# Towards a Unified Observations Monitoring System

**Mohamed Dahoui & Lars Isaksen**

# Outline

1. Introduction
2. ECMWF Monitoring Systems
3. Towards a unified monitoring system
4. Road Map

# Outline

1. Introduction
2. ECMWF monitoring Systems
3. Towards a unified monitoring system
4. Road Map

# Introduction



- Data monitoring is a crucial component of the data assimilation diagnostic system. It allows the control of the availability and the quality of the observing system (and potentially the data impact)

- Data monitoring consists of the production of statistics over large data samples. Statistics are generally computed for quantities related to the data assimilation

- Monitoring tools are designed to produce statistics according to various data  selection criteria

# Introduction



- Monitoring outputs are important to define and evaluate data usage

- Monitoring results are generally exchanged between NWP centres to evaluate the data usage

- Monitoring results are very valuable for data providers

# Outline

1. Introduction
2. ECMWF monitoring Systems
3. Towards a unified monitoring system
4. Road Map

# ECMWF Monitoring Tools



| Software | Usage | Inputs | Outputs | Mode |
|---|---|---|---|---|
| **SATMON** | Satellite data monitoring | BUFR & ODB | NetCDF (raw stats) + plots | Oper & Research |
| **OBSTAT** | Satellite & conventional data monitoring | ODB & BUFR | Ascii (raw stats) + plots | Oper & Research |
| **MetOps Monitoring tools** | Conventional (and partially satellite data) monitoring | BUFR (being adapted to ODB) | Binary format (raw stats) + plots en demand | Oper |
| **Waves Monitoring tool** | Waves and surface winds | Offline from Waves DA (no feedback data) | Binary format + plots | Research & Oper |
| **Alarm system** | Satellite data | ODB | E-mails + plots | Experimental |

# Outline

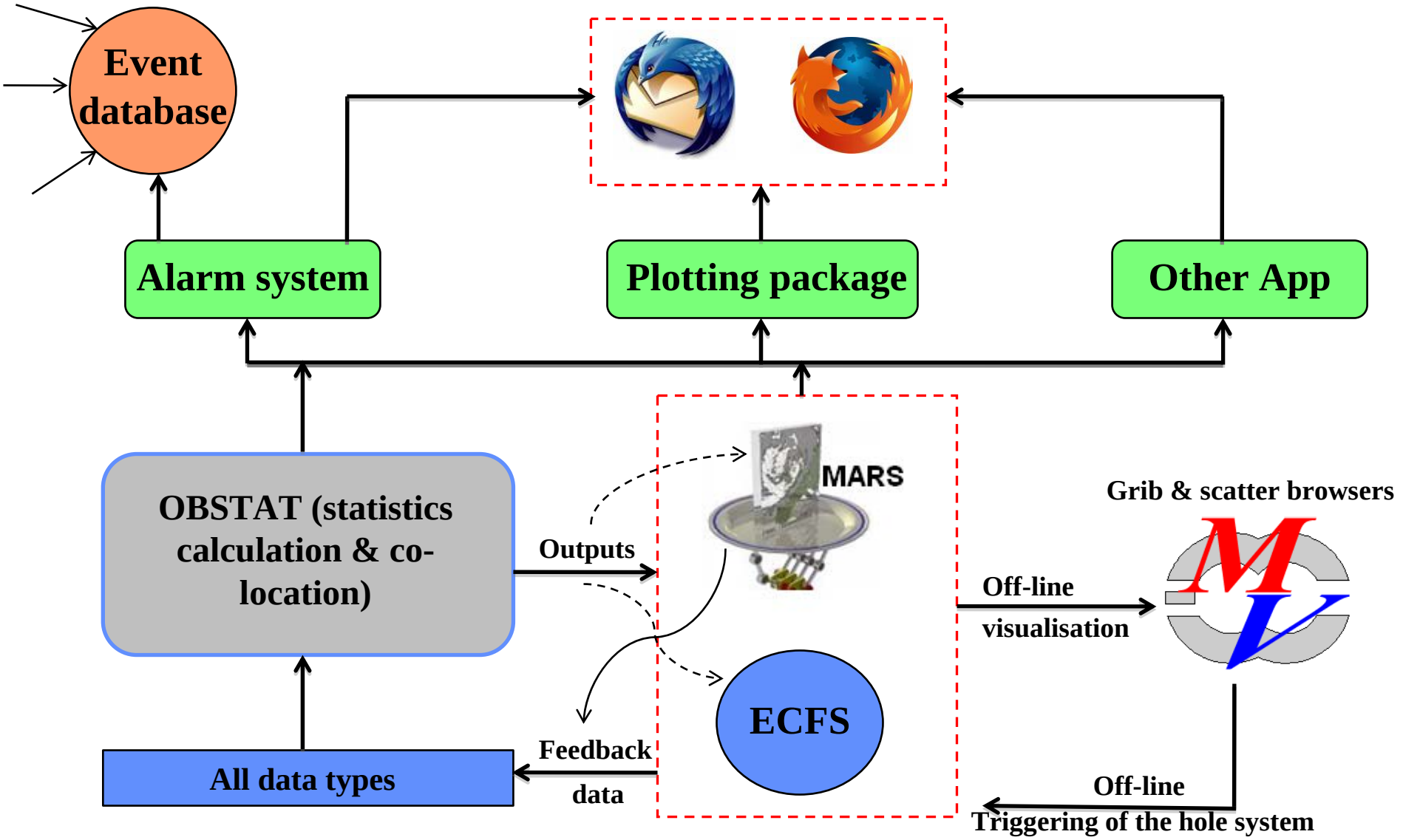
1. Introduction
2. ECMWF monitoring Systems
3. Towards a unified monitoring system
4. Road Map

# Towards a unified system

- Optimization of the maintenance and support
- Handling of all existing data types (conventional and satellite data from all Data Assimilation systems)
- Easy introduction of new data types
- Easy handling and use of the monitoring statistics
- Built-in Alarm system
- Co-location capability

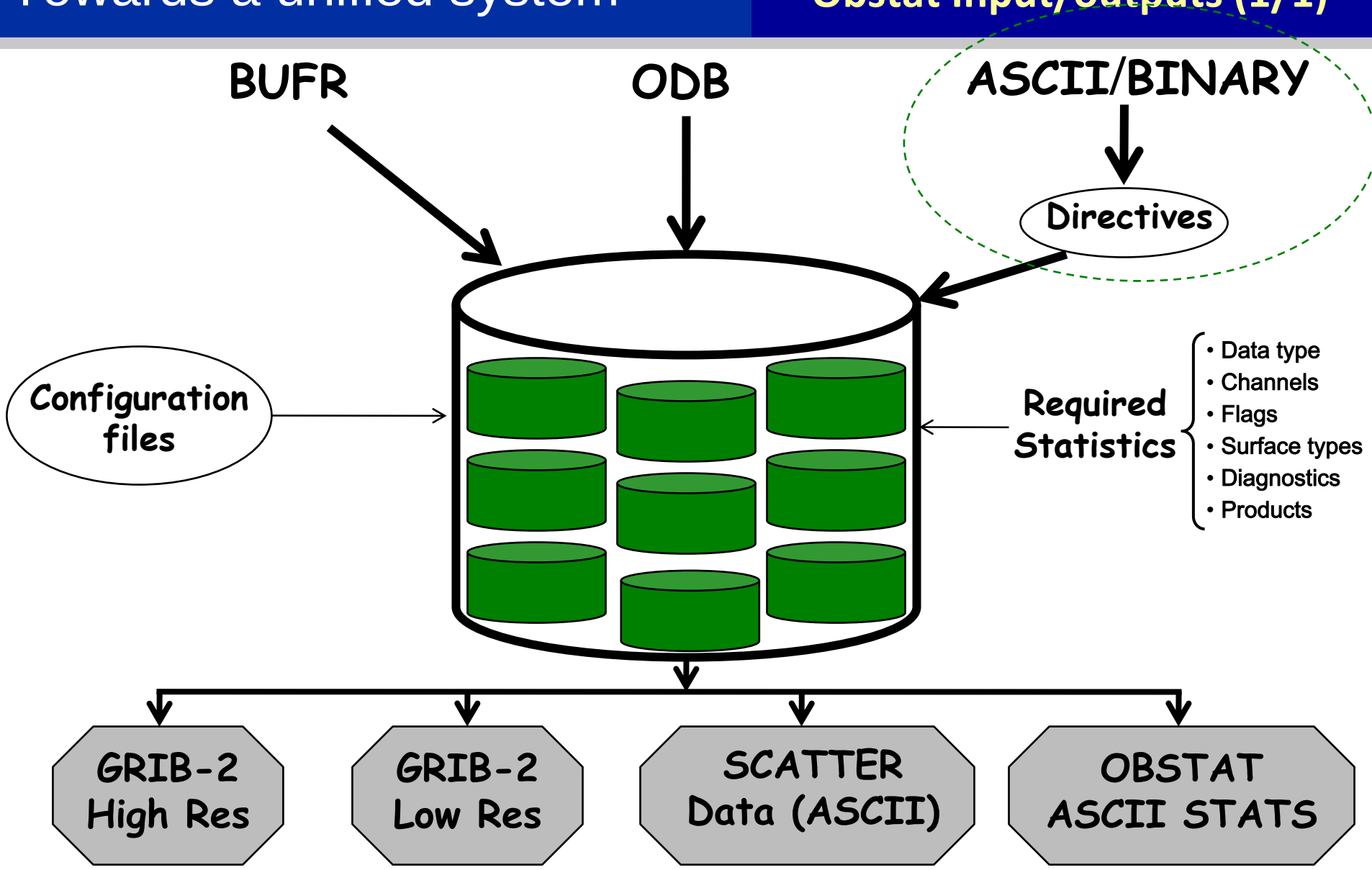
# Towards a unified system

## Unified tool scheme (1/1)



# Towards a unified system

## Obstat Input/outputs (1/1)



# Towards a unified system

## GRIB-2 format

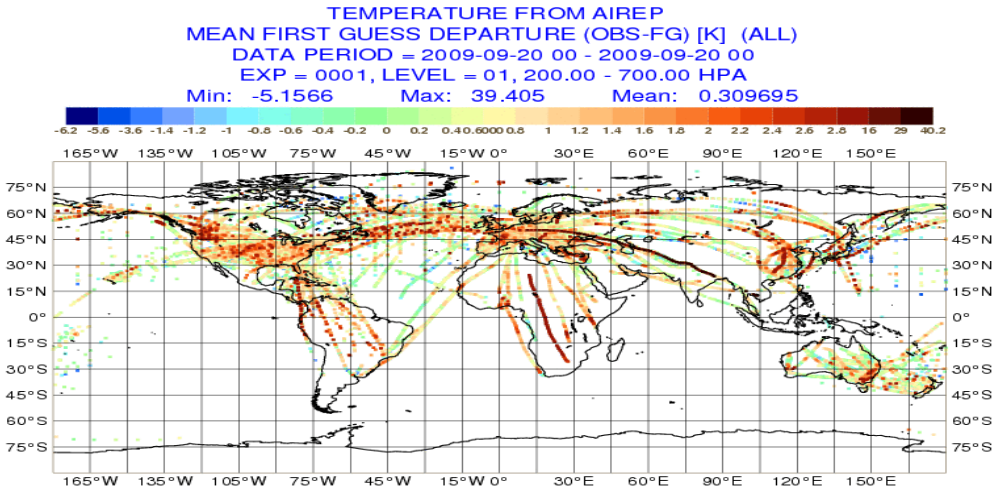
- GRIB-2 is a standard WMO format
- Offers a good compression
- Supported by ECMWF Archiving system (MARS): Archiving with indexation & retrieval with filtering. A new GRIB-2 template is being defined and MARS being extended to cater with the archiving of monitoring statistics
- Well handled by ECMWF graphical tools (Magics, Magics++ and METVIEW)
- Suitable for plotting on demand (if data are quickly accessible)
- Offers large possibilities for statistics exchange and inter-comparison between NWP centres and data providers
- Convertible to other formats (e.g. NetCDF)

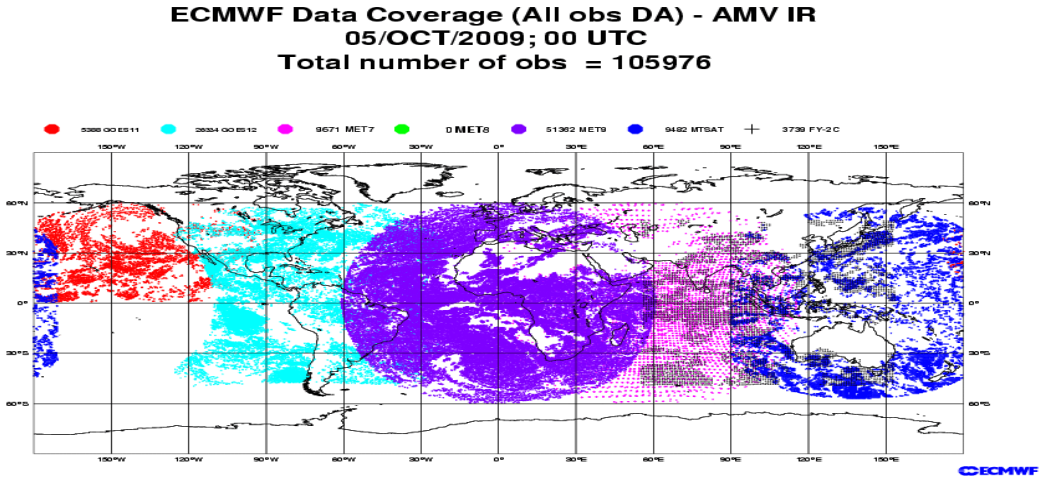
## High Res GRIBs (1/2)

- The production of the high resolution GRIBs is optional (default yes)
- Three grid types supported : REGULAR, GAUSSIAN and REDUCED GAUSSIAN
- For REGULAR GRID, free specification of the resolution (default 1°x1°)
- For GAUSSIAN and REDUCED GAUSSIAN : N16, N24, N32, N48, N64, N80, N108, N128, N160, N200, N256, N320, N400, N512, N576, N640 and N1024
- A bitmap is used to handle missing values
- Possibility to normalize statistics over a period (one GRIB produced) or keep statistics per analysis cycle

# Towards a unified system

## High Res GRIBs (2/2)

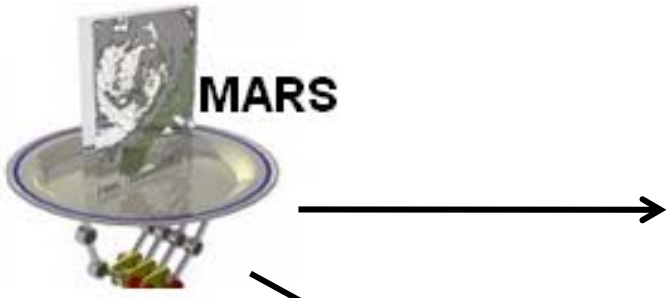



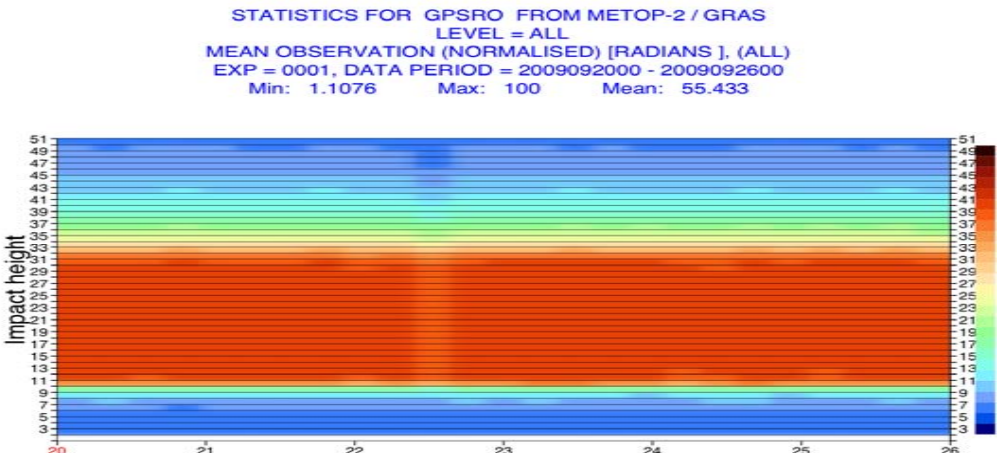
## LOW Res GRIBs (1/3)

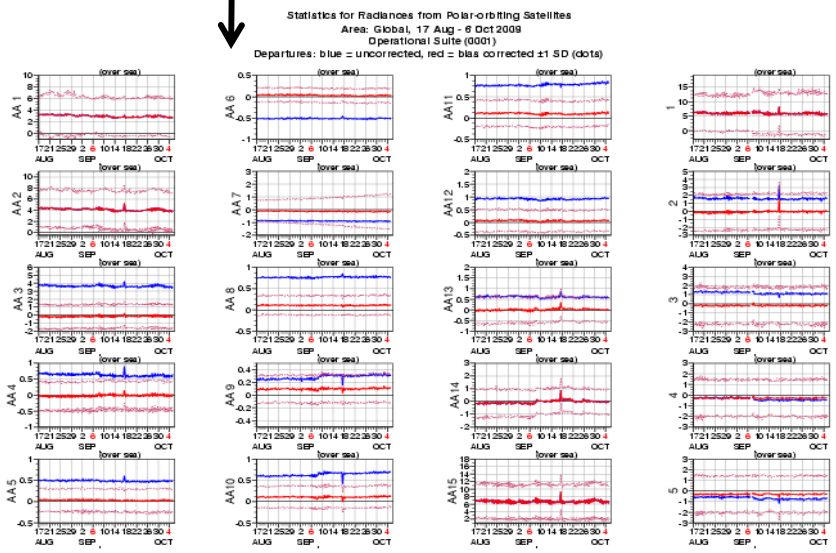
- Low resolution GRIBs are useful for generating quickly large area time series statistics
- The production of the low resolution GRIBs is optional (default yes)
- Only REGULAR grid is supported with a free specification of the resolution (default 2.5°x30°)
- A bitmap is used to handle missing values
- Possibility to produce statistics according to the surface type : land, sea, sea-ice and all surface types

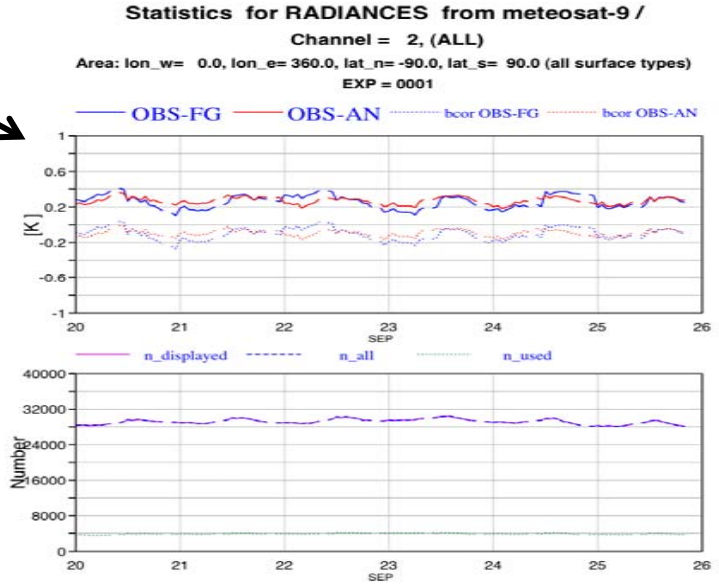
# Towards a unified system

## Low Res GRIBs (2/3)



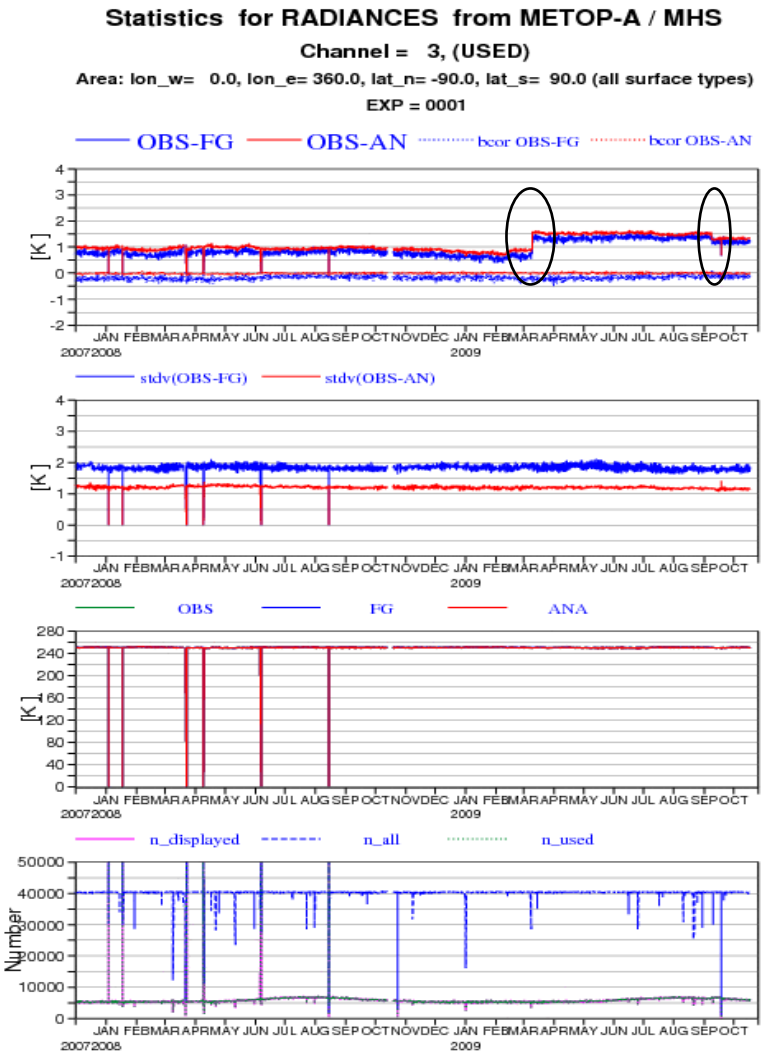



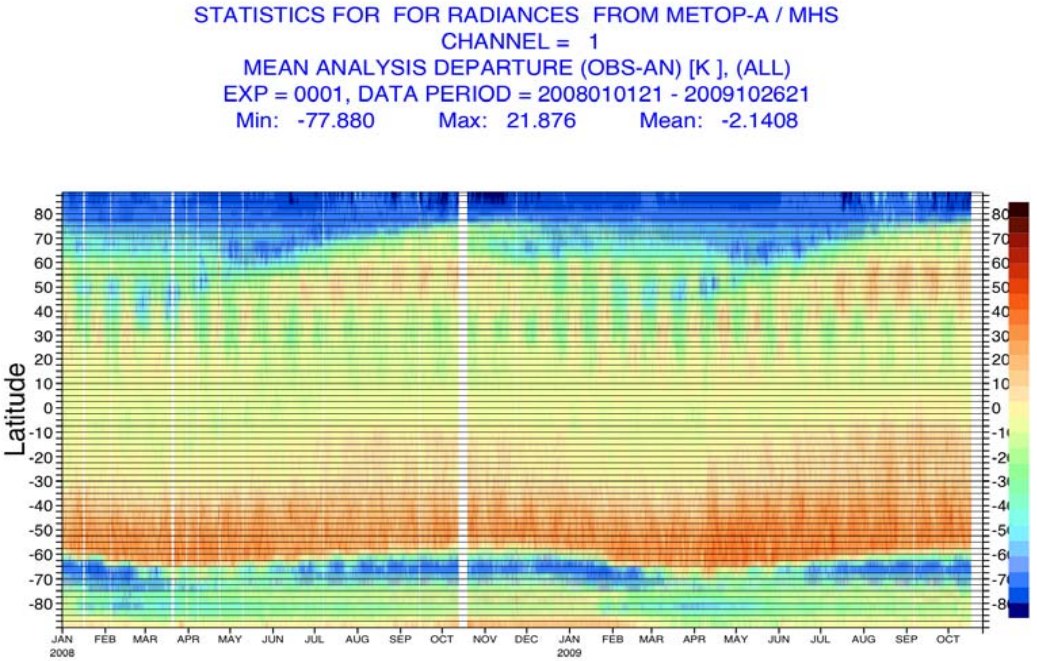



# Towards a unified system

## Low Res GRIBs (3/3)



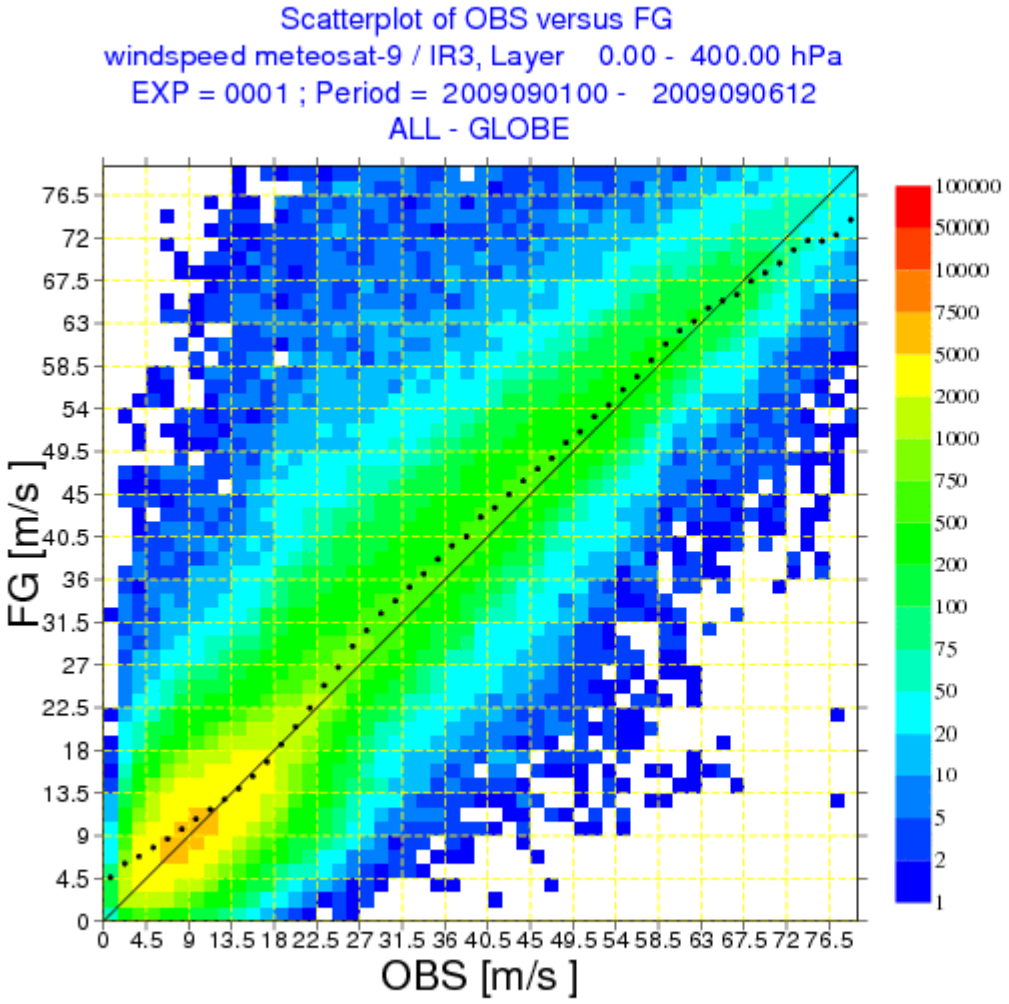



ECMWF

## Scatter data (1/3)

- The production of scatter data is optional (default no)
- Up to 18 couple of variables can be defined for comparison
- Necessity to define the binning to be used for statistics gathering
- Possibility to produce scatter statistics for different areas
- Scatter data can be produced, in one go, for different channels/levels/layers/areas and for different data selection criteria
- Populations are stored in a self-documented ASCII file. Obstat stores global statistics related to both quantities in comparison

## Scatter data (2/3)



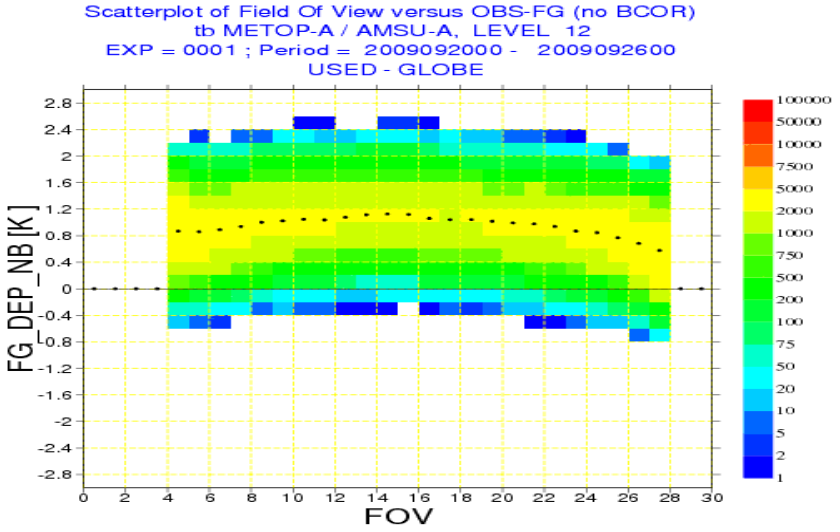

Total number = 468458
Maximum number per bin = 6349

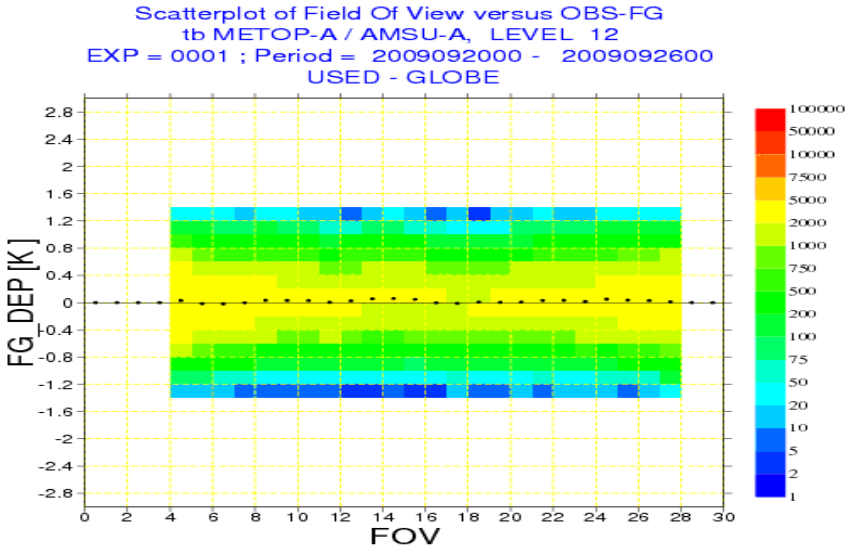
FG min = 0.0 max = 79.5

OBS min = 0.0 max = 79.5

y mean = 20.6 y stdev = 16.4
x mean = 19.4 x stdev = 15.1
BIAS (y-x) = 1.23
RMS = 6.30
corr. coef. = 0.926

## Scatter data (3/3)

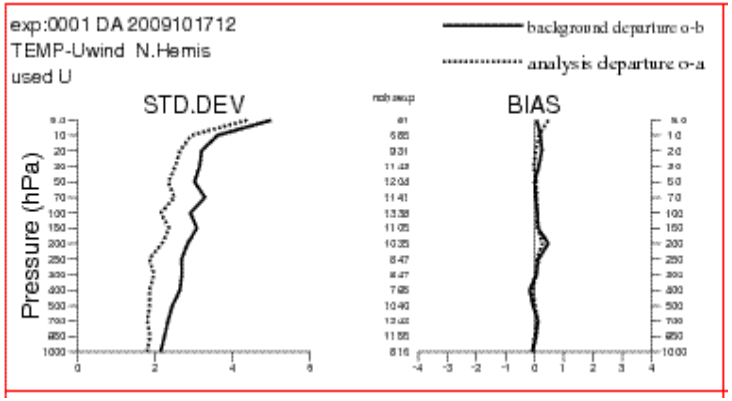


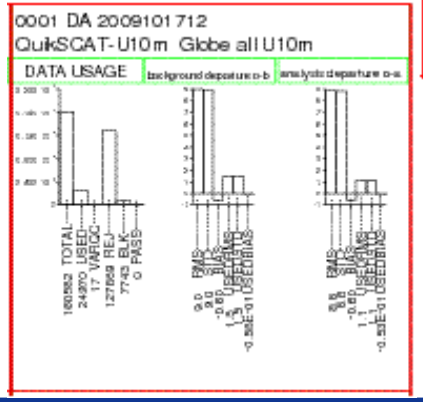



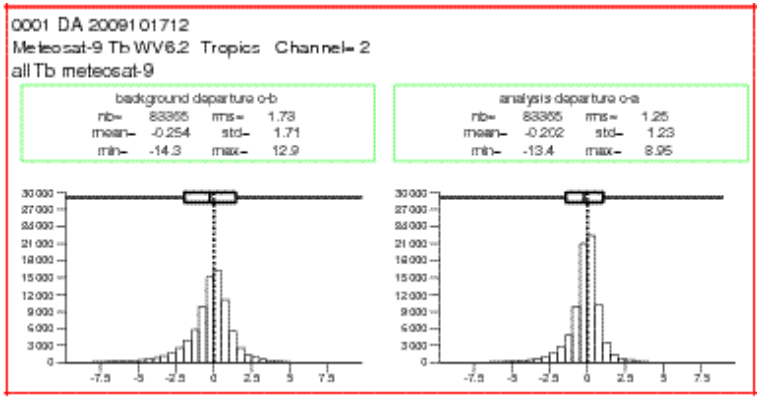
# Towards a unified system

## Classic Obstat products

- Classic Obstat statistics are stored in ASCII format
- Classical Obstat plots are : RMS/STD.DEV plots (vertical profiles), Histograms and Usage charts
- Experiments Superimposition capability
- Possibility to customize the final product by gathering individual plots



- Obstat can produce statistics for a large number of pre-defined quantities
- Pre-defined quantities are either directly extracted from the input database or computed internally inside Obstat
- Users have the possibility to define their own diagnostics without any code change. A configuration file (in ASCII) is available for that

```
#-------------------------------------------------------------------------------------------------------------
# Param     SQLflag      Obstat_Number    Grib_Number    CompMethod       Datalist
# ----------------------------------GE 80-------------------------------------------------------------------
   tausfc   tausfc@body        95              228              0        amsua,amsub,hirs,mhs
  fgvalue         0             80              230        itm_8-itm_9          all
  anvalue         0             90              234        itm_8-itm_10         all
```

- Obstat can produce statistics according to several pre-defined data selection criteria
- Data selection criteria are generally related to data usage flags, surface type and specific data conditions
- Statistics related to several data criteria can be produced in one go
- Users have the possibility to define their own data selection criteria without any code change (if only one channel/level/layer is involved). A configuration file (in ASCII) is available for that

```
#-------------------------------------------------------------------------------
# flag            Number              Definition     Datalist
# -----------------GE 20-----------------------------------------------------

  Daytime          20                 (itm_34>0)       reo3
  Nightime         21                 (itm_34<=0)      reo3
```

# Towards a unified system

## Pre-defined data selection

| Data selection | Valid for |
|---|---|
| All | **All** |
| Active data | **All** |
| Not Active data | **All** |
| Best Active winds | **Scatterometer** |
| Used data | **All** |
| VarQC rejected data | **All** |
| Blacklisted data | **All** |
| Failed data | **All** |
| Data passed FG check | **All** |

| Data selection | Valid for |
|---|---|
| All non rejected data | **All** |
| VarBC passive data | **All** |
| Data failed FG check but not blacklisted | **All** |
| Data failed FG check and VarQC rejected | **All** |
| Selection according to QI | **AMVs** |
| Selection according to cloud status | **Radiances** |
| Selection according to rain status | **Microwave obs** |
| Good ozone data | **Ozone** |
| Day and Night data | **Ozone** |

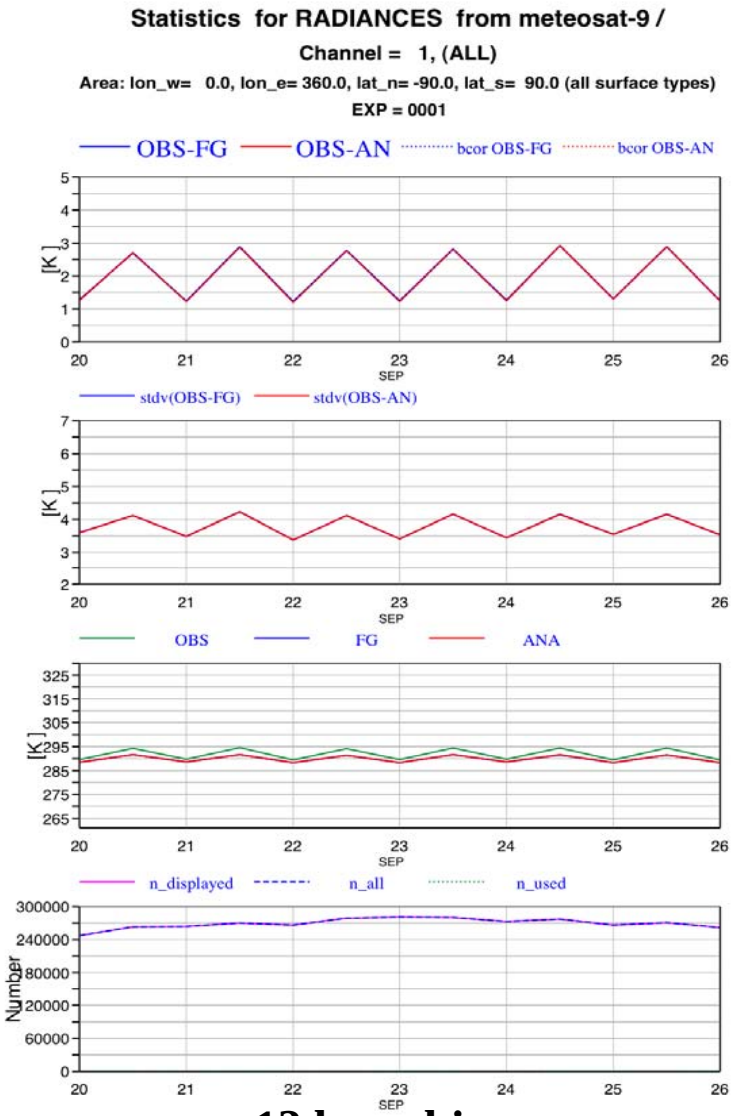
- Obstat can handle all existing vertical coordinate : Channels, Pressure levels and pressure layers
- Obstat offers several possibilities for pressure layers binning :
  - A standard pressure binning is defined in Obstat (through configuration files)
  - For each data type, the user can define a different pressure binning (which has the priority over the standard one)
  - Obstat offers the possibility to define pressure layers which are overlapped (or embedded within each other)

## Time binning (1/2)

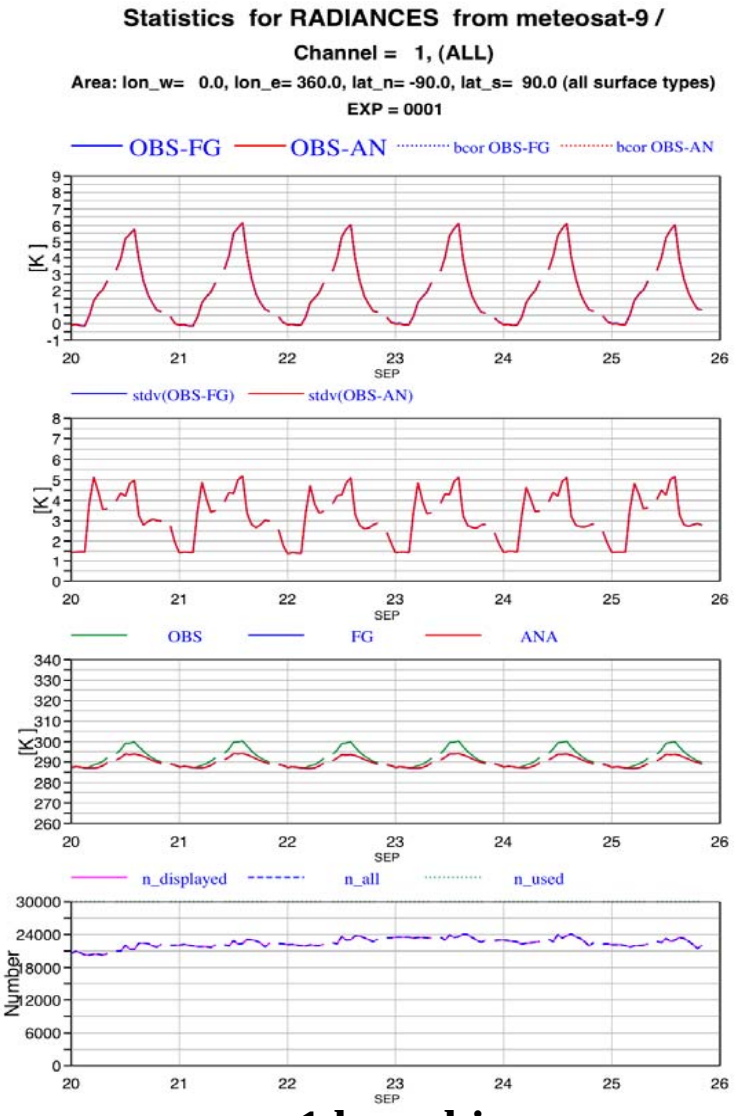
- Obstat offers the possibility to define the time binning of data (from 1 hour to DA window length)
- The default time binning is the DA window length (12h for DCDA and 6 for DA)
- Very useful for data types with high variability, in time, of statistics (e.g CSR)

# Towards a unified system

## Time binning (2/2)




12 hour bins




1 hour bins



ECMWF

## Database Auto browsing

- By default the user has to specify and customize the definition of wanted statistics. For that the user needs to know : Obstype, Codetype, Varno & instrument
- Obstat offers a slow option to browse the input database and produce generic statistics for all present data
- Generic statistics includes all levels, data types but limited to **all** and **used** data only

## Obstat plotting package

- Obstat offers the possibility to produce **generic** monitoring plots : Time series, Hovmoeller diagrams, geographical maps, scatter plots, RMS plots and histograms

- Obstat plotting package is currently based on Magics but being adapted to Magics++

- The plotting is separated from statistics calculation (requiring heavy data transfers). The users are advised to request the maximum statistics to be produced in one Obstat run. The plotting package can be applied separately to a subset of statistics

# Outline

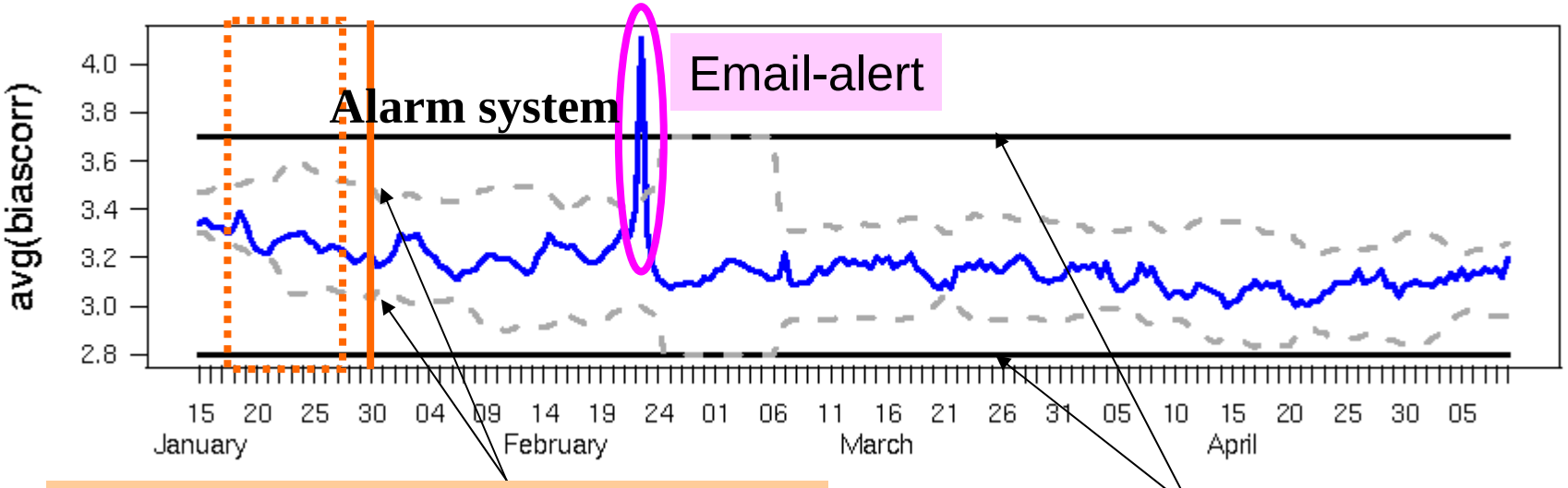
1. Introduction
2. ECMWF monitoring Systems
3. Towards a unified monitoring system
4. Road Map

# Road Map

- Archiving of the monitoring statistics in MARS
- Re-write the plotting programs and scripts to add more possibilities, allow more flexibility and offer full compatibility with Magics++
- Adapt and Plug-in the Automatic alarm system
- Adapt and implement the existing co-location tools
- Implement Obstat as the main operational monitoring software
- Implementation, by the Graphics section, of an interface to Obstat within METVIEW
- Write a comprehensive documentation

Selected statistics are checked against an expected range.

E.g., global mean bias correction for GOES-12 (in blue):




Soft limits (mean ± 5 stdev of statistic to be checked, calculated from past statistics over a period of 20 days, ending 2 days earlier)

Hard limits (fixed)

# Alarm system (2/2)

**http://www.ecmwf.int/products/forecasts/satellite_check/**

**http://www.ecmwf.int/products/forecasts/d/charts/monitoring/satellite/**

Satellite Data Automatic Checking - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ecmwf.int/products/forecasts/satellite_check

ECMWF

Home Your Room Login Contact Feedback Site Map Search:

About Us: Overview, Getting here, Committees
Products: Forecasts, Order Data, Order Software
Services: Computing, Archive, PrepIFS
Research: Modelling, Reanalysis, Seasonal
Publications: Newsletters, Manuals, Library
News&Events: Calendar, Employment, Open Tenders

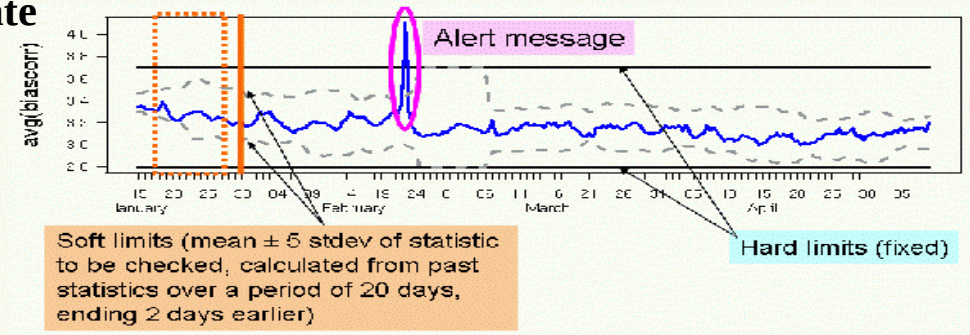
Home > Products > Forecasts > Satellite Data Automatic Checking

## Satellite Data Automatic Checking

Products
- Forecasts
- Data and Software Ordering
- Catalogue
- GTS Products
- Operational Upgrades

An experimental automatic satellite data checking system has been implemented recently at ECMWF. It triggers the production of alarm messages if an anomaly is detected in the quality or the availability of the satellite data assimilated by the model.

Selected statistical parameters (number of observations, bias correction, and mean bias-corrected background and analysis departures) are checked against an expected range. An appropriate alert message (including a time series plot) is generated if statistics are outside the specified ranges. A severity level (slight, considerable, severe) is assigned to each message depending on how far statistics are from the expected values. Two kinds of ranges are used by the automatic checking: Soft and Hard limits. Soft limits are updated automatically using statistics from the last twenty days (extremes are excluded during this process). Hard limits are adjusted manually when required.




Currently, the automatic checking is limited to data passing through the minimisation process (including VarBC passive data). It's being applied, twice a day, to the long cut-off 4D-VAR cycles (DCDA).

- Experimental Satellite Data Checking for 2009102512 DCDA
- Experimental Satellite Data Checking for 2009102500 DCDA

Done

# Thank you for your attention