MARS - Webinar

Webinar Starts at (10:00am UK time)

**PRESENTER:
Cristian Simarro**

Please remember to mute your microphone and camera ☺

**HOST:
Carsten Maass**

ECMWF

MARS – Webinar

Efficient retrievals and new interpolation software

Cristian Simarro and Carsten Maass
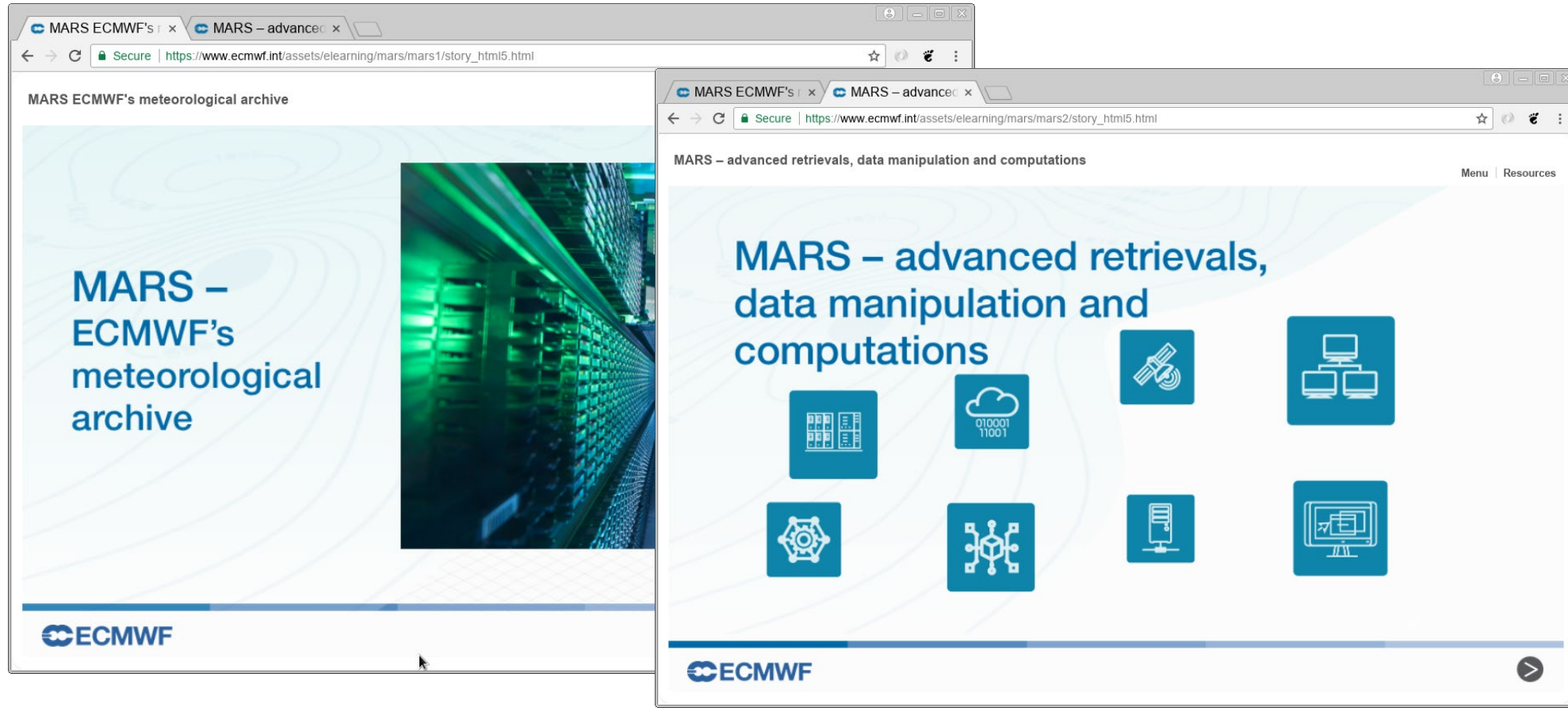
User Support

ECMWF

Content of this Webinar

- MARS Introduction

- How to retrieve large datasets efficiently

- MIR: the Meteorological Interpolation and Regridding library

MARS eLearning modules

https://www.ecmwf.int/en/learning/education-material/elearning-online-resources

Meteorological Archival and *Retrieval* System

- Meteorological data (GRIB: fields, BUFR, ODB: observations)
- Large amount of data (size of archive & number of fields)
- Operational & Research environment
- Complex batch system
- Large number of users with different requirements:

 large datasets rarely ⟷ few fields very often

Introduction – Some figures

- Users
  - +100,000 registered users
  - ~1,000 daily users
- ~1.8 million requests per day:
  - 300 million **retrieved** fields (200 Terabytes)
  - 250 million **archived** fields (230 Terabytes)
- Total volume in tapes 200 PiB in 27,000 tapes
  - In the last 2 years, ~90 PiB have been deleted
- Total volume in disk 5 PiB
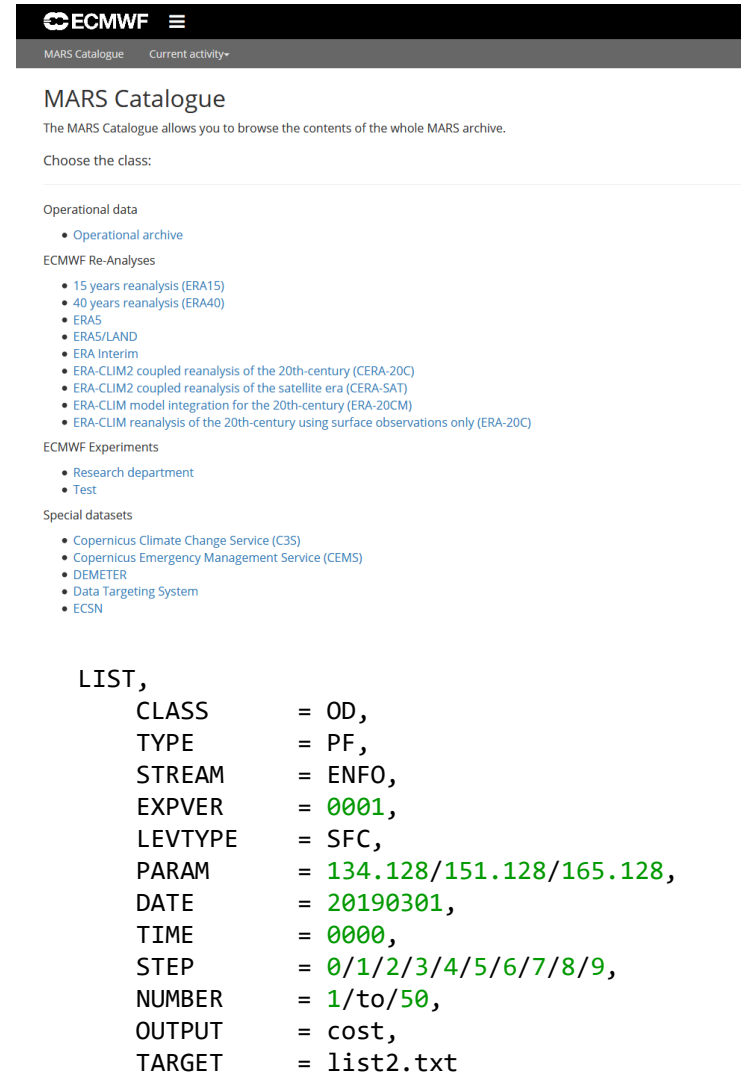- More than $380 * 10^9$ meteorological fields
- Metadata size: 2.5 Tbytes

Introduction – MARS components

- Client/Server architecture

- Clients: workstations, supercomputers, Web API requests …

- 6 instances of MARS servers

- ~50 linux CentOS 7.4

- Tape management software: HPSS 7.5.3
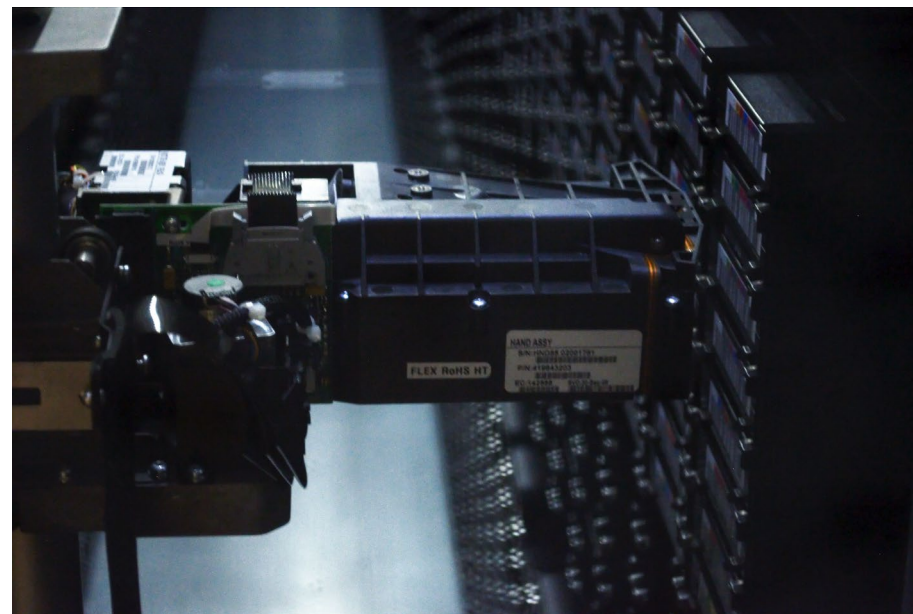
Why do MARS requests need to be efficient?

- ## The resources are shared among many users

- ## The number of tapes drives is limited

- ## Mounting the same tape many times can damage it

- ## Waiting times in the queues might be long

- ## Tools to help creating efficient MARS requests:

  - MARS Catalogue (MARS tree)

  - List verb
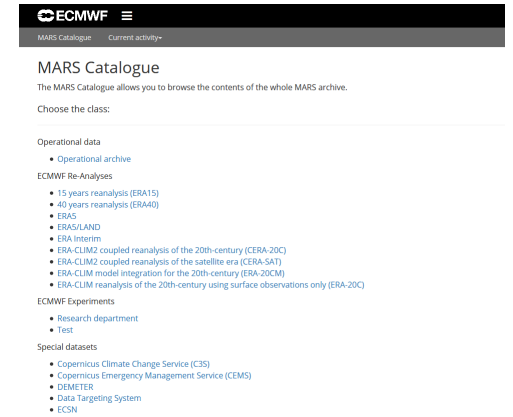
- ## Tool to monitor the activity

  - MARS Activity


```
LIST,
     CLASS       = OD,
     TYPE        = PF,
     STREAM      = ENFO,
     EXPVER      = 0001,
     LEVTYPE     = SFC,
     PARAM       = 134.128/151.128/165.128,
     DATE        = 20190301,
     TIME        = 0000,
     STEP        = 0/1/2/3/4/5/6/7/8/9,
     NUMBER      = 1/to/50,
     OUTPUT      = cost,
     TARGET      = list2.txt
```

MARS Catalogue – http://apps.ecmwf.int/mars-catalogue/

- Content browsing of *every* field in the archive

    - more up to date than static content documentation

- URL based on MARS requests (can be edited & bookmarked)

- Create MARS requests (without checking availability)

- Check availability of data

- Check estimated size of the retrieval

- Retrieval in GRIB and NetCDF for few fields (to so some tests)

Archive license	https://apps.ecmwf.int/archive-catalogue/
Public users	http://apps.ecmwf.int/datasets/

Retrieving data - MARS tree
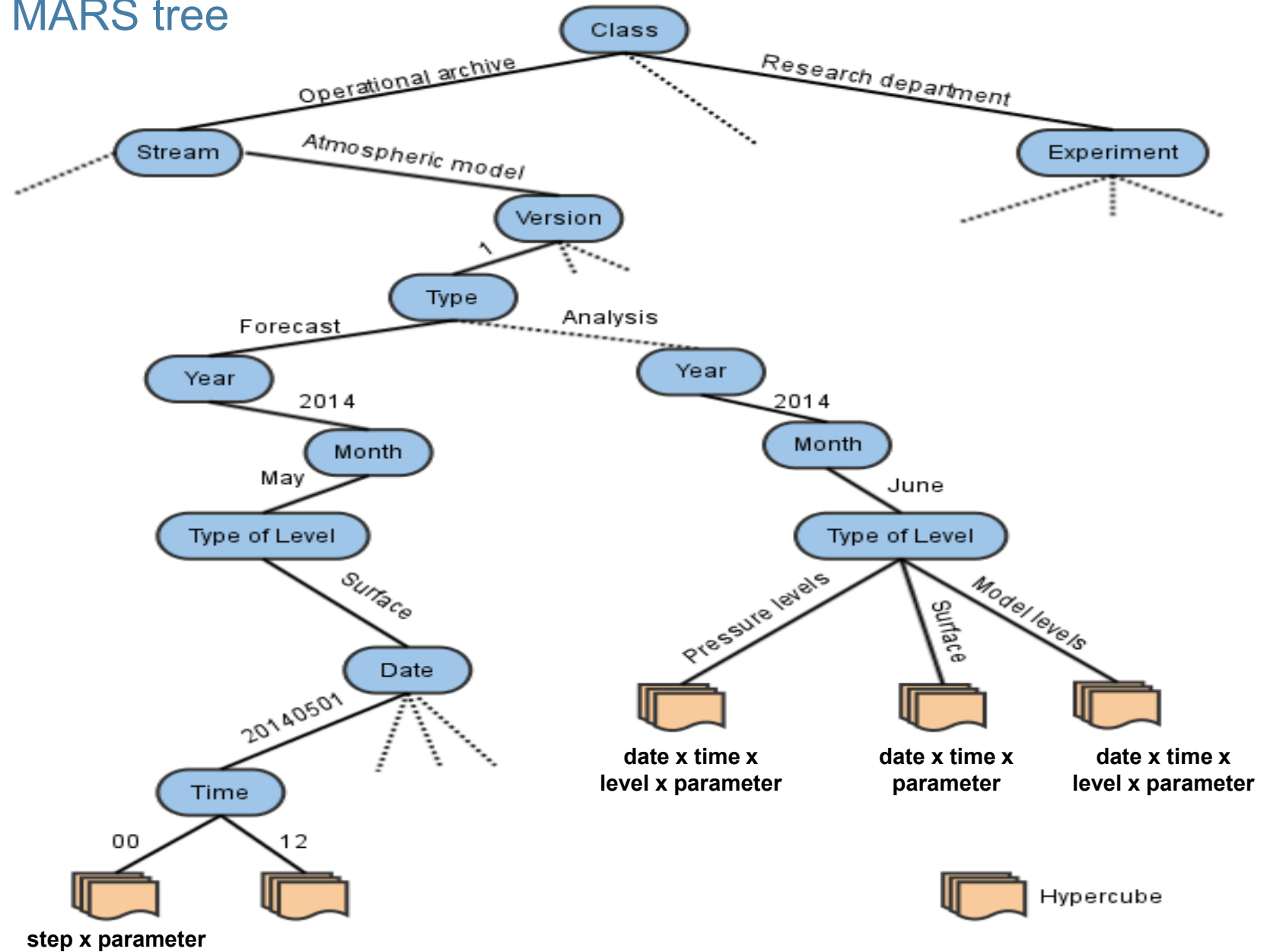
MARS activity – http://apps.ecmwf.int/mars-activity/

Server activity / MARS queue

- Show system activity

- Monitor your requests

- Learn how the queuing system works

  - Reason for queued requests

MARS activity – http://apps.ecmwf.int/mars-activity/

Retrieving data – Efficiency

- Explore data in archive catalogue - collocation

- Estimate amount of data (**list** command)

  - Number of fields (up to tens of thousands / request)

  - Data size (up to several Gigabytes / request)

- **Retrieve as much data from the same tape as possible**

  - Avoid constantly accessing the same tape

- Reduce number of tapes involved (better scheduling)

- Use local target disk (e.g. $SCRATCH for MS users)

MARS Efficient requests use case

- Some users want to retrieve large chunks of data
  - ENS: Atmospheric (enfo) Perturbed Forecast surface
  - HRES: Atmospheric (oper) Model Levels forecast
  - long periods of time
- Depending on how a user retrieves the data
  - Faster extractions
  - Performance of the system

Use case example:

https://software.ecmwf.int/wiki/display/UDOC/Guidelines+to+write+efficient+MARS+requests

Questions about MARS efficiency

Contact:
ServiceDesk@ecmwf.int

... and answers

MIR: the Meteorological Interpolation and Regridding library

- Implementation plan

  - Used in MARS, Metview & product generation (dissemination)

- New features

- Differences between MARS / MIR and MARS / EMOSLIB

  - Interpolation method

  - Parameter classification

  - Spectral to grid transformation

  - Sub-areas

  - Land-sea mask

  - Wave data

Implementation plan: MARS

Beta testing of MARS / MIR available as "mars –m"

MARS / MIR available as "mars –m"

MARS/ MIR as default

MARS / EMOSLIB default

MARS / EMOSLIB available as non-default

We are here

18 October 2017
Availability of beta version of MARS / MIR announced to Comp Reps

6 February 2018
Announced to all users

July 2018
First full release of MIR
- MARS / MIR "ready for production"
- MARS / EMOSLIB still the default MARS

26 January 2019
- **MARS / MIR becomes the default version of MARS**
- **Metview uses MIR by default**

December 2019
MARS / EMOSLIB decommissioned

Dates reflect current expectations and may change in light of actual progress made

New features

- No Land-Sea Mask processing

- Grid to grid interpolation

  - Interpolations from

    - global regular or reduced Gaussian or lat/lon grid to regular or reduced Gaussian or lat/lon grid is supported

- MIR uses an intermediate grid when transforming from Spherical Harmonics to lat-lon grids

  - It can be changed using MARS keywords

- Faster spectral transform (2-3 times)

Interpolation methods

- ## MIR uses by default a linear method (based on a triangular mesh)

  - replaces the EMOSLIB bilinear method (based on 4 points)

- ## MIR computes distance in 3D

  - nearest neighbour method has improve as a result

2 metre temperature @ ground or water surface

225   277   329
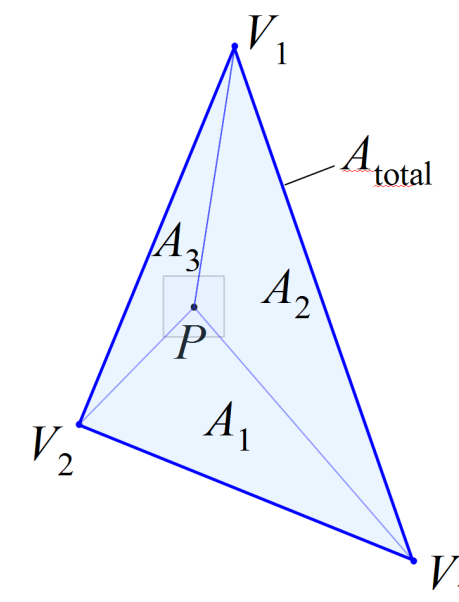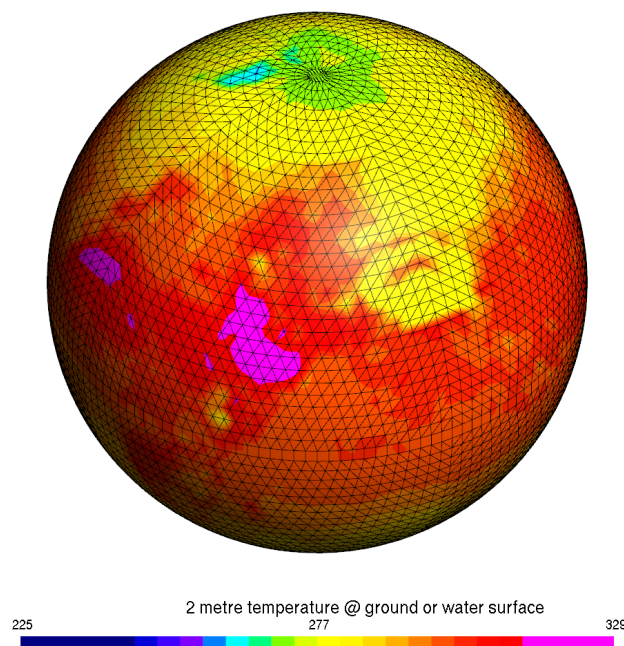
Parameter classification

The parameters have been classified into groups with an associated default interpolation method

- These parameters will be interpolated using nearest neighbour:

  - index parameters: integer values associated to a predefined table

  - wave spectra and associated parameters: intensity as a function of discrete wavelengths

  - vegetation cover ratio parameters

ECMWF

Spectral to grid point transformations (1/3)

- The default transformation with MIR is via an intermediate regular Gaussian grid.

- The post-processing keyword `RESOL` is replaced by two new MARS keywords

  - **TRUNCATION**

    - Describes how to treat the incoming SH, before the transformation to grid points

    - The default TRUNCATION=AUTO is derived from the intermediate grid

  - **INTGRID**

    - Describes the intermediate grid to which the transform is performed

    - The default INTGRID=AUTO uses an intermediate regular Gaussian grid, derived from the output GRID

Spherical Harmonics
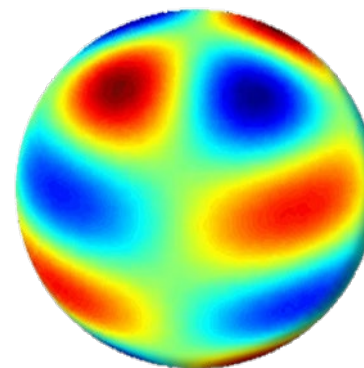
Reduced Gaussian grid O32

ECMWF

Spectral to grid point transformations (2/3)

- Transformation from T1279 to regular lat/lon 1/1

Automatic truncation: via intermediate F90 grid (with an automatic truncation to T179)
This is the **default** behaviour

- `TRUNCATION=AUTO, INTGRID=AUTO, GRID=1/1`

User specified truncation: via intermediate F90 grid (with a user-specified truncation to T179)

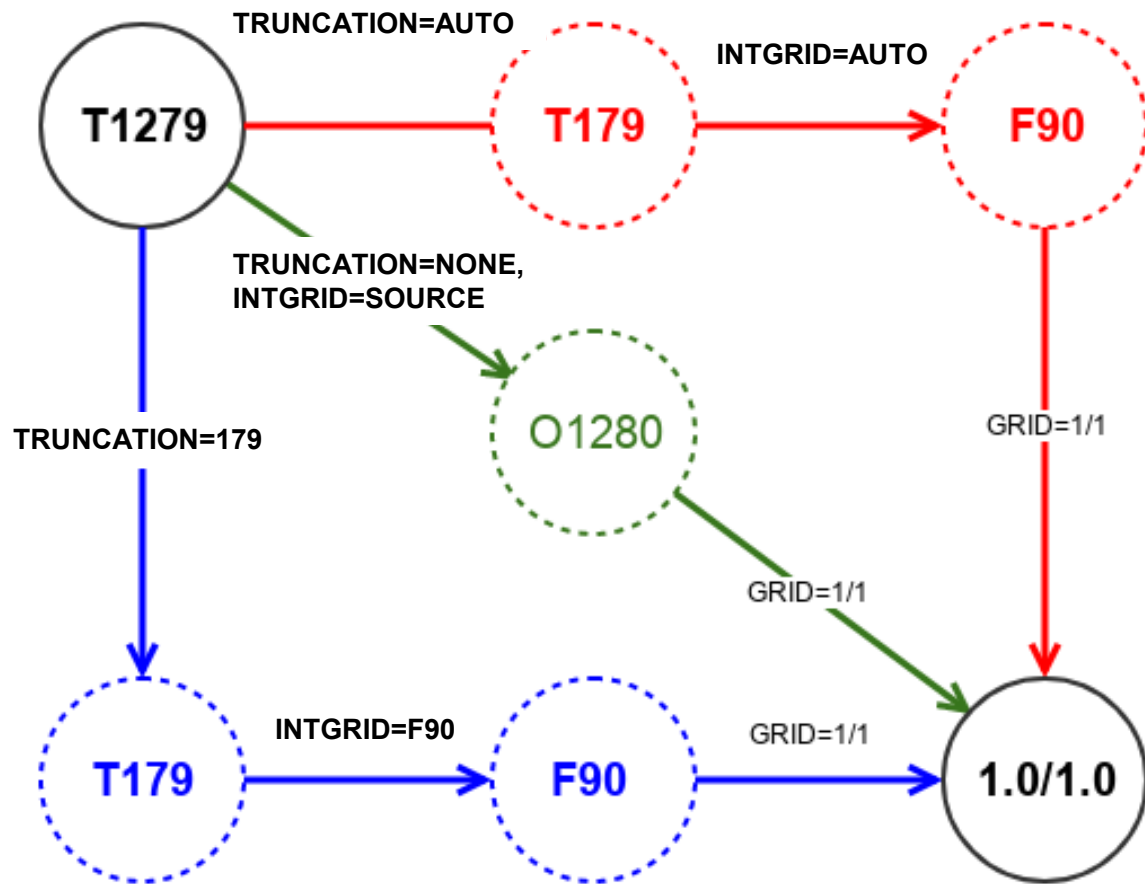- `TRUNCATION=179, INTGRID=F90, GRID=1/1`

No truncation of the SH input field: no truncation of the source, and via intermediate grid based on the source (T1279 corresponds to O1280)
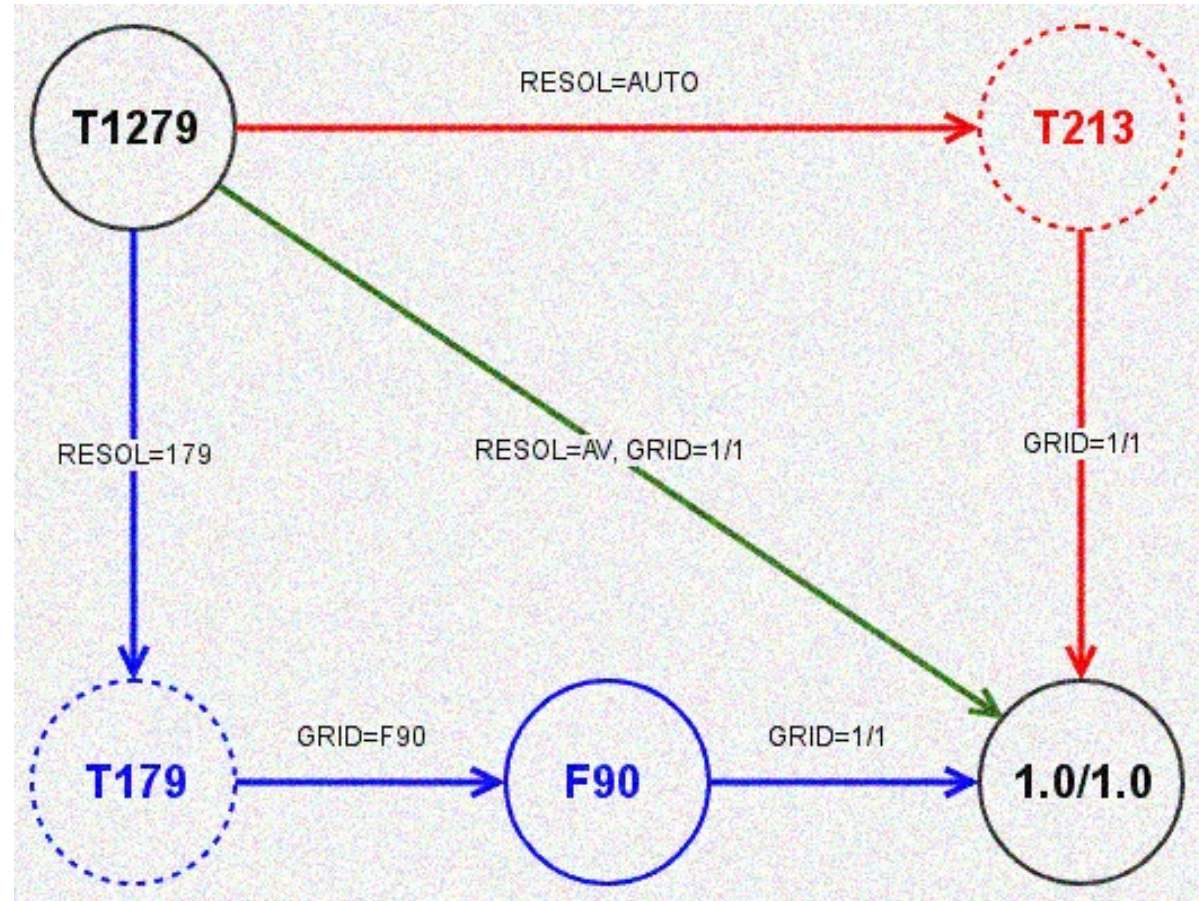
- `TRUNCATION=NONE, INTGRID=SOURCE, GRID=1/1`

Spectral to grid point transformations (3/3)

# Subareas

- Cropping subareas and frames directly from global reduced Gaussian grids, including the octahedral reduced Gaussian grid, is now supported

- MIR can work with staggered grids (no point at latitude=0, longitude=0)

AREA=89.5/0.5/-89.5/359.5,
GRID=1/1

MARS/MIR: honours the user defined area and returns the inner grid

MARS/EMOSLIB: expands the area to fit in the original grid
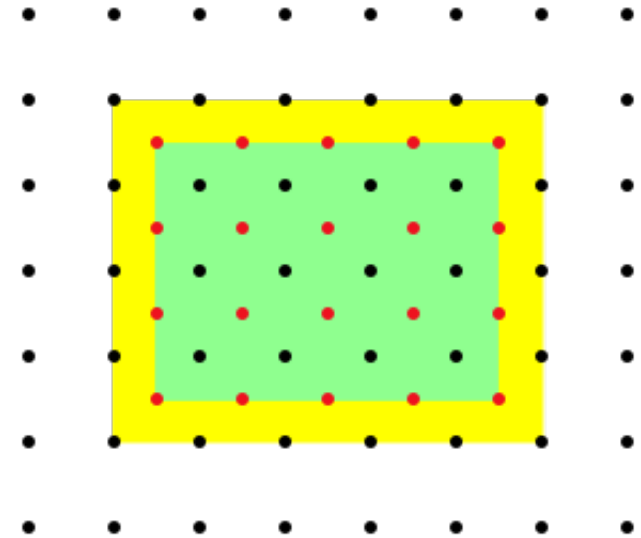
Land-Sea masks

- By default **no** Land-Sea Mask processing

  - LSM processing for both input and output grids can be requested with the new **LSM** keyword

    LSM = ON

  - As usual, **LSM** can be set to value **OFF** in subsequent requests

  - Land-Sea Mask in MIR is based on a much finer 1 km resolution field

- When interpolating the Land-Sea Mask field (paramId= 172) to a regular lat/lon grid, MARS/MIR returns fractional values [0,1]

  lsm ≥ 0.5 : land point

  lsm < 0.5 : sea point

Wave data

- Interpolation from limited-area wave data on a regular lat-lon grid is now supported

- Nearest neighbour points might be different compared to EMOSLIB

- Different treatment of neighbouring missing values produces a smoother interpolation close to coasts

- all "Spectra-related" parameters have been classified and are interpolated in the same way as the 2D spectra, using a nearest neighbour in MIR instead of EMOSLIB's bilinear method

Differences in parameter values between EMOSLIB and MIR

MIR has undergone a thorough validation process and tests to ensure its quality and correctness. You may see some differences when using MARS/MIR compared to MARS/EMOSLIB
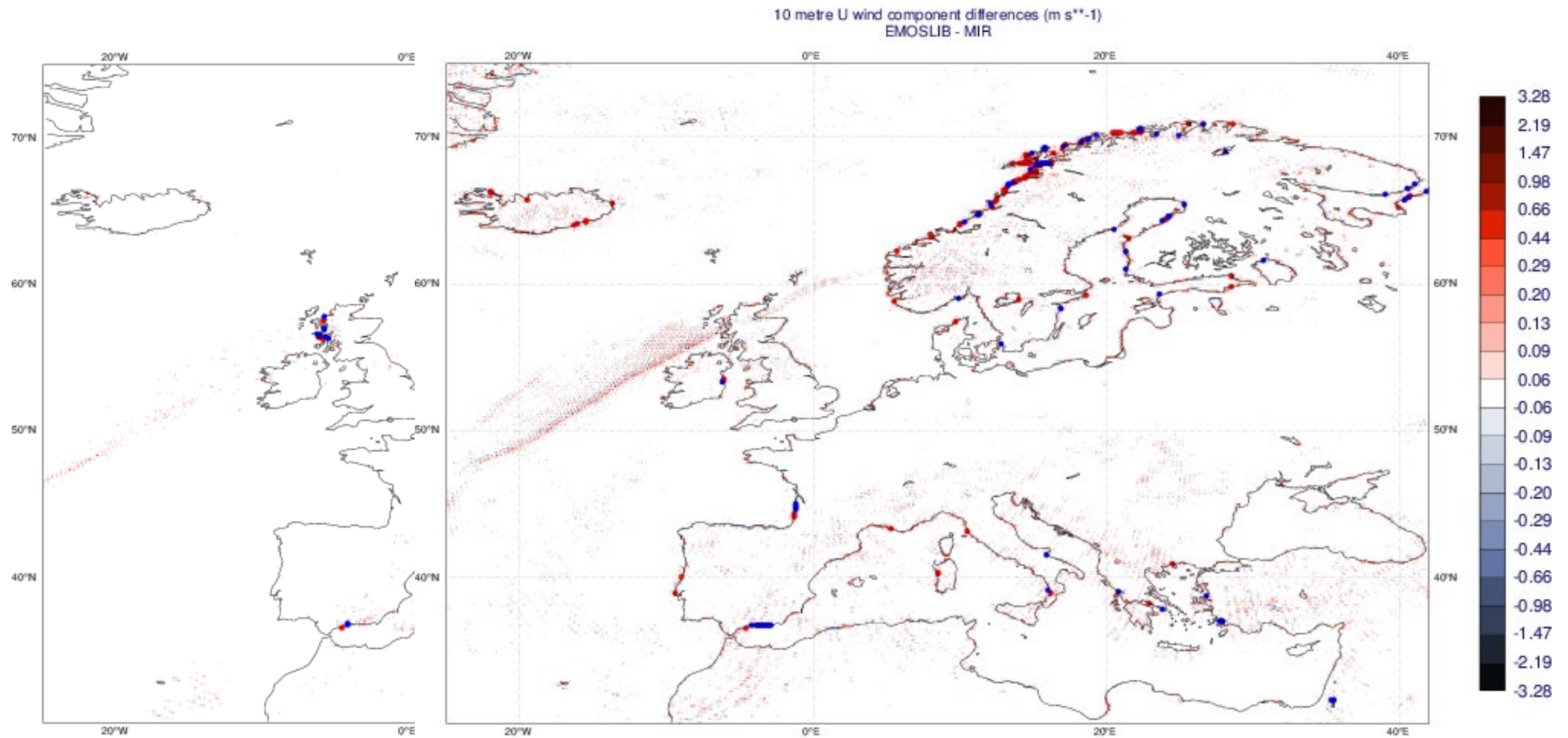
Possible reasons for differences in interpolated fields between MIR and EMOSLIB:

- land sea-mask processing

- default interpolation method

- selection of nearest points

- handling of missing values

For full details see:

https://software.ecmwf.int/wiki/display/UDOC/MARS+interpolation+with+MIR

Examples

10 metre U wind component differences (m s**-1)
EMOSLIB - MIR

MIR article

Contents lists available at ScienceDirect

Computer Physics Communications

Atlas: A library for numerical weather prediction and climate modelling

Willem Deconinck *, Peter Bauer, Michail Diamantakis, Mats Hamrud, Christian Kühnlein, Pedro Maciel, Gianmarco Mengaldo, Tiago Quintino, Baudouin Raoult, Piotr K. Smolarkiewicz, Nils P. Wedi

European Centre for Medium-Range Weather Forecasts (ECMWF), Shinfield Park, Reading RG2 9AX, United Kingdom

ARTICLE INFO

ABSTRACT

The algorithms underlying numerical weather prediction (NWP) and climate models that have been developed in the past few decades face an increasing challenge caused by the paradigm shift imposed by hardware vendors towards more energy-efficient devices. In order to provide a sustainable path to exascale High Performance Computing (HPC), applications become increasingly restricted by energy

https://doi.org/10.1016/j.cpc.2017.07.006

Additional resources

- MARS documentation
  - https://confluence.ecmwf.int/display/UDOC/MARS+user+documentation
- MARS Web Applications
  - http://apps.ecmwf.int/mars-catalogue/
  - https://apps.ecmwf.int/mars-activity/
- Guidelines to write efficient MARS requests
  - https://software.ecmwf.int/wiki/display/UDOC/Guidelines+to+write+efficient+MARS+requests
- eLearning
  - https://www.ecmwf.int/en/learning/education-material/elearning-online-resources
- MARS interpolation with MIR
  - https://software.ecmwf.int/wiki/display/UDOC/MARS+interpolation+with+MIR
- Newsletter article about MIR
  - https://www.ecmwf.int/sites/default/files/elibrary/2017/17439-newsletter-no-152-summer-2017.pdf

Questions …

Contact:
ServiceDesk@ecmwf.int

… and answers