Mine Closure 2026 – AB Fourie, G Boggs, J Heyes & M Tibbett (eds)

doi:10.36487/ACG_repo/2615_118

Artificial intelligence for robust multicriteria analysis in mine closure option assessment

Ethan Hari [a,*]

[a] *BHP, Australia*

Abstract

Multicriteria analysis (MCA) is widely used in mine closure planning to compare closure options across competing objectives such as regulatory compliance, environmental performance, social value and technical effectiveness. When applied well, MCA provides a clear and defensible basis for decision-making. However, long-term closure decisions are often characterised by uncertainty in site conditions, diverse stakeholder perspectives and sensitivity to assumptions that are difficult to fully represent using fixed weights and single point scores alone.

This paper presents an artificial intelligence- (AI) enhanced MCA workflow that strengthens established closure MCA practice by introducing preference learning, stochastic evaluation and explainability. Preference learning captures criteria weights from practical pairwise comparisons across 5 stakeholder groups, allowing different perspectives to be represented transparently rather than collapsed into a single negotiated weight set. Stochastic multicriteria acceptability analysis is used to evaluate option performance across many plausible futures by sampling from uncertainty ranges in scores and, where applicable, weights. Explainability tools based on Shapley value decomposition highlight which criteria most strongly drive rankings and where trade-offs are most significant.

A conceptual case study demonstrates the workflow applied to a legacy closure problem with 3 options assessed against 6 criteria under 3 climate scenarios. Under the regulator perspective and a dry climate scenario, the preferred option achieved rank one in 91% of simulated futures, providing a clear and quantified measure of robustness. The approach is implemented in Excel with AI computation support and retains professional judgement, established governance and existing workflows throughout.

Keywords: *artificial intelligence, multicriteria analysis, mine closure planning, decision support, uncertainty, explainability*

1 Introduction

Mine closure planning involves decisions with long-term environmental, social, regulatory and financial consequences. For most closure problems there is no single correct answer. Instead, several technically feasible options must be compared, each performing differently against competing objectives and stakeholder expectations. These decisions must remain defensible many years after they are made, often under conditions that are uncertain and subject to change (International Council on Mining and Metals [ICMM] 2019).

Multicriteria analysis has become one of the most commonly used decision support methods in mine closure planning. It provides a structured approach for comparing options across multiple criteria and making trade-offs explicit. MCA is well accepted by regulators and stakeholders because it documents how decisions were made and why certain options were preferred (Department of Mines, Industry Regulation and Safety [DMIRS] 2020).

[*] Corresponding author. Email address: ethan.hari@bhp.com

As closure decisions become more complex and long-lived, practitioners increasingly need to consider uncertainty in long-term performance, evolving site conditions and the perspectives of multiple stakeholder groups. The analytic hierarchy process (AHP) introduced by Saaty (1980) provided one of the earliest structured approaches to deriving criteria weights through pairwise comparisons, and this remains a foundation of most closure MCA practice. However, methods such as stochastic multicriteria acceptability analysis (SMAA) (Lahdelma et al. 1998; Lahdelma & Salminen 2001) offer the ability to explicitly represent uncertainty in both criteria measurements and decision-maker preferences, and they have been applied successfully in environmental decision-making contexts (Tervonen & Figueira 2008).

Artificial intelligence (AI) offers an opportunity to enhance existing MCA practice by supporting these aspects while retaining professional judgement, established workflows and existing governance frameworks. In this context, AI is used to strengthen and extend MCA rather than replace it. The purpose of this paper is to describe an AI-enhanced MCA workflow, explain the underlying methods including SMAA and Shapley value-based explainability, and demonstrate the approach through a conceptual case study relevant to mine closure option assessment.

2 Multicriteria analysis in mine closure planning

MCA is a decision support method that evaluates a set of alternatives against a defined set of criteria. Each option is scored against each criterion, and criteria are weighted to reflect their relative importance. The weighted scores are combined to produce an overall value for each option, which is then used to rank the alternatives.

In mine closure planning, MCA is particularly suited because decisions rarely depend on a single factor. Closure options must be acceptable to regulators, protect human health and the environment, respect traditional owner and community expectations, minimise residual liability, and be technically and financially achievable. The ICMM Integrated Mine Closure: Good Practice Guide (2019) identifies the need for structured and transparent decision-making processes that consider multiple objectives and stakeholder perspectives throughout the closure life cycle (ICMM 2019).

When applied well, MCA improves transparency and consistency in decision-making. It helps project teams explain how trade-offs were considered and provides a clear audit trail for regulators, communities and future asset owners. The Western Australian mine closure regulatory framework also expects that closure planning decisions are supported by documented risk assessment and options analysis (DMIRS 2020).

3 Typical multicriteria analysis approach and opportunities for enhancement

A typical closure MCA follows a structured and well-established sequence that is familiar to closure practitioners and widely accepted by regulators and stakeholders. Closure options are first defined and screened for obvious fatal flaws. Evaluation criteria are then selected to reflect regulatory, environmental, social, technical, risk and cost considerations relevant to the closure context. Criteria weights are commonly derived through facilitated workshops, often using pairwise comparisons to explore the relative importance of criteria. Options are then scored against each criterion using predefined scoring scales, and weighted scores are combined to produce an overall ranking.

Pairwise comparison methods are widely used in closure MCA because they are intuitive, practical to apply in workshops and support transparent discussion of trade-offs. The AHP (Saaty 1980) is the most well-known of these methods and has been applied extensively in environmental and resource management decision-making. An example of a typical pairwise weighting approach is shown in Figure 1, where criteria are compared against each other and relative importance is inferred based on how often one criterion is ranked higher than another. The resulting comparisons are converted into a single set of weights that are then applied consistently across all options.

Pairwise Preference Matrix

Decision Matrix	Criterion 1	Criterion 2	Criterion 3	Criterion 4	Criterion 5	Criterion 6	Weighting
Criterion 1	N						22.2%
Criterion 2	S	N					25.9%
Criterion 3	S	S	N				25.9%
Criterion 4	MW	MW	W	N			9.3%
Criterion 5	MW	MW	W	W	N		1.9%
Criterion 6	W	W	W	N	MS	N	14.8%

Note: MS = Much Stronger S = Stronger N = Neutral W = Weaker MW = Much Weaker

Figure 1 Example of traditional pairwise comparison method used to derive fixed criteria weights in a closure multicriteria analysis

This approach provides a clear, auditable and defensible basis for closure decision-making. For complex and long-lived closure decisions, however, there are opportunities to further enhance how well MCA represents uncertainty, stakeholder diversity and decision sensitivity.

Criteria weights derived through workshops typically reflect a negotiated outcome at a particular point in time. In practice, different stakeholder groups such as regulators, traditional owners, communities and asset owners may hold distinct perspectives that are not always explored in detail within a single weighting exercise. As closure plans evolve and new information becomes available, priorities and preferences may also change.

Similarly, option scoring is commonly undertaken using single point values, even though long-term closure performance is subject to uncertainty. Factors such as future climate conditions, material performance, water quality response and long-term management effectiveness are often uncertain and difficult to represent deterministically. Without explicitly testing this uncertainty, rankings may become sensitive to implicit assumptions that are not always visible to decision-makers.

In addition, closure decisions often benefit from a clear explanation of why one option is preferred over another and under what conditions that preference might change. While traditional MCA supports transparency, additional diagnostic insights can further clarify which criteria most strongly influence outcomes and where trade-offs are most significant. These considerations do not reflect the shortcomings of MCA itself, but rather opportunities to enhance an already robust and widely accepted approach when applied to high-consequence, uncertain and long-term closure decisions.

4 Development of an integrated artificial intelligence-enhanced workflow

AI is not used to replace expert judgement in mine closure planning. Instead, it is applied as a decision support tool that strengthens the existing MCA process by better handling complexity, uncertainty and repetition. The intent is to support engineers, regulators and stakeholders in making more informed and defensible closure decisions, not to automate or override professional judgement. Figure 2 illustrates how the

AI-enhanced approach addresses key limitations of traditional MCA while retaining the structured decision-making framework familiar to closure practitioners.

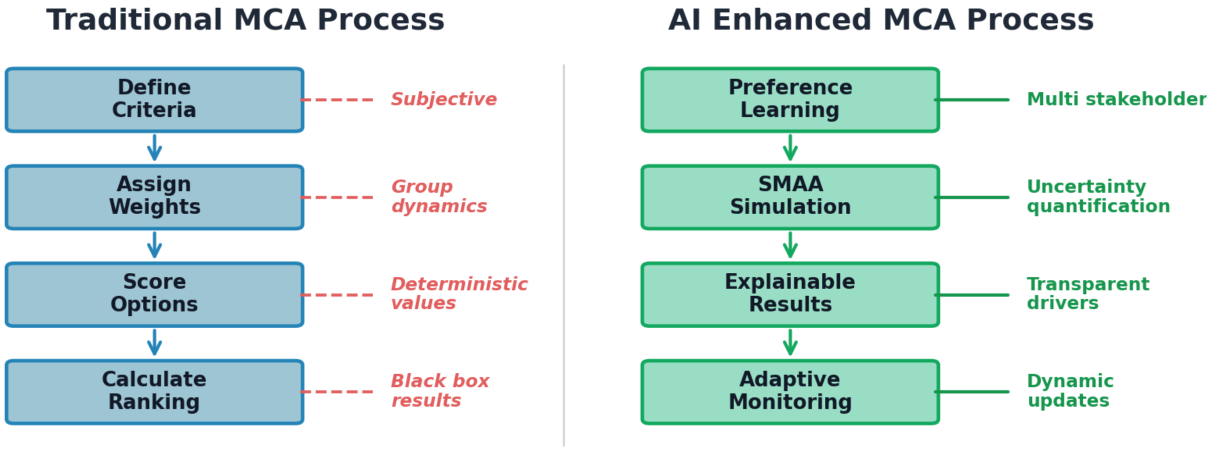

Figure 2 Comparison between traditional multicriteria analysis (MCA) and artificial intelligence- (AI) enhanced MCA, highlighting the improved treatment of stakeholder preferences, uncertainty, transparency and adaptability. The AI enhanced MCA framework incorporates preference learning, Stochastic Multicriteria Acceptability Analysis (SMAA), explainable results and adaptive monitoring.

In Figure 2, each step in the traditional MCA process corresponds directly to an enhanced step in the AI--enhanced workflow. Fixed consensus weights derived from a single workshop are replaced by preference learning, which captures separate weight profiles for each stakeholder group from practical pairwise comparisons. Single point scoring is replaced by uncertainty-based scoring using probability distributions that reflect the project team's confidence in each estimate. Deterministic ranking is replaced by SMAA, which tests how often each option ranks highest across thousands of simulated futures. Limited sensitivity testing is replaced by Shapley value-based explainability, which identifies which criteria drive the ranking and where trade-offs are most significant. Static one-off assessment is replaced by adaptive monitoring with defined triggers that prompt reassessment when site conditions change.

AI adds value to MCA in 3 main ways and also supports adaptive closure planning.

Firstly, it improves how criteria weights are derived by learning stakeholder preferences from limited and practical pairwise input rather than relying on a single fixed-weight set agreed through a workshop. Preferences are captured separately for each stakeholder group, allowing diverse perspectives to be carried through the analysis rather than collapsed into a consensus that may not accurately represent any single group.

Secondly, it allows uncertainty to be explicitly represented by defining option performance scores as probability distributions rather than single point values. This means that each option and criterion combination is described by a mean score and a standard deviation that reflects the project team's confidence in that estimate. Scores are then sampled from these distributions during the stochastic analysis rather than held fixed.

Thirdly, it improves transparency and explainability by using methods such as Shapley value decomposition (Shapley 1953) to show which criteria drive the ranking of each option and how sensitive the results are to key assumptions. This provides clear diagnostic information that supports communication with regulators, communities and internal decision-makers.

In addition, it supports adaptive closure planning by applying constraint checks and linking assumptions to monitoring triggers. This allows the MCA to be reassessed if site conditions change, rather than treating the assessment as a one-off exercise.

It is important to note that all inputs to the workflow including performance scores, uncertainty ranges, stakeholder weights and constraint thresholds are defined by the project team through expert judgement and structured elicitation. The AI enhancement does not generate these inputs. Instead, it handles the computational load of running thousands of stochastic simulations, automates Shapley value decomposition across all option pairs and scenarios, and supports interpretation of results through natural language summaries. The role of AI is to process, analyse and communicate rather than to define the inputs or override professional judgement.

The AI-enhanced workflow draws on 5 core capabilities of AI based computation. Learning is applied through the preference learning algorithm, which derives consistent stakeholder weight profiles from limited and potentially inconsistent pairwise comparison inputs. Reasoning is applied through the SMAA stochastic simulation engine, which evaluates option robustness by reasoning across thousands of plausible futures rather than relying on a single deterministic outcome. Pattern recognition is applied through the Shapley value decomposition, which identifies which criteria most strongly drive rankings and detects patterns of sensitivity across option pairs, stakeholder segments and climate scenarios. Problem solving is applied through constraint filtering and adaptive monitoring, which determine which options meet non-negotiable requirements and when changed site conditions should trigger reassessment of the closure decision. Natural language understanding is applied through the AI reporting and interpretation layer, which generates plain language summaries of simulation outputs, Shapley results and scenario comparisons to support communication with non-technical stakeholders and decision-makers.

The AI-enhanced workflow retains the familiar MCA structure while modifying how key steps are implemented. Criteria selection remains a human-led activity, supported by guidance and consistency checks. Preference learning is used to derive criteria weights from partial pairwise comparisons, allowing different stakeholder perspectives to be represented explicitly. The AI-enhanced MCA workflow applied in this study is shown in Figure 3.

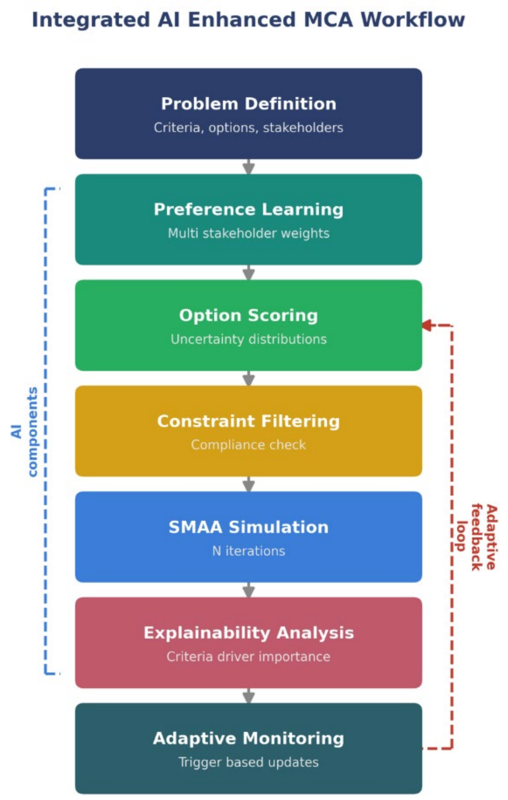

Figure 3 Artificial intelligence-enhanced multicriteria analysis (MCA) workflow showing how established MCA steps are modified to incorporate preference learning, Stochastic Multicriteria Acceptability Analysis (SMAA), explainability and adaptive monitoring

Option scoring is completed using score ranges that reflect uncertainty in long-term performance, climate conditions and regulatory outcomes. Stochastic simulations are then run using SMAA to assess how often each option ranks highest under plausible futures. Explainability tools generate clear outputs showing which criteria drive the rankings and identify threshold conditions where a different option would become preferred. Constraint checking ensures that options failing non-negotiable requirements are excluded. Monitoring triggers link the MCA to closure implementation so that reassessment is prompted if site conditions change.

4.1 Preference learning

In a traditional closure MCA, criteria weights are typically agreed through a facilitated workshop involving a mixed group of stakeholders and technical specialists. The outcome is usually a single set of weights that represents a negotiated consensus. While this approach is practical and well understood, it can mask important differences in how various stakeholder groups view the relative importance of closure criteria.

The AI-enhanced approach addresses this by capturing preferences separately for each stakeholder group through structured pairwise comparisons. Each comparison asks a simple question: for a given pair of criteria, which is more important and by how much? The strength of preference is recorded on a defined scale. These inputs are practical to collect in workshop settings or through structured interviews and do not require stakeholders to assign numerical weights directly.

The pairwise inputs are processed using a preference learning algorithm that derives consistent weight estimates from partial and potentially inconsistent comparisons. Unlike the standard AHP eigenvector approach which requires a complete and consistent comparison matrix, the preference learning method can work with incomplete pairwise data and resolves inconsistencies by finding the weight set that best fits the observed preferences. This makes it practical for real closure applications where not all stakeholder groups may complete every comparison.

If stakeholder priorities change over time, updated pairwise comparisons can be collected and the preference learning algorithm rederives the weight profiles without requiring the model to be rebuilt. The SMAA simulation is then re-run automatically with the updated weights, allowing the assessment to reflect evolving perspectives efficiently.

It is recognised that pairwise comparisons can introduce cognitive biases into the weighting process, particularly if collected at an early stage when information is limited and stakeholder positions are still forming. The framework partly mitigates this by testing robustness across multiple stakeholder perspectives and uncertainty ranges through the SMAA simulation rather than relying on a single weight set. However, the quality of the preference inputs will always depend on the timing and context in which they are collected. Where possible, preference elicitation should be revisited as the project matures and more information becomes available.

4.2 Stochastic multicriteria acceptability analysis

SMAA is the core analytical method used in the AI-enhanced MCA workflow. SMAA was originally developed by Lahdelma, Hokkanen and Salminen (1998) and extended to consider all ranks by Lahdelma and Salminen (2001) as SMAA 2. Tervonen and Figueira (2008) provided a comprehensive survey of the SMAA family of methods and their applications in environmental and public decision-making.

The fundamental idea behind SMAA is that instead of evaluating closure options using a single fixed set of scores and weights, the analysis is repeated many times using values sampled from defined uncertainty ranges. This produces a distribution of outcomes rather than a single deterministic ranking, which gives a much clearer picture of how robust each option is under uncertainty.

For each option and criterion combination, the project team defines a performance score as a probability distribution with a mean value and a standard deviation. Scores are modelled as normal distributions where the mean reflects the team's best estimate and the standard deviation reflects their confidence in that estimate. Where distributions overlap between options on a given criterion, there are plausible futures where one option could score as high or higher than another. This is exactly why stochastic simulation is needed rather than simply comparing single point values.

In this study, normal distributions are used for simplicity to demonstrate the method. However, the framework is not limited to normal distributions. It can accommodate other distribution types such as triangular, lognormal or uniform distributions, depending on the nature of the criterion and the available data. For example, a lognormal distribution may be more appropriate for criteria influenced by cost or environmental performance where outcomes are bounded on one side. The choice of distribution in a real application would be informed by the data characteristics and the project team’s understanding of the uncertainty for each option and criterion combination.

In each simulation iteration, performance scores are randomly sampled from their defined distributions for every option and criterion combination. The sampled scores are then combined with the stakeholder weight profile using a weighted additive value function.

The total value V for option i in simulation t is calculated as:

$$V(i,t) = \Sigma\, w(j) \times s(i,j,t) \tag{1}$$

where:

V(i,t)	=	total weighted score for option i in simulation t
w(j)	=	weight assigned to criterion j under the selected stakeholder segment
s(i,j,t)	=	performance score sampled from the defined distribution for option i against criterion j in simulation t, and the summation is over all j criteria.

The weighted additive value function assumes that criteria are mutually preferentially independent, which is a standard assumption in MCA practice. The interaction diagnostics described in Section 4.5 provide a mechanism for identifying and recording cases where this assumption may not hold. The options are then ranked based on their total values, with the highest value receiving rank 1, the next receiving rank 2 and so on.

The key output of SMAA is the rank acceptability index. This is calculated as:

$$RA(i,r) = (1/N) \times \Sigma\, I(rank(i,t) = r) \tag{2}$$

where:

RA(i,r)	=	rank acceptability index for option i achieving rank r
N	=	total number of simulation
rank(i,t)	=	rank assigned to option i in simulation t
I	=	indicator function that equals 1 when option i achieves rank r in simulation t and 0 otherwise.

The summation is over all t simulations. A rank acceptability index close to 1.0 for rank 1 means the option almost always performs best regardless of which plausible combination of scores is sampled. This provides a direct and intuitive measure of robustness that is easy to communicate to decision-makers and stakeholders.

Before ranking, constraint filtering can be applied to reflect non-negotiable requirements. Any option that scores below a defined threshold on a critical criterion in a given simulation is excluded from ranking for that iteration and automatically assigned the lowest rank. The remaining options are ranked among themselves. This ensures that the stochastic evaluation focuses on outcomes that would be considered acceptable in practice.

Constraint thresholds are not fixed permanently. If stakeholder expectations or regulatory requirements shift over time, the thresholds can be updated and the simulation re-run. Options that were previously excluded because they fell below a threshold would automatically re-enter the ranking if they now meet the updated requirement. This means the framework does not permanently discard options but allows them to be reconsidered as circumstances change.

4.3 Scenario analysis

Long-term closure decisions are particularly sensitive to uncertainty in future environmental conditions. Climate change projections suggest that many mine closure sites in Australia may experience significant shifts in rainfall patterns, temperature extremes and water balance over the coming decades. These changes can directly affect closure performance, including revegetation success, water quality outcomes, erosion rates and long-term material stability.

To represent this uncertainty, the AI-enhanced MCA workflow includes scenario analysis as a core component. Rather than assessing options under a single assumed future, the assessment is repeated under multiple plausible climate scenarios. Each scenario modifies the performance scores for relevant criteria to reflect how option performance would change under different conditions.

In the current framework, scenario analysis adjusts performance scores only. Criteria weights and the criteria themselves remain unchanged across scenarios. This is a deliberate simplification because changing weights would imply a shift in stakeholder values rather than a change in site conditions, and these are treated as separate considerations. However, it is recognised that in practice some scenarios could influence how stakeholders prioritise criteria. For example, under a dry climate scenario, water-related criteria may increase in importance or even become non-negotiable for some groups. The framework can accommodate this by running the SMAA simulation with scenario specific weight profiles or updated constraint thresholds if the project team and stakeholders determine that a scenario warrants a change in priorities. This would be handled as a separate sensitivity run rather than being built into the default scenario logic.

4.4 Explainability and Shapley value decomposition

One of the challenges in traditional MCA is explaining to stakeholders and regulators exactly why one option is preferred over another and what would need to change for that preference to shift. The AI-enhanced MCA addresses this through explainability tools based on Shapley value decomposition.

The Shapley value is a concept from cooperative game theory (Shapley 1953) that provides a fair way to allocate the total outcome of a cooperative effort among the individual contributors. In the context of MCA, the criteria are treated as the players in a cooperative game and the total weighted score of each option is the outcome. The Shapley value for each criterion represents its marginal contribution to the overall ranking of an option, averaged across all possible combinations of criteria. The Shapley value decomposition presented here is applied to the mean performance scores to provide a deterministic explanation of the expected ranking drivers. In practice, the Shapley analysis could also be applied across the simulation iterations to show how criteria contributions vary under uncertainty, though the deterministic approach provides a clearer and more communicable summary for stakeholder engagement.

For a weighted additive value function like the one used in this study, the Shapley value for a criterion simplifies to the weighted difference in performance between two options on that criterion. The contribution of criterion j to the gap between option A and option B is calculated as the weight for criterion j multiplied by the difference in scores between option A and option B on that criterion. A positive contribution means the criterion favours option A, while a negative contribution means it favours option B. The sum of all contributions equals the total weighted score gap between the 2 options.

In a typical closure assessment with multiple options, stakeholder groups and climate scenarios, the number of Shapley decompositions required grows quickly and manual calculation becomes impractical. The AI automates the decomposition across all option pairs, stakeholder segments and scenarios in a single run. It also applies natural language understanding to interpret the numerical outputs and generate plain language summaries that explain in accessible terms which criteria are driving the preference, where trade-offs are most significant and what would need to change for the ranking to shift. This moves beyond a standard sensitivity table by providing a narrative explanation that can be communicated directly to non-technical stakeholders, regulators and community groups without requiring them to interpret raw numerical outputs.

4.5 Interaction diagnostics

The tool also includes an interaction diagnostics sheet that records any synergy or interaction effects between criteria pairs. In a simple weighted additive model, criteria are assumed to be independent. However, in practice, some criteria may reinforce or offset each other. Recording these interactions provides an additional layer of diagnostic transparency that can be reviewed as part of the decision governance process.

4.6 Adaptive monitoring and reassessment

Mine closure decisions are not made once and then left unchanged. Site conditions evolve, monitoring data accumulates and regulatory expectations may shift over time. The AI-enhanced MCA workflow includes a mechanism for linking the assessment to ongoing monitoring and triggering reassessment when conditions change.

The framework includes a set of monitoring triggers linked to specific criteria and closure assumptions. Each trigger defines a measurable parameter, a threshold value, an action to be taken if the threshold is exceeded and the expected effect on relevant MCA criteria.

This approach aligns with adaptive management principles (Williams & Brown 2014) and ensures that the MCA is not treated as a static document but as a living decision framework that responds to real site performance. When a trigger is activated, the relevant performance scores are updated and the SMAA simulation is re-run to determine whether the preferred option has changed or additional intervention is needed.

In the current framework, monitoring triggers cause the relevant performance scores to be updated and the SMAA simulation to be re-run. They do not cause automatic reselection of the assessment criteria. Criteria reselection would represent a more fundamental change to the closure decision framework and would typically require a facilitated review involving the project team and key stakeholders. However, if a trigger reveals a new risk or consideration that is not captured by the existing criteria, this would be flagged as a recommendation for a broader reassessment of the MCA structure rather than handled within the automated loop.

5 Conceptual case study and results

This conceptual case study demonstrates how the AI-enhanced MCA workflow can be applied to a realistic closure option assessment. The purpose is to show how preference learning, uncertainty-based scoring, stochastic evaluation and explainability work together to support a robust and defensible decision without replacing established MCA practice.

5.1 Problem definition and options

The closure problem considered is selection of a preferred closure approach for a legacy waste facility with long-term seepage risk and strong stakeholder interest. Three closure options are assessed.

Option 1, Optimised without, represents a minimum intervention approach that addresses the highest priority risks while relying on long-term monitoring and management for remaining residual risks. option 2 Risk remediation represents a targeted remediation approach that proactively addresses key risks more than option 1 while still leaving some residual risks that require ongoing management. option 3 Low liability represents a higher intervention approach aimed at minimising long-term residual risk and liability and, where practicable, moving towards a walk-away outcome where ongoing active management is no longer required.

The 6 assessment criteria are Likelihood of acceptability to regulators, Social and environmental value, Maintain optionality for future use, Traditional owner acceptance, Effectiveness of technical solution and residual liability, and Operations and state agreement repudiation risk. These criteria reflect typical closure decision drivers and align with common closure options assessment practice. Detailed cost benefit analysis including closure costs, capital implications and balance sheet impacts is typically undertaken as a complementary assessment alongside the MCA rather than embedded within the multi criteria scoring framework.

Each option is scored against each criterion using a 5-point scale where 1 represents the lowest performance and 5 represents the highest performance. A score of 1 indicates that the option performs poorly against the criterion with significant risk or non-compliance. A score of 2 indicates below-average performance with notable gaps or residual concerns. A score of 3 indicates acceptable performance that meets minimum expectations. A score of 4 indicates good performance with minor residual risks. A score of 5 indicates strong performance with negligible residual risk or full alignment with the criterion objective.

5.2 Performance scoring under uncertainty

For each option, criterion and scenario combination, the project team defines a performance score as a normal distribution with a mean and standard deviation. Each option is assessed under 3 climate scenarios, dry, baseline and wet, to test how robust the ranking is to future climate conditions. The wet scenario uses Representative Concentration Pathway 8.5 (RCP8.5), a high greenhouse gas emissions projection. A standard deviation of 0.5 is used across all scores in this case study to represent a consistent level of uncertainty, though in practice this would vary by criterion and option depending on available evidence and confidence.

Under the dry scenario, option 3 Low liability scores consistently higher than options 1 and 2 across most criteria. On Likelihood of acceptability to regulators, option 3 has a mean score of 4.0 compared to 3.4 for option 2 and 2.8 for option 1. On Effectiveness of technical solution and residual liability, option 3 scores 3.8 compared to 3.0 for option 2 and 1.8 for option 1. The one criterion where option 3 scores lower is Maintain optionality for future use, with a mean of 2.8 compared to 3.3 for option 1 and 3.2 for option 2. This reflects that the higher intervention approach reduces flexibility for future alternative land uses. The same scoring pattern is observed under the baseline and wet RCP8.5 scenarios, with slight adjustments to reflect the expected change in performance under each climate condition. Under the wet RCP8.5 scenario, option 3 scores 4.1 on Regulatory acceptability compared to 4.0 under dry and 4.2 under baseline, illustrating the small magnitude of adjustment applied across scenarios.

In practice, performance scores and their uncertainty ranges would be informed by site-specific studies, modelling outputs, monitoring data and structured expert elicitation. The uniform standard deviation of 0.5 used in this case study is a simplification for demonstration purposes. In real applications, different criteria and options would have different levels of uncertainty depending on the available evidence and confidence in delivery.

5.3 Stakeholder weight profiles

In the conceptual case study, 5 stakeholder groups are defined: regulator, traditional owners, operations, community and environment. Each group provides pairwise assessments that reflect its perspective on the relative importance of the 6 evaluation criteria. The pairwise inputs are then converted into a weight profile for each group. For each stakeholder group, the weights across all 6 criteria sum to 1.0. This is consistent across all groups to ensure that the weighted scores are comparable when evaluated through the SMAA simulation.

Under the regulator perspective, the highest weights are assigned to Likelihood of acceptability to regulators (0.25) and Effectiveness of technical solution and residual liability (0.25), reflecting the regulatory focus on compliance and long-term risk management. Under the traditional owners' perspective, Traditional owner acceptance (0.25) and Effectiveness of technical solution and residual liability (0.25) carry the highest weights, reflecting the importance of cultural relationships and confidence in long-term outcomes. The community perspective assigns its highest weight of 0.25 to Social and environmental value, reflecting community interest in broader social and environmental outcomes. The operations segment assigns its highest weight of 0.30 to Effectiveness of technical solution and residual liability and 0.20 to both Likelihood of acceptability to regulators and Operations and state agreement repudiation risk, reflecting the focus on deliverability and long-term management. The environment segment assigns 0.25 to both Social and Environmental Value and Effectiveness of technical solution and residual liability, reflecting the priority on ecological outcomes and long-term risk reduction.

This segmented approach means that the MCA is not evaluated from a single perspective but from multiple perspectives in parallel. The stochastic analysis then tests how robust each option is across these different viewpoints. This makes it possible to identify where stakeholder groups agree on a preferred option and where perspectives diverge, which is valuable information for closure planning and engagement.

5.4 Stochastic evaluation results

Figure 4 illustrates the 3 key components of the SMAA method using values from the conceptual case study. Figure 4a shows the score distributions for the 3 closure options against the Likelihood of acceptability to regulators criterion under the dry climate scenario. Option 3 Low liability has a mean score of 4.0, option 2 Risk remediation has a mean of 3.4 and option 1 Optimised without has a mean of 2.8, all with a standard deviation of 0.5. The overlap between these distributions is important because it shows that even though option 3 has the highest expected score, there are plausible futures where option 2 could score just as high or higher on this criterion.

Figure 4b shows the regulator weight profile across the 6 assessment criteria. Under this perspective, Likelihood of acceptability to regulators and Effectiveness of technical solution and residual liability each carry a weight of 0.25, reflecting the regulatory focus on compliance and long-term risk management. Social and Environmental Value and Operations and State Agreement Repudiation Risk each carry 0.15, while Maintain optionality for future use and Traditional owner acceptance each carry 0.10. These weights directly influence how the sampled scores are combined in each simulation.

Figure 4c shows the rank acceptability results for the regulator segment under the dry scenario based on 200 simulation iterations. Option 3 Low liability achieved rank 1 in 91% of simulated futures, meaning it was the best performing option under the vast majority of plausible score combinations. Option 2 Risk remediation achieved rank 1 in 9% of simulated futures and typically ranked second. Option 1 Optimised without did not achieve rank 1 in any simulation and ranked third in 97.5% of cases. Given the assumptions made in this analysis, these results suggest that the preference for option 3 is robust across a wide range of plausible conditions.

While the published SMAA literature recommends 10,000 or more iterations for final analysis (Lahdelma & Salminen 2010), the 200-iteration implementation used here is sufficient to demonstrate the method and produce stable rank orderings for the 3 option case study. The tool is designed to scale to higher iteration counts for real project applications.

In the case study, a minimum threshold of 3.0 is set for the Likelihood of acceptability to regulators criterion under the regulator segment. Any option that scores below this threshold in a given simulation is excluded from ranking for that iteration. This ensures that the stochastic evaluation focuses on outcomes that would be considered acceptable in practice and reflects how closure decisions are typically screened before detailed trade-off analysis. When an option is excluded by constraint filtering it is automatically assigned the lowest rank for that simulation, and the remaining options are ranked among themselves.

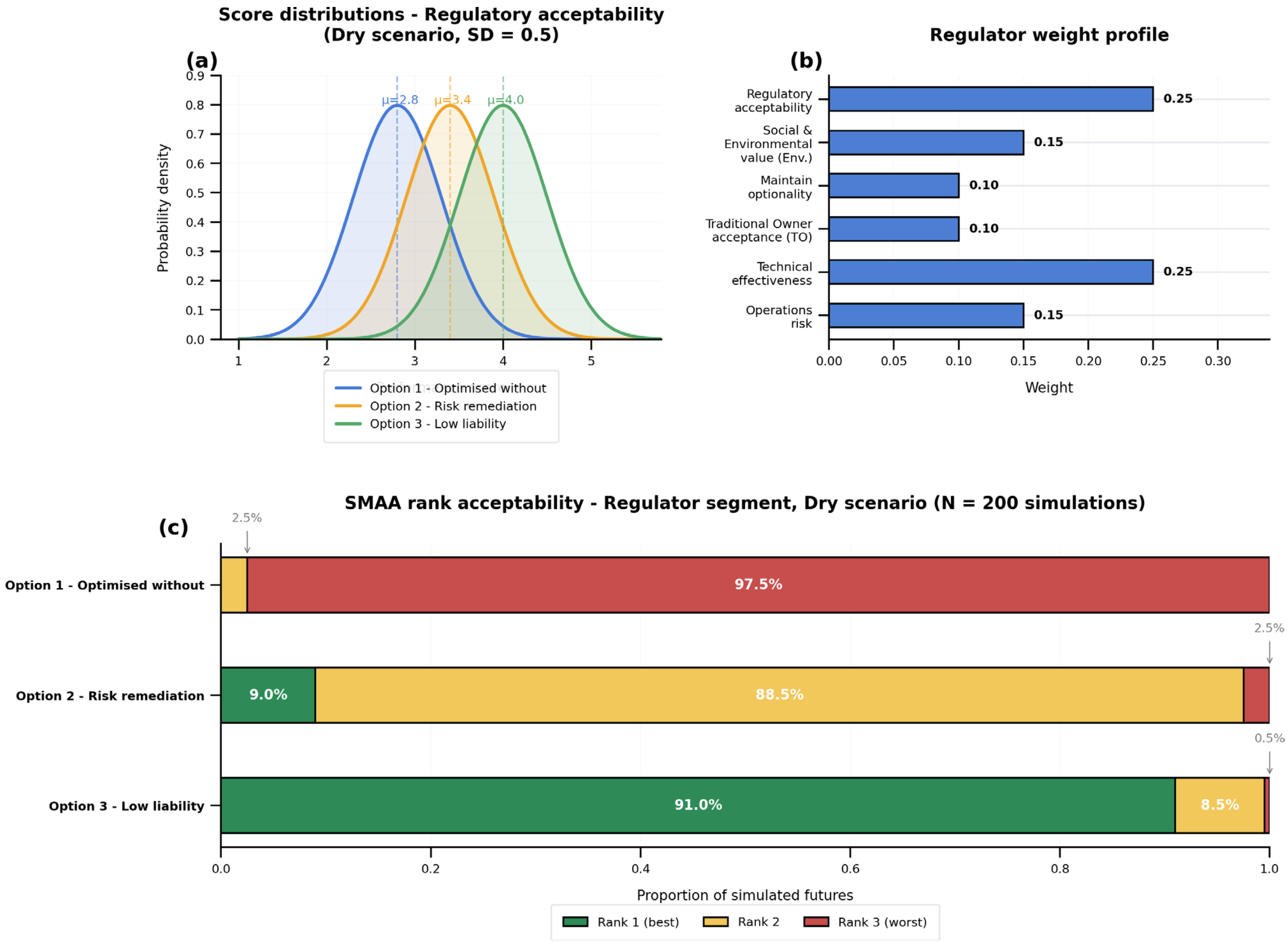

Figure 4 Stochastic multicriteria acceptability analysis (SMAA) methodology illustrated for the regulator segment under the dry climate scenario showing: (a) Overlapping performance score distributions demonstrating uncertainty in one criterion; (b) Criteria weight profile derived from preference learning; (c) Rank acceptability indices across 200 simulated futures

5.5 Scenario analysis results

In the conceptual case study, 3 climate scenarios are defined. The baseline scenario represents current expected conditions based on available site data and modelling. The dry scenario represents a drier climate sensitivity, reflecting conditions where reduced rainfall may affect revegetation success, water availability for treatment and long-term landform stability. The wet RCP8.5 scenario represents a wetter climate sensitivity aligned with a high-emissions climate pathway, reflecting conditions where increased rainfall may affect erosion, seepage rates and water management requirements.

For each scenario, the project team adjusts the mean performance scores for each option and criterion combination to reflect the expected change in performance. For example, under the dry scenario, option 1 Optimised without, scores are lower on Effectiveness of technical solution and residual liability (mean 1.8) compared to the baseline (mean 2.0), reflecting the increased risk associated with relying on long-term monitoring under drier conditions where revegetation and passive treatment may be less effective.

The SMAA simulation is then run separately for each scenario, producing rank acceptability outputs that show how robust each option is under that set of conditions. By comparing results across scenarios, the project team can identify whether the preferred option is robust across a range of climate futures or whether it is sensitive to specific climate assumptions. This is particularly valuable for closure decisions where the time horizon extends well beyond the reliability of any single climate projection.

Under the regulator perspective, the SMAA simulation was run separately for each scenario. Option 3 Low liability ranked first in 96% of simulated futures under the baseline scenario, 91% under the dry scenario and 95% under the wet RCP8.5 scenario. Option 2 Risk remediation achieved rank 1 in 4 to 9% of cases, depending on the scenario, and option 1 Optimised without did not rank first under any scenario. These results confirm that the preferred option is robust across all 3 climate futures and that the closure recommendation is not sensitive to climate uncertainty.

5.6 Criteria driver analysis

Figure 5 shows the Shapley value contributions for the regulator segment under the dry climate scenario, comparing option 3 Low liability against option 1 Optimised without in Figure 5a and against option 2 Risk remediation in Figure 5b.

In Figure 5a the total weighted score gap between options 3 and 1 is 1.11. Technical effectiveness is the single largest driver, contributing +0.50 to the gap. This reflects the large difference in performance scores (3.8 versus 1.8) combined with the high weight (0.25) assigned to this criterion under the regulator perspective. Regulatory acceptability is the second largest driver at +0.30. The only criterion that slightly favours option 1 is Maintain optionality (minus 0.05), reflecting that the minimum intervention approach preserves more flexibility for future use. However, this effect is small compared to the criteria that favour option 3.

In Figure 5b, the gap between options 3 and 2 is smaller at 0.50, which is expected because option 2 represents a higher level of intervention than option 1. Technical effectiveness remains the largest driver at +0.20, followed by Regulatory acceptability at +0.15 and Operations risk at +0.09. Again, Maintain optionality slightly favours option 2 (-0.04), but the effect is negligible relative to the other contributions.

This type of analysis provides clear and practical insights for closure decision-making. It shows that the preference for option 3 is primarily driven by its stronger performance on long-term residual risk management and regulatory acceptability, which are the 2 most heavily weighted criteria under the regulator perspective. If future information changed the expected performance of option 3 on Technical effectiveness, for example due to unexpected material behaviour or treatment failure, the ranking could shift. This is exactly the kind of threshold information that supports adaptive management and informs monitoring priorities.

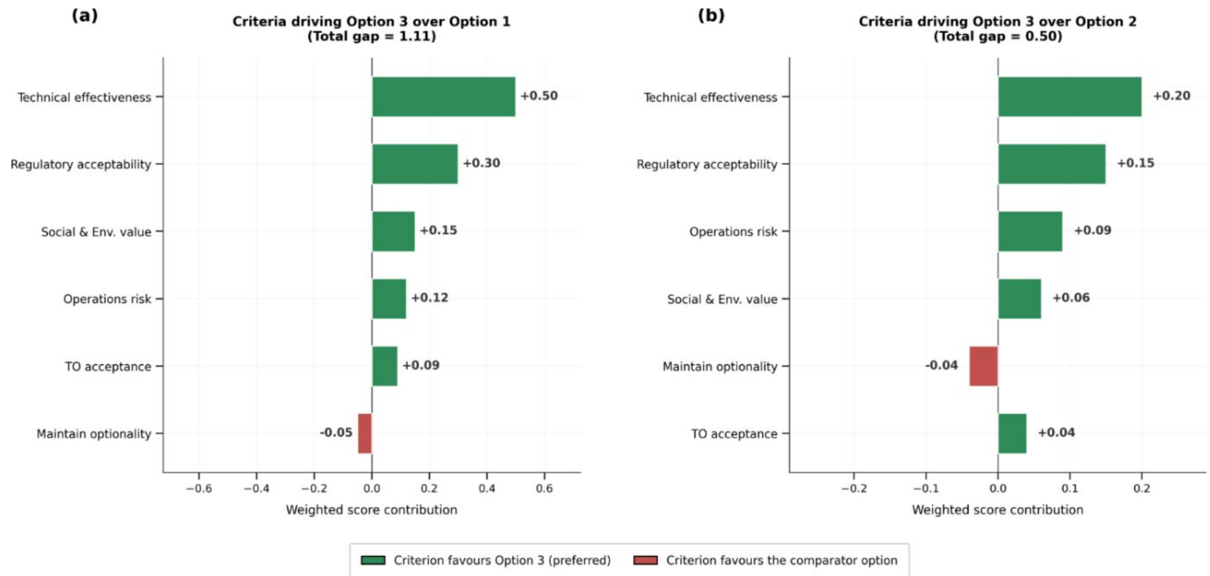

Figure 5 Shapley value decomposition for the regulator segment under the dry climate scenario showing: (a) Criteria contributions to the weighted score gap between options 3 and 1; (b) Criteria contributions to the gap between options 3 and 2

5.7 Multistakeholder perspective

The 5 stakeholder groups defined in Section 5.3 each assign different weights to the 6 criteria based on their priorities. The SMAA simulation was run for each group to test whether the preferred option changes when different perspectives are applied.

Despite these differences in priorities, the SMAA results show strong consensus across all 5 stakeholder groups under the dry scenario. Option 3 Low liability ranks first in 93 to 96% of simulated futures across every perspective. Option 2 Risk remediation achieves rank 1 in 4 to 7% of cases and option 1 Optimised without does not rank first under any stakeholder perspective.

This consensus is a significant finding for closure decision-making. It means the recommendation for option 3 is not sensitive to which stakeholder perspective is applied. Even when priorities differ substantially across groups, the overall preference remains the same. This provides a strong and defensible basis for the closure recommendation and simplifies stakeholder engagement by demonstrating that the preferred option performs well from multiple viewpoints. In more contested closure decisions where stakeholder groups disagree on the preferred option, the framework would provide a transparent basis for understanding why perspectives diverge and which criteria drive the disagreement. This information supports constructive dialogue and helps the project team work towards an agreed path forward.

Similar consensus was observed under the baseline and wet RCP8.5 scenarios, with option 3 ranking first in over 90% of simulated futures across all stakeholder groups and scenarios.

5.8 Monitoring triggers

The following examples illustrate how the adaptive monitoring framework described in Section 4.6 would be applied in the case study context.

If satellite-based ground movement monitoring detects velocity greater than 2 millimetres per week sustained over 3 weeks, the trigger action is to deploy engineering controls and reassess the Effectiveness of technical solution and residual liability criterion. If water quality monitoring detects pH below 6.5 in 2 consecutive samples, the trigger action is to adjust treatment or add alkalinity, with expected positive effects on Regulatory acceptability and Social and Environmental Value criteria. If vegetation survival monitoring shows less than 60% survival at year 2, the trigger action is to amend soil and reseed, with expected positive effects on Social and Environmental Value and long-term residual liability.

6 Conclusion

This paper has described an artificial intelligence- (AI) enhanced multicriteria analysis (MCA) workflow that strengthens established closure option assessment practice by introducing preference learning, stochastic evaluation, explainability and adaptive monitoring. The approach retains the familiar structure of traditional MCA and keeps criteria definition, option development, scoring logic, constraints and decision accountability as human-led activities. The enhancement lies in how stakeholder preferences, uncertainty and sensitivity are represented and examined, and in how repeatable and transparent the assessment becomes.

Preference learning captures diverse stakeholder perspectives through practical pairwise comparisons rather than collapsing them into a single consensus weight set. Stochastic multi criteria acceptability analysis evaluates option robustness by testing performance across many plausible futures rather than relying on a single deterministic ranking. Shapley value-based explainability shows which criteria drive the results and where trade-offs are most significant. Adaptive monitoring triggers link the assessment to real site performance and prompt reassessment when conditions change.

The conceptual case study demonstrated that the preferred option ranked first in 91% of simulated futures under the regulator perspective and dry climate scenario, and it maintained this preference across all 3 climate scenarios, with rank 1 probabilities ranging from 91 to 96%. When tested across 5 stakeholder groups with substantially different weight profiles, the preferred option ranked first in 93 to 96% of simulated futures under every perspective, demonstrating strong consensus. The Shapley analysis showed that Technical effectiveness and Regulatory acceptability were the 2 criteria most strongly driving the preference, providing clear and communicable insight into why the recommendation was made.

The case study presented here is conceptual and uses a simplified 3 option, 6 criteria structure with assumed performance scores and stakeholder weight profiles. In real closure applications, the number of options, criteria, stakeholders and scenarios would typically be larger, and the input data would be informed by site-specific studies, monitoring data and structured stakeholder engagement. Future work should focus on applying the workflow to real closure decisions with larger option sets and more contested ranking outcomes where stakeholder groups disagree on the preferred option. Applying the framework to cases with closer ranked options and stronger trade-offs between criteria would further demonstrate the practical value of the stochastic and explainability components.

The approach is implemented using Excel supported by AI-based computation and reporting assistance. In this implementation, the Excel tool handles data input, simulation logic and results presentation, while AI assistance supports automating repetitive calculations, generating diagnostic reports and interpreting sensitivity outputs through natural language interaction. It does not require specialised software, does not introduce new decision risk and does not remove professional judgement from the process. By making assumptions explicit and decision drivers visible, the AI-enhanced MCA approach strengthens robustness, transparency and defensibility while remaining practical and auditable. As mine closure decisions involve increasing uncertainty, regulatory scrutiny and long-term liability considerations, this approach provides a credible and scalable way to enhance closure option assessment while building on established good practice.

Acknowledgement

The author acknowledges the support of Simon Mollica, Head of BHP WA Asset Projects (WAAP), Scott Sinclair, WA Asset Projects Technical Manager, and John Poole, Discipline Engineering Area Lead, for supporting this work and providing the professional environment that enabled its development. Special thanks go to Abby Richards from the Closure Centre of Excellence for her detailed review and feedback on the framing of multicriteria analysis practice, and to Daniel Meunier from WAAP Project Engineering for his technical support and valuable input to the development and application of the multicriteria analysis (MCA) methodology presented in this paper. The author also acknowledges the contribution of stakeholders whose experience and perspectives on closure decision-making informed the conceptual case study.

References

Department of Mines, Industry Regulation and Safety 2020, *Statutory Guidelines for Mine Closure Plans*, version 3.0, Government of Western Australia, Perth, viewed 18 April 2026, https://www.wa.gov.au/government/publications/statutory-guidelines-mine-closure-plans

International Council on Mining and Metals 2019, *Integrated Mine Closure: Good Practice Guide*, 2nd edn, ICMM, London, viewed 18 April 2026, https://hub.icmm.com/website/publications/pdfs/environmental-stewardship/2019/guidance_integrated-mine-closure.pdf

Lahdelma, R, Hokkanen, J & Salminen, P 1998, 'SMAA: stochastic multiobjective acceptability analysis', *European Journal of Operational Research*, vol. 106, no. 1, pp. 137–143, https://doi.org/10.1016/S0377-2217(97)00163-X

Lahdelma, R & Salminen, P 2001, 'SMAA-2: stochastic multicriteria acceptability analysis for group decision making', *Operations Research*, vol. 49, no. 3, pp. 444–454, https://doi.org/10.1287/opre.49.3.444.11220

Lahdelma, R & Salminen, P 2010, 'Stochastic multicriteria acceptability analysis (SMAA)', in M Ehrgott, J R Figueira & S Greco (eds), *Trends in Multiple Criteria Decision Analysis*, Springer, New York, pp. 285–315, https://doi.org/10.1007/978-1-4419-5904-1_10

Saaty, T L 1980, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, McGraw Hill, New York.

Shapley, L S 1953, 'A value for n-person games', in HW Kuhn & AW Tucker (eds), *Contributions to the Theory of Games II*, Princeton University Press, Princeton, pp. 307–317, https://doi.org/10.1515/9781400881970-018

Tervonen, T & Figueira, J R 2008, 'A survey on stochastic multicriteria acceptability analysis methods', *Journal of Multi-Criteria Decision Analysis*, vol. 15, no. 1–2, pp. 1–14, https://doi.org/10.1002/mcda.407

Williams, B K & Brown, E D 2014, 'Adaptive management: from more talk to real action', *Environmental Management*, vol. 53, no. 2, pp. 465–479, https://doi.org/10.1007/s00267-013-0205-7