Guidance document on applying the Systems Maturity Matrix (SMM) to C3S_511 ECV SPQB dataset assessments

Prepared by: Maynooth University - John Coll

LMU - Benjamin Mueller

Date: 24/09/18

Ref: C3S_511. SMM Guidance_v1

Edit History: Coll 24/09/18

Mueller 24/09/18

Contents

1. Background

1.1 The maturity matrix approach

Data stewardship encompasses all activities that preserve and improve the information content, accessibility, and usability of data and metadata (NRC, 2007). Increasingly stringent regulations levying enhanced stewardship requirements on digital scientific data has increased the need for a formal approach by data stewardship practitioners. Therefore the application of existing maturity assessment models is becoming a key practice component for data service centres.

The system maturity matrix (SMM) for Climate Data Records (CDRs) was developed under the Coordinating Earth Observation Data Validation for Reanalysis for Climate Services project (CORE-CLIMAX) (Schultz *et al*., 2017; Su *et al*., 2018). The aim was to develop a tool to assess different facets of a CDR, and to semi-objectively establish to what extent data record management based on accumulated best practices by the scientific and engineering communities have been met (EUMETSAT, 2014). Since many observing systems were designed to measure weather rather than for climate monitoring, various assumptions, approximations and associated uncertainties have to be characterised as part of the assessment process. In the SMM framework assessments are made in 6 major category areas and a score of 1 to 6 is assigned that reflects the maturity of the CDR with respect to a specific category:

1. Software readiness
2. Metadata
3. User documentation
4. Uncertainty characterisation
5. Public access, feedback, and update
6. Usage

The major categories of the SMM are further subdivided into several minor categories and assessment scores are assigned based on scores in these minor categories. For each of these subcategories an assessment score of 1 to 6 is assigned (and where 1 or 2 can be identical criteria and 6 often not used) with an overall aim of reflecting the maturity of the various aspects of the CDR.

While ascertaining the maturity of climate data records involves many sensitive aspects, the best-practices approach of the CORE-CLIMAX projects has been met with community acceptance (Su *et al*., 2018). Hence the remainder of this document applies the SMM support guidance notes from CORE-CLIMAX V4 documentation (EUM/USC/DOC/13/701166, 19/11/2013).

1.2 SMM adapted for C3S_511

The C3S_511 SMM is derived from the CORE-CLIMAX template and the various categories therein, and which have been adapted to the specific service needs. Thus e.g. while a workbook for 'Software Readiness' and other minor categories have been left included in the Excel file "SMM_CORE_CLIMAX_c3s_Adapted_v5.0.xlsx" as a service/assessor we are not in a position to evaluate these. A further modification for our C3S_511 service is that a set of 'Defensible Traces' cells has been added for the various categories we will assess. The objective of these is to allow each product assessor to support their scoring decisions with reference to any available supporting Literature.

Once all scores of the SMMs are entered we are able to generate an overview of the maturity of the dataset and export these to summary *.csv* file outputs. The overview shows: the variable name, the assessment version, the earth-system which the dataset represents, the original project from where the variable is derived, type of assessment, and the minimum and the maximum scores as well as the time of the assessment. Below specific guidance notes and where to find these in the CORE-CLIMAX SMM Instruction Manual (IM) ("CORE_CLIMAX_MANUAL.pdf") relating to the various categories and sub-categories are supplied as a tabulated summary for our service.

Table 1. Summary guide for C3S_511 assessors: Accessing the CORE-CLIMAX SMM support guidance notes

IM section and notes	Sub-section and topic	Pages
4.1 Software Readiness	4.1.1 Coding Standards	10
	4.1.2 Software Documentation	11
	4.1.3 Portability and Numerical Reproducibility	12
	4.1.4 Security	13
4.2 Metadata	4.2.1 Standards	14
	4.2.2 Collection Level	15
	4.2.3 File Level	16
4.3 User documentation	4.3.1 Formal Description of Scientific Methodology	17
	4.3.2 Formal Validation Report	18
	4.3.3 Formal Product User Guide (PUG)	19
	4.3.4 Formal Description of Operations Concept	20
4.4 Uncertainty Characterisation	4.4.1 Standards	21
	4.4.2 Validation	23
	4.4.3 Uncertainty Quantification	24
	4.4.4 Automated Quality Monitoring	25
4.5 Public Access, Feedback and Update	4.5.1 Access and Archive	26
	4.5.2 Version Control	27
	4.5.3 User Feedback	28
	4.5.4 Updates to Record	29
4.6 Usage	4.6.1 Research	30-31
	4.6.2 Decision Support System	31
Legend:	Currently not assessed	

References

European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) (2014). CORE-CLIMAX System Maturity Matrix Instruction Manual. CC/EUM/MAN/13/002 (v4). Darmstadt, 34pp.

National Research Council (2007). Environmental data management at NOAA: Archiving, stewardship,and access. 130. The National Academies Press. Washington, D.C. DOI: https://doi.org/10.17226/12017
[Available online at: https://www.nap.edu/catalog/12017.html].

Schulz, J., John, V., Kaiser-Weiss, A., Merchant, C.,Tan, D., Swinnen, E., and Roebeling, R.: European climate data record capacity assessment, Geoscience Data Journal, in preparation, 2017.

Su, Z., Timmermans, W. J., Zeng, Y., Schulz, J., John, V. O. and others (2018). An overview of European efforts in generating climate data records. *Bulletin of the American Meteorological Society*, *99*(2), 349-359. DOI: 10.1175/BAMS-D-16-0074.1