STATISTICAL TOOLS FOR THE CALIBRATION OF TRAFFIC CONFLICTS TECHNIQUES

Contribution to Proceedings of the Third International Workshop on
Traffic Conflicts Techniques, organised by the International Committee
on Traffic Conflicts Techniques ICTCT, Leidschendam, The Netherlands,
April 1982.

1. INTRODUCTION

Just before the Second International Workshop on Traffic Conflicts
Techniques, held in Paris, May 1979, an international experiment took
place in Rouen. The purpose of this experiment was to compare the
results of various conflict techniques from different countries. This
experiment showed that, in general, from each technique the same
conclusions were reached with regard to the problems of safety at two
intersections in Rouen. However, it was not always clear how the
observations led to these conclusions. From the discussions at the
Workshop in Paris and extended discussions that took place afterwards,
it was concluded that an international calibration study would be very
informative. This calibration study should be concerned with a
detailed comparison of conflict scores. From such a detailed comparison
conclusions can be drawn about the extent to which different techniques
lead to different results. In order to improve existing techniques, a
comparison of one's own results with the results of other techniques
is of value for each conflict technique. Furthermore, it is important
to know the similarities and dissimilarities of techniques in order to
evaluate validation studies of other techniques and the consequences
of these validation studies for one's own results.
Efforts are made to realise such a calibration study.
This note describes a data-analysis technique that, in the author's
opinion, is an efficient tool for the analysis of the data that will
be collected in such an experiment. An example will be given from
which it is clear that the analysis, which is in fact much more
general and not restricted to the narrow context of the calibration
experiment, will give all the relevant information in this case.

2. GENERAL DESCRIPTION OF DATA AND ANALYSIS

We will not give here a detailed description of the planning of a
calibration study, but comment only on the fundamental structure of
the data and the analysis of this data that should result from such
an experiment.
Essentially, the problem is to measure the same objects (conflict
situations) with different measurement techniques (conflict observation
teams). Each team has to decide whether or not a number of traffic
situations are conflicts, and if so, how serious these conflicts are.
Each team may use its own scale to measure this seriousness (e.g. a
three-point scale or a five-point scale, or even a continuous scale).
Cues that are used by different teams, in order to evaluate the
seriousness of a conflict, may differ from team to team.
A further complication is the lack of an objective norm for the
seriousness of conflicts such as there is in e.g. experiments concern-
ing the estimation of velocities.
Furthermore, in almost all cases where techniques are used in practical
situations, the classification system is more or less subjective and
depends on judgements of observers. Here it is assumed that all teams
measure the seriousness of conflict situations.

Technically speaking each of m teams measures each of n objects (some
values may be missing). This results in a m times n matrix of scores.
We want to investigate to what extent it is possible to scale all n
conflicts on one dimension and at the same time to rescale the response
classes of each team on the same dimension such that maximum homogeneity
is reached between the scores for different teams. This "common"
dimension will be interpreted as the severity dimension of conflicts.
The rescaling of the response classes for each team makes it possible
to compare categories of different techniques with regard to the
seriousness of conflict observations.
If the data will not be described sufficiently by this one-dimensional
representation, then a two-dimensional description may tell us whether
or not severity is judged in a more complicated way.

As mentioned before, a score results from this analysis for each con-
flict. As a second step one can relate these scores to various cues of
the conflicts to find out which cues (that are used explicitly or
implicitly) are relevant and/or redundant.
This last step is not necessary to calibrate techniques. However, it
will be useful to explain agreement and disagreement about the scores
of different teams. We can think of cues such as traffic volume, type
of vehicles involved, velocities, decelerations, manoeuvres, time to
collision, post-encroachment time etc.

3. THE ANALYTICAL TECHNIQUES

3.1. Introduction

As has been described before, the analysis consists of two steps. First
it will be investigated to what extent conflict measurement of teams
can be compared to each other with regard to the severity of conflicts.
The technique that is proposed for this step is a principal components
analysis for classified data. The computer programme is called HOMALS
(homogeneity analysis by means of an alternating least-squares solution).
The analysis results in severity scores for all situations that are
investigated and rescaled values for the severity classes of each team
on the same dimension.
The second analysis relates the severity scores to objective measures
of the situation that will be based on the data collected by means of
video recording and other registration techniques. In this case multiple
regression techniques seem to be useful. If classified data are related
to the severity scores or if we want to relate the objective measures
directly to the classified conflict data of the teams, then the use of
CANALS is proposed. CANALS if a computer programme in which canonical
regression analysis (and multiple regression as a special case) is
generalised to classified data in the same way as HOMALS is a generali-
sation of principal components.

HOMALS and CANALS

Technically speaking, HOMALS is a principal components analysis for
classified data.
If one applies principal components analysis directly to classified data,
one violates the condition that all variables must be measured on an
interval scale. To solve this scaling problem one uses the fact that
this condition is always satisfied for binary data: any rescaling of
two classes into other ones is possible by a linear transformation. If
we use classes of characteristics as if these are itself characteristics
and rescore the objects on these new "characteristics" with one if they
are and zero if they are not in that class of the prior characteristic,
then a new data matrix results with only one's and zero's containing

the same information. E.g. if 25 objects are classified according to
two characteristics having 3 and 4 classes respectively, then the
previous 25x2 data matrix can be rewritten in a 25x(3+4) matrix of
one's and zero's. This matrix is singular: the scores in the last
class of each characteristic can be deduced from those of the previous
classes. Therefore from this matrix a 25x5 matrix of binary scores can
be derived that contains all the relevant information. The principal
components analysis applied to this matrix results in the intended
solution. The weights for the classes can be regarded as scaling fac-
tors.
Therefore, HOMALS may also be regarded as a multi-dimensional scaling
technique.
This generalisation can also be applied to the problem of canonical
regression analysis.
CANALS delivers such a solution for the canonical analysis of classi-
fied data.
HOMALS is a technique that is related to "Analyse des correspondences",
a similar kind of technique developed in France by Benzécri (1973).
A detailed description of these kinds of data-analysis techniques can
be found in De Leeuw (1979) and Gifi (1981).

Example

In order to demonstrate the use of HOMALS we applied this technique
to data from an investigation of Güttinger (1980). He trained observ-
ers to use his conflict technique. Ten observers were asked to score
27 traffic situations on a four point scale. During the training they
got knowledge of results in order to improve their scoring procedure.
We analysed one of the resulting matrices of scores. The data are
taken from Güttinger (1980, Bijlage 12a). The analysis is given in
Appendix 1. We assumed that the observers scored the objects in the
same way and on one simple dimension.
The discrimination measure (a measure of squared correlation, between
the object scores and the rescaled scores for each observer) is
highest for observer 5 (dm = .983) and lowest for observer 9 (dm =
.793).
From this analysis we conclude that the agreement between observers is

rather high. The eigen value λ, the mean discrimination measure representing the degree of homogeneity between observers, is equal to .89. A more-dimensional solution did not add information to this one-dimensional description.

The object scores of the 27 situations with regard to the solution show that situation 4 and 18 are the conflicts which are the least severe and situation 9, 13 and 15 those that are the mose severe. The category scores for each variable show the rescaling for each observer such that the agreement with the common solution is maximal.

A plot of the object scores together with the category scores of observer 5 and 9 is given in Figure 1. The original scores for observer 5 and 9 are also included.

From this plot we see that the scores of observer 5 are in complete agreement with the ordering derived from the solution. The categories of observer 9 (especially the categories 2 and 3) show inconsistencies with this order.

From this analysis we conclude that all observers agree to a large extent. Furthermore, special scoring problems of observers, such as observer 9, also become clear from this analysis.

As a second step we could have related the observation scores to characteristics of the traffic situations in order to investigate with CANALS why the response behaviour of observer 9 differs from the other observers.

References

Benzécri, J.P. a.o. Analyse des données. Paris, Dunod, 1973.

De Leeuw, J. Canonical analysis of categorical data. Thesis. Leyden State University, 1979.

Gifi, A. Non-linear multivariate analysis. Leyden State University, Department of Data theory, 1981.

Güttinger, V.A. Met het oog op hun veiligheid. Thesis. Amsterdam State University, 1980.

observation nr	4,18 2,11 14 20 24,25	6,12,26,27 3,8	17	19 5	22	1	12 7,21	10,16 9,13,15
observation scores	2 2 1 1 2	4 1 1	1	1 1	1	1	1 2	2 3
cat. scores for subject 5	1	2		3			4	
cat. scores for subject 9	1	2,2,2 2	3				4	
original scores for subject 5	1,1,1 1	2,2,2 2,2	2	3 3	3	4	4,4 4,4	4,4,4
original scores for subject 9	1,1,1 3,3	3,2,3,2 3,3	3	3 3	—	4	3	4,3 4,4 4,4

Figure 1. Plot of the 27 traffic situations from Güttinger (1980) on one conflict dimension together with the scoring of observer 5 and 9.

```
   *
   *  VARIABLES
   *
   *
   *      1   2   3   4   5   6   7   8   9  10
OBJECTS  *******************************************
     1 *  4   4   4   3   4   4   4   3   3   3
     2 *  1   1   1   1   1   1   1   1   1   2
     3 *  2   3   3   2   2   2   2   2   2   2
     4 *  1   1   1   1   1   1   1   1   1   1
     5 *  3   3   3   3   3   2   3   3   3   3
     6 *  2   2   2   3   2   2   2   2   2   2
     7 *  4   4   3   4   4   4   3   4   4   4
     8 *  3   2   3   2   2   2   2   2   3   2
     9 *  4   4   4   4   4   4   4   4   4   4
    10 *  4   4   4   4   4   4   4   3   4   4
    11 *  1   1   1   1   1   1   1   1   1   2
    12 *  2   3   3   2   2   1   2   2   3   2
    13 *  4   4   4   4   4   4   4   4   4   4
    14 *  1   2   1   1   1   1   2   1   1   1
    15 *  4   4   4   4   4   4   4   4   4   4
    16 *  4   4   4   4   4   4   4   3   4   4
    17 *  2   3   3   2   2   3   2   2   3   3
    18 *  1   1   1   1   1   1   1   1   1   1
    19 *  3   3   3   3   3   3   2   3   3   3
    20 *  1   1   1   1   1   2   1   1   1   2
    21 *  4   4   4   4   4   4   4   4   3   4
    22 *  4   1   2   4   3   3   4   3   4   4
    23 *  4   4   3   4   4   4   3   3   4   4
    24 *  2   1   1   1   1   1   1   1   3   1
    25 *  2   1   1   1   1   1   1   1   3   1
    26 *  2   1   2   1   2   4   1   2   3   2
    27 *  2   3   2   2   2   2   2   2   2   2
```

MARGINAL FREQUENCIES:

```
          *
          *  CATEGORIES
          *
          *
          *   MISSING     1     2     3     4
VARIABLES *******************************************
     1 *      0          6     8     3    10
     2 *      0          9     3     6     9
     3 *      0          8     4     8     7
     4 *      0          9     5     4     9
     5 *      0          8     7     3     9
     6 *      0          8     6     3    10
     7 *      0          8     8     3     8
     8 *      0          8     7     7     5
     9 *      0          6     3    10     8
    10 *      0          5     9     4     9
```

DIMENSION	EIGENVALUE
1	0.8933

DISCRIMINATION MEASURES PER VARIABLE PER DIMENSION

```
            *
            * DIMENSION
            *
            *
            *           1
VARIABLES ************
        1 *       0.953
        2 *       0.842
        3 *       0.829
        4 *       0.932
        5 *       0.983
        6 *       0.826
        7 *       0.913
        8 *       0.940
        9 *       0.793
       10 *       0.922
```

THE OBJECT SCORES ARE:

```
            *
            * DIMENSIONS
            *
            *
            *.          1
OBJECTS     ************
        1 *       0.94
        2 *      -1.19
        3 *      -0.46
        4 *      -1.24
        5 *       0.20
        6 *      -0.47
        7 *       1.21
        8 *      -0.40
        9 *       1.37
       10 *       1.32
       11 *      -1.19
       12 *      -0.49
       13 *       1.37
       14 *      -1.15
       15 *       1.37
       16 *       1.32
       17 *      -0.23
       18 *      -1.24
       19 *       0.13
       20 *      -1.13
       21 *       1.22
       22 *       0.72
       23 *       1.15
       24 *      -1.06
       25 *      -1.06
       26 *      -0.50
       27 *      -0.49
```

Category	Category quantifications									
	Observer									
	1	2	3	4	5	6	7	8	9	10
1	-1.19	-0.88	-1.16	-1.09	-1.16	-1.08	-1.08	-1.16	-1.19	-1.15
2	-0.59	-0.68	-0.19	-0.41	-0.43	-0.46	-0.45	-0.43	-0.47	-0.70
3	-0.03	-0.22	0.14	0.20	0.35	0.20	0.85	0.82	-0.13	0.26
4	1.20	1.25	1.27	1.23	1.25	1.08	1.20	1.31	1.23	1.23