An Advanced Quality Control System for the CEOP/CAMP *In-Situ* Data Management

Katsunori Tamagawa*, Member, IEEE*, Masaru Kitsuregawa*, Member, IEEE*, Eiji Ikoma, Tetsu Ohta, Steve Williams, and Toshio Koike

***Abstract*—The Coordinated Enhanced Observing Period (CEOP) was proposed in 1997 as an initial step for establishing an integrated observation system for the global water cycle. The Enhanced Observing Period was conducted from October 2002 to December 2004, with satellite data, *in-situ* data, and model output data collected and available for integrated analysis. Under the framework of CEOP, the CEOP Asia-Australia Monsoon Project (CAMP) was organized and provided the *in-situ* dataset in the Asian region. CAMP included 13 different reference sites in the Asian monsoon region during Phase 1 (October 2002 to December 2004). These reference sites were operated by individual researchers for their own research objectives. Therefore, the various sites' data had important differences in observational elements, data formats, recording intervals, etc. This usually requires substantial manual data processing to use these data for scientific research which consumes a great deal of researcher time and energy. To reduce the time and effort for data quality checking and format conversion, the CAMP Data Center (CDC) established a web-based Quality Control (QC) System. This paper introduces this *in-situ* data management and quality control system for the Asian region data under the framework of CEOP.**

***Index Terms*—Data management, observers, quality control.**

I. Background

THE COORDINATED Enhanced Observing Period (CEOP), which was an element of the World Climate Research Programme (WCRP) initiated by the Global Energy and Water Cycle Experiment (GEWEX), was proposed in 1997 as an initial step for establishing an integrated observation system for the global water cycle. Its guiding goal was:

> "To understand and model the influence of continental hydroclimate processes on the predictability of global atmospheric circulation and changes in water resources, with a particular focus on the heat source and sink regions that drive and modify the climate system and anomalies" [1].

Thirty-five reference sites (shown in Fig. 1) were organized by the GEWEX Continental Scale Experiments (CSEs) to characterize global climate variation and to collect the necessary *in-situ* data that make up integrated data sets. The Working Group on Information Systems and Services (WGISS) under the Committee on Earth Observation Satellites (CEOS) member agencies and the Working Group on Information Systems and Services (WGISS) provided the satellite products and integration functions. Eleven Numerical Weather Prediction (NWP) centers provided specialized model products (time series, globally gridded, and reanalysis) for each CEOP reference site.

As shown in Fig. 2, the data from the reference sites, satellites, and the NWP models are archived by the National Center for Atmospheric Research (NCAR) of the USA, the University of Tokyo and Japan Aerospace Exploration Agency (JAXA) of Japan, and the Max-Planck Institute (MPI) for Meteorology of Germany, respectively. In addition the value-added Land Data Assimilation products from both regional and global scale processing systems with emphasis on the Global Land Data Assimilation System (GLDAS) are being developed in the USA as a contribution to CEOP.

The first CEOP Enhanced Observing Period (EOP-1), was conducted from July to September in 2001. The first annual enhanced observing period, (EOP-3), was conducted from October 1, 2002 to September 30 2003. The second, extended annual observing period (EOP-4) was conducted from October 1, 2003 to December 31, 2004. CEOP has begun to assemble a database of common measurements from *in-situ*, satellite remote sensing, model output, and 4-D data analyses for these specified periods [2].

The purpose of this paper is to introduce an advanced quality control system for the CEOP Asia-Australia Monsoon Project (CAMP) *in-situ* data based on the CEOP *in-situ* data management system. Section II describes CEOP *in-situ* data management. Section III describes CAMP data management. Section IV describes the development, application and update of the quality control system. Section V provides a demonstration of the QC system. Section VI provides conclusions and Section VII reviews future plans.

II. CEOP In-Situ Data Management

As mentioned in Section I, there were 35 global reference sites (*in-situ* observation sites) contributing data to CEOP. To manage these data, a CEOP *in-situ* data management procedure was discussed at the CEOP Reference Site Managers Workshop

Manuscript received September 30, 2007; revised May 27, 2008. First published August 1, 2008; current version published September 17, 2008. This work was supported in part by the Science Technology Promotion Program and by the National Key Technology Project "Data Integration and Analysis System," which were both supported by the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

K. Tamagawa is with the Earth Observation Data Integration and Fusion Research Initiative, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: tamagawa@hydra.t.u-tokyo.ac.jp).

M. Kitsuregawa is with the Institute of Industrial Sciences, The University of Tokyo, Tokyo 153-8505, Japan (e-mail: kitsure@tkl.iis.u-tokyo.ac.jp).

E. Ikoma is with the Center for Spatial Information Science, The University of Tokyo, Chiba 277-8568, Japan (e-mail: eikoma@csis.u-tokyo.ac.jp).

T. Ohta and T. Koike are with the Department of Civil Engineering, The University of Tokyo, Tokyo 113-8656, Japan (e-mail: ohta@hydra.t.u-tokyo.ac.jp; tkoike@hydra.t.u-tokyo.ac.jp).

S. Williams is with the Earth Observing Laboratory, the National Center for Atmospheric Research, Boulder, CO 80301 USA (e-mail: sfw@ucar.edu).

Digital Object Identifier 10.1109/JSYST.2008.927710

1932-8184/$25.00 © 2008 IEEE

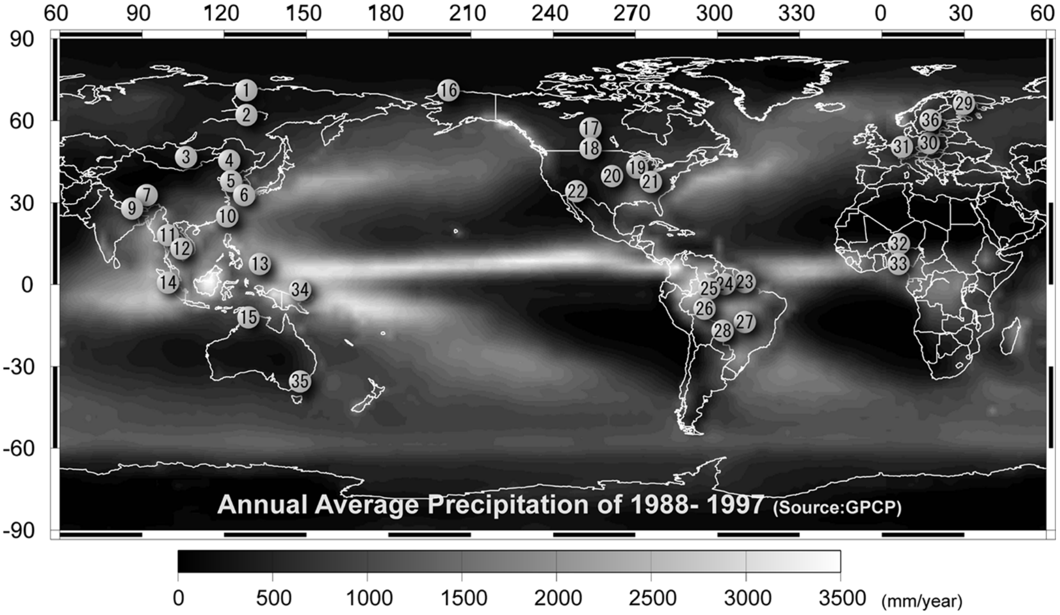


Fig. 1. Thirty-five globally distributed CEOP reference sites.

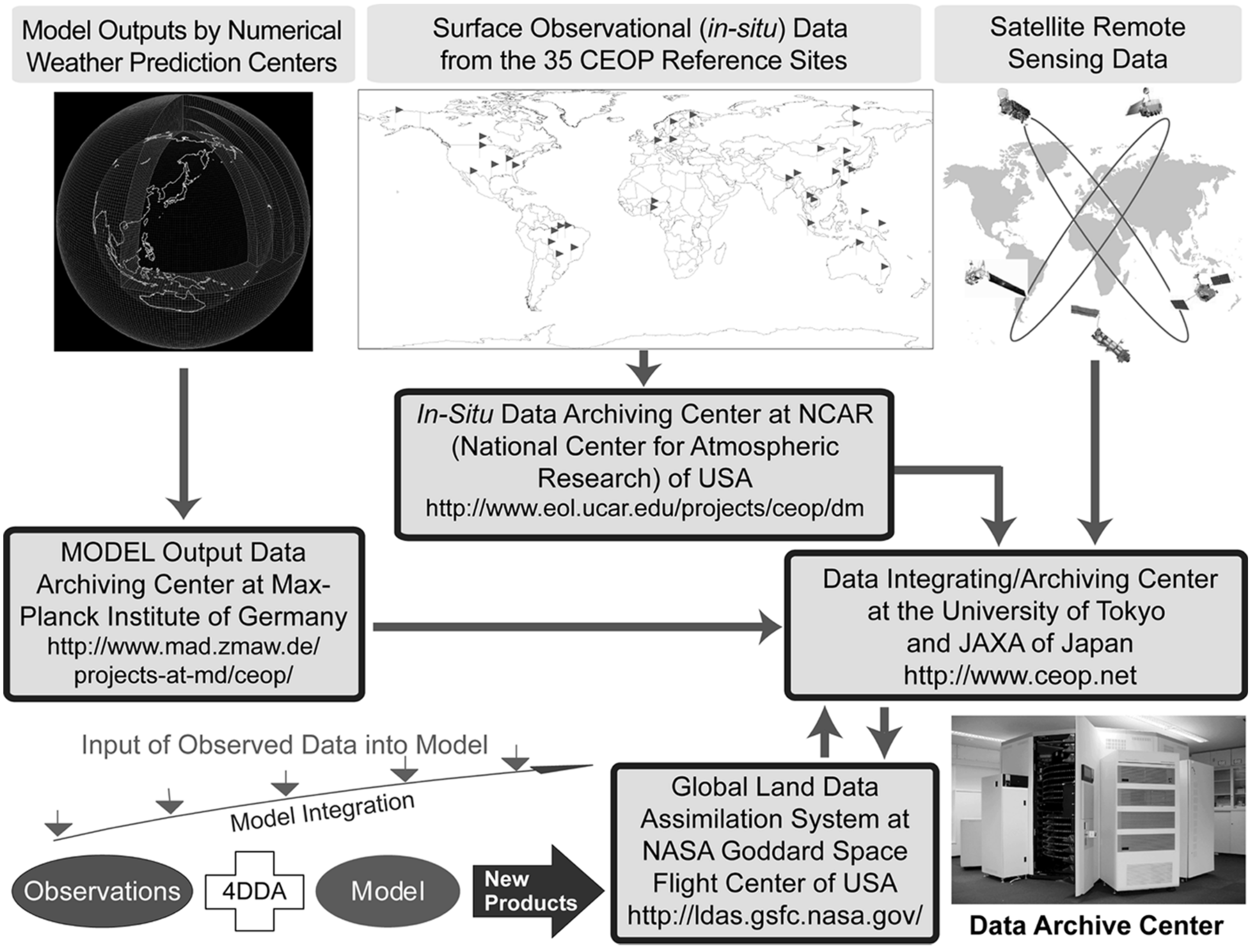

Model Outputs by Numerical Weather Prediction Centers
Surface Observational (*in-situ*) Data from the 35 CEOP Reference Sites
Satellite Remote Sensing Data
In-Situ Data Archiving Center at NCAR (National Center for Atmospheric Research) of USA http://www.eol.ucar.edu/projects/ceop/dm
MODEL Output Data Archiving Center at Max-Planck Institute of Germany http://www.mad.zmaw.de/projects-at-md/ceop/
Data Integrating/Archiving Center at the University of Tokyo and JAXA of Japan http://www.ceop.net
Input of Observed Data into Model
Model Integration
Observations
4DDA
Model
New Products
Global Land Data Assimilation System at NASA Goddard Space Flight Center of USA http://ldas.gsfc.nasa.gov/
Data Archive Center

Fig. 2. First global integrated data sets of the water cycle.

(March 31, 2003 and April 1, 2003, Berlin, Germany) and at the CEOP 2nd Implementation Planning Meeting (April 2–4, 2003, Berlin).

The detailed descriptions such as data format specifications, required parameters (both observed and derived), the units, values specifying missing data values, equations used to derive

TABLE I
CEOP *IN-SITU* DATA FLAG DEFINITIONS

Flag Value	Definition [a]
C	Reported value below or above threshold value
M	Parameter value missing OR Derived parameter can not be computed
B	Bad
I	Interpolated/Estimated/Gap Filled
D	Questionable/Dubious
G	Good
U	Unchecked

parameters, and the definitions of data flags and file naming conventions were discussed in these meetings, as were the metadata requirements.

A. In-Situ Unified Dataset Formats

The CEOP Reference Site data are organized using the following four unified datasets (surface meteorological and radiation data set, flux data set, soil temperature and soil moisture data set, and meteorological tower data set).

- CEOP surface meteorological and radiation dataset (SFC) format contains eight metadata parameters and 38 data parameters and flags.
- CEOP flux dataset format contains nine metadata parameters and eight data parameters and flags.
- CEOP soil temperature and soil moisture dataset format contains nine metadata parameters and four data parameters and flags.
- CEOP meteorological tower dataset format contains 9 metadata parameters and 18 data parameters and flags.

B. Data Flags

The CEOP *in-situ* data flag values and definitions are shown in Table I. Detailed information is available from the CEOP data management web site [3].

III. CAMP DATA MANAGEMENT

Until December 31, 2004, the reference site datasets in the Asia region were managed under the framework of CAMP, which included 13 different reference sites in the Asian region. The CAMP project covered the whole Asian Monsoon area with diverse climate from tundra to tropical and high altitude zones. The 13 CAMP reference sites are shown in Fig. 3 and their characteristics are listed in Table II.

A. CAMP Reference Site Characteristics

The CAMP reference sites are operated by individual researchers for their own research objectives. Therefore, the various sites' data had a great deal of variety in observational elements, data format, recorded intervals, etc.

Processing these data for scientific research usually requires manual data processing such as downloading data from a data logger, checking the number of records, scanning for periods of missing data, checking the data quality, converting between formats, etc. This typically consumes a great deal of a researcher's time and energy to complete this processing.

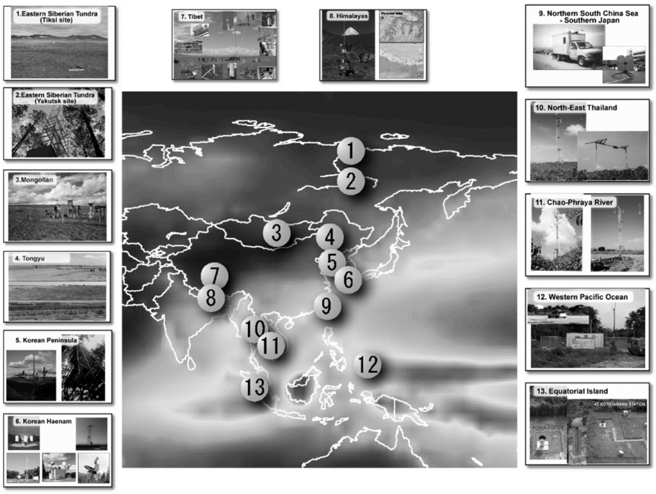


Fig. 3. Thirteen CAMP reference sites.

TABLE II
CHARACTERISTICS OF CAMP REFERENCE SITES

	Reference Site Name	Num. of Stations	Observation
1	Eastern Siberian Tundra	1	Tower
2	Eastern Siberian Taiga	7	Tower
3	Mongolia	18	AWS, ASSH, Radar
4	Tongyu	2	Tower
5	Korean Peninsula	1	Tower
6	Korean Haenam	1	Tower
7	Tibet	14	AWS, Tower, SMTMS, Sonde
8	Himalayas	5	AWS
9	Northern South China Sea - Southern Japan	25	AWS, SMTMS, Sonde
10	Chao-Phraya River	1	Tower
11	North-East Thailand	1	Tower
12	Western Pacific Ocean	2	AWS
13	Equatorial Island	1	AWS

Tower	: Planetary Boundary Layer Tower
AWS	: Automatic Weather Station
ASSH	: Automatic Station for Soil Hydrology
Radar	: Rain Intensity Radar
SMTMS	: Soil Moisture Temperature Measurement System
Sonde	: Radio Sonde
STMS	: Soil Temperature Measurement System

When an observer sees questionable or strange data, he or she is the best person to judge whether the data is good or bad, because he or she is most knowledgeable about the data, the physical phenomena, the local conditions, and the instruments.

To reduce the work and time for data checking and format conversion, the University of Tokyo team established a web-based Quality Control System that data providers can use to check data through a web site.

This system was developed jointly with the Intelligent Information Technology Group of the University of Tokyo.

B. CAMP Data Management Flow

The data management flow for CAMP is shown Fig. 4. The right side of this figure shows the functions of individual data

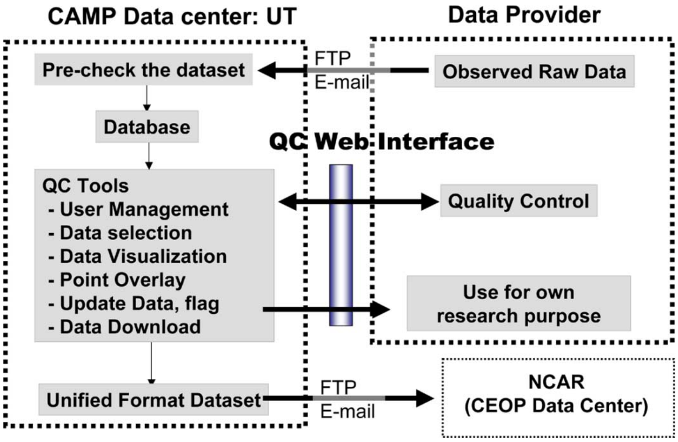


Fig. 4. Data management flow for CAMP data.

providers (data provider) and left side shows the functions of the CAMP Data center. The observed raw data is sent to the CAMP Data Center by FTP or e-mail attachment, where the CAMP data center prechecks and compiles the database.

Each Reference Site's data are protected by a user ID and password, so that no third party has access or use of their unchecked data. When the data are uploaded into the database, the data provider (user) can start with the quality check procedure through their individual account. By using the QC Tools (such as data selection, data visualization update data, flag) the user can correct the values and the flags online, updating the CAMP data center's database.

After the quality check procedure has been completed, the data provider can use their data for their own research by downloading it through the web interface. These data are then sent to CEOP Data Center at NCAR for final archival and dissemination.

IV. Development, Application, and Update of the Quality Control System [4]

A. Development of the Quality Control System

System Structure: The quality control system consists of the following six basic functions.

- *User Management:* The quality control system is used by many data providers. But access to the raw data itself must be protected from others, so this system manages the user's data access rights.
- *Data Selection:* The quality control target data are selected by narrowing down the "Observation Station," "Observation Element," and "Observation Date."
- *Data Visualization:* Selected data are displayed at the data visualization window using different colors.
- *Data Overlay:* The data which are related to the quality control data are also selected and can be overlaid with the target data.
- *Update Data and Flag:* The data and its quality flag are updated using the data update window.
- *Data Download:* Quality controlled data are downloaded from the data download window.

Compiling to Database: The individual site data are compiled to the database by the administrator. Before compiling the database, the data precheck is performed to remove typical errors.

When the data are compiled to the database, all data are compared with threshold values which are specified by the user. If the value is below or above threshold values its flag value is set to [C] automatically. If there is a missing value, its flag value is set to [M], and the unchecked flag value [U] is set for data which have not been checked.

Typical Raw Data Error Pattern: Fig. 5 shows some typical raw data error patterns. Fig. 5(a) shows the deviance of the data between rows. Data is out of alignment on and after line 6. Fig. 5(b) shows the inclusion of a new column. Data suddenly appears in column F beginning with line 6. Fig. 5(c) shows the inclusion of character strings. Two different data values are connected and converted into a false number, sometimes resulting in long strings of data. Fig. 5(d) shows the discontinuity of the time stamp. In this case, the data should be time sequential, but sequential time stamps are missing. Fig. 5(e) show the duplication of a time stamp. The same time stamp, but different values appear on lines 3 and 4.

B. Application of the Quality Control System

This system was provided to the CAMP Reference Site Managers. They provided their observed raw data to the CAMP data center and performed the quality control themselves by using this system.

The First version of the quality control system was utilized for the first half of EOP-3 (October 2002 to March 2003). Based on their experiences, the CAMP data managers provided feedback and made suggestions and comments to improve the operability of the QC system.

C. Update of the Quality Control System

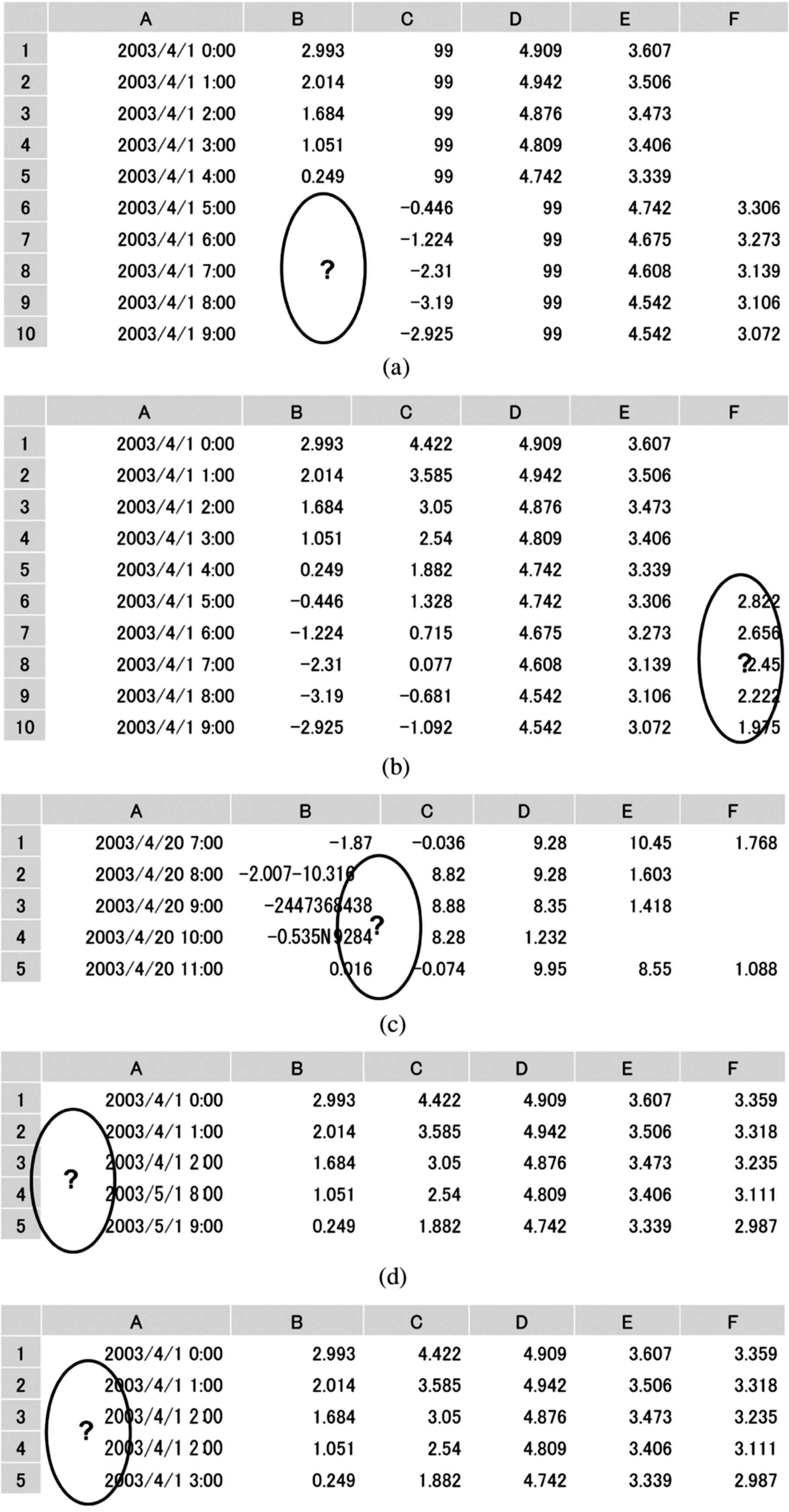
Incorporating these suggestions, the Quality Control System was upgraded to version 2, notably, as follows:

- improvement in the response time by system reconstruction;
- improving the flagging of datasets with a mouse selection function;
- improvement in the operability of graphical data representation;
- task process management.

D. Comparison of Quality Control System Between Versions 1 and 2

With an upgraded version 2 QC system, efficiency of users have improved approximately 26%–43%, demonstrating that the upgraded functions had significantly reduced the time spent by users.

As for the comparison of number of sessions between versions 1 and 2, the number of sessions needed to process a given set of data generally decreased with version 2. On the other hand, the processing time per one session increased by 122%–226%. It is assumed that users could accomplish more per session using version 2, and thus were able to complete their work with fewer sessions. It was also recognized that the number of days required for completion of a data checking job was greatly reduced.



	A	B	C	D	E	F
1	2003/4/1 0:00	2.993	99	4.909	3.607	
2	2003/4/1 1:00	2.014	99	4.942	3.506	
3	2003/4/1 2:00	1.684	99	4.876	3.473	
4	2003/4/1 3:00	1.051	99	4.809	3.406	
5	2003/4/1 4:00	0.249	99	4.742	3.339	
6	2003/4/1 5:00		−0.446	99	4.742	3.306
7	2003/4/1 6:00		−1.224	99	4.675	3.273
8	2003/4/1 7:00		−2.31	99	4.608	3.139
9	2003/4/1 8:00		−3.19	99	4.542	3.106
10	2003/4/1 9:00		−2.925	99	4.542	3.072

(a)

	A	B	C	D	E	F
1	2003/4/1 0:00	2.993	4.422	4.909	3.607	
2	2003/4/1 1:00	2.014	3.585	4.942	3.506	
3	2003/4/1 2:00	1.684	3.05	4.876	3.473	
4	2003/4/1 3:00	1.051	2.54	4.809	3.406	
5	2003/4/1 4:00	0.249	1.882	4.742	3.339	
6	2003/4/1 5:00	−0.446	1.328	4.742	3.306	2.822
7	2003/4/1 6:00	−1.224	0.715	4.675	3.273	2.656
8	2003/4/1 7:00	−2.31	0.077	4.608	3.139	2.45
9	2003/4/1 8:00	−3.19	−0.681	4.542	3.106	2.222
10	2003/4/1 9:00	−2.925	−1.092	4.542	3.072	1.975

(b)

	A	B	C	D	E	F
1	2003/4/20 7:00	−1.87	−0.036	9.28	10.45	1.768
2	2003/4/20 8:00	−2.007−10.316	8.82	9.28	1.603	
3	2003/4/20 9:00	−2447368438	8.88	8.35	1.418	
4	2003/4/20 10:00	−0.535N9284	8.28	1.232		
5	2003/4/20 11:00	0.016	−0.074	9.95	8.55	1.088

(c)

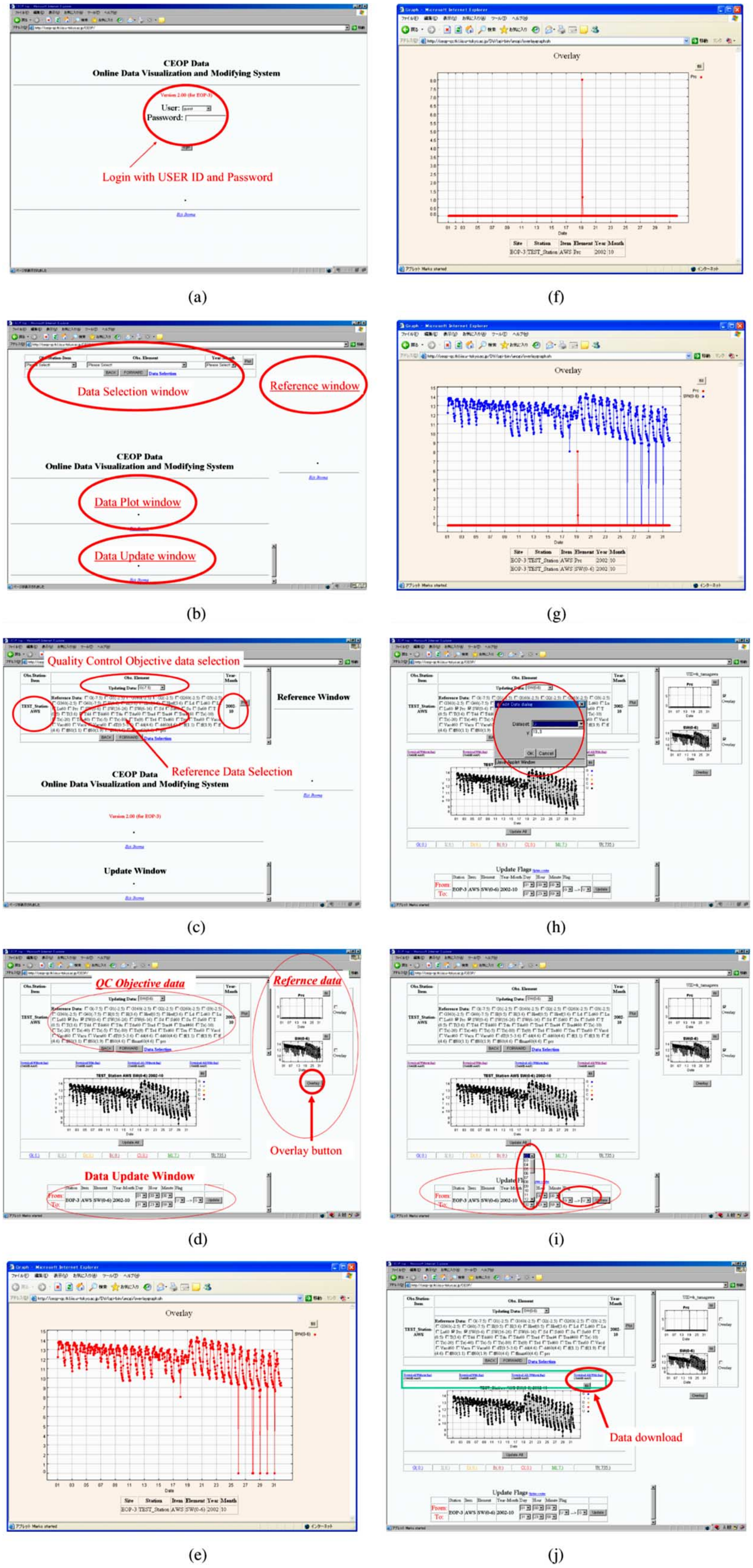
	A	B	C	D	E	F
1	2003/4/1 0:00	2.993	4.422	4.909	3.607	3.359
2	2003/4/1 1:00	2.014	3.585	4.942	3.506	3.318
3	2003/4/1 2:00	1.684	3.05	4.876	3.473	3.235
4	2003/5/1 8:00	1.051	2.54	4.809	3.406	3.111
5	2003/5/1 9:00	0.249	1.882	4.742	3.339	2.987

(d)

	A	B	C	D	E	F
1	2003/4/1 0:00	2.993	4.422	4.909	3.607	3.359
2	2003/4/1 1:00	2.014	3.585	4.942	3.506	3.318
3	2003/4/1 2:00	1.684	3.05	4.876	3.473	3.235
4	2003/4/1 2:00	1.051	2.54	4.809	3.406	3.111
5	2003/4/1 3:00	0.249	1.882	4.742	3.339	2.987

(e)

Fig. 5. Typical raw data error patterns. (a) Deviation of data between rows. (b) Sudden appearance of additional value/row/column. (c) Convolution of character strings/values/data. (d) Discontinuity of time stamp. (e) Duplication of time stamp.



CEOP Data
Online Data Visualization and Modifying System

Version 2.00 (for EOP-3)

User:
Password:

Login with USER ID and Password

(a)

Data Selection window

Reference window

Data Plot window

Data Update window

(b)

Quality Control Objective data selection

Reference Window

Reference Data Selection

Update Window

(c)

QC Objective data

Refernce data

Overlay button

Data Update Window

(d)

Overlay

Site	Station	Item	Element	Year	Month
EOP-3	TEST_Station	AWS	SW(0-6)	2002	10

(e)

Overlay

Site	Station	Item	Element	Year	Month
EOP-3	TEST_Station	AWS	Prc	2002	10

(f)

Overlay

Site	Station	Item	Element	Year	Month
EOP-3	TEST_Station	AWS	Prc	2002	10
EOP-3	TEST_Station	AWS	SW(0-6)	2002	10

(g)

Update Flags

(h)

(i)

Data download

Update Flags

(j)

Fig. 6. Demonstration of the Quality Control System.

V. Demonstration of the QC System

Fig. 6 shows a demonstration of the Quality Control System.

A. Login and Opening Window

Fig. 6(a) shows the login window, where a registered member can login with user ID and password.

After login, the four windows appear on the display [see Fig. 6(b)]: the upper one is the data selection window, the middle one is the data plot window, the lower one is the data update window, and the right one is the reference window.

B. Data Selection

The user can select a set of target data (to be checked and assigned QC flags) according to Obs, Station, Obs Element, and Year-Month, respectively. The reference data are also selected in this window [see Fig. 6(c)].

C. Data Visualization

By selecting the "Plot" button, Quality control objective data graph, and update window, the reference data window appears [see Fig. 6(d)].

D. Reference Data Visualization and Its Use

By selecting the overlay button, an overlaid graph is displayed. It is well known that soil moisture and precipitation have a strong correlation [see Fig. 6(e)] shows the soil moisture distribution data for one month observed at a station. The soil moisture data increases substantially at around the last week and in the middle of the observed month. But the user cannot judge definitively whether this data variation is correct or not from only this figure.

This might be due to the observation sensor experiencing some problems. But if the user selects the precipitation data as reference data as shown in Fig. 6(f) and overlays it with the soil moisture data, one will see that the soil moisture increase accompanies a precipitation event as shown in Fig. 6(g). As a result, the user can judge if this soil moisture data distribution is reasonable and he or she has confidence to set a "Good " flag to the data for this entire period.

E. Data and Data Flag Update

By selecting the data on the graph, the edit data dialog appears [see Fig. 6(h)]. The user can then set their chosen data value and its flag. After that the updated data and its flag is displayed in a different color.

Additionally, the user can change the data flag on a data block by using the update flag window. The user can specify the selected date and flag from the pull down menu [see Fig. 6(i)].

F. Data Download

After applying this process to the complete dataset, the quality control task is finished. The user can then download his/her own data set for their research purposes [see Fig. 6(j)].

VI. Conclusion

Since most of the reference sites in the Asian region are managed by an individual researcher's efforts for their own scientific research purposes, and by their own research funding, it had been very difficult in CEOP to obtain, collect, and quality control data. However, this QC system has been developed and introduced to the reference site managers and they have understood and appreciated the advantages of such a system, with a variety of observation items, observing intervals, and easy data formatting. Gradually they all tried to cooperate with CEOP by using this QC system, and began to provide the data. This resulted in data submission from most sites with greatly improved data quality.

In this research, we tried to solve various problems (especially data loading), and this was achieved mainly due to good cooperation between the IT group and each reference site manager. The two primary results are the following: 1) error correction has improved in loading raw data into the database, by analyzing all the errors systematically and 2) users' comments and suggestions to improve the QC system were collected and are reflected in this system upgrade by the IT group. Thus, the CAMP Data Center provided an effective interface between the IT group and users and contributed to the system's functional upgrading. These efforts should be examined as a possible prototype method for contributing data to GEOSS.

VII. Future Plan

Future tasks and possibilities include how to reduce the time for the processes of data checking, data loading, quality control, data formatting, data converting, etc. There is also the challenge of relevant metadata production at the same time that the data is being loaded, which will be realized by further collaboration with the IT group.

Acknowledgment

This Quality Control System has been established in cooperation with the Information Technology (IT) Group of the University of Tokyo and the CEOP/CAMP reference site managers. The authors would like to thank them for their kind suggestions and comments on this system.

References

[1] IGPO Publication Series, "Coordinated enhanced observing period (CEOP) implementation plan," 2001, vol. 36, p. 4.

[2] T. Koike, "The coordinated enhanced observing period—An initial step for integrated global water cycle observation," *WMO Bulletin*, vol. 53, no. 2, pp. 115–121, 2004.

[3] National Center for Atmospheric Research (NCAR), Earth Observing Laboratory (EOL), Boulder, CO, "Coordinated Enhanced Observing Period (CEOP) data management," 2008. [Online]. Available: http://www.eol.ucar.edu/projects/ceop/dm/

[4] E. Ikoma, K. Tamagawa, T. Ohta, T. Koike, and M. Kitsuregawa, "QUASUR: Web-based quality assurance system for CEOP reference data," *J. Meteorological Society Japan*, vol. 85A, pp. 461–473, 2007.

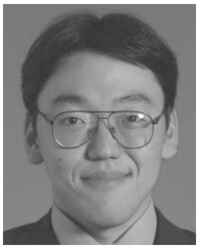
Katsunori Tamagawa (M'08) received the B.A. and M.A. degrees in the field of engineering from the Nagaoka University of Technology, Niigata, Japan, in 1993 and 1995, respectively.

Since 2000, he has been with the University of Tokyo, Tokyo, Japan, where he is currently a Research Associate with the Earth Observation Data Integration and Fusion Research Initiative (EDITORIA). He concurrently serves as Data Manager of the Coordinated Enhanced Observing Period (CEOP) and the Asian Water Cycle Initiative (AWCI) Data Archive Center with the University of Tokyo to mainly handle *in-situ* observation data in the Asian region.

Masaru Kitsuregawa (M'81) received the Ph.D. degree from the University of Tokyo, Tokyo, Japan, in 1983.

He is currently a Full Professor and a Director of the Center for Information Fusion, Institute of Industrial Science, the University of Tokyo. His current research interests cover database engineering, Web archive/mining, advanced storage system architecture, parallel database processing/data mining, digital earth, transaction processing, etc. He served as a program co-chair of the IEEE International Conference on Data Engineering (ICDE, 1999), and served as general co-chair of ICDE'05 (Tokyo). He served as a VLDB trustee and an ACM SIGMOD Japan Chapter chair.

Prof. Kitsuregawa is a fellow of the Information Processing Society of Japan (IPSJ) and Information, Communication, Engineering, Japan (IEICE) and he currently serves a Vice President of IPSJ. He also serves a Science Advisor of Ministry of Education, Culture, Sports, Science, and Technology. He is a member of the IEEE Computer Society.



Eiji Ikoma received the B.S., M.S, and Ph.D. degrees in information and communication engineering from the University of Tokyo, Chiba, Japan, in 1995, 1997, and 2000, respectively.

He is currently an Assistant Professor with the Center for Spatial Information Science, The University of Tokyo. His research interests revolve around database system, user interface, data visualization and data mining. He is now working with a research group about earth observation, water cycle management, climate change, and earth simulation.

Tetsu Ohta received the B.Eng. and M.Eng. degrees from Nagaoka University of Technology, Niigata, Japan, in 1998 and 2000, respectively.

He is currently a Research Associate with the River and Environmental Engineering Laboratory, Department of Civil Engineering, University of Tokyo, Tokyo, Japan. His research interests include passive remote sensing of soil moisture and data processing.

Steve Williams received the B.S. degree in meteorology from the University of Lowell, Lowell, MA, in 1977, and the M.S. degree in atmospheric science from Texas A&M University, College Station, in 1979.

Since 1991, he has been with National Center for Atmospheric Research (NCAR), Boulder, CO, where he is currently an Associate Scientist IV and Head of the Data Management Group of the Computing, Data, and Software Facility, Earth Observing Laboratory (EOL). His research interests include scientific data and information management, data analysis and quality assurance, multi-disciplinary field project planning, and earth observation climatology.

Toshio Koike received the B.A., M.A., and Ph.D. degrees in engineering from the University of Tokyo, Tokyo, Japan, in 1980, 1982, and 1985, respectively.

Since 1999, he has been a Professor with the University of Tokyo. His research interests include the water cycle sciences and their applications to water resources management, which can be classified into the following three components, establishment of satellite remote sensing, development of the data integration and information fusion system, and development of the hydrological down-scaling methods including Land Data Assimilation System (LDAS) and Cloud Microphysics Data Assimilation System (CMDAS).