



## **Progetto di Interesse strategico NEXTDATA**

Rendicontazione scientifica per il primo anno di attività: **01/01/2012-31/12/2012**

### **WP 2.2 - Archivio delle reti osservative marine e ricostruzioni climatiche (Resp. Claudia Fratianni, INGV)**

Partners: INGV

#### **1. Attività prevista e risultati attesi**

L'attività prevista nel primo anno del progetto riguarda lo sviluppo, la realizzazione e l'organizzazione degli archivi di dati storici marini, sottoposti ad uno specifico controllo di qualità per l'assimilazione dati, e di dati atmosferici che verranno utilizzati nella produzione delle "Ricostruzioni-Rianalisi (RR)" del Mar Mediterraneo per i passati 100 anni. Verrà inoltre disegnato il portale specifico e la struttura dell'archivio delle RR, utilizzando la rianalisi già esistente per gli ultimi venti anni. Si organizzeranno incontri con comunità di potenziali utenti per definire le specifiche dell'archivio e dei protocolli di accesso ai dati.

Il milestone definito per il primo anno è il seguente:

**M2.2.1 (PM12):** Incontro per la discussione del Portale specifico delle RR e prodotti specifici. L'incontro si è svolto a Dicembre presso l'INGV di Bologna ed ha visto la partecipazione della CLU s.r.l. che si occuperà nel futuro di strutturare il portale specifico.

#### **2. Deliverables previsti per il periodo di riferimento**

Alla fine del primo anno è previsto il seguente Deliverable:

**D2.2.1 (PM12):** Relazione sull'archivio dati osservativi in situ e da satellite specifici per la Ricostruzione-Rianalisi e i dati atmosferici.

**D2.2.2 (PM12):** Relazione sul disegno del catalogo user-friendly e del database.

### 3. Attività effettivamente svolta durante il periodo di riferimento

Nel corso del primo anno, si è svolta gran parte l'attività prevista. In particolare si è proceduto alla creazione della banca dati di riferimento per i dati osservativi marini e per i dati atmosferici necessari per la produzione delle RR, partendo da ciò che si era già collezionato per la produzione di precedenti versioni di rianalisi, come mostrato in Tabella 1, e procedendo nella ricerca di dataset nuovi e aggiornati per coprire l'intervallo temporale delle RR.

#### 3.1 Attività di ricerca

L'attività è stata svolta in stretta collaborazione con il WP1.3 che ha messo a punto, ove necessario, specifici controlli di qualità del dato, come descritto nel D1.3.1 e nel D1.3.2. Questo ha permesso la creazione di una banca dati di riferimento per la produzione delle RR sufficientemente ampia che contribuirà alla realizzazione di prodotti di qualità crescente.

	MedReanV2 (1985-2007)	MedReanV4(1985-2010)	RR(1912-2011)
Initial conditions	MedAtlas Climatology (obs 1995-1999) (Maillard et al. 2005)	SDNV2aa climatology (obs until1987)	SDNV2aa climatology (obs until1987)
Atmospheric forcings	ERA15 1.125° (1985-1992) ECMWF analyses 0.5° (1993-2007)	ERAInterim 0.75° (1985-2010)	AMIP1.125° (1912-1957) ERA40 1.125° (1958-1978) ERAInterim 0.75° (1979-2011)
Total Cloud Cover	NCEP-NCAR (1985-92) ECMWF analyses 0.5° (1993-2007)	ERAinterim 0.75° (1985-2010)	AMIP 1.125° (1912-1947) NCEP-NCAR(1948-1978) ERAInterim 0.75°(1979-2011)
Precipitations	NCEP-NCAR (monthly climatology)	CMAP (monthly climatology)	AMIP (monthly climatology) NCEP-NCAR (monthly climatology) CMAP (monthly climatology)
SST	SST reconstruction (1985-2007) (Marullo et al.2007)	SST reconstruction (1985-2007) (Marullo et al.2007) DT data (2008-2010)	HadISST (1912-1985) SST reconstruction (1985-2007) (Marullo et al.2007) MyOcean data (2008-2011)
SLA	ERS1, ERS2, EnviSat, TOPEX/Poseidon, Jason1 (Pujol and Larnicol 2005)	AVISO multisensor "UPD" data reprocessed in 2010	MyOcean multisensor "UPD" data reprocessed in 2012
XBT	<1999 MEDATLAS, MATER Ship of Opportunity (Manzella 2007) 2000-2007 MFS observations	SeaDataNet MFS observations (1985-2010)	SeaDataNet MFS - MyOcean observations (1912-2011)
ARGO	MedArgo Program 2001-2007 (Poulain et al. 2007)	All observations 2001-2010	All observations 2001-2011
CTD	MEDATLAS, MATER (<1999) 2000-2007 MFS observations (2000-2007)	SeaDataNet MFS observations (1985-2010)	SeaDataNet MFS - MyOcean observations (1912-2011)

**Tabella 1.** Tabella dei dataset utilizzati nella produzione delle ricostruzioni e delle rianalisi presso INGV. I sistemi sono descritti nel Deliverable D1.3.1.

I dataset archiviati per la produzione delle RR sono i seguenti:

- osservazioni di temperatura alla superficie del mare (SST) da satellite;
- osservazioni di anomalia del livello del mare (SLA) da satellite;
- osservazioni in situ di temperatura e di salinità;
- variabili atmosferiche di superficie;
- condizioni iniziali di temperatura e salinità.

### ***Osservazioni SST***

La creazione del dataset di osservazioni di SST ha richiesto un lavoro di censimento dei dataset disponibili per il periodo precedente il 1985. A partire da questa data infatti esistono prodotti specifici creati per il Mar Mediterraneo consistenti in serie temporali giornaliere di mappe di temperatura media alla superficie del mare ottenute tramite tecniche di interpolazione ottimale dei dati su griglia regolare a 1/16 di grado. Il dataset a partire dal 1985 è il risultato della concatenazione dei seguenti prodotti:

- 1985-20080710: ricostruzioni di SST avvalendosi delle più recenti serie temporali di AVHRR Pathfinder SST (Marullo et al., 2007).
- 20080711-20111231: mappe di temperatura media giornaliera alla superficie del mare prodotte e rilasciate dall'SST-TAC/OSI-TAC (Sea Surface Temperature/Ocean Ocean and Sea Ice Thematic Assembly Center) nell'ambito del progetto MyOcean, come descritto nel D2.2.1.

Per il periodo precedente al 1985, sono stati presi in considerazione diversi dataset tra i quali è stato selezionato il dataset HadISST1 (Met Office Hadley Centre's sea ice and sea surface temperature (SST) data set). Si tratta di mappe mensili globali a una risoluzione spaziale di 1° x 1° a partire dal 1870. (Rayner et al., 2003).

### ***Osservazioni SLA da satellite***

Si è proceduto all'aggiornamento del dataset già esistente estendendo la serie temporale fino all'aprile 2012. Questi dati sono prodotti e forniti dal Sea Level TAC nell'ambito del progetto MyOcean.

Si tratta di anomalie dell'altezza della superficie del mare lungo le tracce di ogni singolo satellite disponibile, calcolate come differenza tra l'altezza della superficie del mare e una media sul periodo 1993-1999.

(<http://catalogue.myocean.eu.org/static/resources/myocean/pum/MY02-SL-PUM-008-001-005-v3.2.pdf>).

La copertura temporale dipende dalla durata della missione del singolo satellite:

- Jason 2: da Ottobre 2008
- Jason 1 (nuova orbita): da Febbraio 2009
- Jason 1: Aprile 2002 - Ottobre 2008
- Envisat (nuova orbita): da Ottobre 2010
- Envisat: Ottobre 2002 - Ottobre 2010
- ERS-1: Ottobre 1992 - Maggio 1995
- ERS-2: Maggio 1995 - Aprile 2003
- GFO: Gennaio 2000 - Settembre 2008
- T/P (nuova orbita): Settembre 2002 - Ottobre 2005
- T/P: Settembre 1992 - Aprile 2002

Ciascun satellite è caratterizzato da un ciclo di ripetizione (“repeat cycle”) e da un intervallo tra una traccia e l’altra (“ground track”) che nel complesso garantiscono una copertura abbastanza uniforme del Mediterraneo.

Le informazioni archiviate per ciascun satellite sono quelle richieste dal sistema di assimilazione, ovvero: tempo espresso in giorni giuliani a partire dal 1 gennaio 1950, longitudine, latitudine, valore di anomalia del livello del mare espresso in metri.

### ***Osservazioni in situ di profili verticali di temperatura e di salinità***

Il dataset è il frutto di un lavoro di omogeneizzazione tra i dati collezionati in diversi progetti quali MedarMedatlas, SeaDataNet, MFS e MyOcean al fine di eliminare la presenza di osservazioni duplicate, come descritto nel D1.3.1 e nel D2.2.1. Si tratta di profili verticali di temperatura e salinità collezionati tramite bottiglie, termometri, XBT, MBT, CTD e ARGO sottoposti ad un controllo di qualità specifico per il loro utilizzo nel sistema di assimilazione dati, come descritto nel D1.3.1. La costituzione del dataset di riferimento per le osservazioni in situ è descritto nel D2.2.1.

### ***Variabili atmosferiche di superficie***

I parametri necessari per forzare il sistema di produzione delle RR sono i seguenti:

1. Pressione media a livello del mare (MSLP)
2. Temperatura a 2 metri (T2m)
3. Componente zonale e meridionale del vento (U10,V10)
4. Copertura totale nuvolosa (TCC)
5. Temperatura a 2 metri del punto di rugiada (D2m)
6. Precipitazione (P)

La creazione del dataset di riferimento richiede la concatenazione di differenti dataset di partenza al fine di prendere in considerazione i suddetti parametri e assicurare la copertura temporale delle RR, come indicato in Tabella 1. Dai dataset di riferimento sono stati estratti i parametri (MSLP,U,V,TCC, T2m,D2m) ed archiviati in files mensili o giornalieri in formato binario che verranno letti dal modello oceanico e che conterranno le suddette variabili alla risoluzione spaziale e temporale del modello originale. La precipitazione invece è stata archiviata in files in formato NetCDF contenenti i valori mensili di precipitazione interpolati dalla griglia originale alla griglia del modello oceanico ad 1/16 di grado.

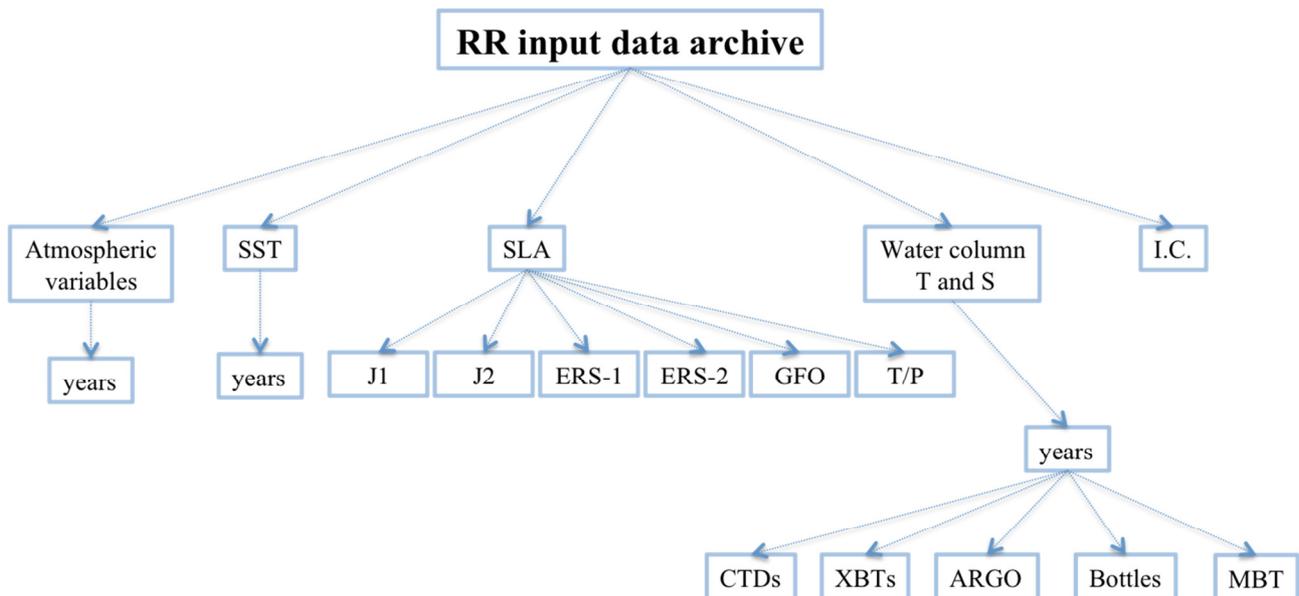
I datasets specifici originali (AMIP, NCEP-NCAR, ERA-40, ERAINTERIM, CMAP) sono descritti nel D2.2.1.

### ***Condizioni iniziali***

Le condizioni iniziali sono archiviate sotto forma di due files distinti in formato netCDF, uno per la salinità ed uno per la temperatura. La metodologia utilizzata per la loro creazione è descritta nel D1.3.1.

## **3.2 Sviluppi applicativi, tecnologici e informatici**

I dati una volta controllati vengono archiviati su filesystem dedicato come mostrato Figura 1.

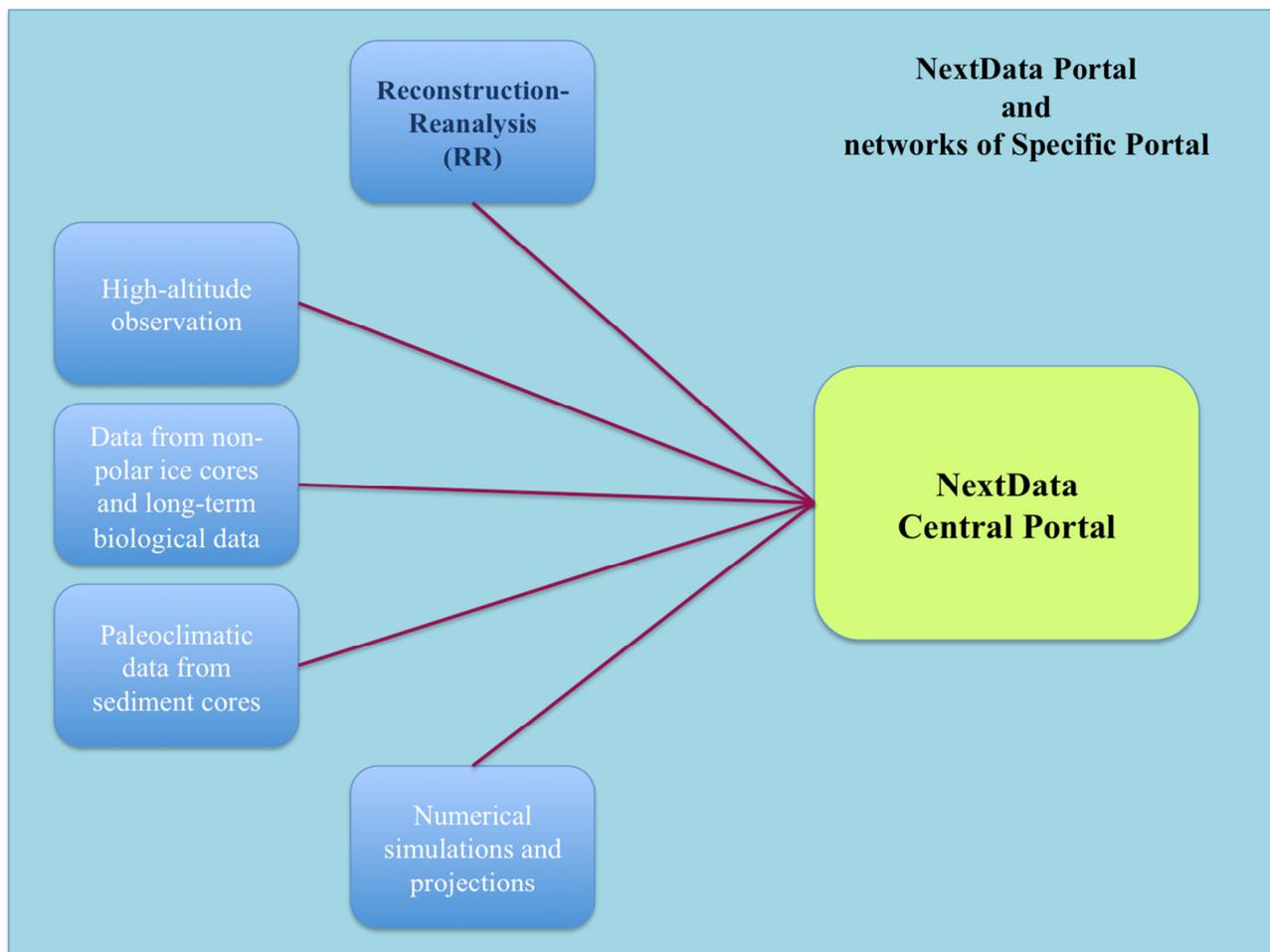


**Figura 1.** Organizzazione delle directories dell'archivio dei dati in input al sistema di RR.

Per ogni dataset si è proceduto ad una archiviazione per anno e/o per strumento. Il formato dei dati è quello richiesto dallo schema di assimilazione OceanVar e dal modello numerico utilizzato nelle RR ed è il seguente:

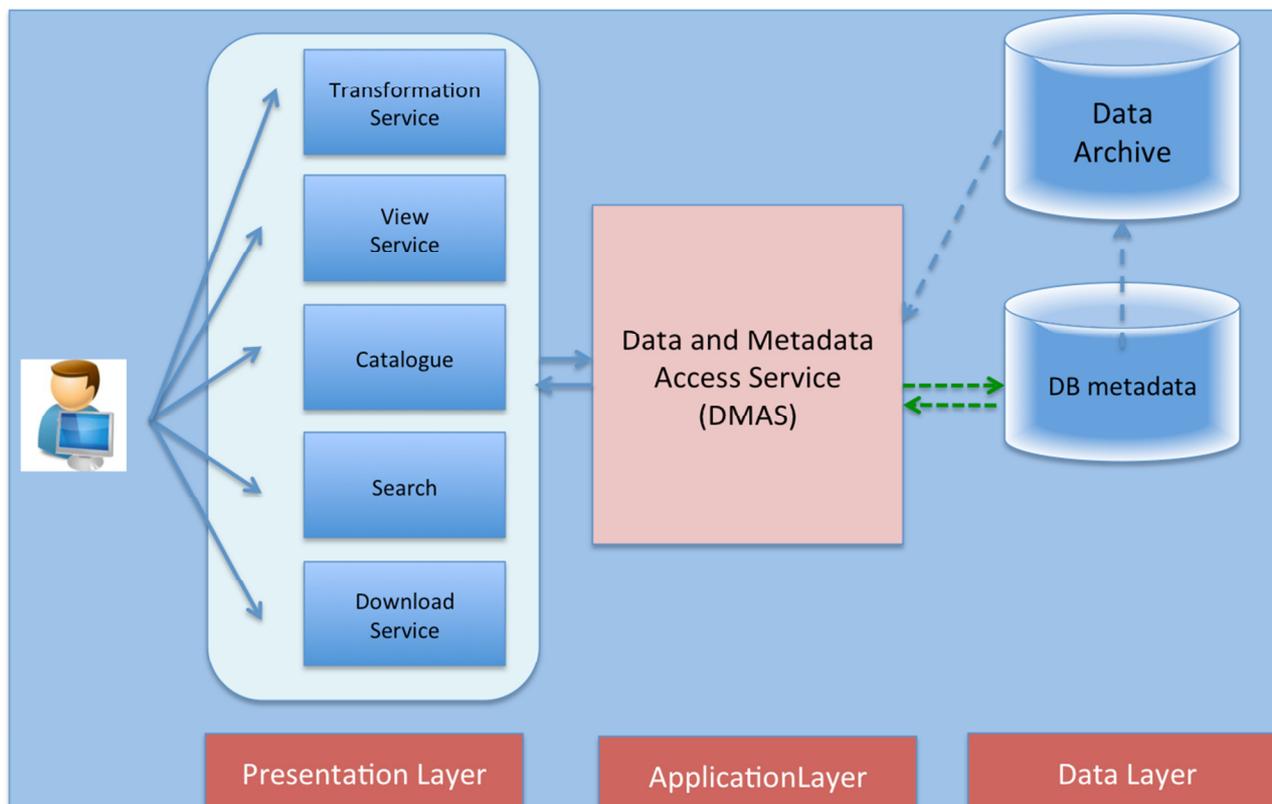
- ASCII: profili verticali di temperatura e salinità; osservazioni di SLA da satellite;
- binario: variabili atmosferiche di superficie;
- netCDF: osservazioni di SST da satellite; condizioni iniziali di temperatura e di salinità.

Parallelemente al lavoro di realizzazione del database di riferimento, è stato disegnato anche il portale specifico per l'accesso ai dati di RR. Nella realizzazione del portale specifico delle RR sono state seguite le più recenti indicazioni della direttiva europea INSPIRE al fine di creare una infrastruttura che sia in grado di integrarsi in un contesto comunitario. Il portale specifico sarà collegato al portale generale del progetto che conterrà tutte le informazioni relative ai metadata dei prodotti che verranno messi a disposizione mentre la proprietà e i servizi di network risiederanno sul portale specifico locale, come mostrato in Figura 2.



**Figura 2.** Schema generale del flusso di informazioni tra il Portale Generale e il Portale Specifico delle RR.

Il portale specifico fornirà i servizi che sono definiti nella direttiva INSPIRE: discovery, view, downloading e transformation, come illustrato in Figura 3.



**Figura 3.** Architettura del portale specifico delle RR.

L'architettura del portale specifico delle RR è costituito da 3 layers strettamente connessi tra di loro:

- **Presentation Layer:** su di esso risiedono i cosiddetti network services (Transformation, View, Search, Download, Catalogue) che si e' deciso di implementare.
- **Application Layer:** comprende il Data e il Metadata Access Service (DMAS).
- **Data Layer:** comprende l'archivio fisico dei dati e il database dei metadata.

L'utente accederà ad una interfaccia grafica (pagina web) che gli permetterà in maniera semplice di accedere al catalogo, ricercare, visualizzare e scaricare i prodotti delle RR.

Il cuore di tutta l'architettura è rappresentata dal DMAS, che svolge una duplice funzione. Da un lato infatti gestisce i metadata utili per effettuare il search e il discovery sul data archive e dall'altro permette di accedere tramite il database dei metadata al data archive. Quest'ultimo rappresenta fisicamente lo storage dove verranno memorizzati i prodotti e consisterà di un filesystem. Questa soluzione è stata adottata visto l'enorme mole di dati con cui si lavora e la necessità di avere uno strumento che sia svincolato da sistemi di gestione di database.

L'utente potrà consultare il catalogo dei prodotti messi a disposizione che comprenderanno i parametri indicati nella Tabella 2. Tale catalogo è stato disegnato sulla base delle rianalisi attualmente prodotte che coprono gli ultimi 20 anni e potrà subire delle modifiche nelle versioni che verranno rilasciate successivamente nel corso del progetto.

Variables	Geographical coverage	Spatial resolution	Number of levels	Temporal resolution	Temporal coverage
Temperature	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Salinity	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Sea Surface Height	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	24 hr average field	1912-2011
Horizontal velocity (meridional and zonal component)	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Surface fluxes	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	6hr	1912-2011
Surface S,T and currents	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	6hr	1912-2011

**Tabella 2:** Elenco dei prodotti di RR che saranno resi disponibili nell'ambito del progetto.

*Campi di temperatura, di salinità e di velocità orizzontale:*

Si tratta di campi tridimensionali, rilasciati come medie giornaliere centrate alle 24 UTC di ogni giorno. I livelli verticali sono in unità di profondità in metri e le coordinate geografiche sono in gradi e decimi di grado.

Si studierà inoltre come filtrare il segnale delle correnti inerziali tramite un filtro passa-alto per poter eliminare il bias che ne deriva nella media giornaliera.

*Campi di temperatura, salinità e velocità orizzontali alla superficie:*

Si tratta di campi bidimensionali, rilasciati come campi istantanei (snapshot) ogni 6hr.

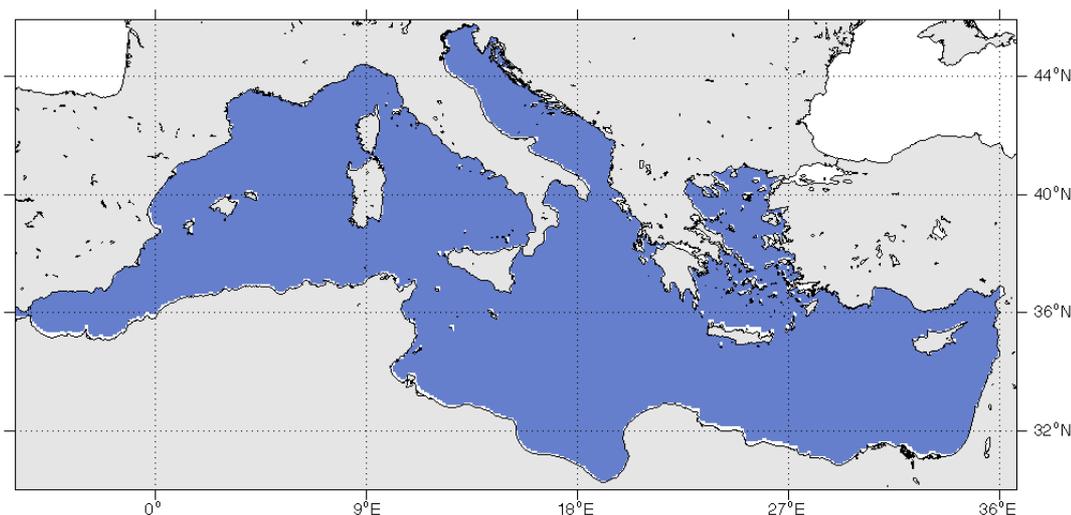
### Campi di elevazione della superficie libera:

Si tratta di campi bidimensionali, rilasciati come medie giornaliere centrate alle 24 UTC di ogni giorno.

### Campi di flussi all'interfaccia aria-mare:

Si tratta di campi bidimensionali, rilasciati come snapshot a 6 ore. Comprendono i flussi di calore con tutte le loro componenti (radiazione d'onda corta netta, radiazione d'onda lunga netta, calore sensibile e calore latente), di momento (sforzo tangenziale del vento zonale e meridionale) e di flusso d'acqua alla superficie (pari alla differenza tra precipitazione, runoff ed evaporazione).

I prodotti delle RR sono su griglia regolare in latitudine e longitudine ed il passo di griglia è di 1/16 di grado pari a circa 6.5 km. Il dominio di riferimento è indicato in Figura 4.



**Figura 4.** Dominio dei prodotti di RR.

I parametri tridimensionali verranno messi a disposizione su 33 livelli che corrispondono ai livelli standard IODE.

Per il viewing interattivo, l'utente avrà a disposizione solo le seguenti opzioni:

- medie mensili
- medie annuali

Questa scelta si è resa necessaria avendo a che fare con un grande volume di dati giornalieri che non permettono di avere buone prestazioni in termini di tempi di richiesta/risposta del Portale. La visualizzazione interattiva sarà consentita solo per campi medi come precedentemente specificato.

I prodotti selezionati potranno essere scaricati in formato NetCDF (Network Common Data Form), formato binario supportato da apposite librerie e largamente impiegato in ambito scientifico ([www.unidata.ucar.edu/software/netcdf/](http://www.unidata.ucar.edu/software/netcdf/)). La convenzione adottata è la CF-1.0.

### **3.3 Attività di formazione**

Nessuna.

### **3.4 Attività di disseminazione e divulgazione**

Nessuna.

### **3.5 Partecipazione a conferenze**

Nessuna.

## **4. Risultati ottenuti durante il periodo di riferimento**

### **4.1 Risultati specifici (banche dati, risultati delle misure, output di modelli, etc)**

È stata creata la banca dati di riferimento dei dati osservativi marini e dei dati atmosferici che verranno utilizzati nella produzione delle RR, come descritto nel paragrafo 3.

### **4.2 Pubblicazioni**

Non erano previste pubblicazioni nel periodo di riferimento.

### **4.3 Disponibilità di dati e output modellistici (formato, supporto, etc)**

Nessun prodotto previsto.

### **4.4 Deliverables completati**

Entrambi i deliverables previsti al primo anno sono stati completati.

## **5. Commento su eventuali scostamenti fra attività/risultati/deliverables previsti ed effettivamente realizzati**

Rispetto all'attività prevista al primo anno, si è avuto uno slittamento al secondo anno degli incontri con le comunità dei potenziali utenti, programmati per la definizione delle specifiche dell'archivio e dei protocolli di accesso ai dati. Nel corso del primo anno sono stati presi i primi contatti con i WP coinvolti nelle attività di archivio e messa a disposizione di prodotti grigliati di rianalisi e simulazioni per individuare elementi comuni nella gestione dei dataset.

## **6. Attività previste per il periodo successivo**

Nel secondo anno la discussione iniziata internamente al progetto, relativa all'organizzazione dell'archivio e della messa a disposizione dei dati di RR, verrà estesa anche al di fuori della comunità dei partecipanti al progetto.

Si procederà all'aggiornamento dei dataset relativi ai dati storici marini in base al rilascio di nuove versioni.

Nel secondo anno del progetto si procederà alla creazione del l'archivio con i dati delle RR prodotte e si implementerà una seconda versione del catalogo sulla base dei nuovi prodotti. Verrà inoltre sviluppato il sistema di visualizzazione e di downloading interattivo dell'archivio.

Sulla base dei risultati ottenuti, si valuterà quando organizzare un incontro con la comunità nazionale degli utenti con il rilascio della versione beta dell'archivio e del portale specifico.