National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

NCAS
Unified Model Introduction

Part 2: UM Data Files

University of Reading, 18-20 March 2015

Contents

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

UM and PP formats

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

There can be confusion about these two types of UM-related files.

UM format

- The format read and written by the model.
- Direct access files that consists of a primary header containing pointers to a series of secondary headers that then point to the data:
  - Implies mixed integer/real arrays conventionally handled with equivalence
  - Assumes that levels for a given variable are contiguous
- Described in document F3:
  http://cms.ncas.ac.uk/wiki/Docs/MetOfficeDocs

PP format

- Produced from UM files for data analysis (**not** used by the model).
- Sequential files (header, data, header, data).
- Files from the BADC and UKMO archive may be in this format.

UM files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Files in	Files out
Initial files (also called start files or dumps)	Restart files (or dumps)
Lateral boundary conditions (LBCs)	Lateral boundary conditions (LBCs)
Ancillary files - forcing - boundary - control	Post processing diagnostic files (also called fields files) Climate means

All model files are in UM format, but there are slight variations between restart files, ancillaries and fields files (see F3).

(CF-)NetCDF

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Network Common Data Form:

- Portable and self-describing data format.
- Various supported and external libraries and tools.
- Widely used and a standard format for many scientific communities.
- http://www.unidata.ucar.edu/software/netcdf/

Climate and Forecast conventions:

- Set of standards for describing data (metadata).
- Definitive explanations of data variables (standard names).
- Definitions for temporal and spatial properties of the data.
- http://cfconventions.org/

NetCDF and the UM

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- When you might come across NetCDF:
  - Input data sets may be in NetCDF (e.g. CMIP5 scenarios).
  - It can be more convenient to use NetCDF for analysing or sharing data.
  - NEMO and CICE read and write NetCDF files.

- There are tools for converting between (CF-)NetCDF and UM/PP formats.

- NCAS are working on a project to write CF-NetCDF directly from the UM.

- UM file types
- **Model input**
- Model output
- Limited Area Models
- NEMO and CICE

UM start dumps

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Start dumps are available on the National Computing Service (ARCHER) for standard resolutions under:

  `/work/n02/n02/hum` (`$UMDIR`)

- Restart dumps from other UM jobs can be used.

- For NWP, particular dates are available from the UK Met Office. Contact the NCAS modelling helpdesk:

  http://cms.ncas.ac.uk/wiki/CmsHelpdesk

Start dates

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- For climate runs the start date is largely irrelevant
- **But for NWP experiments or case studies, the start date matters**

- The UKMO archives only the last ~18 months of analyses which are used as start dumps. To start the UM from other dates, you need to use ECMWF data and convert this to a UM start dump.

- Starting from ECMWF data is available on the national HPC services.
  - For advice on starting from ECMWF data contact the helpdesk.
  - Example reconfiguration jobs are available under the “umui” owner.

The UM reconfiguration (i)

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

The reconfiguration is a standalone program which modifies ("reconfigures") UM atmosphere or ocean start files to produce a new start file.

- This can be on **a new grid** (not for ocean files)
  - with different horizontal area
  - and/or different horizontal resolution
  - and/or different vertical resolution
  - if no new data is included, data from the original start file is interpolated to the new model grid/domain;

- It can also include **different data fields** from ancillary files
  - e.g. new orography data
  - e.g. new ozone data

The UM reconfiguration (ii)

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- The reconfiguration is part of the UM code base but it has its own compilation stage to produce its own executable.
- It runs on multiple processes (the same as or less than the main UM).
- It can be run independently or immediately before a UM model run.

- Use the reconfiguration to:
  - Upgrade a start file from an earlier code version
  - Add fields including user-defined fields
  - Overwrite existing fields
  - Change resolution (not with ocean files)

UM ancillary files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Ancillary fields describe externally prescribed conditions to be imposed on the model fields.
  - These may time varying, be fixed or apply only to the initialisation.
- They may be
  - in the start dump already (standard ancillary fields)
  - replaced in the start dump (modified ancillary fields)
  - or they can be added to the start dump (user ancillary fields)
- The user needs to consider:
  - the horizontal and vertical resolution
  - the time variation
  - the 360 day/Gregorian calendar switch in the file header
- In the UMUI ancillary file names and directories may be specified using environment variable. These are defined in an “ancillary versions file”:

  Atmosphere -> Ancillary and input data files -> In file related options -> Ancillary versions file

Standard ancillary files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Ancillary fields are grouped into files.

Fixed:

- *Atmosphere -> Ancillary and input data files -> Other ancillary files and Lateral Boundary files*
- E.g., orography, land-sea mask, land fraction

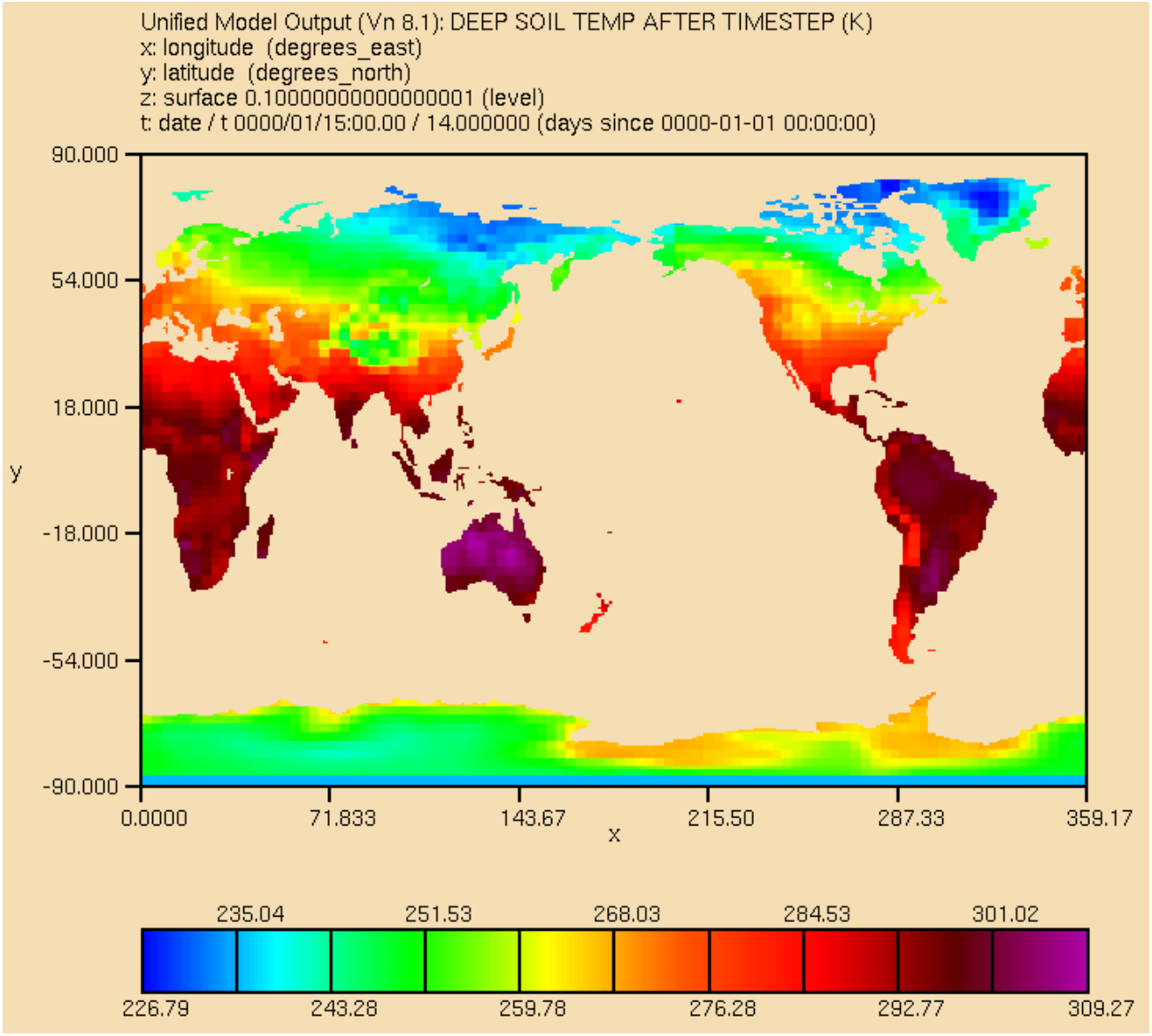
Climatologies:

- *Atmosphere -> Ancillary and input data files -> Climatologies & potential climatologies*
- Standard files are available on the national HPC services:

Ozone	Shine and Li, Cariolle, SPARC
SST / sea ice	GISST, HadISST
Soil fields	AMIP

Example ancillary field

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



Control files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- The SW and LW radiation code use “spectral” files
- Spectral data for the radiation scheme are generated by a separate pre-processing package and are stored in a spectral file.
  - This is not a simple matter and users will not normally generate their own spectral files.
- Standard spectral files are available in the “ctldata/” directory of the UM installation.
  - Different spectral files are available on the national HPC services.
- Spectral files are not in UM format but are Fortran namelists, e.g:

```
&R2SWSP
 L_PRESENT(0)=.TRUE.,
 N_BAND=6,
 N_ABSORB=4,
 N_AEROSOL=13,
 TYPE_ABSORB=1,2,3,7,
 TYPE_AEROSOL=1,2,3,4,6,10,11,12,13,15,16,23,24,
```

- UM file types
- Model input
- **Model output**
- Limited Area Models
- NEMO and CICE

UM restart dumps

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

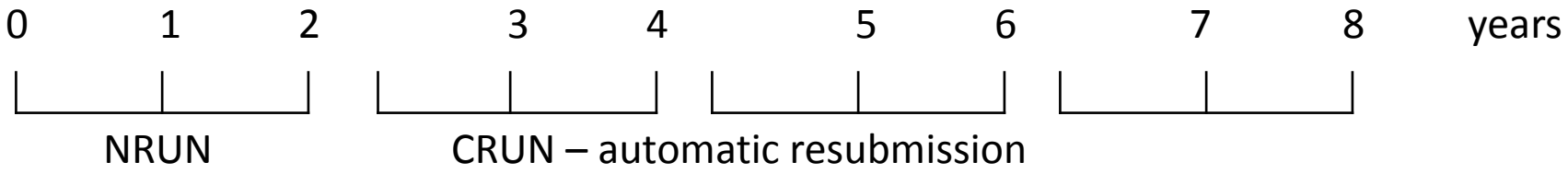
- Instantaneous dumps of the main model variables, written at regular intervals.
  - To restart the model for when things go wrong so that the run can be continued rather than starting again.
  - Start files for running long runs in ‘chunks’
  - The basis for climate meaning.

- UMUI window:
  - *Atmosphere -> control -> post-processing, dumping and meaning*

- All restart dumps will be kept unless you set:
  - *Post processing -> main switch & general questions*
    - automatic post processing : yes
    - delete superseded restart dumps : yes
  - This works with and without archiving

Automatic resubmission

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



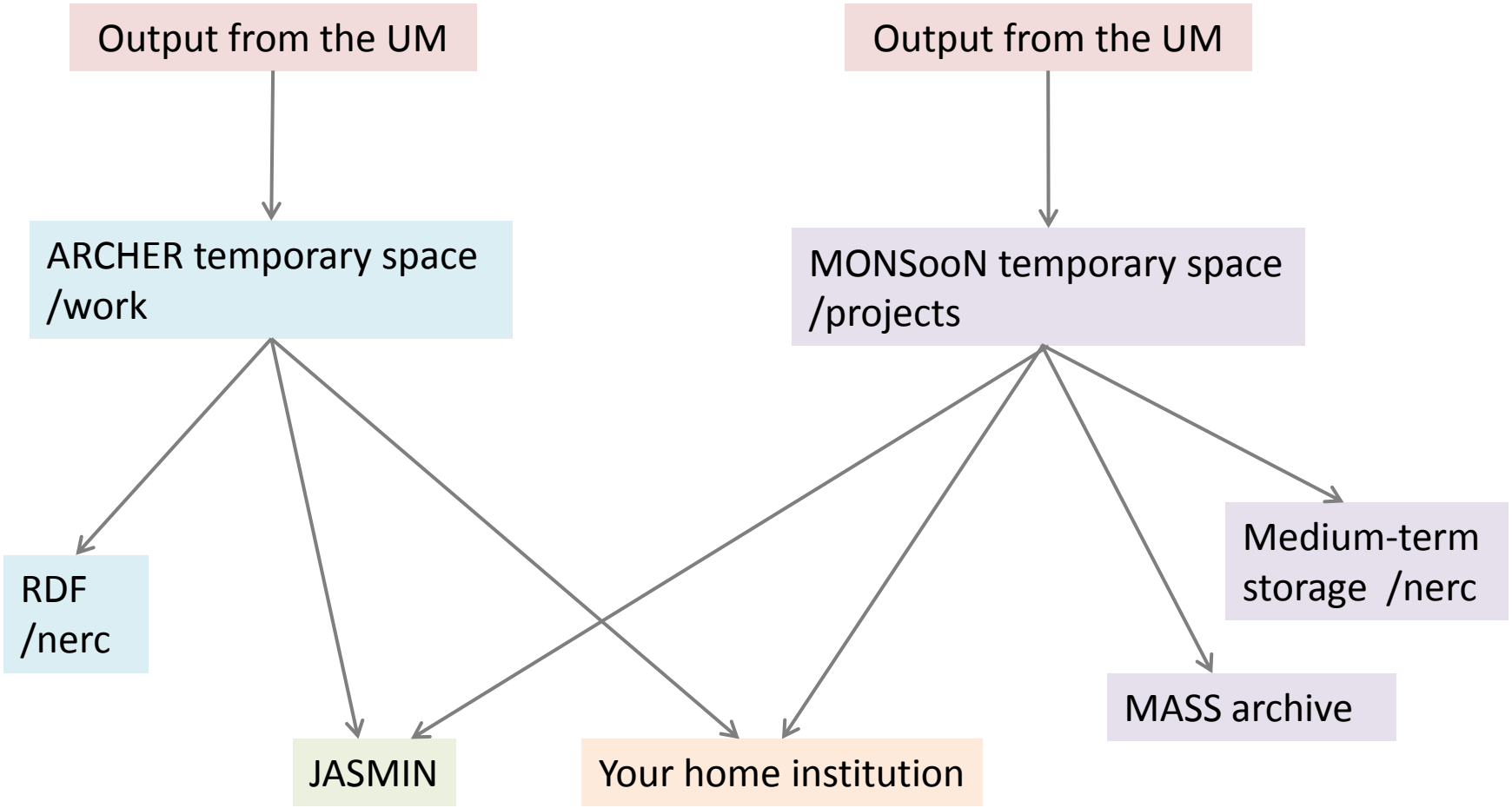
- Always keep the same restart frequency throughout an experiment and for experiment inter-comparisons.

- To run a long climate experiment
  - set the restart dump frequency to a 'reasonable' rate
  - run the first chunk, check results are ok
  - change NRUN to CRUN in SUBMIT (produced by the UMUI)
  - resubmit the job

- Automatic data archiving, based on the UKMO system, has been implemented on the national HPC services.

Automatic archiving

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



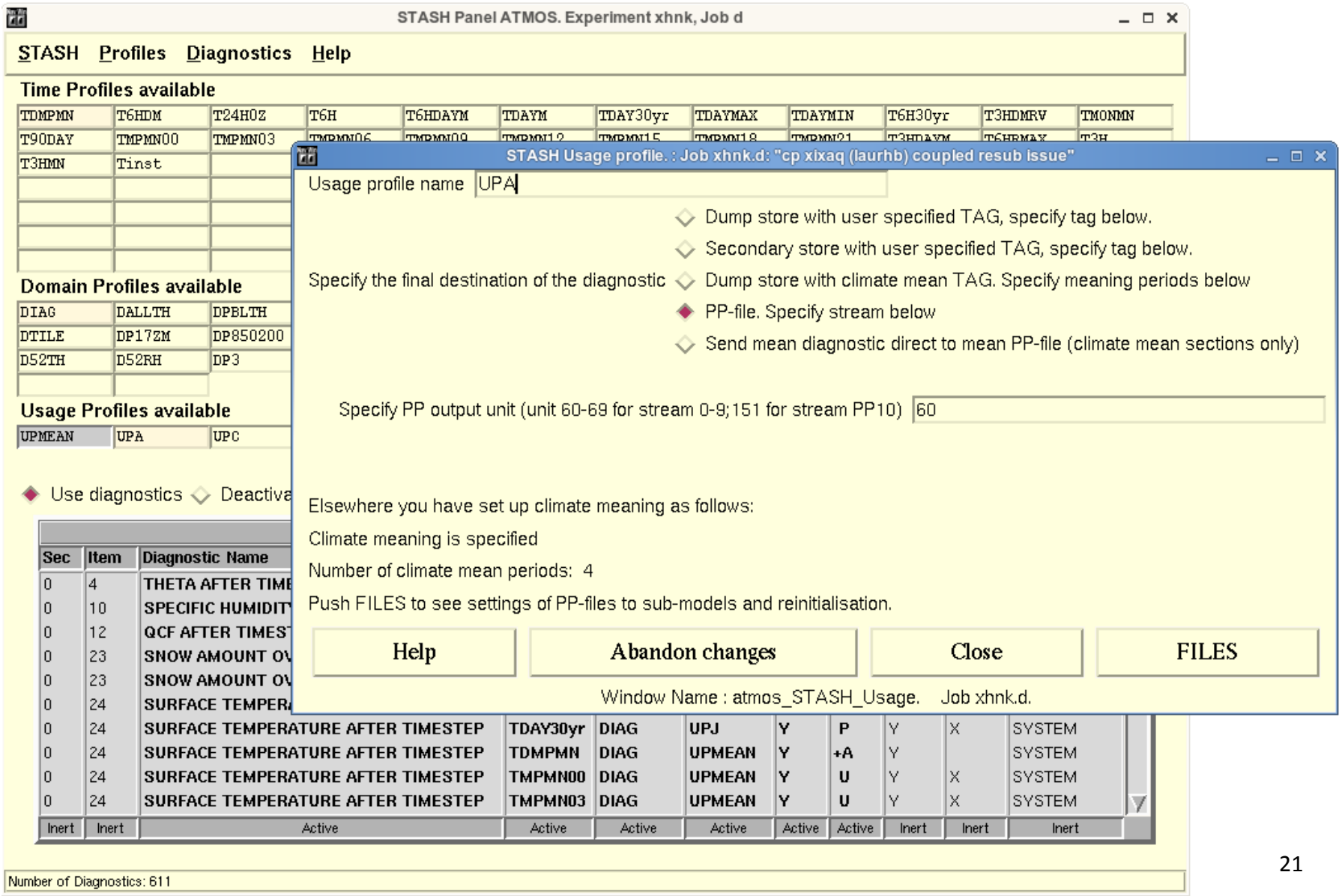
Post processing files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Output in fields file format (soon to be CF-netCDF)
- Diagnostics set up via STASH and written to a "post processing unit":
  - *Atmosphere -> STASH*
- Post processing unit file sizes depend on options selected in the UMUI:
  - *Post processing -> initialisation and processing*
  - Options include
    - Profile 0: unpacked
    - Profile 5: new standard climate packing
      - **Packed files save space when archiving.**
- Reduce file sizes at the end of the run (not if the run is to be continued) using "umpack". This gets rid of padding and any extra headers.

Configuring diagnostics

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



Configuring post processing files

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Define processing and post-processing requirements for the PP output streams.

Define periodic re-initialization for those files which require automatic post processing.

PP Files										
Basics					For re-initialised PP files, also specify					
PP File/Unit	Packing profile	Override size	GRIB FORMAT (Y/N)	Periodic Re-init	Period	Starting	Ending	Time Unit	Sub Model	Archiving
PP0/PA/60	5	16000	N	Y	30	0	-1	DA	A	Y
PP1/PB/61	5	0	N	Y	90	0	-1	DA	A	Y
PP2/PC/62	5	16000	N	Y	90	0	-1	DA	A	Y
PP3/PD/63	5	16000	N	Y	30	0	-1	DA	A	Y
PP4/PE/64	5	16000	N	Y	30	0	-1	DA	A	Y
PP5/PF/65	5	0	N	Y	90	0	-1	DA	A	Y
PP6/PG/66	5	0	N	Y	90	0	-1	DA	A	Y
PP7/PH/67	5	16000	N	Y	30	0	-1	DA	A	N
PP8/PI/68	5	0	N	Y	90	0	-1	DA	A	Y
PP9/PJ/69	5	0	N	Y	90	0	-1	DA	A	Y
PP10/PK/151	5	16000	N	Y	30	0	-1	DA	A	N
Insert	Edit	Edit	Edit	Edit	Edit	Edit	Edit	Edit	Edit	Edit

Time units are: DA=days, H=hours, T=timesteps, RM=real months.

Packing profiles numbers are as defined for mean PP file.

A (Atmosphere) is currently the only valid sub-model.

Unit number for fields files 60 – 69

Files labelled xxaabo.pp4 in $DATAW if not initialised

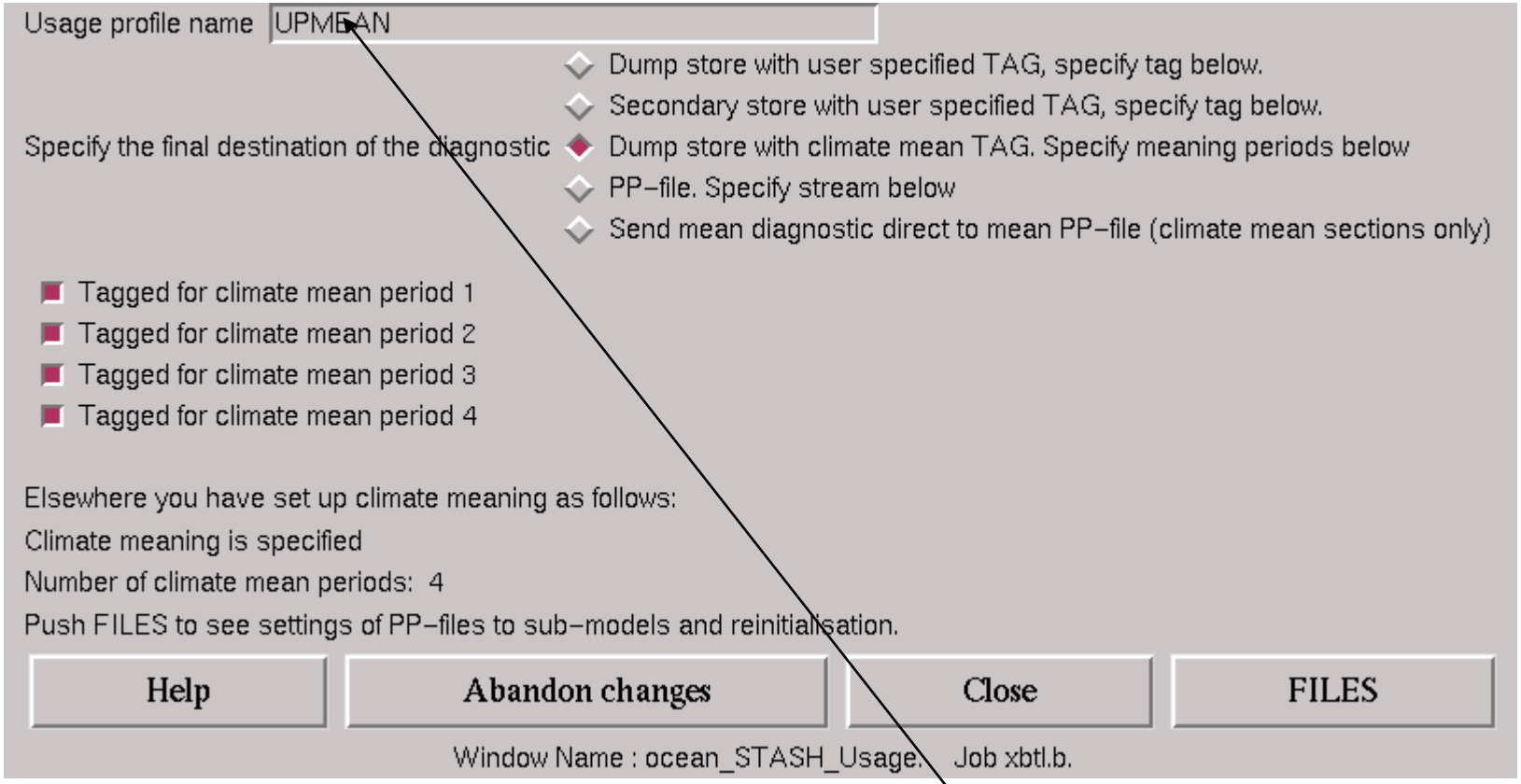
Files labelled xxaabo.pdxxxx in $DATAM if re-initialised

File names calculated from the re-initialisation process

Climate Meaning

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Climate runs using automatic resubmission, meaning the mechanism is safe over restarts.
- Diagnostics can be tagged to be picked up by the climate meaning system via STASH.
- Climate means are related to the dump frequency:
  - *Atmosphere -> Control -> Post-processing, Dumping & Meaning -> Dumping and meaning -> NEXT*
  - For example with a dumping frequency of 10 days, meaning periods of 3, 3, 4 and 10 will give:
    - monthly (3 x 10 day) means
    - seasonal (3 x 30 day) means
    - yearly (4 x 90 day) means
    - annual (10 x 360 day) means
- File names will be of the form $RUNID.mm…, $RUNID.ms…



Climate meaning uses partial sum files to preserve precision across restarts

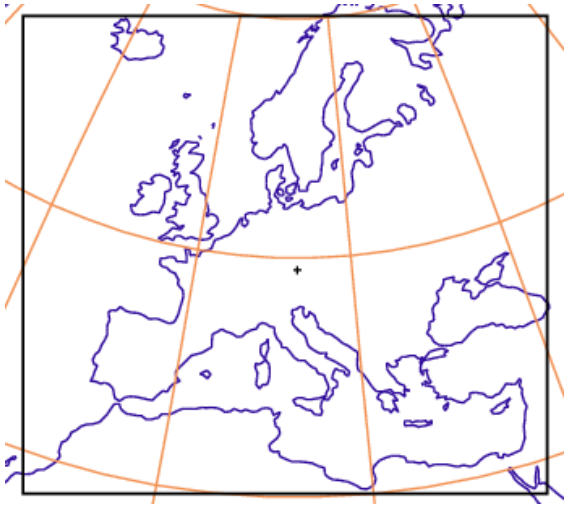
```
xbtlba.dab9cb0   xbtlba.pmb9dec   xbtlbo.dab9cl0   xbtlbo_s2a
xbtlb.apstmp1    xbtlba.dab9cl0   xbtlba_s2a       xbtlbo.dac0110
xbtlbo_s2b
xbtlb.opstmp1    xbtlba.dac0110   xbtlba_s2b       xbtlbo.pdc0c10
xbtlb.phist      xbtlba.pab9dec   xbtlbo.dab9cb0   xbtlbo.pmb9dec
```

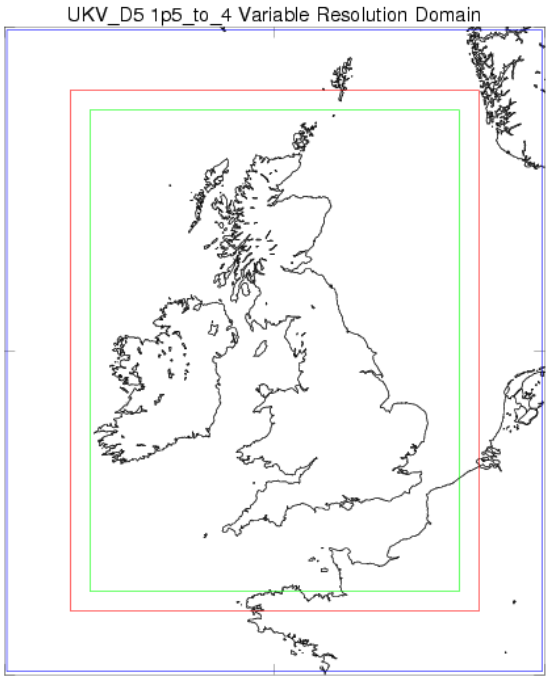
- UM file types
- Model input
- Model output
- **Limited Area Models**
- NEMO and CICE

Current resolutions

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Operational global model: 17 km (previously 25 km)
- Europe LAM: 4 km (operational)
- On-demand LAM: 1.5 km (forecaster requested)
- Storm scale LAM: 1 km (research model)
- Research LAMs: as low as 50 m

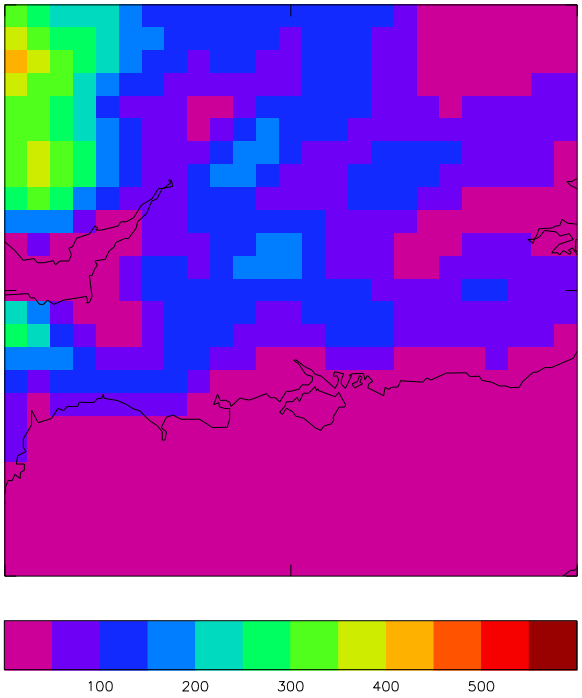




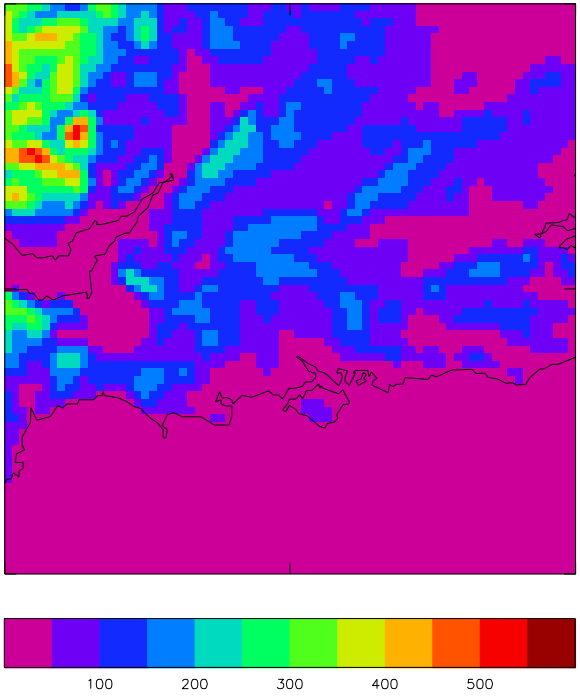
Orography in the UM

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

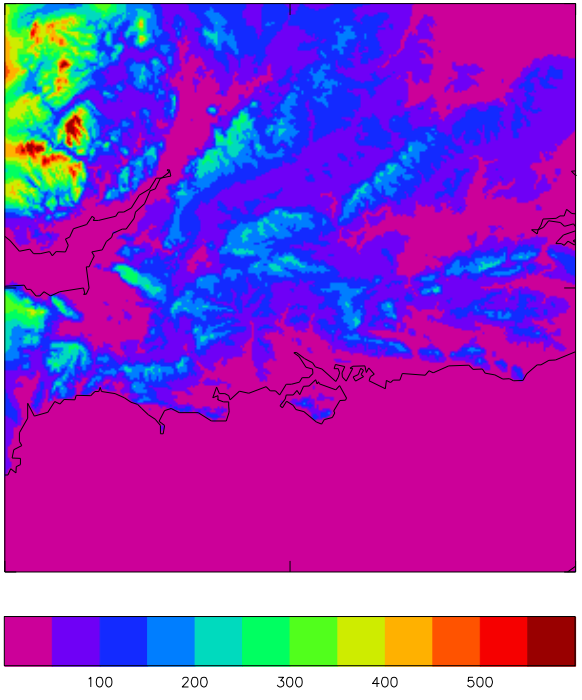
12 km



4 km



1 km



Height of model orography (m)

Lateral boundary conditions

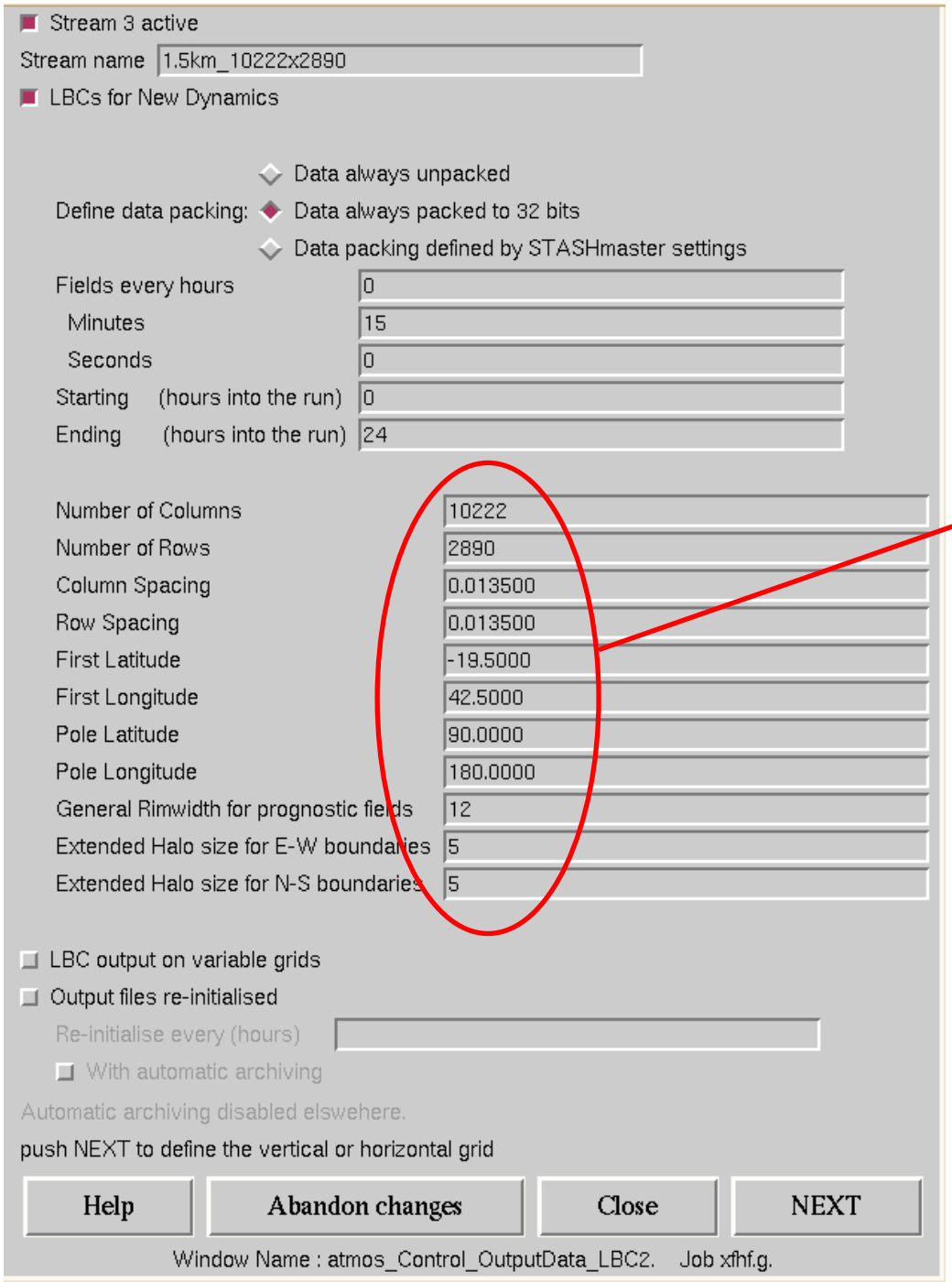
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- LBC files are required by the Limited Area Model (LAM).
- For standard areas (UK LAM and UK mesoscale):
  - Run the global UM to generate LBCs
    - **1024 x 769 L70 – global forecast resolution**

    Atmosphere -> control -> output data files -> LBCs out
  - Run LAM

    Atmosphere -> ancillary and input data -> other ancillary -> LBC
- For higher resolution standard areas:
  - On puma under owner “umui” there are example jobs for
    - 12 km, 4 km and 1 km UK mesoscale
  - These need start and ancillary files, some available on national HPC services.

Outputting LBCs



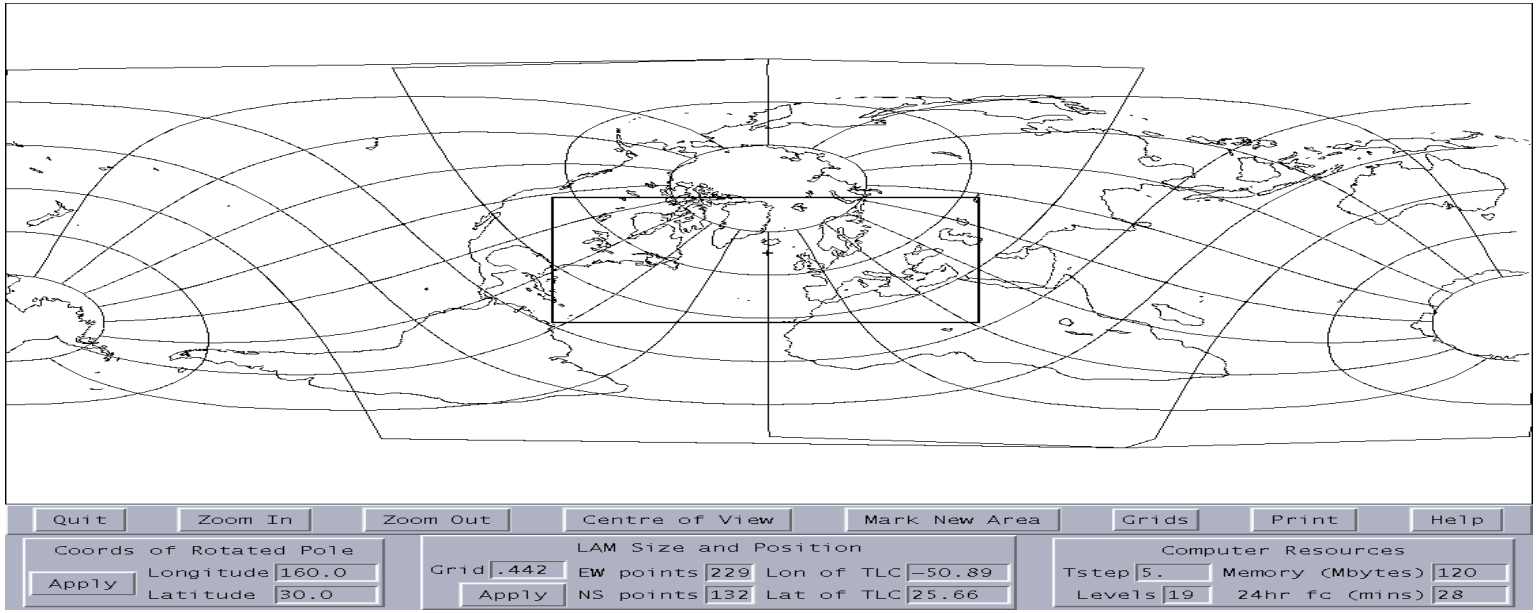
Specification of target grid

LAMPOS

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

For non-standard areas:

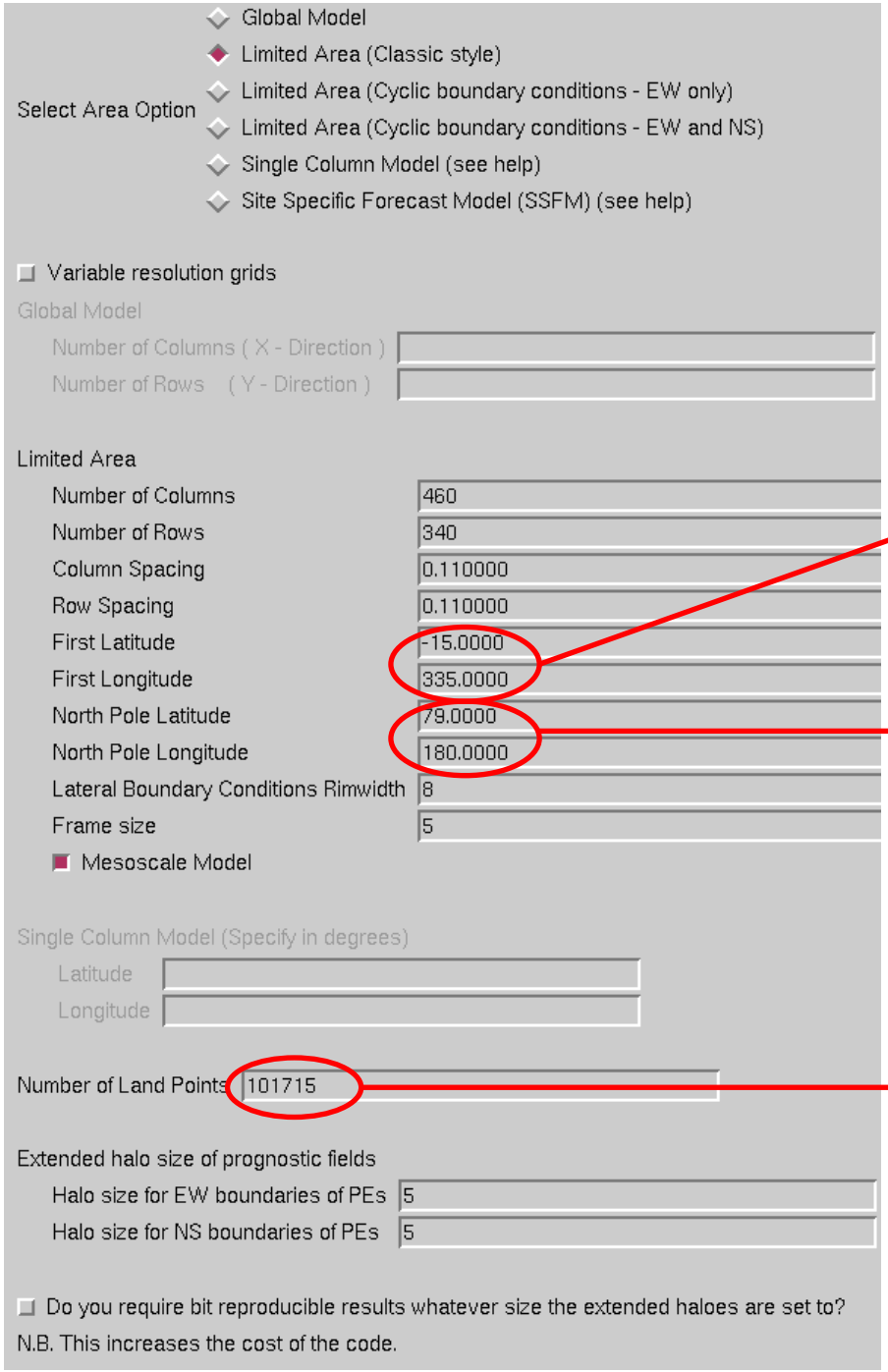
- choose new domain using **LAMPOS** (available on puma)



- Create ancillary files:
  - Use the UKMO ancillary file creation system: **CAP**

    Available via the CMS web page
- Run the global UM or larger LAM to generate LBCs
- Run LAM
- Seek help and advice from NCAS CMS

Horizontal domain specification



Lat, lon in rotated grid

Used to put equator of rotated grid
In the centre of the LAM domain
LAMPOS or BADC Grid utilities

Obtained from CAP
or reconfiguration

- UM file types
- Model input
- Model output
- Limited Area Models
- **NEMO and CICE**

Ocean and sea-ice

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

NEMO files:

- All in NetCDF
- Inputs cover full domain
- By default one output file per process covering only that subdomain.
  - Construct global file after run using **rebuild_nemo**, e.g.:

    rebuild_nemo xdodto_CU150_19780901_19780930_grid_T 16
- *NEMO -> Scientific Parameters and Sections -> Links to NEMO model*

CICE files:

- Restart files in binary format
  - There are tools to edit the start date and convert to NetCDF
  - (contact NCAS-CMS for guidance)
- All other inputs and outputs in NetCDF (global files)
- *CICE -> Scientific Parameters and Sections -> Links to CICE model*

ORCA grids

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Used by NEMO and CICE
- Tripolar grids
- Standard resolutions:
  - ORCA2 (2°)
  - ORCA1 (1°)
  - ORCA025 (¼°)
  - ORCA12 (1/12°)