Available online at www.sciencedirect.com

ScienceDirect

Environmental Modelling & Software 22 (2007) 1557–1571

Environmental Modelling & Software

www.elsevier.com/locate/envsoft

A new approach to testing an integrated water systems model using qualitative scenarios

T.G. Nguyen [a,b,*], J.L. de Kok [a], M.J. Titus [c]

[a] *Water Engineering and Management, Faculty of Engineering Technology, University of Twente, PO Box 217, 7500 AE, Enschede, The Netherlands*
[b] *Faculty of Hydro-meteorology and Oceanography, Hanoi University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*
[c] *Department of Human Geography of Developing Countries, Faculty of Geo Science, Utrecht University PO Box 80115, 3508 TC Utrecht, The Netherlands*

Received 14 January 2005; received in revised form 12 February 2006; accepted 17 August 2006
Available online 16 April 2007

Abstract

Integrated systems models have been developed over decades, aiming to support the decision-makers in the planning and managing of natural resources. The inherent model complexity, lack of knowledge about the linkages among model components, scarcity of field data, and uncertainty involved with internal and external factors of the real system call their practical usefulness into doubt. Validation tests designed for such models are just immature, and are argued to have some characteristics that differ from the ones used for validating other types of models. A new approach for testing integrated water systems models is proposed, and applied to test the RaMCo model. Expert knowledge is elicited in the form of qualitative scenarios and translated into quantitative projections using fuzzy set theory. Trend line comparison of the projections made by the RaMCO model and the qualitative projections based on expert knowledge revealed an insufficient number of land-use types adopted by the RaMCo model. This insufficiency makes the model inadequate to describe the consequences of the changes in socio-economic factors and policy options on the erosion from the catchment and the sediment yields at the inlet of a storage lake.

Keywords: Land use change; Soil loss; Sediment yield; RaMCo; Fuzzy set; Scenario; Validation; Testing

1. Introduction

As every model is an abstraction of a real system, model developers and model users have to struggle with the question of how to develop and evaluate a model (see Jakeman et al., 2006). This methodological problem is argued to be rooted in the controversial debate on justification, verification of scientific theories, and of models in a philosophical perspective (Barlas, 1994; Kleindorfer et al., 1998). The usefulness of the endeavour to prove the validity of any predictive model of a natural system (open system) has been questioned (Konikow and Bredehoeft, 1992; Oreskes et al., 1994). Several authors have suggested that model validity should always be considered within the model's applicability domain or model context (Rykiel, 1996; Refsgaard and Henriksen, 2004). In addition, the purposes of a model are essential in the selection of appropriate validation tests (Nguyen and De Kok, 2003). Depending on different classification criteria, model validation tests can be categorised as qualitative or quantitative, formal or informal, static or dynamic, conceptual or operational, and so on. Traditional statistical methods are proved to have a limited capacity in testing integrated dynamic models (Forrester and Senge, 1980). One of the reasons is that both system dynamics models and integrated water systems (IWS) models do not strive for prediction of future values; that is, not for ''point-prediction''. These models should predict certain aspects of behaviour in the future. Examples include pattern-prediction and event-prediction. Another reason is that statistical tests hardly say anything about the structural errors within a model. The problem of equifinality (Refsgaard and Henriksen,

* Corresponding author. Faculty of Hydro-meteorology and Oceanography, Hanoi University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam. Tel.: +84 4 2173940; fax: +84 4 8583061.

E-mail addresses: giangnt@vnu.edu.vn (T.G. Nguyen), j.l.dekok@ctw.utwente.nl (J.L. de Kok).

1364-8152/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2006.08.005

2004)—structural errors and errors in parameter estimation compensating for each other—is often encountered. This is even more of a problem in the case of integrated models in which many submodels are linked together to predict management variables.

Integrated systems models (ISM) and integrated water systems (IWS) models have been developed over decades, aiming to support decision-makers in the planning and managing of natural resources. Without effective validation, the design of an IWS model remains an art rather than a science. Validation of IWS models is useful for their theoretical improvement. Moreover, validation is necessary prior to any practical implementation of these models. Inherent model complexity, scarcity of field data, and uncertainty over internal and external factors of the real system make the validation of an IWS model a difficult task. Furthermore, the poor predictive ability of the historical data to describe future situations in the complex system involved with social and economic factors hinders the effectiveness of available validation techniques. On the other hand, due to their characteristics, validation tests for IWS models can go beyond the tool kit of available validation tests for conventional process models (Forrester and Senge, 1980; Beck and Chen, 2000). Therefore, the validation of IWS models is likely to depend less on conventional and classical tests, and more on integrated validation tests that are yet to be developed (Parker et al., 2002). In this paper, a new approach for testing IWS models is developed and applied to validate the RaMCo model. The approach is designed to test the capability of the model to describe the dynamic behaviour of system output variables under a variety of possible socio-economic scenarios and policy options. The sediment yield at the inlet of the Bili-Bili dam, one of several state objective variables in the model, is selected as a case example.

This paper is organised as follows. Section 2 starts with a review of the representative frameworks, approaches and techniques for model validation. Following in this section is an overview of a new approach to testing IWS models and a detailed description of this approach. The case study is then introduced in Section 3, in which the conceptual model, the mathematical equations used in RaMCo to model land-use change dynamics, the link to soil loss computation, and the sediment yield at the inlet of the storage lake are explained. Section 4 describes the process of formulating the qualitative experts' scenarios. Translating these qualitative scenarios into quantitative projections of objective variables using fuzzy set theory is demonstrated in Section 5. The comparison of the projections based on experts' knowledge and RaMCo projections in terms of trend lines is presented in Section 6. The paper is concluded with a discussion on the usefulness of the proposed validation approach and recommendations for further improvement of the RaMCo model.

2. Validation methodology

2.1. Literature review

This section presents a review of the representative frameworks, approaches and techniques for model validation which can be found in scientific literature dating back to the 1980s. The models to be validated, which are included in this review, consist of simulation models in operational research (Shannon, 1981; Sargent, 1984, 1991; Balci, 1995; Kleijnen, 1995; Fraedrich and Goldberg, 2000), models in earth sciences (Flavelle, 1992; Ewen and Parkin, 1996; Beck and Chen, 2000), agricultural models (Mitchell, 1997; Scholten and ten Cate, 1999), ecological models (Van Tongeren, 1995; Kirchner et al., 1996; Rykiel, 1996; Loehle, 1997), system dynamics models (Forrester and Senge, 1980; Barlas, 1994; Barlas and Kanar, 1999) and integrated models (Finlay and Wilson, 1997; Beck, 2002; Parker et al., 2002; Poch et al., 2004; Refsgaard et al., 2005). The controversial debate on terminologies for model validation (Oreskes et al., 1994; Oreskes, 1998; Rykiel, 1996; Beck and Chen, 2000) points to the ambiguity and overlap between the terms: model testing, model selection, model validation or invalidation, model corroboration, model credibility assessment, model evaluation and model quality assurance. To counter the ambiguity of the terminology, a clear definition of our approach to testing ISW models is given in Section 2.2.

The most common framework for model validation, which is widely accepted in the modelling community, can be attributed to Sargent's work (1984, 1991). Sargent considered model validation as substantiation that a computerised model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model. In this framework, the validity of a simulation model consists of three dimensions: conceptual validity, operational validity and data validity. To determine the conceptual validity of a model, two supplementary approaches are often used. The first approach is to use mathematical and statistical analyses (e.g. correlation coefficient, Chi-square test) to test the theories and assumptions (e.g. linearity, independence) underlying the model. The second approach is to have an expert or experts evaluate the conceptual model in terms of both the model logic and its details. This approach is often referred to as *peer review*, and is aimed at determining whether the appropriate details, aggregation level, logic, mathematical and causal relationships have been used for the model's intended purpose. Two common techniques used for the second approach are face validation and traces (Sargent, 1984, 1991). Operational validity, in Sargent's term, is primarily concerned with determining that the model's output behaviour has the accuracy required for the model's intended purpose over the domain of its intended application. Three conventional approaches for operational validation based on the comparison of model output and observed data are graphical comparison, hypothesis testing and confidence intervals (Sargent, 1984). In addition, two other comparison approaches, using goodness-of-fit statistics (e.g. root mean square) and residual analysis between model output and observed data, are mentioned by Flavelle (1992). These common approaches based on the comparison between model output and observed data are often referred to as *history-matching* (Beck, 2002). More techniques developed for operational validation, which range from qualitative, subjective, informal tests (e.g. face validity of model behaviour) to quantitative, objective and formal tests (e.g. statistical tests) are described in Sargent (1984), Balci (1995), Kleijnen (1995), Rykiel (1996), Mitchell (1997), Scholten and ten Cate (1999) and Fraedrich and Goldberg (2000). It is important to emphasise that the relevance of the available validation approaches and techniques depends on the availability of field data and the level of understanding of the system studied (or scientific maturity of the underlying disciplines), as recognised by Kleijnen (1995), Rykiel (1996) and Refsgaard et al. (2005). Furthermore, the requirement of validity of a model under a set of experimental conditions under which the model is intended to be used is emphasised and studied by several authors (e.g. Ewen and Parkin, 1996; Kirchner et al., 1996). Ewen and Parkin (1996) proposed a 'blind' testing approach to the validation of the catchment model to predict the impact of changes in land-use and climate, given the limitations of existing approaches, such as the simple split-sample testing, differential split-sample testing, proxy-catchment testing and differential proxy-catchment testing. This 'blind' testing approach, however, does not consider the interactive natural-human systems which are more complex and qualitative in nature.

Another conceptual framework for the validation of system dynamics models has been suggested by Forrester and Senge (1980). Within this framework, validation is defined as the process of establishing confidence in the soundness and usefulness of the model. According to these authors, model validity is equivalent to the user's confidence in the usefulness of a model. The confidence of the model users is gradually built up after each successful validation test. Validation tests are divided into three major groups: tests of model structure, tests of model behaviour and tests of policy implication. Particular validation tests have been proposed, corresponding to each group. The important characteristics of this conceptual framework are: the focus of validation on the structure of the model system, the vital roles of the experts' knowledge/experience and qualitative, informal tests (e.g. extreme condition test and pattern test) in the validation process. These characteristics are reflected by the extensive use of terms such as soundness, plausibility and confidence. Barlas (1994), Barlas and Kanar (1999) separates validation tests into two main groups: direct structure testing and indirect structure (or structure-oriented behaviour) testing. Perceiving that pattern prediction (period, frequencies, trends, phase lags, amplitude) rather than point prediction is the task of system dynamics models, he has developed formal statistics and methods which can be used to compare the simulated behaviour patterns with either observed time series or anticipated behaviour patterns. In line with this philosophical perspective on model validation, Shannon (1981) proposed a similar conceptual framework for the validation of simulation models in operational research. The differences in Shannon's framework are the integration of verification and validation, and an extensive inclusion of the formal, quantitative, statistical approaches to model validation. A closely related framework for the validation of ecosystem models is proposed by Loehle (1997), in which a new version of the hypothesis testing approach is considered to be essential for the validation of ecological models.

As the complexity of integrated models used in decision-making increases, the usefulness of quantitative validation approaches based on the comparison between model output and observed data decreases. This is due to the scarcity and uncertainty of field data for model calibration and validation. The model validation using peer review is also challenged by the conflict of interests of the peers and the limited number of capable peers, due to the multidisciplinary nature of the integrated models (Beck, 2002; Parker et al., 2002). These foster a shift of model validation perspective from scientific theory testing to evaluating the appropriateness of the model as a tool designed for a specified task. In accordance with this view, the two supplementary approaches, which have just begun to develop, are: (i) judging the trustworthiness of the model according to the quality of its design in performing a given task, and (ii) using the information (experience) obtained from the interactions and dialogues between the modellers and a variety of system experts (resource managers, scientific experts) and stakeholders. An example of the former approach is given by Beck and Chen (2000), in which the model quality is judged, based on the properties of internal attributes—the number of key and redundant parameters. Although the need for the latter approach to model validation is recognised (Beck and Chen, 2000; Parker et al., 2002; Poch et al., 2004; Refsgaard et al., 2005) appropriate tools and methods have not been developed yet.

In summary, although the literature on model validation is abundant most of the available techniques and approaches focus on quantitative tests for operational validation (or historical matching), given that the observed data are available. The conceptual validity or structural validity, which is equally important for integrated models, has been neglected. There is a lack of consideration of the uncertain future conditions, under which the model is intended to be used in model validation frameworks. In addition, there is little attention to the qualitative nature of social science, which is often required to be incorporated in integrated systems models to support the decision-making process.

2.2. Overview of the new approach

The design of our new approach was motivated by the three reasons that limit the relevance of the conventional approaches to the validation of IWS models: (i) the limited predictive ability of historical data to describe the future behaviour of interactive natural-human systems, (ii) the qualitative nature of the social science and (iii) the scarcity of field data for model validation. The new approach proposed in this paper is established to determine whether a model is ill or well designed, with regard to the purpose of an IWS model as a tool capable of reflecting the system experts' consensus about the dynamic behaviour of the system output variables, under a set of possible socio-economic scenarios and policy options. The proposed approach acknowledges that we cannot develop any model which is a true representation of the real system. Validation tests should be designed to unravel the incompleteness of or errors in a model in the view of the system experts. The ultimate objective of IWS model validation, according to Forrester and Senge (1980), is to obtain a better model, which has sound theoretical content (model structure) and can fulfil its intended purpose(s). One aspect of model validation is to determine whether a model is ill or well designed for its purpose (Beck and Chen, 2000). The validity of a model cannot be achieved by conducting a single test, but a series of successful tests could increase the users' confidence in the model's usefulness.

The underlying principle of the new approach is that system experts are asked to make an artificial closed system (hypothesised system) with the system's components, prescribed system inputs (drivers), driving mechanisms, and the qualitative response of system's outputs in the form of qualitative scenarios. Fuzzy logic is applied to produce quantitative projections of the output variables from qualitative descriptions of the hypothesised system. The creation of the hypothesised system provides a platform on which "experiments" can be conducted to obtain the system's outputs under the feasible sets of system's inputs. In each experiment, the socio-economic factors and policy options are input by the experts, reflecting one possible future description of the real system. The comparisons of the trend lines between the two systems' outputs under different scenarios are made to arrive at the plausibility of the model structure and the validity of the assumptions. Thus, an obvious difference between the outputs produced by the two systems, in terms of trend lines, can reveal the structural faults of the model system. Otherwise, the model is said to pass the current test. The procedural steps to build an experts' hypothesised system, to use qualitative scenarios and a fuzzy rule-based method to make quantitative projections of system behaviours are presented in the following subsection.

2.3. The detailed description

There are three phases to be taken during the testing process of an IWS model using qualitative scenarios: (1) formulating experts' qualitative scenarios; (2) translating the qualitative scenarios; (3) conducting simulations by the IWS model and comparing the outputs produced by the two systems in terms of trend lines.

2.3.1. Formulating experts' qualitative scenarios

In the context of this paper, a scenario is defined as 'a description of a future situation and the course of events which allows one to move forward from the original situation to the future situation' (Godet and Roubelat, 1996). Qualitative scenarios describe possible futures in the form of words or symbols while quantitative scenarios describe futures in numerical form (Alcamo, 2001). The common understanding is that a scenario is not a prediction of the future, but an alternative image of how the future might unfold. The purpose of scenarios is manifold. Some of them are: illustrating how alternative policy pathways can achieve an environmental target, identifying the robustness of policies under different future conditions, providing the non-technical audience a picture of future alternative states of the environment in an easily understandable form (narrative description), and providing an effective format on which information in both qualitative and quantitative forms can be assimilated and represented. In this paper, scenarios are proposed as testing experiments to test the capability of an IWS model to describe the consequences of possible socio-economic conditions and policy options on the management variables.

A good scenario should be relevant, consistent (coherent), probable and transparent. In principle, only a few substantially different scenarios are needed. Although different authors (Von Reibnitz, 1988; Van der Heijden, 1996; Alcamo, 2001) developed somewhat different procedures and terminologies for the scenario building, these procedures share the same iterative form and have the following steps in common:

(1) Establishing a scenario building team and defining the goals of scenarios
(2) Analysing data and studying literature
(3) Specifying driving forces and driving mechanism (structuring scenarios)
(4) Developing the storylines (scenarios in narrative form)
(5) Testing the internal consistency of scenarios

In applying scenarios for testing IWS models, the composition of the scenario building team (step 1) and testing the consistency of scenarios (step 5) are particularly important, and require more elaboration.

The participatory approach to scenario building, which is widely acknowledged, requires a wide spectrum of knowledge and opinions from multidisciplinary team members (Schwab et al., 2003; Van der Heijden, 1996). In developing scenarios used in international environmental assessment, Alcamo (2001) recommends having two building teams: a scenario team and a scenario panel. The former, which consists of the sponsors of the scenario building exercise and experts, should include around three to six members. The latter, which consists of stakeholders, policymakers and additional experts, should include around 15–25 members. For the purpose of testing IWS models, we propose to distinguish two groups in the scenario building team. The first group includes model developers (they are also interdisciplinary scientists), experts (scientists who may have different views about the model system) and additional analysts (scientists who are not involved in the model building). The second group consists of multidisciplinary experts, resource managers and stakeholders. The second group can play a role both as the fact-contributor and scenario evaluator in the scenario building for the testing of IWS models. Preferably, the stakeholders and resource managers should participate at the beginning of the scenario building process (steps 1–3).

In the iterative scenario building process, the consistency of the scenarios needs to be tested. Van der Heijden (1996) and Alcamo (2001) recommend two similar approaches to establishing the consistency of scenarios, which include two supplementary tests: scenario-quantification testing and actor-testing. Quantification testing comprises quantifying the scenarios and examining the quantitative projections of the system indicators (management variables). Actor-testing diagnoses the inconsistencies by confronting the internal logic of the qualitative scenarios with the intuitive human ability to guess at the logic of the various actors (stakeholders, resource managers and additional experts). We propose to use physical, biological constraints (e.g. the total available area of a watershed) to check the quantitative projections (e.g. the projections of the areas of different land-use types) for quantification testing. In actor-testing, both the narrative descriptions of the scenarios and the quantitative projections of the system indicators should be communicated to the second group (stakeholders, resource managers and additional experts) by means of report papers, workshops and the internet.

2.3.2. Translating qualitative scenarios

For the translation of qualitative scenarios, the application of fuzzy set theory is proposed. Fuzzy set theory was originally developed by Zadeh (1973), based on the concepts of classical set theory. The essential motivation, as he claimed, for the development of fuzzy set theory is the inadequacy and inappropriateness of conventional quantitative techniques for the analysis of mechanistic systems (e.g. physical systems governed by the laws of mechanics) to analyse humanistic systems. The design of a fuzzy system comprises five steps (Mathworks, 2005), which can be reduced to four main steps (De Kok et al., 2000):

(1) Translation of the independent and dependent variables from numerical into the fuzzy domain (fuzzification)
(2) Formulation of the conditional inference rules
(3) Application of these rules to determine the fuzzy outputs
(4) Translation of the fuzzy outputs back into the numerical domain (defuzzification)

In order to test the internal consistency of scenarios, scenario quantification-testing needs to be conducted. Therefore, the process of scenario translation is extended to include step 5 (testing the internal consistency of scenarios). The five steps are demonstrated by the application described in Section 5.

2.3.3. Conducting simulations by the IWS model and comparing the results

After translating the qualitative scenarios into quantitative projections of the output variable, simulations are conducted with the IWS model. A comparison of the output behaviour produced by the two systems in terms of trend lines is carried out. This phase is demonstrated in Section 6.

It is our experience that the interactive communication within the first group (experts, model developers and analysts) should be carried out during all three phases (qualitative scenario building, scenario translating and comparing results). In doing so, any disagreement between model developers and experts can be brought up for discussion at every step. In this way, the experts' bias or inconsistency can be minimised.

3. The RaMCo model

In 1999, a 4-year multidisciplinary programme for sustainable coastal zone management in the tropics was concluded with the presentation of a methodology for integrated policy analysis. In the framework of the project, a Rapid assessment Model for integrated Coastal zone management (RaMCo) was developed (Uljee et al., 1996; De Kok and Wind, 2002). The RaMCo model allows for the analysis and comparison of different management alternatives under various socio-economic and physical conditions, i.e. performing what-if analysis. It is intended to be used as a platform to facilitate discussions between scientists and decision-makers at the intermediate level of analysis (i.e. rapid assessment). The selection of possible sets of measures from larger available alternatives at this analysis level can be followed by the comprehensive analysis, which is not the task of RaMCo (De Kok and Wind, 2002). The coastal zone area of Southwest Sulawesi in Indonesia serves as the study area.

The study area for RaMCo occupies a total area of about 8000 km^2 (80 km $\times$ 100 km), of which more than half is on the mainland (De Kok and Wind, 2002). The offshore part covers the Spermonde archipelago where multi-ecosystems such as coral reef, mangrove and seagrass can be found. On the mainland, the city of Makassar has a fast-growing population of 1.09 million (1995), which is expected to double in 20 years. In the upland rural area, the forest area is rapidly declining, due to the increase in cultivated land. The expansions of urban areas and the conversion of uncultivated to cultivated land are imposing a strong demand on the effective management of water and other ecological systems in the coastal area.

To meet the rapidly increasing demand for water supplies for domestic use, industry, irrigation, shrimp culture and the requirements for flood defence of the city of Makassar, the construction of a multi-purpose storage lake started in 1992. The dam was closed for water storage in November 1997 (Suriamihardja et al., 2001). The watershed of the Bili-Bili dam covers the total area of 384 km^2, which represents the upper part of the Jeneberang river catchment. The dam was designed to have an effective storage capacity of 346 million m^3 and dead storage capacity of 29 million m^3 (CTI, 1994). Its expected lifetime of 50 years was determined by computing the total soil loss due to erosion of the watershed surface. The computation was carried out using the universal soil loss equation (USLE) in combination with the land cover map surveyed in 1992. No future dynamic development of land-use in the watershed area was taken into consideration. Analyses of recently measured sediment transport rates at the inlet of the Bili-Bili dam and land-use maps show an obvious decrease in the storage capacity of the dam, due to increasing sediment input (CTI, 1994; Suriamihardja et al., 2001). This calls for a proper land-use management strategy to minimise the sediment eroded from the watershed surface that runs into the reservoir.

RaMCo quantitatively describes the future dynamic land-use and land-cover changes under the combined inference of socio-economic factors. Then, the resulting soil losses from the watershed surface and the resulting sediment yields at the inlet of the Bili-Bili dam are computed. The following are conceptual and mathematical descriptions of this integrated model.

3.1. Land-use/land-cover change model

3.1.1. Land-use types

During the design stage, a problem-based approach was followed to select relevant land-use-types (De Kok et al., 2001). In RaMCo, a distinction was made between static land-use types (land-use features) and active land-use types (land-use functions). Land-use features such as beach, harbour and airport are expected to be relatively stable in their size and location over the time frame considered. Land-use functions such as industry, tourism, brackish pond culture, rice culture and others are expected to change both in space and over time under the influence of various internal and external driving factors (drivers). In this paper, attention is paid to the two land-use types: nature and mixed agriculture. The model treats the "nature" land-use type as the uncultivated land which is a combination of natural forest, production forest, shrubs and grasses. Mixed agriculture represents food crop culture (other than rice culture) such as maize, cassava and cash crops such as coffee and cacao. These types of land-use predominate in the Bili-Bili catchment and are expected to change rapidly, affecting the amount of sediment transported into the reservoir. In addition to the two defined categories, three other land-use types exist in RaMCo: namely, rural resident, rice culture and inland water.

3.1.2. Drivers of land-use changes: temporal dynamics versus spatial dynamics

The drivers of land-use changes in the RaMCo model can be separated into three categories: (i) socio-economic drivers, such as price, cost, yield, technology development and demography; (ii) management measures, such as reservoir building and reforestation; and (iii) biophysical attributes, such as soil types and road networks. The first two groups of drivers, in combination with the availability of irrigated water and suitable land, determine the rate of land-use change (temporal dynamics), while the final group determines places where the changes take place (spatial dynamics). The rate of change in area for each land-use type is computed by a so-called macro-scale model, which is discussed in more detail below. In the micro-scale model, the spatial allocations of these changes are determined by adopting the constrained cellular automata (CCA) technique. A full description of this technique is outside the scope of this paper. Those who are interested in the details of the CCA approach and the model structure are referred to White and Engelen (1997) and De Kok et al. (2001).

3.1.3. Macro-scale model

As mentioned above, the macro-scale model computes the rates of change, i.e. land demand for different land-use types. Since this paper focuses on land-use change and the resulting soil loss in the Bili-Bili watershed area, only three land-use types are discerned in the following section, namely mixed agriculture, rice culture and nature. Inland water and rural residential land-use types are excluded because of the small portions of land they occupy in the basin and their relative stability in size and location.

For agricultural land-use, following the assumption that land demand is proportional to the net revenue per unit area, the rate of change in land-demand can be computed as (De Kok et al., 2001):

$$\Delta A(t) = \alpha(p(t)y(t) - c(t))A(t)\left[1 - \frac{Z(t)}{Z_{\text{tot}}}\right] \tag{1}$$

where $\Delta A(t)$ and $A(t)$ are the rate of change and area of mixed agriculture at time t, $p(t)$ and $c(t)$ are price and production cost per unit area, and $y(t)$ is the yield which can accommodate technological changes. The growth coefficient α was calibrated using statistical data on the above defined variables. The variable $Z(t)$ is the sum of geographical suitability for agriculture over all the cells occupied by agriculture at time t, and Z_{tot} is obtained by extending the sum over all the cells on the map. The use of these variables ensures that expansion ceases if the maximum suitable area is approached.

For rice culture, Eq. (1) is still applicable, but rice yields are obtained in a different way to account for the irrigation function of the storage lake:

$$y_{\text{rice}}(t) = f(V)\eta(t)y_{\text{irr}} + (1 - f(V)\eta(t))y_{\text{nirr}} \tag{2}$$

In Eq. (2), y_{irr} and y_{nirr} are the maximum yields of rice culture with and without irrigation, respectively. The dimensionless function $f(V)$ has a value ranging from 0 to 1, and reflects the irrigation priority using the actual and maximum volumes of the storage lake. The variable $\eta(t)$ denotes the spatial fraction of rice fields which can be irrigated.

The land demand of "nature" land-use type is computed by:

$$\Delta A_n(t) = \alpha A_n(t)\left[1 - \frac{Z_n(t)}{Z_{n,\text{tot}}}\right] + \delta_n(t) \tag{3}$$

where α is the natural expansion rate of nature (forest), and $\delta_n(t)$ accounts for the area of reforestation at time t, a management variable.

According to these equations, each sector can expand until the maximum suitable area is reached. This allows for a situation where more or less land is allocated to all the sectors taken together than the total available land. Thus, an allocation mechanism has been introduced. If the total computed land demand is less than the available land, the allocated land equals the demand for these sectors. The remainder is assigned to nature (forest). If the total computed land demand for all sectors exceeds the available area, the allocated land for each sector is normalised as follows:

$$A_i(t) = \frac{A_{\text{available}}}{\sum \overline{A_i(t)}}\overline{A_i(t)} \tag{4}$$

where $A_i(t)$ and $\overline{A_i(t)}$ are allocated land and computed land demand for land-use type i, respectively.

3.2. Soil loss computation

To couple the process of land-use changes to predict the sediment yields at the outlet of the Bili-Bili watershed area, the universal soil loss equation (USLE) in spatially distributed form is used. The original USLE (Wischmeier and Smith, 1965) has the following equations:

$$A = R \times K \times L \times S \times C \times P \tag{5}$$

where A is the computed soil loss per unit area, expressed in metric tons/ha; R is rainfall factor, in MJ-mm/ha-h and MJ-cm/ha-h if rainfall intensities are measured in mm/h and cm/h, respectively; K is the soil erodibility factor, in metric tons-h/MJ-cm; C is a cover management factor (−); P is a support practice factor (−); L is the slope length factor, in m; and S is the slope steepness factor. The product of L and S is computed by:

$$LS = \left(\frac{\lambda}{22.13}\right)^m \left(0.0065s^2 + 0.045s + 0.065\right) \tag{6}$$

in which λ is the field slope length, in m, and m is the power factor whose value of 0.5 is quite acceptable for the basin with a slope percentage of 5% or more (Wischmeier and Smith, 1978); s is the slope percentage.

The RaMCo model allows the use of spatial databases to facilitate the computation of soil erosion from individual (400 m × 400 m) mesh cells. Maps containing factors on the right-hand side of Eq. (5) are referred to as factor maps. These factors maps were derived from spatial databases such as topographic maps, geological maps, land-cover maps and isohyetal maps (CTI, 1994). Eqs. (5) and (6) are used to compute soil loss from every cell in the map.

3.3. Sediment yield

To predict sediment yields at the outlet of the watershed, the gross erosion-sediment delivery method (SCS, 1971) is used in combination with the USLE. The gross erosion (E), expressed in metric tons, can be interpreted as the sum of all the water erosion taking place, such as sheet and rill erosion, gully erosion, streambank and streambed erosion as well as erosion from construction and mining sites (SCS, 1971). According to the previous study on sediment in the Jeneberang river (CTI, 1994), the sediment consists mainly of washload caused by sheet and rill erosion. Moreover, sand pockets and Sabo dams were designed to trap coarser sediment resulting from other types of erosion. Thus the neglecting of other erosion types is acceptable with respect to our purpose of estimating the sediment yield at the inlet of the Bili-Bili Dam site. The sediment yield (S_y), the amount of soil routed to the outlet of the catchment in metric tons per ha, can be computed by multiplying the gross erosion (E) by the sediment delivery ratio:

$$S_y = E \times \mathrm{SDR} \tag{7}$$

where SDR is the sediment delivery ratio, which depends on various factors such as channel density, slope, length, land-use, and the area of the catchment. Methods have been proposed in the past to estimate the SDR (SCS, 1971). This research adopts the values established in Morgan's (1980) table (CTI, 1994), which is widely used in Indonesia. In order to identify the areas that are susceptible to erosion for the development of soil conservation strategies, the whole basin was subdivided into eight sub-basins. Eq. (7) is applied to each sub-catchment, and the sediment yields are added together to obtain the total sediment yield running into the reservoir.

4. Formulation of scenarios

The iterative process of qualitative scenario formulation commonly has five steps (Section 2.3). In step 1 (establishing a scenario building team) of this exercise, two groups were distinguished. The first group consisted of a model developer, an expert and an analyst. The second group consisted of around 20 local scientists and potential end-users of RaMCo. Due to practical reasons (e.g. distance, finance), the second group only participated intensively in step 5 (testing the consistency of scenarios) of the current exercise. In step 2, data collection and a historical study were carried out for the study area as well as for other regions (e.g. Yogyakarta and Sumatra) in Indonesia. In this section, steps 3 and 4 (structuring scenarios and developing qualitative scenarios) are described. Since step 5 (testing the consistency of scenarios) is involved with scenarios quantification, it is described at the end of Section 5.

The three qualitative scenarios described here are the accumulated results of research carried out by 12 Dutch MSc students, in collaboration with the local experts in Hasanuddin University (UNHAS) in Makassar. The reports and theses of these students are based on primary data and secondary data collected in the villages, the district capitals and in Makassar, and include the analysis of both household interviews and open interviews with local stakeholders and key persons. The expert only had the final responsibility for formulating the scenarios and related inference rules.

4.1. Structuring scenarios

As mentioned in Section 3, in the Bili-Bili catchment five land-use types were distinguished by modellers, which include nature (forest), agriculture, rice culture, rural residential land and inland water. This categorisation may or may not be sufficient to give a satisfactory description of the real system, given the specified purpose of the model. According to the expert, the separation of nature into forest and shrub and grassland, and the separation of mixed agriculture into dry upland farming and mixed forest garden are necessary to describe the effect of management measures on land-use changes and the resulting dynamic change of the soil erosion from catchment surface. Thus, the new hypothesised land-use system consists of five active types (forest, shrub and grassland, dry upland farming, mixed forest garden, paddy field) and two relatively static types, inland water and rural residential land.

The process of identifying the drivers (stimuli) and driving mechanism was carried out through extensive discussions within the first group (the model developer and the expert). The drivers and driving mechanism of the land-use system which resulted from these discussions are briefly described in Fig. 1.

4.2. Developing qualitative scenarios

Based on the purpose of the scenarios and the insights gained from field research, three qualitative scenarios were formulated for the dynamic land-use system in the Jeneberang catchment. Scenario A reflects an extrapolation of the socio-economic, policy conditions and their effects on the land-use system under the Suharto presidency period (1967–1998). Scenario B represents the post-Suharto period (present situation), in which the forest is more open for logging and is invaded by subsistence farming due to the maximum economic growth objective and the lack of law enforcement from the government. In scenario C, a sustainable development option is projected in which an economic goal is achieved while the environmental issues are kept to a minimum through policy measures such as law, cheap credits and land-conversion programmes.

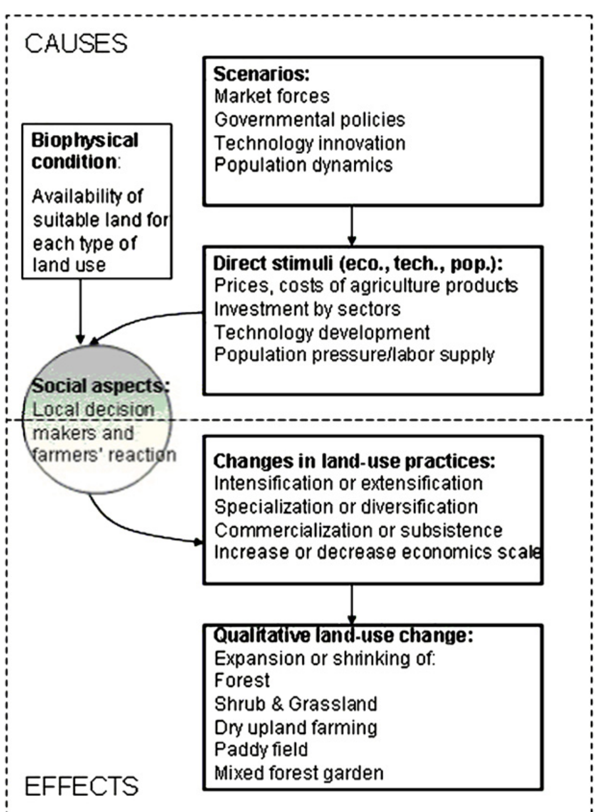

Fig. 1. Reasoning process underlying the scenario-based qualitative projection of the rates of land-use changes.

4.2.1. Scenario A: guided market economy

The guided market economy as developed during the New Order, has been based on strong government interferences and a bureaucratic approach, causing much abuse of power and funds and often leading to misinvestments. On the other hand, it should be acknowledged that government programmes focusing on the boosting of food production, infrastructure, public services (health and education) and industrialisation have had positive impacts in terms of employment creation and income improvements. Environmental conditions (pollution, deforestation and erosion) however, usually have been neglected, as have most issues of regional and social equity. This scenario is assumed to cause the following shifts and changes in land use practices:

- Forest: a gradual retreat of primeval and secondary forest fringes due to the progressive invasion by marginalised upland farmers in search for timber, firewood and land to cultivate food and cash crops
- Shrubs and grasses: expanding in the higher uplands because of the abandonment of exhausted and unproductive dry farming fields left in fallow. Retreating in the lower uplands through their conversion in mixed forest garden.
- Dry upland farming (*tegalan*): expanding *tegalan*-fields in the higher uplands because of land hunger of small peasants and the stimulation of dry food crop cultivation by government programmes.
- Mixed forest gardens: some expansion may occur by planting of lucrative tree crops like cocoa or clove. Most of this expansion will be realised on wasteland areas (shrub and grassland) or marginal *tegalan* fields at lower altitudes (<1000 m).
- Paddy fields (sawah): lack of irrigable land in the Jeneberang Valley and the long dry season are limiting the expansion opportunities for wet rice cultivation beyond the valley bottoms and lower slopes.

4.2.2. Scenario B: maximum growth

The maximum growth scenario is based on the principles of free trade, a facilitating government policy and the attraction of foreign and domestic corporate capital. Through the use of capital and technology, intensive modes of production and increasing economies of scale this will lead to higher levels of productivity and decreasing product prices. In agriculture, this implies that only the bigger farmers are able to draw advantage from this type of development (as only these farmers have enough land, capital and knowledge), whereas the smaller peasants have to revert to subsistence agriculture or labour intensive types of commercial farming with few inputs and low productivity levels.

- Forests: these are increasingly affected by the expansion of subsistence farming and commercial farming in dry upland areas due to processes of marginalisation among landless and small farmers, and the expansion of cash crop cultivation.

- Shrubs and grasses: this type of waste land probably will not change very much in total area for the same reasons as in scenario A.
- Dry upland farming: while there is continuing encroachment of dry upland farming into the forest fringes of the higher uplands, there also is an increase in the productivity of *tegalan* agriculture on existing fields. Total *tegalan* area, however, will only expand slightly due to the intensification of *tegalan* agriculture and the advancement of agro-forestry systems in the lower areas.
- Mixed forest gardens: a similar expansion of agro-forestry systems on the lower slopes and foothills of the Jeneberang Valley would be expected due to the drive for increasing perennial cash-crop production for the export market (i.e. coffee, cacao and clove).
- Paddy fields: few changes can be expected in terms of areal expansion, but productivity of wet rice fields is assumed to rise considerably due to capital investments by richer farmers in high-yielding variety, fertilisers and so on.

4.2.3. Scenario C: sustainable development

This sustainable development scenario is based on a selective operation of the market economy in combination with an active role of the government in securing principles and conditions of sustainability. With respect to agricultural land use this policy requires that farmers are both stimulated and controlled by environmental laws, extension programmes, cheap credits and (initial) subsidies on appropriate inputs. Furthermore, the government should actively support rural economic diversification by improving the rural infrastructure, public services and human resource development, in order to reduce dependency on agriculture and pressures on local natural resources.

- Forests: these will show a recovery, both in area and quality due to more strict regulations and controls on the use of existing forest areas (protected forest and production forest) and the reforestation of waste land areas (shrub and grassland).
- Shrub and grasslands: this wasteland area gradually will be reduced in size and improved by regreening projects. Reduction may also be achieved by converting the waste land areas into agro-forestry systems.
- Dry upland farming: *tegalan* agriculture of annual food crops will become more productive and sustainable through improved cultivation methods, including the integration of animal husbandry, crop diversification and terracing.
- Mixed forest gardens: programs for promoting the sustainable cultivation of perennial cash crops in mixed forest gardens will expand agro-forestry systems in the foothill areas of the valley (i.e. both in the marginal *tegalan* areas and the wasteland areas).
- Paddy fields: the irrigated paddy fields in this scenario will not expand very much for the same reasons as in the previous scenarios. Productivity probably will not increase as much as in scenario B, due to the limited use of chemical inputs.

5. Translation of qualitative scenarios

It is worth noting that depending on the analyst's view on how he interprets fuzziness, the details of step 1 (fuzzification) and 2 (formulation of inference rules) may be substantially different from the present exercise. In the absence of both statistical field data and a number of experts having a good knowledge of the field, the approach adopted by the authors here is presented as if fuzziness is subjective, context-dependent (in accordance with the original idea of Zadeh, 1973) and stems from an individual expert. However, guidelines for the design of these steps based on different views (i.e. fuzziness is objective and stems from a group of experts) are available in the literature, and are given when necessary. The following subsections give the detailed description of the five steps (mentioned in Section 2.3.) applied for this example.

5.1. Fuzzification

The fuzzification, which can be described by the process of establishment of membership functions, requires several steps, consisting of the establishment of ranges in the numerical domains of the variables concerned, the specification of boundaries in the fuzzy domains of associated fuzzy subsets and the selection of the shape of the membership functions (MFs). For the concepts of the MFs (Zadeh, 1973) and (Mathworks, 2005) are referred to. Here, an example is given to describe the steps to establish the MFs for one input variable (food crop price) and one output variable (the rate of change in forest area).

A major problem in establishing the possible numerical range of values for each of the input variables in the respective scenarios is that both prices and costs were subject to a high level of monetary inflation in the late 1990s. Consequently, these values are showing extreme fluctuations over time, which cannot simply be projected in the near future. For this reason we have presented these monetary values in terms of constant prices in 1993 (instead of current prices).

For the ranges of output variables, both statistical and spatial data obtained from survey and satellite images were used. For example, the yearly change in forested area would be negative (e.g. due to logging) or positive (e.g. due to reforestation). Data from the Division of Forestry and Land Conservation of Gowa district (2000) show an estimation of around 10–15% of the Jeneberang watershed area that was converted to other uses in the last 10 years. On the other hand, 2650 ha of forest was rehabilitated through replantation programmes. Taking 15% of the catchment area to represent the deforested area (5760 ha) during these 10 years, the net decrease in forest area is 3110 ha. From that, it is reasonable to have the maximum decrease of forest area for each year at 400 ha/year. The maximum increase due to investment in reforestation, afforestation can be set at 400 ha/year, based on the same information.

In addition to the specification of the numerical ranges of variables, it is necessary to specify the boundaries of the associated fuzzy subsets. For example, from what value to what value can the food crop prices be considered to be "low", "medium" or "high". The boundaries of fuzzy subsets are allowed to have their intersection, i.e. one particular price can belong to both "low" and "medium" fuzzy subsets. These boundaries are often established subjectively from the experience of experts. This is the case adopted in this exercise. A less subjective example of specifying these boundaries, applying the statistical moving average technique, given the data available, was discussed by Draeseke and Giles (2002). Another requirement is the determination of the shapes of the MFs of the input and output variables. There are, in general, no rules for the selection of a shape of a membership function when little data and expert's knowledge about a variable exist. Therefore, the symmetrically trapezoidal, triangular MFs (Aronica et al., 1998) and Gaussian MFs (De Kok et al., 2001) are often chosen. In the present exercise the MFs of independent variables have the Gaussian form, whereas trapezoidal functions are used for dependent variables. Four methods of building MFs using expert knowledge elicitation, if individual expert and groups of experts are present, are described in Cornelissen et al. (2003).

5.2. Formulation of inference rules

A key step in the construction of the fuzzy system is the formulation of inference rules that reflect the mechanisms underlying the qualitative scenarios. For each scenario a set of all possible combinations of independent variables (or direct stimuli) has been defined, which may serve as a basis for assessing their impact on the five major land-use types. From these general sets a number of realistic combinations of independent variables, which are directly relevant for the dynamics in the respective land-use types are derived. The establishment of the direction and intensity of the impacts of these combinations on land-use through expert assessment is then conducted. For practical reasons, the full procedure for scenario A is presented here (Table 1).

In scenario A food prices are maintained at a stable medium (M) level in order to guarantee a sufficient food supply at reasonable prices. This is achieved through import controls, input subsidies and marketing boards. Cash crop prices are fluctuating between low (L) and medium (M) levels, due to the suppressing impact of marketing imperfections on higher price levels. Production costs are gradually rising from L to M through the abolishment of subsidies for agricultural inputs. The labour costs are kept at a low level through the combined impact of a high rural labour surplus and a rigid control of trade union activities. Rural wages, however, may increase near big cities through the impact on increasing rural–urban circulation opportunities. Public investments have been rising from L level to M level through special attention for rural public services, infrastructure and agricultural intensification programmes. But at the end of this period these investments may again decline to the L level, due to the rising importance of the urban-industrial sector. These parameters of the direct stimuli in scenario A are responsible for the fact that only rules 1, 2, 4, 5, 10, 11, 13 and 14 are relevant for this scenario. With this reduced set of rules we will finally assess their impact on the dynamics of the area expansion of the respective types of land use (Table 2).

Table 1
Set of possible combinations of independent variables for scenario A

Rule	Food crop prices	Cash crop prices	Production costs	Public investment
1	M	L	L	L
2	M	L	L	M
3	M	L	L	H
4	M	L	M	L
5	M	L	M	M
6	M	L	M	H
7	M	L	H	L
8	M	L	H	M
9	M	L	H	H
10	M	M	L	L
11	M	M	L	M
12	M	M	L	H
13	M	M	M	L
14	M	M	M	M
15	M	M	M	H
16	M	M	H	L
17	M	M	H	M
18	M	M	H	H
19	M	H	L	L
20	M	H	L	M
21	M	H	L	H
22	M	H	M	L
23	M	H	M	M
24	M	H	M	H
25	M	H	H	L
26	M	H	H	M
27	M	H	H	H

In a similar way we have established the relevant inference rules for the scenarios B and C, as well as their impacts on the areas of the respective land use types (Tables 3 and 4).

5.3. Application of the inference rules

In the next two steps the calculations of the values of the output variable, which are concerned with fuzzy logic operation, are conducted with the Fuzzy Logic Toolbox embedded in MATLAB® (Mathworks, 2005). The method adopted here is referred to as Mamdani inference (Mamdani and Assilian, 1975) and is illustrated as an example. Considering inference rule 27 (Table 4):

Table 2
Reduced set of inference rules in scenario A

Rule	Forest	Shrub and grassland	Dry upland farming	Mixed forest gardening	Paddy fields
1	−	±	±	0	0
2	0	0	0	±	0
4	−	±	+	0	±
5	0	0	±	±	0
10	−	±	0	±	0
11	0	0	−	+	0
13	−	±	±	0	±
14	0	0	0	±	0

Notation: +, strong increase; ±, weak increase; 0, stagnant; —, strong decrease; −, weak decrease.

Table 3
Reduced set of inference rules in scenario B

Rule	Food crop prices	Cash crop prices	Production costs	Public invest.	Forest	Shrub and grassland	Dry upland farming	Mixed forest gardening	Paddy field
4	L	L	M	L	–	0	±	0	0
7	L	L	H	L	—	0	±	0	±
13	L	M	M	L	–	0	0	±	0
16	L	M	H	L	—	0	±	±	±
22	L	H	M	L	–	–	0	+	0
25	L	H	H	L	—	–	±	+	±

IF (food crop price is medium) AND (cash crop price is high) AND (production cost is high) AND (public investment is high) THEN (the rate of change in forest area is strongly increased).

First the fuzzy value for the *rule antecedent*, which is the condition preceding the THEN statement, must be determined by calculating the corresponding membership function. The AND operation is implemented by taking the minimum value of the membership values for the four independent values:

$$\mu_{\mathrm{AND}} = \min[\mu_1(x_1), \mu_2(x_2), \mu_3(x_3), \mu_4(x_4)] \tag{8}$$

where μ_{AND} is the membership value for the rule antecedent and $\mu_1(x_1)$ is the membership value for the food crop price corresponding to numerical value x_1.

5.4. Calculation of the output values

In the next step, the fuzzy value of the THEN part of the rule or the *rule consequent* must be determined. This is done by truncating the MF for the fuzzy output value (the rate of change in forest area) at the value μ_{AND}. The result is a new MF $\mu_{\mathrm{CONS}}(y)$ for the rule consequent, where y is the value for the rate of change in forest area in the numerical domain. This procedure is repeated for each inference rule, after which the results are aggregated to a single MF by taking the maximum value of the membership values for the entire set of inference rules:

$$\mu_{\mathrm{OUT}}(y) = \max[\mu^i_{\mathrm{CONS}}(y)], \quad i = 1, \ldots, n \tag{9}$$

where n is the number of inference rules (for example, $n = 8$ in Table 4).

The result is a single MF for the output variable which must now be translated from the fuzzy to the numerical domain (defuzzification) to allow for the comparison with the quantitative values produced by RaMCo. This defuzzification can take place in different ways. Here the output corresponding to the centroid of the output MF is used.

5.5. Testing the consistency of the scenarios

To increase the credibility of the outputs produced by the experts' system, both actor-testing and quantification-testing were carried out. First, the land-use types, drivers, driving mechanism and inference rules of the land-use system were presented at a symposium in Makassar city. This symposium was attended by local officials, stakeholders and scientists, ranging from forestry experts, agronomists, economists and sociologists to mathematicians, marine biologists and other natural scientists. These participants indulged in a lively debate on the merits and limitations of the respective scenarios and their assumptions, but in general recognised their local relevance and supported their main lines of reasoning. Second, the physical constraint of the total area of the basin is used to check the consistency of the inference rules and the numerical ranges of the outputs. The differences between the basin area and the total of computed land demands do not exceed 10% of the basin area in any of the three scenarios. To use the quantitative changes in the micro-scale model (mentioned below), Eq. (4) in Section 3 is used to scale up and down so that the total computed land demands are always equal to the basin area.

6. Results

Quantitative changes in all land-use types projected by the experts' system need to be spatially allocated in order to

Table 4
Reduced set of inference rules in scenario C

Rule	Food crop prices	Cash crop prices	Production costs	Public invest.	Forest	Shrub and Grassland	Dry Upland Farming	Mixed Forest Gardening	Paddy field
14	M	M	M	M	0	0	0	±	0
15	M	M	M	H	±	/	/	+	0
17	M	M	H	M	±	0	0	0	±
18	M	M	H	H	+	/	/	±	0
23	M	H	M	M	±	/	/	+	0
24	M	H	M	H	+	/	/	+	0
26	M	H	H	M	+	/	/	±	±
27	M	H	H	H	+	/	/	+	0

compute the total soil loss and the sediment yield at the inlet of the Bili-Bili dam. However, experts in socio-economic sciences have difficulty in speculating with regard to the locations where the changes should take place. This is due to the fact that the spatial distribution of land-use changes depends on the biophysical aspects of the basin, such as geomorphology and transportation networks. The Research Institute for Knowledge Systems (RIKS) has recently developed a generic tool, GEONAMICA (Engelen et al., 2004), which aims to represent spatially the quantitative changes of land-use systems in land-use maps. GEONAMICA, which adopts the constrained cellular automata approach, makes flexible use of the minimum available information such as: suitability maps, zoning maps, accessibility maps and cellular automata transition rules. The use of this tool allows the same spatial distribution mechanism as that adopted by RaMCo and the hypothesised system.

Maps of the land-cover changes produced by RaMCo and the experts' system under three scenarios were used to compute soil losses and sediment yields using the approach mentioned in Section 3. The final results of the dynamic development of sediment yields—the information needed by storage lake managers—are presented in Fig. 2.

It can be seen from Fig. 2 that RaMCo can produce the trend lines of increasing sediment yields in scenarios A and B. However, for scenario C, RaMCo gives results which are contradictory to those produced by the experts' system, in terms of trend lines. It also means that RaMCo is incapable of differentiating between the consequences of scenario B and C, which, according to the experts' system, are opposite in direction.

To determine whether the aggregation level of land-use types adopted by the RaMCo model is the cause of the problem, a further analysis is conducted. The five active land-use types classified by the expert are quantitatively aggregated into three land-use types classified by RaMCo. In comparisons between the two predictions in scenario C, the land demand of land-use type ''nature'' is underpredicted, while the land demand of ''mixed agriculture'' is overpredicted by the RaMCo model. An attempt was made to reduce the growth coefficient (α) in Eq. (1), which was originally assumed to be constant. The reason to adjust it is that the growth coefficient originally reflected the stakeholders' reaction to the change in the net benefit obtained per unit area. This should take into account the control exerted by the government through environmental law, giving credits to farmers to convert from upland farming to mixed forest garden, and launching intensification of agriculture programmes. It turned out that when α in scenario A is reduced slightly and α in scenario C is reduced strongly in comparison with the value used in scenario B, the land-use projections made by the RaMCo model are mostly the same as the land-use projections made by the hypothesised system (after five land-use types are aggregated into three land-use types). The new land demands produced by the re-calibrated RaMCo model are put to the micro-scale model and then to the USLE to compute the sediment yields. The new results are presented in Fig. 3.

It can be seen from Fig. 3 that even with the same quantitative changes in the three land-use types, produced by the two systems, the RaMCo model is still unable to produce the trend line which was produced by the experts' system, in scenario C. The trend differences are a direct result of the conceptual differences with the expert system on issues regarding land-use dynamics. The question arises to what extent the difference in the erosion trend is significant, given the intrinsic uncertainties in both RamCo and the expert system. The difference between the erosion trends of RamCo and the expert system

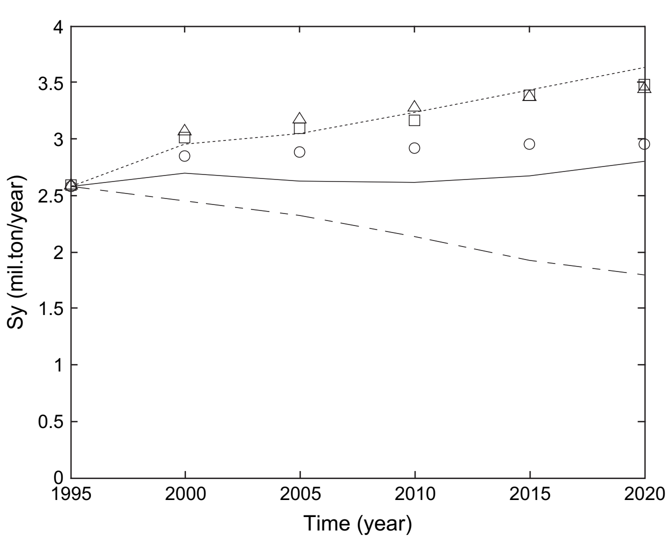

Fig. 2. Comparisons between the sediment yields computed by RaMCo (with fixed α) under three scenarios: A, guided market economy (○); B, maximum growth (□); C, sustainable development (△) and the sediment yields computed by the hypothesized system for scenarios A (solid line), B (dotted line) and C (dash-dotted line).

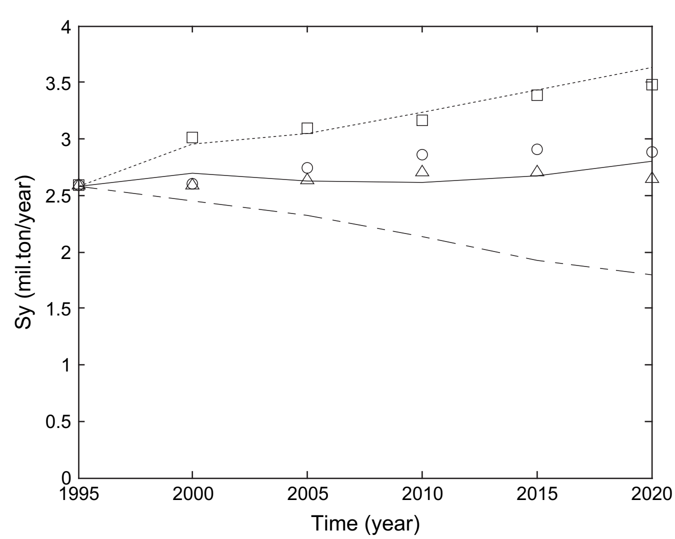

Fig. 3. Comparisons between the sediment yields computed by RaMCo (with adjusted α) under three scenarios: A, guided economy (○); B, maximum growth (□); C, sustainable development (△) and the sediment yields computed by the hypothesized system for scenarios A (solid line), B (dotted line) and C (dash-dotted line).

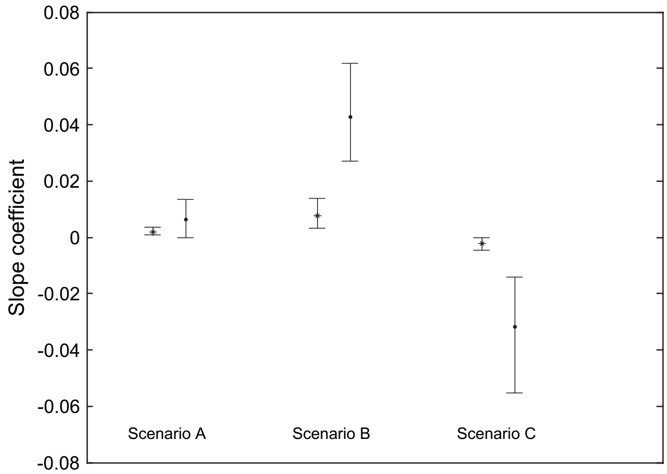

Fig. 4. Comparison of the projected trends in the sediment yield as with RaMCo (*) and the hypothesised system (•) under the three scenarios A, B, C. The error margins reflect the 5th and 95th percentiles in the trends.

originate from differences in the description of land-use dynamics and the erosive properties of the land-use types used. The differences in the simulated land use (the main intermediate variable) can be attributed to both structure- and parameter-related uncertainties. Additional differences emerging in the erosion model relate to model parameter uncertainties only, this part of the model being identical for both approaches. Fig. 4 illustrates the contribution to the uncertainty in the erosion trend from this parameter uncertainty in the erosion model. The figure illustrates that the trend differences cannot be attributed to uncertainties in this part of the model. The uncertainty range has been determined by means of Monte Carlo analysis with variation in all the parameters of Eq. (5) and the SDR of Eq. (7).

In Fig. 4 we observe significant differences between the slope coefficients of the erosion trends projected by the expert system for the three scenarios. For the RaMCo model the differences between the scenarios are much smaller, and difficult to detect given the uncertainty. In addition the differences between RaMco and the expert system are significant for both scenario B and for scenario C.

It can be concluded that the too coarse aggregation level of land-use makes the RaMCo model fall short of describing the consequences of the change in socio-economic factors and policy options produced by the experts' system on the long-term trend in the sediment yield.

7. Discussion

The application of the new approach to validation of the RaMCo model suggests several interesting points about the validity of the RaMCo model in particular and about IWS models in general. First, the RaMCo model is able to describe the dynamic developments of the three aggregated land-use types under three scenarios if the growth coefficient α in Eq. (1) is adjusted according to each scenario. The argument for this adjustment is that it should be dependent on additional policy variables, such as: environmental law, agricultural intensification programmes and cheap credits which have not been explicitly included in Eq. (1). Second, without a refinement of land-use types the RaMCo model fails to produce satisfactorily the consequences of policy options on sediment yields. The lesson learned is that a model can be valid for one purpose but invalid for another. Therefore, the validity of any IWS model should be assessed in accordance with a clearly specified management variable.

With regard to the adjustment of the growth coefficient α in the RaMCo model, the choice to let this coefficient depend on each scenario can be explained by the insights gained from the qualitative scenarios and the sensitivity analysis of RaMCo. All the values of the input variables related to mixed agriculture (the most sensitive to changes in the land-use system in the watershed), such as prices, production costs and yields were kept the same to input into the model system and experts' system. Public investment can be taken into account by RaMCo through the yearly reforestation area, which is relatively insensitive to the changes in the area of all the land-use types, in comparison with the growth coefficient for the mixed agriculture. In the current exercise, the growth coefficient was adjusted according to each scenario. This coefficient is shown to be reversely proportional to level of public investment described in the three scenarios. Under each scenario the value of this coefficient was kept constant over time in the re-calibrated RaMCo model, reflecting the stakeholders' reaction under only one socio-economic regime. In reality, this coefficient may vary temporally.

The advantage of the proposed approach is that it opens a new direction for the validation of IWS models using qualitative hypotheses formulated by system experts on future trends for which conventional techniques fail. It makes the assumptions and reasoning processes that lead to experts' judgments more transparent to modellers. Thus, not only can the final quality of the model be assessed but also possible structural errors can be unravelled. It helps to reduce the possible bias of the experts through the process of documentation and communication between modellers, system experts and stakeholders.

However, some limitations of the new approach should also be mentioned. The first practical difficulty is that it is difficult to find system experts who are knowledgeable about both field and scientific research. These experts, who should be acknowledgeable beyond the limits of their own discipline, are responsible for formulating the inference rules. In this application, to counter the expert's bias, a workshop with the participation of local scientist experts, resources managers and stakeholders, was held at the step of testing the consistency of the scenarios. In the situation where multiple system experts (scientist experts and resources managers) and stakeholders are available at the earlier stage (e.g. structuring scenarios), several techniques could be implemented, which may facilitate the process of identifying key drivers and the driving mechanisms (reflected by the inference rules) underlying the system studied. Elicitation techniques such as analytical hierarchy process (Zio, 1996), adaptive conjoint analysis (Van der Fels-Klerx et al., 2000) and simple or weighted average technique (Nguyen and de Kok, 2007.) can be applied to elicit expert's and stakeholders' opinions on key system drivers. A data mining technique applied to establish the inference rules from a multiple experts' opinions was presented by Kawano et al. (2005). The second difficulty is that the estimation of the quantitative ranges of the inputs and outputs in the hypothesised system is difficult when quantitative data are lacking and surrounded with uncertainty.

8. Conclusions

In the field of integrated systems modelling, researchers often encounter the dilemma that the model should not be so complex that it is unmanageable in terms of data collection, uncertainty propagation, yet not so simple that it cannot give useful information to the decision-makers.

In this paper, a novel approach to testing integrated systems models using qualitative scenarios has been presented. The approach can be used to determine whether a model is ill or well designed, with regard to the purpose of an IWS model as a tool capable of reflecting the system experts' consensus about the dynamic behaviours of system output variables, under a set of possible socio-economic scenarios and policy options. The design of this approach was motivated by the three reasons that limit the relevance of the conventional approaches to the validation of IWS models: the limited predictive ability of historical data to describe the future behaviour of interactive natural-human systems, the qualitative nature of the social sciences and the scarcity of field data for validation.

From a philosophical perspective, the current approach acknowledges that the process of communicating, persuading and convincing groups of modellers, experts and end-users plays a vital role in the process of validating IWS models (Pahl-Wostl, 2002; Poch et al., 2004). The complexity of the environmental problems makes necessary the development and application of new tools capable of processing not only the numerical aspects, but also the experience of experts and wide public participation, which are all needed in the decision-making process. In parallel to this development, the use of the historical data and comparing them with model outputs (empirical test) is of vital importance. This comparison should be included when possible. However, new methods, focusing, for example, on the trend comparison might be promising and so need to be further developed for the validation of IWS models. This new approach to testing IWS models may be useful in both situations where measured data are unavailable and where data are available for the empirical test.

Model credibility can be enhanced by proper modeller-manager dialogues, rigorous validation tests against independent data, uncertainty assessment and peer reviews of the model at various stages throughout its development (Refsgaard et al., 2005). In our opinion, the new approach to testing an integrated water systems model using qualitative scenarios may be used for the different steps of the overall model cycle, such as designing the conceptual model, validating the site-specific model, analysing the model domain of applicability, and assessing the uncertainty of the future conditions.

Acknowledgements

The authors are grateful to RIKS for supplying the GEONAMICA software. They are indebted to various researchers from UNHAS University for sharing opinions on land-use scenario formulation. The research was partially supported by The Netherlands Foundation for The Advancement of Tropical Research (WOTRO). Dr Maarten S. Krol of the Department of Water Engineering and Management of the University of Twente is thanked for his useful suggestions concerning the role of uncertainty. The constructive and critical comments of the four reviewers have improved the structure and content of this paper. These reviewers are gratefully acknowledged.

References

Alcamo, J., 2001. Scenarios as tools for international environmental assessments. Environmental Issue Report, No. 24. Experts' Corner Report. Prospects and Scenarios No. 5. European Environment Agency, Copenhagen, Denmark.

Aronica, G., Hankin, B., Beven, K., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Advances in Water Resources 22 (4), 349–365.

Balci, O., 1995. Principles and techniques of simulation validation, verification, and testing. In: Alexopoulos, C., Kang, K., Lilegdon, W.R., and Goldsman, D. (Eds.), Proceedings of the 1995 Winter Simulation Conference, pp. 147–154.

Barlas, Y., 1994. Model validation in system dynamics. Proceedings of the 1994 International System Dynamics Conference. Methodological Issues. Stirling, Scotland, pp. 1–10.

Barlas, Y., Kanar, K., 1999. A dynamic pattern-oriented test for model validation. Proceedings of 4th Systems Science European Congress, Valencia, Spain, Sept. 1999, pp. 269–286.

Beck, B., 2002. Model evaluation and performance. In: El-Shaarawi, A.H., Piegorsch, W.W. (Eds.), Encyclopedia of Environmetrics, vol. 3. John Wiley, Chichester, pp. 1275–1279.

Beck, M.B., Chen, J., 2000. Assuring the quality of model designed for predictive purposes. In: Saltelli, A., Chan, K., Scott, E.M. (Eds.), Sensitivity Analysis. Wiley, Chichester, pp. 401–420.

Cornelissen, A.M.G., Berg, J.V.D., Koops, W.J., Kaymak, U., 2003. Elicitation of expert knowledge for fuzzy evaluation of agricultural production systems. Agriculture, Ecosystems and Environment 95 (1), 1–18.

CTI, 1994. The detailed design of environmental improvement works and raw water transmission main in the Bili-Bili multipurpose dam project (phase 2). Design Report. Part 1. Report prepared for the Government of the Republic of Indonesia. Ministry of Public works, CTI Engineering Co. Ltd., Japan.

De Kok, J.L., Titus, M.J., Wind, H.G., 2000. Application of fuzzy sets and cognitive maps to incorporate social science scenarios in integrated assessment models: a case study of urbanization in Ujung Pandang, Indonesia. Integrated Assessment 1 (3), 177–188.

De Kok, J.L., Engelen, G., White, R., Wind, H.G., 2001. Modeling land-use change in a decision-support system for coastal zone management. Environmental Modeling and Assessment 6, 123–132.

De Kok, J.L., Wind, H.G., 2002. Rapid assessment of water systems based on internal consistency. Journal of Water Resources Planning and Management 128 (4), 240–247.

Draeseke, R., Giles, D.E.A., 2002. A fuzzy logic approach to modeling the New Zealand underground economy. Mathematics and Computers in Simulation 59, 115−123.

Engelen, G., White, R., Uljee, I., 2004. The MOLAND model for Urban and Regional Growth. Draft Final Report of contract no. 21512-2003-12 F1SP ISP nl. submitted to European Commission Joint Research Center, Institute for Environment and Sustainability, Ispra, Italy. Research Institute for Knowledge System BV, PO Box 463, 6200 AL Maastricht, The Netherlands (http://www.riks.nl).

Ewen, J., Parkin, G., 1996. Validation of catchment models for predicting land-use and climate change impacts. 1. Method. Journal of Hydrology 175, 583−594.

Finlay, P.N., Wilson, J.M., 1997. Validity of decision support systems: towards a validation methodology. System Research and Behavioral Science 14 (3), 169−182.

Flavelle, P., 1992. A quantitative measure of model validation and its potential use for regulatory purposes. Advances in Water Resources 15, 5−13.

Forrester, J.W., Senge, P.M., 1980. Tests for building confidence in system dynamics models. In: Legasto Jr. A.A., Forrester, J.W., Lyneis, J.M. (Eds.), System Dynamics. TIMS Studies in Management Sciences, vol. 14, pp. 209−228. North Holland.

Fraedrich, D., Goldberg, A., 2000. A methodological framework for the validation of predictive simulations. European Journal of Operational Research 124, 55−62.

Godet, M., Roubelat, F., 1996. Creating the future: the use and misuse of scenarios. Long Range Planning 29 (2), 164−171.

Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software 21, 602−614.

Kawano, S., Huynh, V.N., Ryoke, M., Nakamori, Y., 2005. A context-dependent knowledge model for evaluation of regional environment. Environmental Modelling & Software 20, 343−352.

Kirchner, J.W., Hooper, R.P., Kendall, C., Neal, C., Leavesley, G., 1996. Testing and validating environmental models. The Science of the Total Environment 183, 33−47.

Kleijnen, J.P.C., 1995. Verification and validation of simulation models. European J. Operational Research 82, 145−162.

Kleindorfer, J.B., O'Neill, L., Ganeshan, R., 1998. Validation in simulation: various positions in the philosophy of science. Management Science 44 (8), 1087−1099.

Konikow, L.P., Bredehoeft, J.D., 1992. Growth-water models cannot be validated. Advance in Water Resources 15, 75−83.

Loehle, C., 1997. A hypothesis testing framework for evaluating ecosystem model performance. Ecological Modeling 97, 153−165.

Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Study 7, 1−13.

Mathworks, 2005. Fuzzy Logic Toolbox User's Guide (For Use with MATLAB®). http://www.mathworks.com/access/helpdesk/help/pdf_doc/fuzzy/fuzzy.pdf.

Mitchell, P.L., 1997. Misuse of regression for empirical validation of models. Agricultural Systems 54 (3), 313−326.

Nguyen, T.G., De Kok, J.L., 2003. Application of sensitivity and uncertainty analysis for validation of an integrated systems model for coastal zone management. In: Post, D.A. (Ed.), Proceedings of the International Congress on Modelling and Simulation, MODSIM 2003. Modelling & Simulation Society of Australia & New Zealand Inc., Townville, Australia, pp. 542−547.

Nguyen, T.G., de Kok, J.L., 2007. Systematic testing of an integrated systems model for coastal zone management using sensitivity and uncertainty analyses. Environmental Modelling & Software 22 (11), 1572−1587.

Oreskes, N., 1998. Evaluation (not validation) of quantitative models. Environmental Health Perspectives 106 (6), 1453−1460.

Oreskes, N., Frechette, K.S., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263, 641−646.

Pahl-Wostl, C., 2002. Towards sustainability in the water sector - the importance of human actors and processes of social learning. Aquatic Sciences 64, 394−411.

Parker, P., Letcher, R., Jakeman, A., Beck, M.B., Harris, G., Argent, R.M., Hare, M., Pahl-Wostl, C., Voinov, A., Janssen, M., et al., 2002. Progress in integrated assessment and modeling. Environmental modelling & Software 17, 209−217.

Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M., Cortés, U., 2004. Designing and building real environmental decision support systems. Environmental Modelling & Software 19, 857−873.

Refsgaard, J.C., Henriksen, H.J., 2004. Modelling guidelines-terminology and guiding principles. Advances in Water Resources 27, 71−82.

Refsgaard, J.C., Henriksen, H.J., Harrar, W.G., Scholten, H., Kassahun, A., 2005. Quality assurance in model based water management - review of existing practice and outline of new approaches. Environmental Modelling & Software 20, 1201−1215.

Rykiel, E.J., 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90, 229−244.

Sargent, R.G., 1984. A tutorial on verification and validation of simulation models. Proceedings of the 1984 Winter Simulation Conference, pp. 115−121.

Sargent, R.G., 1991. Simulation model verification and validation. Proceedings of the 1991 Winter Simulation Conference, pp. 37−47.

Scholten, H., ten Cate, A.J.U., 1999. Quality assessment of the simulation modelling process. Computers and Electronics in Agriculture 22, 199−208.

Schwab, P., Cerutti, F., von Reibnitz, C., 2003. Foresight-using scenarios to shape the future of agriculture research. Foresight 5 (1), 55−61.

SCS, 1971. Sediment sources, yields, and delivery ratios. In: National Engineering Handbook, Section 3: Sedimentation. US Department of Agriculture, Soil Conservation Service.

Shannon, E.R., 1981. Tests for verification and validation of computer simulation models. Proceedings of the 1998 Winter Simulation Conference, pp. 573−577.

Suriamihardja, D.A., et al., 2001. Study of Integrated Management on Jeneberang Watershed, Phase 2. Final report made by the cooperation program between Canadian International Development Agency and Center for Environmental Study. Hasanuddin University, Makassar, Indonesia.

Uljee, I., Engelen, G., White, R., 1996. RAMCO Demo Guide. Modelling and Simulation Research Group, Research Institute for Knowledge Systems BV, PO Box 463, 6200 AL Maastricht, The Netherlands.

Van der Fels-Klerx, H.J., Horst, H.S., Dijkhuizen, A.A., 2000. Risk factors for bovine respiratory disease in dairy youngstock in The Netherlands: the perception of experts. Livestock Production Science 66, 35−46.

Van der Heijden, K., 1996. Scenarios: The Art of Strategic Conversation. John Wiley & Sons, Chichester.

Van Tongeren, O.F.R., 1995. Data analysis or simulation model: a critical evaluation of some methods. Ecological Modeling 78, 51−60.

Von Reibnitz, U., 1988. Scenario Techniques. McGraw-Hill, Hamburg.

White, R., Engelen, G., 1997. Cellular automata as the basis of integrated dynamic regional modelling. Environment and Planning B: Planning and Design 4, 235−246.

Wischmeier, W.H., Smith, D.D., 1965. Predicting rainfall-erosion losses from cropland east of the Rocky Mountains: Guide for Selection of Practices for Soil and Water Conservation. Agriculture Handbook No. 282. US Printing Office, Washington.

Wischmeier, W.H., Smith, D.D., 1978. Predicting rainfall-erosion losses. Agriculture Handbook No. 537. US Department of Agriculture, Washington, DC.

Zadeh L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man, Cybernetics, SMC-3, 28−34.

Zio, E., 1996. On the use of the analytical hierarchy process in the aggregation of expert judgments. Reliability Engineering and System Safety 53, 127−138.