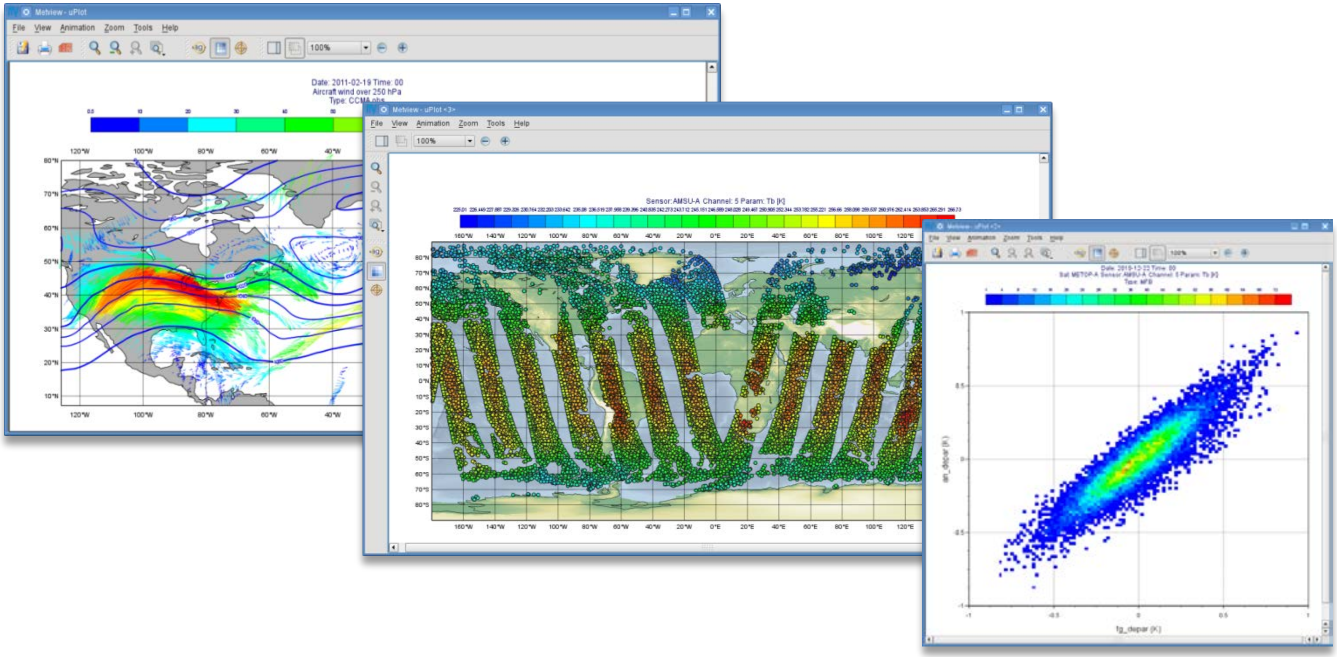
Using ODB at ECMWF


Metview - uPlot
File View Animation Zoom Tools Help
100%
Date: 2011-02-19 Time: 00
Aircraft wind over 250 hPa
Type: CCMA obs
Metview - uPlot <3>
Sensor: AMSU-A Channel: 5 Param: Tb [K]
Metview - uPlot <3>
Sat: METOP-A Sensor: AMSU-A Channel: 5 Param: Tb [K]
Type: MFB
an_depar (K)
fg_depar (K)

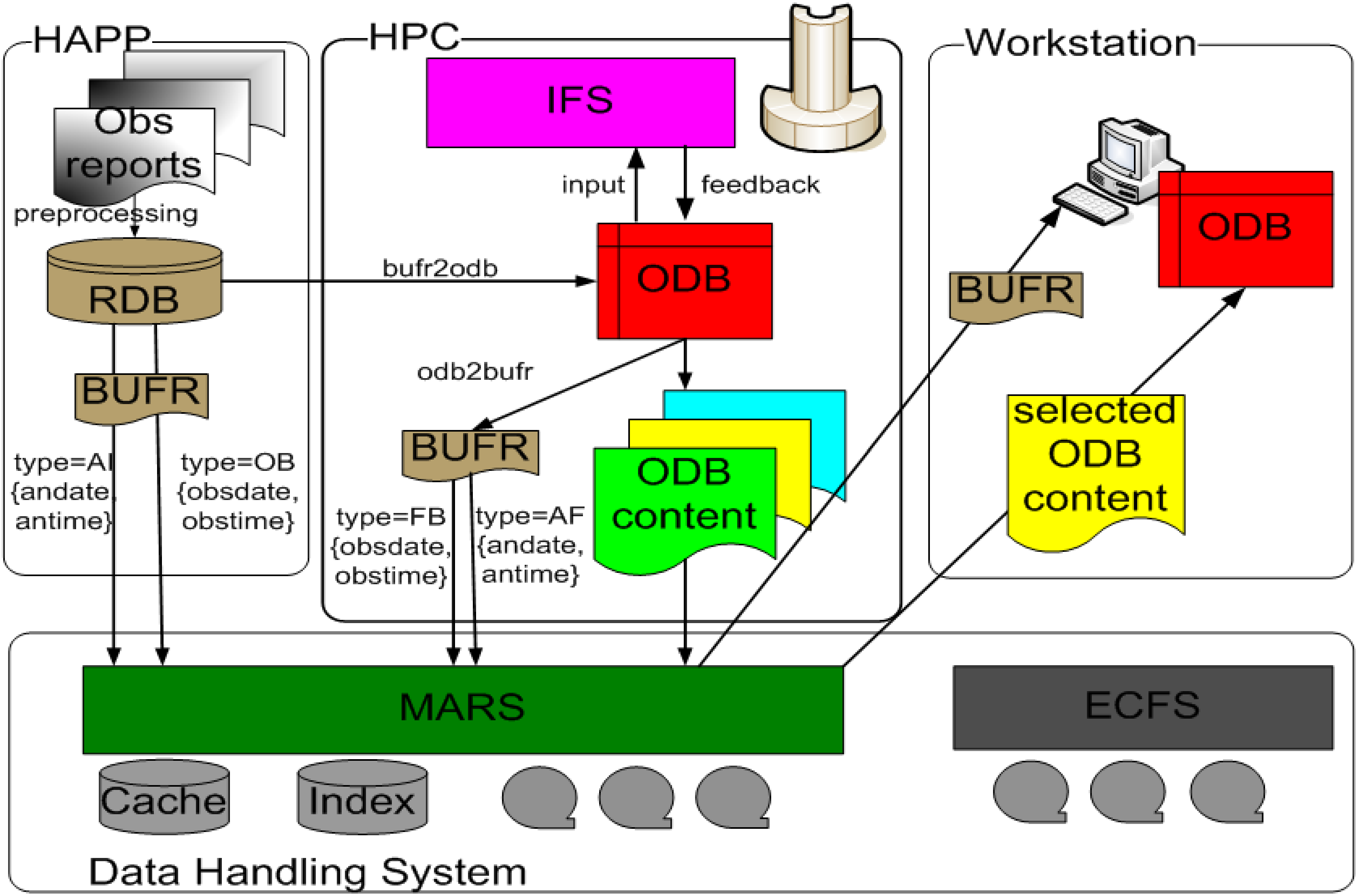
Piotr Kuchta
Sándor Kertész

Development Section
ECMWF

History of ODB in a nutshell

- **1998 – 2008, Sami Saarinen**
  - Database of observations for 4D-Var IFS, parallel Fortran API, MPI
  - Relational databases concepts: SQL for querying/filtering, DDL for schema definition
  - Implementation: C and Fortran, SQL compiled to C
- **2008 – now, Anne Fouilloux, Baudouin Raoult, Piotr Kuchta, Tomas Kral**
  - ODB feedback archiving in MARS, ODB Server
  - New data format suitable for archiving
  - OO, C++ library for encoding/decoding & SQL filtering (ODB API)
  - new SQL engine: fast interpreter, no intermediate C code, can be extended to work with different data formats
  - API: C++, C, Fortran, Python, command line tools
  - Used by: MARS, ODB Server, Metview, Magics++, Obstat

Observational data flow at ECMWF



ODB Feedback archive

- **MARS (server & client) extended to handle ODB data**
  - **Flexible indexing and physical organization on disks/tapes allows for efficient retrieval and processing of long time series**
  - **Any observational data can be identified with a single number: Report Type classification http://data-portal.ecmwf.int/odbgov/ReportType/**
  - **SQL filtering / transformations / statistics computations on the fly in MARS client**
  - **Archive requests validated against data (both server and client)**
- **ODB feedback (TYPE=OFB) and Monitoring (active observations) (TYPE=MFB) archive operational since Nov 2011 http://www.ecmwf.int/services/archive/d/catalog**
  - **Operational data > Deterministic forecasts >… > Observations > MONDB feedback | ODB feedback**

ODB Archive format

- New file format for tabular, numeric data
  - Simple, well defined, machine independent, fast to encode and decode
  - Self-described (meta data in header), handles data types of ODB
  - Extendable: new codecs (compression alg.) can be added
  - Small memory footprint needed to process large datasets (streaming)
- An ODB archive file represents a flat table
  - Data layout: report oriented (one report stored in one row)
  - Every column has: name, type (REAL, INTEGER, STRING, …), value of missing data indicator (NULL), codec (invisible to the user)

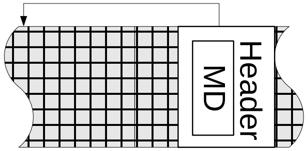
reptype	class	andate	antime	lat	lon	statid	obsvalue	blacklist	status
14	‘od’	20090730	0	0.364777	0.986071	223	235.009995	0	01100
14	‘od’	20090730	0	0.367758	0.972175	223	235.949997	0	01100
14	‘od’	20090730	0	0.370347	0.959772	223	236.880005	0	01100

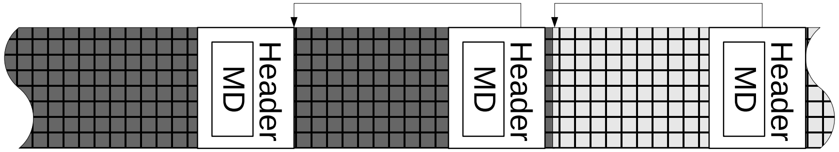
ODB Archive format features (2)

- **New data can be appended to existing files**
  - **Efficient usage of storage medium (tapes)**
  - **Can retrieve long time series in one request**

reptype	class	andate	antime	lat	lon	statid	obsvalue
14	‘od’	20090730	0	0.364777	0.986071	223	235.009995
14	‘od’	20090730	0	0.367758	0.972175	223	235.949997
14	‘od’	20090730	0	0.370347	0.959772	223	236.880005

reptype	class	andate	antime	lat	lon	satid	sensor	obsvalue
14	‘od’	20090731	0	0.364777	0.986071	222	3	235.009995
14	‘od’	20090731	0	0.367758	0.972175	222	3	235.949997
14	‘od’	20090731	0	0.370347	0.959772	222	3	236.880005





MARS language extensions

- **SQL incorporated into MARS syntax**

SQL engine used for filtering on the client

```
RETRIEVE,
TYPE = OFB,
OBSGROUP = hirs,
REPORTTYPE = 13,
DATE = 20110728,
TIME = 0,
SENSOR = 0,
FILTER =  “select *
           where distance(lat, lon, 0, 0) < km(50)
                 and value > 375.2“,
TARGET = "ECMA.hirs.13.odb"
```

ODB API SQL examples

- **Which station identifiers are in the ODB file?**

```
$ odb sql 'select distinct statid' -i $odbfile
```

- **How many manual land SYNOP reports ?**

```
$ odb sql 'select count(*) where reportype=16002' -i $odbfile
```

- **Count the number of temperature records per station identifier, where the observation values are not missing. Sort by the count.**

```
$ odb sql 'select count(*), statid where varno=2 and obsvalue is not
NULL order by 1' -i $odbfile
```

C++ and Python API examples

- **C++ API based on concept of iterators, mimics STL containers**

```
string sql = "select lat, lon, obsvalue from \"in.odb\""
             " where obsvalue > 0.5 ";
odb::Select  sel(sql);
for (odb::Select::iterator it = sel.begin(); it != sel.end(); ++it)
     cout << it->data(0) << " " << it->data(1) << " “ << it->data(2) << endl;
```

```
odb::Reader  reader("file.odb");
for (odb::Reader::iterator it = reader.begin(); it != reader.end(); ++it)
     cout << it->data(0) << " " << it->data(1) << " “ << it->data(2) << endl;
```

- **Python**

```
import odb
for r in odb.open('conv.odb'):
   print r['lat','lon','obsvalue']
```

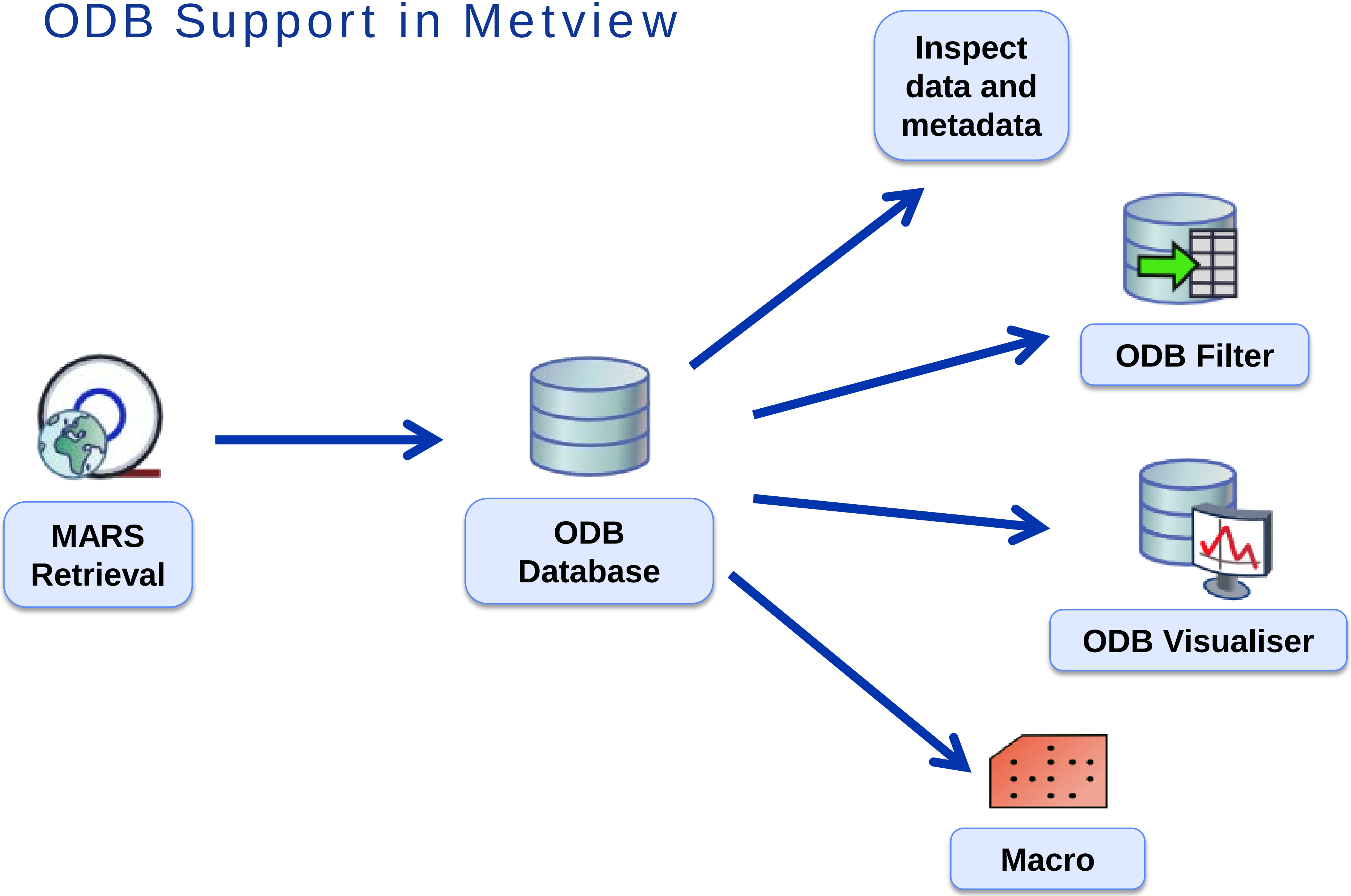
Tools for working with ODB data

- **Command line tools available via odb binary**
  - **Execute SQL SELECT and print results to standard output or save in a new ODB file**
  - **Examine meta data of a file (odb `header`)**
  - **Split data so e.g. data coming from one station, report type, satellite id, etc, is saved in a separate file**
  - **Compare two ODB files**
  - **Import data from a text file (CSV)**
  - **Merge two or more files row by row**
  - **Describe data on file in MARS language syntax (`oda2request`)**
- **Metview, Magics++, Obstat**

Plans for ODB development

- **Parallel (MPI/OpenMP) functionality will be redeveloped in C++**
- **Further development and improvements of SQL engine and other functionality as required by**
  - users,
  - COPE (Continous Observation Processing)
  - OOPS (Object Oriented Prediction System)
  - MARS
  - Metview
  - Magics++
  - OPS (Met Office)

ODB Support in Metview



MARS Retrieval

ODB Database

Inspect data and metadata

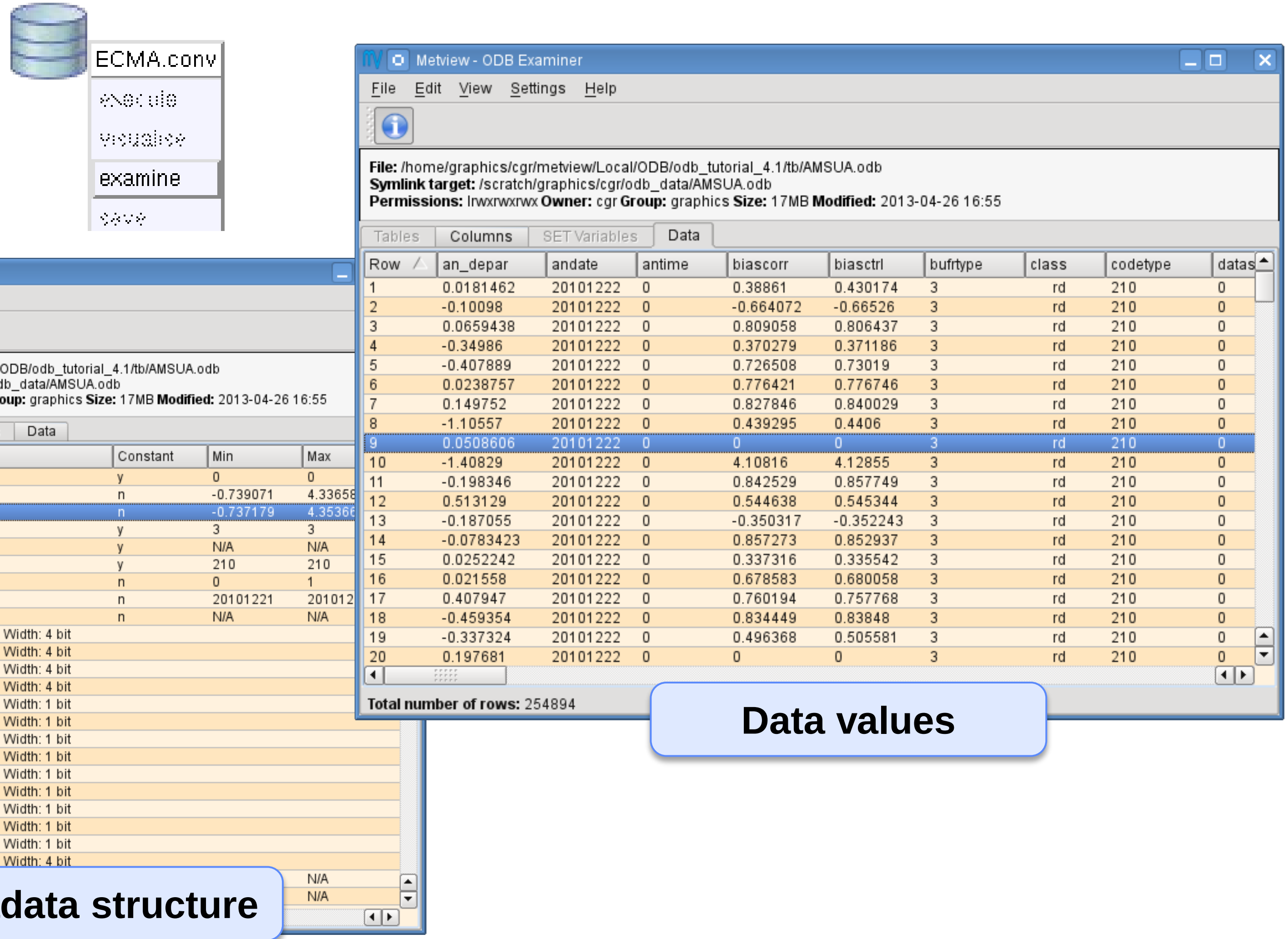
ODB Filter

ODB Visualiser

Macro

Examine ODB Metadata and Data

- **Right-click “Examine”**



File: /home/graphics/cgr/metview/Local/ODB/odb_tutorial_4.1/tb/AMSUA.odb
Symlink target: /scratch/graphics/cgr/odb_data/AMSUA.odb
Permissions: lrwxrwxrwx **Owner:** cgr **Group:** graphics **Size:** 17MB **Modified:** 2013-04-26 16:55

Name	Type	Constant	Min	Max
antime@desc	int	y	0	0
biascorr@body	float	n	-0.739071	4.33658
biasctrl@body	float	n	-0.737179	4.35366
bufrtype@hdr	int	y	3	3
class@desc	string	y	N/A	N/A
codetype@hdr	int	y	210	210
datastream@sat	int	n	0	1
date@hdr	int	n	20101221	201012
datum_anflag@body	bitfield	n	N/A	N/A
blacklist	Pos: 16 Width: 4 bit			
depar	Pos: 08 Width: 4 bit			
fg	Pos: 04 Width: 4 bit			
final	Pos: 00 Width: 4 bit			
uh2	Pos: 24 Width: 1 bit			
uhu	Pos: 22 Width: 1 bit			
ups	Pos: 20 Width: 1 bit			
urr	Pos: 26 Width: 1 bit			
usn	Pos: 27 Width: 1 bit			
usst	Pos: 28 Width: 1 bit			
ut2	Pos: 23 Width: 1 bit			
uv1	Pos: 25 Width: 1 bit			
uvt	Pos: 21 Width: 1 bit			
varqc	Pos: 12 Width: 4 bit			
datum_event1@b				N/A
datum_rdbflag@b				N/A

File: /home/graphics/cgr/metview/Local/ODB/odb_tutorial_4.1/tb/AMSUA.odb
Symlink target: /scratch/graphics/cgr/odb_data/AMSUA.odb
Permissions: lrwxrwxrwx **Owner:** cgr **Group:** graphics **Size:** 17MB **Modified:** 2013-04-26 16:55

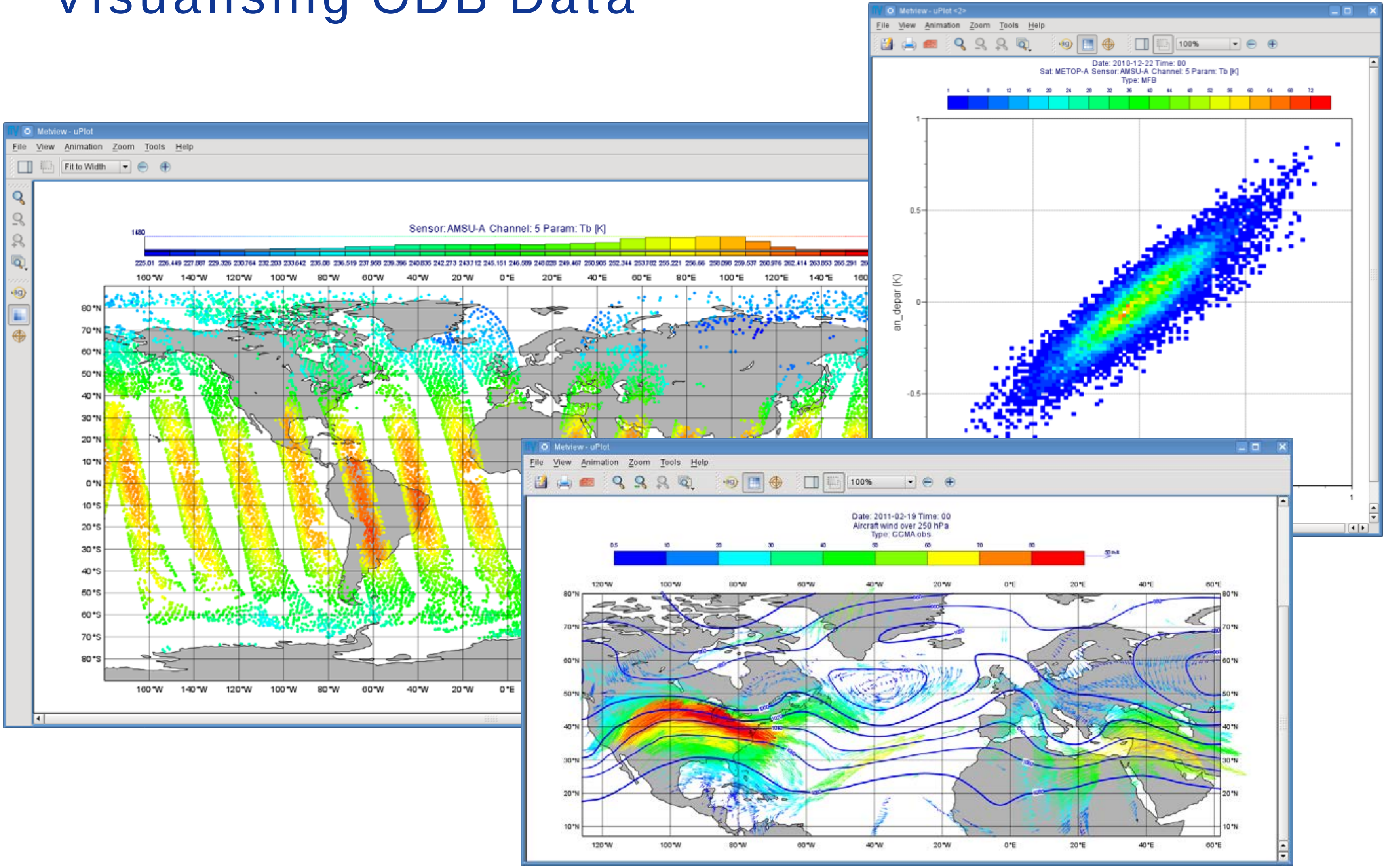
Row	an_depar	andate	antime	biascorr	biasctrl	bufrtype	class	codetype	datas
1	0.0181462	20101222	0	0.38861	0.430174	3	rd	210	0
2	-0.10098	20101222	0	-0.664072	-0.66526	3	rd	210	0
3	0.0659438	20101222	0	0.809058	0.806437	3	rd	210	0
4	-0.34986	20101222	0	0.370279	0.371186	3	rd	210	0
5	-0.407889	20101222	0	0.726508	0.73019	3	rd	210	0
6	0.0238757	20101222	0	0.776421	0.776746	3	rd	210	0
7	0.149752	20101222	0	0.827846	0.840029	3	rd	210	0
8	-1.10557	20101222	0	0.439295	0.4406	3	rd	210	0
9	0.0508606	20101222	0	0	0	3	rd	210	0
10	-1.40829	20101222	0	4.10816	4.12855	3	rd	210	0
11	-0.198346	20101222	0	0.842529	0.857749	3	rd	210	0
12	0.513129	20101222	0	0.544638	0.545344	3	rd	210	0
13	-0.187055	20101222	0	-0.350317	-0.352243	3	rd	210	0
14	-0.0783423	20101222	0	0.857273	0.852937	3	rd	210	0
15	0.0252242	20101222	0	0.337316	0.335542	3	rd	210	0
16	0.021558	20101222	0	0.678583	0.680058	3	rd	210	0
17	0.407947	20101222	0	0.760194	0.757768	3	rd	210	0
18	-0.459354	20101222	0	0.834449	0.83848	3	rd	210	0
19	-0.337324	20101222	0	0.496368	0.505581	3	rd	210	0
20	0.197681	20101222	0	0	0	3	rd	210	0

Total number of rows: 254894

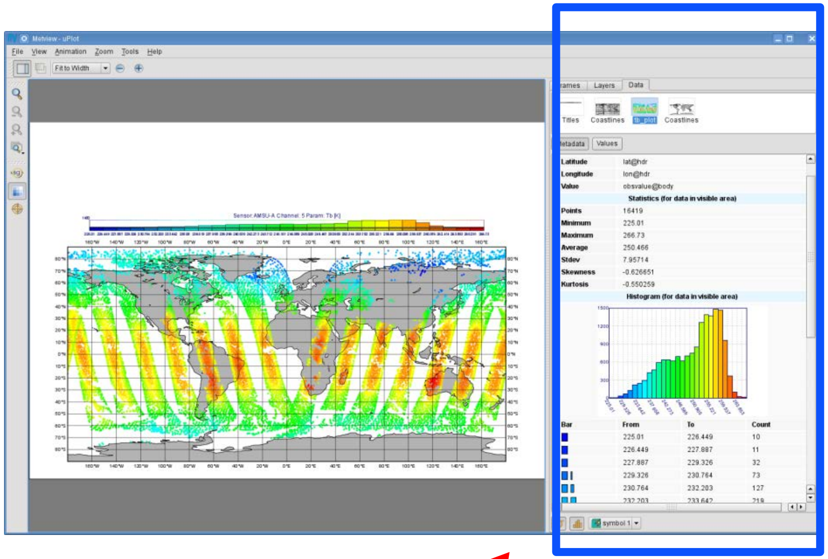
Data values

Metadata structure

Visualising ODB Data

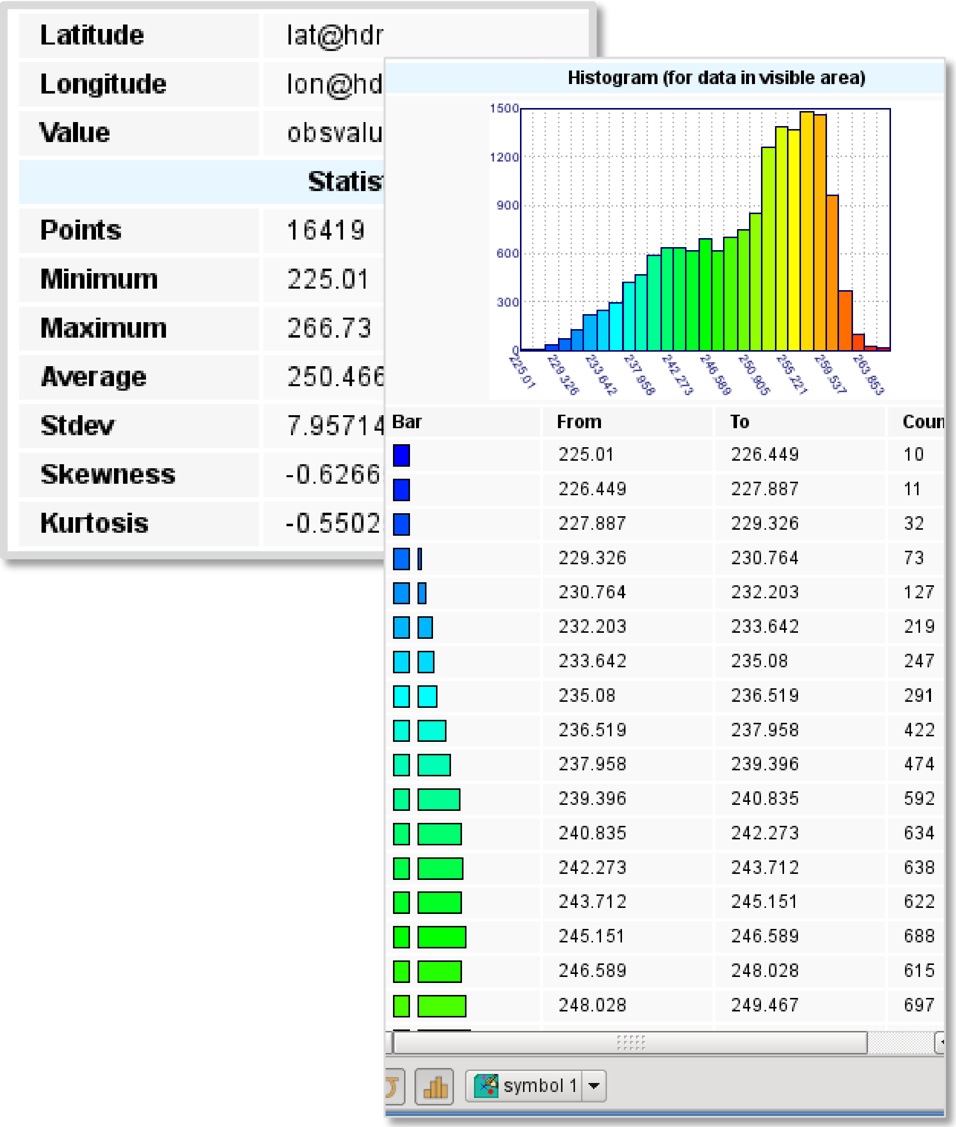

Metview - uPlot
File View Animation Zoom Tools Help
Fit to Width
Sensor: AMSU-A Channel: 5 Param: Tb [K]
160°W 140°W 120°W 100°W 80°W 60°W 40°W 20°W 0°E 20°E 40°E 60°E 80°E 100°E 120°E 140°E 160
80°N 70°N 60°N 50°N 40°N 30°N 20°N 10°N 0°N 10°S 20°S 30°S 40°S 50°S 60°S 70°S 80°S
Metview - uPlot <2>
File View Animation Zoom Tools Help
100%
Date: 2010-12-22 Time: 00
Sat: METOP-A Sensor: AMSU-A Channel: 5 Param: Tb [K]
Type: MFB
an_depar [K]
1
0.5
0
-0.5
Metview - uPlot
File View Animation Zoom Tools Help
100%
Date: 2011-02-19 Time: 00
Aircraft wind over 250 hPa
Type: CCMA obs
120°W 100°W 80°W 60°W 40°W 20°W 0°E 20°E 40°E 60°E
80°N 70°N 60°N 50°N 40°N 30°N 20°N 10°N

Statistics for the Visualised ODB Data



Metview - uPlot
File View Animation Zoom Tools Help
Fit to Width
Frames Layers Data
Titles Coastlines tb_plot Coastlines
Metadata Values
Sensor AMSU-A Channel: 5 Param: Tb [K]
symbol 1

Sidebar with various tabs

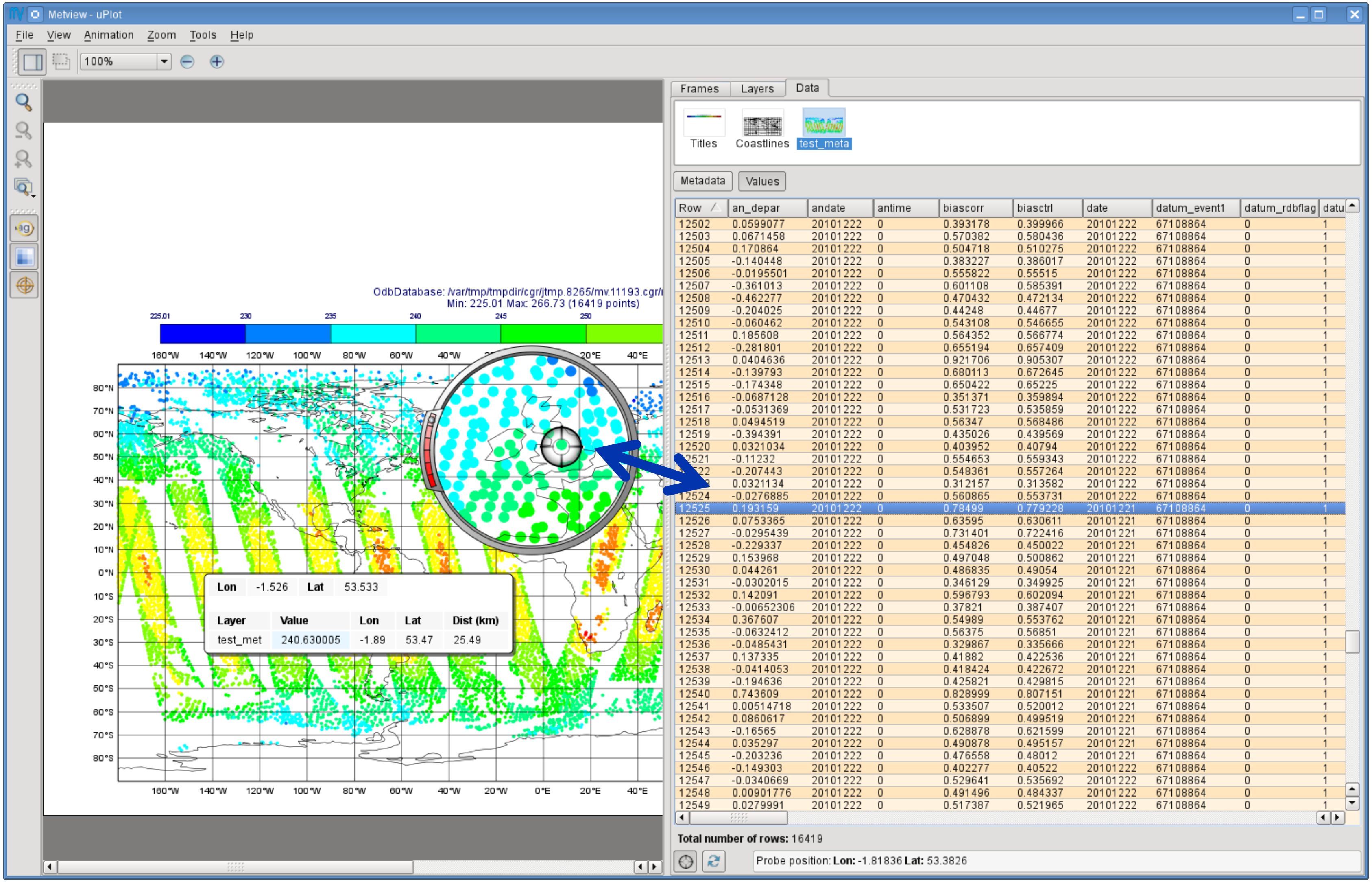


Latitude	lat@hdr
Longitude	lon@hd
Value	obsvalu
	Statis
Points	16419
Minimum	225.01
Maximum	266.73
Average	250.466
Stdev	7.95714
Skewness	-0.6266
Kurtosis	-0.5502

Histogram (for data in visible area)

Bar	From	To	Coun
■	225.01	226.449	10
■	226.449	227.887	11
■	227.887	229.326	32
■	229.326	230.764	73
■	230.764	232.203	127
■	232.203	233.642	219
■	233.642	235.08	247
■	235.08	236.519	291
■	236.519	237.958	422
■	237.958	239.396	474
■	239.396	240.835	592
■	240.835	242.273	634
■	242.273	243.712	638
■	243.712	245.151	622
■	245.151	246.589	688
■	246.589	248.028	615
■	248.028	249.467	697

Inspecting Data Values – Data Probe



OdbDatabase: /var/tmp/tmpdir/cgr/jtmp.8265/mv.11193.cgr/i
Min: 225.01 Max: 266.73 (16419 points)

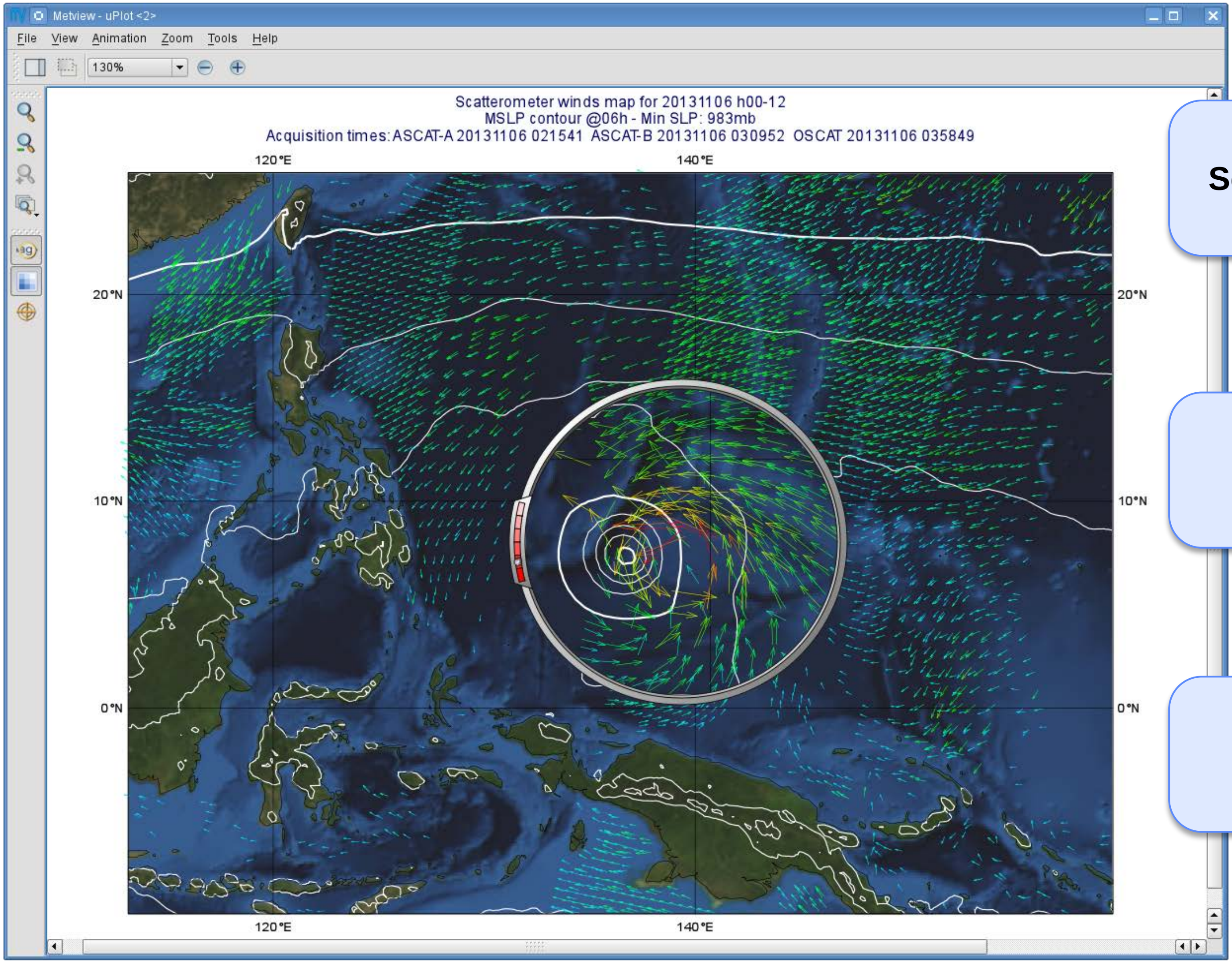
Lon	-1.526	Lat	53.533	
Layer	**Value**	**Lon**	**Lat**	**Dist (km)**
test_met	240.630005	-1.89	53.47	25.49

Row	an_depar	andate	antime	biascorr	biasctrl	date	datum_event1	datum_rdbflag	datu
12502	0.0599077	20101222	0	0.393178	0.399966	20101222	67108864	0	1
12503	0.0671458	20101222	0	0.570382	0.580436	20101222	67108864	0	1
12504	0.170864	20101222	0	0.504718	0.510275	20101222	67108864	0	1
12505	-0.140448	20101222	0	0.383227	0.386017	20101222	67108864	0	1
12506	-0.0195501	20101222	0	0.555822	0.55515	20101222	67108864	0	1
12507	-0.361013	20101222	0	0.601108	0.585391	20101222	67108864	0	1
12508	-0.462277	20101222	0	0.470432	0.472134	20101222	67108864	0	1
12509	-0.204025	20101222	0	0.44248	0.44677	20101222	67108864	0	1
12510	-0.060462	20101222	0	0.543108	0.546655	20101222	67108864	0	1
12511	0.185608	20101222	0	0.564352	0.566774	20101222	67108864	0	1
12512	-0.281801	20101222	0	0.655194	0.657409	20101222	67108864	0	1
12513	0.0404636	20101222	0	0.921706	0.905307	20101222	67108864	0	1
12514	-0.139793	20101222	0	0.680113	0.672645	20101222	67108864	0	1
12515	-0.174348	20101222	0	0.650422	0.65225	20101222	67108864	0	1
12516	-0.0687128	20101222	0	0.351371	0.359894	20101222	67108864	0	1
12517	-0.0531369	20101222	0	0.531723	0.535859	20101222	67108864	0	1
12518	0.0494519	20101222	0	0.56347	0.568486	20101222	67108864	0	1
12519	-0.394391	20101222	0	0.435026	0.439569	20101222	67108864	0	1
12520	0.0321034	20101222	0	0.403952	0.40794	20101222	67108864	0	1
12521	-0.11232	20101222	0	0.554653	0.559343	20101222	67108864	0	1
12522	-0.207443	20101222	0	0.548361	0.557264	20101222	67108864	0	1
12523	0.0321134	20101222	0	0.312157	0.313582	20101222	67108864	0	1
12524	-0.0276885	20101222	0	0.560865	0.553731	20101222	67108864	0	1
12525	0.193159	20101222	0	0.78499	0.779228	20101221	67108864	0	1
12526	0.0753365	20101222	0	0.63595	0.630611	20101221	67108864	0	1
12527	-0.0295439	20101222	0	0.731401	0.722416	20101221	67108864	0	1
12528	-0.229337	20101222	0	0.454826	0.450022	20101221	67108864	0	1
12529	0.153968	20101222	0	0.497048	0.500862	20101221	67108864	0	1
12530	0.044261	20101222	0	0.486835	0.49054	20101221	67108864	0	1
12531	-0.0302015	20101222	0	0.346129	0.349925	20101221	67108864	0	1
12532	0.142091	20101222	0	0.596793	0.602094	20101221	67108864	0	1
12533	-0.00652306	20101222	0	0.37821	0.387407	20101221	67108864	0	1
12534	0.367607	20101222	0	0.54989	0.553762	20101221	67108864	0	1
12535	-0.0632412	20101222	0	0.56375	0.56851	20101221	67108864	0	1
12536	-0.0485431	20101222	0	0.329867	0.335666	20101221	67108864	0	1
12537	0.137335	20101222	0	0.41882	0.422536	20101221	67108864	0	1
12538	-0.0414053	20101222	0	0.418424	0.422672	20101221	67108864	0	1
12539	-0.194636	20101222	0	0.425821	0.429815	20101221	67108864	0	1
12540	0.743609	20101222	0	0.828999	0.807151	20101221	67108864	0	1
12541	0.00514718	20101222	0	0.533507	0.520012	20101221	67108864	0	1
12542	0.0860617	20101222	0	0.506899	0.499519	20101221	67108864	0	1
12543	-0.16565	20101222	0	0.628878	0.621599	20101221	67108864	0	1
12544	0.035297	20101222	0	0.490878	0.495157	20101221	67108864	0	1
12545	-0.203236	20101222	0	0.476558	0.48012	20101221	67108864	0	1
12546	-0.149303	20101222	0	0.402277	0.40522	20101222	67108864	0	1
12547	-0.0340669	20101222	0	0.529641	0.535692	20101222	67108864	0	1
12548	0.00901776	20101222	0	0.491496	0.484337	20101222	67108864	0	1
12549	0.0279991	20101222	0	0.517387	0.521965	20101222	67108864	0	1

Total number of rows: 16419

Probe position: **Lon:** -1.81836 **Lat:** 53.3826

Overlay with Other Data Types



Scatterometer winds map for 20131106 h00-12
MSLP contour @06h - Min SLP: 983mb
Acquisition times: ASCAT-A 20131106 021541 ASCAT-B 20131106 030952 OSCAT 20131106 035849

ODB: Scatterometer wind

GRIB: IFS MSLP analysis

WMS: map image from NASA

Plans for the Metview ODB interface

- **Time series support**
- **Provide contouring for ODB data**
- **Visualise ODB data in thermodynamic diagrams**

For more information …

- **ODB API**
  - **https://software.ecmwf.int/wiki/display/ODB/ODB-API (soon)**
  - **Email to: peter.kuchta@ecmwf.int**

- **Metview**
  - **https://software.ecmwf.int/metview**
  - **Email to: metview@ecmwf.int**