

## Linguistic Form of Multicriterial Evaluation

B e ě v á ř Václav

Research Institut of Water Management, Prague  
Výzkumný ústav vodohospodářský v Praze  
Rohanský ostrov, 186 00 Praha 8 - Karlín  
ČSSR

Implementation of different variants of water economy measures /just as implementation of other technical measures/ will bring about a different response in systems whose part it is going to become - in preparation, during the implementation proper and during the subsequent exploitation. The best variant is a variant which accomplishes its mission and at the same time causes an optimum complex response in the said systems.

In a decision-making system it is also necessary to solve the question "how and with which tools can we characterize the response of individual evaluated areas - of the systems in question?". If the requirement that the characteristics are sufficiently representative is accompanied with a requirement that these partial characteristics be pooled to facilitate complex assessment, it is obvious that the search for such a unit which can be pooled starts from the system of evaluation areas and from the way of priority determination; this way is conditioned by the basic criterion of evaluation and choice. In the System of Complex Variants Evaluation /SCVE/ elaborated in the Research Institut of Water Management in Prague, these input connections resulted in scoring scales application. Moreover, a question was put which influenced further development of this evaluation tool: "How does the expert reach his scoring answer?".

The answer is that the basis of expert experience and delivered documents and/or other calculations and deliberations has in its primary shape a verbal form. In function of expertise requirements this primary /primary in the sense of transformation but not in the sense of cognition/ view is "translated" or transformed into the form of scoring or otherwise. Within SCVE deliberations the possibility of a primary use of these original expertise views was studied. Findings from the elaborated application of theory of linguistic variable were used. On their basis, fundamental research was carried out about significance of the words used in the given relationships.

Two sets of verbal evaluation were established. The set A consists of statements and terms characterising the quality of the solution - e.g. outstanding, good, bad, catastrophic, etc. The set B consists of statements and terms characterising the size of total positive or negative effects of the evaluated solution, e.g. huge, small, negligible, none, etc.

About 200 persons were asked to answer. The respondents determined themselves how they conceived the relationship between the individual statements and the auxiliary 50-point scoring scale. The subjective perception of the adherence, appurtenance function  $A_x$  /  $A_x \in \langle 0; 1 \rangle$  / of scoring scale adhesion to individual words was graphically illustrated.

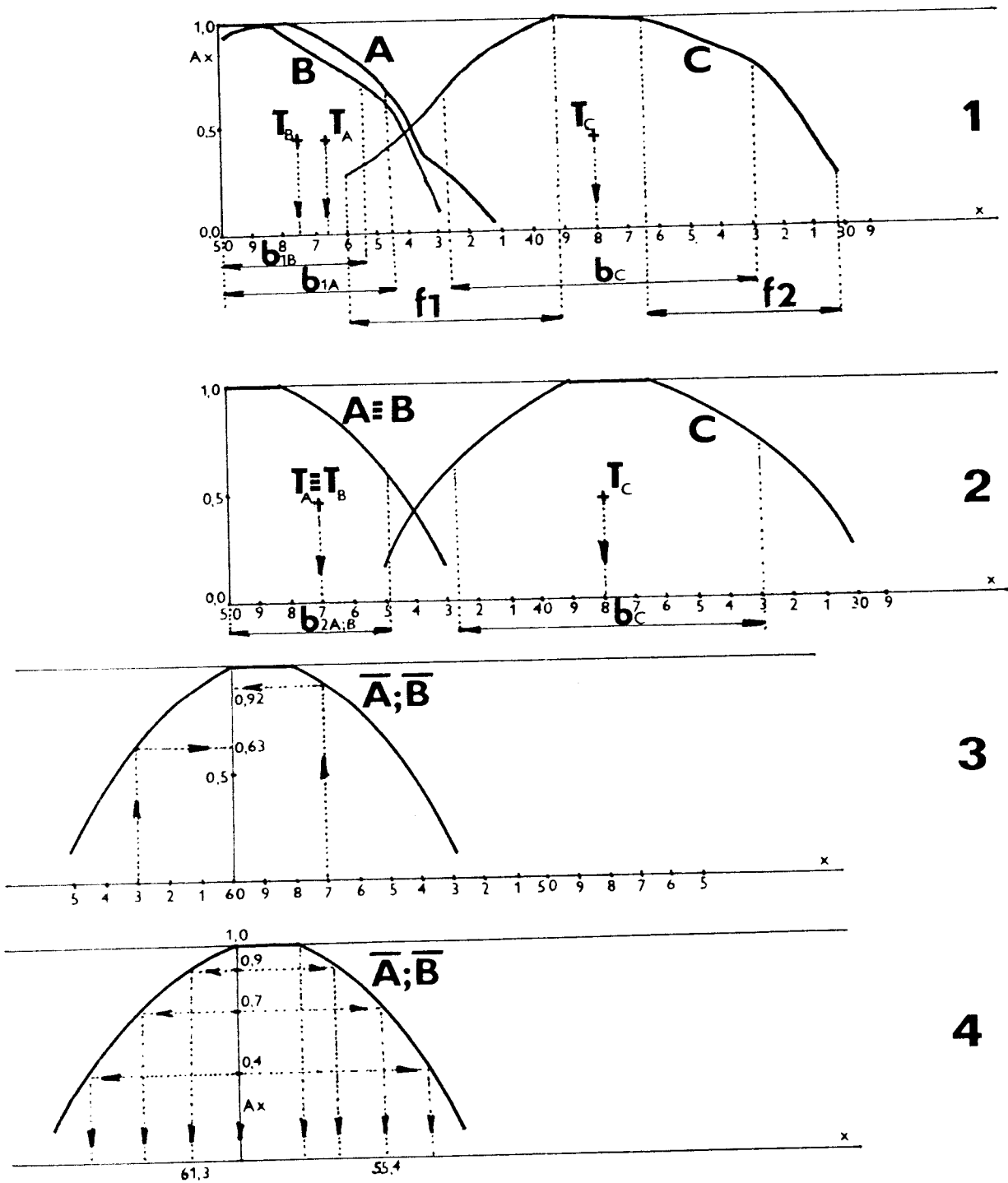
During further processing the examination centred on the question whether there existed understanding and use of words meaning conditioned by age, profession or other group adhesion /e.g. men - women/. The result was negative. It was established that the "location" of words on the auxiliary scoring scale, in width and shape of the respective appurtenance function curves there exist differences. Differences existed in one and the same person during the course of time - in the sense of absolute measurement. On the other hand, in relative comparison an unexpected agreement appeared /order of statements, mutual differences and similarities, etc./. This result is a confirmation of the initial hypothesis - "translation" of verbal elements into a quantification scale. It is carried out individually and without any standard. There is no reason to doubt about the internally identical understanding and use of words at a sufficiently accurate level; everyday interhuman communication is convincing enough.

During further results processing the first task was to complete the existing statements and terms so that they cover uniformly and fully the entire auxiliary scoring scale, to calculate average parameters of individual statements - terms /the centre point of terms -  $\underline{T}$ , the average width of terms -  $\underline{b}$  and the width of terms in area with function of appurtenance  $A_x < 1$  -  $\underline{f}$ /; fig. 1.

Suppressed were also some minor anomalies and deviations from a continuous form of curves of the function of appurtenance as well as negligible differences among some concrete terms - fig.2. All this

was based on the assumption that there exists a dependence of these unimportant anomalies and differences on a concrete and single group research.

The last correction resulting from the needs of computation processing /SCVE is programmed as an interactive system using the HP 1000/40 computer/ was the replacement of S-curves or marginal terms by a uniform bell-like shape - fig.3. For that reason, the scoring



scale  $\langle 0;50 \rangle$  was replaced by a scale from 10 to 60 scores. The interval limits are more or less crossed by individual terms. The advantages of a uniform shape of curves and the possibility to determine on each of them formally adequate points is quite considerable from the point of view of computation - fig.4.

A kind of phraseological semi-products was created. Their task is to permit simultaneous use of two random terms /statements/ with the objective to express aspects of contradiction, divergence, undecision or intermediacy when assessing and evaluating more complex realities. Up till now we have 9 sentences some of them looking as follows: "From one important point of view -----, in most of the cases, however -----" or "Rather ----- than -----". Instead of the symbol ----- the experts will introduce concrete terms or statements.

The linguistic variable prepared for the SCVE differ from illustrative forms of these variables taken from theoretical examples of application.

The principle of primary terms and of the terms derived from them by linguistic operators or linguistic operations is not maintained. In the SCVE a set of terms was chosen which belong together by their meaning. In all these terms their meaning was determined just as in the primary terms. It was expressed in the form of fuzzy sets of elements of the auxiliary variable. The effort was to create a linguistic variable only from terms currently used in real conversations.

Another distinction from theoretical linguistic variables is an intentional use of synonyms. It is expected that in the future the experts will have the possibility of quite free expression.

The third difference is the number of assessment possibilities. In the first set we have up till now 28 terms and in the second one 23 terms. Together with the 9 sentences this represents theoretically 6 832 or 4 577 possibilities of different statements and expressions of the expert. The number of practically usable combinations is, naturally, lower, but it is still sufficient for the expert not to know the precise structure of the linguistic tool. The expert will use a term on the basis of its meaning and will not be influenced by its position in a free scale created by the terms. Only such an approach justifies the sense of using linguistic forms of evaluation

and assessment. In the way, least distorted evaluations with the highest wealth of nuances are obtained from the expert.

Thus the linguistic assessment satisfies all the requirements for representative and pooled characterisation of the foreseen reactions in areas affected by the implementation of the proposed measure. It can be regarded as an ideal evaluation technique in complex assessments; it accomplishes fully its role and thanks to its natural substance it enables the expert to express nearly accurately all his views with fine nuances and not to distort the multidimensional character of the evaluated reality.

Available experience of the RIWM Prague with the SCVE application and, in particular, with the use of linguistic assessment is, on the whole, good. However, it is necessary to take into consideration the absolutely strange form of the method and the mistrust resulting from technical orientation at figures and quantification. It is, therefore, quite advisable to instruct the experts during a discussion with adroit defence and explanation of the principles of linguistic assessment. The acceptance of the method is more a psychological issue than a matter of real material need.

In conclusion, one last idea concerning the interpretation of results. The verbal statements of experts have only a limited accuracy. They represent the forecast reality only in an unfocused, "fuzzy" way. This, however, corresponds to the real quality of cognition. In the best case the results should have the same, i.e. the "fuzzy" character which is most accurate way of pooling the available findings and knowledge; consequently, it is the only real starting point for final decision making. A categorical interpretation of numerical results is not only false, but it is, at the same time, an unserious and confusing abuse of the formal aspect of the results; it endows the relations among variants of undefendably accurate value as well as the decision-maker with a false feeling of higher certainty.