Elicitation of decision parameters for thermal comfort on the trains

Mammeri Lounes[+*], Bouyssou Denis[*], Galais Cedric[+], Ozturk Meltem[*], Segretain Sandrine [+], Talotte Corine[+],

+: SNCF, email: <nom>@sncf.fr

*: Lamsade/CNRS, Universite Paris Dauphine, email: <nom>@lamsade.dauphine.fr,

Abstract. We present in this paper a real world application for the elicitation of decision parameters used in the evaluation of thermal comfort in high speed trains. The model representing the thermal comfort is a hierarchical one and we propose to use different aggregation methods for different levels of the model. The methods used are rule-based aggregation, Electre Tri and 2-additive Choquet. We show in this paper the reasons of the choice of such methods and detail the approach used for the elicitation of the parameters of these methods.

1 Introduction

Comfort is one of the main raison of the choice of trains for long trips. In this paper we are interested in one of the composant of global comfort which is the thermal one. We show how we define the thermal comfort using physical evaluations (temperature, air speed, etc.) in order to be as close as possible to the comfort perception of train passengers. In the following section we present how we establish our model. Our model requires different aggregation steps, in Section 3 we introduce these aggregation steps by presenting in a brief way their formulations, the raisons of their choice and specially the approach that we used for the elicitation of their decision parameters. We conclude our paper by some recommendations for elicitation approaches.

2 Thermal comfort model

Existing methods used for the evaluation of the thermal comfort on high speed trains are based on the Fanger's model ([4]), initially developed for office buildings. Fanger's model uses two indices, the *PMV* (Predicted Mean Vote) and the *PPD* (Predicted Percentage of Dissatisfied). The *PMV* is calculated using five criteria : clothing, metabolic rate (activity of the subject), temperature, air velocity and humidity, and is devised on the basis of tests conducted on a large group of subjects. Once the *PMV* value has been established from tables, it is then possible to determine the *PPD*. Fanger's model is devoted to static situations with long time exposure. The climatic environment parameters and activities of subjects are supposed to be constant. For these reasons its use for trains is not always very adequate. Moreover, some recent research done by the SNCF ([22], [16]), specially some surveys with the passengers on the train, showed that the results of the Fanger's model are not always correlated with the perception of the passengers. Figure 1 presents an example of responses of five passengers to the question "How do you evaluate the thermal conditions in this train?" and the evaluation given by the *PMV*. The first part of the figure represents the answers of passengers and the second part the results obtained by the *PMV*.

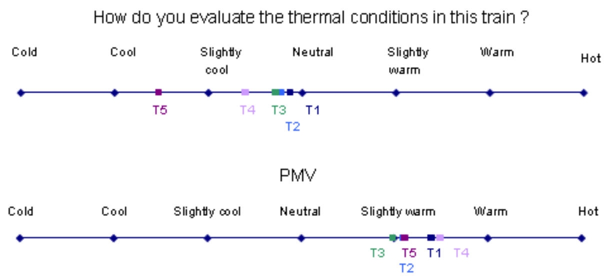

Figure 1. Difference between observations and the *PMV* results

A bibliographical summary of previous projects and research studies on evaluation and perception of thermal comfort was carried at the SNCF [26]. Some of these studies show that there are some perceptive parameters, missed in Fanger's model, which must be taken into account in the evaluation of thermal comfort.

1. The thermal comfort is a subjective notion, the perception can change from a subject to another one and this variability is not taken into account by the *PMV* index. Indeed, although this variability may be estimated by the *PPD*, it is not possible to estimate the thermal comfort of a given subject or a group of subjects sharing the same perception of the comfort.
2. The thermal preferences of a passenger may change with the season.

Other research studies done by the SNCF ([21], [27]) showed that the comfort on the trains is closely related to two perceptions:

1. there must be *no unpleasant sensations* caused by climatic parameters during the journey,
2. there must not be a *discontinuity of ideal thermal conditions*. Such discontinuity is generally caused by the variations of outside temperature, drafts air and the gap between outside and inside temperature.

Another result of these studies is that the most important climatic parameters are temperature and air speed.

Using these results and after several meetings with thermal experts we propose the following model presented in Figure 2 for the evaluation of thermal comfort on high speed trains:

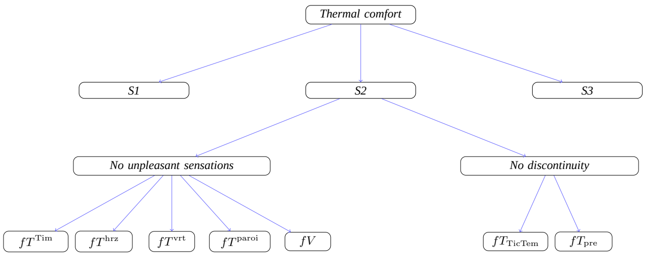

Figure 2. *Thermal comfort model*

In Figure 2,

- S_i represents a season,
- fT^{Tim} represents the weighted mean gap between the average inside temperature t and the reference inside temperature t_{ideal},
- fT^{hrz} represents the weighted mean gap between the maximum and minimum inside temperature on a horizontal section,
- fT^{vrt} represents the weighted mean gap between the maximum and minimum inside temperature measured at head, chest and legs of the passenger,
- fT^{paroi} represents the weighted mean gap between the temperature next to the windows and the average inside temperature,
- fV represents the mean of normalized gaps between the reference range of air speeds and the inside air speeds,
- fT_{TicTem} represents the average rate of change of inside temperatures according to the outside temperatures,
- fT_{pre} represents the gap between the inside temperature and the reference inside temperature when a passenger enters into a train.

3 Agregations

Each node of the thermal comfort model needs an aggregation method, the aggregations are done from the bottom to the top of the model.

3.1 Procedure to find the most appropriate aggregation method

In order to find the most appropriate aggregation method for each node, we ask the following questions to the experts:

Q_1: do we need a ranking or a classification of trains on this node?
Q_2: do we have ordinal or quantitative data?
Q_3: are there any dependance between the subcriteria of this node?
Q_4: is it acceptable to have a compensation between the subcriteria of this node?
Q_5: are there some important subcriteria of this node which may have a veto power (it means that such a subcriteria may put a veto for a good global evaluation if the evaluation of the train is not sufficient for this special subcriteria)?
Q_6: are there too many subcriteria in this node?

Table 3.1 presents a quick analysis of three aggregation methods (rule-based aggregation, Electre tri, Choquet) in relation to the previous questions. These methods will be presented in the following subsections. The answers are given in a very general way, some of them may be different with additional studies (for instance if we have ordinal data, we can translate the ordinal evaluation to utilities with a good elicitation method, ...).

Question	Rule-based	Electre	Choquet
Q_1	classification	classification	ranking
Q_2	ordinal/quantitat.	ordinal/quantitat.	quantitat.
Q_3	dependance	no dependance[1]	dependance
Q_4	no compensat.	no compensat.	compensat.
Q_5	veto[2]	veto	no veto
Q_6	not too many[3]	5-6 criteria[4]	5-6 criteria[5]

Table 1. methods and their properties

The choice of the aggregation is done in accordance with the answers to the questions presented above but there are also two other points that we have to take into account. The method must:

1. provide results in accordance with the preferential expectations of thermal experts and the answers of passengers to the surveys,
2. be easy to understand. It means that if there are many aggregation methods with expected properties, we may chose the most intuitive one in order to facilitate the use and the acceptation of the method by thermal experts. For thermal experts rule-based method is the most intuitive one between the three aggregation methods of Table 3.1 (they are used to have logical rules for the evaluations). However, the logical rules which will be used must be easy to interpret, it is not acceptable to have a big number of rules which have not intuitive meaning but correspond to the answers of the passengers to the survey. After rule-based method the experts fell more comfortable with Electre tri method since all the parameters (weights, indifference thresholds, veto thresholds, etc.) are present in a transparent way while some important indices of Choquet integrals (dependance and importance indices) are not very transparent in the beginning of the evaluations.

These two last points are important for our approach. The first point may be used in the validation step of our approach by comparing the theoretical results with passengers answers. It also says us that we can use some preferential examples in order to determine the parameters of the chosen aggregation method. This point is central for the following section where we will present the elicitation methods. The last point says us that we have to see first of all if we can use rule-based aggregation with simple rules if not we have to try Electre tri and finally we have to test Choquet integrals.

In the following we will present the aggregation method used in each node of the model. For confidentiality purposes we can provide neither the real examples that we used during the elicitation steps nor the real values of decision parameters.

3.2 No discontinuity

In this node we have two criteria $fT^{\text{TicTem}}(e, S)$ and $fT^{\text{pre}}(e, S)$ (see Figure 2), evaluated on cardinal scales, to be minimized. The experts stated that they just need to have a classification into three ordinal categories “no discomfort”, “mild discomfort” and “discomfort”. Our idea is to find a simple aggregation procedure like a small set of rules because of the small number of criteria and categories.

The classical rule-based methods in multicriteria decision making (MCDA) have their roots in rough sets theory [23] which aims at providing a set of rules $R = \{R_1, R_2, .., R_k\}$ (“if <conditions> then

<decision>") from a learning set of decision examples provided by the DM. This learning set is a set of alternatives A evaluated on a set of attributes $Q = \{q_1, .., q_m\}$ for which, decisions (the assignment of alternatives into categories $\{C_1, .., C_p\}$) were taken in the past by the DM (some fictitious examples may be also used if there are no previous decisions). The difficulty of the classical rough sets approach for MCDA is that it can not deal with preference order on the elements of Q and the categories $\{C_1, .., C_p\}$, and thus may violate *the monotonicity* of preferences. For this purpose, Greco, Mattarazzo and Slowinski proposed a generalization of the classical approach by proposing what they called Dominance-based Rough Set Approach (DRSA) [7, 8]. In our application we do not need DRSA since the induction of rules could be performed directly with the DM because of the small number of criteria, categories and also because of some limit values that the experts used to have. However, we should ensure that the set of rules satisfies three properties : the exclusiveness (each alternative must be assigned at most to one category), the monotonicity (the set of rules must be coherent with the dominance principle) and the exhaustivity (each alternative a must be assigned to a category by a rule) [6].

It turned out that the rules inducted with the experts by using a small set of fictitious examples, make use of the *minimum* aggregator. Indeed, for the experts, each criterion has two thresholds (s_1 and s_2 for $fT^{\text{TicTem}}(e, S)$ and s'_1 and s'_2 for $fT^{\text{pre}}(e, S)$) separating three comfort categories (reflecting three levels of comfort like in Tab. 2).

Level of comfort	Ordinal values
No discomfort	3
Mild discomfort	2
Discomfort	1

Table 2. *The coding of comfort categories*

Let $Cl^{\text{TicTem}}(e, S)$ and $Cl^{\text{pre}}(e, S)$ be the translations of $fT^{\text{TicTem}}(e, S)$ and $fT^{\text{pre}}(e, S)$ in terms of levels of comfort :

$$Cl^{\text{TicTem}}(e, S) = \begin{cases} 3 & \text{if } fT^{\text{TicTem}}(e, S) \le s_1 \\ 2 & \text{if } s_1 < fT^{\text{TicTem}}(e, S) \le s_2 \\ 1 & \text{if } fT^{\text{TicTem}}(e, S) > s_2 \end{cases} \qquad (1)$$

$$Cl^{\text{pre}}(e, S) = \begin{cases} 3 & \text{if } fT^{\text{pre}}(e, S) \le s'_1 \\ 2 & \text{if } s'_1 < fT^{\text{pre}}(e, S) \le s'_2 \\ 1 & \text{if } fT^{\text{pre}}(e, S) > s'_2 \end{cases} \qquad (2)$$

After that a train is assigned to one of three comfort categories for *No discontinuity* using the minimum operator:

$$Cl^{\text{NoDisc}}(e, S) = \min\left\{Cl^{\text{TicTem}}(e, S), Cl^{\text{pre}}(e, S)\right\}$$

3.3 No unpleasant sensations

In this node we have five criteria (see Figure 2) evaluated on cardinal scales, to be minimized. The experts stated that here again they just need a classification into three ordinal categories "no discomfort", "mild discomfort" and "discomfort" (see again Table 2).

After some discussions with the experts on comfort and on rolling stocks about these criteria, they claim that the first and the last criterion ($fT^{\text{Tim}}(e, S)$ and $fV(e, S)$) are by far the most important and can not be compensated by the three others for reducing the discomfort sensation.

The fact that we need an ordinal classification by using five criteria (it is too much to have intuitive logical rules) and that we have some type of veto (and/or no compensation), are the basic motivations of the choice of Electre Tri method in this node.

Electre Tri is a multicriteria sorting method developed by B. Roy [24]. Its principle is to assign alternatives to predefined and strictly ordered categories (from the worst to the best): $C_1, C_2, ...C_{p+1}$. The assignment of an alternative $a \in A$ in a category is based on the comparison of a with the profiles $b_1, b_2, ..., b_p$ (which separate these categories) on m criteria $g_1, g_2, .., g_m$. A profile b_h is a fictitious alternative which is considered as the lower limit of the category C_{h+1} for $h = 1, .., p$. The comparison of an alternative a with a profile b_h is performed with an outranking relation S, whose meaning is *"a is at least as good as b"*. The assertion aSb is validated if and only if the two following conditions are satisfied: a "majority" of criteria is in favor of a (the weighted sum of criteria in favor of a is greater than a threshold) and none of the criterion which is in favor of b should be against (put a veto) this assertion.

The parameters that can be inferred for a Electre Tri model are:

- The weights of the criteria $k_1, .., k_m$
- The profiles $g_i(b_h)$ $\forall i$ and $\forall h$
- The veto thresholds v_i $\forall i$ (if there are)
- The indifference q_i and preference p_i thresholds $\forall i$ (if there are)

These preferential parameters can be either provided by the DM himself, which rarely happens, or inferred by aggregation/desaggregation methodologies. In these methodologies, the DM is asked to provide a holistic judgment about a subset of potential alternatives $A^* \subset A$ by assigning them in predefined categories. Often, a mathematical programming is solved in order to obtain the estimated parameters that best restore the assignment proposed by the DM, we can have two possible cases for that:

- the mathematical programming can restore the assignment: then the DM can see the results of assignment of other potential alternatives by the inferred model, which can help him to provide further informations or
- the mathematical programming can not restore the assignment: then the DM can see which preferences are inconsistent with the model (but not necessary with its reasoning), so, he may either modify (or withdraw) them or decide that these preferences are so important that the model of Electre Tri must be dropped.

The main difficulty when inferring an Electre Tri model with a mathematical programming is that we can not infer all the parameters simultaneously because the corresponding constraints are non-linear and non-convex. Therefore some parameters must be inferred directly with the DM.

In the literature, the first methodology for inferring Electre Tri parameters by mathematical programming was performed by V. Mousseau and R. Slowinski [19]. In 2001, V. Mousseau, J. Figueira and J.P. Naux proposed a linear programming formulation for inferring the weights [18]. A. Ngo The and V. Mousseau in 2002 proposed an elicitation of the category limits [28]. Besides the aggrega-

[6] Vanderpooten and Azibi [1] have proposed an approach to check if the rule base satisfies the three previous requirements, provided that the rules have a particular structure. This approach consists on transforming the rules from logical to algebraic representation which allows to solve a series of linear programming in order to check the three requirements. We can also identify with this approach, alternatives which are not covered by rules satisfying these requirements.

tion/desegregation methodologies, direct methods was performed for inferring the Electre Tri model, like SRF ([5]).

We used the following procedure for our application:

- Since we need three categories, we just need to define two profiles. The profiles b_1 and b_2 were determined directly with the experts because they are used to work with some comfort levels defined by the limit evaluations on criteria.
- The weight of criteria were elicited by identifying all subsets of minimum coalitions of criteria $\underline{F} \subseteq F$ in favor of an alternative such that the alternative remains at least as good as the profile for the experts, without taking into account the veto power of criteria (for instance the expert says that it is sufficient to have a better than b for the three first criterion in order to say that a is at least as good as b, ...)
- Ones the weights are determined, we considered a set of learning alternatives in order to elicitate the veto thresholds.

$$A^* = \bigcup_{\underline{F} \subseteq F} (A_{\underline{F}}(1) \cup A_{\underline{F}}(2))$$

built from $\underline{F}$ such that for $p = \{1, 2\}$:

$$A_{\underline{F}}(p) = \{a \in A : \forall i \in \underline{F} : g_i(a) > g_i(b_p) \text{ and } \forall j \in F \backslash \underline{F} : g_j(a) \leq g_j(b_1)\}$$

We focused then on some alternatives $a \in A^*$ for which, we increase progressively the value of $g_j(a)$ (we decrease the performance) $\forall j \in F \backslash \underline{F}$, keeping the same performances in the remaining criteria, until the assertion aSb_p being not valid. Let g_j^* be the smallest value such that aSb_p is not valid. The veto threshold is thus:

$$v_j(g_j(a)) = g_j^* - g_j(a)$$

3.4 Comfort in a given season

In this note the evaluations on the *"No unpleasant sensations"* and the*"No discontinuity"* will be aggregated. These two criteria are evaluated on an ordinal scale with three grades. The experts stated that here again they just need to have a classification into three ordinal categories "no discomfort", "mild discomfort" and "discomfort" (see again Table 2). The small number of criteria and grades allows us to use rule-based methods in a very similar way as in Subsection 3.2. When asking the experts about the importance of the criteria, they stated that the *no unpleasant sensations* criterion is more important. Our second questioning was to know if the overall discomfort in a given season is greater than the maximum discomfort arising from *"No unpleasant sensations"* and *"No discontinuity"*. The answer was negative and besides, they stated that the overall discomfort is close enough to the maximum discomfort arising from the two criteria (the smallest category among the two criteria).

On the basis of this preferential information, we thought it could be useful to keep the same number of categories (the discomfort does not increase) as an evaluation of thermal comfort on each season. The principle of the set of rules is to assign the alternatives to the worst category among the categories corresponding to the two criteria in witch the alternative is assigned excepted when the alternative is in the *category* 3 for *"No unpleasant sensations"* and *category* 1 for *"No discontinuity"*, in which case it is assigned to the *category* 2. This set of rules can be summarized in the following formula:

$$Cl^{\text{season}}(e, S) = \begin{cases} 2 \text{ if } Cl^{\text{NoUnplSns}}(e, S) = 3 \text{ and } Cl^{\text{NoDisc}}(e, S) = 1 \\ \min\{Cl^{\text{NoUnplSns}}(e, S), Cl^{\text{NoDisc}}(e, S)\} \text{ otherwise} \end{cases} \quad (3)$$

3.5 Thermal comfort

The aggregation in this node will provide the global evaluation of thermal comfort. We have to aggregate three evaluations, each of them being ordinal with three grades representing the thermal comfort in season S_i. We began our discussion with experts by trying to define an aggregation which will provide three ordinal classes as in other nodes. We tried first of all rule-based methods. Intuitively, we though that the *minimum* operator would be a good candidate. However, some pairwise comparison examples that we showed to our experts proved that the *minimum* operator was not adequate. Moreover, it was not possible to find simple rules in accordance with their preferences. Then, we tried to see if we could use another classification method such as Electre Tri. The main difficulty of such an approach was the fact that for the experts it was very difficult to define a semantic for the categories. Futhermore, during the discussions with experts we noticed that there may be some dependancies between the three seasons. For that reason we decided to test Choquet integrals by proposing some pairwise comparisons to our experts.

Choquet integral in MCDA is an aggregation operator developed by T. Murofushi and M. Sugeno at the end of the eighteenth [25, 20]. Since, many studies and applications of Choquet integral in MCDA have been carried mainly for building the theoretical foundations [13, 12, 11, 15] and eliciting the parameters [6]. Choquet integral is a generalization of the most known scoring methods: *The weighted sum, the ordered weighted average* [29], *the weighted minimum and the maximum* [3]. Choquet integral of an alternative a, evaluated on the family of criteria F is given by the following formula:

$$C_\mu(a) = \sum_{i=1}^{n} \left[a_{\sigma(i)} - a_{\sigma(i-1)}\right] \mu(\sigma(i), ..., \sigma(n)) \quad (4)$$

Where $a_i = u_i(g_i(a))$.
$u_i(.) : X_i \mapsto [0, 1]$ are non decreasing utility functions.
σ is a permutation on F such that: $a_{\sigma(1)} \leq a_{\sigma(2)} \leq ... \leq a_{\sigma(n)}$.
$\mu(.)$ is a *capacity* on F

Definition 1 *A capacity (or a fuzzy measure) μ on F is a set function*

$$\mu(.) : \mathcal{P}(F) \mapsto [0, 1]$$

satisfying the following conditions:

(i) $\mu(\emptyset) = 0, \mu(F) = 1$
(ii) $\forall S, T \subseteq F : S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$

The capacity $\mu(S)$ of a subset of criteria S can be interpreted as the weight importance of the coalition of criteria of S. It allows to consider more preferential information than the scoring methods mentioned below, like the interactions among criteria or the mutual dependence of criteria.

Choquet integral provides also some numerical indices for analyzing the preferential information like *the Shapley value* $\Phi_\mu(i)$ for measuring the importance of a criterion and the interaction index $I_\mu(S)$ for measuring the interaction among the criteria belonging to $S \subseteq F$.

$$\Phi_\mu(i) = \sum_{T \subseteq F \backslash i} \frac{(n - |T| - 1)!|T|!}{(n)!} [\mu(T \cup i) - \mu(T)] \quad (5)$$

$$I_\mu(S) = \sum_{T \subseteq F \backslash S} \frac{(n - |T| - |S|)!|T|!}{(n - |S| + 1)!} \sum_{L \subseteq S} (-1)^{|S| - |L|} \mu(T \cup L) \quad (6)$$

where n is the cardinality of F. The application of Choquet integral in MCDA requires the elicitation of utility functions u_i and the capacities μ. The main requirement when eliciting the utility functions u_i is the *commensurability* of the scales. The MACBETH approach [2] is often used for eliciting the utility functions (by assuming that the DM is able to give information using intensity of preference) by building an interval scale u_i in order to encode the attractiveness of the elements of subsets $\overline{X}_i = \{(0_1, .., 0_{i-1}, x_i, 0_{i+1}, .., 0_m)/x_i \in X_i\}$ s.t. 0_i and 1_i are the worst and the best values in X_i. The commensurability is ensured by fixing the scales: $u_i(1_i) = 1$ and $u_i(0_i) = 0 \; \forall i$.

Regarding the elicitation of the capacities μ, several methodologies and algorithms have been developed in the literature. The general idea of these methodologies is to ask the DM to express his preferences on a set of learning alternatives A^*. These preferences from which the capacities will be elicited, can be a partial ranking of:

- Alternatives of A^*
- Differences of intensity of preferences of some alternatives in A^*
- Importance of criteria,
- Interactions between criteria
- ...

This preferential information is then translated into mathematical constraints such as:

- $a \succ b \Rightarrow C_\mu(a) \geq C_\mu(b) + \delta_1$
- $a \succ b$ more than $c \succ d \Rightarrow C_\mu(a) - C_\mu(b) \geq C_\mu(c) - C_\mu(d) + \delta_1$
- The criterion i is more important than the criterion $j \Rightarrow \Phi_\mu(i) \geq \Phi_\mu(j) + \delta_2$
- The criteria i and j are complementary $\Rightarrow I_\mu(ij) \geq \delta_3$
- ...

Where δ_1, δ_2 and δ_3 are preference thresholds which must be defined with the DM. It is also possible to fix the number of criteria which may interact.

Definition 2 (k-additivity) *A capacity μ on F is k-additive if there is no interaction among criteria of every subset $S \subseteq F$ whose cardinality is greater than k, ie., $\forall S \subseteq F$ s.t. $|S| > k$, $I_\mu(S) = 0$.*

Most of the methodologies in the literature [14, 10, 9, 17] use an optimization problem with the previous constraints for identifying the capacities. The objective function may differ from a methodology to another.

If a solution is found to this optimization problem then the DM can analyze the results corresponding to the Choquet integral with the identified capacities, he may add further preferential information and thus solves again a new optimization problem. Such a process is performed iteratively until finding a satisfactory model.

If no solution is found to the optimization problem then either the DM preferences are not consistent with the theoretical properties of Choquet integral (transitivity of preferences, monotonicity...etc.) or the number of parameters to be identified is not sufficient to restore the DM preferences. In the first case, inconsistencies must be detected and the DM must change its preferences. In the second case we increase progressively the additivity of the capacities until a solution is found.

The inference of the parameters of the Choquet integral model consists on the elicitation of the utility functions and the capacities:

Elicitation of utility functions The utility functions $u_S(Cl^{\text{season}}(e, S))$ (where $S \in \{S1, S2, S3\}$) corresponding to the criteria of our problem Cl^{season} which will be aggregated with Choquet integral must be commensurate. This requirement leads to put for each criterion $u_S(1) = 0$ (the worst evaluation) and $u_S(3) = 1$ (the best evaluation). We put then $u_S(2) = 0.5$ after ensuring that the difference of attractiveness $u_S(2) - u_S(1)$ is equivalent to $u_S(3) - u_S(2)$ for $S \in \{S1, S2, S3\}$. We thus have:

$$u_S(Cl^{\text{season}}(e,S)) = \begin{cases} 1 \text{ if } Cl^{\text{season}}(e,S) = 3 \\ 0.5 \text{ if } Cl^{\text{season}}(e,S) = 2 \\ 0 \text{ if } Cl^{\text{season}}(e,S) = 1 \end{cases}$$

Let us remark that for our application the elicitation of the utility functions was not problematic because the criteria are evaluated on homogeneous scales and the set X_i of the possible values of the criteria is small. In general cases, this step is more difficult.

Elicitation of the capacities The elicitation of the capacities was performed as follows:

1- **Collecting the preferential information**: We first asked the experts and the DMs to provide an order representing the importance of criteria (which season is important?) in order to build a "relevant" set of learning examples. We built 14 fictitious trains. This set of learning examples was a set of pairwise comparisons of some of these trains. We asked then the experts and the DMs to give their preferences related to this set.

2- **Interactions among criteria**: We have transformed the set of pairwise comparisons into linear constrains and we tried to find additive capacities ($k = 1$) which corresponds to the weighted sum. When solving a linear programming with these constraints, we found no solution. The reason of such failure may be the presence of some types of interactions among criteria. Hence we decided to test a 2-additive model.

3- **Aggregation/disaggregation procedure**: In order to find the capacities which best restore the preferences, we have used an aggregation/disaggregation procedure. We have first fixed the additivity to 2 (interactions only among pairs of criteria) and we tried to find 2-additive capacities by an approach proposed by Marichal and Roubens [14] which aims at solving a linear programming where the objective function to be maximized is the minimal difference between the Choquet integrals of the compared alternatives. The linear programming have the following form:

$$(LP) \begin{cases} \max f = \epsilon \\ C_\mu(a_1) - C_\mu(b_1) \geq \delta_1 + \epsilon \\ C_\mu(a_2) - C_\mu(b_2) \geq \delta_2 + \epsilon \\ . \\ . \\ . \\ \mu(\emptyset) = 0 \\ \mu(F) = 1 \\ \mu(S) \leq \mu(T), \forall S \subseteq T, \forall T \subseteq F \end{cases} \quad (7)$$

The linear programming gave us several feasible solutions representing the capacity values of the Choquet integral. We chose the first solution and used it in order to obtain a total weakorder of our 14 fictitious trains. We asked then the experts and the DMs if this weakorder was in accordance with their preferences (see Fig. 3). The answer was negative because:

- there were trains which were dominated by others while they had the same overall value (the alternatives O2 and O9 in Fig. 3),

- there were some trains (on which the DMs were not asked to express their preferences) having different overall values while they were considered as equivalent by the experts (the alternatives O2 and O8 in Fig. 3).

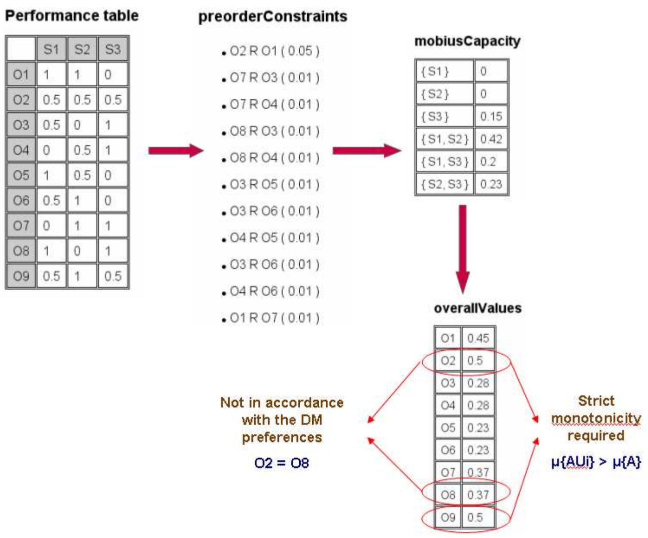

Figure 3. *The first step of the aggregation/disaggregation procedure*

Ones the inconsistencies were identified, we added the corresponding constraints and we solved a new linear programming (LP). Figure 4 represents the results of this second step. The new capacity values obtained by solving (LP) provided to a new ranking of trains which was in accordance with the DMs preferences.

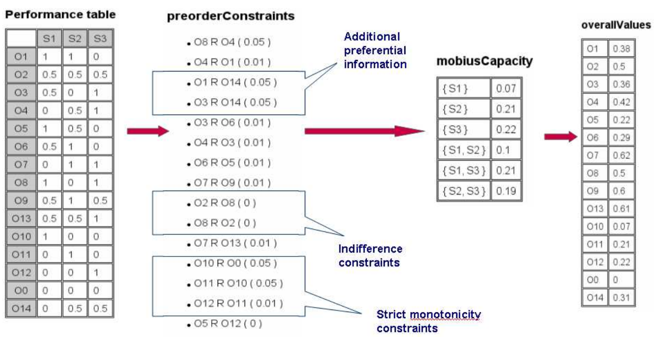

Figure 4. *The second step of the aggregation/disaggregation procedure*

After identifying the capacities, one can proceed to an analysis of the preferential information by computing the *interaction Indices* I_μ and *Shapley values* Φ_μ in order to better understand the nature of interactions among criteria or to have an idea about the intensity of the importance of criteria. We can see in Fig. 5 that all interaction indices are positive which means that the criteria are rather complementary. The interaction between $S1$ and $S3$ is the most important. We can see also that the thermal comfort in the season $S3$ is the most important one and $S1$ is less important than $S2$ and $S3$. Our experts said that this analysis corresponded very well to their intuition.

interactionValues

{S1, S2}	0.1
{S1, S3}	0.21
{S2, S3}	0.19

shapleyValues

S1	0.22
S2	0.36
S3	0.42

Figure 5. *The interaction indices and Shapley Values*

4 Conclusion

In this paper we presented a real world application. Our application shows how we constructed the hierarchical model representing the thermal comfort in the high speed trains, how we chose the aggregation methods and how we elicited the parameters of these methods.

We wanted to point out that a real world application could need several aggregation steps and each step could require a different aggregation method. The choice of the aggregation method must be done with the DMs and experts using a guided approach.

The SNCF insisted on the fact that the results of the application must be in accordance with the perception of train passengers. We thought that for this purpose an elicitation method using some comparison examples coming from surveys with passengers is very adequate.

Sometimes the results obtained by eliciting all the parameters using some examples may provide some unexpected results. For instance if all the parameters of Electre Tri (the weights, the profiles, the thresholds) are elicited all together, one can obtain importance weights which are in contradiction with the intuition of the DMs since they depend also on the profiles. For this reason we think that if some of the parameters can be elicited directly with experts, we have to use these elicitations and then complete them by using more sophisticated methods based on comparison examples.

An aggregation/disaggregation approach (step 1 : proposing comparison examples, step 2 : using them in order to determine some parameters, step 3: presenting some new results to the DM using the results of the second step, step 4 : integrating the comments of the DMs on the step 3 in order to better determine parameters, ...) is appreciated by the DMs and the experts.

REFERENCES

[1] R. Azibi and D. Vanderpooten. Construction of rule-based assignment models. *European Journal of Operational Research*, 138(2):274–293, April 2002.

[2] C.A. Bana, E. Costa, and J.C. Vansnick. A theoretical framework for Measuring Attractiveness by a Categorical Based Evaluation TecHnique (MACBETH). In *Proc. XIth Int. Conf. on MultiCriteria Decision Making*, pages 15–24, August 1994.

[3] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory. *Information Sciences*, 24:205–210, 1986.

[4] P.O. Fanger. *Thermal comfort*. McGraw-Hill, New York, 1973.

[5] J. Figueira and B. Roy. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139:317–326, 2001.

[6] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of methods for capacity identification in choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, 186(2):766–785, 2008.

[7] S. Greco, B. Matarazzo, and R. Slowinski. A new rough set approach to evaluation of bankruptcy risk. In C. Zopounidis, editor, *Operational tools in the management of financial risks*, pages 121–136. Kluwer Dordrecht, Dodrecht, 1998.

[8] S. Greco, B. Matarazzo, and R. Slowinski. The use of rough sets and fuzzy sets in MCDM. In T. Gal, T. Stewart, and T. Hanne, editors, *Advances in MCDM models, Algorithms, Theory, and Applications*, pages 14.1–14.59. Kluwer Academic, Dordrecht, 1999.

[9] I. Kojadinovic. Minimum variance capacity identification. *European Journal of Operational Research*, 177(1):498–514, 2007.

[10] I. Kojadinovic. Quadratic distances for capacity and bi-capacity approximation and identification. *4OR: A Quarterly Journal of Operations Research*, 5(2):117–142, 2007.

[11] I. Kojadinovic, J-L. Marichal, and M. Roubens. An axiomatic approach to the definition of the entropy of a discrete choquet capacity. *Information Sciences*, 172:131–153, 2005.

[12] Ch. Labreuche and M. Grabisch. The choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137:11–26, 2003.

[13] J-L. Marichal. An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE Tr. on Fuzzy Systems*, 8(6):800–807, 2000.

[14] J-L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124:641–650, 2000.

[15] B. Mayag, M. Grabisch, and Ch. Labreuche. A representation of preferences by the choquet integral with respect to a 2-additive capacity. *Theory and Decision*, 71(3):297–324, 2011.

[16] C. Melltet and M. Hernandez. Projet ACONIT, phase 3 - Analyses des donnes : valuation des indices de confort et recherche d'interaction entre modalits, rapport d'essais AEF, tomes 1 6. Internal report, SNCF, AEF, 2006.

[17] P. Meyer and M. Roubens. Choice, ranking and sorting in fuzzy multiple criteria decision aid. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 471–503. Springer Verlag, New York, 2005.

[18] V. Mousseau, J. Figueira, and J.P. Naux. Using assignment examples to infer weights for ELECTRE TRI method: Some experimental results. *European Journal of Operational Research*, 130(2):263–275, 2001.

[19] V. Mousseau and R. Slowinski. Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, 12(2):157–174, 1998.

[20] T. Murofushi and M. Sugeno. An interpretation of fuzzy measure and the choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets and Systems*, 29:201–227, 1989.

[21] M. Mzali. Projet ACONIT, Synthse phase 2: Exploration du confort global. Internal report, SNCF, Direction de l'Innovation et de la Recherche, Novembre 2005.

[22] M. Mzali and C. Talotte. Projet ACONIT, phase 3 - Analyse des corrlations entre les mesures perceptives et mesures physiques - Volume 2 : Bilan de l'analyse des vnements de confort et du confort global et perspectives. Internal report, SNCF, Direction de l'Innovation et de la Recherche, 2009.

[23] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.

[24] B. Roy. The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*, 31(1):49–74, 1991.

[25] M. Sugeno and T. Murofushi. Choquet integral as an integral form for a general class of fuzzy measures. pages 408–411, Tokyo, Japan, 1987. 2nd IFSA Congress.

[26] S. Sgrtain. Confort thermique dans les trains, Synthse du projet de recherche exploratoire. Internal report, SNCF, Direction de l'Innovation et de la Recherche, 2009.

[27] S. Sgrtain. Projet ESTHER-synthse des lot 1 et 2: Diagnostic perceptif et physique du confort thermique dans les trains. Internal report, SNCF, Direction de l'Innovation et de la Recherche, 2011.

[28] A. Ngo The and V. Mousseau. Using assignment examples to infer category limits for the ELECTRE TRI method. *Journal of Multi-Criteria Decision Analysis*, 11(1):29–43, 2002.

[29] R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18(1):183–190, 1988.