NMM XML

Numerical Model Metadata XML

Exeter Metadata Workshop
May 9-12, 2006

L. Steenman-Clark, K. A. Bouton, R. Hatcher
NCAS Centre for Global Atmospheric Modelling,
University of Reading, UK

NMM Overview

- is an *evolving* metadata standard intended for the exchange of information about numerical models or codebases, and the simulations done using them.
- provides a metadata standard to describe the numerical codebase and its associated simulations, extending and refining the researcher's ability to understand how a resulting output data was produced.
- built and based on XML and associated technologies.

NMM Goal

- to provide a metadata standard and structure which is clear, well-defined and flexible.
- able to describe climate and forecast numerical codebases (e.g. Met Office Unified Model).
- able to describe the experiments/runs (e.g. xaabc) enabling scientists to search, retrieve, difference and compare.
- provide a better understanding of the provenance of the model data output.
- Note – presentation concentrates on NMM for an model/exp.

NMM Goals cont'd

- It is envisioned that the scientist would create the NMM XML file, most likely using a tool to automatically create it while, or after, the run is in process. The automatic process would 'capture' the required elements of the codebase and components being used and all the settings and values used to create the model.
- Ideally the NMM Model XML file would be submitted to the data centre with the model output data as part of the data description.

For this Meeting

Interoperabilibity between frameworks will be determined by the metadata available.

- Semantic metadata describes what it really *means*. This is the type of metadata required for initial discovery and are often more descriptive, and use a standard/controlled vocabulary.
- Syntactic metadata describes what the data *looks* like. This is the type of metadata required for coupling frameworks
- From: http://marinemetadata.org/guides/metadatatypes/

NMM Structure

- Information Properties
- Technical Properties
- Numerical Properties
- Science Properties
- Input/Output Properties

NMM Model Attributes

- Which NMM Version was used to create the file?
- What was the NMM creation date?
- Who was the NMM author?
- Assign a unique NMMModel_ID

Information Properties

- Purpose: To give some general descriptive information about the model such as name, description of why it was run, contact, documentation, history, references etc.
- These properties will be common for many other metadata standards and should be able to be culled directly from a NMM file.

Technical Properties

- Purpose: To provide some minimal details of the machine properties it was run on, with an eye to creating performance statistics, or re-creating the run. What machine was it run on, type of compiler used, what libraries, optimization, code parallelisation etc

- Many of the properties are shared with coupler metadata requirements, while they have limited use for data metadata requirements.

Numerical Properties

- Purpose: To provide high level descriptive information on the horizontal and vertical resolutions and extent of a model, and time resolution of the model.

- Semantic field in NMM
    - e.g “regular lat long grid”
- Semantic fields in framework metadata
    - PRISM: PRISM_reglonlatvrt
    - ESMF: ESMF_GRID_TYPE_LATLON
    - Gridspec : logically rectangular grid

Semantic Example from APE

Group	Location	Model	Resoln	Features
AGU for APE	Japan (consortium)	AFES v.1.15	T39 L48	Spectral, eulerian
CGAM	Reading, UK	HadAM3	N48 L30	3.75º x 2.5º grid
CSIRO	Aspendale, Australia	CCAM		~220km conformal cubic grid
DWD	Mainz, Germany			icosahedral-hexagonal grid
ECMWF	Reading, UK	IFS Cycle 29r2	T_L159 L60	Spectral, semi-lagrangian
FRCGC	Yokohama, Japan	NICAM	7km L54	icosahedral grid, non-hydro.
GFDL	Princeton, USA	AM2p14	N72 L24	2.5º x 2º grid (IPCC)
GSFC	Maryland, USA	NSIPP-1	N48 L34	3.75º x 3º grid
K1-Japan	Japan (collaboration)	CCSR/NIES 5.7	T42 L20	s-l moisture and cloud
LASG	Beijing, China	SAMIL	R42 L9	Spectral, eulerian
MGO	St. Petersburg, Russia	MGO-gcm	T30 L14	Spectral
MIT	Cambridge, USA	MIT-gcm	C32 L40	~280km cubed sphere
NCAR	Boulder, USA	CCSM-CAM3	T42 L26	Spectral, eulerian
UKMO	Exeter, UK	pre-HadGAM1	N96 L38	1.875º x 1.25º grid, s-lagrangian

Reg lat long grid

http://www.met.reading.ac.uk/~mike/APE/

Science Properties

- Purpose: To provide details on the science parameters and setting used to create the model

Input/Output Properties

- Purpose: To provide details on the initial conditions, boundary conditions, physical constants used to create the model, and point to the output data files (which have their own metadata e.g. CF)

NMM and the IPCC Models

- IPCC model documentation extremely varied. Was it standardizable?
- Created a NMM XML file for each IPCC Climate Model Documentation and recreated table based on the standard
- Table enabled cross comparison

IPCC Example

View Table create by XSL from the many model NMM XML files

To Do

- Numerical Properties – describing grid specs a la wondergrid
- Can we create a PRISM XML file from NMM XML – what would be missing?
- Apply to other models
- Tools such as a personal notebook
- Get the word out

Information Properties Details

localName
longName
description
changeHistory
contacts
references
resourceConstraints
descriptiveKeywords
browseGraphics

Technical Properties Details

machine - machineName, description
compiler - name, version, description
optimization
libraries
dateCompiled
changeHistory
environmentVariables
code_parallelisation – processorsUsed, haloSize

Numerical Properties Details

HorizontalRepresentation

coverage

schemeType spectral, grid, mixed, isohedral etc

resolution

longitude (longitudeDegrees, extent, description)

latitude (latitudeDegrees, extent, description)

northPole

changeHistory

VerticalRepresentation

resolution- numberOfLevels

coverage

extent

formulaValues

changeHistory

TimeIntegration

Resolution - timeStepsPerDay

changeHistory

Science Properties Details

standardName
localName
definition
modifiedFromStandard
references
parameter
  localName
  value
  perturbed
  modified
forcing internal (or) external (or)component
changeHistory

Input/Output Properties Details

input(s)
- initialCondition
- mode – internal, external or component
- description
- initialConditionType
- perturbed
- changeHistory
- references

physicalConstant(s)
- standardName
- localName
- value
- references

boundaryCondition
- description
- boundaryConditionType
- references
- boundaryDataset

CouplingRequirements
- PMIOD

output
- outputFile_ID
- outputFileFormat
- description
- changeHistory
- references

PRISM AD.dtd

PRISM/OASIS	NMM
application	NMMModel
local_name	Information-localName
long_name	Information-longName
start_mode	
coupling_mode	
nbr_procs	Could be in NMM Component Technical Properties
platform	TechnicalProperties-machine-machineName Plus more-compilers,options,libraries,history etc
component->local_name	NMMComponent_ID localIdentifier
component->long_name	
component->nbr_procs	See above

PRISM SCC.dtd

SCC	NMM
experiment driver->nbr_proc start_date end_date run with multiple applications/components	NMMModel Could be in Technical Properties Numerical Properties another NMMModel with multiple components

PRISM PMIOD.dtd

PRISM/OASIS PMIOD	NMMCodebase
code->contact	contacts
code->documentation	references
Fortran_Units	TechnicalProperties-codingLanguage
Gridfamily	HorizonalRepresentation-schemeType-grid
physical space/sampled space	
compute_space	
Transient local/standard	
units,datatype,computation	ScienceProperties standard and local scheme and parameters
intent	

PRISM SMIOC.dtd

PRISM/OASIS SMIOC	NMMModel
code->contact	contacts
code->documentation	references
Fortran_Units	TechnicalProperties-codingLanguage
Gridfamily	NumericalProperties-HorizonalRepresentation-schemeType-grid
physical space/sampled space	
compute_space	ScienceProperties standard and local scheme and parameter and value
transient	
local/standard	
units,datatype etc	
intent	