Use of Multi-Criterial Involvement Processes (MIP) to Enhance Transparency and Stakeholder Participation at Bergen Harbour, Norway

† ‡ Magnus Sparrevik, § David N. Barton, †Amy Oen, ‖Nagothu Udaya Sehkar and #Igor Linkov

†*Norwegian Geotechnical Institute, PO Box 3930 Ullevål Stadion, NO-0806 Oslo, Norway, magnus.sparrevik@ngi.no,

‡ Department of Industrial Economics and Technology Management. Norwegian University of Technology, 7491 Trondheim, Norway

§ Norwegian Institute for Nature Research, Gaustadalléen 21, NO-0349 Oslo, Norway

Bioforsk Norwegian Institute for Agricultural and Environmental Research Fr. A. Dahlsvei 20, N-1432 Ås Norway

‖ U.S. Army Engineer Research and Development Center, 696 Virginia Rd, Concord, MA 01742, United States.

ABSTRACT

Use of participatory stakeholder engagement processes could be important to reduce the risk of potential conflicts in managing contaminated sites. Most stakeholder engagement techniques are qualitative in nature and require experienced facilitators. This study proposes a multi-criterial involvement process (MIP) to enhance transparency and stakeholder participation and applies it to a contaminated sediment management case study for Bergen Harbor, Norway. The suggested MIP builds on the quantitative principles of MCDA, and also incorporates group interaction and learning through qualitative participatory methods. Three different advisory groups consisting of local residents, local stakeholders and non-resident sediment experts were invited to participate in a stakeholder engagement process to provide consensual comparative advice on sediment remediation alternatives. For stakeholders or residents to be able to embrace a complex decision such as selection of remediation alternatives, the involvement process with lateral learning, combined with MCDA giving structure, robustness and transparent documentation was preferable. Additionally, MIP results in consistent ranking of remediation alternatives across residents, stakeholder and experts, relative to individual intuitive ranking without MIP.

KEYWORDS: Stakeholder involvement, contaminated sediment management, multi-criteria decision analysis, citizens' jury

INTRODUCTION

Emerging environmental challenges coupled with increased stakeholder awareness and concerns call for more effective stakeholder involvement processes for environmental management. A structured stakeholder involvement process could help in overcoming disagreements and result in better management alternatives (Slob et al., 2008). Examples of qualitative involvement processes include focus groups with facilitated communication between parties to reach consensus (Kitzinger, 1995). Co-operative discourse methods are also described by Renn (1999) involving establishment of development criteria and alternatives using value trees elicited by stakeholders and experts in round table meetings. Group Delphi is another systematic, interactive forecasting method which relies on a panel reaching consensus through sequential use of questionnaire and intermittent discussions.

Multi-criteria decision analysis (MCDA) has been proposed as a method to enhance stakeholder involvement in sediment management and to facilitate decision making of complex problems (Kim et al., 2010; Linkov et al., 2005; Yatsalo et al., 2007). The purpose of MCDA in these studies has been to support evaluation and selection among management alternatives in an interactive process involving, decision makers, stakeholders, scientists. Methodologically, MCDA requires developing hierarchy of criteria and metrics to compare management alternatives and subsequent elicitation of weights to quantify relative importance of criteria as well as scoring of alternative performance based on these criteria. The MCDA approach overcomes the limitations of unstructured individual and group decision-making by providing decision transparency and focusing discussion on assessing the weights and scores. Thus MCDA may be valuable in quantitative decision making; however, focus on participatory aspects in the involvement processes for sediment management is also warranted (Sparrevik et al., 2010).

In order to enhance the value of participatory stakeholder involvement in environmental management we propose a multi-criterial involvement process (MIP) which builds on the quantitative principles of MCDA, and also incorporates group interaction and learning through qualitative participatory methods. The process bears resemblance to earlier proposed MCDA processes for sediment management (Alvarez-Guerra et al., 2010; Oen et al., 2010; Kiker et al., 2005; Hong et al., 2010). However, this process also addresses recruitment and includes an involvement and learning step inspired by deliberative decision making using citizens juries (Soma, 2010; Smith & Wales, 2000). The application of the MIP is illustrated using sediment remediation alternatives for Bergen harbour in Norway as well as conducting the MCDA for three different advisory groups including local residents, local stakeholder and non-resident sediment experts. A comparison of individual versus group consensus-based ranking of alternatives is also presented.

THE MULTI-CRITERIAL INVOLVMENT PROCESS (MIP)

Stakeholder Involvement in Contaminated Site Management

A project execution process for managing contaminated sediments typically proceeds through specific project phases involving different actors in the process, Figure 1. Problem owners are usually active in the *problem formulation* and the *approval* phase where the selected concept is being approved by regulatory authorities (Oen et al., 2010). Consultants are normally active in the *concept evaluation* phase collecting lines of evidence and evaluating different concepts of remedial solution based on these data (Sparrevik & Breedveld, 2010). This often also includes preparing permit applications or environmental impact assessments (EIA). A simplified linear process is illustrated in Figure 1; however, in reality the process may be highly iterative with several decision processes conducted in parallel.

Stakeholders, here defined as people, organisations or groups who are affected by the issue and who have the power to make, support or oppose the decision (Susskind et al., 1999) tend normally to be involved late in the process, often prior to the *decision* making as a part of formal hearings. Individuals in their capacity as concerned residents are often not at all involved in the formal decision process insofar as they are not directly affected by project impacts and therefore not active in the formal hearings. The proposed remedial solution circulated for hearing is often designed based on technical feasibility, budget, time and political perspective. The manoeuvring space for changes at this stage of a project is often limited, which may cause problem owners to defend the solution instead of allowing a constructive stakeholder dialogue (Kasperson & Kasperson, 2005). This unfortunate situation may lead to significant opposition that could lead to increased costs and delays in the execution phase of contaminated site remediation projects (Sparrevik et al., 2010).

Description of the MIP Methodology

The MIP methodology as shown in Figure 2 uses multi-criteria decision analysis and consensus-based deliberation to structure the involvement process;

Objectives and alternatives

The 1st step of the process includes formulation of the project objectives, selection of alternatives and study recruitment methods. The problem owner responsible for the decision is presumed to be highly active in this phase. The 2nd step consists of recruitment of the advisory group. Depending on the project objectives defined in previous step, inhabitants representing public interests or stakeholders representing specific business interests may be recruited. The method used to recruit participants is important with regards to representation of different standing/values in relation to the project impacts. The project team (including researchers or consultants) is recommended to be responsible for this step.

Criteria metrics

The 3rd step consists of identification of criteria. The project team is assumed to be heavily involved in this step, possibly in cooperation with the problem owner. In the 4th step, impacts/scores for each criterion are assessed (based on available technical information and expert judgment). We recommend that the project team carry out this task, since it requires detailed technical knowledge about the alternative performances for selected criteria.

Measuring performance

The intention of the 5th step is to allow the advisory group to discuss the analysis with the project team and invite experts with specialized expertise to clarify questions. The advisory group should also evaluate and alter the criteria earlier proposed if necessary. The 6th step includes criteria weighting.

Information synthesis

In the 7th step data are processed by the project team and results are presented to the advisory group. In the 8th and last step, the advice from the project is presented to the problem owner.

MIP APPLICATION TO THE BERGEN HARBOUR SEDIMENT REMEDIATION STUDY

Objectives, Alternatives and Methods

Bergen harbour study objectives

The harbor area of Bergen is contaminated due to previous industrial activities such as naval shipyards and manufacturing industries, earlier releases of municipal sewerage and urban run-off from diffuse sources. One of the major contributors to harbor contamination is polychlorinated biphenyls (PCBs), originating from paint on house facades (Jartun et al., 2008). The area is 1 of 17 fjords in Norway prioritized for remedial actions by the Norwegian

Government (MD, 2006; MD, 2002). The area also has a dietary advisory for fish consumption based on PCB and mercury concentrations in fish. The work with contaminated sediment management has been progressing for several years focusing on site investigations, risk assessments and preparations of management plans. At present, complementary archeological investigations, as well as plans for field trial experiments to assess remediation methods are being executed. The objectives with the MIP in this case was to provide valuable advice to the problem owner on how advisory groups perceive hypothetical remediation alternatives distilled from the recommendations laid out in the management plans.

Remediation alternatives

Five alternatives for remediation of the contaminated sediments in Bergen harbor were suggested on as an outcome from the discussions with the problem owner. Alternative 1 constitutes natural recovery. The sources of contamination in Bergen have significantly decreased due to reduced industrial activity, better emission control and waste water treatment. It is estimated that background values for contaminant fluxes to the water from the sediments will be reached within a time span of 50 years due to natural deposition of clean sediments on top of the contaminated sediments (FM, 2005). Alternative 2 consists of an active reduction of the contaminant flux by capping a 1.5 km^2 area in the inner fjord basin with a 30 cm layer of clean material, as indicated in Figure 3. Capping has proven to be efficient to reduce contaminant transport from contaminated sites (Eek et al., 2008) but since it reduces sailing depth, it may have practical limitations (Palermo, 1998). A combined alternative was therefore formulated consisting of capping the majority of the area, combining with dredging areas with the highest concentration of PCB (hot-spot areas) and areas where sailing depth could be an issue, Figure 3. This alternative was then divided into 3 sub-alternatives based on the handling of the dredged material. In alternative 3, near shore disposal facilities were assumed to be constructed with the possibility to reclaim land for property development, whereas alternative 4 and 5 consisted of land disposal in local and national waste disposal facilities respectively. The transport distances to disposal was 1 km (near shore disposal), 12 km (local disposal) and 800 km (national disposal site) respectively.

Recruitment of Advisory Group

The MIP process requires involvement of advisory groups in the management process. Three advisory groups were created in this case: local residents, local stakeholders, and non-resident sediment experts as summarised in Table 1.

The local residents were randomly recruited based on the census lists in order to represent the general community interest in Bergen. The local residents were asked to act both in their capacity as individuals, promoting the private values of their household, and also as community representatives in a citizens' jury setting (Soma and Vatn 2010) promoting values representing the community needs. The local residents were selected to have fair gender and age representation and cover residential location in the immediate vicinity and at a distance (>3 km) from the Bergen harbor area. None of the persons recruited had previously been directly involved in contaminated sediment management. In total 20 participants were invited to participate in 3 sessions where of 17 (85%) participated in all sessions. At the last meeting, 4 residents' panel groups (5 participants in each group) were established for consensus-based deliberative evaluation.

The local stakeholder group was recruited to reflect specific interests in the Bergen harbor remediation case; based on a review of available documents, commentaries to the prepared management plan, media interest and correspondence with the problem owner. In total, 103 potentially interested parties for the group were identified. From this list, a subset of 14 stakeholders were selected based on mapping in an influence-interest grid (Chinyio & Olomolaiye, 2010). *Influence* was defined as the potential to affect the process either through formal legislative rights or by informal mobilisation through media and financial instruments. *Interest* was defined as the potential level of benefits or losses the stakeholder could experience from the process. Like influence, interest was categorized into formal interests such as regulatory issues and informal interests such as gain or loss of image and popularity. Discussions with 2 of the selected stakeholders and the problem owner were conducted, and as a result from these discussions the list was expanded from 14 to 23. These individuals were invited to participate in advisory meetings, but also to potentially function as an advisory group in forthcoming project stages in Bergen harbour. 16 people out of the 23 agreed to participate in a stakeholder group and 11 people (48%) participated in the advisory meetings.

The non-resident sediment experts were scientists / consultants and regulators working with contaminated sediments on a daily basis. The experts were recruited from the Oslo area without particular connections to Bergen. One session with 4 participants was organised with non-project researchers from one of the participating institutes in the study; a separate session with 8 participants (19 invited) was conducted using web based recruitment specifically targeting consultants, researchers and regulators working with contaminated sediments in the Oslo region.

Selecting Weighing Method and Identification of Criteria

Evaluation of different MCDA weighing methods for use in sediment management has been investigated earlier and all methods have strengths and weaknesses. The selection of appropriate methods will then be a choice between the accuracy of the utility or value based methods, the user friendliness of the AHP method or the simplicity of the outranking methods (Linkov et al., 2007). The use of MCDA with an advisory group early in the project process places the onus on finding methods that are simple and user friendly. Later stages of using MCDA for decision making should emphasis consistency and robustness as well. The AHP method (Saaty, 1987), was selected for this study based on its advantages in scoring and user friendliness. AHP completely aggregates the decision problem into a single objective function and uses a compensatory optimization approach. AHP uses a quantitative comparison method that is based on pair-wise comparisons of decision criteria, rather than utility and weighting functions.

The pair-wise comparison may be performed on different levels in a decision tree allowing people to compare criteria in pairs avoiding cognitively more challenging multiple simultaneous comparisons. In this study a hierarchical decision tree was used, organizing criteria in 3 levels reflecting the different pillars of sustainable development; environmental, societal and economical aspects (UN, 1987). Under each criterion, sub-criteria were added. The advisory groups were able to discuss and comment on the criteria, but only a limited number of alterations were performed in order to assure consistency between the groups.

Assessing Criteria Weights

Criteria weights to each criterion were set based on the EIA as presented in the management plan for the harbor (FM, 2005) and consultations with sediment experts. The criteria and the criteria weights are provided in Table 2 and are briefly described below;

Environmental criteria

The environmental risk was expressed as reduction in flux of PCB from the contaminated sediments, compared to today's baseline scenario. It was assumed that both capping and dredging will be very efficient in reducing the flux of contaminants from the sediments. The effect of dredging is slightly lower due to re-sedimentation of dredged material on top of the dredged seabed after the operation (FM, 2005).

The reduction in human health was assessed based on 10% exceedance of the maximum tolerable risk (MTR) (Baars, 2001), compared to the percent exceedance calculated for the current situation. The calculation of MTR is mainly based on consumption of fish which is exposed to contaminated sediments and to some degree from direct exposure to water and sediment during bathing. The calculation is based on a fish consumption of 15 meals pr month from the contaminated area, which is conservative, since no participants indicated that they had consumed fish from the harbor area during the past year.

Greenhouse gas house emissions were calculated based on the vessel transport distances from shore to capping area for alternatives 2-5, also including the distances from the dredging area to shore and lorry transport to disposal sites for alternatives 3-5. Emission data from Statistics Norway (www.ssb.no) was used for the calculations. In order to illustrate the magnitude of emission values to the advisory group participants, the figures were normalized against emissions from the estimated yearly emission from a private car (1530 kg CO_2-eq, www.naturvern.no).

Societal criteria

The construction impact was assessed as an ordinal number proportional to the surface area impacted during remediation. It was also assumed that capping would be both faster and cause less need to reorganize maritime activities in the area than dredging.

The disposal location was also addressed as an ordinal number where a local solution was considered more favorable than using a national disposal site. This is based on an assumption that sediment storage is best handled by local solutions (Breedveld, 2007). The criteria was also used to investigate whether disposal solutions close to residential areas were disfavoured (NIMBY-effects) (Dyer & Sarin, 1982).

The wharf area, Bryggen, is defined as a world heritage site on the UNESCO list (whc.unesco.org) and thus was an important aspect to be addressed in the MCDA. It was assumed that all marine operations will negatively affect the preservation of marine cultural heritage for the future, with dredging resulting in more negative impact than capping.

The possibility of land reclamation is only relevant for the dredging and near seashore alternative where it was assumed that construction of a confined disposal will establish land for property development. The area of reclaimed land was used as the criteria weight.

Economic criteria

Economic criteria were developed to observe how the distribution of cost, local (municipal financing) and national (governmental financing) were evaluated by the advisory groups. It was assumed that government would finance 25% of the cost for all alternatives and the remaining costs would be shared by local enterprises and the municipality. It was further assumed that enterprises would partially finance the dredging operation, since they would benefit from port development. Initially, private household financing through municipal taxation was addressed in the MCDA. However this criterion turned out to be problematic for the non-resident expert group and was not used for data evaluation.

Involvement and Learning

Three meetings were conducted with the residents and the stakeholders, whereas non-resident experts were invited to 1 session only. The involvement and learning step began in the 1st meeting and included familiarization, general discussions about the study and contaminated sediments and distribution of written material including description of the MCDA method, remediation alternatives and how the consequence criteria weights were estimated. The non-resident expert group received this information by e-mail. The content of the documents was explained in the 2nd meeting (1st meeting for experts) and sufficient time was allowed for questions and comments. Based on the presented material and their interest, the residents and stakeholder groups themselves requested expert witnesses to clarify and address specific topics related to the issue. The intention was to introduce a deliberative discussion valuable to both the advisory groups and the expert witnesses (Renn, 2006) using an approach base on citizens' jury methodology (Soma and Vatn 2010).

Weighing of criteria

Participants were asked to score the consequence criteria weights using questionnaires and were asked to weigh the criteria and adjacent sub criteria in pairs. The scoring was performed in the meeting for the resident and stakeholder groups, whereas the experts were asked to perform a preliminary scoring via e-mail before the meeting. Based on earlier experience (Soma, 2010), the original 9 value scale (Saaty, 1987) was replaced with a less comprehensive scale. We used in total 3 values guided by the text "strong weight" to emphasize high relative importance of the criteria and "neutral weight" to emphasize equal weighing. For participants not answering the question or marking all alternatives, the neutral score was used in the data presentation step. In this study participants were also asked to perform an intuitive ranking of the alternatives directly.

Data Presentation and Discussion

The software DEFINITE (Janssen R & Herwijnen MV., 2007) was used to process the data. For the residents, the results of the MCDA based on individual weighting of criteria were presented and discussed in the group as a whole, before they were divided into residents' panels. Results of individual versus residents' panel based weighting were then compared and discussed. The stakeholders performed their scoring in the last meeting and thus were not presented to the group in the meeting, due to lack of time. The experts where presented preliminary results from the scoring performed before the meeting, but they were allowed to change the scoring based on information given in the meeting. In all cases, the final weighting results were used for ranking the results presented in this study.

Evaluation of advice

Representatives from the residents' panels were invited to present their recommendations to the problem owner in a separate meeting. During the stakeholder meetings, the problem owner was actually present. This local stakeholder group is continuing to follow the process in Bergen with regular meetings, allowing them to contribute to on-going discussions about sediment remediation in the harbor.

RESULTS AND DISCUSSION

Results from the MCDA

The descriptive statistics for the groups show only significant variance between the groups with respect to years of residence in Bergen, Table 1. This was intentional since this group was recruited from persons not living in the Bergen area. The stakeholder group had mostly male participants, whereas the households/residents group was recruited to achieve a balanced male/female representation. It is especially interesting to observe that the differences in initial risk perceptive values relating to contaminated sediments are not significantly different between the groups. Since experiential beliefs also may influence the analytical outcome of a decision (Slovic et al., 2004) sharing the same initial beliefs about the subject may facilitate unbiased advice.

Table 3 shows the results from the scoring of sediment remediation alternatives. The results have first been calculated individually for each participant and subsequently integrated to mean values for each advisory group as illustrated in the table. A grand mean has also been calculated summarizing results from all groups. In addition, the results from a hypothetical scenario with equal score on all weights are presented to illustrate the influence of weighing on the results. Table 3 illustrates that the MCDA results center on alternatives 2 and 3. A t-test shows that the difference between alternative 1 ($t=-17.3$; $df=3$; $p<0.05$), 4 ($t=-5.5$; $df=3$;

$p<0.05$) and 5 ($t=-5.5$; $df=3$; $p<0.05$) against the grand mean of alternative 3 is significant. The difference between alternative 2 and 3 is non-significant ($t=-0.5$; $df=3$; $p=0.63$). It is also clear that the advisory group weighing has significantly affected the results compared to a hypothetical scenario with equal scores which results in a preference of natural recovery.

Additional information may be extracted from the MCDA by analyzing how participants score the criteria using centered weight analysis (Tervonen et al., 2009). This method normalizes each criteria against a scenario where all criteria score equal. In this case a positive value indicates that participants weigh this criterion higher than average and a negative value indicates that the criterion is weighted less than average. As seen in Figure 4, the averaged criteria weights are higher than normal for the reduction in human and environmental risk. This observation could explain the low score on a natural recovery scenario, since this alternative has lower weights for the reduction in human and environmental risk than the other remediation alternatives.

In addition to mean values, standard deviation is also presented in the figure 4. An analysis of variance (ANOVA) concludes that there exists a significant difference between advisory groups for 2 of the scored criteria, construction impacts and marine archaeological preservation. This indicates that for these criteria the differences between groups are significantly larger, than the variance within the group. Stakeholders are significantly more concerned by construction impacts ($F=6.0$; $df=3$; $p<0.05$), than the other groups. This is natural since this group includes representatives from organisations with close ties to Bergen harbour such as harbour authorities, boat owner associations, etc. For marine archaeological preservation ($F=3.9$; $df=3$; $p<=0.05$), stakeholders are again significantly more occupied with the subject than the non-resident sediment experts. It is also interesting to observe that when the residents respond as individuals they are less occupied with marine archaeological preservation, than when they act on behalf of the community in a residents' panel. This finding is consistent with social science theory and the beliefs in differences between individual normative and social normative values. Soma & Vatn (2010) also observed this phenomena where a deliberative citizen jury panel setting favoured mobilisation of social values in decision-making, rather than individual values.

A histogram of the inconsistency scores for the weighing of criteria is presented in Figure **5**. Normally a value below 0.1 is considered to be a consistent scoring (Saaty & Vargas, 1984). The figure shows that 43 out of 120 frequencies of inconsistency scores are below 0.1 with a mean value of 0.22. However, higher values and outliers are observed, indicating inconsistent scoring for some participants.

In general, statistical analysis of MCA weighting is expected to show high variances as the method is designed for decision-makers representing broader interests. Large standard deviation in some cases may also be explained by participants giving inconsistent weights, as documented by some of the high inconsistency scores.

Comparing MCDA to intuitive ranking

Table 4 summarizes the number of participants that indicate a specific remediation alternative as their preferred alternative either through MCDA or via intuitive ranking. The results given in Table 4 indicate that both methods suggest that alternative 2 and 3 are the most preferred. It is also clear that while the intuitive ranking shows that some of the participants also select other alternatives as their preferential choice, the MCDA clearly deselects these other alternatives as preferential.

Figure 6 illustrates the standard deviation between intuitive ranking and MCDA. This comparison provides valuable information regarding the robustness of the process. The results show that in most cases, the standard deviation using MCDA is lower compared to intuitive ranking.

Use of Results for Management Advice

Although the MCDA and the intuitive ranking result in the same most preferred remediation alternatives for all groups with alternative 2 (capping) and alternative 3 (capping combined with near shore disposal) being the overall most preferred, the MCDA is more equipped to sort out the “worst alternatives” than intuitive ranking. This potential use of MCDA has also been documented in earlier studies (Linkov et al., 2007). In addition, the use of MCDA in this study also results in lower standard deviation compared to intuitive ranking, which also is in line with findings in earlier studies (Linkov et al., 2009). For the residents group, the lower variance may also be a result of the deliberative consensus based weighting in the residents’ panels, relative to individual based non-deliberative weighting.

The centered weight analysis of the scoring responses indicates a strong focus on human and environmental risk reduction. The standard deviation presented in figure 4 is in some cases relatively high, which usually is problematic for interpretation of results. In this case it may however be a valuable piece of information, since it may indicate potential disagreement and therefore should be specifically addressed in the management process.

MIP Methodology Evaluations

<u>The role of MCDA in the process</u>

One may argue that some of the data obtained through the structured involvement process with MCDA may be obtained by other less resource intensive survey based methods. Willingness to pay studies have previously been used to map preferences and to map household and recreational users willingness to pay for sediment remediation in Norway (Barton et al., 2010). The advantage of such a large sample survey based approach is representation of public opinion about a project – a typically neglected interest in impact evaluation. Stated preference survey based valuation methods are also the only valuation methods that address socalled non-use and existence values. However, the contingent valuation method has been questioned regarding inflated willingness-to-pay for contingent on hypothetical project alternatives, versus actual willingness to pay once real alternatives are on the table. In Norway, low sample response rates have also been shown to affect representativeness of the affected population (Barton et al., 2010). Choice experiment surveys (CE), another stated preference survey-based method, have been proposed as an alternative with cost-saving, small-sample and preference elicitation advantages over contingent valuation (Bateman et al., 2002). CE can be used as a formal approach to 'map' stakeholder's individual preferences for remediation alternatives. However, a condition for using survey-based valuation methods is that the choice of sediment remediation alternatives can be described in terms of their component attributes in a survey setting. The experiences from our study question this possibility, since a highly interactive process seemed to be necessary for participants to be able to understand the relationship between alternatives and their impact criteria, and then to weight alternatives. However, survey-based studies may have a role to play in confirming preferences for a small set of specific project design criteria, once e.g. MCDA has narrowed alternatives and identified contentious criteria; for example willingness to accept reduced accessibility to certain beach locations during a remediation period. For stakeholders or residents to be able to embrace a complex decision such as selection of remediation alternatives, an involvement process with lateral learning, combined with MCDA giving structure, robustness and transparent documentation is preferable and certainly a necessary step before conducting stated preference survey's to evaluate the representativeness of stakeholder preferences in the population.

Use of AHP as weighing method

One of the main advantages using a compensatory weighing approach as AHP is the pair-wise weighing. This builds on the assumption that decision makers are more relaxed with making relative comparisons between the objectives than to score in absolute values (Baron, 1997). This assumption is to a large degree supported by this study. Even though, especially local stakeholders and non-resident experts criticized the choice of criteria and related weights, the majority of the participants were able to perform the weighing with logical results. This indicates that AHP is a suitable method for performing MCDA in advisory settings. It is however important to note that use of compensatory methods also has disadvantages to outranking methods. The standardization of the utility function is one obvious challenge. In this study, as in many studies, the standardization of the criteria weights was assumed to be linear for all criteria. According to prospect theory, this is not correct (Kahneman & Tversky, 1979). This theory argues that the value curve is asymmetrical to the reference point, i.e. people generally put more emphasis on "losses" than on "gains". Thus the asymmetrical value curve is steepest at the reference point subsequently over-emphasizing small losses compared to larger losses. In our case this means that criteria involving negative aspects, such as human and environmental risk, should be standardized differently to gains, such as land reclamation, Figure 4. This standardization requires a reweighting of impact scores. In our methodology, the impact assessment scoring document was explained to the groups, and despite substantial simplification, 2 rounds of expert witnesses were required to clarify its complexities for the residents' panel. In our opinion, further reweighting of impacts based on assumptions about individual's aversion to risk would probably have confused rather than clarified understanding of the EIA information used in the MCA. We therefore elected to disregard more advanced impact scoring methods. Another challenge with the AHP weighing approach is the hierarchical structure, since uneven distribution of sub-criteria between the main criteria will "dilute" the importance of the sub-criteria in groups having a greater number of criteria. In outranking methods these issues are not present. Therefore outranking methods may be the preferable choice for decision making where simplicity and robustness is favored to user friendliness in criteria weighing. Use of outranking also allows application of stochastic multi-criteria acceptability analysis (SMAA) to the results. In this method, criteria weights may be entered as distributions and a probabilistic approach is used to arrive at the most preferred alternative (Alvarez-Guerra et al., 2010). In this study the focus was to initiate a structured involvement process facilitated with MCDA. Criteria were selected to reflect the interest of residents, stakeholders and experts, rather than to be a comprehensive baseline for a decision since the project is in an early exploratory phase. The aspect of uncertainty evaluation was therefore not highly prioritized in this study. It is evident that as the project process advances, the selection of MCDA methodology should evolve, possibly focusing more on uncertainties in the criteria thus requiring selection of other weighing methods.

Evaluating Advisory Group Perception of the MIP

Within each session of the advisory group meetings the participants were asked with the aid of questionnaires, how they perceived the session and the involvement process. The general impression from both discussions and the results of the questionnaire were that people were in agreement that the involvement and learning process was positive in terms of information exchange between both expert witnesses and the advisory groups. These results are encouraging for the future application of MIP as it indicates a successful exchange of information and that both residents and stakeholders can produce valuable advice for the management process, with results well in line with what experts suggest.

There are however differences between the dynamics in the groups. The residents participated using the information and methods available well within the available time of 3 meetings. The non-residential experts performed the MIP within 1 meeting reflecting their familiarization with the subject; however, they were occupied with the assumptions made when assessing criteria and criteria weights. The work flow with the stakeholder group was different. Although the meetings were constructive, more time was spent on familiarization, clarification of the mandate for the group and questions relating to the MCDA process. The stakeholders also to some degree questioned the objectives of the study and focused on their role in the forthcoming project process. These discussions about roles and expectations are common in stakeholder involvement processes (Gerrits & Edelenbos, 2004) and emphasize the need to invest time in familiarization and formulation of objectives when addressing stakeholders in the involvement processes (Sjöberg L & Drottz-Sjöberg B.M, 2008). It is also important to consider these differences when deciding what kind of advisory groups to engage in the MIP.

CONCLUSION

In this paper we propose a multi-criterial involvement processes (MIP) to enhance participatory involvement processes. The evaluation of the Bergen harbor case study supports the feasibility of this method for these processes. This statement is mainly supported by 2 findings in the study. The results show that using MCDA as an integral step in the MIP adds structure and robustness to the involvement process and provides good documentation of criteria to be further addressed by the problem owner. Secondly, we perceive involvement and learning as important for the participants to be able to perform the MCDA in selection of remediation alternatives.

There also challenges to use MCDA in an advisory process. Firstly there are considerations to make regarding use of MCDA method. In this case the analytical hierarchy process (AHP) was selected due to user friendliness, but other settings may require use of other methods which have been shown in the literature to produce more robust results. Secondly this study showed that the quantitative scoring was perceived as problematic and was questioned especially by stakeholders and experts. The interactions and the qualitative information from the different advisory group discussions as suggested in the MIP, are therefore important in order to reduce misunderstandings and misinterpretations.

Finally, it is important to remember that the emphasis on method and process should be balanced using both quantitative and qualitative methods as proposed in the MIP. By including MCDA in the MIP, the structure and documentation of the process is ensured thus providing quantifiable results that can be replicated by third parties. By engaging group interaction and learning through participatory methods, the quality of the involvement process from recruitment to final discussions is preserved thus setting the stage for successful sediment remediation projects for both stakeholders and problem owners.

ACKNOWLEDGMENTS

The participation from people in the advisory group is this work is greatly appreciated. The authors are also grateful for the cooperation from Bergen Municipality and the sediment remediation project in Bergen. The authors would also like to thank the Norwegian Research Council for financing the work. Anonymous reviewers have performed excellent reviews of the paper.

REFERENCES

Alvarez-Guerra, M., Canis, L., Voulvoulis, N., Viguri, J. R., and Linkov, I. 2010. Prioritization of sediment management alternatives using stochastic multicriteria acceptability analysis. *Science of the Total Environment* 408:4354-4367.

Baars, A. J. R. M. C. Theeelen P. J. C. M. Janssen J. M. Hesse M. E. van Apeldoorn M. C. M. Meijerink L. Verdam and M. J. Zeilmaker. Re-evaluation of human-toxicological maximum permissible risk levels. 711701025 . 2001. Bilthoven, National Institute of Public Health and the Environment .

Baron, J. 1997. Confusion of Relative and Absolute Risk in Valuation. *Journal of Risk and Uncertainty* 14:301-309.

Barton, D., Navrud, S., Bjørkeslett, H., and Lilleby, I. 2010. Economic benefits of large-scale remediation of contaminated marine sedimentsΓÇöa literature review and an application to the Grenland fjords in Norway. *Journal of Soils and Sediments* 10:186-201.

Bateman, I., Great, B., and Dept.for Transport (2002). *Economic valuation with stated preference techniques*

a manual. Cheltenham, UK: Edward Elgar.

Breedveld, G. D. (2007). *Forurensning krever lokal løsning (contamination requires local solutions)*. http://www.forskning.no/Artikler/2007/oktober/1192712348.85: Norwegian Research Council Newsletter.

Chinyio, E. and Olomolaiye, P. O. (2010). *Construction stakeholder management*. Chichester, U.K: Wiley-Blackwell.

Dyer, J. S. and Sarin, R. K. 1982. Relative Risk-Aversion. *Management Science* 28:875-886.

Eek, E., Cornelissen, G., Kibsgaard, A., and Breedveld, G. D. 2008. Diffusion of PAH and PCB from contaminated sediments with and without mineral capping; measurement and modelling. *Chemosphere* 71:1629-1638.

FM. Bergen harbour management plan phase II (In Norwegian). 2005. Regional environmental body Hordaland.

Gerrits, L. and Edelenbos, J. 2004. Management of sediments through stakeholder involvement. *Journal of Soils and Sediments* 4:239-246.

Hong, G. H., Kim, S. H., Suedel, B. C., Clarke, J. U., and Kim, J. 2010. A decision-analysis approach for contaminated dredged material management in South Korea. *Integr Environ Assess Manag* 6:72-82.

Janssen R and Herwijnen MV. *DEFINITE 3.1. A System to Support Decisions on a Finite Set of Alternatives* (software package and user manual). 2007. Institute for Environmental Studies (IVM), Vrije Universiteit: Amsterdam.

Jartun, M., Ottesen, R. T., Steinnes, E., and Volden, T. 2008. Runoff of particle bound pollutants from urban impervious surfaces studied by analysis of sediments from stormwater traps. *Science of the Total Environment* 396:147-163.

Kahneman, D. and Tversky, A. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47:263-291.

Kasperson, J. X. and Kasperson, R. E. (2005). *The social contours of risk*. London: Earthscan.

Kiker, G. A., Bridges, T. S., Varghese, A., Seager, T. P., and Linkov, I. 2005. Application of Multicriteria Decision Analysis in Environmental Decision Making. *Integr Environ Assess Manag* 1:95-108.

Kim, J., Kim, S. H., Hong, G. H., Suedel, B. C., and Clarke, J. 2010. Multicriteria decision analysis to assess options for managing contaminated sediments: Application to Southern Busan Harbor, South Korea. *Integr Environ Assess Manag* 6:61-71.

Kitzinger, J. 1995. Qualitative Research - Introducing Focus Groups. *British Medical Journal* 311:299-302.

Linkov, I., Satterstrom, F. K., and Fenton, G. P. 2009. Prioritization of capability gaps for joint small arms program using multi-criteria decision analysis. *J Multi-Crit Decis Anal* 16:179-185.

Linkov, I., Satterstrom, F. K., Yatsalo, B. I., Tkachuka, A., Kiker, G., Kim, J., Bridges, T., Seager, T., and Gardner, K. H. 2007. *Comparative assessment of several multi+criteria desicion analysis tools for management of contaminated sediments*. In I. Linkov, G. Kiker, and R. Wenning (Eds.), Environmental Security in Harbors and Coastal Areas (pp. 233-249). Springer Netherlands.

Linkov, I., Varghese, A., Jamil, S., Seager, T., Kiker, G., and Bridges, T. 2005. *Multi-Criteria Decision Analysis: A Framework for Structuring Remedial Decisions at Contaminated Sites*. In I. Linkov and A. Ramadan (Eds.), Comparative Risk Assessment and Environmental Decision Making (pp. 15-54). Springer Netherlands.

MD. White paper 12 "Rent og rikt hav" (In Norwegian). 2002. Norwegian Ministry of Environment.

MD. White paper 14 "Sammen for et giftfritt miljø - forutsetninger for en tryggere fremtid" (In Norwegian). 2006. Norwegian Ministry of Environmwent.

Oen, A. M. P., Sparrevik, M., Barton, D. N., Nagothu, U. S., Ellen, G. J., Breedveld, G. D., Skei, J., and Slob, A. 2010. Sediment and society: an approach for assessing management of contaminated sediments and stakeholder involvement in Norway. *Journal of Soils and Sediments* 10:202-208.

Palermo, M. R. 1998. Design considerations for in-situ capping of contaminated sediments. *Water Science and Technology* 37:315-321.

Renn, O. 1999. A Model for an Analytical Deliberative Process in Risk Management. *Environmental Science & Technology* 33:3049-3055.

Renn, O. 2006. Participatory processes for designing environmental policies. *Land Use Policy* 23:34-43.

Saaty, R. W. 1987. The analytic hierarchy process--what it is and how it is used. *Mathematical Modelling* 9:161-176.

Saaty, T. L. and Vargas, L. G. 1984. Comparison of eigenvalue, logarithmic least squares and least squares methods in estimating ratios. *Mathematical Modelling* 5:309-324.

Sjöberg L and Drottz-Sjöberg B.M 2008. *Attitudes towards nuclear waste and siting policy: experts and the public*. In Arnold P.Lattefer (Ed.), Nuclear Waste Research: Siting, Technology and Treatment (pp. 44-47). Nova Science Publishers, Inc.

Slob, A. F. L., Ellen, G. J., and Gerrits, L. 2008. *Sediment management and stakeholder involvement*. In N. O. Philip (Ed.), Sustainable Management of Sediment Resources Sediment Management at the River Basin Scale (pp. 199-216). Elsevier.

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. 2004. Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis* 24:311-322.

Smith, G. and Wales, C. 2000. Citizens' Juries and Deliberative Democracy. *Political Studies* 48:51-65.

Soma, K. 2010. Framing participation with multicriterion evaluations to support the management of complex environmental issues. *Env Pol Gov* 20:89-106.

Soma, K. and Vatn, A. 2010. Is there anything like a citizen? A descriptive analysis of instituting a citizen's role to represent social values at the municipal level. *Env Pol Gov* 20:30-43.

Sparrevik, M. and Breedveld, G. D. 2010. From Ecological Risk Assessments to Risk Governance. Evaluation of the Norwegian Management System for Contaminated Sediments. *Integr Environ Assess Manag* 6:240-248.

Sparrevik, M., Ellen, G. J., and Duijn, M. 2010. Evaluation of Factors Affecting Stakeholder Risk Perception of Contaminated Sediment Disposal in Oslo Harbor. *Environmental Science & Technology*null.

Susskind, L., McKearnan, S., and Thomas-Larmer, J. (1999). *The consensus building handbook*

a comprehensive guide to reaching agreement. Thousand Oaks, Calif: Sage Publications.

Tervonen, T., Figueira, J. R., Lahdelma, R., Dias, J. A., and Salminen, P. 2009. A stochastic method for robustness analysis in sorting problems. *European Journal of Operational Research* 192:236-242.

UN. Report of the World Commission on Environment and Development. 12-11-1987. United nations.

Yatsalo, B. I., Kiker, G. A., Kim, J., Bridges, T. S., Seager, T. P., Gardner, K., Satterstrom, F. K., and Linkov, I. 2007. Application of Multicriteria Decision Analysis Tools to Two Contaminated Sediment Case Studies. *Integr Environ Assess Manag* 3:223-233.

LIST OF TABLES

LIST OF FIGURES

cates that the criterion is scored lower than average. The error bars represent one standard deviation.

Figure 5 Histogram of inconsistency scores for the different advisory groups. Scores for choice between, environmental, societal and economical criteria as well as scores within the environmental and societal criteria are given in the figure. The economical criterion is only 1 pair and has therefore no inconsistency

Figure 6 Analysis of the standard deviation within the groups for the different alternatives for both intuitive ranking and MCDA