribución casi simétrica respecto de la mediana y el recorrido de los datos (Figura 5.7).





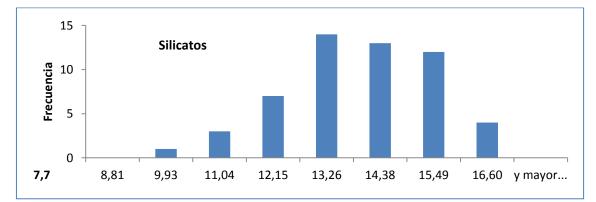
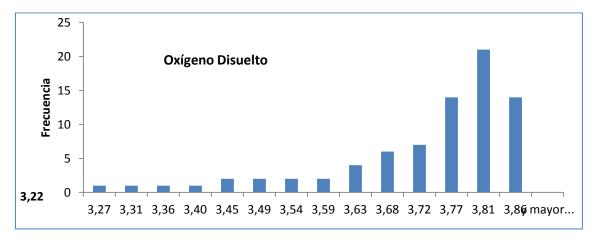


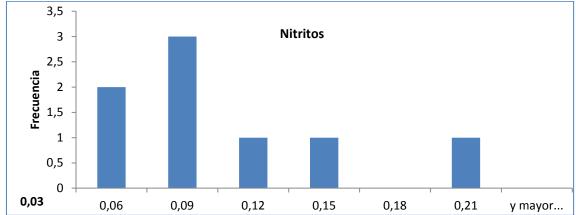
Figura 5.7. Histogramas variables salinidad, fosfatos y silicatos RAPROCAN (2006-2013) Pres=1000dbar.

El resto de las variables, observamos algunos valores extremos en su distribución que afectan directamente a la media (Figura 5.8). Estos valores no alteran el valor central del conjunto de datos pero originan colas en la distribución que se traducen en una asimetría respecto del recorrido del conjunto de datos. Algunos de estos datos extremos se justifican con los datos que han sido identificados en el proceso control de calidad y marcados con flag "3".









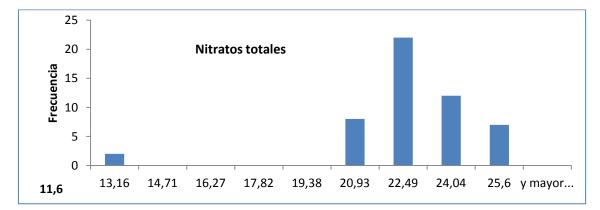


Figura 5.8. Histogramas variables RAPROCAN (2006-2013): oxígeno disuelto, nitritos, nitratos, Pres=1000dbar.

A la vista de los resultados estadísticos y los histogramas de frecuencia, podemos afirmar que los perfiles de distribución promedio obtenidos a partir de los datos validados, siguen una distribución homogénea a una profundidad dada. Además, la escasa disponibilidad de valores de referencia previos refuerza la importancia de estos resultados, ya que supone que el valor promedio para cada perfil de las variables puede tomarse como referencia para validaciones en campañas posteriores.





5.4. Perfiles promedio: zona continental, zona del archipiélago v zona oceánica.

EL radial RAPROCAN se extiende desde el canal de Lanzarote (13°W) hasta el oeste de La Palma (25°W). Dada las distribuciones que presentan los perfiles y las distintas características del radial a medida que nos desplazamos en longitud se ha considerado relevante zonificar el radial en 3 grandes grupos en función de la longitud y observar así si existen diferencias relevantes en la distribución de las variables (cfr. Anexo VI).

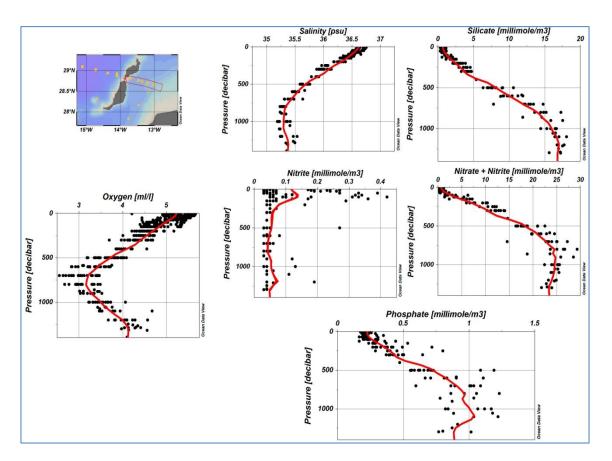


Figura 5.9. Perfiles promedio de la zona continental.

Por su cercanía al extremo occidente del continente africano, se engloban las estaciones de muestreo ubicadas desde la costa y hasta el canal de Lanzarote, denominando esta selección como "zona continental". Se representan los perfiles promedio para las variables y se representa la curva de distribución (Figura 5.9).





Todas las estaciones situadas cerca de las islas, forman lo que se ha denominado en este trabajo como "zona del archipiélago" y sus distribuciones promedio quedan reflejadas en la Figura 5.10.

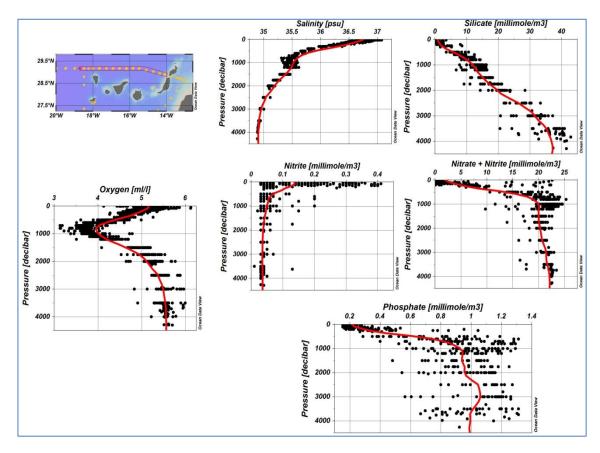


Figura 5.10. Perfiles promedio de la zona del archipiélago.

Por último, se agrupan las estaciones cuya ubicación pertenece a aguas abiertas del Océano Atlántico formando la "zona oceánica" (Figura 5.11).





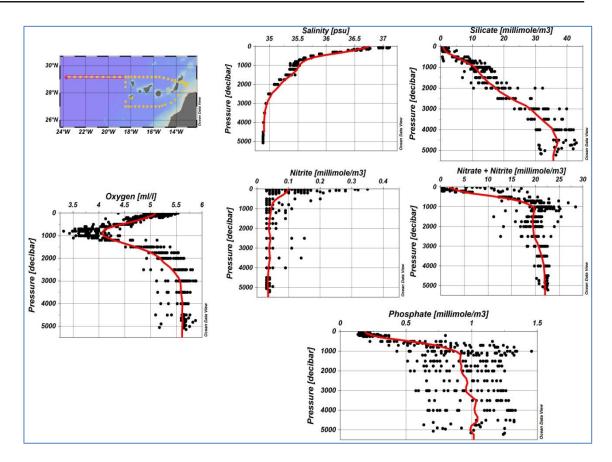


Figura 5.11. Perfiles promedio de la zona oceánica.

La representación simultánea de las curvas de distribución de cada variable para las distintas zonas en las que se ha dividido el perfil, facilita la interpretación de los patrones de distribución (Figura 5.12). A partir de los resultados obtenidos se observa una mayor variabilidad de los datos en la zona continental. La batimetría, la cercanía de la costa y los afloramientos asociados al borde continental, influyen en las concentraciones de las variables en la zona continental. Las curvas de la zona del archipiélago y la zona oceánica presentan una similitud muy marcada para la distribución de las variables en la columna de agua a lo largo del perfil. En el anexo VI se presentan numéricamente los datos promedio correspondientes a estas figuras con el fin de que puedan ser consultados como valores de referencia en futuras campañas.





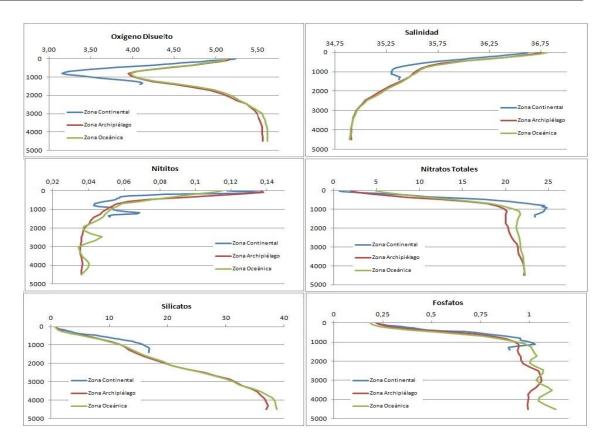


Figura 5.12. Perfiles promedio elaborados para las distintas zonas.

5.5. Generación de secciones promedio

La representación de los datos validados como secciones promedio para cada variable, nos muestra la variación de la concentración de la variable correspondiente tanto a lo largo del radial como en profundidad. Puesto que ODV permite diversos tipos de interpolación, se ha realizado una pequeña comparativa y un posterior análisis de los resultados.

Mediante la interpolación *quick gridding* se obtiene una malla regular equidistante y mediante un algoritmo de promedio ponderado optimizado en todos los nodos de la malla. Se obtiene en pocos segundos una representación objetiva de los registros que puede apreciarse en la Figura 5.6 para el promedio de cada variable del conjunto de campañas RAPROCAN. Esta interpolación está recomendada en casos en los que el volumen de datos es muy superior al caso del Radial de Canarias. Sin embargo, esta visualización permite identificar con facilidad las zonas de la sección cuyos perfiles presentan puntos de muestreo carentes de registro, bien porque no se tomó la





muestra, bien porque los valores se han codificado como no válidos mediante el flag "4" o como flag "9", según la tabla de criterios de codificación aplicada (Figura 5.12).

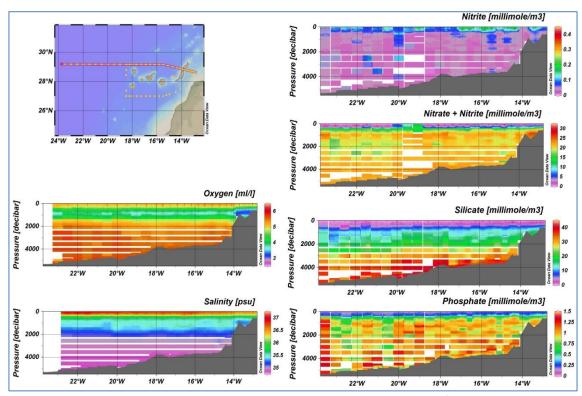


Figura 5.12. Secciones de las variables RAPROCAN 2006-2013. Interpolación Quick gridding.

A diferencia que en el caso anterior, en las interpolaciones *Weighted-avarage gridding* y *Data-Interpolating Variational Analysis (DIVA)* el promedio ponderado se calcula mediante una malla rectangular de resolución variable, de forma que en la malla se producen variaciones en la distancia a lo largo de los ejes "X" e "Y" en función de la densidad de datos existentes. Para áreas con densidades elevadas de registros se obtiene una malla de distancias pequeñas entre los nodos que se traduce en una elevada resolución. En la Figura 5.13 se observan como resultado de la interpolación *weigthted-avarage*, secciones promedio cuya representación muestra una malla más completa que en el caso anterior (Figura 5.12). El mayor número de registros pertenecientes a la columna superior de agua justifica una mayor resolución en esta zona. En zonas profundas la resolución es más baja pero se aprecian cambios que mejoran la visualización de las secciones permitiendo una mejor interpretación de la distribución de las variables a lo largo del radial.





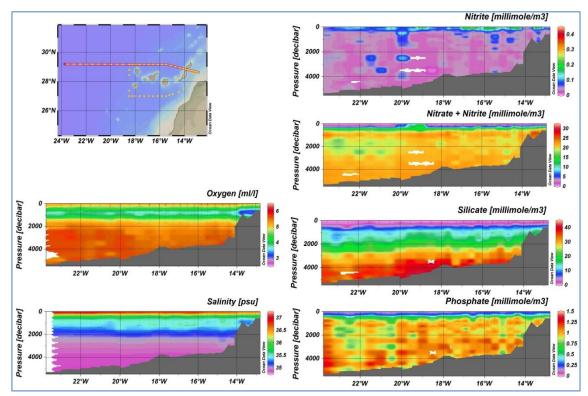


Figura 5.13. Secciones variables RAPROCAN 2006-2013. Interpolación Weighted-avarage gridding.

Mediante la interpolación DIVA los datos son analizados e interpolados buscando una malla óptima donde se tienen en cuenta las líneas costeras y las características de batimetría para estructurar y subdividir el dominio en el que se realiza la estimación. Los cálculos se realizan sobre una malla de elementos finitos adaptado a los dominios específicos para la generación de la malla. En este caso se obtienen secciones completas (Figura 5.14) ya que se estiman valores para las zonas donde inicialmente no hay registros. Con la representación de la sección promedio para cada variable mediante esta interpolación, se observa un patrón de distribución muy estratificado en función de la profundidad, que se mantiene a lo largo del radial.



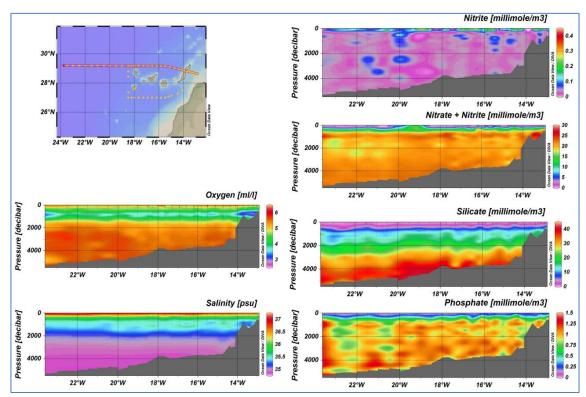


Figura 5.14. Secciones de las variables RAPROCAN 2006-2013. Interpolación DIVA.

A partir de los resultados obtenidos con las tres interpolaciones, se pone de manifiesto que la aplicación de DIVA permite obtener una malla completa del perfil minimizando errores asociados a falta de datos y mejorando la visualización de la distribución de las variables consideradas. Así se facilita el análisis e interpretación de los datos objeto de estudio.





6. Conclusiones.

Este trabajo se inicia a partir de los datos de oceanografía química registrados en las sucesivas campañas realizadas en el RAPROCAN, entre los años 2006-2013 con el objetivo de aplicar un control de calidad unificado de todos los datos disponibles y establecer valores de referencia de las variables en la zona de estudio. Tras la aplicación de la metodología propuesta para hacerlo, se establecen las siguientes conclusiones.

- Los criterios en la nomenclatura empleada en el registro de las variables y en el formato de registro de los datos se han ido modificando a lo largo del tiempo de modo que ha sido necesaria una labor de unificación. En este momento los datos siguen un criterio único de nomenclatura que está en concordancia con los vocabularios y unidades de medida seguidos internacionalmente. Estos vocabularios están accesibles libremente vía web¹⁰.
- Los ficheros definitivos en formato Medar/Medatlas-SDN obtenidos en este trabajo cumplen con los estándares admitidos internacionalmente para su almacenamiento y difusión, por lo que han sido incluidos en la red paneuropea de gestión de datos marinos y oceánicos (SDN) desde el centro de datos del IEO.
- Los perfiles promedio obtenidos presentan una distribución acorde con los patrones conocidos de distribución para cada variable a lo largo de la columna de agua.
- Se recomienda considerar siempre distintos métodos de interpolación para obtener los mapas de distribución de las variables, minimizando errores asociados a la escasez de datos. La aplicación del método DIVA para la representación gráfica y posterior análisis de los datos ha arrojado buenos resultados en este caso.

¹⁰ Vocabularios SeaDataNet P01 y P02 (http://seadatanet.maris2.nl/v_bodc_vocab_v2/welcome.asp.)

_





- En los valores procedentes de la determinación analítica de los nitratos totales se observa una dispersión importante en los datos iniciales. La escasa concentración de los mismos en el agua de mar, la complejidad y limitaciones de la técnica analítica utilizada y problemas asociados a la conservación de las muestras hasta proceder a su determinación analítica en laboratorio, pueden ser responsables de dicha dispersión.
- Los perfiles promedio obtenidos en este trabajo pueden establecerse como perfiles de referencia para utilizarse en campañas futuras de cara a validaciones locales y controles de calidad de datos.

La incorporación sistemática de datos del radial que se realicen en futuras campañas permitirá ir mejorando los perfiles promedios y rellenando las lagunas de información, con lo que se podrán generar secciones y perfiles promedio con mayor significación estadística. Estos datos además son de gran interés para el conocimiento de los ciclos climáticos y su variabilidad, especialmente de cara al estudio del cambio climático y su mitigación.





7. Bibliografía

- CARRITT, D. E. CARPENTER, J. H. (1966). "Comparison and evaluation of currently employed modifications of the Winkler method for determining dissolved oxygen in seawater" a NASCO report. J. Marine Res., 24: 286-318.
- EDWARDS, M. O. (1989). "Global gridded elevation and bathymetry on 5-minute geographic grid (ETOPO5)". NOAA, National Geophysical Data Center, Boulder, Colorado, USA.
- FRANK, J. (2005) "Chemical Oceanography", Third Edition. CRC.
- GARCIA, M^a, NIKOULINE, A. (2010) "User's Guide: Quality Control of Oceanographic Data". Madrid: Instituto Español Oceanografía.
- HANSEN, H. P., KOROLEFF, F. (1999). "Determination of nutrients in Methods of Seawater Analysis", Third Edition (eds K. Grasshoff, K. Kremling & M. Ehrhardt), Wiley-VCH Verlag GmbH, Weinheim, Germany. doi:10.1002/9783527613984.ch10
- PELEGRÍ, J. L., HERNÁNDEZ GUERRA, A., LÓPEZ LAATZEN, F., MACHÍN, F., de ARMAS, D. (2001). "Water masses, circulation and transport in the eastern boundary current of the North Atlantic subtropical gyre" en Sciencia Marina. SCI MAR., 65 (Suppl. 1):177-186.
- MURPHY I, RILEY J. P. (1962) "A single solution method for the determination of soluble phosphate in sea water". J.Ma,.Biol.,4sxn.UK37:9-14.t958
- SCHLITZER, R. (2010). "ODV User's Guide". Bremerhaven, Germany. Alfred Wegener Institute.
- MATLAB & SIMULINK (1995). "User's Guide, The MathWorks", Inc., Massachusetts.





- WESSEL, SMITH, (1996), "Global Self-consistent, Hierarchical, High-resolution Shoreline Database"
- WINKLER, L. S. (1888). "The determination of dissolved oxygen". Ber. Dtsche.
 Chem. Ges. 21: 2843-2855
- INSTITUTO ESPAÑOL OCEANOGRAFÍA (IEO). Campañas oceanográficas http://: www.oceanografia.es/raprocan [Consulta: 10 Noviembre de 2014]