Task-Team on Climate Data Models (TT-CDM)

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

What are the specifications for a Climate Data Model ?

WMO OMM

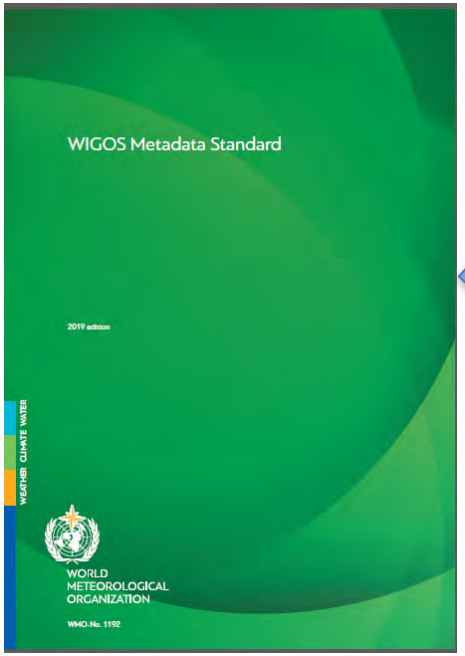
World Meteorological Organization
Organisation météorologique mondiale

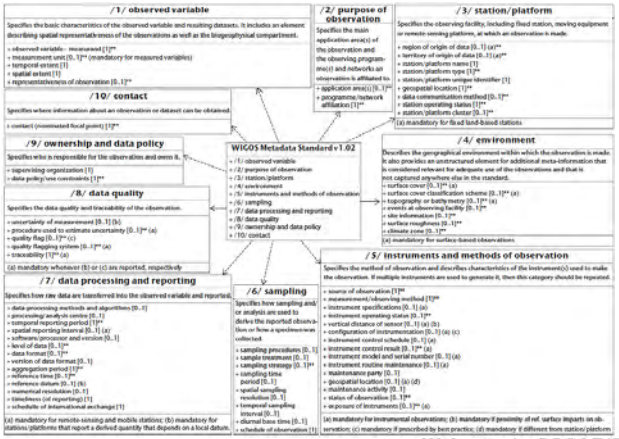
Expert-Team on Data Requirements for Climate Services
ET-DRC- 16 March 2022 - Denis Suber

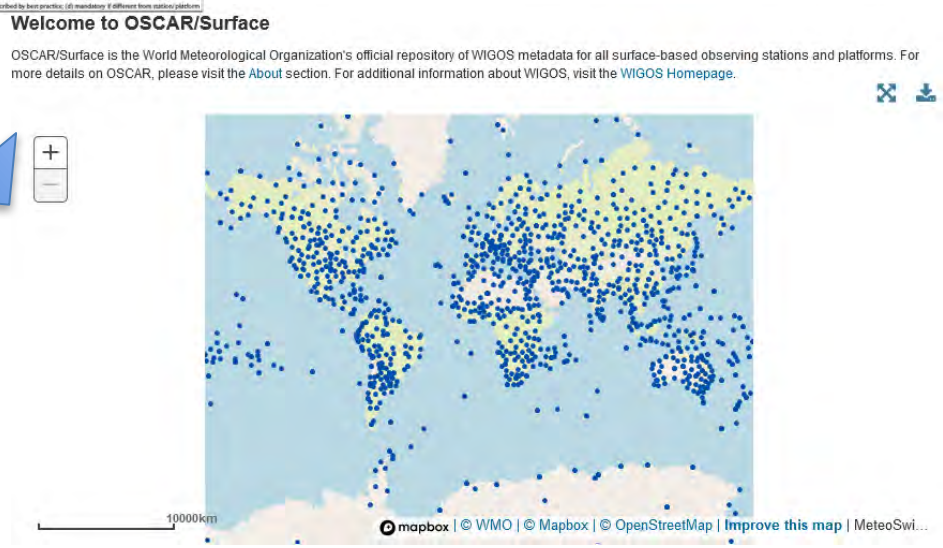
What is a Climate Data Model ?

Is a representation of interactions between information in order to meet the requirements of the climate data community (WMO) and to facilitate the work of users.

Generally, a Climate Data Model is the foundation of a Climate Data Management System (**CDMS**)





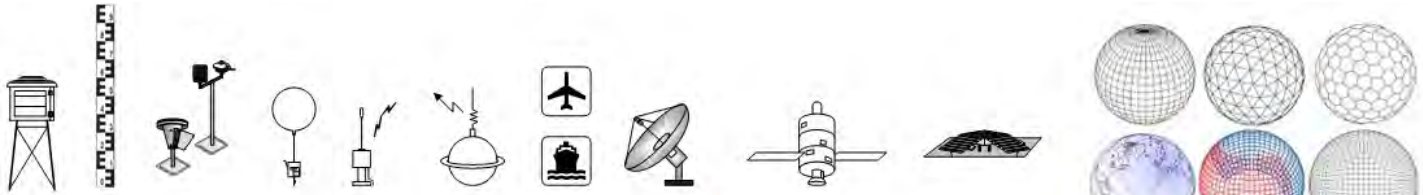


Which Data to manage?

Essential
Climate Variable
Atmosphere, Land, Ocean

Station Data

Spatial Data



Data Collection Platform

Data Processes Platform
ASCII TAC BUFR GRIB NETCDF etc.

Climate Data Management System (CDMS)
Acquisition, Management, Quality, Generation, Product, Services, etc.

Metadata
WIGOS Compliant

Which Data to manage?

Data

- Station Data & Metadata
- Spatial Data
- Images
- ...
- National & WMO Formats (TAC, BUFR,...)
- Other CDMS
- Emerging Format
- 3rd part data

Acquisition Process

- Ingest
  Automatic + key entry
- Climate Data Analysis
  Decode Data, Generate Statistics
- Transform
- Monitoring Data, Task, System, Observation Network
- Extract



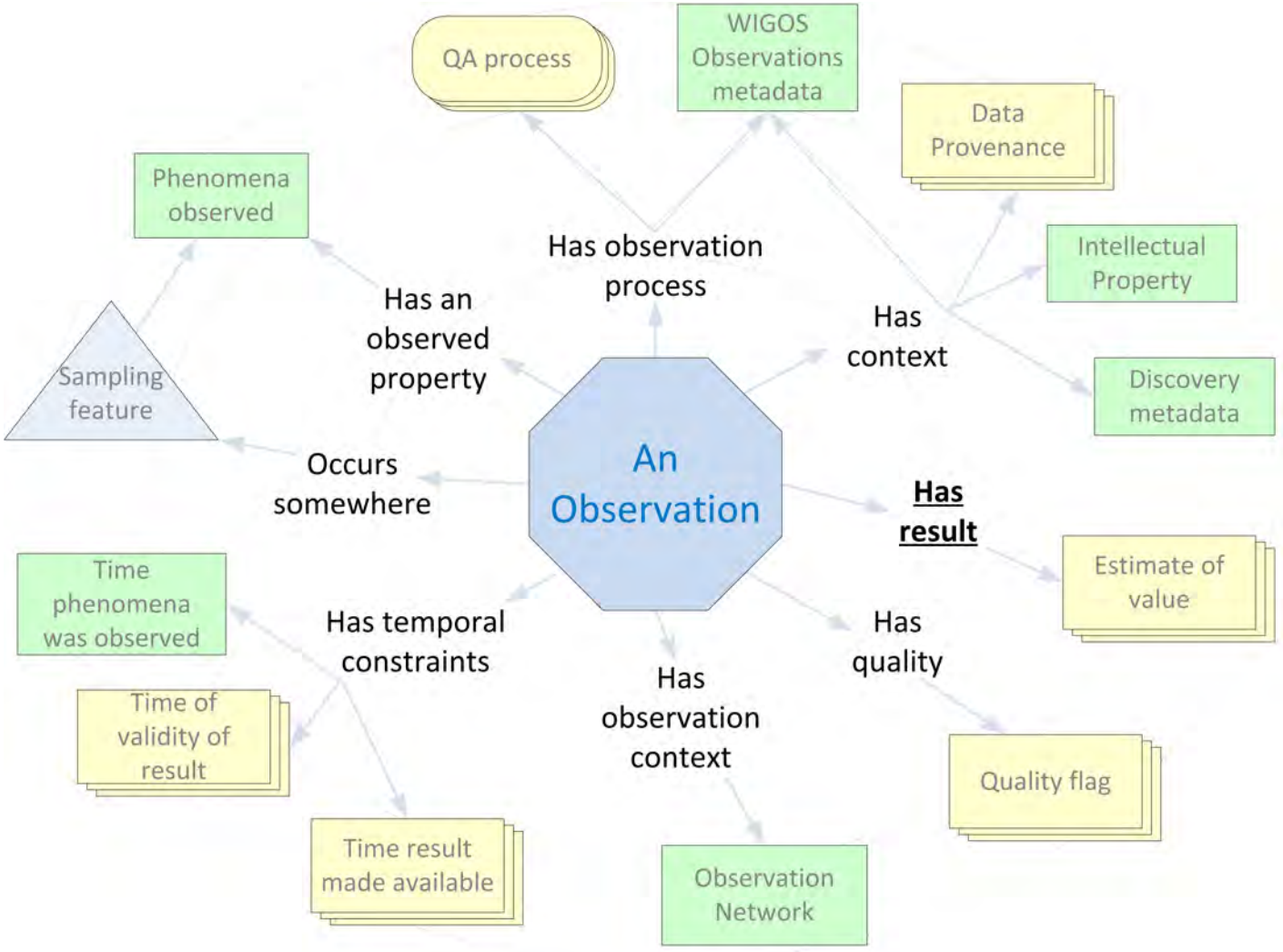
Climate Data Collection

- Observation Data & Metadata
- Gridded Data
- Vector Data
- Images
- Data Quality
- Data Provenance
- Vocabularies

What are the characteristics of a Climate Data Base

- ✓ **Traceability of any data, metadata and process is the main characteristic of a Climate Data Base** (GCOS Monitoring Principles, WMO-No. 1238, WMO-No.1131)
- ✓ **Data, Observation Metadata** and **Processes** are always subject to modification, whatever the date is (current date or old date)
- ✓ **Shall be compliant with National Practices** (Data Policy, Acquisition, Management, Quality, Control, Generation, Product, ...)
- ✓ **Quality** of the Data, of the Metadata and on the overall DataBase (especially Data Completeness) is essential
- ✓ Shall be compliant with **WMO requirements** in term of Metadata, Data Rescue, Data Acquisition, Data Management, Data Product, Data Services
- ✓ Shall help the climatologist and any other user in its every day work, whether it is in a developed or developing country

Data and Metadata



CDMS: An Observation
Revision 0.4 30 Nov 2016

Bruce Bannerman and Denis Stuber

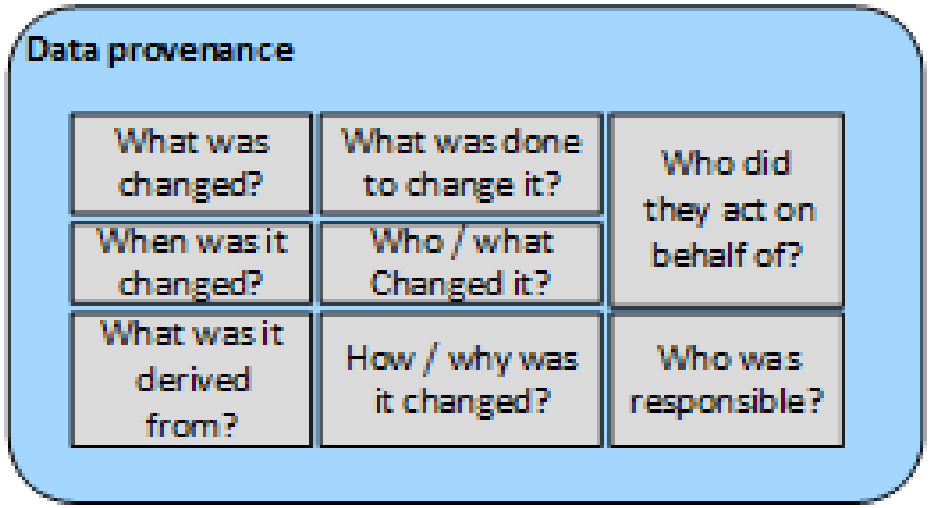
Maintaining the integrity of the historical climate data record

Tracing the life cycle of climate data

- ✓ Maintaining knowledge of the input data into the climatological system
- ✓ Knowledge of the quality control processes carried out on the Data and their results (including the algorithm used)
- ✓ The traceability of modifications made to the Data and their justification
- ✓ The traceability of modifications made to the Metadata and their justification

e.g.

- ✓ All locations of a station during its life
- ✓ All instruments used by the station at any period of its life
- ✓ All different values that a variable may have taken following modifications



Quality

Tracing the life cycle of climate data control

Control Data Types
Format
Constraint
Consistency
Statistical
Completeness
Spatial
Homogenization
Models & Remote System
Consistency after modification
Heuristic / Artificial Intelligence

Observation Quality Assessment
WMO Siting Classification
WMO Measurement Quality Classification

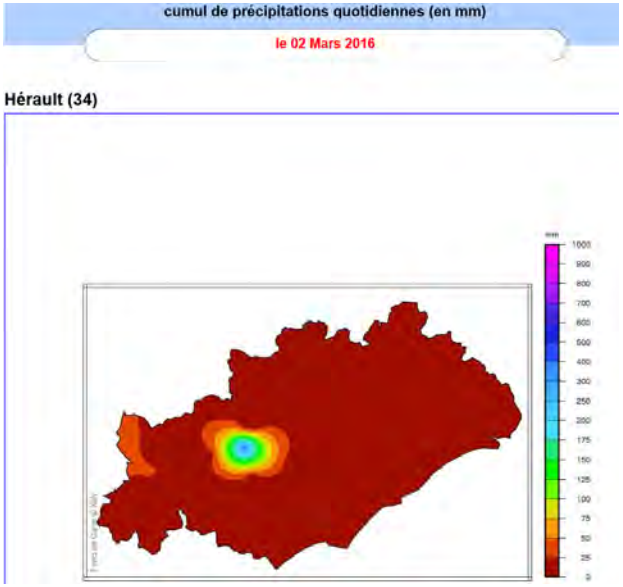
Multi Layer Quality Flag
Origin of the data (Acquisition way)
Result of the control
Data definition (Original, Estimated, Computed, etc.)

Dataset Quality
WMO Stewardship Maturity Matrix (Dataset Quality)

Quality Assurance Metrics
Statistics on the Data Controls Process (false positive, false negative, etc.)
Dashboard by station, by parameter, by control, etc.

Day to Day work in a Climatological Department

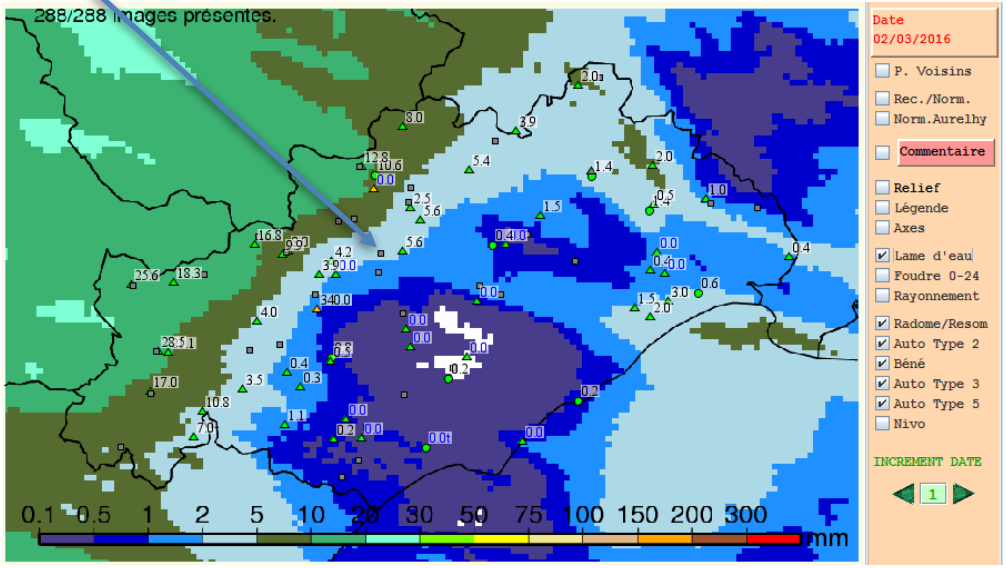
num	Poste	nom	20	19	18	17
10 372 001	2	81 SOULAINES				
11 016 003	2	31 ARQUETTES-EN-VAL				
11 028 001	2	31 BELCAIRE				
11 069 001	0	10 CARCASSONNE				
11 076 001	2	31 CASTELNAUDARY				
11 081 003	2	31 CAUNES-MINERVOIS				
11 124 003	2	31 DURBAN-CORBIERES				
11 168 001	1	12 GRANES		1		
11 181 002	2	30 LABECEDE LAURAGAIS_SAPC	3			
11 202 001	0	10 LEUCATE				
11 203 004	1	12 LEZIGNAN-CORBIERES				
11 221 004	1	12 LES MARTYS	11	5		
11 262 005	1	12 NARBONNE				
11 333 001	2	31 ST BENOIT				
11 412 001	2	31 VILLARDEBELLE		1		
12 005 001	1	12 ALPUECH				
12 014 001	2	30 AURELLE-VERLAC SAPC				



HNT	Month											
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1992	744	696	744	720	744	720	744	744	720	744	720	744
1993	744	672	744	720	744	720	744	744	720	744	720	744
1994	8	2			354	720	744	744	720	744	720	744
1995	739	672	744	57	712	285	57	49	0	2	0	3
1996	3	16	0	0	1	0	0	0	0	0	0	9
1997	21	0	0	0	0	0	0	0	0	0	0	11
1998	2	1	0	0	0	0	0	0	0	0	4	14
1999	8	16	0	2	0	0	0	0	0	0	7	9
2000	0	4	1	0	0	0	0	0	0	0	0	0
2001	7	6	6	0	0	0	0	0	0	0	1	11
2002	20	0	0	0	1	0	0	0	0	0	0	0
2003	12	7	0	0	0	0	0	0	0	0	0	1
2004	13	5	3	0	0	0	0	0	0	0	1	4
2005	7	16	12	0	0	0	0	0	0	0	6	10
2006	2	7	12	0	1	0	0	0	0	0	0	0
2007	0	0	3	0	0	0	0	0	0	0	0	0
2008	8	29	31	30	31	30	31	31	30	31	30	31
2009	744	30	31	30	31	24	31	31	30	31	30	744
2010	1	0	0	0	0				2			
2011	738	668	740	720	744	712	723	736	720	741	715	744
2012	175											

Total Percentage 20.5%

no value | less than 50% | between 50 and 95% | more than 95% | 100%



- ✓ Completeness
- ✓ Data Control & Human action
- ✓ Data Generation, Data Product, etc.

Quality : Flagging System

FHNTAGS	Quality Flag															
Year	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43
1992	8784	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1993	8760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1994	5500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1995	3320	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1996	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1997	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1998	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1999	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2000	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2002	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2003	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2004	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2005	51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2006	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2007	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2008	343	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2009	361	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2010	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2011	8703	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2012	175	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- ✓ Validation Phase
- ✓ Action Phase

10	original basic data not already checked
11	calculated data elaborated with only original basic data among which at least one was not checked
12	basic data that has been corrected and that is not already checked
13	calculated data elaborated with data among which at least one was corrected and at least one was not checked
20	original basic data declared doubtful by the controls
21	calculated data elaborated with only original basic data among which at least one was declared doubtful by the controls
22	basic data that has been corrected and that has been declared doubtful by the controls
23	calculated data elaborated with data among which at least one was corrected and at least one was declared doubtful by the controls
30	original basic data checked doubtful (by the automatic controls) but validated (by human expertize)
31	calculated data elaborated with only original basic data among which at least one was declared doubtful but all of them have been validated by human expertize
32	basic data that has been corrected and declared doubtful by the controls and finally validated by human expertize
33	calculated data elaborated with data among which at least one was corrected and at least one was declared doubtful but all of them have been validated by human expertise
40	original basic data validated by the controls
41	calculated data elaborated with only original basic data that have been validated by the controls
42	basic data that has been corrected and then validated by the controls
43	calculated data elaborated with data among which at least one was corrected but all of them have been validated by the controls

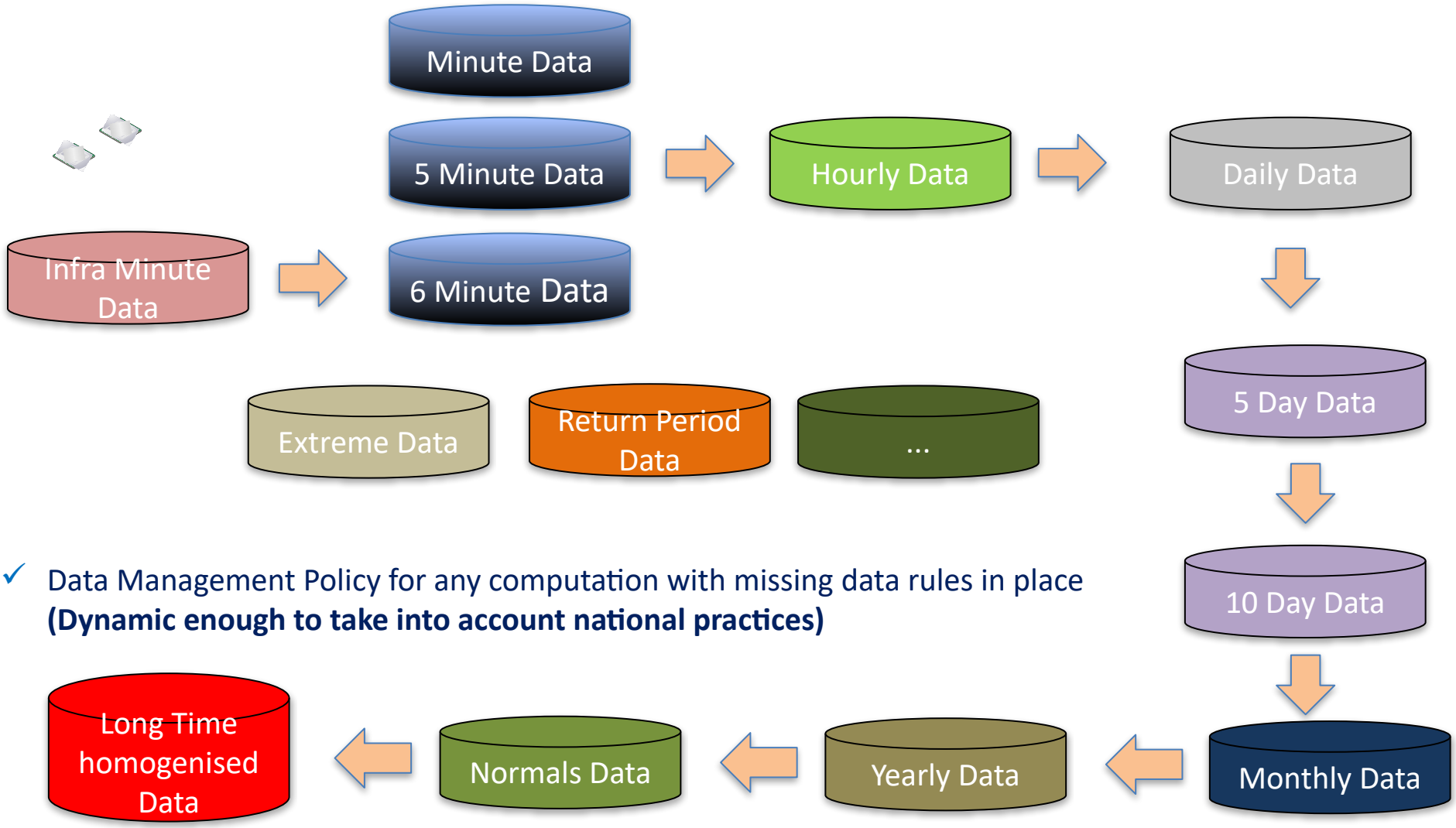
Parameter	Beginning period with data	Ending period with data	Number of values	% for 8/24	% for 24/24	Minimum value	Maximum value
FHNTAGS	1992-01-01	2012-01-06	36197	61.9%	20.6%		

Quality : Data Control Indicators

Anomaly	# Anomalies	# Min for 1 station	# Max for 1 station	# impacted stations	# good detections	# false positive	...
Control 1	4167	1	643	55	3523	644	
Control 2	1298	3	86	61	1280	18	
Control 3	530	5	6	435	136	348	
Control 4	433	2	23	221	510	0	
Control 5	3521	2	5	309	203	230	
Control 6	2658	3	16	95	67	64	
...	...						

Monthly Report

Coherency

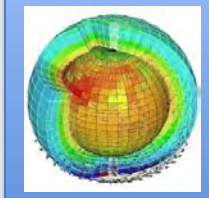


- ✓ Data Management Policy for any computation with missing data rules in place **(Dynamic enough to take into account national practices)**

- ✓ Each data with its origin, provenance, quality code, etc.
- ✓ When a modification is made, all data with a greater temporal reference is supposed to be re-computed

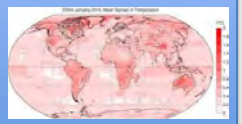
A crucial link between observed data and modeling

Modeling

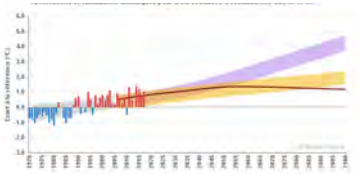


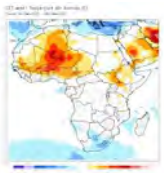
Global

Regional



Reanalysis

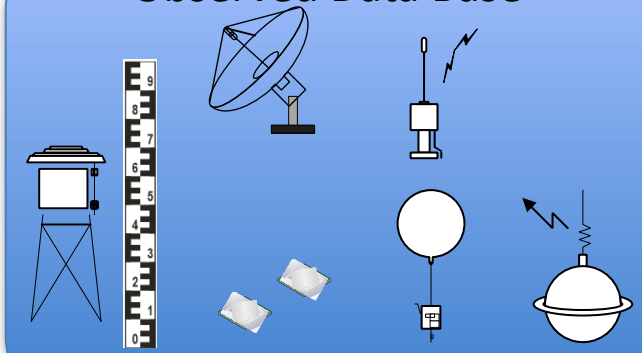




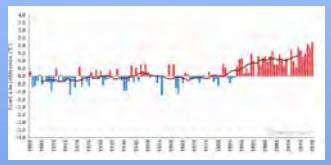
Data Control

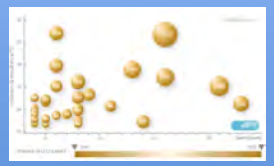
Model Calibration

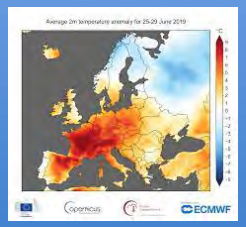
Observed Data Base

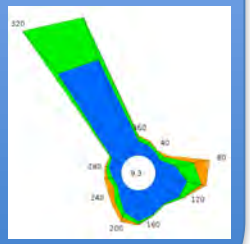


Products and Services







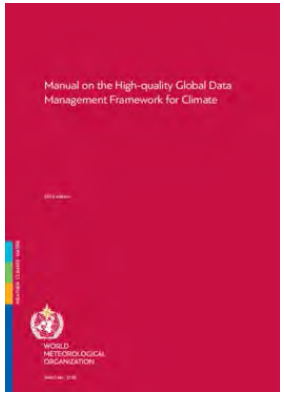


Users

Requirements Summary

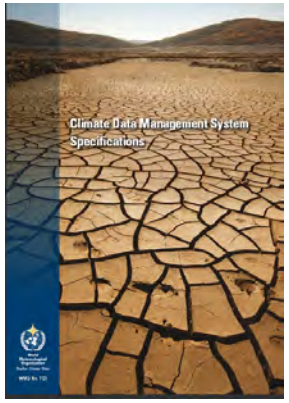
The definition of a **Climate Data Standard**:

- ✓ Is fully Compliant with the **WIGOS Metadata Standard**
- ✓ Manages Station Data & Metadata, **Spatial Data** and **Images**
- ✓ **Traces** all Data , Metadata and Processes
- ✓ Includes the **WMO Siting Classification**
- ✓ Includes the WMO **Measurement Quality Classification**
- ✓ Includes the WMO **Stewardship Maturity Matrix** (Dataset Quality)
- ✓ Uses the **WMO Codes Registry** and other WMO standards (e.g. METCE)
- ✓ Proposes a **Data Quality Policy** with quality information on each data
- ✓ Allows the use of **Modeling Data** for Data Control and Data Products and Services
- ✓ Facilitates the user work: Acquisition, Control & Modification, Computation, Product, Service

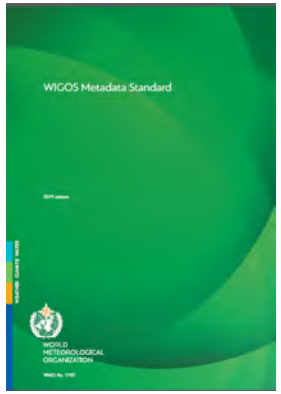


Manual on the High-quality Global Data Management Framework for Climate, WMO-No.1238

Guide to Climatological Practices, WMO-No.100



Climate Data Management Specifications , WMO-No.1131



WIGOS Metadata Standard, WMO-No.1192

Task-Team on Climate Data Models (TT-CDM)

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

Thank you
Merci

Denis Stuber – SERCOM ET-DRC

WMO OMM

World Meteorological Organization
Organisation météorologique mondiale