Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software 22 (2007) 1667–1678
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

An appropriateness framework for the Dutch Meuse decision support system

Yue-Ping Xu [a,*], Martijn J. Booij [b], Arthur E. Mynett [c,d]

[a] Department of Civil Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
[b] Water Engineering and Management, Faculty of Engineering, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
[c] WL|Delft Hydraulics, P.O. Box 177, 2600 MH Delft, The Netherlands
[d] UNESCO-IHE, P.O. Box 3015, 2601 DA Delft, The Netherlands

Received 24 February 2006; received in revised form 13 December 2006; accepted 5 January 2007
Available online 23 March 2007

Abstract

Models are essential in a decision support system for river basin management. In a decision support system for integrated planning and management, the use of appropriate models is important to avoid models being either too simple or too complex. In this paper, appropriate models refer to models that are good-enough-but-not-more-than-that to obtain an acceptable ranking of river engineering measures under uncertainty. A systematic approach called 'appropriateness framework' is proposed to determine appropriate models that can be used in a decision support system. The approach is applied to a decision support system for the Dutch Meuse River. One important component of this decision support system, flood safety, is used in this paper to demonstrate how this approach works. The results show that the approach is very useful in helping to determine appropriate models. Potential applications of the approach in other decision support systems are discussed. The approach presented in this paper is designed as a tool to stimulate the communication between decision makers and modelers and to promote the use of models in decision-making for river basin management.

Keywords: Appropriateness framework; Uncertainty analysis; Sensitivity analysis; Decision support systems; Flood safety; Dutch Meuse River; Uncertainty reduction

1. Introduction

Decision-making in river basin management is becoming more complex because of the integration of surface water and groundwater, issues of water quality, competitions among stakeholders and increasing uncertainty due to effects of environmental change. During the last couple of decades, there has been growing interest in the use of computer-based decision support systems (DSSs) to address the increasingly complex river basin decision-making problems. Keen and Scott Morton (1978) state that a DSS couples the intellectual resources of individuals with the capabilities of the computer to improve the quality of decisions and it is a computer-based system for decision makers who deal with semi-structured problems. The incorporation of DSS techniques in river basin management started in the mid-1980s (McMahon et al., 1984; Barnwell et al., 1986). In recent years, several DSSs have been developed, such as AQUATOOL (Andreu et al., 1996), NELUP (Dunn et al., 1996), RiverWare (Zagona et al., 2001), DSS Large Rivers (Schielen and Gijsbers, 2003) and GAMES (Dorner et al., 2007) to support river basin management.

For an integrated model-based DSS, the capabilities of the models are fundamental in facilitating the major, if not the whole, decision-making process. Models with different complexities are often used in a DSS and there is a gradation of

* Corresponding author. Present address: Institute of Water Resources, School of Architecture and Civil Engineering, Zhejiang University, Hangzhou, China. Tel.:+86 57188273652.
E-mail address: yueping1126@gmail.com (Y.P. Xu).

1364-8152/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2007.01.002

model complexity. Related to this, recently, there have been considerable discussions about the use of simple and complex models (Jakeman and Hornberger, 1993; Nihoul, 1994; Chwif and Paul, 2000). Simple and complex models both have their advantages and disadvantages. In general, modelers often tend to develop more complex models to better understand the system under study while decision makers focus more on practical applications and tend to use simpler models. According to Perrin et al. (2001) and Vreugdenhil (2002), in many cases, especially in planning and strategic management, simple models can be more appropriate than complex ones. For model-based DSSs, the models used are supposed to fit decision makers' use but without leaving out essential mechanisms associated with management problems. Therefore, to stimulate the application of models in decision-making, the use of appropriate models is essential and has been discussed by Rogers (1978), Fread (1985), Vreugdenhil (2002, 2006) and Booij (2003). In these studies, however, the appropriateness of models is only related to the accuracy or uncertainty of model outputs.

The importance of uncertainty in decision-making has been widely recognized (Ministry of Public Housing, Physical Planning, and Environmental Protection, 1985; Reckhow, 1994; European Commission, 2000; Caminiti, 2004; Schlüter and Rüger, 2007). It has been argued that the current state of knowledge and natural variability in the system behavior often precludes sufficiently accurate model outputs to distinguish or rank among potential river engineering measures (Reichert and Borsuk, 2005). To tackle this decision problem in water management, Reda and Beck (1997) and Duchnese et al. (2001) ranked storm water control measures mainly based on the mean values, or the mean values plus standard deviations, or the values of cumulative probability functions of decision variables (model outputs). De Kort and Booij (2007) proposed a ranking procedure based on the significance of the difference between model output distributions to distinguish different measures. In life cycle assessment, Basson and Petrie (2007) investigated whether the uncertainty in decision variables is likely to make it impossible to distinguish measures by using Principal Component Analysis. They emphasized that it would be of little value to commence detailed preference modeling for the evaluation of measures unless some measures are distinguishable from others despite the uncertainty in decision variables. Preference modeling is the procedure of giving different preferences or weights to different decision variables in multi-criteria decision analysis.

Decision support systems are often used to help develop and evaluate management measures (Sojda, 2007), particularly find their way in evaluating river engineering measures (De Kort and Booij, 2007; Giupponi, 2007). A DSS for integrated river basin management encompasses a number of submodels, such as models for flood safety, ecology, tourism, recreation and navigation. For such complicated DSSs, dealing with the right problem, the adequate level of complexity of models and the adequate uncertainty in model outputs is particularly important, while addressing the right problem is most fundamental (Jakeman et al., 2006; Vreugdenhil, 2006). The missing link in past research is that the determination of appropriate models in a DSS is rarely connected with the practical ranking problems and the uncertainty in model outputs at the same time. Appropriate models used in a DSS in this paper are therefore not only related to the uncertainty in model outputs, but also aim to solve the ranking problem through the incorporation of decision makers' attitudes to uncertainty. The objective of this paper is to develop a systematic approach called 'appropriateness framework' and apply it to a DSS to investigate if the approach finally works out. This appropriateness framework can be used as a tool to help to determine appropriate models in a DSS for integrated river basin management and is proposed under the assumption that decisions are made on a rational basis.

In this paper, a DSS for the Dutch Meuse River is developed and used as a case study (Xu, 2005). The proposed appropriateness framework is applied to one of the most important components of the DSS, flood safety, which itself can be regarded as a DSS. The structure of this paper is as follows. Section 2 first proposes the appropriateness framework that will be used to help to determine appropriate models and its main components are introduced. Section 3 gives a brief description of the case study followed by the results of the application of the framework in Section 4. Section 5 discusses several aspects related to the appropriateness framework. Finally, conclusions based on the Dutch Meuse DSS case study are summarized in Section 6, along with recommendations for future applications of the appropriateness framework.

2. Appropriateness framework

2.1. Criterion for appropriate models

The determination of the appropriateness of models used in a DSS generally not only depends on the practical ranking problem and the uncertainty in model outputs. Basically, good practice is the platform for pursuit of model quality or appropriateness (Jakeman et al., 2006). Furthermore, other aspects such as user-friendliness, flexibility of models and computation time will play an important role as well (Xu, 2005). However, in this paper, the authors argue that the ranking problem and the uncertainty are the most critical aspects for determining appropriate models. Therefore, the concept of appropriateness focuses on these two aspects. The core of this concept is to use good-enough-but-not-more-than-that models (of the processes involved) to solve a decision-making problem for obtaining an acceptable ranking of different river engineering measures accounting for uncertainty in model outputs. Here 'good enough' indicates that a model used in a DSS should have a sound statistical or physical meaning and that the most important and relevant processes should be included. The concept of appropriateness is consistent with the Occam's Razor Theorem, which states that 'entities should not be multiplied beyond necessity' (Brooks and Tobias, 1996). Fig. 1 shows a schematic diagram of the decision problem where two measures M_1 and M_2 are used for illustration. The error

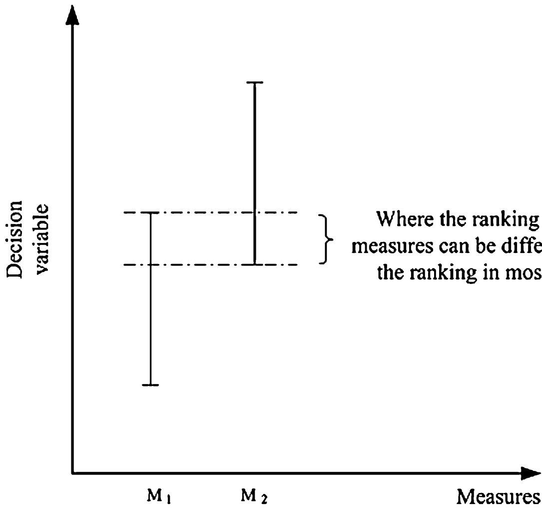

Fig. 1. Schematic diagram of the decision problem.

bars in this figure indicate the uncertainty ranges of model outputs (simulation results) from the two measures. As shown in Fig. 1, in most cases, M_2 produces higher model outputs than M_1 (the probability is more than 50%). If a higher value indicates a better measure, in most cases, M_2 is better than M_1. However, still situations exist where M_1 could be better because of the uncertainty involved. The dashed lines in Fig. 1 indicate where the ranking of the two measures can be different from the ranking in most cases. This paper therefore focuses on how to obtain an acceptable ranking of measures under uncertainty, viz. when can the ranking of M_2 being better than M_1 be accepted? If this ranking can be accepted by decision makers, the models used in the DSS are regarded as 'appropriate'.

To analyze the appropriateness of models quantitatively, the risk of obtaining an unacceptable ranking is adopted. This risk is defined with reference to the classical concept of risks, which usually consider the occurrence probability of a hazard and the consequences of that hazard (Kaplan and Garrick, 1981). Here, in the context of measure comparisons, the risk of obtaining an unacceptable ranking is defined as the product of the mean difference between the model outputs of two measures and the probability of obtaining an unacceptable ranking, that is, the ranking of two measures being different from the ranking in most cases. The consideration of the mean difference in this definition is not only because the value of the difference of the two measures on average is important to decision makers (which indicates which measure is better on average) and regarded as the loss when an unacceptable ranking is adopted, but also because the difference is used here as a scaling factor. The above-defined risk utilizes the mean difference to scale the effect of probability and is independent of measures with different levels of effects on the system. Therefore, the definition of risk in the context of measure comparisons is not exactly the same as the classical one.

Let Y_1 be the model outputs (simulation results) from M_1 and Y_2 be the model outputs from M_2. As shown in Fig. 1, in most cases M_2 is better than M_1. The probability of obtaining an unacceptable ranking is then:

$$P = \Pr(Y_2 < Y_1) = \Pr(Y_2 - Y_1 < 0) \quad (1)$$

A parametric method to calculate the probability can be found in Karl (1999). The mathematical equation of the risk of obtaining an unacceptable ranking then becomes:

$$R = \overline{Y} \times P \quad (2)$$

where R is the risk of obtaining an unacceptable ranking for the two measures; $\overline{Y}$ is the mean difference of model outputs for the two measures; and P is the probability of obtaining an unacceptable ranking for the two measures.

If the risk calculated is acceptable to decision makers, the ranking of measures is regarded as acceptable and models are considered to be appropriate. In the situation with multiple measures, the models are regarded as appropriate if the risks for all combinations of measures are acceptable.

2.2. Appropriateness framework

A systematic approach called appropriateness framework is proposed here to determine appropriate models in a DSS (see Fig. 2). This approach is designed for stimulating the communication between decision makers and modelers (experts) and promoting the use of models in decision-making. For the development, choice or use of such appropriate models, expertise will play an important role. The models used in a DSS should be able to address the main problems and achieve the

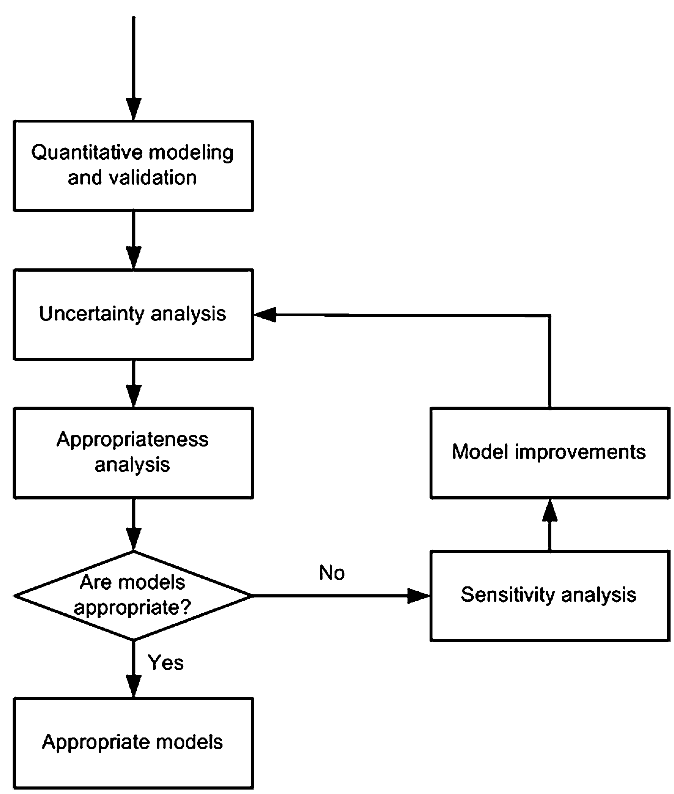

Fig. 2. Appropriateness framework.

objectives. In the meantime, they should include the essence of the system, e.g., should be steady or dynamic, or should be linear or non-linear. Building or selecting a model with good statistical or physical meaning often reduces the later efforts for determining appropriate models. Model building and selection is so called 'quantitative modeling and validation' (see Fig. 2). During this phase, expertise is essential in helping to identify problems and objectives and to carry out appropriate modeling activities. Quantitative modeling is the premise of the appropriateness framework and it ensures that models developed or selected are reasonable.

Therefore, the appropriateness framework generally starts from simple but reasonable models, which include the most important and relevant processes of the system to be modeled. Fig. 2 shows the main components of the framework, including uncertainty analysis, appropriateness analysis, sensitivity analysis and model improvement by uncertainty reduction. In the following, these main components are described consecutively.

Uncertainty analysis is used to analyze sources of uncertainty in the data and models and propagation of these uncertainties into model outputs. In the decision-making process, there are two types of uncertainty. One is decision uncertainty that includes uncertainty on agenda-setting, uncertainty about costs and benefits of measures and uncertainty on goals and preferences of decision makers (Van Asselt, 2000). The other type is outcome uncertainty, which originates from measurement errors of model inputs, uncertainty in scenarios (external factors), uncertainty in model structures, uncertainty in model parameters and natural variability (Morgan and Henrion, 1990; Van Asselt, 2000). To propagate the uncertainty into decision variables, many methods are available in the literature, including error propagation equations, Monte Carlo simulations (e.g., Latin Hypercube simulation) and response surface methods (Bevington and Robinson, 1992; Saltelli et al., 2000). In this paper, Latin Hypercube simulation is used because of its capability to deal with non-linear and non-monotonic systems, and its high efficiency.

However, it is known that the true uncertainty is often not equivalent to the simulated Monte Carlo distribution, although that will be always the case in practice. Even when all other important uncertainties such as model structural uncertainty are incorporated in the uncertainty analysis, there will be still uncertainty in the uncertainty distribution. Therefore, the Monte Carlo simulation can only give an approximate estimation of the uncertainty.

Appropriateness analysis mainly concerns the process of analyzing whether the models are appropriate or not by investigating if the ranking of measures is acceptable. The criterion explained in Section 2.1, viz. the risk of obtaining an unacceptable ranking is used in this study as the indicator for the appropriateness analysis.

To determine the appropriateness of models in the DSS, an acceptable risk level R^* is needed. This value is usually determined by decision makers and is indicative of the lower limit of making a sound decision under uncertainty according to decision makers' experience and opinions. The acceptable risk can be interpreted as the expected loss of different decision variables (such as the expected loss of money). Different from the objective risk R, this acceptable risk is a subjective one. One feasible way to determine this acceptable risk is through decomposing the acceptable risk into the acceptable probability and the acceptable mean difference (see Eq. (2)). The main role of the acceptable risk is to be able to incorporate decision makers' attitudes to uncertainty into the appropriateness framework. The models are determined to be appropriate if the risk R is smaller than R^* for all combinations (pairs) of measures (in the case of multiple measures), i.e.:

$$R < R^* \quad \text{for all combination of measures} \tag{3}$$

Measures here offer opportunities to change the system from the initial undesired situation into a desired one. They are often developed through brainstorming to ensure that adequate measures are included. During measure generation, it is very important that a set of creative and feasible measures is generated, and these measures should be different from one another (Howard, 1988; Matheson and Matheson, 1998; Keisler, 2002). In this sense, the null measure (uninformative model) is regarded to be irrelevant in this approach.

Sensitivity analysis explores the effect of variations of inputs, parameters and models on model outputs to identify the most important model components. If the risk calculated is unacceptable, the models are determined as inappropriate. In order to obtain appropriate models, the models initially developed for a DSS need to be improved by reducing the uncertainty in the data and models. An efficient way to do this is to first identify the most important sources of uncertainty and then put efforts into reducing the uncertainty in these most important sources. Different sensitivity analysis methods can be used for this purpose, such as the Morris method, the 'Method of Sobol', Fourier Amplitude Sensitivity Tests and the Bayesian sensitivity analysis method (Saltelli et al., 2000). In this paper, the Morris method is used because of its efficiency in dealing with a large number of uncertainty sources in the simulation.

Model improvements are carried out, whenever possible, after identifying the most important sources of uncertainty by sensitivity analysis. Model improvements can be achieved by reducing the uncertainty in these most important sources. Two main directions of model improvements are: (i) the reduction of data uncertainty (e.g., model inputs) and (ii) the reduction of model uncertainty (e.g., model parameters, model equations, or spatial and temporal resolution of data). Possible ways for reducing uncertainty include:

- collecting more high quality observation data;
- using better data processing methods;
- collecting expert opinions;
- increasing spatial or temporal resolution of data;
- improving model structures (mainly model equations).

However, one has to realize that not all uncertainties identified in the uncertainty analysis can be reduced due to, for example, the current knowledge limitations, or to the fact that it is not worthwhile to put great efforts in reducing uncertainty due to time and economical constraints. One example is that the uncertainty caused by natural variability, such as the uncertainty in the river discharges (as inputs for flood modeling) can be reduced only by obtaining more observation data, which may not be possible to achieve within the decision-making time frame. If the uncertainty cannot be reduced under current knowledge and resource limitations, measures with ranking problems have to be regarded to have same effects on the system. This implies that decision makers must live with some uncertainty because of current knowledge limitations.

3. Case study

3.1. Flood safety component in Dutch Meuse DSS

A DSS was developed for the Dutch Meuse River located in the Netherlands (Xu and Booij, 2004; Xu, 2005). This DSS seeks answers to strategic questions related to different problems in a long-term context like flood damage, water quality and navigation. Detailed information about this DSS can be found in Xu (2005). The Dutch Meuse River consists of three sections: Border Meuse, Sand Meuse and Over Meuse (Rhijnsburger, 1997). In this paper, one of the most important components in the DSS, flood safety, is selected for the further development of the appropriateness framework as described in Section 2. Fig. 3 shows the system diagram for the flood safety component (models). This diagram mainly includes two subsystems: the physical subsystem and the economic subsystem. The interrelationships between the subsystems are indicated by arrows. The left-hand side of the diagram contains the inputs, including river discharges, channel cross section data, land height and district information. The right-hand side is the decision variable, which is the net present value (NPV) of the flood damage. The flood safety models are applied to the river sections – Border Meuse and Sand Meuse – in the Province of Limburg. Economic growth is the only external factor considered.

3.2. Brief model description

For the purpose of further development of the appropriateness framework, this case study starts from a rather schematic problem with simple models, if compared to many existing and sophisticated models used in river basin management. The main sub-models used in the physical subsystem (see Fig. 3) include a flood frequency model, a hydraulic model, an inundation model, a flood damage model and a risk model while the main model in the economic subsystem is a cost–benefit model.

The primary objective of the flood frequency model is to relate the magnitude of extreme events (flood flows) to their frequency of occurrence through the use of probability distributions. In this analysis, the Gumbel Extreme Value distribution is used (Shaw, 1994). The hydraulic model is used to calculate water levels in the river channel for design floods, based on steady non-uniform flows. The inundation model is used to calculate inundations (the difference between water level and land height) in the floodplains. The damage model is used to calculate the economic damage in the floodplains based on the inundation depth, the land use type and the number of units of that land use type in a cell (De Blois, 2000). The objective of the risk model is to calculate the expected annual damage by using the US Army Corps' method (National Research Council, 2000). The cost–benefit model is used to calculate the NPV of the expected annual damage, which is the decision variable in this case study. In order to quantify the benefits resulting from different measures, often the annual reductions in damage are used as a decision variable for deterministic situations (Shaw, 1994). Due to the consideration of uncertainty, an alternative way is used here based on the sum of the costs of measures and extra benefits, e.g., sand and gravel extraction (see Fig. 3) to the present value of flood damage. This doesn't affect the ranking of the measures, and in this way, the larger the net present value becomes, the less desirable the measures are.

3.3. Data and measures

Most data are obtained from the Dutch Directorate for Public Works and Water Management (RWS) (see Fig. 3). Daily mean discharges are from 1911 to 1997 (RWS, 1997). Channel cross section data are obtained from the RWS (2001) and Folkertsma (personal communication, 2003). Land use and land height data are obtained from De Blois (2000). There are 10 types of land use in the flood safety component, namely households, industry, construction, trade and recreation, services, agriculture, greenhouses, institutions, road and water. The spatial resolution of land use and land height data is 150 m. Three damage classes are used in the damage functions, namely 'low', 'medium' and 'high', which are defined for 33 different municipalities according to their economic importance. Damage factors in damage functions are given by De Blois (2000). For a specified time horizon (50 years), a discount rate of 5% per year is chosen and the economic growth is assumed to be 2% per year (Ministry of Transport, Public Works and Water Management, 1994a). These inputs are indicated on the left-hand side of the system diagram in Fig. 3.

As shown in Fig. 3, three categories of measures are identified: measures aiming at increasing the discharge capacity of river channels, embankments and spatial planning measures. In this paper the current situation (do-nothing) is regarded to be one measure as well. Five measures are included: (i) Measure 0: current situation (M_0); (ii) Measure 1: broadening the main channel of the Border Meuse by 25 m (M_1); (iii) Measure 2: deepening the main channel of the Sand Meuse by 1 m (M_2); (iv) Measure 3: embankments (M_3); and (v) Measure 4: spatial planning measures (M_4).

The cost and benefit data of the measures are obtained from the Ministry of Transport, Public Works and Water Management (1994b).

3.4. Uncertainty in model inputs and parameters

As mentioned before, besides input and parameter uncertainty, other uncertainties can be important as well, such as

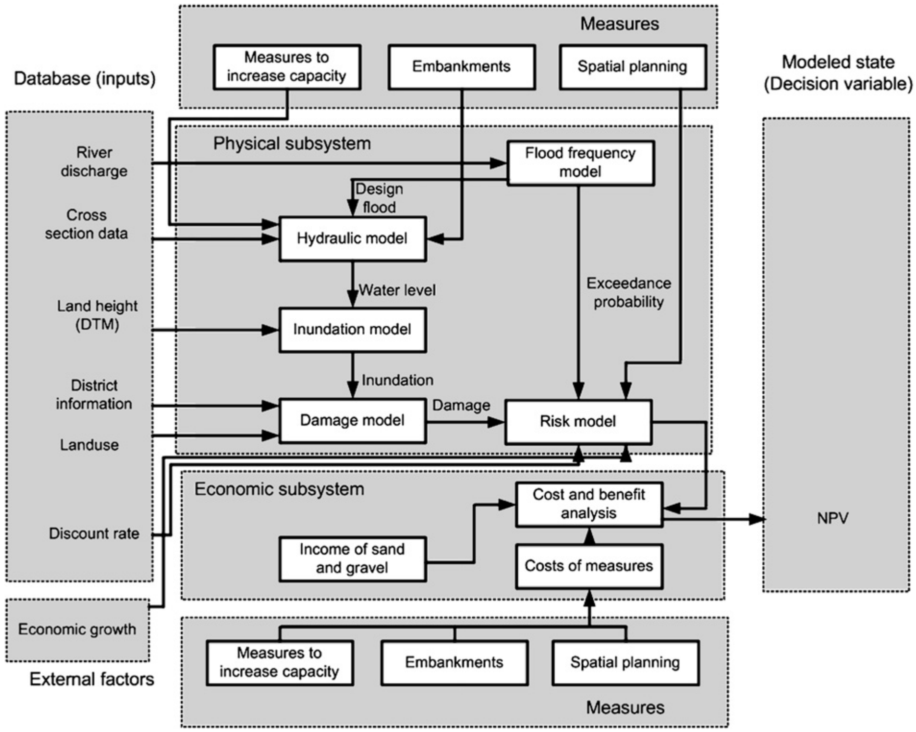

Fig. 3. System diagram for flood safety.

model structural uncertainty. However, input and parameter uncertainty are two important sources (but not all sources) and it is reasonable to start with them for demonstrating and developing the approach. Other important uncertainties such as model structural uncertainty should be included when actually applying the approach, which is the case in the second case study (Xu and Booij, submitted for publication).

In total, there are 112 inputs and parameters in the models (Xu, 2005). Most of them are factors from the damage functions for different land use types. The estimation of uncertainty in the sample's mean and sample's standard deviation of the flood flows is based on the assumption that both parameters are normally distributed. The estimation of the uncertainty in the hydraulic parameters is obtained from experts and assumed to be normally distributed as well. The distributions of other inputs and parameters are set to be uniform in shape, because insufficient data are available to infer any particular type of distribution (Campolongo and Saltelli, 1997). Ranges of uncertainty are selected either according to the information available, or in absence of such information, assuming 20% of uncertainty in input variables and parameters (nominal value $\pm 20\%$). Detailed uncertainty information can be found in Xu (2005).

4. Results from applying the appropriateness framework

4.1. Results from uncertainty analysis

The uncertainties in inputs and parameters are propagated into the model outputs (NPV) for each measure by Latin Hypercube simulation. The number of simulations is 1000. Table 1 gives the mean values and standard deviations of the model outputs associated with the five measures. In this table, Y_i represents the model outputs from M_i, $i = 0,\ldots,4$. The results in this table show that the mean NPV for M_1 is significantly lower compared with other measures and so is the standard deviation. This implies that M_1 is the best choice if the NPV is the only decision variable considered. Furthermore, if decision makers are only interested in the mean values of model outputs, Table 1 also indicates a ranking of the measures: $M_3 < M_2 < M_0 < M_4 < M_1$. M_2 and M_3 rank after M_0 (the current situation) because of their high costs.

However, if uncertainty is considered as an important factor in decision-making, the above ranking doesn't hold all the time, especially for M_0, M_2 and M_3, which produce close mean values and high standard deviations. Some standard deviations are even much higher than the differences in mean values. For example, for measure M_0 and M_4, the mean difference in the NPV is 23 million euros, while individual standard deviations are around 32 and 27 million euros, respectively. This indicates that it is difficult to distinguish these two

Table 1
Mean NPVs and standard deviations for model outputs from five measures (Y_i represents the model outputs from measures M_i, $i = 0,\ldots,4$)

Measures	Mean NPV (million euros)	Standard deviations of NPV (million euros)
M_0 (Y_0)	138	32.2
M_1 (Y_1)	54.0	17.6
M_2 (Y_2)	143	23.3
M_3 (Y_3)	150	20.9
M_4 (Y_4)	115	26.8

measures. In other words, M_4 is not so effective in improving the flood situation, especially in the presence of high uncertainty in the model outputs (M_0 is the current situation).

4.2. Results from appropriateness analysis

Goodness-of-fit tests show that the model outputs from Latin Hypercube simulation for the five measures are log-normally distributed. For risk calculations, the natural logarithms of the model outputs (ln) are used so that the parametric method of Karl (1999) can be employed to calculate the probability of obtaining an unacceptable ranking defined in Eq. (1). Fig. 4 shows the superimposed normal density based on the histograms of the log-normally transformed model outputs for the five measures. This figure shows large overlaps among the five distributions.

According to Eq. (2), the differences between model outputs from measures are of importance to calculate the risk of obtaining an unacceptable ranking. Since the log-normally transformed model outputs are normally distributed and assuming the model outputs from the five measures are independent from each other, the differences for each combination of measures are normally distributed as well (Karl, 1999). The original mean differences before the log-normal transformation are presented in Table 2.

The probability distributions of the differences for 10 combinations of measures are shown in Fig. 5. Y_i and Y_j are different model outputs from M_i and M_j, respectively, $i = 0,\ldots,4$, $j = 0,\ldots,4$, $i \neq j$. It is assumed that, in most cases, the elements from Y_i are bigger than the elements from Y_j. This means that the probability of $Y_i - Y_j > 0$ is higher than 50%. Therefore the probability that $Y_i - Y_j < 0$ is the probability defined in Eq. (1), which corresponds to the negative area of the distributions in Fig. 5. The bigger the negative areas, the more difficult it will be to distinguish the measures. For example, Fig. 5 shows that the pairs (M_0 & M_2), (M_0 & M_3) and (M_2 & M_3) are the most difficult ones to distinguish, because their negative areas are almost half of the entire distributions with respective probabilities of 0.42, 0.35 and 0.40 as shown in Table 2.

Table 2
Mean differences of model outputs, probabilities and risks of obtaining an unacceptable ranking for 10 different combinations of measures

Measures compared	Mean difference (million euros)	Probability	Risk (million euros)
M_0 & M_1	83.9	1.1×10^{-2}	8.95×10^{-1}
M_0 & M_2	5.61	0.42	2.38
M_0 & M_3	12.6	0.35	4.40
M_0 & M_4	23.3	0.28	6.60
M_1 & M_2	89.5	4.2×10^{-3}	3.78×10^{-1}
M_1 & M_3	96.5	2.4×10^{-3}	2.27×10^{-1}
M_1 & M_4	60.6	3.1×10^{-2}	1.89
M_2 & M_3	7.02	0.40	2.84
M_2 & M_4	28.9	0.20	5.69
M_3 & M_4	35.9	0.14	5.02

Table 2 also shows that, in this case study, lower mean differences correspond to higher probabilities and vice versa, for example, the values (5.61, 0.42) for M_0 & M_2 pair and (60.6, 3.1×10^{-2}) for M_1& M_4 pair. On the basis of Eq. (2), the risks for each combination of measures are calculated and presented in Table 2 as well. The risks are calculated as a result of the combined effects of the mean differences and the probabilities. The results also show that these two aspects counteract each other.

The acceptable risk needed to determine the appropriateness of models in the flood safety component is usually determined by decision makers. In this case study, a value of 6 million euros is selected for a preliminary analysis, assuming that this value represents the consensus of decision makers. Table 2 shows that most of the risks calculated are lower than this acceptable value and only the risk for the

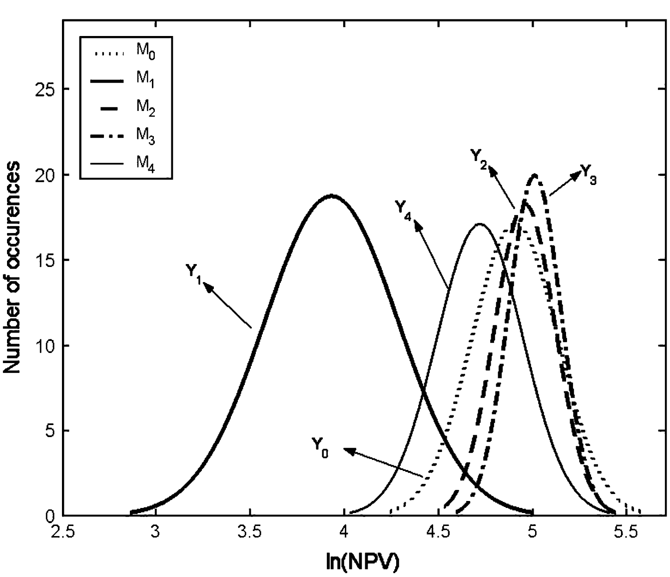

Fig. 4. Superimposed normal probability density function for log-normally transformed model outputs for five measures.

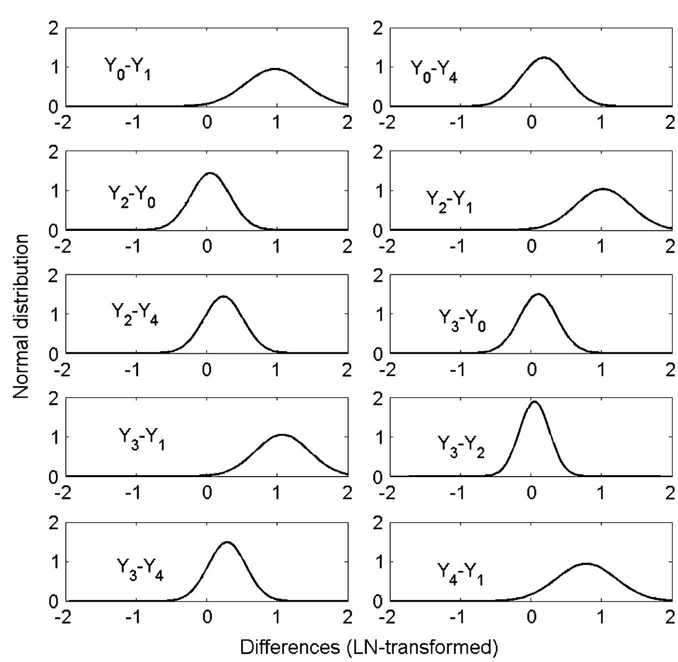

Fig. 5. Probability distributions of differences of log-normally transformed model outputs for 10 combinations of measures.

combination M_0 & M_4 is higher. The occurrence of an unacceptable risk means that the ranking of M_4, although being better than M_0 on average, is not acceptable because of the high levels of uncertainty in model outputs. Therefore, in this case it is concluded that the models are inappropriate because of the unacceptable risk. This also means that it is not possible to obtain an acceptable ranking due to the high levels of uncertainty in the decision variable. Thus, according to the appropriateness framework, the models need to be improved.

4.3. Results from sensitivity analysis

In this section, model improvement is implemented by reducing the uncertainty in the most important sources. To achieve this purpose, the Morris method is used to determine the relative importance of the 112 inputs and parameters based on their respective contributions to the output uncertainty (Morris, 1991). The mean and standard deviation (SD) of the distribution of elementary effects for each input or parameter are calculated (Morris mean and SD). The elementary effect EE is the indicator used in the Morris method to evaluate the effects of ith input or parameter factor on the outputs. It is defined as:

$$\mathrm{EE}_i(X) = \frac{Y(x_1, \ldots, x_{i-1}, x_i + \Delta, x_{i+1}, \ldots, x_k) - Y(X)}{\Delta} \tag{3}$$

where X is any selected value in the region of experiment, Y is the model output, k is the number of inputs and parameters and Δ is a pre-determined multiple representing variations of inputs and parameters. The order of importance of inputs and parameters is finally obtained by calculating the Morris distance, which is the Euclidean distance from the coordinate (Morris mean, Morris SD) to the origin (0, 0) (Campolongo and Saltelli, 1997). The larger the Morris distance, the more important the inputs and parameters are.

Table 3 shows the Morris means, Morris SDs and Morris distances, along with the order of importance, for the top 15 inputs and parameters in the flood safety models. It can be observed that the most important inputs and parameters come from the hydraulic model while damage factors are less important.

4.4. Results from model improvements

For the purpose of illustration of improving the models, only the first 14 inputs and parameters shown in Table 3, whose Morris distances are higher than 10^7, are regarded to be important. On the basis of this list, possible suggestions about realistic uncertainty reductions can be derived. For example, as Table 3 shows, the slope and bed level factors (i_s, i_b, a_s and a_b) used in calculating the bed levels are among the most important parameters. This suggests that a reduction of uncertainty in bed level calculation would be very helpful in reducing the uncertainty in model outputs. A realistic reduction of uncertainty could be obtained by improving the model structure, for example, by adding a morphological model, or by obtaining more high quality measurements of bed levels. In this study, however, only hypothetical reductions of uncertainty in parameters and inputs are used to investigate how uncertainty reductions affect the model outputs and the risk of obtaining an unacceptable ranking. Two cases are proposed. Case 1 deals with a reduction of uncertainty in several of the most important inputs and parameters. The uncertainty in the slopes, bed level factors and the depths of the main channel (i_s, i_b, a_s, a_b, h_b and h_s) is reduced from 5% to 1% and the uncertainty in the Nikuradse coefficients (K_b and K_s) in the floodplains is reduced from 20% to 10%. Case 2 assumes that the 14 important inputs and parameters are deterministic (i.e. without uncertainty). The use of two hypothetical cases aims to investigate the effects of different levels of uncertainty reduction on the appropriateness analysis of models.

Goodness-of-fit tests show that the model outputs (NPV) after uncertainty reductions are still log-normally distributed. The superimposed normal densities based on histograms of the differences for different pairs of measures (log-normally transformed) are shown in Fig. 6. Fig. 6a shows the distributions for all measures for Case 1 and Case 2 and Fig. 6b shows

Table 3
Inputs and parameters in order of importance according to Morris distances (here SDs mean Morris standard deviations)

Inputs and parameters	Descriptions	Morris means	Morris SDs	Morris distances	Order of importance
i_s	Slope of Sand Meuse	10^{13}	10^{14}	10^{14}	1
i_b	Slope of Border Meuse	10^{13}	10^{13}	10^{13}	2
K_s	Nikuradse of Sand Meuse floodplains	10^8	10^9	10^9	3
K_b	Nikuradse of Border Meuse floodplains	10^7	10^9		4
a_s	Bed level factor for Sand Meuse	10^7	10^8	10^8	5
h_b	Depth of Border Meuse main channel	10^8	10^8		6
h_s	Depth of Sand Meuse main channel	10^8	10^8		7
a_b	Bed level factor for Border Meuse	10^8	10^8		8
B_{b1}	Width of Border Meuse main channel	10^7	10^7	10^7	9
C_s	Chézy coefficient of Sand Meuse main channel	10^6	10^7		10
B_{b2}	Width of Border Meuse floodplains	10^6	10^7		11
B_{s2}	Width of Sand Meuse floodplains	10^6	10^7		12
C_b	Chézy coefficient of Border Meuse main channel	10^6	10^7		13
B_{s1}	Width of Sand Meuse main channel	10^5	10^7		14
$\widehat{\mu}$	Sample mean of flood flow	10^5	10^6	$<10^7$	15
Others					16–112

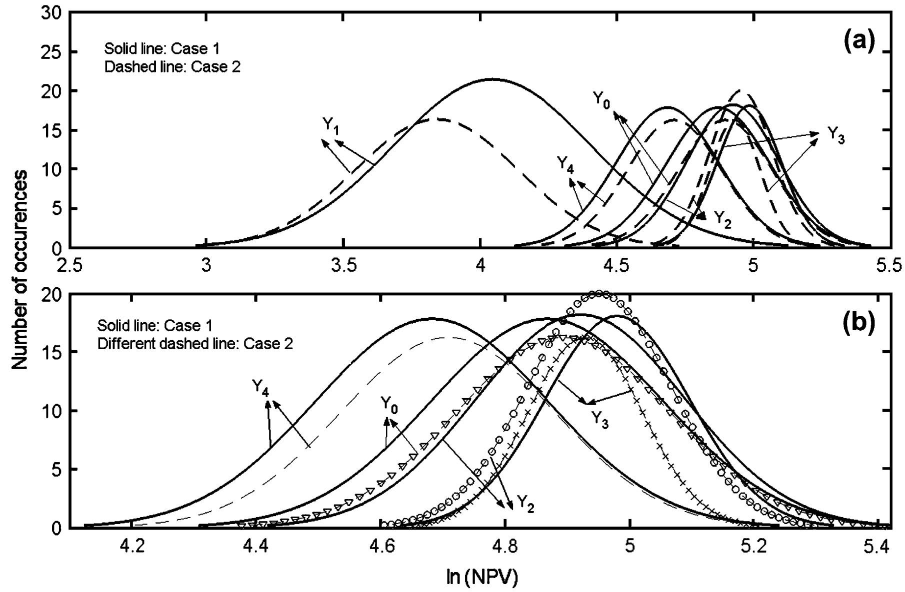

Fig. 6. Superimposed normal probability density function of ln(NPV) for Case 1 and Case 2 for five measures (a) and for measures M_0, M_2, M_3 and M_4 (b).

the distributions for all measures except Measure 1 for Case 1 and Case 2. This figure shows that the distributions of ln(NPV) from both cases are narrower than those in Fig. 4, slightly for Case 1 and obviously for Case 2. This indicates that the uncertainty in the model outputs has decreased after reducing the uncertainty in the most important inputs and parameters. Fig. 6 also shows a shift of the distributions (mainly represented by a change in mean NPVs) when the uncertainty is reduced. These shifts of distributions mainly resulted from the model behavior (e.g., non-linearity and non-monotony) of the flood safety model.

For the convenience of comparison, the situation before model improvements is represented by Case 0. The mean differences for each pair of measures are shown in Table 4 for Case 0, Case 1 and Case 2. This table shows that the mean differences change when the uncertainty is reduced. For example, the mean difference for the pair (M_2 & M_3) changes from 7.02 million euros for Case 0, to 7.38 for Case 1, to 3.50 for Case 2. Fig. 6b shows the switch of distributions (shown by the switch of mean values) from M_2 and M_3 for Case 1 and Case 2.

Table 4 also shows the probabilities for different pairs of measures for the three cases. The results indicate that reducing uncertainty in the most important inputs and parameters has significant effects on the ranking of measures, especially for Case 2, which shows an obvious decrease of the probabilities. Small numbers of probabilities increase, such as the probabilities related to M_1, because of the shift of the distributions. The results from Table 4 also indicate that M_0, M_2 and M_3 are comparable with high probabilities (>30%), especially for M_2 and M_3. Even if the uncertainty is reduced a lot, it is still difficult to distinguish them. As shown in Fig. 6b, the distributions for M_2 and M_3 switched for Case 1 and Case 2. This situation indicates that reducing uncertainty is not worthwhile anymore because too much effort would be needed to reduce the probability even only by 1%. Hence, this aspect should be considered when the appropriateness of models is analyzed. From this point of view, M_2 and M_3 can be regarded to have the same effects on the decision variable.

Finally, Table 4 presents the risks for different pairs of measures for the three cases. The table shows that most risk values decrease while a few slightly increase because of the shifts of the distributions. The effects of model behavior, such as nonlinearity, need to be investigated in future studies.

The analysis shows that the models improved after reducing the uncertainty in the most important inputs and parameters. All risks calculated are lower than the acceptable risk (6 million euros) for Case 1 and Case 2, and it is more obvious for Case 2. Therefore, the models are judged to be appropriate under the reduced uncertainty situation, for both Case 1 and Case 2. Meanwhile, the measures can be ranked as: $M_3 = M_2 < M_0 < M_4 < M_1$, which can be accepted by decision makers. This case study shows that, if realistic reduction of uncertainty to the level in Case 2 is difficult or even impossible to achieve, Case 1 is good enough to achieve appropriate models.

5. Discussion

This paper addressed the problem of determining appropriate models in the Dutch Meuse DSS by applying an appropriateness framework approach developed specifically for this purpose. In this approach, the risk of obtaining an unacceptable ranking for each combination of measures incorporates the mean differences of the model outputs from pairwise measures ('consequences'), which are used as a scaling factor. Another option to represent the consequence of obtaining an unacceptable ranking may be the value at 95% probability of the difference between two measures. This needs further investigations in future.

The appropriateness framework approach starts from simple but reasonable models. One advantage of starting from

Table 4
Mean differences of model outputs, probabilities and risks of obtaining an unacceptable ranking for 10 different combinations of measures

Measures compared	Mean difference (million euros)			Probability			Risk (million euros)		
	Case 0	Case 1	Case 2	Case 0	Case 1	Case 2	Case 0	Case 1	Case 2
M_0 & M_1	83.9	71.5	86.7	1.1×10^{-2}	2.1×10^{-2}	1.0×10^{-3}	8.95×10^{-1}	1.51	8.97×10^{-2}
M_0 & M_2	5.61	6.88	6.93	0.42	0.41	0.39	2.38	2.85	2.70
M_0 & M_3	12.6	14.3	3.42	0.35	0.30	0.43	4.40	4.32	1.46
M_0 & M_4	23.3	22.3	22.8	0.28	0.24	0.22	6.60	5.38	5.10
M_1 & M_2	89.5	78.4	93.6	4.2×10^{-3}	1.3×10^{-2}	2.3×10^{-4}	3.78×10^{-1}	1.06	2.15×10^{-2}
M_1 & M_3	96.5	85.8	90.1	2.4×10^{-3}	6.6×10^{-3}	2.1×10^{-4}	2.27×10^{-1}	5.68×10^{-1}	1.97×10^{-2}
M_1 & M_4	60.6	49.2	63.9	3.1×10^{-2}	5.8×10^{-2}	5.5×10^{-3}	1.89	2.83	3.54×10^{-1}
M_2 & M_3	7.02	7.38	3.50	0.40	0.38	0.44	2.84	2.84	1.54
M_2 & M_4	28.9	29.2	29.7	0.20	0.17	0.12	5.69	4.97	3.58
M_3 & M_4	35.9	36.6	26.3	0.14	0.09	0.13	5.02	3.19	3.40

these models is to save efforts and time with respect to the development of very complex models. Nevertheless, there is still a possibility that these models miss some components in the system, which may contribute to model outputs. Carrying out an uncertainty analysis (as performed in this case study) cannot display what the missing components really are because only the uncertainty in inputs and parameters was considered. This should be investigated by comparing model outputs from models with different model complexity. It can become apparent whether complex models are more appropriate than simpler ones.

The appropriateness framework in this paper is mainly designed for planning and strategic management. Therefore, the models used are supposed to be simpler than those for other purposes. However, there is still a chance that achieving an acceptable ranking could be difficult. The models initially developed or chosen may not be adequate. Then many efforts are often put on collecting more data or even building a new model. Both may cost a lot of time and money. Furthermore, three aspects may hinder the ranking. The first is model behavior, such as model non-linearity. The effect of such model behavior on risk calculations has been shown in the case study. The second aspect again relates to the cost and time involved to reduce the uncertainty. Sometimes it is not worthwhile to spend a huge amount of money and time to get completely distinguishable measures. The third aspect is that some uncertainties in the models cannot be reduced, for example the uncertainty due to current knowledge limitations and the stochasticity in the natural world. Under such conditions, as mentioned in Section 2.2, the indistinguishable measures have to be regarded to have similar effects on the decision variable and decision makers must live with some uncertainty. Finally, the attitudes of decision makers (indicated by acceptable risks) will play a significant role in determining the efforts needed to achieve appropriate models. In practical applications, the determination of acceptable risks is rather important in applying the appropriateness framework. This has been further investigated in the second case study (see Xu and Booij, submitted for publication).

In this case study, the model outputs from different measures were assumed to be independent of each other. This means that the uncertainties originating from the model inputs and parameters are independent of each other (Reichert and Borsuk, 2005). Another assumption could be that all uncertainties are fully dependent, thus the distributions of model outputs from different measures will be highly dependent. Under this extreme assumption, the risks calculated will be nearly zero and a definite ranking can be obtained. These two different assumptions mean that the interdependency among model outputs from different measures will play an important role in calculating the risks of obtaining an unacceptable ranking. However, in reality, both assumptions are only extreme situations. In a future study, a more reasonable assumption is needed to consider some degree of interdependency, so that the correlations between model outputs from different measures could be taken into account. However, modeling dependencies among the model outputs is not a trivial thing. According to Reichert and Borsuk (2005), a possible way to take the dependencies into account is through the construction of dependency structures by copulas. This is often difficult for the case with many uncertainty sources (which is the case in this paper). One may need to first figure out several major sources of uncertainty and it is also difficult to take into account the effects of model structural uncertainties. The independent case in this paper in fact gave more conservative results compared to the dependent case (a smaller risk), which is in favor of the application of the approach.

Besides the ranking information of the measures, the distributions of differences for each combination of measures can be very important to decision makers. The knowledge about the distributions of differences turns out to be useful for evaluating the effectiveness of river engineering measures in the presence of high levels of uncertainty (see Fig. 5). This is often ignored in the real decision-making process.

The models used in a DSS according to the developed systematic approach are fit for decision maker's use because on the one hand, models chosen or built by modelers are based on decision makers' problems and objectives and on the other hand, the ranking problem is solved through the use of the acceptable risk, which is determined by decision makers as well. Therefore, the appropriateness framework can be used as a tool to stimulate the communication between decision makers and modelers, promote the use of models in decision-making under uncertainty, avoid possible efforts to develop over-complex models and improve the quality of decision-making.

6. Conclusions and recommendations

This paper proposed a systematic approach to develop appropriate models which can be used in a DSS for integrated river basin management. The Dutch Meuse case study gave a good demonstration on how the appropriateness framework can be applied to a DSS.

However, the Dutch Meuse case study was mainly used for the further development of the appropriateness framework and focused on how the proposed approach worked out. In future applications of this approach to other DSSs for river basin management, some aspects need to be further investigated. First of all, since the value of the acceptable risk is subjective and could vary among different decision makers, it is recommended that this value is determined in a more delicate and careful way in a future case study (through a collective view and consensus of relevant decision makers). Second, realistic uncertainty reductions using the techniques introduced in Section 2 should be implemented. Third, the dependencies among model outputs need to be investigated, so does the model structural uncertainty. Finally, the practical issues of costs and time for achieving appropriate models should be taken into account. A second case study is therefore proposed to take these aspects into account and should be used particularly for validation purposes. A DSS for the River Elbe (the part in Germany) has been selected for this purpose and it has been investigated if the approach can be applied to a more realistic situation that is different from the one it is designed for (Xu and Booij, submitted for publication).

Acknowledgements

The work described in this paper was part of a research project funded by both the Royal Netherlands Academy of Arts and Sciences (KNAW) and the University of Twente. The authors would like to thank Siebolt Folkertsma from the Dutch Directorate for Public Works and Water Management in The Netherlands for discussions and providing data for the Dutch Meuse River and Prof. C.B. Vreugenhil from the University of Twente for interesting discussions about the use of appropriate models in river basin management. The authors also own great thanks to Prof. Y.K. Tung from Hong Kong University of Science and Technology for his valuable contributions to the completion of this paper. Finally, the authors would like to thank the two anonymous reviewers for their contributions to improve the quality of this paper significantly.

References

Andreu, J., Capilla, J., Sanchis, E., 1996. AQUATOOL, a generalized decision-support system for water-resources planning and operational management. Journal of Hydrology 177, 269–291.

Barnwell, T.O., Brown, L.C., Marek, W., 1986. Development of a Prototype Expert Advisor for the Enhanced Stream Water Quality Model QUAL2E. Internal report. Environmental Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Athens, GA.

Basson, L., Petrie, J.G., 2007. An integrated approach for the consideration of uncertainty in decision making supported by life cycle assessment. Environmental Modelling & Software 22 (2), 167–176.

Bevington, P.R., Robinson, D.K., 1992. Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill.

Booij, M.J., 2003. Determination and integration of appropriate spatial scales for river basin modelling. Hydrological Processes 17, 2581–2598.

Brooks, R.J., Tobias, A.M., 1996. Choosing the best model: level of detail, complexity, and model performance. Mathematical and Computer Modelling 24 (4), 1–14.

Caminiti, J.E., 2004. Catchment modelling – a resource manager's perspective. Environmental Modelling & Software 19 (11), 991–997.

Campolongo, F., Saltelli, A., 1997. Sensitivity analysis of an environmental model: a worked application of different analysis methods. Reliability Engineering & System Safety 57, 49–69.

Chwif, L., Paul, R.J., 2000. On simulation model complexity. In: Proceedings of the 2000 Winter Simulation Conference, Orlando, FL, USA.

De Blois, C.J., 2000. Uncertainty in Large-scale Models for Decision Support in Water Management. University of Twente, Enschede, The Netherlands.

De Kort, I.A.T., Booij, M.J., 2007. Decision making under uncertainty in a decision support system for the Red River. Environmental Modelling & Software 22 (2), 128–136.

Dorner, S., Shi, J., Swayne, D., 2007. Multi-objective modeling and decision support using a Bayesian network approximation to a non-point source pollution model. Environmental Modelling & Software 22, 211–222.

Duchnese, S., Beck, M.B., Reda, A.L., 2001. Ranking stormwater control strategies under uncertainty: the River Cam case study. Water Science and Technology 43, 311–320.

Dunn, S.M., Mackay, R., Adams, R., Oglethorpe, D.R., 1996. The hydrological component of the NELUP decision support system: an appraisal. Journal of Hydrology 177, 213–235.

European Commission, 2000. Communication on the Precautionary Principle, COM (200) 1. European Commission, DGR, RTD. Eurobarometer 55.2, Europeans, Science and Technology.

Fread, D.L., 1985. Applicability Criteria for Kinematic and Diffusion Routing Models. Laboratory of Hydrology, National Weather Service, NOAA, U.S. Department of Commerce, Silver Spring, MD.

Giupponi, C., 2007. Decision support systems for implementing the European Water Framework Directive: the MULINO approach. Environmental Modelling & Software 22, 248–258.

Howard, R., 1988. Decision analysis: practice and promise. Management Science 34 (6), 679–695.

Jakeman, A.J., Hornberger, G.M., 1993. How much complexity is warranted in a rain-fall–runoff model? Water Resources Research 29 (8), 2637–2649.

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.

Kaplan, S., Garrick, B.J., 1981. On the quantitative definition of risk. Risk Analysis 1 (1), 11–27.

Karl, V.B., 1999. Statistical Distributions in Engineering. Cambridge University Press, UK.

Keen, P.G.W., Scott Morton, M.S., 1978. Decision Support Systems: An Organizational Perspective. Addison-Wesley Inc., Reading, MA.

Keisler, J.M., 2002. Attribute-based differentiation of alternatives. Journal of Multi-criteria Decision Analysis 11, 315–326.

Matheson, D., Matheson, J., 1998. The Smart Organization. Harvard Business School Press, Boston, MA.

McMahon, G.F., Fitzgerald, R., McCarthy, B., 1984. BRASS model: practical aspects. Journal of Water Resources Planning and Management, ASCE 110 (1), 75–89.

Ministry of Public Housing, Physical Planning, and Environmental Protection, 1985. Handing Uncertainty in Environmental Impact Assessment. Milieu-Effect Rapportage, 18, The Netherlands.

Ministry of Transport, Public Works and Water Management, 1994a. Research on the Meuse Flood. Section 9, Damage Modeling. Delft, The Netherlands.

Ministry of Transport, Public Works and Water Management, 1994b. Research on the Meuse Flood. Section 10, Embankments, Dikes and Special Constructions. Delft, The Netherlands.

Morgan, M.G., Henrion, M., 1990. Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33, 161–174.

National Research Council, 2000. Risk Analysis and Uncertainty in Flood Damage Reduction Studies. Committee on Risk-Based Analysis for Flood Damage Reduction, Water Science and Technology Board. National Academy Press, 2101 Constitution Avenue, N.W. Washington, DC.

Nihoul, J.C., 1994. Do not use a simple model when a complex one will do. Journal of Marine Systems 5, 401–410.

Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. Journal of Hydrology 242, 275–301.

Reckhow, K.H., 1994. Importance of scientific uncertainty in decision-making. Environmental Management 18, 161–166.

Reda, A.L.L., Beck, M.B., 1997. Ranking strategies for stormwater management under uncertainty: sensitivity analysis. Water Science and Technology 36 (5), 357–371.

Reichert, P., Borsuk, M.E., 2005. Does high forecast uncertainty preclude effective decision support? Environmental Modelling & Software 20 (8), 991–1001.

Rhijnsburger, E., 1997. Of Triangles Proportion to Three-unity. An Investigation to the Administrative Bottlenecks in the Meuse. Report No. 5. University of Maastricht, Maastricht, The Netherlands.

Rogers, P., 1978. On the choice of 'appropriate model' for water resources planning and management. Water Resources Research 14 (6), 1003–1010.

RWS(Dutch Directorate for Public Works and Water Management), 1997. Daily Borgharen Discharge Data. RWS, Maastricht, The Netherlands.

RWS(Dutch Directorate for Public Works and Water Management), 2001. Propagation Speed for High Water Wave in the River Meuse. VVM-rapport nr. 7. RWS, Maastricht, The Netherlands.

Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity Analysis. John Wiley & Sons Ltd., England.

Schielen, R.M.J., Gijsbers, P.J.A., 2003. DSS-large rivers: developing a DSS under changing societal requirements. Physics and Chemistry of the Earth Part B – Hydrology Oceans and Atmosphere 28 (14–15), 635–645.

Schlüter, M., Rüger, N., 2007. Application of a GIS-based simulation tool to illustrate implications of uncertainties for water management in the Amudarya river delta. Environmental Modelling & Software 22, 158–166.

Shaw, E.M., 1994. Hydrology in Practice. T.J. International Ltd., Padstow, Cornwall, Great Britain.

Sojda, R.S., 2007. Empirical evaluation of decision support systems: needs, definitions, potential methods, and an example pertaining to waterfowl management. Environmental Modelling & Software 22, 269–277.

Van Asselt, M.B.A., 2000. Perspectives on Uncertainty and Risk. The PRIMA Approach to Decision Support. Kluwer Academic Publishers, The Netherlands.

Vreugdenhil, C.B., 2002. Accuracy and reliability of numerical river models. Journal of American Water Resources Association 38 (4), 1083–1095.

Vreugdenhil, C.B., 2006. Appropriate models and uncertainties. Coastal Engineering 53, 303–310.

Xu, Y., 2005. Appropriate Modelling in Decision Support Systems for River Basin Management. Ph.D. thesis, University of Twente, Enschede, The Netherlands.

Xu, Y., Booij, M.J., 2004. Appropriate modelling in DSSs for river basin management. In: Pahl-Wostl, C., Schmidt, S., Rizzoli, A.E., Jakeman, A.J. (Eds), Complexity and Integrated Resources Management. In: Transactions of the Second Biennial Meeting of the International Environmental Modelling and Software Society, vol. II. 14–17 June 2004, Osnabrück, Germany, pp. 611–616.

Xu, Y., Booij, M.J. Validation of a model appropriateness framework using the Elbe decision support system. Environmental Modelling & Software, submitted for publication.

Zagona, E.A., Terrance, J.F., Richard, S., Timothy, M., Goranflo, H.M., 2001. RiverWare: a generalized tool for complex reservoir systems modeling. Journal of Water Resources Planning and Management, ASCE 37 (4), 913–929.