CEOS Water Portal Project

Implementation Plan

March 10th, 2015

Version 2.0

Revision History

Version	Date	Page	Reason	Comment
1.0	Mar 2, 2012	-	(v1.0 released)	
1.1	Feb 2,2013	6	Update to incorporate new data partners	
		10	Update of schedule	
1.2	Sep 4, 2013	4	Update to incorporate new data partners	
		6	OpenSearch client capability added	
		7	Update to incorporate new data partners	
		11	Update of schedule	
2.0	Mar 10, 2015	5	Update of water portal introduction	
		6	Fig.3.1-3 CEOS Water Portal Architecture added	
		7	Update to incorporate new data partners	
		11	Update of schedule	

APPENDIX

1 Introduction

For the past several years in CEOS WGISS, WGISS Test Facility for CEOP (WTF-CEOP) project was tackled with the lead of JAXA to meet the demand of CEOP user community to develop the data integration services. With the fulfillment of the task, the project ended in 2009.

The WTF-CEOP successfully having demonstrated the advantages of data integration for CEOP, the request for enhanced services for water communities was motivated by Dr.Toshio Koike, the lead scientist of CEOP, to push the data integration forward and improve the accessibility to the hydrological data, or water relevant data in broader sense, which are distributed globally.
JAXA has taken the roll to expand the current "WTF-CEOP" prototype system into "CEOS Water Portal ", in which more variety of datasets and features are to be explored. Another essential element of consideration is that of the use by not only scientists but also by non-researchers who are dealing with those data in their daily work, such as river administrators etc.

The lead organization representing the CEOS WGISS agencies in the CEOS Water Portal Project is JAXA, with participation by NASA and NOAA. Other WGISS members and WGISS affiliates may join at a future date. The project lead is Satoko Miura, JAXA.

1.1 Purpose of the Implementation Plan

The purpose of this Implementation Plan is to provide the background needed to understand the CEOS Water Portal Project, and to describe a plan for organizing, developing and executing the project.

1.2 Purpose of CEOS Water Portal Project

The purpose of the CEOS Water Portal Project is to provide assistance to the water relevant scientists and non-researchers (e.g. river administrators) in the development of data services associated with data integration and distribution.

2 Water Communities and Data Partners

Water communities are one of the main target users of the portal, and they also collaborate with the portal as data partners which provide the portal open access to their data archives.
Some of the primary water communities involved are listed in this chapter.

2.1 CEOP

The Coordinated Energy and Water Cycle Observation Project (CEOP) was initially motivated by the World Climate Research Program (WCRP) Global Energy and Water Cycle Experiment (GEWEX) international efforts focusing on the measurement, understanding and modeling of water and energy cycles within the climate system.

CEOP data are being archived at 3 archive centers: in-situ data at UCAR (US); satellite data at University of Tokyo (Japan); and model output and MOLTS data at MPI (Germany).

CEOP will achieve a multi-scale framework for its observation and data collection activities by identifying three different scales.

Reference Sites: Well-instrumented locations of small to intermediate scale areas (104 km^2 or less) distributed around the globe in different climatic regimes will provide the data needed on a mesoscale or smaller scale for research in land area and hydrology processes and model validation. There are 35 CEOP geographically distributed Reference Sites.

Mesoscale Regional Areas: Large scale areas ranging from the size of a sub-area of a Continental Scale Experiment (>105 km^2) to a whole continent and surrounding areas as a function of the research objectives in the specific region.

Global Scale: Includes all land and ocean areas of the Earth. These multi-scale, disparate data sets are collected and organized by several data management centers.

Schedule: The observation and data collection phase for CEOP phase-1 extends from 1 July 2001 to 31 December 2004. The implementation of this phase is divided into four Enhanced Observing Periods (EOPs). The four periods are designed to start at a relatively low level for EOP-1 as an enhanced seasonal observing period focusing on a selected set of reference sites 1 July to 30 September 2001. EOP-2 from 1 October 2001 to 30 September 2002 entails a coordinated “Buildup Period” in which CEOP participants begin to make contributions as their capability for model output and satellite data is implemented. The primary focus is on the collective 2-year data set beginning with EOP-3, starting in October 2002, which covers the first of two annual cycles with emphasis on a data set suitable for a synoptic climatology case study. EOP-4 covers the second annual cycle with provisions for some intensive water and energy-cycle experiments using coordinated Intensive Observing Periods (IOPs) as part of the major activities.

CEOP has now entered a phase 2 (2007.1.1-)for further development. Phase 1 (through 2004) had an emphasis on observation and data development (collecting, archiving and preliminary analysis of CEOP data). Phase 2 will have an emphasis on research and system expansion to meet commitments to the CEOS/IGOS-P Water Theme, WCRP/COPES and to the Global Energy and Water Cycle Observation System of Systems (GEOSS).

2.2 Asian Water Cycle Initiative (AWCI)

The Asian Water Cycle Initiative's goal is to better understand the mechanism of Asian water cycle variability and to improve its predictability, and further more to interpret the information applicable to various water environments in different countries, then to help the mitigation of water-related disasters and the efficient use of water resources.

The Asian Water Cycle Initiative has been established in cooperation with national governments, institutes and research communities in Asia.
The objectives for AWCI are defined as follows:

- to develop Integrated Water Resources Management (IWRM) approaches;
- to share timely, quality, long-term information on water quantity and quality, and their variation as a basis for sound national and regional decision making;
- to construct a comprehensive, coordinated and sustained observational system of systems, such as prediction systems and decision support capabilities, under the GEOSS;
- to develop capacity building for making maximum use of globally integrated data and information for local purposes as well as for observation and collecting data.

AWCI has formally agreed to share the data with CEOP from all the Basin sites that are participating.

2.3 The Global Terrestrial Network – Hydrology (GTN-H)

The Global Terrestrial Network – Hydrology (GTN-H) links existing networks and systems for integrated observations of the global water cycle.

The main objective of the GTN-H is to make available data from existing global hydrological observation networks and to enhance their value through integration. GTN-H thus underpins the generation of datasets suitable for:

1. Research in the areas of global and regional climate change
2. Environmental Monitoring
3. Hydrology and water resource management

The network was established in 2001 as a network of networks to support a range of climate and water resource objectives, building on existing networks and data centers, and producing value-added products through enhanced communications and shared development.
Network partners include :

· The City College of New York/City University of New York
· European Space Agency Soil Moisture and Ocean Salinity (ESA SMOS) mission
· FLUXNET
· Food and Agriculture Organization of the United Nations (FAO)
· Global Environment Monitoring System - Water (GEMS/Water)
· Global Precipitation Climatology Centre(GPCC)
· Global Runoff Data Centre (GRDC)
· International Atomic Energy Agency/Global Network of Isotopes in Precipitation (IAEA/GNIP)
· International Data Centre on the Hydrology of Lakes and Reservoirs (HYDROLARE)
· International Groundwater Resources Assessment Centre (IGRAC)
· US National Climatic Data Center (NCDC)
· US National Snow and Ice Data Center (NSIDC)
· US World Glacier Monitoring Service (WGMS)

The GTN-H is a joint project of the Global Climate Observing System (GCOS), the World Meteorological Organization / Climate and Water Department (WMO/CLW), and the Global Terrestrial Observing System (GTOS).

2.4 The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)

CUAHSI defines its mission as :
CUAHSI enables the water science community to advance understanding of the central role of water to life, Earth, and society. CUAHSI focuses on water from bedrock to atmosphere, from summit to sea and from the geologic past, through the present and into the future.

CUAHSI will support the community to advance water science and to improve societal well-being by . . .
- developing, supporting, and operating research infrastructure;
- improving access to data, information and models;
- articulating priorities for community level water-related research and observations;
- facilitating interactions among the diverse water research community;
- promoting interdisciplinary education centered in water science; and
- translating scientific advancements into effective tools for water management and policy.

3 Project Details

3.1 JAXA's approach to CEOS Water Portal Project

3.1.1 WTF-CEOP

Former WTF-CEOP task team within CEOS WGISS, in partnership with the CEOP science community, tailored and developed tools to access over the internet the various data collections with the data services needed to support data integration. To meet various CEOP science objectives, the CEOP science community required data integration services that allow it to access and inter-compare diverse data types from multiple sources. The WGISS agencies' prototypes offered a variety of capabilities towards this goal.

The JAXA team developed a prototype distributed data integration system to provide user-friendly services for discovery, analysis and visualization of CEOP (in-situ, satellite and model output) data to water cycle research scientists globally. The prototype system provides users with menus for selection of data and services through a standard web browser, and for access to the CEOP data from the 3 CEOP data archive centers. The system is distributed in the sense that while the main server is located in Tokyo, the data is located in archive centers, which are globally distributed.
The system is integrated in multiple ways. First, while the data may be created and archived in a variety of formats the system masks the data formats and presents the data in a uniform style. Second, the system knows the geolocation and time of all the data sets, and coordinates the selection, retrieval and display of the various types of data both temporally and geospatially. Thirdly, the system supports selection of the data through a uniform set of menus, by data type, data providing agency, reference site and station, and supports sub-setting according to time, area and height/depth. The system provides user-friendly services to view plots and graphs of the data, to view data values on the screen, or to download data onto the user's (local) computer.

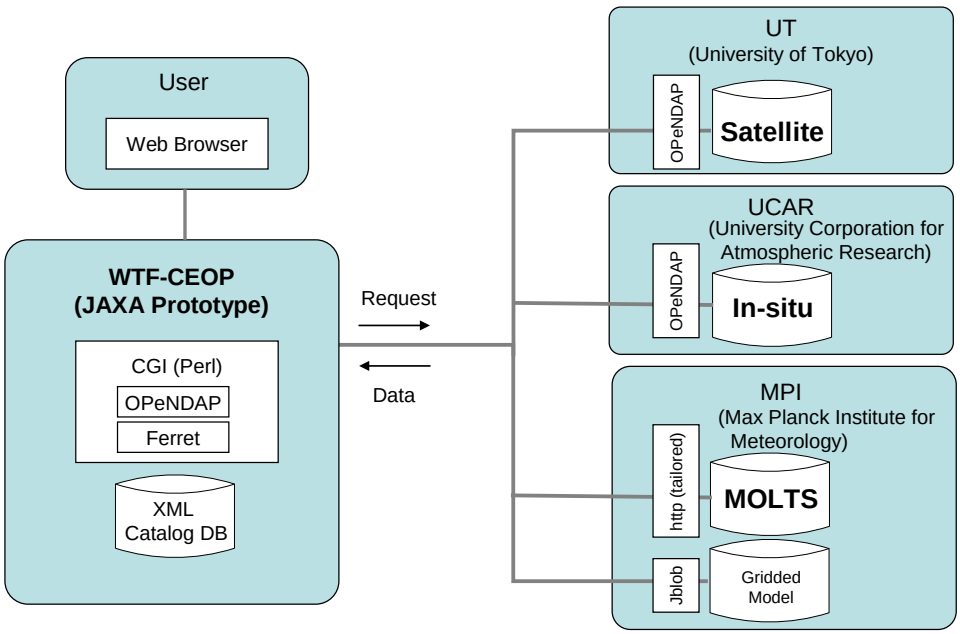

Fig.3.1-1 Diagram of WTF-CEOP JAXA Prototype and CEOP Archive Centers

3.1.2 Transition to the CEOS Water Portal

JAXA's portal concept leverages their WTF-CEOP system which has been implemented with data integration features required by CEOP.

3.1.3 CEOS Water Portal

The purpose of the CEOS Water Portal Project is to provide assistance to scientists and general users (or non-researchers) from the water domain in the development of data services associated with data integration and distribution.

The CEOS Water Portal is built around several key concepts:

- The Water Portal will make different types of datasets (satellite, in-situ, and model output) from various water communities accessible.
- Portal users will include researchers in the hydrological domain as well as operational workers in water-related fields, such as river administrators.
- The portal is connected to each data center using a standard interface protocol for real-time retrieval of the catalog and the data. (Part of the catalog is collected and stored in the portal locally.)
- User registration and sharing of use cases/research summaries is encouraged.

With that said, the concept of CEOS Water Portal is illustrated below.

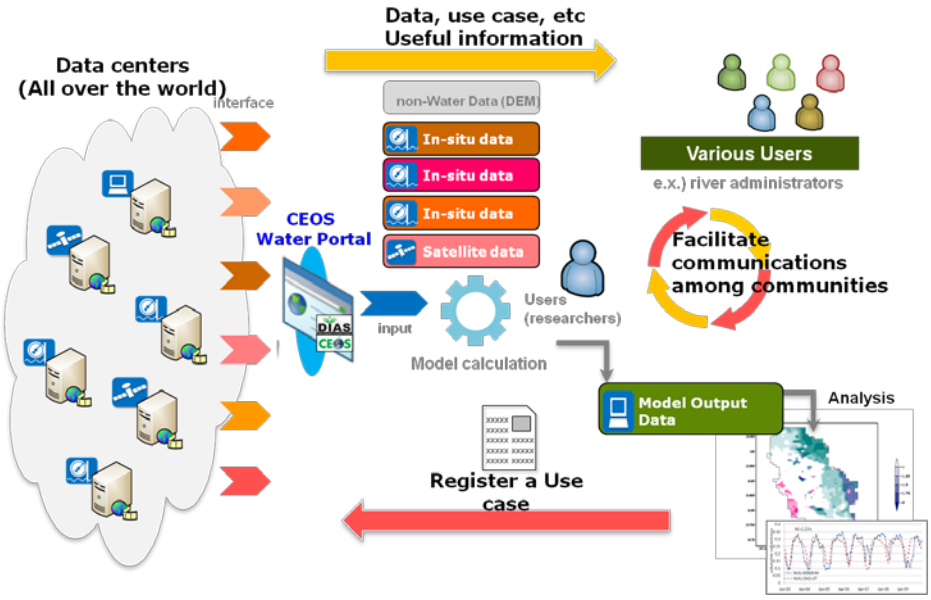

Fig.3.1-2 Concept of CEOS Water Portal

Search and access model in CEOS Water Portal

The catalog of CEOS Water Portal focuses on leveraging an open source catalog broker software (GI-cat) as well as outside catalog servers (or brokers), which will greatly simplify the total architecture and reduce the labor of catalog management without sacrificing service quality for users.
The portal also has a capability to search catalog through a web service, OpenSearch. There are more catalog web services available than before that enable a client to access their catalog in a machine-to-machine manner. This increases the number of available data exponentially in the water portal.

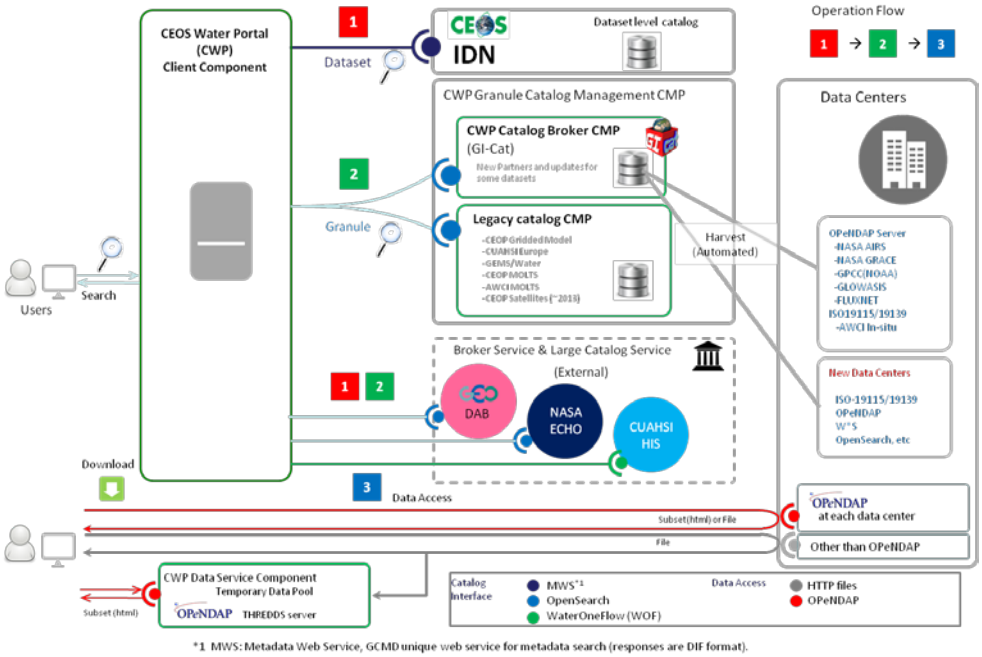

Fig.3.1-3 CEOS Water Portal Architecture (Interface View)

Collaboration with Data Partners

Data Partners and the portal are collaborated in the way that datasets archived at the distributed data servers can be searchable and accessible from the portal.
The portal and servers of data partners are connected by application interface by the use of standard interface protocol, such as OPeNDAP.
The following table lists Data Partners with corresponding data types, physical server locations, and interface methods.

Table 3.1-1 Data Partners

Data Partners	Data Types	Server Locations	Interface Protocol
CEOP	Satellite	University of Tokyo (Japan)	OPeNDAP
	Model (MOLTS)	MPI (Germany)	Jblob
	Model (Gridded)	MPI (Germany)	OPeNDAP
	In-situ	UCAR (USA)	OPeNDAP
AWCI	Satellite	University of Tokyo (Japan)	OPeNDAP
	In-situ	University of Tokyo (Japan)	OPeNDAP
	GIS	University of Tokyo (Japan)	TBD
NASA	Satellite	NASA GSFC (USA)	OPeNDAP, ftp, http
NOAA (GPCC)	In-situ	NOAA (USA)	OPeNDAP
NOAA NCDC	In-situ	NOAA (USA)	OPeNDAP
GEMSWater(GTN-H)	In-situ	Canada	WFS
GLOWASIS	Satellite	Europe	OPeNDAP
	In-situ	Europe	OPeNDAP
	Model	Europe	OPeNDAP
CUAHSI	In-situ	USA	WOF SOS
GEO DAB (GRDC)	In-situ	Unknown	OpenSearch, SOS
GRDC(GTN-H)	In-situ	Germany	TBD
PC Portal	Satellite	NASA (USA)	TBD

Contribution from WGISS Agencies

(1) NASA
NASA contributes to the water portal by providing open access to their data through OPeNDAP servers.

NASA has AIRS/Aqua Level 3 data on the OPeNDAP server at GSFC which has been running for the past years. The AIRS/Aqua Level 3 datasets include the following.

- AIRX3STM (AIRS/Aqua Level 3 monthly standard physical retrieval product (Without HSB))
- AIRX3STD (AIRS/Aqua Level 3 daily standard physical retrieval product (Without HSB))
- AIRX3ST8 (AIRS/Aqua Level 3 8-day standard physical retrieval (AIRS+AMSU))

Also, following datasets are also being served from OPeNDAP or THREDDS servers and included as NASA's contribution.

- Fluxnet Marconi Conference GAP-Filled Flux and Meteorology Data, 1992-2000 (NASA ORNL DAAC)
  Details on this dataset can be found at :
  http://daac.ornl.gov/FLUXNET/guides/marconi_gap_filled.html

- GRACE Level 3 Data (NASA/JPL PO.DAAC)
  Details on this dataset can be found at :
  http://gracetellus.jpl.nasa.gov/

(2) NOAA
NOAA has been providing scientists with GPCC (Global Precipitation Climatology Center) data from their OPeNDAP server in their own project. As it was revealed that the GPCC data among other GTH-H data was on the top list of users' data request, the collaboration was discussed and NOAA agreed to become a data partner with the portal.
The GPCC data are described in the table below.

Variable	units	Statistic
Precipitation	mm/month	Mean (1.0x1.0 combined)
Precipitation	mm/month	Mean (1x1 full data product)
Precipitation	mm/month	Mean (0.5 full data product)
Precipitation	mm/month	Mean (1x1 monitoring)
Precipitation	mm/month	LTM (0.5 full monitoring)
Precipitation	mm/month	LTM (1.0 full monitoring)

3.2 Functionality and Use Case

The portal system aims to provide one-stop-service and access to a variety of hydrological and water relevant data and also supports data integration.
Data integration is essential to users, as they need to access multiple types of data such as in-situ data, satellite data and model output data from many sources. By aggregating the multiple types of data, the data will have more value than they did individually.

Development of the portal (and the former WTF-CEOP) system has primarily been driven by the concepts and requirements of data integration, which can be described as a series of aspects. The most basic aspect is that the various types of data are in common and compatible data formats, data organization and structure and file naming conventions. The next aspect is that the data presentation, structure and formatting will support comparing the data in space (i.e. latitude, longitude and height) and time.

The system of services is required to be flexible, to adapt to multiple and changing user tasks and projects. The way a user will use these services and apply it to the individual datasets depends on a user's purpose for examining the data. Thus, providing datasets in the integrated manner will be useful to a wide range of users for a wide range of purposes.

3.1.1 Functionality

(1) Search and access datasets by Category or Map
The portal allows users to search and access datasets with a single user interface. In the search feature, the portal provides two kinds of searches to facilitate easy access, category search, and map search. Users can either start from a category list or a world map to get to their data of interest. In either search, users can reach any data in the portal and they may :
- Preview data in image format (GIF)
- Download data [available formats: NetCDF, ascii, and GRIB (only or Model Output))
- Compare data (e.g. Model output data and In-situ data)

(2) Share Use Case

The portal has a function that enables users to register use cases. Users can register the results of their research obtained by using data via the portal as a use case, which then becomes available for other users to reference in their data search on the portal. This facilitates users feedback and encourages communication within and across communities.

Datasets Parameters

In addition to CEOP data, a wide spectrum of hydrological datasets from data partners are being added as new elements of datasets, such as river discharge, water quality data, ground water monitoring data, etc..

3.1.2 Use Cases

By using the portal, users may achieve different types of goals. Possible scenarios for the use of the portal, for example, would be:

1) Drought index mapping in Asia by integrating the GPCC datasets, the CEOP Model outputs, the CMIP3 climate model outputs and AWCI hydrological data.
2) Ground water monitoring in Africa by integrating the IGRAC datasets, the GPCC datasets, and the CEOP satellite data and assimilation products.

4 Points of Contact

Responsible Agency
JAXA

Participating Agencies
NASA
NOAA

Points of Contact

Project lead:
Satoko Miura
Contact : miura.satoko AT jaxa.jp

NASA Point of Contact
Christopher Lynnes
Contact : christopher.s.lynnes AT nasa.gov

NOAA Point of Contact
Martin Yapur (TBD)
Contact : martin.yapur AT noaa.gov

5 Schedule

Portal development will continue through the end of March, 2016, at which point it becomes fully operational from April, 2016 onward.

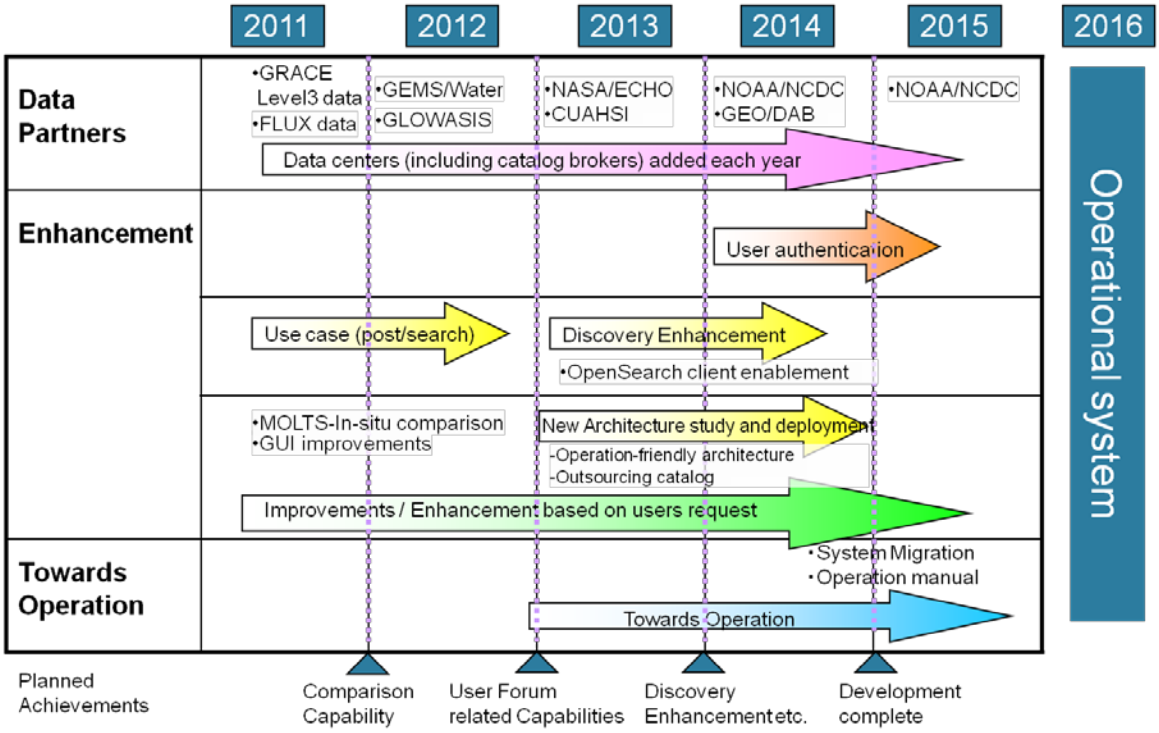

Fig.5-1 Development Schedule

Appendix A – Description of CEOP Data

A.1 In-situ Data

In-situ data is gathered at Reference Sites globally (see Fig.A-1,A-2). These data sets contain 30 minute and hourly data that is archived at the University Corporation for Atmospheric Research/ Joint Office for Science Support (UCAR/JOSS).

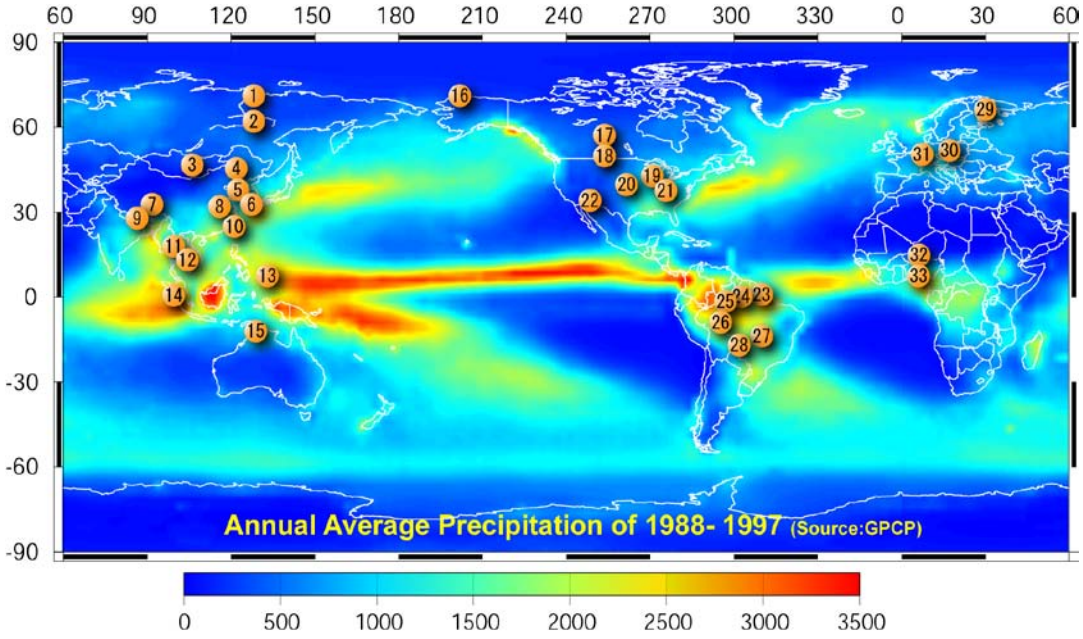

Fig A-1. CEOP Reference Sites (CEOP Phase-1)

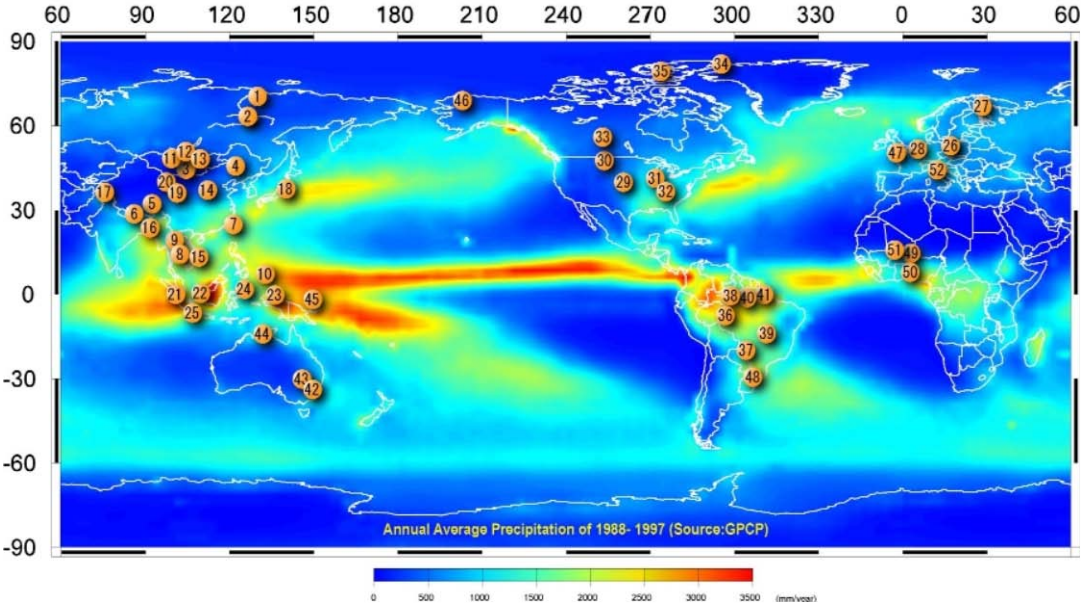

Fig A-2. CEOP Reference Sites (CEOP Phase-2)

There are four types of CEOP in-situ data sets:

- Surface Meteorological and Radiation Data Set
- Meteorological Tower Data Set
- Soil Temperature and Soil Moisture Data Set
- Flux Data Set

A.2 Satellite Data

Satellite data is being supplied to CEOP from JAXA, NASA, ESA and EUMETSAT. This includes full scenes of data over CEOP focused research areas (e.g. Asian Monsoon region), Level 3 global gridded data and subset scene data.
The satellite subset scenes are derived from the full scenes of data. The subset scenes were designed to be 250 km by 250 km squares with the data on a latitude longitude grid. Subset scenes over each of the CEOP Reference Sites are being produced, with the center of the 250 km by 250 km square over the Reference Site.

A.3 MOLTS and Model Output Data

Nine agencies contributes NWP (Numerical Weather Prediction) model output data and MOLTS (Model Output Location Time Series) data to CEOP. The model output data are 2D and 3D gridded data. There are Reference sites where in-situ data will be recorded (and where satellite subset scenes will be created). Values from the NWP model output data at the grid points which are closest to the reference sites will be saved as MOLTS data. The 2D and 3D gridded model output data are generally saved on a 3 hour or 6 hour basis. In general the MOLTS data are saved on an hourly basis.

Appendix B – Software Tools

This chapter will list software tools that are being used by this project.

E.1 OPeNDAP (Open source Project for a Network Data Access Protocol)

The OPeNDAP server (Distributed Oceanographic Data System) is free, open source software (http://www.opendap.org/). This server will provide distribution of the CEOP data. The OPeNDAP/DODS server can also subset the data. There are several types of DODS servers, for different types of data. Only the GrADS-DODS server can serve GRIB format data. The CEOP climate model output data is in GRIB format.

E.2 Ferret

Ferret (http://ferret.pmel.noaa.gov/Ferret/ferret_home.html) is a client, similar to GrADS, that can access OPeNDAP/DODS servers. Ferret is free, open source software supported by PMEL (Pacific Marine Environmental Laboratory). Ferret is a visualization and analysis system for analyzing large and complex gridded data sets.

E.3 Jblob

Jblob is a command-line based program for downloading data from the MPI's(Max-Planck Institute) database.

Appendix C - List of Acronyms

CEOS	Committee on Earth Observation Satellites
CEOP	Coordinated Enhanced Observing Period
CMIP3	phase 3 of the Coupled Model Intercomparison Project
ESA	European Space Agency
GEOSS	Global Energy and Water Cycle Observation System of Systems
IGOS	Integrated Global Observing Strategy
JAXA	Japan Aerospace Exploration Agency
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
TBD	To Be Determined
WGISS	Working Group on Information Systems and Services
WTF	WGISS Test Facility