How to efficiently deal with model quality documentation work during the model development process?

Report of a workshop held at the Wageningen Model & Data Day – 17 October 2024.

Number of participants: +/- 25

Vincent Hin (WMR), Thomas Hagenaars (WBVR), Peter Hobbelen (WBVR)

Background and outline of the workshop

The knowledge base (Dutch: *kennisbasis*) project *Models to Assess Synergies and Trade-offs* (MAST) aims to further develop models, modelling tools and model collaboration within Wageningen Research (WR). Within the MAST project, the work package on model development has the generic task to facilitate the process of model development by creating a more structured approach to model development that incorporates the model quality criteria. These criteria are collected in a *model quality checklist* defined within the Wageningen Modelling Group (WMG) and can be used to (self-)assess the quality of a model. The process of model development is described in the WMG wiki pages[1] as a 'modelling cycle' with five different model development steps and a number of surrounding processes. Each step has a number of good modelling practises (GMPs) guidelines. Following these GMP guidelines will ultimately deliver models that are fit-for-purpose. Although the GMPs and quality checklist items are related, the quality criteria and their associated requirements are not necessarily directly linked to the different steps of the modelling cycle. Within the model development work package of the MAST project, the requirements as set forward by the model quality criteria were explicitly linked to the different stages of the modelling cycle. This linkage was done by creating an extended version of the modelling cycle, which was presented and discussed during the workshop held at the WMDD 2024.

The workshop consisted of three presentations, followed by a plenary discussion. The first presentation by Peter Hobbelen was an introduction to the model quality checklist. Thereafter, Thomas Hagenaars presented the model development cycle according to the GMP wiki. Vincent Hin then introduced the extended model cycle that linked the quality checklist items to the model development cycle. After the presentations there was opportunity for questions and the audience was asked for feedback and input on the proposed extended model development cycle.

Model quality checklist (Peter Hobbelen)

Peter Hobbelen introduced the model quality checklist, which can be found on the intranet page of the Wageningen Modelling Group[2]. The model quality checklist is a collection of requirements for achieving a model quality status. It consisted of two tiers, one for status A and an additional set of requirements for status AA. The requirements for status A require specific actions and analysis that are related to proper documentation of the model, and various aspects surrounding the model, such as model development, model management, and documentation of model usage. The model quality checklist is not limited to models, but can also be applied to indicators, maps, datasets and databases.

There are a number of different aims of the quality checklist. The first and foremost aim is to improve the fitness-for-purpose of the models developed within WR. Here, fitness-for-purpose can be described as the ability of a model to answer the research questions that motivated the development of the model. Other purposes are to increase work flow efficiency and create models that are FAIR: Findable, Accessible, Interoperable and Reusable. Adhering to the requirements of the checklist also helps to improve value creation, model transparency and interpretation of model output.

---

1 https://ictheek.wurnet.nl/spaces/viewspace.action?key=GMP

2 https://intranet.wur.nl/Project/WRModellingToolbox/Pages/YGim_IMb5keB0iMxf2Lj5A

Table 1: Overview of the model quality checklist perspective, topics and number of requirements per topic

3 Perspectives	7 Topics	Number of requirement	Requirements
Scientific rigor & Technology	1. Description	2	.1 General description; .2 conceptual and formal model
	2. Technical implementation	3	.1 Implementation documentation; .2 Technical environment description; .3 Model testing
	3. Parameters, variables, input and output	4	.1 Parameters & variables; .2 Calibration of parameters; .3 Input & output; .4 Origin of input data;
	4. Evaluation	5	.1 Sensitivity analysis; .2 Uncertainty analysis; .3 Model validation; .4 Use monitoring; .5 General assessment
Management	5. Planned development	2	.1 Development plan; .2 Version control
	6. The organization	4	.1 Model metadata; .2 Management plan; .3 Dependencies; .4 External use
User friendliness	7. User documentation	2	.1 Interpretation guidance; .2 User manual

The quality checklist consists of 22 requirements at two tiers (A and AA), which are divided into 7 topics and 3 perspectives (Table 1). Each requirement consists of a number, and list of keywords and a short description of the requirement. An example is given in Box 1. It should be noted that the quality checklist only describes the important steps and requirements to achieve a certain quality status (status A or AA), but does not explicitly describe how to achieve this, e.g. there is no explicit format that structures the model documentation.

Box 1: Model quality checklist requirement 7.1 of perspective "User friendliness", topic 7: User documentation.

7.1 Interpretation guidance is provided

A *Goal * area of application * theoretic framework * summary of evaluations * general public*

Interpretation of model output or dataset contents is in general not trivial. The interpretation implies understanding of the theoretic framework, conceptualisations and formalisations of the model - i.e., the assumptions and simplifications - and of the outcome of model evaluations. Guidance should be supplied on what these mean for the value of the outcome and on when (not) to use the model/dataset. This interpretation guidance should be readable for more general public than the scientific community.

AA *Reflection on goal, area of application, structure, complexity*

The interpretation guidance is extended with a reflection on the tension between goal, area of application and complexity of the model, limitations in the implementation, data quality and availability and the realised model performance.

Model development according to GMPs (Thomas Hagenaars)

The good modelling practises (GMPs) are described for all phases of the modelling processes. These phases are outlined in the modelling cycle (Figure 1). Central to the cycle are data, which connect to all phases of the cycle. The most outward layer of the cycle are surrounding processes that interact with one or more phases of the modelling cycle.

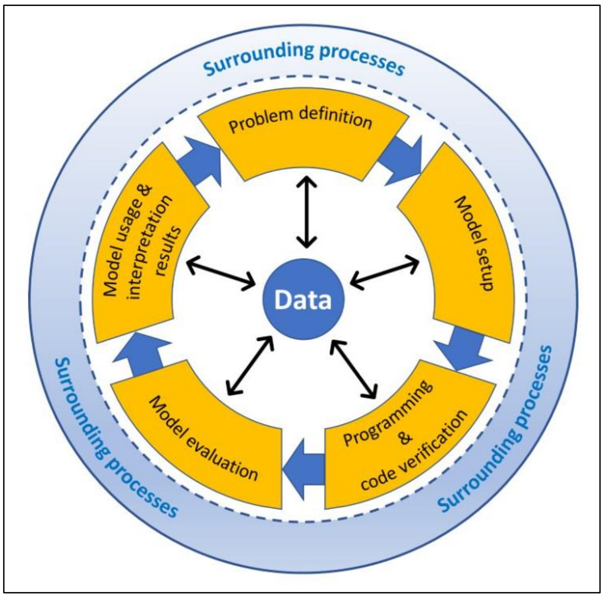

Figure 1 The modelling cycle as defined in the GMP wiki page. Source: https://ictheek.wurnet.nl/spaces/viewspace.action?key=GMP

Although one can run through multiple iterations of the cycle, the most obvious phases for newly developed models is the 'problem definition' phase. In this phase the research questions are defined and the modelling approach should be motivated. A first definition of model inputs and outputs should also be made in this phase. The model setup phases involves model formulation, model formalization and parameter estimation. This phase identifies the model components and system boundaries, model interrelationships (model flow diagram) and system variables. Parameter estimations should consider the need to additional data collection and include an estimation of parameter uncertainty. The phase of programming and code verification is used to create an actual implementation of the model and should be done after the model is formalized (model setup phase). For the programming and code verification step various GMPs are defined, which improve the (re)usability and interpretability of the model code. The model evaluation phase involves model calibration, model validation, sensitivity and uncertainty analysis and a general assessment of whether the model is fit-for-purpose. The last phase, model usage and interpretation, relates to all aspects of using the model, such as the repeatability and storage of model output, and the guidance for the interpretation of model results. Finally, there are six surrounding processes:

1. Model governance
2. Model management
3. Data management
4. Quality assurance
5. Stakeholder interaction
6. Reporting results

These processes can be connected to different parts of the modelling cycle. Model quality assurance is formalized by the model quality checklist as previously discussed and should be part of each phase of the modelling process.

Extension of model development cycle (Vincent Hin)

The modelling cycle (Figure 1) was combined with the model quality criteria (Table 1). This resulting in the extended modelling cycle that is shown in Figure 1Figure 2. This version of the modelling cycle was presented during the last presentation in the workshop. This version of modelling cycle makes explicit which model quality requirements should be met during which phase of the modelling process. Although model development is a dynamic, as opposed to a static process, we believe that fitness-for-purpose of models created at WR and the efficiency of model development will benefit from an integration between model quality criteria and the model development process. Adhering to model criteria guideline of each

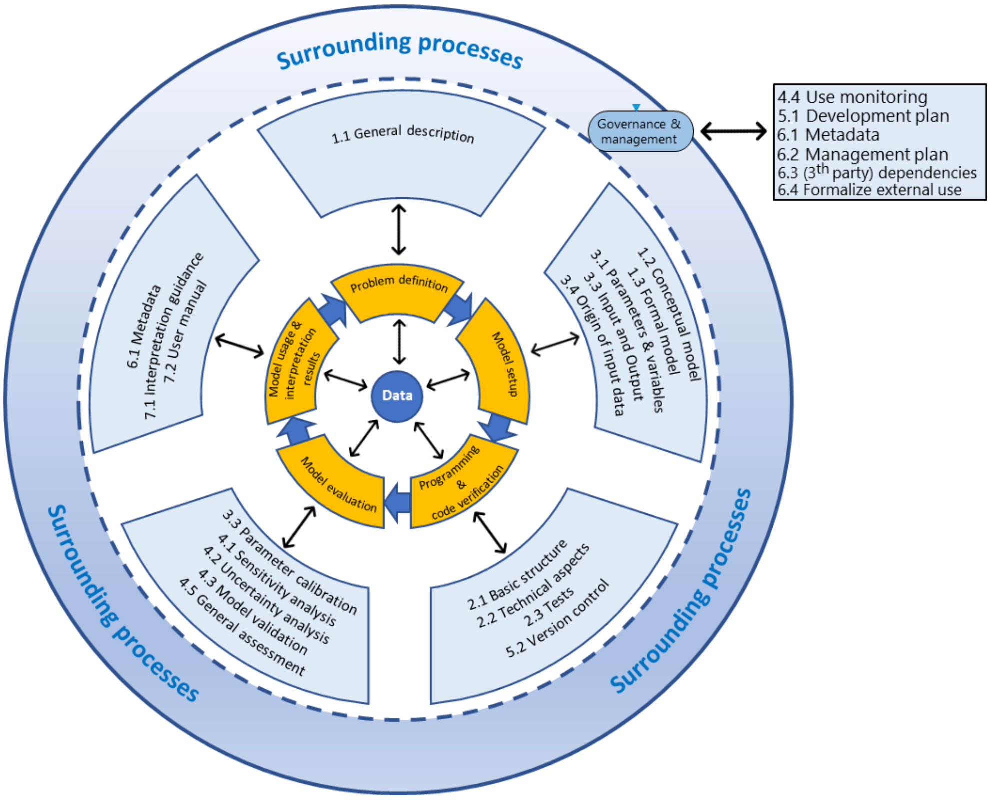

Figure 2:The extended modelling cycle that integrated the model development processes and the model quality criteria.

specific phase of the modelling process will likely result in increased efficiency of the consecutive phases in model development. For example, a proper model formalization and clear definition of model variables and parameters, including a consistency check of their dimensions during the model setup phase, will increase with efficiency in the subsequent programming and code verification phase, and likely also reduce the chance of programming errors.

In the extended modelling cycle (Figure 2) the numbering of the original model quality criteria items was retained (Table 1). Because the criteria are reordered according to the natural process of model development, their numbering does not necessarily match the ordering the cycle. For example, several criteria related to 'parameter, variables, input & output' (topic 3) are placed in the model setup phase and therefore appear before criteria 2.1 'Basic structure' and 2.2 'technical implementation' from topic 2, which appear in the 'programming & code verification' phase. Model quality criteria from different topics appear in the 'surrounding processes' layer.

Discussion and feedback from the audience

After the presentation, the audience was given the opportunity to ask questions and was asked to provide feedback on the extended modelling cycle. Several questions involved clarification of the purpose and approach used, which were elaborated on. There were also several questions related to the generality of the approach. For example, in case of the use of proprietary software or legacy code/models that were developed by other people and for which documentation or user manuals were missing. From a good modelling practise perspective it is to be recommended not to use such code any longer, as the implementation and verification of the code cannot be guaranteed. It was noted by participants that the modelling cycle and the model quality criteria are very much biased towards mechanistic, or process-based models, which aim to model how a system works. This is opposed to data-driven models, or AI models, in which a mechanistic representation is often lacking. The original modelling cycle, and therefore also its extended counterpart, might not properly reflect the natural model development process for data-driven or AI models. It was noted that this doesn't devaluate the current effort. Instead, it should inspire colleagues that work on such models that come up with a model development cycle that will better match data-driven or AI models.

Overall, the workshop participants were enthusiastic about the proposed extended modelling cycle, e.g. considering it as the proper ‘design methodology for model construction’. It was also made clear that many modellers in the WR community struggle with the model quality criteria and how to fill in the related self-assessment. They thought that the proposed structuring according to the model development process could help their own model development by increasing its efficiency. In this context it was also noted that keeping a log during model development is the most crucial good practice to be able to comply with the model quality criteria. Finally, there was a brief discussion about scientific software development and the possibility, if the maturity level of the model is high, to involve software developers (from WUR IT unit).