MEXT MINISTRY OF EDUCATION, CULTURE, SPORTS, SCIENCE AND TECHNOLOGY-JAPAN

THE UNIVERSITY OF TOKYO

DIAS Data Integration & Analysis System

DIAS

as an open science platform for Earth Informatics

Eiji Ikoma , Akiyuki Kawasaki

Toshio Koike, Masaru Kitsuregawa

The University of Tokyo,JAPAN

Our Talk

Leader of DIAS Platform Development team,
DG of National Institute of Informatics, Japan

- Ass.Prof. Eiji Ikoma/Prof.Masaru Kitsuregawa

@**Information Engineering**, The University of Tokyo

Introduction of DIAS system as an open science platform for Earth Informatics

Leader of DIAS Project
DG of Global Center of Excellence for Water Hazard and Risk management, Japan

- Ass.Prof. Akiyuki Kawasaki / Prof. Toshio Koike

@**Civil Engineering**, The University of Tokyo

Introduction of DIAS applications in flood and draught analysis

My Talk

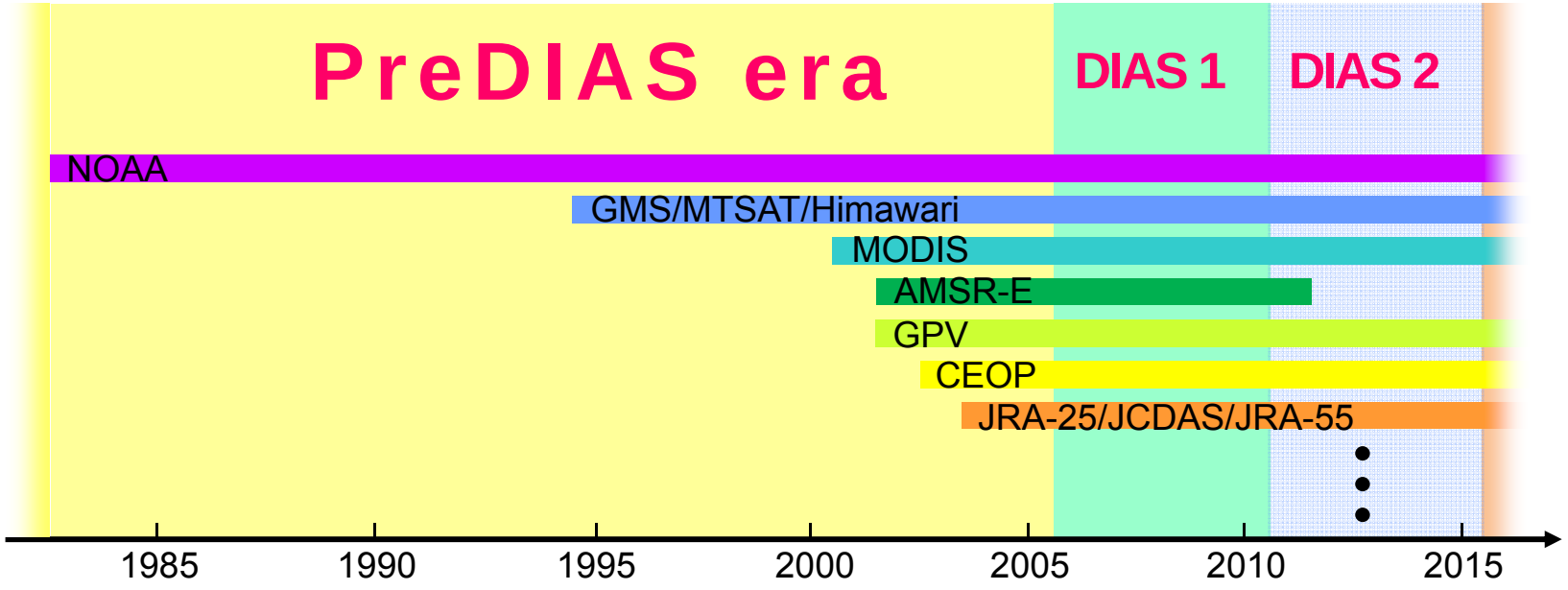
- History of DIAS
- Introduction of DIAS
  - System Structure
  - Power Saving Function
  - Challenges to 4V
  - Network Infrastructure
- “DIAS Value”
- Summary and Future plan

Dr.Kawasaki’s talk

History of DIAS

35years：History of Developing Earth Environmental Data Repository on IIS, The University of Tokyo

DIAS： 1/3 of total development period



Belief of 30 years ago

The era of
「Data is source of power」
will surely come.

30 years ago….

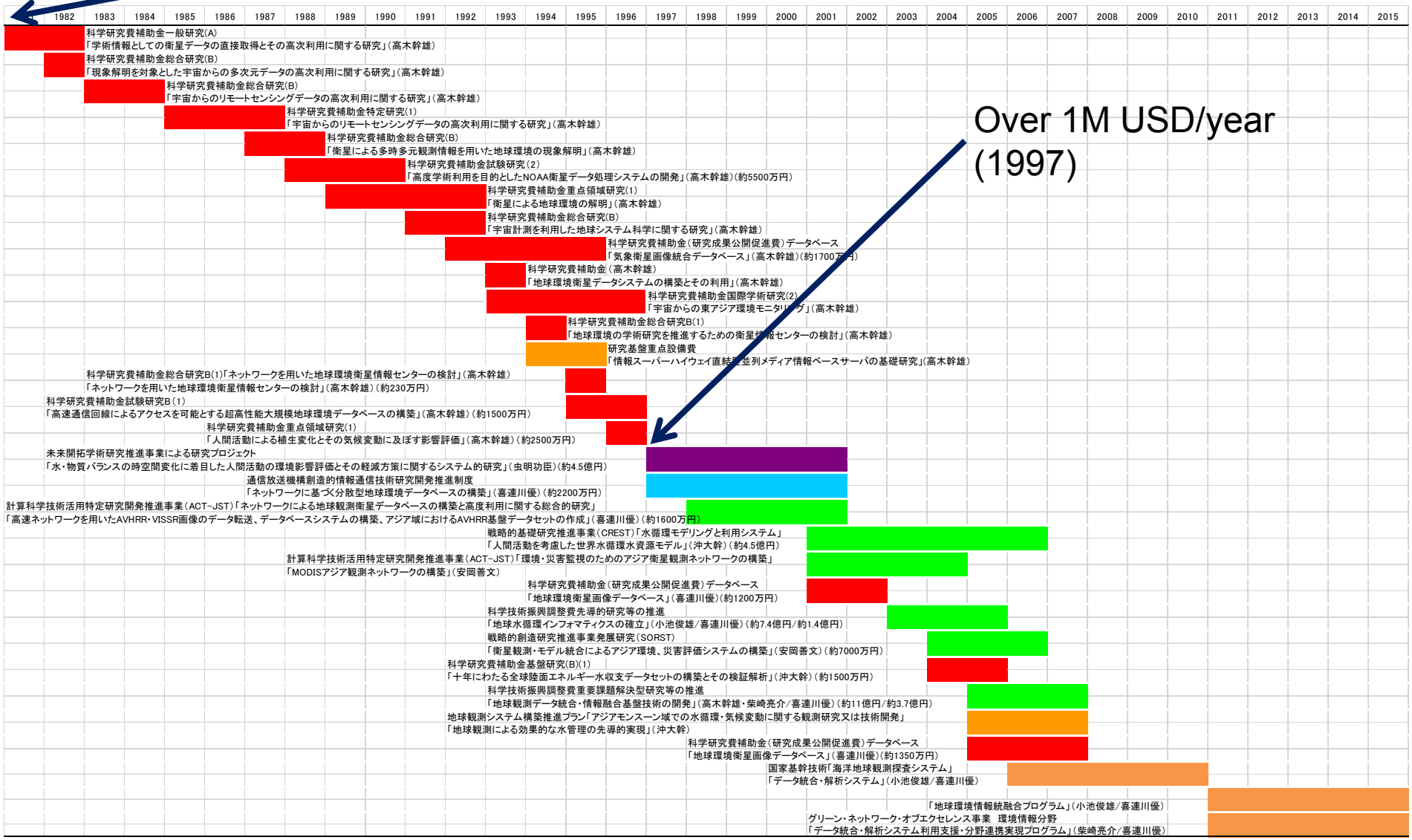
Main Memory 2MB, NOAA data 100MB

Power of CPU is 1/1,000,000

Looking back first 20 years..

Our budget about earth environmental informatics at IIS, U-TOKYO

Start on 1981



Satellite Antenna for NOAA

（Installed in 1980 at Roppongi Campus, Trial operation from 1981 and full operation from 1983）

The late Professor, Mikio Takagi

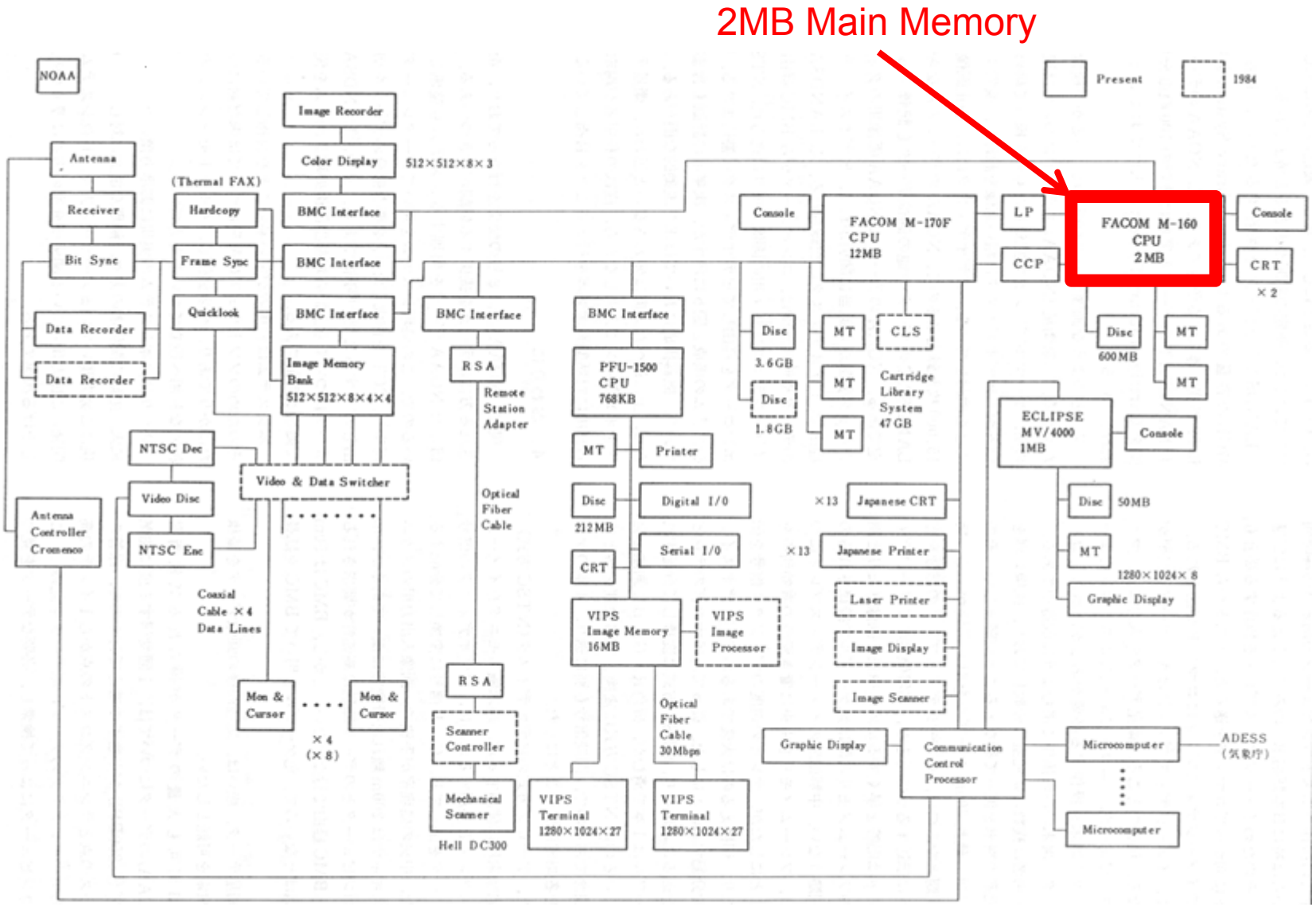
Hand made Receiving Station

bit synchronizer,
frame synchronizer 1981-

Analog Data Recoder 1982-

Mainframe Machine (FACOM M160/170)

System Configuration（1984）



Mass Storage
8mm tape archive **1992-**

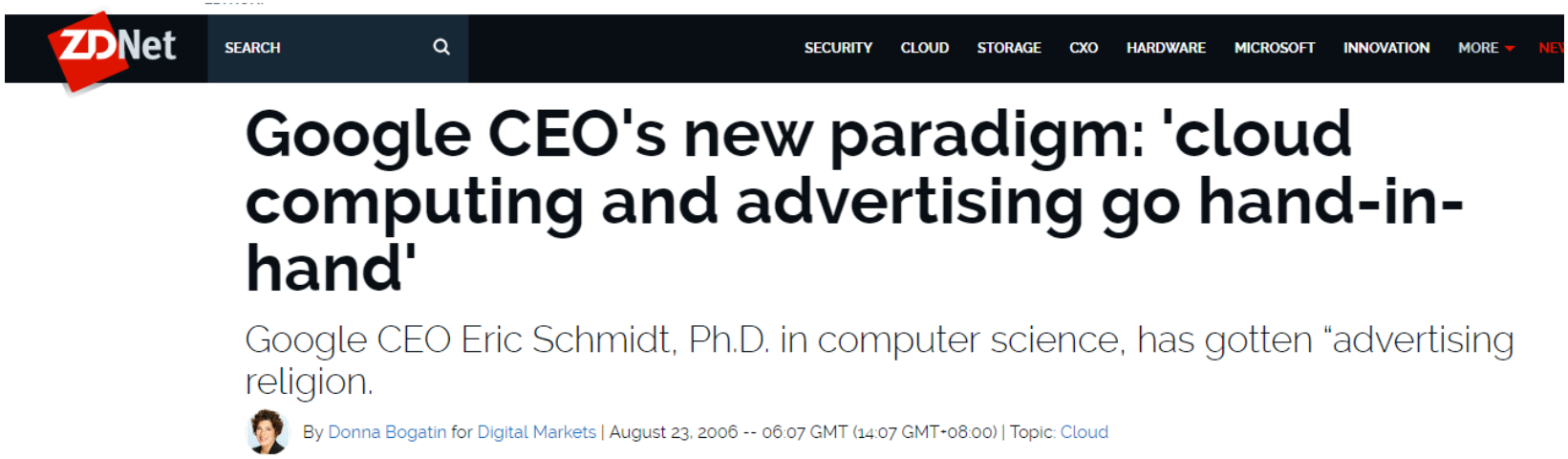
STK 9310 (Powder Horn)

High End Storage(6000 tapes）2001-

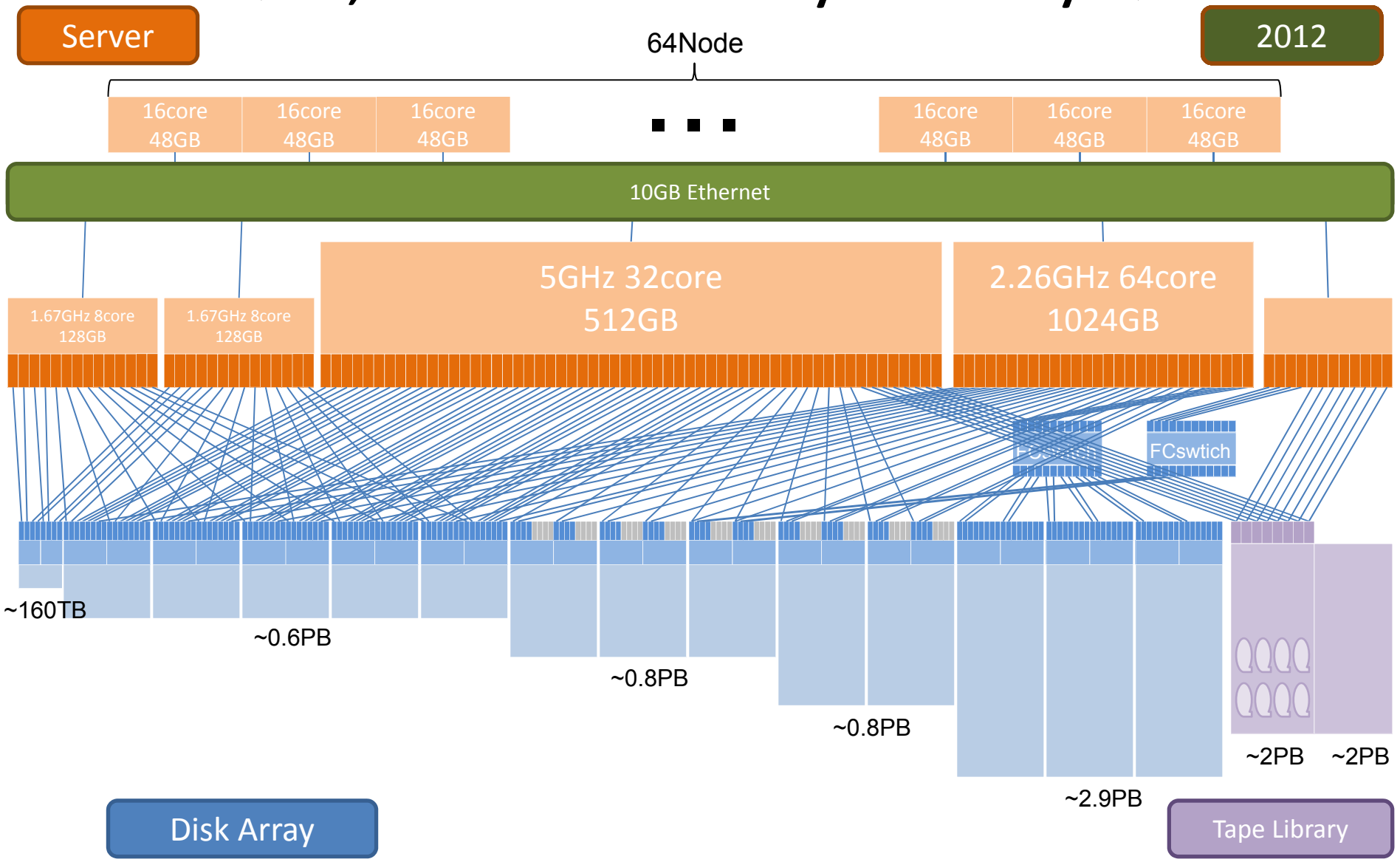
DIAS started in 2006

What we thought 10 years ago (2006)

Cloud-Shift
“Platformer will win”

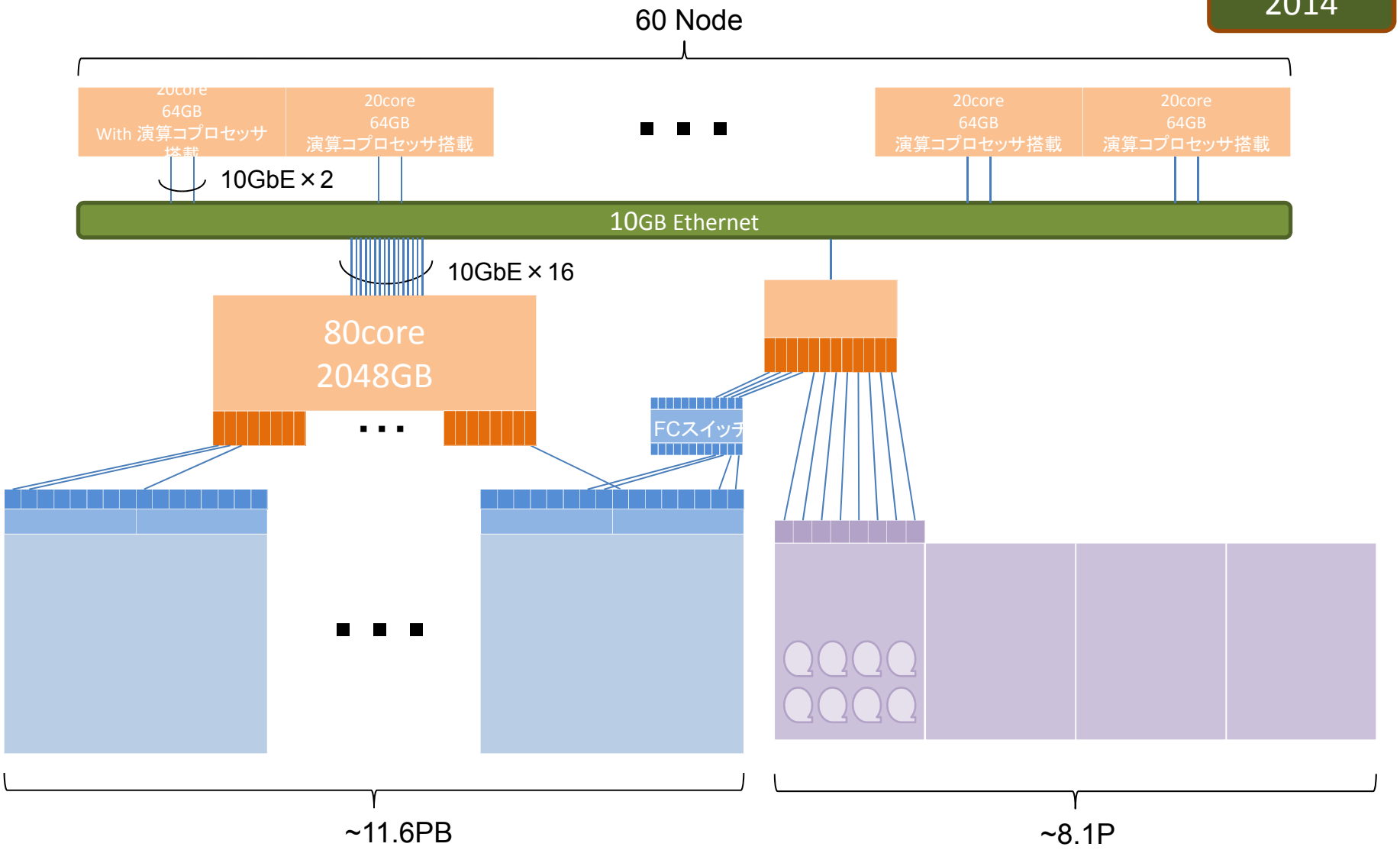


Server-Storage Coupled System
（IIS, The University of Tokyo）



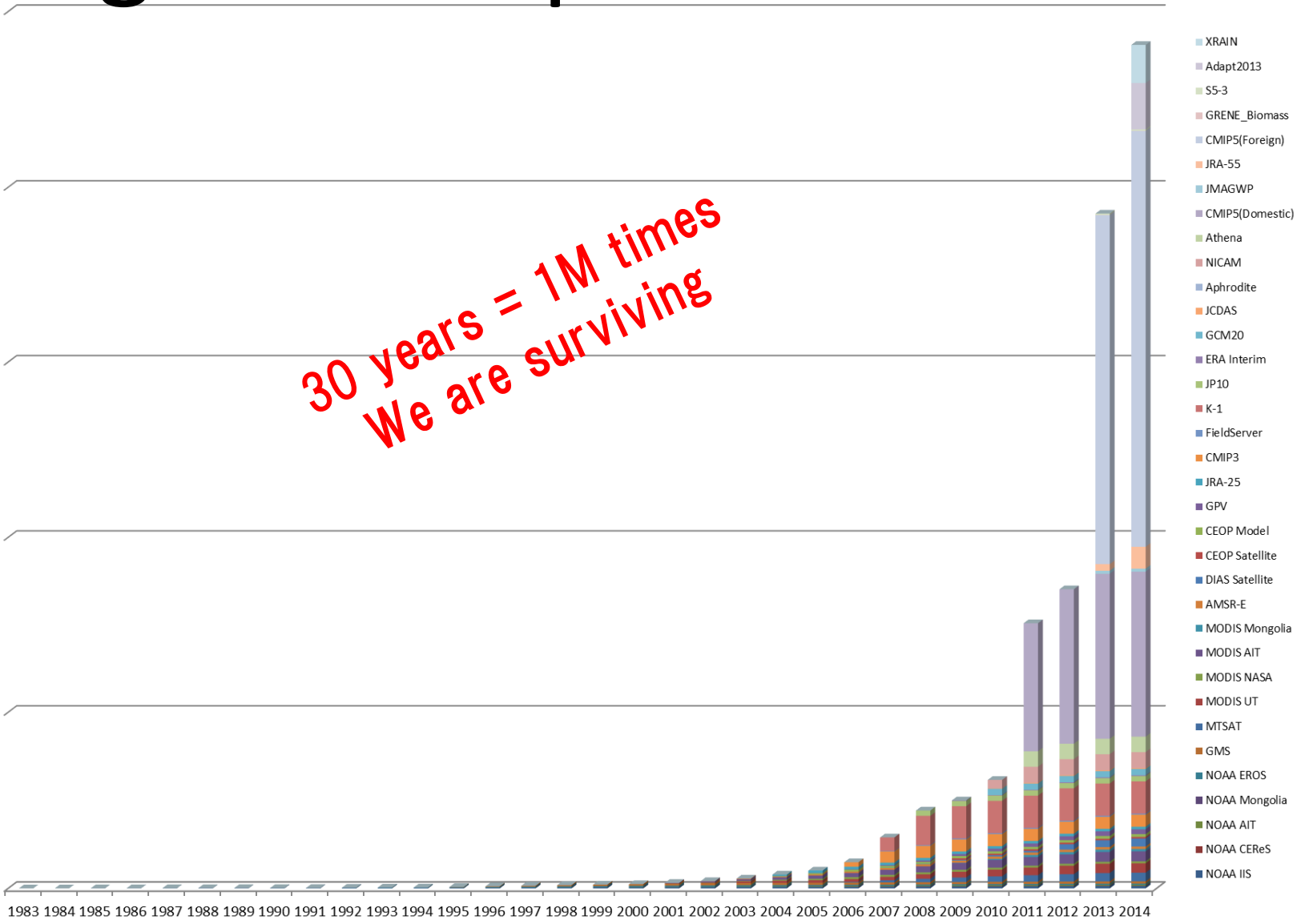
System Structure
（National Institute of Informatics）

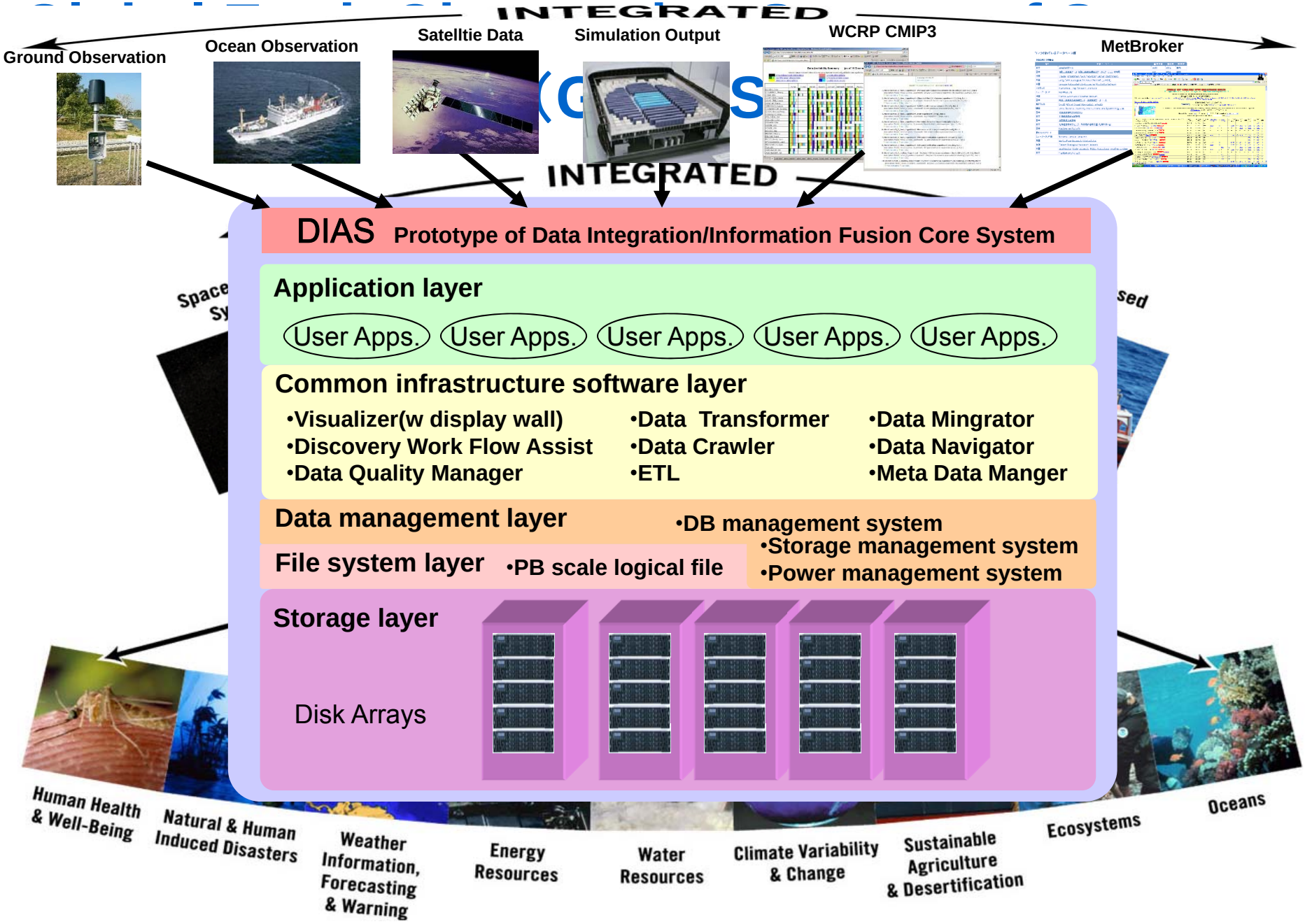
2014



DIAS Today
disk + tape > 25PB

Digital Info-plosion in DIAS




INTEGRATED
Ground Observation
Ocean Observation
Satelltie Data
Simulation Output
WCRP CMIP3
MetBroker
INTEGRATED
DIAS Prototype of Data Integration/Information Fusion Core System
Application layer
User Apps.
User Apps.
User Apps.
User Apps.
User Apps.
Common infrastructure software layer
•Visualizer(w display wall)
•Discovery Work Flow Assist
•Data Quality Manager
•Data Transformer
•Data Crawler
•ETL
•Data Mingrator
•Data Navigator
•Meta Data Manger
Data management layer
•DB management system
File system layer
•PB scale logical file
•Storage management system
•Power management system
Storage layer
Disk Arrays
Human Health & Well-Being
Natural & Human Induced Disasters
Weather Information, Forecasting & Warning
Energy Resources
Water Resources
Climate Variability & Change
Sustainable Agriculture & Desertification
Ecosystems
Oceans

DIAS is really “on premise” cloud

Why we thought “DIAS should be a cloud ”

We should support that researchers in each domain can proceed their original research without thinking “halfway” IT

This is the “driver” of today’s cloud shift

Our feeling of direction in those days is almost corecct

Data Centric Cloud

(Centralization of IT
for huge scale data)

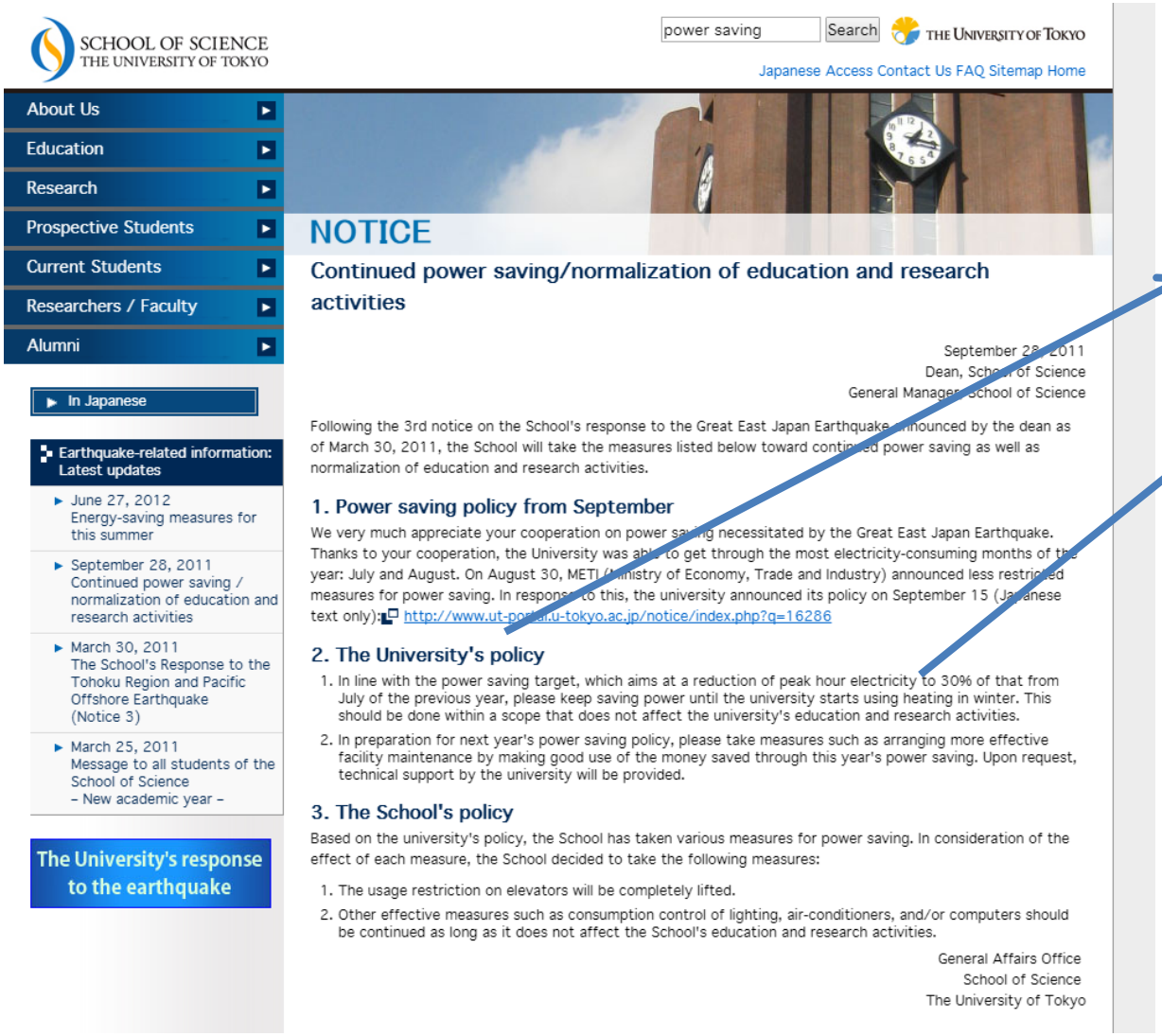
Another feature of DIAS:

Power saving

2011.3.11 Great Earthquake

The university of Tokyo is the largest power consumer in Tokyo

Power Saving @ U-Tokyo



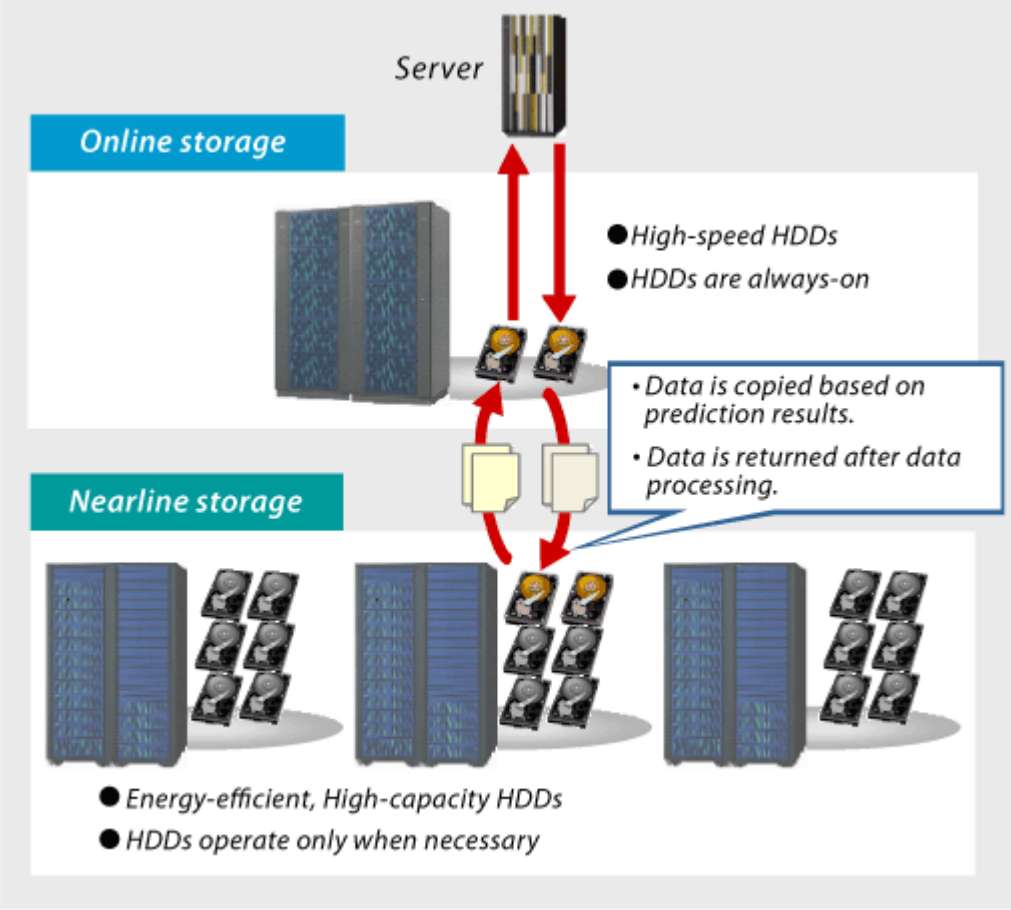
2. The University's policy

electricity to 30%

>-30%!!

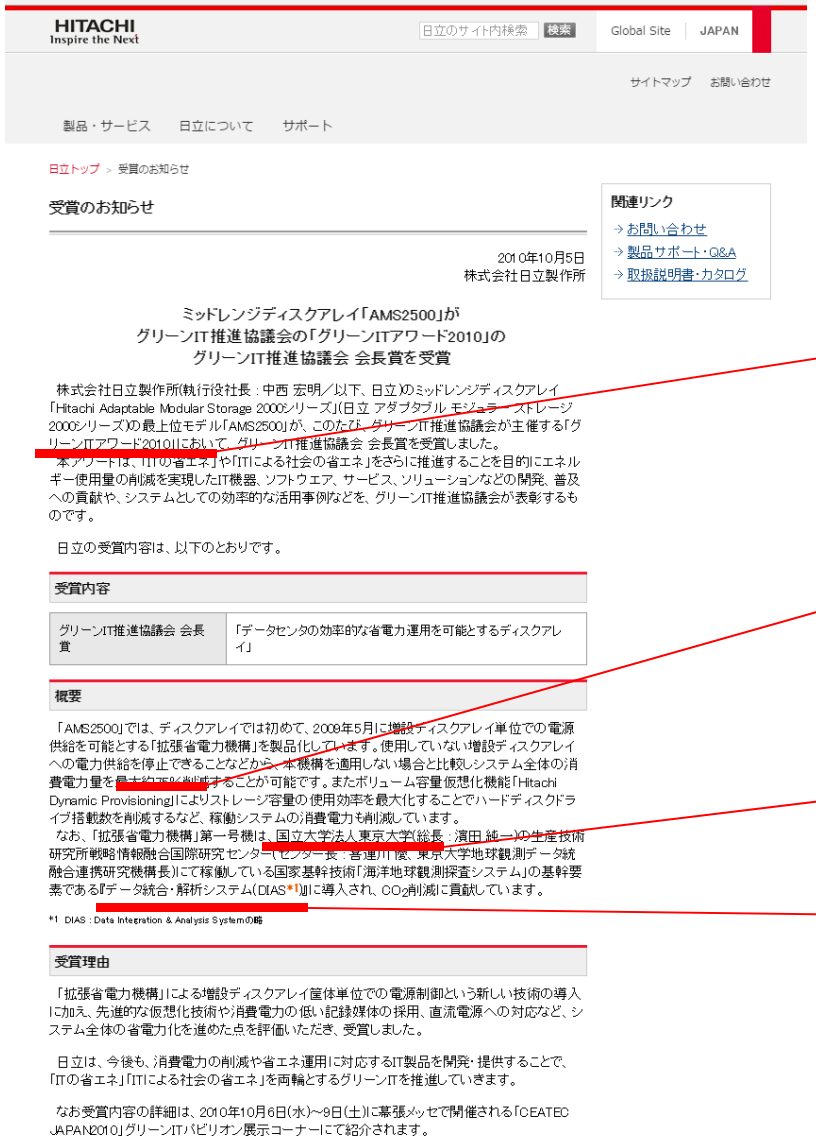
From 2011

Our research about Green IT(2009-)



Tiered storage (case: power-aware proactive data-allocation)

Press release on 2009/05/11



HITACHI
Inspire the Next

日立のサイト内検索 検索 | Global Site | JAPAN

サイトマップ お問い合わせ

製品・サービス 日立について サポート

日立トップ > 受賞のお知らせ

受賞のお知らせ

2010年10月5日
株式会社日立製作所

関連リンク
- お問い合わせ
- 製品サポート・Q&A
- 取扱説明書・カタログ

ミッドレンジディスクアレイ「AMS2500」が グリーンIT推進協議会の「グリーンITアワード2010」の グリーンIT推進協議会 会長賞を受賞

株式会社日立製作所(執行役社長：中西 宏明／以下、日立)のミッドレンジディスクアレイ「Hitachi Adaptable Modular Storage 2000シリーズ」(日立 アダプタブル モジュラーストレージ2000シリーズ)の最上位モデル「AMS2500」が、このたび、グリーンIT推進協議会が主催する「グリーンITアワード2010」において、グリーンIT推進協議会 会長賞を受賞しました。

本アワードは、「ITの省エネ」や「ITによる社会の省エネ」をさらに推進することを目的にエネルギー使用量の削減を実現したIT機器、ソフトウエア、サービス、ソリューションなどの開発、普及への貢献や、システムとしての効率的な活用事例などを、グリーンIT推進協議会が表彰するものです。

日立の受賞内容は、以下のとおりです。

受賞内容

グリーンIT推進協議会 会長賞	「データセンタの効率的な省電力運用を可能とするディスクアレイ」

概要

「AMS2500」では、ディスクアレイでは初めて、2009年5月に増設ディスクアレイ単位での電源供給を可能とする「拡張省電力機構」を製品化しています。使用していない増設ディスクアレイへの電力供給を停止できることなどから、本機構を適用しない場合と比較しシステム全体の消費電力量を最大約[redacted]削減することが可能です。またボリューム容量仮想化機能「Hitachi Dynamic Provisioning」によりストレージ容量の使用効率を最大化することでハードディスクドライブ搭載数を削減するなど、稼働システムの消費電力も削減しています。

なお、「拡張省電力機構」第一号機は、国立大学法人東京大学(総長：濱田 純一)の生産技術研究所戦略情報融合国際研究センター(センター長：喜連川 優、東京大学地球観測データ統融合連携研究機構長)にて稼働している国家基幹技術「海洋地球観測探査システム」の基幹要素である『データ統合・解析システム(DIAS*1)』に導入され、CO_2削減に貢献しています。

*1 DIAS：Data Integration & Analysis Systemの略

受賞理由

「拡張省電力機構」による増設ディスクアレイ筐体単位での電源制御という新しい技術の導入に加え、先進的な仮想化技術や消費電力の低い記録媒体の採用、直流電源への対応など、システム全体の省電力化を進めた点を評価いただき、受賞しました。

日立は、今後も、消費電力の削減や省エネ運用に対応するIT製品を開発・提供することで、「ITの省エネ」「ITによる社会の省エネ」を両輪とするグリーンITを推進していきます。

なお受賞内容の詳細は、2010年10月6日(水)～9日(土)に幕張メッセで開催される「CEATEC JAPAN2010」グリーンITパビリオン展示コーナーにて紹介されます。

Green IT Award

Save 75%(max)

University of Tokyo

DIAS

Green IT AWARD 2010
Green IT Promotion Council

DIAS expand to North

Server（Cluster）
- •8node
- •CPU16core/node
- •Mem48GB/node

Disk Array
- •~1.4PB

Hokkaido Univ.

Kitami Univ.

Server（Cluster）
- •8node
- •CPU16or12core/node
- •Mem48GB/node

DiskArray
- •~0.7PB

Server
- •CPU 64core
- •Mem1024GB

Server
- •CPU32core
- •Mem512GB

Server（Cluster）
- •64node
- •CPU 16core/node
- •Mem48GB/node

Disk Array
- •~5.1PB

Server
- •CPU80core
- •Mem2048GB

Server（Cluster）
- •60node
- •CPU20core/node
- •Mem64GB/node
- •FPU/GPU

Disk Array
- •~11.6PB

Tape Library
- •~8.1PB

IIS, Univ. of Tokyo

National Insitute of Informatics

DIAS-3 started in 2016

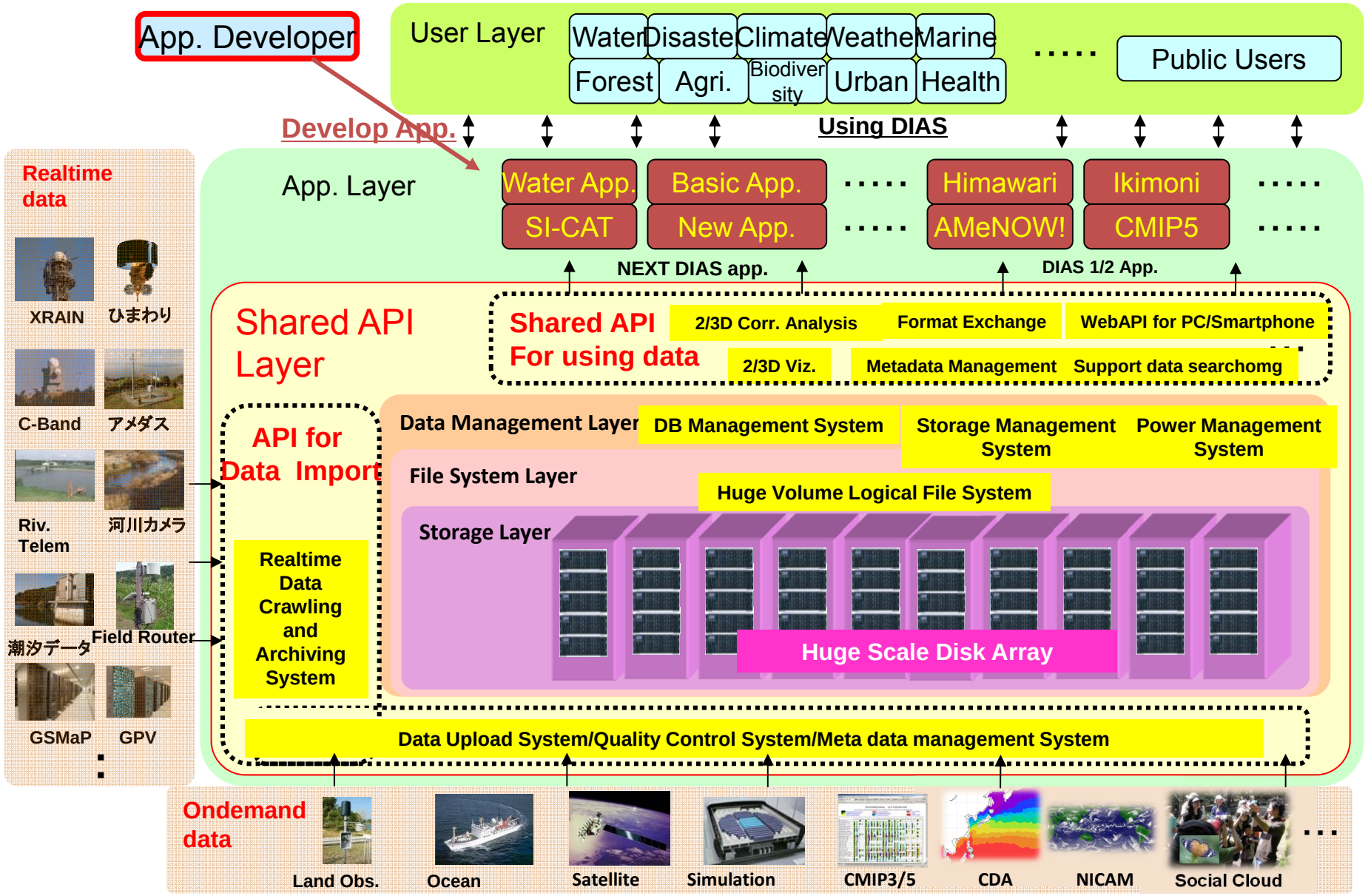
DIAS is evolving to be a open science platform for Application

Data Intensive System with application platform

+

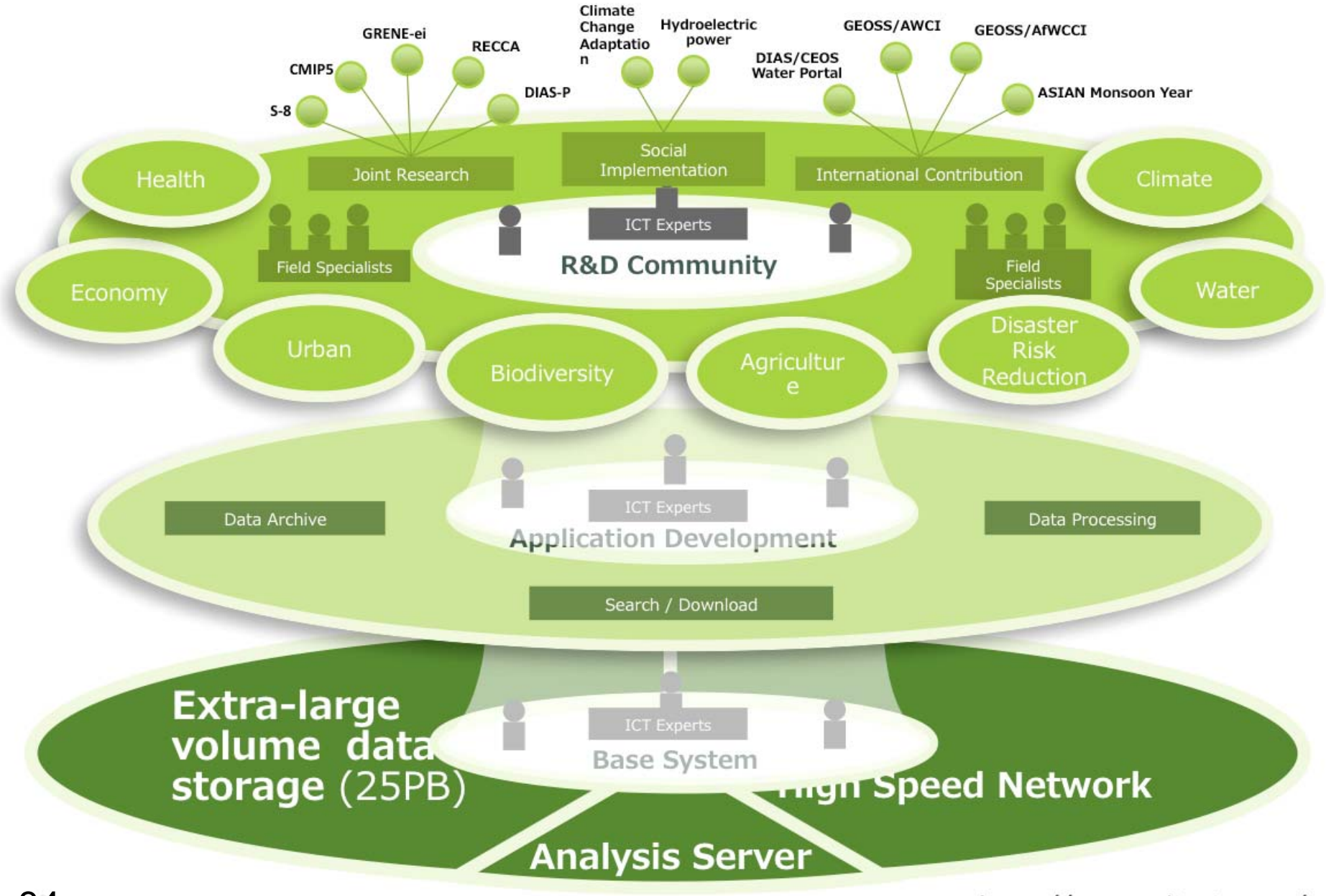
High-Speed Network

Common Base Platform for Application



Data Integration and Analysis System (DIAS)

Challenges to 4V



http://www.diasjp.net/en

Challenges to 4V

- volume
- variety
- veracity
- velocity

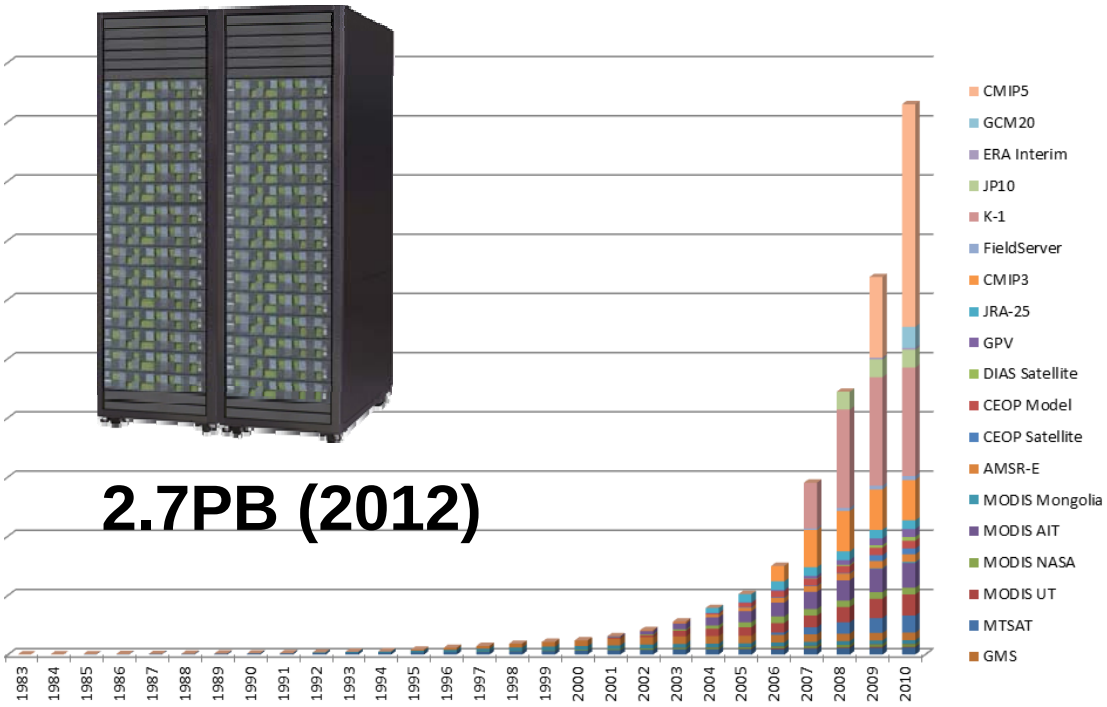
Data Integration and Analysis System

a legacy for Japan's contributions to GEOSS

tackling a large increase in volume of the Earth observation data.

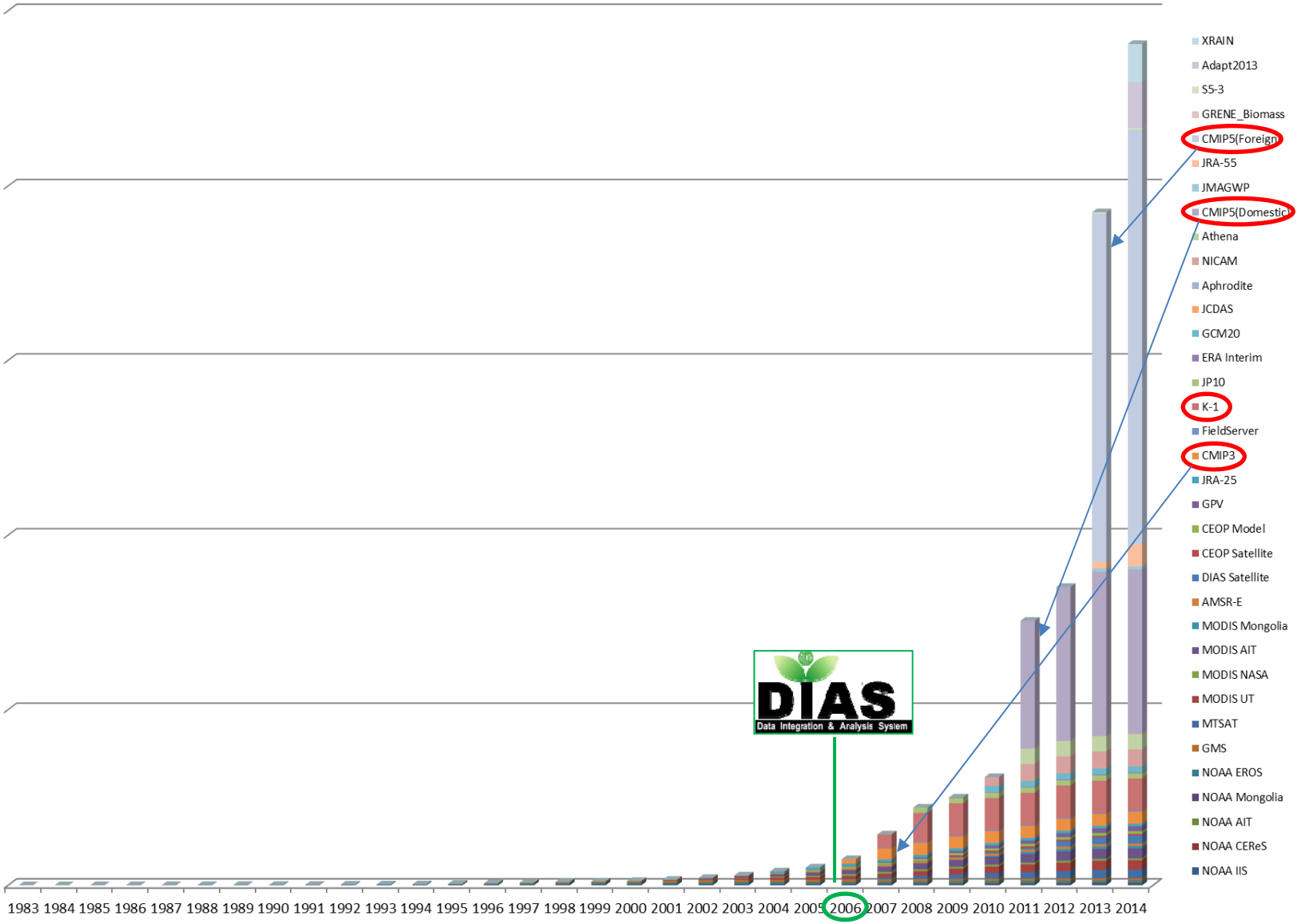
IPCC AR4 (2007): 40TB → IPCC AR5 (2012): 2.6PB

600TB (2007)



Analysis and Visualization System of CMIP5 data in DIAS

Data Volume Trends



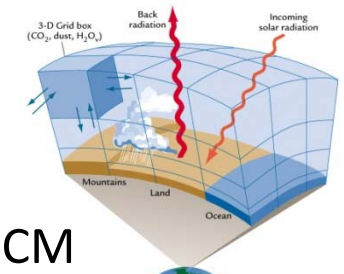
… a fairly large portion of the data volume is occupied by climate change prediction data.

Current Status of CMIP5 archived to DIAS

- Dataset Size: 1.7PB (as of November 2016)
- # of Models: 61
- # of Variables: 732
  - aerosol: 78
  - atmos: 286
  - land / landIce: 91
  - ocean / seaIce: 156
  - ocnBgchem: 121
- # of Experiments: 101
- # of Data Files: ~4 million

CMIP5 Data Analysis and Visualization System

System Overview



GCM Output

Observation

User

Web Server / CGI

Visualization

Analysis

CMIP5, JRA55, GPCP, etc.

Data Storage

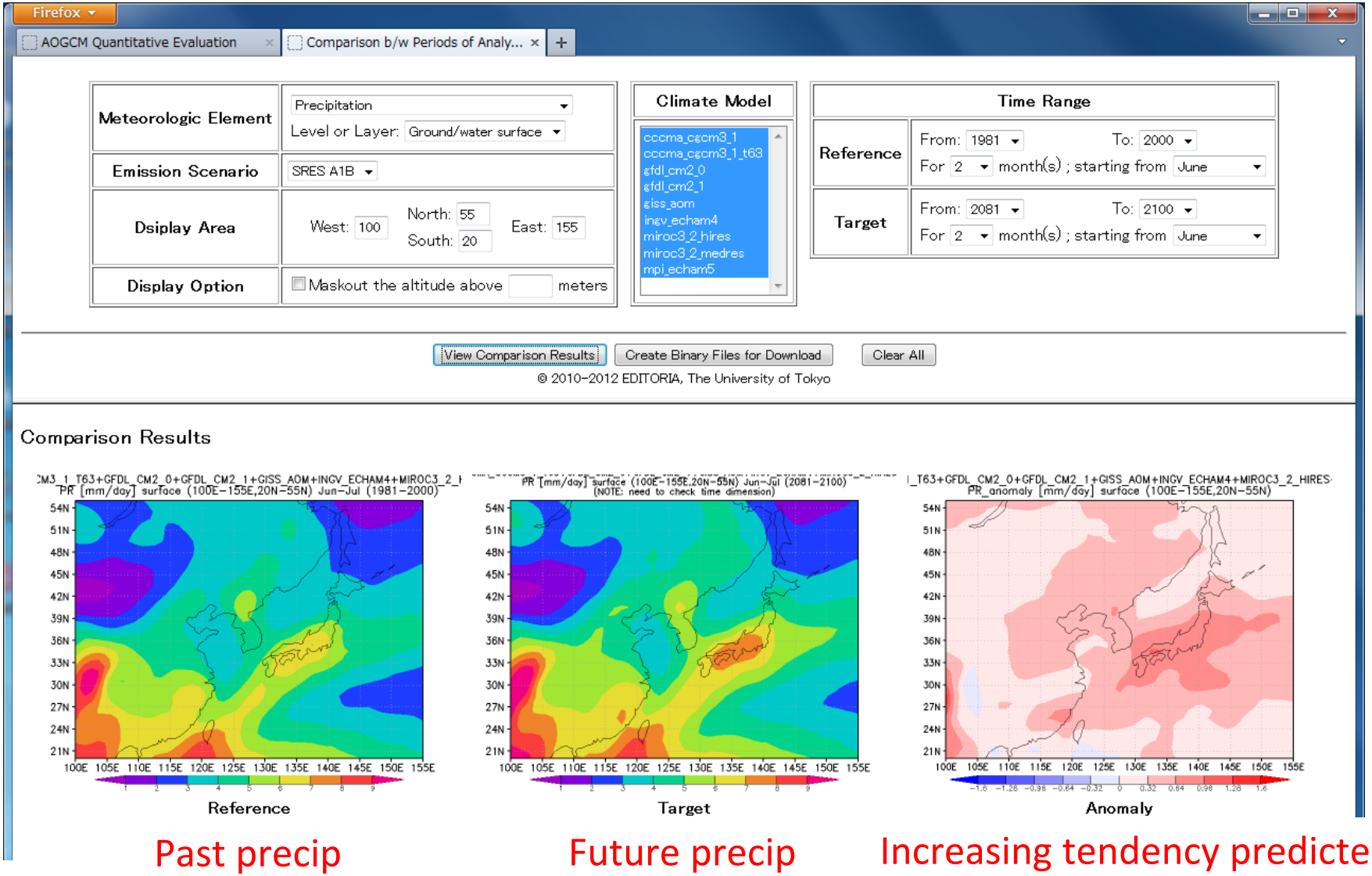
DIAS Core System

Implemented Functions List

- Model output variables
  - Precipitation
  - Surface Temperature
  - Outgoing Longwave Radiation
  - Sea Level Pressure
  - Sea Surface Temperature
  - Air Temperature
  - Geopotential Height
  - Specific Humidity
  - Zonal/Meridional Wind
  - …
- Visualization
  - Time series
  - 2-D mapping
  - Sectional view
    - lat-/long-time
    - lat-/long-height
  - Wind vector, horizontal divergence, and vorticity
- Analysis
  - Spatial correlation, RMSE
  - Time variation
  - (Downloading of results)

Application Example (1):

Multi-model Ensemble Prediction



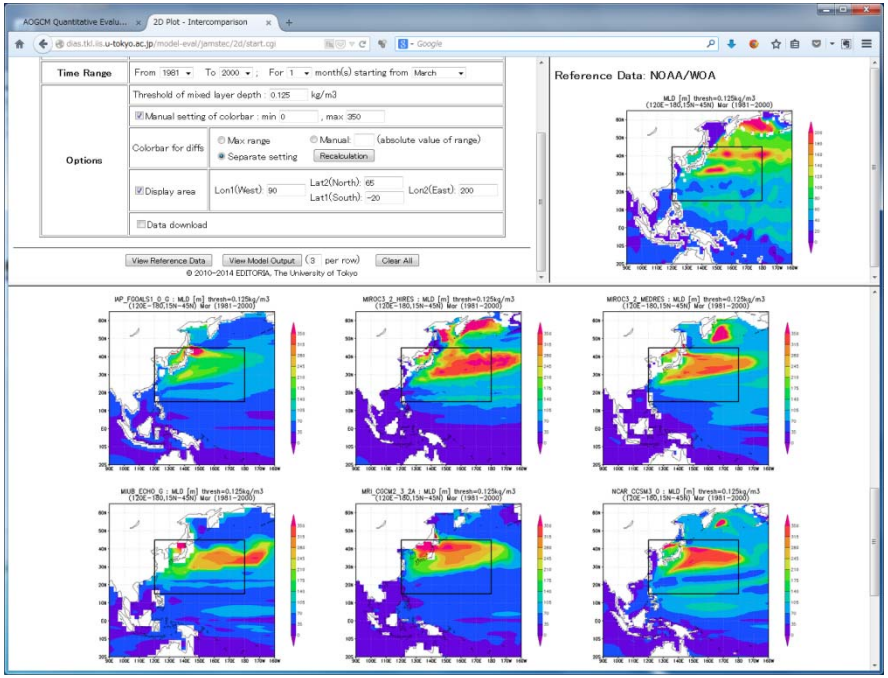
Past precip

Future precip

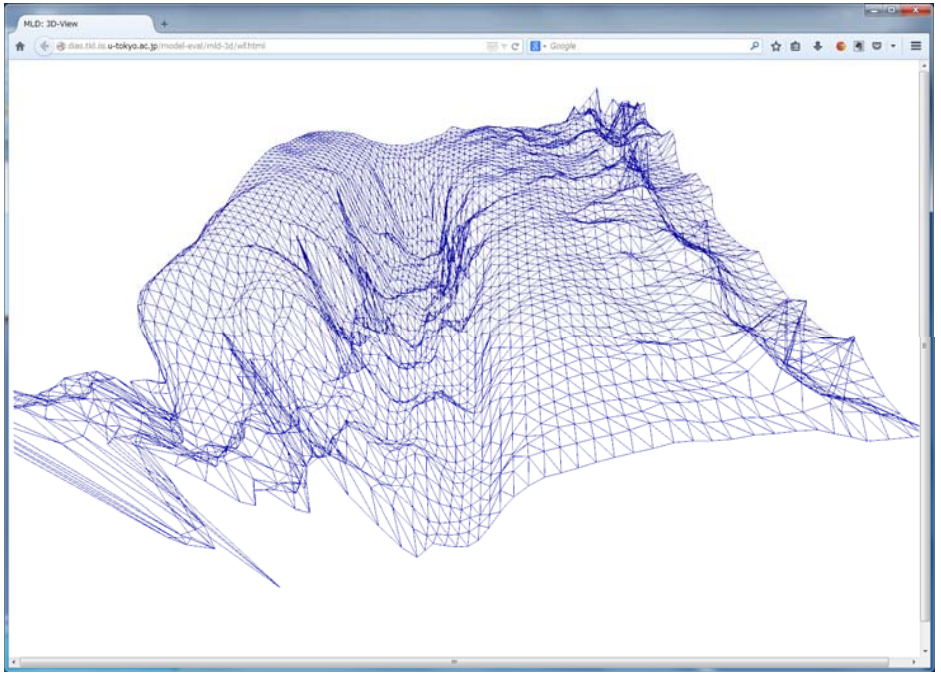
Increasing tendency predicted

Application Example (2):

Fishery Habitat Prediction



Mixed Layer Depth



3D Visualization of Mixed Layer Surface

→ Applicable to predict fishery habitat of neon flying squid

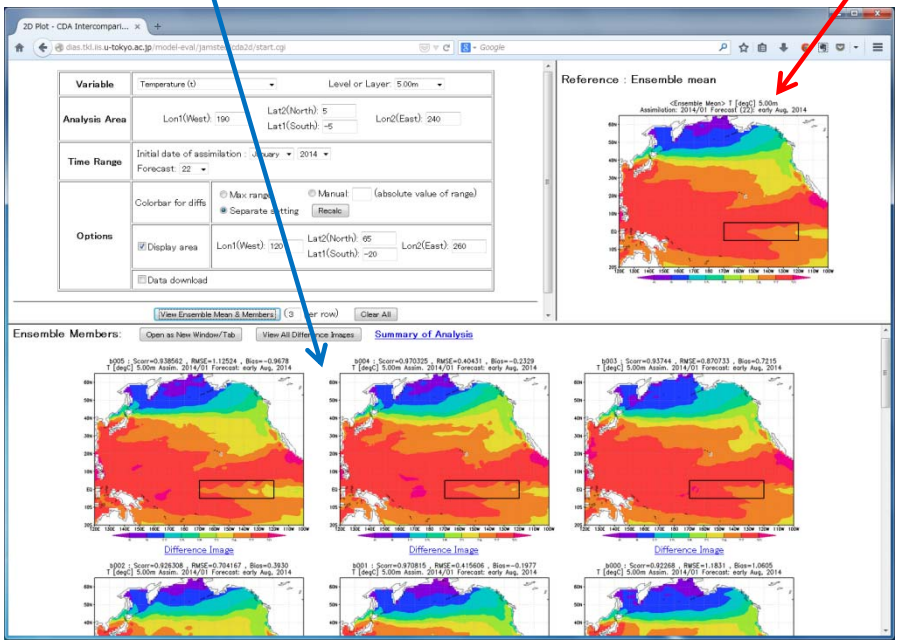
Application Example (3):

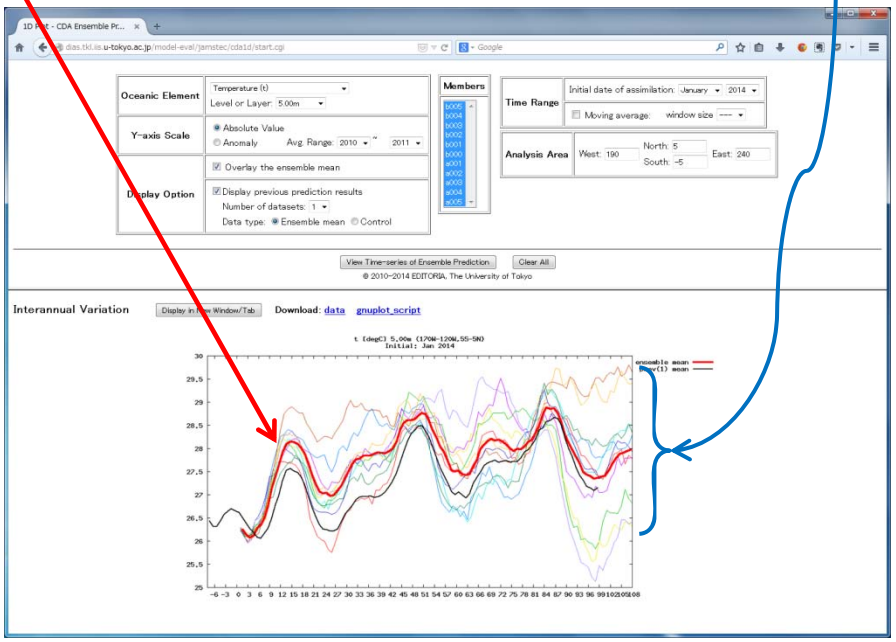
SST Monitoring and Prediction

Ensemble members

Ensemble mean

Ensemble members





2D Distribution of Ensemble Prediction

Time Series of Regional Ensemble Prediction

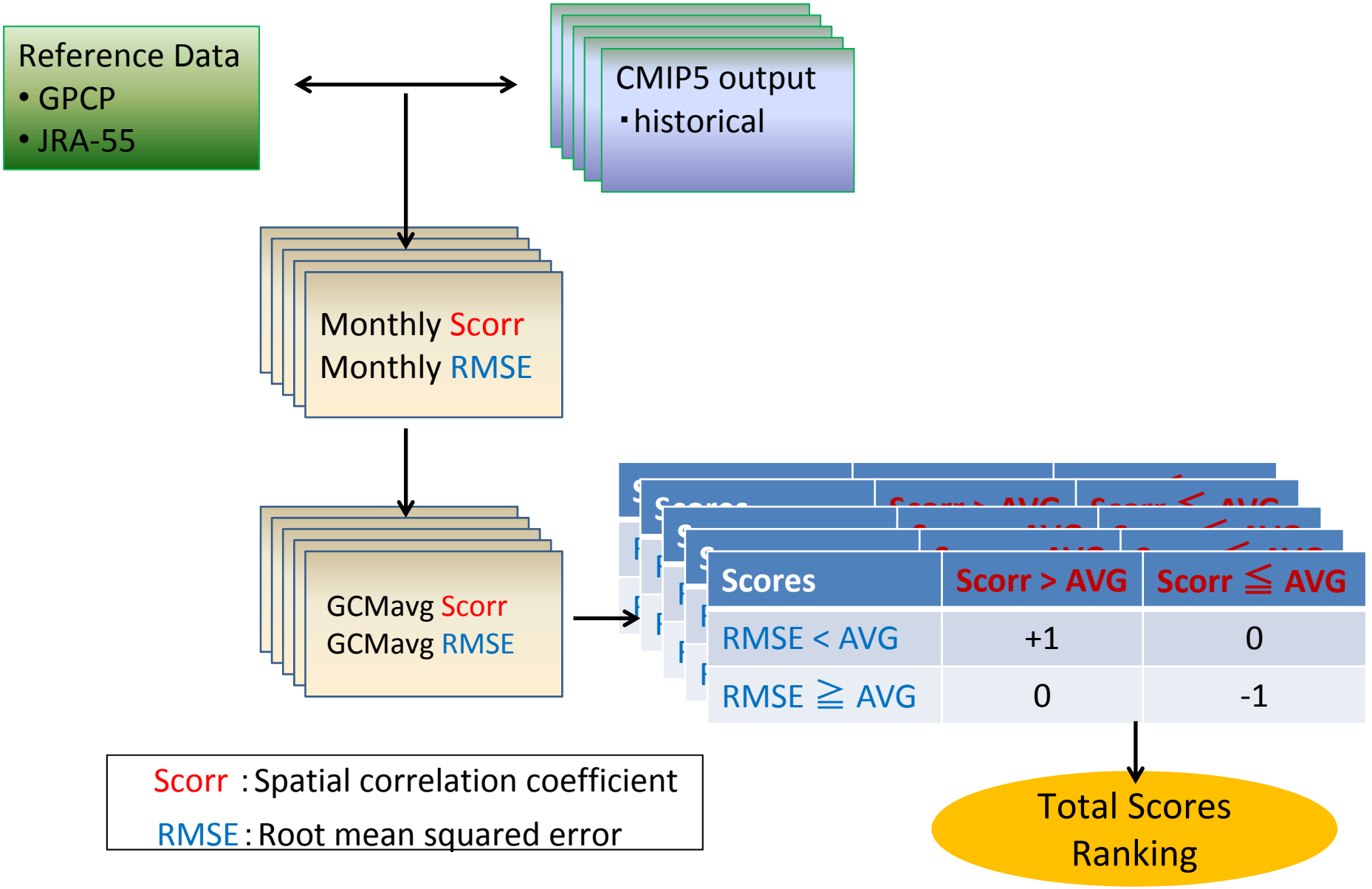
Application Example (4):

Bias Correction of Daily Precipitation

(1) Model Selection

- Pick out good models based on the following statistical measure
  - Spatial correlation coefficient
  - Root mean squared error

Stats Calculation and Ranking

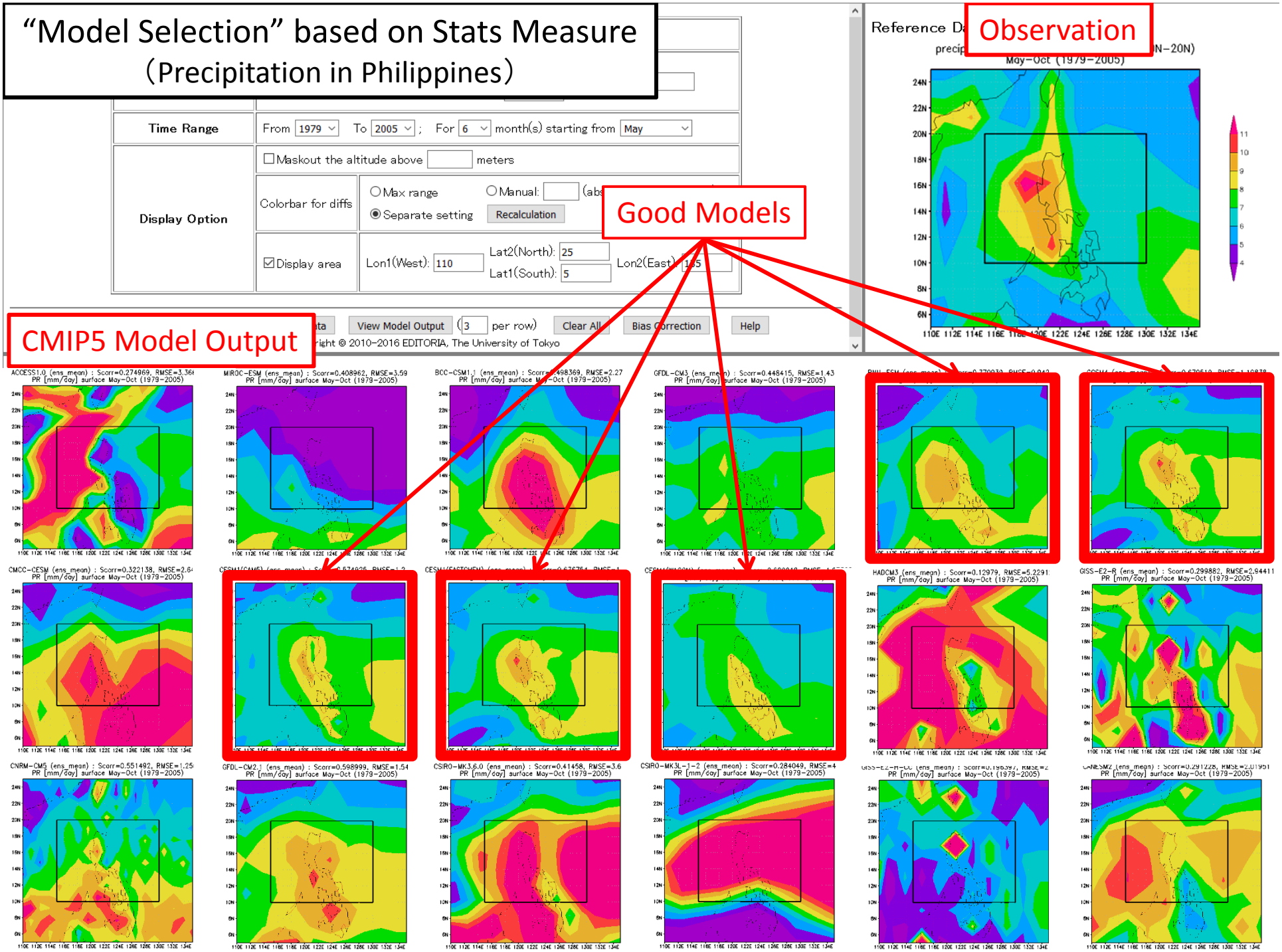


“Model Selection” based on Stats Measure
（Precipitation in Philippines）

Observation

Good Models

CMIP5 Model Output



2D Plot - Intercomparison - Mozilla Firefox

AOGCM Quantitative Evaluation | 2D Plot - Intercomparison

Summary of Analysis Results

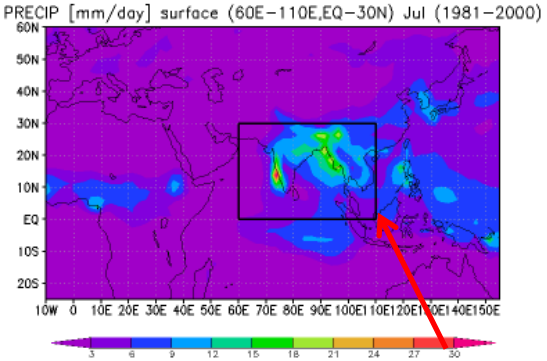
Download : CSV file

Model	Scorr	RMSE
bccr_bcm2_0	0.560668	3.88737
cccma_cgcm3_1	0.489887	4.15787
cccma_cgcm3_1_t63	0.479107	4.21173
cnrm_cm3	0.522781	4.0102
csiro_mk3_0	0.602058	4.07465
csiro_mk3_5	0.462052	4.52435
gfdl_cm2_0	0.722062	3.28271
gfdl_cm2_1	0.670102	3.47762
giss_aom	0.297611	5.14706
giss_model_e_h	0.433207	6.63656

Summary of Analysis Results;
CSV file downloadable

Reference Data: GPCP

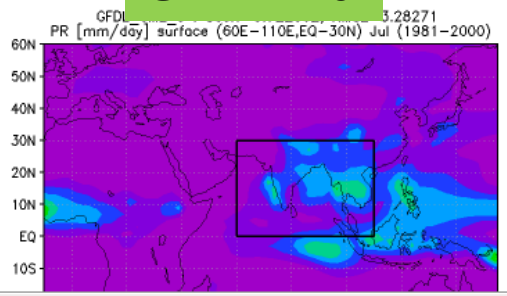
GPCP



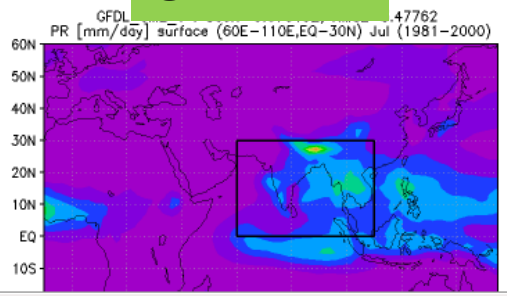
Analysis Area

Spatial Correlation (Scorr) and RMSE are good metrics of model reproducibility.

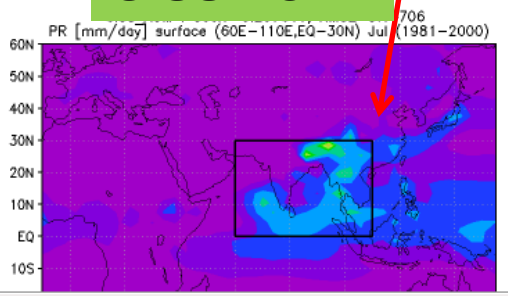
GFDL 2.0



GFDL 2.1

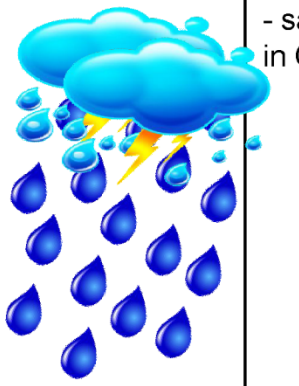
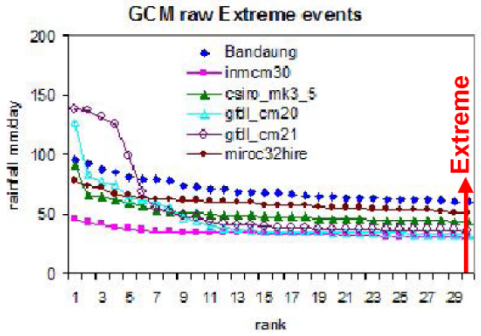
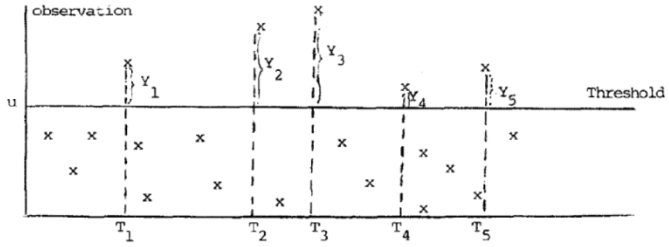
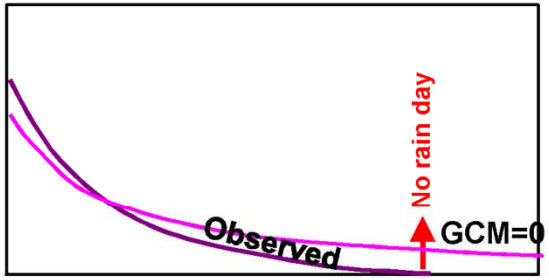
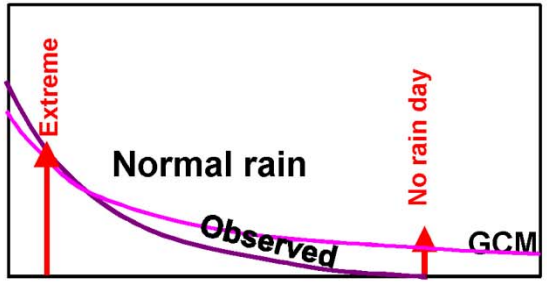
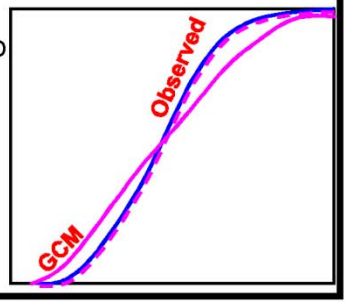


GISS AOM

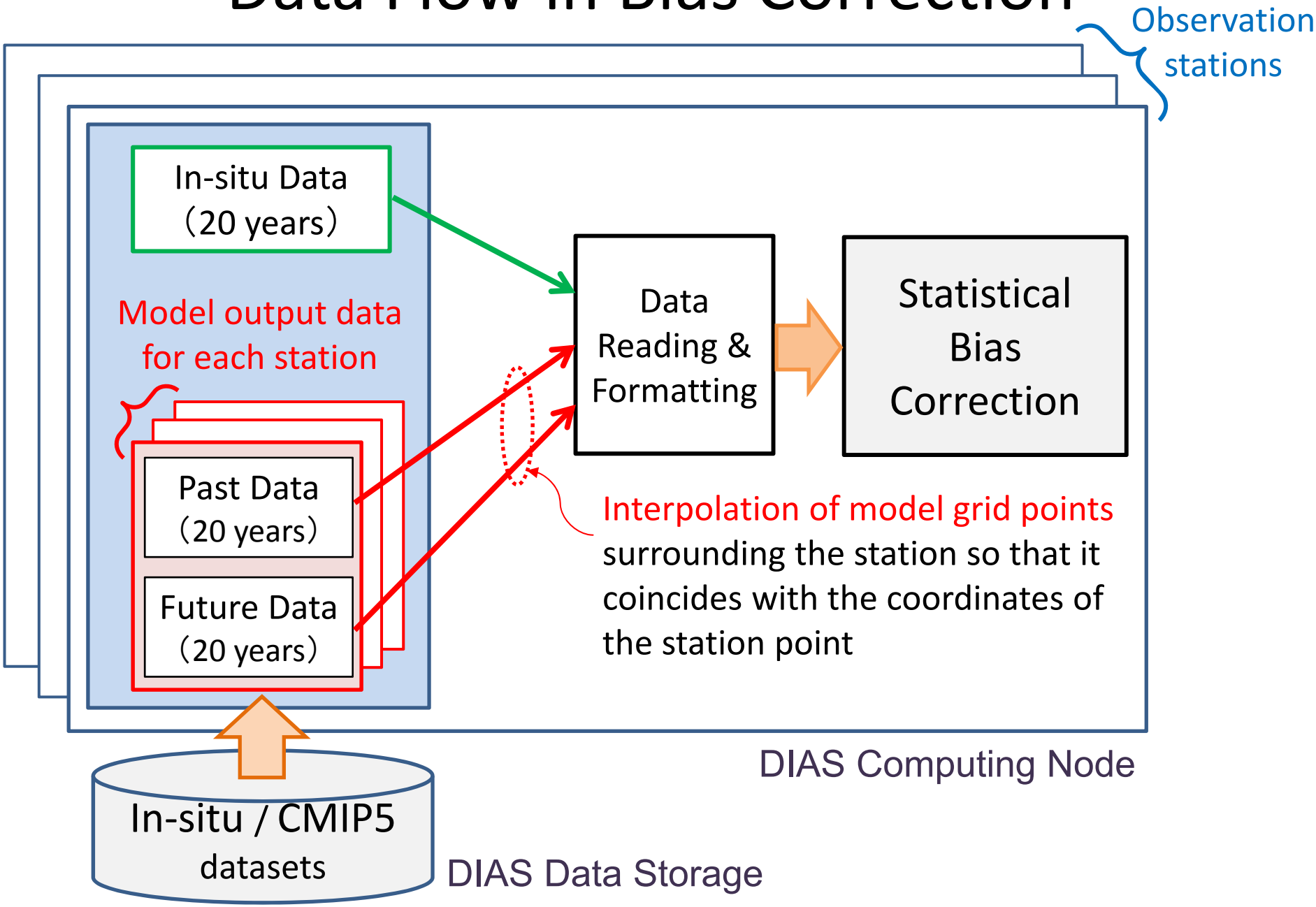


完了

(2) Statistical Bias Correction

Rain Type	Threshold	Correction
Extreme 	- > 99% of daily precipitation during analysis period - same frequencies of extreme as insitu station as in GCM 	**Generalized Pareto Distribution** -Non every year statistics -Extreme (long or short tailed) fitting -Peak over threshold method  Fig. 2. Illustration of threshold model.
No rain day 		**Ranking order statistics** - frequency of no rain day in GCM is same as station - less than no rain day threshold change zero rainfall.
Normal 		**Gamma Distribution** - monthly CDF of GCM mapping to monthly CDF of station - inverse of Gamma CDF in each month is corrected rain 

Data Flow in Bias Correction

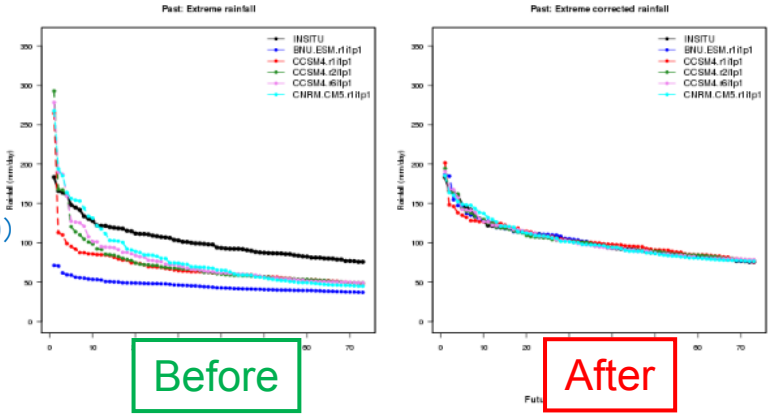


Example of Bias Correction Result

CMIP5 Bias Correction (in-situ)

Extreme Rainfall

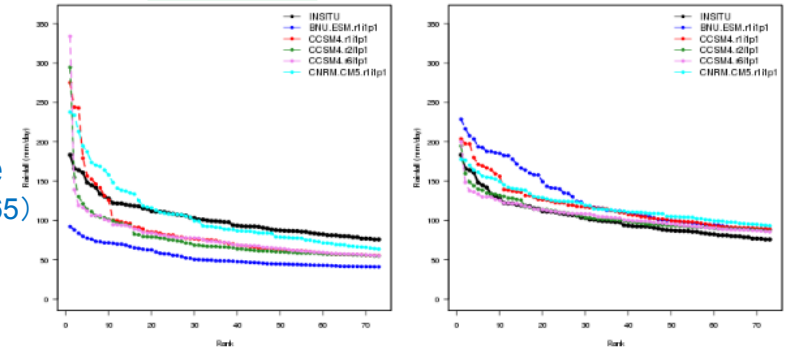
Past
(1981-2000)



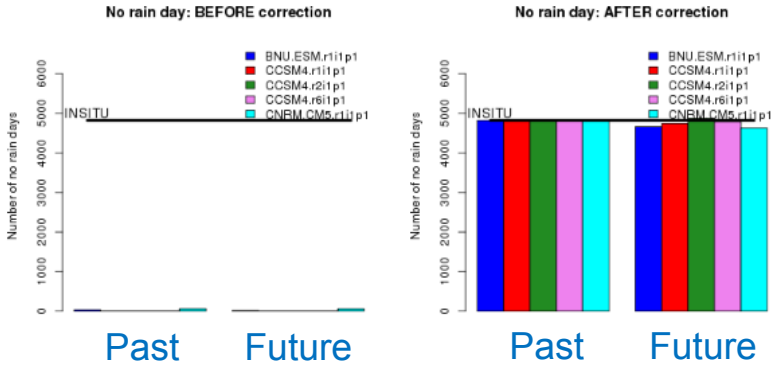
Before

After

Future
(2046-2065)



No-rain Days

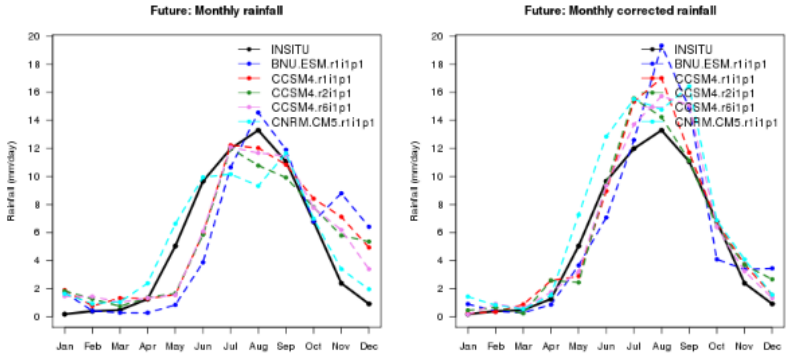


Before

After

Future Seasonal Rainfall

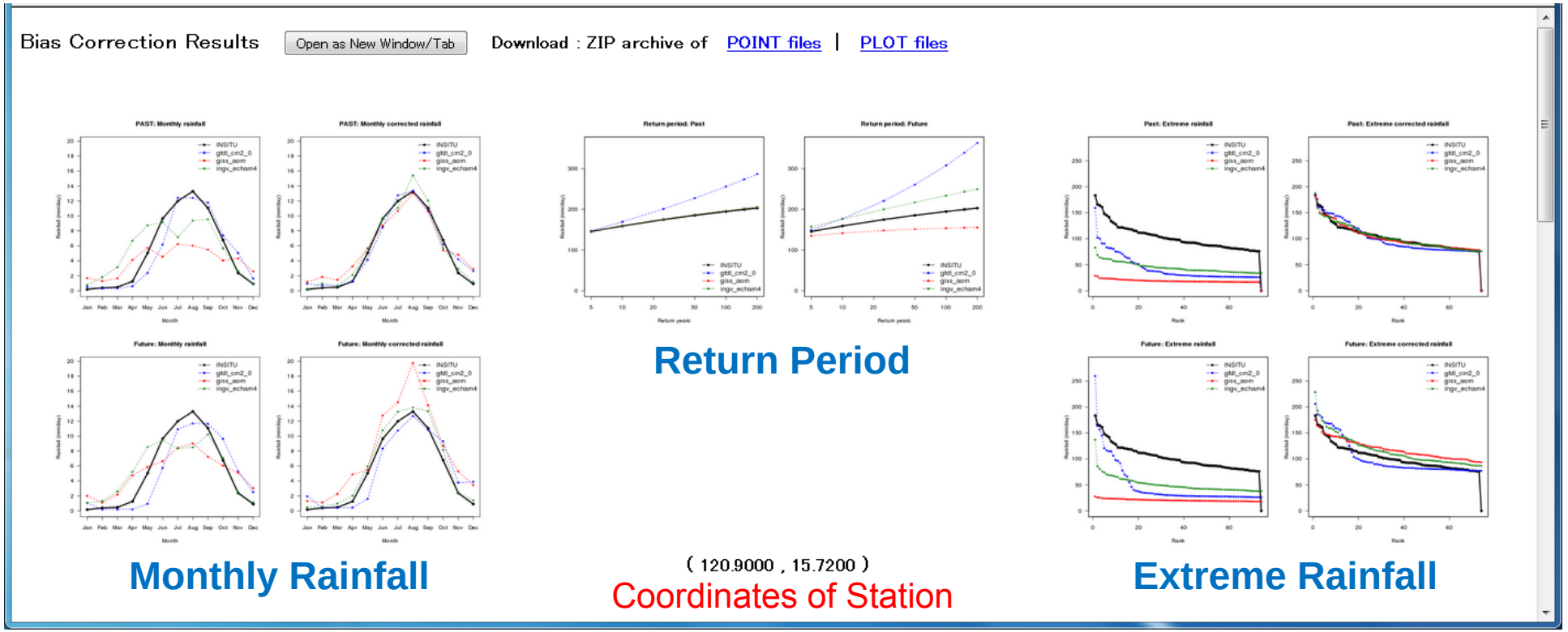
（2046~2065）



Before

After

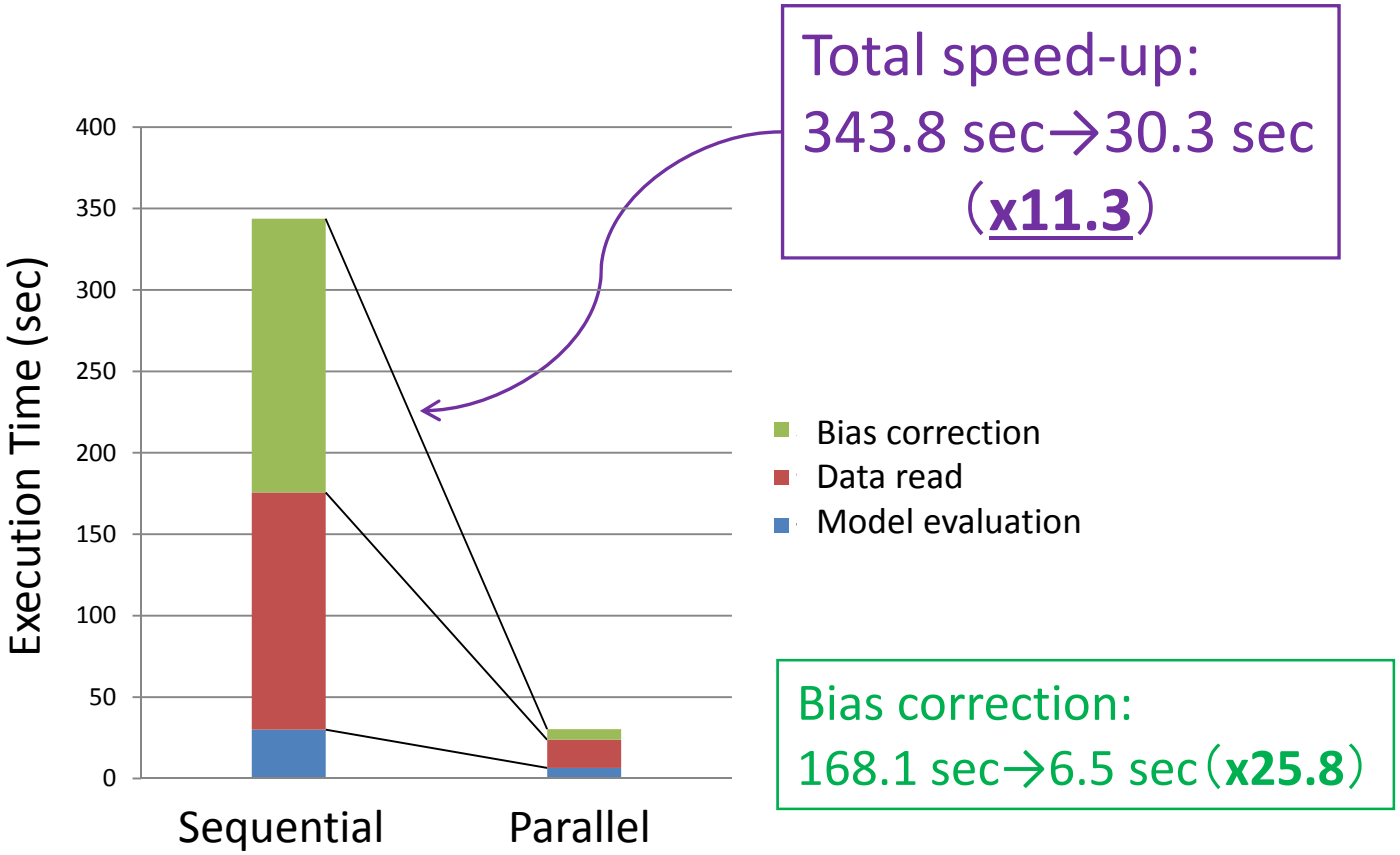
Example of Bias Correction Result



Speed-up by Parallel Processing

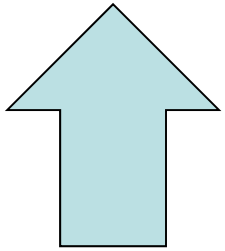
- Model Selection
  - Worker processes created for every model evaluation and data read task (**multi-process**)
- Bias Correction
  - Worker processes created for every observation station, and multi-threaded bias correction program executed (**hybrid multi-process multi-thread**; i.e., running multiple multi-threaded processes)

Performance Evaluation



- Shared-memory Parallel Computer (64 core)
- 16 models; 8 observation stations

BIAS Correction of CMIP5 Data



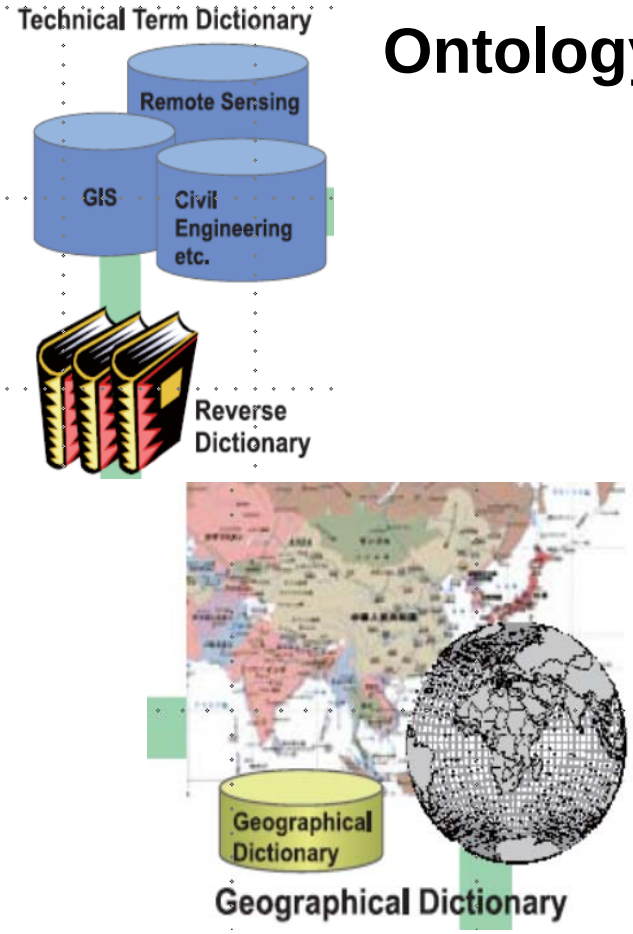
Observed and simulated data
are archived
is one of “DIAS Value”

Data Integration and Analysis System

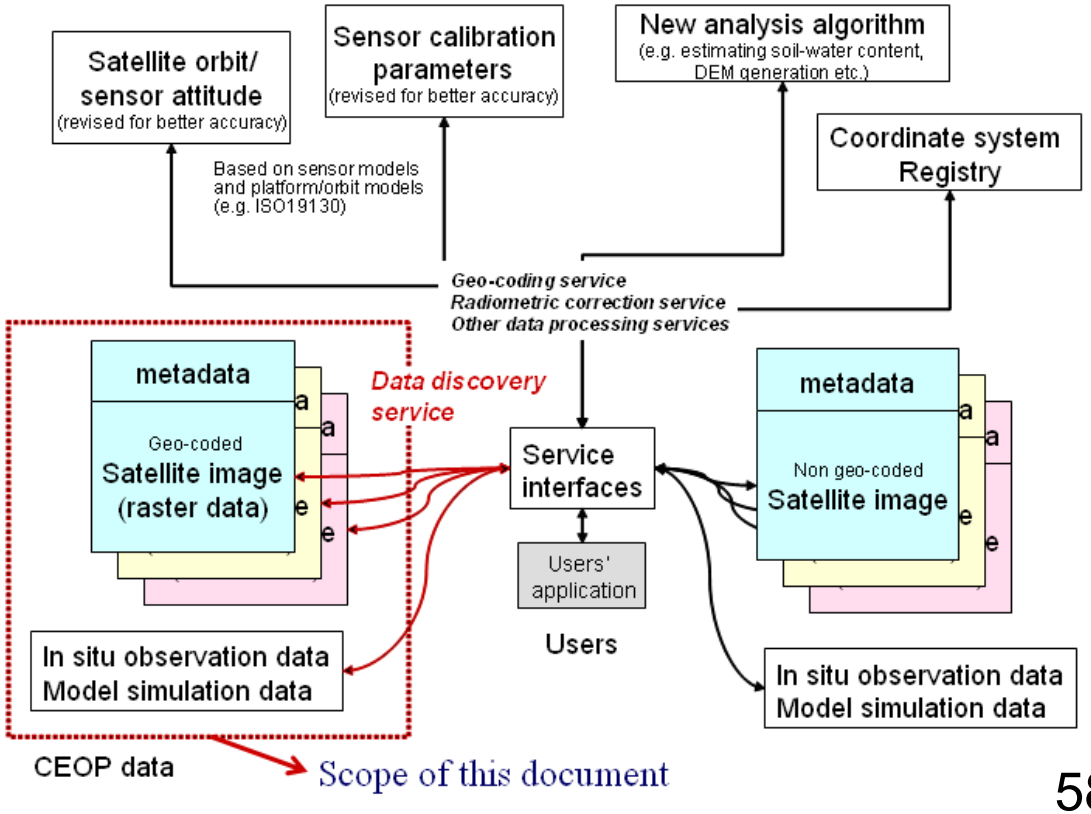
a legacy for Japan's contributions to GEOSS

tackling a large increase in variety of the Earth observation data.

Ontology



Meta data Design



METADATA FOR DIAS DATASETS

DIAS Data Integration & Analysis System 地球環境情報統融合プログラム

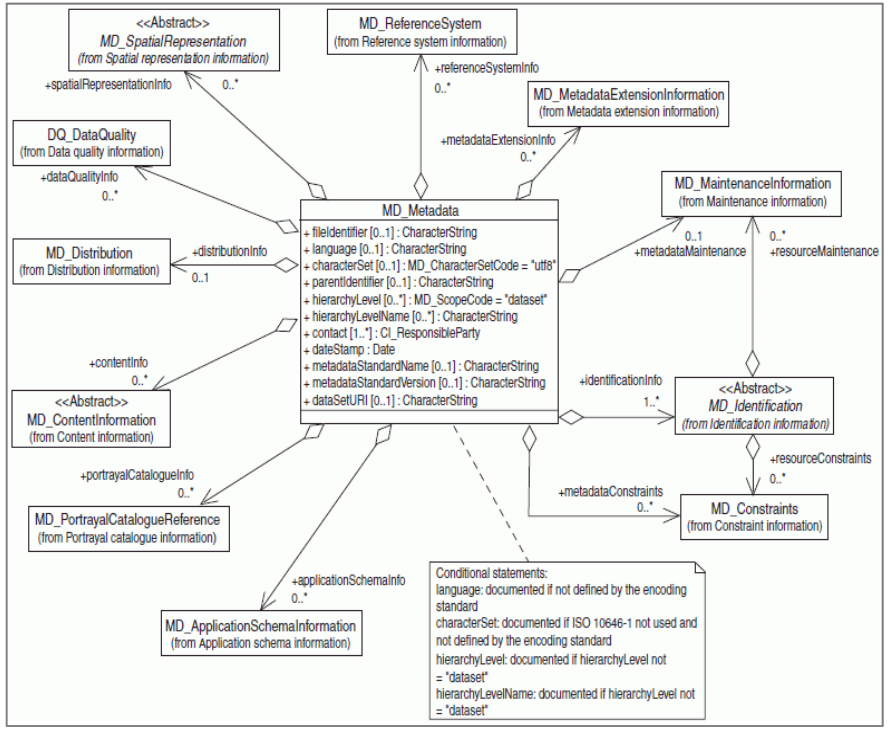
DIAS Metadata

- Adopt the XML metadata used in geographic information system ISO19115 (ISO19139)
- Metadata about dataset : Document-metadata
  - For the purpose of search for and comparison of dataset
  - The granule of dataset is coarse, in general, and is decided after discussion with data provider
  - Files which include both data and metadata (such as those in netCDF) are not included in our target datasets.
- Once metadata is created, dataset documents are automatically generated in HTML, PDF.
- Data providers have to publish datasets with Document-metadata.

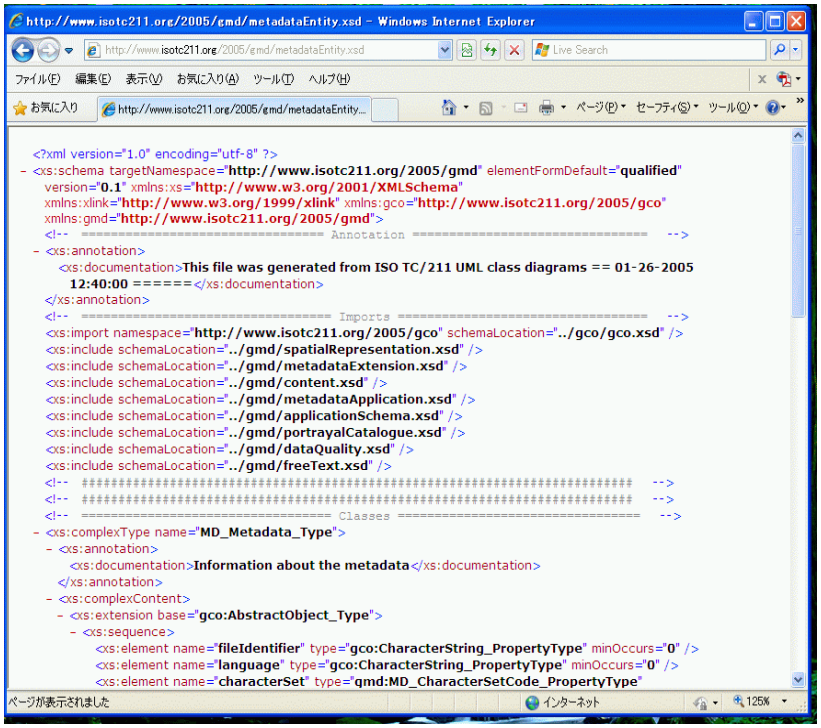
DIAS Data Integration & Analysis System 地球環境情報統融合プログラム

ISO19115 / ISO19139

- Most of earth environmental data commonly have spatial and temporal attributes such as the covering geographic scope or the created date.
- Accordingly, DIAS metadata is developed with basing on ISO/TC211 metadata standards.



ISO19115

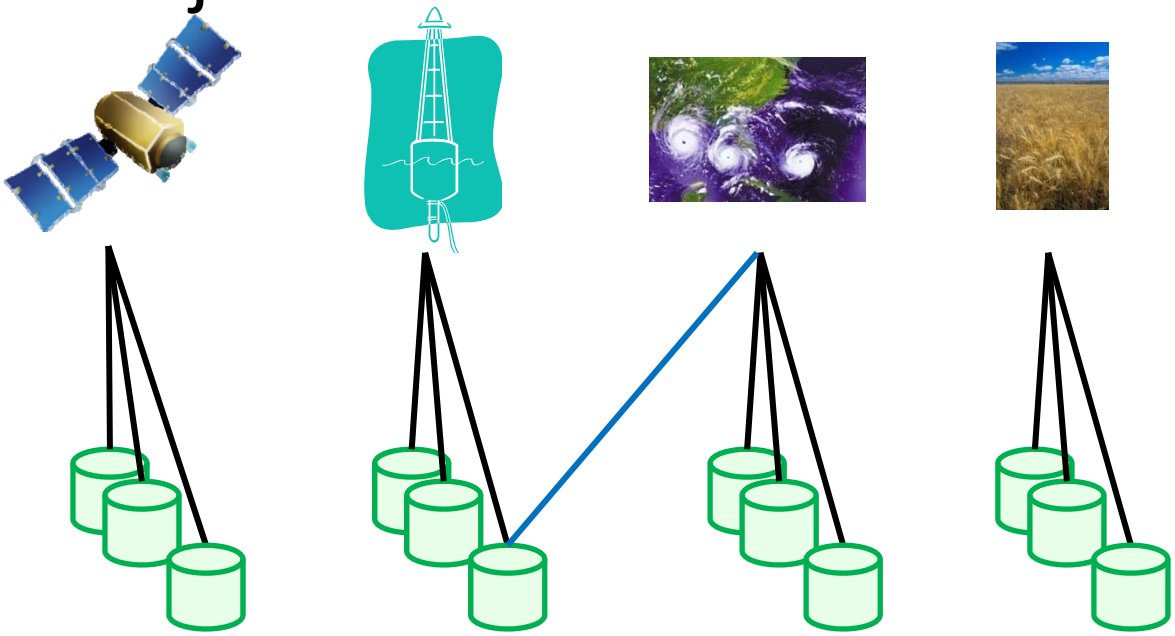


ISO19139

DIAS Data Integration & Analysis System 地球環境情報統融合プログラム

DIAS Metadata =

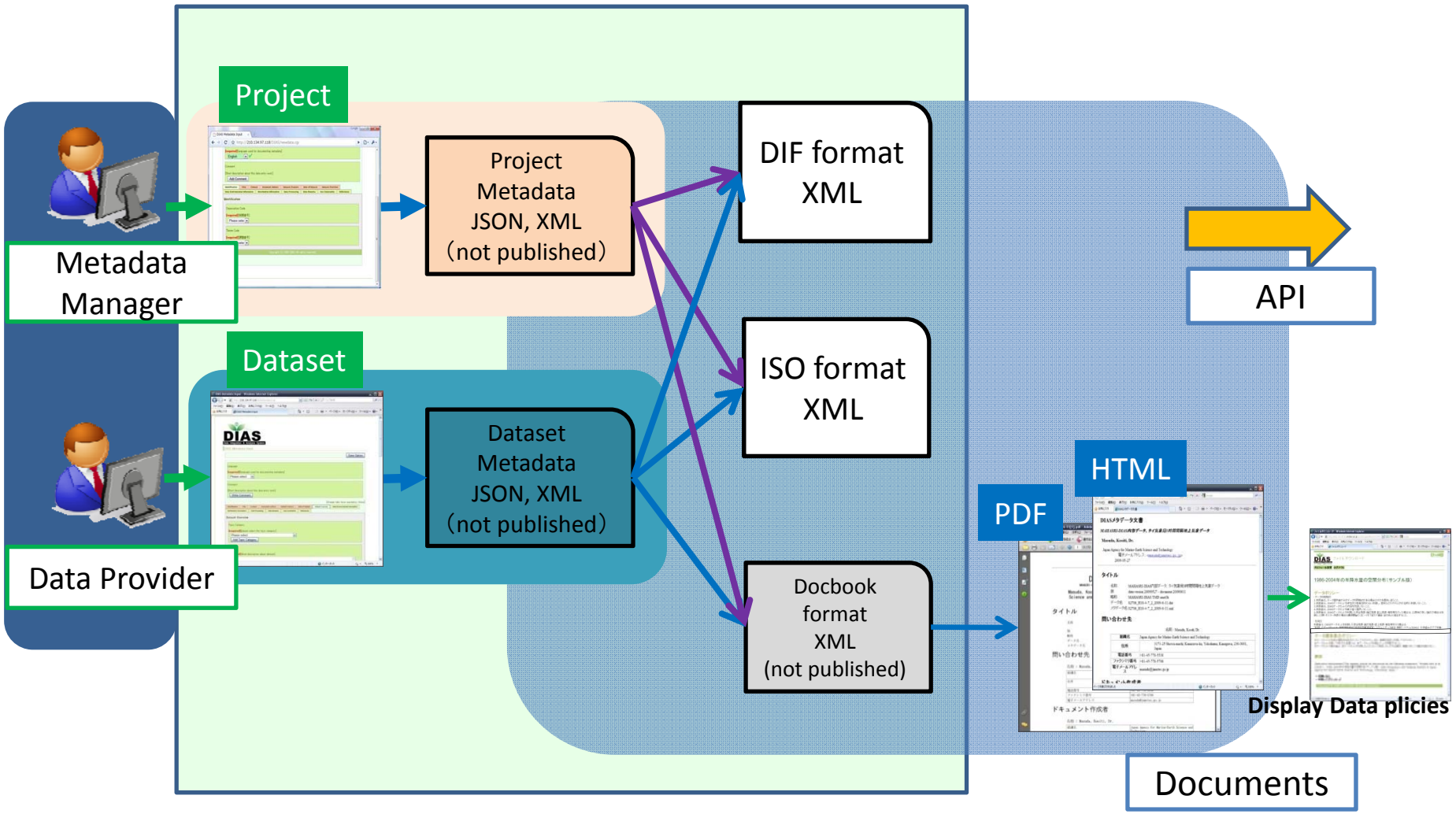
Project metadata + Dataset metadata



- ***Project Metadata***: Project name, Contact address, Keywords, Web site address, Project Data policy, Data Disclaimer, and Acknowledgement
- Datasets that have been created over several projects can be managed by our metadata management system.

DIAS
Data Integration & Analysis System
地球環境情報統融合プログラム

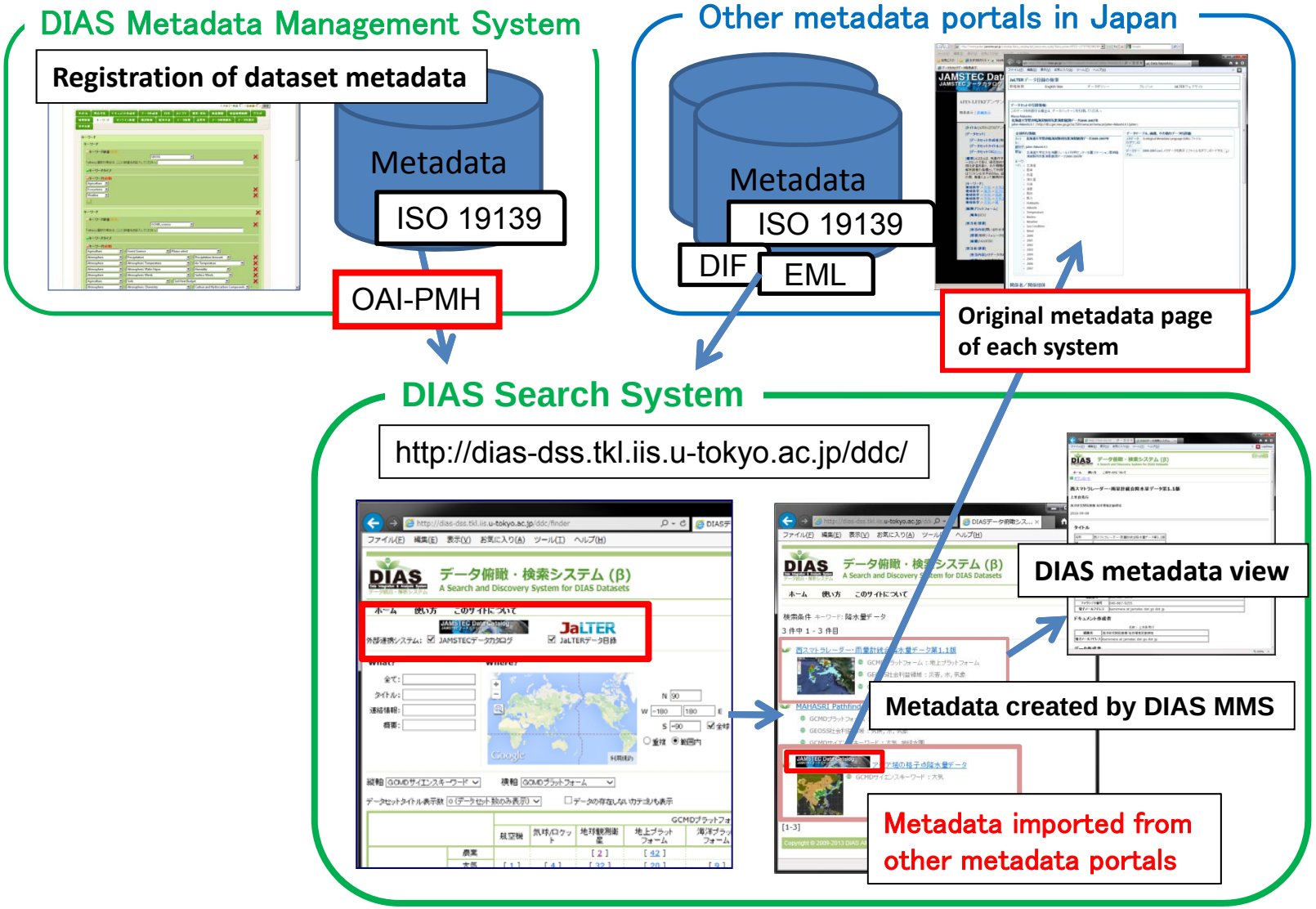
Project and Dataset Metadata



DIAS METADATA SYSTEMS

DIAS Data Integration & Analysis System 地球環境情報統融合プログラム

Architecture of Metadata System



DIAS
Data Integration & Analysis System
地球環境情報統融合プログラム

The Workshop for DIAS metadata input

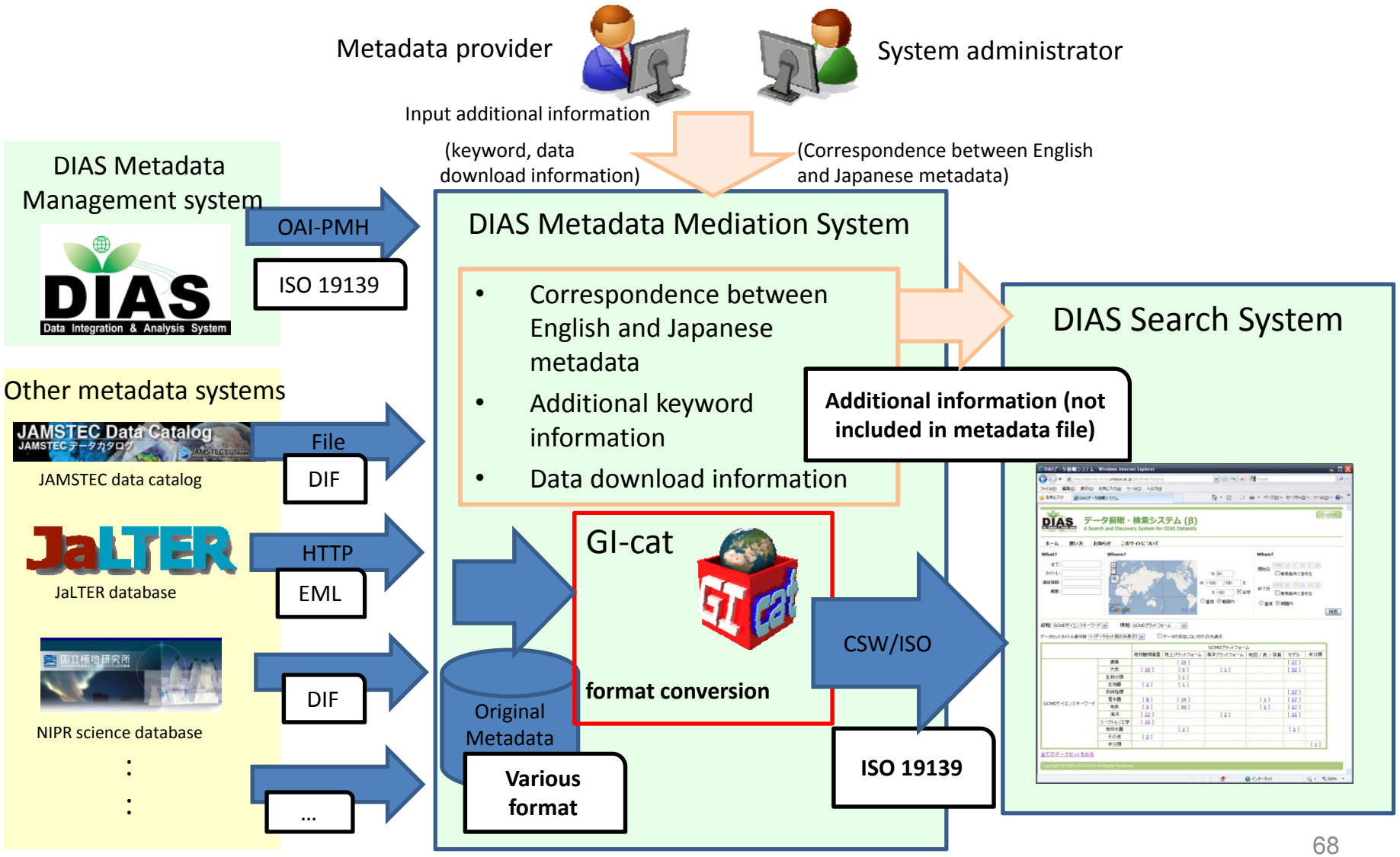
The contents of metadata cannot know except the data provider.
we have held the workshop about DIAS metadata input
 in order to understand what kind of information
should be inputted for each field of metadata.
The workshop has held on July 2014, November 2014 and February 2015.

DIAS Data Integration & Analysis System 地球環境情報統融合プログラム

Metadata Cooperation with external systems

	Metadata format	URL
JAMSTECデータカタログ	DIF	http://www.godac.jamstec.go.jp/catalog/data_catalog/
JaLTERデータ目録	EML	http://db.cger.nies.go.jp/JaLTER/
極地研究所学術データベース	DIF	http://scidbase.nipr.ac.jp/
極地研究所北極圏データアーカイブ (作業中)	ISO19139, DIF	https://ads.nipr.ac.jp/
国土地理院クリアリングハウス (作業中)	JMP2.0	http://ckan.gsi.go.jp/

Plan of the DIAS Metadata Mediation System



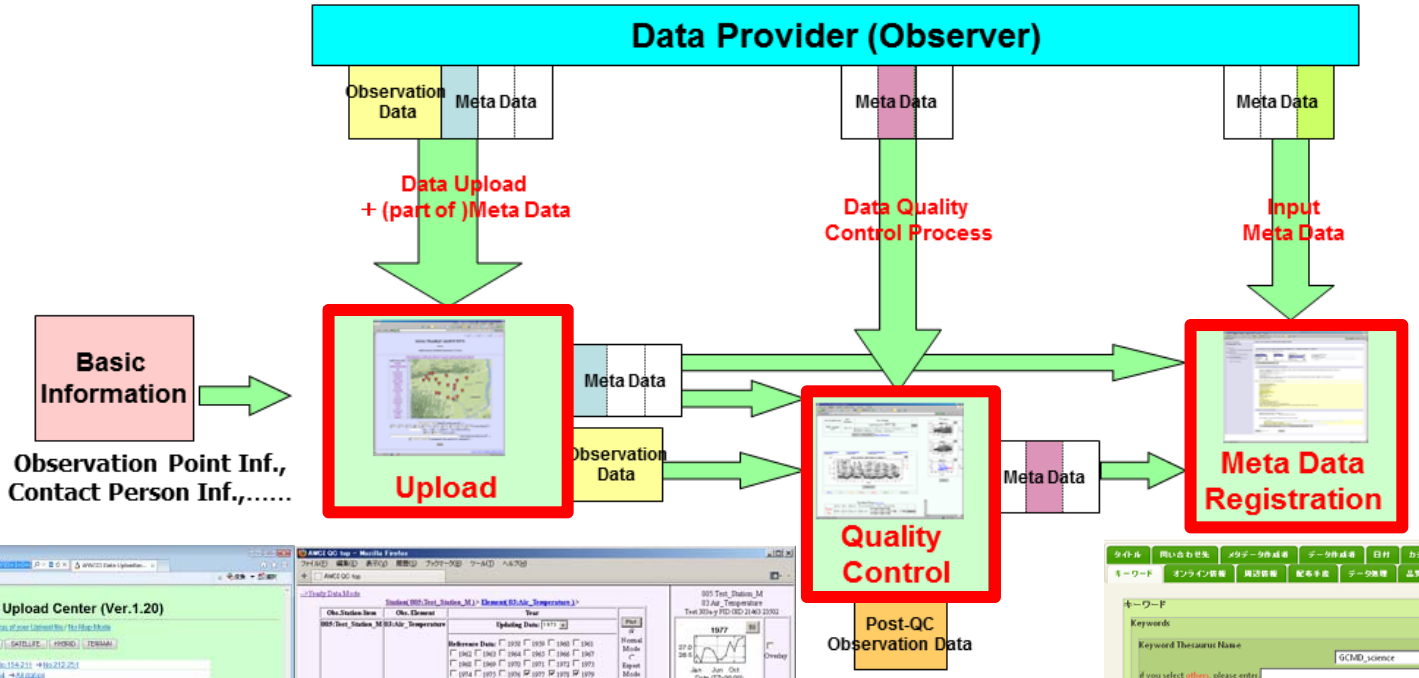
Metadata management system
support
the quality of data in DIAS

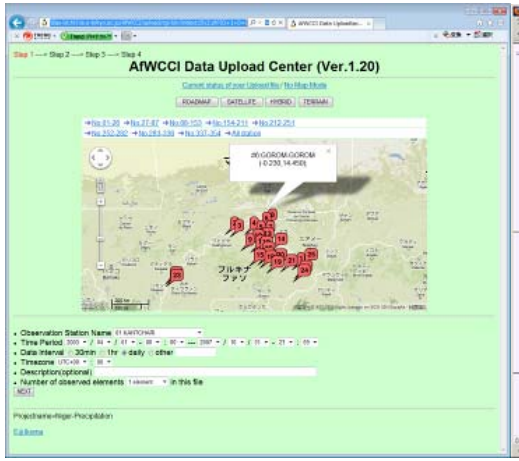
one of “DIAS Value”

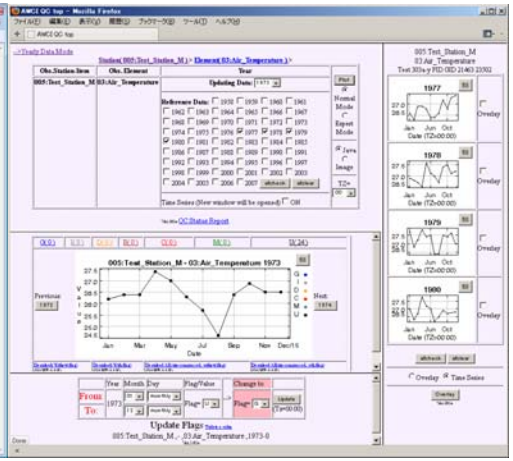
Data Integration and Analysis System

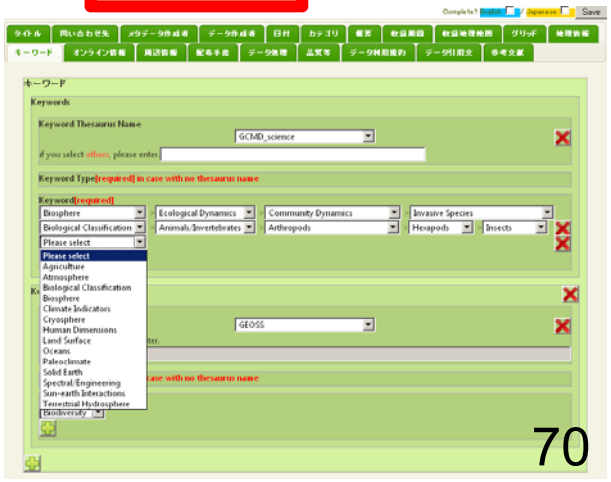
a legacy for Japan's contributions to GEOSS

accelerating data veracity, including data loading, QC and metadata registration







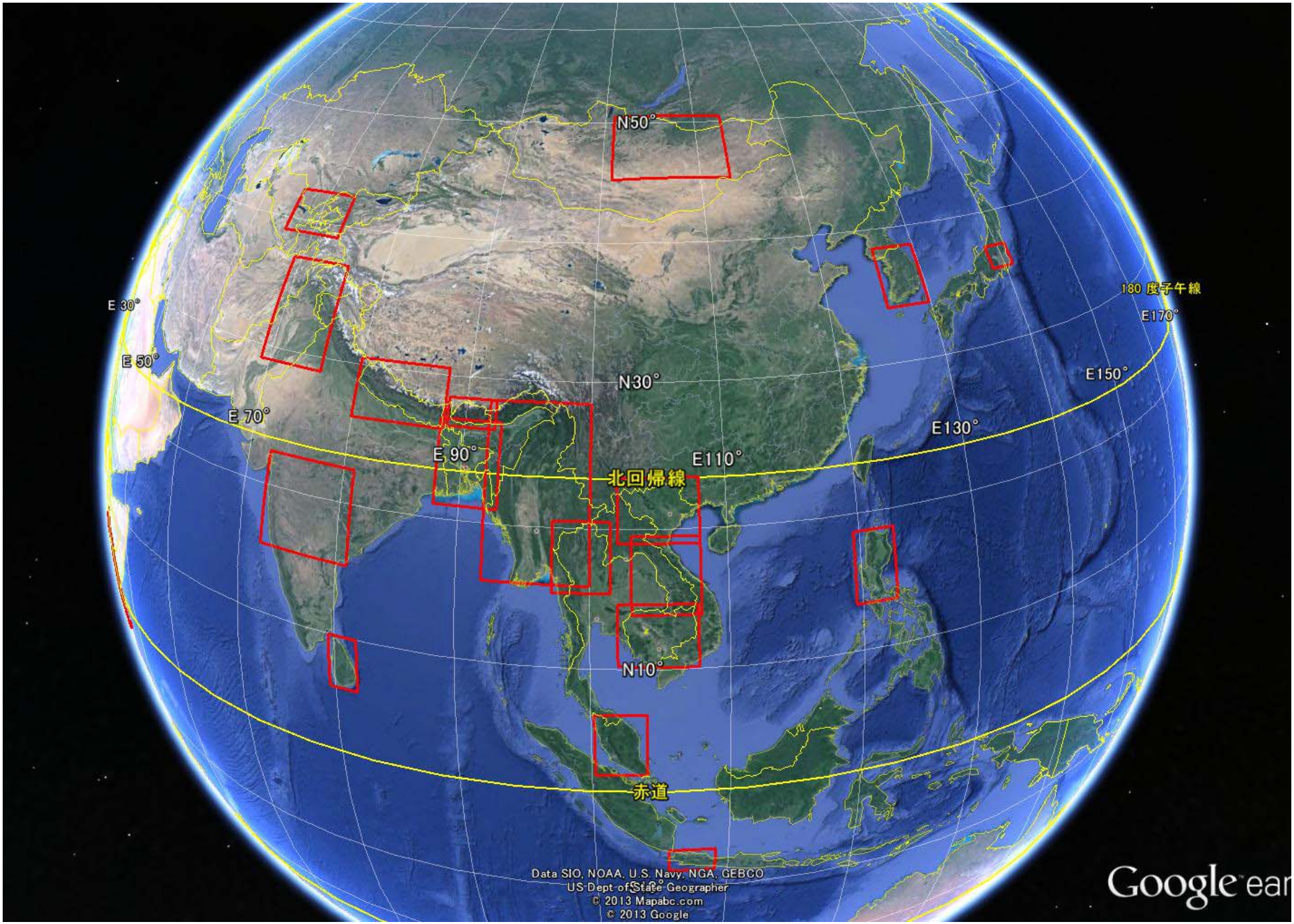


In-situ Data Management System on DIAS

- This system is running with some international research project:
  - AWCI
  - AfWCCI
  - AMY
- 4-component: Data Upload, Quality Control, Metadata-management, Data Download

AWCI :

Asian Water Cycle Initiative

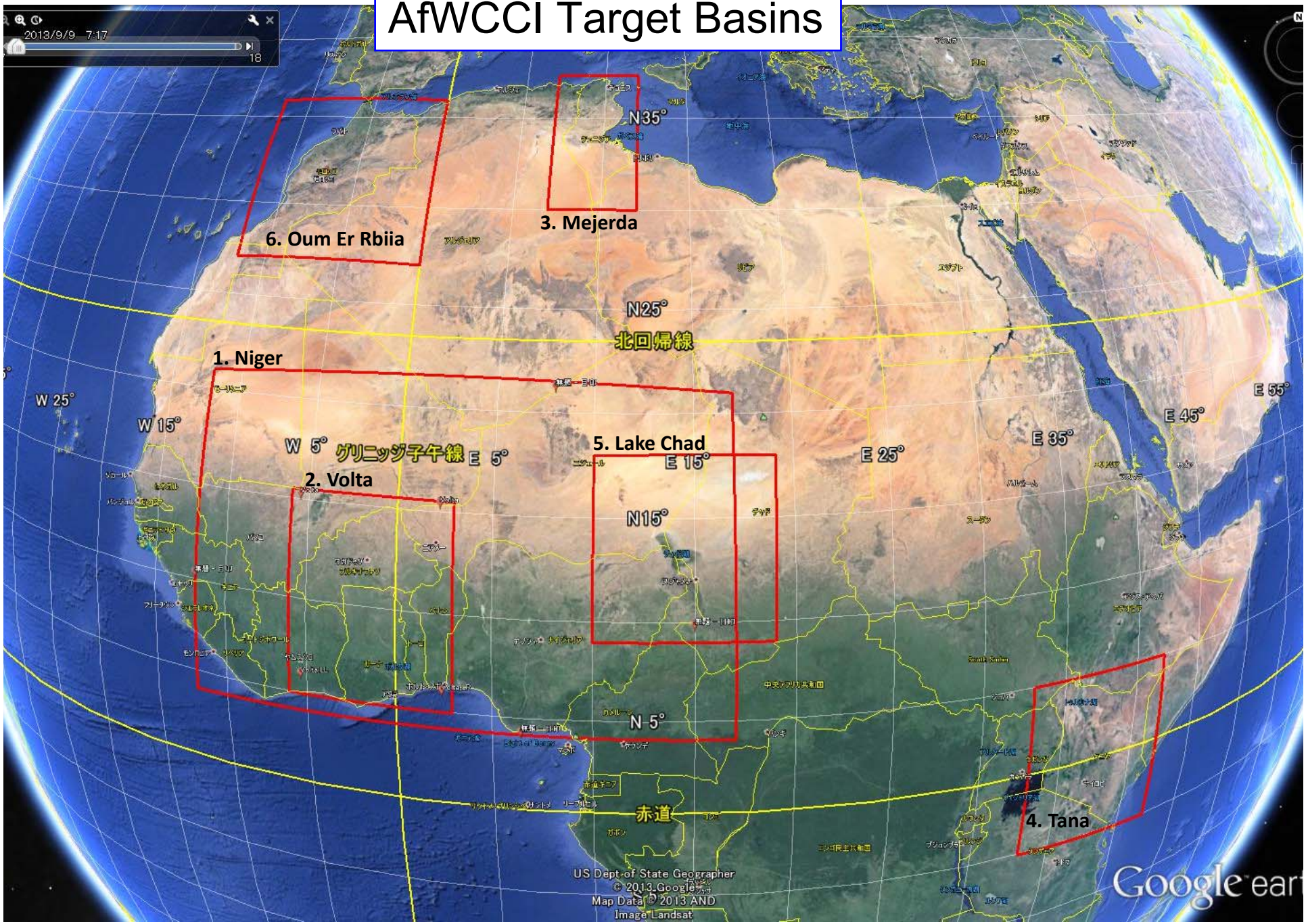

N50°
E 30°
E 50°
E 70°
E 90°
N30°
E110°
北回帰線
E130°
E150°
180 度子午線
E170°
N10°
赤道
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
US Dept of State Geographer
© 2013 Mapabc.com
© 2013 Google
Google ear

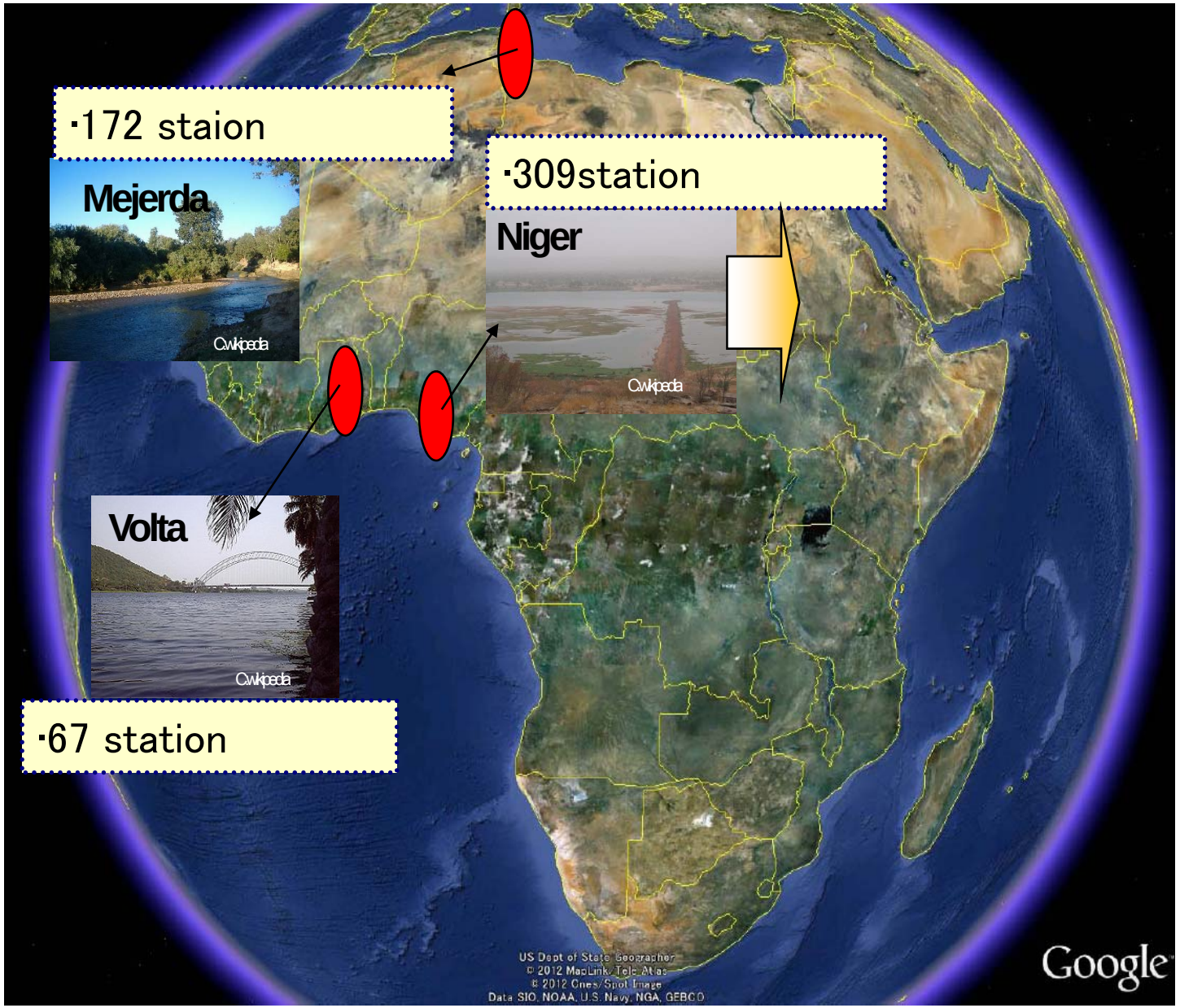
AWCI: 18 Demonstration Basin

AfWCCI :

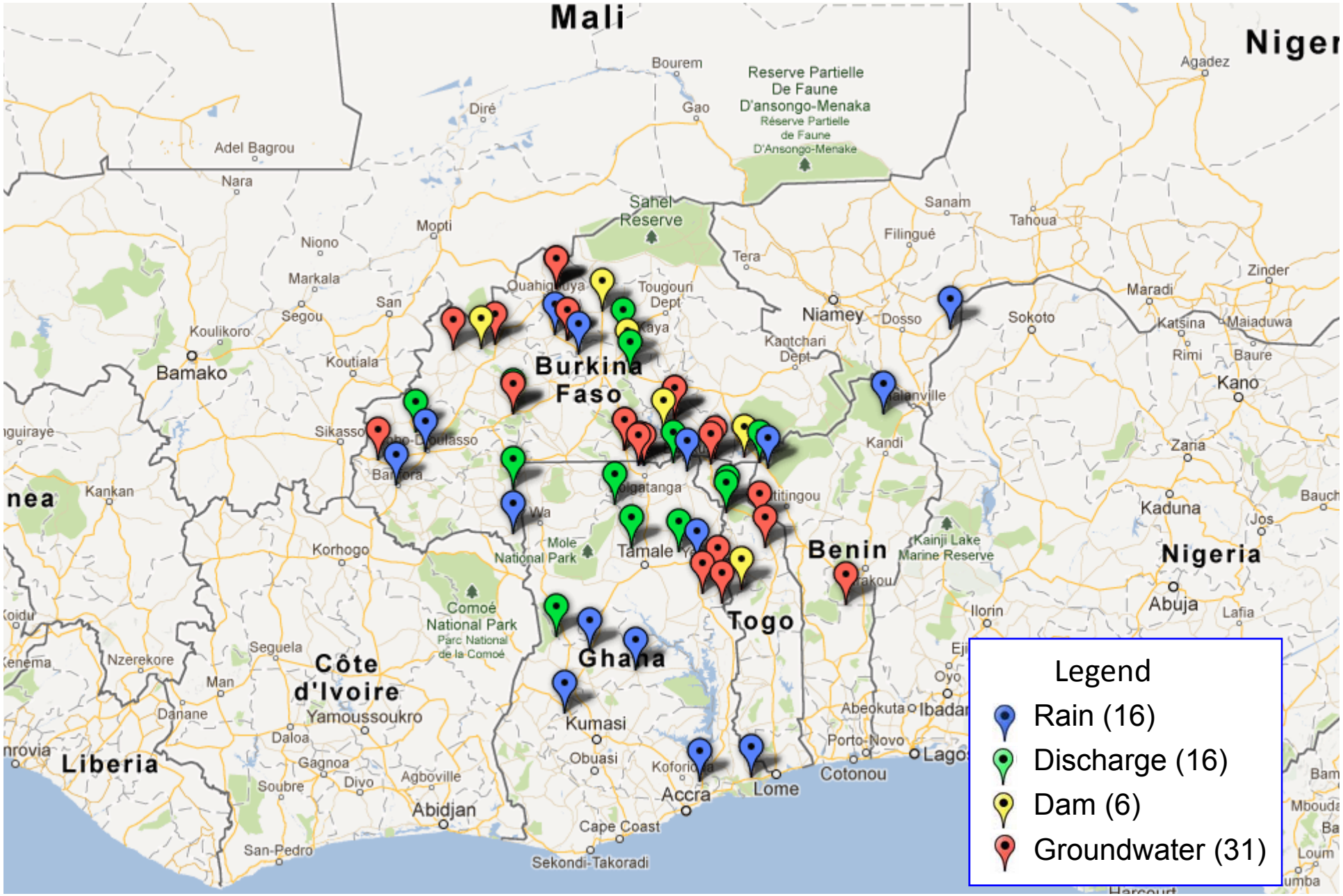
African Water Cycle Coordination Initiative

AfWCCI Target Basins

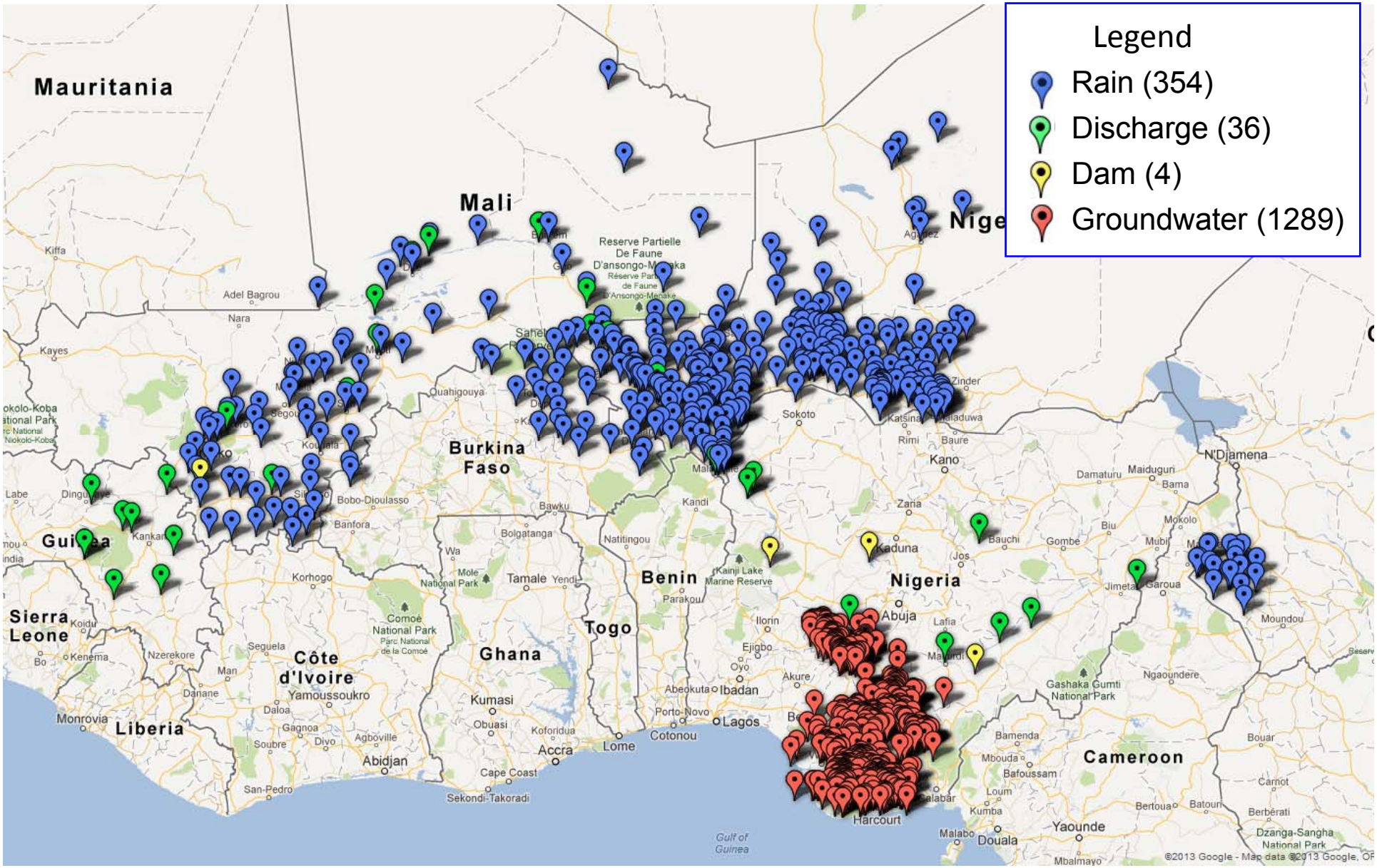



·172 staion
Mejerda
Cwikpeda
·309station
Niger
Cwikpeda
Volta
Cwikpeda
·67 station
US Dept of State Geographer
© 2012 MapLink/Tele Atlas
© 2012 Cnes/Spot Image
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Google

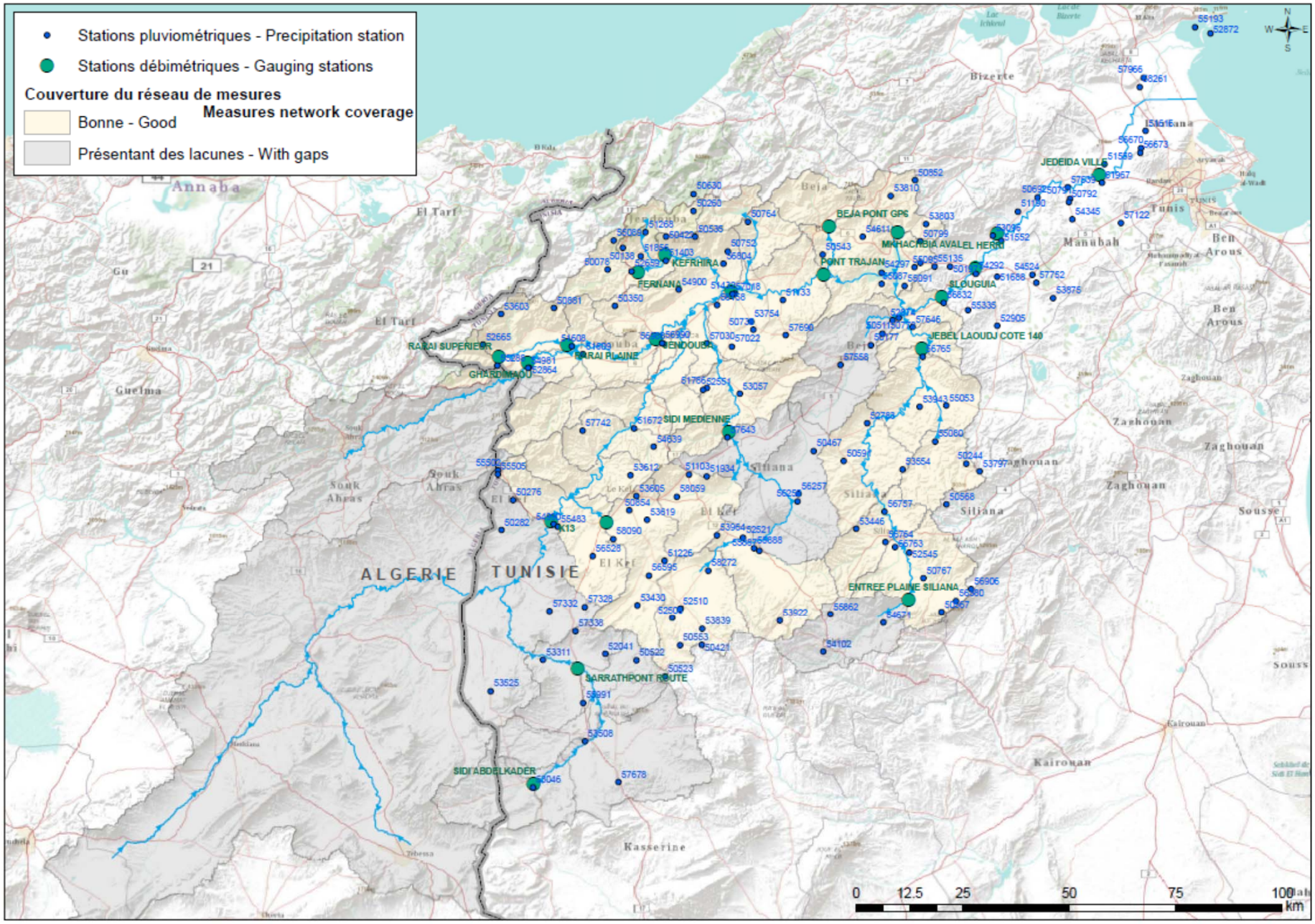
AfWCCI: Volta River : Station Map



AfWCCI: Niger River : Station Map



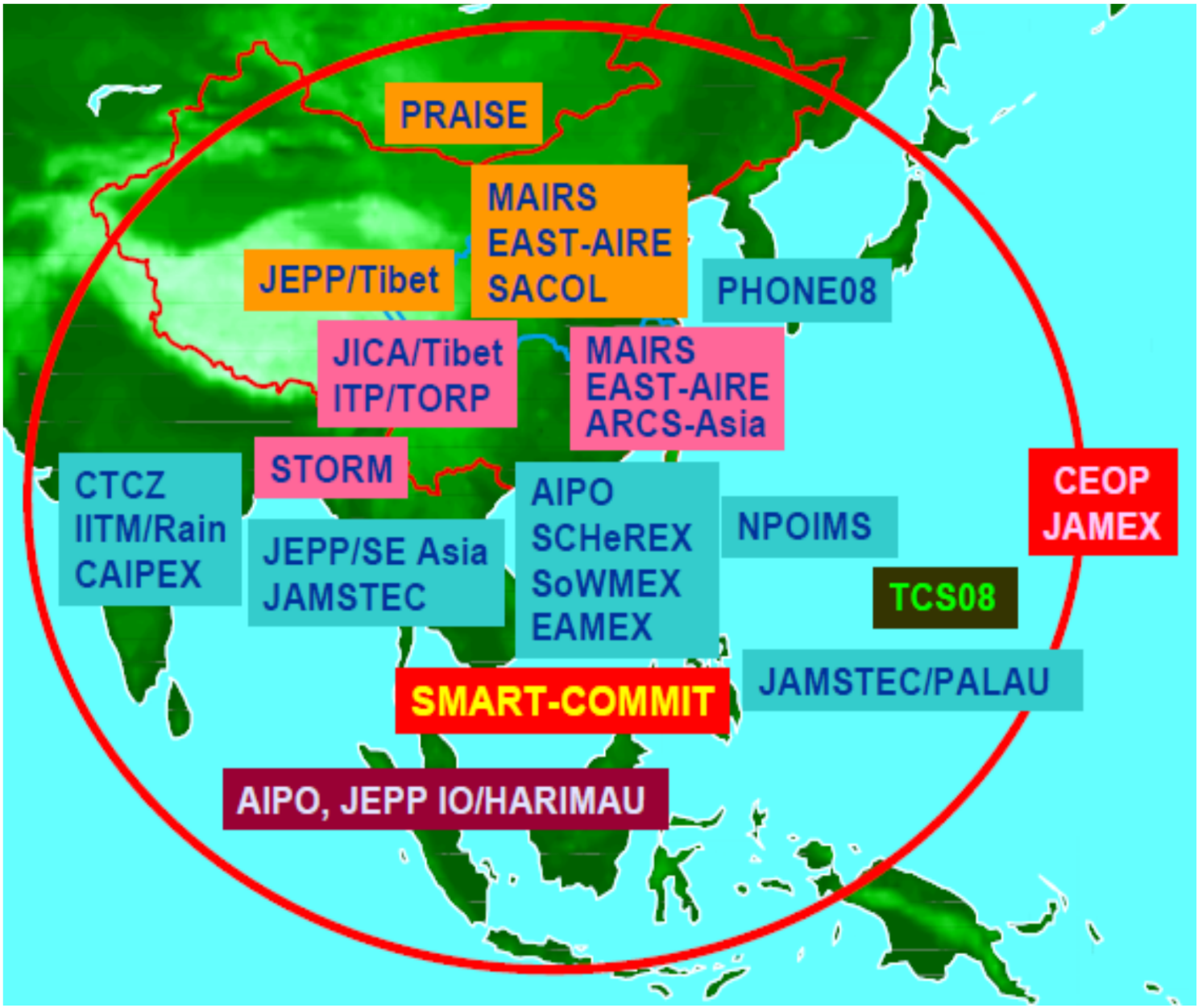
AfWCCI: Mejerda River : Station Map



AMY :

Asian Monsoon Year

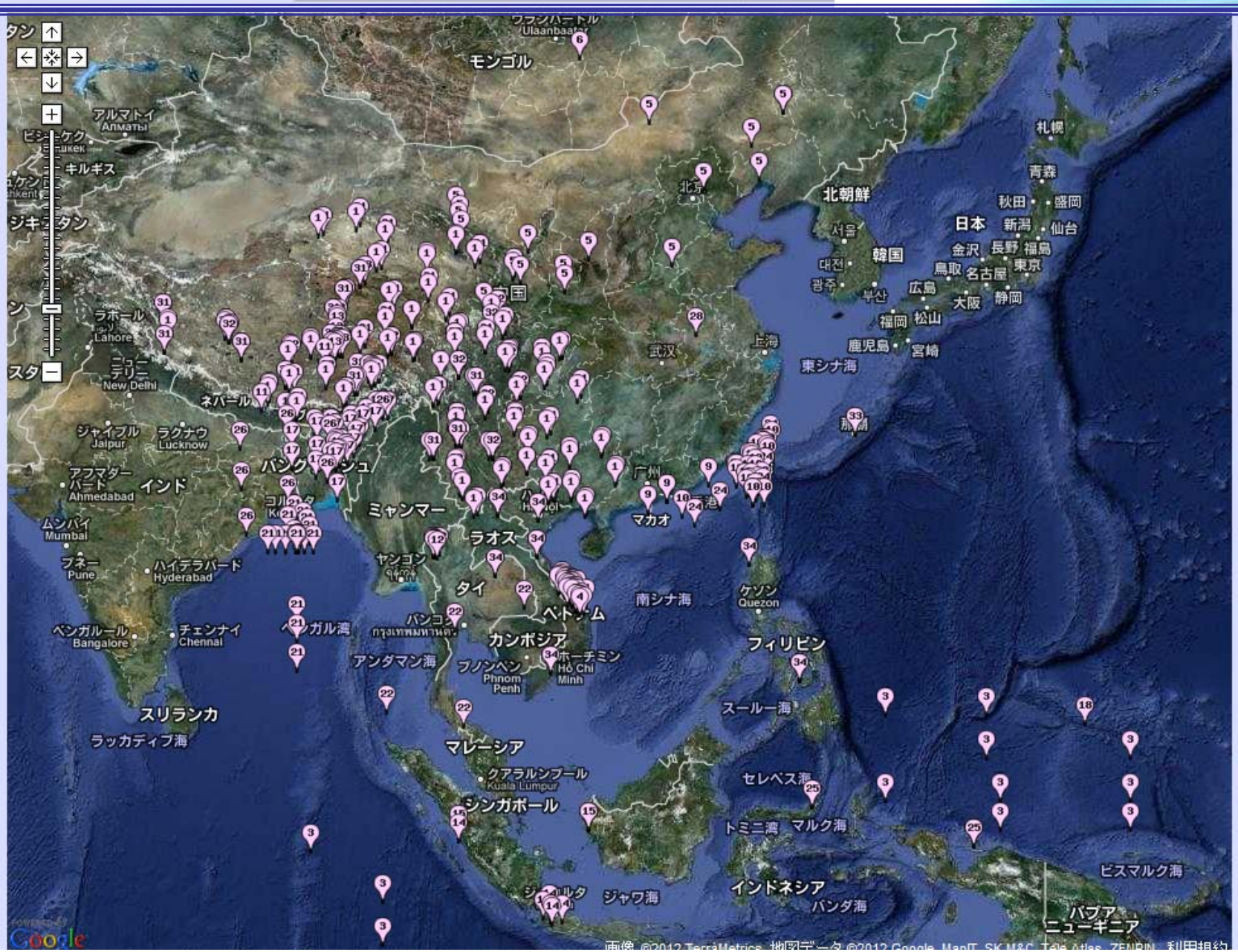
AMY: Project Map



GEWEX/AMY In-situ data Management status

Asian Monsoon Years (AMY 2007-2012)

Observation stations of AMY Projects

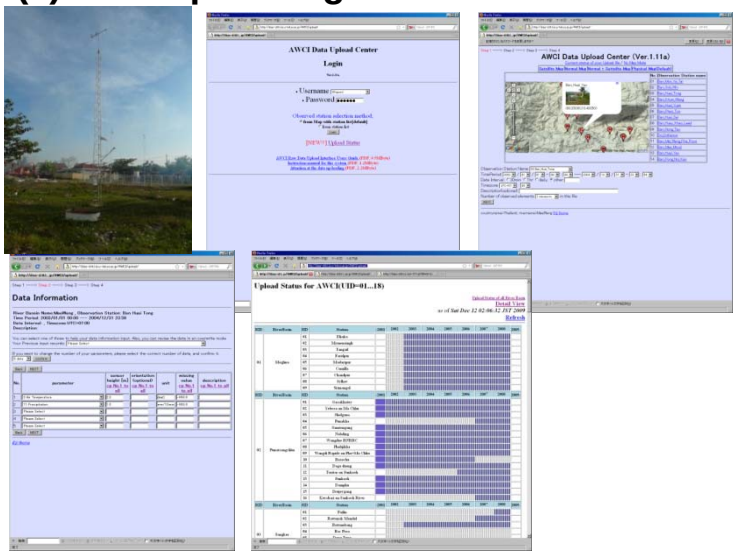


4 Components

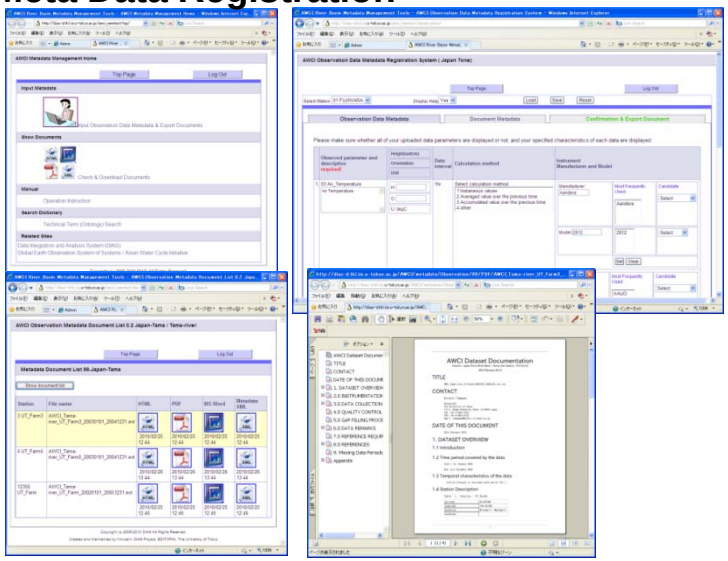
1. Data Upload
2. Quality Control
3. Meta-Data Input
4. Data Download

4 Components of In-situ data management

(1) Data Uploading

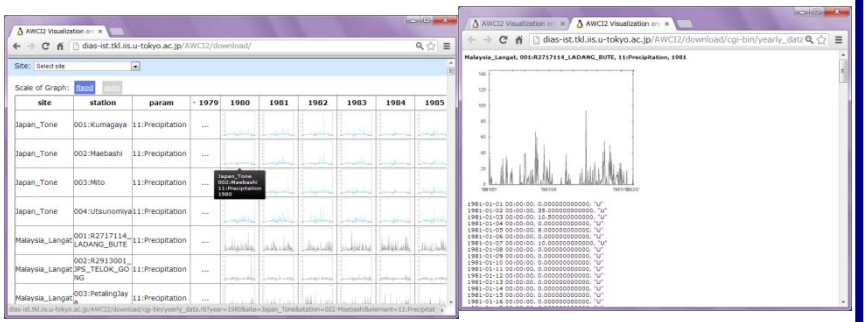


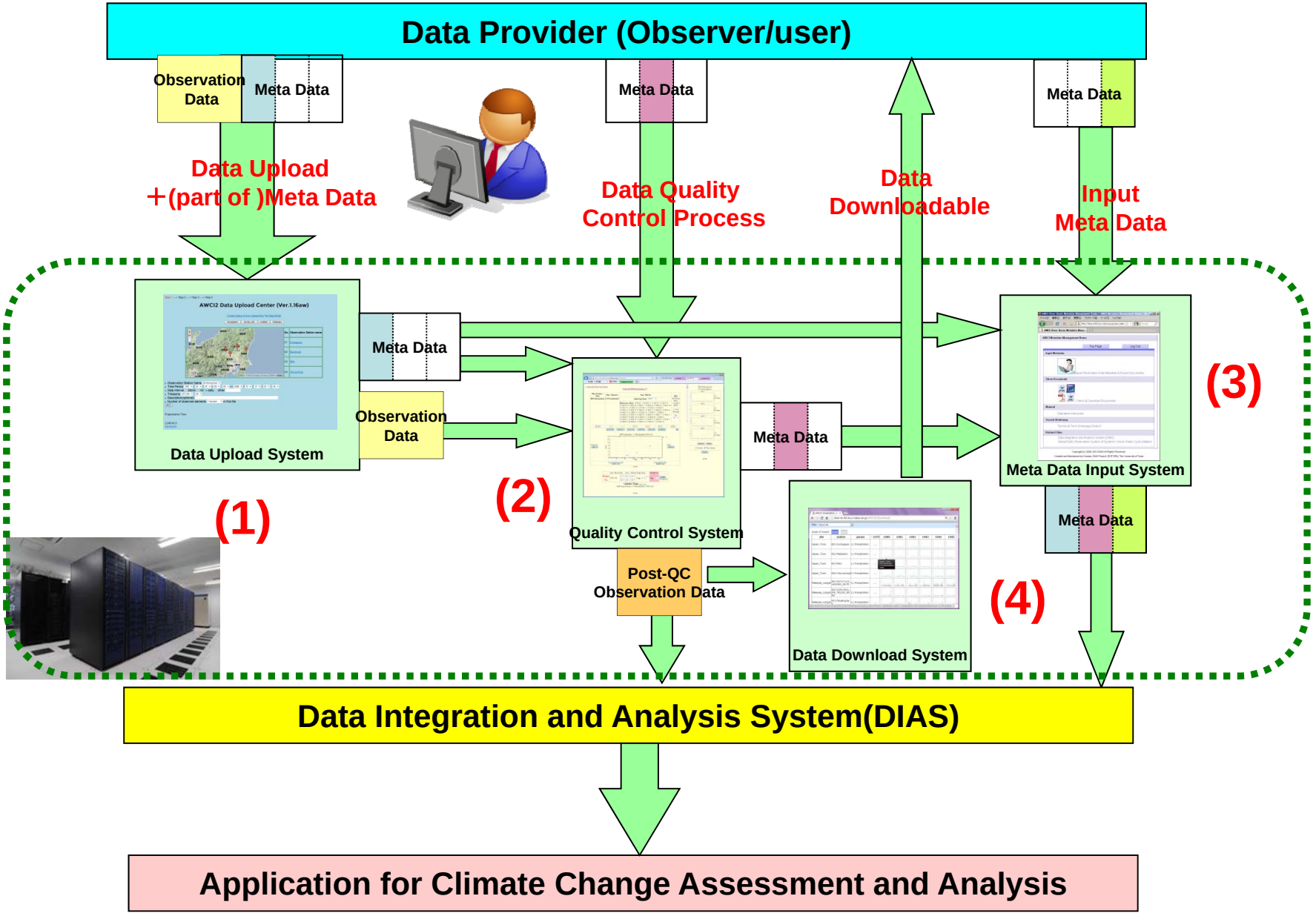
(3) Meta Data Registration



(2) Quality Controlling

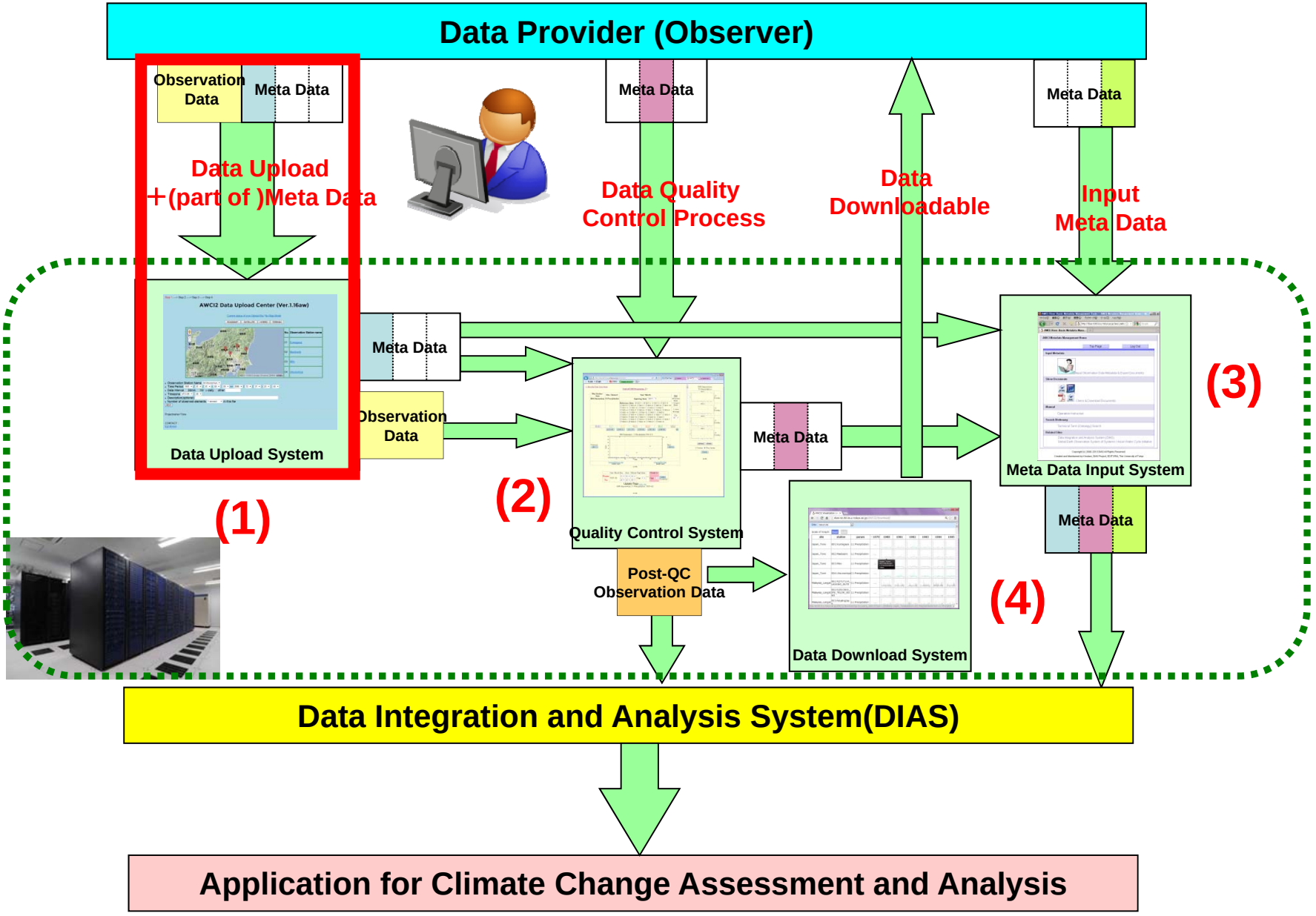
(4) Data Downloading



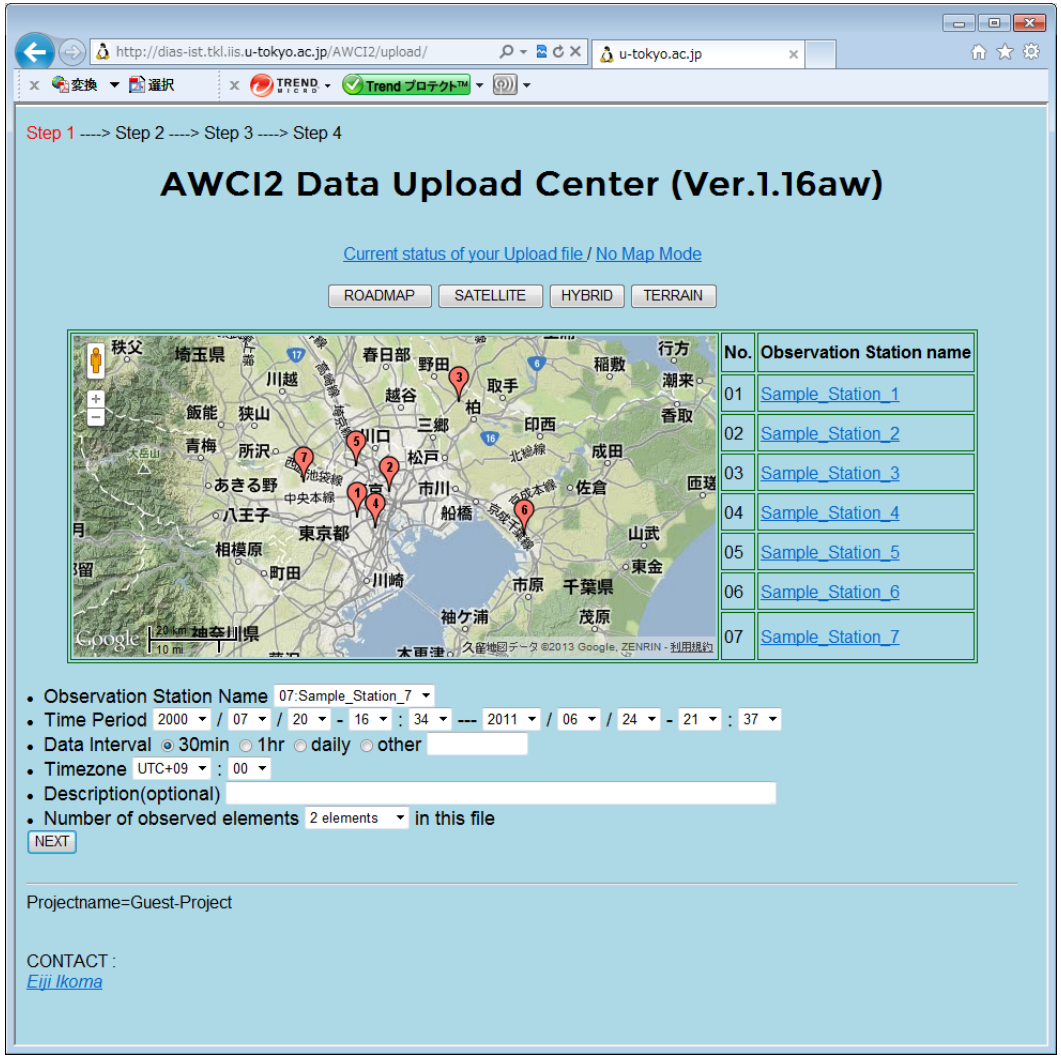

Data Provider (Observer/user)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Meta Data
Data Quality Control Process
Data Downloadable
Meta Data
Input Meta Data
Data Upload System
Meta Data
Observation Data
(1)
Quality Control System
(2)
Meta Data
Meta Data Input System
(3)
Meta Data
Post-QC Observation Data
Data Download System
(4)
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

4 Components

1. Data Upload
2. Quality Control
3. Meta-Data Input
4. Data Download

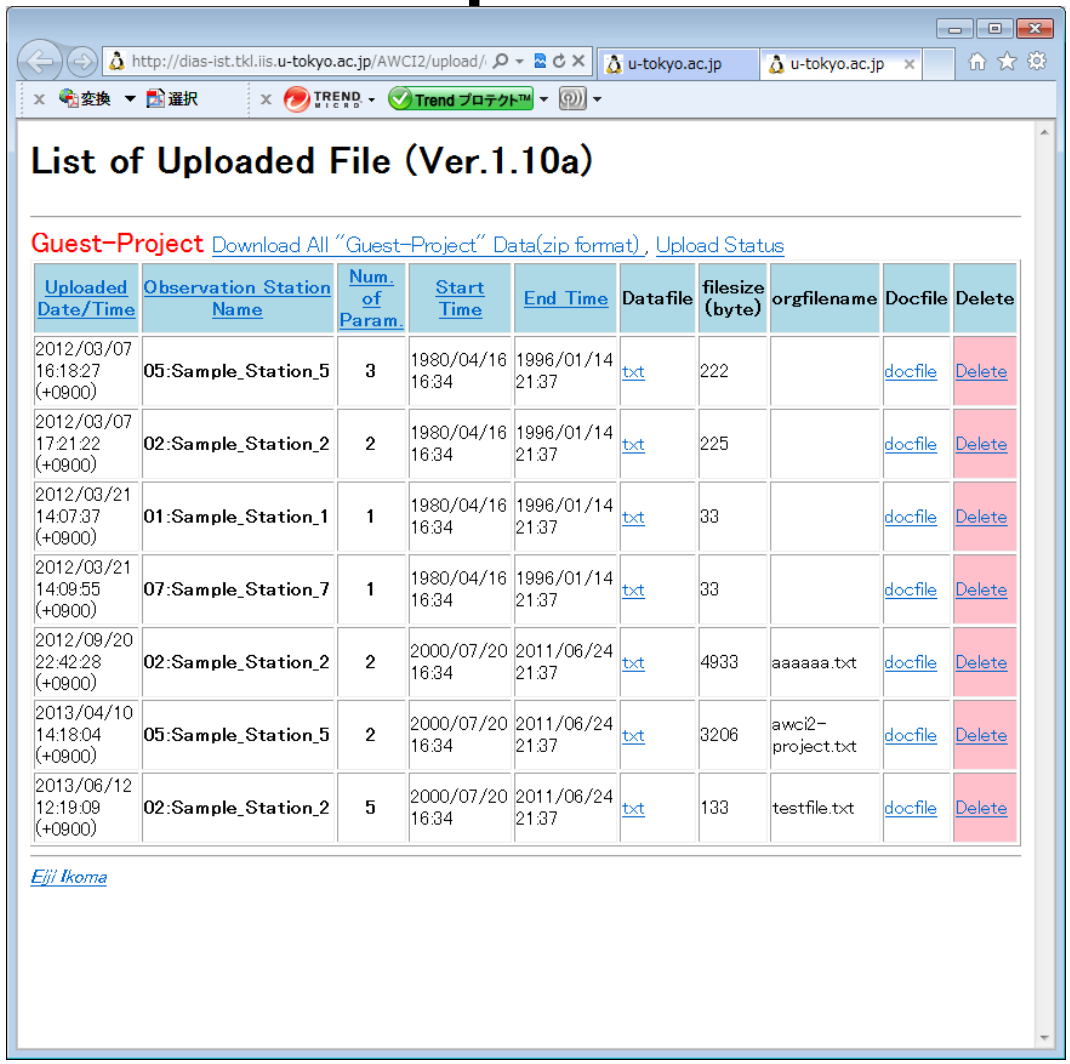

Data Provider (Observer)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Data Upload System
(1)
Meta Data
Observation Data
Meta Data
Data Quality Control Process
Quality Control System
(2)
Post-QC Observation Data
Meta Data
Data Downloadable
Data Download System
Meta Data
Input Meta Data
Meta Data Input System
Meta Data
(3)
(4)
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

Web-based data upload system



- Observation Point(Map/List)
- Time Period
- Data Interval
- Timezone
- Description (optional)
- Num. of observed elements

Upload Status Page



List of Uploaded File (Ver.1.10a)

Guest-Project Download All "Guest-Project" Data(zip format), Upload Status

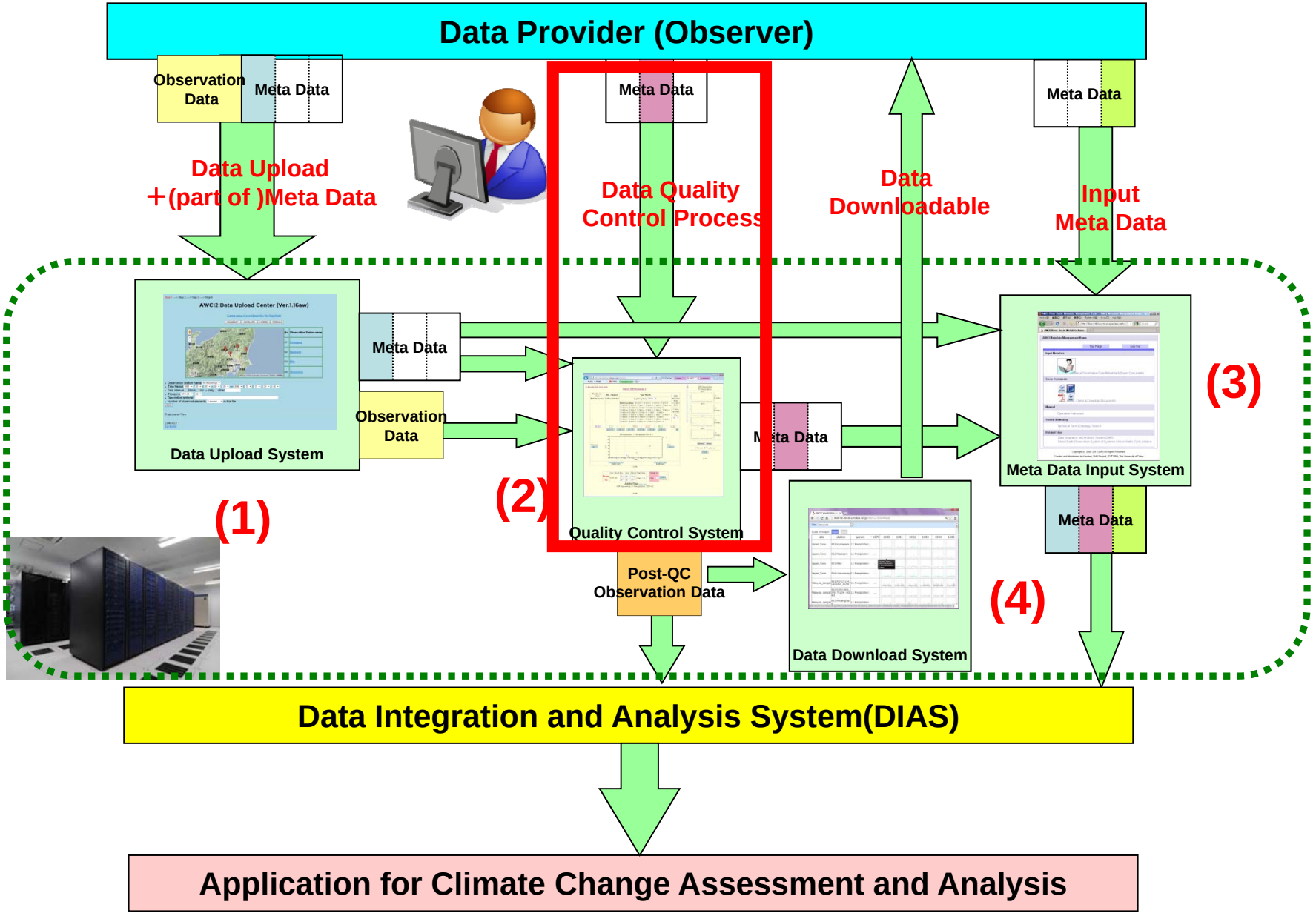
Uploaded Date/Time	Observation Station Name	Num. of Param.	Start Time	End Time	Datafile	filesize (byte)	orgfilename	Docfile	Delete
2012/03/07 16:18:27 (+0900)	05:Sample_Station_5	3	1980/04/16 16:34	1996/01/14 21:37	txt	222		docfile	Delete
2012/03/07 17:21:22 (+0900)	02:Sample_Station_2	2	1980/04/16 16:34	1996/01/14 21:37	txt	225		docfile	Delete
2012/03/21 14:07:37 (+0900)	01:Sample_Station_1	1	1980/04/16 16:34	1996/01/14 21:37	txt	33		docfile	Delete
2012/03/21 14:09:55 (+0900)	07:Sample_Station_7	1	1980/04/16 16:34	1996/01/14 21:37	txt	33		docfile	Delete
2012/09/20 22:42:28 (+0900)	02:Sample_Station_2	2	2000/07/20 16:34	2011/06/24 21:37	txt	4933	aaaaaa.txt	docfile	Delete
2013/04/10 14:18:04 (+0900)	05:Sample_Station_5	2	2000/07/20 16:34	2011/06/24 21:37	txt	3206	awci2-project.txt	docfile	Delete
2013/06/12 12:19:09 (+0900)	02:Sample_Station_2	5	2000/07/20 16:34	2011/06/24 21:37	txt	133	testfile.txt	docfile	Delete

Eiji Ikoma

- Download each/all data
- Check meta-data
- Delete uploaded data

Components

1. Data Upload
2. Quality Control
3. Meta-Data Input
4. Data Download


Data Provider (Observer)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Meta Data
Data Quality Control Process
Data Downloadable
Meta Data
Input Meta Data
Meta Data
Observation Data
Data Upload System
Quality Control System
Meta Data Input System
Meta Data
(1)
(2)
(3)
(4)
Post-QC Observation Data
Data Download System
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

Our Data Quality Control System

- First version of our QC system was born in 2002.
- Ver.0.x(2002-2003) → Ver.1(2004-2005) →Ver.2(2005-2006) → Ver.3(2007-)
- Web based UI, Easy-to-use and light operation
- Post-QC Data Download, Progress management system is also avilable
- Ver 3.05a are now running for AWCI2(2012-)

To control data quality is..

- Check the data one by one
- Add a “flag”, which shows the quality level of data

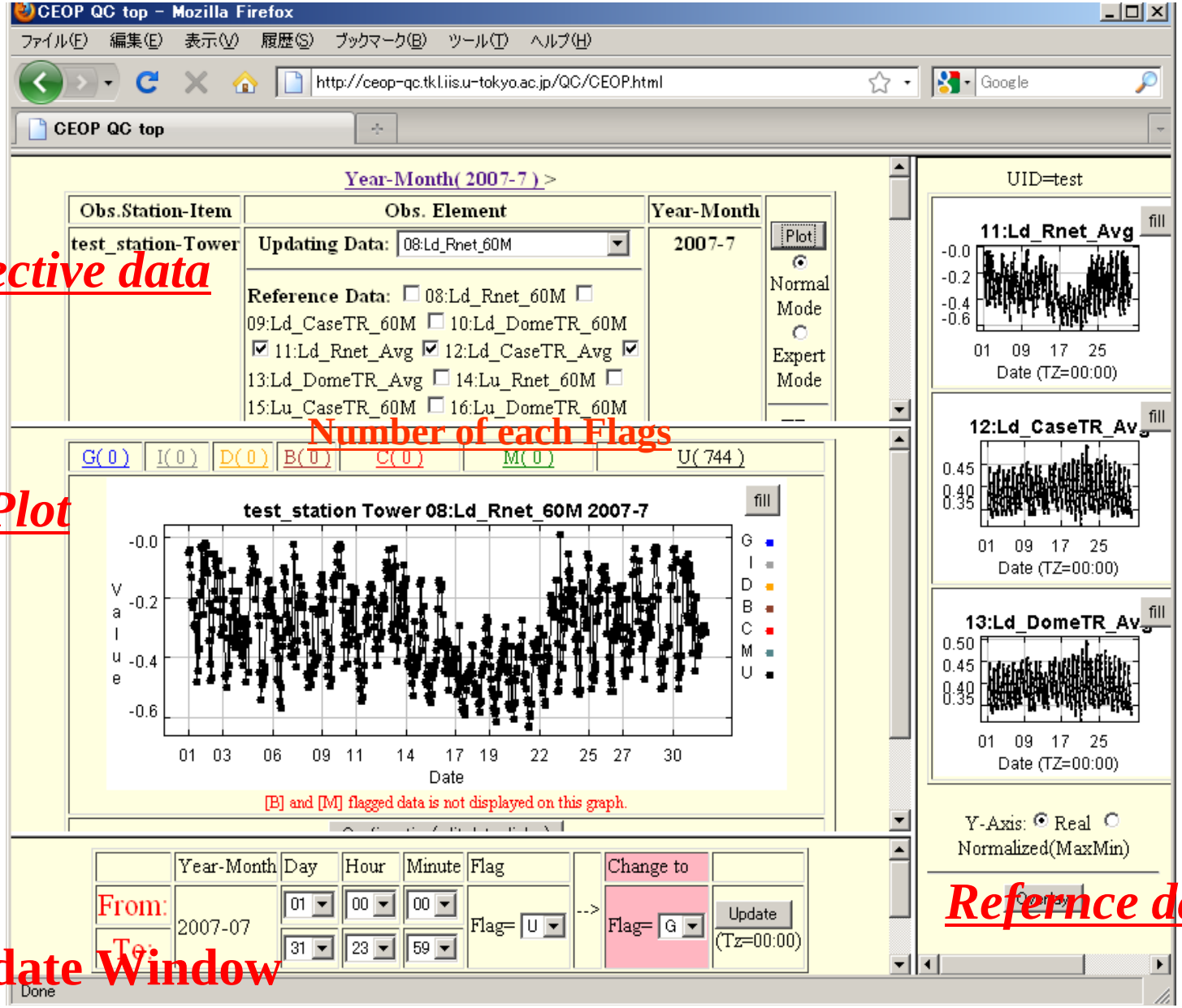
Quality control flag definitions

Flag Definitions

G: Good
I : Interpolated
D: Dubious/Questionable
B: Bad
C: minus precipitation or Abnormal value
M: Missing
U: Unchecked

http://www.eol.ucar.edu/projects/ceop/dm/documents/refdata_report/data_flag_definitions.html

Data Quality Checking System



QC Objective data

Number of each Flags

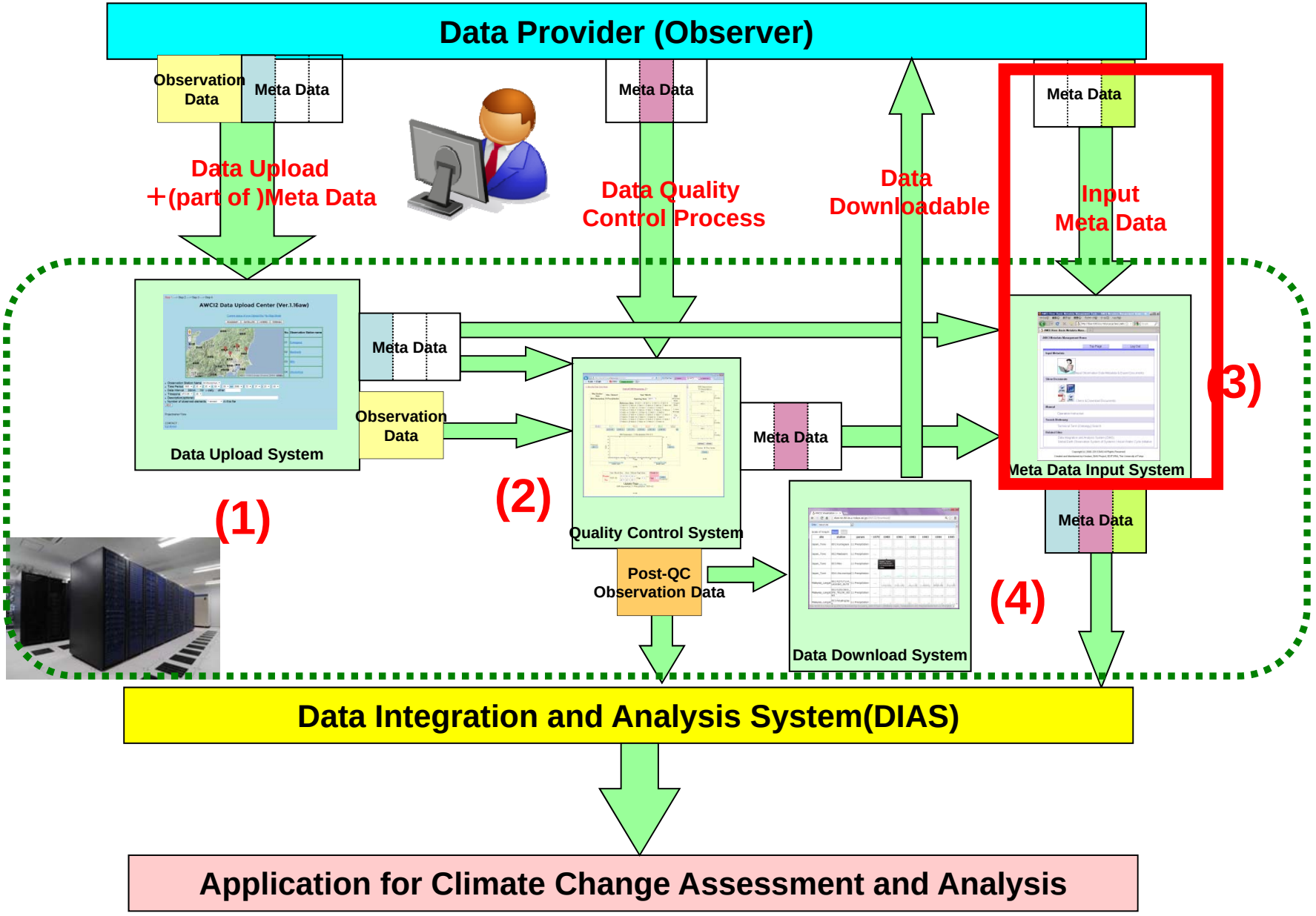
Data Plot

Refernce data

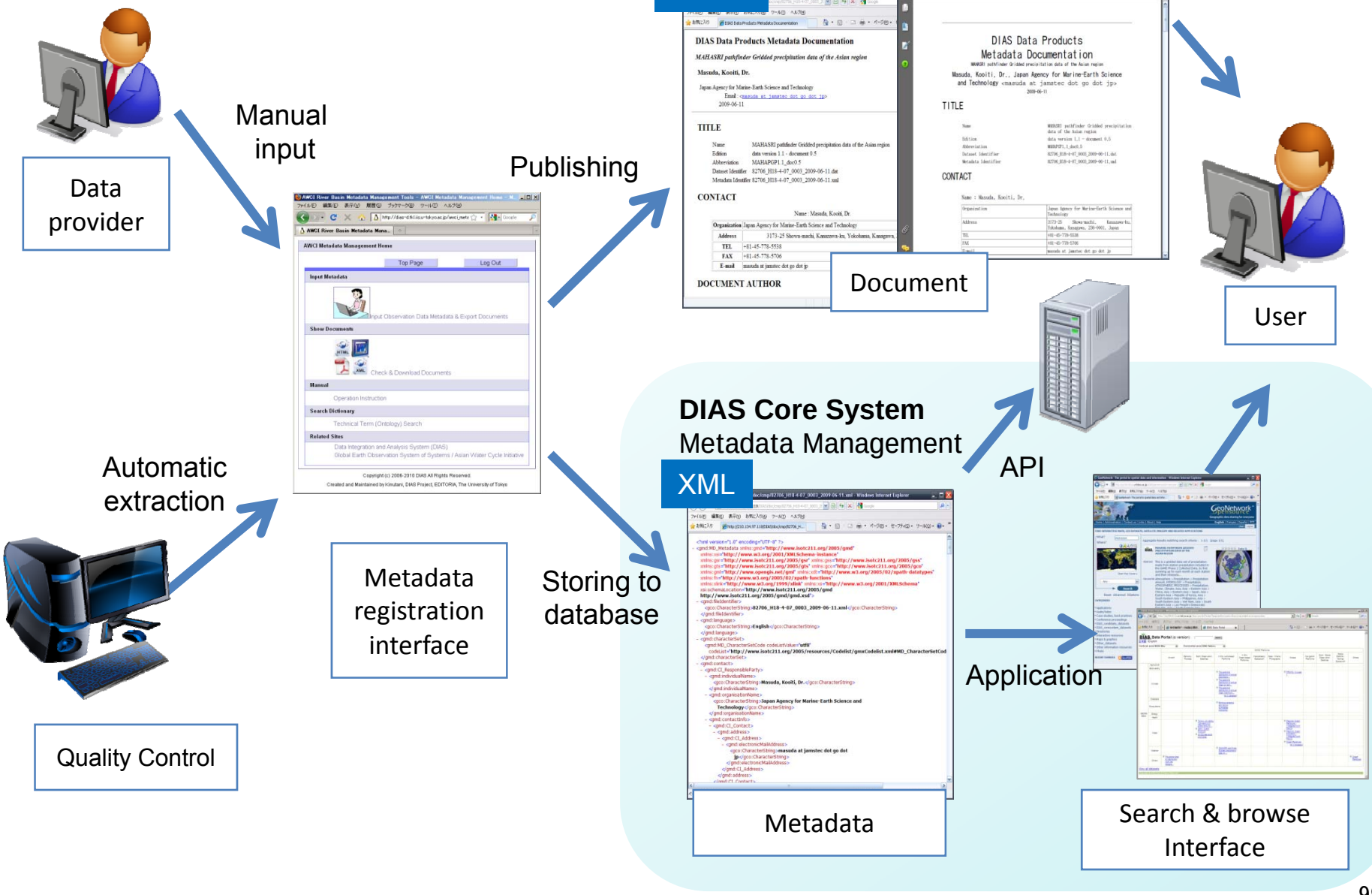
Flag Update Window

Components

1. Data Upload
2. Quality Control
3. Meta-Data Input
4. Data Download

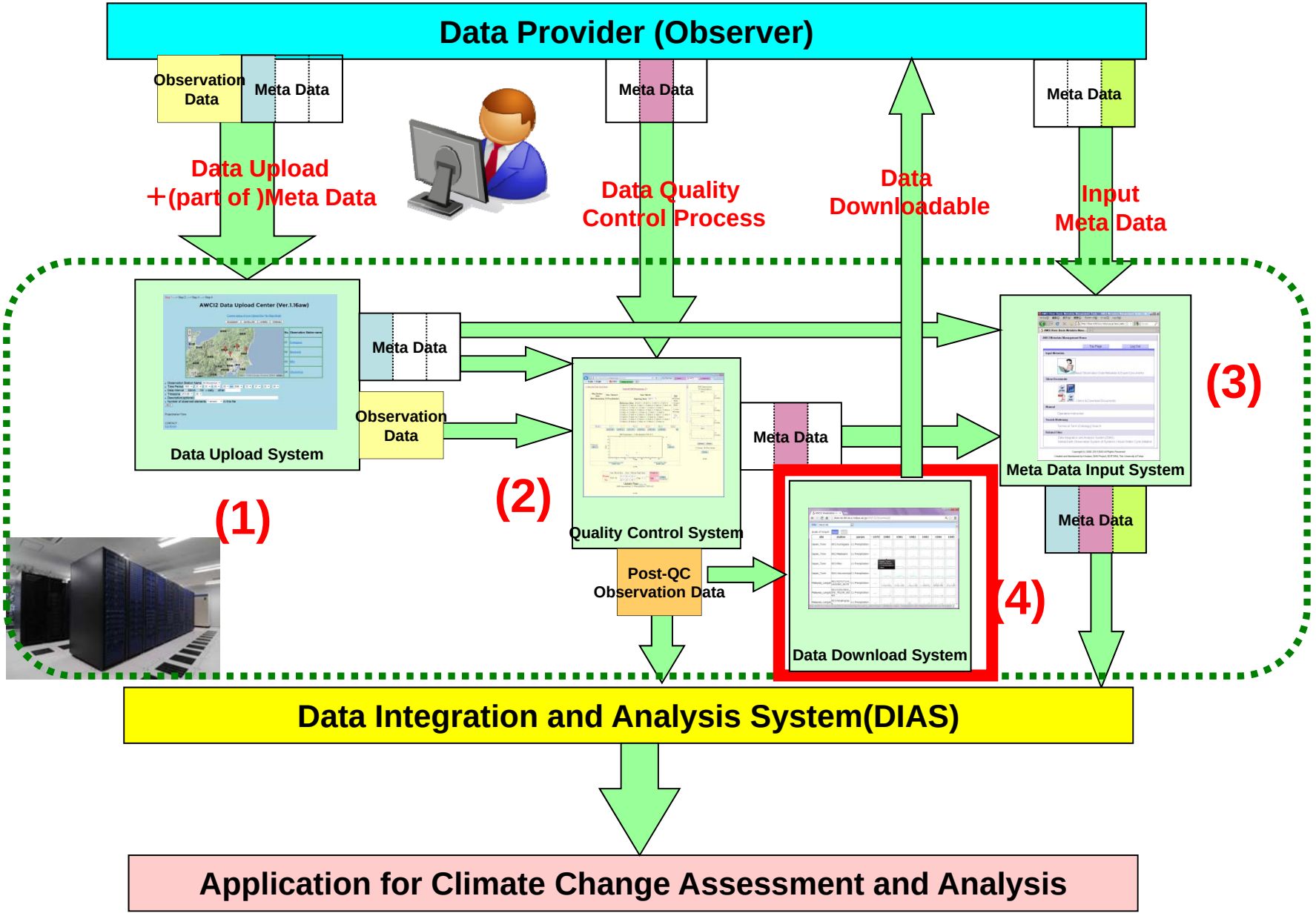

Data Provider (Observer)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Meta Data
Data Quality
Control Process
Data
Downloadable
Meta Data
Input
Meta Data
Data Upload System
Meta Data
Observation Data
(1)
(2)
Quality Control System
Meta Data
Meta Data Input System
(3)
Meta Data
Post-QC
Observation Data
Data Download System
(4)
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

Role of Metadata & Dataset Documents

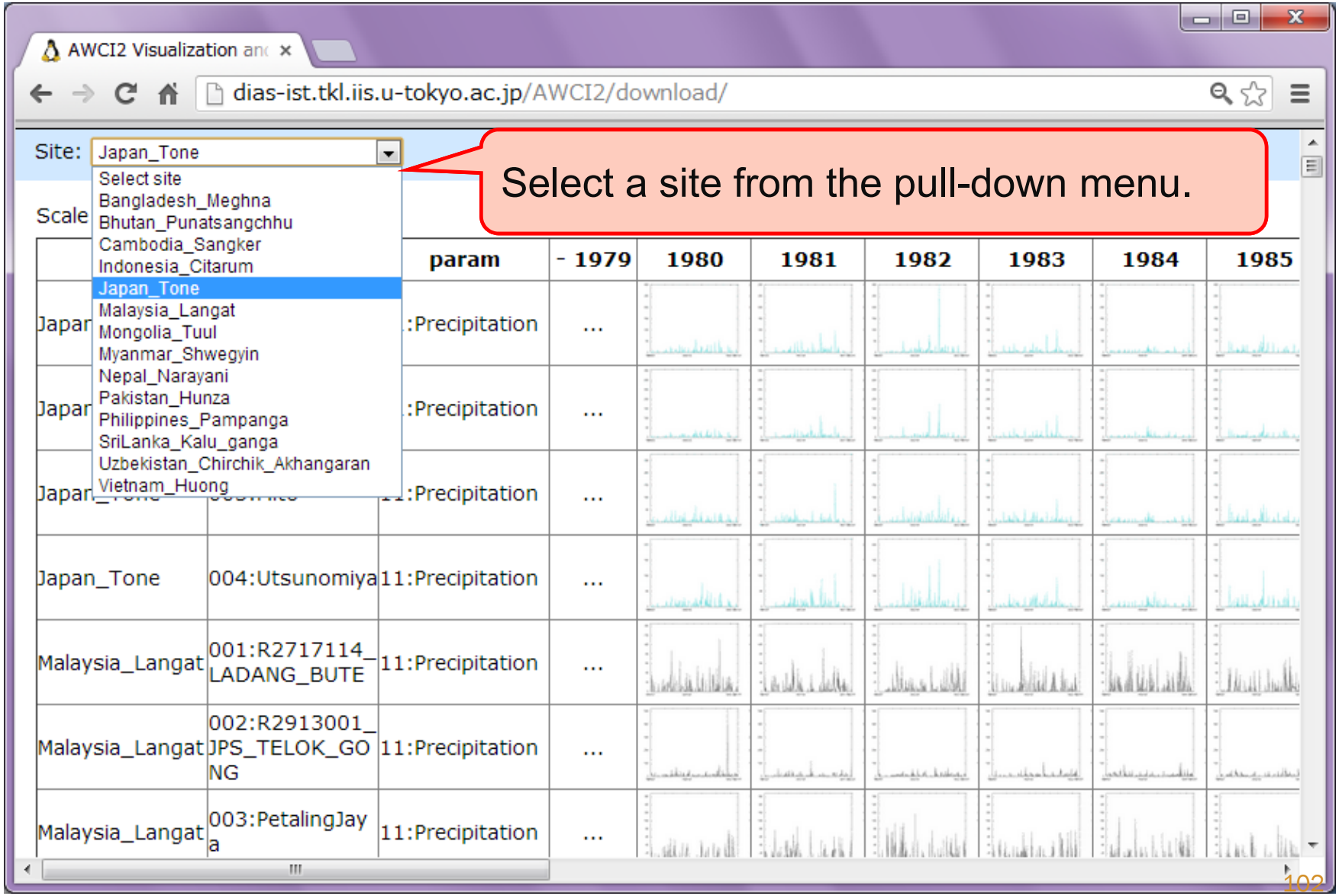


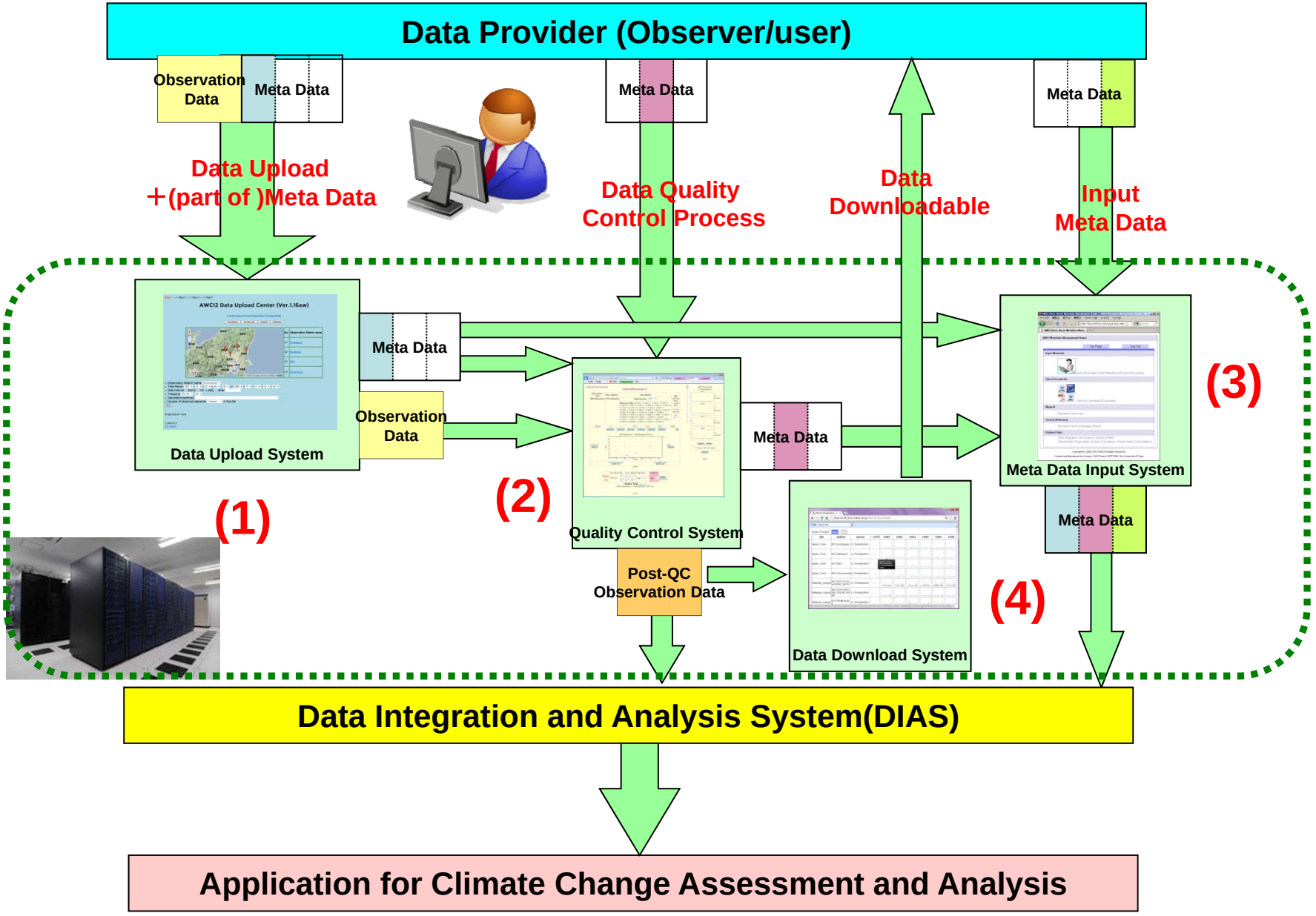
Components

1. Data Upload
2. Quality Control
3. Meta-Data Input
4. Data Download


Data Provider (Observer)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Meta Data
Data Quality Control Process
Data Downloadable
Meta Data
Input Meta Data
Data Upload System
Meta Data
Observation Data
(1)
Quality Control System
(2)
Meta Data
Post-QC Observation Data
Data Download System
(4)
Meta Data Input System
(3)
Meta Data
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

Step. 2 Data Visualization

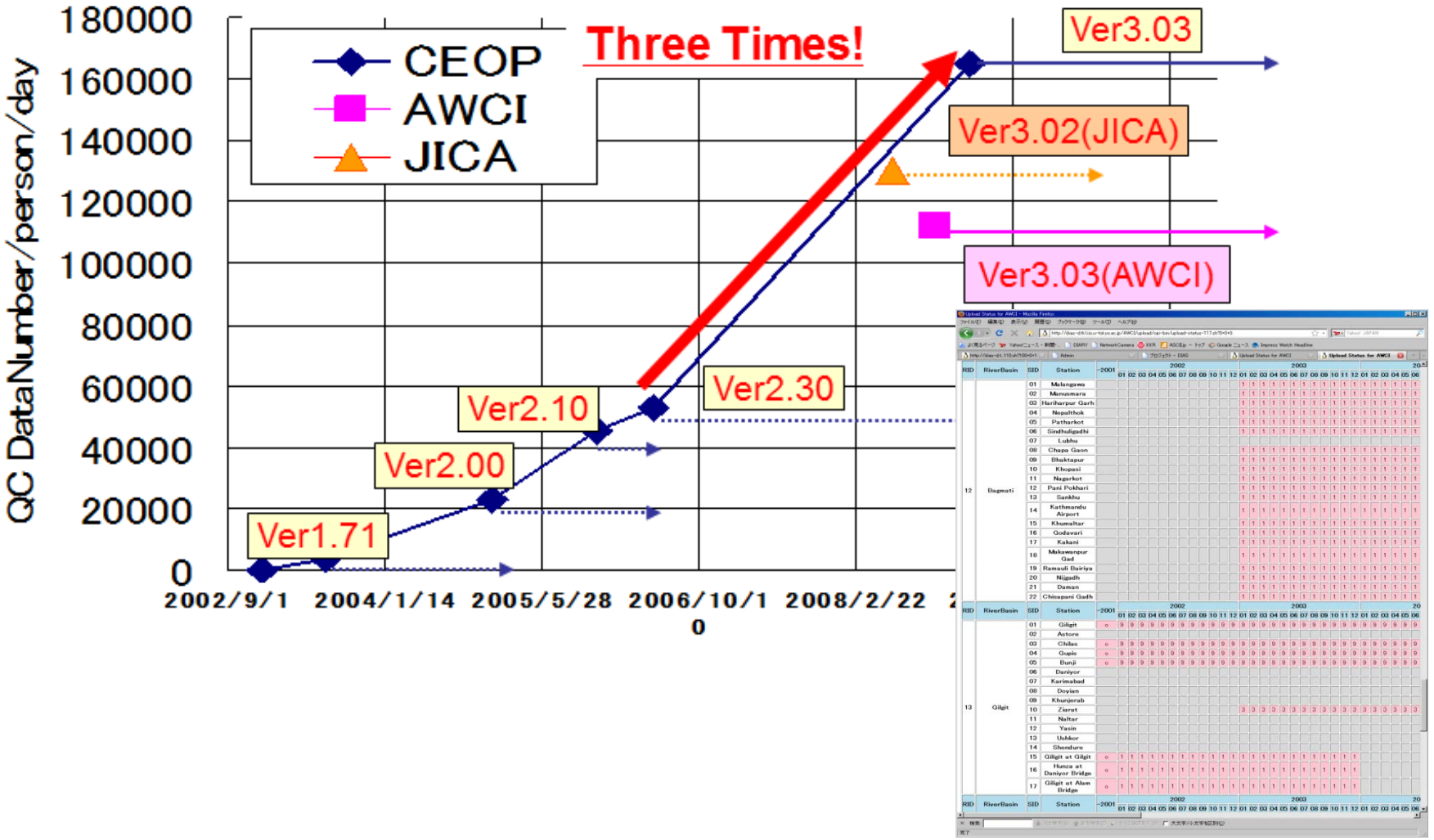



Data Provider (Observer/user)
Observation Data
Meta Data
Data Upload
＋(part of)Meta Data
Meta Data
Data Quality Control Process
Data Downloadable
Meta Data
Input Meta Data
Data Upload System
Meta Data
Observation Data
(1)
Quality Control System
(2)
Meta Data
Meta Data Input System
(3)
Meta Data
Post-QC Observation Data
Data Download System
(4)
Data Integration and Analysis System(DIAS)
Application for Climate Change Assessment and Analysis

Data Integration and Analysis System

a legacy for Japan's contributions to GEOSS

accelerating data archiving, including data loading, QC and metadata registration

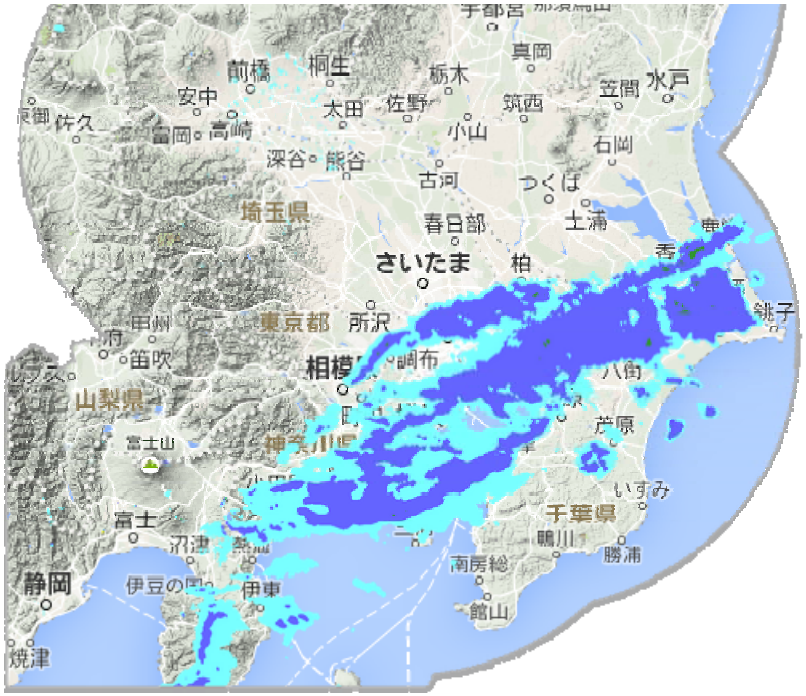


High-Quality data
with complete metadata
is one of **“DIAS Value”**

Data Integration and Analysis System

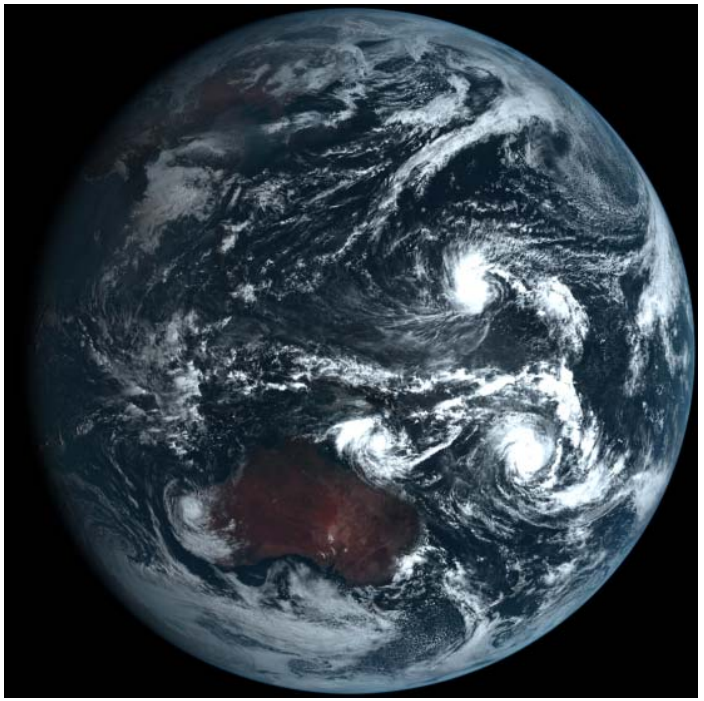
a legacy for Japan's contributions to GEOSS

archiving, analyzing and disseminating data and information with high velocity.



X-band MP Radar
- 250 m grid
- Every 1 min.

500GB/day



New Gestational Satellite
- 0.5 km grid
- Every 2.5 min.

500GB/day

Realtime archiving data on DIAS

MLIT


Live Camera


River Telemetry


C-band Radar


X-band Radar

/10 min.

/1 min.
250m mesh
14area

/5 min.
1km mesh
All area


Live Camera Images

/10min.

Local government etc.

NOAA,GMS,MTSAT,
MODIS,AMSR-E,
GMS8


Satellite Data


DIAS Core System

/1 hour,
1300 points


AMeDAS


GPV

JMA


GSMaP

JAXA

/1 hour
0.1deg. mesh
Global(60S-60N)

/5 min.
91 points

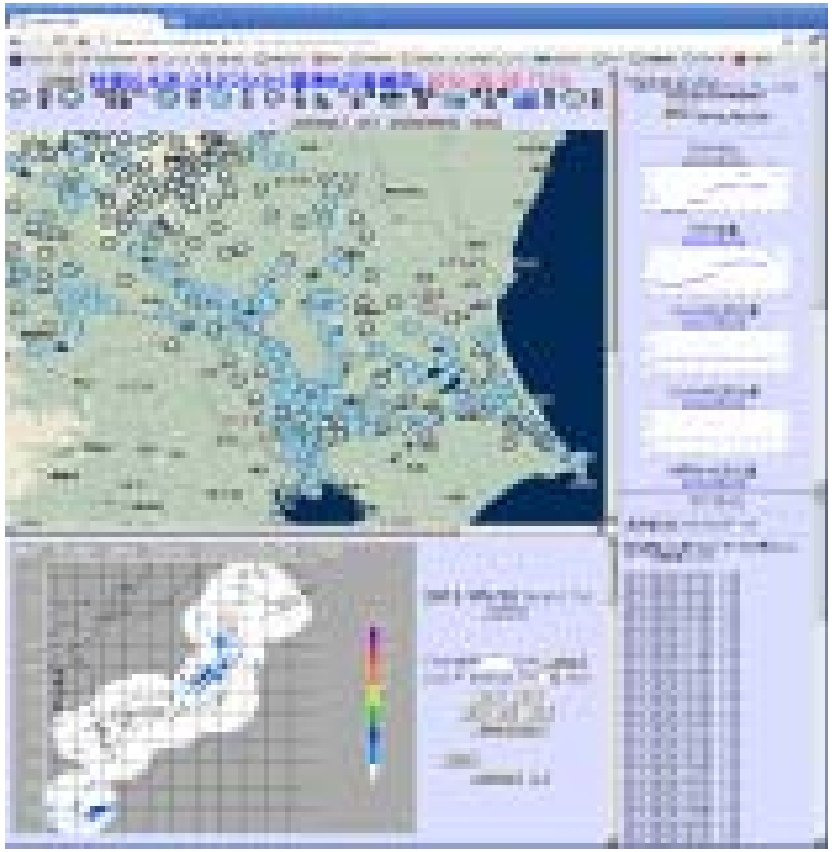
/1-3 hour
84-264hr. forecast
0.2～1deg. mesh
Global, Japan area

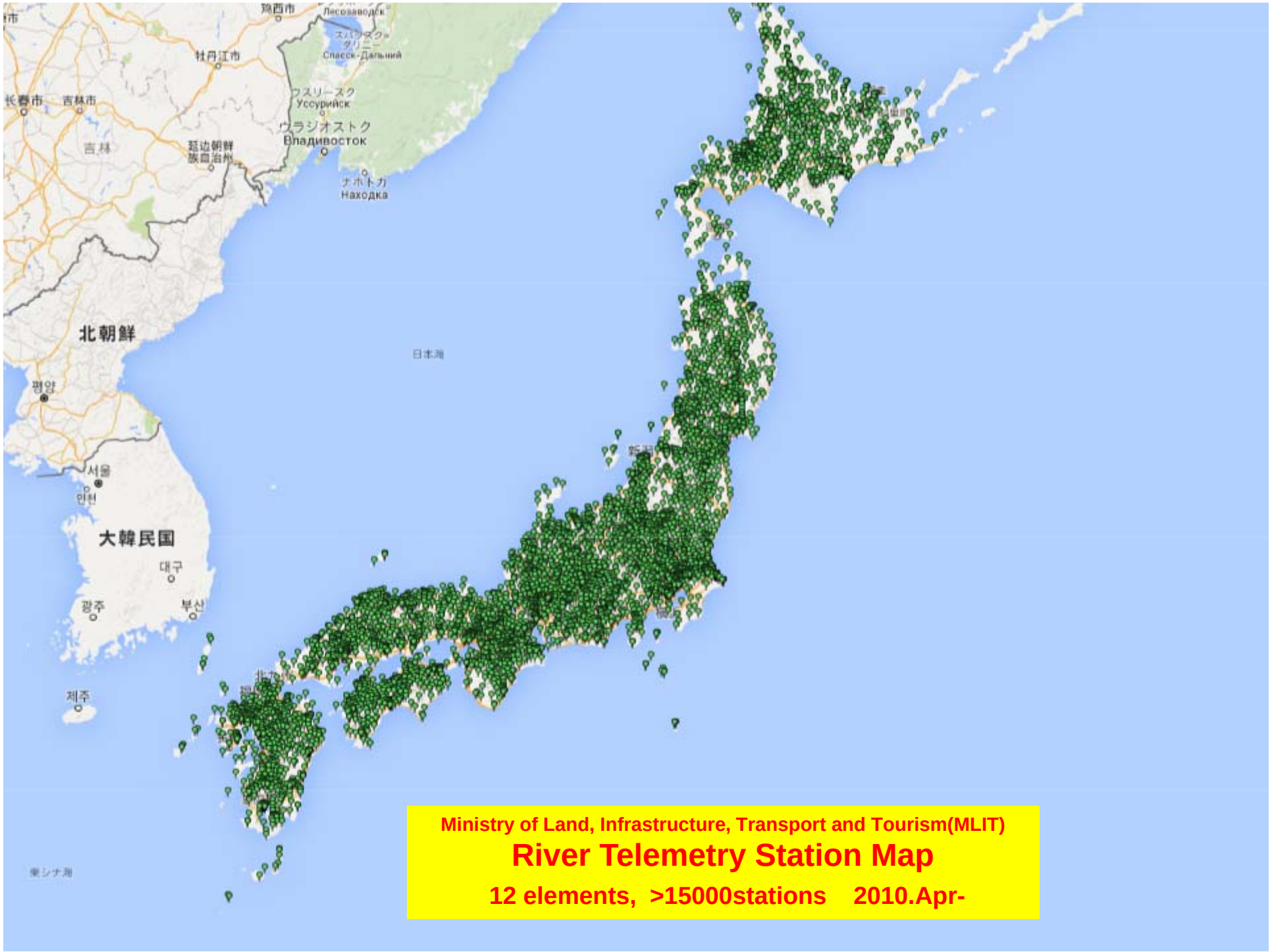

Tidal level

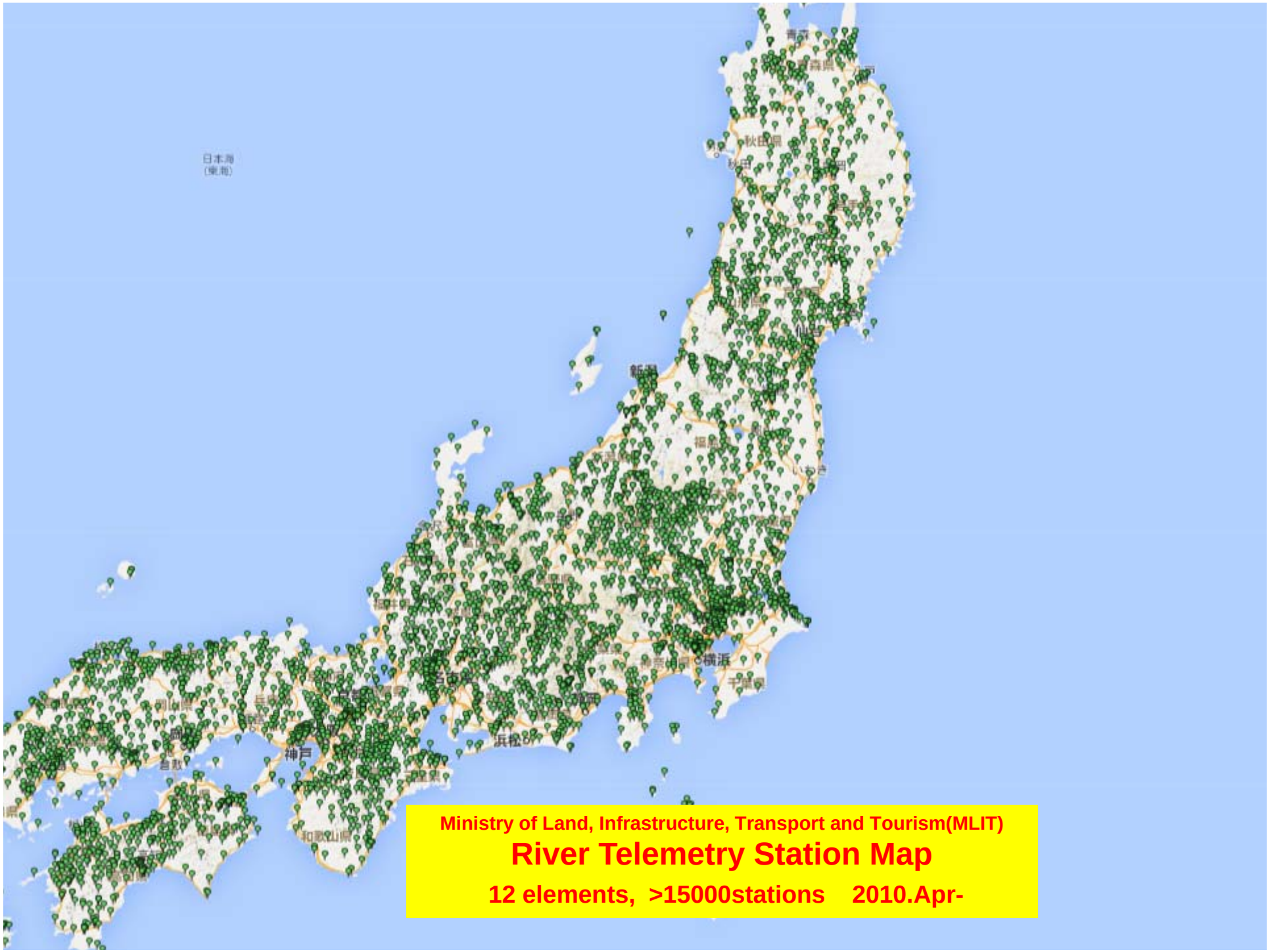
Japan Coast Guard

Telemetry data – river information

- 10 division of MLIT, 480 river system, 15000 observation points /10min. Telemetry data
- 12 categoly 300 data (-2014.3)　3 categoly 30 data(2014.4-)
- Main System: Komaba-campus
- Sub System: Chiba-NII = redundant system
- Realtime distribution to application (Web-DHM on Tonegawa)
- Under developping for Web-DHM on Tsurumigawa
- Archved and service from 2010.4-



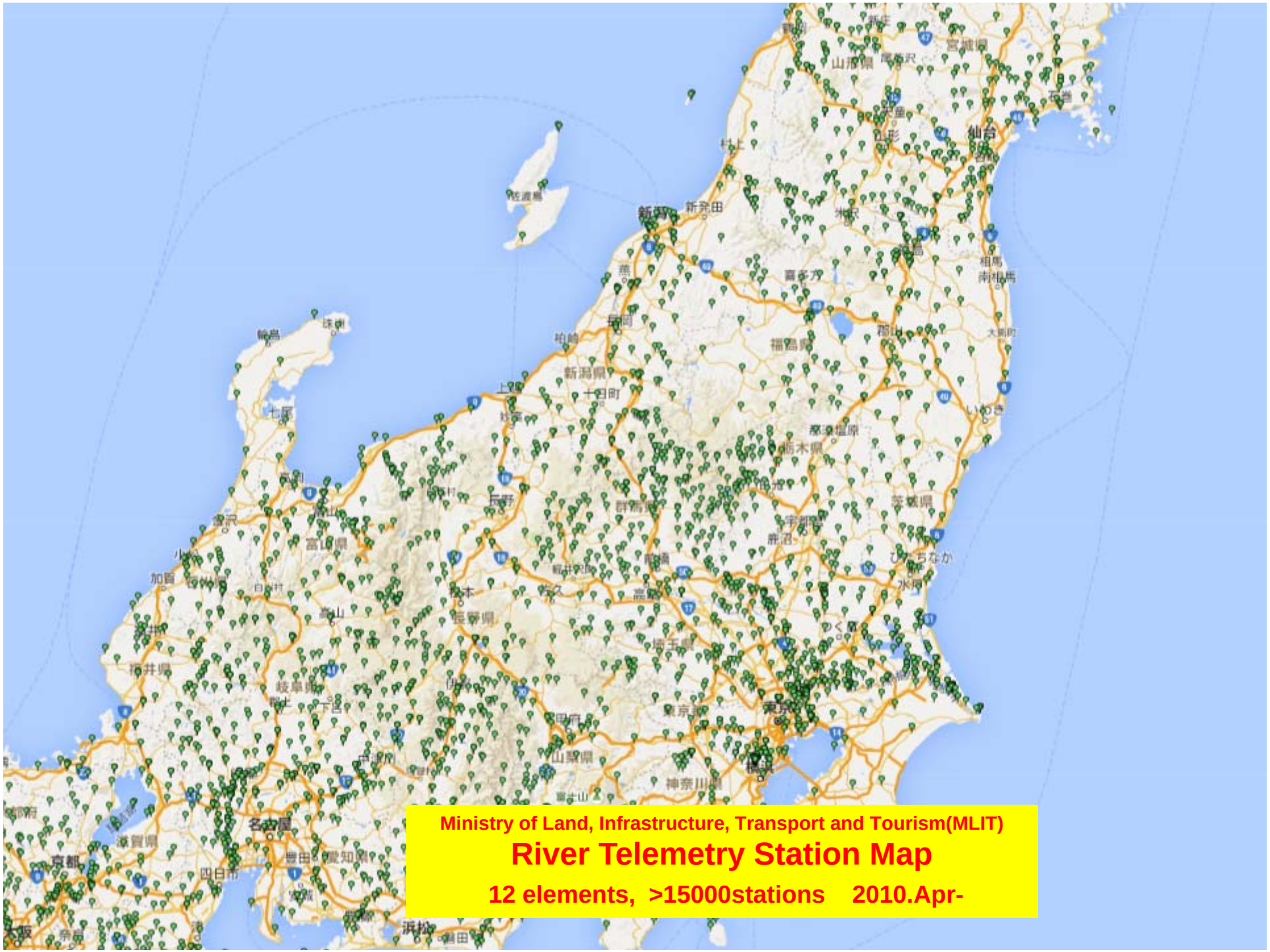

北朝鮮
大韓民国
日本海
Ministry of Land, Infrastructure, Transport and Tourism(MLIT)
River Telemetry Station Map
12 elements, >15000stations 2010.Apr-

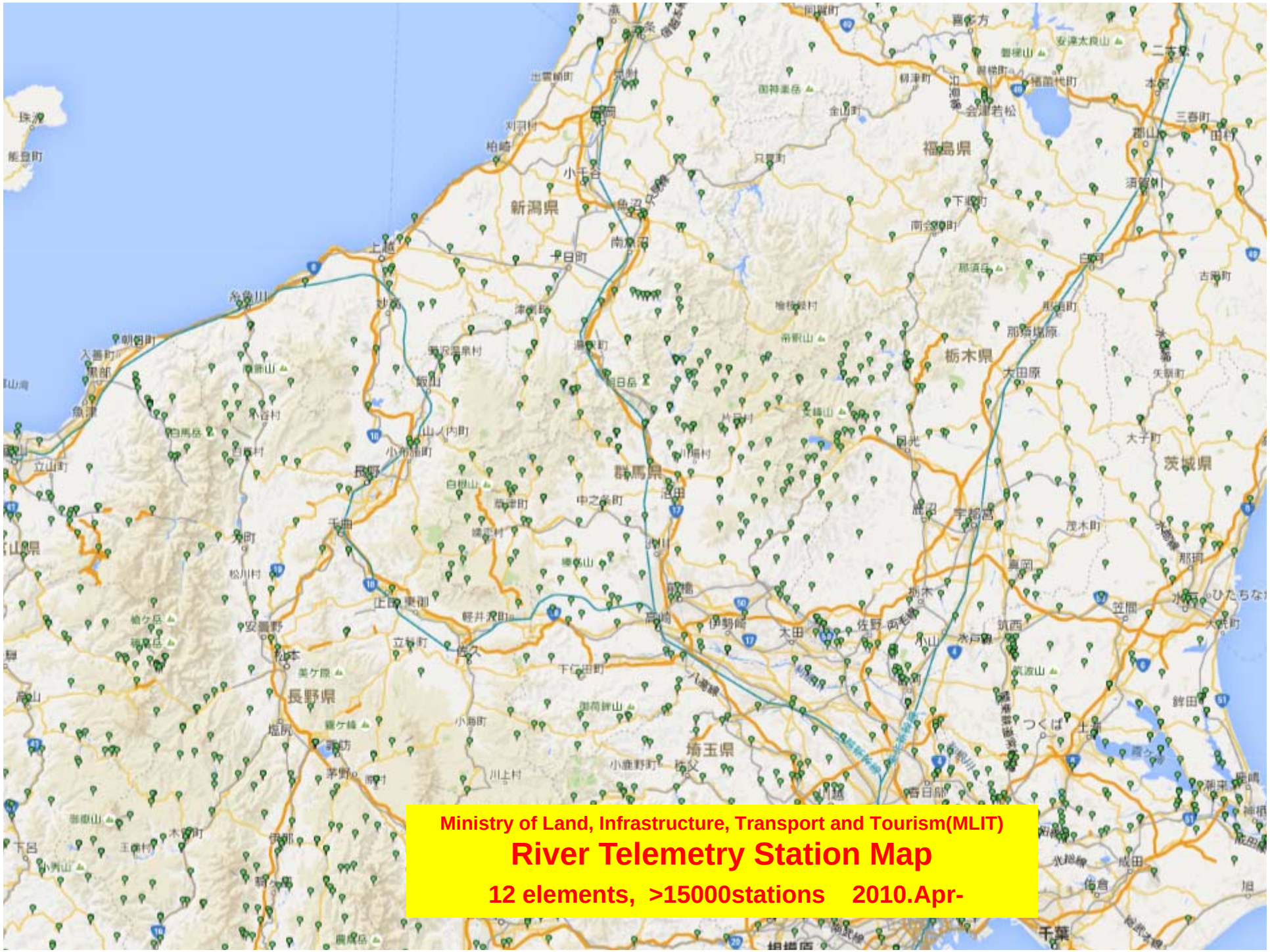


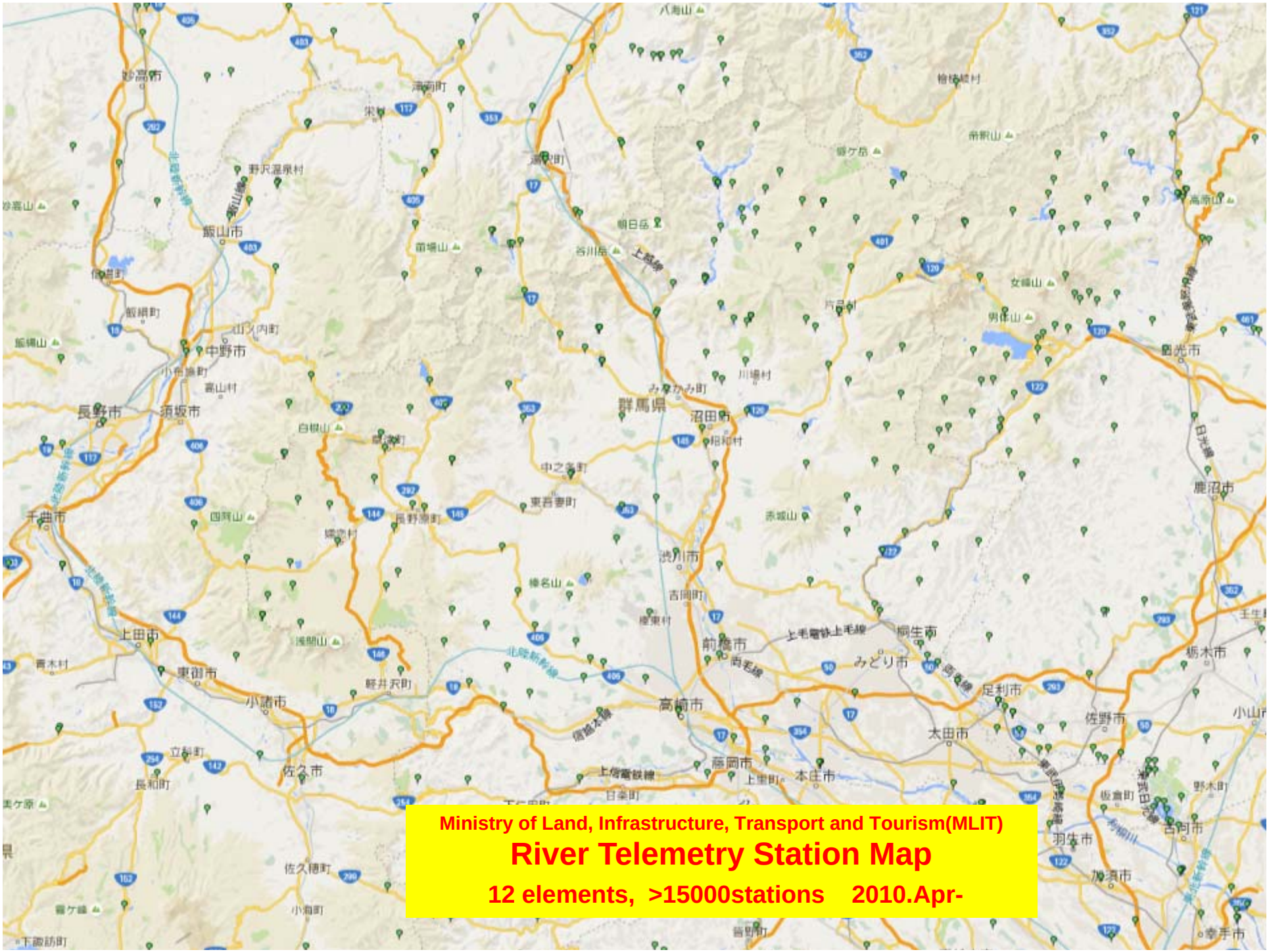
Ministry of Land, Infrastructure, Transport and Tourism(MLIT)

River Telemetry Station Map

12 elements, >15000stations 2010.Apr-

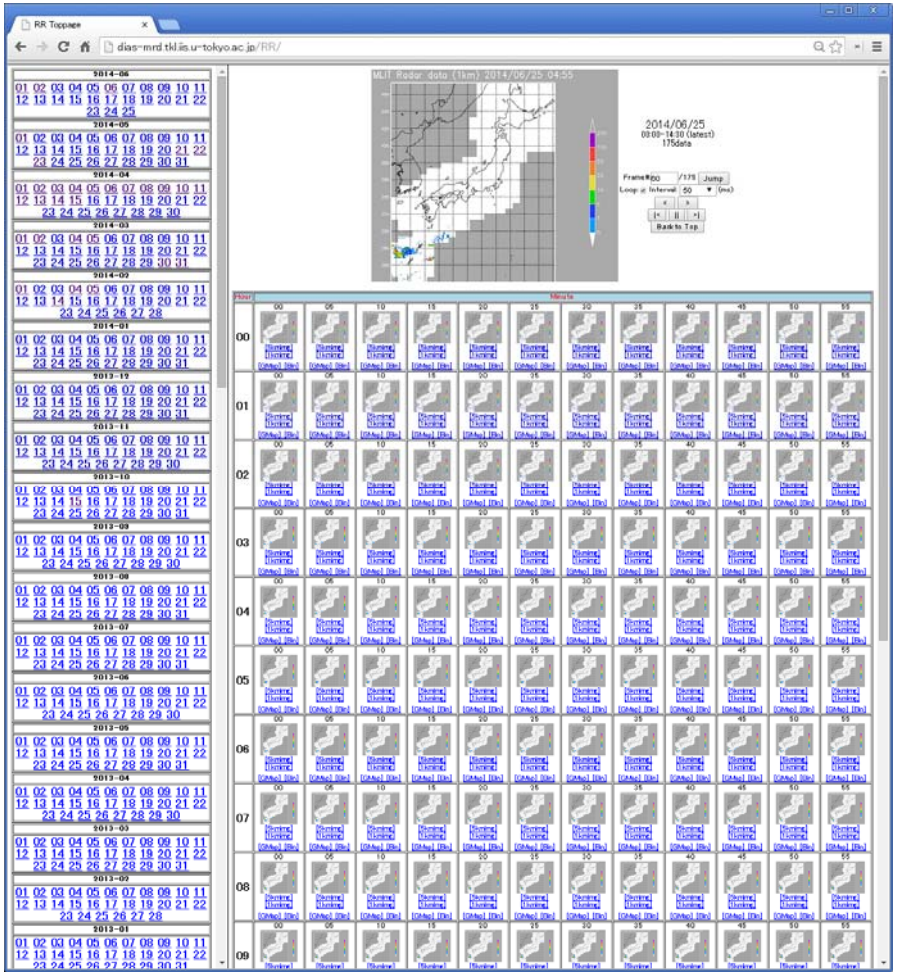

Ministry of Land, Infrastructure, Transport and Tourism(MLIT)
River Telemetry Station Map
12 elements, >15000stations 2010.Apr-


Ministry of Land, Infrastructure, Transport and Tourism(MLIT)
River Telemetry Station Map
12 elements, >15000stations 2010.Apr-


Ministry of Land, Infrastructure, Transport and Tourism(MLIT)
River Telemetry Station Map
12 elements, >15000stations 2010.Apr-

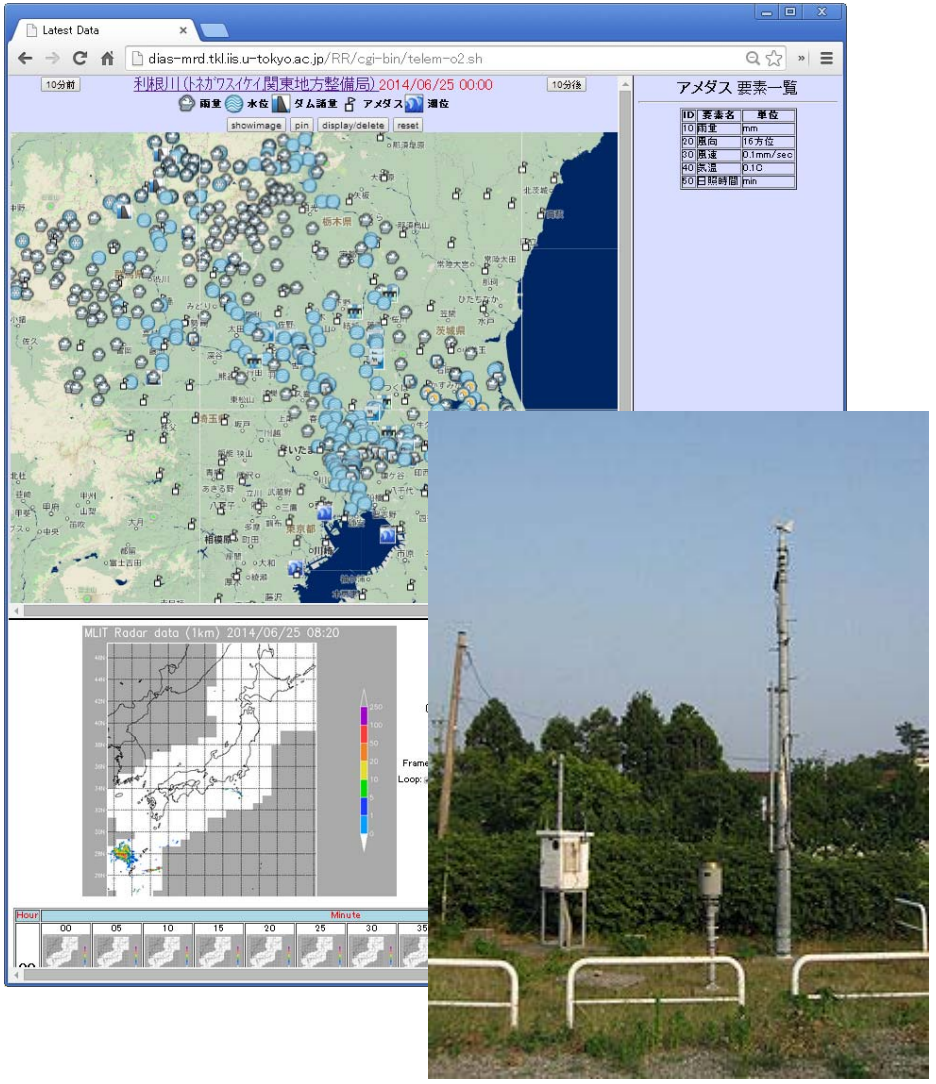
C-Band Radar

- 1km mesh, /5min. rainfall radar data（/10min. -2014.3)
- Raw data, 5km quicklook image, 1km image, 1km map-overlay system
- Komaba x 2, Chiba-NII x 1 = 3 redundant system
- Realtime distribution to application (Web-DHM on Tonegawa)
- Under developping for Web-DHM on Tsurumigawa
- Archive and service from 2010.4



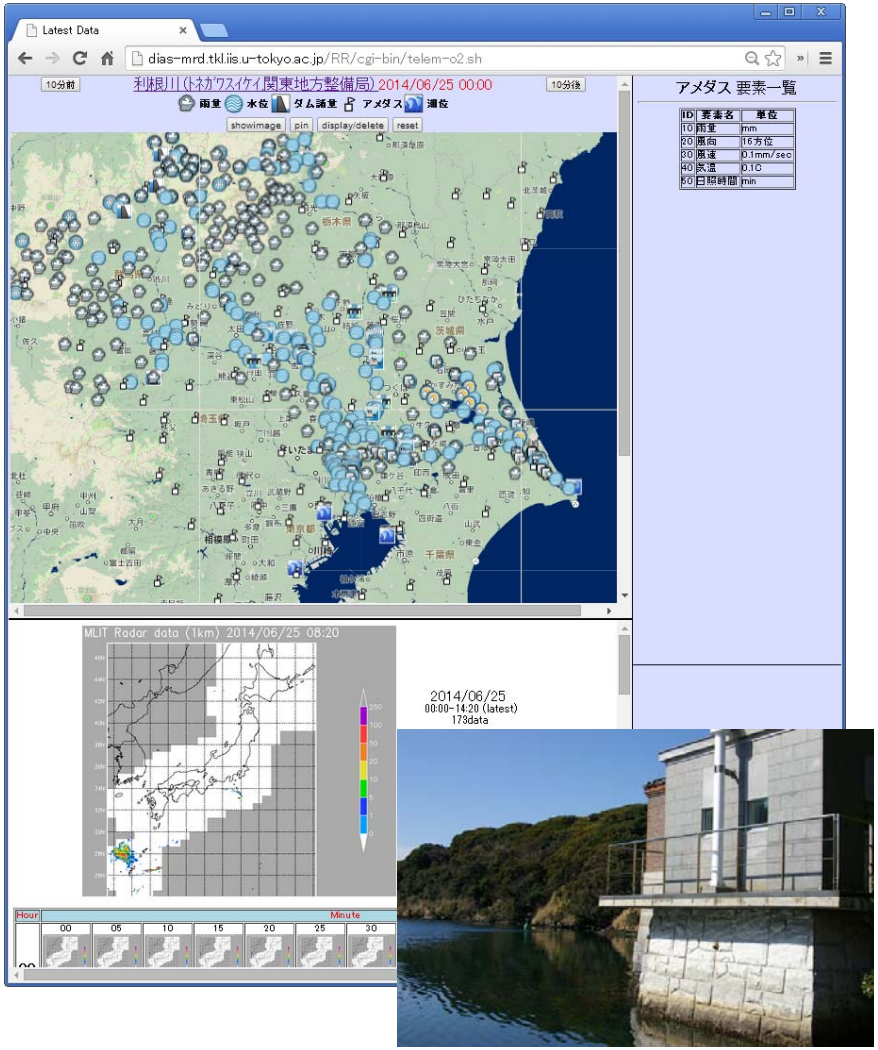
AMeDAS

- JMA's 1300 observation point, /1hr update.
- Precipitation, Wind direction, Wind speed, Temperature, sunshine duration
- Visualize with river telemetry data on DIAS
- Komaba x 1, Chiba-NII x 1 = 2 redundant system
- Realtime distribution to application (Web-DHM on Tonegawa)
- Under developping for Web-DHM on Tsurumigawa
- Servicr from 2012.04-
- Archived from 1999.08



Tidal Level data

- 91 observation point maintained by Japan Coastal Guard
- /5min. Tidale level and air pressure
- Visualize with river telemetry data on DIAS
- Archived and service from 2014.06
- Planning to distribute to WebDHM and other model

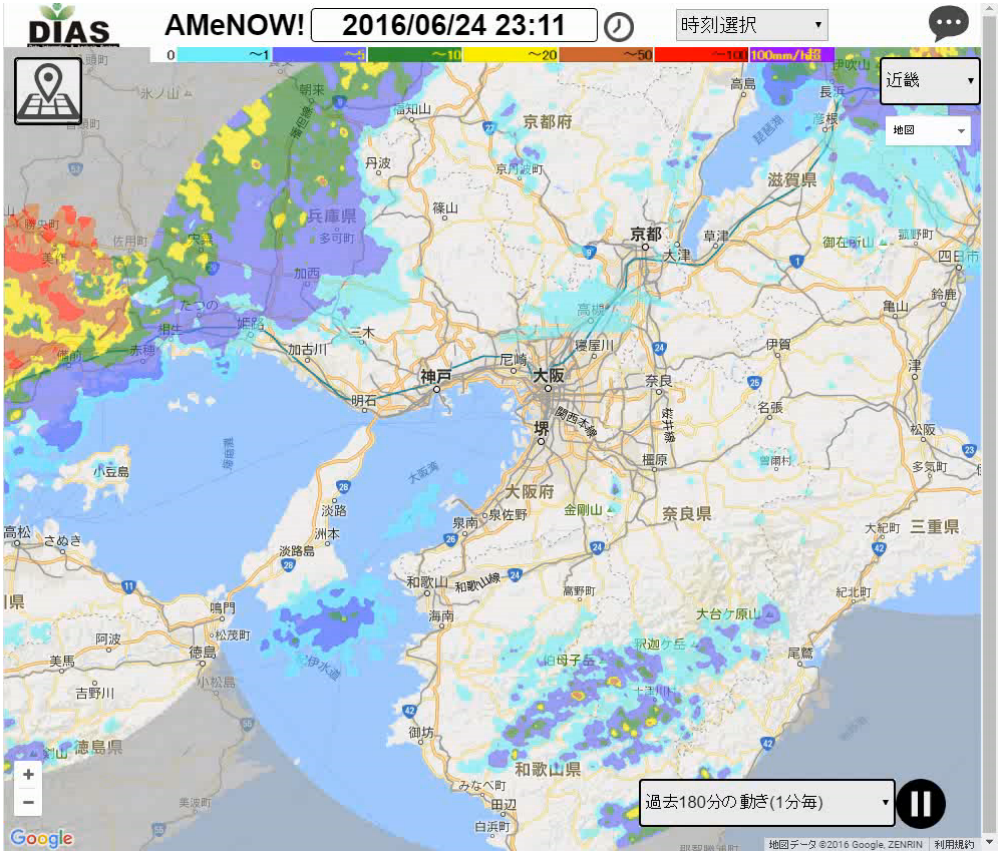


Live camera images

- Local office of MLIT, Local government etc. install and maintain many cameras
- /5～10min realtime camera image
- Archived data close to water-level observation point

AMeNOW!
Realtime rainfall information on Smartphone

- Since 2015/10



http://rain.diasjp.net/

Real-time Huge scale data
is one of “DIAS Value”

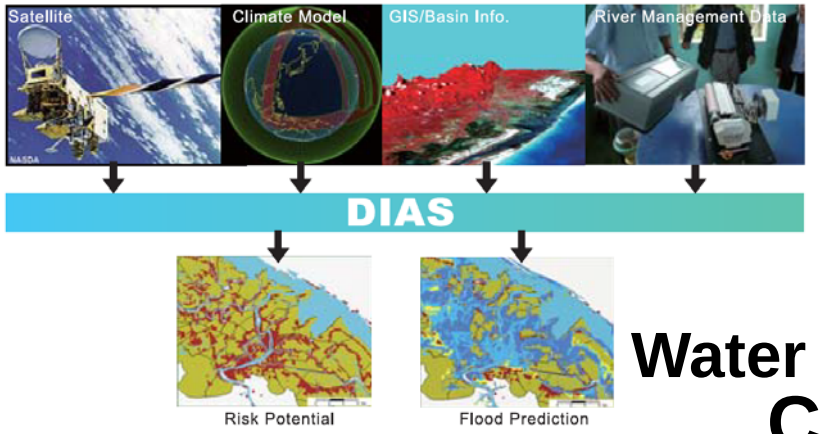
Challenges to 4V

- volume
- variety
- veracity
- velocity

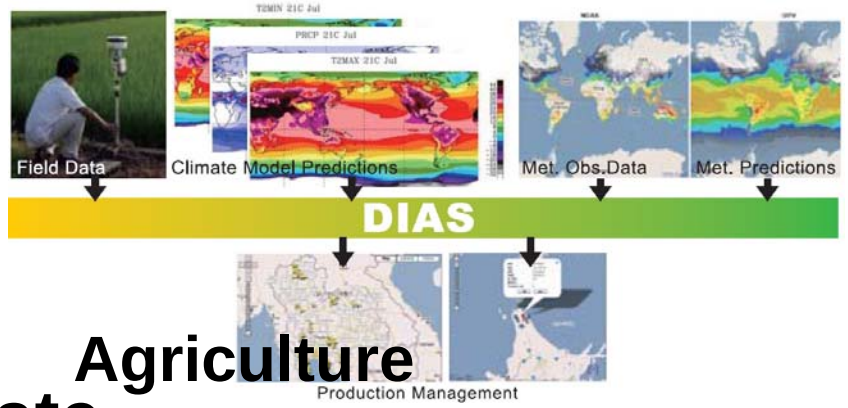
Data Integration and Analysis System

a legacy for Japan's contributions to GEOSS

enabling us to do integrated research and to realize inter-disciplinarity



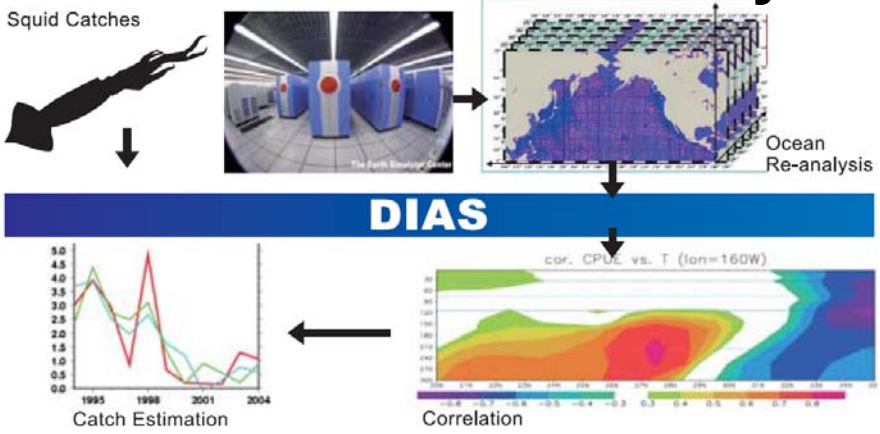
Water



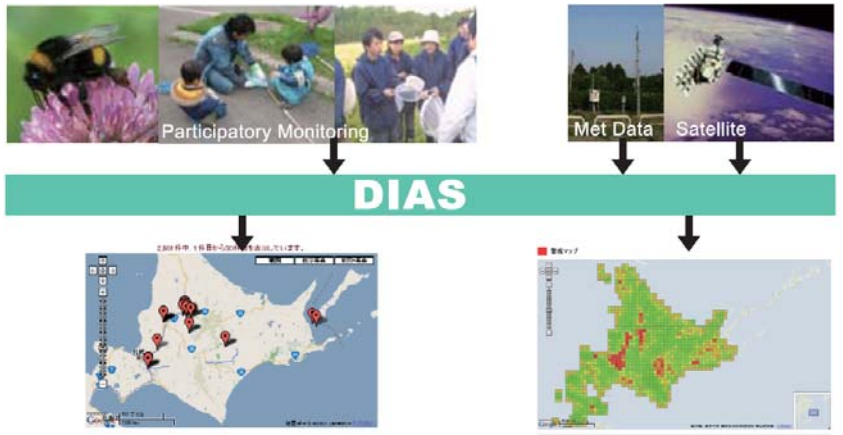
Agriculture

Climate

Fishery



Biodiversity



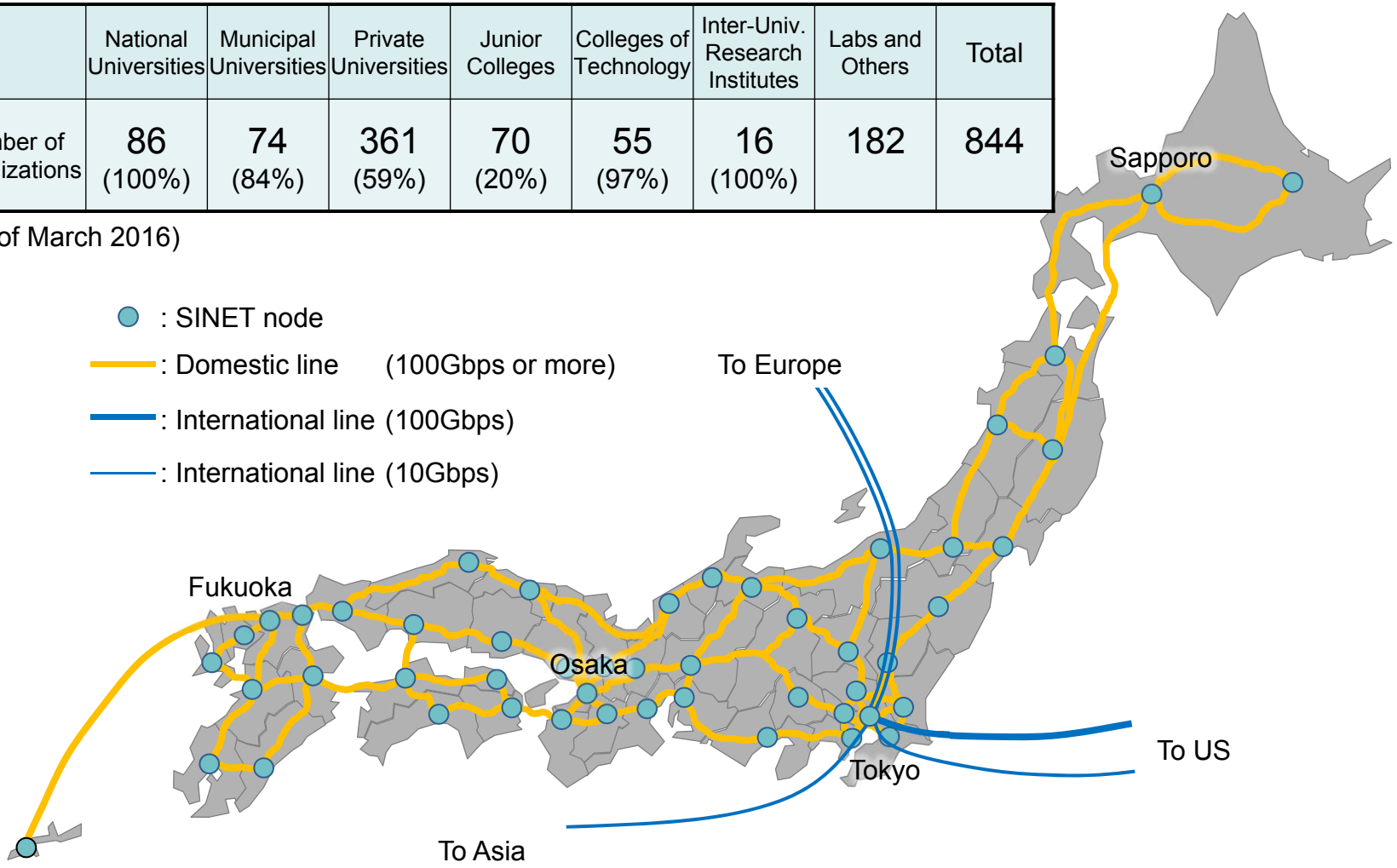
+ High-Speed Network

Science Information Network (SINET)

- SINET is a Japanese academic backbone network for more than 800 universities and research institutions, and for about 3 million users.
  - SINET covers 100% of national, 84% of municipal, and 59% of private universities.

	National Universities	Municipal Universities	Private Universities	Junior Colleges	Colleges of Technology	Inter-Univ. Research Institutes	Labs and Others	Total
Number of Organizations	86 (100%)	74 (84%)	361 (59%)	70 (20%)	55 (97%)	16 (100%)	182	844

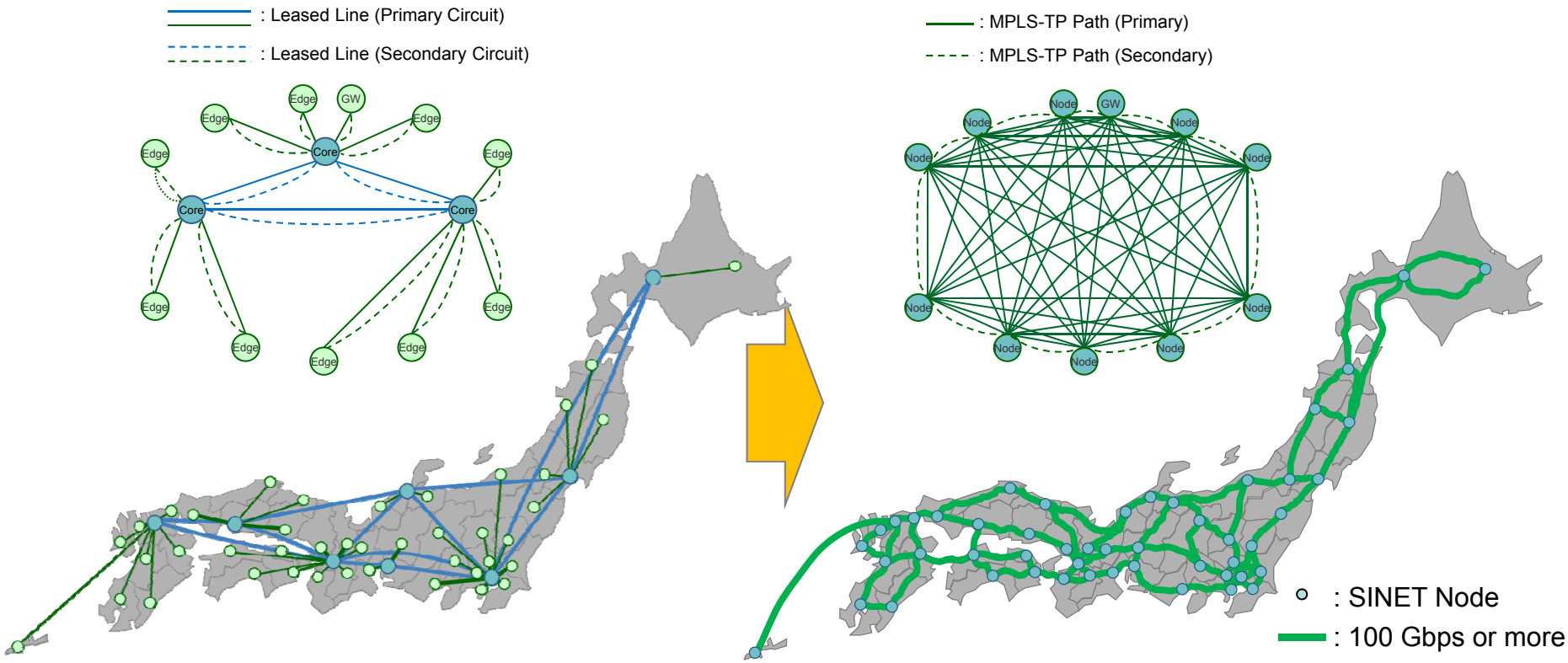
(As of March 2016)



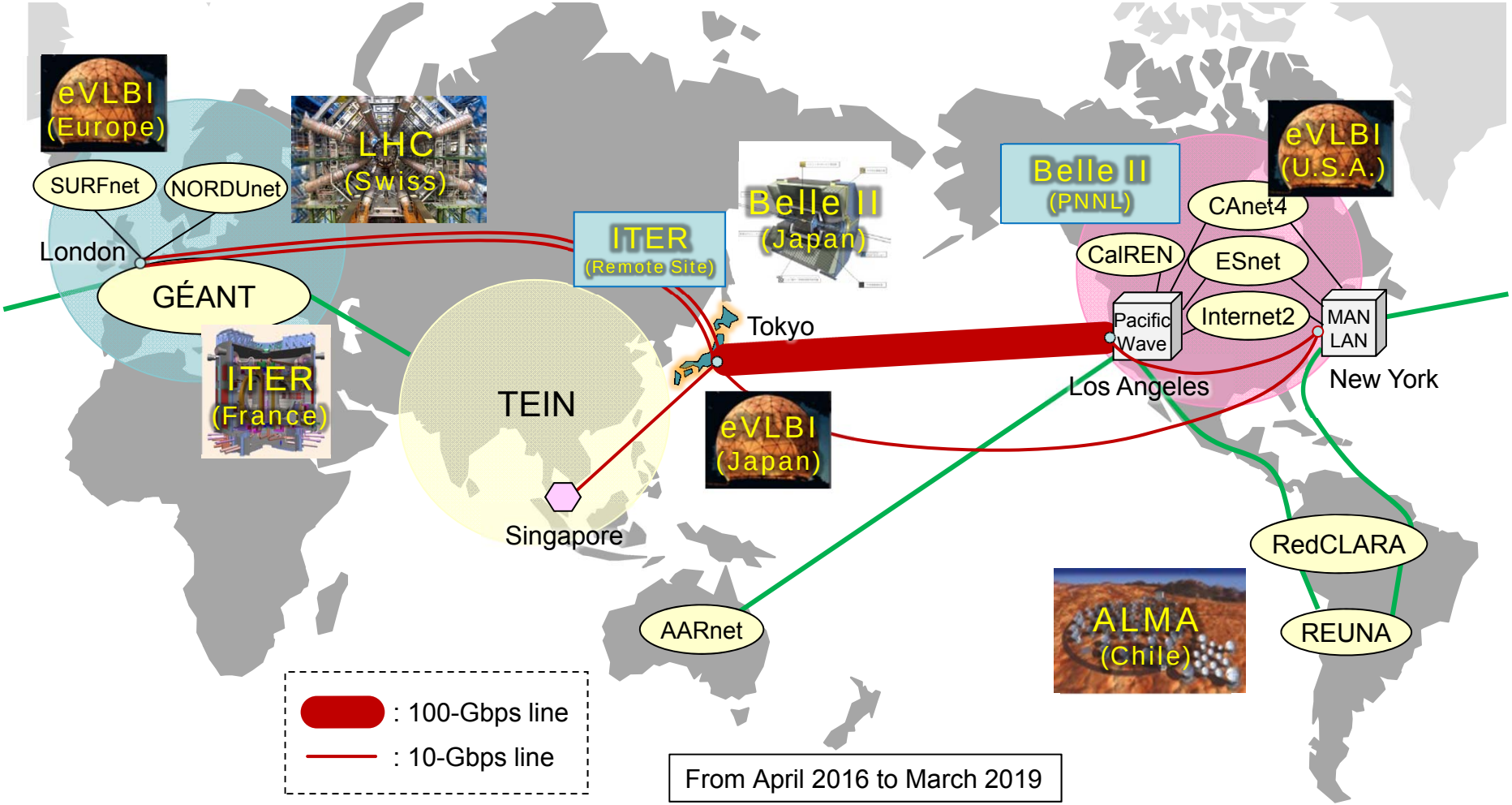
High-Performance, Reliable, and SDN-friendly

- SINET5 directly connects each pair of IP routers by the smallest-latency MPLS-TP path and the disjoint path to it. This fully-meshed topology creates a high-performance, reliable, and SDN-friendly backbone network.

SINET4	SINET5
• Star-like topology	• Fully-meshed topology with redundancy



- SINET5 has direct international lines to USA, Europe, and Asia.
  - USA: 100-Gbps line to Los Angeles, 10-Gbps line to New York, and 10-Gbps backup line
  - Europe: Two 10-Gbps lines to London for small latency
  - Asia: 10-Gbps line to Singapore



Collaboration with HPC

- K computer

- Earth Simulator

Summary

- History of DIAS
- Today's DIAS = challenges to 4V
- DIAS is supported by top-level IT
  - Power saving, High-Speed Network
- "DIAS Value"
  - Model + Observed data
  - Huge scale, Realtime data
  - Quality controlled data with metadata

Data Center → Application Platform

Future Direction of DIAS

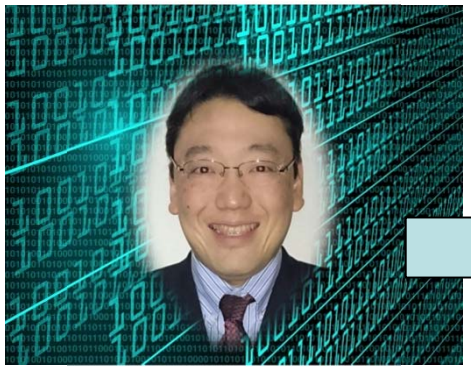
- DIAS=Research Platform

  → Collaboration with commercial sector

- DIAS=data intensive platform

  →Collaboration with more HPC

Thank you for your attention.



Ikoma

Kawasaki