

IMOS Data Workshop

AIMS 12 – 14 August 2008

ANMN Report

Lindsay Pender and Craig Steinberg

Introduction

On 12 – 14 August 2008, teams from AIMS, eMII and CSIRO met for a 3 day workshop to discuss various aspects of data management within the context of IMOS facilities. This report concentrates on those aspects of the workshop specifically relating to the Australian National Mooring Network (ANMN). A separate report is being prepared by eMII personnel. From an ANMN perspective, the objectives of the workshop were to review current progress, work on an agreed framework for moorings data management to be used nationally within the ANMN facility and to break down the tasks to be undertaken into manageable units with consensus on who undertakes which development task. The workshop agenda is attached as appendix A.

Progress since First Workshop

The first ANMN data management workshop was held at UNSW on 11 – 12 February 2008. At this workshop we concentrated on ensuring that instrument, sensor and deployment metadata would capture everything that it was felt would enable the data to be of value well into the future. There was also discussion on calibration guidelines and demonstrations of existing software packages (most notably the AIMS Deployment Database (field deployable Access database) and Clive Holden's Excel packages) used at ANMN nodes. Limitations of these packages were discussed in the ANMN context. The AIMS database was seen as a good starting point for capturing much of the metadata required and AIMS undertook to develop the package for the wider ANMN community.

Since then, AIMS have extended the deployment database so that it now interfaces with Matlab (for data download tools) and a centralised Oracle database. Work is progressing to map deployment database metadata to centralised database fields, and to output this data to SensorML for entry into MEST.

The AIMS Data Centre have also made significant progress with their real time processing and web viewer for handling real time data transmitted from their automatic weather stations (AWS – now operational <http://www.aims.gov.au/docs/data-centre/weatherstations.html>) and GBROOS sensor network (FAIMMS). This system was demonstrated at the workshop and it was agreed that it represented a good approach to handling ANMN real-time data.

Data Flow

At the workshop, the data flow for both the delayed mode, e.g. sub-surface instruments, and real-time systems was broken into blocks utilising where ever possible existing packages, and clearly identifying tasks still to be undertaken.

The delayed mode data flow is shown in figure 1. To the left is a simplified view of the structure contained within the Deployment Database. It is intended that this database capture all metadata (both from the perspective of MEST and node management) required for all instruments, their individual sensors, calibration and maintenance, mooring design and deployment details. There is mapping of this metadata into a centralised Oracle database which then maps to MEST via SensorML. It is envisaged that ultimately eMII will host the Oracle database currently under development by AIMS. The moorings metadata within MEST will have a hierarchical structure containing at the top level, mooring location information, down through deployment details, instruments and sensors. The detail of how actual data are linked to the sensor metadata is still to be finalised.

Although packages for taking the data from individual instruments currently exist, there is no common and consistent system available to handle all ANMN instruments. The aim is to have, where appropriate, common tools producing data products in a common format for all instruments and to be used by all ANMN nodes. The detail of QC and flagging standards, and data formats is still to be finalised.

There was also discussion on the provision of valued added data products, e.g. tidal removal, alternate data formats, web based data viewers and disseminators etc., however, no specific recommendations were made.

The real-time data flow is shown in figure 2. This system is to utilise the same Deployment Database as used for the delayed mode system and associated metadata will follow the same path. Given the similarity between the requirements for the AIMS AWS and FAIMMS systems, and the real-time component of the ANMN, it was decided that the ANMN should utilise, where ever feasible, the AIMS system. AIMS have agreed to make this system available to the wider ANMN community. There was discussion as to the feasibility of eMII hosting a common automated system or whether it was more appropriate for each node to host their own. General feeling was that it would be better for the system to be deployed regionally, with ownership and development remaining with AIMS. As with the delayed mode system, detail of QC and flagging standards, and final data formats are yet to be determined.

On-going and Future Work

- Finalise first ANMN version of the Deployment Database (AIMS).
- Finalise mechanism for the upload of the centralised Oracle database metadata to MEST (AIMS).
- Decide on QC flagging standards to be adopted (CSIRO with ANMN feedback).

- Decide on final QA/QC procedures to be adopted (CSIRO with ANMN feedback).
- Decide on netCDF conventions to be adopted for the standard data product (CSIRO with ANMN feedback).
- Further development of the AIMS real-time automatic processing and data dissemination system for use by ANMN (AIMS).
- Development of the delayed mode data processing system utilising where appropriate existing algorithms (CSIRO).
- Development of web based data viewing tools. (AIMS and eMII).

Recommendations

The following is a summary of the recommendations arising from the workshop.

- The AIMS Deployment Database is further developed and deployed throughout the ANMN facility.
- The AIMS Oracle database with links to the Deployment Database and MEST is further developed and deployed centrally within eMII.
- The AIMS real-time data acquisition and processing system as used by the AWS and FAIMMS is further developed and deployed regionally where required.
- CSIRO investigates QC and netCDF standards and makes a recommendation to the ANMN community.
- CSIRO with assistance from eMII undertakes the development of the delayed mode data processing system.
- The principle developer of the data processing system is to assist ANMN nodes with training and implementation.
- Lindsay Pender and/or Simon Allen to investigate options for gaining assistance (financial or personnel) from eMII for undertaking the above two dot points.

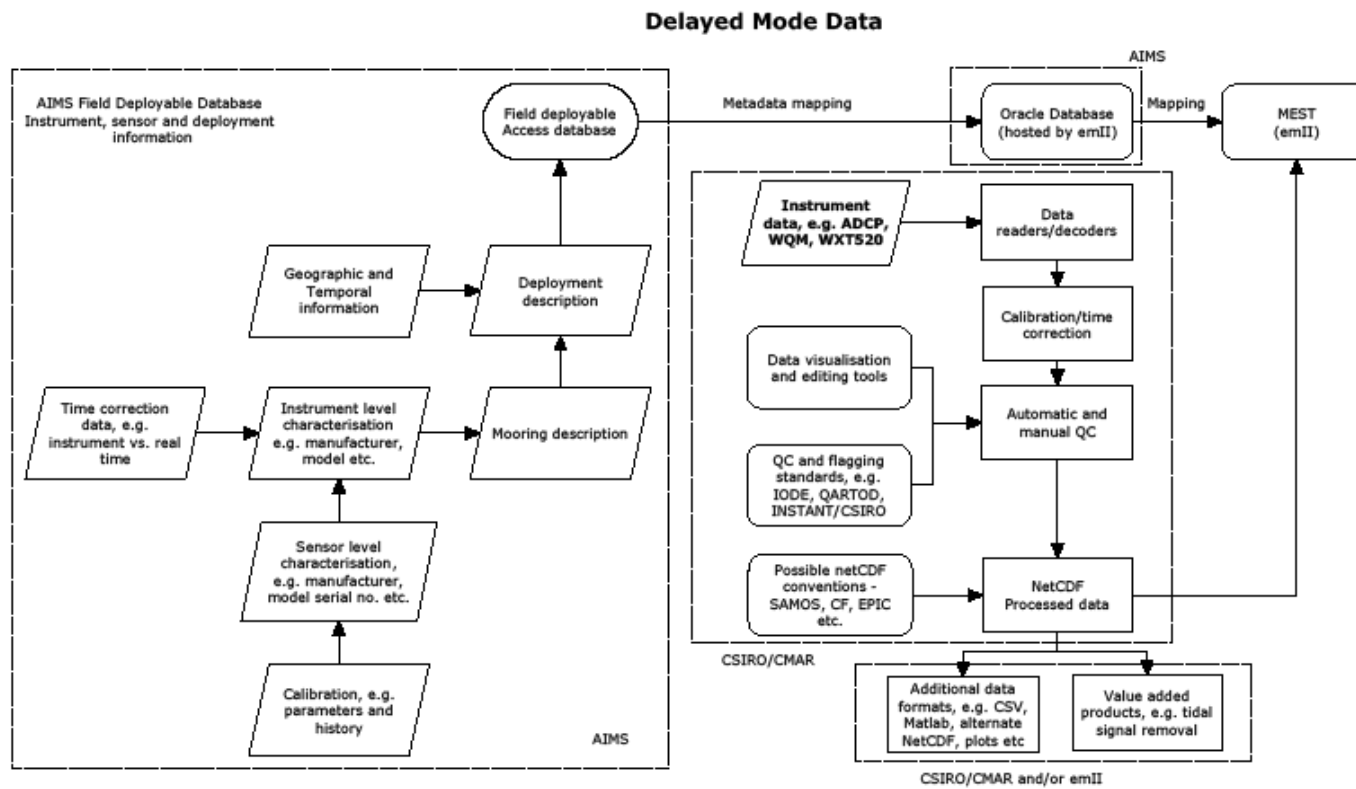


Figure 1.

REAL Time Data

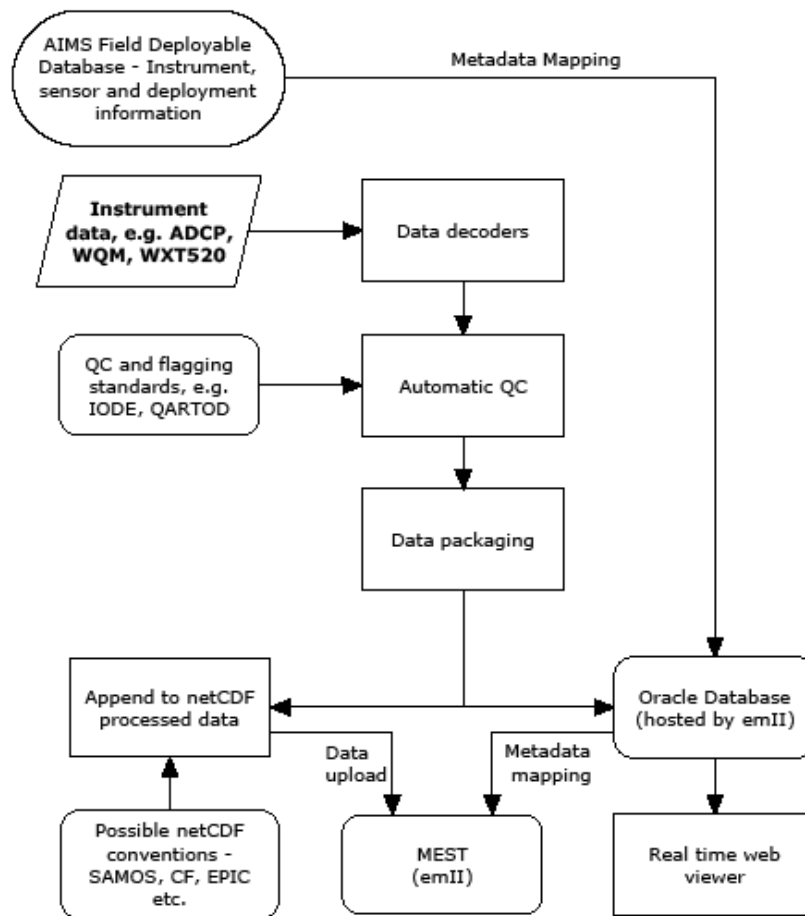


Figure 2

Appendix A

IMOS Data Workshop

AIMS August 12-14 2008

Agenda

Tuesday 12 August

0930	Welcome	Scott Bainbridge
1000	AIMS Data Centre capability	Mark Rehbein
	GBROOS Moorings data processing	Craig Steinberg
	eMII overview	eMII

Discuss structure for the remainder of the visit:

- ANMN issues
- eMII discussions with IMOS sub-facility data custodians

1200 Lunch

1300 Discuss ANM Moorings work plan:
Options for breaking the project down into manageable blocks.
Produce a requirements/specification document for each block.
Decide what software already exists and how it can be utilised.
Decide who will undertake the specific developments required.
Decide how we will manage data before the full system is implemented.

Decide on the following:

Metadata requirements
QC standards
QC processing what tests, results, outcomes
Resourcing

1630 End of day

Wednesday 13 August

0900 Breakout to meet with data custodians of the IMOS sub-facilities within GBROOS

Thursday 14 August

0900 Breakout to meet with data custodians of the IMOS sub-facilities within GBROOS

1500 Seminar DR JACK A GILBERT, Plymouth Marine Laboratory
Western Channel Observatory - A UK Marine Observation System

Attendees

Lindsay Pender	ANMN CSIRO
Roger Proctor	eMII (Apologies)
Beth Strain	eMII
Regina Magierowski	eMII
Katherine Tattersall	eMII
Scott Bainbridge	GBROOS Project Leader & FAIMMS
Craig Steinberg	GBROOS ANMN (Moorings) sub-facility leader
Felicity McAllister	GBROOS Moorings Database and processing
Irena Zagorskis	AIMS SOOP

Miles Furnas	AIMS SOOP
Mike Mahoney	AIMS Remote Sensing & SOOP
Mark Rehbein	AIMS Data Centre (ADC)
Alex Hendry	AIMS real-time weather station QA/QC& GBROOS
Valonna Baker	ADC MEST
Kevin Gunn	ADC
Greg Coleman	ADC LTM
Gavin Ericson	ADC
David Crossman	ADC