ECMWF Data Services

Baudouin Raoult, Data and Services Section, ECMWF

Data and Services

Dissemination	Products catalogue
Real time	
Archive	
MARS	Data Services

Data and Services Section

- Providing access to the Centre's data
  - To internal users
  - To the member states
  - To the research community
  - To the general public

- Managing the Centre's catalogues:
  - Real-time
  - Dissemination
  - Archive
  - Software

Background

- Services
  - Cater for different needs
  - Have evolved independently
  - Were designed by different persons
  - Often have a different understanding of the data

- Our goals
  - Unify vocabulary
  - Unify catalogues
  - Unify request language
  - Unify interfaces

Web based Data Services

- Web MARS
  - Access to the archive for registered users
- Data Server
  - Public access to a selected datasets for research purposes
- Data Services
  - Sale of archived data
  - Public access
- Costing of products from the real-time catalogue
  - Restricted access
- Dissemination (in development)
  - Edition of dissemination requirements, restricted access

Web MARS – What is it?

- Web interface to MARS
  - Monitoring
  - Catalogue browsing
  - Retrieval
  - Plot
  - Data finder
  - Changes in the archive
- Targeted at internal and Member State users

Web MARS – Issues

- User identification
  - Access controls
- Large catalogue (metadata)
  - Consistent Navigation

- Large amount of retrieved data
  - Very long transactions
  - Large data transfers
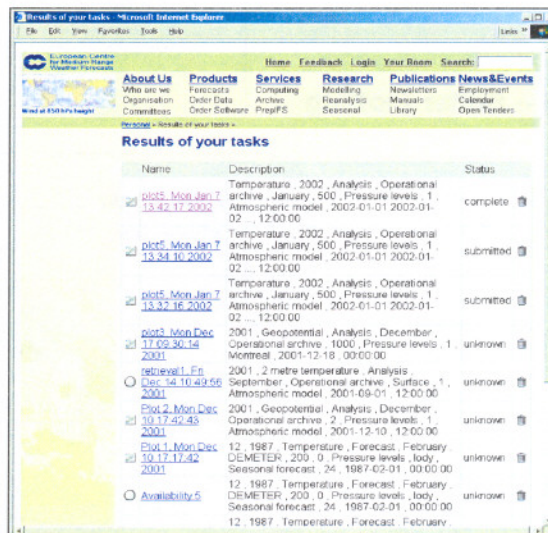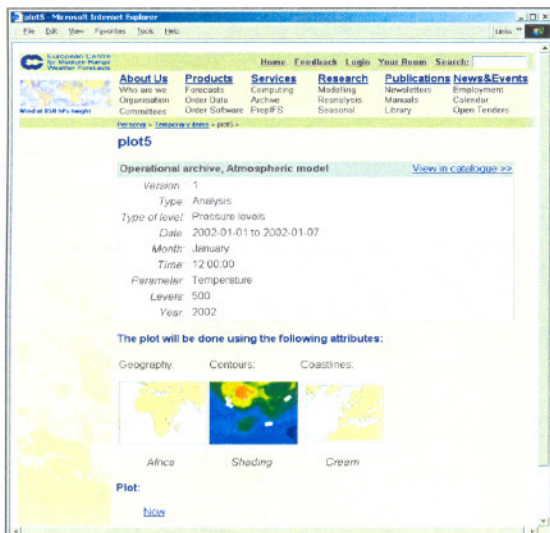- Garbage collection
  - Local disk management
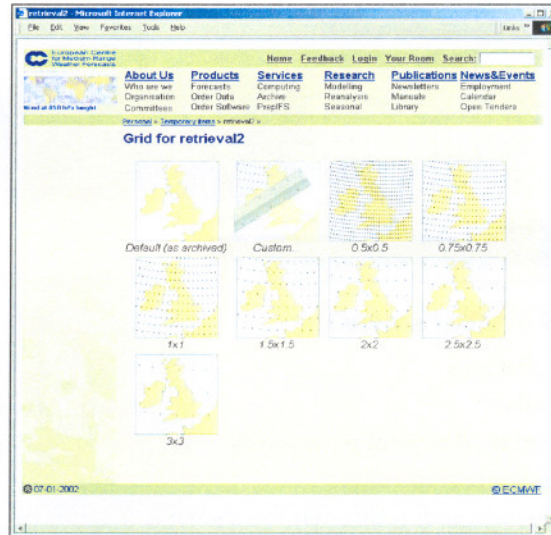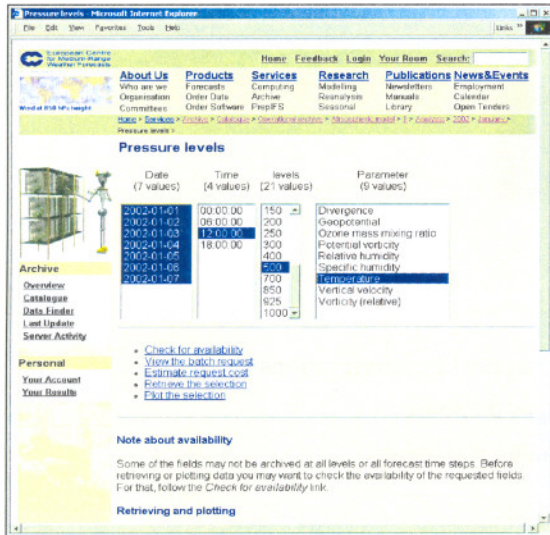
Web MARS – Catalogue

- Based on metadata
- Tree hierarchy
- Selecting fields
- Availability
- Costs
- Retrieval
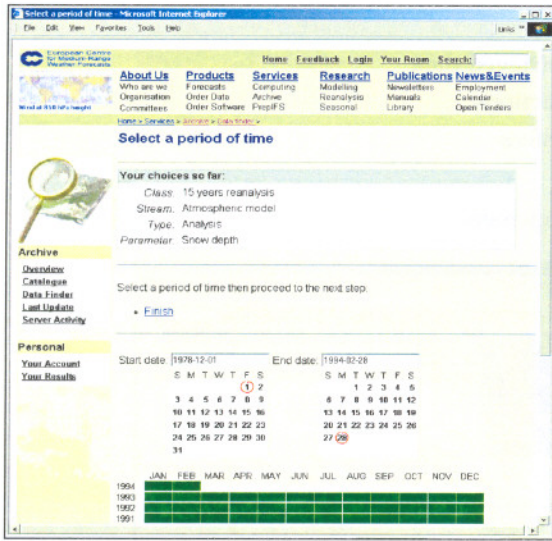- Plot Location Other Choices

Web MARS – Retrieval

- Settings
- Area
- Grid
- Contouring
- Projection
- Persistence
- User settings are saved Attributes Request

Web MARS – Plot

Web MARS – Tasks

- Asynchronous
- Supervised
- Persistence

Web MARS – Data Finder

- Task oriented
- Where can I find cloud fields?
- What period is covered by the 40 years reanalysis?
- When did ECMWF start to forecast ozone?
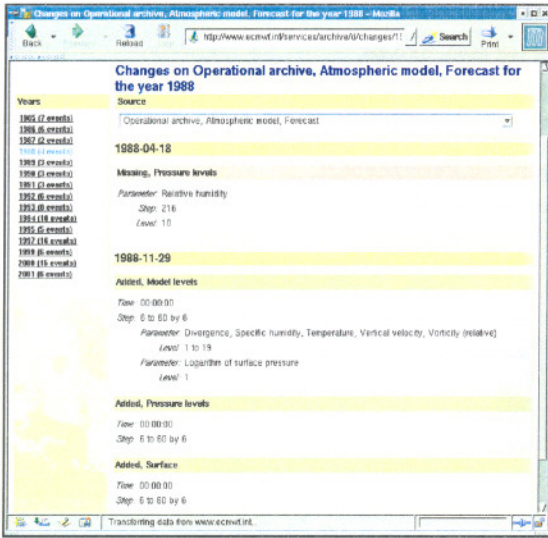
Web MARS – What's new?

- Better plotting
  - Winds
  - Overlay of several parameters
  - Bases on set of rules
  - New navigation through the frames

- Basic support for observations
- Changes database
- Parameter database
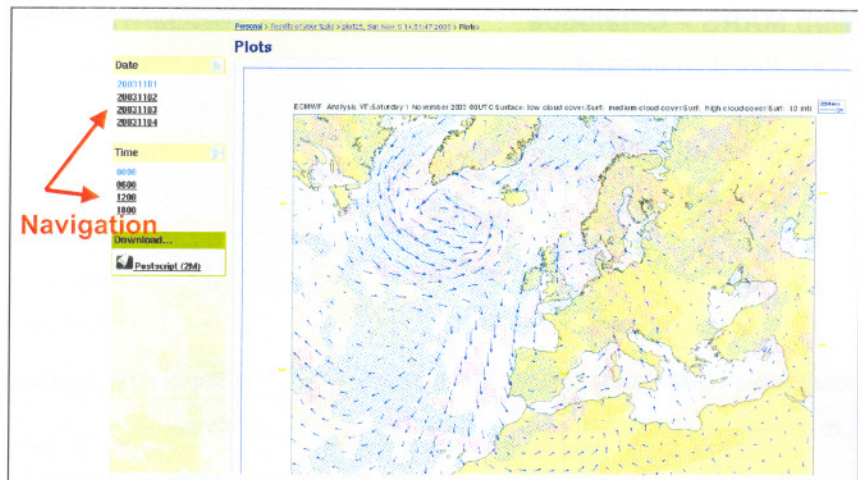
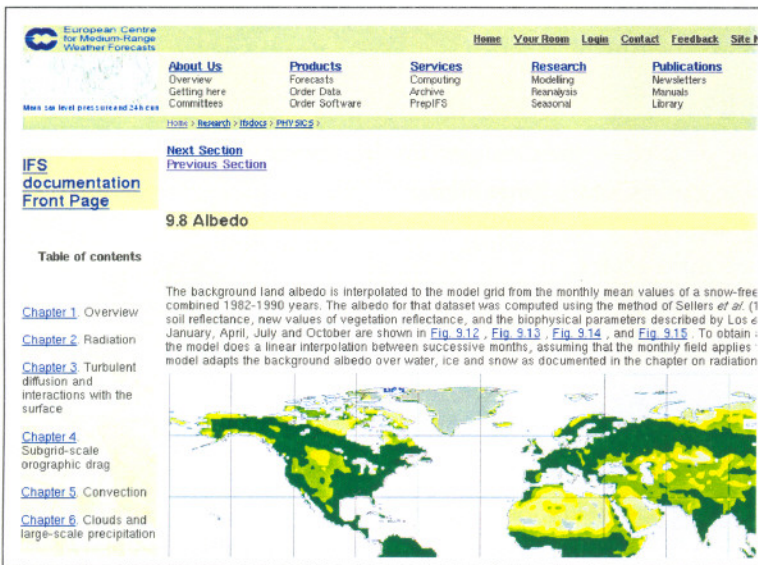
Web MARS – Better Plotting
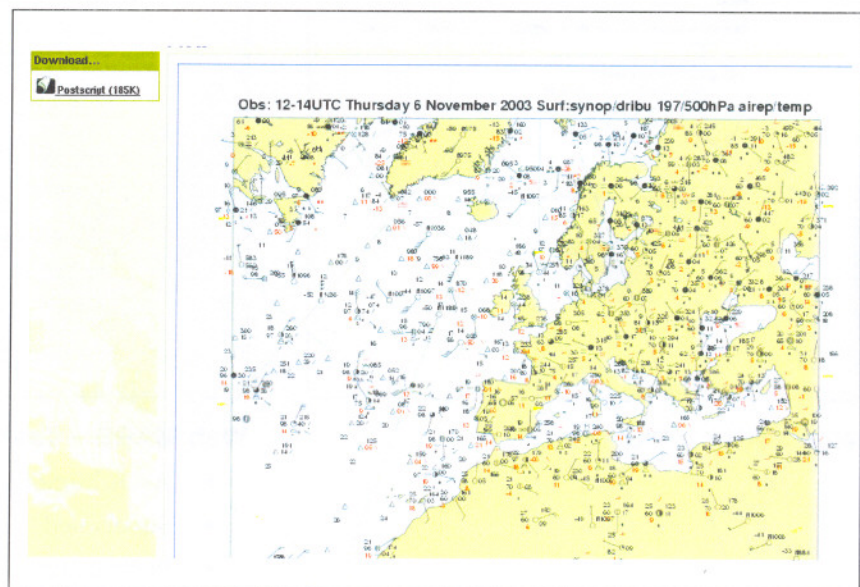
Web MARS – Observation support

Web MARS – Changes in the archive

- Archive history (events)
- Added data
- Discontinued data
- Missing data
- Scan all server metadata
- Takes several days to complete
- Updated daily Data source Events Description

Web MARS – Parameter database

Data Services – Issues

- Sale on archived data research and commercial users
- Large catalogue, many datasets
- Many changes in the archives
  - Resolutions
  - Missing data
  - Additions of new parameters, models, . . .

- Complex costing algorithm
  - Based on volume, number of fields, interpolation,. . .
- Time consuming for Data Services staff
  - Answer 'What if?' scenarios
  - Several interactions between Data Services and client

On-line Data Services – What is it?

- On-line ordering system
- Allow users iterate through:
  - Find the data
  - Apply post-processing
  - Get a quote
  - Refine

- Components of Web MARS:
  - Changes in the archive
  - Own Data Finder
- Targeted at external users (extensive help)

Data Server – What is it?

- Public (non-commercial) distribution of data
  - Self-registration
- Datasets
  - DEMETER, 24 GBytes
  - ERA-15, 1 GByte
  - ERA-40, 400 GBytes, 2.5 degrees
- Based on Web MARS
  - Disk-only MARS server
  - MARS client, Metview, SMS
  - fastCGI, Perl, MySQL
- A standalone PC-Linux system outside the firewall

Data Server – Conclusion

- User interface
  - different for each dataset
  - built dynamically from MySQL contents
- NetCDF is produced on-the-fly from GRIB Experimental
- Very popular since ERA-40 made available
  - About 1000 users
  - Download about 1 TB/month
- Useful tool for sharing data among project collaborators

Costing of real-time products – What is it?

- Provides a centralised costing facility to be used by the *Catalogue Contact Points.*
- Embody the rules decided by Council (e.g. mandatory times steps)
- Avoid misunderstandings on data availability
- Avoid costing discrepancies between Member States
- Allow browsing and searching real-time catalogues easily
- Provide extensive help and up-to-date documentation
- Allow *'What if?'* scenarios

WebAPPS Framework

- Apache, fastCGI
- Perl
  - Object Oriented
  - Rapid prototyping/development
- User identification
  - Certificates, Internet domain, Cookies
  - Access control
- Persistence
  - Very long transactions
  - Keep track of user requests, results, preferences, ...
- MySQL
  - Different viewsof (meta)data

WebAPPS Framework: Software components

Web MARS	Data Services	Real-time costing	...
WebAPPS Framework			
PERL			
Metview		CGI	SQL
MARS	MAGICS	SMS	ImageMagick
EMOSLIB			Apache

What needs to be done?

- Next: dissemination
- Interoperability
  - Define a request for dissemination, get it directly with MARS
  - Cost the requirements for a client, add it to dissemination requirements
  - Create a plot in Web MARS, add it to 'Your room'
- Develop further the parameter database
- More support for observations
- More controls of graphical attributes

Conclusions

- WebAPPS framework
  - Re-use of components, different applications
  - View of a very complex system in a simple way
  - Provide tools for consistency and interoperability
- A few URLs:
  - ECMWF: http://www.ecmwf.int/
  - Web MARS: http://www.ecmwf.int/services/archive
  - Data Server: http://data.ecmwf.int/data
  - Data Services: http://www.ecmwf.int/products/data