INTERNATIONAL MARINE DATA AND INFORMATION SYSTEMS CONFERENCE
IMDIS 2010

29-31 March 2010
Muséum National d'Histoire Naturelle, Paris France

Book of Abstracts

Editors

Michèle Fichaut, Vanessa Tosello

Organised by

IFREMER/SISMER

EU SeaDataNet project

International Oceanographic Commission of UNESCO

International Oceanographic Data and Information Exchange IODE

A regional oceanographic database for the Nordic Seas: from observations to climatic dataset

Alexander Smirnov, *State national center Arctic and Antarctic Research Institute (SI "AARI"), avsmir@aari.nw.ru (Russian Federation)*

Alexander Korablev, Nansen Environmental and Remote Sensing Center (NERSC), alexander.korablev@nersc.no

Observational datasets and climatologies are essential tools for the monitoring of the ocean climate change and models simulation/validation. Besides the fact that in situ observations at the sea are difficult, expensive and scarce the existing generalized datasets are still not complete and derived climatologies suffer from serious disadvantages. The sources of deficiency can be divided into three main categories – incomplete initial dataset, inadequate quality control and limitations resulting from objective interpolation. Therefore improvement of datasets for climate research should include efforts in all directions simultaneously.

Region under study located northward of 60o parallel and includes Iceland, Norwegian, Greenland and Barents Seas commonly known as the Nordic Seas (NS). This is one of the most sampled regions of the World Ocean of primary importance for shaping of the global climate due to strong air-sea interaction and intermediate and deep water formation. Current version of the database contains observations at more than 450,000 oceanographic stations (bottle, CTD, profiling floats) for 60o-82oN, 45oW-70oE geographic area since 1870 up to 2009.

Original database for the NS was compiled in the Arctic and Antarctic Institute (AARI), nearly 2.5 million stations were added form more than 30 sources. Station metadata and profiles are being permanently updated by comparison with original cruise reports, data reports and manuscripts under GODAR project.

Special software was developed for the data storage and processing. Its functionality generally follows to the main stages in the data processing and includes three main applications. First application (ODB3ALoad) is designed to produce merged quality-duplicate controlled observed level database and interpolated level database with assigned quality control flags. Merging procedure was designed to create an 'operational' observed level dataset with reduced instrumental and vertical biases. The duplicates identification algorithms allow automatic detection and metadata/profiles merging to generate complete oceanographic station composition from multiple duplicates. The application contains convertors from original formats to common database format. The second (ODB3A) represents a graphical user interface for data access, visualization and analysis including advanced data selection, data editing, import, export, modules for spatial and temporal statistics. Application supports data catalog and interactive selection from a map. Embedded module allows preparing ASCII files for objective analysis. The last application (ODB3AClimate) provides access to objectively analyzed fields and related services. It allows computing of mean climatological fields, anomalies fields and different climatologies comparison. On the whole, the applications provide complete technology for oceanographic data processing and analysis. Flexible modular structure of the applications allows quick modification and extension.

Applied quality control (QC) on observed data had several stages. It started from checking of the variables physical range. Observed depth levels were controlled by four procedures. The vertical resolution for CTD profiles was reduced with keeping original upper and last levels. Then, depth inversions were detected and unordered levels were deleted. For depth duplicates correction both automatic and expert control was used. The last algorithm applied on observed depths was the elimination of stations with number of levels for temperature or salinity less then three. After depth levels control, the profiles were sent to vertical stability check. The next applied algorithm on integrated database was the checking of variables against range of standard deviations (SD) computed for a sample selected around an observation. After the QC flags were assigned, the integrated database was ready for vertical interpolation. Applied QC algorithms were intended to preserve the regional variability and to produce a dataset which further can be used for computing of climatology fields with high spatial resolution. The QC procedures were applied on temperature, salinity and oxygen while database contains additional chemistry variables. Instrumental and vertical biases in data were reduced by excluding profiles measured by low accuracy instruments and surface only observations.

Original hydrographic stations are irregularly distributed in space and time while many practical tasks in oceanography and environmental sciences require data represented on regular grid with high resolution. Wide range of geo-statistical methods was developed to solve spatial interpolation problem. It is complicated by non-stationary behavior of oceanographic data when simple methods cannot be applied. As a result there is no common methodology accepted for data mapping in oceanography. To overcome the difficulties and reduce uncertainties most efforts were focused on production of long-term mean climatological fields while variability on different time scales was regarded as noise. We are concentrated on opposite approach where climatological fields are derived from individual grid fields with data type dependent temporal step. Three most common data types in oceanography are repeated vertical sections, horizontal fields and three-dimensional surveys with appropriate time steps. For repeated section it is equal to time difference between consecutive surveys, for horizontal fields most appropriate step is one month. Actually, selection of the data integration step is always a compromise between a depth dependent variable variability and data quantity/spatial coverage.

Geo-statistical modeling of temporally integrated data make high demands for interpolation algorithm. Error of interpolation should be controlled and not exceed natural variability. Different models were tested including ordinary kriging (OK), universal kriging (UK) and global trend modeling. While in some cases UK and global trend modeling get a good results in situations with complicated trend the methods cannot be applied. Intrinsic kriging (IK) model was utilized instead by two reasons. Trends are tested within defined neighborhood prior interpolation and model parameters are adjusted automatically. It is allow processing of large datasets in batch mode. The IK model is embedded into ISATIS commercial software that was extensively used in the study. Kriging standard deviation (KSD) accompanies a variable estimation in each grid point, depends on samples density (relative location) and represents quality of the estimation. The arbitrary selected KSD threshold was used to eliminate grid points with low quality estimations in data scarce regions.

We present an example of dataset for ocean climate change study in the NS. It contains monthly gridded horizontal fields on standard levels in June for 1900-2009 computed on 0.25ox0.5o latitude-longitude grid by means of block variant of Intrinsic Kriging model. Derived climatologies fields can be computed for any period taking into account quality of variable estimations. We use 1957-1990 as a reference period for anomalies computing since most uniform station coverage and good observation quality. Difference between 1900-2009 and 1957-1990 means shows warmer (~0.5o) and saltier (~0.1 psu) conditions for century-long averaging as result of warmer/saltier condition during the first half of 20th century and the late 1990s, early 2000s. Abnormality of each year can be estimated now. Periods with uniform anomalies pattern represent stable hydrographical regime and call for driving mechanisms. Periods with warm/salty conditions during 1958-1963, convection intensification and Atlantic Water sinking during 1967-1972 under strong atmospheric cooling, the Great Salinity Anomaly (GSA) propagation during 1976-1981, mid-1990s low salinity anomaly and generally warm/salty regime after 1998 are most consistent examples of stable regimes in the NS.

Presented methodology of a climatic dataset compilation get access not only to climatological fields representing mean state of ocean variables for a certain period but also to dynamic of spatial pattern of anomalies. Similar datasets can be constructed for any repeated observations along standard sections to access more detailed regional variations. Interpolation method can be improved by utilization co-kriging technique and three dimensional (3D) modeling. The 3D approach gives a chance to skip profiles vertical interpolation which a source of considerable errors. Since unknown amount of observation still not in the public access, more efforts are needed to collect complete initial datasets.