2.5 DISTRIBUTED DATA INTEGRATION WTF-CEOP
(WGISS Test Facility for CEOP)

March 12, 2007
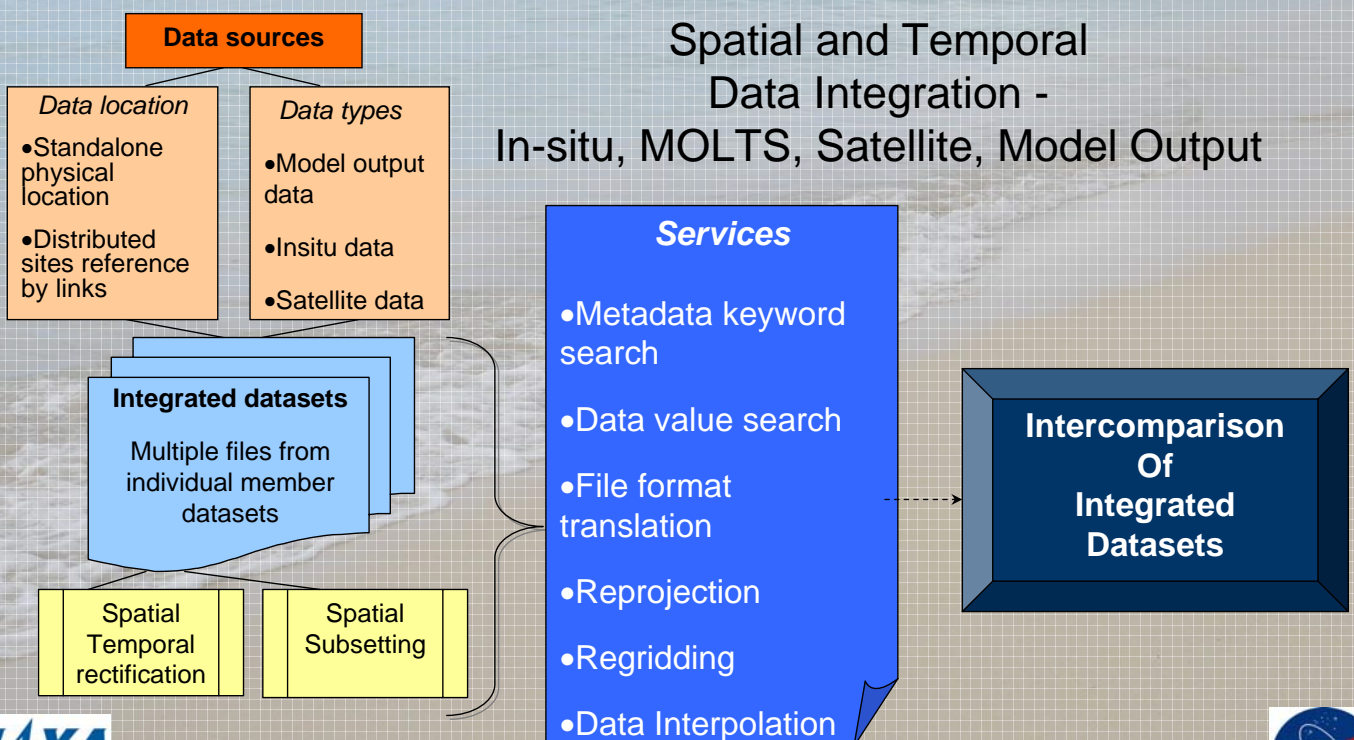
Satoko H. Miura (JAXA)

Ben Burford (JAXA/RESTEC)

Ken McDonald (NOAA)

Yonsook Enloe (NASA/SGT)

1

WTF-CEOP
Distributed **Data Integration** Services

Spatial and Temporal
Data Integration -
In-situ, MOLTS, Satellite, Model Output

Data sources

Data location
- Standalone physical location
- Distributed sites reference by links

Data types
- Model output data
- Insitu data
- Satellite data

Integrated datasets

Multiple files from individual member datasets

Spatial Temporal rectification

Spatial Subsetting

Services

- Metadata keyword search
- Data value search
- File format translation
- Reprojection
- Regridding
- Data Interpolation

Intercomparison Of Integrated Datasets

2

WTF-CEOP: **Phase 1 Prototype**

Web Browser

Live Access Server (LAS)

Ferret

WCS Server (satellite data) | **NASA Bridge**

CSDIC Archive [Satellite] (Japan)

NCAR Archive [In-situ] (U.S.)

MPI Archive [MOLTS, Model Output] (Germany)

CEOP Data Archives

▶ **OPeNDAP Interface**
▶ **Custom Interface**

Users can access;
(1) The data located in **distributed archive centers**
(2) The data which is **temporally and geospatially coordinated**

JAXA Prototype:
"Distributed Data Integration Prototype System for CEOP "

NASA Prototype:
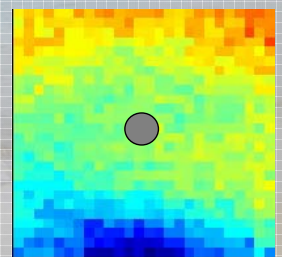A bridge between DODS formatted data requests and OGC WCS servers

---

JAXA Prototype: **Phase 1 Status**
= Available Data =

(i) Satellite Data
- 250 km x 250 km subset scenes, regridded to Lat/Lon grid, centered over the 35 in-situ data sites
- Band selection, period selection and download multiple scenes available

(ii) In situ Data
- Four types of CEOP in-situ data from 35 Reference Sites
  Surface (19 variables), Tower (9 variables), Soil (temperature, moisture), Flux (Sensible Heat, Latent Heat, CO2, Soil Heat Flux)
- Download of QC flags available

(iii) Model Output Data
- NWP – Numerical Weather Prediction data
- MOLTS – Model Output Location Time Series
(time series at model grid point nearest center of in-situ data sites)
- Gridded Model Output – 2D/3D time series.

Annual Average Precipitation of 1988- 1997 (Source:GPCP)
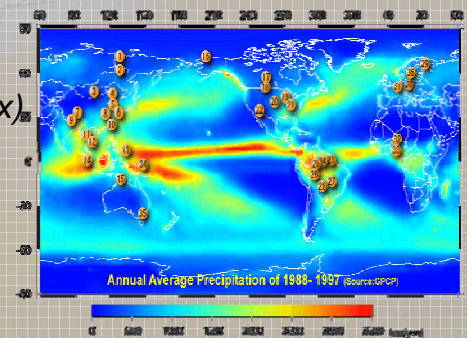
JAXA Prototype: **Phase 1 Status**
= Main Functions =

- *Access* data from distributed archive sites
- *Select*
  - ✓ data type
  - ✓ location (Reference Site, Station name)
  - ✓ variable name
  - ✓ height or depth
  - ✓ time range (start time / stop time)
- *"See the data"*
  - ✓ plot
  - ✓ view values on the screen
- *Download*
  - ✓ NetCDF
  - ✓ Ascii
  - ✓ Spreadsheet format
- Registration (to satisfy CEOP data policy)

5

JAXA Prototype: **Phase 1 Status**
=Examples=

Please visit and use our system at:
http://jaxa.ceos.org/wtf_ceop/

Top Page

Tutorial DVD will be distributed tomorrow.

MOLTS data plot in time series

plot in 2D view

Demonstration is planned on March 17th (during IGWCO workshop)

Satellite data plot in 2D view

In-situ data in NetCDF format with QC flags

QC flags   Observed Data

- ☀ **Adding other types of data and information**
  - ➢ Maps
    - flood risk maps, etc.
  - ➢ Socioeconomic data
    - population density
    - land use, etc.
- ☀ **Prototyping capabilities for more advanced data search and discovery functions**
- ☀ **Access of data from other server technologies**
  - ➢ OGC (Open Geospatial Consortium) WCS (Web Coverage Servers)
  - ➢ OGC WMS (Web Map Servers), etc.

CEOP Data Access

- CEOP science community uses OPeNDAP based clients to access in-situ and model data

- JAXA assisting CEOP central archives in making their data available via OPeNDAP
  - In situ and satellite data

- NASA and JAXA providing satellite data subsets to Univ. of Tokyo

- NASA also developing a tool to improve access to distributed satellite data collections
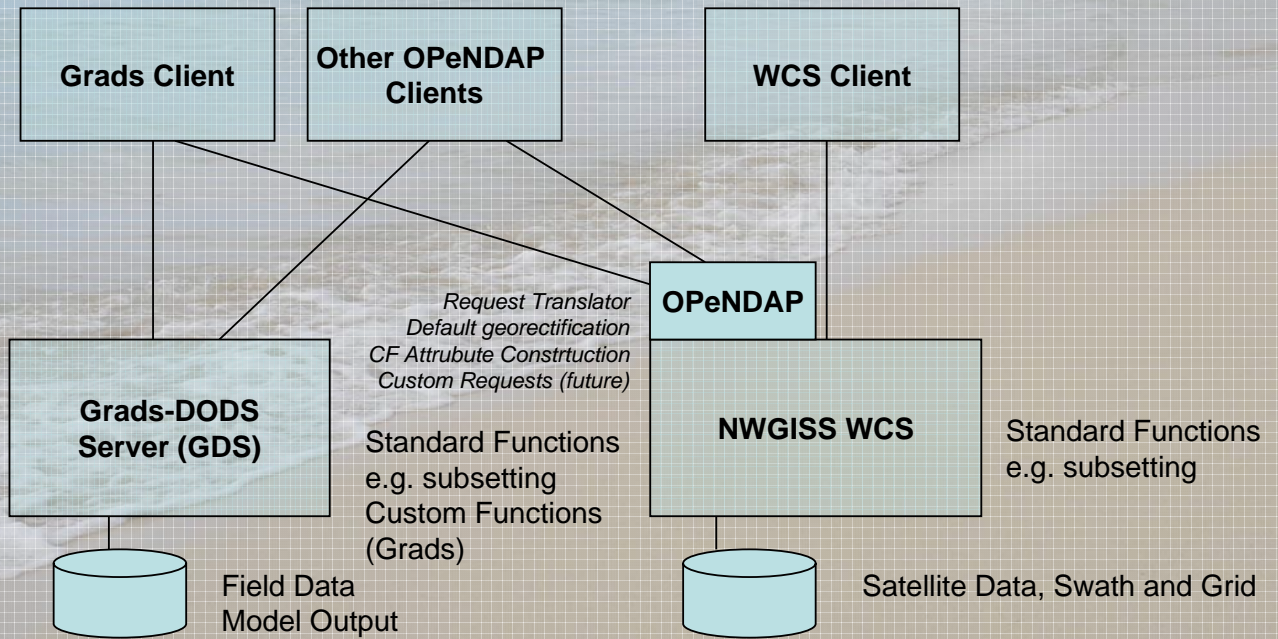
NASA Satellite Data Status

- Request for selected MODIS and AIRS products for EOPs submitted to NASA.
- Products will be provided by multiple facilities.
  - AIRS – Goddard DAAC
  - MODIS Atmosphere, Land – MODAPS
  - MODIS Ocean – Goddard ODPS
- Technical discussions to refine request underway
  - Metadata specifications
  - Data service requests – format, projection, etc.
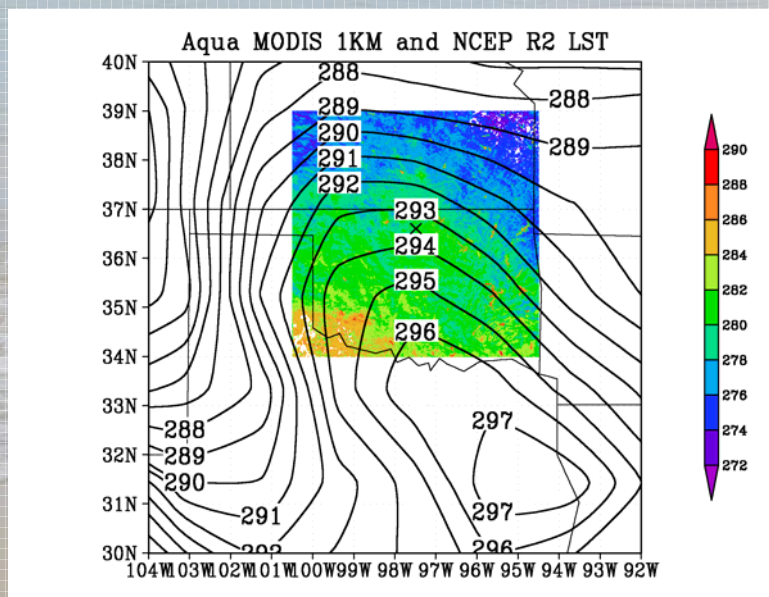 - Delivery schedules

NASA's WTF-CEOP Contribution

- CEOP scientists also desire direct access to full range of satellite data products and services from base providers
  - Latest versions of products
  - Flexible access to data services

- Growing use of Open Geospatial Consortium (OGC) specifications by satellite data providers

- NASA developed a CEOP Satellite Data Server prototype that demonstrated OPeNDAP client (Grads) access to the OGC-WCS server capabilities
  - Complement to the CSDIC at Univ. of Tokyo

- Prototype has been demonstrated at CEOP and CEOS meetings

NASA Prototype: Satellite Data Server

Grads Client

Other OPeNDAP Clients

WCS Client

Request Translator
Default georectification
CF Attrubute Constrtuction
Custom Requests (future)

OPeNDAP

Grads-DODS Server (GDS)

NWGISS WCS

Standard Functions e.g. subsetting
Custom Functions (Grads)

Standard Functions e.g. subsetting

Field Data
Model Output

Satellite Data, Swath and Grid

Integration of Satellite and Model Data

Aqua MODIS 1KM and NCEP R2 LST

Prototype vs. Real Need

- Prototype - limited, proof of concept
  - Demonstrated ability to access WCS data from OPeNDAP client
    - No catalog services
    - Default data services
- Science need -
  - Data discovery / data search
  - OPeNDAP client access to the full range of WCS services
- Support from NASA's Advancing Collaborative Connections for Earth-Sun System Science (ACCESS) Program
  - Combined with second proposal - OGC to OPeNDAP Gateway

Gateway Functions/Approach

- Transformation of WCS metadata to be consistent with other CEOP data sources
  - Climate and Forecast Metadata Conventions
  - Must handle various coordinate reference systems
  - Augment with metadata on requested WCS operations
- Enable request of WCS operations
  - Construct DAP request to include WCS services
    (Georectification, reprojection, subsetting…)
  - Transform DAP request to WCS request
  - Issue request to WCS
  - Transform response to DAP data object

OPeNDAP & ECHO

- ECHO is the metadata clearinghouse for NASA satellite data – contains metadata of all EOS missions
- OPeNDAP and ECHO teams worked together to provide direct search and access to ECHO via OPeNDAP enabled clients (e.g. Matlab) using ECHO's web services.
- OPeNDAP Group developed a stand alone Java client to search ECHO metadata. This stand alone client can be invoked by Matlab and other OPeNDAP enabled clients.
- ECHO and OPeNDAP teams are working together to provide operational access of ECHO metadata search using OPeNDAP enabled clients such as Matlab.

Status and Plans

- Project initiated last spring
  - Finalized work plan for both gateways
- Prioritization of WCS services discussed with CEOP science representative
- Preliminary design of gateway completed at end of summer
- First release scheduled for June, 2007.
- Plan to deploy the Satellite Data Server at NASA and other sites with satellite data of interest to the CEOP community

Issues for CEOP Science: From Sept 2006 JAXA/NASA meeting

- Identify data sources of interest for data discovery/search (Max Planck, UCAR, University of Tokyo, JAXA, NASA, Eumetsat, ESA, NOMADS, ...)
- Discuss the CEOP need for broader data discovery
- Identify data type – product and geophysical parameter names – have this info from Prof. Koike already (list of satellite products he sent out last year) but the list of NASA satellite products do not have the geophysical parameters listed like those of the ESA products
- JAXA team will ask CEOP scientists what additional LAS system functions they would like to have implemented in the JAXA prototype