A MULTI-CRITERIA SOLAR DESIGN METHOD

Inger Andresen, Ph.D.
Department of Building Technology, Faculty of Architecture, Planning and Fine Arts, Norwegian University of Science and Technology, N-7496 Trondheim, Norway, +4773592628, +4773598285, Inger.Andresen@civil.sintef.no

Abstract – This paper describes results of a Ph.D. study and work within the IEA SHCP Task 23: “Optimization of Solar Energy Use in Buildings”. The background for the work was the assumption that the success of solar buildings relies on the assessment and integration of all the different design objectives or criteria. These criteria are often quite complicated to deal with (e.g. environmental loading) and may be conflicting. The different design issues and the many different available energy technologies call for different areas of expertise to be involved in the design of solar buildings. This makes it difficult to evaluate the overall “goodness” of a proposed design solution, and the communication between design professionals and the client becomes complicated. The goal of this work was therefore to produce a means for the design team and clients to be able to better understand and handle holistic solar design. The main hypothesis was that a structured approach for evaluating design alternatives might be a means to this end. Thus, a method for evaluating solar design alternatives was developed, based on design theory, decision analysis, and behavioral science. The method has been tested in actual building design processes in Norway and in other countries participating in the IEA project. These test cases suggested that the multi-criteria decision-making method helped organize the design work and facilitated carefully balanced and integrated evaluations. It provided a common reference for the design team to synthesize all judgements and values in the early design, and promoted documentation of the choices.

1. INTRODUCTION

Building design involves the consideration of a wide range of complex issues, requiring expertise within fields ranging from structural engineering and environmental sciences, to architectural and psychosocial issues. An integrated approach to building design seeks to incorporate all the important aspects in a holistic synthesis. It views the individual systems not as isolated entities, but closely connected and interacting with the rest of the building. In an integrated approach, all the different design criteria need to be focused on simultaneously and traded off against each other in order to optimize the “overall goodness” of the design.

Unfortunately, solar and energy saving measures have often been implemented and “optimized” without taking the whole building performance into account. The “optimization” procedures employed usually only include energy and economic performances, at the most. There are numerous examples of how comfort issues, environmental issues and aesthetics have not been properly valued in the design of energy efficient buildings.

This paper describes results of a Ph.D. study at the Norwegian University of Science and Technology. The work was also done in part within the IEA SHCP Task 23: “Optimization of Solar Energy Use in Buildings”. The background for the work was the assumption that the success of solar buildings relies on the assessment and integration of all the different design objectives, called criteria. These criteria are often quite complicated to deal with (e.g. environmental loading) and may be conflicting. The different design issues and the many different available energy technologies call for different areas of expertise to be involved in the design of solar buildings. This makes it difficult to evaluate the overall “goodness” of a proposed design solution. Also, the communication between design professionals and the client becomes complicated.

The goal of this work was therefore to produce a means for the design team and clients to be able to better understand and handle holistic solar design. The main hypothesis was that a structured approach for evaluating design alternatives might be a means to this end.

2. DESIGN PROCESS THEORY AND PRACTICE

In order to specify an approach that would fit into the building design process, an analysis of design process theory and building design practice was carried out. Also, special solar design issues were investigated. This analysis resulted in the following conclusions:

• Most building design processes start out with no clearly defined goals or criteria of success. The design criteria are refined and discovered through evaluation and feedback on alternative design proposals.

• Design involves a lot of subjective value judgements, and decisions are often based on experience, “gut feeling”, or intuition. Design options are evaluated based on quantitative and qualitative performance
measures. There exists no objective optimal design solution.

- It is possible to identify some main activities that are common to most design processes. These are categorized into 4 main tasks: problem formulation, generation of alternatives, performance prediction and evaluation. The activities are very much overlapping and dependent on each other.

- Decision-making in design happens mainly through evaluation of proposed design solutions.

- Close cooperation between the design team and the client is imperative for the success of solar design. It is important that solar energy considerations are included in the early design stages because the decisions taken here are essential for achieving a well-integrated solar energy system.

This led to the conclusion that a structured approach for evaluation of alternative design options would be useful, and important, but it has to be seen in relation to the other design tasks, i.e. put into a building design framework.

3. STRUCTURED EVALUATION IN DESIGN

The next task was therefore to search for useful ideas for structuring evaluation in the early design stage. The search was concentrated on the building industry and the field of decision analysis. The field of decision analysis was included because evaluation is an important task within this field.

The result of the search was that some structured approaches to evaluation were found within environmental planning and the emerging field of “green building” assessments. These methods resembled somewhat the Multi-Criteria Decision-Making (MCDM) methods that were found within the field of decision analysis.

The next task was then to evaluate these methods with respect to use in building and solar systems design. This evaluation was based on the findings about the design process characteristics.

The first conclusion was that a structured side-by-side comparison of alternative solutions seemed to be essential to evaluation. However, a simple side-by-side comparison of alternatives with respect to the different criteria, will in most cases not be sufficient to reveal the best alternative. This is due to the fact that the goals are often conflicting or apparently incommensurate. Therefore, trade-offs need to be made, and some sort of aggregation of the performance measures into an overall measure of “goodness”, would be useful. From the survey of multi-criteria evaluation approaches, the Simple Additive Weighting (SAW) approach appeared to be most suitable. Various applications of SAW approaches were found in the “green building” and environmental impact assessment tools. The SAW-based methods were also found to be the most simple and intuitive of the MCDM approaches. The main advantage of the SAW model is that it makes value judgements explicit, thereby acting to increase mutual understanding among the design team participants and the client about what is important to focus on. It is also important that both qualitative and quantitative values can be incorporated in the model. The model is also quite flexible, thus it can be tailored to individual needs. A problem of the method is that it is quite difficult to create commensurate measurement scales and elicit representative weights. Therefore, scaling and weighting techniques were investigated in more detail. It was concluded that it might be possible for the design participants to create scales and weights to be used in an evaluation based on a SAW approach that is adapted for the building design process. However, appropriate guidance is needed, especially concerning how to create commensurate measurement scales.

4. A FRAMEWORK FOR MULTI-CRITERIA DECISION-MAKING IN SOLAR DESIGN

A structured evaluation approach based on the SAW method was then developed and put into a building design framework. The framework was separated into 4 parts, problem formulation, generation of alternatives, performance prediction and evaluation (Figure 1). Although the main emphasis was placed on the evaluation part, the other tasks were also included since the different tasks are very much interconnected and overlapping.

Figure 1. The Framework in relation to the design phases.

The framework is primarily directed towards the multidisciplinary design team and the client, but the methodology should be a common point of reference to
which all stakeholders\textsuperscript{1} can relate. However, there is a need for a central decision-maker that can guide the process. This may be the architect, a representative of the client, or a specially appointed project leader.

Below follows the specification of the different tasks in the framework.

4.1 Problem Formulation

A hierarchy of design criteria should be developed using the principles of value focused thinking (Keeney 1992). Although it is difficult to specify the objectives before any concrete design alternatives are present, it is important to focus on the values first, not to impose any destructive constraints to the creative process. In value focused thinking, the design criteria are specified using a top-down approach. This means that one starts with defining the higher-level objectives, i.e. the general, strategic objectives of the decision-makers. Then, the more precise, lower-level criteria are specified. Although, it is recommended to start with a top-down approach to eliciting criteria, a bottom-up approach may be useful in the next step of the process. In a bottom-up approach, potential solutions are used to help specifying the criteria.

It is not possible to give a general recipe of what criteria to consider or to what extent they should be included. The number and nature of the criteria will vary from case to case. Generally, it may be useful to focus on the reasons why a criterion is important, and to document this. This will also help clarify the objectives for the different members of the design team and the other stakeholders. For solar and low energy buildings, the life cycle perspective is important. The stakeholders of the building and their present and future needs and requirements have to be projected. The number of criteria and sub-criteria and the level of detail should be chosen based on an estimation of the available time and resources, and the needed accuracy of the evaluation. The following general guidelines may be given:

The criteria hierarchy should be complete and exhaustive. This means that all important performance attributes deemed relevant to the final decision should be represented. The real value attributes should be separated from topics people are mainly interested in. One should focus the design effort on those issues that make a major difference in the quality of the outcome.

The criteria hierarchy should contain mutually exclusive items. When eliciting criteria the same concept may arise under different headings. If both are included in the analysis then it is likely that as a consequence the concept will be attributed greater importance. If the hierarchy is restricted to mutually exclusive items, it would permit the decision-maker to view the attributes as independent entities among which appropriate trade-offs may later be made.

In addition, the number of main criteria and sub-criteria must be considered. If there are a large amount of criteria, the problem may be difficult to handle, and one may loose the overview. Also, with a large number of criteria, the importance weights end up small and thus the meaningfulness of the weights is blunted. On the other hand, if there are only a small number of criteria, it may lead to an oversimplification of the real world. The number of criteria may affect the “value balance” of the hierarchy. The greater the level of detail pertaining to an objective, possibly reflected in the number of sub-criteria, the more likely it is that it will be attributed a high level of importance. Miller’s theory (Miller 1956) suggests that it may be wise to limit the number of major and sub-criteria to approximately seven. Miller claimed that seven plus or minus two represents the greatest amount of information that an observer can give us about an object on the basis of an absolute judgement. This theory is widely accepted within behavioral sciences.

Figure 2 gives an example of a hierarchy that may be used to evaluate potential solutions for building energy systems. At the top is the overall goal of the design; “the optimal energy system” for this particular building. Then, at the next level follows the main criteria, which are divided into sub-criteria, sub-sub-criteria, and so on. The sub-criteria are presented in separate entities due to space limitations. Also, the criteria are indicated using key words only, e.g. full criteria statements are “minimum environmental loading”, “maximum comfort”, etc. This hierarchy may have been derived from a hierarchy of fundamental objectives as described by Keeney (1992). Note that some of the sub-criteria are crossed out; this is to show how the final hierarchy has been deducted from a wider, more general one.

\textsuperscript{1} A stakeholder of a business is defined as any individual or group who has an interest in the business because he can effect, or is affected by the activities of the business.
Finally, a simple assessment of the relative importance of the criteria should be carried out in this phase. This will provide a basis for the next two tasks, to help see what is important to focus on when generating alternatives and predicting their performance. This may also help specifying the level of detail for the different branches in the hierarchy. A simple ranking of criteria weights may be sufficient.

4.2 Generation of Alternatives
Since the generation of alternatives is mainly a craft, little formal guidance can be given for this task. However, it is important that the alternatives are generated based on the criteria and their relative importance. Techniques to promote creativity may be used, e.g. brainstorming techniques. It may be wise to start out wide to test the extremes of the criteria and to be sure that a wide range of possibilities has been considered. Different screening techniques may be helpful, for example simple ranks or the concept of dominance.

It may be useful to categorize the alternative solutions into main groups and subgroups, according to their purpose or main principles. For example, the main categories for façade systems may be Thermal Collectors, PV collectors, and Windows.

4.3 Performance Prediction
The performance prediction should be done in accordance with the criteria hierarchy. The criteria hierarchy and the preliminary weights dictate the level of detail needed in the performance predictions for the different criteria. In this way, an appropriate and balanced performance prediction may be carried out.

The result of the performance prediction may be presented in a matrix structure to facilitate side-by-side comparison. The matrix should contain descriptions of the alternatives’ performances with respect to the different criteria. Different types of performance indicators may be used, i.e. numbers may be used for the quantitative criteria and verbal statements for the qualitative criteria. Uncertainties should be specified, for example by indicating the plausible ranges of deviation from the expected value. At this stage, the results should be presented as objectively as possible, i.e. as “raw data” without any relation to reference values.

A so-called “star diagram” may be used to present the overall performance of an alternative, see the figure below. In this diagram it is possible to show multiple dimensions, thus all the individual performance measures can be gathered into one picture. Each “finger” represents the scale for one criterion. The performance on each criterion is plotted on each “finger” (the center of the star usually designates the lowest/worst performance). However, there is a limit to the number of dimensions that can be presented without making the graphics look cluttered. Following Miller’s theory (Miller 1956), a maximum of 9 fingers is probably the limit. Only the main criteria should be plotted in the star diagram. The sub-criteria may possibly be indicated below the main criterion.

![Star diagram](image)

Although the star diagram may be used to give an indication of the overall performance of an alternative, it should be used with caution. This is because the performance scales for the criteria are not commensurate; individual scales are used, and no weights are applied to make them comparable. Therefore, it is recommended to show the individual scales on each “finger”, to clarify that the scales are not directly comparable.

The next task is to prepare for the evaluation. Evaluation involves measuring and comparing the alternative solutions with respect to the objectives. To measure and compare something, you need to have a reference. The choice of reference is important, because it may have a big influence on the results. Common references are minimum, maximum or mean values. In the case of building design, these may be “least acceptable”, “standard/common practice”, and “best practice”. Alternatively, one of the proposed solutions may serve as a reference. The reference performances may be indicated in the star diagram, as shown in Figure 4.

![Star diagram with individual criteria scores](image)

Figure 4. Star diagram with individual criteria scores indicated.
4.4 Performance Evaluation

The evaluation process is separated into 4 steps:

1. **Scaling and scoring**: The performances of each attribute are transformed into a common scale.
2. **Weighting**: The criteria are weighted based on the chosen scales.
3. **Aggregation**: The weights and the scores are aggregated to obtain an overall evaluation of each alternative.
4. **Sensitivity analysis**: Sensitivity analyses are performed to test the robustness of the results. Conclusions are drawn and recommendations are made.

It may be necessary to do several iterative loops in this process. Firstly, the weighting may lead to a need to redefine the scales. Secondly, the sensitivity analysis may reveal a need to change both the scales and the weights.

**Step 1: Scaling and scoring**

The common measurement scale need not have a large number of intervals. This is because fine gradations do not make sense for the qualitative criteria; these criteria can only be described verbally, and humans have a limited vocabulary to express qualitative gradations. Also, for the quantitative criteria, it does not make sense to have a very fine scale, partly because of the large uncertainties of the performance predictions in the early design phase, and partly because they should be compatible to the qualitative ones. A 9- or 10-point scale seems to be the maximum usable gradation. When comparing alternatives, we have 5 main terms to express the relative attractiveness of one alternative to the other, see Table 1. The categories in between can be used when the decision-maker hesitates between the neighboring qualifications. The 10-point scale is usually reduced to a 7-point scale because the performances that are assigned a score lower than 4 are so poor that they cannot be compensated elsewhere. The 7-point scale has also been widely used within behavioral sciences. Thus, a 7-point scale, with scores ranging from 4 to 10 seems to be most appropriate.

<table>
<thead>
<tr>
<th>relative judgement</th>
<th>direct judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>reference</td>
</tr>
<tr>
<td>9</td>
<td>excellent</td>
</tr>
<tr>
<td>8</td>
<td>somewhat less attractive</td>
</tr>
<tr>
<td>7</td>
<td>less attractive</td>
</tr>
<tr>
<td>6</td>
<td>much less attractive</td>
</tr>
<tr>
<td>5</td>
<td>vastly less attractive</td>
</tr>
</tbody>
</table>

The qualitative criteria are rated directly on this scale. When constructing scales for the quantitative criteria, the qualitative scale should be used as a basis. First, the end points should be decided. The end points should be realistic, that is, it should be easy to imagine that some of the alternatives being considered might realistically score at the maximum and minimum value. The scale should be wide enough that it is probable that new alternatives do not fall outside the boundaries, and narrow enough so that the alternatives are well spread out in order to be able to differentiate among them. This can be done by considering the acceptable range instead of the possible range, i.e. the range that one would be willing to consider. For example, if the criterion is “energy consumption”, one could ask the following question to the decision-makers: “What is the best performance of energy consumption you would expect in a potential design solution?” and “What is the least you would accept?”

The scale needs to be divided into intervals that are felt to be equal. There are three different approaches that may be used to determine the points in between the extremes. The easiest approach is to apply a linear scale. However, such a scale implies that our preference follow a linear progression, which is probably not likely, at least not for all criteria. Nevertheless, a linear scale may be a useful approximation. Some experiments suggest that straight-line approximations to curved utilities make little difference to the final decision (Edwards and Newman 1982) (von Winterfeldt and Edwards 1986).

The second approach is to manually subdivide the scales so that the intervals are felt to be equal, e.g. the difference between the first and the third, and the third and the fifth scalar pairs have to be subjectively equal. This, however, might be a cumbersome and time-consuming process.

The third approach is based on the assumption that our preferences follow a geometric progression. In his recent work, Lootsma (1999) presents convincing arguments based on behavioral sciences and psycho-physics, which supports this. He shows that our sensory systems for the perception of sound, light, smell, taste and time all follow geometric sequences that are very similar. Then he shows that this enables us to categorize quantitative values on scales using verbal expressions like fair, good, excellent, etc. The qualitative attributes can be rated directly on this scale. Thus, a link between the quantitative and qualitative scale is established. A prerequisite for this approach is that the decision-maker considers the range of acceptable performance data from the so-called desired target ($P_{min}$ in Table 2).
Table 2. The geometric progression scale, based on Lootsma (1999).

<table>
<thead>
<tr>
<th>Numerical performance indicator (e.g. price)</th>
<th>Qualitative verbal and symbolic scales</th>
<th>Score, grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{min}$ reference</td>
<td></td>
<td>20000</td>
</tr>
<tr>
<td>$P_{h} = P_{min} + c_{h}$</td>
<td>desirable</td>
<td>excellent</td>
</tr>
<tr>
<td>$P_{i} = P_{min} + c_{i}$</td>
<td>good</td>
<td>++</td>
</tr>
<tr>
<td>$P_{j} = P_{min} + c_{j}$</td>
<td>somewhat less desirable</td>
<td>good</td>
</tr>
<tr>
<td>$P_{k} = P_{min} + c_{k}$</td>
<td>fair/good</td>
<td>--</td>
</tr>
<tr>
<td>$P_{l} = P_{min} + c_{l}$</td>
<td>unsatisfactory/fair</td>
<td>---</td>
</tr>
<tr>
<td>$P_{m} = P_{min} + c_{m}$</td>
<td>much less desirable</td>
<td>unsatisfactory</td>
</tr>
</tbody>
</table>

The choice of scaling and scoring approach must be viewed in connection to the weighting approach.

Step 2: Weighting
A weight assigned to a criterion is essentially a scaling factor that relates scores on that criterion to scores on all other criteria. This is a subtle idea, and is not always easy to see. Therefore, it may have to be explained to the decision-makers, e.g. by giving an example. Using the first two scaling approaches described above, the link between the scales and the weights need to be established. Thus, consistency checks have to be made using trade-off questions or indifference questions. The link to the scales may be established by showing the scales and asking the decision-makers about their trade-offs or indifference between the criteria. For example, a set-up like the one in Figure 5 may be used. The scales for the different criteria are drawn side-by-side next to the common scale of grades. The trade-offs between the criteria given by the ratio of weights are then checked by indifference questions such as the ones below the scales in the figure below.

Figure 5. Set-up for linking the scales to the weights.

An alternative is to use an approach based on the preference theories of Freerk A. Lootsma (1999). The criteria weights are determined on a 7-point scale similar to the one used for scoring the performances, see Table 2 above. The decision-maker can express the importance of criteria in grades on the scale 4.5,...,10, as follows (grades lower than 4 are possible but will practically eliminate the corresponding criteria):

Table 3. Grading scale for determining weights.

<table>
<thead>
<tr>
<th>grade $h$</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Preferred criterion (reference)</td>
</tr>
<tr>
<td>9</td>
<td>Somewhat less preferred</td>
</tr>
<tr>
<td>8</td>
<td>Less preferred</td>
</tr>
<tr>
<td>7</td>
<td>Much less preferred</td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

The most important criterion receives a grade of 10. All the other criteria are compared to this, e.g. if a criterion is felt to be somewhat less important than the most important one, it receives a grade of 8. The normalized weight of criterion $C_i$ is calculated as follows:

$$c_i = \left(\frac{f_i}{\sum f_i}\right)^{\frac{1}{h}}$$

Using this approach, the link between the scales and the weights is embedded in the procedure, and does not have to be established explicitly.

Step 3: Aggregation
The simple additive weighting (SAW) model is proposed to aggregate the performance grades and the criteria weights:

$$f_j = \sum_{i=1}^{n} c_i g_{ij}, \quad j = 1,...,n$$
where \( f_j \) is the final score for alternative \( A_j \), \( n \) is the number of alternatives, \( m \) is the number of criteria, \( c_i \) is the normalized weight of the performance criterion \( C_i \), and \( g_{ij} \) is the performance grade (score) for alternative \( A_j \) with respect to criterion \( C_i \).

The calculations needed are so simple that a hand calculator is sufficient to produce the result (or it may even be done manually). However, a simple worksheet setup might be useful because the process can be more automated and it gives the possibility to produce graphics to illustrate the results. The results of the evaluation should be presented in such a way that it is possible to trace the weighing and scoring outcomes. Figure 6 shows an example of such a presentation where the total scores for the alternatives are shown (top) together with the performance scores with respect to the different criteria (bars – different colored sections), and the weights of the criteria (pie-chart).

![Figure 6](image)

Figure 6. Graphical presentation of the aggregation of weights and scores. The scores are indicated on the vertical axis. The numbers above the bars are the total weighted scores for each alternative. The weights are shown in the pie chart at the top right.

**Step 4: Sensitivity study**

A sensitivity study is a test of the robustness of the results with respect to variations in the assumptions used in the evaluation, i.e. the criteria weights and the performance measures. It would probably be wise to start with investigating the sensitivity to weights. The criteria weights are the essence of value judgements; they very much determine the outcome. Also, the weights are more likely to be in dispute than the performance measures (scores). It is usually sufficient to vary the weights at the higher level of the trees, because the lower-level weights have much less effect on the result. Obviously, one should vary the weights that are most dubious. The next task is to investigate variations of the performance measures. This is only needed if some performance measures discriminate much between alternatives. The result may be tested with extreme values for different future scenarios.

**5. CONCLUSIONS**

The evaluation approaches were then exemplified and tested in 3 different building projects. The first case was a project that aimed at designing an a-Si PV facade concept for office buildings. The participants defined their own measurement scales. The second case project was a test of the evaluation approach in a meeting between researchers and practitioners working on the preliminary design of a school building. In this case, the participants used pre-defined linear measurement scales. The third case included an implementation of the multi-criteria evaluation approach in an ongoing design project, using the logarithmic type scales.

The test cases suggested that the multi-criteria decision-making framework helped organize the design work and facilitated carefully balanced and integrated evaluations. It provided a common reference for the design team to synthesize all judgements and values in the early design, and promoted documentation of the choices. However, further testing is needed in order to streamline the approach and to confirm whether the framework is an efficient means to promote holistic solar design.

**REFERENCES**


