Subway Station Diagnosis Index (SSDI): A Condition Assessment Model

Nabil Semaan

A Thesis
In the Department of
Building, Civil, and Environmental Engineering

Presented in Partial Fulfilment of the Requirements
For the Degree of Master of Applied Sciences at
Concordia University
Montreal, Quebec, Canada

December 2006

Your file Votre référence
ISBN: 978-0-494-28892-4
Our file Notre référence
ISBN: 978-0-494-28892-4

ABSTRACT

Subway Station Diagnosis Index (SSDI): A Condition Assessment Model

Nabil Semaan

Condition assessment of subway stations is a major issue facing public transit authorities worldwide. In 2002, the *Société de Transport de Montreal* (STM) valued its subway station replacement cost at 2.6 CAD Billions. While its stations are becoming aged, the STM requires a rehabilitation budget of 643.6 CAD Million between the year 2006 and 2010. Nevertheless, the STM lacks a planning strategy reflecting this increase. The principal obstacle to the development of effective planning strategies is the lack of condition assessment models of subway stations.

This research develops a condition assessment model, the 'Subway Station Diagnosis Index (SSDI)', and a scale. The SSDI model is used to diagnose a specific subway station and assess its condition using an index (0 to 10). Based on the SSDI, the condition scale describes the station's condition state, its deterioration level (%), and proposed consequent actions.

The new model identifies and evaluates the different functional/operational criteria for subway stations; mainly structural, architectural, mechanical, electrical, security and communications criteria. It uses specific decision analysis tools in order to evaluate a 'Functional Diagnosis Index' (FDI), and a global 'Station Diagnosis Index' (SDI). In other words, the SSDI model uses the Analytical Hierarchy Process (AHP) in order to determine the criteria weights. It also utilizes the Preference Ranking Organization METHod of Enrichment

Evaluation (PROMETHEE) in order to aggregate the multi-criteria. Finally, the SSDI applies the Multi-Attribute Utility Theory (MAUT) to determine the FDI and the SDI values.

Data were collected from experts through interviews, phone calls and questionnaires as well as STM inspection reports. The targeted interviewees were transit authority experts in both Canada and U.S.A. Statistical and sensitivity analyses were performed on the collected data. Analyses show that structural and security/communication criteria are the most important (36.1% and 27.3%, respectively).

The newly developed model is applied to seven stations from the STM network. Results show that these stations are deficient, with an average SDI of 4.4 out of 10. Ranking of the seven stations is compared to that of PROMETHEE, which shows similar results. In addition, the SDI values are confirmed by STM engineers with 80% agreement. This research is relevant to industry practitioners (management, engineers, and field inspectors) and researchers, since it develops, a multi-criteria condition assessment model and scale for subway stations.

Acknowledgement

I wish to express my deepest gratitude and my most sincere appreciation to my supervisor Dr. Tarek Zayed, for his solid support, valuable advice, constructive guidance, and inspiration throughout all the stages of this research. His influence and dedication far exceeds the scope of the work presented in this thesis.

I also would like to thank Mr. Mazen Farran for helping me to get in contact with STM. A special thanks to Mr. Bramo Girgi for helping me in the development of the model automated tool.

Very special thanks to my parents, brother and sisters, for their constant motivation, unfailing support and love.

Last but not least, I would like to express my deep gratitude to my wife Tereza, and two kids Michael and Theodora, for having been my source of strength and inspiration over the last few years. When the going got tough, their encouragement was the reason I kept going. Their patience in putting up with the lifestyle of a student is deeply appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

1.1 Problem Statement and Research Motivation

The goals of every subway transit authority are to augment the level of reliability, public safety, and achieve a better level of service. The goal behind these objectives is to attract more users and to ensure their safety. The 'Société de Transport de Montréal' (STM) has estimated the replacement value of its network at 4.6 CAD Billion in 2002, out of which 2.6 (56.5%) CAD Billion are only assigned to stations. Therefore, stations represent a major section of any subway transit network. A significant number of subway stations are aging and hence surpassing their functional life. If stations are showing serious deterioration, they become unsafe to the public. In 2002, the STM estimated a maintenance budget of 81 CAD Million in order to maintain the stations at a minimum level of repair. According to the STM president, Mr. Pierre Vandelac, in an interview in La Presse newspaper (2006), it was confirmed that the STM stations "will deteriorate rapidly in the next five years, with a decrease in the level of public safety, coupled with more metro stoppages".

The major problem that faces STM and most transit authorities is the lack of proper rehabilitation planning for their stations. This includes budget allocation, investment plans, and financing. The lack of proper rehabilitation planning is directly linked to the lack of assessment tools stations' condition. Although transit authorities have typical inspection reports, models for interpreting these reports and assessing stations' condition are not available. Previous research in this field has provided ranking methods for the stations, prioritizing stations for rehabilitation, but these methods fail to provide a condition index (level of deterioration) to each station. Therefore, there is an urgent need to develop an index in order to assess the condition of subway stations, and rank them according to a unified and universal scale.

1.2 Research Scope and Objectives

The research focuses on the functional condition assessment of subway stations, which is applicable to any transit authority. The main objective of this research is to develop a condition assessment model and scale for subway stations: the Subway Station Diagnosis Index (SSDI). In order to fulfill this main objective, the following sub-objectives are identified:

1. Identify the functional/operational criteria.
2. Develop a model to evaluate the weights of these criteria.
3. Design a Station Diagnosis Index (SDI) and Functional Diagnosis Index (FDI) models.
4. Perform sensitivity analysis for the elements of the developed model.
5. Develop an automated tool to implement the new developed model.

The new developed model has a key role in managing maintenance and repair activities for subway stations. It is developed in such a manner that it is easy and fast to implement.

1.3 Methodology of the Research

The methodology adopted in this research is divided into four main steps, as illustrated in Figure 1.1:

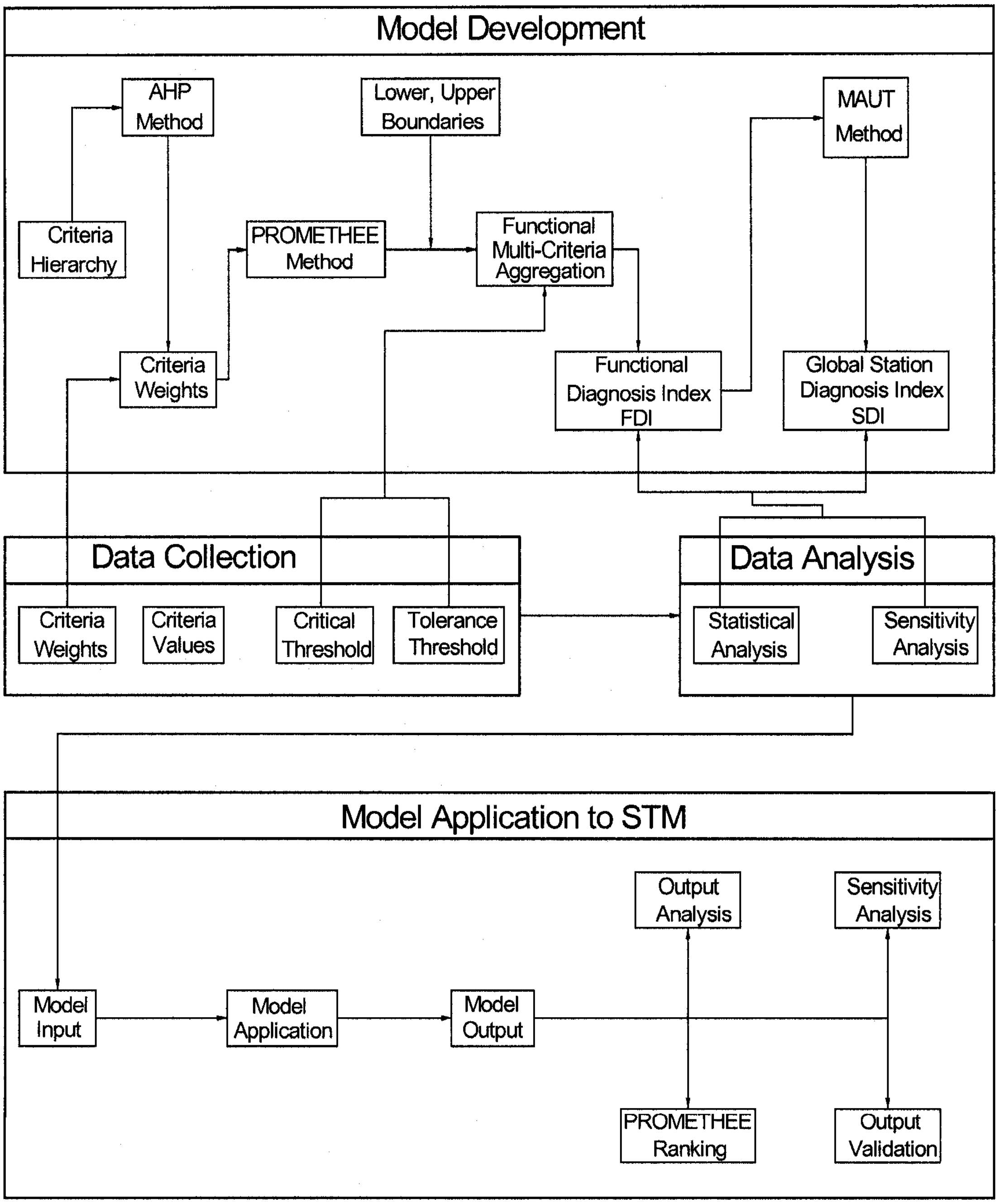


Figure 1.1 Research Methodology

I. Model development:

   a. Define the criteria and their hierarchy.

   b. Apply the Analytical Hierarchy Process (AHP) in order to evaluate the criteria weights.

   c. Apply the Preference Ranking Organization METHod of Enrichment Evaluation (PROMETHEE) outranking method. PROMETHEE results in a rank on an ordinal scale.

   d. Introduce lower and upper limits. This transforms the floating ranking of PROMETHEE into a rank between two absolute limits

   e. Aggregate the functional multi-criteria according to the PROMETHEE method and the two boundaries. This comprises the use of the pseudo-criteria concept, which involves the Critical Threshold and the Tolerance Threshold and a Generalized Preference Function.

   f. Evaluate the Functional Diagnosis Index (FDI) according to the PROMETHEE method.

   g. Apply the Multi-Attribute Utility Theory (MAUT) to FDI and evaluate the global Station Diagnosis Index (SDI).

II. Data Collection:

   a. Collect criteria values and weights from inspection reports, and questionnaires, respectively.

   b. Collect critical and tolerance thresholds from the same questionnaires.

III. Data Analysis:

a. Perform statistical and sensitivity analyses.

b. Revisit and complete the FDI and SDI condition scale tables.

IV. Model application to STM:

a. Introduce the data collected in the SSDI model application as the input.

b. Apply the SSDI model to seven STM stations.

c. Analyse the results of the SSDI model for the seven stations.

d. Compare the ranking of the seven stations from the SSDI model to the ranking from the PROMETHEE outranking method.

e. Perform sensitivity analysis.

f. Verify the SSDI model.

1.4 Thesis Overview

The thesis is divided into six chapters:

- The first chapter consists of the introduction.
- The second chapter is the literature review. It is divided into two parts. It describes some of the existing condition assessment methods for different transit authorities, as well as for bridges, pavement and underground water pipes. In addition, it describes the different methods used in the SSDI model concerning the decision analysis, chiefly AHP, PROMETHEE, and MAUT.
- Chapter three illustrates, in detail, the SSDI model. It identifies and defines the functional criteria of the model, and the corresponding hierarchy. This

is followed by the evaluation of the criteria importance weights using AHP. In addition, the mathematical process of PROMETHEE is described with the introduction of the boundaries. This is followed by an evaluation of the ranking procedure. Finally, using MAUT the evaluation procedure of Diagnosis indices is performed. The chapter concludes with the SSDI theory.

- Chapter four analyzes the data collected for the SSDI model. The data were collected from questionnaires, interviews, phone calls, inspection reports, and a description of the STM rehabilitation programs. The collected data comprise the criteria weights, the Critical Threshold and the Tolerance Threshold. This chapter also presents statistical and sensitivity analyses.
- Chapter five comprises the application of the SSDI model to seven STM stations, and an analysis of the results. A comparison to PROMETHEE is performed and model verification is presented.
- Chapter six presents the SSDI model as an automated tool.
- The thesis ends with the conclusions and recommendations chapter.

2 Literature Review

2.1 Condition Assessment Models for Transit System

Although metro stations in major cities worldwide are aging, a unified condition assessment model has yet to be developed. Each transit authority, depending on its need, developed preliminary rating methods according to their own management plans. An overview of the condition assessment strategies is listed in the following sections, for the most important cities in Canada, U.S.A. and Europe.

2.1.1 Société de Transport de Montréal (STM)

The maintenance of STM stations is performed through a specialized team, known as the 'planification team'. This team connects between a maintenance unit composed of different trade technicians and an engineering unit. It also runs day-to-day minor inspections related to trades (architectural, plumbing, mechanical, electrical, water, etc…) and planify the daily work of the maintenance technicians unit. The engineering unit amends the bigger scale inspection to outside consultants. It also designs, plans and coordinates with the planification team for larger maintenance and/or renovation jobs. This configuration of responsibilities can change from one year to another and depends greatly on the decisions taken by the planification and engineering teams.

The Montreal metro system is showing signs of aging, and needs proper rehabilitation planning and execution. In the 1990's, a program was implemented to renovate the old stations done in the 60's only. This program was called 'Réno-Stations I', and was managed solely by the engineering unit. Its main purpose was the structural and architectural renovation of the stations. A ranking was needed at this time, since the program included all stations. Another program has started in 2005, which represents a continuation of 'Reno-Station I", and is named 'Réno-Station II'. This latter program consists of the renovation of the remaining stations (24 in total). It is also mainly a structural/architectural renovation work directed by the engineering unit only, and coordinated with the planification team. It neither has a condition assessment procedure, nor ranking norms between the different stations (Reno-Station report, 2005).

Simultaneously, another group in the STM runs a 'Réno-Systèmes' program. This program calls for the repair or replacement of all major equipment (mechanical stairs) and systems directly related to operations (communication and security systems). These systems are controlled in a centralized location, where all information relating to underground network activities converge via a communication system (Reno-Systèmes report, 2004). On occasion, these two programs interfere with each other, or conflict in the prioritization of their rehabilitation plans. Hence, an important action in one program is delayed due to a need for the implementation of another urgent job in the other program.

STM maintenance and operation planification is not based on any condition assessment method, but simply a selection procedure depending solely on the age of the station. In addition, the *Ministry of Transportation of Quebec* (MTQ) has no condition assessment guideline for its metro systems.

2.1.2 California Train Transit System (Cal Train)

The Cal Train transit system, inaugurated in 1864, is one of the oldest systems in the United States. In the 1990's Cal Train had set objectives to improve its stations and thus initiate the station planning process. In 1994, Cal train has developed a specific system for the evaluation of stations and ranking from *excellent* to *poor*: (1) Excellent; (2) Good; (3) Average; (4) Below average; (5) Poor.

The criteria used for the evaluation of the stations are:

- i) Ease of access to and from the station.
- ii) Location of the station and proximity to amenities.
- iii) Availability of parking capacities.
- iv) Ability to use other modes of transportation.
- v) Appearance and cleanliness of the stations.
- vi) Physical and structural condition of the stations.
- vii) Public information, signs, telephones.
- viii) Ticket vending machines.
- ix) Security.
- x) Safety.

The evaluation method adopted was a weighted average of the criteria values (Abu-Mallouh, 2001).

2.1.3 Metropolitan Transit Authority of New York Transit (MTA NYCT)

The MTA NYCT was built in 1904. It is the largest transit authority in the eastern United States (Abu-Mallouh, 2001). In 1995, the MTA NYCT faced many problems with respect to some of its old stations, which forced it to re-evaluate the aging infrastructure develop a ranking system for condition assessment. Each station is ranked in order of priority, by allocating points to each of the factors considered, depending on a rating system:

i) Structural conditions (up to 51 points).

ii) Daily usage (up to 25 points).

iii) Felonies (up to 2 points).

iv) Terminal station (up to 2 points).

v) Intermodal American Disabled Agreement ADA (up to 2 points).

vi) Automatic Fare Control AFC (up to 2 points).

vii) Secured outside funding (up to 2 points).

viii) Potential developer funding (up to 2 points).

ix) Point of interest (up to 2 points).

Points for each factor are added by for each station. The condition of the station assignment depends on total points, as shown in Table 2.1 (Abu-Mallouh, 2001).

Table 2.1 MTA NYCT Stations Condition Point Allocation

Scale	Condition Assignment	Maximum Points
5	Severe deterioration	51
4	Deteriorated condition	41
3	Moderate deterioration	31
2	Minor deterioration	20
1	No repair required	0

Abu-Mallouh (2001) improved the condition assessment point allocation model of MTA NYCT, and developed a Model for Station Rehabilitation Planning (MSRP). The MSRP considers functional factors (structural, mechanical, communications, water condition, and safety) and social factors (daily usage, safety, and Level Of Service). MSRP uses the Analytical Hierarchy Process (AHP) to assign weights for each station, and then uses Integer Programming IP to optimize the fund allocation for rehabilitation. Stations that have a certain weight and budget above certain thresholds assigned by management are eligible for instant rehabilitation (Abu-Mallouh, 2001). MSRP is a model for ranking the stations and not evaluating an index for the station on a fixed condition scale. And since it considers many factors, it is very lengthy to implement. Furthermore, Abu-Mallouh's application of MSRP in his research work was based on purely fictitious data, and thus it was never validated nor applied to real life.

2.1.4 London Transport

In 1990, London Transit's main objective was to improve its system. It developed the Key Performance Indicator (KPI), which evaluates the performance of the station from the point of view of its customers (Tolliver, 1990). Surveys and interviews were performed in order to obtain a direct evaluation of customer satisfaction. Customers were asked to rate 23 items on a scale from 0 to 10, based on the following criteria:

i) Cleanliness.

ii) Information services.

iii) Information on trains, i.e. station services (ticket gates, ease of access to platforms, buying a ticket and the degree of platform crowding.

iv) Safety and security.

v) Train services (crowding, journey time, smoothness of the ride…).

vi) Staff helpfulness and availability.

KPI is an overall weighted average of the 23 measures of evaluation based on the user's satisfaction (Abu-Mallouh, 2001).

2.1.5 Paris Rapid Transit Authority (RAPT)

In 1982, RAPT made a considerable effort in order to develop a selection procedure of the stations that should be renovated. A study was delegated to LAMSADE, University of Paris-Dauphine in France (Roy et *al.*, 1986). The study resulted in a selection procedure that used seven criteria:

i) Platform users.

ii) Transit passengers.

iii) Coordination of works.

iv) Maintenance of wall and roof tiles.

v) Visual aspect of the station.

vi) Level of discomfort.

vii) Environment (RAPT wish to favour stations in rapidly changing and low-income areas).

LAMSADE used the ELECTRE III decision support model and software to rank the stations according to the criteria listed above (Roy et al. 1986). The result of the study is again a rank and not a fixed index for condition assessment.

2.2 Condition Assessment Models for Pavement

Pavement diagnosis and assessment as part of Decision Support System DSS has been studied, developed and even established by authorities. The US army corps of Engineers introduced the 'Pavement Condition Index' PCI, and it received wide acceptance and has been formally adopted as standard procedure by many highway agencies (Sinha and Knight, 2004).

The PCI is a numerical condition-rating index, ranging from 0 for a failed pavement to 100 for a pavement in excellent condition. PCI is evaluated by considering the degree of pavement deterioration, distress severity, and amount of density of distress. Assigning one index that considers the three factors is difficult since they are interconnected, so 'deduct values' are introduced. 'Deduct values' are a type of weighting factor to account for efforts caused by the combination of the three factors. The 'deduct values' are used to derive the PCI (Shahin, 1994), (Sinha and Knight, 2004).

2.3 Condition Assessment Models for Bridges

The Condition assessment of bridge infrastructure has always been an integral part of the '*Bridge Management System*' (BMS). BMS has the purpose of predicting: 1) future bridge conditions, 2) maintenance and improvement needs, and 3) the state of bridge deterioration. The bridge condition assessment aims at evaluating the degree of damage in the bridge components, and their effect in the overall performance of the bridge (DeStephano, 1998). In the U.S.A., the Federal HighWay Association FHWA in its 'recording and coding guide' formulates a scale called the Sufficiency Rating (SR) (from 0 to 100 percent) for the appraisal of selected structural and functional elements of each bridge. Bridges with SR less or equal to 50% are considered for replacement or rehabilitation, while bridges with SR between 50% and less or equal to 80% are eligible for rehabilitation only (Infrastructure Condition Assessment, 1997). In this bridge condition assessment system, the bridge is classified according to the design type: steel girder, steel truss, monolithic concrete and separate concrete. Each type is divided into different components: superstructure, substructure, deck and wearing course. Afterwards, each component in turn comprises different elements: a) wearing surface, b) deck elements; c) superstructure (primary and secondary elements), d) substructure (abutments, piers, columns, pier cap beam, etc.). Finally, for each these elements, a condition rating is assigned as indicated in Table 2.2.

Table 2.2 Bridges Condition Rating

Condition Rating	Description
1 or 2	Potentially hazardous
3	Serious deterioration
4	-
5	Minor deterioration
6	-
7	Excellent

Several models are used to combine the hierarchy of factors described above with the Condition Rating, in order to produce the SR (DeStephano, 1998).

2.4 Condition Assessment Models for Underground Pipes

The underground pipeline infrastructure is most susceptible to decay. This is due to insufficient quality control, little inspection and maintenance and a general lack of uniformity and improvement in design, construction, and operation practices (Yan and Vairavamoorthy, 2004). The main constraint in pipeline infrastructure is that reconstruction of a piping system is not financially realistic. And for this reason, research concentrated in developing a system of monitoring the condition of underground pipes. Therefore, reliable cost-effective pipeline assessment methods are a must in order to develop long-term cost effective maintenance and repair programs (Infrastructure Condition Assessment, 1997). After the inspection, an assessment is done and then a condition rate is assigned to each pipe. The goal of a condition rating system is to objectively rate, by means of a scoring system the current condition of pipelines. It may seem easy, but it is not the case, since a standard procedure has not been developed (as in the case of pavement). An attempt has been made at the Center for Advanced Trenchless Technology, CCATT, to develop a rating system for the municipal sewer system. The CCATT rated the sewer condition based on general defect criteria including crack pattern (transverse or longitudinal, minor or major), joint conditions (minor, major, or multiple), lateral conditions and structural defects (sagging, collapsing, or crushed).

Yan and Vairavamoorty (2004) developed a procedure to screen the conditions of pipelines using Fuzzy Set theory (Zadeh, 1965). The criteria used in their model were: 1) piping age; 2) pipe diameter; 3) pipe material; 4) road loading; 5) soil condition; and 6) surroundings.

Al-Barqawi and Zayed (2006) developed another condition assessment model, using an Artificial Neural Network (NeuroShell software), applied to water mains in Canada. The authors used 8 factors similar to the above: 1) type of soil; 2) type of road surface, 3) pipe cover, 4) pipe diameter, 5) pipe material, 6) pipe age, 7) number of breaks, and 8) the C-factor. His work resulted in a pipe condition scale table.

2.5 Multi-Criteria Decision Analysis Tools

The *Multi-Criteria Decision Making* (MCDM) process consists of the following (Fulop, 2004):

1. Identifying the problem.
2. Establishing the goal.
3. Identifying the function-based factors or criteria.
4. Establishing the rules or choosing the most appropriate mathematical *Multi-Criteria Decision Analysis* (MCDA) tool that best aggregates the criteria with the goal.
5. Validating the solution adopted.

This *Multi Criteria Decision Making* (MCDM) process is not always simple. The complex process of decision aiding goes beyond pure mathematics. It aims at comprehensive support in order to reveal the subjective preferences between distinct criteria scores. This underlying subjectivity is naturally beyond a strictly logical or mathematical analysis (Geldermann and Rentz, 2000). Figure 2.1

presents the different trends of Multi Criteria evaluation methods depending on the criteria aggregation procedure (Petrie et al. 2006).

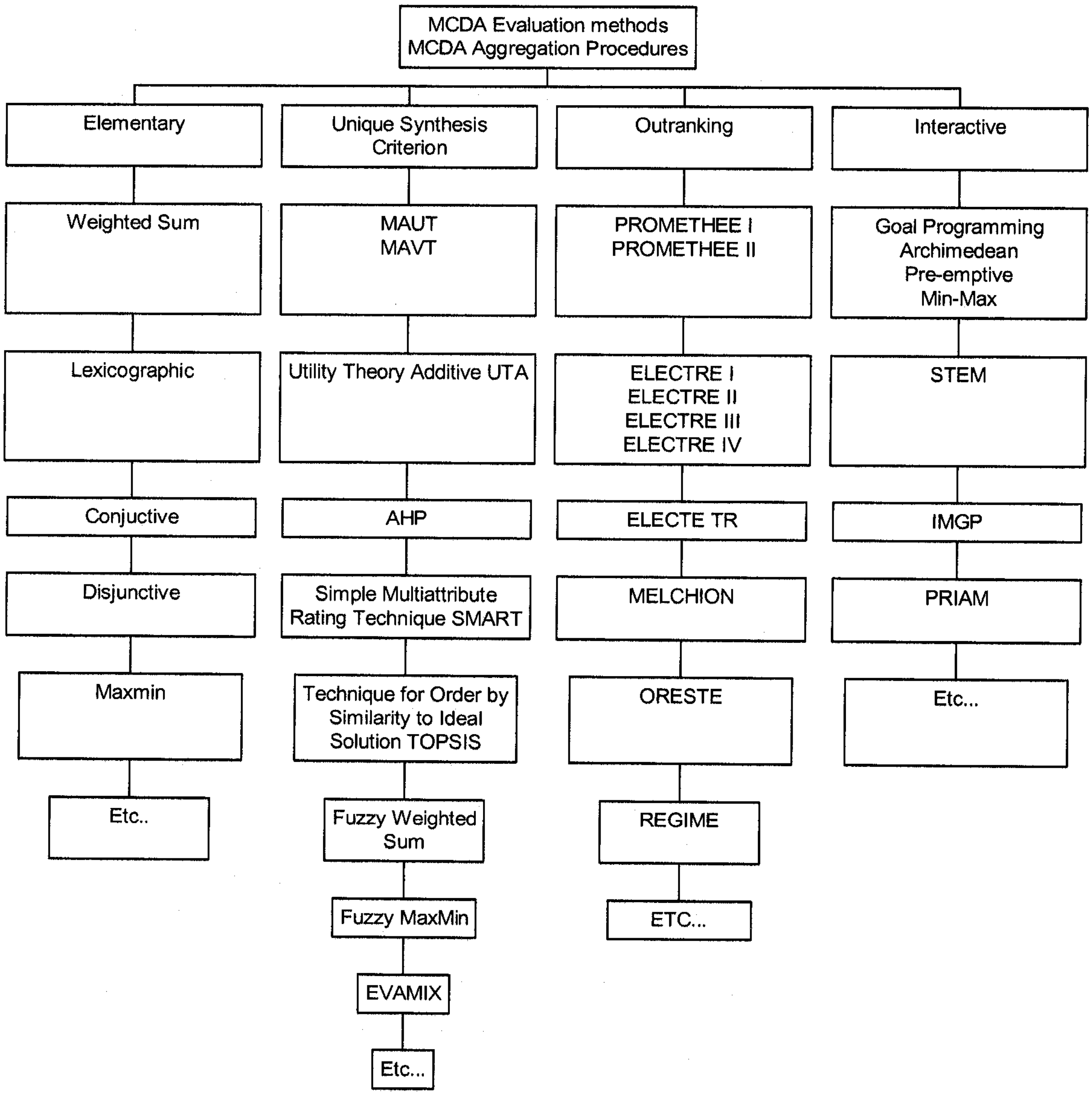


Figure 2.1 Schools of MCDA

Recent research has been done on MCDA tools, including Multi Attribute Utility Theory, Outranking methods, Fuzzy set theory outranking methods, Neural Network methods and statistical methods. This thesis concentrates on the Multi Attribute Utility Theory (MAUT), the Analytical Hierarchy Process (AHP), and

PROMETHEE II (Preference Ranking Organization METHod of Enrichment Evaluation).

According to researchers (Geldermann and Rentz 2000; Kangas et *al.* 2001; Vincke 1992; Doumpos and Zopounidis 2002; Belton and Stewart 2002), two philosophies are being distinguished for MCDA, the North American School, and the European school. The North American school assumes that the decision maker has an exact conception about the utility of the scores and the weights of the different criteria, which are to be discovered and to be interpreted by the means of decision support. Well-known approaches are the Multi Attribute Utility Theory MAUT, the Multi attribute Value Theory MAVT, first developed by Keeny and Raifa (1976), and the Analytical Hierarchy Process (AHP) first developed by Saaty (1980). On the other hand, the European school supposes that the preferences are not apparent to the decision maker. Therefore, decision support is necessary for structuring the decision situation and for giving insight into the consequences of different weightings in the decision problem. The emphasis here is on the recognition of the limits of objectivity. Thus, the researcher should help to build a value judgement model by seeking working hypotheses for making recommendations. The method PROMETHEE is among the most prominent examples for this philosophy.

The formation and the existence of these two philosophical schools is often subject of fierce discussion, mainly because researchers have not been able to agree on a shared view about a philosophical 'correct' method of modelling human value judgement. It has been stated that without a generally accepted paradigm, there are only competing schools and sub-schools, where each researcher feels obliged to build up his work from anew, from the foundation (Tompkins, 2003). In addition, cross-cultural differences in management and decision styles are given as a reason for the formation of the different schools of MCDA.

2.5.1 European School of MCDA: Outranking Methods

As an outcome of the European school, outranking methods serve as one alternative for approaching complex MCDA procedures. Outranking is based on the degree of dominance of one alternative over another. In outranking methods, it is not necessary to assume that a utility function exists, or that it can be described with in a certain functional form. The main question is whether there is enough information to state that one alternative is at least as good as another. Outranking takes into account that preferences are not constant in time, or are unambiguous, and are independent of the process of analysis. Therefore, outranking could be thus defined: alternative a_t outranks a_k, if there is a sufficiently strong argument in favour of the assertion that a_t *is as good as* a_k, from the decision maker's point of view. Accordingly, the outranking relation is the result of pairwise comparison between the alternatives with regard to each criterion. The ability to deal with uncertain and fuzzy information is an indisputable advantage of outranking methods. The criteria are treated as so called *pseudo-criteria* (Brans et al. 1986; Hokkanen and Salminen 1997). The values of the pseudo-criteria may be descriptive and/or quantitative, thus containing uncertainty or fuzziness. The pseudo-criteria can be evaluated either by probability distributions (random variation) or fuzzy zones (uncertainty due to ignorance, etc). In order to reflect this concept of pseudo-criteria, European researchers defined two thresholds for each criterion, the indifference and the preference thresholds. The *Indifference Threshold* for any criterion is a difference beneath which the decision maker is indifferent between two management alternatives. The *Preference Threshold* for any criterion is a difference above which the decision maker strongly prefers one management alternative to another. Between these two thresholds, there is a zone where the decision maker hesitates between 'indifference' and 'strong preference', therefore creating a zone of weak preference (Kangas et *al.* 2001).

Outranking methods have many limitations, especially when dealing with the techniques by which the preference information is calculated, which are complicated and hard to explain to non-specialists. In addition, as with many of MCDA tools, rank reversal is a common problem associated with outranking methods. Nevertheless, flexibility and ease of use, coupled with an understanding of the method, and interpretation of the results, are important qualities of these methods. The outranking method PROMETHEE offers a means of MCDA characterised by simplicity and clarity to the decision maker.

2.5.2 PROMETHEE Method

The *Preference Ranking Organization METHod of Enrichment Evaluation* (PROMETHEE) technique, developed by Brans and Mareschal (1986), belongs to the class of outranking approaches. It is one of the best-known and most widely applied outranking method because it follows a transparent computational procedure and can be easily understood by actors and decision makers. This is evident by its widespread use in decision-making situations in Europe (Hyde et *al.* 2003) such as assessing water resource management problems (Abu-Taleb and Mareschal, 1995; Al-Kloub et al. 1997; Al-Rashdan et al., 1999; Al-Shemmeri et al., 1997; Ozelkan and Duckstein, 1996; Raju et al., 2000; Raju and Pillai, 1999), energy planning (Georgopolou et al. 1998; Haralambopoulos and Polatidis, 2003), and waste management (Hokkanen and Salminen, 1997).

The MCDA process utilizing the PROMETHEE technique has the following sequence:

1. Identifying decision makers (final decision makers), actors (people involved in the decision analysis process), and stakeholders (anyone who might be affected by the decision).
2. Selecting criteria.
3. Formulating management alternatives.
4. Weighting the criteria.
5. Assessing the performance of alternatives against the criteria.
6. Selecting a generalized preference function and associated indifference and preference thresholds for each criterion.
7. Applying PROMETHEE aggregation of criteria.
8. Analysis of results and making the final decision.

The foremost difference between the PROMETHEE method and other outranking MCDA techniques is the utilization of generalized preference functions. Figure 2.2 shows an outline of the PROMETHEE method. The PROMETHEE outranking method allows only the ranking of the alternatives according to their strength (best or worst condition) relative to each other, thus provides a ranking solution not on a final ordinal scale.

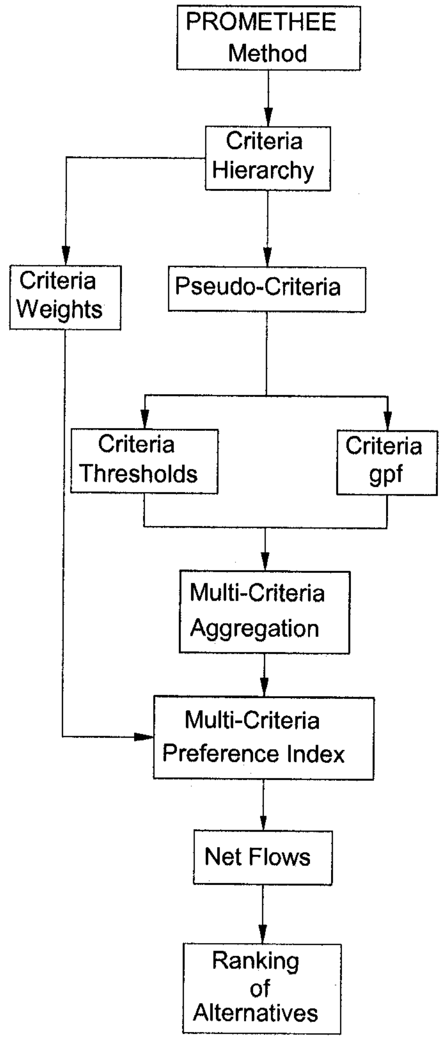


Figure 2.2 PROMETHEE Outline

2.5.2.1 Definition of the decision makers and actors.

The decision makers consist of management group or individual. The actors consist of the engineers, inspectors and the decision analysts, in the case of a condition assessment, for example.

2.5.2.2 Selection of criteria.

The criteria (C_i, *i = No. of criteria = 1…n; n= total No. of criteria*) are all the factors that affect the decision or choice of alternatives. The management and/or the actors select them by analysing the physical aspects of the problem.

A hierarchy of the criteria is the best way to present the different aspects in the decision problem, i.e. to build a hierarchy table of upper level criteria and lower level sub-criteria, as illustrated in Figure 2.3.

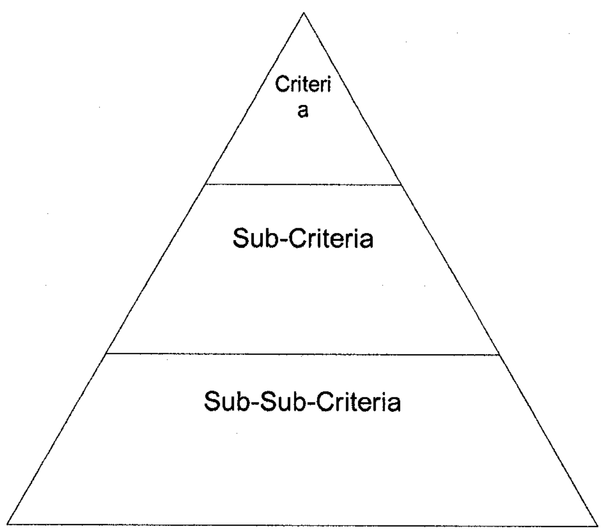


Figure 2.3 Criteria Hierarchy

2.5.2.3 Formulation of the management alternatives.

The alternatives A_j (*j= No. of alternatives = 1…k…m; m=Total no. of alternatives*) are set by the management itself. The alternatives indicate the different solutions to the problem.

2.5.2.4 Evaluation of the weights of the criteria.

The importance weight of each criterion W_i [C_i (A_j)] can be directly assigned either by the decision makers (management) themselves or by some of the actors (engineers, inspectors, or decision analysts). Brans and Mareschal (1986), the researchers behind the PROMETHEE, did not specify any definite method to evaluate the weights. Nevertheless, their only condition was that the sum of the weights, for a given alternative A_j, should always be a unity, as in Equation 2.1:

$$\sum_{i=1}^{n} W_i\,[C_i\,(A_j)] = 1 \qquad (2.1)$$

The literature provides several ways to evaluate weights. Hokkanen and Salminen (1997), who worked intensively with PROMETHEE, suggest assigning a score from 1 to 7 to the criteria. The least important criterion has a score of 1, and the rest are assigned scores relative to the least. The scores are then divided into their sum, in order to normalize the raw weights.

Roberts and Goodwin (2002) have summarized their extensive research on weight elicitation in decision-making, and have found three main streams. The first is the direct rating method, developed by Von Winterfeldt and Edwards (1986). It uses a 'direct numerical ratio judgement of relative criterion importance'. There are number of ways of implementing this method; one of them is the scale 1 to 7 proposed by Hokkanen and Salminen. An alternative approach involves arbitrarily assigning a raw weight of 100 to the criterion where switching from the worst to the best on that criterion is most desirable. The desirability of making similar worst-to-best switches on each of the other criteria is then assessed relative to this, yielding raw weights on a scale with a maximum of 100. Finally, the weights are normalized to sum to either 1 or 100 (Goodwin and Wright 1998).

The second method consists of point allocation. Here, the decision maker has a 'budget' of points to allocate between the criteria in a way that reflects their relative importance. For example, the decision maker may be asked to allocate 100 points between the five criteria that are relevant to a particular decision. Clearly, in this method, there is no need to normalize the weights since the sum of 100 is already prescribed. Doyle et *al.* (1997) declared that the point allocation method is a more difficult task since it is easier to take 100 as the weight for the most important criterion and then allocate weights relative to this 100 starting point as the weight of successive criteria. Doyle et *al.* (1997) pointed out that although this method of determining weights and the direct rating method would seem to be minor variants of each other, in practice they produce different profiles of decision weights, a result that surprised Jia et *al.* (1998).

The third method is the rank ordering of criteria. Baron and Barrett (1996) proved that evaluating weights is not an exact science, but the weights generated by most of the methods previously used are influenced by the method itself, and hence there is no way of directly identifying the 'true' weights. They argued that decision makers are more comfortable in ranking the importance of the criteria rather than assigning weights. This has led to the development of a number of methods that allow for importance ranking to be translated into 'surrogate' weights. These methods include the 'Rank Order Centroid (ROC)', the 'Rank Sum (RS)', and the 'Rank Reciprocal (RR)' weights. In all three methods (ROC, RS, RR) the weight is a function of the reciprocal of both the rank and the total number of criteria (Roberts and Goodwin 2002).

Kangas et *al.* (2001) and Macharis et *al.* (2004) are the sole researchers who indicated that weights incorporated in the PROMETHEE outranking method, could be evaluated using the Analytical Hierarchy Process AHP. Nevertheless, they did not present a real application. Macharis et *al.* (2004) noted that AHP may be used to evaluate the criteria weights for the PROMETHEE method as long as the pairwise comparison matrices are consistent, and the whole decision-making problem is defined in a hierarchical model.

2.5.2.5 Evaluation of the criteria.

The PROMETHEE method gives the freedom to evaluate the criteria (or the performance of the criteria) in any available manner (Fernandez-Castro and Jimenez, 2005). Thus the evaluation can be quantitative for objective criteria, i.e. using expert formulae and/or field test results. It can be also a qualitative evaluation for subjective criteria. So the actor here has the freedom to evaluate the subjective criteria using Fuzzy Set theory, developed by Zadeh (1965). There exists a significant difficulty in the estimation of the required qualitative criteria. The Fuzzy Set theory, as developed by Zadeh (1965), allows the representation of a qualitative description in a quantitative fashion. Zadeh (1965) demonstrated a remarkable method of dealing with vague and approximate situations when he identified classes of objects, which may not be compartmentalized into watertight sections. Zadeh called such sets 'Fuzzy', and defined them as those sets "which have a continuum of grades of membership ranging between zero and one". Goumas and Lygerou (1998) introduced the use of Fuzzy Set Theory in the evaluation of the criteria using the PROMETHEE outranking method. They proved that subjective criteria can be defined as a fuzzy number and then in order to use it in PROMETHEE, it needs only to be de-"fuzzified", using proper methods in Fuzzy Set Theory. The use of the Yager (1981) Index is strongly suggested.

The actor may also use a scale to measure the performance of subjective criteria, and this method is powerful for field inspection. It is important to note that PROMETHEE allows for the use of different means of criteria evaluation in one decision-making model. And this is viewed as an advantage, since it leaves the actors the freedom and flexibility to use available simple input for every criterion separately.

2.5.2.6 Evaluation of the pseudo-criteria.

The PROMETHEE method uses the concept of pseudo-criteria. This concept consists in transforming the true criteria performance into pseudo-criteria. This concept is used in the outranking method, in order to take into account the three following phenomena that affect the criteria performance value:

1. Imprecision: because of the difficulty of determining the value, even in the absence of random fluctuation.

2. Indetermination: because its method of evaluation results from a relatively arbitrary choice between several possible definitions.
3. Uncertainty: because the value involved varies with time.

Various solutions exist for modeling any one of these phenomena, including probability distributions, interval of confidence and fuzzy numbers. The concept of the pseudo-criterion and its two thresholds allow all three phenomena to be taken into account (Roy 1987). Pseudo-Criteria principle enhances the preference or performance when comparing each criterion between two alternatives (A_j and A_k). And this is translated mathematically in PROMETHEE into a preference threshold, an indifference threshold and a general preference function.

PROMETHEE considers the simplest method in comparing the same criterion for two alternatives, by taking their difference defined in Equations 2.2a and 2.2b:

$$\Delta C = C_i(A_j) - C_i(A_k) \tag{2.2a}$$

Or more generally $$\Delta C = \min [C_i (A_j), C_i (A_k)] \tag{2.2b}$$

The preference threshold 'p_i (C_i)' points to the limit where the value of the criteria C_i in alternative A_j is strictly preferred over the value of the same criteria C_i in alternative A_k, defined in Equation 2.3:

$$C_i(A_j) - C_i(A_k) \geq p_i (C_i) \tag{2.3}$$

Means that A_j is strictly or strongly preferred to A_k regarding the criterion C_i only.

The indifference threshold 'q_i (C_i)' points to the limit where the value of the criteria C_i in alternative A_j is strictly indifferent to the value of the same criteria C_i in station A_k, defined in Equation 2.4:

$$C_i(A_j) - C_i(A_k) \leq q_i\ (C_i) \tag{2.4}$$

Means that A_j is strictly indifferent to A_k regarding the criterion C_i only.

Now p_i and q_i are related in a mathematical function expressing the preference model. The purpose of the function is to facilitate the inclusion of the inherent uncertainty in the criteria performance values in the decision analysis process. However, the selection of the function for each criterion is a complex and ambiguous task for decision makers and actors and therefore adds another element of uncertainty into the decision analysis process. Thus the generalized preference function gpf $[C_i\ (A_j)]$ can be defined for any set of criteria either fuzzy or crisp as follows (Goumas and Lygerou, 1998):

A. For a general fuzzy definition where C_i is a fuzzy set of criteria:

- *For y in C_i set;* gpf [x,y] *is a non-deceasing membership function of x*
- *For x in C_i set;* gpf [x,y] *is a non-increasing membership function of x*
- *For z in C_i set;* gpf [z,z] = 1

Thus, gpf $[C_i]$ *is a reflexive, complete, semi-transitive fuzzy semi-order, and every α-cut of* gpf $[C_i]$ *is a crisp semi-order* (Fodor, 2000) and the general formula would be in Equation (2.5) and shown in Figure 2.4:

$$\text{gpf}\ [C_i(A_j), Ci(A_k)] = \frac{p_i(C_i(A_j)) - \min[C_i(A_k) - C_i(A_j), p_i(C_i(A_j))]}{p_i(C_i(A_j)) - \min[C_i(A_k) - C_i(A_j), q_i(C_i(A_j))]} \tag{2.5}$$

Where j and k from 1 to m; m = total No. of stations

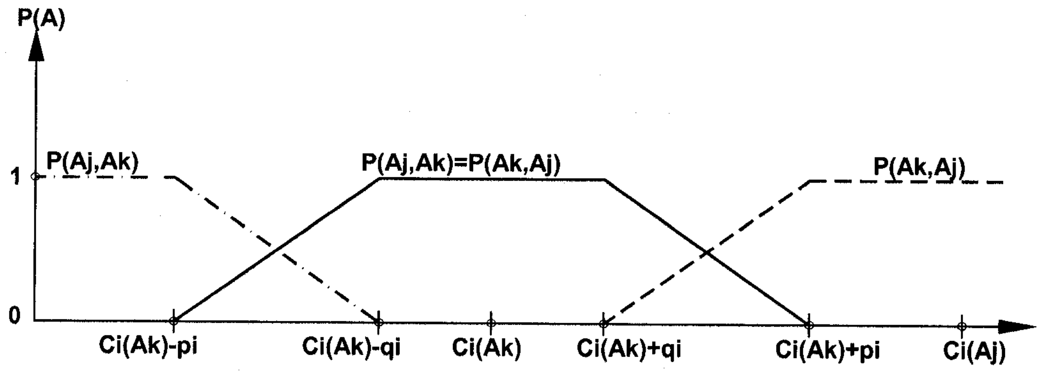


Figure 2.4 Generalized Preference Function

B. For a crisp set of criteria, Brans (1986) has developed 6 types (or 6 forms) of general preference functions (Figure 2.5) that represent most of the preferences types in decision-making, and left for the actor to decide which best apply to a specific criterion, and even the flexibility for the actor to create his own general preference function (Geldermann and Rentz 2000).

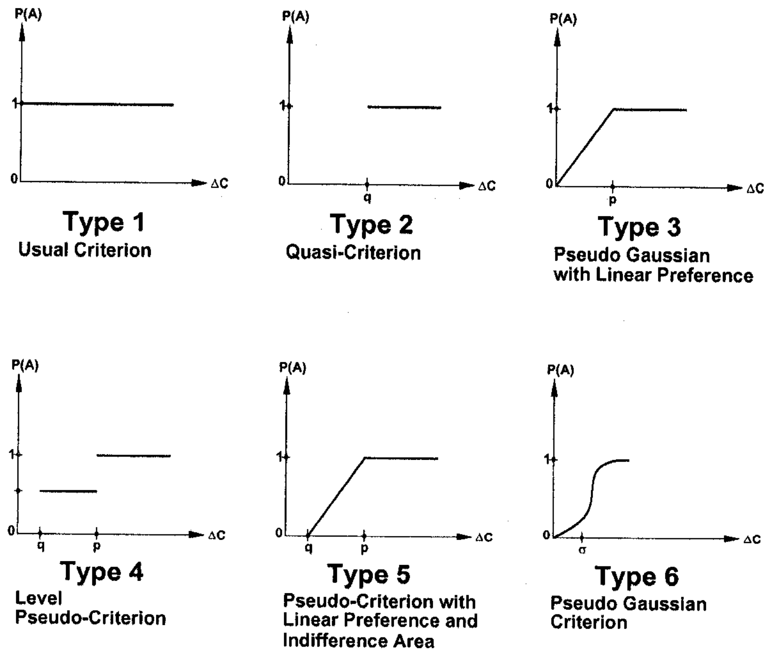


Figure 2.5 Types of gpf

The preferences are defined as follows:

P = Strong Preference; A_k **P** A_j = *A_k strictly preferred to A_j*

Q = Weak Preference; A_k **Q** A_j = *A_k is just slightly better than A_j*

I = Indifference; A_k **I** A_j = *A_k is as good as A_j*

Thus, the general performance function is defined in Equation 2.6:

$$P[C_i(A_j,A_k)] = \begin{cases} 1 & \text{if} \quad C_i(A_k)-C_i(A_j) \geq p_i(C_i) \Rightarrow \quad A_kPA_j \\ 0 & \text{if} \quad C_i(A_k)-C_i(A_j) \leq q_i(C_i) \Rightarrow \quad A_kIA_j \\ \text{gpf}\,[C_i(A_k)-C_i(A_j)] & \text{otherwise} \quad \Rightarrow \quad A_kQA_j \end{cases} \quad (2.6)$$

It should be noted that the type 1 and type 3 gpf could also be modelled by customising the type 5 gpf, setting the parameters p and/or q equal to zero. With p=q≠0, also the type 2 (Quasi-Criterion) could be modelled. σ in type 6 stands for the standard deviation.

2.5.2.7 Aggregation of criteria in PROMETHEE.

PROMETHEE performs a pairwise aggregation of the criteria between the alternatives. Hence, a $P_i[C_i(A_j,A_k)]$ is calculated for every criteria, and then a *multi-criteria preference index* (Π) is evaluated as per Equations (2.7) and (2.8):

$$\Pi_i\,[A_j\,,A_k] = \sum_{i=1}^{n} W_i \,.\, P_i(A_j\,,A_k) \qquad (2.7)$$

Where $0 \leq \Pi_i[A_j\,, A_k] \leq 1$ (2.8)

The multi-criteria preference index considers the weighted preference of alternative A_j to alternative A_k in regards to all the criteria.

From the multi-criteria preference index, a measure of the strength of alternative A_j to all other alternatives is evaluated, indicating how much alternative A_j is preferred to all the others, (i.e. the leaving flow $\Phi^+(A_j)$). Similarly, a measure of the weakness of alternative A_j to all the other alternatives is also calculated, (i.e. the entering flow $\Phi^-(A_j)$). Figure 2.6 describes the entering and leaving flows. And Equations (2.9) and (2.10) define the formulae of the flows.

The leaving flow measures the strength of A_j:

$$\Phi^+(A_j) = \sum_{j=1}^{m} \Pi[A_j, A_k] \tag{2.9}$$

The entering flow measures the weakness of A_j:

$$\Phi^-(A_j) = \sum_{j=1}^{m} \Pi[A_k, A_j] \tag{2.10}$$

Where m = total number of alternatives

Thus, the net flow is defined in Equation (2.11):

$$\Phi^{net}(A_j) = \Phi^+(A_j) - \Phi^-(A_j) \tag{2.11}$$

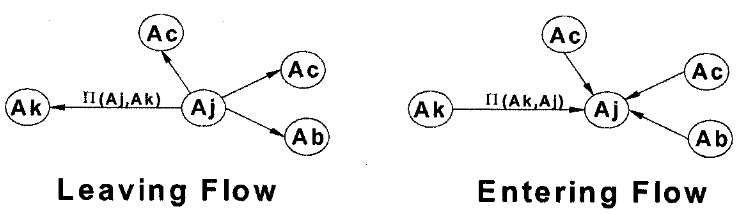


Figure 2.6 Flow Diagram

The calculation of both, the Multi-Criteria Preference Index and the flows, can be tabulated in Table 2.3 in a matrix form for easier representation.

Table 2.3 PROMETHEE Aggregation Table Format

	A_1		**A_j**		**A_k**		**A_m**	**Φ^+**
A_1	0		$\Pi(A_1,A_j)$		$\Pi(A_1,A_k)$		$\Pi(A_1,A_m)$	$\Sigma=\Phi^+(A_1)$
		0						
A_j	$\Pi(A_j,A_1)$		0		$\Pi(A_j,A_k)$		$\Pi(A_j,A_m)$	$\Sigma=\Phi^+(A_j)$
				0				
A_k	$\Pi(A_k,A_1)$		$\Pi(A_k,A_j)$		0		$\Pi(A_k,A_m)$	$\Sigma=\Phi^+(A_k)$
						0		
A_m	$\Pi(A_m,A_1)$		$\Pi(A_m,A_j)$		$\Pi(A_m,A_k)$		0	$\Sigma=\Phi^+(A_m)$
Φ^-	$\Sigma=\Phi^-(A_1)$		$\Sigma=\Phi^-(A_j)$		$\Sigma=\Phi^-(A_k)$		$\Sigma=\Phi^-(A_m)$	

Where A = Alternatives.

$\Pi(A,A)$ = Multi-criteria preference index.

$\Phi^+(A)$ = Leaving flow = Sum of rows of the matrix.

$\Phi^-(A_m)$ = Entering flow = Sum of Columns of the matrix.

2.5.2.8 Ranking of alternatives.

PROMETHEE finally ranks the alternatives according to the flows. Two methods were developed by Brans and Mareschal (1986): PROMETHEE I and PROMETHEE II. PROMETHEE I allows for a partial ranking of alternatives when the ranking is performed with the leaving flow Φ^+ only. While PROMETHEE II ranks the alternatives according to the net flow Φ^{net} as in Equation 2.12, this results in a unique ranking and not a partial one.

$$\Phi^{net}(A_j) \geq \Phi^{net}(A_k) \Rightarrow \quad A_j \text{ Outranks } A_k \text{ in all criteria} \tag{2.12}$$

PROMETHEE I allows two alternatives to have the same rank, whereas PROMETHEE II attributes a sole rank to each alternative as illustrated in Figure 2.7.

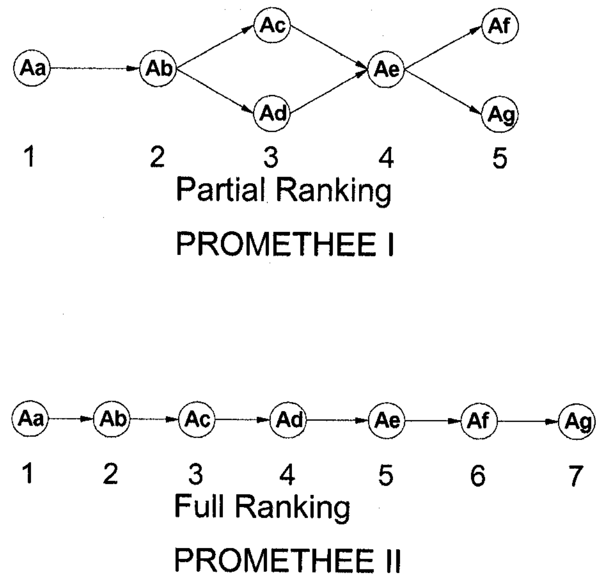


Figure 2.7 PROMETHEE Ranking

2.5.2.9 Remarks on PROMETHEE method.

The following are some general remarks on the PROMETHEE outranking method, underlying some of its advantages and disadvantages.

- The PROMETHEE outranking method is a very flexible, simple and transparent tool to rank alternatives. However, it does not provide an ordinal or fixed-scale final condition of the alternatives. The method result is a rank of the alternatives from best to worst or vice versa (depending on the criteria measurement). Hence, the rank is not an absolute one and is not a fixed measure of the status of the alternative. Thus PROMETHEE is useful in prioritization of the alternatives only.

- One of its advantages is that it indicates at the end of the analysis the strength and weakness of the alternative. The net flow is a term not found in most, if not all of the MCDA methods.
- PROMETHEE also has no limitations in the evaluation of the criteria. Hence any scientific method may be used for the evaluation.
- PROMETHEE considers uncertainty both in the performance and preference of the criteria by transforming them into pseudo-criteria.

2.5.3 The North-American school: Multi Attribute Utility Theory

The *Multi Attribute Utility Theory* (MAUT) is one of the most widely applied multi-criteria methods. Although it most represents the North American school of decision analysis, its origin is Anglo-Saxon (Vincke and al. 1992). From even the early stages of the MCDA field, MAUT has been one of the cornerstones of the development of MCDA and its practical implementation. Directly or indirectly all other MCDA approaches employ the concepts introduced by MAUT (Doumpos and Zopounidis 2002).

2.5.3.1 Principles of MAUT

MAUT is based on developing a utility function representing the decision maker's system of preferences. The theory is founded on the following fundamental axiom: any decision maker attempts unconsciously (or implicitly) to maximize some function 'U' aggregating all the different points of view which are taken into account. In other words, if the decision maker is asked about preferences, his answers will be coherent with a certain unknown function U, which has a general form of Equation (2.13)

$$U = U(C_1, C_2, \ldots, C_m) \tag{2.13}$$

Where m = total No. of alternatives, and C are the criteria involved in the decision-making problem.

The role of the researcher is to try to estimate that function by asking the decision maker some well-known chosen questions. Essentially two types of problems are studied in the frame of this theory:

1. What properties must the decision maker's preferences fulfil in order to be able to represent them by a function U with a given analytical from (additive, multiplicative, mixed, etc.)?
2. How can such functions be built and how can the parameters to be chosen in an analytical form be estimated?

It is also important to insist on that utility theory concerns functions of criteria based on true criteria. At present, very little research has considered generalizations to other types of criteria such as a fuzzy set of criteria or pseudo-criteria. (Roy and Bouyssou 1987).

2.5.3.2 MAUT functions

Generally, the utility function is either a non-linear or a linear function defined on the criteria space, such that:

$U(A_1) > U(A_2) \leftrightarrow A_1 > A_2$ (alternative A_1 is preferred to A_2)

$U(A_1) = U(A_2) \leftrightarrow A_1 = A_2$ (alternative A_1 is indifferent to A_2)

The simplest (and most commonly used) analytical form is the additive form, as shown in equation (2.14):

$$U(A) = \sum_{i=1}^{n} u_i\,[C_i(A)] \tag{2.14}$$

Where n = total no. of criteria

Weights of criteria can also be included in the function, as in Equations (2.15) and (2.16):

$$U(A) = \sum_{i=1}^{n} W_i . u_i [C_i (A)] \tag{2.15}$$

$$U(A) = W_1 . u_1 (C_1) + W_2 . u_2 (C_2) + \ldots + W_n . u_n (C_n) \tag{2.16}$$

The u_i are strictly increasing real functions (their only purpose is to transform the criteria in order for them to follow the same scale: this avoids problems of units and ensures that the summation makes sense). The main assumption underlying the use of the additive utility function involves the mutual preferential independence condition of the evaluation criteria, described in Equation 2.17:

If $C_i (A) = C_i (B)$

$C_i (C) = C_i (D)$

And $C_i (A) = C_i (C)$

$C_i (B) = C_i (D)$

Then $U(A) - U(B) = U(C) - U(D)$ (2.17)

i.e. *A is preferred to B ↔ C is preferred to D.*

Where A, B, C, and D are various alternatives.

The global utility of the alternatives, estimated on the basis of the developed utility function, constitutes an index used for choice, ranking or classification purposes. This index can be represented on an ordinal scale (depending on the global utility), and this is the real power of MAUT.

The weights included in the MAUT function can be evaluated using several tools, either by the tools described in the PROMETHEE method, or by probabilistic modelling, or by simulation (Etezadi-Amoli et al. 1983).

The additive model can be mathematically transformed into a multiplicative one, in Equations 2.18 and 2.19:

$$U'(A) = e^{U(A)} \tag{2.18}$$

Thus $$U'(A) = \prod_{i=1}^{n} u'_i(C_i(A)) \tag{2.19}$$

The multiplicative utility function is efficient when a critical criterion dominates the decision.

2.5.3.3 Trade-off points

Other forms of global utility functions can also be constructed depending on the nature, representation and type of the decision-making problem, alternatives and criteria. The most difficult task in the MAUT is to choose the utility function, and researchers have developed miscellaneous methods to build the function (mainly the additive one). All of these methods are based on defining criteria trade-off points (maximum C_{max}, minimum C_{min}, middle point C_{mid}, etc.).

Three general and widely known methods will be indicated; other methods can be derived from these three.

1. Method 1:

Ask the decision maker to determine the state C_{mid1} to be considered midpoint of C_{max} and C_{min}, then state C_{mid2} midpoint of C_{min} and C_{mid1}, and C_{mid3} midpoint of C_{mid1} and C_{max}, and so on. We get, in Equation 2.20:

$$\begin{cases} U(C_{mid1}) = \frac{1}{2}\{U(C_{min}) + U(C_{max})\} \\ U(C_{mid2}) = \frac{1}{2}\{U(C_{min}) + U(C_{mid1})\} \\ U(C_{mid3}) = \frac{1}{2}\{U(C_{mid}) + U(C_{max})\} \\ \ldots \end{cases} \tag{2.20}$$

2. Method 2:

The decision maker is asked to determine state C_i such that he considers it equivalent to:

- Obtain C_i;
- Obtain C_{min} with probability ½ and C_{max} with probability ½.

Therefore, in Equation 2.21:

$$U(C_i) = \frac{1}{2}\{U(C_{min}) + U(C_{max})\} \tag{2.21}$$

Thus, this continues with (C_{min}, C_i), with (C_i, C_{max}).

3. Method 3:

The decision maker is asked to determine the state C_i *(p)* such that he considers it equivalent to:

- Obtain C_i *(p)*;
- Obtain C_{min} with probability '*p*' and C_{max} with probability '*1-p*'.

Therefore, in Equation 2.22:

$$U(C_i) = p \,.\, U(C_{min}) + (1-p) \,.\, U(C_{max}) \tag{2.22}$$

Figure 2.8 shows general forms of MAUT functions

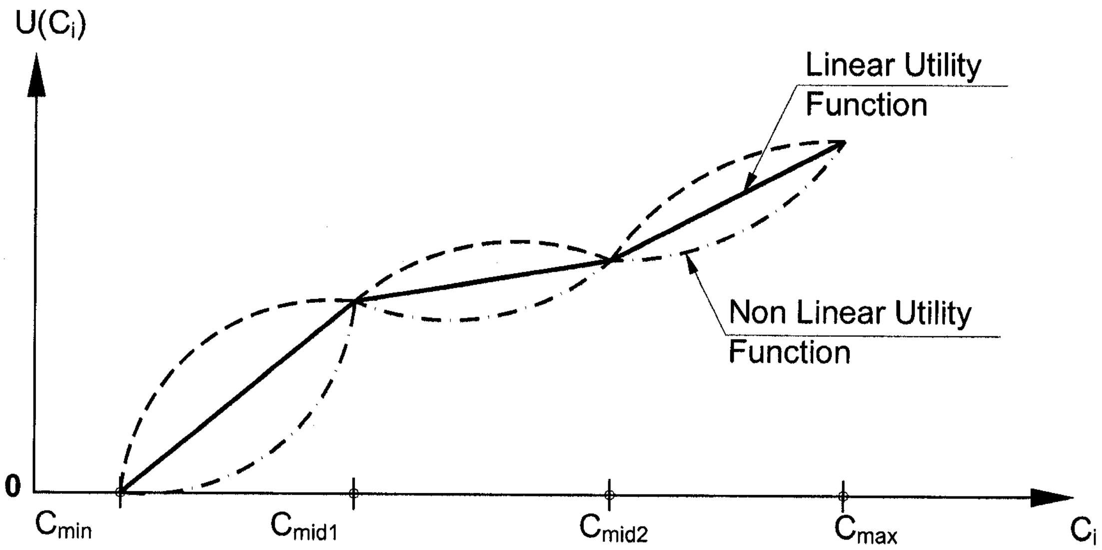


Figure 2.8 General Form of MAUT Functions

2.5.3.4 Multi-Attribute Value Theory MAVT

A simpler form of MAUT is the *Multi-Attribute Value Theory* (MAVT). The MAVT utilizes the same principles of the utility theory, save that the criteria are evaluated on an ordinal scale (criteria performance), and hence the value function is evaluated also on the same scale. Value function methods synthesize assessments of the performance of alternatives against individual criteria, together with inter-criteria information reflecting the relative importance of the different criteria, to give an overall evaluation of each alternative indicative of the decision maker preferences.

The general form of MAVT is defined in Equation (2.23):

$$V(A) = \sum_{i=1}^{n} W_i . v_i \left[C_i (A)\right] \qquad (2.23)$$

2.5.3.5 Remarks on MAUT and MAVT

The following are some remarks on the MAUT, indicating some advantages and disadvantages:

- Several theoretical and empirical studies have shown that the additive model provides a reasonable approximation to the 'true' aggregate utility function even if additive utility independence does not hold (Lam et al. 1997).
- The MAUT and MAVT are widely used and powerful, since it gives a measurable decision. MAUT is normative in nature in that it tells the user what should be done, based upon measurements of utility for different criteria, alone or in combination.
- The theory is not descriptive, in the sense that it does not provide a good prediction or approximation of actual behaviour, but it has the advantage that it does allow experimental testing of the theory itself (Vignaux 2005).
- Imprecision in MAUT parameters can be attributed to assessment errors as well as to vague and ambiguous preferences (Lam et *al.* 1997).
- MAUT can consider uncertainty in the model if a probability can be introduced in the function.
- MAUT is based on using a common scale of local to global utility, and this could be used in some decision-making problems a constraint.

2.5.4 The Analytical Hierarchy Process

The *Analytical Hierarchy Process* (AHP) is one of the methods based on the MAUT. The AHP developed at the Wharton School of business by Thomas Saaty (1980) allows decision makers to model a complex problem in a hierarchical structure. The hierarchy shows the relationships of the goal, criteria, sub-criteria, and alternatives as illustrated in Figure 2.9.

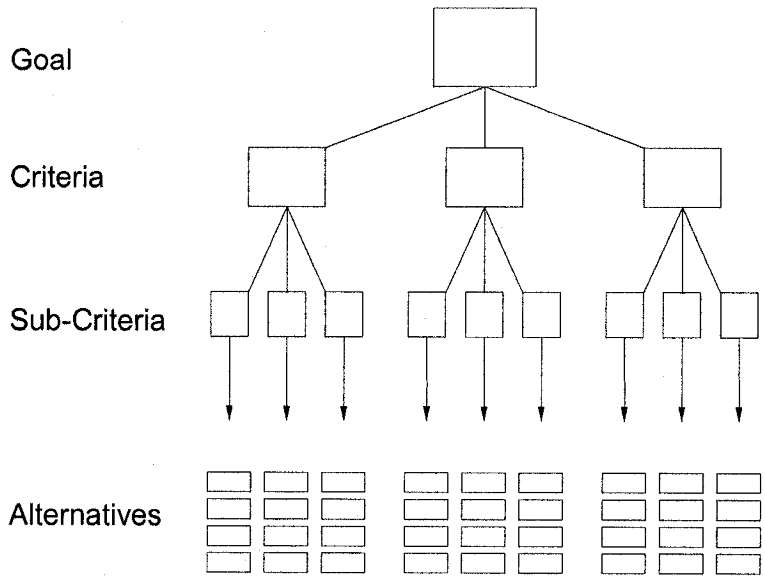


Figure 2.9 Decision Hierarchy

AHP allows for the application of data, experience, insight and intuition in a logical and thorough way. It enables decision-makers to derive rational scale priorities or weights as opposed to arbitrarily assigning them. In so doing, AHP not only supports decision-makers by enabling them to structure complexity and exercise judgment, but allows them to incorporate both objective and subjective considerations in the decision process. It is a compensatory decision methodology because alternatives that are deficient with respect to one or more criteria can compensate based upon their performance with respect to other criteria. AHP is composed of several previously existing but un-associated concepts and techniques such as hierarchical structuring of complexity, pairwise comparisons, redundant judgements, an eigenvector method for deriving weights, and consistency considerations. Although each of these concepts and techniques were useful in and of themselves, Saaty's synergistic combination of the concepts and techniques (along with some new developments) produced a process whose power is indeed far more than the sum of its parts.

2.5.4.1 The principles and axioms of AHP

AHP is based on three principles: decomposition, comparative judgements, and hierarchic composition of priorities (Saaty 1994). The decomposition principle is applied to structure a complex problem into a hierarchy of clusters (criteria), sub-clusters (sub-criteria), sub-sub clusters (sub-sub-criteria) and so on. The principle of comparative judgements is applied to construct pairwise comparisons of all combinations of elements in a cluster with respect to the parent cluster. These pairwise comparisons are used to derive 'local' priorities of the elements in a cluster with respect to their parent. The principle of hierarchic composition, or synthesis, is applied to multiply the local priorities of elements in a cluster by the 'global' priority of the parent element, producing global priorities throughout the hierarchy and then adding the global priorities for the lowest level elements (the alternatives).

All theories are based on axioms. The simpler and fewer the axioms, the more general and applicable is the theory. Originally AHP was based on three relatively simple axioms. The first axiom, the reciprocal axiom, requires that, if a_{ij} is a paired comparison of elements *i* and *j*, representing how many times more the element *i* possesses a property than does element *j*, then $a_{ji} = 1/a_{ij}$. For example, if *i* is 5 times larger than *j*, then *j* is 1/5 as large as *i*. The second, or homogeneity axiom, states that the elements being compared should not differ too much, as this will tend to larger errors in judgement. When constructing a hierarchy of criteria, one should attempt to arrange elements in a cluster so that they do not differ by more than an order of magnitude. The AHP verbal scale ranges from 1 to 9, or about an order of magnitude. Judgements beyond an order of magnitude generally result in a decrease in accuracy and increase of inconsistency. The third axiom states that judgement about, or the priorities of, the elements in a hierarchy do not depend on lower level elements. This axiom is required in order for the principle of hierarchic composition to apply. While the first two axioms are always consistent with real world applications, this axiom requires careful examination, as it is not uncommon for it to be violated. Thus, while the preference for alternatives is almost always dependent on higher-level elements (the criteria), the importance of the criteria might or might not be dependent on lower level elements (the alternatives). There are two basic ways to proceed in those situations where this axiom does not apply, that is, where there is feedback. The first involves a formal application of feedback and a super matrix calculation for synthesis (or aggregation) rather than hierarchic composition. This approach is called the Analytical Network Process (ANP). For simple feedback between adjacent levels only, this is equivalent to deriving priorities for the alternatives with respect to each criterion. The resulting priorities are processed in a super matrix, which is equivalent to the convergence of iterative hierarchical compositions. While this approach is extremely powerful and flexible, a simpler approach that usually works well, is to make judgements for lower levels of the hierarchy first, then to higher levels (or to reconsider judgements at the upper levels after making judgements at the lower level). In so doing, the brain performs the feedback function by considering what was learned at lower levels of the hierarchy when making judgements for upper levels. Thus, an important rule of thumb is to make judgements in a hierarchy, from the bottom up, unless one is sure that there is no feedback, or one already has a good understanding of the alternatives and their tradeoffs. Even if this is done, adherence to the fourth axiom of the AHP (below) as well as the process notion of AHP, can usually lead to appropriate judgements, since an examination of the priorities after a first iteration of the model will highlight those areas where judgements should be revised based on what has been learned. A fourth axiom, introduced by Saaty, says that individuals who have reasons for their beliefs should make sure that their ideas are adequately represented for the outcome to match these expectations. While this axiom might sound a bit vague, it is very important because the generality of the AHP makes it possible to apply it in a variety of ways and adherence to this axiom prevents applying the AHP in inappropriate ways. These axioms will now be described in a mathematical form.

2.5.4.3 Mathematics of the AHP

The first step in the AHP is to arrange the decision-making problem in a hierarchical fashion. The next step is to establish priorities, i.e. to perform pair wise comparisons. Pair wise comparisons of the sub-criteria and criteria (according to the hierarchy) are made in terms of either:

- Importance: when comparing criteria with respect to their relative importance.
- Preference: when comparing the preference of criteria for alternatives with respect to an objective.
- Likelihood: when comparing uncertain events or scenarios with respect to the probability of their occurrence.

Pair wise comparisons are fundamental to the AHP methodology. When comparing a pair of criteria, a ratio of relative importance, preference or likelihood of the criteria can be established. This ratio need not be based on some standard scale but merely represents the relationship of the two criteria being compared. While researchers developed several graphical and numerical scales (Leskinen, 2000), Table 2.4 shows the original comparison scale developed by Saaty (1980).

Table 2.4 Saaty's Scale of Measurement

Value	Definition
1	Equally important or preferred
3	Slightly more important or preferred
5	Strongly more important or preferred
7	Very strongly more important or preferred
9	Extremely more important or preferred
2, 4, 6, 8	Intermediate values to reflect compromise

The comparison matrix is as shown in Figure 2.10 below (4x4 sample matrix).

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} \quad 1 \quad & \quad 2 \quad & \quad 3 \quad & \quad 4 \quad \end{array} \\ \left\{ \begin{array}{|c|c|c|c|} \hline 1 & a_{12} & a_{13} & a_{14} \\ \hline 1/a_{12} & 1 & a_{23} & a_{24} \\ \hline 1/a_{13} & 1/a_{23} & 1 & a_{34} \\ \hline 1/a_{14} & 1/a_{24} & 1/a_{34} & 1 \\ \hline \end{array} \right\} \end{array}$$

Figure 2.10 AHP Sample Matrix

Where A is the comparison matrix, and a_{ij} is the pairwise comparison scale.

Assuming that the comparison matrix is an nxn matrix A [nxn], hence the weight (W) is defined in Equation (2.24):

$$A.W = n.W \tag{2.24}$$

In evaluating W, we calculate the weight of each element.

Furthermore, AHP allows for inconsistency, but provides a measure of the inconsistency in each set of judgements. This measure is an important by-product of the process of deriving priorities based on pairwise comparisons. It is natural for people to want to be consistent. Being consistent is often thought of as a prerequisite to clear thinking, however the real world is hardly ever perfectly consistent and we can learn new things only by allowing for some inconsistency with what we already know. Research has indicated several causes of inconsistency. The most common cause of inconsistency is a clerical error. When entering one or more judgements into a computer, the wrong value or perhaps the inverse of what was indeed is entered. Another important source of inconsistency is the lack of information. If one has little or no information about the factors being compared, then judgements will appear to be random and a high inconsistency ratio will result. Another cause is the lack of concentration during the judgement process. Still another cause is the actual lack of consistency in whatever is being compared especially since real life problems are rarely consistent. A final cause of inconsistency is inadequate model structure. Thomas Saaty, in 1982, has derived formulae to measure the degree of inconsistency in order to control it. For a consistency matrix, $\lambda max = n$, where λmax is the largest eigen-value of the reciprocal matrix of order n. The Consistency Index can be calculated in Equation 2.25:

$$CI = (\lambda max - n)/n-1 \qquad (2.25)$$

And the Consistency Ratio, in Equation 2.26:

$$CR = CI / RCI \leq 10\% \qquad (2.26)$$

The RCI (Random Consistency Index) refers to the average consistency for different order random matrices (Saaty 1982).

A Consistency Ratio of $CR \leq 0.1$ or 10% is an acceptable evaluation of the consistency of the judgement.

2.5.4.3 Remarks on AHP

The following are some general remarks on AHP:

- Due to the presence of inconsistency, and the use of a non-ordinal numerical comparison scale, several researchers in the decision analysis field have fought AHP.
- The main disadvantage of AHP is the rank or weight reversal.

- The AHP also does not produce a final ordinal scale of weights or priorities but a relative comparative one.
- The AHP is a suitable process to estimate relative weights of the decision elements and culminates into their aggregation in order to arrive at the outcome.
- Triantaphyllou (1994) has pointed out that as the number of alternatives increases, the more the weight of the alternative becomes sensitive to the numerical rating (9-point scale) assignment in the matrix. Therefore, the solution suggested was to minimize the criteria and alternatives in one matrix.

2.6 Sensitivity Analysis

Some values of the multi-criteria decision models are often subjective. The weights of the criteria and the scoring values of the alternatives against the subjective (judgmental) criteria always contain some uncertainties. It is therefore an important question as to how sensitive the ranking values of the alternatives are to the changes of some input parameters of the decision model.

Regarding the sensitivity of the weights of the criteria used in a wide class of multi-criteria decision models, Mareschal (1988) studied how to determine the stability intervals or regions for the weights of different criteria. These consist of the values that the weights of one or more criteria can take without altering the results given by the initial set of weights, all other weights being kept constant. Wolters and Mareschal (1995) proposed a linear programming model to find the minimum modification of the weights that is required to make a certain alternative rank first.

Triantaphyllou and Sanchez (1997) presented an approach of a more complex sensitivity analysis, complete with changes to the scores of the alternatives against the criteria. Mészáros and Rapcsák (1996) presented a general and comprehensive methodology for a wide class of MAUT models. In their models, the aggregation was based on generalized methods, including the additive and multiplicative model. In this approach the weights and the scores of the alternatives against the criteria can change simultaneously in given intervals. The models address the following questions:

- What are the intervals of the final ranking values of the alternatives with the restriction that the intervals of the weights and scores are given?
- What are the intervals of the weights and scores with the restriction such that the final ranking of the alternatives does not change?
- Consider a subset of alternatives whose ranking values are allowed to change in an interval. In what intervals are the weights and scores allowed to vary, and how will these modifications affect the ranking values of the entire set of alternatives?

They pointed out that these questions lead to the optimization of linear fractional functions over rectangles and proposed an efficient technique to solve these problems. Some of the results of Mészáros and Rapcsák (1996) were recently extended by Ekárt and Németh (2005) for more general decision functions.

2.7 Summary

This literature review has shown that research has emphasized two schools of MCDA, the North American and European. The North-American school is based on the complete aggregation of the criteria, i.e. allotting a function of utility to each attribute. For each alternative, a mathematical function incorporates the various specific partial utilities to each criterion. One thus obtains a synthetic answer, which is unique. These methods authorize the compensation of the judgements, which are transitive, between the various criteria. A defect in these methods comes owing to the fact that the determination of the function of utility is sometimes very complex. The European school of MCDA is based upon partial aggregation of the different criteria, which consists of comparing the alternatives two by two, criteria by criteria. This makes it possible to establish the relationship of outclassing that exists between them. A synthesis of these relationships between the various alternatives is the made in order to carry out a sorting, an arrangement or to determine the best alternative in the batch. These methods admit the postulates of the incomparability and of the intransitivity. They authorize a greater richness in the relationship between the alternatives. The results are sometimes not very clear, which can be perturbing for the decision maker who wants to receive a clear and final answer (Tille and Dumont 2003).

Research related to combining both schools of MCDA together has been non-existent. Combining both schools would make use of their counterpart's advantages in order to strengthen their weaknesses.

3 Subway Stations Diagnosis Index (SSDI) Model

3.1 Introduction

Prioritization of Maintenance and Repair (MR) expenditures represents a major task in the management of an aging infrastructure. As a subway station approaches its design life, there is an increasing demand for maintenance and repair projects to extend its design life, minimize the potential for loss of function and reduce or eliminate danger to public transit users. It is necessary to develop rational prioritization schemes for these expenditures because not all MR projects can be funded in a given fiscal year. This involves the development of condition assessment methodologies, consisting of sets of rules that convert visual inspection, and/or engineering judgement (instrumentation readings, operational information and engineering computations) into numerical values known as condition indices.

A new model of condition assessment designed for subway stations is developed, called 'Subway Station Diagnosis Index SSDI'. The SSDI model assigns a condition index or a *Diagnosis Index* (DI) to the subway station, considering the functional-operational criteria only. The new developed model has to be transparent in usage in order to manage the MR planning of the subway stations. The SSDI model has to use the most recent data; and it has to take into account the fact that data are often incomplete or inaccurate.

3.2 The SSDI Model Outline

The SSDI model is based on the *Multi-Criteria Decision Analysis* (MCDA) theories described in the literature review chapter. Its main mathematical decision analysis tool is the PROMETHEE method, which belongs to the European school of MCDA. As a matter of fact, the PROMETHEE tool, since its development, was used in North America in the economics field (Veilleux, 1999), forest management (Cox, 2003) and (Abu-Taleb, 1993), and water management (Alam, 1992).

PROMETHEE is chosen since it uses pseudo-criterion instead of a usual criterion (refer to literature review). It considers the uncertainties, inaccuracies, and incompleteness of information in the overall evaluation of criteria. Furthermore, the evaluation of the criteria performance in PROMETHEE is flexible; i.e. each criterion can be evaluated differently than the other. Moreover, the PROMETHEE aggregation of criteria is transparent enough, so that the decision-maker can easily track the corner stones of his decision and thus of the index itself. Since PROMETHEE, as all the outranking methods, results in a rank on a cardinal scale, and not an index measured in an ordinal scale, two concepts were introduced: 1) fictitious conditions fixed lower and upper datum and 2) MAUT to transform the rank into an index by a certain function.

Figure 3.1 illustrates the general outline of the SSDI model. Based on Figure 3.1, the SSDI model considers one station only S_k for the assessment of its condition. The model then structures the criteria into a hierarchy, and utilizes AHP in order to evaluate the sub-criteria weights and the functional weights.

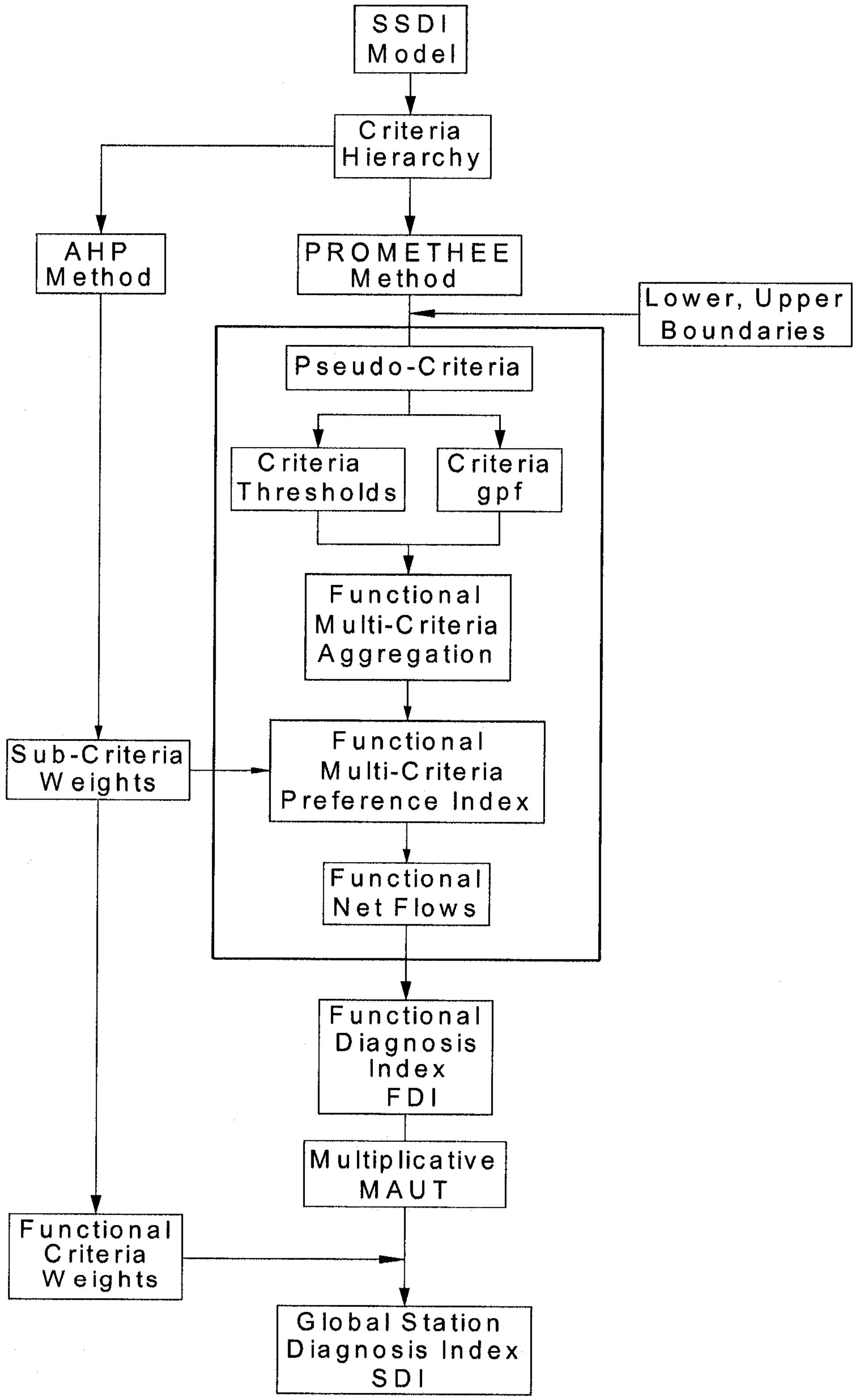


Figure 3.1 The SSDI model Outline

The SSDI model introduces two fictitious stations, which are two boundaries, or datum. These two boundaries serve in defining the ultimum worst and best station condition, and hence provide the limits of the index ordinal scale. The usual criteria performances are then transformed into pseudo-criteria by using the criteria thresholds and the general preference function. The functional multi-criteria aggregation is then performed and the functional multi-criteria preference index is evaluated. From the latter evaluation, the functional net flows are calculated. The functional Diagnosis Index (FDI) is then evaluated (for each function separately) from the net flows. Finally, using the MAUT (multiplicative form), the global Station Diagnosis Index (SDI) is calculated.

3.3 Criteria Definition

A criterion is a qualitative or a quantitative expression that permits the decision maker to judge the consequence or the performance of an alternative for an object or a constraint of the considered project (Tille and Dumont 2003). A criterion must be useful and reliable. It is very important to accurately define the criteria or the factors that come into play in the decision. It is also preferable for most MCDA problems to define them in a hierarchical structure. The choice of the criteria must be coherent, which is checked against the three conditions:

1. Exhaustiveness
2. Coherence, which introduces the local preferences of each criterion and the total preference.
3. Independence, i.e. there should not be redundancy between the criteria (Tille and Dumont 2003).

Based on the general problems found in subway stations, several types of condition assessment goals can be identified:

1. Level of Service LOS: the goal of this type of condition assessment is to increase the LOS. The LOS objective is to improve the comfort of travelling, and to prevent passengers being lost to private transportation, and to attract new categories of clients to public transportation.
2. Functional/Operational: The goal of this type of condition assessment is to maintain and rehabilitate or renovate the functions of the subway stations, mainly due to the old age of the stations itself. The functions concerned are the structural, architectural, mechanical, environmental, communication, security system, control system, electrical factors or components that affect directly the operation of the subway.
3. Sustainable Development: The goal of a condition assessment could be to increase the sustainable development of the subway system. The factors would include social, environmental and economic aspects of the station.
4. Against Terror: Increase in security, especially against terror, a condition assessment goal could be done in order to evaluate the security aspect of the station, and enhance plans to increase it.

The research considers the functional/operational aspect of the subway station, and thus studies the many factors – directly linked to the MR of the station - or the criteria that affect the condition assessment. The Maintenance and Repair (MR) planning is the major goal of any functional condition assessment, and this is the most important task for the management of any transit authority, especially when its facilities are beyond their design life. Aging infrastructure maximizes the potential for loss of function and increases danger to public transit users. In order to perform such a functional assessment, i.e. rate in terms of its ability to perform an intended function, the authority management should first define the functional criteria that directly affect the condition of the facility, and then provide the guidelines and methods for the functional inspection. This current research, based on both the previous literature and direct coordination with STM management, has defined 5 major functional criteria that affect the condition assessment of subway stations, as illustrated in Figure 3.2:

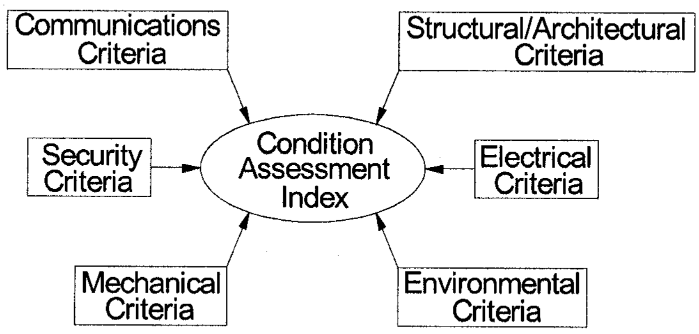


Figure 3.2 The SSDI Model Criteria

The five independent categories of functional criteria directly affect the operation of the station and hence the MR plans. The research has developed a hierarchy of criteria and sub-criteria for each functional category. The sub-criteria should be meaningful, independent and easy to evaluate by use of visual inspection or proper engineering methods. In order to measure the performance of each criterion for every station, the sub-criteria for each criteria group were defined for the new subway station model SSDI as follows:

1. Structural / Architectural criteria:

   a. Global Structure sub-criterion: indicates the global performance of the structural function of the station. It includes the evaluation of the different structural elements, such as slabs, quay platforms, columns, walls, beams, frames, etc., against cracks, leakage, corrosion, deflection and other defects. This sub-criterion can be even divided or hierarchized into structural elements or sub-sub-criteria. Then an average or minimum value of the performance of each one will be chosen as the global structural sub-criterion performance.

   b. Global Architecture sub-criterion: indicates the global performance of the architectural function of the station. It includes the evaluation of the different architectural elements, such as paint, joints, windows, doors, ceiling, walls, etc. This sub-criterion can also be hierarchized in a similar fashion as in the global structure sub-criterion.

   c. Concrete (Fixed) Stairs sub-criterion: indicates the condition or performance of the concrete (fixed) stairs, such as the condition of tiling, cracks in stairs and slab, joints, handrails, etc. It considers the combined architectural and structural condition performance of the stairs. This sub-criterion can also be hierarchized into sub-sub criteria.

2. Mechanical criteria:

   a. Mechanical Stairs sub-criterion: indicates the performance of the mechanical stairs. This sub-criterion includes the motor, the steel frame, the steel steps, etc. This sub-criterion can also be hierarchized into sub-sub criteria.

b. Pipes and Mechanical Equipments sub-criterion: indicates the performance of the pipes and the mechanical equipments in the subway station, such as the pumps, manholes, drainage pipes, sewer pipes, water heaters, etc. This sub-criterion can also be hierarchized into sub-sub criteria.

c. Ventilation, A/C, Heat System sub-criterion: indicates the performance of the air system, such as the ventilators, chillers, ducts, pumps, manholes, etc. This sub-criterion can also be hierarchized into sub-sub criteria.

d. Fire Stand Pipes sub-criterion: indicates the performance of the fire standpipes network, including the pipes, the alarm, the pumps, etc. This sub-criterion can also be hierarchized into sub-sub criteria.

3. Electrical criteria:

   a. Lighting sub-criterion: indicates the performance of the lighting system, such as the lamps, emergency lights, lighted signs, etc. This sub-criterion can also be hierarchized into sub-sub criteria.

   b. Cables sub-criterion: indicates the performance of the electrical cable network in the subway station. This sub-criterion can also be hierarchized into sub-sub criteria.

   c. Panels, Transformers and Breakers sub-criterion: indicates the performance of the electrical panels, transformers, and breakers, i.e. the electrical equipments in the electrical room. This sub-criterion can also be hierarchized into sub-sub criteria.

4. Communication / Control and Security system criteria:

a. Security System sub-criterion (Alarm, Smoke detectors sub-criterion): indicates the performance of the security system connected to a control room transmitting the security information, mainly the alarm and smoke detectors. This sub-criterion can also be hierarchized into sub-sub criteria.

b. Communication / Control system sub-criterion (Telemetry and Sign Boards sub-criterion): indicates the performance of the operation control system, mainly the telemetry system and the communication system. This sub-criterion includes telephone lines network, the information system, and thus the communication system connected to the control room. This sub-criterion can also be hierarchized into sub-sub criteria.

5. Environmental Criteria:

a. Air Quality sub-criterion: indicates the condition and quality of air in the subway station. STM does not consider air quality as one of the condition assessment criteria, on the basis that air quality is controlled mainly by the ventilation system.

b. Noise Pollution sub-criterion: indicates the condition and quality of noise in the subway station. Similar to the air quality criterion, the STM does not consider noise control as a sub-criterion affecting the condition of subway stations.

The environmental criteria are presented in the above hierarchy description, but will be ignored afterwards in the SSDI model.

Figure 3.3 shows the hierarchy of the criteria used for the SSDI model.

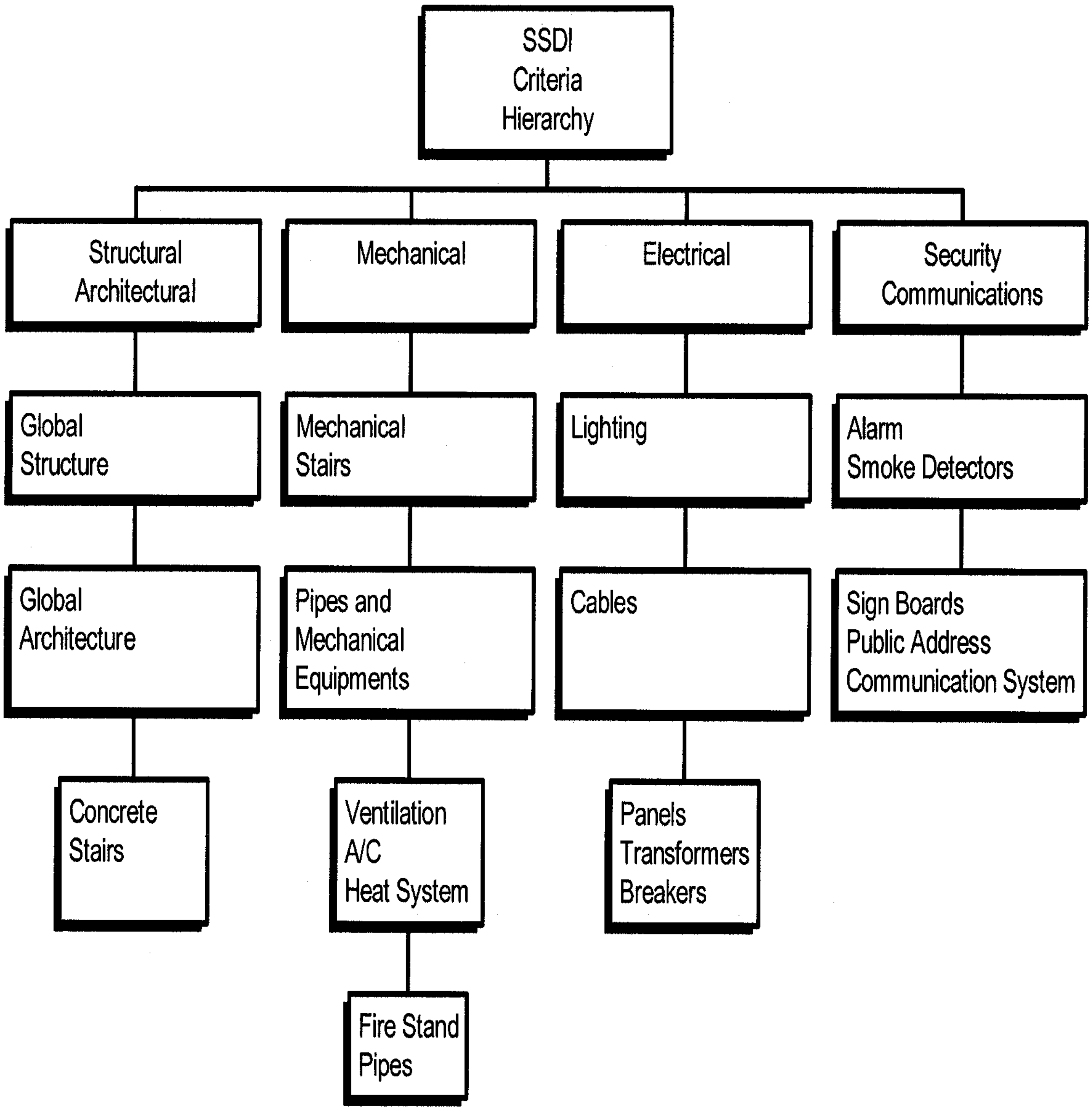


Figure 3.3 The SSDI model Criteria Hierarchy

3.4 Criteria Performances Evaluation

The SSDI model allows for the measurement of the sub-criteria in any proper method: a scale, descriptive manner, formulae, etc., as described in the literature. As a procedure to show a type of measurement, with conformity to STM and MTQ, the performances of sub-criteria C_i (*i= 1 to 14, total No. of criteria*) are measured on the following scale in Table 3.1.

Table 3.1 Field Inspection Scale

Scale	Description
1	Critical Condition
2	Deficient Condition
3	Poor Condition
4	Acceptable Condition
5	Good Condition

It should be noted that the SSDI is not bound to this performance measurement scale; the above scale is only a guideline. In order to get a better representation of the difference of the criteria performance between two stations, especially in the aggregation process, it is found that it is better to have the criteria performance (or simply value) $V_i[C_i]$ in a scale from 0 to 100. Thus if the criteria values – performances - are measured in the direction from 'Bad' to 'Good' – i.e. 0 means critical condition, 100 means good condition – then, Equation 3.1 applies:

$$V_i[C_i] = 25 \,.\, C_i - 25 \qquad (3.1)$$

Where C_i means the sub-criterion value from the filed inspection scale (Table 3.1).

If the values are measured in the direction from 'Good' to 'Bad' condition, then, Equation 3.2 applies:

$$V_i[C_i] = -25 \,.\, C_i + 125 \qquad (3.2)$$

Thus $V_i[C_i] = 0$ means 'Good' condition, and $V_i[C_i] = 100$ means 'Critical' condition. The SSDI model uses the latter evaluation of the criteria values $V_i[C_i]$; i.e. Equation 3.2.

3.5 Criteria Weights Evaluation

The criteria hierarchy allows for the use of the AHP for the evaluation of the criteria relative importance or weights. The AHP has been proven by previous researches as one of the most efficient MCDA tools for the evaluation of criteria weights, as long as they are arranged in a hierarchy. Table 3.2 illustrates sample matrices of AHP, according to the hierarchy of the criteria. The values of the matrices are the values of the first questionnaire used for data collection (see Chapter 4).

Table 3.2 AHP Sample Matrices for Levels 1 and 2

Hierarchy level 1: Criteria

Level 1 Matrix	Structural / Architectural	Mechanical	Electrical	Security / Communication	Function Weights w_{fj}
Structural / Architectural	1	3	2	1	0.33
Mechanical	1/3	1	1	1/5	0.11
Electrical	1/2	1	1	1/3	0.14
Security / Communication	1	5	3	1	0.42
$\Sigma =$	2.8	10.0	7.0	2.5	1.0

Hierarchy level 2: Sub-criteria

Structural / Architectural		C1	C2	C3	Sub-Criteria Weights w_{ci}
C1	Global Structure	1	3	4	0.62
C2	Global Architecture	1/3	1	1/2	0.16
C3	Concrete (Fixed) Stairs	1/4	2	1	0.22
	$\Sigma =$	1.6	6.0	5.5	1.0

Mechanical		C4	C5	C6	C7	Sub-Criteria Weights w_{ci}
C4	Mechanical Stairs	1	2	3	5	0.45
C5	Pipes and Mechanical Equipments	1/2	1	1/3	1/3	0.11
C6	Ventilation, A/C, Heat	1/3	3	1	1/5	0.16
C7	Fire Stand Pipes	1/5	3	5	1	0.28
	$\Sigma =$	2.0	9.0	9.3	6.5	1.0

Security / Communication		C11	C12	Sub-Criteria Weights w_{ci}
C11	Alarm / Security / Smoke Detectors	1	5	0.83
C12	Sign Boards / Public Address / Communication System (Telemetry)	1/5	1	0.17
	$\Sigma =$	1.2	6.0	1.0

Electrical		C8	C9	C10	Sub-Criteria Weights w_{ci}
C8	Lighting	1	2	1	0.41
C9	Cables	1/2	1	1	0.26
C10	Panels, Transformers, & Breakers	1	1	1	0.33
	$\Sigma =$	2.5	4.0	3.0	1.0

All the axioms of AHP can be applied to the SSDI model. No criteria weight reversal will be encountered, as long as the total number of criteria and sub-criteria is fixed. If any transit authority changes the criteria, new weights must be evaluated. As a general formula:

For level 1: Functional criteria

Where f_1 = Structural / Architectural functional criterion

f_2 = Mechanical functional criterion

f_3 = Electrical functional criterion

f_4 = Security / Communication functional criterion

Thus, the sum of functional criteria weights should be a unity as in Equation 3.3:

$$\sum_{j=1}^{4} W_{fj} = 1.0 \qquad (3.3)$$

Where j = 1 to 4

And for level 2, the sum of sub-criteria (C_i) weights should be a unity, as in Equations 3.4, and 3.4a to 3.4d

$$\sum_{i=1}^{n} W_{ci} = 1.0 \qquad (3.4)$$

$$W_{C1} + W_{C2} + W_{C3} = 1 \qquad (3.4a)$$

$$W_{C4} + W_{C5} + W_{C6} + W_{C7} = 1 \qquad (3.4b)$$

$$W_{C8} + W_{C9} + W_{C10} = 1 \qquad (3.4c)$$

$$W_{C11} + W_{C12} = 1 \qquad (3.4d)$$

Where C_i = sub-criteria, i = to 12

And, thus, Equation 3.4e combines the sub-criteria weights and functional weights as follows:

$$(W_{C1}+W_{C2}+W_{C3}).W_{f1}+ (W_{C4}+W_{C5}+W_{C6}+W_{C7}).W_{f2} + (W_{C8}+W_{C9}+W_{C10}).W_{f3} + (W_{C11} + W_{C12}).W_{f4} = 1 \quad (3.4e)$$

Therefore, the global weight of the criteria can be defined in Equation 3.5:

$$W_{g1} = W_{C1} . W_{f1}$$

Or, generally:

$$W_{gi} = W_{ci} . W_{fj} \quad (3.5)$$

Where W_{gi} = global weight of sub-criterion C_i

3.6 Lower and Upper Datum

The weakness of any outranking MCDA method (PROMETHEE for example) is that the final decision is on a cardinal scale, thereby resulting in fluctuations. The interest of the new model is to determine an index on an ordinal scale (a fixed scale between a '0%' state and a '100%' state). The SSDI model overcomes the final scale problem by introducing a lower datum S_0 and an upper datum S_{100}. These boundaries will only appear in the algorithm and aggregation, but will be hidden in the input and the final output.

SSDI considers two fictitious Stations with Lower and Upper bound Datum:

S_0 = Lower Datum Station = Very Good Station, defined in Equation 3.6:

$$V_i [C_i (S_0)] = 0 \quad (3.6)$$

S_{100} = Upper Datum Station = Very Bad Station, defined in Equation 3.7:

$$V_i [C_i (S_{100})] = 100 \quad (3.7)$$

The lower and upper datum allow the outranking of a specific station between these two fictitious, and at the same time real, extreme cases. Fictitious, because they are not physical stations, and they appear only in the mathematical calculation. Nevertheless, these stations are real because they form concrete boundaries to the Diagnosis Index. Furthermore, the introduction of the boundaries stations allows SSDI to overcome the problem of rank reversal inherited in any outranking method.

3.7 Pseudo-Criteria Evaluation

Pseudo-criteria allows SSDI to consider uncertainties and imprecision in criteria performance. Thus, applying a preference for each of the criterion to specific conditions, states, or performances of the criterion itself does this. These conditions or states are the thresholds, which would be interpreted in a generalized preference function (Hyde et al., 2003).

3.7.1 Thresholds Evaluation

The PROMETHEE general algorithm uses Indifference and Preference thresholds for each sub-criterion. These two thresholds indicate the indifference limit and the strong preference limit of one criterion performance between two alternatives.

With the introduction of the lower datum station whose criteria performance is always equal to 0, and the upper datum station, whose criteria performance is always equal to 100, the preference and indifference thresholds are transformed into two physical limits as follows:

- The Critical Threshold (CT) is the limit value of each criterion C_i beyond which it is considered dangerous or critical.
- The Tolerance Threshold (TT) is limiting value of each criterion C_i below which it is considered tolerable or safe.

- Thus, the Preference Threshold (PT), measures the preference of the criterion relative to the lower and upper datum, and is equal to CT, if criteria are measured in the 'Bad' Condition direction; or equal to 100-TT, if criteria are measured in the 'Good' condition direction.
- The Indifference Threshold (IT), measures the indifference of the criterion relative to lower and upper datum, and is equal to TT, if criteria are measured in the 'Bad' Condition direction; or equal to 100-CT, if criteria are measured in the 'Good' condition direction, as shown in Figure 3.4.

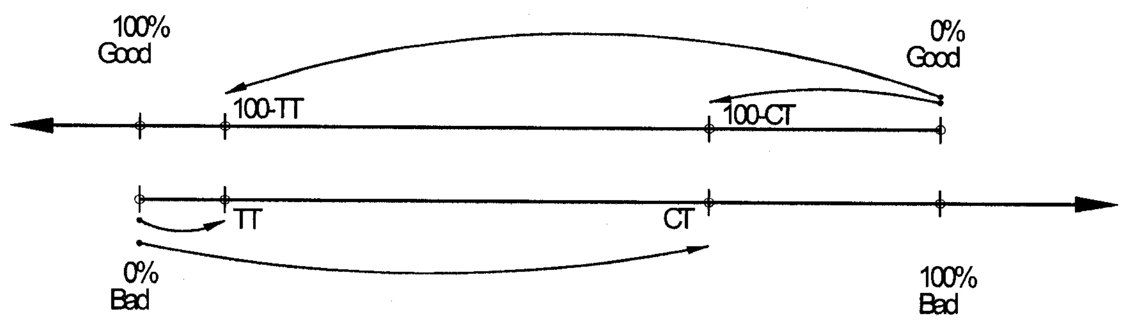


Figure 3.4 The SSDI Model Thresholds

The SSDI model considers S_0 as an extremely good station, and S_{100} as an extremely bad station (refer to section 3.6), thus Equations 3.8a and 3.8b are defined as:

$$CT = PT \tag{3.8a}$$

$$TT = IT \tag{3.8b}$$

3.7.2 Generalized Preference Function

The *Generalized Performance Function* (gpf), P[C], is the function that best represents the criteria preference to both the lower datum and upper datum. The gpf trade-off points are the critical (preference to extremely good state S_0) and the tolerance (indifference to extremely good state S_0), as shown in Figure 3.5 (type 5 in Fig. 2.5).

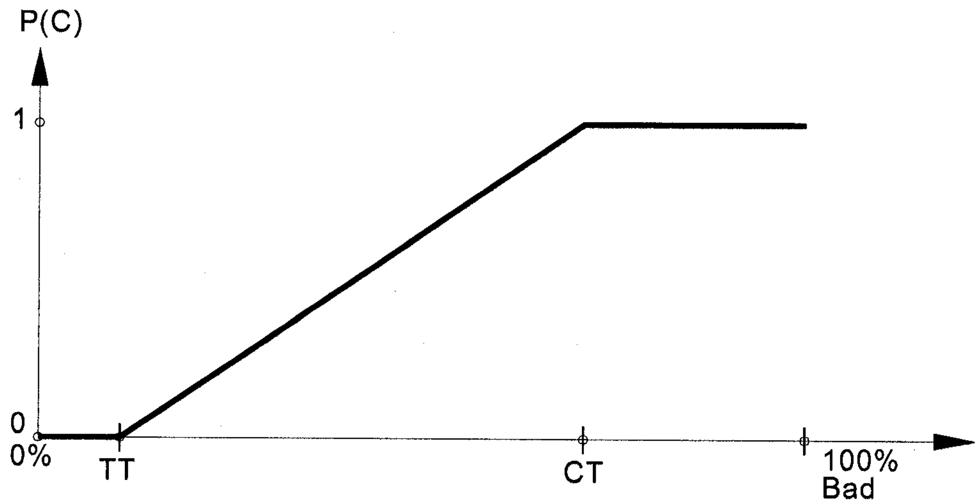


Figure 3.5 The SSDI Model Generalized Preference Function (gpf)

In a mathematical form, Equation 3.9 illustrates the definition of the SSDI gpf:

$$P_i\,[V_i\,[C_i(S)]] = \begin{cases} 0; & \text{if } V\,[C_i(S)] < TT \\ 1; & \text{if } V\,[C_i(S)] > CT \\ (V_i\,[C_i(S)] - TT) / (CT\text{-}TT); & \text{if } TT \le V\,[C_i(S)] \le CT \end{cases} \quad (3.9)$$

Where i = 1 to n; n = total No. of criteria; and S = given Station.

- P=0 means the criterion performance is tolerable or indifferent to the lower datum station, or extremely good station S_0.

- P=1 means the criterion performance is critical, or 100% preferred than the lower datum, extremely good station S_0.

3.8 Multi-Criteria Aggregation

In this step, the SSDI model uses the principles of PROMETHEE II criteria aggregation. It outranks any specific Station S_k to S_0 and S_{100} (Fixed datum boundaries) as illustrated in the following series of Equations 3.10:

$V_i\,[C_i(S_0)] - V_i[C_i(S_0)] = 0 \quad\Rightarrow\quad P_i\,(S_0, S_0) = 0 \Rightarrow \quad S_0\ I\ S_0$

$V_i[C_i(S_0) - V_i[C_i(S_k)] = -V_i\,[C_i\,(S_k)] < 0 \quad\Rightarrow\quad P_i\,(S_0, S_k) = 0 \Rightarrow \quad S_k\ P\ S_0$

$V_i[C_i(S_0) - V_i[C_i(S_{100})] = 0 - 100 = -100 < 0 \Rightarrow P_i\,(S_0, S_{100}) = 0 \quad\Rightarrow\quad S_{100}\ P\ S_0$

$V_i[C_i\,(S_k)] - V_i[C_i(S_0)] = V_i\,[C_i\,(S_k)] \quad\Rightarrow\quad \mathbf{P_i\,(S_k)}$

$V_i[C_i(S_k)] - V_i[C_i(S_k)] = 0 \quad\Rightarrow\quad P_i\,(S_k, S_k) = 0 \Rightarrow \quad S_k\ I\ S_k$

$V_i[C_i(S_k)] - V_i[C_i(S_{100})] = V_i\,[C_i\,(S_k)] -100 < 0 \Rightarrow P_i\,(S_k, S_{100}) = 0 \Rightarrow \quad S_{100}\ P\ S_k$

$V_i[C_i(S_{100})] - V_i[C_i(S_0)] = 100 \quad\Rightarrow P_i\,(S_{100}, S_0) = 1 \quad\Rightarrow\quad S_{100}\ P\ S_0$

$V_i[C_i(S_{100})] - V_i[C_i(S_k)] = 100 - V_i[C_i(S_k)] \quad\Rightarrow\quad \mathbf{P_i\,(100\text{-}S_k)}$

$$V_i\,[C_i(S_{100})] - V_i[C_i(S_{100})] = 0 \quad\Rightarrow\quad P_i\,(S_{100}, S_{100}) = 0 \Rightarrow S_{100}\ I\ S_{100} \tag{3.10}$$

All the preference values are zeros except for $P_i(S_k)$ and $P_i(100\text{-}S_k)$, and of course $P(S_{100}, S_0)$ which is equal to 1.

3.9 Multi-Criteria Preference Index

3.9.1 General Definition

Generally, the multi-criteria preference index Π, between any two stations S_1 and S_2, Π $[S_1,S_2]$, indicates the weighted preference of S_1 to S_2 in all criteria, as defined in Equation 3.11:

$$\Pi\,[S_1,S_2] = \sum_{i=1}^{n} W_{Ci} * P_i(S_1,S_2) \qquad (3.11)$$

Where $0 \leq \Pi[S_1,S_2] \leq 1$

And i = *1 to n* is the total number of sub-criteria in one functional criteria set (refer to 3.9.2).

3.9.2 SSDI Model Application

For any station S_k compared to S_0 and S_{100}, we have the following multi-criteria indices in the series of Equations 3.12, calculated for each function separately:

$$\Pi\,[S_0,S_0] = \sum_{i=1}^{n} W_{ci} \, . \, P_i\,(S_0,S_0) = 0$$

$$\Pi\,[S_0,S_k] = \sum_{i=1}^{n} W_{ci} \, . \, P_i\,(S_0,S_k) = 0$$

$$\Pi\,[S_0,S_{100}] = \sum_{i=1}^{n} W_{ci} \, . \, P_i\,(S_0,S_{100}) = 0$$

$$\Pi\,[S_k,S_0] = \sum_{i=1}^{n} W_{ci} \, . \, P_i\,(S_k,S_0)$$

$$\Pi\,[S_k,S_k] = \sum_{i=1}^{n} W_{ci} \, . \, P_i\,(S_k,S_k) = 0$$

$$\Pi\,[S_k,S_{100}] = \sum_{i=1}^{n} W_{ci}\,.\,P_i\,(S_k,S_{100}) = 0$$

$$\Pi\,[S_{100},S_0] = \sum_{i=1}^{n} W_{ci}\,.\,P_i\,(S_{100},S_0) = 1$$

$$\Pi\,[S_{100},S_k] = \sum_{i=1}^{n} W_{ci}\,.\,P_i\,(S_{100},S_k)$$

$$\Pi\,[S_{100},S_{100}] = \sum_{i=1}^{n} W_{ci}\,.\,P_i\,(S_{100},S_{100}) = 0 \qquad (3.12)$$

These above calculations are the echoes of the preference function calculation in the previous paragraph.

3.10 Station Outranking

3.10.1 General Definition

The station outranking is evaluated depending on the flow calculation. The flow calculation in PROMETHEE is as follows:

The measure of strength of S_1, as defined in Equation 3.13:

$$\Phi^{+}(S_1) = \sum_{i=1}^{n} \Pi i(S_1,S_2) \qquad (3.13)$$

The measure of weakness of S_1, as defined in Equation 3.14:

$$\Phi^{-}(S_1) = \sum_{i=1}^{n} \Pi i(S_2,S_1) \qquad (3.14)$$

Thus, the Net Flow is defined in Equation 3.15:

$$\Phi^{net}(S_1) = \Phi^{+}(S_1) - \Phi^{-}(S_1) \qquad (3.15)$$

And finally, the ranking of S_1 to S_2 is as defined in Equation 3.16:

$$\Phi^{net}(S_1) \geq \Phi^{net}(S_2) \qquad (3.16)$$

$\Rightarrow$ S_1 Outranks S_2 in all criteria

3.10.2 SSDI Model Application

In the SSDI model, the flows of S_k, S_0 and S_{100} are evaluated in the following equations. Note that the evaluation is performed for each function separately.

The measure of strength of S_k is defined in Equations 3.17a, 3.17b and 3.18:

$$\Phi^+(S_j) = \Pi\,[S_k,S_0] + \Pi\,[S_k,S_{100}] + \Pi\,[S_k,S_k] \tag{3.17a}$$

$$\Phi^+(S_k) = \Pi\,[S_k,S_0] + 0 + 0 = \Pi\,[S_k,S_0] \tag{3.17b}$$

$$\Phi^+(S_k) = \sum_{i=1}^{n} W_{ci} * P_i(S_k) = \sum_{i=1}^{n} W_{ci} * P_i(V_i[C_i(S_k)]) \tag{3.18}$$

The measure of weakness of S_k is defined in Equations 3.19a, 3.19b and 3.20:

$$\Phi^-(S_k) = \Pi\,[S_0,S_k] + \Pi\,[S_k,S_k] + \Pi\,[S_{100},S_k] \tag{3.19a}$$

$$\Phi^-(S_k) = 0 + 0 + \Pi\,[S_{100},S_k] \tag{3.19b}$$

$$\Phi^-(S_k) = \sum_{i=1}^{n} W_{ci} * P_i(100,S_k) = \sum_{i=1}^{n} W_{ci} * P_i(100 - V_i[C_i(S_k)]) \tag{3.20}$$

The Net Flow is defined in equations 3.21 and 3.22:

$$\Phi^{net}(S_k) = \Phi^+(S_k) - \Phi^-(S_k) \tag{3.21}$$

$$\Phi^{net}(S_k) = \sum_{i=1}^{n} W_{ci} * P_i(V_i[C_i(S_k)]) - \sum_{i=1}^{n} W_{ci} * P_i(100 - V_i[C_i(S_k)])$$

$$= \text{Measure of strength} - \text{measure of weakness} \tag{3.22}$$

The net flows of the two datum stations S_0 and S_{100} are evaluated in the same manner as S_k. The absolute range of net flows is the reason behind introducing the two datum stations. For the absolute range flows calculation, only S_0 and S_{100}

are considered. Therefore, the Φ^{net} of S_0 with S_{100}, and similarly S_{100} with S_0 is defined in the following Equations 3.23 and 3.24; and thus is true for all functions.

$$\Phi^{net}(S_0) = -1 = \text{Minimum value} \tag{3.23}$$

$$\Phi^{net}(S_{100}) = 1 = \text{Maximum value} \tag{3.24}$$

The above values are not variable, but fixed values and are independent from the status of S_k, and from its function.

Now, the rank according to the net flows of S_k to S_0 and S_{100}, for each function, is shown in Equations 3.25 and 3.26:

$$\Phi^{net}(S_0) \leq \Phi^{net}(S_k) \leq \Phi^{net}(S_{100}) \tag{3.25}$$

$$-1 \leq \sum_{i=1}^{n} W_{ci} * P_i(V_i[C_i(S_k)]) - \sum_{i=1}^{n} W_{ci} * P_i(100 - V_i[C_i(S_k)]) \leq 1 \tag{3.26}$$

Hence, the outranking of S_k to S_0 and S_{100} for each function becomes a fixed value between fixed lower and upper limits [-1,1]; i.e. an ordinal scale.

3.11 Functional Diagnosis Index (FDI)

The use of the net flows is important, because it can be treated as an attribute to the station. By using S_0 and S_{100}, the net flow of a specific station becomes a multi-criteria attribute of the different functions of the station itself. This attribute is physically meaningful, since:

- It considers the preference of each sub-criterion to the thresholds (critical and tolerance).
- It considers the relative weights of each sub-criterion.
- It takes into account the uncertainty in the sub-criteria performance.

Therefore, the multi-criteria aggregation in the SSDI model (with S_0 and S_{100}) can be transformed into a simple additive MAUT function. The SSDI model transforms the 'multi-function' attribute (ranging from [-1,1]) into a *Functional Diagnosis Index* (FDI) scale (ranging from [10,0]) by a straight-line Equation 3.27 as follows:

$$FDI\ (S_k) = -5\ .\ \Phi^{net}\ (S_k) + 5 \qquad (3.27)$$

Figure 3.6 shows the FDI versus the Net Flow; i.e. the FDI equation plot.

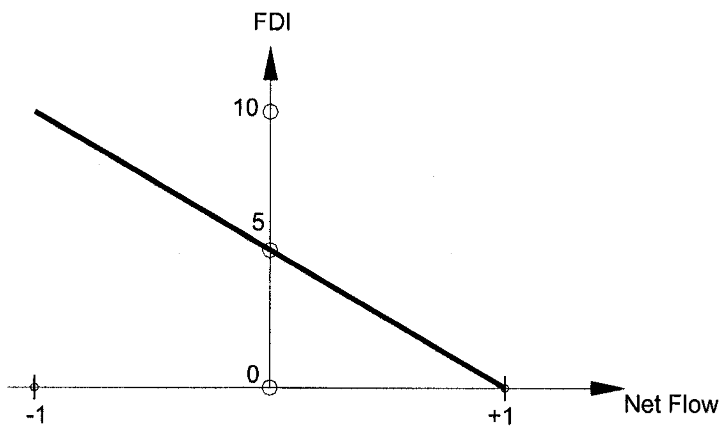


Figure 3.6 FDI Equation

As illustrated in Equation 3.27 and Figure 3.6:

- For S_0, which is an absolute 'Very Good Station', Equation 3.28 applies:

$$\Phi^{net} = -1 \qquad \Rightarrow \qquad FDI\ (S_0) = 10 \qquad (3.28)$$

- For S_{100}, which is an absolute 'Very Bad Station', Equation 3.29 applies:

$$\Phi^{net} = +1 \qquad \Rightarrow \qquad FDI\ (S_{100}) = 0 \qquad (3.29)$$

- Thus, for any station S_k, Equations 3.30 and 3.31 apply:

$$-1 \leq \Phi^{net}\ (S_k) \leq +1 \qquad (3.30)$$

Hence: $10 \leq FDI\ (S_k) \leq 0$ (3.31)

Or: Extremely Good $\leq S_k \leq$ Extremely Bad

3.12 Global Station Diagnosis Index (SDI)

In order to transform the *Functional Diagnosis Index* (FDI) into a global *Station Diagnosis Index* (SDI), the theory of MAUT can be used. FDI is a measure (an index) in an ordinal scale with fixed limit values (0 and 10), similar to trade-offs points. Thus the FDI can be treated as an attribute of each function. A function or relation between these attributes can determine a global attribute. This global attribute will be the *Station Diagnosis Index* (SDI). The choice of a particular function is one of the most difficult tasks. As a matter of fact, the choice of an MAUT function is dependent on the physical condition logic. The actual management logic supposes that if a station is failing in the structural function ($V[C_{structure}] >> CT_{structure}$), it needs instant rehabilitation measures (FDI is very low). On the other hand, the other functions may be in a good condition, and they need no urgent rehabilitation measures (FDI is high). Since the physical situation is obvious in the FDI analysis, however it should be reflected in the global SDI result.

An additive function will not adequately reflect the above-mentioned situation, because the global index would be much higher than the physical condition of the station. Thus a multiplicative form of MAUT function is the best choice for the SSDI model, the formula of the function would be defined in Equations 3.32 and 3.33:

$$SDI = \prod_{j=1}^{4} FDI_j^{Wfj} \tag{3.32}$$

Or $$SDI = FDI_{str}^{Wfstr} * FDI_{mech}^{Wfmech} * FDI_{elec}^{Wfsmech} * FDI_{comm}^{Wfscomm} \tag{3.33}$$

Where W_{fi} = Functional Weight.

The transit authority now has the choice to build its MR planning based on FDI, SDI or both. This flexibility is a big advantage for the transit authority decision makers (managers) for budget allocation (crew planning, funding, etc.) since it is easy to update.

3.13 FDI and SDI Condition Scales

After calculating the FDI and the SDI for the global station condition, the decision maker or management has to interpret theses indices in a proper manner. SSDI provides a methodical way in determining what these indices mean. The research, in close coordination with STM - via their input, inspection reports, description of rehabilitation programs, and verification of the SSDI model - has developed a SSDI (FDI and SDI) condition scale with the percentage deterioration for each scale, and proposed actions. The scale chosen for both FDI and SDI is from 0 to 10 (refer to sections 3.12 and 3.13), a universally accepted, a logical, and an easy to interpret scale. Table 3.3 describes the FDI scale. The values of CT and TT for each function will be discussed later on. The same table will be completed at the end of next chapter.

Table 3.3 FDI Condition Scale

FDI	Description	Deterioration Level *(%)*	Proposed Action
8 < FDI ≤ 10	*Good*	<TT For each function	Long Term: * Expertise < 2 years * Physical < 5 years Review in 2 years
6 < FDI ≤ 8	*Medium*	>TT & <¼(CT-TT) for each function	Medium Term: * Expertise < 1 year * Physical < 2 years Review in 1 year
3 < FDI ≤ 6	*Deficient*	>¼(CT-TT) & <¾(CT-TT) for each function	Short Term: * Expertise < 6 months * Physical < 1 year Review in 6 months
0 ≤ FDI ≤ 3	*Critical*	>CT for each function	Immediate: Physical intervention Now

The values of ¼(CT-TT) and ¾(CT-TT) in the '*deterioration level*' column will be confirmed after the sensitivity analysis section (next chapter). The '*description*' column reflects the inspection reports input scale, while the '*proposed action*' column is taken directly from (but not bound to) the STM inspection reports, the 'Réno-Stations' and the 'Réno-Système' reports. The '*proposed action*' used by STM may be generally used as proposed in the SSDI model. Additional descriptions will be added to this table in the next chapter. The 'physical intervention' comprises physical rehabilitation measures, while the 'expertise intervention' indicates an expert opinion (or report), testing and/or measures.

Table 3.4 describes the SDI scale, which has the same parameters of the above FDI scale table.

Table 3.4 SDI Condition Scale

SDI	Description	Deterioration Level *(%)*	Proposed Action
8 < SDI ≤ 10	*Good*	<TT Structural or, <TT Communications or, <TT Electrical or, <TT Mechanical	Long Term: * Expertise < 2 years * Physical < 5 years Review in 2 years
6 < SDI ≤ 8	*Medium*	>TT & <¼(CT-TT) Structural or, >TT & <¼(CT-TT) Communications or, >TT & <¼(CT-TT) Electrical or, >TT & <¼(CT-TT) Mechanical	Medium Term: * Expertise < 1 year * Physical < 2 years Review in 1 year
3 < SDI ≤ 6	*Deficient*	>¼(CT-TT) & <¾(CT-TT) Structural or, >¼(CT-TT) & <¾(CT-TT) Communications or, >¼(CT-TT) & <¾(CT-TT) Electrical or, >¼(CT-TT) & <¾(CT-TT) Mechanical	Short Term: * Expertise < 6 months * Physical < 1 year Review in 6 months
0 ≤ SDI ≤ 3	*Critical*	>CT Structural or, >CT Communications or, >CT Electrical or, >CT Mechanical	Immediate: Physical intervention Now

From Table 3.4, it is clear that for a 'Critical' SDI for example, one or several of the functions can be critical, while the rest could be in a better condition. This requires the immediate intervention for the specific function(s) only as described in the '*proposed action*' column. In this case, FDI should be considered in order to adequately manage the MR for the subway station.

3.14 Analysis of the SSDI Model Theory

The SSDI model is developed based on MCDA tool – PROMETHEE – inherited from the European school of decision analysis. The PROMETHEE method ranks alternatives (or stations) on a cardinal scale. With the introduction of the lower and upper boundaries to PROMETHEE, SSDI allows the transformation of the rank from a fluctuating to a fixed one. The SSDI, thus, ranks a specific station between two fictitious but fixed boundaries. Hence, the SSDI model transforms an outranking method into a pure multi-attribute method, where the net flows can be considered as multi-criteria attributes for each station.

On the other hand, the advantages of SSDI over any MAUT method are numerous. Firstly, it leaves the flexibility for measuring the performance or value of each sub-criterion. Secondly, SSDI uses the pseudo-criteria concept. This concept allows SSDI to consider the uncertainties, imprecision, and lack of information for the criteria performance. In addition, it measures the significance of each sub-criteria performance by comparing it to the physical critical and tolerance thresholds. The use of the physical thresholds is crucial since it allows the decision maker, using SSDI, to define engineering expert 'trade-off points' of the preference function.

While MAUT relies completely on experts to define theoretical trade-off points and shape of function, SSDI uses physical attributes ($\Phi^{net}{}_{S0}$ and $\Phi^{net}{}_{S100}$ based on CT and TT) in order to set the function of both FDI and then of SDI. The engineer or manager has to evaluate the physical thresholds for each sub-criterion. And that is relatively an easier job than estimating the trade-off points and shape of function for a regular MAUT method. The critical and tolerance threshold can be estimated in reference to the factors of safety used in the design of different functional sub-criteria or elements. Hence, SSDI is flexible, while directing the transit management to the diagnosis of the subway station. It also indicates the level of deterioration, and proposes actions for intervention. The SSDI model has proved that on outranking method can be transformed into a MAUT method by just introducing fixed lower and upper values. In addition, the pseudo-criteria concept can serve as a guide in defining the function of a MAUT model.

4 Data Collection and Analysis

4.1 Introduction

The SSDI model as illustrated in the previous chapter is based on crucial data, and primarily comprises the input for the model. These data are criteria performances, criteria weights and pseudo-criteria thresholds. Collecting data and analysing them is an important step in implementing SSDI into real life. Data were collected from different sources. A statistical analysis is needed in order to prove the relevance of the values of data collected. Afterwards, a sensitivity analysis of the data collected was performed on the SSDI model in order to study the effect of the imprecision in the collection of the data. Finally, the data collected was incorporated into the SSDI model; and the final FDI and SDI condition scale tables were completed.

4.2 Data Collection and Statistical Analysis

The data for the SSDI model are divided into three categories. The first category includes the criteria (or more precisely the sub-criteria) performance values. The second category comprises the AHP matrices for the criteria weights. And the third one consists of the pseudo-criteria thresholds (critical and tolerance). The data were collected through questionnaires, interviews, inspection reports, and MR planning reports. The questionnaires targeted the subway stations practitioners (engineers, inspectors and management). The inspection reports were provided by the STM rehabilitation team (engineering unit) and the MR reports were provided by the STM planning unit. Each category of data were analysed separately, although two of them – criteria weights and pseudo-criteria thresholds - shared the same source (questionnaires).

4.2.1 Criteria Performances

The criteria performances are the kind of data specific to every transit authority. STM engineers and managers, responsible for the rehabilitation planning, were contacted via interviews. The interviews were spread for almost a year, and the data are mainly from four sources: 1) inspection reports for stations under the 'Réno-Station' program, 2) 'Réno-Station' general rehabilitation description report, 3) 'Réno-Système' program description report, and 4) verbal input for the STM engineers. For confidentiality reasons, the STM insisted not to transmit information to the public, especially names of stations with specific criteria condition. For this reason, this research uses real data (criteria values) without assigning them the real names of the stations (arbitrary names were chosen).

The criteria performance is measured in STM by an inspection performance scale (1 to 5) explained in the SSDI model chapter (see Table 3.1). Moreover, the inspection scale (1 to 5) is transformed into a scale from 0 to 100 (see Equation 3.2). The inspection reports and 'Réno-Station' reports have provided data for all the criteria, except the mechanical stairs, the communication system and security system. These latter ones are part of the 'Réno-Système' program report. 'Réno-Système' specifies that the communication and security system, in addition to the mechanical stairs and their rehabilitation, since they are old and in a 'poor' condition. An inspection performance equal to 2 (out of 5) is thus assigned to the corresponding sub-criteria. The criteria performances collected are used in the SSDI model application to STM stations (see next chapter).

4.2.2 Criteria Weights

The SSDI model utilizes AHP to evaluate the weights of the criteria. A questionnaire was developed and distributed to most of the North-American transit authorities engineers (refer to Appendix A for a sample questionnaire). 40 questionnaires were sent to the STM, and 10 for each of the TTC, MTA NYCT, Massachusetts Bay Transit Authority (MBTA), Cal Train Authority and Chicago Transit Authority. In addition, questionnaires were sent to engineers directly linked to the inspection of subway stations in Montreal (STM provided the names). Only 24 questionnaires were received, 16 from the STM and engineers linked to it, 4 from TTC, 2 from MTA NYCT, and 2 from MBTA. For less than 30 questionnaires, statistical analysis is required in order to prove that the results from the 24 questionnaires have a certain level of confidence (Montgomery and Runger, 2007).

A statistical analysis was performed on the weights gathered from the questionnaires. The average (mean) values of the weights, standard deviation, standard error, lower and upper values for a 68% level of confidence, and similarly for a 95% level of confidence were evaluated. Statistical tests – the Chi-Squared, the Anderson-Darling, and the Kolmogorov-Smirnov tests - were performed (Zayed and Nosair, 2006). 'Best Fit' software, consisting of a family of @Risk software, was used to best fit the questionnaire values for the best probability density function.

Table 4.1 shows the results of the statistical analysis for the criteria weights, evaluated from the AHP of 24 questionnaires.

Table 4.1 Criteria Weights Statistical Analysis Results

Criteria Weights		C1	C2	C3	C4	C5	C6
Average	μ (%)	18.8	5.1	12.2	3.9	4.1	4.4
Standard Deviation	σ (%)	5.9	2.4	5.0	2.2	2.0	2.5
Standard Error	ε (%)	1.2	0.5	1.0	0.5	0.4	0.5
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ (%)	17.6	4.6	11.2	3.4	3.7	3.9
	$\mu+\sigma/n^{\dagger}$ (%)	20.0	5.6	13.2	4.3	4.5	4.9
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ (%)	16.4	4.1	10.2	3.0	3.2	3.4
	$\mu+2\sigma/n^{\dagger}$ (%)	21.2	6.1	14.2	4.8	4.9	5.4
Normal Distribution	Skewness	-0.24	0.55	0.70	0.20	0.18	0.23
	Kurtosis	2.09	2.09	2.30	1.40	1.84	1.58
Chi-Sq Test††	Test Value	2.25	6.47	2.7	8.08	2.7	2.25
	P-Value	0.69	0.17	0.62	0.09	0.62	0.69
A-D Test†††	Test Value	0.26	0.72	0.94	1.19	0.41	0.77
	P-Value	0.25	0.1	0.03	0.005	0.25	0.05
K-S Test††††	Test Value	0.1	0.18	0.15	0.19	0.11	0.16
	P-Value	0.15	0.1	0.15	0.05	0.15	0.1

† n = Total No. of data, i.e. 24

†† Chi-Sq statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

††† A-D (Anderson-Darling) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

†††† K-S (Kolmogorov-Smirnov) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

Table 4.1 (Continued) Criteria Weights Statistical Analysis Results

Criteria Weights		C7	C8	C9	C10	C11	C12
Average	μ (%)	5.4	5.0	6.3	7.6	21.2	6.0
Standard Deviation	σ (%)	2.4	2.5	3.1	3.8	8.3	3.3
Standard Error	ε (%)	0.5	0.5	0.6	0.8	1.7	0.7
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ (%)	4.9	4.5	5.7	6.8	19.5	5.3
	$\mu+\sigma/n^{\dagger}$ (%)	5.9	5.5	7.0	8.4	22.9	6.7
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ (%)	4.4	3.9	5.1	6.0	17.9	4.7
	$\mu+2\sigma/n^{\dagger}$ (%)	6.4	6.0	7.6	9.1	24.6	7.4
Normal Distribution	Skewness	0.19	0.005	-0.01	0.33	0.26	0.45
	Kurtosis	2.02	1.83	2.05	1.86	1.68	2.04
Chi-Sq Test††	Test Value	2.7	2.7	1.4	2.25	2.25	3.9
	P-Value	0.62	0.62	0.84	0.51	0.69	0.42
A-D Test†††	Test Value	0.32	0.37	0.17	0.25	0.65	0.54
	P-Value	0.25	0.25	0.25	0.13	0.1	0.25
K-S Test††††	Test Value	0.12	0.11	0.07	0.15	0.16	0.12
	P-Value	0.15	0.15	0.15	0.15	0.15	0.15

Table 4.1 shows that C11 (Alarm, Security and Smoke Detectors), C1 (Global Structure), and C3 (Concrete Stairs) are the most important criteria, with 21.2%, 18.8% and 12.2% importance weights respectively. Together they form 52.2% of the total weights of the SSDI model. Figure 4.1 shows the distribution of the criteria weights in a bar chart, while Figure 4.2 shows the function weights distribution in a pie chart. The rest of the criteria are approximately similar in importance, a little more for C10 (Panels, Transformers and Breakers), with 7.5%. This result is logical, since the security and structural conditions are the

most factors that affect the public security. However, the rest of the criteria affect public security, but to a lesser degree.

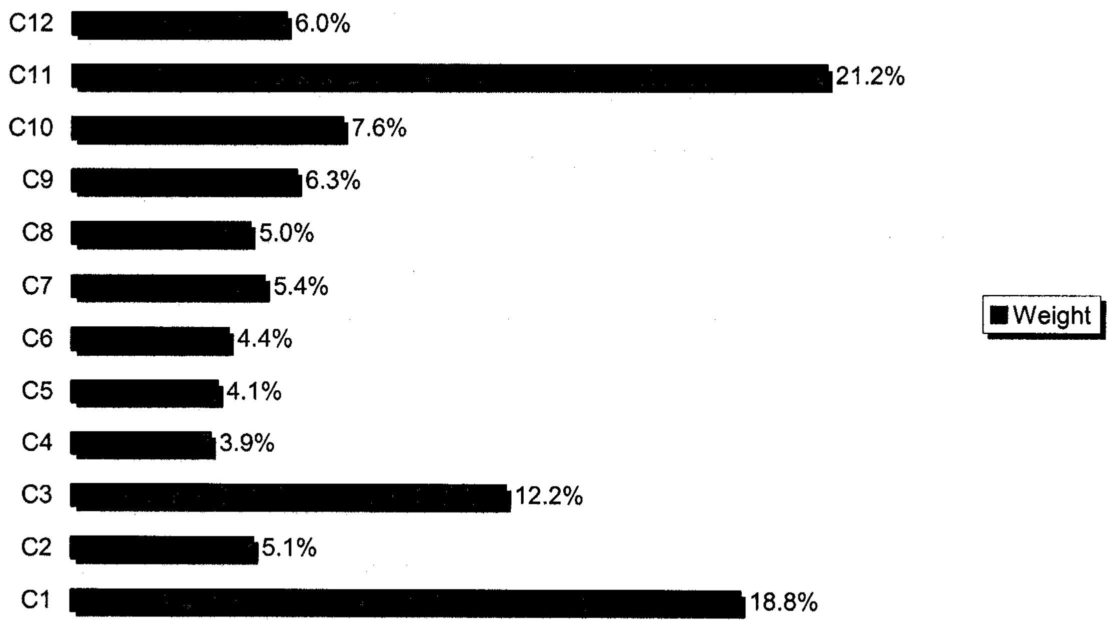


Figure 4.1 Criteria Weights Distribution Bar Chart

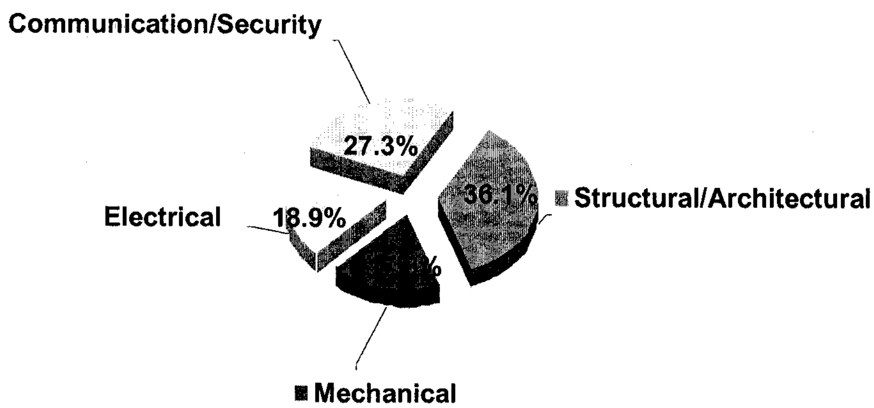


Figure 4.2 Functional Weights Distribution Pie Chart

Figure 4.2 indicates that the Structural/Architectural function is the most important function in the SSDI model, followed by the Communication/Security function. Also the Mechanical and Electrical function have almost the same importance.

Depending on the statistical tests in Table 4.1, the normal distribution is not the best fit for all the criteria, though it is for the criteria that have the highest weights. The statistical figures of the criteria show that the most important criteria C11, C1 and C3 follow a normal distribution density function (Chi-Sq, A-D, and K-S test values are low, while the P-Values are close to 1). The test for the rest of the criteria is acceptable although not as confident as the three above. Figure 4.3 shows a sample normal distribution density curve for C11.

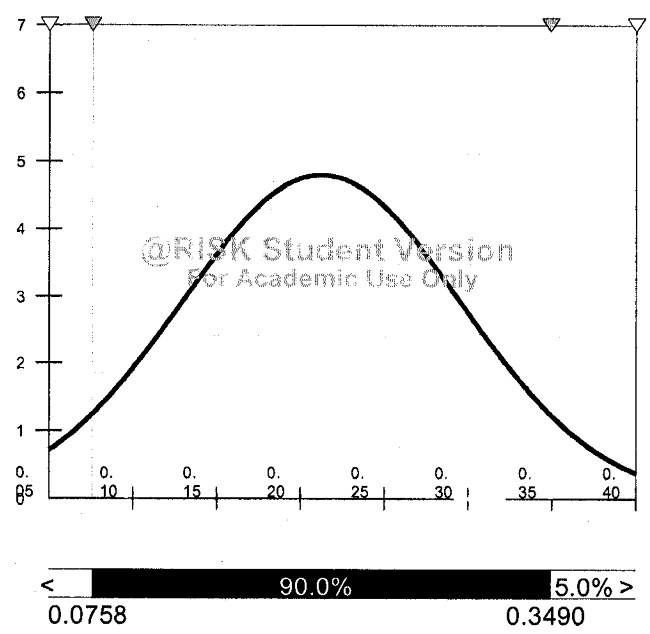


Figure 4.3 Criterion C11 Normal Distribution Density Curve

The standard deviation of the weights for all the criteria is not high. In addition, the standard error is small, and the range for both the 68% and 95% confidence levels is narrow. For example, for a 68% level of confidence, C1 weight (18.8%) may vary between 17.6% and 20%, while for a 95% level of confidence, it varies between 16.4% and 21.2%. This implies a difference in change from the mean by -6.38% and +6.38% for 68% level of confidence, and similarly by -12.7% and +12.7% for 95% level of confidence. If it can be proven by sensitivity analysis that a difference from the mean value of ±12.7% does not affect the Diagnosis Index, the average results from the questionnaires is satisfactory. Figure 4.4 shows the range for each criterion weight, for a 95% level of confidence.

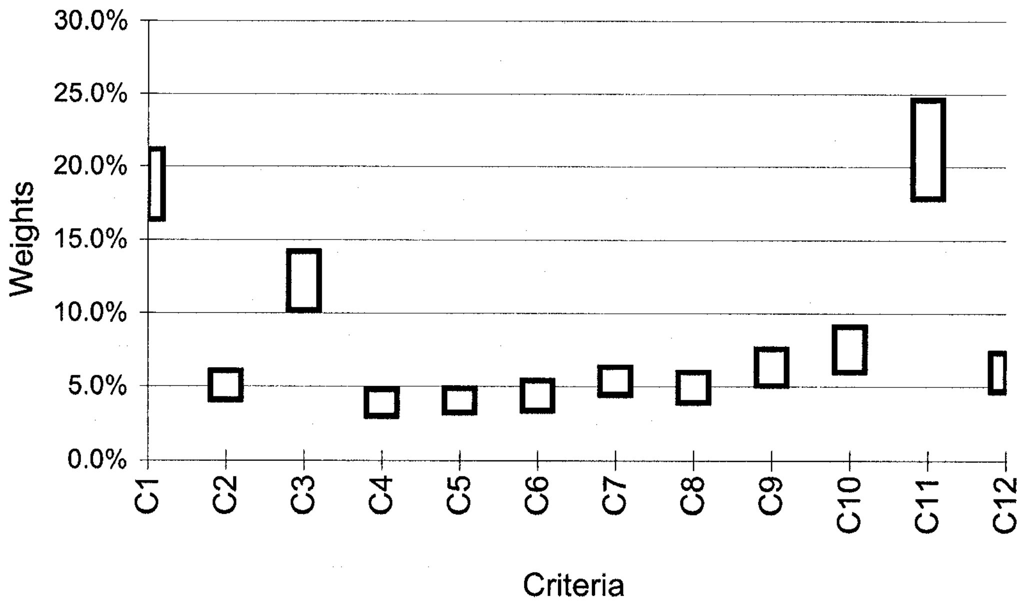


Figure 4.4 Range of criteria weights for 95% level of confidence

It should be noted that the range widens in proportion to the higher weight (important criterion). Figure 4.5 illustrates the range for each criterion weight, for a 68% level of confidence. The 68% level of confidence provides a narrower range of values for each weight, since the reliability is less.

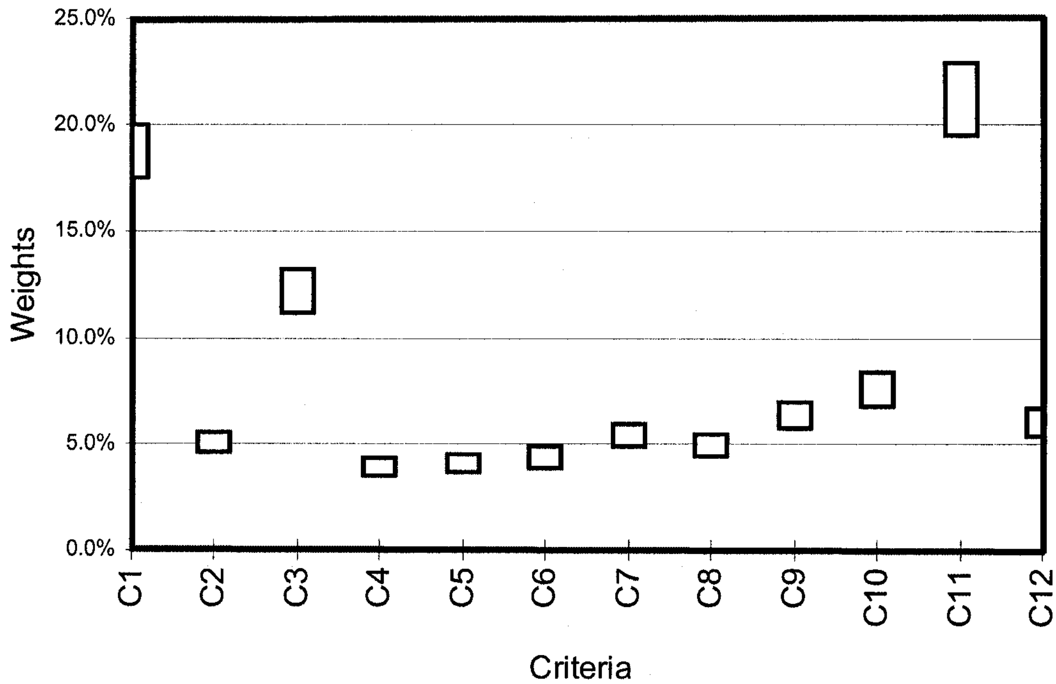


Figure 4.5 Range of criteria weights for 68% level of confidence

The above data, which resulted from the questionnaires, are verified with weights provided by the STM head manager for rehabilitation. The STM manager (head of engineering unit of STM) did not assign direct weights to the criteria, but filled the AHP matrices. His AHP results are shown in Table 4.2, and compared to the results obtained from questionnaires.

The C1 weight was given less importance than the questionnaires, and the difference is distributed to the others. C11 and C3 have almost the same weights, since the difference is very small. More importantly is that the rank of importance is the same, C11, followed by C1 and C3, and the rest follows.

Table 4.2 STM vs. Questionnaires Criteria Weights Verification

Criteria	STM Manager Weights	Questionnaires Weights	Difference
C1	11.4%	18.8%	-7.4%
C2	2.3%	5.1%	-2.8%
C3	11.4%	12.2%	-0.8%
C4	6.4%	3.9%	+2.5%
C5	6.9%	4.1%	+2.8%
C6	4.7%	4.4%	+0.3%
C7	6.9%	5.4%	+1.5%
C8	8.3%	5.0%	+3.3%
C9	9.5%	6.3%	+3.2%
C10	7.2%	7.6%	-0.4%
C11	20.8%	21.2%	-0.4%
C12	4.2%	6.0%	-1.8%

Thus, the criteria weights from the questionnaires can be used with a considerable level of confidence in the SSDI model. Nevertheless, the SSDI model leaves the transit authority (decision maker) the freedom to choose its own criteria weights, and by any preferred method (AHP or other), on the condition that the summation of the weights is unity or 100%.

4.2.3 Pseudo-Criteria Thresholds

The second part of the same questionnaires is used to collect data related to the pseudo-criteria thresholds, (i.e. the Critical Threshold (CT) and the Tolerance Threshold (TT)). The decision maker was asked to indicate which percentages he believes beyond which the conditions of the criteria are considered dangerous (CT), and which percentages below which the conditions of the criteria are considered tolerable (TT). Furthermore, the same statistical analysis is performed on the data collected, for each criterion. Tables 4.3 and 4.4 show the results of the statistical analysis for CT and TT respectively.

Table 4.3 Criteria Critical Thresholds Statistical Analysis Results

Critical Thresholds		C1	C2	C3	C4	C5	C6
Average	μ *(%)*	36	62	40	39	48	45
Standard Deviation	σ *(%)*	13	14	13	13	13	16
Standard Error	ε *(%)*	3.0	3.0	3.0	3.0	3.0	3.0
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ *(%)*	34	59	38	36	46	42
	$\mu+\sigma/n^{\dagger}$ *(%)*	39	65	43	41	51	48
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ *(%)*	31	57	35	33	43	39
	$\mu+2\sigma/n^{\dagger}$ *(%)*	42	68	46	44	54	52
Normal Distribution	Skewness	0.24	0.23	0.37	0.002	0.21	0.04
	Kurtosis	2.09	1.26	1.74	1.51	1.29	1.48
Chi-Sq Test[††]	Test Value	5.6	21.4	1.0	11.42	17.3	8.91
	P-Value	0.23	0.002	0.91	0.022	0.02	0.06
A-D Test[†††]	Test Value	0.59	2.37	1.13	1.03	2.14	1.09
	P-Value	0.15	0.005	0.005	0.1	0.05	0.01
K-S Test[††††]	Test Value	0.18	0.31	0.20	0.22	0.26	0.17
	P-Value	0.05	0.01	0.025	0.01	0.01	0.1

[†] n = Total No. of data, i.e. 24

[††] Chi-Sq statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

[†††] A-D (Anderson-Darling) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

[††††] K-S (Kolmogorov-Smirnov) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

Table 4.3 (Continued) Criteria Critical Thresholds Statistical Analysis Results

Critical Thresholds		C7	C8	C9	C10	C11	C12
Average	μ *(%)*	30	47	40	30	23	55
Standard Deviation	σ *(%)*	11	16	14	14	9	13
Standard Error	ε *(%)*	2.0	3.0	3.0	3.0	3.0	3.0
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ *(%)*	28	43	37	27	21	52
	$\mu+\sigma/n^{\dagger}$ *(%)*	32	50	43	33	24	57
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ *(%)*	25	40	34	24	19	49
	$\mu+2\sigma/n^{\dagger}$ *(%)*	35	53	46	36	26	60
Normal Distribution	Skewness	0.39	0.09	0.003	0.29	0.48	0.03
	Kurtosis	1.36	1.62	1.76	1.55	1.39	1.22
Chi-Sq Test††	Test Value	19.7	1.4	0.58	1.41	18.9	11.4
	P-Value	0.001	0.84	0.96	0.84	0.001	0.02
A-D Test†††	Test Value	2.4	0.76	0.72	1.3	2.8	1.9
	P-Value	0.005	0.05	0.1	0.05	0.005	0.005
K-S Test††††	Test Value	0.27	0.17	0.13	0.22	0.30	0.22
	P-Value	0.01	0.1	0.15	0.01	0.01	0.01

† n = Total No. of data, i.e. 24

†† Chi-Sq statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

††† A-D (Anderson-Darling) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

†††† K-S (Kolmogorov-Smirnov) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

Table 4.4 Criteria Tolerance Thresholds Statistical Analysis Results

Tolerance Thresholds		C1	C2	C3	C4	C5	C6
Average	μ *(%)*	13	31	17	14	16	18
Standard Deviation	σ *(%)*	6	12	9	8	9	10
Standard Error	ε *(%)*	1.0	3.0	2.0	2.0	2.0	2.0
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ *(%)*	12	29	15	13	14	16
	$\mu+\sigma/n^{\dagger}$ *(%)*	15	34	18	16	18	20
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ *(%)*	11	26	13	11	13	14
	$\mu+2\sigma/n^{\dagger}$ *(%)*	16	36	20	17	20	23
Normal Distribution	Skewness	0.71	0.29	0.10	0.28	0.36	0.36
	Kurtosis	2.38	1.83	1.70	1.52	1.72	1.93
Chi-Sq Test††	Test Value	12.2	1.83	3.08	3.9	5.58	7.25
	P-Value	0.01	0.76	0.54	0.42	0.23	0.12
A-D Test†††	Test Value	1.46	0.76	0.79	1.37	1.27	0.91
	P-Value	0.05	0.05	0.05	0.05	0.05	0.02
K-S Test††††	Test Value	0.30	0.16	0.20	0.22	0.25	0.18
	P-Value	0.01	0.1	0.02	0.01	0.01	0.05

† n = Total No. of data, i.e. 24

†† Chi-Sq statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

††† A-D (Anderson-Darling) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

†††† K-S (Kolmogorov-Smirnov) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

Table 4.4 (Continued) Criteria Tolerance Thresholds Statistical Analysis Results

Tolerance Thresholds		C7	C8	C9	C10	C11	C12
Average	μ *(%)*	10	19	17	11	9	25
Standard Deviation	σ *(%)*	6	8	8	6	4	10
Standard Error	ε *(%)*	1.0	3.0	2.0	1.0	1.0	2.0
68% Confidence Level	$\mu-\sigma/n^{\dagger}$ *(%)*	9	17	15	27	21	23
	$\mu+\sigma/n^{\dagger}$ *(%)*	11	20	19	33	24	28
95% Confidence Level	$\mu-2\sigma/n^{\dagger}$ *(%)*	8	15	14	9	19	21
	$\mu+2\sigma/n^{\dagger}$ *(%)*	12	22	20	13	26	30
Normal Distribution	Skewness	0.67	0.29	0.05	0.59	0.46	0.03
	Kurtosis	2.09	1.59	1.71	2.02	1.79	1.88
Chi-Sq Test††	Test Value	17.3	3.08	14.7	12.2	19.7	5.6
	P-Value	0.01	0.54	0.05	0.01	0.01	0.2
A-D Test†††	Test Value	1.7	1.2	1.12	1.7	2.2	0.9
	P-Value	0.01	0.05	0.05	0.05	0.01	0.02
K-S Test††††	Test Value	0.27	0.19	0.23	0.27	0.28	0.20
	P-Value	0.01	0.05	0.01	0.01	0.01	0.02

† n = Total No. of data, i.e. 24
†† Chi-Sq statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.
††† A-D (Anderson-Darling) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.
†††† K-S (Kolmogorov-Smirnov) statistic test (for a normal distribution) = the Test Value should be close to 0, and the P–Value close to 1 to have the most confidence level that the data follow a normal distribution.

From the above statistical tests (Chi-Sq, A-D, and K-S tests), the CT and TT do not fit a normal distribution with a high level of confidence, while the standard error is higher than the one for the weights. Nevertheless, the range for 68% and 95% levels of confidence is not wide. This result can be justified by considering that the assignment of CT and TT for criteria is a difficult, complicated and highly subjective particularly since this is the first research work that has attempted this. The CT and TT are two physical values, and are highly affected by many factors. An obvious factor would be the *Factor of Safety* (FS) in the design of the different functional elements that differs from one code to another (and from one region to another). Another factor would be the different perspective of every practitioner in assigning the thresholds. A manager would have a different perspective from an inspector or an engineer, for example. Figures 4.6 and 4.7 show the normal distribution curve for the C3 Critical Threshold and Tolerance Threshold.

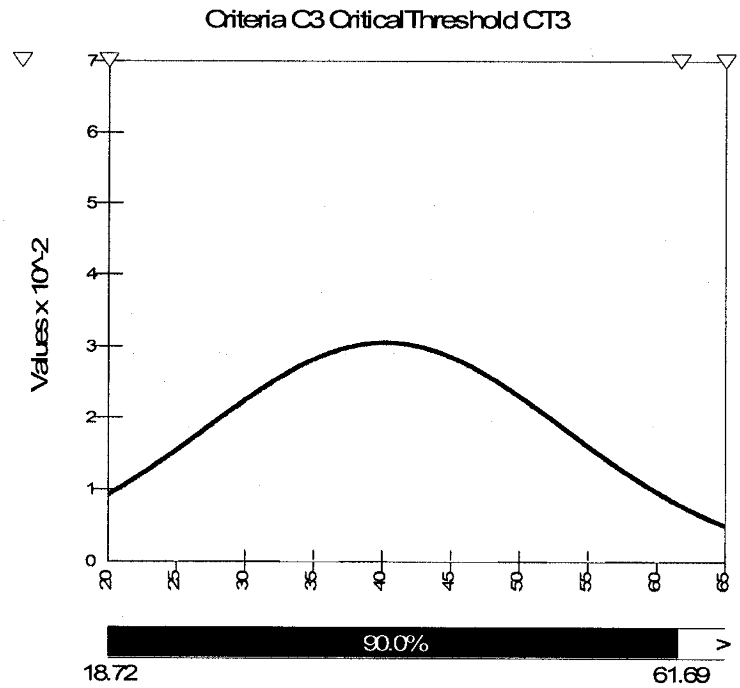


Figure 4.6 Criterion C3 Critical Threshold Normal Distribution Curve

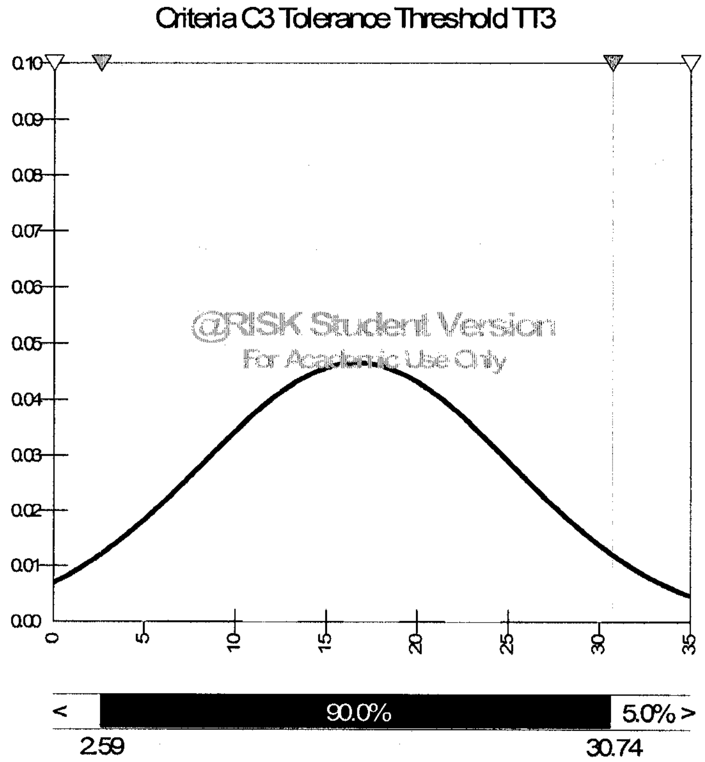


Figure 4.7 Criterion C3 Tolerance Threshold Normal Distribution Curve

The average values of CT and TT may be used, with a certain level of confidence (at least 68%), if it can be proven by the sensitivity analysis, that a difference in the range of 68% level of confidence slightly affects SDI, to an acceptable degree.

The values of CT and TT are directly incorporated in the *generalized preference function* (gpf) of the pseudo-criteria. Figure 4.8 illustrates a gpf for criterion C3 that will be used in the SSDI model. Criterion C3 has a CT=36% and a TT=13%.

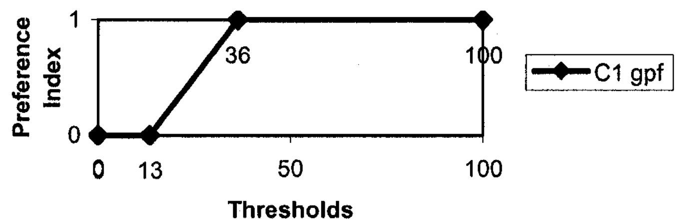


Figure 4.8 Criterion C1 Generalized Preference Function (gpf) Graph

In the case of CT and TT, any input from the same STM manager would not be significant since he has also his own subjective view of the problem.

Next, the weighted average of CT and TT are calculated for the functions (level 1 of the criteria hierarchy), and these values are used in the final SDI and FDI condition table. Table 4.5 shows the weighted average of CT and TT.

Table 4.5 Weighted Functional CT and TT

Functions	Function Weight	Critical Threshold	Tolerance Threshold
Structural / Architectural	36.08%	41%	17%
Mechanical	17.77%	40%	14%
Electrical	18.89%	38%	15%
Communication / Security / Control	27.26%	30%	12%

Critical and tolerance thresholds are two important figures that can be used, in addition to the SSDI model. Thus, an interesting field of research would be to develop a general method in evaluating the critical and tolerance thresholds of the functional elements (or criteria), a field that is still in its infancy.

4.3 Sensitivity Analysis

As already repeated, SSDI considers the numerous uncertainties in the performance of criteria, the pseudo-criteria concept and the weights. The effect of changing the criteria performance, the criteria thresholds, and the criteria weights on the SSDI model will be studied. The data collected show a range of values for 68% and 95% levels of confidence (refer to section 4.2).

4.3.1 Criteria Performances

The sensitivity analysis will consider first of all the effect of the criteria performance on the SSDI model. For this case, six different states of fictitious stations S_1 to S_6 are considered:

S_1 = Extremely Bad Station $\Rightarrow$ $V_i\ [C_i\ (S_1)] = 100\% > CT$

S_2 = Deficient Station $\Rightarrow$ $V_i\ [C_i\ (S_2)] \approx \leq CT$

S_3 = Average Station $\Rightarrow$ $TT < V_i\ [C_i\ (S_3)] < CT$

S_4 = Good Station $\Rightarrow$ $V_i\ [C_i\ (S_4)] \approx \geq TT$

S_5 = Extremely Good Station $\Rightarrow$ $V_i\ [C_i\ (S_5)] < TT$

S_6 = Special Station $\Rightarrow$ $V_{Structure}\ [C_{Structure}\ (S_6)] > CT$,

$V_{Mechanical}\ [C_{Mechanical}\ (S_6)] < TT$,

$V_{Electrical}\ [C_{Electrical}\ (S_6)] < TT$,

And $V_{Communication}\ [C_{Communication}\ (S_6)] < TT$

The exact values of CT and TT are taken from the previous section of this chapter. They can be even arbitrary, but in this case they have to be the same for all stations. Table 4.6 indicates the results of SSDI for these different stations.

Table 4.6 Criteria Performance Sensitivity Analysis Results

	$FDI_{Structure}$	$FDI_{Mechanical}$	$FDI_{Electrical}$	$FDI_{Communication}$	SDI_{Global}
S_1	0	0	0	0	0
S_2	2.5	2.5	2.5	2.5	2.5
S_3	5	5	5	5	5
S_4	7.5	7.5	7.5	7.5	7.5
S_5	10	10	10	10	10
S_6	0.5	10	10	10	1.5

It is obvious from the above table that if the criteria performances are equal to or just slightly larger than TT, the FDI would be approximately 2.5, if rounded, it is 3.0. This appears on the scale (0-10) of FDI and SDI in the previous chapter (see Tables 3.3 and 3.4). If the criteria performances are equal to CT (or just less than) CT, then the FDI would be around 7.5, if rounded it is 8.0. This appears also in the scale (0-10) of FDI and SDI in the previous chapter (Tables 3.3 and 3.4). These above-mentioned values justify the use of ¼(CT-TT), and ¾(CT-TT) in the deterioration level column, because they correspond to 2.5/10 = ¼, and 7.5/10 = ¾ of the thresholds. Furthermore, S_1 and S_5 confirm the use of the two datum stations S_0 and S_{100}.

The special case of S_6 describes how the SSDI takes a deteriorated function into account, even if the function's condition seems tolerable. Figure 4.9 illustrates the graph the relation of the criteria performance to the Functional Diagnosis Index FDI.

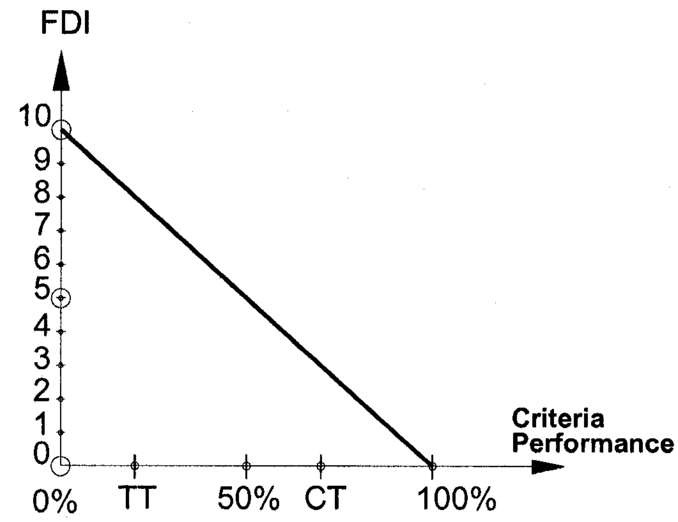


Figure 4.9 FDI versus Criteria Performances Graph

4.3.2 Pseudo-Criteria Thresholds

The statistical analysis has proven that both CT and TT are not definite values for each criterion but could vary in a range, depending on the level of confidence or reliability. For this reason, a sensitive analysis to study the effect of both CT and TT on SDI is necessary. This sensitivity analysis considers the change of CT alone, keeping TT as the average value; and then considers the change of TT alone, keeping CT as the average value. The SSDI model was done on Excel sheet for a specific station that has an arbitrary criteria performance. Two cases were studied, the first one using the weights of the data collected from the questionnaires, and the second one using equal weights for all the criteria. For the two cases, the critical and tolerance thresholds were changed in a range of –40% to +40% of the original value used and the results were tabulated. Then, the

'Sensitivity Analysis' option in @Risk software was performed, in the excel sheets, changing each time the CT and TT and recording SDI. The study was done on SDI only, and not on the different FDI of the model. Figure 4.10 illustrates the SDI sensitivity to CT with variable weights. Figure 4.11 shows the SDI sensitivity to CT with equal criteria weights.

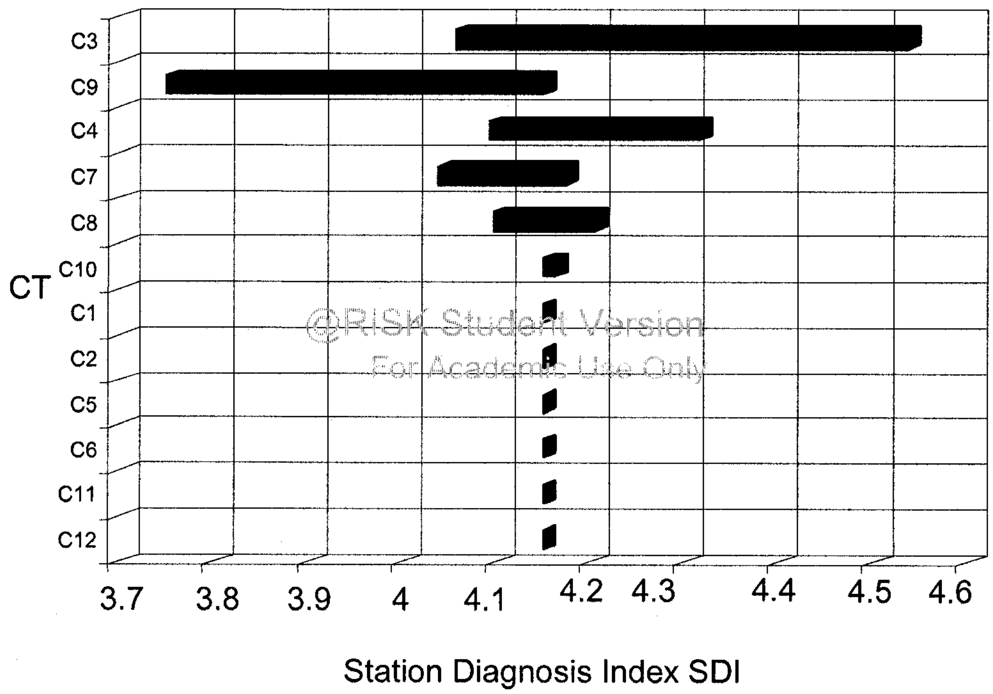


Figure 4.10 Sensitivity of SDI to CT (Variable Weights)

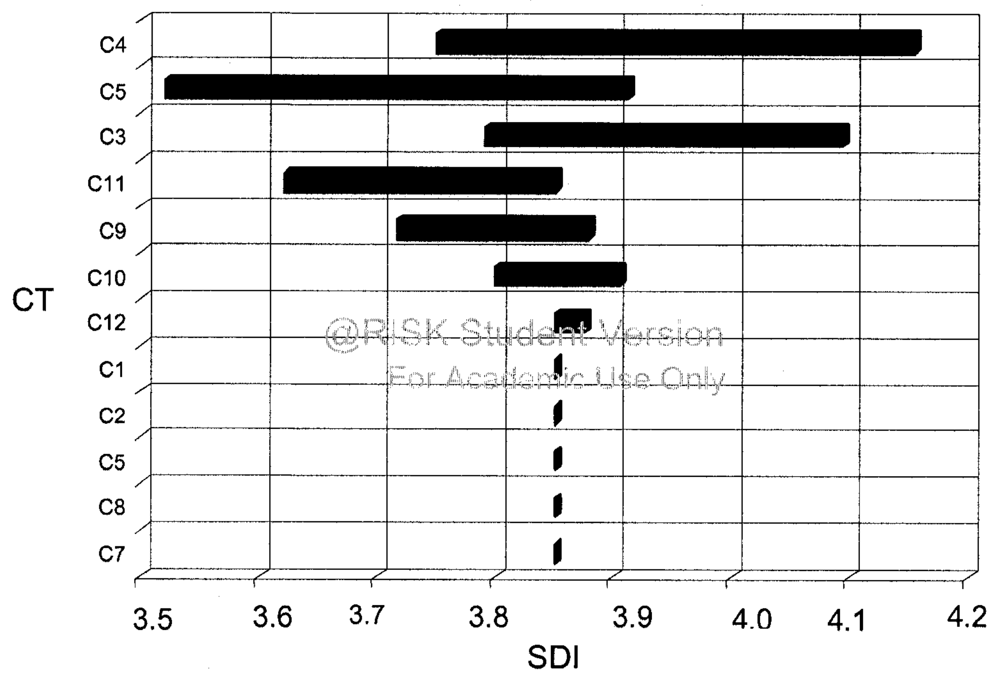


Figure 4.11 Sensitivity of SDI to CT (Equal Weights)

The SSDI model output for the previous cases has given that the SDI, for variable criteria weights, is 4.15, while for equal criteria weights, it is 3.85. From Figures 4.10 and 4.11, if CT changes from -40% to +40% of its original value, the SDI changes from 3.75 to 4.5 for variable criteria weights. And similarly, the SDI changes from 3.5 to 4.1 for equal criteria weights. Thus the percent change of SDI is around ±10% for variable criteria weights, and ±6% for equal criteria weights. Hence, for a 95% level of confidence, where the CT changes in the range of ±20% for C3 for example, SDI changes to a lesser degree (for variable weights, see Figure 4.10), which is around ±3%. Furthermore, the same sensitivity analysis is performed for the same station, changing this time the value of TT from -40% to +40%. Since TT has a lower percent value from CT, the case of equal criteria weights is only considered for simplicity. Figure 4.12 shows the sensitivity of TT to SDI, for equal weights.

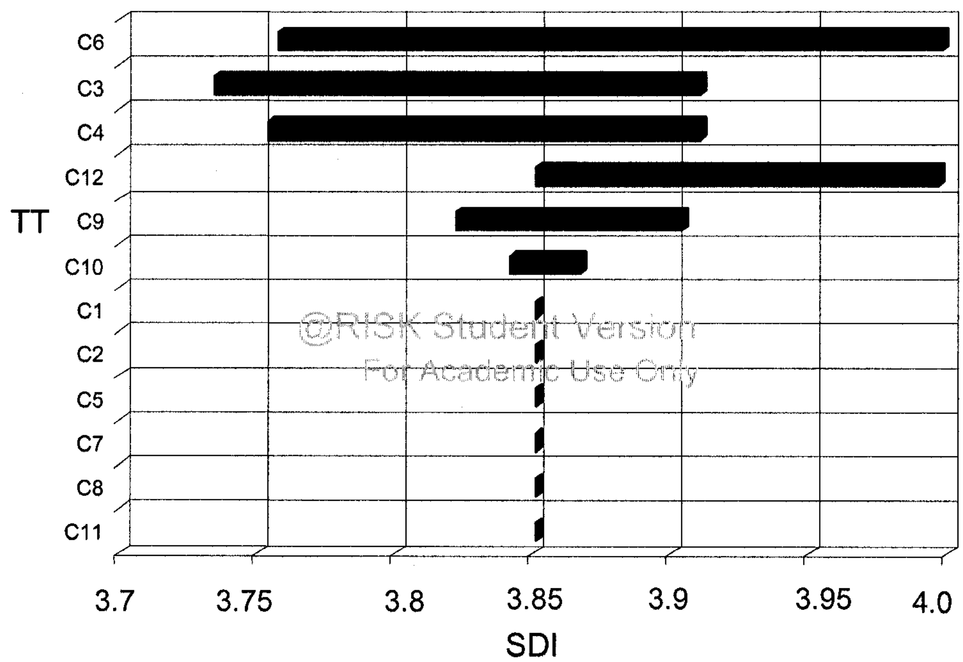


Figure 4.12 Sensitivity of SDI to TT (Equal Weights)

From the above figure, the percent change for TT versus SDI is much less than the one of CT, which is around ±5%. Thus, SDI is less sensitive on TT than CT. Finally, the use of the values of CT and TT from the questionnaires is justified, since their sensitivity analysis has proved that their change for a 95% level of confidence affect the SDI with a small degree.

4.3.3 Criteria Weights

The criteria weights are of most importance to any decision-making method especially to the SSDI model. It has shown that the SSDI model can use any of the methods discussed in the literature review chapter to evaluate the weights of its criteria. Nevertheless, the SSDI model suggested AHP as one of the most appropriate methods. The statistical analysis of the weights, evaluated from the questionnaires, has proved that the weights are not a 'fixed' value, but can vary depending on the level of confidence, because AHP is a very subjective method, however scientific. Thus, a sensitivity analysis is very crucial to the overall analysis of the SSDI model. The goal of the sensitivity analysis is to study the effect of the change of weights to SDI. The same fictitious station, studied for the thresholds sensitivity analysis (SDI=4.15, and with variable weights), is used for the weights sensitivity analysis as well. Furthermore, another fictitious station, that is an average station (all the criteria performances are 50%, thus SDI=5.0), is also analyzed. Below, Figures 4.13 and 4.14 illustrate the sensitivity graphs of SDI to the criteria weights, for the two cases listed above.

The sensitivity analysis for both stations has shown that the SDI is more sensitive to the criteria weights than the thresholds. From Figure 4.14, the most important criteria C11, C1 and C3 have the most effect on SDI. For example, a change of ±40% for C11 weight changes SDI from 4.3 to 5.7, i.e. a change of ±14% (Figures 4.13 and 4.14). Actually, the range of the weights of these criteria (C1 for example) varies ±20% (see Table 4.1), and the SDI changes in the range of ±8% (Figures 4.13 and 4.14 for C1). This result is acceptable.

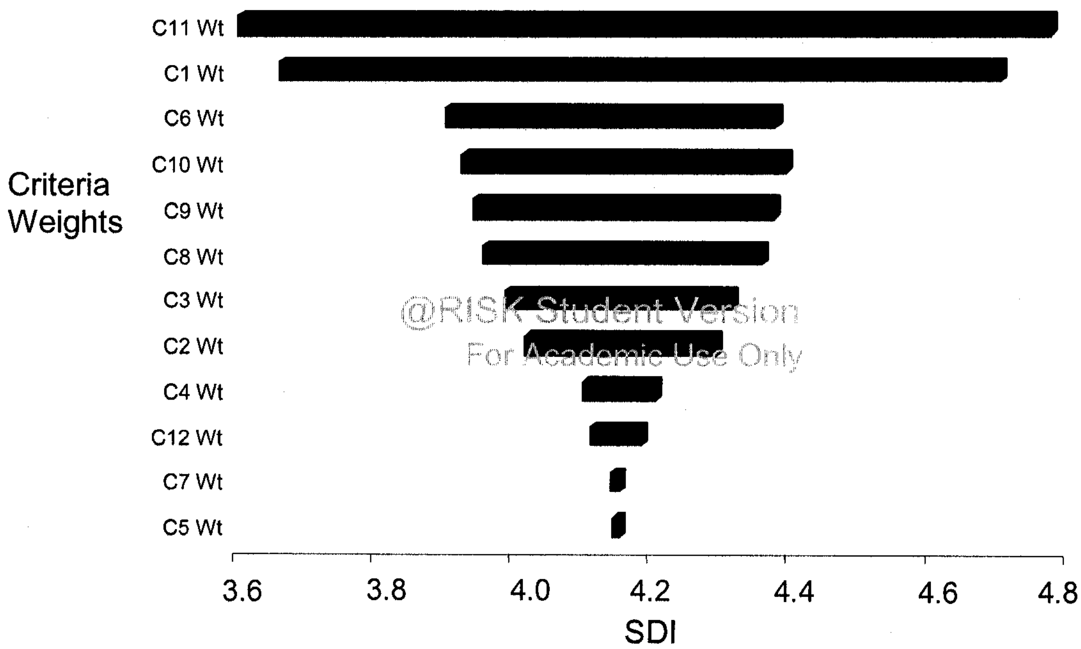


Figure 4.13 Sensitivity of SDI to Criteria Weights

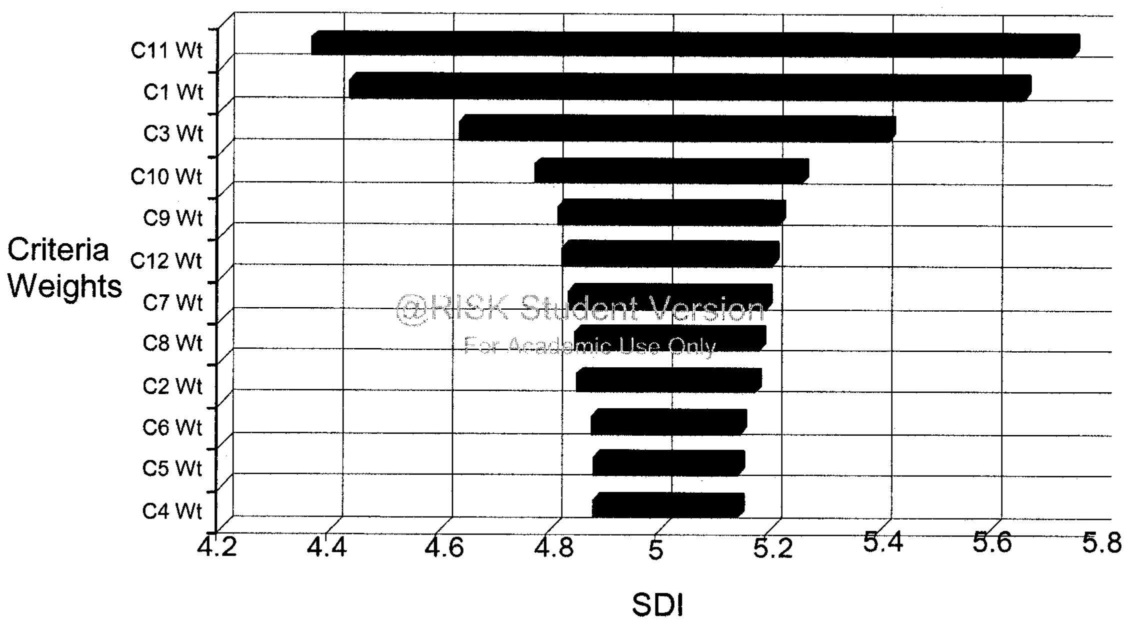


Figure 4.14 Sensitivity of SDI to Criteria Weights (Average Station)

Thus, the use of the average values of criteria weights is justified with a 95% confidence level. It can be concluded that the weights, either from the questionnaires or from the STM manager, can be used in the SSDI model.

4.4 Data Incorporation in SSDI model

After collecting the data, and performing statistical and sensitivity analysis, the SSDI condition table must be revised and completed. The average weights and thresholds are incorporated in the FDI and SDI scale. The final SDI condition scale is shown in Table 4.7. The FDI condition scale can be modified accordingly.

Table 4.7 SDI Final Condition Scale Table

SDI	Description	Deterioration Level *(%)*	Proposed Action
8 < SDI ≤ 10	*Good*	<17% Structural or, <12% Communications or, <15% Electrical or, <14% Mechanical	Long Term: * Expertise < 2 years * Physical < 5 years Review in 2 years
6 < SDI ≤ 8	*Medium*	>17% & <23% Structural or, >12% & <17% Communications or, >15% & <21% Electrical or, >14% & <21% Mechanical	Medium Term: * Expertise < 1 year * Physical < 2 years Review in 1 year
3 < SDI ≤ 6	*Deficient*	>23% & <35% Structural or, >17% & <26% Communications or, >21% & <33% Electrical or, >21% & <34% Mechanical	Short Term: * Expertise < 6 months * Physical < 1 year Review in 6 months
0 ≤ SDI ≤ 3	*Critical*	>41% Structural or, >30% Communications or, >38% Electrical or, >40% Mechanical	Immediate: Physical intervention Now

5 SSDI Application to STM Stations

5.1 STM Stations

The SSDI model described in detail in Chapter 3 needs to be applied to real life stations in order to prove its functionality. The STM has been contacted to get inspection reports. They have provided inspection reports for 7 out of the 24 stations that are part of the 'Réno-Station' program (The oldest stations of the STM network). The twenty-four that comprise the Montreal metro system are shown below:

1.	Atwater	2.	Frontenac	3.	Rosemont
4.	Guy-Concordia	5.	Bonaventure	6.	Beaubien
7.	Peel	8.	Square-Victoria	9.	Jarry
10.	McGill	11.	Place d'Armes	12.	Cremazie
13.	Place des Arts	14.	Champs-de-Mars	15.	Sauve
16.	St Laurent	17.	Sherbrooke	18.	Henri-Bourassa
19.	Beaudry	20.	Mont-Royal	21.	Jean-Drapeau
22.	Papineau	23.	Laurier	24.	Longeuil

For reasons of confidentiality within the authority, the STM did not allow for the identification of the seven stations, and hence arbitrary names S1 to S7 were assigned to these real stations. It should be noted that only the names are fictitious, but the values used for the criteria performances are taken from the inspection reports.

5.2 Criteria Definitions and Performances

The first step in the SSDI model is to define the criteria and establish their hierarchy. In the application, the criteria already defined in the model illustration (Chapter 3) are used. Moreover, the criteria definitions are confirmed by the inspection reports.

5.2.1 Criteria Definitions and Hierarchy

The inspection reports gave assessments of structural, architectural, mechanical and electrical conditions. The remaining criteria (communication and security) are taken from the 'Reno-Systèmes' report. Table 5.1 summarizes the criteria used for the SSDI application.

Table 5.1 Criteria Definition

I.D	Description	Function
C1	Global Structure	Structural Architectural
C2	Global Architecture	
C3	Concrete (Fixed) Stairs	
C4	Mechanical Stairs	Mechanical
C5	Pipes and Mech. Equipments	
C6	Ventilation, A/C, Heat	
C7	Fire Stand Pipes	
C8	Lighting	Electrical
C9	Cables	
C10	Panels / Transformers / Breakers	
C11	Alarm / Security / Smoke Detectors	Communication Security
C12	Sign Boards / Public Address Communication System (Telemetry)	

5.2.2 Criteria Performances

The values related to criteria performances are taken directly from the inspection reports, except for the '*mechanical stairs'* sub-criterion, and the communication/security criteria. The latter were taken from the 'Réno-Système' program report. This report specifies that, for all STM stations, mechanical stairs, security system, control system and communication system are labeled as in '*Bad*' condition, and thus require '*short-term*' intervention. The values of performance criteria for C4, C11 and C12 are assigned the scale of 2 as described in the 'Réno-Système' report. The inspection reports assess the condition of different elements by a field inspection scale already described in Chapter 3 (Table 3.1), and then assign a global scale for some functions such as architectural or structural only. This global structural assessment scale is based on the lowest scale assigned to any element and not as an average value. Table 5.2 shows the criteria performance values for the seven stations.

Table 5.2 SSDI Application to STM Criteria Performances Input

Criteria	S1	S2	S3	S4	S5	S6	S7
C1†	1	3	3	3	3	3	2
C2	4	4	3	4	3	4	4
C3†	3	4	2	1	1	1	4
C4 *	2	2	2	2	2	2	2
C5	3	2	1	2	4	2	1
C6	1	3	4	1	3	3	3
C7	4	3	1	3	4	2	1
C8	5	5	5	5	5	5	5
C9	4	4	4	4	4	4	4
C10	4	4	4	4	4	4	4
C11† *	2	2	2	2	2	2	2
C12 *	2	2	2	2	2	2	2

* Based on 'Réno-Système' report

† Most Important Criteria

The performance values are then transformed into criteria values as per Equation 3.2. The criteria values are then measured on a scale from 0 to 100%.

5.3 Criteria Weights

The hierarchical structure of the criteria allows the SSDI model to use AHP for the weights evaluation. The weights used in the application are the average values of the data collected from the 24 questionnaires. In reference to Chapter 4, AHP matrices in the questionnaires were filled in and the results provided average values of the criteria weights. It has also been confirmed, through statistical and sensitivity analysis, the average values can be used with 95% reliability. Table 5.3 summarizes the criteria weights values.

Table 5.3 SSDI Application Criteria Weights

I.D.	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}
C1	52.6%	19.0%	36.1%
C2	13.9%	5.0%	
C3	33.5%	12.1%	
C4	22.5%	4.0%	17.8%
C5	23.0%	4.1%	
C6	23.6%	4.2%	
C7	30.9%	5.5%	
C8	27.5%	5.2%	18.9%
C9	32.8%	6.2%	
C10	39.7%	7.5%	
C11	78.4%	21.4%	27.3%
C12	21.6%	5.9%	

The relation between the '*local*' criteria, the global and the functional weights are defined in Equations 3.3 to 3.5 shown in Chapter 3. The global criteria weights are the values taken from the questionnaires. The functional weight is the sum of the global criteria weights for each function. While the '*local*' criterion weight is the global sub-criterion weight divided by the functional weight. For example, the average global weight of C1 from the questionnaires is 19.0%, thus the functional weight for Structure/Architecture would be equal to 19.0%+5%+12.1% = 36.1%. Therefore the '*local*' C1 weight would be 19.0%/36.1% = 52.6%.

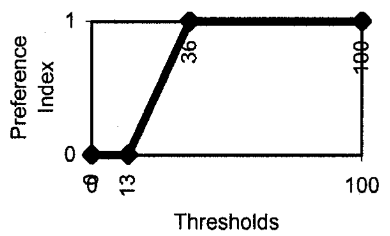
5.4 Pseudo-Criteria Thresholds

Similar to the criteria weights, the thresholds used in the application are also the average values taken from the questionnaires. The statistical and sensitivity analyses performed in Chapter 4 confirm their usefulness and applicability with 95% reliability. Table 5.4 summarizes the threshold values.

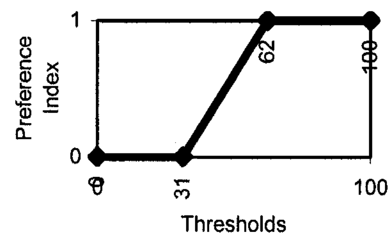
Table 5.4 SSDI Application Thresholds

I.D	Description	CT (%)	TT (%)
C1	Global Structure	36	13
C2	Global Architecture	62	31
C3	Concrete (Fixed) Stairs	40	17
C4	Mechanical Stairs	39	14
C5	Pipes and Equipments	48	16
C6	Ventilation, A/C, Heat	45	18
C7	Fire Stand Pipes	30	10
C8	Lighting	47	19
C9	Cables	40	17
C10	Panels / Transformers / etc..	30	11
C11	Security	23	9
C12	Communication	55	25

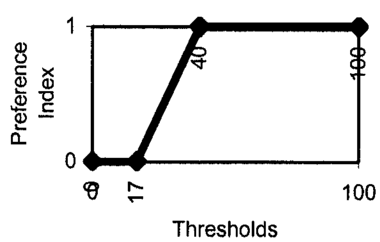
The thresholds are used to model the generalized preference function (gpf). Figure 5.1 defines the gpf for the twelve criteria.



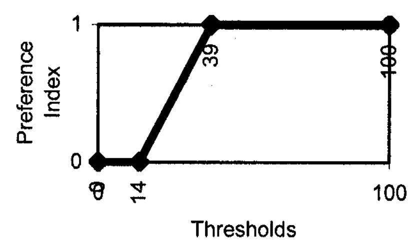
A - C1 gpf



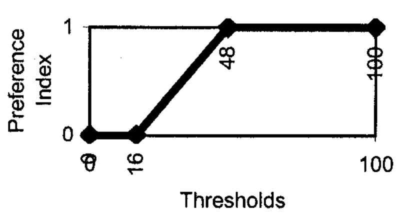
B - C2 gpf



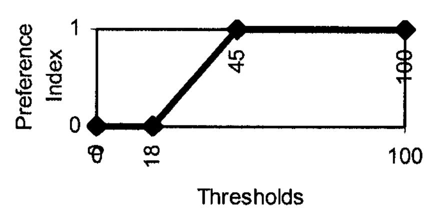
C - C3 gpf



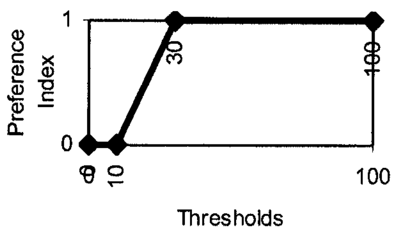
D - C4 gpf



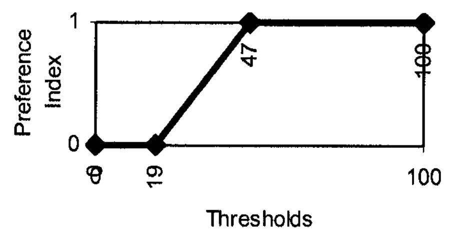
E - C5 gpf



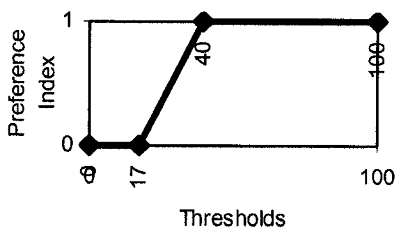
F - C6 gpf



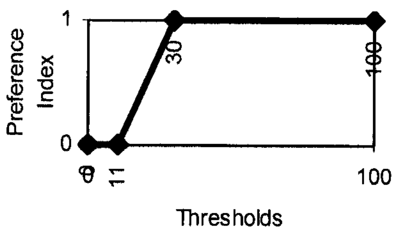
G - C7 gpf



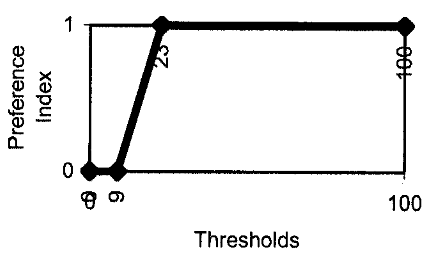
H - C8 gpf



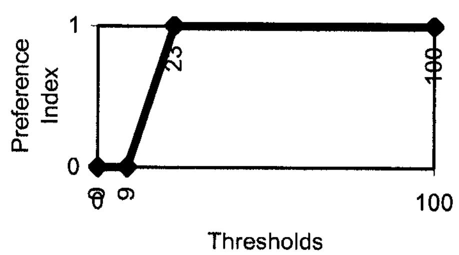
I - C9 gpf



J - C10 gpf



K - C11 gpf



L - C12 gpf

Figure 5.1 SSDI Application Criteria gpf

The criteria gpf will be used to evaluate the criteria preference index as per Equation 3.9. The criteria gpf transforms the criteria into pseudo-criteria. The gpf slope indicates the risk degree in the decision.

5.5 Multi-Criteria Aggregation

The next step in the SSDI model application is criteria aggregation. The criteria performances will be compared to the performances of the boundary stations S_0 and S_{100}. This step consists of evaluating the difference between the criteria values ($\Delta V[C]$) and S_0 and S_{100}, which are 0 and 100 respectively (refer to Equations 3.6 and 3.7). Table 5.5 shows the criteria aggregation for S1 as per Equation 3.10. The remaining stations (S2 to S7) have a similar aggregation. Furthermore, both the criteria aggregation and the gpf are used to evaluate the preference index. The pseudo-criteria preference index (P[V]) is a value between 0 and 1. Table 5.6 shows the pseudo-criteria preference index for S1 only. For stations S2 to S7, the same calculation is also performed.

Table 5.5 SSDI Application Criteria Aggregation for S1

I.D	S_0-S_0	S_0-S_1	S_0-S_{100}	S_1-S_0	S_1-S_1	S_1-S_{100}	S_{100}-S_0	S_{100}-S_1	S_{100}-S_{100}
C1	0	-100	-100	100	0	0	100	0	0
C2	0	-25	-100	25	0	-75	100	75	0
C3	0	-50	-100	50	0	-50	100	50	0
C4	0	-75	-100	75	0	-25	100	25	0
C5	0	-50	-100	50	0	-50	100	50	0
C6	0	-100	-100	100	0	0	100	0	0
C7	0	-25	-100	25	0	-75	100	75	0
C8	0	0	-100	0	0	-100	100	100	0
C9	0	-25	-100	25	0	-75	100	75	0
C10	0	-25	-100	25	0	-75	100	75	0
C11	0	-75	-100	75	0	-25	100	25	0
C12	0	-75	-100	75	0	-25	100	25	0

Table 5.6 SSDI Application Pseudo-Criteria Preference Index for S1

I.D	S_0-S_0	S_0-S_1	S_0-S_{100}	S_1-S_0	S_1-S_1	S_1-S_{100}	S_{100}-S_0	S_{100}-S_1	S_{100}-S_{100}
C1	0	0	0	1.00	0	0	1	0	0
C2	0	0	0	0.00	0	0	1	1	0
C3	0	0	0	1.00	0	0	1	1	0
C4	0	0	0	1.00	0	0	1	0.44	0
C5	0	0	0	1.00	0	0	1	1	0
C6	0	0	0	1.00	0	0	1	0	0
C7	0	0	0	0.75	0	0	1	1	0
C8	0	0	0	0.00	0	0	1	1	0
C9	0	0	0	0.35	0	0	1	1	0
C10	0	0	0	0.74	0	0	1	1	0
C11	0	0	0	1.00	0	0	1	1	0
C12	0	0	0	1.00	0	0	1	0	0

Based on Table 5.5 and 5.6, it is clear that the series of Equations 3.10 are confirmed. The aggregation is actually for S_k-S_0 and S_{100}-S_k (where S_k is any station). The rest are all zero, except for S_{100}-S_0, which is always equal to 1.

5.6 Multi-Criteria Preference Index

The next step in the SSDI model is to evaluate the multi-criteria preference index (Π), based on the preference index. The multi-criteria preference index (Π) is a weighted average of the criteria preference index (P[V]) and is defined in Equations 3.11 and 3.12.

Equation 3.12 shows that the multi-criteria preference index are all zero, with the exception of $\Pi[S_k$-$S_0]$, of $\Pi[S_{100}$-$S_k]$, and $\Pi[S_{100}$-$S_0]$. The above observation is obvious in Table 5.7. Table 5.7 shows the multi-criteria preference index calculation for station S1. The same calculation is performed for stations S2 to S7.

Table 5.7 SSDI Application Multi-Criteria Preference Index for S1

I.D	S_0-S_0	S_0-S_1	S_0-S_{100}	S_1-S_0	S_1-S_1	S_1-S_{100}	S_{100}-S_0	S_{100}-S_1	S_{100}-S_{100}
C1	0	0	0	0.19	0	0	0.19	0.00	0
C2	0	0	0	0.00	0	0	0.05	0.05	0
C3	0	0	0	0.12	0	0	0.12	0.12	0
C4	0	0	0	0.04	0	0	0.04	0.02	0
C5	0	0	0	0.04	0	0	0.04	0.04	0
C6	0	0	0	0.04	0	0	0.04	0.00	0
C7	0	0	0	0.04	0	0	0.06	0.06	0
C8	0	0	0	0.00	0	0	0.05	0.05	0
C9	0	0	0	0.02	0	0	0.06	0.06	0
C10	0	0	0	0.06	0	0	0.08	0.08	0
C11	0	0	0	0.21	0	0	0.21	0.21	0
C12	0	0	0	0.06	0	0	0.06	0.00	0

5.7 Functional Diagnosis Index

The multi-criteria functional preference index leads to the evaluation of the Functional Diagnosis Index (FDI), which is the first result of the SSDI model. However, it is not a single step, but requires an evaluating of the flows (Φ^+, Φ^-, Φ^{net}), as per Equations 3.17 to 3.22. The next step is to calculate the FDI as per Equation 3.27. These two steps are done for the station itself, in addition to S_0 and S_{100} (Refer to Equations 3.23, 3.24, 3.28 and 3.29).

5.7.1 Flow Evaluation

The first step in evaluating the Functional Diagnosis Index (FDI) is the calculation of the flows. The flows indicate the strength and weakness of the multi-criteria of the station. The flows of the seven stations are evaluated, in addition to S_0 and S_{100}. The latter are the same for all the stations. Table 5.8 shows the flow calculation for S1, as an example.

Table 5.8 SSDI Application Flow Evaluation for S1

I.D	Φ^+_{S1}	Φ^-_{S1}	Φ^{net}_{S1}	Φ^+_{S0}	Φ^-_{S0}	Φ^{net}_{S0}	Φ^+_{S100}	Φ^-_{S100}	Φ^{net}_{S100}
C1, C2, C3	0.86	0.47	0.39	0.0	1.0	-1.0	1.0	0.0	1.0
C4, C5, C6, C7	0.92	0.64	0.28	0.0	1.0	-1.0	1.0	0.0	1.0
C8, C9, C10	0.41	1.00	-0.59	0.0	1.0	-1.0	1.0	0.0	1.0
C11, C12	1.00	0.78	0.22	0.0	1.0	-1.0	1.0	0.0	1.0

Table 5.8 confirms the basis of the flow calculation for S_0 and S_{100}. The net flow of S_0 is always equal to –1.0, while the one of S_{100} is always equal to +1.0. Furthermore, the net flow for any station falls always between these two boundary values, and this is true regardless of the function.

5.7.2 FDI Evaluation

The functional Diagnosis Index (FDI) is a simple application of Equation 3.27. This equation linearly transforms the net flows (-1 to 1) into the FDI index scale (0 to 10). Table 5.9 shows the FDI calculation for station S1. The FDI of S_0 and S_{100} are per Equations 3.28 and 3.29.

Table 5.9 SSDI Application FDI Evaluation for S1

<table>
<tr><th>I.D</th><th>Φ^{net}_{S1}</th><th>FDI_{S1}</th><th>Φ^{net}_{S0}</th><th>FDI_{S0}</th><th>Φ^{net}_{S100}</th><th>FDI_{S100}</th></tr>
<tr><td>C1</td><td rowspan="3">0.39</td><td rowspan="3">3.1</td><td rowspan="3">-1.0</td><td rowspan="3">10</td><td rowspan="3">1.0</td><td rowspan="3">0.0</td></tr>
<tr><td>C2</td></tr>
<tr><td>C3</td></tr>
<tr><td>C4</td><td rowspan="4">0.28</td><td rowspan="4">3.6</td><td rowspan="4">-1.0</td><td rowspan="4">10</td><td rowspan="4">1.0</td><td rowspan="4">0.0</td></tr>
<tr><td>C5</td></tr>
<tr><td>C6</td></tr>
<tr><td>C7</td></tr>
<tr><td>C8</td><td rowspan="3">-0.59</td><td rowspan="3">8.0</td><td rowspan="3">-1.0</td><td rowspan="3">10</td><td rowspan="3">1.0</td><td rowspan="3">0.0</td></tr>
<tr><td>C9</td></tr>
<tr><td>C10</td></tr>
<tr><td>C11</td><td rowspan="2">0.22</td><td rowspan="2">3.9</td><td rowspan="2">-1.0</td><td rowspan="2">10</td><td rowspan="2">1.0</td><td rowspan="2">0.0</td></tr>
<tr><td>C12</td></tr>
</table>

Table 5.9 clarifies the above-mentioned equations. The FDI for S_0 is always 10, while the FDI for S_{100} is always 0.0, regardless of the function. Figure 5.2 illustrates the FDI values for station S1.

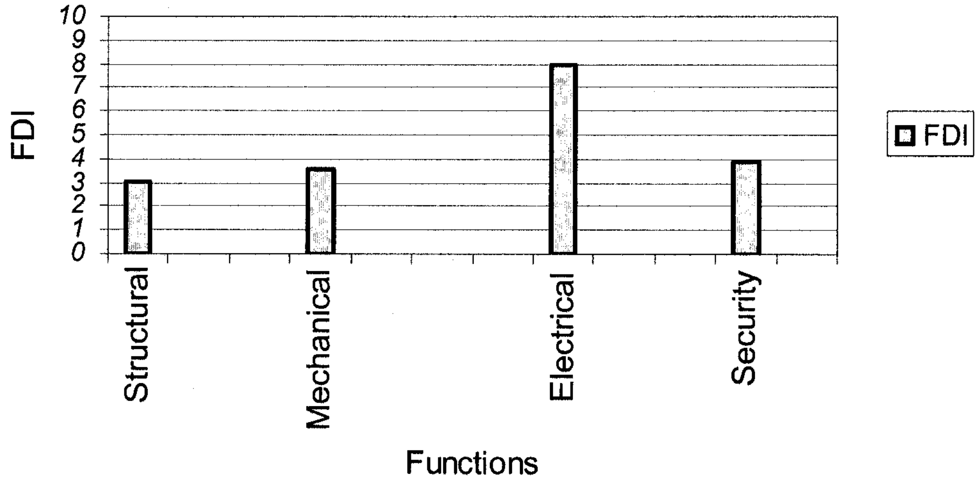


Figure 5.2 SSDI Application FDI for S1

The same calculation is performed for the seven stations (Refer to Appendix B). Tables 5.10 and 5.11 illustrate the FDI output for the seven stations, with a respective description and deterioration level in accordance with FDI condition scale (Refer to Tables 3.3 and 4.7).

Table 5.10 SSDI Application FDI Output (1 of 2)

Station	$FDI_{Structure}$	Description	Deterioration Level	$FDI_{Mechanical}$	Description	Deterioration Level
S1	3.1	Deficient	>23%&<35%	3.6	Deficient	>21%&<34%
S2	6.8	Medium	>17%&<23%	3.5	Deficient	>21%&<34%
S3	3.9	Deficient	>23%&<35%	2.5	Critical	>40%
S4	4.0	Deficient	>23%&<35%	2.4	Critical	>40%
S5	3.3	Deficient	>23%&<35%	5.6	Deficient	>21%&<34%
S6	4.0	Deficient	>23%&<35%	3.2	Deficient	>21%&<34%
S7	5.5	Deficient	>23%&<35%	1.7	Critical	>40%

Table 5.11 SSDI Application FDI Output (2 of 2)

Station	$FDI_{Electrical}$	Description	Deterioration Level	$FDI_{Communcication}$	Description	Deterioration Level
S1	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S2	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S3	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S4	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S5	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S6	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%
S7	8.0	Medium	>15%&<21%	3.9	Deficient	>17%&<26%

The FDI for the structural condition for all stations were labelled as 'Deficient' except for S2. The FDI for the mechanical function show 'Critical' and 'Deficient' conditions and the FDI for the electrical function reveals that the condition is 'Medium' for all stations. Regarding communications and security, the FDI was determined to be 'Deficient' for all seven stations. It should be noted that a global rank will be ambiguous if the decision is based solely on the FDI. Each function has to be studied separately. The FDI results leave the STM the freedom to base their MR planning, budget allocation, and funding planning on the function they need.

5.8 Station Diagnosis Index

The final result of the SSDI model is the global Station Diagnosis Index (SDI). This step requires the application of Equations 3.32 and 3.33. Once the FDI are evaluated, the calculations remaining for this step are minor. Table 5.13 shows the SDI for the seven stations, with the appropriate description (Refer to Table

4.7). The corresponding deterioration level is a recapitulation of what was showed in Tables 5.10 and 5.11.

Table 5.12 SSDI Application SDI Output

Station	SDI_{Global}	Description
S1	4.0	Deficient
S2	5.4	Deficient
S3	4.2	Deficient
S4	4.1	Deficient
S5	4.5	Deficient
S6	4.4	Deficient
S7	4.4	Deficient

Figure 5.3 illustrates the FDI and SDI values for the seven stations.

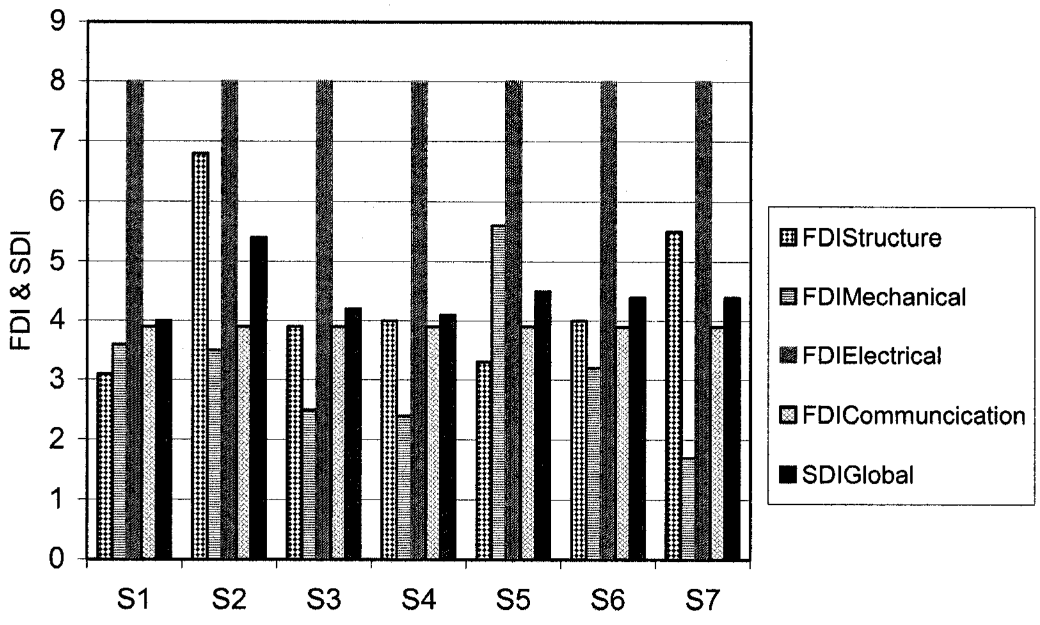


Figure 5.3 SSDI Application Result Graph

The main observation from the above table and graph is that the seven stations are found to be 'Deficient'. This justifies the STM choice for the 'Réno-Station' rehabilitation program. In addition, it is found that S2 is the best station, while S1 is the worst. For station S2, the SDI is 'Deficient'; although it has an FDI of 'Medium' for structural and electrical functions. This also justifies the use of the multiplicative form of MAUT in the formula for SDI evaluation. It should be noted that the good electrical condition of the seven stations does not increase the value of the SDI, since its functional weight is less than the others.

5.9 The SSDI model vs. PROMETHEE Ranking

Since the SSDI model is based mainly on the PROMETHEE outranking method, the rank of the stations according to their condition should theoretically be the same for both models. The seven stations were analysed with the PROMETHEE method and the results are tabulated. The utilization of the PROMETHEE application to the seven stations consisted of several assumptions:

- The PROMETHEE Preference Threshold is identified with the Critical Threshold of the SSDI model.
- The PROMETHEE Indifference Threshold is identified with the Tolerance Threshold of the SSDI model.
- The same SSDI gpf shape is used in PROMETHEE.
- The lower and upper datum stations are not used in PROMETHEE, and therefore only stations S1 to S7 are considered.
- Both the SSDI model and PROMETHEE used the same criteria weights

- PROMETHEE does not consider functional a multi-criteria preference index and flow calculation.

Table 5.13 illustrates the multi-criteria preference index and the flow calculation matrix in PROMETHEE.

Table 5.13 PROMETHEE Application to STM Matrix

Multi-Criteria Preference Index Π								
Π	S1	S2	S3	S4	S5	S6	S7	Φ^+
S1	0.0	0.3	0.2	0.2	0.3	0.2	0.2	1.36
S2	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.15
S3	0.1	0.2	0.0	0.1	0.1	0.1	0.1	0.68
S4	0.2	0.2	0.1	0.0	0.1	0.0	0.2	0.76
S5	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.42
S6	0.2	0.2	0.1	0.0	0.1	0.0	0.1	0.67
S7	0.1	0.2	0.1	0.2	0.2	0.2	0.0	0.90
Φ^-	0.79	1.09	0.54	0.48	0.84	0.49	0.71	

The seven stations are compared to each other. The result of PROMETHEE is a rank of the stations according to the net flow. The net flows, as well as the Best-to-Worst station rank are tabulated in Table 5.14.

Table 5.14 PROMETHEE Application to STM Rank

Station	S1	S2	S3	S4	S5	S6	S7
Φ^{net}	0.57	-0.93	0.14	0.28	-0.43	0.18	0.19
Rank	7	1	3	6	2	4	5

The above table shows S2 as the best station, and S1 as the worst station. Table 5.15 is a comparison table of the ranking of the SSDI model and the PROMETHEE application.

Table 5.15 STM Stations Ranking

Station	S1	S2	S3	S4	S5	S6	S7
SSDI Rank	7	1	5	6	2	3	4
PROMETHEE Rank	7	1	3	6	2	4	5

The rank comparison proves that S2, S5, S4 and S1 have the same rank. The remaining stations (S3, S6, and S7) are very close to each other. In the SSDI model, the SDI difference between the latter stations is minimal, and the same for S6 and S7. The net flow in PROMETHEE for these stations is also very close (0.14, 0.18, and 0.19). It can be said that the SSDI model and PROMETHEE result in the same station ranking. This justifies the use of the lower and upper boundaries in the SSDI model.

5.10 Sensitivity Analysis

In order to confirm the results of the SSDI application to the STM stations, a sensitivity analysis is required. This sensitivity analysis is used to study the effect of the criteria weights and the two thresholds on the SDI result. The range of the change in the weights and the thresholds is the same range used in the statistical analysis (95% reliability range of values). The range for the weights is ±20% and the range for the thresholds is ±15%.

The sensitivity analysis is performed using @Risk software. Figure 5.4 shows the SDI change versus the weight change for S1. S1 is chosen since it is the worst station (SDI is the lowest).

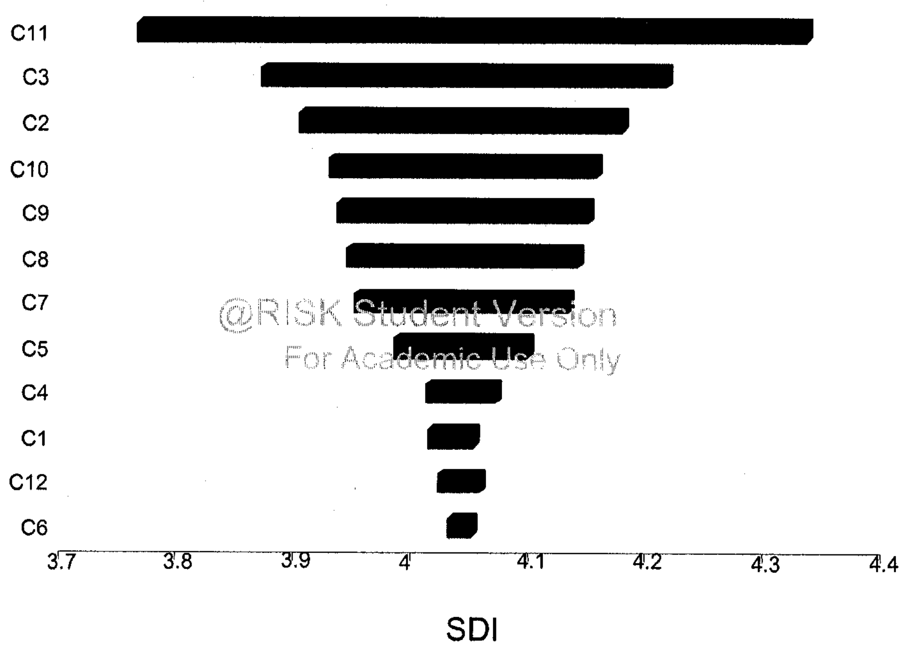


Figure 5.4 Sensitivity of SDI to Criteria Weights for S1

The percentage change in SDI for a ±20% is ±5.0%, which is very small. The sensitivity analysis is repeated for S2, and is found to have the same percent of change (5%). Therefore, the weight range for 95% reliability does not affect the SDI result. Special attention should be paid for the stations that have a SDI at the limit between two condition states (SDI=5.8, for example). Here, an increase of 5% gives a SDI of 6.1. In this particular case, the FDI analysis is more significant. The second sensitivity analysis is performed on the Critical Threshold. Figure 5.5 shows the SDI vs. the CT change for S1.

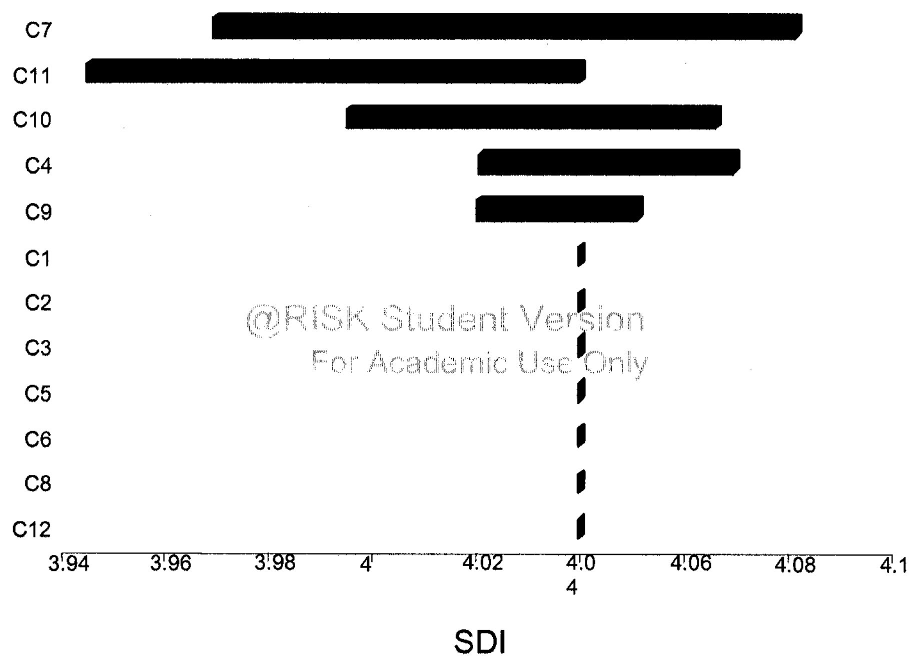


Figure 5.5 Sensitivity of SDI to Criteria CT for S1

The above graph indicates that the change in the Station Diagnosis Index (SDI) is even less than the weights with an SDI change of ±3%, if the Critical Threshold (CT) changes in the 95% reliability range. For Criteria C7 and C11, the SDI changes from 3.94 to 4.08. This does not change the condition of station S1. A 15% change in the Critical Threshold means a degree less in the Factor of Safety in the design of the elements related to these criteria.

The same analysis is performed for the Tolerance Threshold. This threshold varies ±15% for 95% reliability. Figure 5.6 illustrates the SDI change versus the Tolerance Threshold change for station S1.

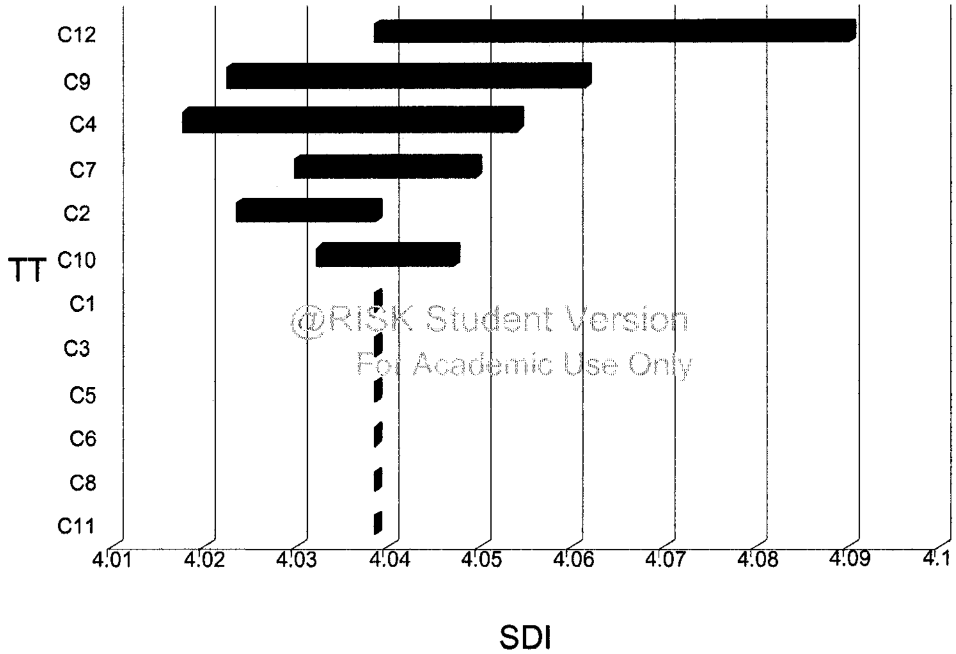


Figure 5.6 Sensitivity of SDI to Criteria TT for S1

The Tolerance Threshold (TT) has a minimal affect on the SDI. A 15% change in the TT only changes the SDI 1% for one criterion only, C12. The SDI change is less for the remaining criteria. Therefore, the SDI is relatively 'not' sensitive to the Tolerance Threshold.

The sensitivity analysis proves that the SDI results for the STM stations (S1 to S7) are true, even if the criteria weights and thresholds are variable in a 95% range of confidence.

5.11 Verification of Results

The STM manager was asked again to assign indices for each of the seven stations according to the SDI scale. The 'STM' indices are then used to verify the

SSDI model results. There are several methods to verify a model based on statistic tests. This thesis adopts a straightforward method, which consists of calculating a Verification Factor (VF) (Zayed and Halpin 2004). The Verification Factor (VF) is defined in Equation 5.1:

$$VF = MR / AR \tag{5.1}$$

Where *MR = SSDI model results = SDI*

And *AR = Assigned Results = STM index*

Table 5.16 indicates The SSDI model SDI index versus the STM assigned index, with the respective description, for S1 to S7, in addition to VF:

Table 5.16 The SSDI Model Verification

	SSDI Model		STM Assignment		Verification
Station	SDI	Description	Index	Description	VF
S1	4.0	Deficient	5.0	Deficient	0.80
S2	5.4	Deficient	6.6	Medium	0.82
S3	4.2	Deficient	5.0	Deficient	0.84
S4	4.1	Deficient	5.0	Deficient	0.82
S5	4.5	Deficient	5.5	Deficient	0.82
S6	4.4	Deficient	5.5	Deficient	0.80
S7	4.4	Deficient	5.5	Deficient	0.80

Next, Figure 5.7 illustrates the above table in a graph form.

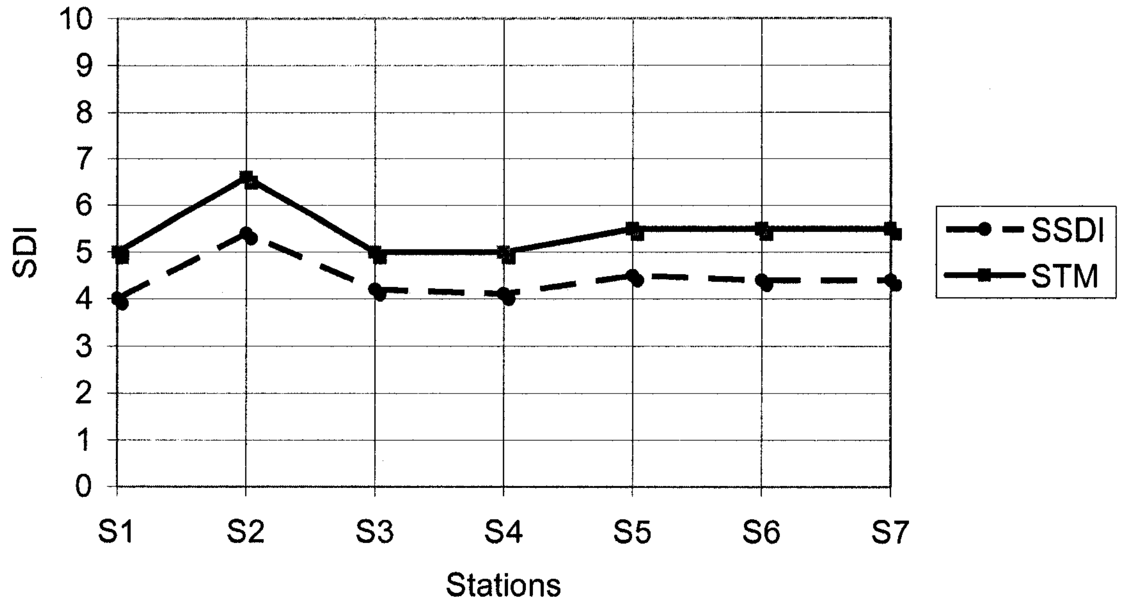


Figure 5.7 The SSDI model Verification Graph

The indices from the STM are higher than the ones calculated from the SSDI model. The STM does not have a decision support system that can be used to assign or to evaluate the overall condition of its stations, although they have the field inspection scale (1 to 5, see Chapter 3). This scale was only applicable to separate elements or functions. The STM also does not have any weight allocation methodology. For these two reasons, any index obtained from the STM, prior to the SSDI model, will be purely subjective, and solely based on the manager's point of view rather than scientific input. This does not underweight the STM manager's decision index, since it is based on his experience in interpreting the field inspection reports. Nevertheless, the condition assessment is '*Deficient*' for all stations, except S2, since the S2 SDI index falls close to the edge between '*Deficient*' and '*Medium*'. This again justifies the short term need for intervention in the 'Réno-Station' MR program.

The VF range is approximately 0.80, which is close to 1.0 and this is meaningful. The most probable explanation for this is that the SSDI model, by using a multiplicative form of MAUT, reduces the SDI if an important function has a low FDI. Any regular subjective decision assignment will tend to use the additive form (closest to human reasoning), which is not true in this case.

6 The SSDI Automated Tool

6.1 The SSDI Automated Tool Steps

An automated SSDI modelling tool is developed in order to make it accessible to transit authorities. The original SSDI model calculations are performed in an Excel file. 'Macros' are added to the SSDI model Excel file, in order to transform it into an automated tool. The SSDI automated tool is illustrated in the following flow diagram in Figure 6.1.

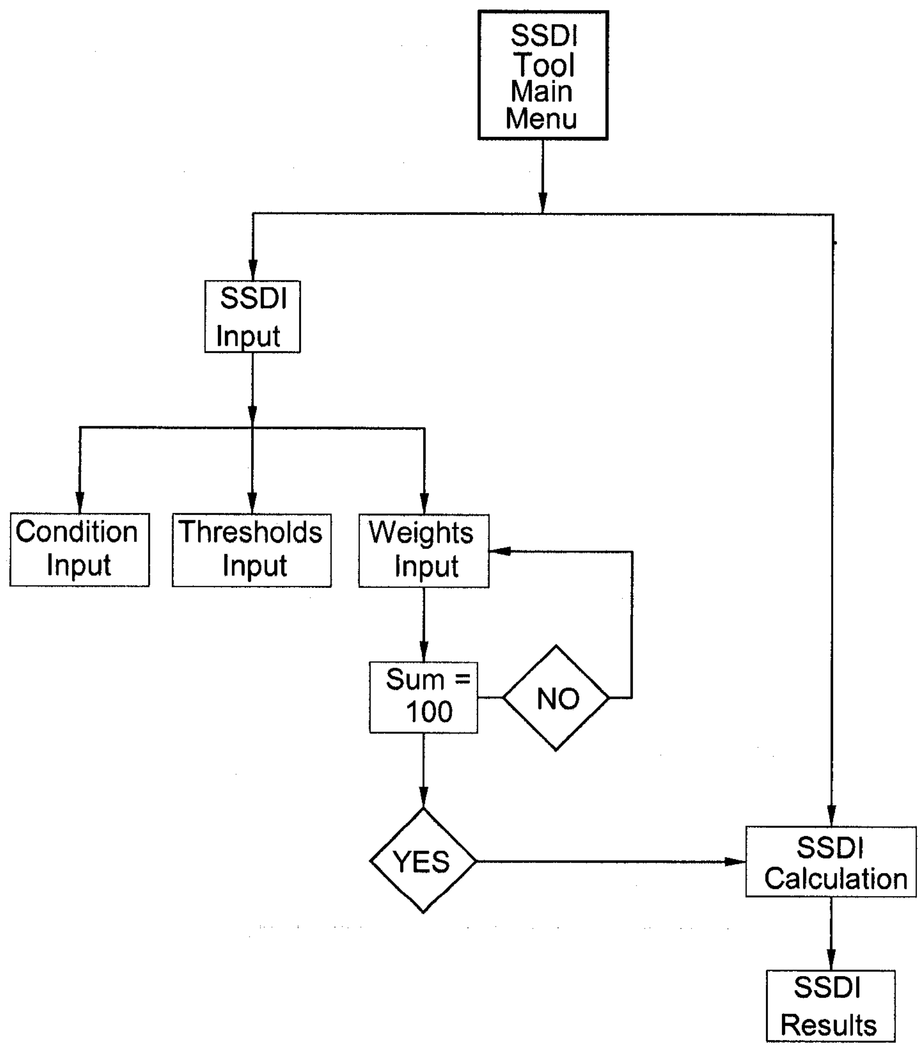


Figure 6.1 The SSDI Model Automated Tool Flow Diagram

The SSDI model automated tool file is entitled 'SSDIApp.xls', which stands for '*SSDI Application'*. When the 'SSDIApp.xls' opens, the user will be shown a picture of a subway station, and the '*SSDI Main Menu*' window display, with two options, '*SSDI Inputs*' and '*SSDI Calculation*'. The main menu for the tool is shown in the following Figure 6.2:

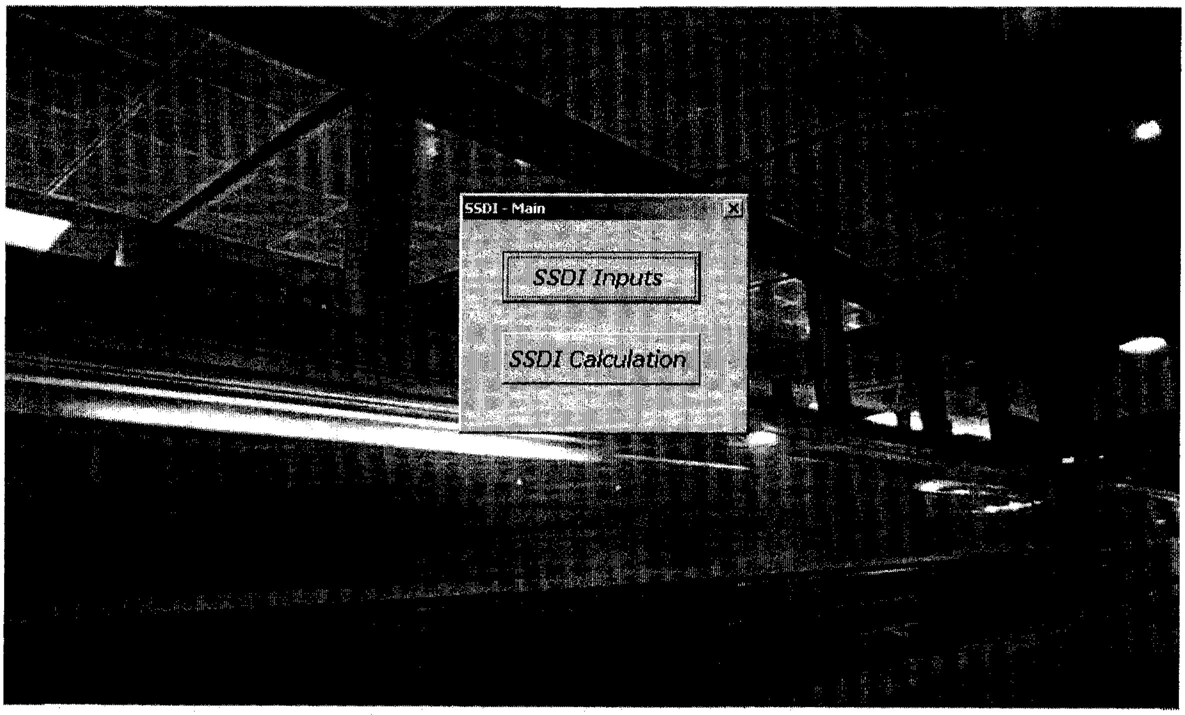


Figure 6.2 The SSDI Automated Tool Main Menu

The user may choose between the two options (dialog boxes). The last input is always stored in the '*SSDI Inputs*' window.

6.2 The SSDI Model Automated Tool Input

When the user chooses the '*SSDI Inputs*' button, a new window opens, which is the '*Input Menu*'. The '*Input Menu*' consists of a series of input fields required for the SSDI namely the weights, the thresholds (Critical and Tolerance), and the inspection scale.

The criteria in the SSDI tool have already been defined in the thesis. They are fixed in number, while their hierarchy is not apparent in the program. Figure 6.3 shows the '*Input Menu*'.

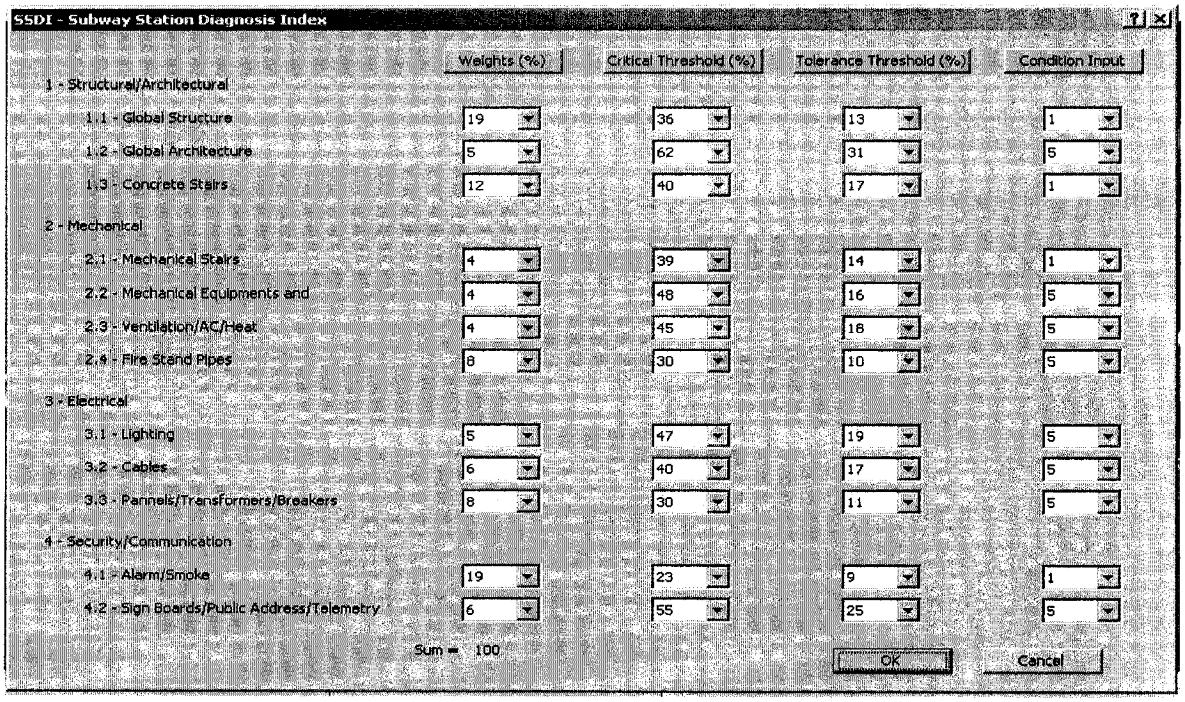


Figure 6.3 The SSDI Automated Tool Input Menu

Following is the list of the four lines of entry in the SSDI tool input:

1. The first line of entry is the criteria weights. The SSDI tool leaves the decision maker the freedom to enter the weights he or she deems best for the analysis. Thus it does not bind the decision maker by using AHP for the weights. Therefore. if the user wants to run AHP, it can be done from external software, either Expert Choice (widely found in the market) or any Excel spreadsheet. Figure 6.4 illustrates the '*Weights Description*' window, when the '*Weights%*' dialog box is chosen.

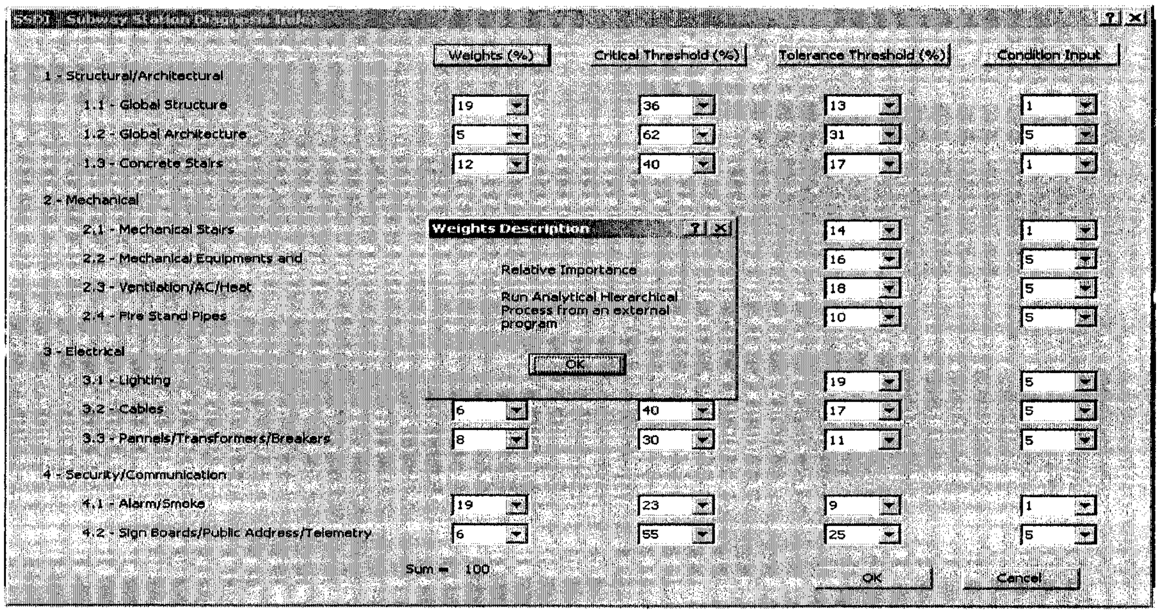


Figure 6.4 The SSDI Automated Tool Weights Description Window

A sole condition with AHP, regarding the Weights input, is that their sum should be equal to 100%. Therefore, the user should make sure that the weights add up to 100%. If it is not the case, an '*ALERT*' window will display. Figure 6.5 illustrates the '*ALERT*' window.

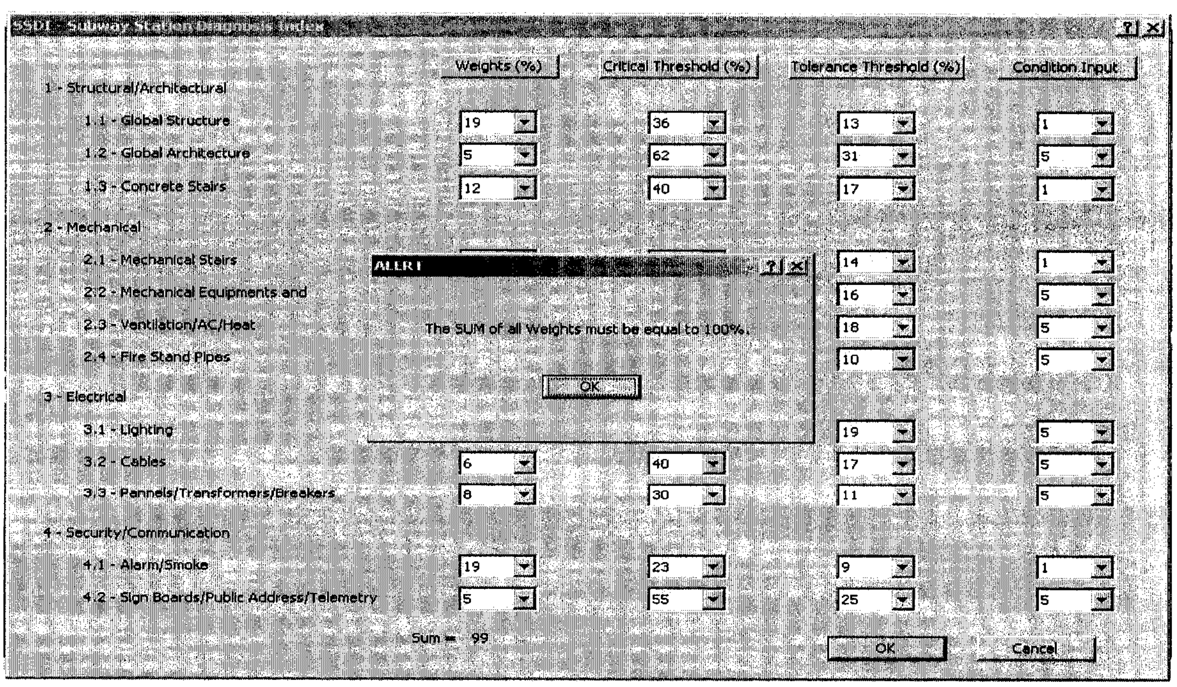


Figure 6.5 The SSDI Automated Tool ALERT Window

2. The second line of entry is the Critical Threshold. A click on the '*Critical Threshold %*' button displays a window describing the meaning of the threshold. Figure 6.6 shows the '*Critical Threshold Description*' window.

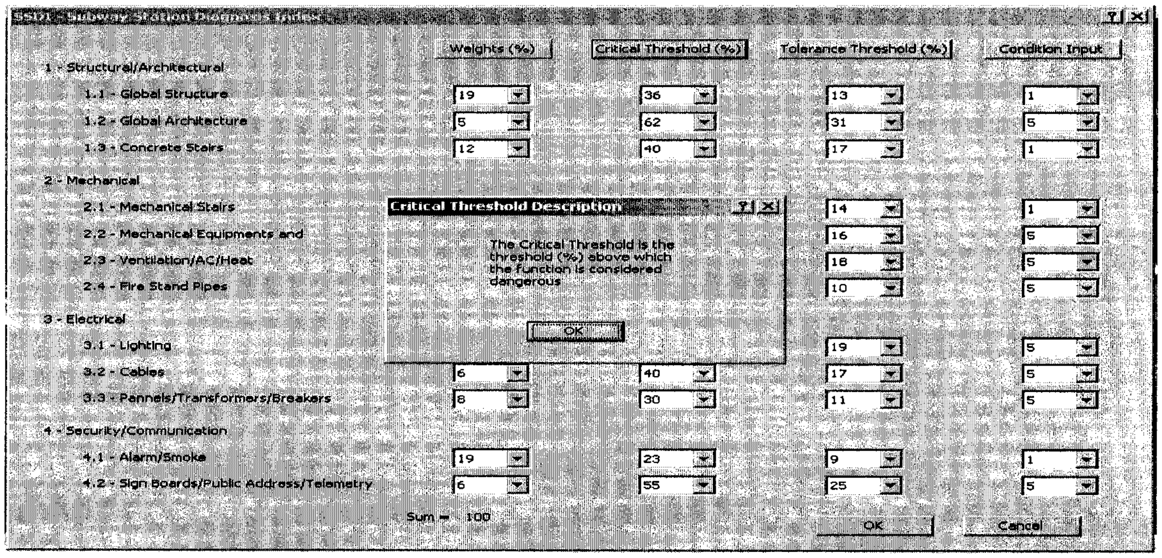


Figure 6.6 The SSDI Automated Tool Critical Thresholds Description Window

3. The third line of entry is the Tolerance threshold. Similar to the Critical threshold, a description is provided upon clicking on the '*Tolerance Threshold %*' button, as illustrated in Figure 6.7.

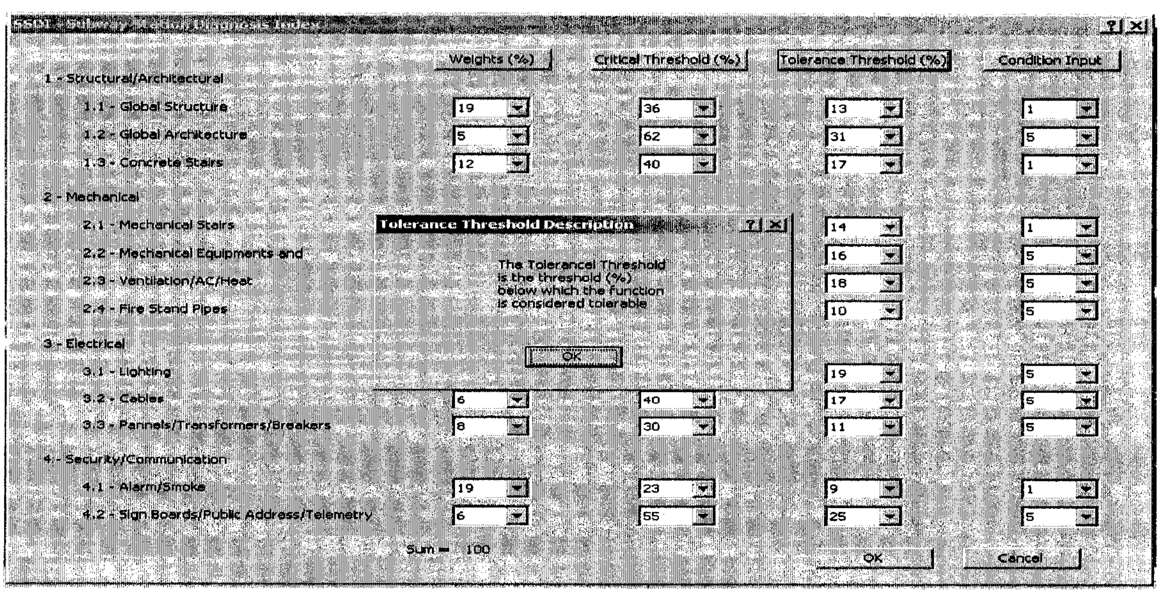


Figure 6.7 The SSDI Automated Tool Tolerance Thresholds Description Window

4. The fourth entry is the criteria performances, or the 'condition Input' values. The criteria input scale is utilized by the STM and MTQ in their assessments. This scale can be used universally, but if a transit authority uses another scale it can perform a transformation to the program scale. Figure 6.8 illustrates the '*Condition Input*' description window.

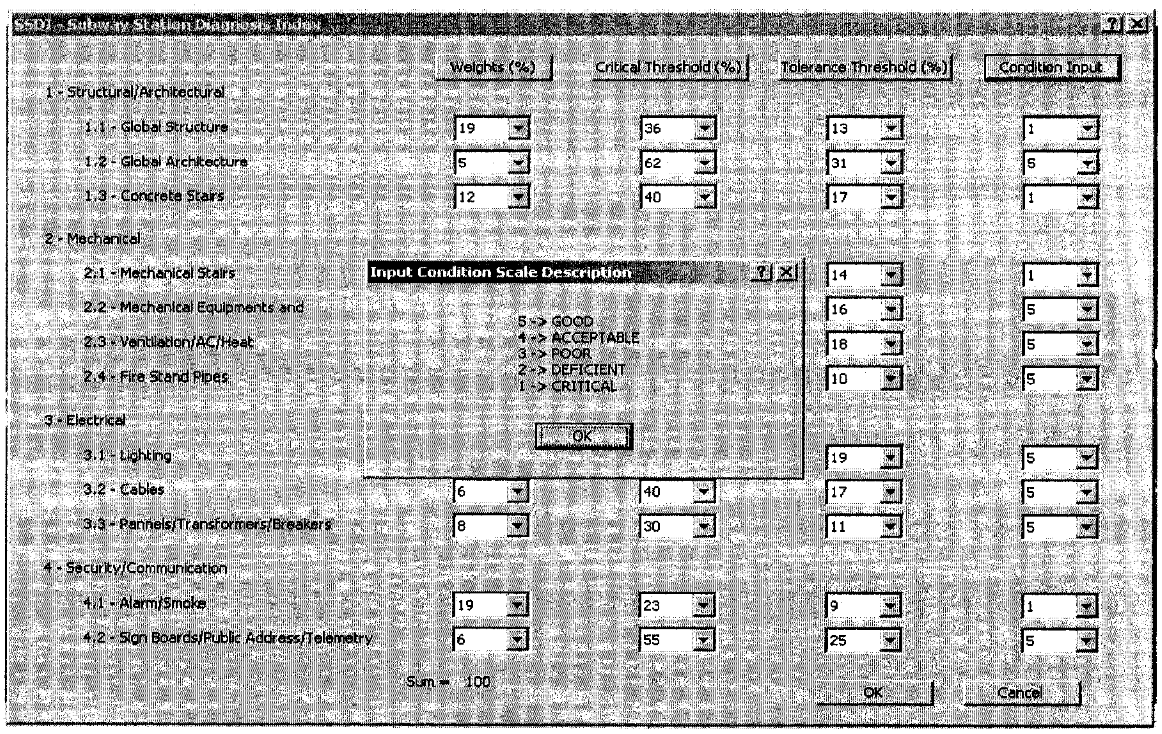


Figure 6.8 The SSDI Automated Tool Condition Input Window

6.3 The SSDI Automated Tool Output

After the input has been completed, the user clicks on the '*OK*' button, the '*Main Menu*' reopens (Figure 6.2) and the user has to click on the '*SSDI Calculation*' button to perform the calculation. All calculations are performed in the background. After completion the '*SSDI Results*' window opens, as shown in Figure 6.9. The user may quit the SSDI tool, or do another input. If the user chooses to quit, the information is saved directly.

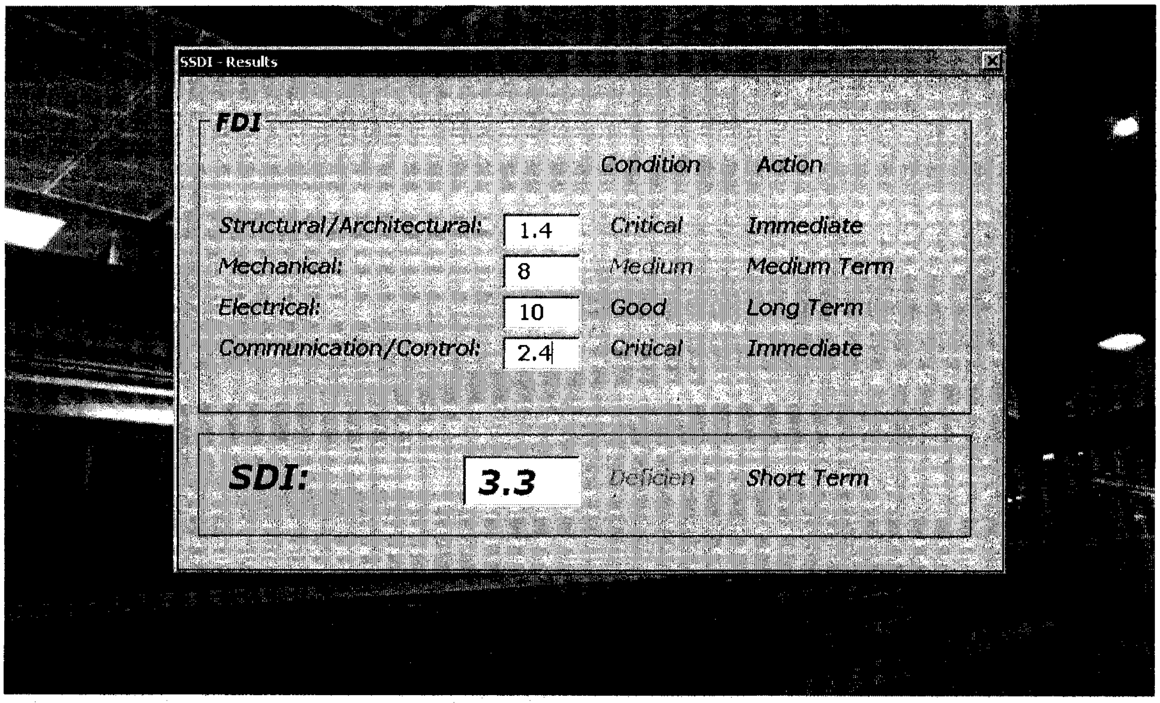


Figure 6.9 The SSDI Automated Tool Result Window

6.4 Remarks on the SSDI Automated Tool

Some remarks on the SSDI automated tool are as follows:

- The SSDI tool fixes the criteria and does not leave the choice of adding new criteria.
- The SSDI tool does not run AHP to evaluate the criteria weights, but allows the user the freedom to input weights.
- The SSDI tool is not flexible in the input of the criteria performances, but binds it to an input scale from 1 to 5.
- The SSDI tool is not transparent in showing the calculation (preference index, strength and weakness of criteria).

Enlarging the SSDI automated tool to include all of these aspects requires a huge effort in programming, and is outside the scope of this research. The SSDI tool delivers the research results in an easy way for application purposes only.

7 Conclusions and Recommendations

7.1 Conclusions

A new model for evaluating the condition, or performing a diagnosis, of subway stations was developed, the Subway Station Diagnosis Index, SSDI. The new model satisfies the need of a well-defined and unified condition assessment model and index for subway stations. Based on this research, the following conclusions may be drawn:

- The operational condition assessment of subway stations is mainly affected by the structural and security functions. The structural function weights (36.1%) and the security weights (27.3%), totalled 63.4% of the total functional weights.
- Public security in the subway stations is directly affected by the security, structural and concrete stairs criteria. The security criterion is most important, with a weight of 21.2%. The structural criterion follows, with a weight of 18.8%. The concrete stair condition is the third most important criteria with a weight of 12.2%. The remaining criteria are less important to the decision process.
- The criteria weights evaluated from the questionnaires using AHP are similar to the weights evaluated by the STM manager.
- The critical thresholds for the different functions are between 30% and 41%, while the tolerance thresholds are beyond 10% and 17%. This indicates that a deterioration level for any element of approximately 10%

is still tolerable. But when the deterioration level becomes greater than 30%, the element becomes dangerous.

- Based on the SSDI model and the statistical and sensitivity analyses performed on the collected data, the Diagnosis Index SDI value of 0 – 3 indicates a 'critical' subway station, and requires 'immediate' physical intervention. The SDI value of 3 - 6 indicates a 'deficient' station and requires a 'short term' intervention and its value of 6 - 8 indicates a 'medium' station and requires a 'medium term' intervention. Finally, the SDI value of 8 - 10 indicates a 'good' station and requires a 'long term intervention'.
- An 'extremely bad' subway station has an SDI equal to 0, while an 'extremely good' station has an SDI equal to 10. An 'average' station has an SDI equal to 5.
- The application of the SSDI to seven STM stations proves that these stations are 'deficient', although their electrical function is 'medium', with a FDI of 8. The result observed from the SSDI model justifies the choice of the seven stations in the 'Reno-Station' rehabilitation program.
- Ranking of various stations using the SSDI model and the PROMETHEE method is similar.
- The SSDI model is not considerably sensitive to the criteria weights. A change of ±20% in the weights affects the SDI by only ±5%. The new model is even less sensitive to the thresholds. Hence, a change of ±15% in both thresholds affects the SDI by a maximum of ±3%.

- The SSDI model results verification shows a discrepancy of 20% from the indices assigned by the STM. This clearly enhances the difference between a scientific decision-making model and a subjective decision-making assignment.

7.2 Research Contributions

The SSDI model solves the problem of condition assessment related to subway stations. It provides several contributions to the MCDA field of research, which are identified as follows:

- Criteria hierarchy for the condition assessment of subway stations.
- The SSDI model that introduces the idea of absolute boundaries to an outranking method, thus transforming it to a specific MAUT method, while maintaining its mathematical skeleton. It combines an outranking method (PROMETHEE) with MAUT method. In addition, it considers uncertainty and lack of information from the field data (inspection reports) by using the concept of pseudo-criteria.
- Introducing the concepts of Critical and Tolerance Thresholds to the infrastructure management field whereas they can be used in several other applications.
- The condition assessment scale for subway stations.
- The SSDI model contributed to the field of asset management for sustainable infrastructure.

7.3 Research Limitations

The developed SSDI model has several limitations, such as:

- It does not consider group decisions. Criteria weights and thresholds, as well as indices are evaluated from an individual point of view. Group decision comprises the development of a 'consensus' decision, and a social interaction procedure.
- It considers functional criteria solely.
- The SSDI model, by using AHP for criteria weights, assumes that the criteria are fixed. So if new criteria are introduced, new matrices have to be formed, and new weights have to be evaluated.
- The SSDI model is applied and validated for seven stations only, and for one transit authority. It has to be applied and validated for more stations in several transit authorities, preferably worldwide.

7.4 Future Research Potential

The SSDI model enhances existing and encourages future studies in decision analysis and therefore facilitates the development of new ideas to this field of research.

7.4.1 Research Enhancements

North American research on outranking methods has, until now, been limited in the field of infrastructure management. The SSDI paves the way for research in outranking methods for North American managers and researchers. The SSDI model itself is the first step for the management of rehabilitation programs in transit authorities and it can be expanded to consider more criteria. The SSDI automated tool can also be further developed to consider more criteria, to include AHP, to be flexible in the condition input of criteria performance, and be transparent in the calculation.

7.4.2 Research Extensions

- Further research can concentrate on SSDI critical and tolerance thresholds, since these are important figures useful for many management applications. More studies are needed to get standard figures for the thresholds, since they could change from one country to another and from the use of one code to another.
- Group decision-making can be introduced into the SSDI model, since it represents a real life situation with respect to some of the transit authorities' policies. The thresholds and criteria weights can be evaluated using a 'consensus' method and these values can be revised and adjusted from social interaction procedures.
- The SSDI model can be used to assess the condition of other infrastructures, by altering criteria and thresholds. Its output can be compared to models in bridges, pavement, underground pipes, etc.
- Concepts such as fuzzy set theory can be introduced to SSDI model in order to generalize the evaluation of criteria performance.

References

Abu-Mallouh M., *Model for Station Rehabilitation and Planning (MSRP)*, (1999), PhD Thesis, Polytechnic University, Construction Engineering and Management Department, Civil Engineering.

Abu-Taleb M., (1993), *Optimization of Non Renewable Groundwater Development and Related Conveyance Systems Design using Multi-Criteria Analysis*, PhD. Thesis, The Catholic University of America, USA.

Abu-Taleb M., and Mareschal B., (1995), "Water Resources Planning in the Middle East: Application of the PROMETHEE V Multi-Criteria Method", *European Journal of Operational Research*, Vol. 81, pp. 500-511.

Alam M., (1992), *Multi Objective Optimization and Multi-Criteria Decision Analysis for Reservoir Operations*, PhD. Thesis, Colorado State University, USA.

Al-Barqawi H., and Zayed T., (2006), "Condition Rating for Underground Infrastructure Sustainable Water Mains", *Journal of Performance of Constructed Facilities*,Vol. 20, Issue 2, p. 126 .

Al-Kloub B., Al-Shemmeri T., Pearman A., (1997), "The Role of Weights in Multi-Criteria Decision Aid, and Ranking of Water Projects in Jordan", *European Journal of Operational Research*, Vol. 99, pp. 278-288.

Al-Rashdan D., Al-Shemmeri T., Al-Kloub B., Dean A., (1999), "Environmental Impact Assessment and Ranking the Environmental Projects in Jordan", *European Journal of Operational Research*, Vol. 118, pp. 30-45.

Al-Shemmeri T., Al-Kloub B., Pearman A., (1997), "Computer Aided Decision Support System for Strategic Planning in Jordan", *European Journal of Operational Research*, Vol. 102, pp. 455-472.

Baron F.H., and Barret B.E., (1996), "Decision Quality using Ranked Attributes Weights", *Management Science*, Vol. 42, No. 11, pp. 1515-1523.

Belton V., and Stewart T., (2002), *Multiple Criteria Decision Analysis*, An Integrated Approach, Kluwer Academic Publishers, Norwell, MA., USA.

Brans J.P., and Mareschal B., (1986), *How to decide with PROMETHEE*, ULB and VUB Brussels Free Universities, http://smg.ulb.ac.be

Brans J.P., Mareschal B., Vincke P., (1986), "How to select and how to rank projects: the PROMETHEE method", *European journal of Operational Research*, Vol. 24, pp. 228-238.

Bryson N., and Joseph A., (2000), "Generating Consensus Priority Interval Vectors for Group Decision-making in the AHP", *Journal of Multi-Criteria Decision Analysis*, Vol. 9, pp. 127-137.

Cox G.L., (2003), *Applying Multi-Criteria Decision Analysis in Participatory Planning for Sustainable Management of the New York State Forest*, PhD. Thesis, Rensselaer Polytechnic Institute, USA.

DeStephano P., (1998), *Performance Prediction and Decision Analysis in Bridge Management*, PhD Thesis, Civil Engineering, Rensselaer Polytechnic Institute, USA.

Doumpos M., and Zopounidis C., (2002), *Multi Criteria Decision Aid Classification Methods*, Kluwer Academic Publishers, Norwell, MA. USA.

Doyle JR., Green R.H., Bottomley P.A., (1997), "Judging Relative Importance: Direct Rating and Point Allocation are not Equivalent", *Organizational Behaviour and Human Decision Processes*, Vol. 70, No. 1, pp. 55-72.

Ekart A., and Nemeth S.Z., (2005), "Stability Analysis of Tree Structured Decision Functions", *European Journal of Operational Research*, Vol. 160, pp. 676-695.

Engineering Statistics Handbook, http://www.itl.nist.gov/div898/handbook

Etezadi-Amoli J., and Ciampi A., (1983), "Simultaneous Parameter Estimation for Multiplicative Multi Attribute Utility Model", *Organizational Behaviour and Human Performance*, Vol. 32, pp. 232-248.

Fernandez-Castro AS., and Jimenez M., (2005), "PROMETHEE: an extension through fuzzy mathematical programming", *Journal of the Operational Research Society*, Vol. 56, pp. 119-122.

Fodor J., (2000), "*Preference Relations in Decision Models*", Research Paper, Department of Biomathematics and Informatics, Szent Istvan University, Budapest, Hungary.

Fulop J., (2005), "*Introduction to Decision Making Methods*", Laboratory of Operations Research and Decision Systems, Computer and Automation Institute, Hungarian Academy of Sciences.

Geldermann J., and Rentz O., (2000), "Bridging the Gap between American and European MADM-Approaches", 51st Meeting of the European Working Group Multi-Criteria Aid for Decisions, Madrid.

Georgopoulou E., Sarafidis Y., Diakoulaki D., (1998), "Design and Implementation of a group DSS for sustaining renewable energies exploitation", *European Journal of Operational Research*, Vol. 109, pp. 483-500.

Good I., (2006), *Resampling Methods: a Practical Guide to Data Analysis*, 3rd Ed., Birkhouser, Boston, MA., USA.

Goodwin P., and Wright G., (1998), *Decision Analysis for Management Judgement*, 2nd Ed., Wiley Publishers, N.Y., USA.

Goumas M., and Lygerou V., (2000), "An Extension of the PROMETHEE method for Decision Making in Fuzzy Environment: Ranking of alternative exploitation projects", *European Journal of Operational Research*, Vol. 123, pp. 606-613.

Haralambopoulos D.A., and Polatidis H., (2003), "Renewable Energy Projects: Structuring a Multi Criteria Group Decision Making Formwork", *Renewable Energy*, Vol. 28, pp. 961-973.

Hokkanen J., and Salminen P., (1997), "Locating a Waste treatment Facility by Multi Criteria Decision Analysis", *Journal of Multi-Criteria Decision Analysis*, Vol. 6, pp. 175-184.

Hyde k., Maier H., Colby C., (2003), "Incorporating Uncertainty in the PROMETHEE MCDA Method", *Journal of Multi-Criteria Decision Analysis*, Vol. 12, pp. 245-259.

Infrastructure Condition Assessment: Art, Science, and Practice, (1997), *Proceedings of the Conference*, sponsored by the Facilities Management Committee of the Urban Transportation Division of the American Society of Civil Engineers, August 25-27, Boston, MA, USA.

Interview with Pierre Vandelac, (2006), "Le Service est interrompu…", *La Presse Newspaper*, Montréal, Canada.

Jia J., Fischer GW., Dyer JS., (1998), "Attribute Weighting Methods and Decision Quality in the Presence of Response Error, a simulation study", *Journal of Behavioural Decision Making*, Vol. 11, No. 2, pp. 85-105.

Kangas A., Kangas J., Pykalainen J., (2001), *"Outranking Methods as Tools in Strategic Natural Resources Planning*", Silva Fennica Research Articles, Vol. 35, pp. 215-227.

Keeney R.L., and Raiffa H., (1976), *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley and Sons, N.Y., USA.

Kolli SS., (1992), *A Multi-Criteria Methodology for the Economic Evaluation of Flexible Manufacturing Systems*, Master's Thesis, University of Louiseville, USA.

Lam P., Moskowitz H., Eppel T., Tang J., (1997), "Decomposition, Interdependence and Precision in Multi Attribute Utility Measurements", *Journal of Multi-Criteria Decision Analysis*, Vol. 6, pp. 25-40

Lamsade-Dauphine, Paris, http://www.lamsade.dauphine.fr

Leskinen P., (2000), "Measurement Scales and Scale Independence in the Analytical Hierarchy Process", *Journal of Multi-Criteria Decision Analysis*, Vol. 9, pp. 163-174.

Macharis C., Springael J., De Brucker K., Verbeke A., (2004), "PROMETHEE and AHP: The Design of Operational Synergies in Multi Criteria Analysis. Strengthening PROMETHEE with ideas of AHP", *European Journal of Operational Research*, Vol. 153, pp.307-317.

Mahmoud M.R., (1995), *Multi-Criteria Location and Sizing of Water Resource Projects by Dynamic Programming and GIS*, PhD. Thesis, Colorado State University.

Meszaros C. and Rapcsak T., (1996), "On Sensitivity Analysis for a Class of Decision Systems", *Decision Support Systems*, Vol. 16, pp. 231-240.

Montgomery D., and Runger G., (2007), *Applied Statistics and Probability for Engineers*, 4th Ed , Wiley and Sons, N.J., USA.

Olzekan EC., and Dukstein L., (1996), "Analysing Water Resources Alternatives and Handling Criteria by Multi Criterion Decision Techniques", *Journal of Environmental Management*, Vol. 48, No. 1, pp. 69-96.

Palisade, (2000), *@Risk Analysis Add-In for Microsoft Excel*, 4.0.5 (Student Version).

Petrie J., Stewart M., Basson L., Notten P., Alexander B., (2006), "*Structured Approaches to Decision Making for Cleaner Products and Processes*", CRESTA, Center for Risk, Environment and Systems Technology and Analysis", Department of Chemical Engineering, University of Sydney, Sydney, Australia.

Raju K.S., Duckstein L., Arondel C., (2000), "Multi Criterion Analysis for Sustainable Water Resources Planning: a case study in Spain", *Water Resources Management*, Vol. 14, pp. 435-456.

Raju K.S., and Pillai C., (1999), "Multi Criterion Decision Making in River Basin Planning and Development", *European Journal of Operational Research*, Vol. 112, pp. 249-257.

Roberts R., and Goodwin P., (2002), "Weight Approximations in Multi-Attribute Models", *Journal of Multi-Criteria Decision Analysis*, Vol. 11, pp.291-303.

Roy B, Present M., Silhol D., (1986), "A programming method for determining which Paris Metro Stations should be renovated", *European Journal of Operational Research*, Vol. 24, Issue 2, pp. 318-334.

Roy B., (1990), "Decision-Aid and Decision Making", *European Journal of Operational Research*, Vol. 45, No. 2-3, pp. 324-331.

Saaty T., (1980), *The Analytical Hierarchy Process*, McGraw Hill, New York, USA.

Salminen P., and Hokkanen J., (1998), "Comparing Multi-Criteria Methods in the context of environmental problems", *European Journal of Operational Research*, Vol. 104, pp. 485-496.

Shahin M.Y., (1994), *Pavement Management for Airports, Roads, and Parking Lots*, Chapman and Hall, New York, USA.

Sinha S., and Knight M., (2004), "Intelligent System for Condition Monitoring of Underground Pipelines", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 19, pp. 42-53.

Société de Transport de Montréal, (2005), Bureau de projets Réno-Stations, "*Programme de Maintien des Stations Réno-Stations II*", Rapport Justificatif, Synthèse, Décembre.

Société de Transport de Montréal, Bureau de projets Réno-Système, (2004), "*Programme de Maintien du Patrimoine des équipements fixes du métro Réno-Système*", Rapport Justificatif, Synthèse, Juillet.

Tille M., and Dumont A.G., (2003), "Methods of Multi-Criteria Decision Analysis within the road projects like an element of sustainability", *Proceedings of the 3rd Swiss Transport Research Conference*, STRC 03 Conference paper, Session Decision Support, Monte Verita.

Tolliver Heidi, (1996), *A Tail of Four Cities, 'Transit Operation Massachusetts Transit'*, pp.22-74, Boston, USA.

Tompkins E., (2003), "*Using Stakeholders Preferences in Multi-Attribute Decision Making: Elicitation and Aggregation Issues*", Center for Social and Economic

Research on the Global Environment and the Tyndall Center for Climate Change Research, University of East Anglia, Norwich, United Kingdom.

Triantaphyllou E., (1999), "Reduction of Pairwise Comparison in Decision Making via a Duality Approach", *Journal of Multi-Criteria Decision Analysis*, Vol. 8, pp.299-310.

Triantaphyllou E., and Sanchez A., (1997), "A Sensitivity Analysis Approach for some deterministic multi-criteria decision making methods", *Decision Sciences*, Vol. 28, pp. 151-194.

Triantaphyllou E., (2001), "Two New Cases of Rank Reversal when the AHP and some of its Additive Variants are used that do not occur with the multiplicative AHP", *Journal of Multi-Criteria Decision Analysis*, Vol. 10, pp. 11-25.

Veilleux M., (1999), *Une Approche Multicritère a la Composition de Portefeuilles*, PhD. Thesis, Université de Laval, Laval, Canada.

Vignaux G. A., (2005), "Multi-Attribute Decision Problems", March 7,8 pp., available at: www.mcs.ac.nz/courses/OPER251/2004T1/Lecture-Notes/multi.pdf.

Vincke P., Gassner M., Roy B., (1992), *Multi Criteria Decision-Aid*, Wiley and Sons, N.Y., USA.

Von-Winterfeldt D., and Edwards W., (1986), *Decision Analysis and Behavioural Research*, Cambridge University Press, Cambrige, MA, USA.

Wolters W.T.M., and Mareschal B., (1995), "Novel types of Sensitivity analysis for additive MCDM methods", *European journal of Operational Research*, Vol. 81, pp. 281-290.

Yager R.R., (1981), "A Procedure for Ordering Fuzzy Subsets of the unit interval", *Information Science*, Vol. 24, pp. 143-161.

Yan J.M., and Vairavamoorthy K., (2004), "Fuzzy Approach for Pipe Condition Assessment", *Journal of Construction Engineering and Management*, Vol. 7, pp. 466-476.

Zadeh L., (1965), "Fuzzy Sets", *Information and Control*, Vol. 8, No. 5, pp. 338-353.

Zayed T., and Halpin D., (2004), "Quantitative Assessment for Piles Productivity Factors", *Journal of Construction Engineering and Management*, Vol. 130, Issue 3, pp. 405-414.

Zayed T., and Nosair I. (2006), "Cost Management for Concrete Batch Plant using Stochastic Mathematical Models", *Canadian Journal of Civil Engineering,* NRC Canada, August, 33: 1065-1074.

Appendix A

Subway Station Diagnosis Index Questionnaire Sample

1. General Information

This survey is part of a research done at Concordia University, Construction and Engineering Management graduate program, under the title: **Condition Assessment Model of Subway (Metro) Stations**. The purpose of this survey is to analyse the effect of different criteria on the condition assessment of different subway stations, and rate their importance relative to each other.

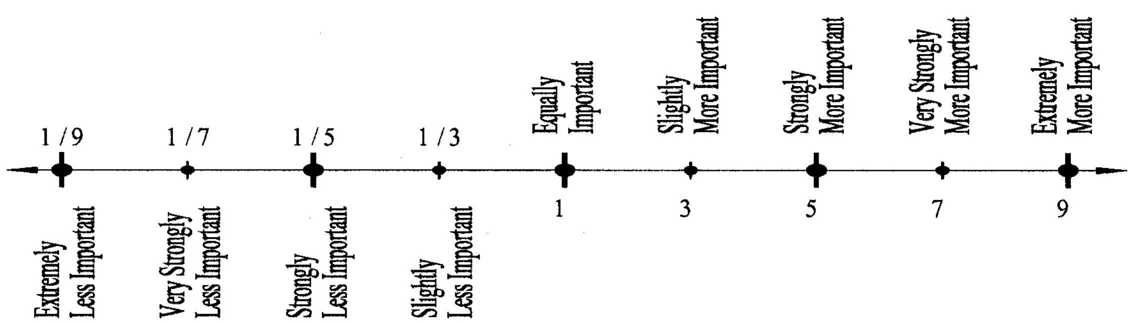
In order to ensure confidentiality, your company information will not be linked in any way to the questions in the subsequent sections. You may respond anonymously if you wish. *Would you like your company name acknowledged as a participant in this research?*

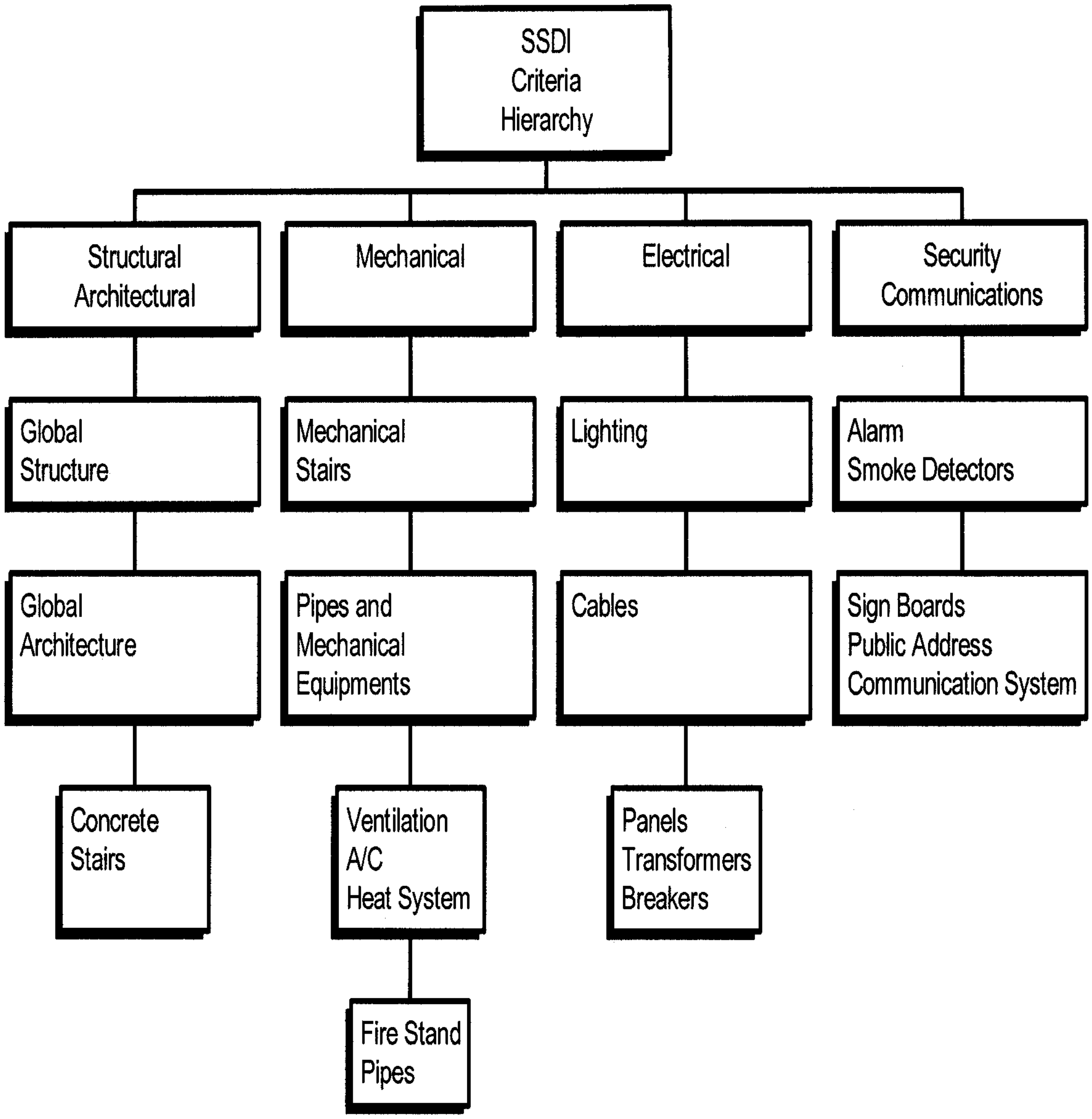
Yes † *No*

Name of the Company	
Location of the Company	
Name of the Respondent	
Title of the Respondent	

2. Comparison between the criteria

The information gathered from this part of the survey will be used to model the importance of each criterion (Level 1) and sub-criterion (Level 2) relative to the whole set of criteria and sub-criteria, respectively. The following questions require a pair-wise comparison between the different criteria (Levels 1 & 2) using the importance scale shown below. The criteria are shown in tables-matrices; using the scale of importance please assign subjectively at each relevant space in the table-matrix a number scale, which describes the relative importance of a specific criterion with respect to the others.





A. **Level 1**: Compare the following criteria with respect to the others:

▶	Structural/Architectural	Mechanical	Electrical	Security / Communication
Structural/Architectural	1			
Mechanical		1		
Electrical			1	
Security / Communication				1

B. Level 2: Compare the following sub-criteria with respect to the others:

Structural / Architectural Sub-Criteria:

►	Global Structure	Global Architecture	Concrete Stairs
Global Structure	1		
Global Architecture		1	
Concrete Stairs			1

Mechanical Sub-Criteria:

►	Mechanical Stairs	Pipes &Mechanical Equipments	Ventilation, A/C, Heat	Fire Stand Pipes
Mechanical Stairs	1			
Pipes &Mechanical Equipments		1		
Ventilation, A/C, Heat			1	
Fire Stand Pipes				1

Electrical Sub-Criteria:

►	Lighting	Cables	Panels, Transformers, Breakers
Lighting	1		
Cables		1	
Panels, Transformers, Breakers			1

Security / Communications Sub-Criteria:

►	Alarm, Security, Smoke Detectors	Sign Boards, Public Address, Communication system (Telemetry)
Alarm, Security, Smoke Detectors	1	
Sign Boards, Public Address, Commun. System		1

3. Criteria Thresholds

The information gathered from this part of the survey will serve to set critical Threshold and tolerance Threshold of each of the sub-criteria (Level 2).

The critical indicates the value (percentage) above which the criterion value is considered critical or dangerous. Whereas the tolerance threshold is the value (percentage) below which the criterion value is considered acceptable, or not dangerous at all. Assuming that criteria values are expressed as percentage, use a scale from 0% to 100% and assign a percentage value for the critical and tolerable thresholds respectively

Criteria	Threshold *(From 0% to 100%)*		Criteria	Threshold *(From 0% to 100%)*	
	Critical	Tolerance		Critical	Tolerance
Global Structure			Alarm, Security, Smoke Detectors		
Global Architecture			Public Address, Sign Boards, Communication System		
Concrete stairs					
Mechanical stairs					
Pipes and Mechanical Equipments					
Ventilation, A/C, Heat					
Fire Stand Pipes					
Lighting					
Cables					
Panels, Transformers, Beakers					

Thank you for your participation…

Appendix B

SSDI Model Application to STM Stations Outputs

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S1}	Φ^{-}_{S1}	Φ^{net}_{S1}	FDI_{S1}
C1	1	100	36	13	*52.6%*	19.0%	*36.1%*	0.86	0.47	0.39	*3.1*
C2	4	25	62	31	*13.9%*	5.0%					
C3	3	50	40	17	*33.5%*	12.1%					
C4	2	75	39	14	*22.5%*	4.0%	*17.8%*	0.92	0.64	0.28	*3.6*
C5	3	50	48	16	*23.0%*	4.1%					
C6	1	100	45	18	*23.6%*	4.2%					
C7	4	25	30	10	*30.9%*	5.5%					
C8	5	0	47	19	*27.5%*	5.2%	*18.9%*	0.41	1.00	-0.59	*8.0*
C9	4	25	40	17	*32.8%*	6.2%					
C10	4	25	30	11	*39.7%*	7.5%					
C11	2	75	23	9	*78.4%*	21.4%	*27.3%*	1.00	0.78	0.22	*3.9*
C12	2	75	55	25	*21.6%*	5.9%					

Input *Input* *Input*

SDI_{Global} = 4.0

$FDI_{Structure}$ = 3.1

$FDI_{Mechanical}$ = 3.6

$FDI_{Electrical}$ = 8.0

$FDI_{Communcication}$ = 3.9

Station S1

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S2}	Φ^{-}_{S2}	Φ^{net}_{S2}	FDI_{S2}
C1	3	50	36	13	52.6%	19.0%	36.1%	0.64	1.00	-0.36	6.8
C2	4	25	62	31	13.9%	5.0%					
C3	4	25	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	1.00	0.71	0.29	3.5
C5	2	75	48	16	23.0%	4.1%					
C6	3	50	45	18	23.6%	4.2%					
C7	3	50	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					
	Input		Input			Input					

SDI_{Global} = **5.4**

Station S2

$FDI_{Structure}$= 6.8

$FDI_{Mechanical}$= 3.5

$FDI_{Electrical}$= 8.0

$FDI_{Communcication}$= 3.9

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S3}	Φ^{-}_{S3}	Φ^{net}_{S3}	FDI_{S3}
C1	3	50	36	13	52.6%	19.0%	36.1%	0.95	0.73	0.22	3.9
C2	3	50	62	31	13.9%	5.0%					
C3	2	75	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	0.83	0.33	0.49	2.5
C5	1	100	48	16	23.0%	4.1%					
C6	4	25	45	18	23.6%	4.2%					
C7	1	100	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					
	Input		Input			Input					

SDI_{Global} = **4.2**

Station S3

$FDI_{Structure}$ = 3.9
$FDI_{Mechanical}$ = 2.5
$FDI_{Electrical}$ = 8.0
$FDI_{Communcication}$ = 3.9

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S4}	Φ^{-}_{S4}	Φ^{net}_{S4}	FDI_{S4}
C1	3	50	36	13	52.6%	19.0%	36.1%	0.86	0.66	0.20	4.0
C2	4	25	62	31	13.9%	5.0%					
C3	1	100	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	1.00	0.47	0.53	2.4
C5	2	75	48	16	23.0%	4.1%					
C6	1	100	45	18	23.6%	4.2%					
C7	3	50	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					

Input *Input* *Input*

SDI_{Global} = **4.1**

Station S4

$FDI_{Structure}$ = 4.0
$FDI_{Mechanical}$ = 2.4
$FDI_{Electrical}$ = 8.0
$FDI_{Communcication}$ = 3.9

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S5}	Φ^{-}_{S5}	Φ^{net}_{S5}	FDI_{S5}
C1	3	50	36	13	52.6%	19.0%	36.1%	0.95	0.61	0.34	3.3
C2	3	50	62	31	13.9%	5.0%					
C3	1	100	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	0.76	0.87	-0.12	5.6
C5	4	25	48	16	23.0%	4.1%					
C6	3	50	45	18	23.6%	4.2%					
C7	4	25	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					
	Input		Input			Input					

SDI_{Global} = **4.5**

Station S5

$FDI_{Structure}$ = 3.3
$FDI_{Mechanical}$ = 5.6
$FDI_{Electrical}$ = 8.0
$FDI_{Communcication}$ = 3.9

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S6}	Φ^{-}_{S6}	Φ^{net}_{S6}	FDI_{S6}
C1	3	50	36	13	52.6%	19.0%	36.1%	0.86	0.66	0.20	4.0
C2	4	25	62	31	13.9%	5.0%					
C3	1	100	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	1.00	0.63	0.37	3.2
C5	2	75	48	16	23.0%	4.1%					
C6	3	50	45	18	23.6%	4.2%					
C7	2	75	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					
	Input		Input			Input					

SDI_{Global} = **4.4**

Station S6

$FDI_{Structure}$= 4.0

$FDI_{Mechanical}$= 3.2

$FDI_{Electrical}$= 8.0

$FDI_{Communcication}$= 3.9

I.D	Input Scale	Criteria Performance	CT	TT	Criteria Weight w_{ci}	Global Weight w_{gi}	Functional Weight w_{fi}	Φ^{+}_{S7}	Φ^{-}_{S7}	Φ^{net}_{S7}	FDI_{S7}
C1	2	75	36	13	52.6%	19.0%	36.1%	0.64	0.75	-0.11	5.5
C2	4	25	62	31	13.9%	5.0%					
C3	4	25	40	17	33.5%	12.1%					
C4	2	75	39	14	22.5%	4.0%	17.8%	1.00	0.33	0.67	1.7
C5	1	100	48	16	23.0%	4.1%					
C6	3	50	45	18	23.6%	4.2%					
C7	1	100	30	10	30.9%	5.5%					
C8	5	0	47	19	27.5%	5.2%	18.9%	0.41	1.00	-0.59	8.0
C9	4	25	40	17	32.8%	6.2%					
C10	4	25	30	11	39.7%	7.5%					
C11	2	75	23	9	78.4%	21.4%	27.3%	1.00	0.78	0.22	3.9
C12	2	75	55	25	21.6%	5.9%					

Input *Input* *Input*

SDI_{Global} = 4.4

Station S7

$FDI_{Structure}$ = 5.5

$FDI_{Mechanical}$ = 1.7

$FDI_{Electrical}$ = 8.0

$FDI_{Communcication}$ = 3.9