User Experience and Requirements

Renate Hagedorn

European Centre for Medium-Range Weather Forecasts

Definitions

User:

- user of “put” infrastructure ➘ producer of data
- user of “get” infrastructure ➘ consumer of data

Experience:

- DEMETER data producer / consumer

Requirements:

- Response to questionnaire / draft proposal

DEMETER:

Development of a European multi-model ensemble system for seasonal to interannual prediction

- DEMETER goals:
  - investigate single-model vs. multi-model concept
  - explore utility of end-user applications in probabilistic context

- DEMETER system:
  - 7 coupled global circulation models, each 9 ensemble members
  - Production and diagnostic of re-forecasts for 1958 - 2001
  - Downscaling of DEMETER data set
  - Use of DEMETER data as input for application models

Past Experience

Important steps in the creation of the DEMETER data set

- Agree on what to archive
  - which parameter, level, frequency, ocean data
- Agree on how to archive
  - units, accum./avg. periods, format
- Agree on where to archive
  - central archive (MARS) for original data
  - public data server for reduced data set on 2.5 deg

Sounds trivial… but the devil lies in the detail!

What to archive

Common output for DEMETER (atmosphere)

Output every 24 hours at 00 GMT

Pressure levels (instantaneous): Z (129, m^2/s^2), T (130, K), U (131, m/s), V (132, m/s), Q (133, kg/kg) for 850, 500, 200 hPa

Surface (instantaneous):
- 2m Tmax - daily (201, K)
- 2m Tmin - daily (202, K)
- Total soil moisture (229, m of water)
- Snow depth (141, m of water)
- Sea surface temperature and/or some temperature over land (034 or 139, depending on the value over land, K)
- Mean sea level pressure (151, Pa)

Surface (accumulated):
- Total precipitation (228, m)
- Downward surface solar radiation (169, Ws/m^2)
- Downward surface longwave radiation (175, Ws/m^2)
- Surface net solar radiation (176, Ws/m^2)
- Surface net longwave radiation (177, Ws/m^2)
- Top net solar radiation (178,Ws/m^2)
- Top net longwave radiation - OLR (179, Ws/m^2)

Output every 6 hours at 00, 06, 12, 18 GMT

Surface (instantaneous):
- Total cloud cover (164, [0,1])
- 10m wind - U (165, m/s)
- 10m wind - V (166, m/s)
- 2m T (167, K)
- 2m dewpoint (168, K)

- GRIB codes and variable units are indicated in parentheses, the ECMWF local table version being 190
- Daily accumulations
- Integrations starting at 00 GMT
- Archive on model grid
- Q stands for specific humidity, Z for geopotential (and not for geopotential height)
- Archive monthly means of the same fields into MARS (convert accumulated fields to fluxes, units: m/s for precipitation and W/m^2 for radiation fields)
- All relevant constant fields should also be archived into MARS. These include land-sea mask, orography, roughness length, underlying albedo, field capacity, wilting points, etc, taking into account their seasonal variations (if any). Values for one year might be enough (e.g. 1987).

Common output for DEMETER (ocean)

Monthly mean output

Horizontal:
- T (129, °C), S (130, PSU), U (131, m/s), V (132, m/s), at the surface
- D20 (163, m), 20°C isotherm depth
- SL (145, m) sea level
- TAV300 (164, °C) average potential temperature in the upper 300m, proxy to heat content
- Mixed layer depth (148, m)

Vertical:
- T(z), S(z), U(z), W(z) (section along equator)
- T(z), S(z), U(z), V(z), W(z) (meridional sections along 140W, 165E, 30W, and 60E)

Ocean analyses

Same fields as above, but stored as instantaneous fields

- GRIB codes and variable units are indicated in parenthesis, the ECMWF local table version being 151
- Archive on Levitus grid

How to archive

- Special attention to
  - non instantaneous output parameter
    - total precipitation
    - heat and radiation fluxes
    - min/max temperatures, wind gust,…
  - model specific output parameter
    - soil temperature / moisture in different levels
    - surface / skin temperature
    - field capacity, wilting points,…

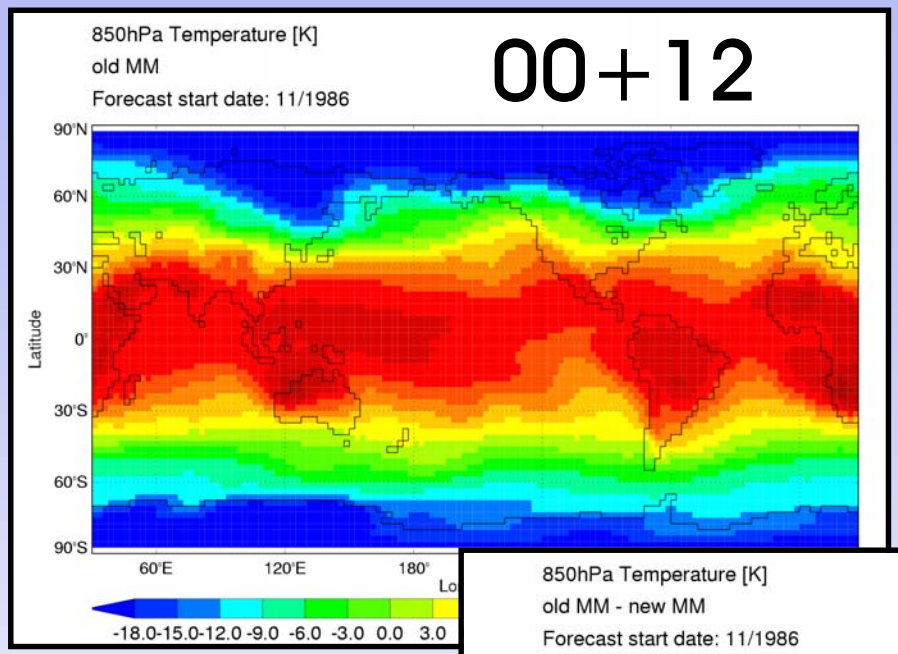
Exact definition of content, aggregation and units is essential

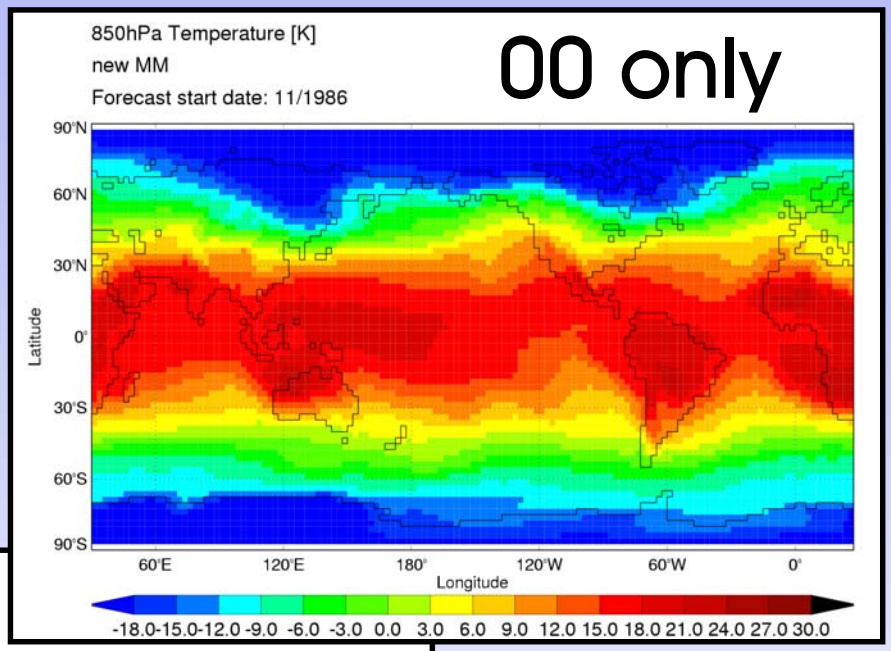
Example of what can go wrong

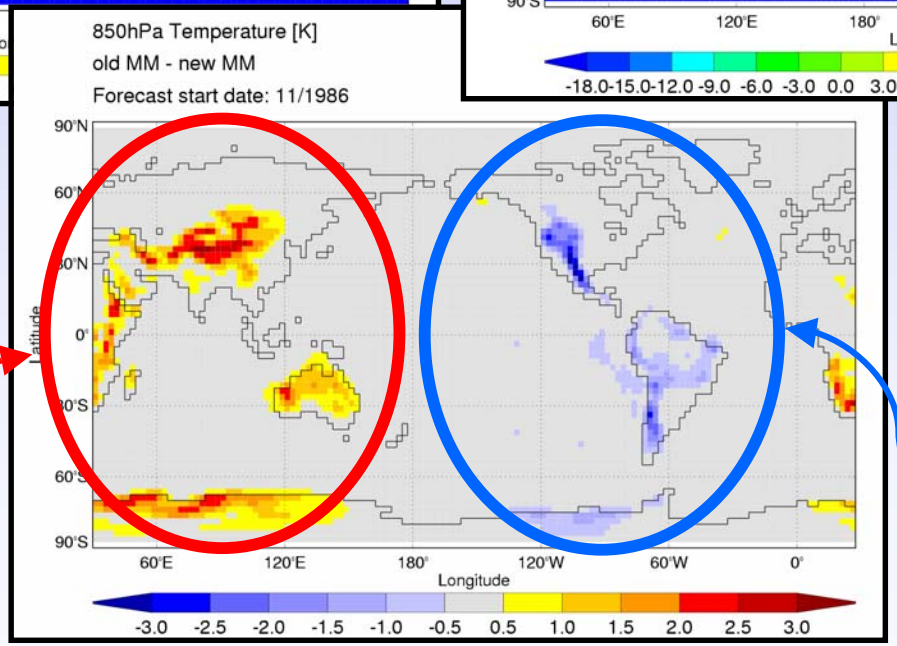
- agreed construction of monthly means:
  - take output at 00 only

- ECMWF construction of monthly means:
  - take output at 00 + 12

Monthly Means: 00+12 vs. 00 only





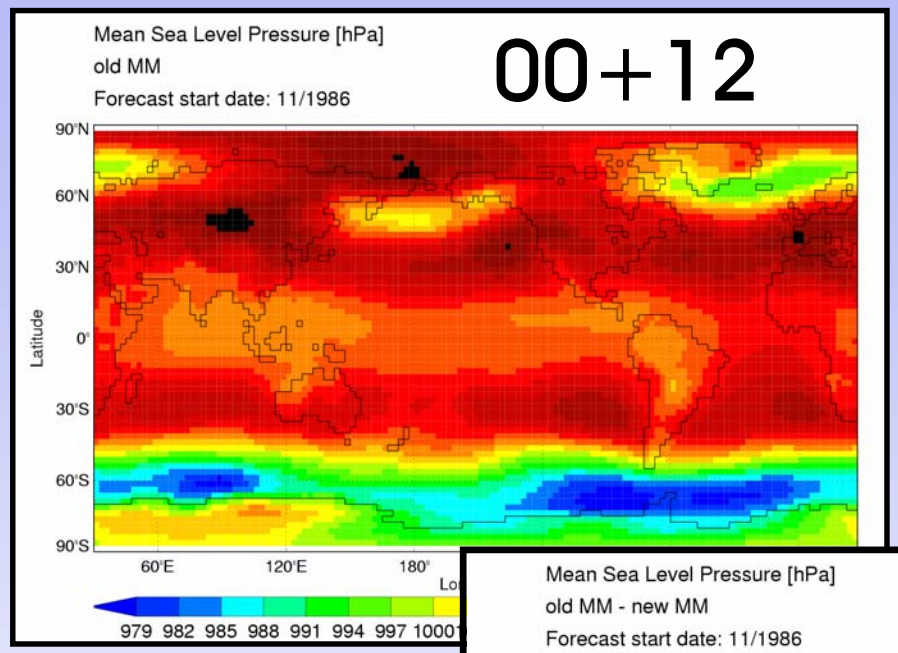


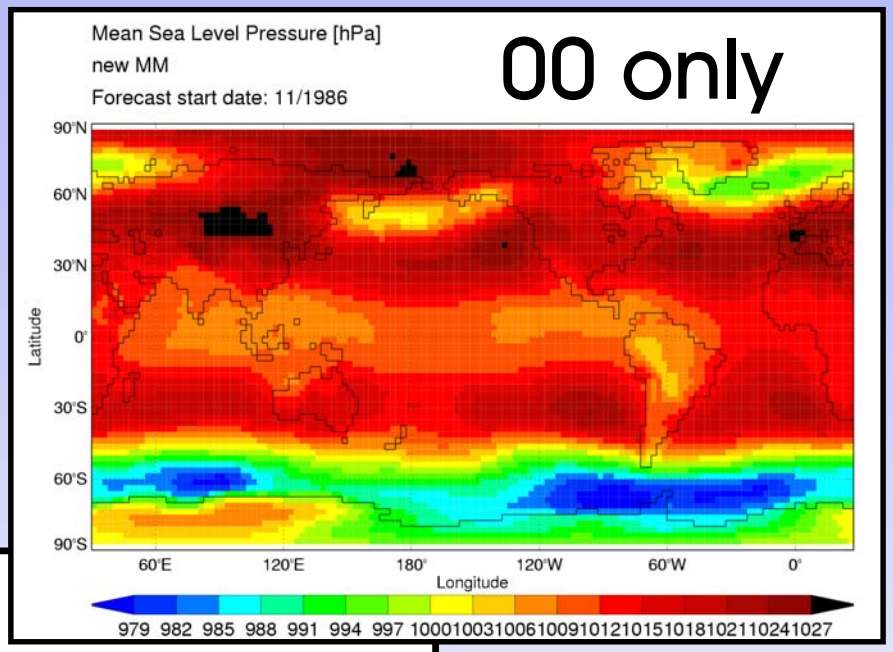
00+12
MM's are:

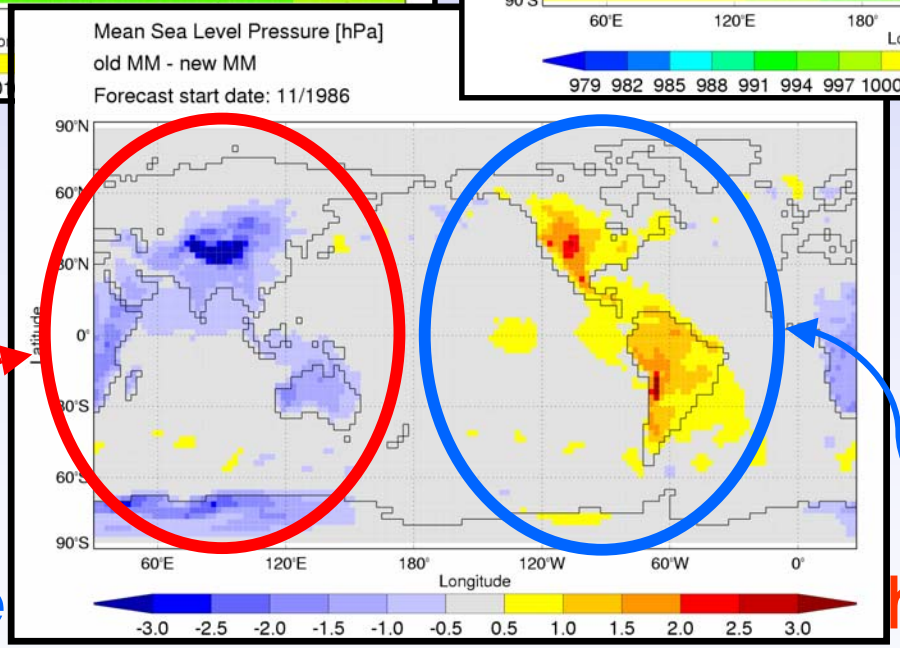
warmer

colder

Monthly Means: 00+12 vs. 00 only







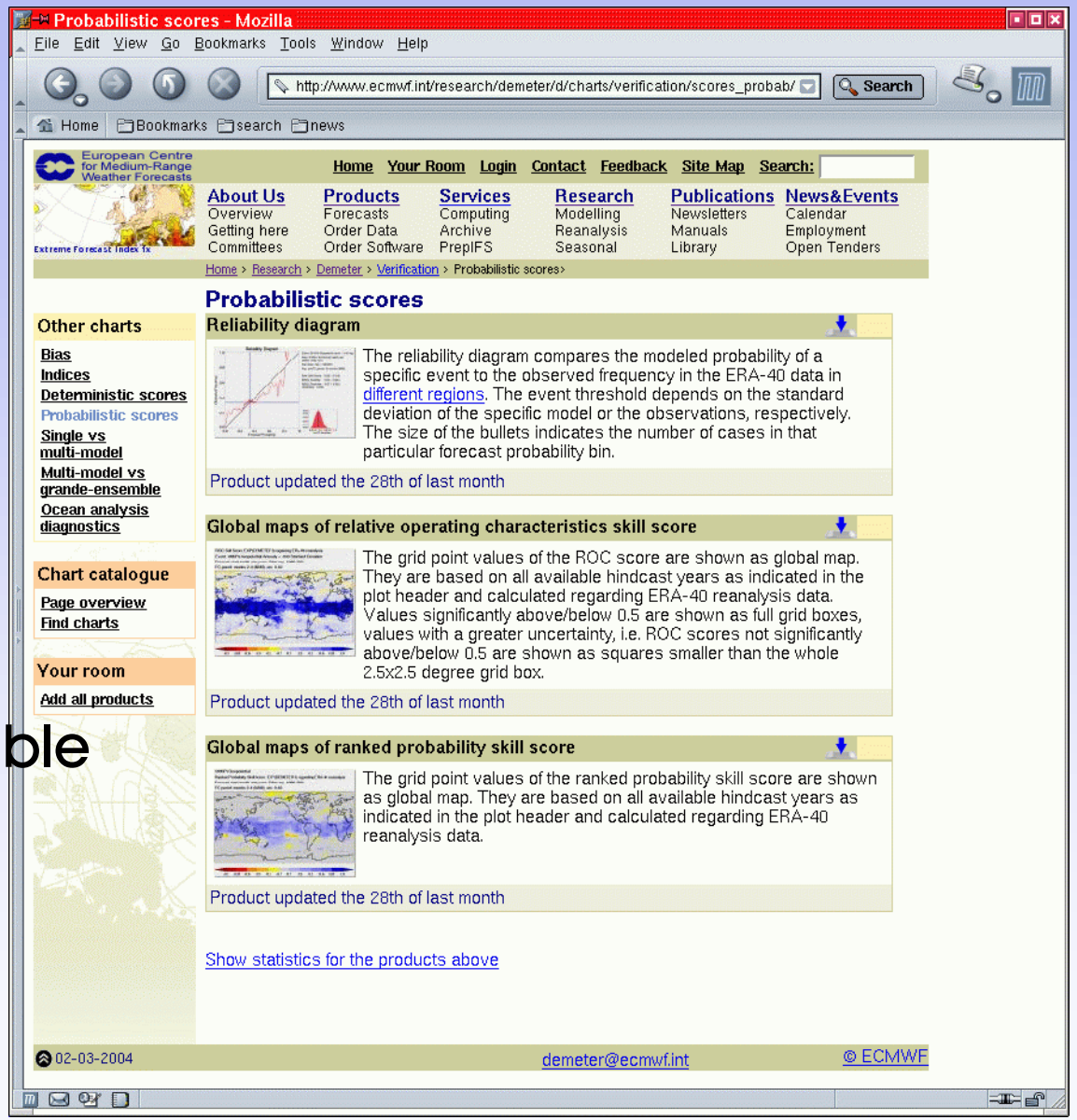
00+12
MM's are:

warmer
lower pressure

colder
higher pressure

Verification Suite

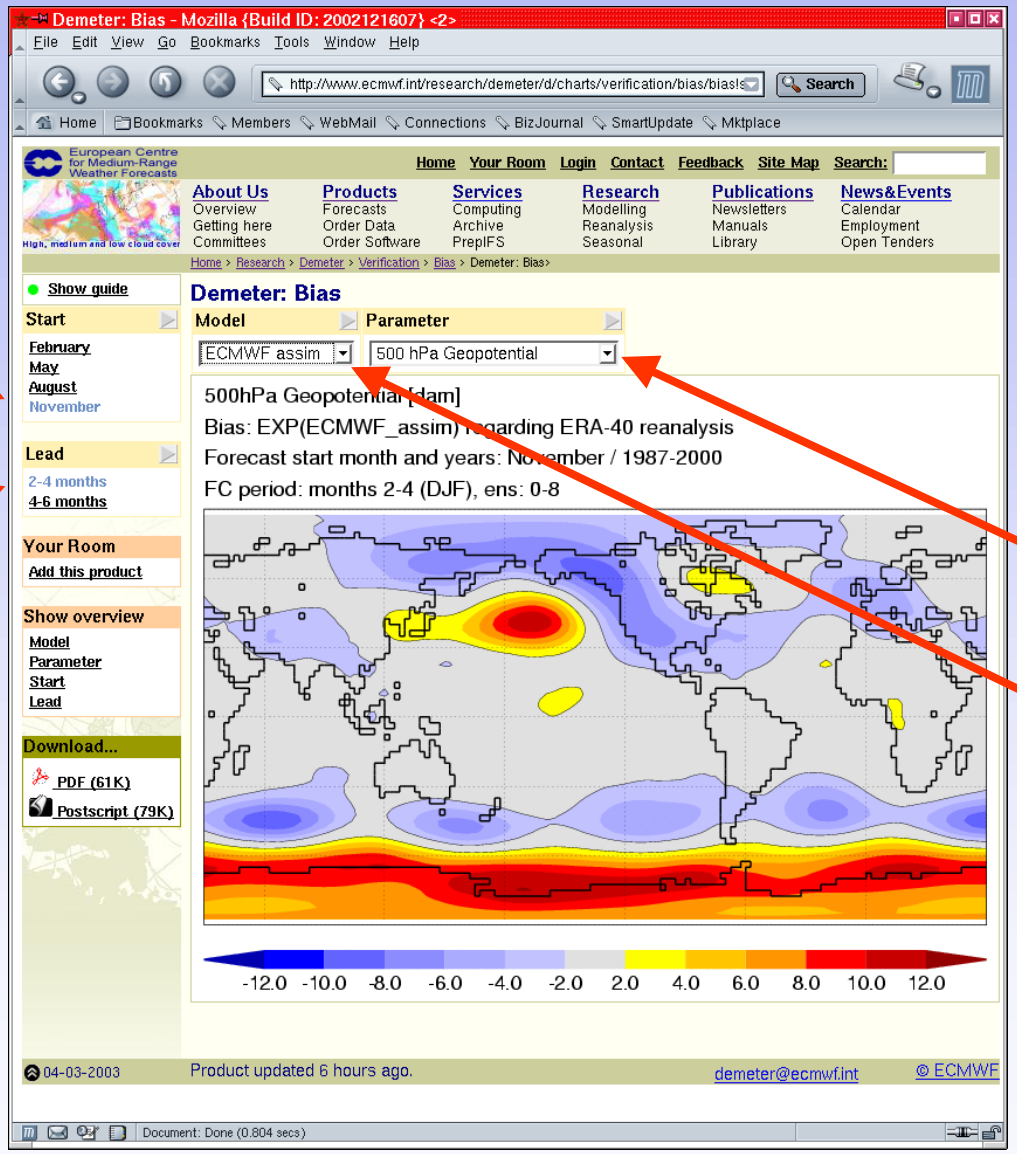
- Bias
- Indices
- Deterministic Scores
- Probabilistic Scores
- Single vs. multi-model
- MM vs. grande ensemble
- Ocean diagnostics



Verification Suite

Start date

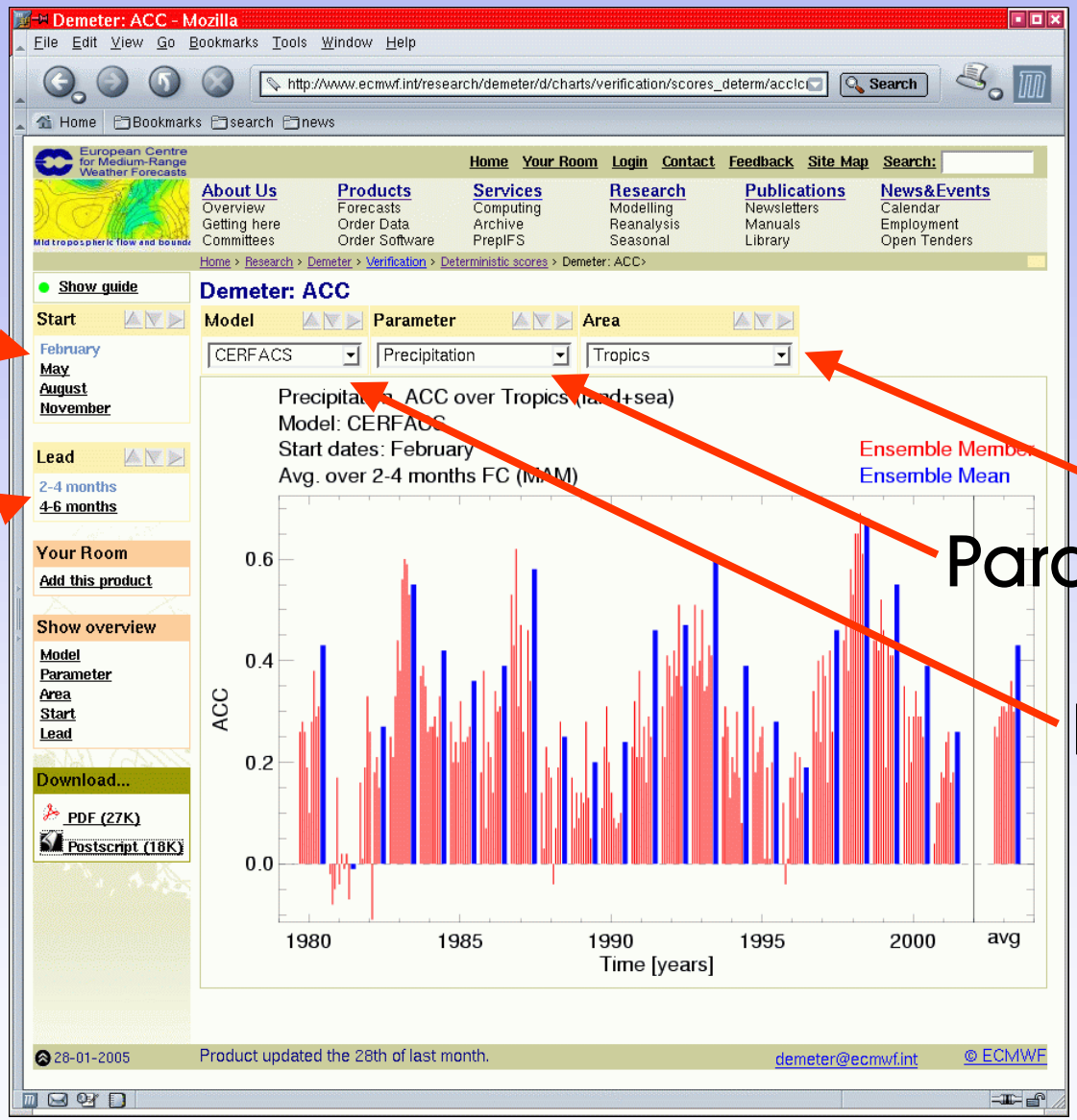
Lead time



Parameter

Model

Verification Suite



Start date

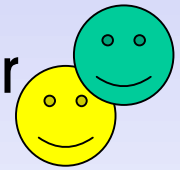
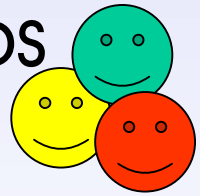
Lead time

Area

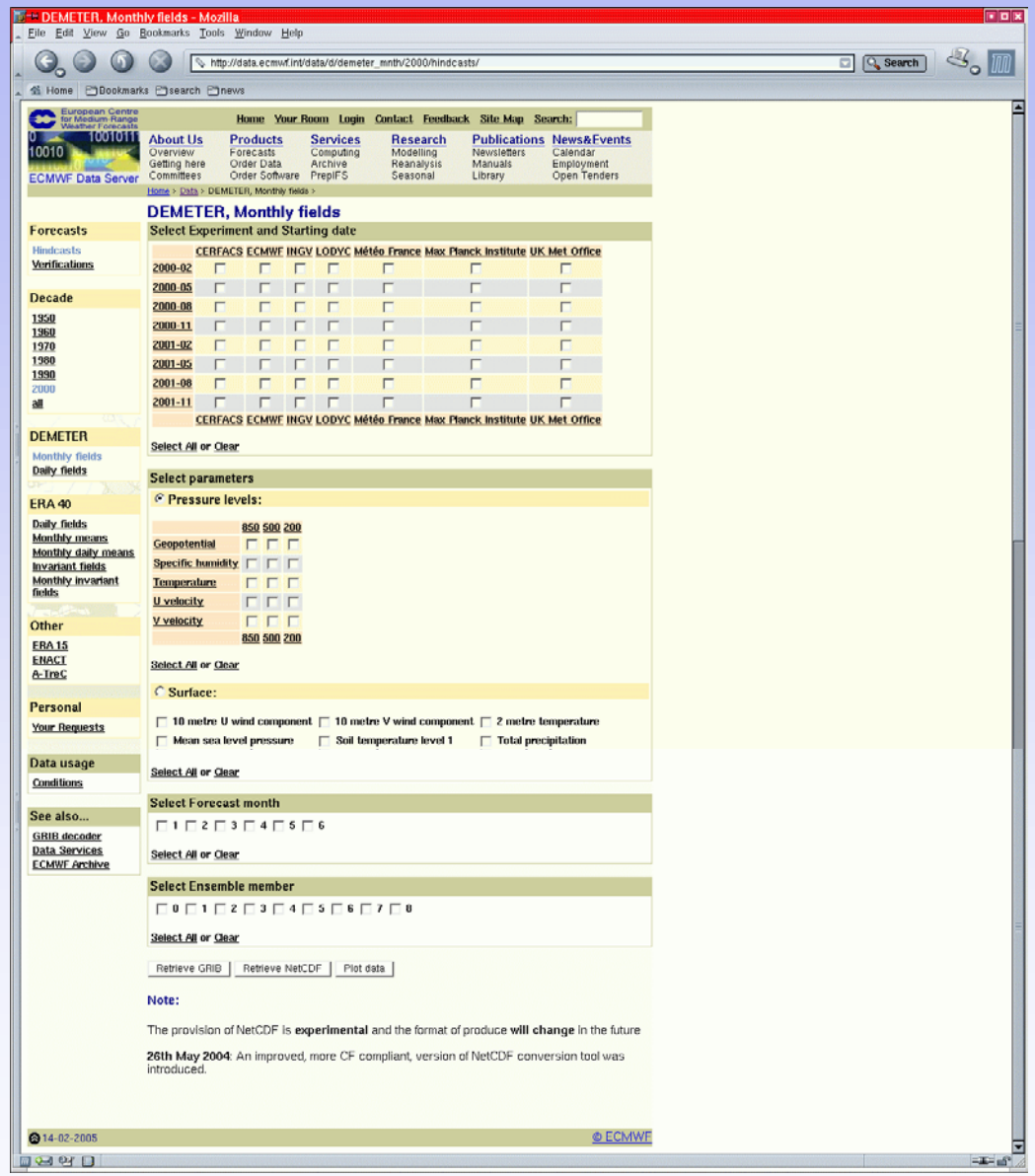
Parameter

Model

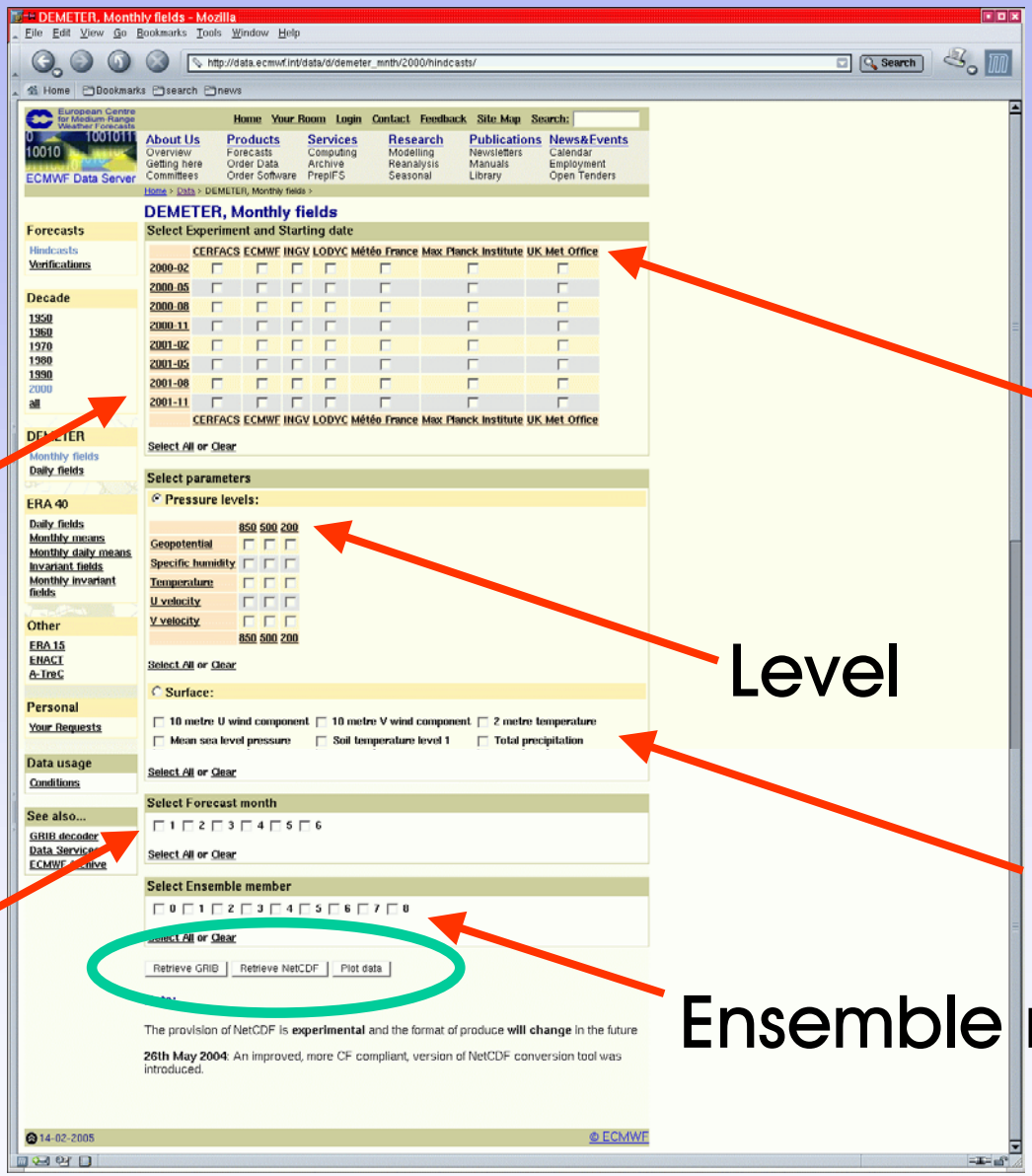
Where to archive

- Past/current experience from other multi-model projects demonstrates importance of common data base
  - PROVOST: central archive 
  - DEMETER: central “QC” archive + public data server 
  - ENSEMBLES: central archive(s) + public server + DODS 

Public data server



Public data server



Data consumers want…

- Use the data as input:
  - to produce calibrated / combined new data set
  - for statistical / dynamical downscaling
  - for application models (decision-making tools)
- Fast, easy and free access
- Quality controlled data
- NetCDF or even ascii data – additionally to GRIB
- Applications installed at archive centre

Definitions for TIGGE archive

What and How? Where and When?

- Contributing models (BMRC, CPTEC, ECMWF,…)
- Number of ensemble members (minimum requirement?)
- Resolution (minimum requirement?)
- Length of forecasts
- Number of forecast starts per day
- Which fields, on which levels, in what units, at what frequency
- Data format
- “Dual archive” concept during research phase of TIGGE?

The Questionnaire

- Some background information and 8 specific questions
  - Concept of data exchange
  - Content of archive
  - Tools and applications
- Response from 12 centres
- Compiled paper version for WG-4 (or any other interested)

User requirements - access

- Do we want to embark on “dual archive” strategy,
  i.e. quick access / exchange archive plus extended archive?

Response:

- No-one is keen on tapes!
- For research purposes no “quick” access necessary
- For operational demonstration projects “quick” exchange has to be realised for “weather parameters”
- To reduce data transfer: tailored regional subsets

User requirements - archive

- A TIGGE archive priority list?

Response:

top, medium, low priority

	ECMWF	SAWS	MetNO	CPTEC	UCI	JMA	Met Office
ens	50			all	all		
res	T255			original	original		
sfc	12 par			u,v,q,T, p, RR	6 var	5 par	prime interest
p-level	5 par 5 levels			5 par 5 levels		5 par 7 levels	
pv-level	1 par 1 level						
o-freq.	6h	sh-r FC	3h / 6h	6h /12h	1h / 24h	6 sfc, 12	
base-time	00z, 12z			00, 12z		1 per d	
m-level	none	LAM on p-level	for LAM's				

Fields - Levels – Units - Frequency

Parameter	Abbrev	Level	Unit	Output frequ.	Comments
Mean sea level pressure	MSL	surface	Pa	6h	inst
10m U-velocity	10U	10m	m s**-1	6h	inst
10m V-velocity	10V	10m	m s**-1	6h	inst
2m temperature	2T	2m	K	6h	inst
2m max temperature	MX2T	2m	K	6h	det_lo
2m min temperature	MN2T	2m	K	6h	det_lo
2m dew point	2D	2m	K	6h	inst
Total precipitation (liquid+frozen)	TP	surface	m	6h	acc_st
Total cloud cover	TCC	surface	0-1	6h	inst
Snow fall	SF	surface	m of water equivalent	6h	acc_st
Wind gust at 10m	10FG	10m	m s**-1	6h	det_lo
Land/sea mask	LSM	surface	0-1	Once	inst

Parameter	Abbrev	Level	Unit	Output frequ.	Comments
Temperature	T	L5	K	6h	inst
Geopotential	G	L5	m**2 s**-2	6h	inst
U-velocity	U	L5	m s**-1	6h	inst
V-velocity	V	L5	m s**-1	6h	inst
Specific Humidity	Q	L5	kg kg**-1	6h	inst

DRAFT: WG-4 has to work on this!

User requirements - archive

- Typical requests to the archive will be

  Response:
  - regional fields
  - local timeseries (e.g. LOC_PAR_LEAD_MM_2006-2008)
  - global fields

- The data format hit-list

  Response:
  - GRIB
  - NetCDF, in particular for local timeseries
  - ascii

User requirements - tools

- Priority list for tools and applications

  Response:

  - Tools for reading / converting data
  - Basic quality control / validation
  - Visualisation
  - Statistical tools
  - Interpolation
  - Basic data postprocessing (ens mean, spread,…)
  - Advanced toolbox for calibration / combination of models

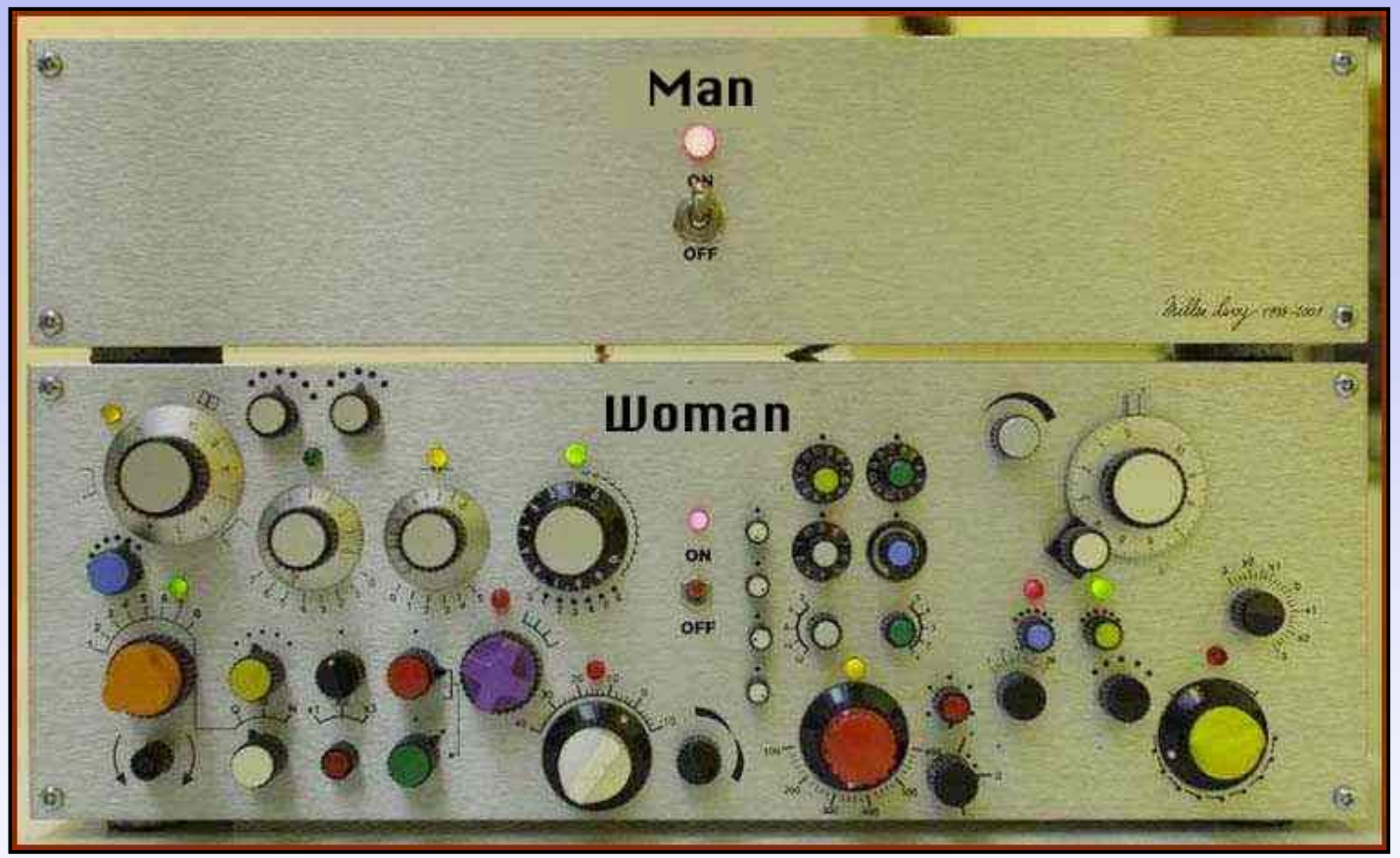
- Running user applications at the archive centre?

  Response: Yes!

Summary

- Be very precise in technical definition of archive content
- Define operational vs. research requirements
- Content of archive “more or less” undisputed, but…
  agreement on output frequency and how to deal with LAM
- Provide tools for “end user” needs

How can we avoid an overload of information…



…which – it seems - not everyone can deal with?!