Improving the accessibility and re-use of environmental models through provision of model metadata – a scoping study

Andrew Riddick (1), Andrew Hughes (1), Quillon Harpham (2), Katherine Royse (1), and Anubha Singh (1)
(1) British Geological Survey, United Kingdom (atr@bgs.ac.uk), (2) HR Wallingford, United Kingdom

There has been an increasing interest both from academic and commercial organisations over recent years in developing hydrologic and other environmental models in response to some of the major challenges facing the environment, for example environmental change and its effects and ensuring water resource security. This has resulted in a significant investment in modelling by many organisations both in terms of financial resources and intellectual capital. To capitalise on the effort on producing models, then it is necessary for the models to be both discoverable and appropriately described. If this is not undertaken then the effort in producing the models will be wasted.

However, whilst there are some recognised metadata standards relating to datasets these may not completely address the needs of modellers regarding input data for example. Also there appears to be a lack of metadata schemes configured to encourage the discovery and re-use of the models themselves. The lack of an established standard for model metadata is considered to be a factor inhibiting the more widespread use of environmental models particularly the use of linked model compositions which fuse together hydrologic models with models from other environmental disciplines. This poster presents the results of a Natural Environment Research Council (NERC) funded scoping study to understand the requirements of modellers and other end users for metadata about data and models.

A user consultation exercise using an on-line questionnaire has been undertaken to capture the views of a wide spectrum of stakeholders on how they are currently managing metadata for modelling. This has provided a strong confirmation of our original supposition that there is a lack of systems and facilities to capture metadata about models. A number of specific gaps in current provision for data and model metadata were also identified, including a need for a standard means to record detailed information about the modelling environment and the model code used, to assist the selection of models for linked compositions.

Existing best practice, including the use of current metadata standards (e.g. ISO 19110, ISO 19115 and ISO 19119) and the metadata components of WaterML were also evaluated. In addition to commonly used metadata attributes (e.g. spatial reference information) there was significant interest in recording a variety of additional metadata attributes. These included more detailed information about temporal data, and also providing estimates of data accuracy and uncertainty within metadata.

This poster describes the key results of this study, including a number of gaps in the provision of metadata for modelling, and outlines how these might be addressed. Overall the scoping study has highlighted significant interest in addressing this issue within the environmental modelling community. There is therefore an impetus for on-going research, and we are seeking to take this forward through collaboration with other interested organisations. Progress towards an internationally recognised model metadata standard is suggested.