Integrated Environmental Assessment and Management — Volume 9, Number 4—pp. 652–664

Case Study

Use of Multi-Criteria Decision Analysis in Regulatory Alternatives Analysis: A Case Study of Lead Free Solder

Timothy F Malloy,†‡ Peter J Sinsheimer,‡ Ann Blake,§ and Igor Linkov‖#*

†UCLA School of Law, Sustainable Technology and Policy Program, Los Angeles, California, USA
‡UCLA School of Public Health, Sustainable Technology and Policy Program, 21-293 Center for Health Sciences, Los Angeles, California, USA
§Environmental and Public Health Consulting, Alameda, California, USA
‖US Army Engineer Research and Development Center, US Army Corps of Engineers, Concord, Massachusetts, USA
#Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

(Submitted 29 June 2012; Returned for Revision 1 October 2012; Accepted 7 May 2013)

ABSTRACT

Regulators are implementing new programs that require manufacturers of products containing certain chemicals of concern to identify, evaluate, and adopt viable, safer alternatives. Such programs raise the difficult question for policymakers and regulated businesses of which alternatives are "viable" and "safer." To address that question, these programs use "alternatives analysis," an emerging methodology that integrates issues of human health and environmental effects with technical feasibility and economic impact. Despite the central role that alternatives analysis plays in these programs, the methodology itself is neither well-developed nor tailored to application in regulatory settings. This study uses the case of Pb-based bar solder and its non-Pb-based alternatives to examine the application of 2 multi-criteria decision analysis (MCDA) methods to alternatives analysis: multi-attribute utility analysis and outranking. The article develops and evaluates an alternatives analysis methodology and supporting decision-analysis software for use in a regulatory context, using weighting of the relevant decision criteria generated from a stakeholder elicitation process. The analysis produced complete rankings of the alternatives, including identification of the relative contribution to the ranking of each of the highest level decision criteria such as human health impacts, technical feasibility, and economic feasibility. It also examined the effect of variation in data conventions, weighting, and decision frameworks on the outcome. The results indicate that MCDA can play a critical role in emerging prevention-based regulatory programs. Multi-criteria decision analysis methods offer a means for transparent, objective, and rigorous analysis of products and processes, providing regulators and stakeholders with a common baseline understanding of the relative performance of alternatives and the trade-offs they present. *Integr Environ Assess Manag* 2013;9:652–664. © 2013 SETAC

Keywords: Alternatives analysis   Alternatives assessment   Chemicals   Multi-criteria decision analysis   Regulation

INTRODUCTION

Chemical regulation in its various forms relies primarily on a risk management paradigm in which use of a chemical is permitted so long as exposures are kept below acceptable levels. Acceptable exposure levels are based on a variety of standards. Some rely largely on the performance of best available control technologies, others are based more heavily on health concerns and risk assessment. The risk management paradigm has been subject to criticism on a variety of grounds (Leadership Council 2011). A different paradigm has begun to emerge in Europe and some states in the United States. Rather than asking what level of exposure to the subject chemical is acceptable, this prevention-based paradigm asks whether there are viable alternative chemicals that are safer.

Today in California and Maine, regulators are beginning to implement new programs that require manufacturers of products containing certain chemicals of concern to identify, evaluate and adopt potential safer alternatives (Assemb. B. 1879 2008; Maine 2011). In the European Union, the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) program imposes similar obligations on certain particularly dangerous listed chemicals (REACH 2006). Although such prevention-based programs ostensibly avoid the thorny issues associated with establishing acceptable exposure and risk levels, they raise a potentially more difficult question for policymakers and regulated businesses: which alternatives are "viable" and "safer"?

To address that question, prevention-based programs use "alternatives analysis," a newly emerging methodology that integrates issues of human health and environmental effects with technical feasibility and economic impact. Alternatives analysis is a scientific method for prioritizing different courses of action—in this case for determining the viability of safer substitutes for existing products or processes that use hazardous substances. Alternatives may include drop-in chemical substitutes, material substitutes, changes to manufacturing operations, and changes to component or product design (Sinsheimer et al. 2007). The methodology compares the alternatives to the regulated product and to one another across a variety of attributes, typically including public health, environmental, technical, and economic. It is a value-based balancing of the respective attributes (e.g., lower toxicity vs higher cost) of the regulated product and of the alternatives. Its goal is to

All Supplemental Data may be found in the online version of this article.
* To whom correspondence may be addressed: malloy@law.ucla.edu
Published online 23 May 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/ieam.1449

essentially rank the regulated product and alternatives in the relative order of how well each option fits the decision criteria guiding the evaluator.

Despite the central role that alternatives analysis plays in prevention-based regulation, the methodology itself is neither well-developed nor tailored to application in regulatory settings. The California and Maine statutes and the REACH regulation provide little guidance. To a limited degree, businesses, researchers, and nongovernmental organizations (NGOs) have developed and implemented various forms of alternatives analysis (Zhou and Schoenung 2008; Kuczenski and Geyer 2010). However, these efforts were not undertaken in a regulatory context in which the methodology must be applied consistently and transparently across a range of settings (Koch and Ashford 2006). In the regulatory setting, agencies have engaged in the comparison of chemical alternatives to limited degrees. The European Chemical Agency has generated guidance for alternative analysis of certain regulated chemicals under REACH, but those methods stop short of providing an integrated decision platform and have not yet been widely implemented (ECHA 2007).

Multi-criteria decision analysis (MCDA) has proven useful in many prioritization and environmental decision-making contexts (Pohekar and Ramachandran 2004; Huang et al. 2011; Linkov and Moberg 2012). Multi-criteria decision analysis consists of a variety of methods and approaches, all of which offer a systematic approach for evaluating and ranking alternatives (Giove et al. 2009). The various MCDA methods handle the data and rank the alternatives in different ways based on disparate theoretical perspectives (Kiker et al. 2005). This article demonstrates the use of 2 different MCDA methods as part of alternatives analysis in a regulatory context. A linear multi-attribute value model (Belton and Stewart 2002) was used to ensure that the analysis logically incorporated tradeoffs among the wide range of alternative chemicals and considerations at the multiple levels. For comparison, an outranking model was also used (because of the limited scope of the project, other MCDA methods such analytic hierarchy process were not examined). The application of the 2 MCDA methods as part of a regulatory alternative analysis is illustrated in a case study of Pb-based solder used in consumer electronics and its non-Pb-based alternatives.

The case study uses the framework of the California statute and informal draft regulations issued under that statute as the regulatory setting. The case of Pb solder used in consumer electronics was selected for the richness of the data set around both environmental and human health impacts, functional performance of the various alternatives, and the life cycle approach taken to the analysis. Our primary source for the data set was the Life Cycle Assessment produced by the US Environmental Protection Agency (USEPA) as part of its Design for the Environment (DfE) program; the methodology and data are described in Geibig and Socolof (2005). Other published analyses identified in the Supplemental Data provided additional data regarding bar Pb solder and its alternatives (Lau et al. 2000; Hwang 2005; Handwerker et al. 2007).

METHODOLOGY

Alternatives analysis consists of 2 separate yet related components (Sinsheimer and Malloy 2009). The first, alternatives assessment, includes 1) identification of the key criteria for comparison (e.g., technical, health and safety, environmental, and economic attributes) and the metrics for measuring performance on those criteria, 2) identification of potentially viable alternatives, and 3) collection and compilation of data regarding performance of the regulated product and alternatives for each criterion. The second component of alternatives analysis is alternatives evaluation, largely conducted after the alternatives assessment is completed. It involves the development of weights for the relevant criteria and the balancing of the performance of the regulated chemical or product and the alternatives, respectively, across the criteria to develop a ranking. In this project, the balancing is facilitated through application of multi-attribute utility theory (MAUT) and outranking decision analysis support tools.

Alternatives assessment

Model building: criteria and metrics. We created a generic alternatives assessment model to provide a uniform method for comparing a regulated hazardous product to one or more alternatives to determine the safety and viability of these substitutes. As shown in Figure 1, the generic alternatives assessment model consists of 7 major comparison criteria: Physical/Chemical Hazards, Human Health Impact, Environmental Impact, Ecological Impact, Technical Feasibility, and Economic Feasibility. These 7 criteria were derived from a set of factors required under the California statute to be considered in performing alternatives analyses. The statutory factors were organized and consolidated based on existing alternatives assessment frameworks, the experience and expertise of project team members, and the draft regulations promulgated under the California statute. Each of the 7 major criteria have one or more levels of subcriteria. Some—such as Physical/Chemical Hazards and Technical Feasibility—Pb directly to measures having metrics against which alternatives can be scored. For example, one measure associated with Physical/Chemical Hazards is flashpoint, which can be quantified by means of specified tests. Others have intermediate subcriteria that in turn lead to measures. The criteria, subcriteria, and measures are integrated into a "performance matrix," which is populated with data specific to the case under study.

The Physical/Chemical Hazards criterion encompasses 5 measures: flammability, flashpoint, explosivity, auto-ignitability, and oxidizing properties. Each of these hazards is associated with fire and explosions, which can result in harm to both human health and ecological harm. Nonetheless, the human health or ecological harm caused by these physical chemical hazards does not result from direct exposure to the chemical but rather indirectly from fire or explosion. Accordingly, physical/chemical hazards were retained as a separate major criterion.

Human Health Impact links 2 subcriteria to human health impact—Toxicity and Human Exposure. Toxicity is linked to 8 measures typically addressed in alternatives assessment. Human Exposure is linked to 6 measures drawn from the California statute as well as from guidance issued by the European Chemicals Agency (ECHA) under REACH (ECHA 2010). Placement of Human Exposure as a subcriterion parallel to Toxicity was based on traditional notions of risk; namely, that the ultimate impact of chemical depends on both its inherent hazard and the level and nature of exposure. For example, in comparing 2 alternatives, a decision maker may be more concerned about a moderately toxic chemical with high exposures than about a more toxic chemical having extremely

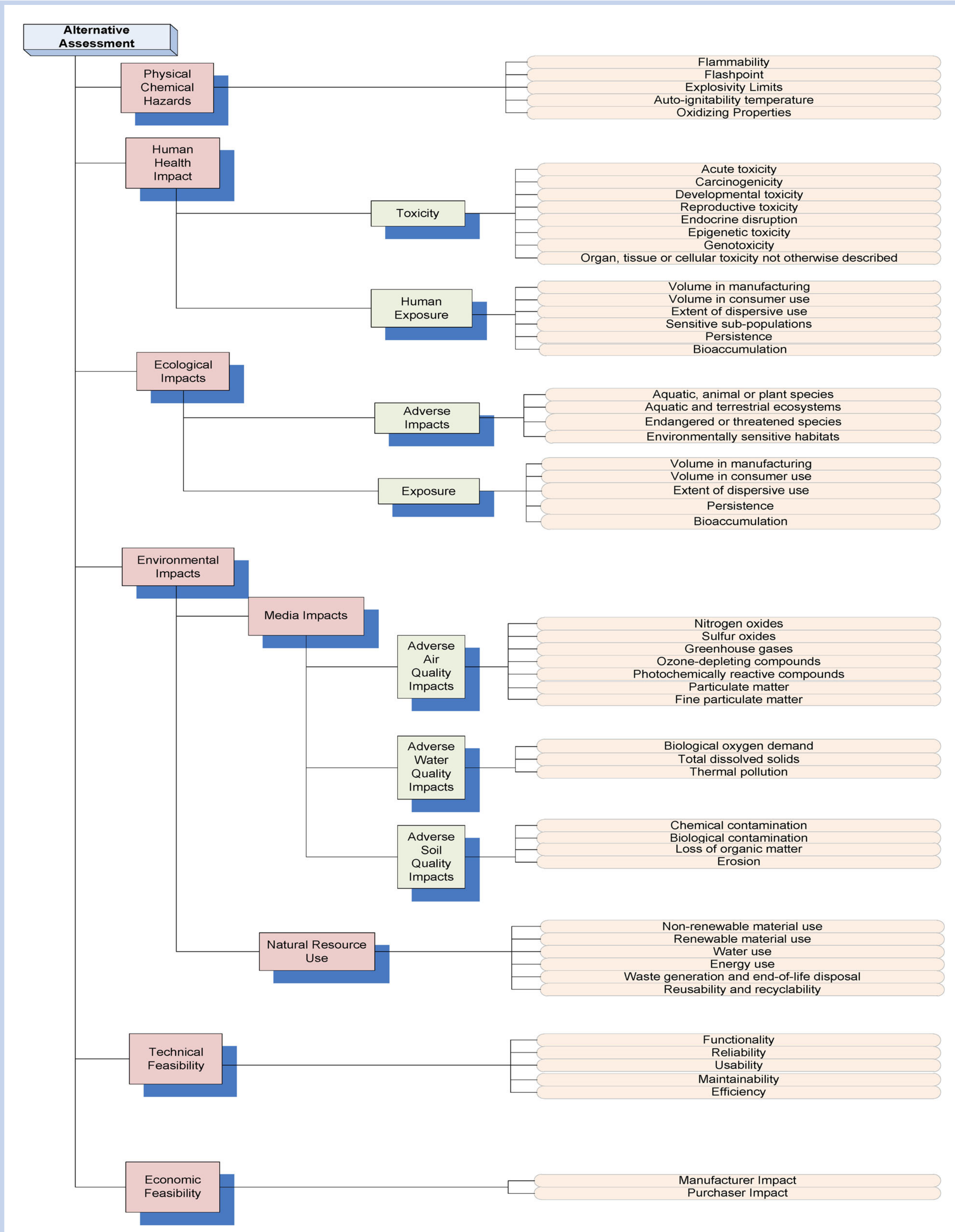

Figure 1. Generic alternatives assessment model.

low exposure. That said, this design treats the relationship between hazard and exposure very generally; it does not link a specific exposure concern (e.g., ingestion by an infant) to the hazard criterion with which it is associated (e.g., developmental toxicity). An alternative approach worthy of future consideration would link exposure concerns more directly to the specific hazards, perhaps by developing an integrated metric reflecting both.

For exposure measures, we used the volume of chemicals in manufacturing, volume of chemicals in consumer use, and extent of dispersive use as surrogate indicators of the likelihood of occupational, bystander, and consumer exposure (ECHA 2010). Persistence and bioaccumulation were selected as human exposure measures because an increase in either increases the likelihood and scope of exposure to humans. The potential exposure of sensitive subpopulations was placed under the broader Exposure criterion due to the increased concern evident in the statute regarding such populations.

The Ecological Impacts criterion focuses on impacts to animals, plants and the ecosystems in which they exist. It incorporates the same notions of hazard and exposure described for human health impacts, above, through 2 subcriteria: Adverse Impacts and Exposure. Adverse Impacts were linked to 4 measures: Aquatic, animal, or plant species; Aquatic and terrestrial ecosystems impacted; Endangered or threatened species; and Environmentally Sensitive Habitats. Exposure had 5 measures: volume of the chemical in manufacturing, volume in consumer use, extent of dispersive use, persistence, and bioaccumulation.

The generic model separates Environmental Impact from Ecological Impact to distinguish between negative impacts to animals, plants and the particular ecosystems in which they exist on the one hand, and other generalized adverse impacts on the conservation and quality of natural resources more broadly (issues of potential overlap are addressed below in *Discussion and Conclusions*). Environmental Impact encompassed 8 of 15 criteria specified in the California statute, broken into 2 distinct subcriteria: Natural Resource Use and Media Impacts. Media Impacts included Adverse Air Quality Impacts, Adverse Water Quality Impacts, and Adverse Soil Quality Impacts. Natural Resource Use was linked to 6 measures: nonrenewable material use, renewable material use, water use, energy use, waste generation and end-of-life disposal, and reusability and recyclability.

Technical Feasibility is evaluated using 5 measures: Functionality, Reliability, Usability, Maintainability, and Efficiency. These criteria were defined as follows: Functionality—a set of functions that satisfy the stated or implied need; Reliability—attributes that affect the capability of the product to maintain its level of performance; Usability—attributes that bear on the effort needed to use the product; Maintainability—the ability to identify and fix faults in the product; and Efficiency—attributes that bear on the resources needed to use the product.

The Economic Feasibility criterion was defined by 2 measurement criteria: Manufacturer Impact and Purchaser Impact. Manufacturer Impact refers to the extent to which expected revenues in selling a product are greater than expected costs of manufacturing the product, taking into account the manufacturer's ability to pass on increased costs to the consumer or to its suppliers. Purchaser Impact refers to the increased price paid by the consumer for the end product (ECHA 2007).

Identification of potential alternatives. In the context of regulatory alternatives analysis, the decision maker will almost always be comparing an existing product or industrial process to a set of potential alternatives, seeking to identify one or more viable, safer alternatives. In some cases, just a few potential alternatives may be available; in others, they may be many. In the latter case, the regulatory program must include a threshold screening step in which a tractable number of potential alternatives are identified for more in-depth assessment and evaluation. For example, under the REACH authorization program, analysts are essentially directed to use best professional judgment to identify alternatives for further consideration, taking into account functionality, technical and economic feasibility, and capacity to reduce overall risk (ECHA 2011). Other screening approaches might rely on a streamlined MCDA method or a set of minimum specifications (Belton and Stewart 2002). In any event, care should be taken to avoid eliminating alternatives too early in the process (Stewart and Losa 2003).

We did not engage in screening of potential alternatives in this case study, relying instead on the set of alternatives assessed in the DfE life cycle assessment. In that DfE process, the alternatives were selected by the project participants based largely on consideration of current trends and performance studies. In other words, the alternatives were viewed as technically viable substitutes with some existing market penetration (Geibig and Socolof 2005). Table 1 describes the alternatives considered in this case study.

Alternatives evaluation

As discussed above, the alternatives evaluation component of alternatives analysis (analogous to "model application" in decision theory) includes the development of weights for the decision criteria, and the evaluation of the relative performance of the regulated product and alternatives regarding those criteria. This section describes the methods the project used in performing those 2 steps, beginning with weighting.

Weighting and stakeholder elicitation. Generally speaking, criteria weights can be established in 1 of 3 ways: use of existing generic weights, calculation of weights using objective criteria, or elicitation of weights from experts or stakeholders (Zhou and Schoenung 2007). Generic weighting can be as simple as applying equal weights to all criteria, or relying on

Table 1. Bar solder alternatives

Alternative	Description
Sn/Pb (baseline regulated product)	Solder alloy composed of 63% Sn/37% Pb
Sn/Ag/Cu (water quenched)	Solder alloy composed of 95.5% Sn/3.9% Ag/0.6% Cu; water quenching is used to cool and harden solder
Sn/Ag/Cu (air cooled)	Solder alloy composed of 95.5% Sn/3.9% Ag/0.6% Cu; air is used to cool and harden solder
Sn/Cu (water quenched)	Solder alloy composed of 99.2% Sn/0.8% Cu; water quenching is used to cool and harden solder
SnCu (air quenched)	Solder alloy composed of 99.2% Sn/0.8% Cu; air is used to cool and harden solder

publically available sets of criteria weights developed by third parties through calculation or elicitation. For example, the National Institute of Standards and Technology (NIST) life cycle impact assessment software for the selection of environmentally preferred building product provides 3 generic weighting sets from an NIST-convened panel, the Environmental Protection Agency (EPA) Science Advisory Board (SAB) and a Harvard University study, respectively (Gloria et al. 2007). We chose not to use existing generic weights for 2 reasons. First, neither the NIST impact categories nor those of other existing weighting schemes map well onto the criteria used in our generic alternatives assessment model. Second, the composition of the panels used by NIST does not reflect the diversity of opinion we were seeking.

Weights can also be calculated using objective measures. For example, in distance-to-target methods, all criteria are initially assumed to be of equal importance, and then each is weighted by reference to the magnitude of variance between the desired conditions and existing conditions for each criterion (Soares et al. 2006; Zhou and Schoenung 2007). Distance-to-target methods have been subject to significant criticism when used alone, primarily because of the underlying assumption that, but for the level of variance from the desired condition, all criteria were of equal concern (Finnveden et al. 2009). Another calculation approach is monetary evaluation, in which weighting is made on the basis of costs related to environmental consequences. Although the specific valuations methods vary, they generally rely on some objective calculation of the costs of responding to or avoiding the criterion's impact. The more significant the monetary value of an assessment criterion, the greater significance it will take on (Soares et al. 2006). We rejected the use of these 2 objective methods because of their limited focus to variance from target and costs, respectively, and their complexity given the limited scope of this project.

The third major approach, the elicitation of weights from experts or stakeholders directly, includes a wide variety of methods, such as public opinion surveys, facilitated group consensus-based procedures (e.g., the modified Delphi technique), and various weighting procedures used in multi-criteria decision analysis models. Given the limited scope of this pilot project, we did not use the resource intensive survey or facilitated consensus-based procedures. Multi-criteria decision analysis approaches typically use 1 of 3 weighting methods: direct rating, pairwise comparison (used in analytical hierarchy process [AHP] and in outranking methods such as PROMETHEE), and "swing weighting" (typically used with MAUT methods) (Bottomley and Doyle 2001; Belton and Stewart 2002; Linkov and Moberg 2012). We chose direct rating because we needed a simple and transparent weighting approach that could be used for both outranking and MAUT methods. The direct rating approach we used—Max100—is relatively straightforward to apply, and an empirical evaluation of the method demonstrated it to be reliable and preferred by interview subjects (Bottomley and Doyle 2001).

In light of the narrow scope of this pilot study, the stakeholder elicitation was not intended to develop weightings that were statistically representative of the respective stakeholder groups. Rather, it was designed as an initial exploration of differences among various stakeholder groups regarding weighting, and of the impact of any such differences on the rank order of alternatives. Four stakeholder groups were considered: Environmental Nongovernmental Organizations, Industry, Policymakers, and Consumers. We created a pool of interview subjects by reviewing the rule-making docket for the California regulations and considering participation (as members and as commentors) in meetings of California's Green Ribbon Science Panel, an interdisciplinary advisory panel created by the California statute. We selected 3 subjects from each of the 4 stakeholder groups and conducted an individual interview with each subject lasting approximately 1.5 h.

The interviewer provided the subject a copy of the conceptual model (see Figure 1) showing the criteria and subcriteria and implemented the following procedure for each level of criteria and subcriteria in sequence. For each level, the subject received a set of cards (in random order), with each card containing the name of a criteria and a brief definition and example. The subject then organized the cards of criteria by placing them in order, with the most important criteria on the top to the least important criteria on the bottom. If the subject felt that specific criteria should have equal ranking, they paper-clipped those cards together.

Next, the subject indicated the relative importance of the criteria by rating them along a 100 point scale. The interviewer presented the subject with a 100 point scale, noting that the subject's most important criterion (i.e., the first card) would be placed at the 100 mark. Beginning with the subject's most important criterion, the subject then placed each of the criteria along the scale. The subject repeated the card sorting and scaling procedures for each level of subcriteria until all levels were completed. For each group we used the criteria weights obtained from the interviews to develop the set of average criteria weights using the following equation:

$$w_{i,\text{average}} = \frac{\sum_{j=1}^{m} w_{i,j}}{m}$$

where m is the number of interviewees and i is the criterion index. Normalization was done using the following equation:

$$w_{i,norm.by100} = \frac{w_{i,\text{average}}}{\text{MAX}(W_{\text{average}})} \times 100.$$

Evaluation using MCDA as a decision-aid tool. This project focused on 2 commonly used MCDA methods: multi-attribute utility theory (MAUT) and outranking. We used the software package DECERNS for both MAUT and for outranking. With respect to outranking, DECERNS uses the outranking method known as preference ranking organization method for enrichment evaluations (PROMETHEE) (Brans et al. 1986; Linkov and Moberg 2012). The discussion that follows provides more detail regarding these 2 methods. Although this project compares the application of these 2 methods, selection of one recommended MCDA method in the context of regulatory alternatives analysis is beyond the scope of this project. For discussion of some factors relevant to method selection, see Guitoni and Martel (1998) and Seager (2004).

MAUT is an optimization approach, meaning that it represents the decision maker's preferences as utility functions, and attempts to maximize the decision maker's overall utility. MAUT is premised on the assumption that the decision maker has a fairly well-defined set of preferences that can be represented on a dimensionless utility scale. It also assumes that the decision maker is rational (i.e., they prefer more utility rather than less) and is consistent in those preferences. Last, it assumes that preferences are stable and transitive (that is, if the decision maker prefers alternative A to alternative B, and alternative B to alternative C, then they will prefer alternative A to alternative C) (Linkov et al. 2004; Gass 2005). In the context of this project, therefore, we generated a utility function for each criterion, which reflects how a decision maker's preference changes for different values of that criterion. This utility function spans from 0 to 1, with a utility of 1 being assigned to the value of the best (or highest) alternative score for that criterion and 0 being assigned to the value of the worst (or lowest) alternative score. In this case, a linear utility function was used, which assumes that increases in utility are directly related to increases in the alternative's score for the criterion in question. We used the linear utility function as a default; in some scenarios below we explore the use of different utility functions that reflect more realistic preferences under some circumstances.

Having converted an alternative's performance on a particular criterion to a score on a 0 to 1 utility scale, we determined the weighted score for that criterion by multiplying the score by the weight assigned to that criterion. Thus, for example, if an alternative performed better than all others on acute toxicity, it would receive a score of 1. If acute toxicity had an overall weight of 0.2, the alternatives score on that criterion would be 0.2. The total score for the alternative is simply the sum of all weighted scores received for all criteria by the alternative. Accordingly, if an alternative performed the best for all criteria, it would receive an aggregate score of 1. Because alternatives that are deficient with respect to 1 criteria can compensate by their good performance with respect to other criteria, MAUT is a "compensatory" method.

Outranking models are not premised on expected utility theory, and thus do not create utility functions. Instead, outranking approaches are based on the principle that alternatives may exhibit differing degrees of dominance over one another (Linkov et al. 2004). Accordingly, outranking approaches directly compare the performance of 2 alternatives at a time, in terms of each criterion, to identify the extent to which one alternative out performs the other. Two types of thresholds are used to construct the outranking relationships by defining preferences with respect to a single criterion. The indifference threshold defines the difference in a criterion that is deemed insignificant. The preference threshold is the smallest difference that would change the expert preference (Brans et al. 1986). Between these 2 lay a zone of "hesitation" or indifference. The outranking approach then aggregates that information for all possible pairings to rank the alternatives based on overall performance on all criteria. Generally speaking, the PROMETHEE method used in the project creates a "preference index" for each alternative, which is calculated by reference to the alternative's positive flow (i.e., those instances in which the alternative outperforms another alternative on a given criterion) and negative flow (i.e., those instances in which the alternative is outperformed by another alternative). The value awarded for winning a particular pairing is weighted, meaning it is adjusted to reflect the value placed on that criterion by the decision-maker. Thus, outperforming another alternative in a minor criterion is worth less than outperforming it with respect to a more highly weighted criterion.

As a default in PROMETHEE and most other outranking methods, any difference in performance—however small—will result in an increase in positive flow for the better performing alternative. As in MAUT, PROMETHEE recognizes that a decision-maker may be indifferent to how alternatives perform on certain criteria until certain levels are met or after certain levels are exceeded. For example, taking the example again of acute toxicity, regulators may be indifferent to differences in LD50 levels above a certain level. Likewise, with respect to economic impact—particularly for materials such as solder that make up a small portion of total production costs—decision-makers may be indifferent to cost impacts until they exceed a certain threshold. PROMETHEE allows the decision-maker to incorporate such considerations into the generation of the preference index through a variety of preference functions (Belton and Stewart 2002).

Because outranking techniques aggregate the results of pairings for all criteria, they allow superior performance on some criteria to compensate for inferior performance on other criteria. However, they do not necessarily reflect the magnitude of relative underperformance in a criterion versus the magnitude of overperformance in another criterion, and vice versa. In other words, if Alternative A is marginally worse than Alternative B in 1 criterion, but substantially better with respect to another, outranking may not fully "compensate" Alternative A for its overall better performance. Therefore, outranking models are known as "partially compensatory." Whether a fully compensatory, partially compensatory, or noncompensatory approach should be used is addressed in the Discussion section below.

RESULTS

Alternatives assessment

The primary data source for populating the baseline performance matrix was an August 2005 EPA document entitled "Solders in Electronics: A Life-Cycle Assessment" that used life cycle impact assessment (LCIA) methodology to generate point estimates for a number of human health, ecological, and environmental impacts comparing SnPb solder with a range of non-Pb alternatives (Geibig and Socolof 2005). For human health impacts, the LCIA method used in this study focused on 2 population groups—occupational and the broader public—as well as 2 impact categories—carcinogenicity and chronic noncancer effects. Chronic noncancer effects included reproductive toxicity, developmental toxicity, neurotoxicity, immunotoxicity, behavioral effects sensitization, radiation effects, and chronic effects to other specific organs or body systems (Geibig and Socolof 2005). The study did not generate impact scores for each individual noncancer impact, providing only a single score for the group as a whole. In addition, occupational scores were based on all hazardous chemicals used throughout the respective product life cycles of the Pb solder and its alternatives (Geibig and Socolof 2005).

The EPA study on solder in electronics also provided data on technical performance and economic impact, but these data were extremely limited, and additional sources listed in the supplemental materials fill some data gaps. In addition, whereas the human health, ecological, and environmental effects were compiled over the life cycle of the product, performance and economic impacts were limited to the production of the printed circuit board.

Generally speaking, SnCu outperformed both SnPb and SnAgCu on human health impacts and ecological hazards, with SnPb performing the worst on most measures. In terms of environmental impacts, SnPb performed better than the alternatives with respect to impacts on environmental media whereas SnCu and SnPb performed best concerning natural resource use. Overall SnAgCu outperformed SnPb and SnCu on technical feasibility. Regarding economic feasibility, SnPb was the least expensive in terms of solder cost. However, given the small contribution that solder cost makes to the overall cost of production of the electronic component, a manufacturer using any of the solders would achieve a positive return on equity.

The complete baseline performance matrix for the bar solder case study is set out in the supplemental materials. The matrix sets out the metrics used for each measure, as well as the relevant sources and notes. For 35 of the 75 measures, there was no data for any of the alternatives. For those 35 measures, we used identical default data points for the SnPb bar solder and all of the alternatives. Performance matrices reflecting other data scenarios discussed below are available from the authors.

Alternatives evaluation

This section presents the results of the alternatives evaluation, beginning with presentation of the weighting regimes derived from the stakeholder elicitation. Next, it describes the results from the evaluation of the "baseline" performance matrix for the case study. It then examines the outcomes from a series of variations from that baseline; namely, variations of certain data assumptions, of weighting, and of the decision-making model itself.

Weighting and stakeholder elicitation. The average weights derived for the first level criteria for each of the 4 groups and for all interviewees are set out in Table 2.

As Table 2 demonstrates, at least at this level, there were not substantial differences across the groups. On average, all stakeholder groups (except for Industry) placed more weight on human health than on ecological hazards and environmental impact criteria. Industry and Policymakers assigned more weight to technical feasibility as compared to consumer and environmental NGOs. Industry placed more weight on economic feasibility than the other 3 groups. As discussed above, however, the sample sizes for the stakeholder groups were quite small (3 in each group), with the goal of getting a sense of the potential differences across and within groups. Except as specifically identified below, the evaluation uses the overall average weights derived from the stakeholder solicitation.

Baseline scenario. One primary goal of the case study is to demonstrate and examine the operation of 2 MCDA techniques: MAUT and outranking. The baseline scenario compares the performance of MAUT and outranking in the context of the baseline performance matrix. Table 3 displays the performance of Sn/Pb solder and each alternative under MAUT and outranking, respectively. Complete agreement between MAUT and outranking in the Pb solder case was lacking. Under MAUT, the air-cooled SnCu solder and water-cooled SnCu solder had the highest scores, 0.7878 and 0.7673, respectively (a higher score reflects a better overall performance). They were followed by SnPb solder, water-cooled SnAgCu, and air-cooled SnAgCu in that order. However, the SnPb and SnAgCu solders were quite close in scores.

The order was somewhat different under outranking. As in MAUT, the 2 forms of SnCu solder were the best performers in the outranking analysis, as indicated by their high net flows. Unlike MAUT, in outranking SnPb solder took the last position behind both forms of SnAgCu solder. However, the differences between SnPb, SnAgCu (air cooled) and SnAgCu (water cooled) were relatively small, and there was a noticeably larger gap in performance between this group of 3 on the one hand and the top 2 performers on the other.

The value of MCDA runs beyond simply generating an ordering of alternatives; the methods also enable decision-makers to understand the basis of the ordering. For example, Figure 2 breaks down the MAUT score for each alternative to indicate the relative contribution of each criterion to the overall score of the alternative.

Effect of variation in data conventions. These baseline results incorporate a number of decisions regarding data conventions, such as dealing with missing data or choosing between continuous versus categorical data. Each of these data conventions, taken separately or in combination, can be tested within MCDA to see if a different convention would make a difference in the final rank order of alternatives. By way of example, we focused on the treatment of missing data (i.e., measures for which there is no data for any of the alternatives). The baseline inserted identical default values for missing data for each of the alternatives. To identify the impact of missing data on the outcome, we removed all criteria for which there is missing data from the generic model.

In the MAUT analysis, the removal of those criteria had a pronounced effect on both scores and ranking of the 3rd-through 5th-ranked technologies. Although both forms of SnCu solder remained the best performers with relatively small changes in scores, Figure 3 illustrates that the ordering and scores of SnPb solder and the SnAgCu forms of solder changed substantially, making SnPb the lowest ranked alternative. Under the outranking analysis, although the scores for each alternative changed somewhat, criteria removal resulted in no change in ordering of the alternatives.

Table 2. Weighting by stakeholder group

	Environmental NGO (%)	Industry (%)	Consumer (%)	Policymaker (%)	Overall average (%)
Physical chemical hazards	15.22	11.04	15.21	13.12	13.75
Human health impact	21.14	18.07	20.28	24.75	20.83
Ecological hazards	18.60	18.67	19.68	18.07	18.75
Environmental impacts	18.60	20.08	19.68	14.11	18.33
Technical feasibility	14.38	16.47	11.56	16.58	14.58
Economic feasibility	12.05	15.66	13.59	13.37	13.75

NGO = nongovernmental organizations.

Table 3. Baseline outcome under MAUT and PROMETHEE

MAUT		Outranking	
Alternative	**MAUT score**	**Alternative**	**Net flow**
SnCu (air)	0.7878	SnCu (air)	0.09
SnCu (water)	0.7673	SnCu (water)	0.05
SnPb	0.7125	SnAgCu (water)	−0.03
SnAgCu (water)	0.7017	SnAgCu (air)	−0.04
SnAgCu (air)	0.6997	SnPb	−0.07

MAUT = multi-attribute utility theory; PROMETHEE = preference ranking organization method for enrichment evaluations.

Effects of stakeholder weighting on ranking. This set of scenarios considers the impact of adjustments to criteria weights. As noted above, in our baseline and other scenarios we used the average weights derived from the stakeholder elicitation process. In this set of scenarios, we systematically varied the weights to test the robustness of the outcomes under different stakeholder weighting regimes. Rather than using the average weighting from all individual stakeholders, in these scenarios we used the average weighting for each of the 4 stakeholder groups separately: Industry, Environmental NGO, Policymaker, and Consumer groups, respectively.

In MAUT, there was some small variation in scores across the different weighting regimes. The ordering of the alternatives was consistent across the regimes, however, with the exception of the policymaker weighting, which moved SnPb solder from the third position to the last position in terms of overall performance. Under outranking, none of the 4 weighting regimes affected the original ordering that resulted from the baseline scenario's use of average weights.

Sensitivity analysis. MCDA methods also allow for sensitivity analysis, a technique in which a variable is systematically modified to determine its impact on the outcome. In this case, we demonstrated the use of sensitivity analysis, using it to modify the weighting for technical feasibility, a criterion on which both SnAgCu forms of solder outperformed SnPb solder. With a baseline weighting of 14.5% for technical feasibility, SnPb solder is ranked above the SnAgCu solders in MAUT when all criteria are considered. As the weighting placed on technical feasibility (and thus its importance to the outcome) is increased, SnPb's performance vis-à-vis SnAgCu solders in MAUT deteriorates. Indeed, after the weighting for technical feasibility is increased by just over 4 percentage points to 18.7%, SnPb solder drops to the last position. Finally, when the weighting of technical feasibility rises to 24.1%, SnCu (water quenched) is displaced from 2nd position in the rank ordering by SnAgCu (water quenched).

Sensitivity analysis of changes to the weighting for technical feasibility was also carried out in the outranking approach with generally similar results. In the baseline scenario, SnPb solder

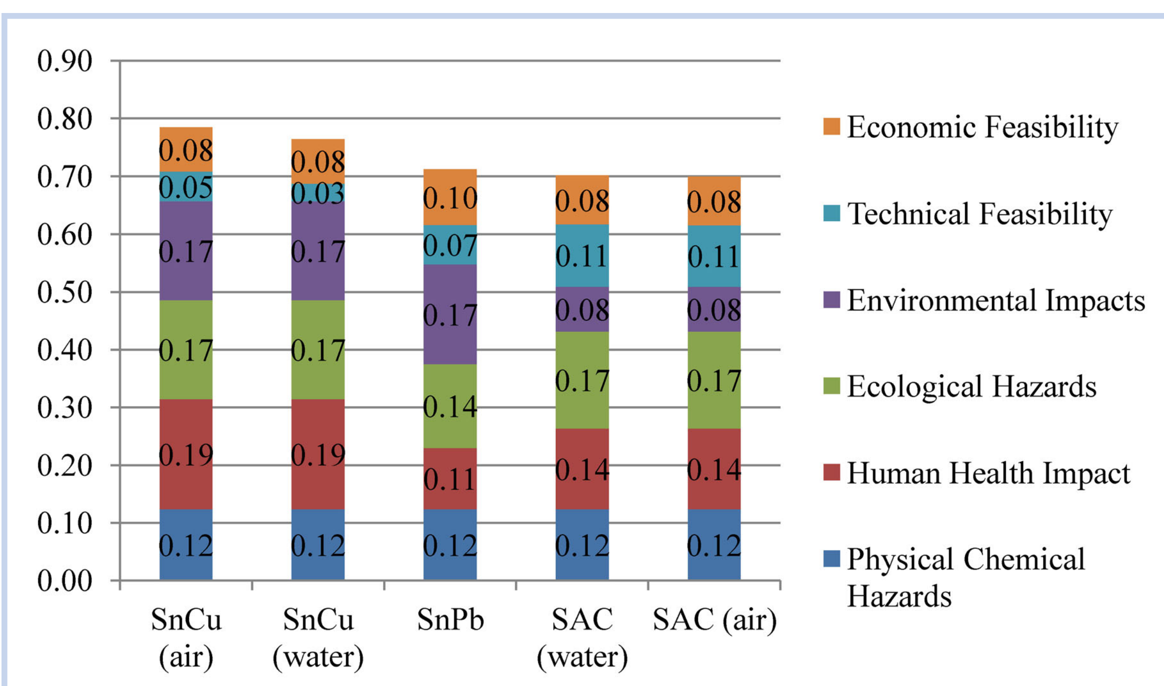

Figure 2. Criteria contribution to MAUT scores.

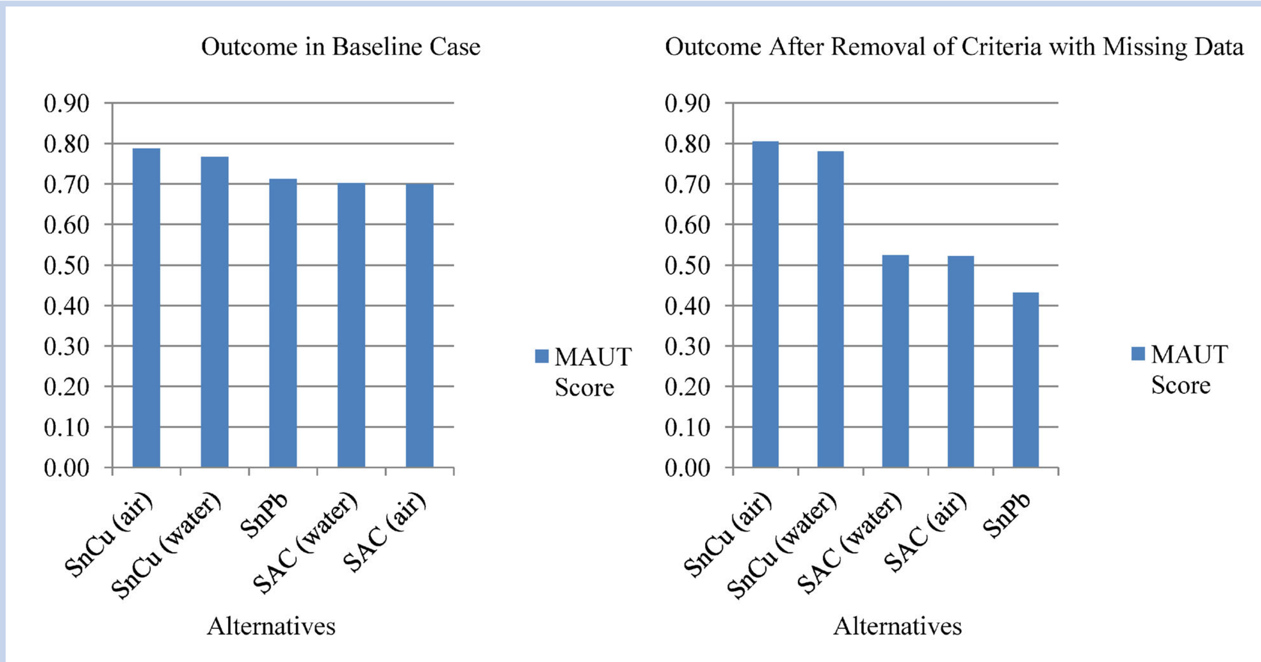

Figure 3. Missing data criteria removal under MAUT.

was ranked last, with the SnAgCu solders immediately above it. As in MAUT, increased weighting for technical feasibility led to deterioration of SnPb solder's performance vis-à-vis the SnAgCu solders. SnCu (water quenched) drops below SnAgCu (water quenched) to 3rd place in the ranking when the weighting for technical feasibility reaches 21.5%.

Decision model variations. The final set of scenarios tests the robustness of the outcomes under different design parameters, 2 relating to the broader decision framework and 1 relating to specific modeling assumptions regarding utility functions under MAUT. With respect to the decision framework, our basic assumption was that even though the various criteria may have differing importance to the decision-maker, all criteria would be considered in comparing the alternatives. In other words, no single criterion would operate as a threshold factor to screen out alternatives from further consideration. The first 2 scenarios of this set—the Sequential Decision Model and the Modified Sequential Model—alter that assumption, examining the outcomes if various sets of criteria are used as threshold screening factors. The remaining scenario in this set introduces a nonlinear utility function designed to reflect potential decision-maker preferences more realistically.

The Sequential Decision model scenario first evaluated the alternatives using just the Physical Chemical Hazards, and Human Health, Environmental, and Ecological impacts. It subsequently evaluated the top 3 alternatives (including PbSn solder if it is in the top 3) from that screening using the Technical Feasibility and the Economic Feasibility criteria. In the first part of the Sequential Decision model under MAUT, the top 3 performing alternatives regarding Physical/Chemical Hazards, Human Health, Environmental, and Ecological criteria were SnCu (water), SnCu (air), and SnPb in that order. Those 3 were carried on for further evaluation under the Technical Feasibility and the Economic Feasibility criteria, which resulted in identification of SnPb as the best performing alternative. The same rankings were obtained under outranking using the Sequential Decision model.

The Modified Sequential Decision model examines the outcome if Technical and Economic Feasibility were used as an initial screen, followed by consideration of the Physical Chemical Hazards, Human Health, Environmental and Ecological criteria. The top 3 performing alternatives concerning Technical and Economic Feasibility were SnAgCu (water), SnAgCu (air), and SnPb in that order. Those 3 alternatives were further evaluated under Physical/Chemical Hazards, Human Health, and Environmental, and Ecological criteria, resulting in the top ranking for SnAgCu (water). The same rankings were obtained under outranking using the Modified Sequential Decision model.

Table 4 illustrates that the insertion of a screening step using a subset of criteria in the Sequential Decision model and Modified Sequential Decision model resulted in significantly different rankings of the alternatives than that obtained in the baseline scenario. A decision maker such as a regulator who places substantially greater importance on health and environmental performance might rely on the Sequential Decision model to avoid trade-offs in which technical or economic performance compensate for poor performance on the latter criteria. Conversely, a decision maker with a strong interest in ensuring robust financial or technical performance might use the Modified Sequential Decision model to prevent other criteria from shielding poor performance in those areas.

The last scenario in this set modifies the default linear utility function to demonstrate the capacity of MAUT to account for different types of preferences. We evaluated the alternatives as in the baseline scenario, with one change. In this scenario, we modified the linear utility function for reproductive toxicity, developmental toxicity, endocrine disruption, genotoxicity, and other organ damage. Each of these measures were scored using a qualitative scale ranging from high(4) to medium(3) to low(2) to very low(1), with a score of "1" being the most desirable. Here we assumed that once the effect reached the

Table 4. MAUT outcome by decision framework

Baseline scenario		Sequential decision model		Modified sequential decision model	
Alternative	**Score**	**Alternative**	**Score**	**Alternative**	**Score**
SnCu (air)	0.79	SnPb	0.66	SnAgCu (water)	0.78
SnCu (water)	0.77	SnCu (air)	0.52	SnAgCu (air)	0.78
SnPb	0.71	SnCu (water)	0.45	SnPb	0.77

MAUT = multi-attribute utility theory.

relatively high level of 3 (i.e., an "unsafe" level), decision-makers would be indifferent to further increases in the level. Use of the modified utility function resulted in changes to each of the scores, but no change to the overall ordering of alternatives under MAUT.

DISCUSSION AND CONCLUSIONS

This case study had 2 primary goals with respect to alternatives analysis: 1) explore the development of a rigorous alternatives assessment methodology for use in a regulatory setting, and 2) examine whether MCDA methods may be appropriate for regulatory alternatives evaluation. Although the study used a regulatory framework provided by the California statute, the specific approaches and methods developed and applied as part of the study were not intended to be applied directly to alternatives analysis under that law. Rather they were designed to inform and enhance the development of regulatory alternatives analysis generally. This feasibility study developed a workable, comprehensive alternatives analysis model, and demonstrates the promise of MCDA as a robust method to assist in alternative analysis. Even though 2 MCDA methods were illustrated (MAUT and Outranking), other MCDA methods and tools can be implemented for alternatives analysis and would be useful in different regulatory settings. The use of a specific MCDA method should be selected based on specifics of the problem, time, and resource availability as well as analyst and decision maker experience and preferences as described below. That said, a number of steps are needed to take this from a feasibility stage to a generally applicable methodology that is rigorous enough to provide consistent results but flexible enough to adapt to the variety of products-of-concern likely to be regulated.

Alternatives assessment

Regarding alternatives assessment, the study created a generic alternatives assessment model consisting of broad upper level criteria each of which was defined by specific subcriteria and measurement criteria. Because MCDA tools are specifically designed to handle complex data sets, the generic alternatives assessment model was constrained by neither the number of criteria and subcriteria nor the form of the data—categorical, ordinal, continuous, or nominal.

The breadth of the criteria provides for a comprehensive evaluation of alternatives, significantly broader in reach than the set of indicators typically used in life cycle impact assessment. The study established generalized but measurable subcriteria for technical and economic feasibility, a significant enhancement to existing alternatives analysis methods currently available. However, the comprehensive nature of the criteria creates commensurately greater data collection and management efforts. The amount of missing data in the study's performance matrices highlights the need for meaningful data generation and collection elements in the regulatory program.

The generic alternatives assessment method presented here requires further refinement is a number of areas. In terms of model building, principles must be developed for determining which criteria to include in any particular analysis. In this case study, the generic alternatives assessment model included societal level environmental impacts, such as atmospheric ozone and greenhouse gas emissions, but did not include societal level economic impacts, such as sector-wide impacts on labor or costs imposed on government agencies. This issue has arisen in existing regulatory programs. California's statute appears to adopt a broad perspective, requiring that "economic impacts" be taken into account in the analysis of alternatives, whereas the guidance for European's REACH regulations reflects a narrower focus on the economic impact to individual firms. Similarly, development of specific principles for determining where to place criterion in an alternatives assessment model are needed. In this case, such decisions were based on the judgment and experience of the project team, but a more systematic approach should be used for regulatory purposes.

Further refinement of standards for dealing with overlaps between criteria is needed. Generally speaking, we allowed overlap where the same attribute in question exhibited distinct impacts in different areas. For example, energy use associated with an alternative could have both energy conservation impacts (primarily a societal concern) and energy cost impacts (primarily an individual facility concern). Conversely, where the 2 impacts associated with an attribute were directly linked, we selected 1 measure to capture those impacts. Take the case of chemicals that are listed as hazardous air pollutants. One could reach this through both a measure under air quality and a measure under human health impacts. In that case, we chose human health impacts so as to avoid double-counting the human hazards. The general decision rule we used here should be more fully articulated and tested in other circumstances.

Improvement in accounting for interaction of variables is also needed. The assumption used in the generic model was that criteria and subcriteria within the same level and at different levels were independent of one another. Further work is required to test this assumption across all categories, and to refine the model in cases in which there is some interaction. The hazard and exposure criteria highlight this concern. For many stakeholders, assessment of overall impact requires consideration of the specific hazard and the related exposure together, a perspective that our model and others do not fully capture. Possible approaches include the use of MCDA tools that allow for interaction among criteria, or the development of hybrid criteria that merge the related criteria and reflect their interaction.

Alternatives analysis should also account for differences in the quality of data. The generic model treats all data as essentially of the same quality. Existing models tend to highlight data quality issues without integrating those issues into the analysis. Further work is needed to integrate relative data quality into the model. For example, consider the use of continuous data. In theory this allows greater resolution of differences between alternatives than categorical or nominal classification. That said, even when a precise point estimate is available the quality of the data may relatively low and therefore one may consider using a more qualitative categorical or nominal (yes/no) classification to reflect the uncertainty regarding the point estimate. Other more formal approaches to uncertainty such as stochastic multi-criteria analyses may be appropriate in some circumstances (Linkov and Moberg 2012).

Alternatives evaluation

The study demonstrates the potential viability of MCDA methods to assist in the evaluation of complex alternatives assessment data. In particular, the case study shows that the MCDA models can provide decision makers and stakeholders with a transparent evaluation of such data. For these purposes, transparency means that interested parties should have access to the methods, information, assumptions, and data underlying the outcomes (Drew and Nyerges 2004). Transparency serves multiple purposes. It reflects normative views about the right of the public to be engaged, pragmatic interests of securing legitimacy for the ultimate outcome, and substantive beliefs that knowledgeable public engagement can improve the outcomes (National Research Council 2008). Transparency does not require that methods be simple enough that a lay person without specialized training could evaluate the methods or their application, but rather that engaged, knowledgeable participants could do so.

In this case, use of MCDA provides transparency at multiple levels. The systematic collection and management of the data, and the use of a formal method with clearly specified assumptions, provides enough information for a skilled analyst to be able to follow all the reasoning and replicate the results. Moreover, the types of reports generated from the analysis allow a broad range of interested parties to "follow the main arguments and understand the overall process of analysis and its conclusions" (National Research Council 2009). For example, both methods present a ranking of alternatives accompanied with explanations of how the alternatives' performance on various criteria affected that ordering. The methods also allow parties to understand how their weighting affects outcomes. The methods also permit decision makers to adjust the MCDA method's assumptions regarding the nature of their preferences. For instance, by altering the shape of utility functions, MAUT can capture situations in which a user is less concerned about a criterion where performance on that criterion is above or below certain ranges. Outranking methods are nonlinear and typically allows flexibility in incorporating user values by varying preference thresholds.

The study also provided useful insights on several aspects of alternatives evaluation. First, the stakeholder weighting process suggests that there was relatively little difference across groups with respect to the relative weight they placed on the criteria. This was driven home by the lack of substantial differences in the ranking of alternatives in either case study by the 4 groups. Of course, the elicitation process was limited in terms of sample size; nonetheless this outcome raises interesting questions regarding just how much disagreement there will be regarding weighting among the groups as the regulatory process moves forward.

Second, the results regarding missing data demonstrate that how a method handles missing data can significantly affect the outcome of the evaluation. In the Pb solder case, there was a substantial amount of missing data, and the ranking of alternatives differed significantly depending on whether we identical default values to fill in the missing data or simply removed those criteria for which there was no data. Conceptually, it is clear that inserting default data for all alternatives (such as a mid-point or worst-case value) will dilute the impact of other criteria for which there is data as compared to removing the criteria for which data is missing. For compensatory methods in which good performance on 1 criterion can offset bad performance on another, inserting identical default values for all alternatives would have an equalizing effect. Accordingly, it is not clear that inserting a default value—even a very "conservative" worst case value—would necessarily have an overall protective effect. However, in cases in which certain criteria are simply not relevant to the alternatives, it would be appropriate to remove those criteria from the model at the outset rather than filling them with uniform default values for all alternatives.

Third, the study's outcomes indicate that the 2 MCDA approaches are fairly robust, particularly in identifying the top ranked alternatives across a range of scenarios that shifted weighting, data assumptions and model design. The results of the weighting sensitivity analysis in the Pb solder case offers a cautionary note, however, demonstrating that, depending on the specifics of a given case, shifts in how criteria are weighted can have significant impacts on relative ranking. Additionally, further development of weighting methods for regulatory purposes is necessary. The project engaged in limited stakeholder elicitation and weighting. Research regarding alternative weighting methodologies specific to the regulatory context and the integration of weighting into the regulatory framework would be useful.

As with alternatives assessment, the alternatives evaluation component also requires additional development in a variety of areas. For example, further consideration is needed regarding the approach to normalizing performance scores in MAUT. This study used an internal normalization approach to convert an alternative's scores on each criterion to dimensionless units ranging from 0 to 1. Internal normalization in this case assigned a utility of 1 to the score of the best performing alternative and a utility of 0 to the score of the worst performing alternative. By contrast, external normalization would assign utilities based on an absolute scale derived from data or normative preferences external to the set of alternatives (Norris 2001). Although commentators have identified potential drawbacks to internal normalization—including magnitude sensitivity and the possibility of rank reversal on the addition of new alternatives—the normative and pragmatic advantages of external normalization are subject to continuing debate (Bare 2010; Prado et al. 2012).

Further evaluation of additional MCDA methods would be useful. Although the project focused on 2 leading MCDA methods, other methods with additional useful features are available; for example stochastic MAUT approaches could be used to deal with uncertainty in weights or performance scores (Linkov and Moberg 2012). Likewise, policymakers may

Table 5. MCDA selection principles

Principle	Description
Cognitive comfort	The extent to which the decision maker is comfortable with making pairwise comparisons as opposed to making trade-offs based on utility functions (Guitouni and Martel 1998)
Ranking	The extent to which the decision maker is seeking a complete or partial ranking of alternatives (Guitouni and Martel 1998)
Characteristics of data	Whether the MCDA method is suitable for the type, quality, and quantity of input information involved (Guitouni and Martel 1998; Løken 2007).
Compensation degree	The extent to which the method allows good performance on one criteria to offset inferior performance on another (Guitouni and Martel 1998; Benoit and Rousseaux 2003; Prado et al. 2012)
Suitability of theoretical principles	The extent to which the theoretical basis for the MCDA method is consistent with the decision context (Guitouni and Martel 1998; Seppala et al. 2002).
Practicality and tractability	The relative ease of use, level of resources required, and transparency of the method (Benoit and Rousseaux 2003; Løken 2007).

MCDA = multi-criteria decision analysis.

instead consider an entirely different framework. For example, under "goal aspiration" or threshold decision frameworks, the decision maker does not seek to identify the best option or even rank the options. Rather, such methods identify those options that achieve (or come closest to achieving) some minimum level of performance for 1 or more criteria (Linkov et al. 2004). In addition, other MCDA approaches such as analytic hierarchy process should likewise be evaluated for regulatory alternative analysis purposes.

Moreover, standards for selecting the most appropriate MCDA method for various regulatory alternatives analysis applications should be developed. The project identified some differences in outcomes between MAUT and outranking. For example, although the removal of all criteria with missing data in the Pb solder case study resulted in significant reordering of alternatives ranking in MAUT, in outranking it had little effect. Depending on a specific decision context, this may or may not be important. A recent literature review revealed a lack of specific guidance for selection of MCDA methodologies for particular application contexts. In practice choice of MCDA approaches appears to be driven by availability of specific expertise and familiarity of software tools (Huang et al. 2011). Nonetheless, Guitouni and Martel (1998) and other commentators have identified general principles (set out in Table 5) that are of some use in judging the appropriateness of MCDA methods in a regulatory setting.

A comprehensive comparison of MAUT and outranking across these principles is beyond the scope of this article. That said, 2 points are worth noting. First, with respect to several of these principles, a comparison must await articulation of specific structure, standards, and goals of the relevant regulatory program for which the methods may be used. For example, the level of cognitive comfort, the importance of ranking, and the suitability of the method's theoretical principles are heavily dependent on the actual design of the regulatory program. Second, given the centrality of the protection of human health and the environment to most chemical regulation programs, it is quite possible that policymakers may reject a fully compensatory MCDA approach such as that used for the MAUT baseline scenario in this project. For example, extremely poor performance on some criteria (such as impact on endangered species, technical feasibility, or carcinogenicity) may be so disfavored by a regulatory agency or other stakeholder that a partially compensatory or even noncompensatory approach may be selected (Prado et al. 2012). Thus, an environmental agency may wish to exclude any alternative using a known carcinogen. It is important to note, however, that it is possible to allow for noncompensatory aspects of decision analysis even when using MAUT or outranking by structuring the analysis into several steps, one of which excludes alternatives presenting unacceptable performance on 1 or more criteria (Stewart and Losa 2003). The Sequential Decision Model and the Modified Sequential Model illustrate such a strategy.

Although additional research and refinement concerning the application of MCDA methods to chemical alternatives analysis is required, MCDA can play a critical role in emerging prevention-based regulatory programs. MCDA methods offer a means for transparent, objective and rigorous analysis of products and processes, providing regulators and stakeholders with a common baseline understanding of the relative performance of alternatives and the trade-offs they present.

Acknowledgment—The authors would like to acknowledge the 3 anonymous peer reviewers for their helpful comments, and the research and analytical support of Carrie Gibson and Elina Nasser, which was central to the success of this project, and the research assistance of Courtney Havens. The authors also acknowledge the support of the DECERNS team (lead by Professor Boris Yatsalo) for allowing the use of the DECERNS software. Funding was provided by the Public Health Trust, a program of the Public Health Institute, through a defendant's settlement of a complaint brought by the State of California. Funding was also provided by the Robert Wood Johnson Public Health Law Research Program. Protocols for interviews with human subjects were approved by the UCLA Institutional Review Board.

Disclaimer—The peer-review process for this article was managed by the Editorial Board without the involvement of I Linkov.

REFERENCES

Assemb B. 1879, 2007-2008 Regular Session, 2008-Pamphlet No. 6A Cal. Adv. Legis. Serv. pp. 1–5 (LexisNexis).

Bare JC. 2010. Life cycle impact assessment research developments and needs. *Clean Technol Environ Pol* 12:341–351.

Belton V, Stewart TJ. 2002. Multiple criteria decision analysis: An integrated approach. Norwell (MA): Kluwer Academic Publishers. 400 p.

Benoit V, Rousseaux P. 2003. Aid for aggregating the impacts in life cycle assessment. *Int J Life Cycle Ass* 8:74–82.

Bottomley PA, Doyle JR. 2001. A comparison of three weight elicitation methods: Good, better, and best. *Int J Manage Sci* 29:553–560.

Brans JP, Mareschal B, Vincke P. 1986. How to select and how to rank projects: The PROMETHEE Method. *Eur J Oper Res* 24:228–238.

Drew CH, Nyerges TL. 2004. Transparency of environmental decision making: A case study of soil cleanup inside the Hanford 100 area. *J Risk Res* 7:33–71.

[ECHA] European Chemicals Agency. 2007. Guidance for the preparation of an annex XV Dossier for restrictions. Helsinki, Finland: ECHA.

[ECHA] European Chemicals Agency. 2010. General approach for prioritisation of substances of very high concern (SVHCs) for inclusion in the list of substances subject to authorisation. Helsinki, Finland: ECHA.

[ECHA] European Chemicals Agency. 2011. Guidance on the preparation of an application for authorisation. Helsinki, Finland: ECHA.

Finnveden G, Hauschild MZ, Ekvall T, Guinee J, Heijungs R, Hellweg S, Koehler A, Pennington D, Suh S. 2009. Recent developments in life cycle assessment. *J Environ Manage* 91:1–21.

Gass SI. 2005. The great debate: MAUT versus AHP. *Interfaces* 35:308–312.

Geibig JR, Socolof ML. 2005. Solders in electronics: A life-cycle assessment. Cincinnati (OH): US Environmental Protection Agency. EPA 744-R-05-001.

Giove S, Brancia A, Satterstrom FK, Linkov I. 2009. Decision support systems and environment: Role of MCDA In: Marcomini A, Suter GW, Critto A, editors. Decision support systems for risk-based management of contaminated sites. Netherlands: Springer Science & Business Media. p 53–73.

Gloria TP, Lippiatt BC, Cooper J. 2007. Life cycle impact assessment weights to support environmentally preferable purchasing in the United States. *Environ Sci Technol* 41:7551–7557.

Guitouni A, Martel JM. 1998. Tentative guidelines to help choosing an appropriate MCDA method. *Eur J Oper Res* 109:501–521.

Handwerker CA, Kattner U, Moon K, Bath J, Bradley E, Snugovsky P. 2007. Alloy selection. In: Bradley E, Handwerker C, Bath J, Parker R, Gedney R, editors. Lead-free electronics. Hoboken (NJ): John Wiley & Sons. p 9–46.

Huang IB, Keisler J, Linkov I. 2011. Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. *Sci Total Environ* 409:3578–3594.

Hwang J. 2005. Lead-free implementation and production: A manufacturing guide. New York: McGraw-Hill. 473 p.

Kiker GA, Bridges TS, Varghese A, Seager TP, Linkov I. 2005. Application of multicriteria decision analysis in environmental decision making. *Integr Environ Assess Manag* 1:95–108.

Koch L, Ashford N. 2006. Rethinking the role of information in chemicals policy: Implications for TSCA and REACH. *J Cleaner Prod* 14:31–46.

Kuczenski B, Geyer R. 2010. Chemical alternatives analysis: Models, methods and tools. Santa Barbara (CA): Department of Toxic Substance Control. Contract report 08-T3629.

Lau JH, Wong CP, Lee NC, Lee SW. 2000. Electronics manufacturing with lead-free, halogen free and conductive-adhesive materials. New York (NY): McGraw-Hill.

Leadership Council. 2011. Addressing public health and chemical exposures: An action agenda. National Conversation on Public Health and Chemical Exposures. Atlanta (GA): US Centers for Disease Control and Prevention, Agency for Toxic Substances and Disease Registry.

Linkov I, Varghese A, Jamil S, Seager TP, Kiker G, Bridges T. 2004. Multi-criteria decision analysis: A framework for structuring remedial decisions at contaminated site. In: Linkov I, Ramadan AB, editors. Comparative risk assessment and environmental decision making. Netherlands: Kluwer Academic Publishers. p 15–54.

Linkov I, Moberg E. 2012. Multi-criteria decision analysis: Environmental applications and case studies. Boca Raton (FL): CRC Press. 204 p.

Løken E. 2007. Use of multicriteria decision analysis methods for energy planning problems. *Renew Sust Energy Rev* 11:1584–1595.

Maine Toxic Chemicals in Children's Products Act, Me. Rev. Stat. Ann. tit. 38 §§ 1691–1699 (West Supp. 2011).

Norris GA. 2001. The requirement for congruence in normalization. *Int J Life Cycle Ass* 6:85–88.

[NRC] National Research Council. 2008. Public participation in environmental assessment and decision making. Washington DC: National Academies Press. 322 p.

[NRC] National Research Council. 2009. Science and decisions: Advancing risk assessment. Washington DC: National Academies Press. 424 p.

Pohekar SD, Ramachandran M. 2004. Application of multi-criteria decision making to sustainable energy planning—A review. *Renew Sust Energy Rev* 8:365–381.

Prado V, Rogers K, Seager TP. 2012. Integration of MCDA tools in valuation of comparative life cycle assessment. In: Curran MA, editors. Life cycle assessment handbook. Salem (MA): Scrivener Publishing. p 413–431.

[REACH] Registration, evaluation, authorisation and restriction of chemicals. 2006. Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). *Official J Eur Union* 396:1–849.

Seager TP. 2004. Understanding industrial ecology and the multiple dimensions of sustainability. In: Bellandi R, editor. Strategic environmental management for engineers. Hoboken (NJ): John Wiley & Sons. p 17–70.

Seppala J, Basson L, Norris GA. 2002. Decision analysis frameworks for life-cycle impact assessment. *J Ind Ecol* 5:45–68.

Sinsheimer PJ, Grout C, Namkoong A, Gottlieb R, Latif A. 2007. The viability of professional wet cleaning as a pollution prevention alternative to perchlorethylene dry cleaning. *J Air Waste Manage* 57:172–178.

Sinsheimer PJ, Malloy TF. 2009. Integrating safer alternatives into chemical policy: Developing a regulatory framework for AB 1879. Los Angeles (CA): Sustainable Technology and Policy Program, UCLA. 13 p.

Soares SR, Toffoletto L, Deschenes L. 2006. Development of weighting factors in the context of LCIA. *J Cleaner Prod* 14:649–660.

Stewart TJ, Losa FB. 2003. Towards reconciling outranking and value measurement practice. *Eur J Oper Res* 145:645–659.

Zhou X, Schoenung JM. 2007. An integrated impact assessment and weighting methodology: Evaluation of the environmental consequences of computer display technology substitution. *J Environ Manage* 83:1–24.

Zhou X, Schoenung JM. 2008. An integrated impact assessment and weighting methodology: Evaluation of the environmental consequences of lead-free solder alternatives. May 18, 2008. Proc IEEE Electron Environ. San Francisco, CA. 6 p.