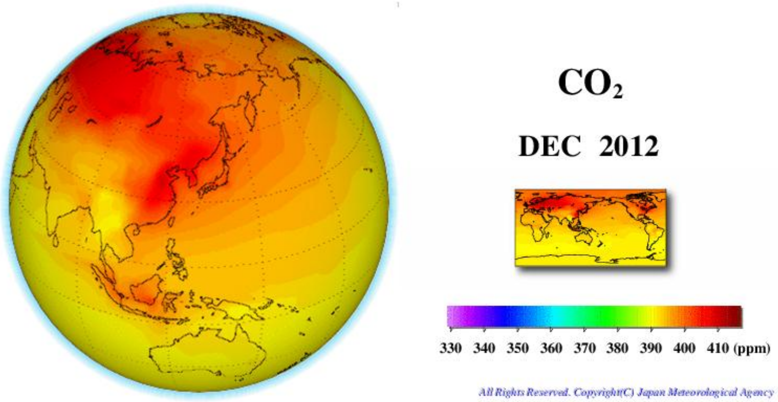
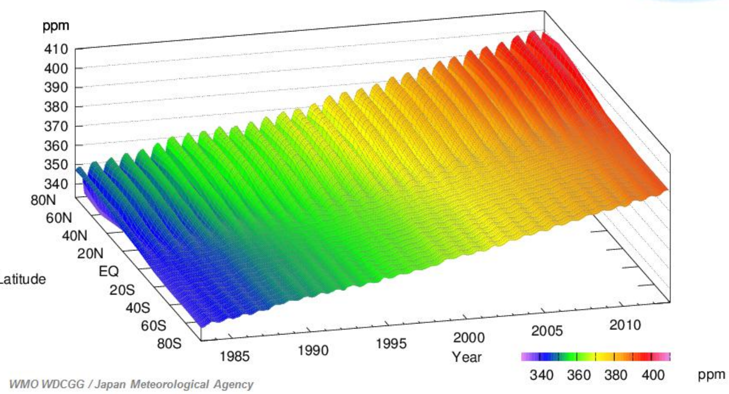
Quality Control and Quality Assurance from WDCGG viewpoint



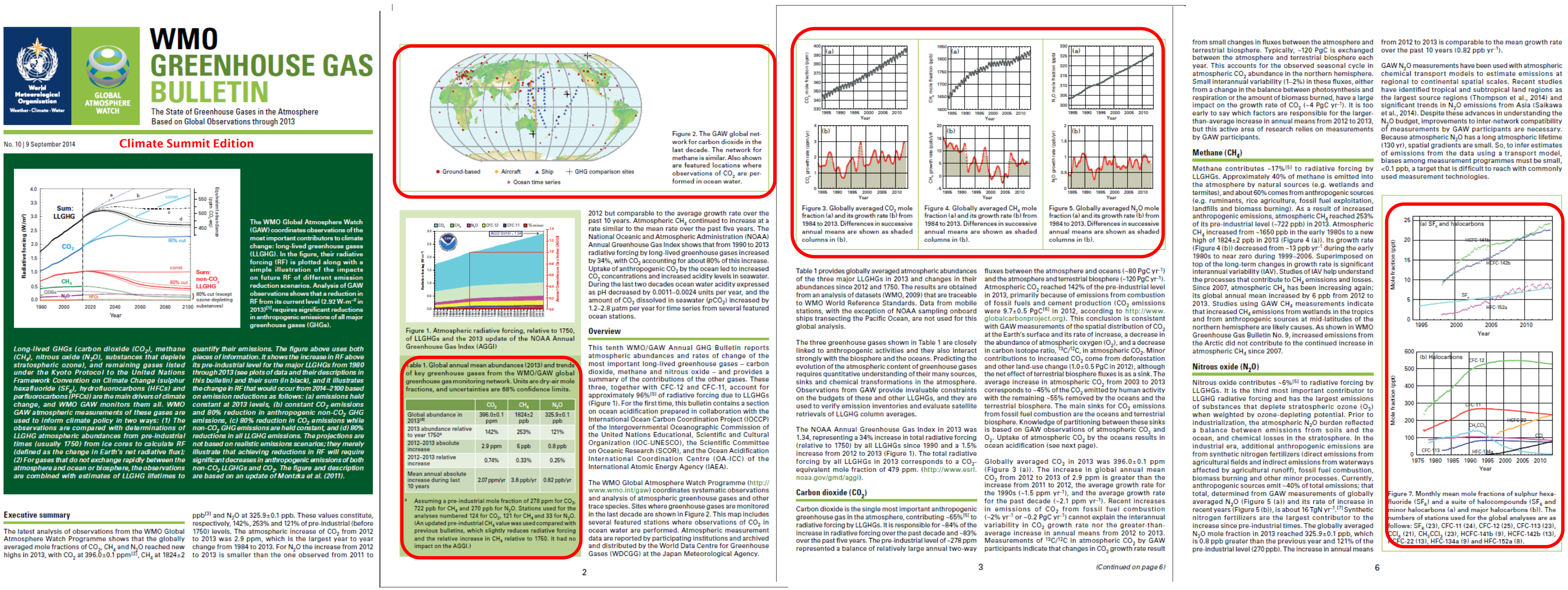


Hiroshi Koide, with colleagues in ad-hoc data archive management team (GGMT-2013), RG-SAG and ET-WDC

Topics

1. How to maintain reliable historical time series
2. How to ensure the propagation of correct metadata to users
3. Discussion on Flagging
4. A common software for Quality Assurance Workflow at each laboratory

World Meteorological Organization (WMO)
Global Atmosphere Watch (GAW)
World Data Centres



WMO Greenhouse Gas Bulletin No.10
for UN Climate Summit on 23 Sep. 2014

Future Perspectives

WMO OMM GAW

WIS/WIGOS, Other Scientific Community

GAW Data Policy
GAW Strategic Plan

GHG-SAG

RG-SAG

Scientific Advisory

WDCGG

JMA

- **Commit to align the needs of users and submitters alike.**
- **Permanent maintenance of DATA archives**
- **Quality assurance and control for scientific accountability**
- **Better notification and compliance of the data policy**
- **Enhance interoperability**

Contribution to
the International Science Community

for Data Providers

Monitoring: Data Registration Number

- Simplification of reporting procedure
- Preparation of the user Information for submitters
- Feedback Information on Characteristics of Data
- Enhance the relationship between submitters and the data centre

for Data Users

Monitoring:
Download Number
Google Scholar Hit Number

- Improved interface
- Consolidated flagging
- Tools for better data use
- Preparation of ISO compliant metadata
- Provision of reliable products (Data assimilation)

IPCC
UNFCCC

Future Perspectives

WMO OMM GAW

WIS/WIGOS, Other Scientific Community

GHG-SAG

RG-SAG

Why reform?

- Financial pressure to GHG & RG measurements
- Model and Assimilation
- Cyber Infrastructure
- Stabilize the WDCGG services

Our ultimate goal

- To serve better to users and contributors alike

…ion to
…ence Community

…Users

…umber
…olar Hit Number

…erface
… flagging
…er data use
…f ISO
…tadata
…eliable
…a

for Data P…

Monitoring: Dat…

- Simplificatio… procedure
- Preparation … Information
- Feedback In… Characterist…
- Enhance the … between sub… and the data centre

interoperability

IPCC
UNFCCC

A Tiny Group for Data Archive Management

A Tiny Group for Data Archive Management


Lynn Hazan, ICOS/ATC


Ken Masarie, NOAA/ESRL


Ludwig Ries, UBA Germany

A Tiny Group for Data Archive Management


Lynn Hazan, ICOS/ATC

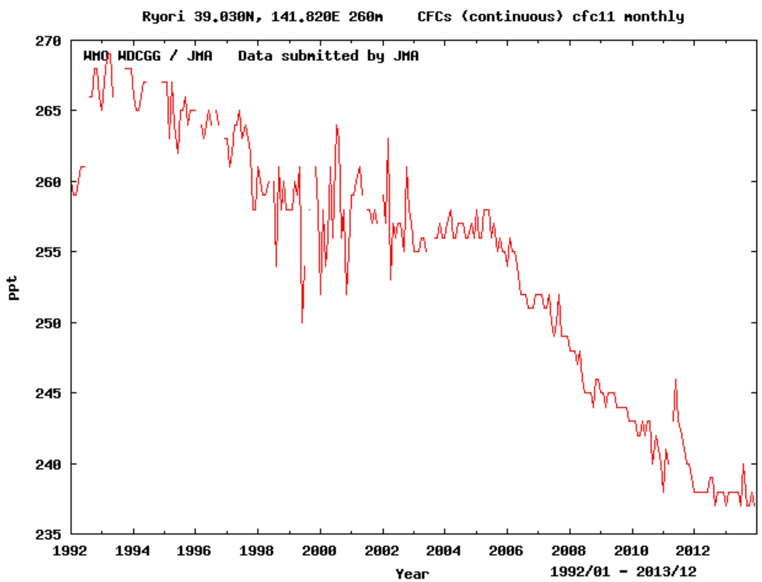

Ludwig Ries, UBA Germany

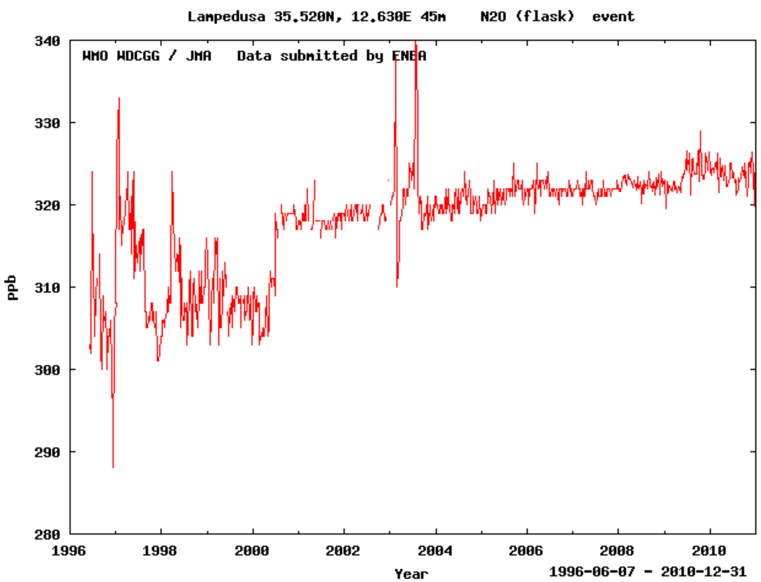
Table 2 - WMO GAW members who have offered to share expertise

Name	Contact Email	Lab	Location	Area of Expertise
WDCGG	hkoide@met.kishou.go.jp	JMA	Japan	Data management in WDCGG
Lynn Hazan	lynn.hazan@lsce.ipsl.fr	LSCE	France	Data management
Paul Krummel	paul.krummel@csiro.au	CSIRO	Australia	Quality control, non-CO2 scale conversions, inter-comparisons
NOAA data team	kenneth.masarie@noaa.gov	NOAA	United States	Data management, quality control, scale conversion
Ludwig Ries	ludwig.ries@uba.de	UBA	Germany	Data acquisition, management and quality control. Software solutions available for data acquisition, instrument control, calibration processing, interactive data preparation and validation.
Martin Steinbacher	martin.steinbacher@empa.ch	EMPA	Switzerland	Data acquisition and processing with commercially available and custom-built software
Doug Worthy	Doug.worthy@ec.gc.ca	EC	Canada	Near real-time data processing via GC, NDIR, and CRDS technologies

Consolidated Time Series

Q : How you can properly merge the historical time series when you replace your instrument?





Parallel Data Streams

Q : How you can utilize parallel data streams at the same site and for the same parameter, for example, by two different instruments or with both in situ measurement and flask sampling?

Parallel Data Streams

Two Controversial Approach

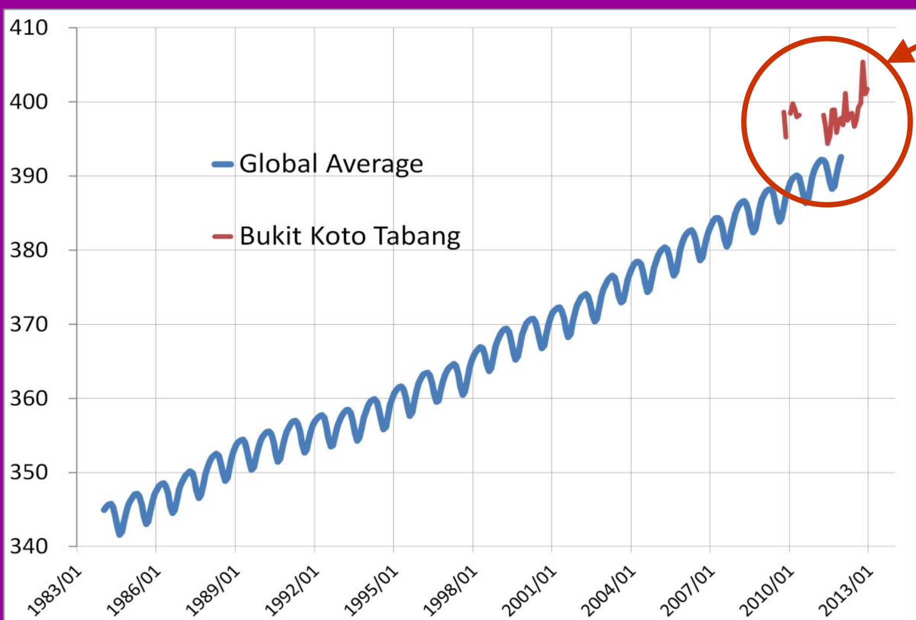
1. Somehow to merge the multi-streams into one consolidated representative time series per station per parameter

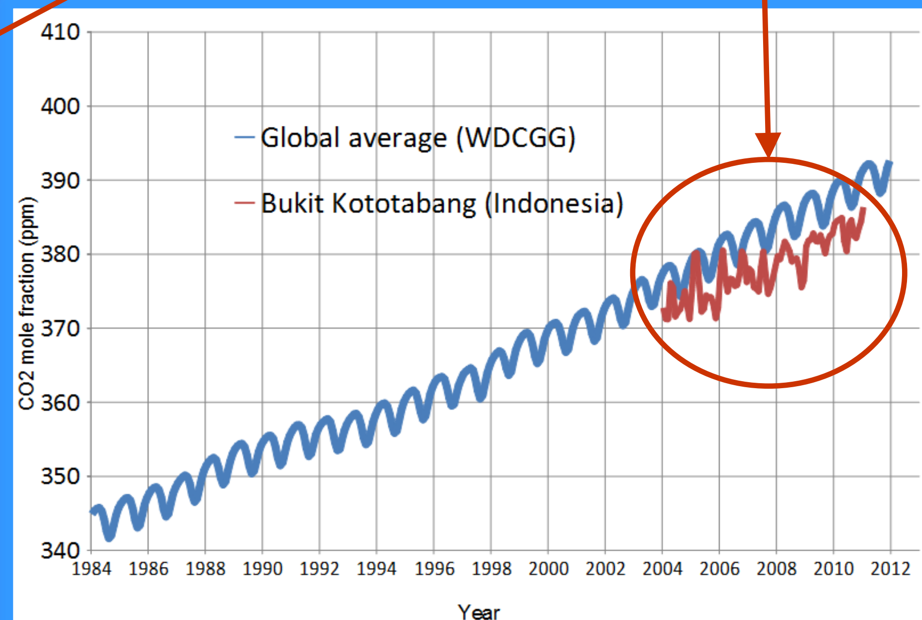
2. The original information from different data streams is precious and useful.

Parallel Data Streams

From the WDCGG Archive

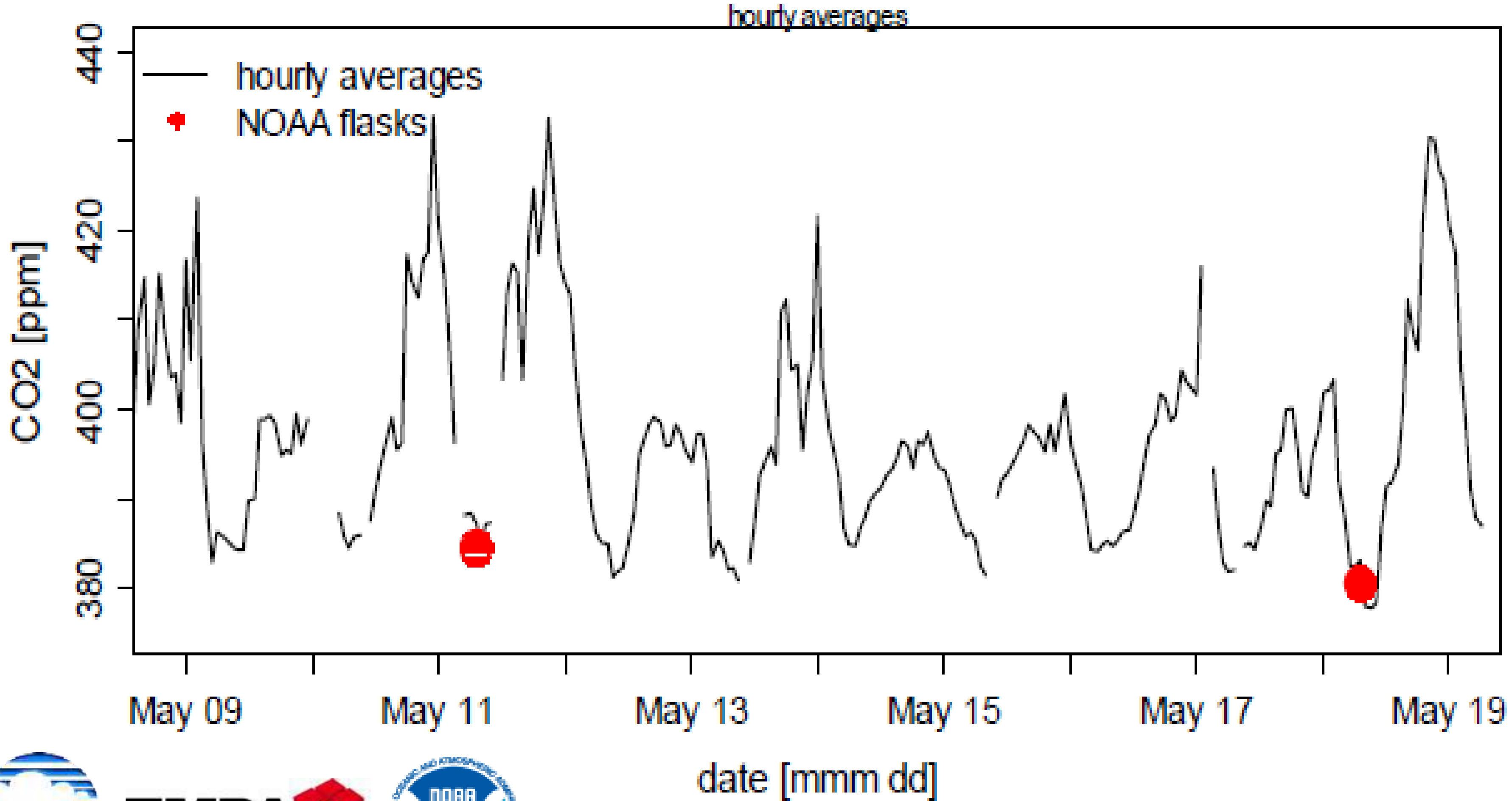
Bukit Koto Tabang *	Indonesia	BMKG,Empa	CH_4*, CO, CO_2*, O_3, MET
Bukit Koto Tabang *	Indonesia	NOAA/ESRL	$^{13}CO_2$, $C^{18}O_2$, CH_4*, CO, CO_2*, H_2, N_2O*, SF_6, VOCs, MET





Charcteristics of GHG data in pristine rain forest

Continuous (by EMPA) vs Flask (by NOAA)



BMKG

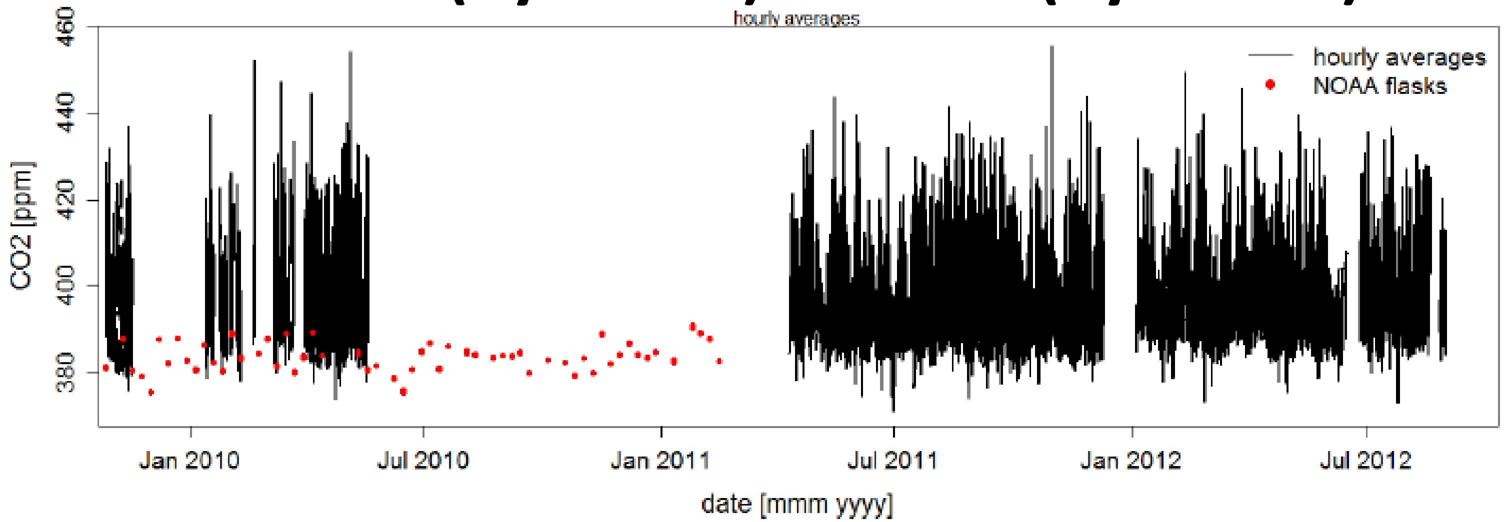
EMPA Materials Science & Technology

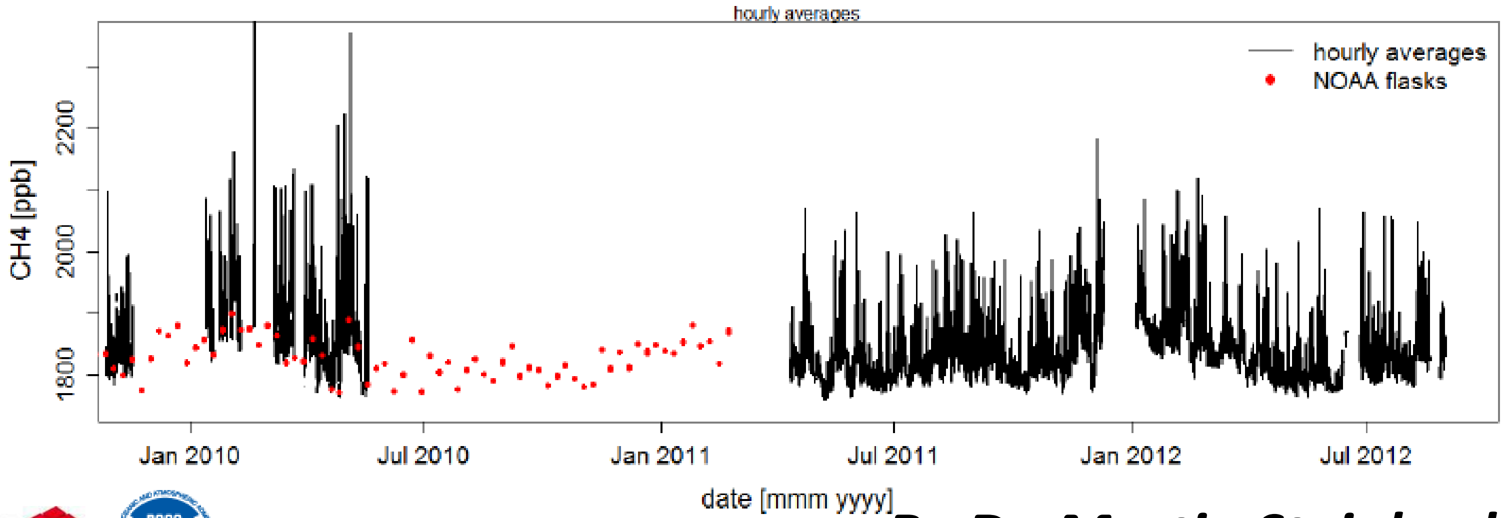
NOAA

By Dr. Martin Steinbacher

Charcteristics of GHG data in pristine rain forest

Continuous (by EMPA) vs Flask (by NOAA)





BMKG EMPA Materials Science & Technology NOAA

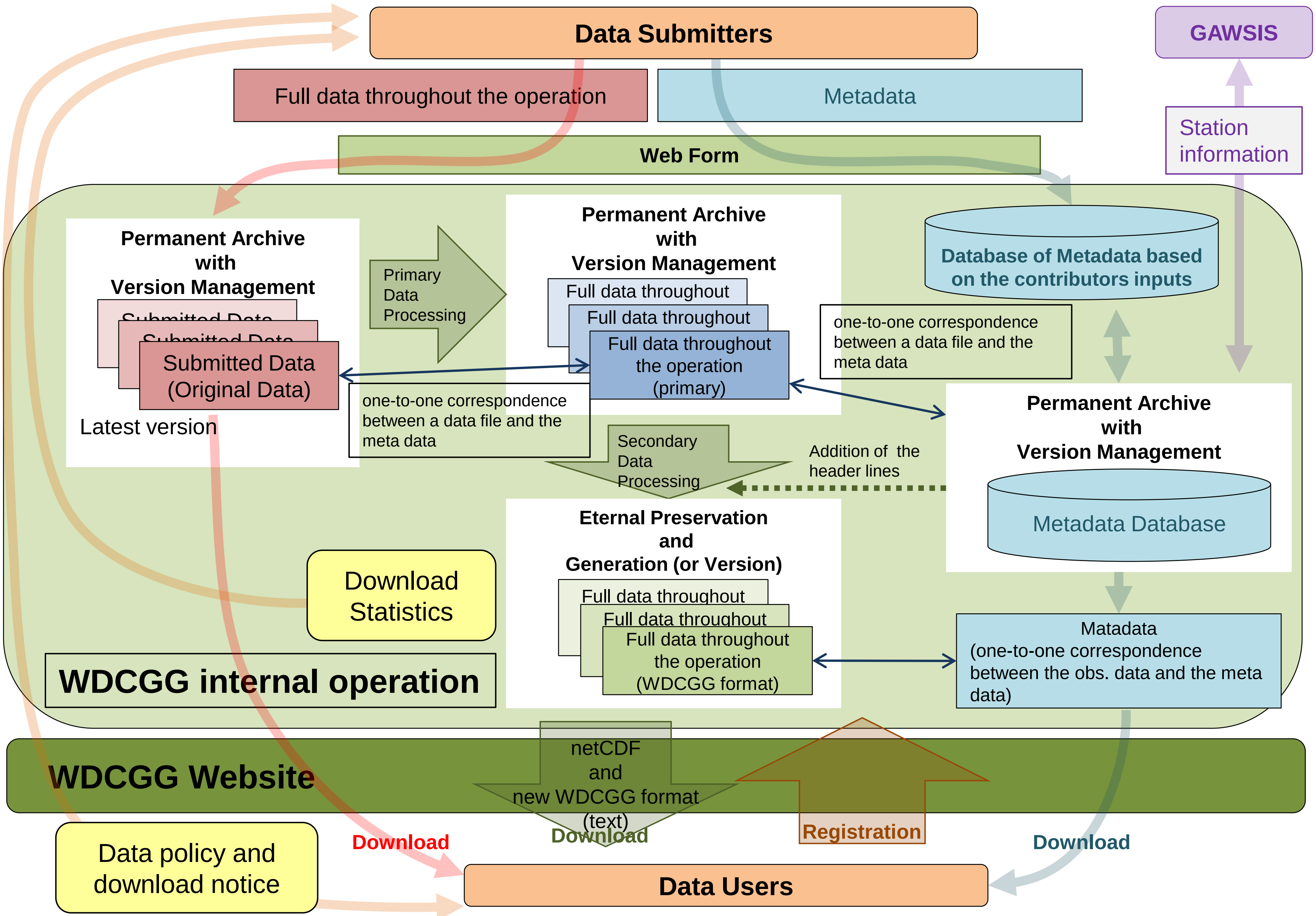
By Dr. Martin Steinbacher

Metadata is Crucial

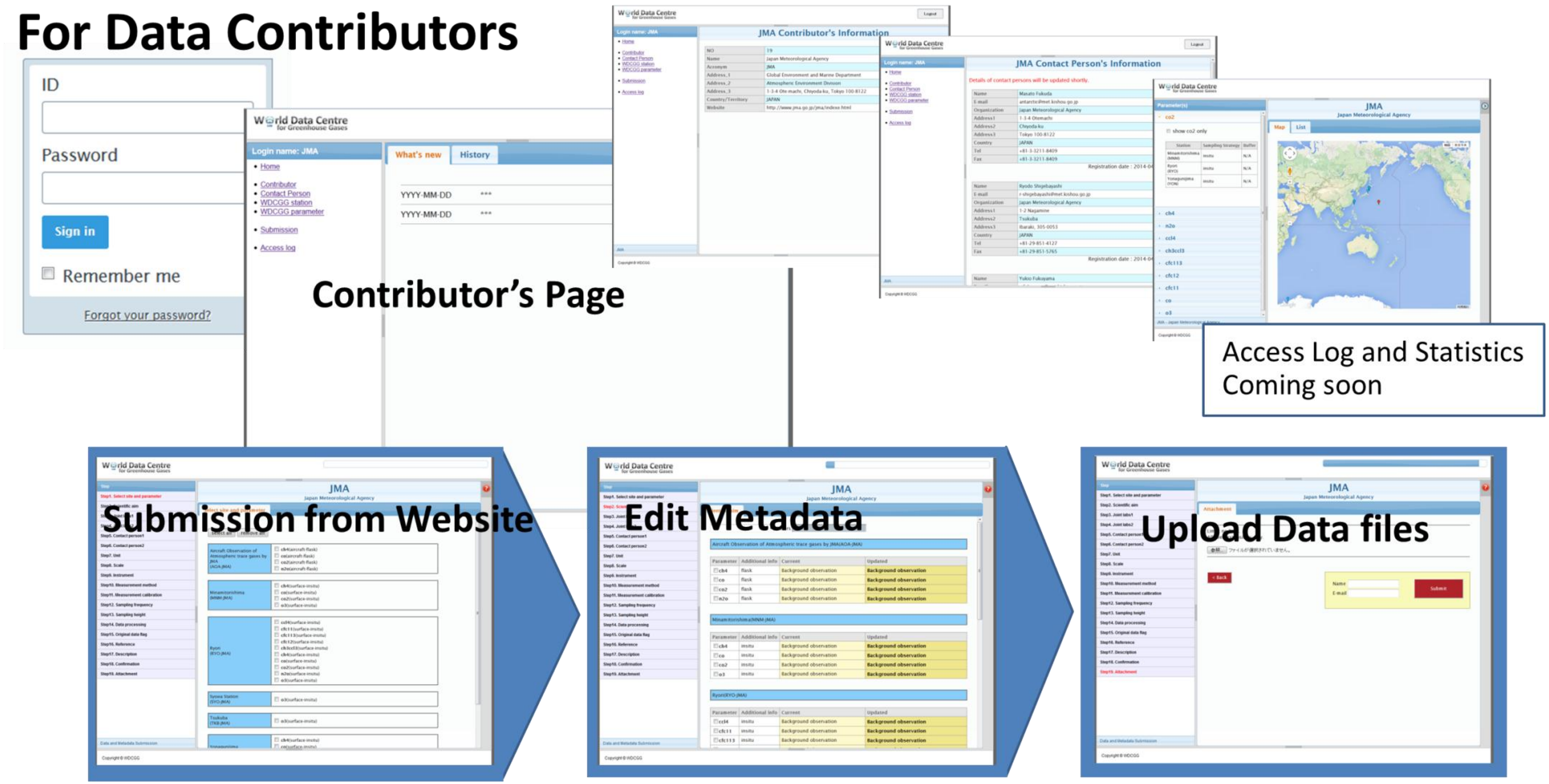
Q: How to ensure the propagation of necessary metadata information from data providers to data users?

A: Data submission interface for always one-to-one combined data and metadata

The concept of the new WDCGG



Metadata is Crucial



Interface design for data and metadata submission

Flagging

WORLD METEOROLOGICAL ORGANIZATION

Doc. 6.2
(2013-12-26)

MEETING OF THE
Expert Team on WORLD DATA CENTRES
(JMA, Tokyo, Japan, 21-23 January 2014)

Harmonisation of data flagging?

(submitted by M. Schultz and H. Koide)

"The purpose of data flagging is to obtain, at the end of the processing of the measurements, a time series of mole fractions that represents ambient conditions and to clearly identify artefacts as such. No entries should ever be removed from the original (raw) data set. Samples designated as not representing atmospheric composition should be identified […]. The periods of automatic and/or manual calibration or maintenance as well as instrument problems should be clearly flagged. Instrumental problems are sometimes not obvious and identifying them in the time series may require significant experience." (from GAW report 209 – ozone measurement guidelines)

1. Motivation

Currently, there is no standardized, GAW-wide data flagging scheme, and the amount and quality of data quality information varies widely between data centers and even within data

644	V	Low instrument precision and/or calibration issues
641	I	Aerosol filters installed incorrectly
640	V	Instrument internal relative humidity above 40%
635	I	Internal temperatures too far off target value, considered

Group 5: Chemical problem

599	I	Unspecified contamination or local influence
593	I	Industrial contamination, considered invalid
591	I	Agricultural contamination, considered invalid
578	I	Large sea salt contribution (ratio between marine and excess sulphate is larger than 2.0). Used for old data only. For newer data use 451/450.
568	I	Dust contamination, considered invalid
567	I	Insect contamination, considered invalid
566	I	Bird droppings, considered invalid
565	I	Pollen and/or leaf contamination, considered invalid
559	V	Unspecified contamination or local influence, but considered valid
558	V	Dust contamination, but considered valid
557	V	Insect contamination, but considered valid
556	V	Bird droppings, but considered valid
555	V	Pollen and/or leaf contamination, but considered valid
549	I	Impure chemicals

Flagging

Two Categories in Flag Sets

- Internal Flagging (Providers)
- Universal (External) Flagging (Users)

Color coding	Value	Meaning
(white)	0	no quality control applied
(green)	1	acceptable data
(yellow)	2	questionable data
(red)	3	erroneous data
(gray)	9	missing values

Flagging

- **Internal Flagging (Providers)**

Conversion Scheme

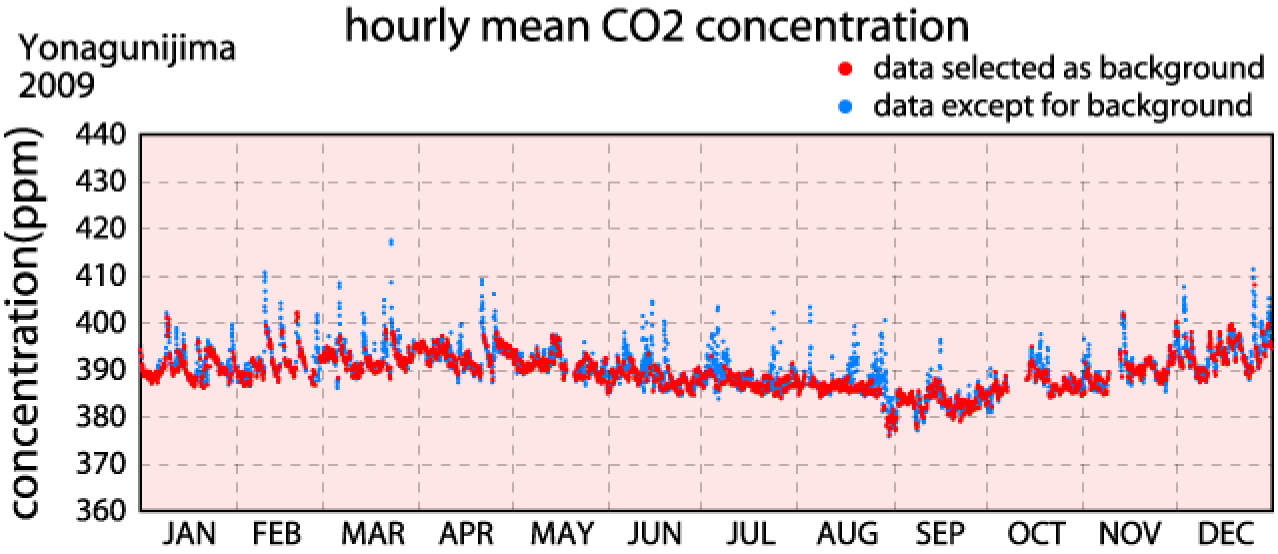
- **Universal Flagging (Users)**

Other (separated) Information

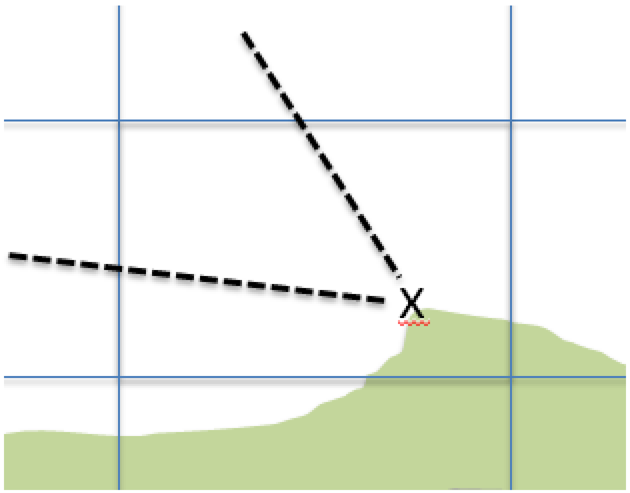
- **Change history**
- **Status (Raw/Processed Data)**

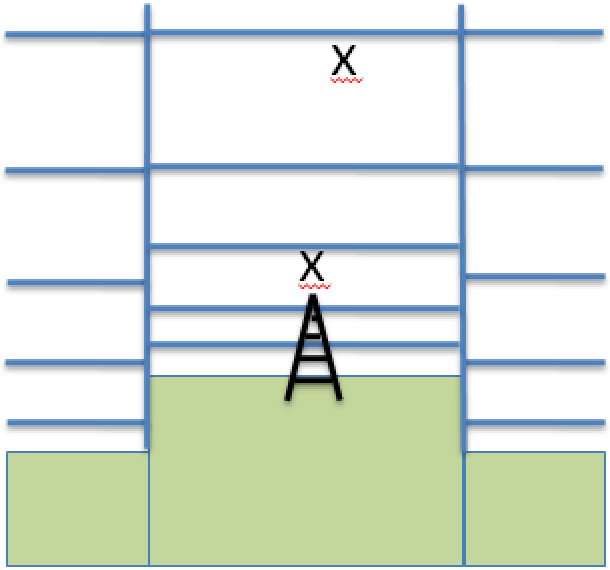
JMA's data screening process

1. Hourly data contain half or more of the number of data that can be measured in the hour.
2. Standard deviation of hourly data doesn't exceed defined thresholds.
3. Differences from both of the adjacent hourly data don't exceed defined thresholds.



Different Criteria for Screening





Figures by Sander Houweling, Netherland

On a Cape: by Wind Direction

On a Summit: by Diurnal Cycle

Quality Assurance in GAW

Q : How you can practically implement the GAW reports presenting recommendations on Quality Assurance into hundreds of GAW stations worldwide?!

A: Common Single Software for Workflow (from inlet to users)

Standardization Approach with an Interactive Software

Dafit: Data acquisition files integration tool

by Ludwig Ries, Germany (*l.ries@web.de*)

For standardized **Workflow on data** preparation, flagging and validation

on Windows 7 & 8.1, CSV output, conformable to DBMS

Internet download now in preparation on gawstat.de

check the detail with his paper in GGMT-2013, Beijing

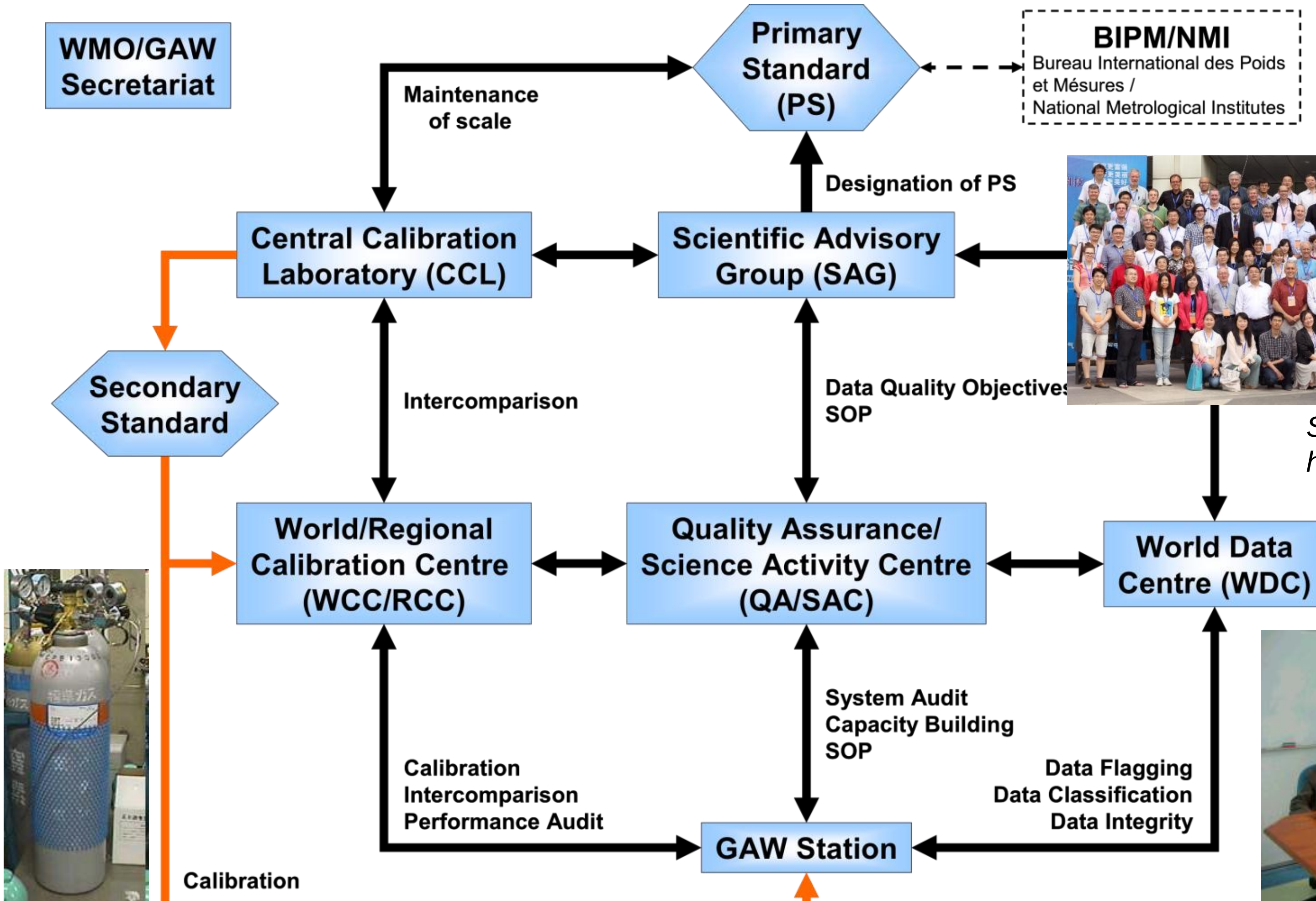
http://ggmt-2013.cma.gov.cn/dct/page/70029

Thank you for your attention!

경청해 주셔서 감사합니다.

Framework of the GAW Quality System

- **GAW Quality System is supported by CCLs, QA/SACs and WCCs, as well as SAGs in ensuring data quality.**



CO_2 Expert Meeting > GGMT

Source: GGMT-2013 website
http://ggmt-2013.cma.gov.cn

Source: GAW Strategic Plan: 2008 – 2015