Argo workstation: a key component of operational oceanography

Mingmei Dong, Shanshan Xu, Qingsheng Miao, Xinyang Yue, Jiawei Lu and Yang Yang

National Marine Data and Information Service, Tianjin 300171, China

1 yupoyunhun@163.com

Abstract. Operational oceanography requires the quantity, quality, and availability of data set and the timeliness and effectiveness of data products. Without steady and strong operational system supporting, operational oceanography will never be proceeded far. In this paper we describe an integrated platform named Argo Workstation. It operates as a data processing and management system, capable of data collection, automatic data quality control, visualized data check, statistical data search and data service. After it is set up, Argo workstation provides global high quality Argo data to users every day timely and effectively. It has not only played a key role in operational oceanography but also set up an example for operational system.

1. Introduction

Operational oceanography is essential to mitigate natural hazards and was proposed in 1989. With the start of the GOOS (Global Ocean Observing System) program, the outline of operational oceanography is gradually clear. Parrilla pointed out that operational oceanography is the activity of systematic and long-term routine measurements of the seas and oceans, their interpretation and dissemination [1] in order to fulfil upcoming needs from many different sectors: industry, service, policy making, etc., besides the scientific one. In Nowlin’s opinion, operational oceanography aims to provide oceanographic information and data in a routine manner from observations and/or models for regular use [2]. Over the past decade, under the framework of GODAE (Global Ocean Data Assimilation System) [3], a significant amount of effort has been devoted to the development of global and regional operational capability. NOAA operates and keeps upgrading Hurricane Weather Research and Forecast Model (HWRF), which is a regional, dynamical TC model and became operational in 2007 after 5 years of development at the Environmental Modeling Center, in collaboration with NOAA’s Geophysical Fluid Dynamics Laboratory (GFDL) and the University of Rhode Island [4]. In Europe, European Coastal Sea Operational Observing and Forecasting System (ECOOP) aims to build up a timely, quality-assured marine service (including data, information products, knowledge and scientific advice in European coastal-shelf seas [5]. The Spanish Institute of Oceanography Observing System (IEOOS) provides quality-controlled data and information about Spanish surrounding waters and consists of several subsystems.[6] In Asia, China distributed Chinese Global operational Oceanography Forecasting System (CGOFS) in 2013, which is configured in three levels of nested grids from global ocean, open ocean to offshore[7]. The Korea Operational Oceanographic System (KOOS) was developed at the Korea Institute of Ocean Science and Technology (KIOST) to produce real-time forecasting and simulation of interdisciplinary multi-scale oceanic fields. The system became operational running in March 2012[8].
Operational oceanography critically depends on the near-real-time availability of high-quality in situ and satellite data with sufficiently dense spatial and temporal sampling. Argo is a global array of profiling floats measuring every 10 days temperature and salinity throughout the deep global oceans, down to 2000 m. Argo was initially developed as a joint venture between GODAE and CLIVAR. It has been an outstanding achievement and a second revolution in oceanography. In November 2007, Argo reached its initial target of 3000 profiling floats. More than 30 countries are involved in the development and maintenance of the array. By providing timely quantity and high-quality data, Argo has brought remarkable advances in ocean forecasting capability and became essential for operational oceanography [9-12].

In order to meet the needs of operational oceanography, we established Argo business platform, called Argo workstation. The system can not only automatically obtain Argo daily data, but it can also provide operational oceanography with marine research and forecast system after the format conversion and quality control.

2. System functional requirements
The functional design of the system needs to meet the requirements of the operational work (Figure 1 the work-flow of China Argo data center). The basic requirement of the operational system is to be stable and strong. Besides, the system needs to process the real-time data timely, to management the data effectively and distribute the Argo data to the end user in real time. The specific needs are as follows:

1. Acquire, process and quality-control Argo data automatically.
2. Provide various and timely data services to meet the data application requirements of forecasting, disaster prevention and mitigation, scientific research and so on.
3. Provide management, monitoring, and statistical information in a timely manner; provide visualization tool for manager and operator
4. Being structural and flexible; the functional module of the system can be used as part of the whole and also be used alone.

Figure 1. Work-flow of China Argo data center

3. System designs and capabilities
The software system adopts Win32 architecture, adopts standard C / C++ language as programming language, and selects Visual Studio 2010 as the project development tool. The whole system is divided into five modules: data receiving module, data processing module, automatic quality control module, visualization check module and data management module. Each module can run independently and correlate with each other to form a relatively complete management system. In addition to the data
receiving module, the rest of the modules are developed based on database, which is convenient to discharge and manage data. System function structure is shown in Figure 2:

![ArGo workstation structure](image)

**Figure 2.** Structure of Argo workstation

### 3.1. Data receiving module

In order to guarantee the timeliness of Argo data, this module scans file server every 20 seconds and downloads the data files as soon as there is a new one. The module adopts the way of multi-thread processing, and the working thread starts automatically once the computer is powered-on. The module has four working threads to complete the entire data receiving operation: searching thread, testing thread, downloading thread and processing thread.

1. **Searching thread**
   
   Searching thread is designed to search the original file server at fixed time interval which is set by the operator, the thread scans the server every 20 seconds default.

2. **Testing thread**

   Testing thread works through the search thread to obtain the server file status, then makes comparison between files in local database and the remote server to determine the list of files needing to be downloaded.

3. **Downloading thread**

   Downloading thread can download the files on the file-list produced by testing thread. The destination of the files could be set up by operators.

4. **Data processing thread**

   Data processing thread is strictly not in the data receiving module, but the data processing thread is equivalent to the data processing module message pump, which is responsible for driving the data processing module to handle the data.

### 3.2. Data processing module

The data processing module is capable of generating standard format data and uploading the data to the database for the next-step. It consists of 5 sub-modules: file consistency check module, file cycle redundancy check (CRC) module, duplicate check module, data decoding module and format conversion module.

1. **File consistency check module**

   File consistency check module reads the number of rows of data from satellites record and makes sure all the message packets for one Argo profile are received.

2. **Cyclic Redundancy check**

   Cyclic Redundancy check is used to verify whether the information of the packet is complete during data transmission. The module calculates a number from the received message according to the formula provided in the float manual to see if it is matched with the original number recorded in the message. If the data fails in cyclic redundancy check then the data will not be used.

3. **Duplicate check module**

   Duplicate check module is designed to delete the duplicate data in the same profile caused by satellite repeating transmission. It removes the duplicate part and keeps the data for the same float cycle complete and unique.

4. **Data decoding module**

   Data decoding module decodes the data according to the buoy manual, converting the hexadecimal characters into physical pressure temperature salinity and so on.
(5) Format conversion module produces two types of standard format. One is ASCII format file defined by China Argo Data Center(Figure 3). It has simple structure that is easy to read and understand. The other is the NetCDF format file required by the international Argo Data Management Group.

---

**Figure 3. Example of ASCII data.**

| CADC ARGOS DATA FORMAT VERSION 1.0 (2006) |
|-----------------------------------------|
| PROFILE HEADER ------------------------|
| PLATFORM_NUMBER = 2901615               |
| CYCLE_NUMBER = 2                        |
| DIRECTION = A                          |
| JULD = 22481.18745400                  |
| REFERENCE_DATE_TIME = 19500101000000    |
| OBSERVATION_DATE_TIME = 20110721042900  |
| JULD QC = 1                            |
| LATITUDE = 28.821                      |
| LONGITUDE = 140.735                    |
| POSITION QC = 1                        |
| N_PARAM = 3                            |
| STATION_PARAMETERS = PRES, TEMP, PSAL  |
| N_LEVELS = 86                         |
| DATA_MODE = R                         |
| DATA_CENTRE = NM                       |
| PROFILE_PRES_QC = A                    |
| PRES QC ADJUSTED QC ADJUST_ERR         |
| 1.0 1 99999.0 9 99999.0                |
| 6.0 1 99999.0 9 99999.0                |
| 17.0 1 99999.0 9 99999.0               |
| 26.0 1 99999.0 9 99999.0               |

3.3. **Automatic quality control module**

This module divides the real-time quality control of Argo profile into 19 sub-modules according to the latest Argo data quality control manual issued by the international Argo data management group. Each sub-module performs a test including the deepest pressure test, the platform identification code test, the date test, Buoyancy test, global range test, regional range test (red sea, Mediterranean), climatology test, pressure increasing test, spike test, gradient test, digit rollover test, stuck value test, density inversion test, grey list test, gross salinity or temperature sensor drift test and frozen profile test. Operator can make a flexible choice of what tests are applied and what is the parameter of the specific test through the interface shown in Figure 4.

3.4. **Visualization check module**

The manual audit module is the most important part of the whole system, and it is also the module with the highest requirements for the manual interaction. Because the auto quality control cannot detect all the errors, subjective visual inspection of the float values by an experienced operator is very necessary. The module provides flexible tools to assist operator in identifying the bad data points and setting appropriate flags for these data. The module is a multi-disciplinary integration platform, which can process variety graphics such GIS geographic database, computer graphics and remote sensing image(Figure 5). The module implements the following functions:

1. Data inquiring: Operators apply this module to select the data needing inspection. It is user-friendly and provides many to items to define the data.
2. Data display: Display all the relevant information of the float such as the location, time of observation, and the sensor and float maker information.
3. Data editing: The wrong QC flags can be manually modified and saved in this section. And Each time the QC flags change, the program automatically recalculates and rewrites the entire profile QC flags. And all the modification will be recorded in case there is a need to trace back.
Figure 4. Data quality control parameter set interface

(4) Graphics drawing and export: The module can draw various graphics such as float station map, single / multiple profiles of the temperature profile map, salinity profile map and the trajectory map of the float. All the pictures can be zoomed in and out and be exported. By clicking on the map, the relevant data will be showed in the data table.

(5) Data comparison: This module allows the operator to make comparison between data by drawing temperature and salinity profiles on the same time and the same view. Data can be selected by drawing a region directly from the map by setting center point and radiation of a circle or by specific observation time and platform number. Comparison between data change and the surrounding waters in a certain period of time will provide clear and sufficient reference information to determine whether the data is correct.

(6) Results saving: What have been done to the data or to the QC flags will be saved. The records will be saved as historical record into the database for tracing back.

(7) Data distribution: Data can be extracted into designated FTP server or specific folder as soon as the processing is finished. Users can define the format and directory structure of the data.
3.5. **Data management module**

Data management module is used to understand the status of the data and also the whole work flow, and it consists of three parts.

1) Float status monitoring: all the floats will be displayed on the screen so that users can clearly find the floats status. The background uses 2D / 3D GIS system, and updates automatically with the changing database.

2) Data query: Data query is based on float’s attribution, such as platform number, float type, delivery country, observation time, geographical location and so on. The querying results will be displayed on the map so that the user can click anywhere to view the profile’s specific information and draw graphics.

3) Query results saved: Query including float status, data and statistics could be saved and exported. The result can be saved as picture or word documents according to users. Data can be saved in ASCII or NC format, and be stored by date or by platform number.

4. **System features**

Argo workstation is designed for operational oceanography and has the following characteristics:

1) Good stability: The workstation is invoked by computer system, which means it will never shut down only if the computer is powered off. And each time Argo workstation restart, it will continue with the work before.

2) Effectiveness and timeliness: data is distributed to the end user as soon as the processing is finished.

3) High modularization: the functions of Argo workstation are divided into small independent modules, making it easy to maintain and improve the system.

4) Interactivity: The interface design is friendly, flexible and high interoperability, especially for the data browsing and manual inspection. It is extremely convenient.

5) Automation: In addition to manual review, all the other functions operate automatically without human participation, saving manpower and resources.
5. Conclusion
The Argo workstation system is operated and run by China Argo data center, realizing the real-time processing, management and distribution of the Argo data. The rapid and effective processing is the premise of deep application and sharing of information. In conclusion, the design and implementation of the system have theoretical and practical significance as well as broad application prospects.

Acknowledgement
This work is supported by the National Key R&D Program of China, (No. 2016YFC1401901)

References
[1] Parrilla, G.: Oceanografia operacional: un nuevo reto, Fisica de la Tierra, Servicio Publicaciones Univ. Complutense Madrid, Madrid, Spain, 167–179, available at: http://revistas.ucm.es/index.php/FITE/article/view/FITE0101110167A/11942
[2] Nowlin, W.D., T.C. Malone 2003 Research and GOOS Marine Technology Society Journal 37 42-46
[3] Bell, M.J., M. Lefèvre, P.-Y. Le Traon, N. Smith, and K. Wilmer-Becker 2009 GODAE: The Global Ocean Data Assimilation Experiment Oceanography 22(3) 14 ~ 21
[4] Richard M. Yablonsky, Isaac Ginis, Biju Thomas et al. 2015 Description and Analysis of the Ocean Component of NOAA’s Operational Hurricane Weather Research and Forecasting Model (HWRF) Journal of Atmospheric & Oceanic Technology 32(1) 144-163
[5] Proctor Roger, Howarth John 2008 Coastal Observatories and Operational Oceanography: A European Perspective Marine technology society journal 42(3) 10-13
[6] Elena Tel, Rosa Balbin et al 2016 IEOOS: the Spanish Institute of Oceanography Observing System Ocean Science 12(2) 345-353
[7] Wang Hui, Wan Liying, Qin Yinghao et al. 2016 Development and application of the Chinese global operational oceanography forecasting system[J] Advances in Earth Science 31(10) 1090-1104
[8] Kwang-Soon Park, Ki-Young Heo, et al. 2015 Development of the Operational Oceanographic System of Korea Ocean Science Journal 50(2) 353-369
[9] Richard D West 2002 Operational oceanography in 2001 Sea Technology 43(1) 37-37
[10] Berx, B, M Dickey-Collas, M D Skogen, Y H De Roeck, H Klein, R Barciela, R M Forster, E Dombrowsky, M Huret, M Payne, Y Sagarminaga and C. Schrum 2011 Does operational oceanography address the needs of fisheries and applied environmental scientists? Oceanography 24(1) 166 ~ 171, doi:10.5670/oceanog.2011.14
[11] Le Traon, P Y, G Larnicol, S Guinehut, et al. 2009 Data assembly and processing for operational oceanography: 10 years of achievements Oceanography 22(3) 5-69 doi:10.5670/oceanog.66
[12] P Y Le 2013 From satellite altimetry to Argo and operational oceanography: three revolutions in oceanography Traon_Ocean Sci. 9 901-915 www.ocean-sci.net/9/901/2013/ doi:10.5194/os-9-901-2013