



University
of Southampton



Agenzia nazionale per le nuove tecnologie, l'energia
e lo sviluppo economico sostenibile

Università degli Studi di Parma
Dipartimento di Ingegneria e Architettura
Dipartimento di Scienze della Terra

MASTER UNIVERSITARIO INTERSEDE
IN SCIENZE COSTIERE APPLICATE
ANNO ACCADEMICO 2010-11

*Un sistema informativo ambientale
per la gestione del mare:
uso del “free software” Ocean Data View*

Relatore

*Chiar.mo Prof.
Giuseppe M. R. Manzella*

Candidata

Dott.ssa Chiara Romano



University
of Southampton



Agenzia nazionale per le nuove tecnologie, l'energia
e lo sviluppo economico sostenibile

Università degli Studi di Parma
Dipartimento di Ingegneria e Architettura
Dipartimento di Scienze della Terra

MASTER UNIVERSITARIO INTERSEDE
IN SCIENZE COSTIERE APPLICATE
ANNO ACCADEMICO 2010-11

*Un sistema informativo ambientale
per la gestione del mare:
uso del “free software” Ocean Data View*

Relatore

Giuseppe M. R. Manzella
Giuseppe M. R. Manzella
ENEA-CRAM S. Teresa

Candidata

Chiara Romano
Dott.ssa Chiara Romano

INDICE

1. Importanza dei dati per la gestione dell'ambiente costiero	pag. 2
1.1. Strumenti per la raccolta dei dati	pag. 3
1.2. Procedura controllo di qualità dei dati	pag. 9
2. Informazioni che accompagnano i dati: i metadati	pag. 11
2.1. SeaDataNet: un esempio di infrastruttura Pan-Europea	pag. 14
2.2. Uso dei vocabolari controllati	pag. 17
3. Format condivisi	pag. 21
4. Ocean Data View	pag. 22
5. Conclusioni	pag. 31
APPENDICE 1	pag. 36
APPENDICE 2	pag. 38
Bibliografia	pag. 40

1. Importanza dei dati per la gestione dell'ambiente costiero

I dati sono di vitale importanza sia per la ricerca marina e oceanografica, sia per applicazioni nel campo marittimo (la costruzione di navi, per esempio, necessita di conoscenze di fluidodinamica e quindi di dati). La loro raccolta è necessaria per interpretare il moto ondoso, studiare i meccanismi degli scambi tra il mare e l'atmosfera (e quindi il clima), per studiare i parametri biologici e fisico-chimici del mare e poter fare delle valutazioni sull'ecosistema marino. L'interpretazione dei dati ci permette di fare previsioni sullo stato del mare, da utilizzare per la pesca, per la navigazione, per la salute degli ecosistemi, etc.

Oltre gli aspetti pratici che vi sono nel campo marittimo, i dati, in definitiva, permettono di studiare le complesse relazioni non lineari tra fenomeni fisici, chimici e biologici, dalla cui conoscenza si ottengono quelle informazioni che permettono di gestire in maniera ottimale l'ambiente marino. Ma vi è anche un aspetto, che è necessario sottolineare, in un mondo in cui uso del territorio, attività economiche e salute pubblica sono strettamente correlati, secondo quei concetti di sviluppo sostenibile che talvolta non sono compiutamente capiti nei suoi aspetti partecipativi. Questi infatti richiedono che ogni cittadino sia coinvolto nel processo decisionale e quindi abbia gli strumenti cognitivi per poter fornire il proprio parere. Questa è la società della conoscenza che i legislatori Europei intendono rafforzare e alla cui base vi è l'accesso ai dati.

L'importanza dei dati ambientali marini sta nella loro unicità e non riproducibilità. Il dato raccolto in quel momento e in quelle date condizioni meteo-marine sarà unico e la sua conservazione è importante per le ragioni sopra esposte, ma anche perché si può disporre di una serie temporale che permette di valutare i "trend" e quindi conoscere se (per esempio) le azioni fatte per la difesa ambientale sono efficaci o no.

Negli ultimi anni l'interesse della comunità mondiale verso le problematiche ambientali quali il riscaldamento globale, la salvaguardia dell'ambiente e i cambiamenti climatici, è aumentato enormemente e così anche la richiesta di risposte dalla comunità scientifica, tutto questo implica la costruzione di modelli che possano prevedere quello che succede.

Negli anni la tecnologia e la metodologia per la raccolta dei dati è stata migliorata, cercando di renderli il più accurati e precisi possibile, sono aumentate anche le organizzazioni o gli enti per la raccolta; sono stati redatti dei protocolli da seguire per uniformare la raccolta dati a livello mondiale e sono nate delle banche dati per archivarli. Nei paesi confinanti con i mari europei, più di 1000 laboratori scientifici, sotto la guida di organizzazioni governative e dell'industria privata, raccolgono dati utilizzando vari sensori a bordo di navi da ricerca, sottomarini, piattaforme fisse e alla deriva, aerei e satelliti per misurare parametri fisici, geofisici, geologici, biologici, chimici, e altro.

Molte lezioni sono state imparate da chi osserva il mare e lo studia. Innanzitutto, per poterlo misurare è importante conoscere il mare in tutte le sue componenti. Sembra quasi un ossimoro per cui si vuole osservare qualche cosa che bisogna conoscere. In effetti questo è il risultato proprio dell'incremento delle conoscenze, che a sua volta ci porta a ipotizzare l'esistenza di fenomeni prima non prevedibili e quindi a raccogliere nuovi dati. Ma questo processo richiede spesso la costruzione di nuovi sensori e nuove piattaforme così come lo sviluppo di nuove metodologie di campionamento. Tutte le informazioni relative a tecnologie e metodologie devono essere conservate assieme ai dati, passando quindi dal concetto di 'banca dati' a quello di 'sistema informativo'.

Questa tesi trova le sue ragioni in una serie di iniziative condotte a livello Europeo per la gestione dell'ambiente marino. Nella comunicazione della Commissione Europea COM(2007) 575 “**An Integrated Maritime Policy for the European Union**”, meglio conosciuta con il nome di Blue Book, viene affermato: *Availability and easy access to a wide range of natural and human-activity data on the oceans is the basis for strategic decision-making on maritime policy. Given the vast quantity of data collected and stored all over Europe for a wide variety of purposes, the establishment of an appropriate marine data and information infrastructure is of utmost importance.*

La gestione dei dati quindi risulta uno dei tasselli della Politica Integrata Europea Marino-Marittima, che ha gli obiettivi di:

- creare le condizioni ottimali per un uso sostenibile degli oceani e dei mari, permettendo la crescita del settore marittimo e delle regioni costiere;
- costruire una base di conoscenza ed innovazione per la politica marittima;
- permettere la più alta qualità di vita nelle regioni costiere;
- promuovere una leadership Europea negli affari marittimi internazionali;

Nel 2010 la DG MARE (Directorate-General for Maritime Affairs and Fisheries) ha lanciato la comunicazione Marine Knowledge 2020 che propone di liberare il potenziale di conoscenza dell'ambiente marino europeo. Esso mira principalmente a migliorare la nostra comprensione dei mari e degli oceani europei, rendendo l'utilizzo di dati marini più semplice e meno costoso e favorire la competitività tra gli utenti dei dati marini.

1.1. Strumenti per la raccolta dei dati.

L'acqua marina è composta per il 96,5% da acqua ed il 3,5% da sali disciolti quali Na^+ , Cl^- , Mg^{2+} , SO_4^{2-} , Ca^{2+} , K^+ ; i flussi di calore, l'evaporazione, le piogge, i fiumi e lo

scioglimento dei ghiacci vanno ad influire su quelli che sono i parametri fondamentali dell'acqua di mare: temperatura, salinità, densità, pressione, ossigeno disciolto, nutrienti, torbidità, corrente.

La misurazione di questi parametri, che non avviene per tutti in maniera diretta perché alcuni di essi sono ricavati da altri, fornisce dei numeri, espressi in unità di misura, e questi numeri sono i dati.

Gli strumenti utilizzati per la raccolta dei dati sono diversi a seconda delle informazioni che si vogliono avere, a seconda di chi raccoglie i dati e della disponibilità economica.

Le misure servono a determinare il valore di una particolare grandezza e la misurazione è soltanto una stima che diventa completa (assume significato) solo se si aggiungono informazioni sulla taratura degli strumenti e sull'incertezza, cioè quanto è giusta la misura.

Prima di effettuare la misura bisogna analizzare il sistema di misura stesso, questo consente di determinare cosa può influenzare e modificare i risultati ottenuti da esso e lo si può così utilizzare al meglio. Conosciuti i limiti dello strumento si può determinare il modo di utilizzo tenendo conto dei requisiti di accuratezza, delle condizioni in cui si usa lo strumento, se sono statiche o dinamiche, e di eventuali integrazioni con altri strumenti.

Le condizioni del sistema in cui si effettua la misura sono molto importanti perché le condizioni statiche implicano variabili diverse dalle condizioni dinamiche. Quando si lavora in condizioni statiche l'errore nelle misurazioni è dato dal valore riportato meno il valore vero, quest'errore può essere di due tipi:

- Sistematico: un errore che si ottiene per cause intrinseche del sistema di misura e lo si può quantificare e compensare con delle correzioni; questo errore influenza la media delle misure, la differenza tra il valore di riferimento e la media delle misure è detta accuratezza, si esprimerà quindi il valore della misura con affianco il valore \pm dell'accuratezza.
- Casuale (stocastico): è un errore che non si può prevedere, non si può correggere, ma lo si può ridurre aumentando il numero delle misurazioni. L'errore casuale si calcola considerando tutte le misurazioni (i-esime) e trovando il valore medio di esse, da cui si ricaverà la deviazione standard, ovvero di quanto i valori si discostano dal valore medio.

Uno dei vantaggi del lavorare in condizioni statiche è la ripetibilità della misura, un sistema statico consente di riavere quelle stesse condizioni più e più volte. Un sistema statico è inoltre un sistema lineare e questo permette una correzione della misura in maniera semplice. Al contrario in un sistema dinamico la prima variabile da considerare è la transitorietà del sistema, il dato raccolto in quel momento non lo si riavrà mai più, il sistema dinamico come dice il nome stesso, è qualcosa in movimento, qualcosa che cambia di continuo. Altra variabile è il tempo di risposta cioè il tempo necessario ad avere dei segnali di risposta significativi; infine la frequenza di campionamento ovvero il numero di volte con cui viene effettuata la misura.

Con qualsiasi strumento venga effettuata la misura, e per qualsiasi tipo di misura, è necessario calibrare lo strumento: lo si regola in modo da migliorarne l'accuratezza. Quest'operazione richiede il confronto con delle misure di riferimento prodotte utilizzando uno strumento campione.

L'utilizzo degli strumenti è legato al parametro di cui si vogliono avere i dati:

- **Temperatura.** È un parametro facile da misurare, ma difficile da definire; il valore che si misura è un concetto relativo che implica il riferimento ad uno standard contenuto nella ITS 90 (International Temperature Scale). Il valore della temperatura si misura con il termometro e nel campo dell'oceanografia ne esistono di diversi tipi: termometria standard e termometri a rovesciamento, funzionano con la dilatazione della colonna di mercurio al variare della temperatura e sono usati per il controllo in mare di altri sistemi o di misure puntiformi (Bottiglia a rovesciamento o *Niskin*-rosette sampler system); termometri termoresistivi-metallici e semiconduttori detti anche termistori, usano la variazione delle caratteristiche elettriche (resistenza) al variare della temperatura, sono utilizzati per la misura della temperatura in funzione della profondità e del tempo (catene di termistori, XBT, CTD); termometri all'infrarosso o radiometri, misurano la radiazione emessa dai corpi caldi e trovano applicazione nella misura remota della temperatura superficiale (radiometri satellitari).

Profilatore CTD: le misurazioni CTD sono basilari nell'oceanografia per caratterizzare le diverse masse d'acqua, grazie ai dati raccolti forniscono diagrammi di temperatura e salinità. Misurano simultaneamente in-situ conducibilità, temperatura e profondità generalmente abbinabili a sensori per la misura di altre grandezze bio-oceanografiche (ossigeno, pH, fluorimetro, torbidità, etc.) di cui forniscono i profili continui lungo tutta la colonna d'acqua al variare della profondità. Si può avere un unità singola di CTD (CTD profiler) oppure abbinata ad un campionatore d'acqua (CTD package). Il profilatore può essere operante su navi in movimento e trainato superficialmente e/o oscillare, può essere a perdere (XCTD), può essere posizionato su boe ancorate o su strutture sul fondo, può essere attaccato su un sistema a caduta libera. Lo standard attuale dei CTD è il modello SEABIRD 911 plus. L'acqua richiamata dalla pompa in aspirazione passa prima nel sensore della temperatura e successivamente nella cella della conducibilità (per il calcolo della salinità). Il flusso dell'acqua è costante e la costante di tempo di C è data dal numero di giri della pompa (ovvero i giri al secondo). Il sensore della conducibilità è totalmente interno e a tre elettrodi, invece quello della temperatura è un termistore invecchiato.

Sistema XBT (Expendable Bathythermograph): il batitermografo è una sonda che misura solamente la temperatura, mentre la profondità viene calcolata sulla base del tempo che la sonda impiega a profilare. Ha un'accuratezza scarsa. Lo XBT è composto da un sensore di temperatura a perdere (probe), un sistema di lettura-acquisizione dati e un lanciatore di bordo. Fornisce i profili della temperatura, in funzione della profondità, da una nave in movimento; il probe a perdere è dotato di una sonda con termistore ed è

collegato a bordo da un sottile cavo bipolare che si srotola da due bobine (bordo-sonda) disaccoppiando il movimento nave dalla discesa in caduta libera della sonda. La speciale forma idrodinamica della sonda garantisce una velocità di discesa controllata e nota e, misurando il tempo necessario alla discesa e conoscendo la velocità, si ricava la profondità. Tutto questo è effettuato dal sistema di bordo che misura la variazione termica attraverso il termistore della sonda, in collegamento via cavo; quando arriva sul fondo, o al termine del cavo, la sonda viene persa con la rottura del cavo stesso. Nei vecchi modelli i dati venivano registrati in forma analogica su carta, ora sono acquisiti in forma digitale, analizzati e restituiti anche in forma grafica attraverso un PC. I vantaggi degli XBT sono innanzitutto la facilità di utilizzo, le misure senza operazioni restrittive per la nave, le operazioni possibili su navi di opportunità e senza personale specializzato e si può effettuare il lancio anche in condizioni di mare mosso. Presenta anche degli svantaggi quali la scarsa accuratezza ($\pm 0,1^\circ\text{C}$) a causa della deriva del tempo dei termistori, ha problemi di invecchiamento, per indagini oceanografiche può essere necessaria una calibrazione pre-lancio, infine hanno un costo non trascurabile soprattutto per le acque profonde.

Immagini da satellite: dal satellite vengono eseguite misure di temperatura superficiale, colore del mare (normalmente associato a clorofilla), altezza del mare con metodologie che differiscono da parametro a parametro e queste misurazioni da lontano (remote sensing, telerilevamento) danno la possibilità di campionare la ricchezza di strutture nel mare, mostrando l'importanza della circolazione, l'impatto sul clima e la biologia del mare. Si può avere il risultato della misura per la temperatura con il ritardo di un giorno, del livello del mare con ritardo di qualche giorno, dei venti atmosferici sempre con un ritardo di qualche giorno e la presenza della clorofilla in acqua.

- **Salinità.** La sua definizione è cambiata negli anni e recentemente è stata adottata la cosiddetta “Salinità Assoluta” che è una misura di concentrazione basata sulla conducibilità di KCl. Originariamente la sua definizione era il peso in grammi di tutti i sali esistenti in un kg di acqua marina, quando tutti i carbonati erano stati convertiti in ossidi, il bromo e lo iodio sostituiti dal cloro, e tutte le sostanze organiche completamente ossidate. Praticamente si misurava la clorinità conoscendo la proporzionalità dei componenti: $\text{salinità} = 0.03 + 1.805$. La misurazione può avvenire in maniera elettrica (conducibilità) o chimica (titrazione). La misurazione elettrica può essere di due tipi: induttivo o ad elettrodi (conduttivo). Nel tipo induttivo si ha una cella basata su un doppio trasformatore di forma toroidale assemblato coassialmente, il primario del primo trasformatore riceve il segnale di eccitazione (AC), il secondario del secondo trasformatore lo rileva. L'accoppiamento tra i due avviene attraverso la spira conduttrice generata dall'immersione della cella nell'acqua marina che agisce da avvolgimento secondario nel primo trasformatore e da primario nel secondo; ogni variazione di resistenza elettrica della spira, quindi di conducibilità dell'acqua marina, genera una corrispondente variazione del segnale di uscita. I vantaggi di questo metodo sono la robustezza e la facilità costruttiva, in più si hanno pochi problemi di corrosione e

biofouling. Gli svantaggi comprendono innanzitutto una cella costruttivamente troppo grande e quindi una risoluzione spaziale limitata, in più la cella è parzialmente aperta ed è influenzabile da oggetti posti nelle vicinanze o limitati volumi di acqua circostante, infine necessità di bagni di calibrazione con volumi di acqua considerevoli (100 litri). Il metodo conduttivo invece si basa sulla misura diretta della conducibilità dell'acqua attraverso la legge di Ohm, rilevando il rapporto tra la corrente che passa nell'acqua della cella e la corrispondente caduta di tensione. La cella è generalmente costituita da due o più elettrodi (4 normalmente: due per la corrente e due per la tensione). Il profilatore CTD Sea Bird ne è un esempio con la sua cella a campo totalmente interno, il Neil Brown ha invece le celle parzialmente a campo esterno. Il metodo ad elettrodi è comunque il più usato per le misure in-situ e in laboratorio (Salinometro), con il metodo Neil Brown si raggiunge un'elevata miniaturizzazione e quindi un'elevata risoluzione spaziale per gli studi sulla microstruttura. Sono strutture parzialmente sensibili al campo esterno e necessitano di notevoli volumi di acqua per la calibrazione in laboratorio, solo le celle a campo interno sono di dimensioni ridotte e richiedono ridotti volumi di acqua. La costante di tempo è data dal tempo di lavaggio della cella e nel caso di profilatore CTD dal tempo di discesa.

- **Corrente.** È una grandezza vettoriale e si esprime in cm/s o nodi se si parla di velocità, e in gradi o radianti se si vuole esprimere la sua direzione. Si usano due metodi: il metodo Langrangiano e il metodo Euleriano. Il metodo Langrangiano usa dei sistemi derivanti, i *Drifters*: sostanze traccianti come la rodamina e la fluorescina; la *crociera* o il *paracadute* collegati con un cavo ad una boa che ne segnala il movimento; l' *Holey sock* è il tracciamento superficiale tramite posizionamento satellitare ARGOS; lo *Swallow floats* ovvero un pinger acustico a bassa frequenza racchiuso in una sfera di vetro prebilanciato di modo che una volta lanciato dalla superficie raggiunge un posizionamento in profondità, in equilibrio con la massa d'acqua circostante di cui segnala il movimento, è quindi un tracciamento acustico basato su un intervallo di tempo necessario al segnale per coprire la distanza emettitore-ricevitore; *ALACE* (Autonomous Langrangian Circulation Explorer) è un sistema derivante profondo capace di posizionarsi alla profondità voluta cambiando il proprio volume (densità), ciclicamente ritorna in superficie dove viene localizzato e trasmette i dati che ha immagazzinato (profili CTD e di temperatura) attraverso il sistema satellitare ARGOS, il tracciamento della deriva si ricava interpolando le varie posizioni (partenza ed arrivo) in superficie. Nel metodo Euleriano sono utilizzati: i *correntometri tradizionali*, a elica o a rotore, con i quali si misura l'intensità della corrente rilevata dal numero di giri del rotore in un determinato intervallo di tempo e la direzione dalla orientazione di un timone rispetto al nord magnetico; *correntometri acustici* col sistema Neil Brown basati sulla misura della differenza di fase tra i segnali acustici emessi-ricevuti dalle coppie di trasduttori, noti C ed L, aumenta la risoluzione nella misura della velocità aumentando il percorso del segnale acustico (riflettore acustico) senza cambiare la distanza tra i trasduttori (risoluzione spaziale); *ADCP* (Acoustic Doppler Current Profiler) si basa sull'effetto doppler, cioè sfrutta gli echi di ritorno (backscatters) delle particelle sospese in acqua e

permette di misurare il vettore della velocità della corrente nei vari strati della colonna d'acqua sopra/sottostante, lo si può installare sulla chiglia di una nave o su una boa subacquea.

- **Pressione.** È la forza per unità di superficie, la sua unità di misura in SI è il Pascal = Newton/m^2 . Viene misurata per conoscere parametri oceanografici legati alla sua variazione quali onde, maree e profondità. Secondo l'equazione idrostatica $P = \rho g z$, con g accelerazione di gravità, z profondità e ρ densità, e assumendo che il mare ha una densità media costante, esiste una proporzionalità tra pressione e profondità; in oceanografia l'unità di misura che si usa per la pressione è il decibar che equivale a 10^4 Pascal e che corrisponde a circa 1 m di profondità. Per la misurazione sono utilizzati sia strumenti analogici che digitali. Le *misure analogiche* sono effettuate o tramite elementi sensibili alla pressione, basati su colonna di liquido di densità conosciuta oppure elastici (diaframma, soffiato, tubo di Bourdon), oppure sono accoppiati con elementi di rilevazione e misura di tipo potenziometrico, strain gauge, capacitivo, induttivo o ottico. Le *misure digitali*, invece, utilizzano elementi a circuito risonante di tipo vibrante meccanici (corde, cilindri, diaframmi) o elettronici (capacità, induttanze, quarzi).

Boe off-shore: sono un insieme di strumenti. Considerate piccole isole, sono in realtà delle piattaforme che misurano lunghe serie temporali sia a livello climatico sia a livello di cambiamenti ambientali; misurano e analizzano, a bordo, il fenomeno dell'iterazione aria-acqua; serie temporali di dati sui parametri dell'acqua nel punto di ancoraggio; effettuano misure ondometriche in-situ; forniscono dati sulla corrente marina in mare aperto per gli studi sui movimenti delle masse d'acqua; effettuano indagini biologiche circa le comunità animali e vegetali che si insediano attorno e sulla boa stessa. Nel Mar Ligure, a 72 km da Genova, è ancorata la W1-M3A (west Mediterranean Multi Mooring) detta anche ODAS Italia 1 che è gestita dal CNR. Questa boa è ancorata a 1377 m di profondità e la parte galleggiante è lunga 51,22 m, pesa 12 tonnellate e a bordo ha un piccolo laboratorio ed una piccola stazione meteorologica posizionati nella "testa" della boa. I sensori per le indagini a mare forniscono dati su temperatura, conducibilità, pressione, ossigeno disciolto, clorofilla A totale, torbidità, altezza onde e direzione, profili sulla corrente e nutrienti. I dati sono raccolti a bordo e registrati in un data base ogni tre ore (00, 03, 06, 09, 12, 15, 18, 21); vengono trasmessi tramite segnale GSM a terra ogni sei ore quando ci sono almeno 2 data record disponibili. La comunicazione tra la W1-M3A, la stazione in superficie, e la boa ODAS in profondità avviene tramite una coppia di modem acustici della LinkQUEST. Ogni due ore il modem installato sul fondo a 1000 m di profondità, invia al modem della stazione in superficie segnali su conducibilità, temperatura e profondità acquisiti grazie ad un probe CTD. Queste informazioni sono integrate con tutti i dati collezionati dalla boa e trasmessi alla stazione a terra.

1.2. Procedura Controllo di Qualità dei Dati.

Un fattore importante nella diffusione dei dati è la velocità con cui questi sono trasmessi. È necessario trasmettere dati in tempo reale affinché vengano assimilati in modelli numerici che forniscono un'analisi più precisa dello stato del mare e previsioni più realistiche delle caratteristiche future dell'oceano.

Prima che i dati siano diffusi deve essere assicurata la loro qualità. Normalmente l'attività che assicura una buona qualità viene effettuata secondo due processi che prendono il nome di Quality Assurance e Quality Control.

La Quality Assurance è formata da quelle procedure da applicare prima e durante la raccolta dei dati ed include la preparazione del personale, i test degli strumenti e la loro calibrazione, nonché il controllo dei dati e degli strumenti durante l'acquisizione.

Il Quality Control si effettua dopo la raccolta dei dati ed include anche dei controlli di tipo statistico.

I criteri e le procedure per i controlli di qualità vengono definiti a livello internazionale nell'ambito dei programmi dell'Intergovernmental Oceanographic Commission dell'UNESCO (ed in particolare nel programma International Oceanography Data Exchange – IODE) in collaborazione anche con una altra organizzazione internazionale denominata Scientific Committee for Ocean Research (SCOR) .

Controlli di qualità possono essere fatti in “tempo reale” o “a posteriori”. Nel primo caso il controllo di qualità avviene con un minore numero di procedure di lavoro rispetto ai dati archiviati.

In generale, dati di buona qualità possono essere ottenuti solo seguendo precise metodologie e protocolli durante tutte le fasi della raccolta, dalla preparazione delle crociere all'utilizzo degli strumenti (es. preparazione del personale, il controllo della strumentazione, calibrazione) sino all'elaborazione e all'invio degli stessi.

Effettuata l'acquisizione esistono dei principi generali basati su 7 passaggi fondamentali per i dati in tempo reale:

1. Ogni osservazione in tempo reale trasmessa alla comunità oceanografica deve essere accompagnata da un buon descrittore di qualità.
2. Tutte le osservazioni dovrebbero essere automaticamente oggetto di un test per la qualità del dato.
3. I test di qualità devono essere sufficientemente descritti nei metadati.
4. Gli osservatori dovrebbero verificare o calibrare, in modo indipendente, il sensore prima della distribuzione.
5. Gli osservatori dovrebbero descrivere il loro metodo/taratura in tempo reale nei metadati.

6. Gli osservatori dovrebbero quantificare il livello di precisione di calibrazione e i relativi limiti di errore.
7. Sono necessari controlli manuali sulle procedure automatizzate, i dati raccolti in tempo reale e lo status del sistema di osservazione, devono essere forniti da un osservatore, su una opportuna scala temporale, per garantire l'integrità del sistema di osservazione.

2. Informazioni che accompagnano i dati: i metadati.

Fino a qualche anno fa tutti i dati provenivano da fonti diverse e soprattutto venivano raccolti solo su richiesta di specifico uso: previsioni climatiche, studio di un ecosistema, etc. Esistevano delle banche dati centralizzate dove solo pochi partecipavano. Lo scambio dei dati avveniva su richiesta e venivano spediti o inviati su floppy disc o CD-ROM da parte dell'ente o dell'organizzazione che li possedeva. La situazione col passare del tempo è cambiata per vari motivi, tra cui la sempre più crescente richiesta di dati da parte di ricercatori, operatori ambientali e cittadini, e dalla quantità enorme di dati che viene prodotta con le odierne tecnologie (il "diluvio di dati - data deluge"). Vi è stata una esigenza ed una forte richiesta affinché i dati fossero organizzati, raccolti, conservati e inseriti all'interno di banche dati che ne rendessero disponibile l'uso ai più.

I risultati della raccolta dei dati sono foto, fogli di calcolo, mappe, grafici, file di dati, ecc, questi oggetti sono utili ma generalmente non contengono informazioni su come, dove e da chi i dati sono stati raccolti.

Innanzitutto, bisogna distinguere tra "dato" ed "informazione". I dati sono osservabili, valori grezzi che sono il risultato di misure; questi valori possono essere numerici (come nella misurazioni di temperatura o di salinità) o nominali (come nei censimenti di specie per particolari regioni). Il termine informazione è comunemente usato per quei dati che sono già stati processati e si è giunti ad una conclusione.

I dati da soli potrebbero essere interpretati in maniera non corretta se non si conoscono (per esempio) le condizioni ambientali in cui sono state svolte le misure, le tecnologie, le piattaforme usate, eccetera. Tuttavia, se si forniscono le giuste informazioni descrittive, i dati possono essere trasformati in "informazioni" corrette. Le ulteriori informazioni possono includere la latitudine e la longitudine, la data di raccolta, la precisione della misura, la persona da contattare con domande sui dati, o il tipo di attrezzature utilizzate. La comunità scientifica chiama tutte le informazioni necessarie per meglio interpretare il dato con il nome di "metadati". I metadati sono file di testo, associati ai dati raccolti, che forniscono informazioni sui dati stessi; possono contenere diversi tipi di informazioni e costituiscono uno strumento per catalogare i file ed avere sempre a disposizione una scheda che documenti la natura del dato e le operazioni compiute su di esso. (Fig. 1.1.)

The image shows a text editor window titled 'TIG_COR1_136_SC - Blocco note'. It contains two sections of metadata, each enclosed in a red box. The first section is for Depth data, and the second is for Currents, Temperature, and Conductivity data. Below the metadata is a table of data points.

Depth Metadata:

```

//<subject>SDN:LOCAL:Depth</subject><object>SDN:P021:64:MPMN</object><units>SDN:P061::ULAA</units>
//<subject>SDN:LOCAL:Currents E-W</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
//<subject>SDN:LOCAL:Currents N-S</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
//<subject>SDN:LOCAL:Temperature</subject><object>SDN:P021:64:TEMP</object><units>SDN:P061::UPAA</units>
//<subject>SDN:LOCAL:Conductivity</subject><object>SDN:P021:64:CNDC</object><units>SDN:P061::OHMM</units>

```

Currents, Temperature, and Conductivity Metadata:

```

//<subject>SDN:LOCAL:Depth</subject><object>SDN:P021:64:MPMN</object><units>SDN:P061::ULAA</units>
//<subject>SDN:LOCAL:Currents E-W</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
//<subject>SDN:LOCAL:Currents N-S</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
//<subject>SDN:LOCAL:Temperature</subject><object>SDN:P021:64:TEMP</object><units>SDN:P061::UPAA</units>
//<subject>SDN:LOCAL:Conductivity</subject><object>SDN:P021:64:CNDC</object><units>SDN:P061::OHMM</units>

```

Data Table:

Station	Instrument	Type	Status	Time	Lat	Lon	Depth [m]
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T11:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T12:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T13:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T14:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T15:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T16:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T17:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T18:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T19:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T20:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T21:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	18/04/1997T22:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T10:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T11:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T12:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T13:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T14:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T15:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T16:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T17:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T18:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T19:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T20:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T21:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	19/04/1997T22:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T01:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T02:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T03:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T04:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T05:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T06:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T07:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T08:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T09:00:00	9.3025	44.3083	12
TIG_COR1_136_SC_136	ENVIREG02	TIGSC	ST	20/04/1997T10:00:00	9.3025	44.3083	12

Fig. 1.1.: metadati di un set di dati in formato ODV.

Metadati è una parola che è stata menzionata raramente in passato e da qui nasce un incertezza sulla qualità delle osservazioni storiche. Ora i metadati sono critici nella registrazione della provenienza di un set di dati. Una volta pubblicati, i dati devono rimanere disponibili e possono essere usati anche dopo la fine dei progetti che li hanno raccolti. Così, per capire i set di dati si rende necessaria una descrizione su:

- come gli strumenti di acquisizione sono stati progettati e costruiti (*HOW*);
- ciò che è stato acquisito (*WHAT*);
- quando, dove e come i set di dati sono stati raccolti e pre-trattati (*WHEN, WHERE*);
- una descrizione accurata delle fasi di elaborazione (*HOW*);
- chi ha acquisito ed elaborato i set di dati (*WHO*);
- perché i set di dati sono stati acquisiti e trattati (*WHY*).

Ci sono ancora molti ostacoli alla fornitura di un set di dati per gli utenti intermedi e finali, alcuni dei quali sono legati alle tecnologie dell'informazione (ad esempio la codifica semantica, modello di dati).

La codifica semantica è necessaria in quanto occorre definire ogni variabile con un nome unico e attribuire a tale nome un significato preciso.

Il modello di dati è relativo al cosiddetto 'format' con cui sono scritti. Nella Fig. 1.1. si dà un esempio di format Ocean Data View, ma ne esistono altri (inclusi i semplici formati tabellari ottenuti con EXCEL).

Per accedere ai dati occorre avere un catalogo di essi. Tuttavia questi cataloghi (che permettono la selezione sulla base di richieste specifiche) possono essere costruiti in maniera diversa. Quindi per accedere a dati residenti in diverse banche occorre mettersi d'accordo (tra l'altro) sul contenuto del catalogo, sulla semantica e sul formato dei dati. In questo modo si arriva a rendere 'interoperabili' sistemi differenti.

L'interoperabilità, intesa come la possibilità di accesso ed uso dei dati per chiunque, è stata definita come una delle priorità al fine di fornire agli utenti finali un più facile accesso ai diversi tipi di prodotti. L'interoperabilità non è un concetto nuovo nella scienza dell'informazione, l'idea è sempre stata quella di garantire che i dati e le informazioni possano fluire, nel modo più efficiente possibile, tra persone e sistemi. Quindi, i dati gestiti da più fonti eterogenee richiedono un'elaborazione complessa come la mediazione dei contenuti e l'aggregazione. Questo non è un compito semplice, a causa della diversa gestione dei dati e dei diversi sistemi di elaborazione.

La Direttiva 2007/2/CE del Parlamento Europeo e del Consiglio del 14 Marzo 2007 ha istituito un'infrastruttura per l'informazione territoriale nella Comunità Europea (Direttiva INSPIRE). Tale direttiva fornisce le linee guida per la diffusione dei dati e dei metadati basandosi sui seguenti punti: "La politica della Comunità in materia ambientale mira ad un elevato livello di tutela tenendo conto della diversità delle situazioni nelle varie regioni della Comunità. Le informazioni, comprese quelle territoriali, sono necessarie anche per la formulazione e l'attuazione di questa e di altre politiche comunitarie, che devono integrare disposizioni di protezione dell'ambiente, come sancito dall'articolo 6 del trattato. Per realizzare tale integrazione occorre istituire misure di coordinamento tra gli utilizzatori e i fornitori delle informazioni, per poter combinare le informazioni e le conoscenze disponibili in vari settori diversi."

"I problemi relativi alla disponibilità, alla qualità, all'organizzazione, all'accessibilità e alla condivisione delle informazioni territoriali sono comuni a molte tematiche politiche e categorie di informazioni e si riscontrano a vari livelli dell'amministrazione pubblica. Per risolvere tali problemi sono necessarie misure in materia di scambio, condivisione, accesso e utilizzo di dati territoriali e di servizi relativi ai dati territoriali interoperabili tra i vari livelli dell'amministrazione pubblica e tra i vari settori. Occorre pertanto istituire un'infrastruttura per l'informazione territoriale nella Comunità."

"La notevole diversità di formati e di strutture in cui vengono organizzati e resi accessibili i dati territoriali nella Comunità ostacola la possibilità di formulare, attuare, monitorare e valutare in maniera efficiente la normativa comunitaria che incide, direttamente o indirettamente, sull'ambiente; per questo è necessario disporre di misure di attuazione per agevolare l'utilizzo dei dati territoriali provenienti da fonti diverse in tutti gli Stati membri. Tali misure dovrebbero essere tali da consentire l'interoperabilità dei set di dati territoriali e gli Stati membri devono garantire che i dati o le informazioni necessari per il conseguimento dell'interoperabilità siano disponibili secondo condizioni che non ne restringono l'utilizzo per il suddetto scopo. Le disposizioni di esecuzione dovrebbero essere basate, ove possibile, su standard internazionali e non dovrebbero comportare costi eccessivi per gli Stati membri."

All'interno della direttiva, come detto prima, vengono fornite anche le istruzioni per la compilazione dei metadati; nel regolamento CE n. 1205/2008 della Commissione del 3 dicembre 2008 si stabilisce che: "È necessario definire un set di elementi di metadati per consentire l'individuazione della risorsa di informazione per la quale sono creati i metadati, la loro classificazione, la loro ubicazione geografica e il riferimento temporale, nonché la

loro qualità e validità, la conformità rispetto alle disposizioni di esecuzione sull'interoperabilità dei set di dati territoriali e dei servizi, le limitazioni d'uso e di accesso e l'organizzazione responsabile delle risorse. Gli elementi di metadati legati al registro dei metadati stesso sono altresì necessari per controllare che i metadati creati siano aggiornati e per individuare l'organizzazione responsabile della loro creazione e del loro aggiornamento. Si tratta del set minimo di elementi di metadati necessari per conformarsi alla direttiva 2007/2/CE e non esclude la possibilità per le organizzazioni di documentare le risorse di informazioni in modo più ampio con elementi supplementari derivanti da norme internazionali o metodi di lavoro esistenti nella loro comunità di interesse. Non è esclusa neanche la possibilità di adottare linee guida stabilite e aggiornate dalla Commissione, in particolare quando è necessario garantire l'interoperabilità dei metadati.”

La Commissione Europea (DG MARE in particolare) ha varato vari regolamenti per implementare la Direttiva INSPIRE. Di seguito si presenta la lista dei regolamenti e gli standard definiti in sede internazionali.

- Commission Regulation 1205/2008
 - Metadata implementing rules (lineage ISO19101, metadata elements ISO19115, metadata on metadata ISO8601)
- Commission Regulation 976/2009
 - Network Services (ISO 19128, QoS, Discovery, View)
- Commission Regulation 268/2010
 - Response Time
- Commission Regulation 1088/2010
 - QoS (Quality of Service)
 - Downloading Services
 - Transformation Services
- Commission Regulation 1089/2010
 - Interoperability (ISO 19103, ISO 19107, ISO 19108, ISO 19111, ISO 19115, ISO 19127, ISO 19139)

2.1. SeaDataNet: un esempio di infrastruttura Pan-Europea.

Un esempio di infrastruttura Pan-Europea è il progetto finanziato dall'Unione Europea, ancora in corso di svolgimento (2006-2011; 2011-2014), SeaDataNet. Questo progetto è stato creato per la gestione, l'indicizzazione e l'accesso a set di dati, raccolti dalle crociere da ricerca e altre attività di osservazione nei mari europei e negli oceani. È sotto il controllo dei National Oceanographic Data Centres (NODCs) e dei servizi di informazione marini di importanti istituti di ricerca, provenienti da 35 paesi costieri che si affacciano sui mari europei. La rete SeaDataNet include anche Satellite Data Centers, centri esperti nella

costruzione di modelli e la Commissione Oceanografica Intergovernativa (COI) dell'Unesco, il Consiglio per l'esplorazione del Mare (ICES), e l'EU Joint Research Centre (EU-JRC). SeaDataNet, quindi, dà una panoramica sulle varie organizzazioni marine in Europa e sul loro impegno nei progetti di ricerca marina, su come gestiscono i dati raccolti e come avviene l'acquisizione dei dati attraverso crociere di ricerca e programmi di monitoraggio nelle acque Europee e negli oceani del mondo.

La Fig. 2.1. mostra l'insieme di informazioni gestite e disponibili in SeaDataNet. Esse comprendono:

- European Directory of Marine Organisations (EDMO) che presenta la lista di circa 1500 organizzazioni che lavorano a livello pan-europeo per le ricerche in campo marino;
- European Directory of Marine Environmental Data sets (EDMED) che descrive circa 3500 data sets raccolti da istituzioni pan-europee;
- European Directory of Marine Environmental Research Projects (EDMERP) che descrive circa 1800 progetti;
- Cruise Summary Reports (CSR) che consiste essenzialmente in un rendiconto delle compagnie di misura, secondo standard approvati in sede internazionale;
- European Directory of the Ocean-observing Systems (EDIOS) che da accesso a piattaforme (punti fissi) di monitoraggio in situ; questa directory è in fase di nuovo sviluppo in ambito di un progetto della DG MARE chiamato EMODnet Physical Parameters.

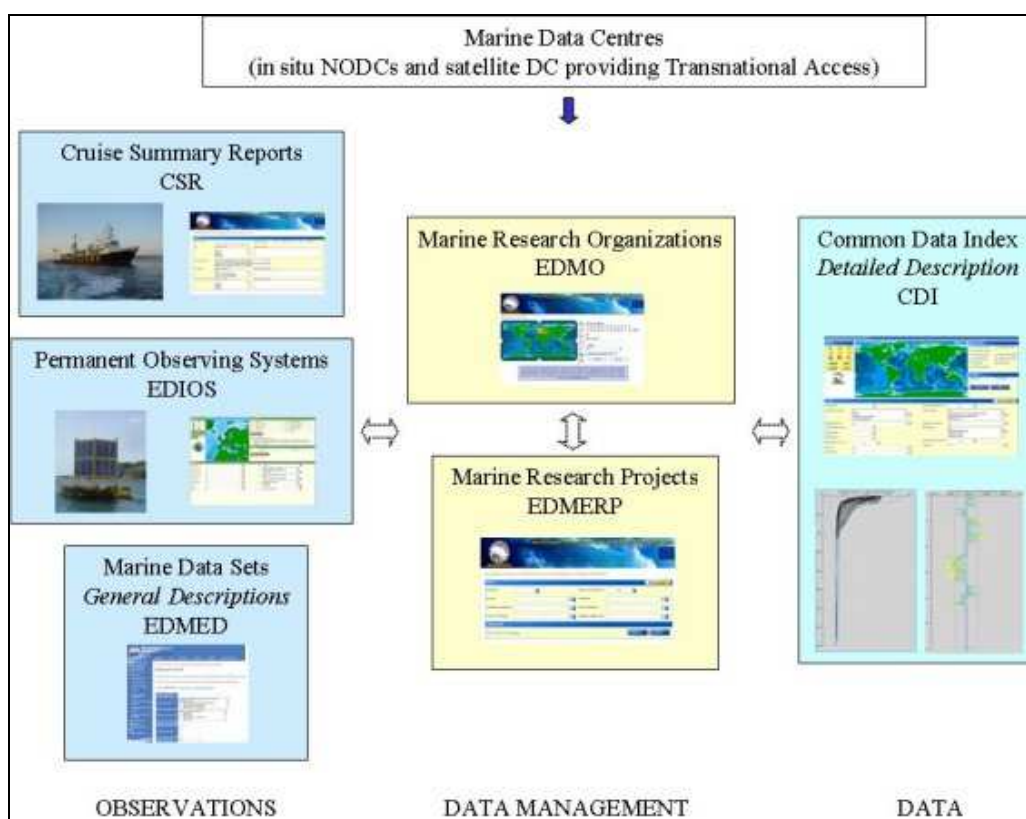


Fig. 2.1.: provenienza delle informazioni e gestione delle stesse in SeaDataNet.

SeaDataNet ha costruito un sistema standardizzato per la gestione dei dati di grandi dimensioni e dei diversi set di dati raccolti dalle flotte oceanografiche e dai nuovi sistemi di osservazione automatici; ha potenziato le infrastrutture già esistenti e messo in rete i vari centri nazionali di dati oceanografici e satellitari, di 35 paesi diversi, attivi nella raccolta di dati. L'interoperabilità di SeaDataNet consiste nel:

- uso di vocabolari comuni;
- adozione dello standard ISO per i metadati;
- fornire servizi di validazione per il controllo della qualità della manutenzione dei metadati;
- usare dei formati di trasformazione dei dati per la loro divulgazione;
- usare un protocollo comune per il controllo dei dati;
- fornire dei software standard;
- fornire un Registro Centrale per gli Utenti;
- usare i web services nella struttura di SeaDataNet.

L'accesso ai dati in SeaDataNet si basa sull'utilizzo di Common Data Index (CDI) Data Discovery e Access Service, ovvero un indice di set di dati individuali basati sul formato standard ISO. All'utente sono forniti un enorme volume di dati uniformati presi dalla rete SeaDataNet.

La CDI query interface è la schermata dove è possibile inserire i criteri di scelta per il set di dati richiesto, il risultato della query sarà l'elenco di dati e la loro localizzazione su una carta geografica. I dati sono presi dalla rete formata dalle banche dati differenti che formano la rete SeaDataNet. Insieme ai dati vengono forniti i relativi metadati che forniscono inoltre informazioni sul possibile utilizzo dei dati selezionati e sulla loro possibilità di uso pubblico. L'interfaccia utilizzato è una sorta di negozio virtuale che permette di inserire nel carrello della spesa il set scelto. (Fig. 2.2.)

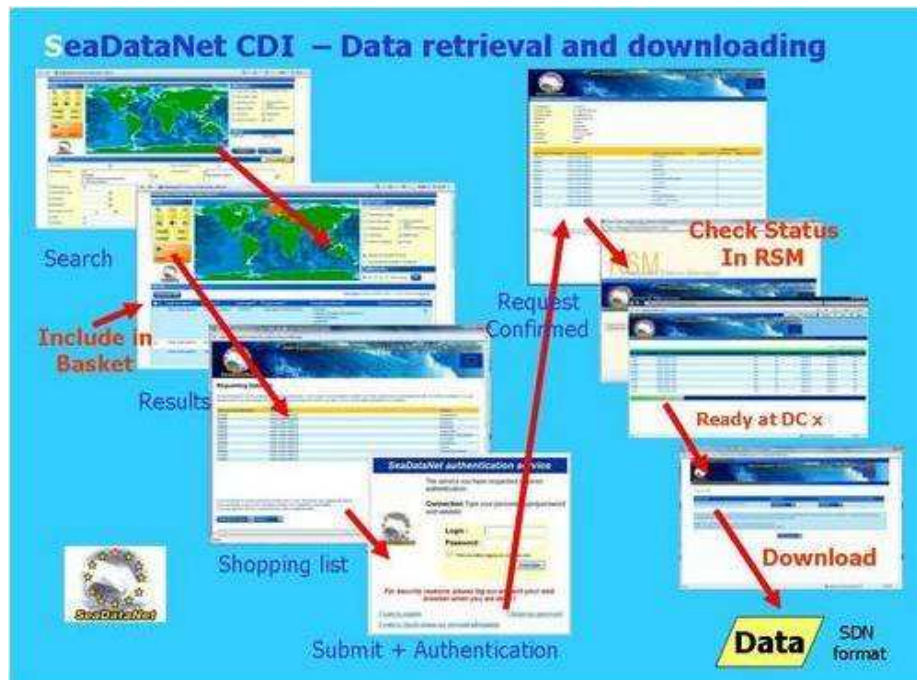


Fig.2.2. : passaggi per l'acquisizione dei dati in SeaDataNet. (dal sito SeaDataNet)

L'utente dei CDI si troverà davanti a 2 interfacce:

- ❖ CDI Quick Search Interface funziona con il metodo drill-down, ovvero la ricerca penetra nei dati anche a diversi livelli di gerarchia; l'utente ha immediatamente una panoramica dell'intero range di dati, sia sulla mappa, sia tramite specifici campi di metadati. Molti dei campi utilizzano termini presi dai vocabolari comuni e ogni termine ha il riferimento per il campo del CDI database in cui compare. Cliccando su ogni campo o termine viene rinnovata la selezione in base alle restrizioni imposte, di fianco l'utente può combinare le richieste con ulteriori criteri spazio-temporali.
- ❖ CDI Extended Search interface funziona con una combinazione di criteri diversi che possono essere impostati dall'utente da un menu a tendina; una volta attivata la ricerca si ottiene il risultato. Non inserendo nessun criterio di ricerca verrà mostrato l'intero CDI file come risultato.

In questo modo si riesce a ricercare meglio e rapidamente set di dati marini, il CDI attualmente ha accesso ad oltre 850000 data set provenienti da più di 300 organizzazioni Europee.

2.2. Uso dei vocabolari controllati.

Quando è partito il progetto SeaDataNet nel 2006, i vocabolari controllati esistevano già all'interno della comunità di gestione dei dati oceanografici Europei, ma non erano descritti bene come erano invece gestiti. Il precursore di SeaDataNet, Sea-Search, riempiva il campo

dei metadati con delle “librerie”. Quale dovesse essere il contenuto dei vocabolari è stato stabilito successivamente dalle governance nella forma del Technical Task Team. Nel frattempo gli esperti nel settore disponibili sono aumentati, e le governance hanno dato un punto d’appoggio più sicuro con la creazione del SeaVoX vocabulary governance group che è sotto l’auspicio dello IODE MarineXML Steering Group.

Da quando esiste SeaDataNet c’è stata un’esplosione nel numero dei vocabolari attraverso l’aggiunta di termini e attraverso la creazione di nuovi vocabolari per riempire i campi dei metadati. Attualmente ci sono circa 100 vocabolari ritenuti interessanti per SeaDataNet e Geo-Seas, che vengono usati per compilare i metadati in EDMED, CSR, EDIOS, e documenti CDI.

I vocabolari comuni sono un insieme di termini (parole, codici, ecc) che vengono utilizzati in una specifica comunità, i vocabolari forniscono un meccanismo per la comunicazione, sia essa scritta, orale o elettronica, purché il significato dei termini sia noto e concordato dai membri della comunità. Quando un vocabolario è formalmente gestito, diventa un vocabolario controllato. In questo caso "gestito" indica che i termini vengono memorizzati e gestiti con procedure concordate, queste procedure dovrebbero esistere per l’aggiunta dei termini, per modificare i termini e, più raramente, per disapprovare dei termini. Ad esempio, qualora volessimo dare un nome ad un parametro dovremmo scegliere fra una serie di nomi già concordati.

Per consentire una gestione ufficiale, un vocabolario controllato può essere organizzato in diversi modi. Ci sono tre grandi categorie di vocabolari controllati: flat, multi-livello e relazionale.

- flat: sono vocabolari controllati che forniscono una serie di termini usati, alcuni vocabolari controllati flat forniscono ulteriori informazioni su ogni termine;
- multi-livello: vocabolari controllati in grado di sviluppare un vocabolario controllato flat assegnando ogni termine di una categoria;
- relazionali: vocabolari controllati che forniscono un insieme di termini, e la modalità con cui sono associati tra loro.

Quindi utilizzando set standardizzati di termini nei metadati e nei dati etichetta, si risolve il problema delle ambiguità associate ai dati di marcatura. Questo apre una serie di possibilità per la manipolazione al computer, la distribuzione e il riutilizzo a lungo termine. Un esempio di come i computer possono trarre beneficio dall’uso dei vocabolari controllati: una serie di dati può avere una colonna denominata "temperatura della colonna d’acqua" e un altro potrebbe avere "temperatura acqua" o anche "temperatura". Per l’occhio umano, la somiglianza è evidente, ma un computer non sarebbe in grado di interpretare questi come la stessa cosa a meno che tutte le opzioni possibili non siano impostate nel suo software.

Ad esempio in formato ODV nei metadati sono elencate le variabili di cui si sono raccolti i dati e per ogni variabile, che viene indicata col suo nome, vi è una sigla di riferimento; di fianco è indicato anche il numero di vocabolario all’interno del quale bisogna andare a ricercare la sigla. Una volta individuato il vocabolario basterà ricercare la sigla al suo interno e ci saranno le informazioni che vogliamo sapere su quella determinata variabile.

Vediamo un esempio:

```
...</subject><SDN:LOCAL:Depth</subject><object>SDN:P021:64:MPMN</object><units>SDN:P061::ULAA</units>
...</subject><SDN:LOCAL:Currents E-W</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
...</subject><SDN:LOCAL:Currents N-S</subject><object>SDN:P021:64:RFVL</object><units>SDN:P061::UVBB</units>
...</subject><SDN:LOCAL:Temperature</subject><object>SDN:P021:64:TEMP</object><units>SDN:P061::UPAA</units>
...</subject><SDN:LOCAL:Conductivity</subject><object>SDN:P021:64:CNDC</object><units>SDN:P061::OHMM</units>
```

Fig. 2.3.: descrizione dei metadati in un set di dati in formato ODV.

nella Fig. 2.3. all'interno del rettangolo rosso si può leggere SDN:LOCAL:Depth questo indica la variabile Profondità e SDN:P021:64:MPMN indica il vocabolario P021, versione 64 e la sigla da ricercare è MPMN. All'interno del rettangolo giallo, invece, è riportata l'unità di misura della variabile con il relativo vocabolario di riferimento. Si può aggiungere che in questa fase di sviluppo dei sistemi informativi i vocabolari vengono aggiornati continuamente e quindi esistono varie versioni.

Se cerchiamo all'interno del vocabolario troveremo (Fig. 2.4.):

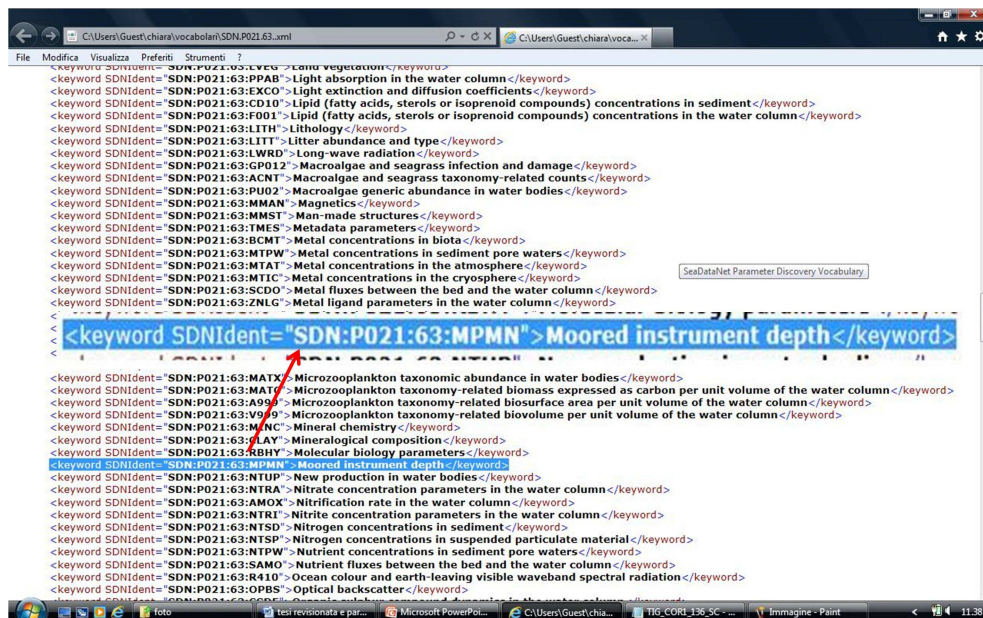


Fig. 2.4.: ricerca all'interno di un vocabolario controllato.

Quindi la sigla MPMN della Profondità si riferisce alla profondità in cui è ormeggiato lo strumento.

The screenshot shows a software window titled 'TIG_CORI_136_SC - Blocco note'. It contains an XML header and a data table. The table has the following columns: LOCAL_CDI_ID, EDMO_code, Cruise, Station, Type, yyyy-mm-ddThh:mm:ss, Longitude [degrees_east], Latitude [degrees_north Bot., and Depth [m]. The data rows show various measurements from cruise TIGCORI136SC, station TIGSC, starting from 18/04/1997112:00:00 to 20/04/1997110:00:00. A red box highlights the first two table headers, and another red box highlights the second two table headers.

LOCAL_CDI_ID	EDMO_code	Cruise	Station	Type	yyyy-mm-ddThh:mm:ss	Longitude [degrees_east]	Latitude [degrees_north Bot.	Depth [m]
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997112:00:00	9.3025 44.3083 20	-2.17	16.249 45.946
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997113:00:00	9.3025 44.3083 20	-2.50	16.291 45.976
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997114:00:00	9.3025 44.3083 20	-0.89	16.291 45.991
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997115:00:00	9.3025 44.3083 20	-0.39	16.273 45.968
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997116:00:00	9.3025 44.3083 20	-0.92	16.273 45.946
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997117:00:00	9.3025 44.3083 20	-0.66	16.273 45.961
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997118:00:00	9.3025 44.3083 20	-0.50	16.273 45.961
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	18/04/1997119:00:00	9.3025 44.3083 20	-0.40	16.273 45.953
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997120:00:00	9.3025 44.3083 20	2.60	16.297 45.983
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997121:00:00	9.3025 44.3083 20	4.55	16.315 46.021
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997122:00:00	9.3025 44.3083 20	4.82	16.333 46.028
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997123:00:00	9.3025 44.3083 20	5.93	16.303 45.990
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997124:00:00	9.3025 44.3083 20	2.69	16.333 46.020
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997125:00:00	9.3025 44.3083 20	2.70	16.333 46.020
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997126:00:00	9.3025 44.3083 20	3.01	16.309 45.998
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997127:00:00	9.3025 44.3083 20	2.24	16.291 45.953
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997128:00:00	9.3025 44.3083 20	2.28	16.309 45.983
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997129:00:00	9.3025 44.3083 20	3.16	16.261 45.931
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997130:00:00	9.3025 44.3083 20	1.90	16.243 45.894
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997131:00:00	9.3025 44.3083 20	2.59	16.237 45.901
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997132:00:00	9.3025 44.3083 20	2.05	16.255 45.931
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997133:00:00	9.3025 44.3083 20	1.82	16.279 45.953
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997134:00:00	9.3025 44.3083 20	1.82	16.279 45.953
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997135:00:00	9.3025 44.3083 20	1.88	16.327 45.983
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997136:00:00	9.3025 44.3083 20	1.74	16.345 46.006
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997137:00:00	9.3025 44.3083 20	1.72	16.315 45.968
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997138:00:00	9.3025 44.3083 20	1.80	16.279 45.938
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997139:00:00	9.3025 44.3083 20	1.84	16.267 45.938
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997140:00:00	9.3025 44.3083 20	1.89	16.237 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997141:00:00	9.3025 44.3083 20	1.82	16.220 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997142:00:00	9.3025 44.3083 20	1.84	16.225 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	19/04/1997143:00:00	9.3025 44.3083 20	1.85	16.208 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997144:00:00	9.3025 44.3083 20	1.88	16.225 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997145:00:00	9.3025 44.3083 20	1.86	16.202 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997146:00:00	9.3025 44.3083 20	1.99	16.202 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997147:00:00	9.3025 44.3083 20	1.98	16.202 45.856
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997148:00:00	9.3025 44.3083 20	1.88	16.273 45.938
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997149:00:00	9.3025 44.3083 20	2.79	16.369 46.035
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997150:00:00	9.3025 44.3083 20	2.65	16.345 46.035
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997151:00:00	9.3025 44.3083 20	2.55	16.345 46.035
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997152:00:00	9.3025 44.3083 20	2.26	16.333 46.013
TIGCORI136SC	136	ENVIREG02	TIGSC	ST	20/04/1997153:00:00	9.3025 44.3083 20	1.88	16.321 46.006

Fig. 2.5.: descrizione dell'intestazione delle colonne di dati in un set di dati in formato ODV.

Anche nell'intestazione delle colonne dei dati si fa riferimento a dei vocabolari (Fig. 2.5.), ad esempio EDMO_code fa riferimento all'European Directory of Marine Organizations. Ciascuna di queste directory ha un up-to-date di nomi e indirizzi delle organizzazioni, che sono coinvolte nell'attività di acquisizione dati, elaborazione e gestione e progetti di ricerca. Pertanto, il repertorio europeo delle organizzazioni Marine - EDMO è stata messo a punto per ottimizzare e gestire questi indirizzi e profili di organizzazioni in modo efficiente e in modo coerente. Attualmente, EDMO elenca e descrive più di 1.000 istituti di ricerca, Data Center Holding, agenzie di monitoraggio, organizzazioni governative e private, che sono in un modo o nell'altro coinvolte in attività di ricerca oceanografica e marina e/o gestione di dati e informazioni o servizi di assistenza. Nel caso dell'ENEA Centro Ricerche Marine di Pozzuolo di Lerici il codice di identificazione EDMO è il 136 come si osserva in Fig. 2.5..

Nel mondo reale, non è sempre possibile o piacevole per i fornitori di dati utilizzare gli stessi termini, in questi casi, il vocabolario controllato può essere usato come un mezzo per cui i data center sono in grado di classificare i propri termini equivalenti.

3. Format condivisi.

Fino a qualche anno fa tutti i dati provenivano da fonti diverse, soprattutto venivano raccolti solo su richiesta di specifico uso: previsioni climatiche, studio di un ecosistema, etc. Non esistevano delle banche dati dove poterli inserire e conservare nel tempo, consultarli ed eventualmente utilizzarli, su richiesta venivano spediti o inviati su floppy disc o CD-ROM da parte dell'ente o dell'organizzazione che li possedeva. La situazione col passare del tempo è cambiata e i dati sono ora molti di più, vengono raccolti, conservati ma soprattutto sono raccolti periodicamente e inseriti all'interno di banche dati che ne rendono disponibile l'uso ai più.

L'importanza del preservare i dati già raccolti non è mai sottolineata abbastanza: innanzitutto l'uso ed il riuso dei dati può essere utile alla società, in più il conservare dei dati può essere meno dispendioso che doverli riacquisire, inoltre i dati servono a dare validità alle pubblicazioni scientifiche, infine è importante conservare dei dati che si rischia di non poter riprendere mai più. La comunità scientifica oceanografica sta compiendo un serio sforzo nell'analisi dei dati ambientali marini e nella valutazione degli errori ad essi associati. Il lavoro che viene fatto è tutt'altro che facile, dal momento che occorre conoscere i metodi e gli strumenti con cui sono stati osservati i dati, la accuratezza, la precisione e il modo in cui gli stessi dati vengono usati. Inoltre occorre conoscere le condizioni in cui essi sono stati raccolti.

Un grosso problema è rappresentato dal dover convertire molti dei dati raccolti dai diversi enti in un formato che sia comune a tutti. Molto spesso i dati vengono raccolti e poi analizzati in formati diversi a seconda di chi li ha raccolti, quindi per poterli diffondere e rendere disponibili per le banche dati è necessario convertirli in formati di uso comune.

Ad esempio per poter accedere ai dati da SeaDataNet è necessario che i files siano in un formato definito come standard SeaDataNet. I formati "standard" sono: Ocean Data View e netCDF. Opzionale, ma non consigliato, è il formato MedAtlas.

L'infrastruttura mette quindi a disposizione, per chi ne avesse bisogno, i programmi di conversione NEMO e Med2MedSDN.

Il software NEMO permette la conversione di tutti i file di tipo ASCII, a SeaDataNet ODV, e a Medatlas ASCII format. In più NEMO utilizza sia vocabolari comuni sia le EDMO directory.

I file di input per NEMO possono essere individuali, profili verticali, serie temporali o traiettorie, o file di raggruppamento di diversi profili verticali, serie temporali o traiettorie (CTD, bottiglie, correntometro, etc...). NEMO è in grado di riformattare tutti questi tipi di file, l'unico requisito è che i file siano ASCII, con le misure dei dati in colonne.

Tutte le informazioni relative a Ocean DataView si trovano nel sito <http://odv.awi.de/>, per netCDF si trovano le informazioni nel sito del programma americano UNIDATA <http://www.unidata.ucar.edu/software/netcdf/>.

4. Ocean Data View.

Ocean Data View (ODV) è un programma che consente l'esplorazione interattiva e grafica di profili oceanografici e/o altri profili georeferenziati, transetti o serie temporali di dati. Il software è disponibile per Windows, Mac OS X, Linux, e UNIX.

Le collezioni di ODV e i file che ne derivano sono piattaforme indipendenti ed è possibile scambiarle con tutti i sistemi che li supportano. Il software permette di analizzare grandi set di dati e può produrre delle mappe con contrassegnate le stazioni di campionamento di alta qualità, questo grazie a delle cartine fornite dal programma ETOPO.

Il National Oceanic and Atmospheric Administration (NOAA) ha sviluppato ETOPO, ovvero un pacchetto che permette la visualizzazione di cartine, delle varie zone del mondo, molto dettagliate (l'ultima versione del programma permette la visualizzazione dei ghiacci antartici e delle "Bedrocks"). ETOPO permette una copertura topografica e batimetrica totale estendendosi in latitudine dal -90° al $+90^{\circ}$ ed in longitudine da -180° a $+180^{\circ}$; è stato sviluppato per supportare i ricercatori nelle previsioni degli tsunami, nella costruzione di modelli per previsioni, nello studio della circolazione marina e per migliorare la visualizzazione della Terra. Le linee di riva, le batimetrie, la topografia, ed i dati sulle Bedrocks sono forniti dalle diverse agenzie del Governo degli Stati Uniti e dalle Istituzioni Accademiche.

Vediamo alcune funzioni generali del programma.

Innanzitutto in Ocean Data View è possibile importare file testo, il programma li analizza e tramite il riconoscimento delle variabili li rende in forma grafica. Quindi una volta raccolti i dati, posso creare un foglio Excel (Fig. 4.1.) dove è possibile modificare i dati aggiungendo o rimuovendo delle colonne o dei singoli dati. Il foglio deve essere salvato con l'estensione .txt affinché possa poi essere aperto in ODV.

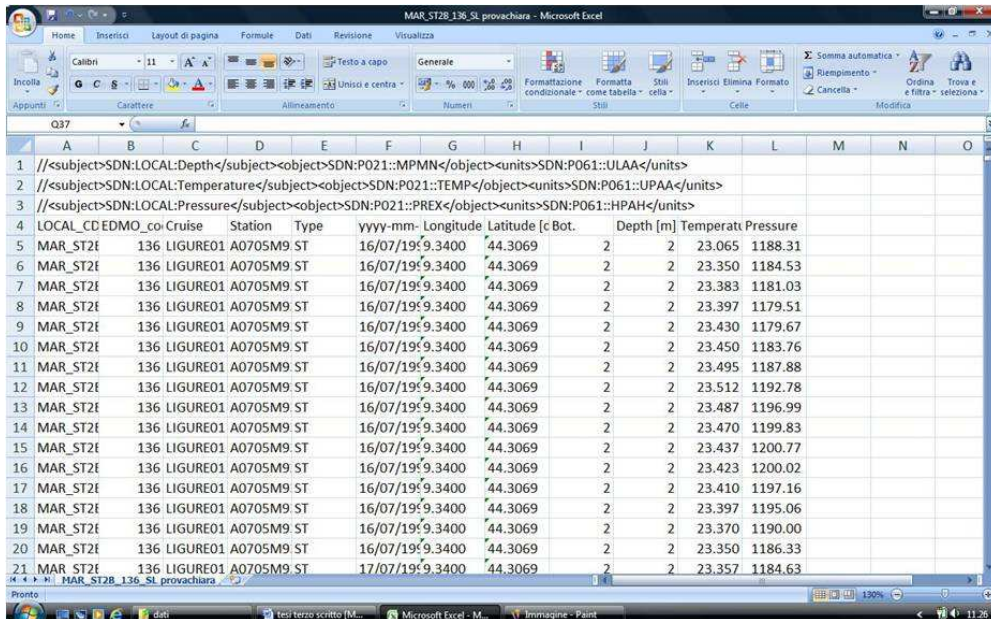


Fig. 4.1.: foglio Excel con in colonna i dati di un set, successivamente potrà essere salvato in formato .txt.

Il foglio Excel salvato in formato testo avrà questa forma (Fig. 4.2.):

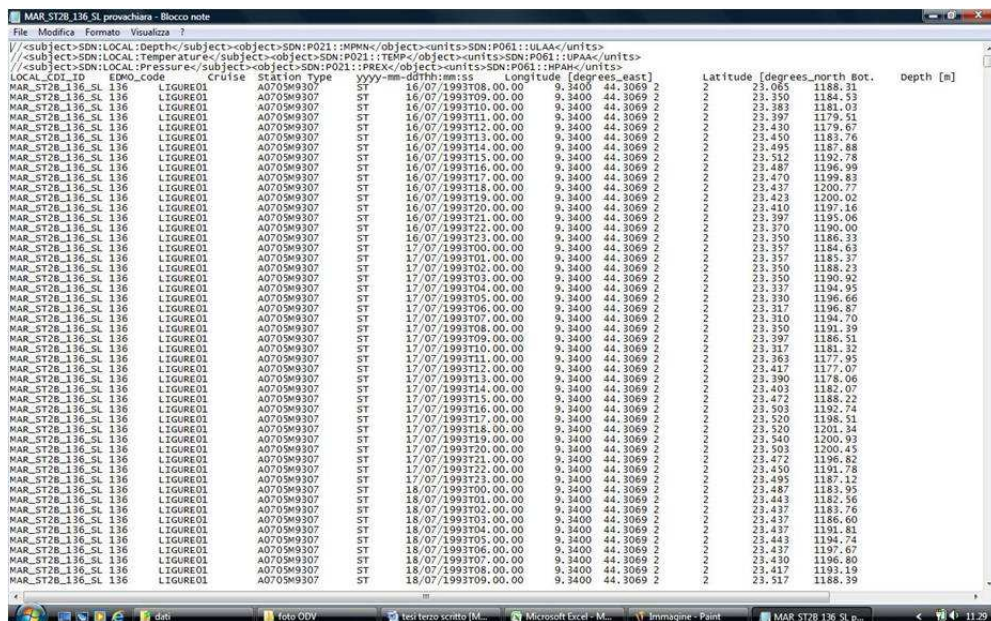


Fig. 4.2.: foglio Excel salvato in formato .txt e successivamente aperto con il programma Microsoft Blocconote.

A questo punto posso importare il file testo in ODV (Fig. 4.3.).

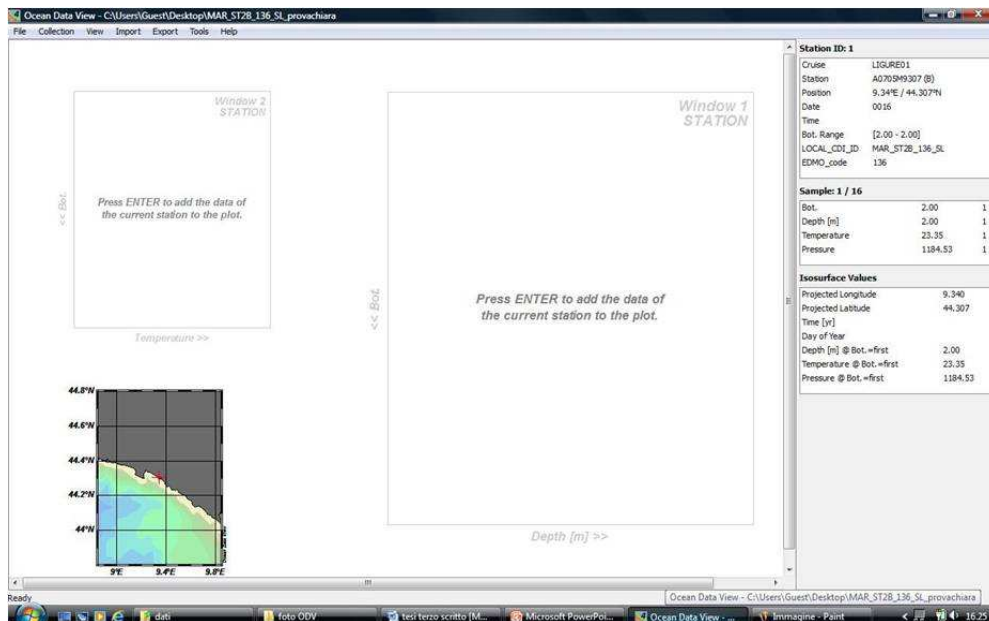


Fig. 4.3.: schermata principale per un file importato in ODV.

Nell'attuale versione di Ocean Data View è possibile importare vari tipi di formati. (Fig. 4.4.)

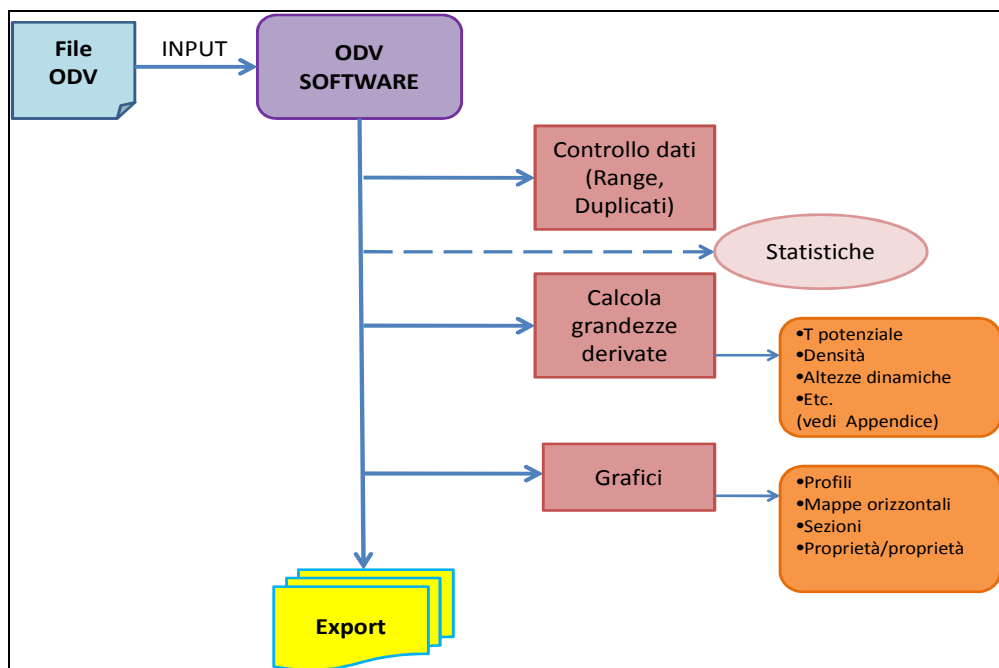


Fig. 4.4.: il diagramma descrive l'inserimento di file di diverso formato nel software Ocean Data View, cosa il software permette di visualizzare e la restituzione di file in formato ODV.

La schermata principale (Fig. 4.5.) è composta da diversi elementi, innanzitutto dal Canvas dove si posiziona sia la cartina per la localizzazione delle stazioni di campionamento sia le

due finestre dove plottare i grafici. Alla destra del canvas sono presenti le finestre descrittive della stazione e del campione.

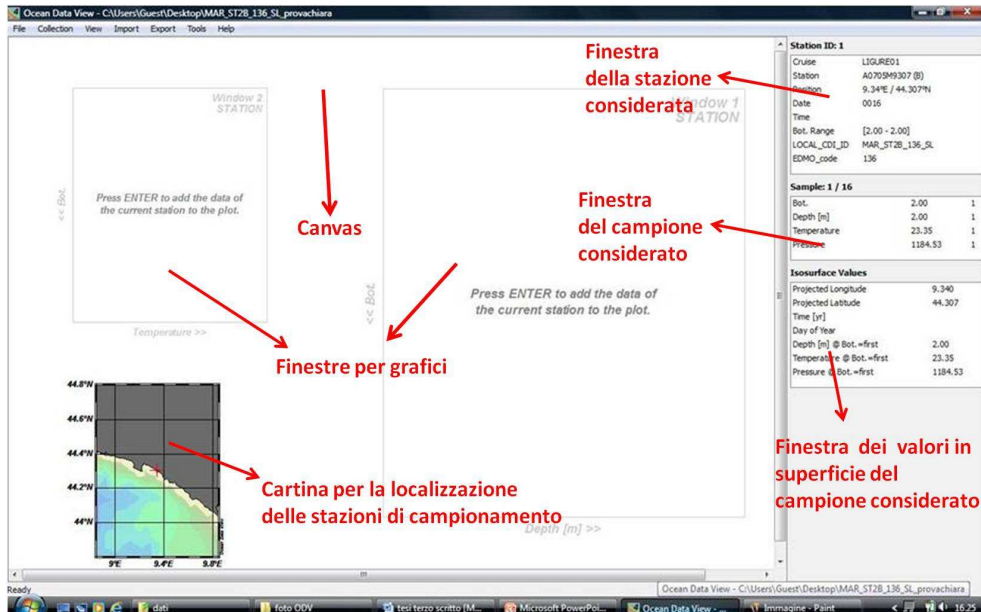


Fig. 4.5.: Canvas del software Ocean Data View.

La cartina della localizzazione delle stazioni di campionamento (Fig. 4.6.) si ottiene dal programma ETOPO, di cui si è parlato prima, e a seconda delle coordinate, latitudine e longitudine, presenti all'interno del file dà la giusta posizione. È possibile anche fare uno zoom in e out della stessa per avere una visione più precisa dei punti sulla mappa.

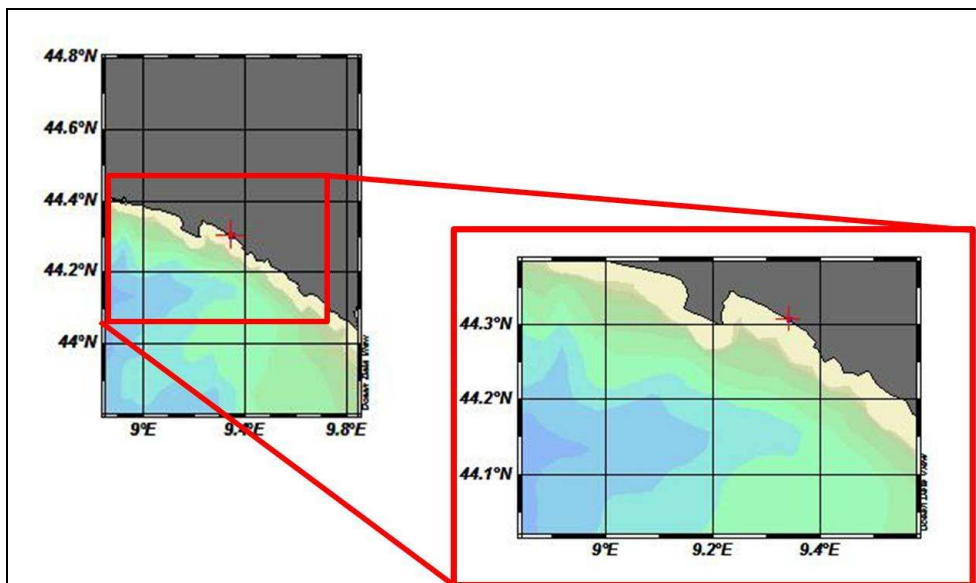


Fig. 4.6.: zoom in della cartina di localizzazione dei punti di campionamento.

Nel primo dei tre riquadri in alto a destra della schermata (Fig. 4.7.) si ha la descrizione della stazione dove è avvenuto il campionamento. Viene dato il nome della crociera, il nome della stazione, la posizione (coordinate: latitudine e longitudine), la data e l'ora in cui è stato

effettuato il campionamento (ma come nel caso della figura 4.7. possono essere anche trascurabili), il range di profondità dalla superficie al fondo, il codice identificativo per il Common Data Index e infine l'EDMO code che identifica chi ha preso il campione.

Nel secondo riquadro, invece, viene descritto il campione che si sta prendendo in considerazione; in questo riquadro si leggono i valori dei dati e le quality flags, ovvero per ogni variabile viene espressa con un numero la procedura del Controllo di Qualità. Nel caso della finestra in Fig. 4.7. il numero 1 di fianco a tutte le variabili indica che la qualità del dato raccolto è sconosciuta.

Nel terzo ed ultimo riquadro abbiamo i valori alla profondità scelta, automaticamente il programma ci darà il valore in superficie, ma gli si può impostare un'altra profondità di cui si può creare anche un grafico.

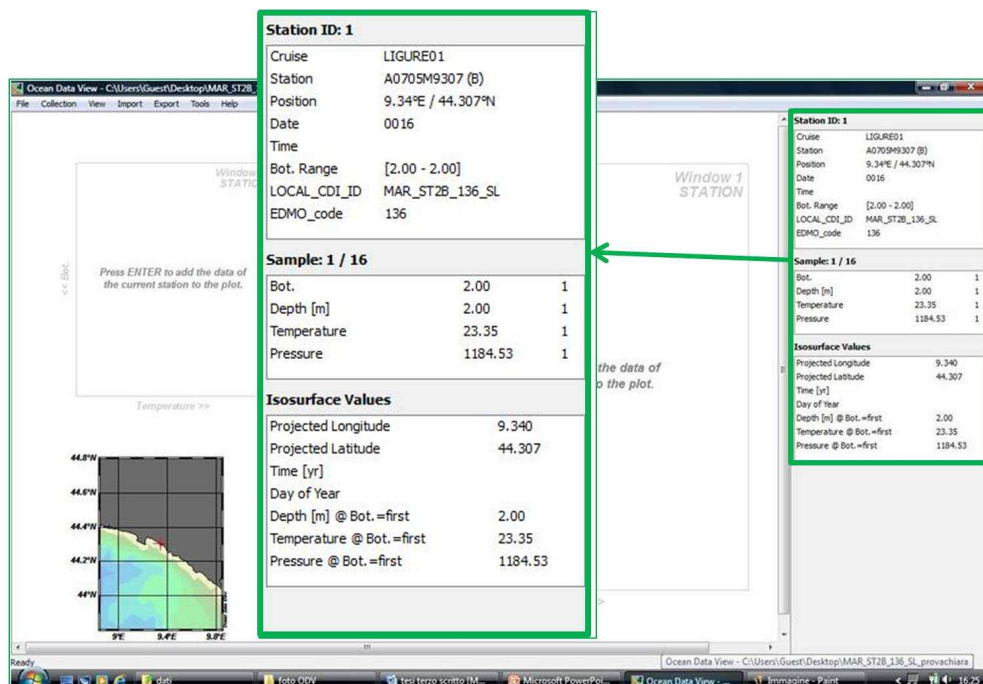


Fig. 4.7.: finestre descrittive della stazione, del campione e dei valori alla profondità scelta.

Nelle finestre per i grafici è possibile crearne diverse tipologie. Nella Fig. 4.8. vi è un esempio di MIXED windows, ogni finestra rappresenta lo stesso campione in maniera diversa e da un diverso punto di vista. Nel caso in figura è rappresentata una sezione evidenziata in rosso sulla cartina di localizzazione delle stazioni e, nell'altra cartina in alto a destra, è rappresentata la temperatura alla superficie e con un punto rosso il campione preso in considerazione; negli altri grafici è rappresentata la temperatura in funzione della profondità sia del singolo campione, sia del singolo campione all'interno del range di temperatura presente in tutto il file dati.

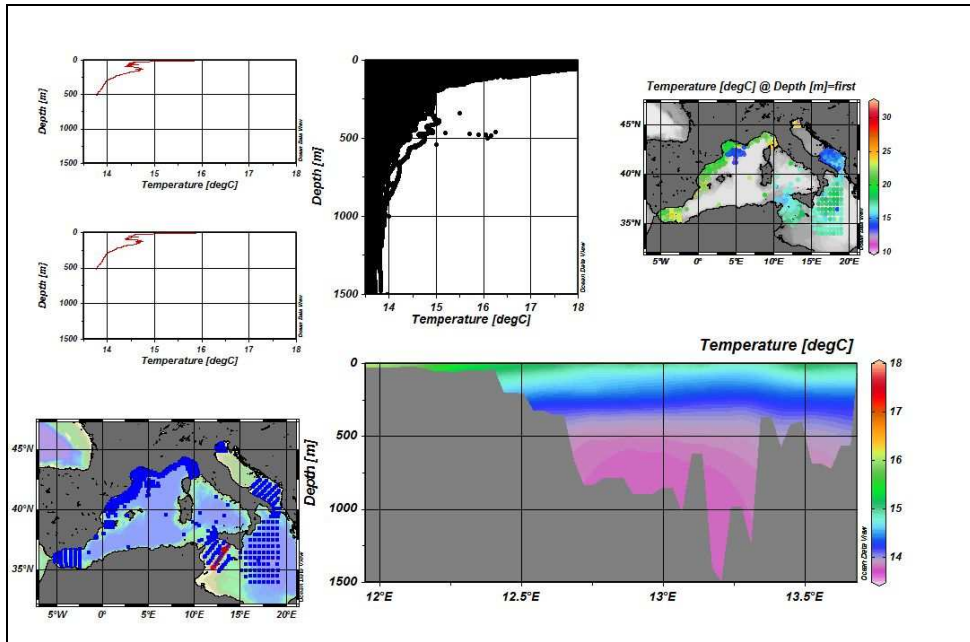


Fig. 4.8.: MIXED window in OceanDataView.

Sulla cartina (Fig. 4.9.) si può osservare la disposizione dei punti di campionamento delle varie campagne.

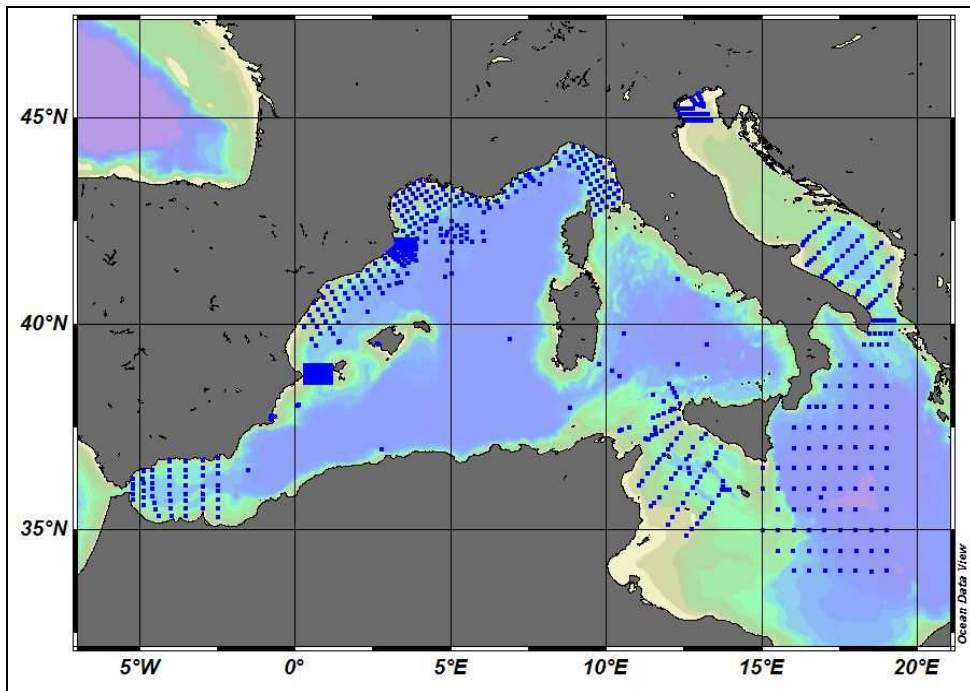


Fig. 4.9.: cartina con contrassegnati i punti di campionamento.

Con l'opzione SCATTER (Fig. 4.10.) il risultato sarà un grafico con tutto il range disponibile della variabile impostata; in figura è rappresentato tutto il range della temperatura disponibile, scegliendo un punto di campionamento all'interno della cartina il profilo dello stesso rientrerà all'interno del range.

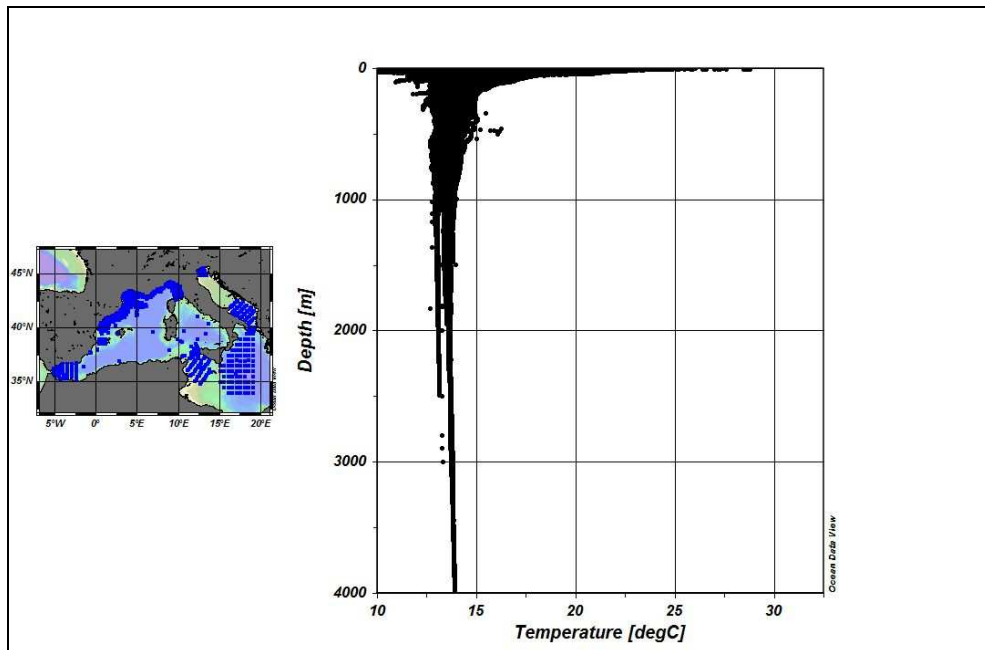


Fig. 4.10.: grafico che si ottiene con l'opzione SCATTER in cui si visualizza tutto il range della variabile scelta.

Con SURFACE (Fig. 4.11.) invece si ottengono i valori della variabile scelta in superficie. Nella figura si hanno i valori della temperatura in superficie delle aree interessate ai campionamenti.

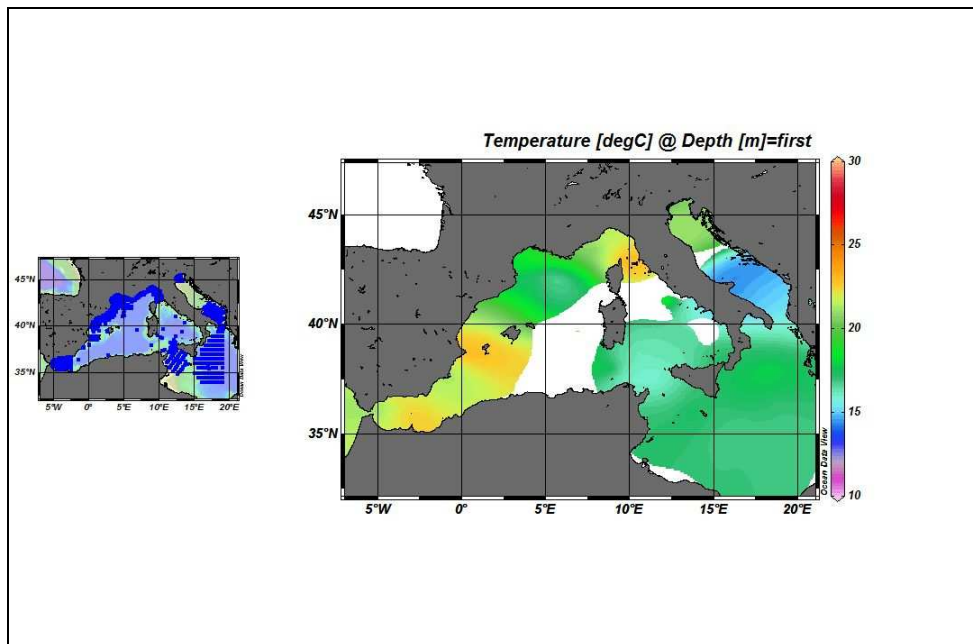


Fig. 4.11.: temperatura alla superficie, le diverse tonalità di colore indicano le aree con le diverse temperature.

L'opzione STATION (Fig. 4.12.) permette di visualizzare il profilo di un punto di campionamento selezionato sulla mappa.

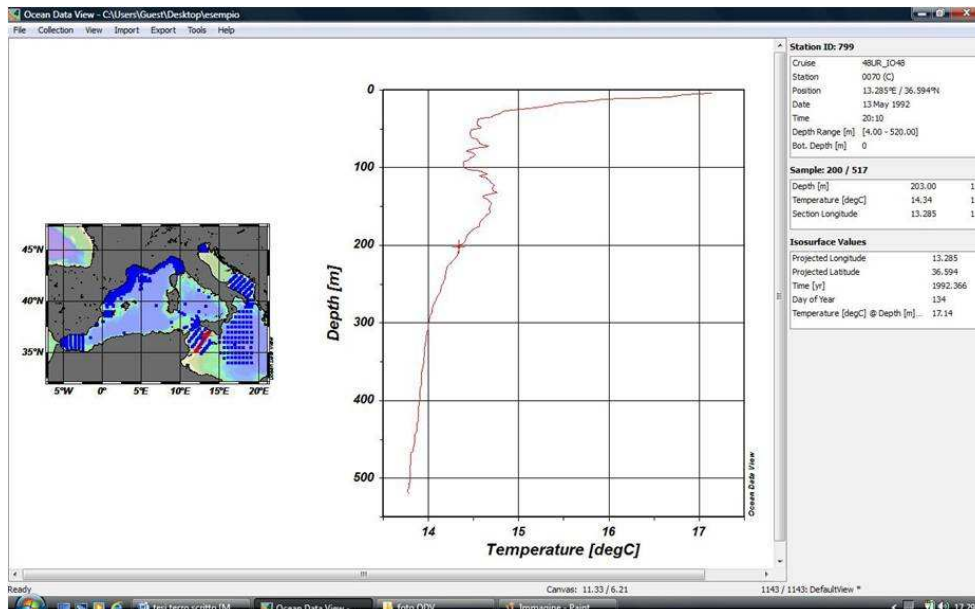


Fig. 4.12.: profilo di uno dei punti di campionamento.

Infine l'opzione SECTION (Fig. 4.13.) permette di selezionare un transetto all'interno dei punti di campionamento e di creare una sezione dello stesso in cui è visualizzata la variabile

scelta. La cartina della batimetria è fornita dal programma ETOPO. Nella Fig. 4.13. è visualizzata la temperatura lungo un transetto.

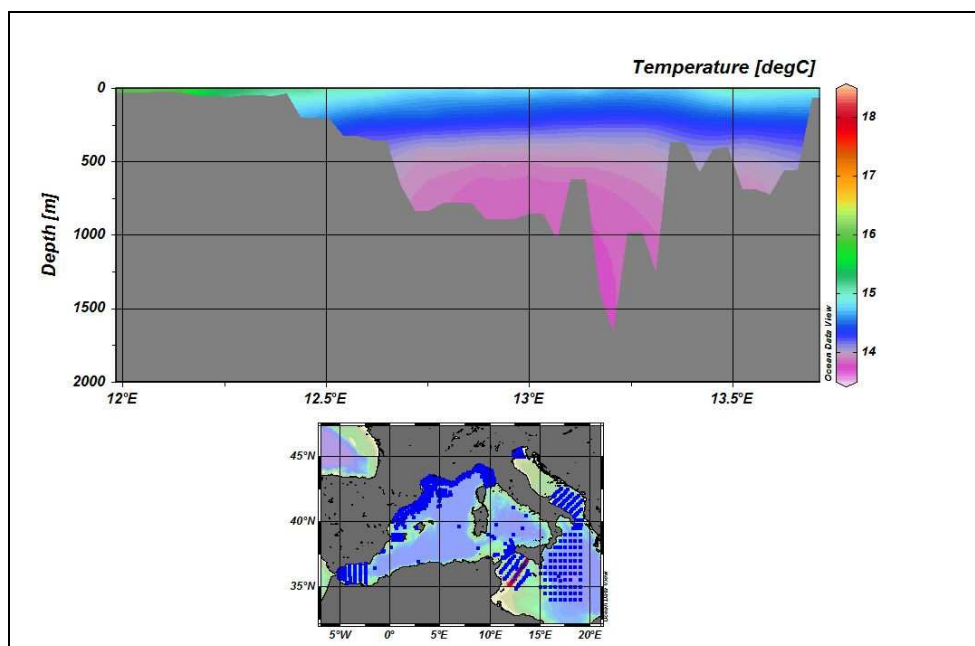


Fig. 4.13.: sezione di un transetto scelto tra i punti di campionamento.

5. Conclusioni.

Un sistema informativo che consenta di gestire, accedere e valutare dati ambientali risulta di notevole utilità per ragioni che sono state presentate nel primo paragrafo di questa tesi e che sono state sintetizzate nella comunicazione Marine Knowledge 2020:

1. ridurre i costi operativi e ritardi per chi utilizza dati marini e quindi:
 - a. Aiutare l'industria privata a competere nell'economia globale e soddisfare la sfida della sostenibilità;
 - b. Migliorare la qualità del processo decisionale pubblico a tutti i livelli;
 - c. Rafforzare la ricerca scientifica marina
2. favorire il mercato aumentando la concorrenza e l'innovazione tra gli utenti e riutilizzatori di dati relativi all'ambiente marino, fornendo un accesso più ampio di qualità controllata a dati marini coerenti;
3. ridurre l'incertezza nella conoscenza degli oceani e dei mari e così fornendo una base più solida per la gestione dei cambiamenti futuri;

Si può anche aggiungere che la disponibilità dei dati può portare allo sviluppo di servizi e quindi a nuova occupazione.

Infatti è chiaro che l'obiettivo di Lisbona di innovazione come driver per lo sviluppo di servizi commerciali basati su dati relativi all'ambiente marino sarà irraggiungibile se è vietato l'accesso ai dati raccolti da enti pubblici o il loro pagamento è irragionevole. Un vecchio studio del 2000 fatto per conto della Commissione Europea - Direzione Generale per la Società dell'Informazione, ha indicato che l'Europa ha investito € 9,5 miliardi all'anno su informazioni per pubblici servizi, di cui una percentuale significativa (37% in Francia e 57% nel Regno Unito) è stato in informazione geografica (catasto, servizi meteorologici, dati ambientali, servizi idrografici, ecc). Il valore economico di tale attività è stato stimato tra i 28 miliardi € ed i 134 miliardi € all'anno. Lo studio ha suggerito che il corrispondente sfruttamento di informazioni negli Stati Uniti era molto più alto con un fatturato annuo di quasi € 800 miliardi, ed una crescita annua compresa tra 11% e 37% dal 1995 fino al 2005. Anche se lo studio veniva fatto in altri periodi, in un momento in cui è necessario creare nuove opportunità la costruzione di infrastrutture per l'osservazione della Terra, elaborazione e accesso a dati ambientali può contribuire a far ingrandire una fiorente industria delle informazioni, creando nuove opportunità di lavoro.

Fin'ora si è visto come la ricerca in ambiente marino sia importante e soprattutto come sia difficile attuarla. Il saper raccogliere i dati in maniera corretta e la loro utilizzazione sono alla base della ricerca, ma anche un fattore importante per lo sviluppo economico. È

necessario rendere disponibili questi dati affinché chiunque possa usufruirne per studiare e capire i fenomeni marini, meteorologici e dell'ambiente più in generale, così come per la costruzione di servizi a supporto delle industrie marittime e dello sviluppo sostenibile. I dati, per essere veramente utili, devono essere sempre accompagnati dalle informazioni su di essi, i metadati. Prima di renderli disponibili i dati sono soggetti a controlli sulla qualità del dato, anche se le procedure sulla qualità cominciano già da prima della misura.

Per facilitare la ricerca e lo studio dei fenomeni naturali, nonché per divulgazione, è necessario che i set di dati possano essere scambiati tra i vari enti o organizzazioni che li hanno raccolti e soprattutto che siano “scritti nella stessa lingua”; a questo scopo la Direttiva INSPIRE della Comunità Europea impone la creazione di infrastrutture che mettano a disposizione i dati, organizzate come banche dati, ma evolute in sistema informativo. La Direttiva impone anche regole sulla struttura sia dei dati che dei metadati. Perché questi dati siano gestiti al meglio e per ottenere nella miglior maniera possibile più notizie da essi, esistono software come Ocean Data View che funzionano da banca dati.

Con le parole “banca dati” viene indicato un archivio informatico comprendente una grande quantità di dati organizzati in modo tale che possano essere gestiti tramite applicazioni software che ne consentano l'aggiornamento e la modifica in maniera semplice.

Il vantaggio fondamentale dell'organizzazione dei dati all'interno di un database consiste soprattutto nel fatto che, quando i dati acquisiti sono molto numerosi (sia come serie storiche che come tipologia di dati diversi che confluiscono nello stesso database), è possibile effettuare in maniera veloce una ricerca mirata eseguendo delle interrogazioni (query) all'interno di tutti i dati a disposizione; da queste interrogazioni, se formulate correttamente, vengono estratte delle risposte molto precise e che consentono di ridurre al minimo il tempo necessario per collegare i dati tra loro.

Mediante l'archiviazione, in maniera standardizzata, del maggior numero possibile di informazioni di tipo ecologico-ambientale, le banche dati costituiscono non solo un contenitore informativo di tutti i dati raccolti durante il monitoraggio dell'area, ma anche di quelli recuperati nel corso delle ricerche effettuate con l'elaborazione di serie storiche dei fenomeni presenti sul territorio e mira ad organizzare tutta l'informazione disponibile.

Per una corretta gestione ambientale è necessario conoscere a fondo la zona in cui si deve intervenire; conoscere le dinamiche naturali che intervengono nella zona e soprattutto i parametri chimico-fisici e biologici che la caratterizzano. In questo sono fondamentali le campagne di monitoraggio periodico e la ricerca. Tramite la raccolta dati e la successiva analisi degli stessi si possono ottenere i trend di una determinata area e cercare di capire cosa succederà in futuro.

Per tutti i tipi di ambiente, e a seconda dei parametri da analizzare esistono delle scale di valori con delle soglie che definiscono lo stato dell'area rispetto a quel parametro. L'uso dei dati attuali e storici dell'area può spiegare cosa succede e se è necessario intervenire o meno.

Ocean Data View è un ottimo strumento per la gestione dei dati, è una banca dati che non mostra elenchi di dati, ma li rappresenta in forma grafica ed essenziale; presenta in forma chiara e intuitiva le informazioni sulle campagne, sulle variabili indagate e sulla localizzazione dei punti di campionamento. Fornisce gli strumenti (diverse tipologie di grafici) per la miglior interpretazione dei dati e permette di creare modelli per previsioni; usa un formato comune e si può intervenire sui set semplicemente usando Excel. È in definitiva un software che può essere utilizzato da tutti, sia dagli esperti nel settore, sia da chi è alle prime armi. Più che una banca dati potremmo definire Ocean Data View un sistema informativo che permette la gestione di un gran numero di set di dati.

APPENDICI

APPENDICE 1. Elenco delle variabili derivate in Ocean Data View.

Variable	Comment
AOU [$\mu\text{mol/kg}$]	Apparent oxygen utilization
Brunt-Väisälä Frequency [cycl/h]	see ODV version 4.3.10 Guide
CFC-11 Saturation [%]	Warner & Weiss, Deep Sea Res., 32,1485-1497,1985
CFC-12 Saturation [%]	Warner & Weiss, Deep Sea Res., 32,1485-1497,1985
CFC-10 Saturation [%]	Bullister & Wisegarver, Deep Sea Res., 45,1285-1302,1998
CFC-113 Saturation [%]	Bu & Warner, Deep Sea Res., 42,1151-1161,1995
CH ₄ Saturation [%]	Wiesenburg & Guinasso, J. Chem. Eng. Data,24,356-,1979
CO*2 (DIC, ALK) [$\mu\text{mol/kg}$]	see ODV version 4.3.10 Guide
CO3 ²⁻ (DIC, ALK) [$\mu\text{mol/kg}$]	see ODV version 4.3.10 Guide
Day of Month (station date)	Decimal day of the Month derived from station date
Day of Month (time variable)	Decimal day of the Month derived from a time variable
Day of Week (station date)	Decimal day of the Week derived from station date
Day of Week (time variable)	Decimal day of the Week derived from a time variable
Day of Year (station date) [days]	Decimal day of the Year [days] derived from station date
Day of Year (time variable) [days]	Decimal day of the Year [days] derived from a time variable
Day Time (station time) [hours]	Decimal time of day [hours] derived from station time
Day Time (time variable) [hours]	Decimal time of day [hours] derived from a time variable
Difference from Reference Data	see ODV version 4.3.10 Guide
Dynamic Height [dyn m]	EOS801) (any reference pressure)
fCO ₂ (DIC, ALK) [μAtm]	see ODV version 4.3.10 Guide
Freezing Temperature [°C]	F. Millero, UNESCO Tech. Papers in the Marine Science, No. 28., 29-35, 1978
HCO ₃ ⁻ (DIC, ALK) [$\mu\text{mol/kg}$]	see ODV version 4.3.10 Guide
In Situ Density [kg/m ³]	EOS801) (using in situ pressure and temperature)
In Situ Temperature [kg/m ³]	EOS801) (using in situ pressure and potential temperature)
Latitude	Decimal latitude derived from station latitude
Longitude	Decimal longitude derived from station longitude
Meta Variable Value	Values of numeric meta-variables
Month of Year (station date)	Decimal month of year derived from station date
Month of Year (time variable)	Decimal month of year derived from a time variable
Neutral Density [kg/m ³]	Jackett & McDougall, J. Phys. Ocean., 237-263, 1997 (more)
OmegaA (DIC, ALK)	see ODV version 4.3.10 Guide
OmegaC (DIC, ALK)	see ODV version 4.3.10 Guide
Oxygen Saturation [%]	Weiss, Deep Sea Res., 17, 721-735, 1970
pCFC-11 [pptv]	Warner & Weiss, Deep Sea Res., 32,1485-1497,1985
pCFC-12 [pptv]	Warner & Weiss, Deep Sea Res., 32,1485-1497,1985
pCFC-10 [pptv]	Bullister & Wisegarver, Deep Sea Res., 45,1285-1302,1998
pCFC-113 [pptv]	Bu & Warner, Deep Sea Res., 42,1151-1161,1995
pCH ₄ [ppbv]	Wiesenburg & Guinasso, J. Chem. Eng. Data,24,356-,1979
pCO ₂ (DIC, ALK) [μAtm]	see ODV version 4.3.10 Guide
pH (DIC, ALK)	see ODV version 4.3.10 Guide

Potential Density [kg/m ³]	EOS801) (any reference pressure)
Potential Temperature [°C]	Bryden, Deep Sea Res.,20,401-408, 1973 (any reference pressure)
Potential Vorticity [10 ⁻¹² m ⁻¹ s ⁻¹]	Planetary potential vorticity (derived from Brunt Vaissala Frequency $Q=f/g*N^2$)
Pressure [db]	Saunders, J. Phys. Ocean., 1981
pSF6 [pptv]	Bullister, Wisegarver & Menzia, DSR-I, 49 (1), 175-187, 2002
Quality Flag Value	ASCII codes of the quality flags of any variable
Ratio	Any two variables
Revelle Factor (DIC, ALK)	(d fCO ₂ / d TCO ₂) / (fCO ₂ / TCO ₂). (details)
SF6 Saturation [%]	Bullister, Wisegarver & Menzia, DSR-I, 49 (1), 175-187, 2002
Sound Speed [m/s]	Fofonoff & Millard, Unesco Tech. Pap. in Mar. Sci., No. 44, 53 pp, 1983
Specific Heat Cp [J/(kg °C)]	F. Millero et al, J. Geoph. Res., 78, 4499-4507, 1973
Specific Volume Anomaly [mm ³ /g]	EOS801) (any reference salinity and temperature)
Spiciness π	Flament, Progr. in Oceanogr., 54 (1-4), 493-501, 2002
Time (station date/time) [yr]	Decimal time [yr] derived from station date and time
Time (time variable) [yr]	Decimal time [yr] derived from a time variable
Vertical Derivative	see ODV version 4.3.10 Guide
Second Vertical Derivative	Any variable
Vertical Integral	see ODV version 4.3.10 Guide
Year (station date)	Integer year derived from station date
Year (time variable)	Integer year derived from a time variable

APPENDICE 2. Siti web consultati.

- **British Oceanographic Data Center** (<https://www.bodc.ac.uk>)
- **European Commission INSPIRE** (<http://inspire.jrc.ec.europa.eu>)
- **European Marine Observation and Data Network** (www.emodnet-chemistry.eu)
- **Intergovernmental Oceanographic Commission** (<http://ioc-unesco.org/>)
- **Istituto Superiore per la Protezione e la Ricerca Ambientale**
(http://www.isprambiente.gov.it/site/_contentfiles/00004200/4234_relazionefinale.pdf)
- **Lockheed Martin** (www.sippican.com)
- **Marine Metadata Interoperability** (<http://marinemetadata.org>)
- **Mediterranean Forecasting System Toward Environmental Prediction**
(<http://www.ifremer.fr/mfstep/>)
- **Mediterranean Forecasting System Volunteer Observing Ships Portal**
(<http://moon.santateresa.enea.it>)
- **National Oceanic and Atmospheric Administration** (www.ngdc.noaa.gov)
- **Ocean Data View** (<http://odv.awi.de>)
- **Ocean Monitoring and Forecasting** (<http://www.myocean.eu.org/>)
- **Pan-European Infrastructure for Ocean and Marine Data Management**
(www.seadatanet.org)
- **Unidata providing data services, tools and cyberinfrastructure leadership**
(<http://www.unidata.ucar.edu/software/netcdf/>)

Bibliografia

Blanc F., Jolibois T., Loubrieu T., Manzella G. M. R., Mazzetti P., Nativi S. (2010).

MyOcean Internal Information System (Dial-P). Geophysical Research Abstracts, 12

Gambetta M., Manzella G. M. R. (2010). Implementation of quality control procedures by means of probabilistic estimate of Mediterranean Sea temperature and its temporal evolution.(Available at http://moon.santateresa.enea.it/documents/MOON-InSitu_QC_VerticalProfiles.pdf)

Keeley R., Woodruff S., Pouliquen S., Conkright-Gregg M., Reed G. Collectors to archives.

(Available at <http://www.oceanobs09.net/proceedings/pp/4C3-keeley-OceanObs09.pp.25.pdf>)

Manzella G. M. R. Information service system based on GOOGLE graphical interfaces.

(Available at http://moon.santateresa.enea.it/MOON-VOS_Information-System.pdf)

Manzella G. M. R. (2011). SeaDataNet: A Pan-European Infrastructure for Ocean and Marine Data Management. Newsletter n. 6, 1-8 (Available at

<http://www.seadatanet.org/>)

Manzella G. M. R., Iona S., Nativi S. (2010). Informatics in oceanography. Advances in Geosciences, 28, 1-2

Maudire G., Maillard C., Fichaut M., Manzella G. M. R., Schaap D. M. A. (2009).

SeaDataNet: Pan-European infrastructure for marine and ocean data management – Project objectives, structure and components. Geophysical Research Abstracts, 11

Nativi S., Mazzetti P., Santoro M., Boldrini E., Manzella G. M. R., and Schaap D. M. A.

(2010). CDI/THREDDS interoperability in the SeaDataNet framework. Adv. Geosci., 28, 17-27

Ringraziamenti

Ringrazio vivamente:

Il Dott. Giuseppe Manzella per la sua pazienza e disponibilità, nonostante i suoi numerosi impegni ha sempre trovato del tempo da dedicarmi.

Il Prof. Renzo Valloni, l'ideatore del Master, per la sua dedizione.

La Dott.ssa Roberta Delfanti, perché ci ha sempre ascoltato e aiutato anche nei momenti più duri.

I ricercatori ed il personale del centro ENEA, ci hanno accolto sempre con un sorriso.

In particolare sento di dover dire grazie a:

Jasmine, perché i nostri caffè, le nostre passeggiate e le nostre serate sono state e sono intense!

Lisa, perché mi rimprovera, ma è sempre pronta ad ascoltarmi e consigliarmi, e sono consigli preziosi!

Walter, perché è il mio migliore amico qui!

Giacomo, ... per tutto! Ed è un "tutto" veramente, veramente grande!

Eleonora, Michela, Laura, Andrea, perché ci siamo divertiti tantissimo!!!

Mamma!

Papà!