Call for a common model documentation etiquette

J. Benz[*], M. Knorrenschild[†]

May 29, 1997

During the past decades a variety of mathematical models for different questions in ecology and environmental protection have been developed. These models represent a large collection of scientific knowledge and experience that is, in many cases, not readily and not completely available to the modelling community.

In far too many cases, model documentations leave a lot to be desired. Even worse, there are a number of models for which no documentation exists at all, so that published research results based on these models are not verifiable. Currently, there are several resources that provide documentation for a wide range of ecological models. For example, the recently published "Handbook of Environmental and Ecological Modelling" by Jørgensen et al. includes documentation on 450 ecological models. There are also several on-line databases that provide model information, accessible via the internet. However, because standard etiquette for documenting models has not yet been adopted, most documentation schemes have different layouts, different degree of detail and are not systematically linked. To carry out model comparisons under such conditions is difficult.

It is our view that the open exchange of information is at the heart of the concept of science and that every scientist must be responsible for completely documenting his or her own work. In addition, we maintain that scientists have a right to complete information on the work of fellow scientists. We believe the present situation in ecological research calls for a significant turnaround towards sharing information both within the scientific community and with the public. In this letter, we propose some measures to improve the current status of documentation of ecological models.

We propose that a common model documentation etiquette be adopted and that the resulting model documentations are **conveniently accessible, complete and mutually comparable**. Without these three features model documentations are of no general use. Convenient access to incomplete information does not help much, nor does information given in an inconsistent or nonstandard format, as comparison with related or previous work would not be possible. For practical reasons, however, a single documentation system with all three features may not be feasible and we may instead need multiple documentation systems, each with different features. For this reason, and based on our experiences with model documentation within our own information systems, we propose a hierarchy of model information systems that includes three different levels of model information, ordered by increasing degree of detail.

The **first level** would include general information about a model such as an abstract, contact address, references, links (URL's) to detailed information about the models and perhaps remarks by the model builder and/or model users. The **second level** of model information would encompass the general information included in the first level plus detailed scientific and technical information and some basic information on the mathematics of the model. A **third and most detailed level** would include all the general, scientific and technical information in the second level, plus detailed information on the mathematics of the model.

[*]University of Kassel, Faculty of Agriculture, International Rural Development and Environmental Protection (FB 11), Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany, e-mail: benz@wiz.uni-kassel.de

[†]GSF – National Research Center for Environment and Health, Ingolstädter Landstr. 1, 85764 Oberschleissheim, Germany, e-mail: knorren@gsf.de

Of course, such model information systems will not make journal publications on models obsolete. Modelling results published in scientific journals do not, in most cases, fit into a model information system. Journal articles tend to focus on the scientific interpretation of the findings, not the model itself. Nevertheless, interpretation of the findings by people other than the authors (e.g. an article's reviewer!) requires detailed knowledge about the model that is not normally included in the article itself (and for good reason). For most ecological models typically in use today, full documentation within a journal article would be impractical as it could ostensibly require far more pages than the scientific discussion of the model. Full documentation should, however, be readily available elsewhere to those who want and/or need it.

However, we do believe that the minimum model information given in journal articles should comprise

- a standard minimum level of information, equivalent to the first level of documentation, described above and
- references to information systems, which provide a detailed documentation of the model,

for a paper on a model to be accepted in a refereed journal. Papers submitted for publication which do not contain this minimal information cannot, in our view, be properly critically reviewed and therefore would have only limited scientific value. Likewise the model documentation in information systems should point to articles in journals, which describe scientific interpretation and applications of the models.

For the specific details on the three levels of documentation we suggest to take advantage of the following model information systems:

- REM - Register of Ecological Models
  which provides information of the level one and the possibility for navigation
  (see http://dino.wiz.uni-kassel.de/ecobas.html)
- UFIS - the environmental research information system
  which satisfies the demand of level two
  (see http://www.gsf.de/UFIS/ufis/)
- ECOBAS - documentation of mathematical formulation of ecological processes
  which gives detailed information about the mathematics
  (see http://dino.wiz.uni-kassel.de/ecobas.html)

These three systems give rise naturally to the following hierarchy of nested model information which satisfies our requirements as outlined above.

1. minimum information and navigation: REM
2. intermediate level of detailed model information: UFIS
3. complete information: UFIS+ECOBAS

For an overview of the information stored in these systems, see Fig. 1.

Modellers can enter information about their models into these systems by filling out questionnaires available on the WWW. Parts of the questionnaires must be filled out using given keywords creating a standardization that allows users of the system to perform quick searches. Because all questionnaires in a single information system have the same layout, models can be easily compared within each system. Completed questionnaires are stored in the information systems and are available to the public on the WWW.

The general information required by REM, that we feel should be a minimum requirement for publication, does not put a heavy burden on the modeller (the questionnaire takes about 15 minutes to fill out) and we feel that this additional work should be considered a service to the scientific community. If we in the ecological modelling community can agree on a common model documentation etiquette, the benefits that the individual modeller would reap due to increased access to model information would more than make up for the relatively small amount of time that the modeller has to invest in documenting his/her own models.

The third level of this information hierarchy satisfies our goals of easy access (WWW), completeness and comparability (by structuring and key wording of the questionnaires). Of course, the time involved to fill out the questionnaire at the third level (UFIS + ECOBAS) is much greater than at the first level (REM). However, experience shows that modellers profit from filling out such a detailed questionnaire as it may give new insight into the structure of a model and the processes being modeled. Such insight can be useful in further process studies and further development of the model.

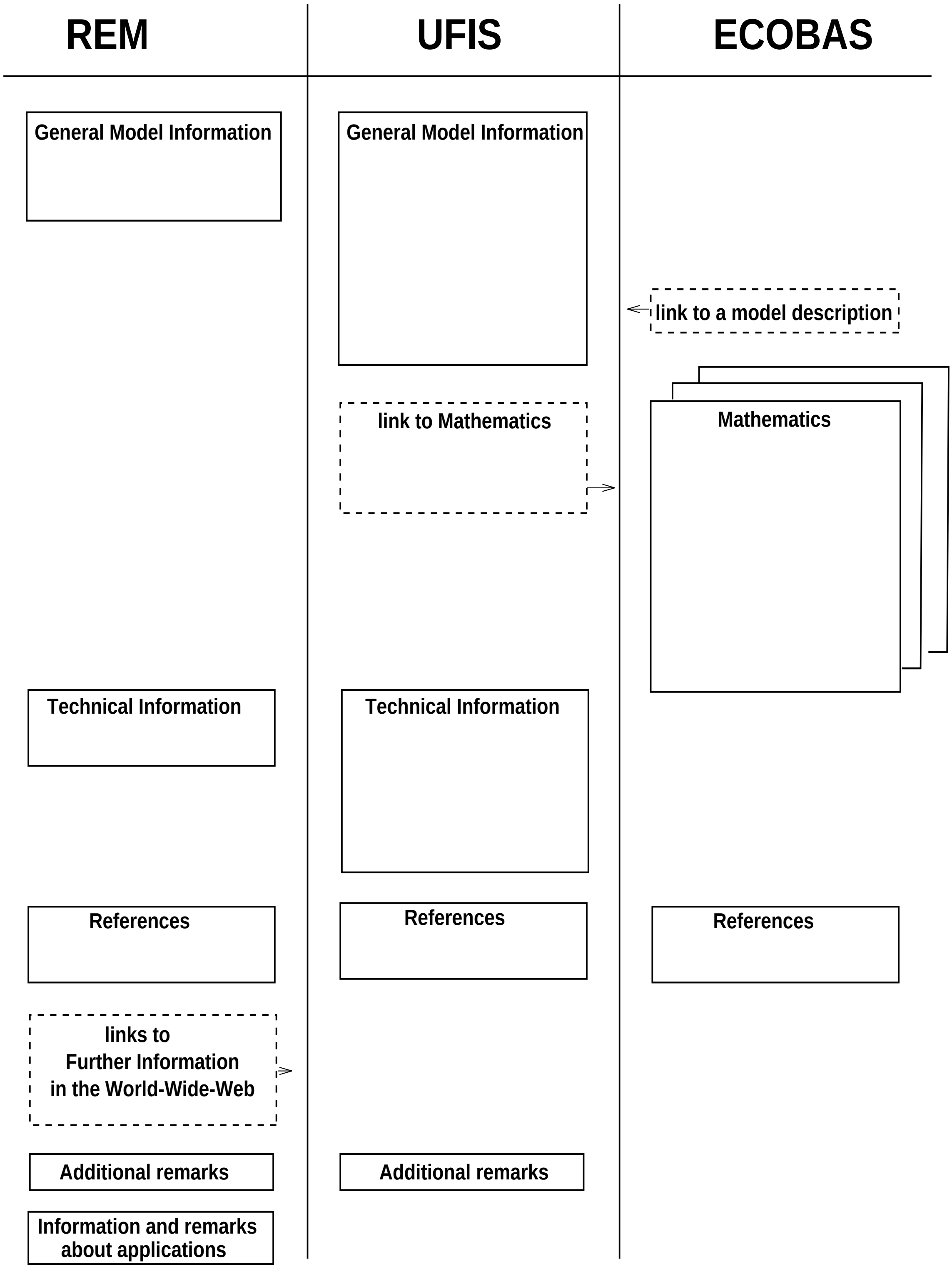

Figure 1: Hierarchy of model documentation in REM, UFIS and ECOBAS