Click Here for Full Article

Multicriterion decision analysis approach to assess the utility of watershed modeling for management decisions

Amin Elshorbagy[1]

Received 15 May 2005; revised 9 May 2006; accepted 6 June 2006; published 14 September 2006.

[1] This paper employs the multicriterion decision analysis (MCDA) technique in a novel way to address the efficiency of watershed instrumentation programs and the efficacy of watershed modeling. A case study of reconstructed watersheds in northern Alberta, Canada, is used to illustrate the proposed usage of the MCDA technique. The watersheds have been disturbed as a result of oil sands mining activities. Assessing the performance of the reconstructed watersheds with regard to restoring the hydrology of the disturbed watershed is a crucial issue for both the mining industry and other stakeholders. The problem is formulated in a multicriterion context. A payoff matrix containing seven evaluation criteria and three different soil covers as feasible alternatives is constructed. The system dynamics watershed (SDW) model is used to simulate the reconstructed watersheds over a period of 61 years using historical meteorological records. Accordingly, 61 payoff matrices that are populated using the results of the SDW model are evaluated. A multicriterion decision analysis framework is implemented to evaluate the different alternatives with respect to the chosen set of criteria. The three alternatives are ranked every year, and accordingly, the probability that a certain alternative dominates others is estimated. The alternative that has the highest probability of occupying the top rank over the period of analysis is indicated as the best alternative. The probability value is called the probability of making the right decision (PMRD). Various types of uncertainty analyses are conducted to evaluate the sensitivity of the final decision to changes in the scores of the evaluation matrices. An index, named the confidence in the PMRD, is developed to quantify the reliability of the results of the watershed model. The results highlight the utility of modeling as a possible alternative to some components of the intensive instrumentation program. Moreover, areas of deficiency and inaccuracies in the watershed model are identified for further improvements.

Citation: Elshorbagy, A. (2006), Multicriterion decision analysis approach to assess the utility of watershed modeling for management decisions, *Water Resour. Res.*, *42*, W09407, doi:10.1029/2005WR004264.

1. Introduction

[2] Since the early 1960s, mathematical hydrologic models that account for the continuous dynamics of hydrologic processes have been in use. In the field of watershed hydrology, there have been a few attempts to provide a comprehensive account of watershed models or at least of the most commonly used models. Although notable efforts have been made in this direction [*Singh and Frevert*, 2002a, 2002b; *Singh*, 1995], the number of developed watershed models is more than could possibly be gathered in a single all-inclusive publication. Why hydrologists have needed so many models is an interesting question that has many possible answers. Some of the possible reasons are (1) the unavailability of a previously developed model at the time of need, (2) a lack of awareness that a suitable previously developed model is available, (3) a lack of understanding of the processes and algorithms considered in the available models, and most significantly (4) a conviction by the developer that available models do not satisfy the conditions of the situation at hand, and therefore the situation under consideration requires a new model.

[3] As *Schaake* [2002] noted, the exercise of modeling is both the art and the science of applying a limited and imperfect understanding of the "real" world. Such an understanding requires knowledge of the physics of hydrologic processes at different spatial and temporal scales, and information on soils, vegetation, topography, and water and energy forcing variables. These time- and location-varying requirements should be the only acceptable motives for developing new models. Calibration and assessment of model performance are two of the most cumbersome processes involved in watershed model development. These two processes require extensive instrumentation and a monitoring program to provide the necessary data. In order to go beyond rainfall-runoff modeling and to reduce the parameter uncertainty, the calibration of the developed model should be based on all simulated processes (e.g., soil moisture, evapotranspiration) [*Elshorbagy et al.*, 2005; *Wooldridge et al.*, 2003], not just runoff.

[1]Centre for Advanced Numerical Simulation, Department of Civil and Geological Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Copyright 2006 by the American Geophysical Union.
0043-1397/06/2005WR004264$09.00

[4] Many experimental watersheds and extensive monitoring programs have been conducted around the world to measure various hydrological and meteorological processes, with the purpose of understanding the real world and thus of developing better watershed models. However, fewer efforts have been devoted to the issue of simplicity vs. complexity [e.g., *Jakeman and Hornberger*, 1993] in watershed models, as well as to the extensiveness of the instrumentation required to carry out the modeling exercise. The desire of engineers and hydrologists to collect as much data as possible regardless of the value of such data is becoming a concern in light of the increasing cost of measurements. Paradoxically, the data collected in certain situations (e.g., long-term monitoring programs) go beyond the capabilities of analysts who are actually handling such data, causing unwarranted confusion rather than an increase in the understanding of the hydrologic systems. The confusion is mostly due to the unmanageable size of data sets that requires tremendous efforts of screening, verification, and documentation.

[5] In this paper, a systematic methodology of evaluating the relative value of watershed modeling and monitoring as two interrelated processes that provide feedback to each other is developed using the multicriterion decision analysis (MCDA) technique. This methodology is applied to a case study of evaluating reclamation strategies for disturbed watersheds. The proposed methodology enables the establishment of an objective-oriented measure to assess the utility of watershed modeling and models. The approach also helps in developing objective guidelines for refocusing and redirection of current monitoring programs.

2. A Problem to Be Addressed: Case Study

[6] The mining of oil sands in northern Alberta, Canada, is the focus of this paper. The processes used leave behind large pits, tailings facilities, and overburden (piles of saline shale layer) in which the natural hydrology of surface and groundwater has been completely disrupted. The oil sands industry is committed to reconstructing functioning landscapes by designing reclaimed watersheds to restore the different functions of nature, such as habitat function (hosting aquatic ecosystems), production function (e.g., biomass), and carrier function (for dissolved and suspended materials). The carrier function plays a central role in land degradation processes such as erosion, sedimentation, and the leaching of nutrients through moving surface and subsurface water [*Falkenmark*, 1997]. The restoration of the above mentioned functions relies first and foremost on the restoration of functioning hydrologic systems, a central feature of which is sufficient water to sustain revegetation efforts.

[7] The mining of oil sands near Fort McMurray, Alberta, involves the stripping of the saline/sodic overburden in order to gain access to the oil-bearing formation. The overburden is placed in large mined-out pits and surface dumps and is recontoured before being capped with a mandated 1.0 m soil cover. The potential for slope instability, subsidence, and salinization resulting from the character of the saline/sodic material and its interaction with fresh water makes it imperative that the amount of precipitation percolating below the root zone be minimized [*Barbour et al.*, 2001].

[8] The oil sands mining industry is conducting large-scale cover experiments at the Mildred Lake mine in order to assess the performance of different reclamation strategies. Four 1-hectare prototype covers were placed on an area referred to as the South Hills Overburden (Figure 1) to study the basic mechanisms controlling moisture movement within the cover systems. Three covers were constructed in 1999 with thicknesses of 1.0 m, 0.50 m, and 0.35 m, each composed of a thin layer of peat (15–20 cm) overlying varying thickness of secondary (till) soil. A fourth study site was established at an area of previously reclaimed (1996) watershed capped with a 1.0 m cover of peat/secondary mix. A field instrumentation program was carried out consisting of detailed monitoring of matric suction, volumetric water content, and temperature within the different soil profiles, as well as measurements of runoff, interflow, and site-specific meteorological conditions [*Meier and Barbour*, 2002].

[9] The ability of the soil covers to maintain sufficient soil moisture during growing seasons is an important indicator of the efficacy of the covers in restoring the production function of the watershed. The minimization of water percolating into the underlying shale is another indicator of cover success [*Boese*, 2003; *Elshorbagy et al.*, 2005]. In this case the cover serves to minimize undesirable future subsidence and salinization, which in turn could affect the carrier and production functions of the watershed. A site-specific system dynamics watershed (SDW) model has been developed by *Elshorbagy et al.* [2005] to perform continuous simulation of the reconstructed watersheds. The developed model helps quantify the above mentioned two indicators of soil moisture retention and minimization of water percolation into the underlying shale and potentially will help the mining industry to develop a comprehensive understanding of the hydrologic performance of the reclamation strategies.

[10] Some of the main concerns of the oil sands industry are the future and the evolution of the reconstructed watersheds over time and whether destroyed forests will regrow and become self-sustaining. Accordingly, the current plan is to continue the monitoring program as a long-term project, but the financial, technical, and human resources needed for long-term commitment to such a program could be prohibitive. As an example of the financial costs, each one of the three eddy covariance systems installed in the field to measure actual evapotranspiration costs roughly $60,000 Canadian dollars, in addition to the annual maintenance costs. Volumetric soil moisture content, matric suction, and soil temperature are recorded every 30 min along varying depths within the peat, till, and shale soil layers. The industry is interested in the answer to the following fundamental question: Which one of the three proposed covers is the best alternative to adopt as the optimal reclamation strategy?

3. Rationalization of Monitoring and Modeling

[11] Developing a model that can mathematically reproduce and simulate the ‘‘real’’ system is valuable because it can be used to replace a significant portion of the collected data. The SDW model [*Elshorbagy et al.*, 2005] was developed for this purpose, and its performance was assessed using traditional error statistics with regard to its

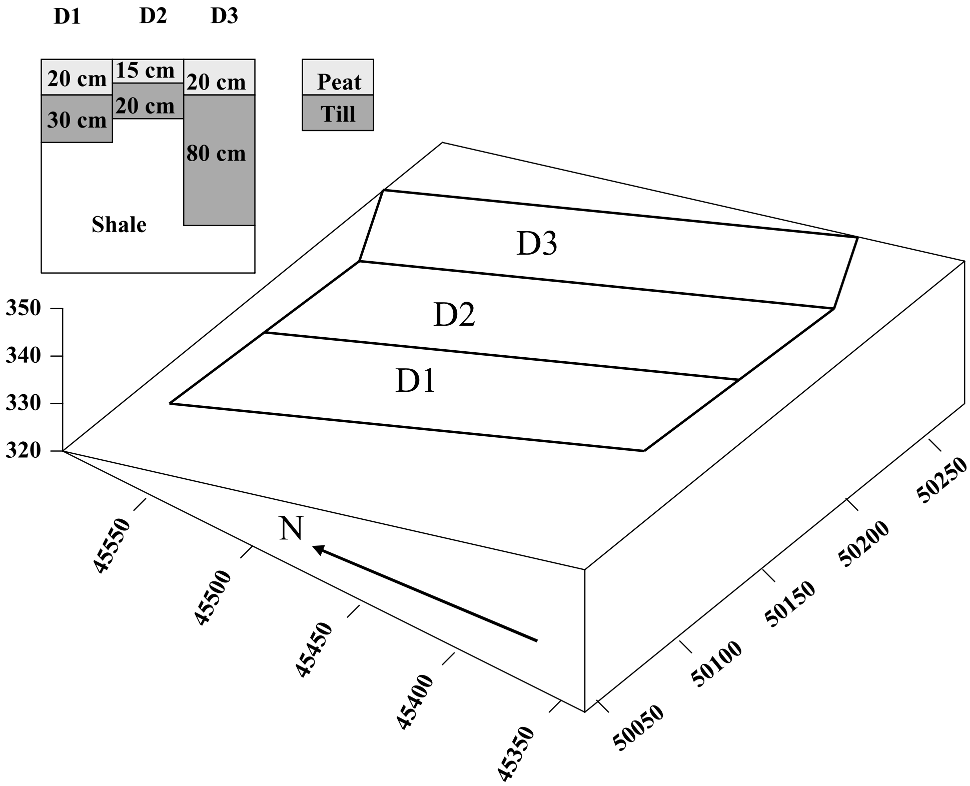

Figure 1. Prototype cover site [*Elshorbagy et al.*, 2005].

ability to simulate soil moisture, runoff, and evapotranspiration. However, the remaining question, which is addressed in this paper, is whether its predictive error and uncertainty could lead to the wrong decision regarding the best reclamation strategy. Because the SDW model was built to be a site-specific model, the importance of assessing its performance with regard to its impact on the final decision cannot be overemphasized. This paper presents, as one of its objectives, an objective-oriented assessment of the model. This information cannot be easily provided using the traditional error measures (e.g., the root mean square error, RMSE) that contrast simulated results against observed values.

4. Proposed Analysis Framework

[12] The methodology proposed in this paper is as follows: (1) identify a set of criteria that can be used to evaluate the hydrologic performance of each soil cover; (2) construct a payoff matrix that consists of the proposed three soil covers in columns and the selected criteria in rows; (3) use the SDW model to quantify the impact of each soil cover with regard to the evaluation criteria (impact assessment or scoring) on an annual basis; (4) evaluate the various soil covers (the alternatives) using a prescriptive decision analysis approach; (5) repeat steps (3) and (4) over multiple years, using available historical records of meteorological data as inputs to the SDW model, to account for the effect of climate variability; (6) repeat step (4) using the observed data rather than the outputs of the simulation model; and (7) conduct sensitivity analysis and uncertainty analysis, and conclude. Apparently, the multicriterion decision analysis (MCDA) technique is a suitable approach to handle the proposed methodology. It is possible to simulate the performance of the soil covers every year and consider the expected values (the mean value of each criterion over the entire period of the analysis) to populate one evaluation matrix. However, this approach will result in a decision that is based on a "virtual" year represented by an evaluation matrix populated with the "mean" scores. This approach obscures the effect of the climatic variability on the decision regarding the best soil cover. In this paper, a more comprehensive approach of constructing and evaluating multiple evaluation matrices; one for each year, is adopted.

5. Multicriterion Decision Analysis Technique

[13] Multicriterion decision analysis (MCDA) traditionally has been used in water resource literature as a major component of decision support systems (DSS) [*Stansbury et al.*, 1991; *Goicoechea et al.*, 1992; *Qureshi and Harrison*, 2001; *Fassio et al.*, 2005]. The central role of the MCDA technique is a trade-off analysis that compares the impacts of each alternative and indicates which alternative most nearly satisfies all relevant concerns (criteria). This is a useful technique, especially when the impacts cannot be readily estimated in monetary terms [*Belton and Stewart*, 2003]. The MCDA technique has been applied to an array of problems in water resources, including water transfer options and reservoir operation [*Stansbury et al.*, 1991; *Ko et al.*, 1992; *Harboe*, 1992; *Bogardi and Duckstein*, 1992; *Roy et al.*, 1992; *Tecle*, 1992; *Srdjevic et al.*, 2004; *Raju and Duckstein*, 2004; *Fassio et al.*, 2005], design of monitoring networks [*Woldt and Bogardi*, 1992; *Harmancioglu and Alpaslan*, 1992], various applications in forestry [*Huth et*

al., 2004; *Lasch et al.*, 2005], wastewater treatment alternatives [*Tecle et al.*, 1988; *Khalil et al.*, 2005], and computer-assisted tools for negotiation of water resources conflicts [*Thiessen and Loucks*, 1992].

[14] Traditionally, the MCDA technique has been used to make a selection or, at least, to aid the decision-making process with respect to the alternative to be adopted. In this study, the MCDA technique is employed in a novel way to address the issue of the utility of watershed modeling, the reliability of the developed simulation model and the efficiency of the monitoring program. Moreover, the MCDA technique is used in the traditional way to evaluate the various alternatives. Nonetheless, issues, such as the selection of the pertinent criteria, the priorities (weights) of the various criteria, and the MCDA methods used to solve the payoff matrix, are still relevant to the application under consideration.

[15] In the case study under consideration, the three alternatives (soil covers D1, D2, and D3) are the only "feasible" alternatives since they already have been imposed and adopted by the industry. The fourth soil cover (the one constructed earlier) has been evaluated by the industry as a dominated and unsatisfactory alternative and therefore is excluded from this study. Since the three alternatives are to be evaluated based on their hydrologic performance, the evaluation criteria should be hydrologic processes-related indicators. Accordingly, scoring (assessing the impacts of the three alternatives on the various criteria) can be achieved by analyzing the observed data or the simulation results. An important issue in MCDA is to be able to determine the relative weights or importance of a collection of criteria. Usually, such values are between 0.0 and 1.0 and they add up to 1.0. Issues related to how relative weights should be assigned and who should assign them have been thoroughly discussed by psychologists and social scientists [*Von Winterfeldt and Edwards*, 1986] and in MCDA literature [*Parlos*, 2000]. In this paper, the results do not rely on a single set of weights; analysis is initiated by assigning two different sets of weights to the criteria followed by weight sensitivity analysis to show the sensitivity of the outcome to changes in the priority structure of the problem.

6. Problem Formulation in a Multicriterion Decision Analysis Context

6.1. Evaluation Criteria

[16] Five key criteria have been selected in this study as the basic indicators with respect to the hydrologic performance of the different soil covers: (1) Soil moisture, assessed as the average depth (mm) of water available in the soil throughout the growing season (mid-May until mid-October). It is calculated as the arithmetic mean of the daily values. In the MCDA context, this is a "benefit criterion," which means the more the better [*Belton and Stewart*, 2003]. This criterion is broken down into two subcriteria: soil moisture in the peat layer and soil moisture in the till layer. The availability of soil moisture is the key indicator of the ability of vegetation to be sustained on the cover; (2) Soil suction pressure, assessed as the average suction pressure (K Pa) that existed in the soil layer throughout the growing season. Similar to soil moisture, it is calculated as the arithmetic mean of the daily values. The suction pressure is expressed with a positive sign, but this is considered a "cost criterion," i.e., the less the better. It is further divided into suction pressure within the peat layer and suction pressure within the till layer. Soil moisture alone cannot be sufficient to evaluate the availability of water for the plants because there are soil layers of various types and thicknesses. Higher levels of volumetric moisture content do not mean that more water is readily available for the roots since root activities tend to seek zones where the energy required to take up water is minimized (i.e., less suction pressure); (3) Shale percolation, assessed as the total depth of water (mm) percolating into the low-permeability shale layer. It is calculated as the sum of the daily values throughout the year. Since one of the reclamation objectives is to minimize deep percolation, this is a "cost criterion"; (4) Interflow, assessed as the total depth of water (mm) flowing as a lateral subsurface flow in the year. Similar to the shale percolation, it is calculated as the sum of the daily values throughout the year. This is perceived as an important indicator of the leaching ability of the soil cover to get rid of the excessive salts with the interflow; it is a "benefit criterion"; and (5) Surface runoff (overland flow), assessed as the total depth of water (mm) running off the surface of the cover without contributing to the soil moisture storage; in this study, surface runoff is considered a "cost criterion" due to the climatic nature of the area as a semiarid region. Surface runoff is lost and becomes unavailable for plants. This could be a disputable criterion because runoff could also beneficially contribute to surface water bodies that need to exist on any natural landscape.

[17] Even though the key criteria selected for this study are five, they were broken down into seven subcriteria, and thus weights need to be assigned to the seven subcriteria for conducting the MCDA. Eventually, the seven subcriteria address two ultimate objectives: (1) the availability of water for plants, which is represented by soil moisture, matric suction pressure, and surface runoff and (2) the quality (health) of the ecosystem, which is represented by the shale percolation and the interflow that leaches excessive salts. Some of the criteria might be slightly correlated on hourly or daily basis. For example, wetting an upper thin layer of the peat soil until it reaches the saturation point could trigger the interflow; however, the temporal scale of the indicators used in this study (e.g., total annual depth of interflow and the average seasonal depth of peat moisture) marginalizes or even eliminates such correlations. Even on daily basis, the relationship between soil moisture content and interflow is nonexistent until a certain threshold (soil saturation) is achieved. Taking into account various years of wet and dry climate, such correlation between the two criteria cannot be identified in a way that makes them interrelated criteria. This argument applies with various degrees to all seven criteria used in this study. It should be also noted that some correlation between various criteria does not disqualify them as evaluation criteria for the MCDM technique if each criterion retains its own individual importance [*Ko et al.*, 1992].

6.2. Impact Assessment (Scoring)

[18] Quantifying the performance of each soil cover with respect to the above mentioned criteria was conducted using the observed values and the SDW model simulation results.

Criteria	D1	D2	D3
Peat moisture (mm)	54	32	40
Till moisture (mm)	60	43	174
Peat suction (KPa)	159	95	39
Till suction (KPa)	73	59	27
Shale percolation (mm)	57	61	96
Interflow (mm)	0.26	0.3	1.24
Surface runoff (mm)	10.3	38	24.4

(a)

Criteria	D1	D2	D3
Peat moisture (mm)	53	29	37
Till moisture (mm)	62	40	175
Peat suction (KPa)	40	216	81
Till suction (KPa)	14	97	6
Shale percolation (mm)	44	55	39
Interflow (mm)	2.8	1.8	4.8
Surface runoff (mm)	18.6	34	13.3

(b)

Figure 2. Payoff matrix showing the values of various criteria for the three alternatives (a) using the observed values and (b) using the simulation results.

Details of the field instrumentation are given by *Boese* [2003], whereas the SDW model development, results, and performance assessment are given by *Elshorbagy et al.* [2005]. The available observed values (2000–2004) were used to populate five separate payoff matrices (one for each year) and the simulated results based on the available record of 61 years of meteorological data (1944–2004) were used to populate 61 separate payoff matrices. Examples of observed and simulated payoff matrices from year 2002 are shown in Figure 2.

6.3. Does the Choice of the Multicriterion Method Matter?

[19] *Hobbs et al.* [1992] posed and addressed the question of how various MCDA methods differ from one another. On the basis of their experiment, it was concluded that it is desirable to apply more than one method as a check. In this study, two MCDA methods, namely additive value functions (AVF) and ELECTRE II, were selected for the trade-off analysis. The AVF, which simply can be perceived as the normalized weighted summation of all criteria scores with respect to each alternative, is an easy to understand method. It is well perceived by water resources planners [*Hobbs et al.*, 1992] and has been used by other researchers in water resources literature [*Qureshi and Harrison*, 2001; *Huth et al.*, 2004; *Fassio et al.*, 2005]. ELECTRE II was selected because it is conceptually different from the AVF in the way it solves the payoff matrix and ranks different alternatives.

6.3.1. Additive Value Function

[20] Each score (s_{ij}) in the matrix is replaced with a value v_{ij} according to the following formula:

$$v_{ij} = \frac{s_{ij} - s_{i-}}{s_{i+} - s_{i-}} \tag{1}$$

where s_{ij} is the impact of an alternative (j) with respect to a criterion (i); s_{i-} is the "worst" score of the criterion (i) with respect to all alternatives, i.e., the worst score in the row (i) of the payoff matrix; and s_{i+} is the "best" score of the criterion (i) with respect to all alternatives, i.e., the best score in the row (i) of the payoff matrix. This way, all scores in the payoff matrix are scaled between values of 0.0 and 1.0. The AVF method is a simplified version of the multiattribute utility function [*Clemen*, 1996]. In the multiattribute utility function method, the risk attitude of the decision maker can be incorporated through a concave (risk averse) or a convex (risk seeker) utility curve. The way scores are normalized in the AVF method makes it a multiattribute utility function method for a decision maker that has a risk-neutral attitude (represented by a straight line instead of a utility curve) [*Clemen*, 1996]. An overall value index (V_j) for each alternative is estimated as follows:

$$V_j = \sum_{i=1}^{n} w_i v_{ij} \tag{2}$$

where w_i is the relative weight assigned to criterion (i) and n is the total number of criteria.

6.3.2. ELECTRE II

[21] This acronym stands for Elimination and Choice Translating Reality [*Roy*, 1996]. Variants of ELECTRE have been successfully used in water resources literature [*Tecle et al.*, 1988; *Hobbs et al.*, 1992; *Roy et al.*, 1992; *Raju and Duckstein*, 2004]. ELECTRE II is a variant of ELECTRE family that produces a ranking of alternatives rather than indicating the most preferred. It outranks based on alternatives that are preferred with respect to most of the criteria and that do not drastically fail with respect to any

Table 1. Statistics of the Historical Precipitation and the SDW Model Simulation Results[a]

	D1			D2			D3		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum
Precipitation	247	442	676	247	442	676	247	442	676
Peat moisture	47	61	81	28	53	62	35	53	83
Till moisture	48	99	134	33	88	104	166	250	357
Peat suction	2	27	496	1	73	1180	2	40	109
Till suction	1	9	452	1	11	266	1	2	13
Shale percolation	0	21	92	0	19	49	0	18	68
Interflow	0	11	76	0	88	150	0	1	9
Surface runoff	0	5	23	0	6	28	0	6	26

[a]Units are in mm except for suction, which is in KPa. All fractions are rounded up to the nearest integer.

one or more criteria. The first attribute is expressed by the "concordance" index and the second by the "discordance" index. Alternative A outranks alternative B if both concordance and discordance indices are satisfied [*Belton and Stewart*, 2003]. The concordance index C(A,B) measures the strength of support in the information given for the hypothesis that A is at least as good as B. The discordance index D(A,B) measures the strength of evidence against this hypothesis. The concordance index is calculated as

$$C(A,B) = \frac{w^+ + w^=}{w^+ + w^= + w^-} \tag{3}$$

where w^+ is the sum of the weights of all criteria where A is better than B; w^- is the opposite case, i.e., the sum of the weights of the criteria where B is better than A; and $w^=$ is the indifferent cases. Using the same concept of value functions expressed in Equation 1, a discordance index can be calculated as follows:

$$D(A,B) = Max(v_{iB} - v_{iA}) \tag{4}$$

where v_{iB} is the value function of the impact of alternative B with respect to criterion (i) and v_{iA} is the value function of the impact of alternative A with respect to criterion (i). In order for alternative A to outrank alternative B, C(A,B) has to be greater than D(A,B), and both of C(A,B) and D(A,B) should be higher than a preset threshold value p and lower than a preset threshold value q, respectively. Moreover, w^+ has to be greater than w^-. For more details on the ELECTRE II method and its use of different pairs of concordance and discordance thresholds, and how to rank the alternatives, readers should refer to *Belton and Stewart* [2003] and *Parlos* [2000].

7. Results and Analysis

7.1. Trade-off Analysis Using the Additive Value Functions

[22] The trade-off analysis among the three alternatives was conducted using the additive value functions (AVF) with two different sets of weights; for the first analysis, each criterion is given equal weight (i.e., 0.143 each) and for the second analysis, the relative weights of peat moisture and interflow were twice the relative weights of each of the rest of the criteria, which were kept at equal relative weights. The relative weights of peat moisture and interflow were 0.222, whereas each of the remaining criteria carried a relative weight of 0.111. Peat moisture and interflow are key indicators of the ability of the soil cover to provide sufficient moisture for evapotranspiration and to prevent the harmful accumulation of excessive salts in the root zone. The AVF ranks the alternatives based on overall scores that indicate the relative distances or closeness of the different alternatives. The meteorological record of 61 years (1944–2004) was input to the SDW model and the results were used to populate 61 payoff matrices. Some of the statistics of the historical annual precipitation and the model results are provided in Table 1.

[23] The results of the ranking analysis are shown in Table 2. For brevity and readability purposes, only the probability of occupying the first position is shown in Table 2. Changing the relative weights of the evaluation criteria does not have a significant effect on the results. D3 is ranked first in at least 65% of the years based on the climatic conditions. However, when five separate payoff matrices were populated using observed values (2000–2004) rather than the simulated ones, D3 occupied the first position in all 5 years. Relying on the observed payoff matrices (as opposed to the long simulated series) may lead to the impression that D3 is the superior alternative all the time. This is apparently a positive contribution of the simulation model toward the comprehensive understanding of the problem at hand. From this analysis, it is evident that cover D3 can be dominated 35% of the time (in case of equal criteria weights) under varying climatic conditions. The balance of 65% is called the probability of making the right decision (PMRD) since a logical decision maker would choose D3, which is dominant, on average 65% of the time. In other words, if the decision maker chooses to adopt alternative D1 for economical reasons (D1 is thinner than D3), the PMRD will be 25% (presuming equal criteria weights). The alternative that is ranked 1 in a year is

Table 2. Probability of Winning the First Position by Various Alternatives (AVF Method)

	Simulated Payoff Matrix	
Alternative	Equal Weights	Differential Weights
D3	65%	69%
D1	25%	20%
D2	10%	11%

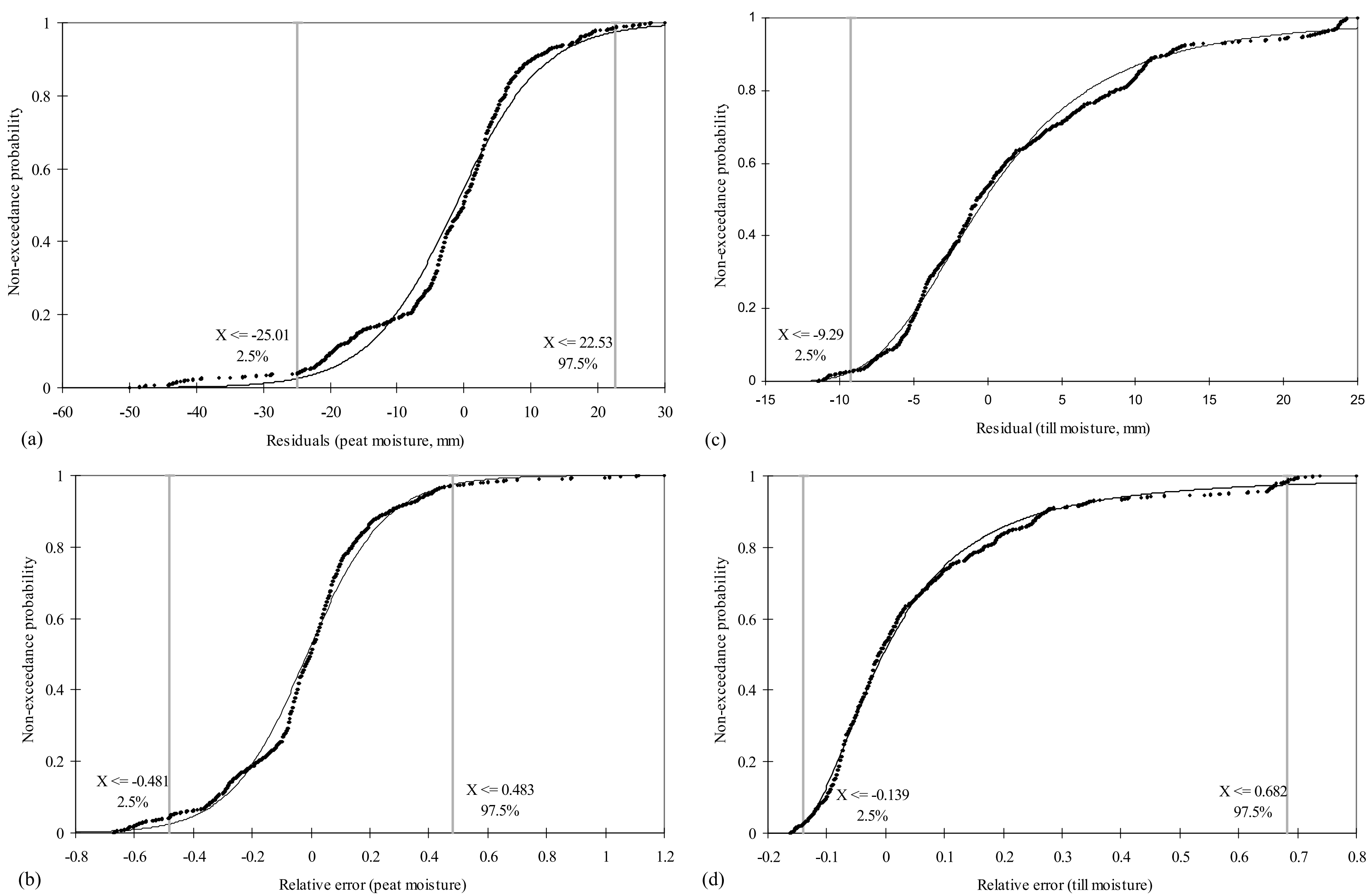

Figure 3. Probability distributions of model residuals: (a) peat moisture residuals, (b) peat moisture relative error, (c) till moisture residuals, and (d) till moisture relative error.

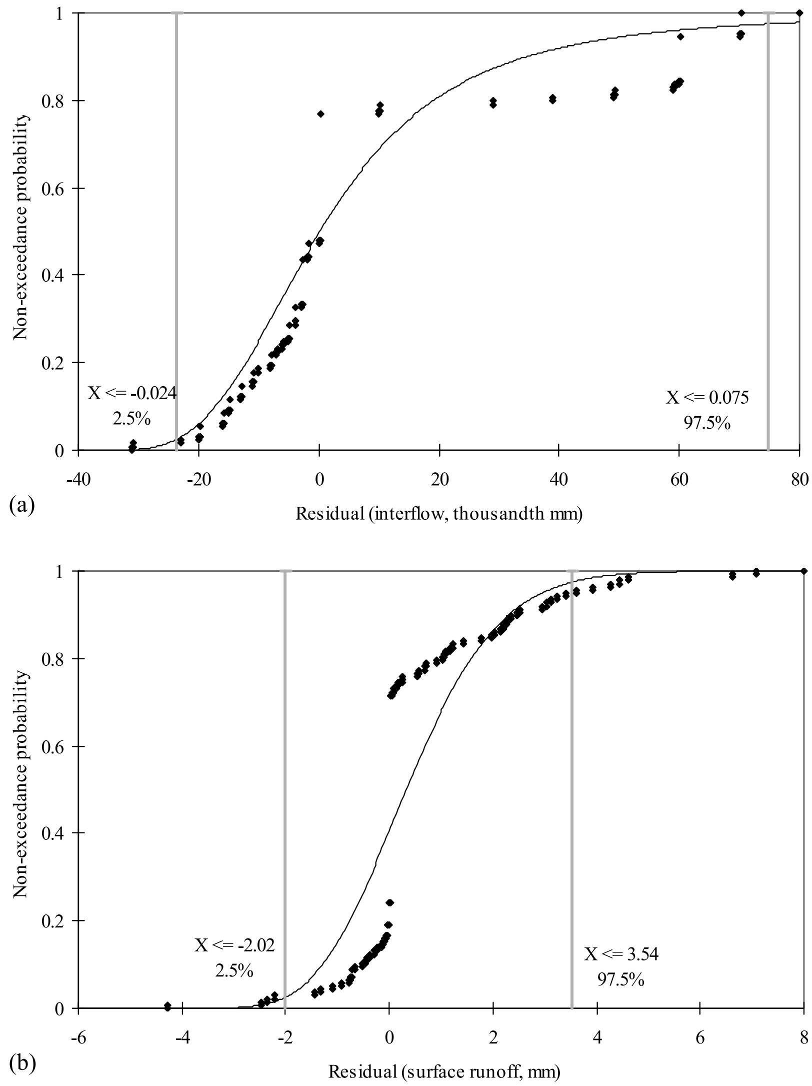

Figure 4. Probability distributions of model residuals: (a) interflow residuals and (b) surface runoff residuals.

assigned a score of 1; otherwise it takes a score of 0. The scores are summed over the analysis period of 61 years. The resulting sum is divided by 61 to obtain the PMRD. This way the PMRD does not take into account the degree of dominance of a certain alternative. Certainly, a weight can be associated with the rank to reflect degree of dominance in a certain year and in this case, one could end up with a weighted PMRD. In this paper, the degree of dominance is ignored.

7.2. Analysis of Score Uncertainty

[24] Uncertainties about the simulation results are inevitable. Such uncertainties could be due to the simulation model parameters, structure, or natural variability [*Bastidas et al.*, 2003]. The uncertainty due to natural variability (e.g., climatic condition) has been accounted for in the calculation of the PMRD index by simulating the various alternatives over a sufficiently long period of time (61 years in this study). Since the SDW model is a custom-built site-specific model that is calibrated and validated for the site under consideration in this study [*Elshorbagy et al.*, 2005], the remainder portion of uncertainty will stem from the predictive uncertainty of the model. Such uncertainty is demonstrated by the fact that the model predictions or simulations are imperfect; i.e., there are residuals representing the differences between observed and simulated data. One of the possible ways to quantify the score uncertainty in the simulated payoff matrices is through accounting for the model residuals [*Borsuk et al.*, 2002]. The model residuals (difference between observed and simulated output) are considered as a representation of the total predictive uncertainty.

[25] The simulated values (2000–2004) were contrasted against the observed values for the same period, and the model residuals (simulated values minus observed values) were calculated. Various probability distributions were tested using @ RISK Software [*Palisade Corporation*, 2004] and the distributions that best fit into the residuals were adopted (Figures 3 and 4) based on the value of chi-square [*Hines et al.*, 2003] as well as on the visual inspection. The best fit distributions were found to be logistic (−1.24, 6.49), log logistic (−8.24, 8.23, 62.61),

Table 3. Typical Ranking Probabilities of Various Alternatives Based on Score Uncertainty Using the Results of the SDW Model (AVF Method)

Alternative	First Position	Second Position	Third Position
D3	86%	14%	0%
D1	14%	86%	0%
D2	0%	0%	100%

log logistic (−14, 13.83, 3.41), log logistic (−0.17, 0.17, 2.26), log logistic (−0.03, 0.04, 3.21), and Gamma (26.03, 0.28) for peat moisture residuals, peat moisture relative error, till moisture residuals, till moisture relative error, interflow residuals, and surface runoff residuals, respectively. The relative error is dimensionless, whereas other variables have units of mm. The parameters (given in parentheses) of the Logistic distribution are the location and scale parameters; for Log Logistic distribution are the location, scale, and shape parameters, respectively; and for the Gamma distribution are the shape and the scale parameters.

[26] For sampling purposes and to perturb the payoff matrix scores, either the residuals directly or the relative error (RE) as defined by equation 5 can be used.

$$\mathrm{RE} = \frac{X_s - X_o}{X_o} \tag{5}$$

where X_s is the simulated value, X_o is the measured value. Using the RE values is more realistic than using the residuals because sampling from the residuals means that an extreme value of error (e.g., large error from the tail of the distribution) can be picked up to perturb the opposite extreme (e.g., low value of peat moisture). Such an error may have not been encountered by the model. Monte Carlo simulation [*Hines et al.*, 2003] was used first to sample values from the relative errors to randomly perturb the scores of the simulated payoff matrices, then to redo the MCDA and rank the alternatives. A set of 5000 different combinations was tested during the Monte Carlo simulation for every payoff matrix (i.e., every year). During each run, the alternatives were ranked. An example of typical results for 1 year is provided in Table 3. Even in that single year, uncertainty is evident about the dominance of D3: there is 86% chance that alternative D3 occupies the first rank, while there is 14% chance that alternative D1 outranks other alternatives. The 86% represents the confidence in the conclusion that D3 is the best alternative in that year. It should be noted that in the case of interflow and surface runoff, it was not possible to calculate relative errors due to the existence of zeros in both observed and simulated daily values. Therefore Monte Carlo simulation was used to sample from the residual distributions directly after multiplying by factors of 40 and 25 for interflow and surface runoff, respectively. The multiplication factors are needed because the residuals represent error in daily values whereas the interflow and surface runoff criteria used in the payoff matrices are total annual values. From the simulations based on the historical records, it was found that the average number of days where interflow and surface runoffs were nonzeros were approximately 40 and 25 days, respectively. This was the only possible way to quantify the uncertainty about the simulated values of interflow and surface runoff. Some unrealistic daily readings and serious doubts about the continuous in situ soil suction measurements have been raised [*Boese*, 2003], which make adopting the probability distribution approach to quantify the score uncertainty about peat and till suction impossible. Since the SDW model uses the developed soil water characteristic curves (SWCC) [*Boese*, 2003; *Elshorbagy et al.*, 2005] to estimate soil suction pressure based on simulated soil moisture, the same percentages of moisture uncertainty in the peat and till layers were used to represent the score uncertainty about suction pressure in both layers, respectively.

[27] Applying this analysis to climatic data spanning 61 years and provided that the PMRD (i.e., D3 occupies the first rank) is 65% (using equal criteria weights), it is possible to quantify the uncertainty about the PMRD using equation (6).

$$R_{1i} = \frac{\text{Number of simulations that the alternative is ranked \# 1 in a year}}{\text{Total number of simulations for one year}} \times 100$$

$$\mathrm{CPMRD} = \frac{1}{N}\sum_{i=1}^{N} R_{1i} \tag{6}$$

To clarify, in individual years where D3 is ranked as the best alternative, Monte Carlo simulations were executed to perturb the scores of one year 5000 times and rank the alternatives. R_{1i}, which is the confidence in the top rank in a given year (i), is, for example, 86%. The average value of R_{1i} over all years (N) where the top rank alternative was the best was calculated and called the confidence in the PMRD (CPMRD). This value is 84% for D3 in case of equal weights. It should be noted that the PMRD reflects the effect of climatic variability on the different soil covers and cannot be increased unless a new soil cover that combines some of the characteristics of D3, D1, and D2 is designed and tested. On the other hand, the CPMRD reflects the effect of the predictive uncertainty of the SDW model and can be increased if the predictive accuracy of the model is improved. For example, a hypothetical situation was tested in which the level of the predictive uncertainty of the model was decreased by half. The resulting CPMRD increased from 84% to 95%. Theoretically, a perfect simulation model would result in a CPMRD value of 100%. If the 84% level is unsatisfactory, then a decision should be made to increase the CPMRD, which means decreasing the level of score uncertainty. Decreasing the level of score uncertainty can be translated into improving the model structure or inputs, or even employing a different model to reduce the predictive uncertainty (i.e., the mean relative error, MRE).

[28] The average value of CPMRD calculated to be 84% is the result of annual values of R_{1i} ranging from 53% to 100%. This might indicate that in some years, the best alternative significantly dominates others such that some uncertainty about the scores does not reverse the rank of the alternatives. In other years, the trade-offs govern the relationship, and thus any change in the criteria scores could reverse the rank. The first case is represented by a higher

Table 4. Ranking Probabilities of Various Alternatives Based on Combined Score and Weight Uncertainty Using the Results of the SDW Model (AVF Method)

Alternative	First Position	Second Position	Third Position
D3	41%	30%	29%
D1	32%	34%	34%
D2	27%	36%	37%

value of R_{1i}, whereas the latter case is represented by a lower value of the R_{1i}.

7.3. Analysis of Priority (Weight) Uncertainty

[29] Assigning relative weights to the various evaluation criteria is an issue that raises serious controversy in the MCDA community. In this study, the purpose of the analysis is to present the framework and the methodology for evaluating the impact of using watershed modeling to assess various reclamation covers. Therefore what really matters is consistency; i.e., the same weights are applied to all payoff matrices. An uncertainty level of 100% is assumed for all criteria weights based on equal relative weights for all criteria. This means that the importance of each criterion may range from nil to twice its original value of 0.143. On the basis of the Monte Carlo simulation, conducted with the simulation payoff matrix, the PMRD is 68%, with a 32% chance that either alternative D1 or alternative D2 occupies the first position. From this analysis, one may conclude that being uncertain about the relative weights of various criteria, provided that scores are certain, does not significantly affect the final decision. In the previous case of equal criteria weights, the PMRD was 65%.

[30] Considering the score uncertainty and weight uncertainty individually may or may not have a notable effect on the PMRD and the CPMRD indices. However, when both uncertainties are combined, the effect on the decision indices could increase significantly. Applying the previously set levels of combined score and weight uncertainties to the 61 simulated payoff matrices results in the values summarized in Table 4. The PMRD index could decrease to 41% in the case of uncertain weights and scores. An important observation could be made based on this analysis; that the deterioration of the PMRD is due to the combined effects of uncertainty. This means that watershed managers should first agree on the priority structure of different criteria. This could increase the PMRD significantly by restricting the uncertainty to score uncertainty only.

7.4. Analysis of Method Uncertainty

[31] Results obtained by any MCDA method can be influenced by the underlying algorithm, the assumptions of that method, and the approach it takes to solve the payoff matrix. Accordingly, the results can be inflicted by another type of uncertainty called method uncertainty. Method uncertainty can be addressed by applying more than one method. The previously discussed trade-off analysis is repeated using the ELECTRE II method. As explained in a previous section and inferred from equations (3) and (4), this method conducts a pairwise comparison between alternatives. First, the extent to which the alternative is preferred above others based on the weights is considered, then the extent to which it is dominated by another alternative based on the scores [*Janssen et al.*, 2003] is taken into account. One can observe that ELECTRE II places more emphasis on the relative weights compared to the scores of the evaluation criteria.

Table 5. Probability of Winning the First Position by Various Alternatives (ELECTRE II Method)

Alternative	Simulated Payoff Matrix: Equal Weights	Simulated Payoff Matrix: Differential Weights
D3	70%	71%
D1	20%	20%
D2	10%	9%

Table 6. Ranking Probabilities of Various Alternatives Based on Score Uncertainty Using the Results of the SDW Model (ELECTRE II Method)

Alternative	First Position	Second Position	Third Position
D3	50%	40%	10%
D1	40%	45%	15%
D2	10%	15%	75%

[32] Using the ELECTRE II method, the ranking analysis conducted using the AVF method was repeated and summarized in Table 5. Alternative D3 has slightly higher values of PMRD index (70% and 71% using equal and differential criteria weights, respectively). The CPMRD index is 86% in both scenarios of criteria weights using the ELECTRE II method. Comparison of results based on Tables 2 and 5 indicates that the problem under consideration is not highly sensitive to the MCDA method adopted. It can be concluded that alternative D3 considerably dominated the other two alternatives.

7.5. Analysis of Score and Weight Uncertainties Using ELECTRE II Method

[33] Both scores and weights were randomly perturbed together using the same uncertainty levels as before and the analysis conducted using the AVF was repeated using ELECTRE II. The results are summarized in Table 6. A set of 5000 different combinations was tested using Monte Carlo simulation as before. The PMRD using the simulation model turns out to be 50% (compared to 41% using the AVF method). There is a 50% chance that either D1 or D2 outranks alternative D3. This confirms the previously drawn conclusion regarding deciding on criteria weights first to restrict uncertainty to the inevitable score uncertainty. The quantitative nature of the criteria considered in this study makes it easy to adopt an MCDA method based on a value function, and therefore the AVF is a potential candidate MCDA method for this study as well as other similar studies.

7.6. Score Interval Analysis

[34] Although the two MCDA methods adopted in this study result in similar preference structures with regard to the three alternatives, it is suggested that the AVF method be adopted for applications similar to the one presented in this study because of the transparency of the method; the final overall scores assigned to the alternatives by the AVF method provide an indication of how close or far two

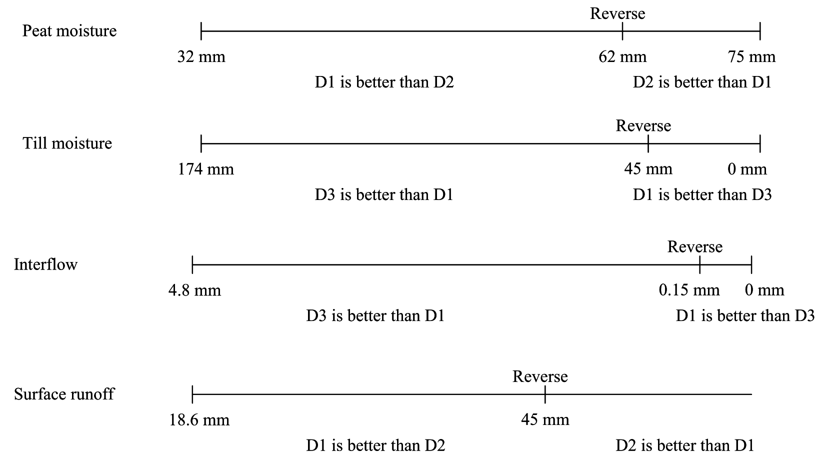

Figure 5. Score intervals for different criteria using the measured values (AVF method).

alternatives are with respect to each other. The algorithm is also simple and easy to understand. Accordingly, further analysis and discussions will be based only on the results of the AVF method.

[35] Score interval analysis is another possible type of MCDA analysis that helps determine the intervals within which the rank order of two alternatives is insensitive to changes in score. This may provide a deeper insight into the robustness of a certain alternative with respect to the scores populated in the payoff matrix. If the ranking is found to be sensitive to the score of a certain criterion, a feedback to the instrumentation program and/or modeling may be deemed necessary to reduce the uncertainty associated with such a criterion score. The results of score interval analysis performed on the observed payoff matrix of year 2002, as an example, are presented in Figure 5. D1 and D2 could exchange their second and third ranking positions if the peat moisture level in cover D2 increases from 32 mm to 62 mm, provided that all other scores are kept unchanged. This means that the average value of the observed moisture content in the 15 cm peat layer of D2 should become significantly higher than that in the 20 cm peat layer of D1. This could be improbable, and therefore the final ranking may be considered insensitive to this score. Similarly, with regard to the till moisture, the till moisture content in cover D3 has to decrease to a level much lower than the wilting point (around 90 mm) in order for a rank reverse to occur. However, in the case of the surface runoff criterion, a rank reverse could occur (i.e., D2 is better than D1) if the surface runoff from the cover D1 increases from the value of 18.6 to 45 mm. Since most of the runoff occurs in the spring due to snowmelt [*Elshorbagy et al.*, 2005] one cannot exclude the possibility of this scenario. Nonetheless, this does not affect the first ranking position of cover D3. On the basis of this type of analysis, and if repeated over all years where observations are available, a general conclusion could be drawn with regard to the instrumentation and monitoring program; i.e., whether the current practices are satisfactory or not. In the case study considered in this paper, inherent inaccuracies and uncertainties are not

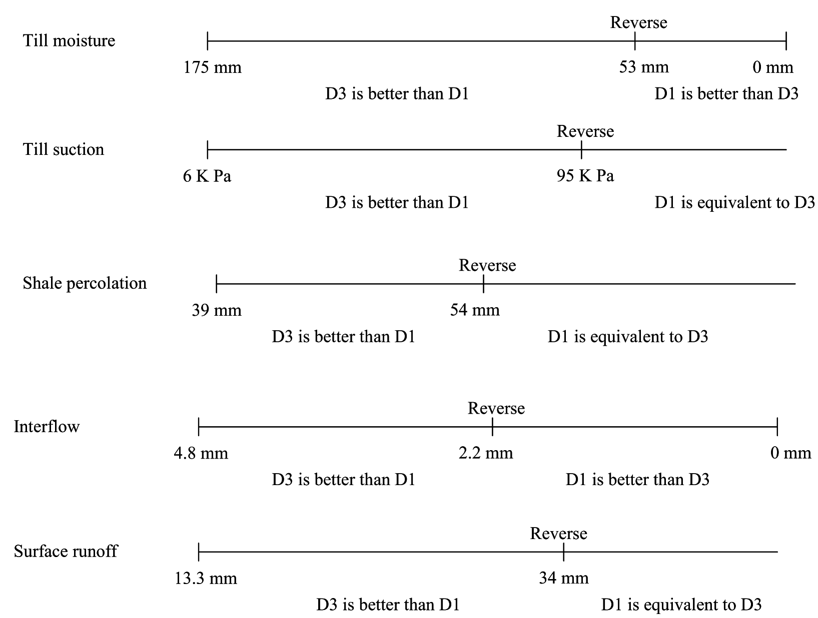

Figure 6. Score intervals for different criteria using the results of the SDW model (AVF method).

expected to have significant effects on the final rankings of the various alternatives. However, attention should be given to the spring runoff in order to have an accurate account of the event.

[36] The score interval analysis is repeated with the simulation payoff matrix of the same year (2002) to assess the efficacy of the SDW model used for simulating the watersheds. The results are presented in Figure 6. No rank reversal is observed with regard to the scores of peat moisture and peat suction (not shown in Figure 6). The cover D1 could be preferred to cover D3 if the soil matric suction pressure in the till layer of the cover D3 increases from 6 K Pa to 95 K Pa. Even though such an increase seems extremely high, it cannot be excluded given the significant discrepancies between observed and simulated suction pressure values. The SDW model relies on the developed soil water characteristic curve (SWCC) to estimate soil suction pressure based on simulated soil moisture content. This score interval analysis indicates that it is worth revisiting the SWCC to ensure higher accuracy and less uncertainty about the moisture-suction relationship. Similarly, cover D1 can share the first ranking position with cover D3 if the amount of water percolating into the shale layer increases from 39 mm to 54 mm. This could be critical since *Elshorbagy et al.* [2005] have reported a low level of confidence with regard to the simulated shale percolation. Moreover, the observed values indicate that shale percolation in cover D3 could be higher than that in D1. Certainly, refinement of the simulation model will be highly desirable in order to obtain more accurate and less uncertain results with regard to the shale percolation criterion. The rank reversal with regard to the remaining criteria is less significant since this would happen only in the following unlikely cases: (1) the moisture content in the till layer of D3 decreases to a level much below the wilting point or (2) the interflow in D3 becomes less than that in D1 in the analyzed year.

8. Discussion

[37] The analysis presented in this paper helps identify the best alternative for reclamation strategy to be adopted by the oil sands industry. It also helps identify areas where the monitoring program needs to be improved and refocused. One main such area is the measuring of surface runoff, especially during spring melt. The weirs constructed on the site to measure the runoff should be properly maintained to avoid ice blockage and flow bypassing the weirs. With regard to the simulation payoff matrix, the MCDA analysis (score uncertainty and score interval) shows that the SDW model results in reliable outputs (a range of 84–86% confidence that the model leads to the right decision). However, to increase the CPMRD index to above 86%, the predictive uncertainties of the model should be properly quantified and kept below the level used in this study for score uncertainty analysis. As an example, when the score uncertainty level used in this study is halved with regard to all criteria, the CPMRD increases to 95%. Special attention should be given to the development of a reliable SWCC to improve the estimation of soil matric suction and the model component that estimates the shale percolation. This conclusion may not be easily obtained by assessing the model performance using traditional error measures such as RMSE and MRE. However, one could say that the SDW model results in reliable results with respect to the peat and the till moisture contents. If the locations and frequency of measurements of such variables have to be reduced to minimize the cost of the instrumentation program, the model can satisfactorily substitute for these missing variables.

[38] In this study, the MCDA technique has been employed in a novel way to address issues related to the efficiency of measuring (monitoring) and the efficacy of watershed modeling. This approach can be extended to evaluate the reliability of various watershed models and the dependability of these models to help make the right decision with regard to a specific case study. For example, various simulation payoff matrices, each populated with scores based on different watershed models (e.g., SDW model, SLURP, HSPF) can be constructed. Apparently, the CPMRD will vary based on the predictive accuracy of the adopted model. Such an index becomes 100% in the hypothetical case of adopting a perfect model. The PMRD index may or may not change when different models are tested. The model that results in the highest value of CPMRD should be adopted. Such an exercise can prove valuable for an industry that needs to choose among various watershed models to use for assessing the reclamation strategies.

[39] It is important to remember that the ranking of various soil covers in this study was based only on hydrologic performance and the results should be taken only as an example of the utility of the MCDA technique. However, a complete evaluation of the various soil covers should include the cost of constructing each cover as one of the evaluation criteria. The monetary cost of laying one cubic meter of soil can be used as the attribute for this criterion. The cost of landscaping is $4.0/m^3, which makes the values of this criterion for different soil covers $2.0, $1.4, and $4.0 per m^2 of landscape for D1, D2, and D3, respectively. The inclusion of the cost criterion in the payoff matrix, assuming equal weights for all criteria, did not affect the final ranking (i.e., D3 is ranked first, followed by D2, then D1). However, the PMRD dropped from 65% to 53%. There is a 47% chance that either D1 or D2 becomes the best option. This is not surprising because the lower cost of D1 and D2 compared to that of D3 provides D1 and D2 with greater chances of being ranked first. The CPMRD remained unchanged.

[40] A final note on the quantification of the predictive uncertainty of the SDW model should be emphasized. In this study, fitting probability distributions to the residuals to quantify the predictive uncertainty was adopted. In the case where model residuals show signs of heteroscedasticity due to state-dependent error (e.g., larger residuals with higher values of the variable under consideration), the Bayesian approach can be adopted to quantify such an uncertainty [*Freer et al.*, 1996]. If the SDW model is to be transferred to other sites, the predictive uncertainty should be more comprehensively quantified by considering uncertainty about the model parameters.

9. Conclusions

[41] The multicriterion decision analysis (MCDA) technique is a useful tool that can be employed outside the

traditional usage of ranking different alternatives with respect to a definite set of evaluation criteria. In this paper, the MCDA has been used to combine watershed simulation with subsequent multicriterion evaluation, and to evaluate various reclamation strategies with the purpose of assessing the value and efficacy of watershed monitoring and modeling exercises. Using the MCDA to evaluate the payoff matrices that were populated with scores based on field measurements revealed that the final ranking could illusively indicate that one alternative is completely dominant. Using the SDW model to simulate a long period of time using historical data helped provide a complete picture of the relative performance of the various alternatives. This has been represented by an index developed in this study and named the probability of making the right decision (PMRD), an index that reflects the effect of the inevitable natural variability (i.e., climatic conditions) on the final decision. Another index, one representing the confidence in the previous index (CPMRD), was developed to represent the reliability of the adopted watershed model to estimate the PMRD index. The developed CPMRD index could be used as a decision variable to decide on the level of refinement needed for a particular watershed model or whether another model should be considered. The ways in which MCDA techniques were used in this paper could be extended to other applications that could be useful to both industry and the scientific community.

[42] **Acknowledgments.** The author acknowledges the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada through the Discovery Grant and the CRD programs. The support of Syncrude Canada Limited is greatly appreciated. The insightful reviews of the anonymous reviewers and associate editor helped reshape and improve this paper.

References

Barbour, S. L., C. Boese, and B. Stolte (2001), Water balance for reclamation covers on oilsands mining overburden piles, in *Canadian Geotechnical Conference*, pp. 313–319, Can. Geotech. Soc., Alliston, Ont., Canada.

Bastidas, L. A., H. V. Gupta, K. Hsu, and S. Sorooshian (2003), Parameter, structure, and model performance evaluation for land-surface schemes, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 229–237, AGU, Washington, D. C.

Belton, V., and T. J. Stewart (2003), *Multiple Criteria Decision Analysis: An Integrated Approach*, Springer, New York.

Boese, K. (2003), The design and installation of a field instrumentation program for the evaluation of soil-atmosphere water fluxes in a vegetated cover over saline/sodic shale overburden, M.Sc. thesis, Univ. of Saskatchewan, Saskatoon, Sask., Canada.

Bogardi, J. J., and L. Duckstein (1992), Interactive multiobjective analysis embedding the decision maker's implicit preference function, *Water Resour. Bull.*, *28*, 75–88.

Borsuk, M. E., C. A. Stow, and K. H. Reckhow (2002), Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development, *Environ. Sci. Technol.*, *36*(10), 2109–2115.

Clemen, R. T. (1996), *Making Hard Decisions: An Introduction to Decision Analysis*, Duxbury, Belmont, Calif.

Elshorbagy, A., A. Jutla, S. L. Barbour, and J. Kells (2005), System dynamics approach to assess the sustainability of reclamation of disturbed watersheds, *Can. J. Civ. Eng.*, *32*, 144–158.

Falkenmark, M. (1997), Society's interaction with the water cycle: A conceptual framework for a more holistic approach, *Hydrol. Sci. J.*, *42*, 451–466.

Fassio, A., C. Giupponi, R. Hiederer, and C. Simota (2005), A decision support tool for simulating the effects of alternative policies affecting water resources: An application at the European scale, *J. Hydrol.*, *304*, 462–476.

Freer, J., K. Beven, and B. Ambroise (1996), Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Res. Res.*, *32*, 2161–2173.

Goicoechea, A., E. Z. Stakhiv, and F. Li (1992), Experimental evaluation of multiple criteria decision models for application to water resources planning, *Water Resour. Bull.*, *28*, 89–102.

Harboe, R. (1992), Multiobjective decision making techniques for reservoir operation, *Water Resour. Bull.*, *28*, 103–110.

Harmancioglu, N. B., and N. Alpaslan (1992), Water quality monitoring network design: A problem of multi-objective decision making, *Water Resour. Bull.*, *28*, 179–192.

Hines, W. W., D. C. Montgomery, D. M. Goldsman, and C. M. Borror (2003), *Probability and Statistics in Engineering*, John Wiley, Hoboken, N. J.

Hobbs, B. F., V. Chankong, W. Hamadeh, and E. Z. Stakhiv (1992), Does choice of multicriteria method matter? An experiment in water resources planning, *Water Resour. Res.*, *28*, 1767–1779.

Huth, A., M. Drechsler, and P. Kohler (2004), Multicriteria evaluation of simulated logging scenarios in a tropical rain forest, *J. Environ. Manage.*, *71*, 321–333.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall- runoff model?, *Water Resour. Res.*, *29*, 2637–2649.

Janssen, R., M. van Herwijnen, and E. Beinat (2003), DEFINITE: Case studies and user manual, report, Inst. voor Milieu Vraagstukken, Amsterdam.

Khalil, W. A., A. Shanableh, P. Rigby, and S. Kokot (2005), Selection of hydrothermal pre-treatment conditions of waste sludge destruction using multicriteria decision-making, *J. Environ. Manage.*, *75*, 53–64.

Ko, S., D. G. Fontane, and J. W. Labadie (1992), Multiobjective optimization of reservoir system operation, *Water Resour. Bull.*, *28*, 111–127.

Lasch, P., F. Badeck, F. Suckow, M. Lindner, and P. Mohr (2005), Model-based analysis of management alternatives at stand and regional level in Brandenburg (Germany), *For. Ecol. Manage.*, *207*, 59–74.

Meier, D., and S. L. Barbour (2002), Monitoring of cover and watershed performance for soil covers placed over saline-sodic shale overburden from oilsands mining, paper presented at National Meeting of the American Society of Mining and Reclamation, Lexington, Ky.

Palisade Corporation (2004), Guide to using @RISK: Advanced risk analysis for spreadsheets, report, Ithaca, N. Y.

Parlos, P. M. (2000), *Multi-criteria Decision Making Methods: A Comparative Study*, Springer, New York.

Qureshi, M. E., and S. R. Harrison (2001), A decision support process to compare riparian revegetation options in Scheu Creek catchment in north Queensland, *J. Environ. Manage.*, *62*, 101–112.

Raju, K. S., and L. Duckstein (2004), Integrated application of cluster and multicriterion analysis for ranking water resources planning strategies: A case study in Spain, *J. Hydroinformatics*, *6*, 295–307.

Roy, B. (1996), *Multicriteria Methodology for Decision Aiding*, Springer, New York.

Roy, B., R. Slowinski, and W. Treichel (1992), Multicriteria programming of water supply systems for rural areas, *Water Resour. Bull.*, *28*, 13–31.

Schaake, J. C. (2002), Introduction, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 1–7, AGU, Washington, D. C.

Singh, V. P. (Ed.) (1995), *Computer Models of Watershed Hydrology*, Water Resour. Publ., Highlands Ranch, Colo.

Singh, V. P., and D. K. Frevert (Eds.) (2002a), *Mathematical Models of Large Watershed Hydrology*, Water Resour. Publ., Highlands Ranch, Colo.

Singh, V. P., and D. K. Frevert (Eds.) (2002b), *Mathematical Models of Small Watershed Hydrology and Applications*, Water Resour. Publ., Highlands Ranch, Colo.

Srdjevic, B., Y. D. P. Medeiros, and A. S. Faria (2004), An objective multicriteria evaluation of water management scenarios, *Water Resour. Manage.*, *18*, 35–54.

Stansbury, T., W. Woldt, I. Bogardi, and A. Bleed (1991), Decision support system for water transfer evaluation, *Water Resour. Res.*, *27*, 443–451.

Tecle, A. (1992), Selecting a multicriterion decision making technique for watershed resources management, *Water Resour. Bull.*, *28*, 129–140.

Tecle, A., M. Fogel, and L. Duckstein (1988), Multicriterion selection of wastewater management alternatives, *J. Water Resour. Plann. Manage.*, *114*, 383–398.

Thiessen, E. M., and D. P. Loucks (1992), Computer assisted negotiation of multiobjective water resources conflicts, *Water Resour. Bull.*, *28*, 163–177.
Von Winterfeldt, D., and W. Edwards (1986), *Decision Analysis and Behavioural Research*, Cambridge Univ. Press, New York.
Woldt, W., and I. Bogardi (1992), Ground water monitoring network design using multiple criteria decision making and geostatistics, *Water Resour. Bull.*, *28*, 45–62.
Wooldridge, S. A., J. D. Kalma, and J. P. Walker (2003), Importance of soil moisture measurements for inferring parameters in hydrologic models of low-yielding ephemeral catchments, *Environ. Modell. Software*, *18*, 35–48.

A. Elshorbagy, CANSIM, Department of Civil and Geological Engineering, University of Saskatchewan, Saskatoon, SK, Canada S7N 5A9. (amin.elshorbagy@usask.ca)