A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units

Promotoren: Dr. Ir. M. Molenaar
Hoogleraar in de Theorie van de Geografische Informatie Systemen en de Remote Sensing

Dr. A. C. Imeson
Bijzonder hoogleraar in de Bodemerosie, in het bijzonder in relatie tot de verwoestijningsproblematiek
Universiteit van Amsterdam

A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units

Nanna Suryana

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Landbouwuniversiteit Wageningen,
Dr. C.M. Karssen,
in het openbaar te verdedigen
op woensdag 22 januari 1997
des namiddags om vier uur in de aula.

The research presented in this thesis was performed at
the Department of Geographic Information Processing and Remote Sensing,
the Wageningen Agricultural University
P.O. Box 339, 6700 AH, Wageningen
The Netherlands

The cover is a closeup photograph of run-off at a fallow field

A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units/Nanna Suryana
Thesis Landbouwuniversiteit Wageningen
With summary in Dutch
ISBN 90-5485-630-0

Printed by : Ponsen and Looijen B.V., Grafisch Bedrijf, Wageningen

"To my people in the village "

NN08201, 2208

Nanna Suryana Warsitakusumah **A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units**

PROPOSITIONS

1. The definition of terrain mapping units (TMUs) is context and scale related. This implies that the thematic and the geometric description of TMUs *can be hierarchically* identified, observed, distinguished and described at different class and aggregation levels. (This Thesis).

2. A TMU can be defined as a natural division of the terrain systems based on *terrain relief characteristics.* This allows the establishment of a repeatable boundary of TMU which is important for the definition of repeatable data acquisition procedures. (This Thesis).

3. An Inductive Erosion Model (IEM) can be built on the basis of erosion studies at farmer's field level and incorporated into GIS as *a region specific inference model.* It can be used in a data and information poor environment that is normally found in developing countries for the establishment of field engineering design plans (FEDP). (This Thesis).

4. The geo-information theoretical approach *facilitates* the establishment of the relationship between an IEM and TMUs to predict the occurrence of erosion severity classes at different aggregation levels. (This Thesis).

5. With special reference to a data and information poor environment, the plausibility reasoning expressed using certainty factor (CF) is *a sufficient means* to assess the quality of information produced by an IEM. The CF values attached to a particular erosion severity class can be interpreted and used to evaluate *"risk"* associated with making wrong decisions based on incomplete information and uncertainties in providing the development options. (This Thesis).

6. Studies of the nature of soil structure are not true scientific studies in the sense that there is no formal definition of soil structure that allows quantification through a repeatable measurement. The real interpretation of soil structure remains more an art than a science (after Letey, 1991. The study of soil structure: Science or art?, Australian journal of soil research, vol. 29, pp. 699-707).

7. With GIS maps can be created that are worth a thousand numbers, maybe more. (J.K. Berry, 1993. Beyond Mapping. Concepts, algorithms, and issues in GIS, GIS World Inc, Colorado).

8. The introduction of a PC-based GIS into the establishment of *a completely decentralised* land rehabilitation and soil conservation programme in Indonesia does not only enhance the ability of *farmers in the study area* to plan sustainable and efficient use of their resources but also assists soil conservation planners and policy makers (*in Jakarta*) to gain an understanding of resource management problems from *the farmer's perspective.* (Personal Experience).

9. We do not come to the discussion of how we obtain reasonable belief in a scientific hypothesis..... already knowing what we mean by such a "reasonable" belief and such a"valid" inference; in stating the conditions which justify inferences, we shall, *ipso facto*, be giving criteria which determine the meaning of the phrases "valid inference" and "reasonable belief" in the case of an inductively established hypothesis. (From H. Mortimer, 1988, after R.B. Braithwaite, Scientific explanation, 1953).

10. Einstein is right when he states that imagination is more important than information, but directed imagination needs the best information it can get. (J.K. Berry, 1993. Beyond Mapping. Concepts, algorithms, and issues in GIS, GIS World Inc, Colorado).

11. Only within the optimal environment, can men be well developed, and only a friendly environment leads men toward the optimal environmental development. (Otto Soemarwoto, 1985. Ekologi, lingkungan hidup dan pembangunan. Penerbit Djambatan, Jakarta).

12. Diversity of opinion about a piece of work shows that the work is new, complex and vital. (Oscar Wilde).

14. Perfect is the enemy of good. (Anonymous).

Abstract

Suryana, N., 1996, *A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units*. Doctorate thesis, Wageningen Agricultural University, Wageningen, The Netherlands, (xxvi) + 235 pp.

Three main aspects of the research, namely the concept of object orientation, the development of an Inductive Erosion Model (IEM) and the development of a framework for handling uncertainty in the data or information resulting from a GIS are interwoven in this thesis. The first and the second aspect of the thesis discuss simultaneously the application of a terrain mapping unit (TMU) in hierarhical observational procedures and an IEM in a GIS environment. These aspects were aimed at providing an alternative solution to the traditional approach to data acquisition, data capture and producing aggregated information for a GIS.

The third aspect discusses the application of standard deviation, probability of misclassification, membership degree and plausibility reasoning for handling error and uncertainty associated with data inputs and information outputs handled by a GIS in general and into and from the Indonesian Field Engineering Design Plan (FEDP) in particular. It is aimed mainly at establishing a framework for representing uncertainty in geographical data manipulation. GIS logical models, the characteristics of logical GIS models, types of uncertainty including error due to variability, imprecision, ambiguity and a proposed conceptual framework based on the concept of certainty factors are discussed.

The research involved the establishment of stable basic mapping units that allow the definition of repeatable and hierarchical observational procedures. This solution was addressed especially to the situation when sophisticated software and good quality data are not available. In this research, TMUs are defined as areas with a particular combination of geology, geomorphology, morphometry and soil characteristics, usually obtained by interpretation of aerial photo or SPOT images. Terrain areas having similar relief characteristics are identified, delineated and verified in the field. The delineated TMUs represent natural divisions of the terrain often with distinct boundaries.

Attributes associated with the established TMUs were selected and used to clasify TMUs. A classification hierarchy of TMU was established in the light of object oriented modelling including abstraction, inheritance, aggregation and association of terrain objects. The hierarchy has three levels, namely level +1 (superclass level refered to as TMU), level 0 (class level refered to as sub TMU) and level -1 (elementary object refered to as subsub TMU). A lower level in the classification hierarchy represents more refined or specialised information.

The well known deductive erosion model, the Universal Soil Loss Equation (USLE) is incomplete in predicting spatial erosion processes. More sophisticated models (i.e. *CREAMS, ANSWERS, EPIC, WEPP, GAMES*) have failed to account for the complexity of erosion processes and there are no means for validation of model predictions. An alternative to the problem is suggested through an inductive (bottom-up) approach. This approach involves an Inductive Erosion Model (IEM), which was built on observations including dynamic (resilience) and static (inertia) site specific erosion influencing factors in one or more sample areas, made on site at the farmer's field level which is the best functional unit to describe erosion class at local level. An IEM model

therefore is region specific. Once an IEM is built and tested for each type of TMU then it can be incorporated within the GIS environment as an acceptable means to predict safely the severity of soil erosion for the entire study area. Erosion severity classes predicted by an IEM are considered as active or dynamic attributes of the established TMUs. By definition TMU provides inherently erosion influencing factors, so called terrain characteristics including morphometry, geology, soil and ground cover. An IEM is intended to predict homogeneous erosion severity classes, related to TMUs at different aggregation or hierarchical levels. The aggregation levels are related to point observations, farmer's field level (FFL) and larger parts of the terrain. The discussion of this aspect is focused on the role of theTMU in the observational procedure providing input for an IEM.

The established hierarchical mapping units served as a basis for inductive erosion modelling, incorporating expert knowledge-based inference rules. The inductive erosion modelling followed a multi-scale approach and was implemented in a GIS environment. Application of the concepts of regionalization, observed pattern, and decision rules in predicting and modelling purposes are discussed. At regional level patterns associated with the main erosive processes such as sheet, rill, gully and ravine features are generally still identifiable on the aerial photos at scale 1 : 50 000. However, more detailed information on these types of active process at local level can be obtained only by more detailed study, i.e., erosion study at the FFL. In this regard, the FFL is considered as a suitable basic functional unit to describe erosion at local level.

Instead of using probability reasoning, which must follow statistical constraints, production rules allow the introduction of a Certainty Factor (CF) for handling both uncertainty in data, models and the resulting information. The CF can be obtained as a subjective judgment made by experts and comes naturally to experts either in inferring underlying processes or estimating quality of data and models being used. With special reference to the situation when all procedures and techniques for determining probability and obtaining quantitative information particularly in data poor environment are unlikely to be performed, this study demonstrated sufficiently the application of the concept of CF.

In the light of evidence theory, an IEM for predicting erosion severity at a specific TMU was built as a function of various certainty factors of spatial erosion influencing factors. The certainty factor has a value between -1 and +1 and its value indicates the estimated change in belief of allocation of a TMU to a particular erosion class as evidence (from maps, air photos, field observations etc.) is gathered, for each contributing factor. The erosion severity class to which a TMU is finally allocated is the one with overall certainty factor closest to +1. It is proposed as a method of handling uncertain information caused by incompleteness such as inferences established and derived by experts from a set of observations including the effect of causal relationships among various uncertain evidences.

Keywords: observational-functional units, inductive modelling, bottom-up approach, geographical information systems, observed patterns, phenomenological process, error and uncertainty, fuzziness, fuzzy measures.

Acknowledgements

Shortly after I returned in 1987 to Indonesia from studying in The Netherlands, I was assigned as a member of the team which was involved in preparing the technical guidance for the establishment of the Field Engineering Design Plan (FEDP). To increase the computerisation of the FEDP production particularly in establishing the timely FEDP, in cooperation between the Ministry of Forestry and the East West Center, Hawaii, I was also involved in testing the effectiveness of the IDRISI-GIS software for data acquisition, data processing and presenting resulting erosion severity classes and land suitability for dryland agriculture in the study area. The introduction of GIS into the FEDP was a very impressive technological breakthrough.

However, for a person who has worked for a considerable amount of time in using and enjoying the application of this relatively new method in rural or urban development in general and in establishing the FEDP in particular, some questions formulated by policy and decision makers may become of daily relevance. The fact that many organisations are using GIS before some important research questions have been answered has caused tension.

With three research questions in my mind, I took the difficult step of taking unpaid leave of absence from my organisation and I went back to The Netherlands to continue my studies. In September 1991, I started to conduct officially my doctorate research project at the Department of Geographic Information Processing and Remote Sensing, the Wageningen Agricultural University, Wageningen.

A research project of this kind could never have been accomplished without the contribution of many people. It is very hard to select names to be mentioned in this restricted context and the author deeply apologizes in advance to anyone who may feel unjustly omitted.

I would like to thank Prof. Dr. Ir. M. Molenaar, head of the department of Geographic Information Processing and Remote Sensing, the chairman of the Centre for Geographic Information Processing and also my promotor for his trust, opportunity, confidence, encouragement, constructive criticism and for his patience in supervising me to get through the most difficult period of my research. He was always very quick to get critically involved in many facets of my research. I still remember in the first year of my research he told me that my research topic appeared very interesting but it would be tedious and very long to get to the light at the end of a long and dark tunnel. I did not give up, because I believed deeply in his trust which was as important to me as my own energy.

I wish to thank my supervisor Prof. Dr. A.C. Imeson (from the University of Amsterdam, UvA), my promotor, who visited and guided me in the field and in developing an inductive approach in modelling soil erosion. His great interest in my research, constructive suggestions, enthusiasm and warm feeling always coloured every discussion. At the end I was always satisfied and enriched in confidence. He taught me to look always at the wider horizon of my research outside of the immediate topic. I find him not only to be a good scientist but also a person who can share his own bread. I wish also to thank his wife (Mrs. J. E. Imeson), my second professional English editor and also their youngest son (Marc Imeson) who looked after me and shared our good time together during our latest international working visits.

I wish to extend my appreciation to my supervisor Prof. Dr. Ir. R. Atmawidjaja, Deputy I of BAKOSURTANAL, the Rector of Pakuan University, Bogor, for his constructive suggestions and encouragement particularly during preliminary study, preliminary remote sensing image interpretation and the selection of my study area and encouragement. He always gave me confidence and comfort from the beginning until the end of my research.

I wish to thank my supervisor Dr. J.E. Drummond, now Senior Lecturer in the Department of Topographic Science and Geography at Glasgow University, Scotland, UK, for her field work supervision, understanding of the soil erosion process and its measurement (outside of her field of expertise), her patience, her motivation and dedication in assisting the developing world. Often I was amazed by her scientific performace, loyalty and generous advice. She inspired me to struggle with a situation when I was not sure whether GIS was an appropriate technology or only a passing fashion. At the end this became the essential part of my research topic.

I would like to thank Prof. Dr. Ir. L. Stroosnijder and Ir. L.A.A.J. Eppink of the Department of Irrigation and Soil and Water Conservation, LUW, for guiding me particularly in the early and last stages of my research. They were very helpful in providing me generously with their time and to put me in their own tight schedules. I also thank friends and colleagues in the United Nations Institute for Training and Research (UNITAR), Geneva, Switzerland, and in the Department of Geographic Information Processing and Remote Sensing, LUW, who have supported and encouraged me, especially Marcel Boisard, Stephen Gold, Christope Nuttal, Jan Hein Loedeman, Dick de Wit, Wim Droesen, Hans van Leeuwen, Sylvia de Hoop, Elisabeth Addink, John van Smaalen, Sytze de Bruin, and the departmental secretary, Ms. Anke Hoeneveld. Several staff members of the ITC, Enschede, especially Prof. Dr.Ir. A. Meijerink, Ir. E.Bergsma, Dr. C. Valenzuela, my Indonesian colleagues particularly Mr. Hermawan Indrabudi, Mr. Tetra Yanuariadi and members of the Cartography Division, have assisted me at various stages. I wish to thank them all.

I wish also to thank members of the Alpine Geomorphology Research Group, the Physical Geography and Soil Science Laboratory, UvA, Amsterdam, especially Prof. Dr. Ir. P. Jungerius, Dr. J. Rufke, Dr. H. Sijmonbergen, Drs. L.W.S. De Graaff who taught me how to map the complexity of the terrain of the Hintere Bregenzerwald Area,Vorarlberg, Austria, and also Dr. G.B.M. Heuvelink for their constructive suggestions and generous discussions in uncertainty aspects of GIS.

I wish to thank the Indonesian Ministry of Forestry (MOF) and the Sub Balai RLKT DAS Citanduy-Cisanggarung who gave me permission and logistical support during my field work periods in Indonesia. I would like also to thank the Director of the International Centre for Education and Research and the Sector of Landuse and Environment, LUW, as well as the Board of Governors of theVan Deventer-Maas Stichting, Den Haag, especially Mr.W.M. Fredriksz, for their trust and financial support which made my stay in The Netherlands much more comfortable.

I wish to thank my house mates, Allan Brown, MA and Drs. Connie Blok, both lecturers in the Cartography Division, Department of Geoinformatics, ITC, Enschede and my cat Putih for sharing the comfort of my roomy house in Enschede. They and my house together created a supportive spirit which sustained me through the frustrating periods of my research, and contributed to my enjoyment when things were going well. For their motivation, affection and loyalty I am morally indebted. Connie Blok taught me to maintain a balance between social life, career and private life. Allan Brown, my first English editor has very deep insight and understanding into all aspects of my research. This made it easier for both him and me to get through my thesis. Therefore, correcting my manuscript was done perfectly patiently. Also, he understood my difficulties in mastering English as a foreign language. So, I forgave him his joke about the typical Indonesian difficulty with 'f' and 'p' ('ploffy disk', for example). I was lucky to be adopted by my cat Putih. I often talked to her while I burned the midnight oil at my computer. When she jumped on the keyboard I used that as a sign to have a rest.

I must give thanks to my beloved Dra. Ida Badriah Adjidji and my beloved son Ryan Aditya Perdana who have been waiting patiently for the completion of my research and who gave me inspiration, encouragement, moral and spiritual support from thousands of miles away. It was a very difficult period of my life and I missed them such a great deal.

Finally I must give thanks to my entire family and my parents particularly my mother who taught me to look always beyond the immediate horizon.

Wageningen, September, 1996.

Nanna Suryana

Contents

Abstract vii
Acknowledgements ix
Figures xix
Tables xxi
Equations xxiii
Executive summary xxv

1 Introduction

1.1 The deductive versus the inductive approach alternative strategies for modelling soil erosion in the humid tropics 3
- 1.1.1 Erosion problems and deductive erosion prediction models
- 1.1.2 An inductive strategy for modelling soil erosion
- 1.1.3 The necessity of a stable basic mapping unit for data acquisition, data capture and representing information

1.2 Thesis organisation 12

CONCEPTUAL FRAMEWORK

2 Soil erosion modelling and inductive reasoning based on observations in terrain mapping units

2.1 Soil erosion as a spatial and phenomenological process 13
2.2 Spatial erosion models : current situation 14
- 2.2.1 Empirical lumped models
- 2.2.2 Empirical distributed models
- 2.2.3 Conceptual lumped models
- 2.2.4 Conceptual distributed models
- 2.2.5 An evaluation of existing physical based erosion models
- 2.2.6 Soil erosion modelling using computer-based image processing

2.2 Inductive reasoning based on TMU 20

3 Terrain object description in Geographical Information Systems

3.1 GIS, geographical and spatial data defined 25
3.2 Formal data structure 26
3.3 Thematic attributes of terrain objects 28
3.4 Geometric attributes of terrain objects 28
3.5 Classification and aggregation hierarchy of terrain objects 29
3.6 The dynamics of terrain objects 30

4 An approach in handling uncertainty and ambiguity in procedure in Geographical Information Systems

4.1 An overview 33
4.2 Types of uncertainty and error in GIS 34
4.3 Attribute accuracy assessment associated with non-continuous data 35
4.3.1 General view
4.3.2 Method and techniques used
4.4 Attribute accuracy assessment associated with continuous data 39
4.4.1 General view
4.4.2 Method and techniques used
4.5 Plausibility reasoning and the inductive approach 46
4.5.1 General view
4.5.2 Notation of the plausibility reasoning
4.5.3 Discussion

CASE STUDY

5 Study area and datasets used

5.1 General description of the study area 49
5.2 Datasets used 57

6 The establishment of functional observation units: TMU classes for describing terrain object characteristics at the sub-regional level

6.1 The establishment of TMUs: the application of aerial photo and SPOT image interpretation 59
6.2 TMU defined for data acquisition, data capture and observation procedure 62
6.2.1 Introduction
6.2.2 TMU defined
6.3 The identification of homogeneous relief TMUs and spatial distribution of accelerated erosion at sub-regional scale 66
6.3.1 Introduction
6.3.2 Spatial and specific patterns associated with TMUs
6.3.3 Conclusion
6.4 The differentiation of classes of TMUs at sub-regional level 69
6.4.1 Method used
6.4.2 Results and discussion
6.5 The repeatability of the boundary of TMUs 74

7 The hierarchical method for data acquisition and for representing erosion information at local and subregional level

7.1 General view 77
7.2 The observational role of the TMU in identification of soil erosion influencing factors 78
    7.2.1 Introduction
    7.2.2 The observational characteristics as the dynamic and static properties of TMUs
7.3 Hierarchical structure for TMUs 81
    7.3.1 Class and aggregation hierarchies of TMUs
    7.3.2 Structuring of the units
    7.3.3 The interpreted TMUs in the study area
7.4 The relationship between TMU and the proposed IEM in erosion modelling at local and sub-regional levels 85

8 Data requirements, methods and techniques used in data acquisition and in building an Inductive Erosion Model

8.1 Introduction 89
8.2 The identification and the definition of an Inductive Erosion Model 89
    8.2.1 Assumption and approach
    8.2.2 The identification of an IEM
    8.2.3 Defining and synthesing an IEM
    8.2.4 Dataset requirements for an IEM
    8.2.5 Selection of sample areas
8.3 Erosion Survey at the FFL and observed patterns 97
    8.3.1 Introduction
    8.3.2 Method and techniques used
    8.3.3 Data analysis
8.4 Results and observed patterns in the study area 100
    8.4.1 Erosion classification under different cover types
    8.4.2 Erosion features under the cover types
    8.4.3 The use of indicators for erosion assessment
    8.4.4 Soil loss estimations
    8.4.5 The relationship between erosion and locational attributes
8.5 Testing and developing an Inductive Erosion Model 112
    8.5.1 Introduction
    8.5.2 Methods and techniques used
    8.5.3 Results
8.6 The incorporation of an IEM into a Geographical Information Systems environment 114
    8.6.1 Introduction
    8.6.2 The incorporation of induction task (rules) into a GIS
    8.6.3 Class generalisation and object aggregation adopted by an IEM
8.7 Conclusions 119

9 The attribute accuracy aspects of the interpreted aerial photo TMUs used in the FEDP for the land rehabilitation and soil conservation programme in Indonesia

9.1 Introduction 121
9.2 Data sources for data input, data capture and error in GIS 121
9.3 Uncertainty aspect of GIS data inputs 122
9.4 Quality assessment of existing categorical maps: non-continuous data type 123
 9.4.1 The application of external testing
 9.4.2 The application of Kappa statistic
 9.4.3 The application of internal testing and the percentage correctly classified at stated confidence level
9.5 Attribute accuracy assessment associated with classified continuous data 127
 9.5.1 Introduction
 9.5.2 Practical examples
9.6 Conclusions 132

10 Handling uncertainty-ambiguity associated with an Inductive Erosion Model in Geographical Information Systems

10.1 Introduction 133
10.2 An approach to the model quality assessment associated with an IEM 133
 10.2.1 General view
 10.2.2 Plausibility reasoning and an IEM
 10.2.3 Synthesizing an IEM in the light of plausibility reasoning perspective
 10.2.4 Certainty Factor
 10.2.5 Practical example
10.3 Conclusions 148

11 Conclusions and recommendations

11.1 Conclusions related to the establishment of the hierarchical and observational procedure using TMU 149
11.2 Recommendations relating to the first research objective 150
11.3 Conclusions related to building and developing the proposed IEM in a GIS environment 151
11.4 Recommendations relating to the second research obejctive 152
11.5 Conclusions related to handling uncertainty in a GIS environment 153
11.6 Recommendations relating to the third research objective 154
11.7 Final conclusions 155
11.8 Final recommendations 155

Appendices

1 Mathematical equations associated with empirical and process-based erosion model 157
2 Land use classification as used by RMI 163
3 Landsat image-based terrain classification as used in Indonesia 165
4 Examples of selected sample areas for building an Inductive Erosion Model 167
5 Erosion survey form 171
6 Data analysis for building an Inductive Erosion Model 177
7 Erosion classification at FFL 187
8 Pascal program for erosion hazard assessment using an IEM 207
9 The prototype uncertainty subsystem 209

Abbreviations and glossary 217
References 221
Samenvatting 231
About the author 234

Figures

1 The establishment of unit lahan adopted in the existing FEDP 9
2 General methodology of the study 11
3 Current strategy for erosion modelling 15
4 Inductive chain of reasoning 22
5 Levels in data modelling 26
6 Formal Data Structure for GIS 27
7 Classification and aggregation of terrain objects 29
8 Fuzzy boundaries of ranked interval 42
9 Overlap probability in a ranked interval 43
10 The three possible cases of ranked interval probability 45
11 Location of the Ciseel Subwatershed 49
12 Transport and annual rainfall in the Ciseel Subwatershed 51
13 Erodibility of surficial geologic materials in the Ciseel Subwatershed 54
14 General landuse and ground cover in the Ciseel Subwatershed 56
15 TMU identification procedure 60
16 (a) Aerial photo interpretation (API) procedure adopted in the establishment of TMUs. (b) Traditional aerial photo interpretation 61
17 TMUs as terrain objects described by thematic and geometric attributes 63
18 Escarpment slope facets of a particular terrain system 64
19 Data acquisition techniques 65
20 Example of specific patterns 68
21 Diagram of separability of TMUs according to slope steepness (%) 72
22 Means of slope steepness (%) mapped into slope length (m) using Scheffes' F test 73
23 TMU boundaries produced by UNPAD and RMI 76
24 (a) Attributes defined at class level and evaluated at object level. (b) Determinant order related to different aggregation levels 82
25 General relationship between the TMU and the proposed IEM 86
26 Stepwise induction steps at different aggregation levels 87
27 The syntax of data models for reprsenting TMUs at different aggregation levels 88
28 The inductive erosion modelling procedure 90
29 Eroded and compacted soil surface 95
30 Pedestals 95
31 Field (sheet) erosion indicators 96
32 The relationship between the estimated and observed erosion class 104
33 (a) The relationship between K-factor and observed erosion; (b) The relationship between K-factor and predicted erosion by the USLE 105
34 (a) The relationship between SL-factor and observed erosion; (b) The relationship between SL-factor and predicted erosion by the USLE 106
35 The relationship between bare soil and C-factor 107
36 The relationship between C-factor and observed soil erosion 108
37 Induced erosion classes in the study area 110
38 Decision tree adopted by an IEM 116
39 Erosion generalisation strategies adopted by an IEM 118

40 Erosion severity classes predicted by the USLE 143
41 (a) Induced erosion severity class: Severely eroded and CFs 144
(b) Induced erosion severity class: Moderately eroded and CFs 145
(c) Induced erosion severity class: Slightly eroded and CFs 146
(d) Induced erosion severity class: Very slightly eroded CFs 147
42 Soil accumulation estimation 179
43 Influence of vegetal canopy height and cover percentage on kinetic energy of throughfall per unit rainfall, assuming bare soil underneath the canopy (Dissmeyer and Foster, 1981). 182
44 Soil reconsolidation effects on erodibility in the form of a reconsolidation factor (Dissmeyer and Foster, 1981) 183
45 Effect of fine roots in top 30-50 mm of soil erodibility, in the form of a fine root subfactor value (Dissmeyer and Foster, 1981) 183

Tables

1 Typical error matrix 37
2 Geologic composition of the Ciseel Subwatershed 52
3 Land units with dominant slope 52
4 Soil types classified by geologic origin and soil erodibility 53
5 Landuse classes in the Ciseel Subwatershed 55
6 Units of analysis within a hierarchy of terrain systems 63
7 Relief attributes associated with TMUs 70
8 Slope steepness classes on the basis of individual measurements 70
9 One-way analysis of variance of slope steepness attributes 71
10 Regrouped and rescaled slope steepness classes 72
11 An example of area difference between two corresponding TMU maps 75
12 Observational characteristics of terrain mapping units 79
13 TMU determinant order 80
14 The interpreted TMUs in the study area 84
15 Cover percentage of the main unit types in the study area 85
16 The occurrence of erosion classes under different cover types 100
17 Erosion features under different cover types 102
18 Observed and predicted soil losses stated in mm/year 104
19 An observed pattern and locational attributes at observation point 13 111
20 Error matrix of an IEM compared to the visually observed (true) class 113
21 Error matrix of the USLE compared to an IEM class 114
22 The error matrix of the present landuse 124
23 The result of internal testing of 100 randomly selected observation points on 10 landuse/landcover maps 126
24 Slope steepness and interval class 128
25 Observation points of the higher accuracy (control) map 128
26 Observation points of the FEDP (test) map 129
27 The probability and the membership degree of observations falling around the lower boundary of the class 130
28 The probability and the membership degree of observations falling around the upper boundary of the class 130
29 The probability and the membership degree of observations falling within the kernel of the class 131
30 Typical occurrence likelihood of erosion and corresponding CFs 137
31 Vegetation cover, soil conservation measure (in Java) 139
32 Slope steepness value 140
33 Average slope length overland flow 140
34 Soil type, soil erodibility 141
35 Example of CFs at observation point 13 142

Equations

1	The equation of the USLE	5
2	Probability of correct classification estimates for entity (class) level	38
3	Overall attribute accuracy estimate for coverage level	38
4	Kappa statistic	38
5	Expected probability due to chance	38
6	The lower and upper limits of the confidence interval	39
7	Standard deviation	41
8	Overlap probability	44
9	Interval discrimination	44
10	Probability of correct ranking for observations falling within the dispersion of the lower boundary and the kernel of the class	44
11	Membership degree of correct ranking for observations falling within the dispersion of the lower boundary and the kernel of the class	45
12	Probability and membership degree of correct ranking for observations falling within the kernel of the class	45
13	Probability of correct ranking for observations falling within the dispersion of the upper boundary and the kernel of the class	45
14	Membership degree of correct ranking for observations falling within the dispersion of the upper boundary and the kernel of the class	45
15	The proportionate decrease in disbelief MB (h,e)	47
16	The proportionate decrease in belief MD (h,e)	47
17	Change in belief	47
18	Erosion severity classes	92
19	Certainty Factor (CF) Model	135
20	The updated parallel combination	135
21	The rainstorm parameter	157
22	The total storm energy	157
23	The computed soil loss per unit area for the selected season (GAMES)	157
24	The seasonal delivery ratio	158
25	The number of downslope fields	158
26	The sediment load of CREAMS	158
27	The distance downslope of WEPP	159
28	The rill erosion rate of WEPP	159
29	The detachment capacity flow of WEPP	159
30	The rill erosion rate (if sediment load greater than transport capacity)	160
31	The total runoff volume	160
32	The rainfall intensity	160
33	The rainfall detachment rate	160
34	The deposition rate of size range class i	161
35	The rate of sediment entainment of size range class i	161
36	The rate of sediment movement in flow	161

37 The rainfall impact detachment rate 162
38 The overland flow detachment rate 162
39 The transport capacity of flow (1) 162
40 The transport capacity of flow (2) 162

Executive summary

The topic of this doctorate thesis is ***"A geo-information theoretical approach to inductive erosion modelling based on terrain mapping units"***. This research was conducted under the framework of the Field Engineering Design Plan (FEDP) for Land Rehabilitation and Soil Conservation Programme, in Indonesia. It was intended to investigate the needs of representing information on spatial or geographic entities at different scales, and also the need to introduce inference facility and uncertainty analysis into a Geographical Information System (GIS).

This thesis explores some interlinked research topics contributing to the building and developing of a rule-based Inductive Erosion Model (IEM) in a GIS environment as an alternative to the USLE (*general objective 1*); the establishment of functional observation units for observing terrain characteristics from which erosion classes can be inferred at different aggregation levels (*general objective 2*); establishing a framework for handling uncertainty associated with a GIS environment (*general objective 3*). In order to achieve these objectives, this research included three phases of study, namely relevant literature research, field work and post fieldwork conclusions.

The literature research was intended to build a strong foundation for obtaining clearly defined concepts. The results are formulated and presented in the form of conceptual frameworks. Furthermore, problem identification and problem definition were formulated. In order to obtain clear ideas related to the bio-physical characteristics of the study area, the available soil, geology, geomorphology, and land use maps from previous studies have been studied and at the same time preliminary aerial photo and satellite imagery interpretation has been completed.

The fieldwork stage was conducted in the Ciseel Subwatershed, West Java, Indonesia. This subwatershed is classified as a vulnerable subwatershed particularly in terms of erosion problems. The fieldwork was conducted in two parts. The intention was to conduct an erosion study at the farmer's field level and to perform an erosion classification adopting erosion features and a semi quantitative approach. In addition to this, the robustness of the proposed Inductive Erosion Model (IEM) was validated.The post fieldwork activities were concentrated on tabulating field data, analysis, mapping, presenting and report writing.

Three main themes or aspects are interwoven in this thesis. The first aspect is the application of the concept of object orientation, using Terrain Mapping Units (TMUs) as objects for data acquisition and producing aggregated information in GIS. The second aspect concerns the development of an IEM and its implementation in a GIS environment. The third main theme is the development of a framework for handling uncertainty in the information resulting from a GIS, caused by uncertainty in the original data inputs and in an IEM model itself. The main results of the research are set out below, organised by theme.

1. The role of the interpreted terrain mapping unit

This aspect of the thesis is concerned with the establishment of the interpreted terrain mapping unit (TMU) used as a stable basic mapping unit for effective and rapid data collection, observation procedures and representing understandable outputs at different aggregation hierarchies. These hierarchical levels associated with TMUs are related to the linkage between observation points, farmer's field level (FFL) and larger parts of the terrain. The aim is to find an observational strategy or hierarchical method for data acquisition, data capture and producing information categories at different aggregation levels with special reference to the situation when good quality data (single thematic maps) at the desired scale are not available.

Although a TMU is only a small part of the terrain, it can nevertheless be a very complex system. By examining at aerial photos at scales 1 :50 000 and 1: 15 000, a particular TMU can be divided into smaller and more detailed units, say Sub TMU and Sub-sub TMU respectively. All the TMUs possess a direct topological (connectivity) relationship. This means that terrain characteristics associated with TMUs can be aggregated (generalised) and disaggregated (specified). In other words the TMU, as the basic mapping unit, can be put at different orders or levels of the classification and aggregation hierarchy. The nature of the TMU, as well as the interpretation process allow the identification of processes, active in the terrain. Therefore, the TMU is considered as a functional unit in describing terrain characteristics, e.g., spatial distribution of soil erosion at local and regional level.

2. The role of an Inductive Erosion Model

Using inputs from the first part of the research, i.e., information on the spatial distribution of soil erosion, an IEM is established. This model adopts the concept of induction, i.e., reasoning from a set of specific premises obtained through field observation to a general conclusion. This is intended to be an alternative solution to problems associated with deductive (i.e. reasoning from a general premise to specific conclusions) erosion hazard modelling in humid tropical countries in general and in the generation of the computerised Field Engineering Design Plan (FEDP) in particular. The proposed IEM is intended to predict the occurrence of soil erosion at different hierarchical levels which are related to observation points, FFL and larger parts of the terrain. Considering factors related to soil erosion (soil erodibility, slope steepness etc.), attached inherently to the main units, subunits and subsub units, and using a unique identifier, these units can be related to other erosion related factors, e.g., land use/land cover, after which the assessment of erosion severity can be performed. Adopting generalisation and classification rules, these units can be used to describe erosion severity information at different levels. In other words, the TMU provides terrain characteristics, e.g., factors related to soil erosion and bio-physics, which can be used to describe soil erosion at different levels of spatial detail. Information on the spatial anthropic soil erosion categories, was obtained by performing a soil erosion classification at the farmer's field level. The classification was established using a semi-quantitative approach that included the identification of soil erosion features, e.g., sheet, gully, rill erosion. Some indicators representing type of soil erosion, e.g., pedestal, sealing, root exposure, soil accumulation etc. have been used to derive the erosion class. By combining these erosion features the erosion class at the farmer's field level was established. At the same time, at this level, all related influencing factors are observed, identified and measured systematically.

From each observation point, observed patterns i.e. the relationship between soil erosion and related influencing factors were obtained. Several observed patterns at this level may produce a similar class of erosion. This induction process will be explained in a formal way in Section 2.3. For example the slightly eroded class may be generated by one or more possible combinations of formation related factors. This leads us to find the most proper combination. The separation of these combinations (observed patterns) was made by applying a decision regression tree developed by Ward (1963), Stanat and McAllister (1977). At the end the best combinations were selected.

This aspect of the research is also intended to contribute to increasing the additional value of a GIS in handling special tasks, thus increasing the usefulness of a GIS. This was achieved by adding (incorporating) the inductive approach and experts' inference into a GIS, formulated in the form of a set of decision rules and used as a means of predicting erosion severity in other uniform geographical units.

3. The role of GIS and associated problems

A geographic entity is an identifiable object of which the location on the earth, at any point in time can be identified and described through establishing spatial or geographical referencing (Goodchild and Gupta, 1988). Some examples of geographic entities may include description of a drainage network within a watershed, the distribution of hydrological processes or environmental characteristics . In this study, these entities are referred to as data classes.

With regard to the description of a geographic entity, the introduction of GIS into planning and decision making processes offers some advantages. The system provides fast and timely data processing, analysis and presentation of the resulting information. A computerised GIS system allows one to inventorise these entities in order to produce newly defined information categories used at different levels of the planning and decision making processes. In short, GIS facilitates the elimination of bottlenecks or technical problems of organising and integrating spatial data and for implementing complex spatial analysis which have been faced and previously could not be solved (de Man, 1988).

Because of the complexities of the environment, and due to fuzziness of the concepts, technical limitation and human errors during data acquisition and data processing, information resulting from handling by GIS has limited precision and accuracy, referred to in terms of uncertainty and errors.

An IEM as developed in this research, classifies predicted soil erosion into four erosion severity classes, namely very slightly eroded, slightly eroded, moderately eroded and severely eroded. These typical linguistic terminologies imply the divisions between these adjacent classes, e.g., very slightly to severely eroded, are not sharp, but consist of transition zones. In other words, these terminologies are referred to as *unsharply defined or fuzziness* of the concept. They are developed on the basis of *the presence* or *absence* of soil erosion in a particular TMU. With regard to building an IEM, the terminology of presence and absence are referred to as *crispness* of the concept. This has been established by identifying and quantifying erosion features, e.g., sheet, gully erosion and performing soil erosion classification at farmer's field level. Thus, an IEM concerns both the crispness and the fuzziness of erosion severity classes. It is considered that an IEM involves ambiguity different to the term fuzziness as applied normally in another logical model, namely a crop suitability model. In this regard, an IEM deals with the selecting of several hypotheses, e.g., area A is very slightly, slightly, moderately or severely eroded. But we are only allowed to choose one hypothesis which is assumed to be correct. The inapplicability of the fuzzy set/fuzzy subset theory in erosion modelling becomes more clear after looking further at the process of soil erosion.

In operational conditions the process of soil erosion can be understood as a phenomenological process which takes the causal relationship between soil erosion influencing factors and soil erosion processes into account. In other words, the severity of soil erosion is not only determined by each single soil erosion formation factor but it is determined by the combination or interaction amongst these influencing factors.

Considering this, each soil erosion formation factor works differently. One factor may accelerate the process and the other may retard the process. Therefore the combination of factors involved takes a very important role in determining the process of soil erosion. On this basis the application of minimum and maximum rules associated with fuzzy operation is not consistent with the erosion process.

4. The role of reasoning with uncertainty

The TMU and IEM research results have been used as inputs to the third aspect of the research which is aimed at establishing a framework for handling uncertainty (errors) associated with data inputs and resulting erosion hazard information processed using an IEM. The present GIS generation does not yet include explicitly either uncertainty analysis , e.g., fuzziness modelling or inference capability.

Data input into the GIS and into FEDP in particular includes existing topographical maps, categorical-coverage maps, visually interpreted imagery, digitally remote sensing image classification and direct field measurements. As stated above when these data are inserted into a GIS or the FEDP they are not error free but include uncertainty. Moreover the uncertainty is propagated in the model used.

For *numerical data* such as a slope steepness map, usually the standard deviation (*sd*) has been used to represent uncertainty, particularly as a measure of the precision and accuracy of numerical data gathered from direct field measurement or interpolation procedures. However, this measure is not applicable to the typical *categorical coverage* land use and soil map resulting from interpretation. Considering this, the quality assessment of data input has been done according to data types, i.e., *continuous and discontinuous*, as used as inputs to the generation of the FEDP.

The uncertainty associated with the continuous data type has to do with the assignment of a set of observations to a class. The basic premise is that divisions between adjacent classes are not sharp, but consist of probability transition zones. In the central part of each class, the kernel, there is no doubt that an observation belongs to that class. However, due to the fact that there may be some unreliability in the observations, e.g., due to measurement errors or stong spatial variations around the observation point, observations with values near the class boundaries have likelihood of actually belonging to the adjacent class. Within each class, therefore, observations may fall in one three zones: within the kernel (*case1*), between the lower class boundary and the kernel (*case2*) or between the kernel and the upper class boundary (*case 3*). The assessment of this likelihood is based on formal statistics including the normal distribution and probability theory, using standard deviation (sd) as a measure of dispersion. Then, the overlap probability, and the probability of correct ranking and the membership degree of observations can be determined.

Concerning an IEM itself and how it is built and operated, this model can be classified as a *logical model* (Suryana, 1993). Thus it is not constructed on the basis of mathematical formulae as in the deductive, traditional, top down approach of the Universal Soil Loss Equation (USLE) erosion model. The uncertainty associated with the mathematical model can be handled using the error propagation technique (Heuvelink, Burrough and Stein, 1989). This technique, however, is not applicable to a logical model such as an IEM.

Considering the ambiguity and the process of soil erosion as described above, the uncertainty associated with an IEM in a GIS environment can be approached using plausibility reasoning associated with evidence theory. This approach is well known as the Certainty Factor (CF) model and concerns maintaining and rejecting an original hypothesis on the basis of given different evidences.

To combine evidences to yield a single combined evidence, this approach is formulated in the form of parallel combination which maintains the concept of linearity and associativity. By this is meant that the order of observation should not affect the final result.

Chapter 1

Introduction

In Indonesia the problems associated with land degradation due to the use of land for the wrong purpose particularly when hilly or mountainous areas are opened up and converted for annual crop cultivation have been identified as a continuous problem from one generation to the next. The environmental effect of rapid change of land cover in the study area of this thesis has been easily and frequently observed by using indicators, e.g., Q (debit) maximum and minimum ratio of the river. This value shows the occurrence of flooding or drought during and after the southwest and northeast monsoons. These disturbed hydrological indicators draw attention to the importance of soil conservation practices in sloping agricultural areas.

In response the Indonesian Government has implemented the Land Rehabilitation and Soil Conservation Programme. For this, the Indonesian archipelago has been divided into more than one hundred watersheds. In the years 1983/1984 the Indonesian Government selected 17 priority and 22 *"superpriority"* watersheds throughout the country. The selection was based on defined criteria, e.g., erosion rate per year per hectare. *"Superpriority "* implies that these watersheds have to be handled and developed completely by the end of the fifth Indonesian five year development plan. In this regard, two types of complementary plans, namely the management plan (*at watershed level*) and the field engineering design plan (FEDP, *at subwatershed level*) have been introduced.

The FEDP is *a detailed medium term plan* for land rehabilitation and soil conservation. Generally the establishment of the FEDP consists of two main activities, namely erosion severity mapping and providing the priority list of recommended soil conservation measures and soil conservation controls for specific areas within each particular subwatershed. The erosion severity classes are described quantitatively with respect to the estimates of the amount of soil loss per year per hectare. Based upon these erosion severity class estimates and the actual soil depth at specific sites, the priority list of recommended soil conservation measures and soil conservation controls are derived. Using the assumption that an erosion severity class linked to a particular combination of erosion influencing factors does exist at a specific site, the traditional establishment of the FEDP adopts *"unit lahan (land unit)"* and the traditional, top-down, deductive erosion model, the Universal Soil Loss Equation (USLE), to estimate the amount of soil loss and to assess erosion severity classes. However, the developer of the USLE has himself warned that the equation may not be suitable in the highly diverse conditions typical of the humid tropics (Wischmeier, 1976). There are data acquisition problems and the results of the model may be unreliable.

These problems become obvious when attempting to computerise the FEDP, by means of the introduction of a PC-based Geographic Information System (GIS). The problems encountered with regard to the development of an alternative strategy in modelling soil erosion gave rise to the following practical issues, which this thesis addresses:

(i) *What observational units should be used for stable, repeatable data acquisition which can maintain also the representation of information at different scales?*
(ii) *How robust and reliable is the model currently used in a GIS?*
(iii) *How to handle and represent data or information quality if the results may be doubtful?*

Questions such as those above are related to the specific performance of a GIS, e.g., its application in an erosion study, which is determined by the availability of good quality data, the reliability of the model used and a method to represent the information categories. The third question is related specifically to the uncertainty associated with the resulting information which is determined or generated by the models used (2nd question) and data inputs gathered from observation units at different levels (1st question). However, the data quality aspect of geographical data can not be discussed without looking at how data are collected and processed by the model used. Thus, these questions in fact are inseparable and interwoven.

GIS users are aware that *"uncertainty"* is important in relation to the consequences of making any wrong decision on the basis of wrong information (Molenaar, 1991). For example it can be very costly for farmers if they are recommended to abandon existing practices in favour of new ones, only to find that the new ones do not produce the expected results because of a mismatch between the real (as distinct from the perceived) situation and the recommended solution. It is necessary to obtain better knowledge of uncertainty in data and in the model used to handle the data. More important is to know how this uncertainty can be handled and represented to GIS users. Clearly, knowledge of the data and model quality would be an advantage particularly for a GIS user who may on the basis of this quality information become aware of unforeseen consequences and be able to provide proper decision options.

Current generations of GISs do not yet explicitly include the facility to develop *"inferences"* on spatially varying relationships between features and their attributes and the facility to produce information on the quality of resulting information. An example is the quality of information associated with the presence and absence of accelerated (anthropic) soil erosion in different parts of a subwatershed.

A typical technique much used in GISs is the overlaying of areas (polygons) which are assumed to have one or more similar terrain characteristics. This technique requires sharply defined or repeatable units of observation. When used with the USLE, areas have to be defined for all of the parameters of the equation. Consequently, the data input requirements are very demanding. Furthermore, the accuracy of the model prediction decreases if small unit areas are aggregated into larger areas, since the larger the area the greater the internal variability. Yet overlaying techniques have been used in GISs, based on the USLE, without considering this aggregation or hierarchy problem.

This thesis addresses the above problems, and has defined the following objectives:

(i) the building and developing of a rule-based Inductive Erosion Model (IEM) in a GIS environment as an alternative to the USLE;

(ii) the establishment of functional observation units for observing terrain characteristics from which erosion classes can be inferred at different aggregation levels;

(iii) establishing a framework for handling uncertainty associated with a GIS environment.

In order to achieve the defined objectives, special research attention is given to:

(i) identifying, defining and analysing an IEM model requirements; identifying, quantifying, classifying and establishing observed patterns (accelerated soil erosion class) and associated locational terrain attributes at Farmer's Field Level (FFL); establishing and inducing a set of decision rules and mapping the spatial distribution of observed patterns into a larger area;

(ii) the definition of Terrain Mapping Units (TMUs) as functional observation units; the identification of observational characteristics of TMUs; the differentiation of TMUs to set up an aggregation and classification hierarchy of TMUs in describing soil erosion at local and sub-regional level;

(iii) setting up methodology for representing parameters expressing the quality of information handled by and within a GIS; evaluating the application of plausibility reasoning in handling the ambiguity associated with the expert's inference model, an IEM, in modelling soil erosion; developing a prototype uncertainty subsystem within a GIS.

The achievement of the above objectives in a GIS environment requires three different interlinked approaches. The first approach is related to *the inductive learning procedure* which is used for inferring soil erosion severity classes. The second is the *geoinformation theoretical approach* which is used to support the establishment of a hierachical procedure and the incorporation of inference rules at different aggregation hierarchies. The third requirement is the formalisation of the fuzzy reasoning techniques including *standard deviation, probability misclassification, membership degree and evidential theory or plausibility reasoning* which are used to assess the quality of data and information produced by the first and the second approaches.

The research study area is the Ciseel Subwatershed, West Java. This subwatershed is an intermontane valley and is classified as a vulnerable subwatershed particularly in terms of erosion problems (Ditkontan, 1985). The field study was conducted in two parts, namely selection of sample areas and building and developing the proposed strategy for modelling soil erosion. The fieldwork was aimed at (i) conducting an erosion study at the farmer's field level; (ii) performing an erosion classification using erosion indicators and a semi-quantitative approach; (iii) building and testing an IEM. With special reference to data poor environment, it is expected that the establishment of a hierarchical procedure of observation and the incorporation of a rule-based erosion model and uncertainty analysis will lead to the improvement of erosion hazard assessment for any specified use in the humid tropics in general, particularly when used in a GIS. The achievement, or otherwise, of the research goals is examined through the eleven chapters of this work as discussed briefly in Subsection 1.2.

1.1 The deductive versus the inductive approach as alternative strategies for modelling soil erosion in the humid tropics

In the introduction above, three research questions related to hierarchical method of observation, current erosion modelling and uncertainty analysis within a GIS, have been formulated. Subsections 1.1.1 and 1.1.2 below discuss problems associated with erosion and with deductive erosion prediction models for the humid tropics implemented in a GIS as well as problems associated with the present GIS generation facililities for handling special tasks and uncertainty analysis in GISs. Subsection 1.1.3 includes a discussion on the establishment of the observation procedure as well as the importance of the stability or repeatability of basic mapping or observational units that can be used in data acquisition and the representation of information at different scales.

1.1.1 Erosion problems and deductive erosion prediction models

The Indonesian soil types are susceptible to erosion and productivity decline (Ambar and Syafrudin, 1979). It is the consensus that soil erosion by water is one the most widespread causes of land degradation. The study of soil erosion has become an issue of considerable importance (Ditkontan, 1985). If this process is not properly controlled it may bring about progressive decline in soil productivity, sediment and nutrient accumulation in streams, lakes and reservoirs, and cause off-stream effects, such as flood damage (Soemarwoto, 1985; Clark, 1985). Many soils in the study area have a shallow effective depth beyond which roots cannot penetrate (UNPAD, 1986). As these infertile layers are progressively brought nearer the surface by accelerated erosion and as the organic-nutrient content and fine particles of the rich top soil are washed away and the microclimate of the soil is altered, the water and nutrient retaining capacity, biomass and productivity inevitably decline (Imeson, 1984; Eppink, 1985; Nortcliff, 1986). Thus, soil erosion is inseparably linked to soil productivity (Stocking, 1984).

That erosion causes a loss in productivity is suggested from various sources. Direct effects are clearly seen in the physical, chemical and structural nature of the soil (Imeson, 1984). In Indonesia, it appears that the decline in productivity is directly related to the severity of erosion. This link has been studied in experiments in which it was found that eroded soil contained twelve times the nutrients found in the remaining soil (Suryana,1981). Maize has shown a 1200 kg/ha product loss on an Oxisol and a 1850 kg/ha product loss on an Ultisol per cm of soil loss (Suwardjo and Abujamin, 1983). Such figures show for Indonesia and other humid tropic countries that the prevention of erosion is a major factor in crop yield increase or maintenance. Moreover, Wiersum (1990) has mentioned that the results of soil deterioration are far-reaching: not only do eroded farmlands have a lower yield per hectare, but cropping area and accessibility may also decrease due to gully formation.

Regarding the problem of declining land productivity caused by erosion as mentioned by Stocking (1984) Williams et al. (1984) and Wiersum (1990), some efforts have been made as described below.

(i) The FEDP for land rehabilitation and soil conservation in Indonesia which was introduced in 1983 (Ditkontan, 1985) is a detailed medium term plan for soil conservation and land rehabilitation programmes. This programme includes practices aimed at increasing crop yield in a sustainable manner. Such a practice is *"farming ... with soil conservation"* (Foster, 1991), so Indonesia's Forestry Ministry has helped farmers with a programme intended to control erosion and flooding, and increase land productivity and local income. At each location the combination of the recommended treatments is determined by soil loss estimated by the USLE (Wischmeier and Smith, 1978).

(ii) Erosion survey and mapping in tropical countries in general and in Indonesia in particular have been conducted by various institutions (Ambar and Syafrudin, 1979). In this regard, the USLE has been the most extensively used model. This predicts the effects of rainwash processes on soil loss from standard plots and was designed to compare treatments.

The equation of the USLE, which was designed to be valid for the United States east of the Rocky Mountains, is as follows:

$$A = RKLSCP \qquad (1)$$

where:

A	: annual soil loss	[tons ha^{-1} yr^{-1}]
R	: erosivity factor	[MJ mm ha^{-1} hr^{-1}]
K	: soil erodibility factor	[tons MJ^{-1} mm^{-1}]
L	: slope length factor	[-]
S	: slope steepness factor	[-]
C	: crop-management factor	[-]
P	: supporting practices factor	[-]

The USLE was statistically calibrated based on data collected in the United States. The USLE is not a process-based erosion or hydrological model. From data on land use cover, topography, soil and precipitation, the annual average long term soil loss is generated. Long term soil losses are calculated, which makes the model inappropriate for single storm events.

Risse et al. (1993) conducted a study to develop a set of statistics that would measure the performance of the USLE. They compared estimates of soil loss with measured values on 208 natural runoff plots to assess the error associated with the USLE predictions. Their conclusions were that the USLE overpredicts soil loss on plots with low erosion rates while the plots with higher rates were underpredicted. The accuracy of the USLE in terms of percentage difference between predicted and expected values increases with increasing values of total soil loss.

Recent studies have concluded that the USLE is of only limited value in the very diverse conditions of humid tropical upland areas (Stocking, 1981; Imeson, 1984; Utomo and Mahmud, 1984; Dissmeyer and Foster, 1985; Millington, 1986; Bergsma and Kwaad,1992; Sukresno, 1990; Roo, 1993). It has been applied to the study of uniform (controlled) erosion influencing factors for example using standard plot observations (Hudson, 1985; Dickinson, Wall and Rudra, 1990). In addition, the nature of the model leads to difficulties for validation and the development of the model, which require many field or plot experiments. This USLE can only be applied to irregular slopes and different soils after various adjustments have been made (Dissmeyer and Foster, 1985; Bergsma et al., 1992).

Attempts have been made to improve the USLE. Several Modified Universal Soil Loss Equations (MUSLE) have been developed. They use the same factors as the USLE, but calculate one or more of them in a different way (Freebairn et al., 1989). Onstad and Foster (1975) combined a runoff erosivity factor with a rainfall erosivity factor to use the USLE for estimating erosion caused by single rainfall events. Williams and Berndt (1972) introduced a delivery ratio related to features of the catchment to extend the application of the USLE to larger catchments.

In 1975 Williams (in Lane et al., 1992) replaced the rainfall erosivity factor by a runoff erosivity factor which improved the USLE estimates for single storm events. Dissmeyer and Foster (1985) adopted the idea of Wischmeier (1975 in Lane et al., 1992) of using the subfactor approach to develop cover-management factors (C) for forested areas. This approach allows the estimation of on-site erosion where no data are available by evaluating the site specific conditions. Elsenbeer et al. (1993) developed an erosivity factor based on daily rainfall information. The erosivity factor commonly used requires a continuous record of rainfall intensity. In many areas the best temporal resolution of rainfall data available is the daily rainfall amount. According to Dissmeyer and Foster (1985), Freebairn et al. (1989), and Albaladejo, Montoro and Stocking (1989) the soil erodibility factor cannot be transferred from one area to another area without modification. It should thus be concluded that many MUSLEs exist.

The Revised Universal Soil Loss Equation (RUSLE) (Renard et al., 1991; Lane et al., 1992); the Soil Loss Equation Model for Southern Africa (SLEMSA) (Stocking, 1981) contains the same factors as the USLE, but all equations used to obtain the factor values have been revised. The RUSLE and SLEMSA models were developed as an update of the USLE and the RUSLE will be substituted by the Water Erosion Prediction Project (WEPP) model in 1995 (Lane et al., 1992). Based on the equations of the USLE and SLEMSA, Van den Berg and Tempel (1995) have developed theSoil and Terrain (SOTER) Water Erosion Assessment Programme, called SWEAP. The WEPP and the SWEAP models are computerized, which makes these models possible to evaluate new conditions. The sediment yield of single storm events still cannot be calculated. Because of the limitations of the USLE approach, there is a reason to look for another erosion prediction model as a more promising alternative to get much better prediction, always taking the validity and applicability into account. Stocking (1981) has mentioned that the need for improved prediction of erosion is documented in lists of highest priority needs in agricultural development. This statement is still valid today. What is needed for conservation planning are appropriate, simple, reliable erosion models which exhibit flexibility and ease of updating (Evan, 1990; Ditkontan, 1985).

Consequently more sophisticated erosion estimation models e.g. the WEPP (Nearing et al., 1989); the Areal Nonpoint Source Watershed Environment Response Simulation (ANSWERS) (Beasley et al., 1980); the Chemical Runoff and Erosion from Agricultural Management Systems (CREAMS) (Knisel, 1980); the Guelp Model for Evaluating Agricultural Management Systems on Erosion and Sedimentation (GAMES) (Dickinson et al., 1990); ROSE (Rose et al., 1983), the Erosion Productivity Impact Calculator (EPIC) (Williams et al., 1984) have been under investigation. These models are based on physical erosion processes. However, these models have limited success in dealing with the complexity of the large number of processes involved, which make their validation a continuing problem.

1.1.2. An inductive strategy for modelling soil erosion

In response to the problem definition associated with erosion and current deductive erosion models as described above, one of the major components of this study was building and developing a suitable alternative strategy for predicting soil erosion. It should be noted that the introduction of any alternative erosion prediction modelling to the Indonesian computerised FEDPs is solely based on the consideration that such models should be applicable, suitable and more reliable than the USLE presently in use.

The existing overlay GIS facilities provide for the extraction and map combination of feature attributes which are generally *deduced* by the user (see also Eweg, 1994). There is, however, a general need for incorporation of error (uncertainty) analysis (Molenaar, 1991) and there is also potential for inductive learning procedures, i.e., reasoning from a set of specific premises obtained through field observation to a general conclusion, to be integrated within a GIS (Walker and Moore, 1988; Suryana, 1992; Aspinall, 1992). Such a procedure can be used to generate testable hypotheses concerning relationships among spatial data sets (Openshaw, 1987).

With regard to the inductive learning procedure, several investigations have been carried out in different fields of application. Walker and Moore (1988) have adopted the concept of induction to map the distribution of the Eastern Grey Kangaroo in Australia. The induction in this research employs classification rules and regression of spatial data describing climatic conditions and the distribution of kangaroos to develop a decision tree and probability maps based on a generalised linear model.

Adopting the application of Bayes's Theorem, Aspinall (1992) has applied the concept of induction in GIS to predict the distribution of red deer in North East Scotland. In response to the the first objective of this thesis, Suryana (1992) and Suryana, Molenaar, Imeson (1996), by adopting the Theory of Evidence, have proposed that the inductive concept be applied to erosion studies in the humid tropics in general and in Indonesia in particular. A region-specific Inductive Erosion Model (IEM) was put forward as an approach but requires further validation (Suryana, 1992). The method used, in which each rule is a split of a decision tree, was originally developed by Stanat and McAllister (1977). In this research the method was used to infer a set of locational erosion influencing factors for a particular erosion severity class.

However, the expert's inferences of the effect of each formation factor involved in an IEM and their causal relationship to erosion class are non-deterministic and uncertain. Therefore, the resulting information predicted by an IEM will also include uncertainty. An IEM is a processing model and the implementation of an inference from erosion influencing factors to a particular erosion severity class follows a rule-based approach. The inference rules adopted by an IEM allow this model to be classified as a logical processing model and not as a mathematical processing model (Drummond, 1991). This will imply that uncertainty associated with information produced by an IEM can not be handled using error propagation techniques nor by the operation of fuzzy subset theory (Heuvelink, 1993; Drummond, 1990; Suryana, 1993). An IEM involves fuzziness associated with fuzzy measures (Klir et.al., 1988). An IEM selects the most likely erosion class, with an overall certainty factor which is a function of the various erosion influencing factors involved. An IEM is a region-specific model and its intention *is not to describe the processes or effects of erosion, but to infer where erosion is taking place and its degree, together with an estimate of the uncertainty of the prediction* as explained briefly below.

An erosion severity class, as an attribute of a terrain object TMU, is derived using a limited set of erosion influencing factors, each of which is present to a certain degree. In this regard, an IEM infers from a set of influencing factors to a particular erosion severity class. This is the inference classification rule adopted by an IEM and it is considered to be similar to the decision rules adopted in crop suitability mapping (see FAO,1976; Rodrigue, 1995). Considering this and referring to Molenaar (1995a) this rule can be reformulated as follows. An inference model R consisting of a set of rules relates a set of erosion severity classes E to a set of terrain descriptions

D. This means that erosion severity inference rules $R : (D_i , S_i) \rightarrow (E_j , U_{ij})$ will be established, that map terrain object descriptions D_i associated with a set of erosion influencing factors *(i)* and with the degree of uncertainty S_i , to erosion severity classes E_j with the degree of uncertainty of U_{ij} . When *(i)* satisfies D_i, then uncertainty U_{ij} extends straightforwardly to a function $U_{ij} = U\{D_i , S_i , E_j : R\}$. This function shows that the uncertainty (U_{ij}) associated with an erosion severity class (E_j) inferred by an inductive erosion model *R*, is determined by the uncertainty (S_i) associated with terrain descrition D_i. See also the discussion in Section 2.3.

Thus, an IEM is an inference rule-based erosion model that can be understood in the sense of the relation between locational erosion influencing factors, referred to as conditions, and the spatial distribution of erosion categories, referred to as conclusions. These decision rules are constructed, incorporated and induced after observing patterns of spatial erosion severity classes, erosion influencing factors and their causal relationship at FFL. One advantage of an IEM approach based on TMU as discussed in Subsection 1.1.3 is that it can still be used, though with lower reliability, when not all the erosion influencing factors are known.

The bottom-up approach adopted in an IEM is considered analagous to the concept adopted in up-scaling of observations in soil investigations as proposed by Bouma and Hoosbeek, (1996) and also in the Desertification Response Unit (DRU) which has been developed and tested in the context of European programmes dealing with land degradation.The DRU approach is directed towards finding ways to a better understanding of desertification in semi-arid areas (Imeson, Cammeraat and De Boer, 1992).

It is *a hypothesis of this thesis* that using an IEM based on farm field research will lead to obtaining a better prediction of soil erosion in the humid tropics than that produced by the USLE. The use of an IEM may lead to a better prediction of the relationship between crop yield and soil loss in a specific area under a specific upland farming system and at the same time it may allow examinaton of the changes in soil fertility status caused by erosion (Bergsma and Kwaad,1992), (Suryana, 1992). Once an IEM is built into the GIS environment, it can be used to infer the severity of soil erosion in other similar areas. Incorporating the proposed IEM into a GIS environment will increase its inference capability and therefore its usefulness. In the case of the inductive approach, for example, an expert system can be based on it for use in a GIS, particularly when the required data may be either incomplete or unavailable.

1.1.3 The necessity of a stable basic mapping unit for data acquisition, data capture and representing information

The reliable and repeatable prediction of soil erosion as described in Subsection 1.1.1 requires a particular fixed unit for assessment. In this regard, the Field Engineering Design Plan (FEDP) for land rehabilitation and soil conservation programme in Indonesia adopts the assessment unit *"unit lahan"* (land unit). It is used as a basic (elementary) mapping unit for data capture, analysing and predicting erosion processes and for proposing recommended soil conservation measures (Ditkontan, 1985). For the time being this unit is considered sufficient to support manually the establishment of the FEDP. For computerisation of the FEDP, however, for example by using GIS and remote sensing technologies, it faces the following problems.

Data requirements in the establishment of a FEDP include data on climate, slope steepness, land use, and soil erodibility. Moreover, these data are collected independently by different sources and institutions assigned to conduct particular studies and research projects in the area. Each *"unit lahan"* adopted in the FEDP and used for data acquisition and data analysis is delineated manually by overlaying techniques using the relevant separate maps, i.e., soil, landuse, isohyet and geomorphology. Each *"unit lahan"* therefore inherently contains data relating to erosion processes.

Using for example a geomorphology map to identify the basic boundary of the unit and adopting the concept of polygon intersection then the *"unit lahan"* which carries homogeneous terrain characteristics can be established. Therefore, each *"unit lahan"* represents inclusively a particular land use, slope steepness, slope length, isohyet and geomorphology. Figure 1 shows how this unit is established.

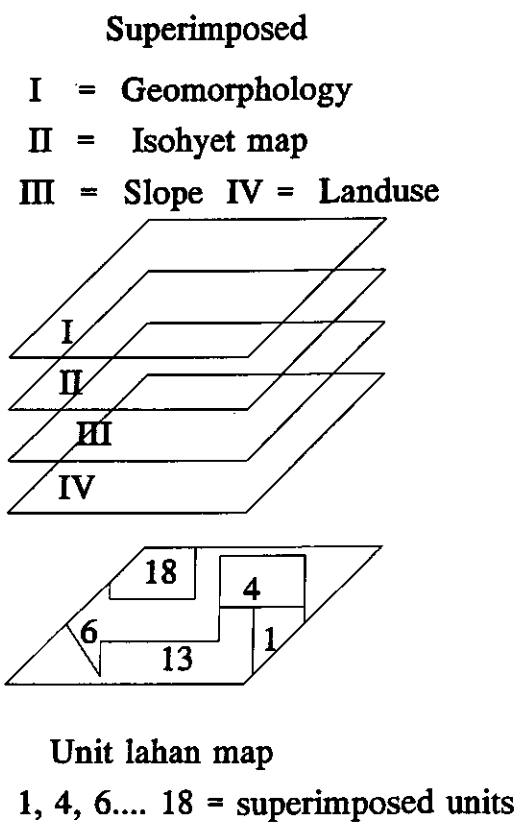


Figure 1: **The establishment of unit lahan adopted in the existing FEDP**

From experience it was found that there is sometimes a significant time interval between data acquisition, data processing, presenting and updating the resulting information. This is identified as a major cause in postponing the establishment of the required FEDP and it reduces its effectiveness. In addition to the needs of improvement of the *"unit lahan"*, the introduction and operation of a GIS to produce information categories such as erosion severity classes may offer some advantages particularly in producing a timely FEDP.

The basic mapping units which have been adopted in the establishment of the FEDP, e.g., units representing soil erosion severity classes, are established on the basis of the attribute values. Thus, the combination of attribute values (themes) defines boundaries of *"unit lahan"*. See again Figure 1. However, it should be remarked that these units are often not defined *a priori* which leads in the end to difficulties in providing better insight into the procedure of how the defined or classified units are obtained. This is caused by the fact that these units do not necessarily represent individual and independent spatial objects possessing identifiable, measurable, repeatable and visible physical boundaries (see further in Sections 6.2 and 6.5).

It is not surprising that this technique is found to have a poor repeatability. Using these units the processes of up-scaling of observations, updating and conducting monitoring of particular changes of land use in general and the effect of land rehabilitation and soil conservation programmes in particular become no longer possible. Another technical problem associated with the overlaying technique is that the boundaries on the independent thematic maps may not be reliable and some spurious *"sliver"* polygons may result after overlaying. The elimination of these polygons may take quite some time and effort (Smith and Campbell, 1989).

For this particular research, the concept of Terrain Mapping Unit (TMU) as introduced by Meijerink (1988) was adopted. This solution was addressed especially to the situation when sophisticated software and good quality data are not available. In order to be able to implement an IEM in observing terrain characteristics and in inferring soil erosion at different aggregation levels (see previous Subsection), *a geo-information theoretical approach* as developed by Molenaar (1995 b) is required. This approach provides better insight into the definition of TMU and into the relationship between these TMUs and the proposed IEM at different aggregation levels.

By adopting a geo-information approach, a stable basic mapping unit TMU that allows the definition of repeatable observational procedures can be established (*the first hypothesis of this particular research*). TMUs as terrain objects are homogeneous relief units that can be delineated by interpretation of aerial photo or SPOT images. In this regard, each delineated TMU represents natural divisions of the terrain usually with *"distinct boundaries"* of geology, geomorphology, morphometry and soil distribution. This implies that each delineated TMU *carries* implicitly relevant and sufficient homogeneous terrain characteristics which permits the use of the TMU approach to derive erosion classes (*the second hypothesis*).

Although a TMU is only a small part of the terrain, it can nevertheless be a very complex system. A particular TMU can often be divided into smaller and more detailed units. This implies that the concept of TMU representing the different aggregation levels and the relations between these units can be modelled through the aggregation and classification hierarchy (Suryana and Molenaar, 1995).

An erosion severity class is a peculiar attribute belonging to a particular TMU. The proposed IEM is intended to infer erosion severity class at different aggregation levels which are related to the observations made at different scales. An IEM takes three different, but interlinked levels of functional units of observation, namely point observations at farmer's field level (FFL), the homogeneous landform level and the level of larger parts of the terrain system. These functional units of observation are identical to Sub-sub TMU, Sub TMU and TMU respectively. Point observations for the proposed IEM are taken at FFL, which is considered to be the most appropriate basic or elementary functional unit to describe erosion over a larger area. This becomes apparent if it is planned to study erosion in relation to the agricultural use of the land.

The nature of the TMU as well as the interpretation process allow the identification of active processes including soil erosion in the terrain. By implication TMU provides information on the relationship between erosion influencing factors and possible erosion classes at each aggregation level. This implies that the abstraction of erosion information will follow simultaneously the abstraction of the TMU information. The TMU is considered therefore to be a suitable observational unit in describing soil erosion at local and sub-regional level. This becomes the ***third hypothesis*** of this aspect of the research.

Considering problems as described in Subsections 1.1.1, 1.1.2 and 1.1.3 as well as in the achievement of the objectives as presented in the introduction, this research was conducted in three stages of study , namely the establishment of observational procedure using TMU(*Stage I*), an IEM model building (*Stage II*) and uncertainty analysis (*Stage III*). An overview of the general methodology is presented in Figure 2 below.

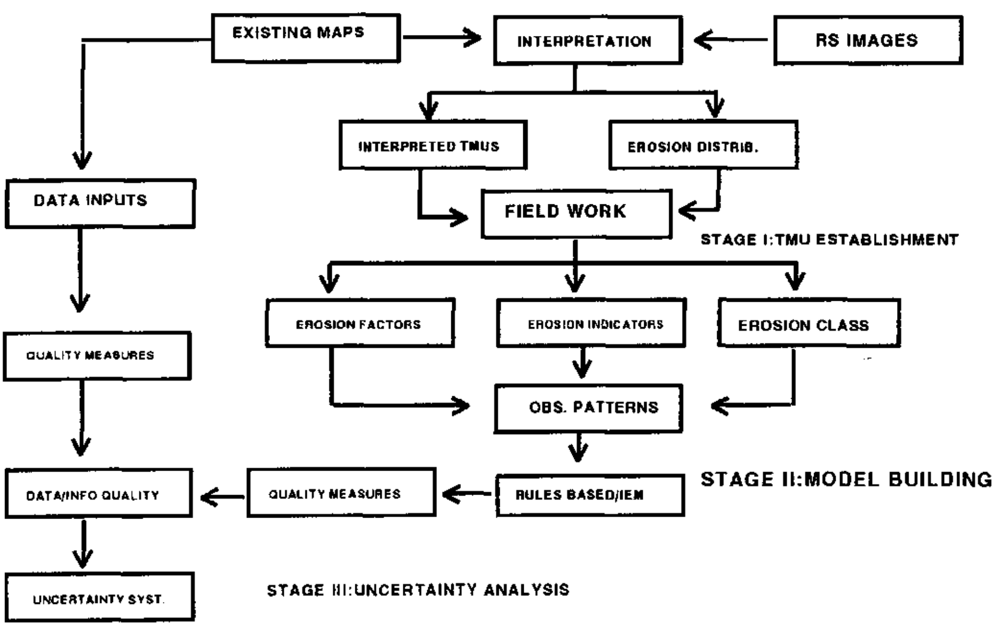


Figure 2: **General methodology of the study**

1.2 Thesis organisation

In order to achieve the objectives of study as described earlier, this thesis is organised generally into 11 chapters as described below.

Chapter 1 discusses general introduction and the objectives of the study, Section 1.1. discusses the inductive approach as an alternative strategy for modelling soil erosion in humid tropics. Subsections 1.1.1 and 1.1.2 discuss problems associated with erosion problems and deductive erosion prediction models implemented in a GIS as well as problems associated with the present GIS generation facililities for handling special tasks and uncertainty analysis in GISs. Subsection 1.1.3 discusses the establishment of the observation procedure as well as the importance of the stability or repeatability of basic mapping or observational units that can be used in data acquisition and the representation of information at different scales. It discusses also the general methodology adopted by this study.

Chapters 2, 3 and 4 discuss the conceptual framework (relevant theory and concepts) related to the underlying problems. Chapter 2 discusses and evaluates the current strategy used in erosion modelling. This chapter also discusses the need for the proposal of the concept of induction and the inductive approach for modelling soil erosion based on TMU. Chapter 3 discusses the geo-information theoretical approach used in the definition of terrain object, data acquisition and observational procedures. Finally, Chapter 4 discusses an approach to handling uncertainty associated with data inputs and resulting information produced by a GIS.

Chapters 5 to 10 discuss the performance of the conceptual framework and were written in the form of case studies of the study area. Chapter 5 gives a general description of the study area and datasets used in this study. It is aimed at providing better insight into environmental characteristics of the study area.

Chapters 6, 7, and 8 discuss the application of the concepts as discussed in Chapter 3 in the establishment of the functional unit TMU (Chapter 6), the hierarchical method for data acquisition, data capture and representing information (Chapter 7), building, developing, testing and the incorporation of the proposed IEM into a GIS (Chapter 8).

Chapter 9 discusses approaches, methods and techniques used in handling uncertainty associated with data inputs into GIS in general and into the FEDP in particular.

Chapter 10 discusses the use of the evidence theory or CF model in handling and representing the ambiguity associated with the proposed IEM.

Chapter 11 presents concluding remarks and recommendations with respect to the adopted method and future perspective of establishing the stable basic mapping unit and inductive modelling in a GIS environment.

Chapter 2

Soil erosion modelling and inductive reasoning based on observations in terrain mapping units

2.1 Soil erosion as a spatial and phenomenological process

Accelerated water erosion is a process by which soil particles are first loosened and broken apart and then transported, rolled or washed away (Morgan, 1974). It is basically caused by the interaction between rainfall as an erosive agent and soil as the medium that is detached and transported. These processes are generally determined by locational factors including climate, soil, relief, vegetation and man-made soil conservation measures. A brief discussion on the effect of these erosion influencing factors is given below.

The potential ability of rainfall to cause erosion, called *rain erosivity*, depends on such characteristics of rainfall as the energy of the falling raindrop impact and the intensity particularly intensity rainfalls within 30 minutes (EI_{30}), the length and the total number of rainstorms. These characteristics determine the ability of raindrops to detach soil particles and the possible occurrence of surface runoff, a primary means for transporting and deposition of detached soil particles (Wischmeier and Smith, 1978 and Hudson, 1981).

The susceptibility of soil to erosion (K-factor), called *soil erodibility*, depends on various soil characteristics, e.g., aggregate stability, transportability of loosened soil particles and infiltration rates. The aggregate stability of a soil determines how easily soil particles are detached. The transportability determines how easily these loosened soil particles can be washed away. The infiltration rate determines surface runoff. The K-factor is obtained using the nomograph of Wischmeier (1976). It is calculated according to the percentage of silt, very fine to fine sand, organic matter, soil structure and soil permeability.

The effects of vegetation cover on erosional processes especially on surface erosion are varied depending on the type of vegetation cover, density, undergrowth cover and litter. These determine the interception loss, absorption of kinetic energy and increasing water infiltration. Land with good cover allows soil retardance to overland flow (Stroosnijder and Eppink, 1993).

In particular circumstances, farmers and their farming practices are the most active geomorphological agents in the erosion process. Their activities include shifting and annual crop cultivation on sloping areas without taking soil conservation practices into account.

Slope steepness and slope length are considered to have a strong relationship to erosional processes. Therefore, both of them can be used for quantitative evaluation and input in inductive erosion modelling.

Erosion may take various forms based on the interaction between the erosive agent and the soil. Erosion caused by water flowing over the soil surface and forming minute chanels is *rill erosion*. Detachment and removal of soil more or less evenly between rills by rainsplash is *interrill or sheet* *erosion*. Other erosion indicators, e.g., *pedestal*, occur when impermeable materials such as rocks, stones or roots provide cover for small areas of the soil and protect them from erosion in the shape of small columns. In gully erosion, the channels formed by water erosion are usually so deep and extensive that the land cannot be used for normal cultivation. Other special types of erosion, i.e., geological erosion, such as piping, pinnacle, streambank erosion, landslides and soil creep, are caused by water agent, but these processes are not treated in this thesis.

2.2 Spatial erosion models: current situation

This part describes briefly the current situation of erosion prediction models. It is intended to give a better insight into the problems of erosion studies in general and in erosion modelling in particular, and therefore to clarify the case for the introduction of new, inductive types of erosion model.

A model is a simplified representation of reality. The reality to be represented can differ considerably even when only erosion models are regarded. The spatial and temporal scales of different models cover a wide range and also the way of representation can vary. Models can be facsimiles of existing areas, in which case they are classified as physical models (Burrough, 1989); analogue models assume a similarity in behaviour between an agent in reality and one in the model; and reality can be represented in formulae, which is the case in mathematical models (De Roo, 1993).

As stated earlier, for land degradation, empirical models have little value since many processes are not represented. Furthermore, changes in the environment cannot be incorporated in the model during the simulation. Land degradation can be considered as a change of ecosystem, i.e., a complex of interacting processes. Modelling this interaction is a prerequisite to analyzing the land degradation.To model the interacting processes, equations of physical, chemical and biological processes may be used. In most soil erosion models no chemical and biological processes are represented.

Figure 3 shows the general strategy of erosion modelling which is currently used. Most of the models adopt either erosion features or erosion hazard approaches. The features approach is intended to map the erosion type that may take place in a particular area. The erosion hazard approach is more concerned with quantitative and qualitative erosion modelling. The qualitative model is intended to model soil erosion on the basis of the geomorphologic and landscape soil erosion relationship. The geomorphology approach assumes that each geomorphologic unit gives a different response to soil erosion. The landscape approach models soil erosion based on environmental-related factors.

In the latest development of quantitative erosion prediction models were created on the basis of the physical laws; they could be incorporated in models giving them a physical basis. This fundamental erosion process includes the processes of detachment by raindrop impact, infiltration, runoff, detachment by flow, transport by raindrop impact, transport, sediment and deposition by flow.

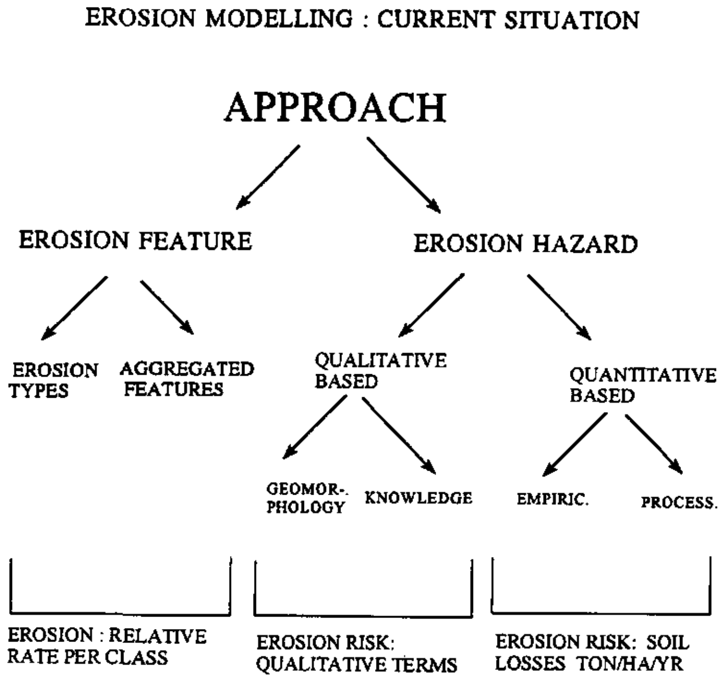


Figure 3: Current strategy for erosion modelling

The total amount of soil loss is expressed as the total of sediment yield which results and accumulates during the process of erosion. Mathematically the sediment load increases or decreases along the slope according to the rate of detachment and deposition. As stated earlier, the operation of these models uses the mathematical and the stochastic model for any single erosion process applied to various areas (scale of study). Some models use an erosion model as a subroutine and generate economic costs or production loss as a final output. No existing models have an ecological output.

The physical based models are classified into four groups according to their spatial characteristics and their physical or empirical basis.

(1) Empirical lumped
(2) Empirical distributed
(3) Conceptual lumped
(4) Conceptual distributed

The equations in conceptual models describing physical processes are deduced from physical laws. In empirical models the equations are based on statistical observations and experiments. Lumped models do not take into account the spatial variability of the variables and parameters used. Average values are used and an average value for the area is given as an output. The distributed models, in contrast, do consider spatial variability.

2.2.1 Empirical lumped models

The most common empirical lumped model is the USLE. Later several models, like MUSLE, SWEAP and RUSLE, have been developed introducing many modifications. See again Subsection 1.1.1.

2.2.2 Empirical distributed models

Very few empirical models have a spatial variability incorporated. An example is GAMES (Dickinson et al., 1986).

GAMES is based on the USLE. GAMES is a modification to estimate erosion over an average season. The model is used to provide the estimates of soil loss by water erosion and subsequent delivery of sediment yield from agricultural catchments. The USLE equation is converted into the GAMES equation in Appendix 1.

The spatial component of GAMES is added for the event that runoff travels across downslope fields prior to entering the stream. The characteristics of different land cells are incorporated in the equation as stated in Appendix 1.

2.2.3 Conceptual lumped models

The conceptual models use the laws of physical processes as a starting point. Since the spatial variability of parameters largely determines the processes, most conceptual models have a spatial component for their parameters. As a result very few models are lumped. CREAMS and EPIC are examples of conceprual lumped models.

CREAMS (Knisel, 1980) was developed when agricultural practices became a serious problem to off-site water quality. The model should evaluate the relative effects of the pollutants from agricultural practices. Since sediment is a major pollutant and moreover a carrier of contaminants an erosion component is included. The area considered is a field-sized watershed, which can range in size from 4 to 400 ha. The variables related to soil, land use and topography are assumed to be uniform (Lane et al., 1992; Silburn and Loch, 1989). The main equation governing both overland flow and channel elements is the steady-state continuity equation for sediment transport (see equation 27 in Appendix 1).

This equation accounts for sediment detachment, transport, deposition and thus sediment yield. The values related to these processes are calculated using several subroutines. According to Hirschi and Barfield (1988) the major drawbacks of the model are its complexity, the intensive data requirements and its reliance on USLE relationships and parameters. They conclude that the model is extremely useful for its intended purposes, although its use is limited for research due to its empiricism. Loch et al. (1989) conclude that the model produces excellent results provided that the erosion processes present under field conditions are represented accurately in the data set used to determine the parameters for the model.

2.2.4 Conceptual distributed models

Most of the more recent models are of a conceptual base and assume a spatial heterogeneity of soil parameters. They are also referred to as 'physically based' models. The advantage of those models when compared to empirical models is their ability to estimate spatial and temporal variations in soil loss. Furthermore, since they are process based, they can be extrapolated to a range of conditions which may not be practical or economical to test in the field (Nearing et al., 1989). Many models of this type exist, some examples are WEPP (Nearing et al., 1989), ROSE (Rose et al., 1983), ANSWERS (Beasley et al., 1980).

WEPP (Nearing et al., 1989) has been developed for use in soil and water conservation and environmental planning and assessment. In 1995 the definitive version became fully operable. Spatial distributions of net soil loss can be calculated, and spatial variability in topography, surface roughness, soil properties, hydrology and land use is taken into account. Simulations are made with a daily time step. Three versions of the model were developed; a watershed, a grid and a profile version (Lane et al., 1992) all dealing with 'field-size' areas (Laflen et al., 1991). This limitation in area is more a limitation of processes considered. Hence, the region involved determines the maximum size of the area.

The watershed version computes erosion within a catchment and simulates erosion in small ephemeral gullies, but not the erosion in classical gullies or continuously flowing streams. The grid version is designed for areas that cannot be represented by the watershed version. This can be because there are several watersheds involved or only a part of the watershed is being studied.

The profile version is used when an area is broken up into elements. The different elements can be combined by the grid version.The model can be used to calculate soil loss caused by single storm events as well as long term averages. The model is based on many processes like infiltration, surface runoff, plant growth, decomposition, management and erosion mechanics. This requires a large amount of data, which may be difficult to obtain (De Roo, 1993).

The WEPP erosion model computes estimates of net detachment and deposition using a steady state sediment continuity equation.The net soil detachment in rills, i.e., rill erosion rate, is calculated for the case when hydraulic shear stress exceeds the critical shear stress of the soil and when sediment load is less than sediment transport capacity (see equations 28 to 33 in Appendix 1).

The ROSE model (Rose et al., 1983; Rose and Freebairn, 1985) is based on three processes, rainfall detachment, sediment deposition and soil entrainment by overland flow. No rill processes are directly represented in the model, although their presence has effect on the entrainment efficiency. The area where the processes are active is a plane land element. The output of the model is sediment flux, which can be calculated for any time and position on the plane. ANSWERS (Beasley et al., 1980) is designed to simulate the hydrological behaviour of primarily agricultural catchments during and immediately after storm events (De Roo, 1993). De Roo et al. (1989) integrated the model within a Geographical Information System (GIS). Data can thus be entered easily and the output data can be displayed as maps and tables.

The model incorporates hydrological components, a sediment detachment/transport model and routing components to describe the movement of water. The catchment is assumed to be composed of square cells. Both spatial and temporal variability can be taken into account. The spatial resolution depends on the cell size, which is about 1 to 4 ha in the latest release. Using the GIS version this can be significantly reduced giving the model a more distributed character. The time step used can be chosen freely, although a better spatial resolution requires a shorter time interval.

The model was designed to compose data files using readily available sources of information. Changes within an existing data file can be easily entered by modifying those areas where they have occurred. The governing equation for the erosion component of the model was developed by Foster and Meyer (1972, in Beasley et al., 1980). See equations 37 to 39 in Appendix 1.

To solve the equation, three steps are necessary. During the first step the available transport rate is compared to the incoming sediment movement rate. The second step consists of finding the impact detachment rate by ANSWERS and comparing this rate with the available (excess) transport rate. If additional transport capacity exists, the rate of flow detachment is calculated and compared to the transport rate as a third step.The detachment of soil particles by raindrop impact is calculated using the relationship described by Meyer and Wischmeier (1969). See equations 40 and 41 in Appendix 1.

2.2.5 An evaluation of existing physical based erosion models

Even though in the foregoing paragraphs four different types of models have been described and several example models were discussed, all of the models described are similar in that they predict the amount of soil loss during a certain period. This period can vary from just one single storm event to an average year, but the outcome is always given in units of weight per area per period.

In the method of calculating the amount of soil loss, however, they do differ considerably. It has been noticed before, that the spatial component of land degradation can be represented or ignored. But more important is that the way the equations have been obtained differs considerably. The empirical models are mostly based on statistical observations and correlations, while the conceptual models are based on the different components active in the soil erosion process. Alternatively formulated, in the empirical models the existing correlations between several parameters are known, the values of the input parameters are determined and subsequently the output value is known. However, in the conceptual models the input values are established and via algorithms based on physical processes and causal relationships the amount of soil loss is calculated. The different components are represented in the conceptual models, which therefore give a more realistic representation of the soil erosion problem than empirical models.

An important factor that influences soil erosion strongly, but that is hardly ever taken into account, are the initial soil conditions. They are determined by for example the soil moisture content, but also the land management of the previous season. It is very labour-intensive to determine the initial conditions, since they are influenced by the complete history of the soil.

A severe disadvantage of most conceptual models is the amount of input data required to run the model. It can be very difficult to obtain all necessary data for a larger area and especially their reliability can cause problems. Several authors (De Roo, 1993; Morgan, 1986) have pointed out the desirability of conceptual models, but due to the unreliability of these models at the moment, they conclude that at present empirical models perform just as well as conceptual models.

A different approach of the land degradation problem could be found in assessing soil erodibility rather than soil erosion. The erodibility depends mainly on the soil structure and its infiltration capacity. With a well developed soil structure, i.e. a strongly aggregated soil, less material will be detached by rain or overland flow, and thus cause a lower erodibility. A higher infiltration capacity yields a larger volume of water that can be stored before overland flow is induced, and thus causes also a lower soil erodibility. Generally, it can be said that a better aggregated soil results in a better infiltration capacity. So if a model could be developed to monitor the aggregation of the soil rather than the whole process of soil erosion, it could be possible to monitor the erosion problem with far fewer input parameters than are necessary for the present conceptual erosion models.

The output of such a model would not be sediment yield, but rather a measure of soil stability. This could be a problem for some applications, where the sediment yield is required. This happens for example in the CREAMS model (Beasley et al., 1980), where the sediment yield is used to calculate the pollution of streams by the sediment. But in many cases, like land management practices and governmental decisions, the outcome of erosion models is used to indicate areas where (severe) erosion could occur. A measure of soil stability is the outcome of a different approach to the problem, but is just as useful for such purposes.

2.2.6 Soil erosion modelling using computer-based image processing

Another promising alternative approach is the quantification of an erosion prediction model by the computer processing of remote sensing (satellite) data (Pickup and Chewings, 1986; Hill, 1993). In the early nineteen seventies digital image processing techniques were introduced in soil survey. These techniques are solely based on the different reflectances of different terrain characteristics (Mulders and Epema, 1986; Hill, 1995). These techniques recently have been used to study or to assess erosion problems in the framework of soil survey (Bocco,1991). The results were supported by Epema who conducted studies in Tunisia (1986) from which he found that the effect of soil erosion could be identified, especially in terms of the degradation of soil texture and structure, and the decreasing content of minerals, nutrients and organic matter. In this regard, remote sensing images provide data of sufficient spatial resolution for large areas with reasonable frequency. The introduction of digital image processing techniques as mentioned above enables the process of updating erosion information that changes rapidly over time and space.

It is proposed that the integration of remote sensing data with GIS data or other ancillary data will further improve the prediction and erosion classification (Bocco and Valenzuela, 1988; Suryana, 1991; Janssen, 1993). The application of remote sensing data combined with the use of GIS has been undertaken. Seubert , Baumgardner and Weismiller (1979) and De Jong (1994) studied ground cover classification using the unsupervised approach and evaluated the usefulness of the approach in delineating severely eroded cultivated land. At the end of their study they concluded that higher reflectance correlates with more severely eroded upland soil. In other words, increasing soil reflectance indicates an increasing degree of erosion.

Even though it is recognized by several authors that remote sensing is the right tool to estimate land degradation (Pickup, 1990; De Jong, 1994), so far only two models based on it have been developed. Pickup et al. (1986) developed a model using Landsat MSS images which generates erosion patterns which evolve when the intensity of erosion increases or when the landscape restabilises. Incorporating this concept into a remote sensing model, Pickup and Chewings (1988) developed a Soil Stability Index (SSI). In a graph the points are plotted according to their ratios of MSS bands 4/6 and 5/6. They appear to cover an area between two parallel lines. The distance to the upper line is the SSI. Points close to the upper line represent severe erosion. Points close to the lower line represent deposition and points in between are transfer zones. In addition, Bocco and Valenzuela (1988) used Landsat MSS images to estimate land use and combined this with a digital elevation model. This combination produced three coefficients of the USLE, the slope gradient, the slope length and cover.

De Jong (1994) developed a model, called the Soil Erosion Model for MEDiterranean (SEMMED) areas. The SEMMED model has been developed for the northern part of the Bas-Vivarais, Ardèche, France, using Landsat TM images which yield an erosion hazard map.

2.3 Inductive reasoning based on TMU

The following description is intended to provide a general insight into the concept of inductive reasoning adopted by the proposed IEM and to the role of the observational procedures based onTMUs in facilitating the estimation of soil erosion severity classes.

Because of the limitations of the USLE approach in very diverse upper watersheds in the humid tropics, there is a reason to look for another erosion prediction model. More sophisticated erosion prediction models have been under investigation. These models have been based on physical erosion processes. However, these later models have limited success in dealing with the complexity of the large number of processes involved, which make the validation of these models a continuing problem (discussed in Subsection 1.1.1).

In response to the problem definition associated with erosion and current deductive erosion prediction models as described above, one of the major objectives of this study was building and developing a suitable alternative strategy for predicting soil erosion which involves an IEM. It should be remarked that the introduction of an IEM to the Indonesian computerised FEDPs is solely based on the consideration that this model is applicable, suitable and more reliable than the USLE presently in use.

According to Bonnet (1985) the deductive approach, e.g., adopted by the USLE, is not well adapted by its very nature to the conditions which actually operate. An alternative which appears to be able to handle local variabilities in, e.g., relationships between animal distributions and their influencing factors, is the use of the inductive approach (Aspinall 1991; Walker and Moore, 1988). Furthermore, Suryana (1992), Suryana and Molenaar (1995) and Suryana et al. (1996) have proposed this approach to be applied to a modelling erosion severity classes. This is a different type of approach as compared to the erosion prediction models that were discussed in Subsection 1.1.1 and Section 2.2.

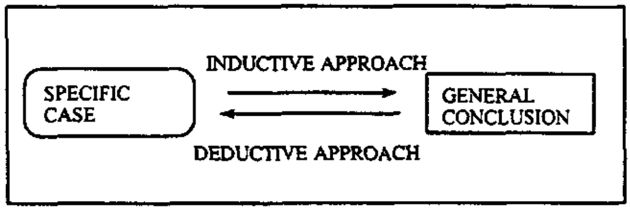
As stated also in Subsection 1.1.2, an induction is a reasoning process by which a general conclusion or inference is drawn from a set of premises, based mainly on field experience or experimental evidence (Mortimer, 1988; Walker and Moore, 1988). Thus, the concept of induction understood in this way is opposed to deduction, i.e., reasoning from a general premise to individual or specific conclusions. This reasoning process can also be found by analogy or statistical inferences in which conclusions are not universal statements. This will imply the importance of the creation of *"rules"* which govern inferences which are needed to formulate criteria and allow recognition of those which may be correct.

Adopting the concept of induction as mentioned above, the proposed IEM is considered to be a region specific model which is intended not to describe the processes or effects of erosion, but to infer where particular degrees of erosion severity occur, together with an estimate of the uncertainty of the prediction (see again Subsection 1.1.2). Starting from the concept of induction, an IEM can be built based on the inference rules which are constructed, incorporated and induced after observing spatial erosion severity classes, erosion influencing factors and their causal relationship at the FFL. An IEM uses these rules to infer a particular erosion severity class from a set of locational erosion influencing factors in similar geographical areas.

Each inference rule of an IEM relates a possible erosion severity class to a set of erosion influencing factors with a certain level of confidence (in Subsection 1.1.2). Let D_i represent a description of a particular terrain situation. A rule R relates then D_i to an erosion severity class E_j with a confidence level (uncertainty) U_{ij}. In general, this level will never have a confidence level of 100%, because an expert will never be completely certain about his/her inference. A final decision will therefore be to infer the erosion class E_j, if for all other possible erosion classes E_k, we have confidence expressed by U_{ij} which is greater than confidence expressed by U_{ik}.

Considering the phenomenological processes of soil erosion, the inference rules adopted by an IEM may also have an overall goal of confirming or discarding inferences on the basis of the facts or evidence introduced by users. Therefore, confidence in this inference is increased progressively or shown to be untenable as more observations on erosion influencing factors are collected and inserted to the inference rules of an IEM. Let original observation be D_i and new evidence be D_I, so that confidence in D_i is less than D_I. This will imply that the new evidence increases certainty of E_j, if confidence expressed by U_{Ij} is greater than confidence expressed by U_{ij} and vice versa. The creation of inference rules adopted by an IEM allows the capability to cope with competing hypotheses and may lead to conflict resolution (Suryana, Molenaar, Imeson, 1996).

In order to be able to infer a particular erosion severity class from a set of erosion influencing factors, the proposed IEM exhibits three distinct requirements as depicted in Figure 4.



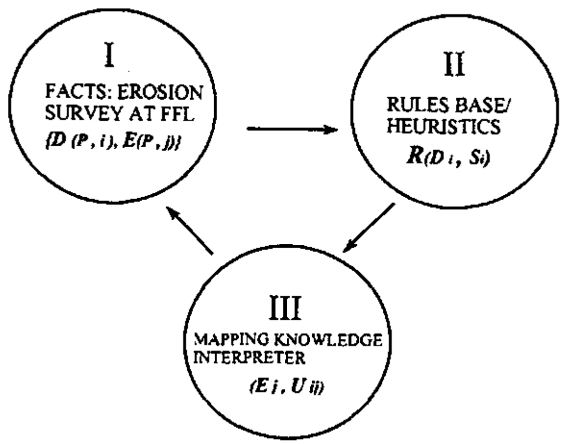


Figure 4: Inductive chain of reasoning

Firstly an IEM requires *facts or definitions of observed patterns* at a particular location **P**, i.e, P $(D_{P,i}, E_{P,j})$, representing the relationship between soil erosion $(E_{P,i})$ and their influencing factors $(D_{P,j})$ at location P. These observed patterns are acquired through a repeated number of FFL observations on soil erosion and under a wide variety of conditions. A more elaborate discussion on the identification and the collection of observed patterns is given in Subsection 6.3.2 and Section 8.3.

The second requirement of an IEM is *the induction of inference rules, i.e.,* $R : (D_i, S_i) \rightarrow (E_j, U_{ij})$ *used in soil erosion classification procedure including the definition of erosion severity classes.* This can be defined generally as a part of the learning systems of an IEM which are developed from evaluation of a set of observed patterns as discussed above. The fact that the suffix $_{(P)}$ has been removed implies that the observed patterns P $(D_{P,i}, E_{P,j})$ are considered to be representative within a specified region for a general pattern P(D_i, E_j) which not only valid at location P, but also at the other locations.This means that if within this region D_i has been observed at some locations, then E_j can be infered.

Based on collection of the observed patterns as mentioned above, it becomes apparent that sets of values of erosion influencing factors are linked to erosion classes. This leads to the last requirement of an IEM that is the incorporation of *a knowledge interpreter in finding a systematic or operational method of inferring, i.e., the application of the same rules in other similar areas.* This method is operated constantly as drawing an inference on particular *"information categories"*, from a set *"data classes"*. The term *"data classes"* refers to a set of erosion influencing factors (D_i) and is expressed using attribute values. The term *"information categories"* refers to erosion severity classes (E_j) describing specific characteristics of the terrain. The process of inferring from data classes into erosion information category is referred to as mapping *(R)* in the sense of relation between subsets and not in the sense of map making.

An IEM in a GIS environment is intended to infer soil erosion severity classes at different hierarchical levels. This implies that the implementation of an IEM requires a fixed or stable observational unit of analysis. In this regard, the concept of TMU as discussed briefly in Subsection 1.1.3 is relevant to the proposed IEM. The TMU facilitates the associate aspects of an IEM in:

(i) *the establishmentof a stable observational unit of analysis;*

(ii) the establishment of an observational procedure in identifying relationships between spatial objects and their attributes at repeated and different scales of observation;

(iii) the establishment of a method for handling uncertainty associated with data inputs and resulting information produced by an IEM.

The operationalisation of the aspects (i) and (ii) requires a geo-information theoretical approach to define TMU and an observational procedure. Chapter 3 discusses the the relevant data modelling concepts and observational procedures based on TMU. The discussion is focused on how the complexities of the terrain object TMU in GIS environment can be defined, understood, modelled and described at different aggregation hierarchies. Chapter 4 discusses fuzzy reasoning concepts including the evidence theory for handling uncertainty in GIS. More elaborate discussions on the application of these concepts with examples are presented in Chapters 6 to 10.

Chapter 3

Terrain object description in Geographical Information Systems

3.1 GIS, geographical and spatial data defined

A Geographical Information System (GIS) could be viewed as a combination of computer hardware and software which is capable of manipulating both locational and nonlocational data. Locational or geographical data form a part of spatial data. The term spatial data refers to any data concerning, e.g., a process in two, three or multi dimensional space. Geographical data are spatial data referring to a particular geographical location. Thus, the nature of geographical data can be presented in either two or many dimensions and differentiates a GIS from other information systems (Fox and Chow, 1988; De Man, 1988; Goodchild and Goval, 1989).

The spatial aspects of geographical data can be obtained by applying the concept of fields or object based concepts (Janssen, 1993; Ehlers, Edwards and Bedard, 1989). The field based concept refers to the situation when thematic data are assigned to spatial units of predetermined shape and size and presented in the form of a raster format. Therefore this concept does not allow the small scale mapping of natural vegetation areas which have irregular shape and size. To avoid this problem, the object based concept defines terrain objects with boundaries that delineate thematically homogeneous units. Spatial objects are generally conceptual entities which are defined within some specific users context.

Basically the spatial datasets stored in a GIS environment represent the real world at different observation levels. Due to human activities and natural processes the real world changes continuously. These changes are considered as the dynamic properties of the real world which can be described by changes of the geometric and thematic attributes. This statement emphasises the importance of spatial data modelling in a GIS related to changes in thematic and geometric aspects.

Geographical data modelling can be implemented in four levels namely spatial, conceptual, logical and physical models as presented in Figure 5 (Molenaar, 1995 b; De Hoop, 1993).

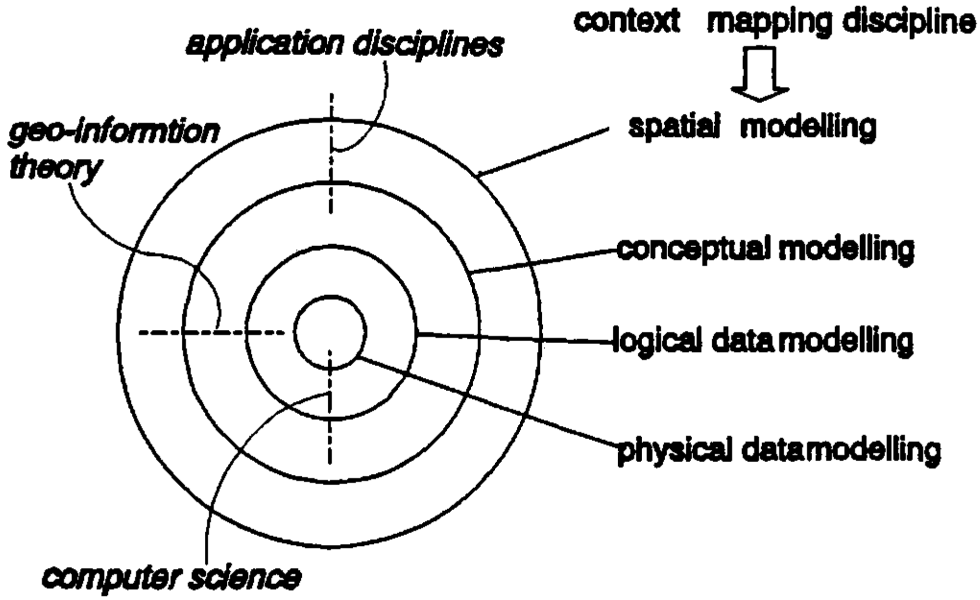


Figure 5: Levels in data modelling (after Molenaar, 1995 b)

The conceptual model describes every single entity and the relationship among them and is addressed to an intended field of application. The nature (level) of the conceptual data model is system independent which allows the formulation procedure to be done without considering the implementation of a particular data base management system or of a particular GIS (Peuquet, 1984).The logical data model is a step further from the conceptual model. It represents the operationalization of the conceptual data model either in the form of an object oriented or relational data base. Thus, the logical model is typically not system independent but depends on the type of selected data base. The physical data model addresses the actual achievement of the logical data model in the computer.

3.2 Formal data structure

In the middle nineteen seventies, the term object orientation was introduced in information modelling including entity relationship modelling, information and semantic modelling. The term object orientation in this thesis refers to geographical or terrain objects.

In this regard, the term object depicts a symbol representing one or more occurrences of a real world entity (Coad and Yourdan, 1991). By this is meant that any occurrence or complex of occurrences in the real world is an object which is described by its geometric and thematic attributes.

The attributes associated with a geographical object are related to the description of substance, characteristics, geocodes (position), temporal information (indicating time or date when characteristics are still valid), lineage (indicating means used in gathering information) and connectivity.

The formal data structure (FDS) developed by Molenaar (1989 b) is a very good example of the representation of a terrain object oriented data model. The FDS concept has been designed for single valued vector maps. According to this concept terrain objects can be described by two semantic levels: firstly the geometric level comprising the metric and topology information of a geometrical primitive (arc and node); secondly the thematic level in which terrain objects are described by their thematic attributes. See the model developed by Molenaar (1989 b), in Figure 6.

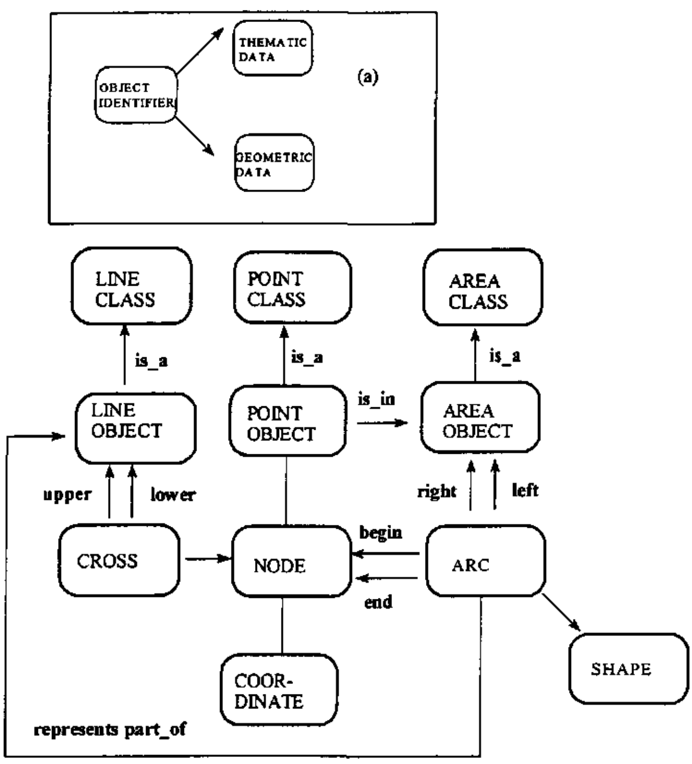


Figure 6: Formal data structure for GIS (after Molenaar, 1989 b).

According to Figure 6, the datasets are represented by rounded boxes. Link types in the FDS are represented by arrows and straight lines. Each arrow indicates *a many_to_one relationship* (e.g. an area feature consists of many arcs) and a line without arrows represents *a one_to_one relationship* (e.g. a point feature only can be represented by one node).

3.3 Thematic attributes of terrain objects

The data base of a particular GIS system has to be semantically meaningful. By this is meant that every single terrain entity stored in a data base is a meaningful object which can be described using its attributes and the position of each point, line, area and shape entity.

As stated in Molenaar's FDS model a thematic and a geometric description are associated with objects as follows: (i) object types; (ii) object classes. On the basis of geometric aspects, object types can be differentiated into point, line and area objects, while object classes are defined by their thematic aspects. The thematic attribute data is a means for describing a particular entity. Examples of such defined object classes are town, roads and tea plantation.

Referring to the above examples and the FDS it becomes apparent that the object classes, e.g., town, roads, tea plantation are mutually exclusive within a well defined context. In addition, *"each class of objects has its own class name or class label and a list of attributes"* which give the thematic characteristics of the defined class (Molenaar, 1993).

The implication of this statement can be restated as follows.

(i). Each object has only one class label or description structure or list of attributes. In other words, terrain objects are distinct by their attribute values in combination with their geometry. Any terrain objects belonging to the same class share the same descriptive (attribute) structure.

(ii). An object class should contain objects of only one geometric type. This implies that the attribute structure as mentioned in (i) is determined by the class to which it belongs. Thus, each object has a list containing one value for every attribute of its class.

3.4 Geometric attributes of terrain objects

The geometric or positional data in a GIS have a function as a tool to link directly different types of thematic data which may be obtained at different times. However, the occurrence of a direct link between the two aspects of a GIS, i.e., thematic and positional attributes occur at the field approach to the data structure. The representation of such a field structured geo-data base requires that the continuum is discretised in the form of points or finite cells in a regular grid (Molenaar, 1995 b). In the object oriented approach, thematic attributes are not linked directly to the positional data but to terrain features using feature identifiers. As stated earlier, the semantic meaning of a spatial data base includes topological relationships among individual entities. It is referred to as the connectivity properties of geographic entities (Molenaar, 1989 a).

Moreover, Molenaar (1995 a) mentioned three levels of topological relationship as follows.

* Low level topology: (i) the relationship between the geometrical primitives as given by the

graph-structure of the vector map; (ii) the linkage of the primitives geometrical elements with the terrain objects.
* High level topology: (iii) the relationship among terrain objects.

The first two levels of topological relationships provide geometrical information about terrain objects and are represented by the relationship between *left and right; begin and end; upper and lower; and is_in relationship*, e.g., point to area connectivity (a well in a grass field). See again Figure 6. The high level topology is a connectivity or relationship among terrain objects. An example is area object to area object connectivity (farm nr. 04 is an orchard and is adjacent to farm nr. 05 which is a farmyard). See Figure 7 (b). In recent GIS developments this topological connectivity has been stated explicitly and stored in the topological codes. In particular situations such topological information can be of interest to GIS users.

3.5 Classification and aggregation hierarchies of terrain objects

The classification hierarchy may result in several levels (Smith and Smith, 1977). The terrain objects are placed at the lowest level of the classification hierarchy (Molenaar, 1995 a; O'Neill, 1988). Therefore, they can be considered as elementary objects within this hierarchy. According to this concept and based on complete common thematic attributes, Figure 7(a) shows that terrain objects like mixed garden, upland farming and farm yard can be grouped into drylands agriculture. Based on partly common attributes these object classes can be grouped into farmland.

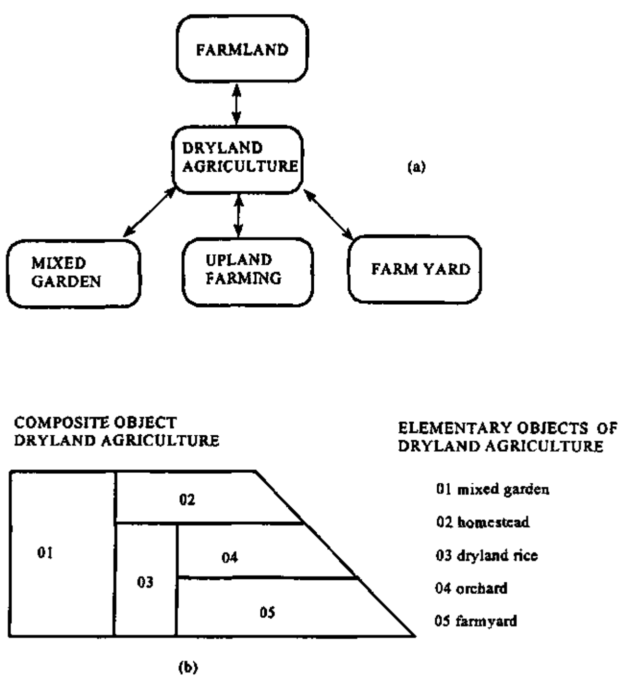


Figure 7: Classification and aggregation of terrain objects (a) class generalisation; (b) object aggregation

Figure 7 represents the generalization and specialization operations on object classes. The upward direction of the arrows represent *is _a_ link*, e.g., mixed garden is *a dryland agriculture* and *is a farmland*, therefore the upward and downward direction of the arrows denote generalisation and the specialisation respectively.

The existence of elementary objects as described above implies the existence of composite objects. The composite objects in an aggregation hierarchy which is defined based on the thematic and geometrical data of the elementary objects adopting *"part_of link relationship"*, e.g., farmyard is a part of a farmland. This implies that the aggregation hierarchy requires direct or indirect topological relationships among elementary objects.

Considering this condition, Molenaar (1995 a,b) introduced the concept of class generalisation (i) and object aggregation (ii) procedure. In this regard, the aggregation of elementary objects only can be performed according to the following procedure:

(i) select any of the (thematic) classes of elementary objects that may possibly be aggregated on the basis of some shared characteristics. For example in Figure 7(a) the elementary objects *upland farming* and *mixed garden* may have shared common attributes.

(ii) select the specific elementary objects that may possibly be aggregated into particular composite objects. At selection or aggregation stage the elementary objects have to have a direct topological relationship. Figure 7(b).

In an aggregation hierarchy it should be noted that the elementary object should only belong to one particular composite object of a particular type. This statement implies that the aggregation hierarchy adopts a many to one relationship. The definition of elementary and composite object depends on the application context.

3.6 The dynamics of terrain objects

Terrain mapping units in this thesis denote typical terrain objects which are defined within a sub-regional context and refer to the representation of real world phenomena.

Due to human activities and natural processes these types of objects are very sensitive to changes both in time and space. These are the dynamic properties of terrain objects that refer to changes in the geometrical and thematic characteristics.

Regarding erosional processes the opening up of ground cover certainly will accelerate these processes. For example a suddenly cleared forest area or changes in farming activities will influence the degree of erosional processes and the natural boundary of the terrain objects. This implies that data pertaining to these objects stored in a GIS need to be updated to sustain a valid representation of the real world, i.e., eroded TMUs.

Considering aggregation and classification hierarchies as described in Section 3.1.5, some changes of terrain objects may take place at the lowest level of the hierarchy. In this regard, Molenaar and Janssen (1993), Molenaar and Janssen (1994) have mentioned that the geometrical and/or thematic data of elementary objects need to be updated. In consequence the redefinition of aggregation structure of composite objects at a higher level of hierarchy is also required as explained in (i), (ii) and (iii) below.

With regard to the dynamics of terrain objects there are three types of changes that can be used to evaluate the dynamic property of the terrain object as follows:

(i) Change in thematic characteristics
Erosion processes will be highly affected by the changes of ground cover type and the percentage of ground cover on a particular TMU, e.g., the conversion of forest cover to rubber plantation or the conversion of secondary forest to upland agriculture. The first example does not affect erosion classification, because these ground cover types have a similar effect on erosion. Thus, the attribute structure of this TMU does not change. The change in the second example may imply a change in the attribute list.

(ii) Change in geometrical characteristics
Changes in thematic characteristics may be followed by changes in size, shape, position and any combination of these three. The implication of changes in these geometrical characteristics may or may not result in changes in topological relationship. The implication resulting in changes in topological relationships is associated with the situation when extreme changes in position, size and shape of typical man-made terrain objects may form new topological relationships completely different to the original topological situation. This implies that the changes can be detected using their geometrical data. The implication without having change in topological relationship is associated with the situation when the changes of position, size and shape are ignorable. These issues are not treated in this thesis.

(iii) Change in aggregation structure
Change in aggregation structure is related to the presentation of erosion information at different levels to form composite eroded TMUs. This may be found in the process of fragmentation which creates new boundaries or in the process of merging and replacement in which old boundaries of eroded TMUs are dissolved. This case represents change of attribute values associated with elementary object TMUs. However, it should be noted that the aggregation of eroded TMUs is not always made within one class hierarchy. It may be possible that a change in aggregation structure may be also a result of changes in both geometrical and thematic attributes. In this case, the change in the aggregation structure implies the necessity of the definition of a complete new set of classes of eroded TMU.

Chapter 4

An approach in handling uncertainty and ambiguity in Geographical Information Systems

4.1 An overview

In general, models used in GIS and the data input are not error free but include uncertainty. This applies also to the particular case of the FEDP. Moreover, the uncertainty is propagated as data are inserted into a model. Therefore, the model outputs may be unreliable.

The concept of uncertainty is defined as a global term which is used to encompass related characteristics of the data, their processing or presentation which may raise concern or doubt in the mind of the user as to the validity of an intended decision (Brimicombe, 1993; Molenaar, 1991; Drummond, 1991). However, in this study, the term *"uncertainty"* refers to the inaccuracies, inexactness, fuzzy values or inadequacies inherent in most spatial data sets (discussed in Chapter 9) and how these may magnify and affect the quality of resulting information produced by fuzzy functions, e.g., change in belief function (discussed in Chapter 10).

Considering this, there is a necessity to obtain knowledge and better understanding of the forms of uncertainty. This and the associated theoretical background may be used as an approach for handling uncertainty in order to estimate the accuracy of spatial data sets (Sections 4.3 and 4.4) and to estimate the quality of resulting information (Section 4.5).

GIS developers and researchers have attempted to handle uncertainty associated with data inputs and the resulting information. See Drummond (1991), Heuvelink (1993), Suryana (1993) and Suryana et al. (1996). Some theories for handling the uncertainty of information have been proposed. For a long time the Bayesian model has been the primary "Numerical" and "Probability" approach for representation and inference with uncertainty. The limitations of the Bayesian approach have led several investigators to try to find another alternative considered suitable for handling different forms of uncertainty.

Also, the uncertainty associated with information may take different types (see Section 4.2). Because of these differences, the probability approach as stated earlier has prompted many mathematical models of uncertainty including the concept of fuzzy measure.

According to Klir and Folger (1988), the concept of fuzzy measure consists of two main types namely *Belief* and *Plausibility measures* and it holds the monotonicity property with respect to the subset relationship. Plausibility is a fuzzy measure associated with belief. The plausibility of an eroded area being present is directly related to the degree of belief that that erosion is not present. In other words, believing the eroded area not to be present strengthens the plausibility of this erosion being absent.

By definition, Plausibility and Belief measures are interrelated from which one of these can be uniquely calculated from the other (see Klir et al., 1988). The combination of these measures forms a theory, called *the mathematical theory of evidence*.

Furthermore, Klir et al. (1988) introduced a special subtype of Plausibility and dual subtype of Belief measures namely Possibility and Necessity measures. Possibility is the plausibility measure applied to a nested family of subsets (e.g., inclusion of a *fuzzy subset A* in a *fuzzy sub set B* in which there is no conflicting evidence). Medical diagnosis is often of this type, as the evidence narrows down to a more specific diagnosis. Necessity is a fuzzy measure associated with the restriction of belief. In the medical example, the necessity of a diagnosis of a particular disease is directly related to the belief of that disease not being present.

In this thesis much attention is given to belief measure, that is a fuzzy measure associated with the type of uncertainty quantifying the degree of evidence supporting membership of several crisp sets, such as in the phenomenological process where two crisp sets are the presence and the absence of a particular erosion severity class that may exist on a particular TMU.

4.2 Types of error and uncertainty in GIS

Burgess and Webster (1983) mentioned that due to human failure, fuzziness of the concept, complexities of the environment, technical limitation during field survey, data acquisition and data processing, information resulting from handling by GIS has limited precision and accuracy, referred to in terms of errors and uncertainty.

Generally speaking the term uncertainty in any area of expertise is associated with statements or decisions as to whether an object (a) is a member ($a \in X$) or not a member ($a \notin X$) of a subset X (Molenaar, 1993). In the particular case of terrain attributes in a GIS, uncertainty is represented by the degree of certainty associated with the assignment of an object to a particular class of objects, e.g., erosion class.

Concerning the process of assignment, attribute uncertainty in a GIS may take different types as described below.

(1) The first type of uncertainty in a GIS may be generated because the information carried by object is not sharply defined and as a consequence the membership of this object cannot be defined either.

(2) The second type of uncertainty in a GIS may be generated because although the object is well defined the membership of the class X is not well defined, i.e., X is fuzzy.

(3) The third type of uncertainty in a GIS may be generated because of the incompleteness (lack) of information. It may be that the object description is precise and that X is well defined, but that the information carried by (a) is not sufficient to decide that $a \in X$.

To handle each type of uncertainty as listed above, there is a particular theoretical framework, technique and approach. The third type of uncertainty arises as *a result of insufficient data (evidence)*, and consequently a particular assumption must be made. This type of uncertainty is closely related to the straightforward application of the Fuzzy Subset Theory (e.g. *supporting subset A as a finite crisp set but with fuzzy member*), as introduced by Kaufmann (1975). See also Doyle (1983) and Drummond (1991). Within the framework of the research of the author, the techniques associated with the first and the second form will be elaborated in Sections 4.3, 4.4 and 4.5 below. The application of approaches and practical examples are given in Chapters 9 and 10.

4.3 Attribute accuracy assessment associated with non-continuous data

4.3.1 General view

This section of the thesis deals with the quality assessment of existing categorical maps used in the FEDP for the Land Rehabilitation and Soil Conservation Programme, in Indonesia. The theoretical insight into the uncertainty associated with data inputs is given.

A categorical coverage map is a map representing a complete mosaic of polygons having different categorical attributes (attribute values). They refer to data classes representing substances, characteristics, variables and values that are grouped into continuous (non-categorical) and discontinuous (categorical) data types at both coverage and entity level.

This section is devoted to approaches in handling uncertainty associated with categorical-discontinuous variables derived from existing categorical coverage maps, visually interpreted imagery and remote sensing image classification. The second part deals with continuous variables resulting from direct measurement, e.g., interpolation. A detailed discussion is given in Sections 9.4 and 9.5.

Considering the process of assignment of some particular attribute values (refers to subset M) to a particular basic terrain mapping unit (refers to subset S) there is considered to be an interdependence or relationship (R) between polygons and the particular attribute contained. This implies that the attribute attached to a particular unit depends completely on attributes associated with these polygons.

In this regard, Molenaar (1995 a) has mentioned the existence of a many to many and a many to one relationship (R) between two different subsets M and S which can be expressed in the form of a binary relation ($R \subset M \otimes S$). The case of the many to one relation implies that the relation R relates 0 (nil) or more elements of M to each element of S, and that the subsets of M generated by the elements of S are disjunct. A many to many relation occurs when elements of the subsets of M overlap to form a particular subset S.

Based on these relationships, it becomes apparent that the assignment of the attributes to polygon(s) or to positions within polygon(s) will be affected by the following relationships.

(1) Boundary accuracy is an aspect of positional accuracy and is not discussed in this thesis.
(2) Control coordinate accuracy is related to the registration of mapping units and the polygon attribute map with respect to each other.
(3) Classification accuracy is related to the quality of attributes associated with designation of the polygons. See also Edwards (1994).

Considering all three aspects listed above, literature study revealed limited explanations with regard to the involvement of these in determining the quality of the resulting information. In one approach to the problem, Hord and Brooner (1976) suggest to take average percentage of the classification accuracy (e.g. 90% accuracy), boundary accuracy (e.g. 95% accuracy) and the control coordinate accuracy (e.g. 97% accuracy) leading to an overall accuracy of 94%. In this section of the thesis much attention is given to the classification accuracy (the third aspect listed above). A more elaborate discussion is given below.

To assess the classification accuracy of discontinuous variables (attributes) associated with categorical coverage maps the following statistical parameters from Greenland, Socher and Thomson (1985), Moellering (1986), and Blakemore (1983) are often used:

(1) the percentage of correctly classified object per class or for a whole set of attributes;
(2) the expected percentage range correctly classified object per class or a whole set, usually stated as confidence level; and
(3) the percentage of correctly classified object per class or the whole set of attributes due to chance. This is related to the Kappa statistic. Its application is discussed in Chapter 9.

The three listed parameters require the percentage correctly classified object for their determination and a specified sample area for the third parameter. The percentage correctly classified can be obtained using external or internal testing and deduction method.

The deduction is closely related to resurveying of existing maps. A practical method for deducing quality is to use values from different but similar surveys that have attribute quality obtained either by external or internal testing.

External testing is completed by comparing the existing coverage map with the ground truth. In internal testing, several independent repeated surveys are involved. Some examples are given in Chapter 9.

4.3.2 Methods and techniques used

The process of attribute quality assessment was outlined as follows. The input for the assessment consisted of a coverage with sample values from representative locations, collected directly by field survey or from a higher accuracy source, e.g., a larger scale map as ground control. Then, this file is compared with the values stored in the system, for the same geographical locations. Through the overlay of the sample coverage with the system coverage to be assessed a two dimensional table is obtained. This table represents the frequency of every combination of values occurring between the sample and the system coverages. An error matrix can then be applied, which disposes the different classes or interval values observed in the sample coverage along the rows, and those observed in the system coverage along the columns (Chrisman, 1990; Greenland et al., 1985). The simplest case is depicted in the Table1.

Table 1: Typical error matrix

	System coverage classes		
Sample coverage level	**Class I**	**Class II**	**Total**
Class I	f_{11}	f_{12}	f_{1+}
Class II	f_{21}	f_{22}	f_{2+}
Total	f_{+1}	f_{+2}	Grand total

where

f_{ij}	:	frequency of combination between:"i"-sample class; "j"-system class
f_{i+}	:	sum of frequencies along row "i"
f_{+j}	:	sum of frequencies along column "j"
Grand total	:	sum of all frequencies, a measure of the sample size.

The error matrix can be obtained both for vector or raster data structures, using a proper overlaying technique. In the vector case, the coverage can be either a point or polygon coverage. If the sample coverage consists of labelled points, then a point-in-polygon search must be provided by the GIS, in order to derive the combination frequency table. If the sample coverage consists of classified polygons, a polygon overlay must be performed, and the areas of the polygon thus obtained can be a better measure to provide values to the misclassification (error) matrix.

For raster data structures a simple raster overlay is applied and the frequency of combinations is calculated on a pixel basis. For example the ILWIS-GIS software provides such a routine, in the module *"crossing"* under the option *"Spatial Modelling."* This routine computes the number of pixels for every occurring combination between sample and system classes, and stores this information in a cross table shown by extension (*. CTD).

The CTD table for the sample and system coverages can also be obtained by rasterizing polygon coverages, and then running the *"Crossing"* module. After this, probabilities concerning each system class accuracy and overall accuracy parameters can be obtained. The following algorithms were applied.

4.3.2.1 Probability of correct classification estimates for entity (class) level

From the misclassification matrix, the sum of pixels occurring in each row, i.e., the total number of pixels classified C_i (f_{i+}) is computed. This represents the sum of the diagonal frequency value (f_{ii}) plus the classification errors (f_{i+} - f_{ii}). A classification error occurs whenever a ground truth (sample) class is misclassified (confused) with different classes on the system coverage. The diagonal value represents the number of pixels correctly classified, for a given class. This means that the sample class coincides with the system class, for the pixel locations concerned. The value for the class probability is calculated as:

$$Pc_i = f_{ii}/f_{i+}\text{ , for every class "i"} \qquad \text{....eq.(2)}$$

where

Pc_i : class "i" probability that class of pixels assigned to c_i is really c_i
f_{ii} : diagonal frequency value
f_{i+} : row sum of frequencies.

Since the classified system coverage consists of polygons and their related attributes, it is easy to assign class probability values to every polygon in the coverage. For that purpose, this procedure creates a file with class probabilities associated with each class value. This last field works as a key field, to join the file produced with the entity attribute table.

4.3.2.2 Overall attribute accuracy estimates for coverage level

The error matrix can be used to obtain the percentage estimates concerning the classification accuracy at coverage level. These measures consist of the overall classification accuracy (Pov), and the Kappa statistic (K) which also can be obtained from this matrix.

The overall accuracy can be obtained using the ratio between the sum of all diagonal frequencies and the total sum of frequencies (the sample size). The formula is given below.

$$Pov = \left(\sum_{1}^{nc} f_{ii} / total \right) * 100 \qquad \text{......eq.(3)}$$

where

Pov : overall classification accuracy
nc : number of classes
total : sum of all frequencies

The value of Pov does not take into account, however, the fact that even if the classification was performed randomly, some values would fall in the diagonal cells of the matrix. To remove the effect of chance from the Pov value, the sums along the rows (f_{i+}) and along the columns (f_{+j}) must be included in eq. (3). Those sums represent the probability of correct classification above the chance agreement, and lead to the Kappa statistic. The formula is:

$$K = \frac{((Pov/100) - Pe)}{(1 - Pe)} * 100 \qquad \text{..... eq.(4)}$$

where Pe: expected probability due to chance (after random classification)

The value of Pe is given by the following formula:

$$Pe = \sum_{1}^{nc} ((f_{i+} / total) * (f_{+j} / total)) \qquad \text{.....eq.(5)}$$

In a particular situation the GIS user may require information quality stated as a confidence level, for example a true map accuracy of 98% if the actual accuracy found from a sample of 300 check points lies between 95.07% (as lower limit) and 99.08% (as upper limit). For this purpose the method developed by Topping (1972), Hord and Brooner (1976) as described below is used.

(1). Determine the confidence level (e.g., 95%) and find the ***"Za"*** value associated with such a percentage of the normal distribution curve (Topping, 1972). For confidence interval of 95%, Za value equals 1.96.
(2). Determine the number (*"N"*) for sample observations.
(3). Determine the proportion of correctly classsified sample check points (*"x"*) in the range 0.0 - 1.0 (e.g., 98%).
(4). Determine the true map accuracy, i.e., lower (y_1) and upper limit (y_2) of confidence interval (y) using the notations defined in (1), (2), (3) as follows:

$$0 < Za^2 * y\,(1-y) - N * x^2 + 2\,N*x*y - Ny^2$$

or

$$0 < (-Za^2 - N) * y^2 + (Za^2 + 2\,N*x) * y - N*x^2 \qquad \text{....eq.(6)}$$

From eq. (6) it becomes clear that confidence interval (y) consists of lower (y_1) and upper limit (y_2). These limits have to be always positive and derived through a calculation associated with a normal quadratic equation.

4.4 Attribute accuracy assessment associated with continuous data

4.4.1 General view

All types of measurement using various types of models either in physics or in other sciences, e.g., soil sciences, are inaccurate to some degree. Based on the concept of accurate value Topping (1979) defined *errors* as the difference or discrepancy between *the true and the observed value* of any physical quantity.

Because of errors, finding an accurate value, for example for a slope steepness measurement used in the FEDP, is very difficult. However, the difference or discrepancy between the expected and true value and observed values have to be made as small as possible by choosing the most accurate value within a confidence interval which can be measured and from which the degree of accuracy can be estimated. In other words, errors in a spatial data can be treated by using any usual statistical measures such as *standard error* or *standard deviation* of observations (Topping, 1979; Offermans, 1986 and Drummond,1987). The degree of errors of observation affects the accuracy and precision of measurement (Wheeler and Lyday, 1990; Lemmens, 1993). The term accuracy concerns the closeness of the measurement to the actual value or to the real value of the physical quantity while the term precision is used to indicate the closeness with which some measurements give identical results, neglecting systematic error.

Considering this, this section is intended to provide better insight into assessing the accuracy of continuous attribute data. The procedure implemented in this thesis performs a post-classification accuracy assessment, assuming that the user has a polygon map and a control sample set. In this study, these are values associated with a classified polygon map.

This type of map is made by *ranking a continuous variable*, e.g., slope steepness between *15 - 25 %*, into many classes followed by *verbal or fuzzy descriptions, e.g., flat, slightly undulating* of these classes and display of the class boundaries on the map as polygons. The map is usually produced in two stages.

(1) Analysis of the data collected at sample points to derive at a suitable classification.

(2) Depending on the type of data, two second stages are possible:

(a) computation of values for a regular grid of points, from the irregularly distributed sample points, followed by interpolation of isolines (which then are treated as polygon boundaries); or

(b) direct interpolation of polygon boundaries. This method is, for example, suited to a slope classification map, in which the polygon boundaries are drawn with reference to a contour map as well as using sample data values. In this type of map, adjacent polygons do not have to be adjacent in rank, unlike the isoline type of map (ICA, 1973).

Based on the concept as described above and considering the whole procedures in the establishment of the isoline or polygon map, the uncertainty in terms of error may be generated by (i) the decision which is made as to which numerical class an interpolated or observed value should be assigned. In other words, this problem is related to the probability of misclassification of observed values due to measurement errors and ranked intervals. Error may also be generated by (ii) class boundaries of slope steepness which are defined arbitrarily by the user to classify observed values on a continuous or fuzzy boundary may cause problems, particularly if the observed values fall near the boundaries of the class. Above all, these show the requirement of a proper measure for handling and representing uncertainty associated with spatial data.

With regard to problem (i), there are several methods which can be applied for the grid and isoline interpolation steps, but the estimation of final errors through normal methods of variance/covariance propagation is rather complex. Ideally within a GIS, there should be a supporting system or module which assists the user interactively in deriving the probability of each observation point to be in the correct ranked interval (as well as its associated membership value). It should also compute mean values, which estimate the overall accuracy of the coverage, provided a sufficient number of well distributed points is being interpolated.

A package developed at Utrecht University, so called ADAM, can actually perform error propagation routines during interpolation, though it works in the raster domain (Wesseling and Heuvelink 1991; Heuvelink, 1993).

With regard to problem (ii), Kandel (1986) and Fisher (1992) suggest to apply the concept of *continuous or fuzzy classification with values varying between 0 and 1*. Associated with this concept they introduced a *membership function*, the *central kernel of the class* and a *parameter* that determines the *width* of the transition zone, around the kernel of the class.

Considering the variability of the terrain, the kernel of a class is often not a single value, but a range of values (Burrough, McMillan and Van Deursen, 1992) which includes the upper and lower transition zones. Considering this, the situation and the membership function introduced by Kandel (1986) was adjusted as depicted in Figures 8, 10 and equations 11, 12 and 14.

The entire process can be done in two phases, as follows: (i) computation of the classified polygon map precision;(ii) computation of the probability of correct ranking. A more elaborate discussion is given in the following Subsubsection.

4.4.2 Methods and techniques used

4.4.2.1 Computation of the slope steepness map precision

A sample *control* file, with attribute values obtained from a higher accuracy source, e.g., field survey or a larger scale map, is compared to slope steepness classes obtained from the slope steepness (*test*) map. As a general rule it can be stated that control point values near the centre of each slope steepness class have a very high probability of being correctly classified. If the classes are sufficiently broad, this probability becomes 1(*one*) , i.e., *certainty*.

Toward the class boundaries, there is an increasing probability of the control point values being incorrectly classified. The greatest probability of incorrect (confusion) classification occurs for values falling exactly on class boundaries.

It should be noted that, for this type of map, while it is generally the case that points falling near the map polygon boundaries have a high probability of being misclassified (confused), the opposite is not necessarily true, i.e., points near the centre of a polygon may also be misclassified.

There are two approaches to measuring the precision of the steepness map. In the first case, only misclassified sample points are noted. For each class boundary, it will be found that the distribution of misclassified points follows a normal distribution with the highest frequency of error for actual slope very near the class boundary. The standard deviation (*sd*) is then used to estimate the probability of misclassification. This probability accounts for the attribute classification only, irrespective of ground distance from the class polygon boundaries.

For the second approach, at the sample locations, the high accuracy slope steepness value are compared to values derived from measurements made on the 1: 50 000 topographic maps (i.e., data collection method used to create the slope steepness). The error distribution, after the removal of any systematic error, will fit the normal distribution. In turn, in a way very similar to the first approach, standard deviation can be used to find the probability of misclassification near the class boundaries.

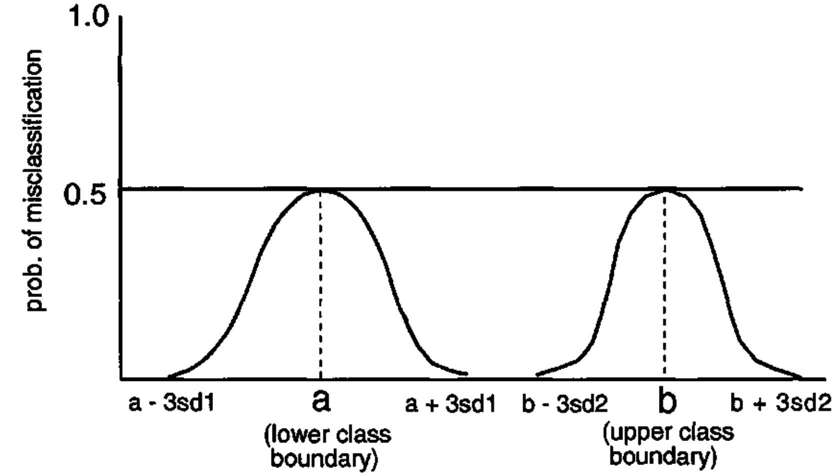
For a sufficiently broad class, these normal curves, associated with the upper and lower boundaries respectively, do not overlap (taking mean more or less equal (3 * sd) as the limit of the curve). The zone between the limits of the two curves is *"the kernel of the class."* See Figure 8a.

$$|Vc_i - Vs_i| = error\ (e_i); \quad \sum_{1}^{n} e_i / n = x_i\ (error\ means)$$

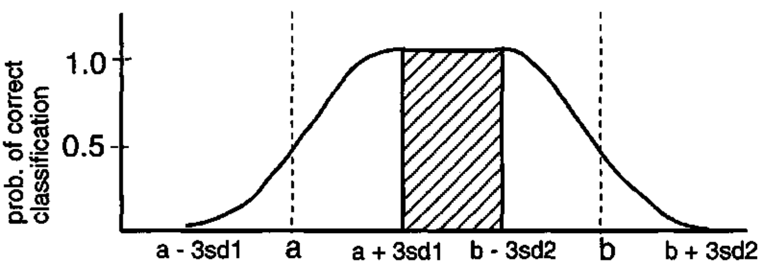
$$sd = [\ ((\Sigma\ (x_i - e_i)^2\ )/n)\ ]^{0.5} \qquad \text{......eq.(7)}$$

where

sd : attribute values standard deviation; vc_i : attribute values of the control samples; vs_i : attribute values obtained from test map; ns : number of samples



a. Fuzzy class boundaries
where : sd 1 = standard deviation around a
sd 2 = standard deviation around b



b. class kernel and transition zone

Figure 8: Fuzzy boundaries of ranked interval

Figure 8b shows that the values *sd1* and *sd2* determine the width of the transition zones on either side of the class kernel. If the value of *sd is zero (0)*, the class kernel coincides exactly with the class itself, yielding a *Boolean membership function* with value 0 (outside the class boundary a or b) or 1 (inside the class boundary a or b). In the general case, where $sd \neq 0$, there exists on each side of the kernel a transition zone described by a membership function whose value decreases from 1 to 0 (see again Figure 8b). This is used to assign a membership degree of observations belonging to this particular class.

The calculation of slope steepness map precision was used in an iterative procedure to define slope steepness class boundaries such that each class contains a kernel with probability one of correct classification.

4.4.2.2 Computation of the probability of correct ranking

After the value of the map precision is provided, the interval range input defined by the user can be tested using an interval performance test. This method considers the resulting computed precision, e.g., the dispersion of values around the lower and the upper bounds may or may not overlap each other. A significance level input by the user will allow also the modification of the extent of possible variation around the interval boundaries.

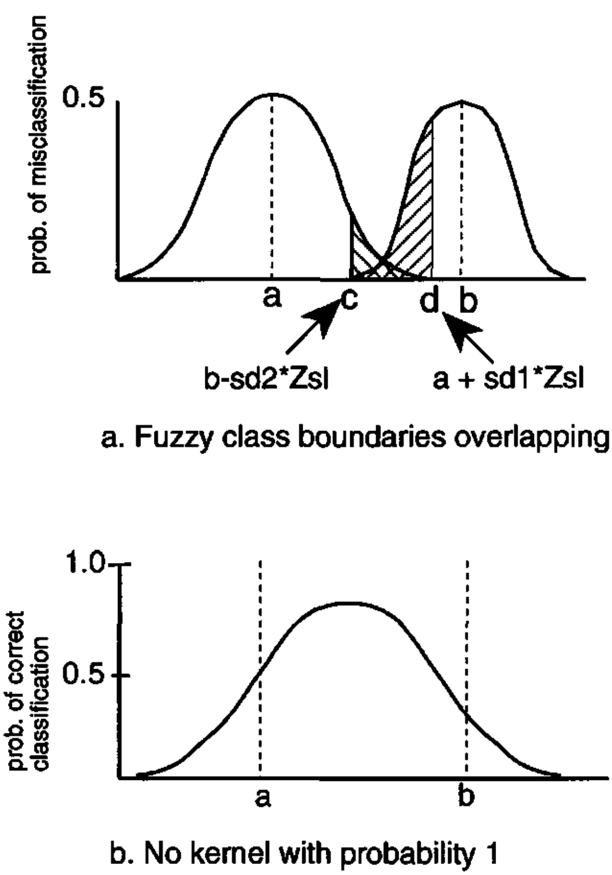


Figure 9: Overlap Probability, in a ranked interval

The overlap, if it exists, is expressed in the probability domain. The situation is described in Figure 9 (a).

The following equation is applied to compute the overlap probability:

$$Pov = 1 - Prob[(b - (sd * Zsl) - a)/ sd] \qquad \text{....eq.(8)}$$

where
a : lower bound of the ranked interval; b : upper bound of the ranked interval; sd: standard deviation around a and b;Zsl : normalized random variable value for the chosen significance level (sl);Prob: function that returns the probability value for the normalized variable computed between brackets.

A conclusion can be drawn for example a value for the overlap probability greater than zero, representing a poor quality isopleth or polygon map. It means that the interval range chosen was too small for the sample density and resolution available.

Therefore if the interval range was appropriately defined, the upper limit of dispersion of values to the right of the lower bound would be represented by a smaller value than the value of the lower limit of dispersion to the left of the upper bound. This is the so-called interval discrimination. It is given as a percentage of the interval range. It is computed using the following formula:

$$Discr = (((b - (sd_2 * Zsl) - (a + (sd_1 * Zsl))) / int.) * 100 \qquad \text{....eq.(9)}$$

where
Disrc : discrimination, in percentage; int : interval size = (b - a)
(other variables have the same meaning as in eq. 8)

In this situation, applying eq. (8) results in zero (0) overlap probability, conversely the equation shows that all the attribute values falling within the discriminated portion of the ranked interval will have a probability of one (i.e., the correct ranked interval).

After the interval performance test, probabilities of correct ranking are computed for each sample attribute value, considering three possible cases (see Figure 10):

(a) The sample value falls *between the lower boundary and the kernel of the class*, i.e., *if* $v_i < (a + 3sd_1)$. The formula used is

$$Pv_i = Prob[(v_i - a) / 3sd_1] \qquad \text{.....eq.(10)}$$

where
Pv_i : probability of correct ranking for sample value "i"
v_i : attribute value for sample "i"

According to Burrough and Heuvelink (1992), Fisher (1989) and Heuvelink (1993) the membership value for the same case is computed as follows:

$$MFv_i = 1 / [1 + (v_i - a - 3sd_1)/3sd_1)^2] \qquad .....eq.(11)$$

where MFv_i: membership function value for sample "i"

(b) The sample value is *placed within the kernel of the class or within the interval discriminated area*, i.e., if $(a + 3sd_1) \leq v_i \leq (b - 3sd_2)$. The probability and membership values are both equal to one, meaning that the value is in the correct rank,

$$Pv_i = 1 \quad ; \quad MFv_i = 1 \qquad ....eq.(12)$$

(c) The sample value falls between the dispersion of the upper boundary and the kernel of the class, i.e., *if* $v_i > (b - 3sd_2)$. The formula for this case is,

$$Pv_i = Prob[(b - v_i) / 3sd_2] \qquad ....eq.(13)$$

and the membership value is given by the formula

$$MFv_i = 1 /[1 + (v_i - b + 3sd_2) / 3sd_2)^2] \qquad ....eq.(14)$$

remark: the variables mean the same as in eqs. (8) to (14).

The graphical representation of the three cases is shown in Figure 10.

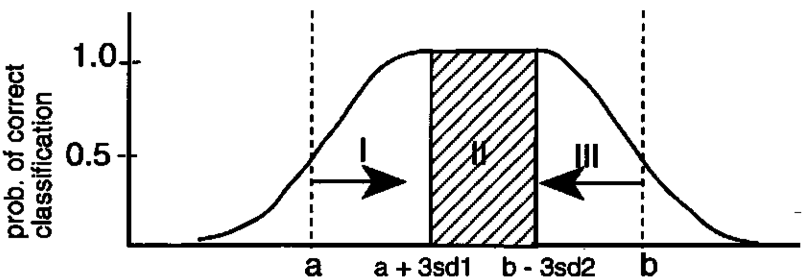


Figure 10: The three possible cases of ranked interval probability[*)]

Note: The "Prob" function computes the area under the normal curve, for normalized variables. That area is defined between the normalized variable and the value of Zsl (defined in eqs. 8,9).

4.5 Plausibility reasoning and an inductive approach

4.5.1 General view

Because of the different forms of uncertainty, the concept of fuzzy measures has been proposed for handling uncertainty associated with spatial and digital geographic information. In this study, these measures are related to the degree of belief in assigning evidence to a particular set.

A relevant example related to the topic of this thesis concerns the representation of a fuzzy measure concerning *"the presence"* or *"the absence"* of a particular erosion class on a particular area. The occurrence of soil erosion either that are *absent* or *present* is considered to have *very distinct, well-defined boundaries*. This illustration implies that the *original degree of belief* in the degree of presence or absence can be *increased* or *decreased* as more evidence is collected. Thus, the fuzzy measures are associated with the situation when one has to search for perfect evidence to establish an underlying process, e.g., presence or absence, in *which full membership in one and only one is allowed*. This makes it different to the concept of Fuzzy Set Theory which is more concerned with assigning a value to each element of a universal set, representing *its degree of membership* in a particular set with unsharp or gradual boundaries. See further in Section 10.2.2.

4.5.2 Notation of the plausibility reasoning

Plausibility is a fuzzy measure associated with belief (Section 4.1). Therefore, plausibility reasoning is *"a specific case of evidential theory "* which is concerned with collecting and selecting *confirming* (positive effect) and *disconfirming* (negative effect) of evidence involved in inferring an underlying process. In order to maintain the commutativity of various evidence some adjustments have been made as described by Heckerman and Horvizt (1986); Bonnet (1985).

The plausibility reasoning in this thesis is related to a *Certainty Factor (CF) Model*. It is considered as a further development of Heckerman's parallel combination model or a confirmation function. With regard to the occurrence of erosional processes plausibility reasoning is related to the measurement of evidential strength using belief and disbelief as units of observation. The unit of disbelief measure is associated with a disconfirmation as an adjunct to a measure for degree of confirmation.

In relation to the occurrence of soil erosion class (*h*) in a particular area with evidence (*e*), the notion of the plausibility measure is defined as follows:

MB [h,e] = x (used as a measure of *increased belief in the hypothesis h,* based on given evidence e)

MD [h,e] = y (used as a measure of *increased disbelief in the hypothesis h,* based on given evidence e)

It has to be noted that evidence e is not always an observed pattern but may be a hypothesis. Thus it may allow one to express MB[hl,h2] to represent *the measure of increase of belief* (or decrease in disbelief) in the hypothesis hl given that the hypothesis h2 is true which than becomes evidence. On the other hand MD[hl,h2] is *the measure of increased disbelief* (decreased belief) in hypothesis hl if hypothesis h2 is true. Based on this understanding one can draw the conclusion that P (h) reflects one's belief in h at any given time. Thus 1 - P(h) can be considered as an estimate of the expert's disbelief regarding the truth of h. If P(h/e) is greater than P(h), the observation e increases the expert's belief in h while decreasing his or her disbelief regarding the truth of h. Based on this understanding then the proportionate decrease in disbelief MB [h,e] can be formulated as follows:

$$MB\,[h,e] = [P(h/e) - P(h)] / [1 - P(h)] \qquad \text{....eq.(15)}$$

The proportionate decrease in belief MD [h,e] can be formulated as follows:

$$MD\,[h,e] = [P(h) - P(h/e)] / P(h) \qquad \text{.....eq.(16)}$$

Related to the situation as described above, the concept of CF facilitates thinking about confirmation and the quantification of degrees of belief used as a measure to obtain the combined effect of MD and MB as follows:

$$CF\,[h,e] = MB[h,e] - MD[h,e] \qquad \text{....eq.(17)}$$

The formula for determining CF [h,e] involves the following determinants:

(1) *Range of degree*

a. $0 \le MB[h,e] \le 1$
b. $0 \le MD[h,e] \le 1$
c. $-1 \le CF[h,e] \le 1$

(2) *Evidential strength and mutually exclusive hypotheses: If h is shown to be certain [P(h/e) = 1]*

a. $MB\,[h,e] = \dfrac{1 - P(h)}{1 - P(h)} = 1$ (*from eq.15*)
b. $MD\,[h,e] = 0$ (it means *no decrease in belief is possible*)
c. $CF\,[h,e] = +1$

If the negation of h is shown to be certain, that is *zero belief*, then P(h/e) = 0

d. $MB\,[h,e] = 0$ (it means *no increase in belief is possible*)
e. $MD\,[h,e] = \dfrac{P(h) - 0}{P(h)} = 1$ (*from eq.16*)
f. $CF\,[h,e] = -1$

(3) *Lack of evidence*

a. MB [h,e] = 0 if h is not confirmed by e or e disconfirms h.
b. MD [h,e] = 0 if h is not disconfirmed by e or e confirms h.
c. CF [h,e] = 0 if e neither confirms nor disconfirms h.

note: h and e are independent

4.5.3 Discussion

The general notation of plausibility reasoning of (1) ,(2) and (3) above using the CF concept as a plausibility measure is relevant in the context of observational procedure and erosion study using the proposed IEM.

Plausibility reasoning is a special case of evidence theory. It was evolved from classical probability theory and includes the proportionate decrease in *disbelief MB[h,e]* and the proportionate decrease *in belief MD* [h,e]. These measures are related to uncertainty associated with *finding, collecting, selecting* and *choosing* perfect evidence to *confirm* or *disconfirm* a particular erosion class.

The accelerated erosion processes in this thesis deal with active processes and phenomena. The combination of spatial pattern and distribution of erosional process associated with a particular TMU is identifiable on remote sensing images and aerial photos (Chapters 6, 7 and 8). These have been studied in the field and through a stereo viewer to determine environmental characteristics such as cover, land use, slope and soil erosion. Concerning the proposed IEM, the adopted observational procedure associated with the concept of TMU, i.e., *stepwise specification* or *generalisation* is considered as a method to find necessary facts or knowledge at different aggregation hierarchies.

In this regard, the knowledge of erosion influencing factors, erosion classes and their relationships on particular area facilitates the dialogue between the expert and the proposed IEM in collecting and evaluating the confirming (MB) and disconfirming (MD) erosion influencing factors. This implies that the conclusion on the presence or absence of soil erosion can only be drawn based on combined evidence.

Considering this and the effect of causal relationships among erosion influencing factors, the CF adopted in an IEM does not allow maintaining a CF value between 0 and 1 (see Table 30), as used in a degree of belief measure. Instead the CF can take any value between +1 (in this case, definite allocation to a soil erosion class) to -1 (definite exclusion from a soil erosion class) to express change of belief. It is based on *subjective judgement by experts,* and its value depends on the original hypothesis. However, in particular circumstances the soil erosion process is very complex and change in belief cannot be easily justified (CF equals to 0). In this regard, more evidence is required for further justification. A more elaborate discussion is given in Chapter 10.

The particular strength of this approach is that it can be used in situations in which there are no quantitative data concerning erosion, for example in tons/hectare/year. It may, of course, be supplemented by detailed, quantitative measurements to find the actual erosion in each class.

Chapter 5

Study area and datasets used

5.1 General description of the study area

Location
The Ciseel Subwatershed is located in the southern part of the island of Java, with its approximate geographical centre, Banjar, lying about 330 kilometers to the southeast of Jakarta (see Figure 11). The basin lies between approximately 7° 20' and 7° 41' South latitude and 108° 26' and 108° 29' East longitude.

The Ciseel Subwatershed is an intermontane valley subwatershed, which in turn, falls within the Citanduy Watershed, and it is bounded on the east by the Sagara Anakan Subwatershed, on the north and west by the Upper Citanduy Subwatershed, on the south by the Indian Ocean, and on the south west by the Ciwulan Subwatershed.

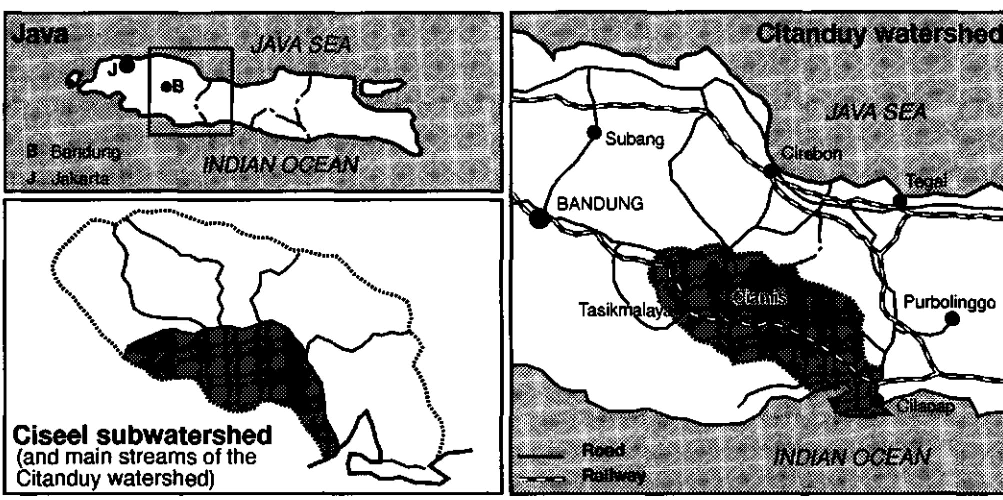


Figure 11: The location of the Ciseel Subwatershed

Area
The total area of the Ciseel Subwatershed basin is 96.500 hectares (965 km^2) consisting of six distinct subsub watersheds, namely: (i) Cikembang (ii) Upper Ciseel (iii) Citalahab (iv) Ciputrahaji (v) Cihapitan and (vi) Cikaso.

For the purpose of the study, the upper watershed was defined as those areas in the basin containing sloping uplands which are naturally more prone to erosion. The lower watershed was defined as those areas consisting of lowlands which exhibit little erosion hazard.

Using these admittedly subjective technical definitions, the Ciseel Subwatershed contains approximately 19,850 hectares of lowland and 76,650 hectares of upland. However, there is no way to neatly demarcate the upper watershed from the lower watershed on the basis of subwatershed boundaries or merely on the basis of elevation.

Climate

The climate in the Subwatershed is generally humid with a relative humidity of about 85 - 90%. The subwatershed also has a high relatively constant temperature averaging 25 degrees Centigrade. However, there is considerable variation among different parts of the region due to differences in elevation and the heavier precipitation that normally occurs at the higher elevations. As a general rule, the average temperature will fall by about four degrees Centigrade for every 500 meters increase in elevation. Thus, for example, Tasikmalaya at about 350 meters elevation is considerably cooler than Banjar, which is only 45 kilometers away but at an elevation of only about 100 meters.

Records from more than 25 stations in the basin show a spatial variation in long-term average rainfall from just above 2,200 mm to 3,600 mm, with an average annual rainfall of about 3,000 mm. Relief and elevation play an important role in determining precipitation at any one location, with rainfall increasing with higher elevations as a general rule. Thus, the lower watershed receives 2,400 - 2,800 mm annual rainfalls while the relatively higher elevations of the upper watershed have an annual rainfall of about 3,000 mm. Although there are no recording stations at the highest point of Mt. Sawal (2,158 m) or in the hills forming the northern boundary of the basin, it is likely that rainfall exceeds 4,000 mm at these elevations. Conversely, there appears to be a rainshadow effect which lowers precipitation in the central part of the basin. See Figure 12.

The rainfall, which is characterized by high intensity storms of short duration and limited areal extent, is determined largely by the influence of the northwest and the southeast monsoons. During the period from approximately November through April the northwest monsoon predominates, picking up large amounts of moisture over the Indian Ocean and bringing to the basin the heaviest precipitation. Although there is considerable variation in the amount and distribution of rainfall from year to year, most places in the basin receive about two-thirds of their rainfall during the November-April period. The period from May to October is the so-called dry season. The southeast monsoon predominates during this period bringing with it smaller amounts of precipitation due to the lower atmospheric moisture caused by lower temperatures in the southern hemisphere at this time of the year. Many areas of the basin receive only about 10% of their precipitation during the July-September period and although data on evaporation is somewhat sparse, information from Tasikmalaya and Cilacap indicates that there is a strong probability of an excess of evaporation over precipitation during these three months, even though the margins tend to be fairly small. However, to lay stress again on the facts that there are significant fluctuations in rainfall from year to year and from place to place within the study area, the point is made that some of the heaviest downpours and the floods that result from them occur during the May-October period.

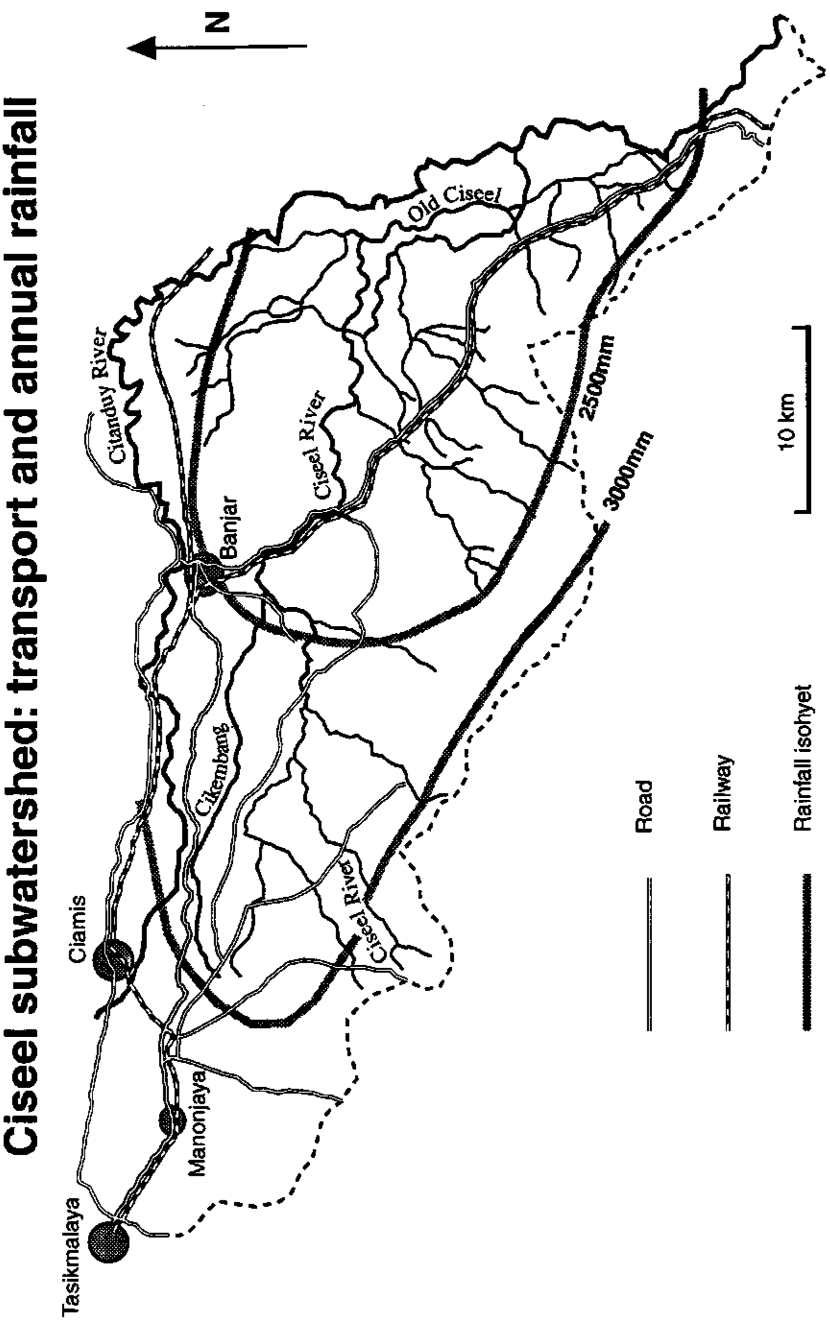


Figure 12: Transport and annual rainfall in the Ciseel Subwatershed

Terrain and geological features of the Ciseel Subwatershed
The Ciseel Subwatershed is formed on the Bandung zone, which consists of a longitudinal belt of the intermontane depression starting approximately at Tasikmalaya (351m) in the west and ending in the Segara Anakan at the south coast of Central Java (Van Bemmelen 1970). The Bandung zone trends northwest/southeast. Volcanoes such as Galunggung, Sawal and Guntur form characteristic features of the skyline. The Bandung zone is predominantly filled with young volcanic, alluvial-colluvial and alluvial deposits.

The lowlands of the Ciseel Subwatershed are dominated by mountains and hills near Banjar at the western edge of the former Lakbok swamp. A low mountain ridge extends from Wanareja to the southeast. These hills and ridges are formed mainly by Miocene sedimentary and some volcanic rocks. Due to the presence of this ridge, a small intermontane valley is formed by the Cikawung River, thereby separating it from the Bandung zone mountains. The geologic composition of the Ciseel Subatershed is summarized in Table 2.

Table 2: Geologic composition of the Ciseel Subwatershed (RMI,1986)

Geologic material	Area (Ha)	Area (%)
Alluvium	30,531	31.63
Recent volcanic rocks	9,768	10.12
Pliocene sedimentary rocks	19,690	20.40
Miocene sedimentary rocks	34,254	35.49
Miocene limestone	2,257	2.36
Total	96,500	100

Factors affecting land utilization

Slope
The data in Table 3 shows that the Ciseel Subwatershed is composed of land units with dominant slopes as follows:

Table 3: Land units with dominant slope (RMI, 1986)

Physiographic description	Dominant slope (%)	Area (ha)	Area (%)
Flat	0 - 8	34,587	35.84
Undulating	>8 - 15	5,741	5.94
Moderately steep	>15 - 25	20,521	21.27
Steep	>25 - 45	25,889	26.84
Very steep	>45	9,762	10.11

About 36% (34,587 hectares) of the basin is composed of land with slope 0 - 8%, which is a generally favourable topography for the cultivation of food crops. An additional 54% (52,151 hectares) of the basin is composed of land that, although soil conservation measures are necessary and terracing, strip planting, contour planting, hillside ditches, etc. must be implemented as appropriate, can still be used for sustained upland agricultural production with only limited hazards and constraints (>8 - 45 %).

However, 10% (9,762 hectares) of the basin is composed of land which requires very careful soil conservation and land management and where diversified tree crops, agro-forestry and forestry are the best land utilization options due to the need for a permanent vegetative cover and strict conservation measures because of the slope factor alone. In fact, the Ciseel Subwatershed mostly consists of predominantly steep slopes (15 - 45%).

Erodibility of geologic material

In the course of field investigations it was established that a close relationship exists between the geologic origin of soil parent material and erodibility of various soils. (Resorces Management International (RMI), 1986).

Three basic levels of soil/surficial geologic material were established:
(i) soils of low and low to medium erodibility - derived from recent and Pliocene volcanic rocks; ii) soils of generally high erodibility - derived from Pliocene sedimentary rocks; and (iii); soils of very high erodibility - derived from Miocene sedimentary rocks.

In accordance with the above established ranking of the erodibility of geologic material, it appears that about 56.201 hectares of the Ciseel Subwatershed or 58 % of the total area is composed of very highly erodible surficial geologic material, (see Table 4 and Figure 13). The soils derived from these Miocene sedimentary parent materials are highly erodible with high on-farm erosion but also high slope instability and river bank and roadside erosion. See also Table 3. At the same time, roadside and river bank erosion are also very significant on soils derived from recent and Pliocene volcanic rocks, although on-farm erosion on these soils is generally low (RMI, 1986).

Table 4: Soil types classified by geologic origin and soil erodibility (UNPAD, 1983; RMI, 1986)

Soil types*)	Soil erodibility	Area (Ha)	Area (%)
Kambisol, gleisol, alluvial	Very low erodibiity	42,120	43.65
Latosol	Low erodibility	11,952	12.39
Mediterran	Low to medium erodibility	6,162	6.39
Yellow red Podsolic	High erodibility	35,612	36.90
Litosol, Renzina, Regosol, Organasol	Very high erodibility	654	0.67

Note: *) Adopted from the Indonesian Soil Research Institute Classification Systems

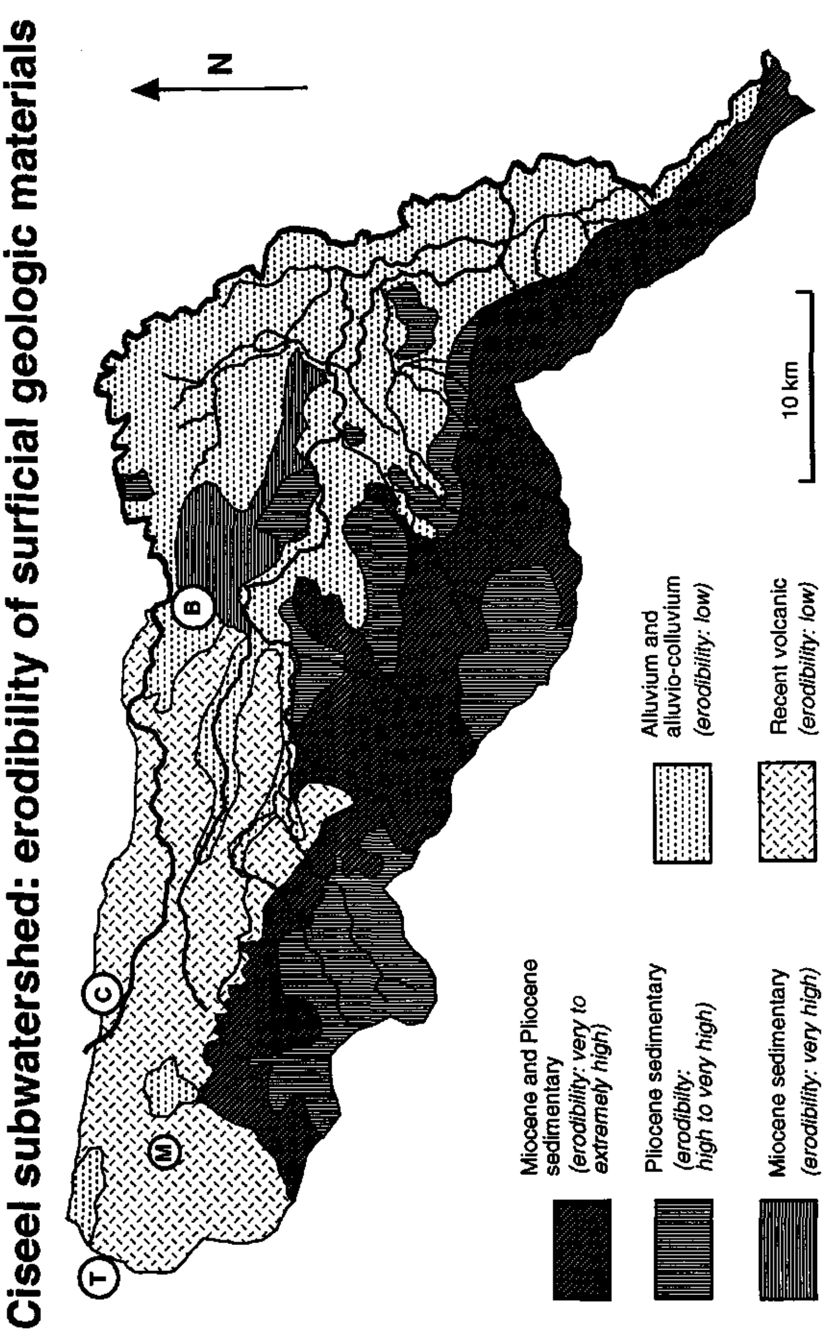


Figure 13: Erodibility of surficial geologic materials in the Ciseel Subwatershed

Present land use

The land use mapping of the study area was done by Resource Mangement International (RMI) (1986) on the basis of the characterization of crops and cropping patterns as identified by satellite imagery and checked with aerial photography and through field visits (Appendix 1). In this respect, mapping units were described so as to identify major individual crops. It was updated using TM 121/65 (1991). The quality of this mapping activity has been assessed as presented in Chapter 9.

The mapping established that about 19 750 hectares or 20.47 % of the basin is under *sawah* cultivation. This figure also includes the rice grown in the uplands in valley bottoms or where rain-fed *padi* with supplementary irrigation is grown on steep slopes outside of the alluvial soils. See Figure 14 and Table 5.

Table 5: Land use classes in the Ciseel Subwatershed (RMI, 1986)

Land Use Classes	Area (Ha)	Area (%)
Rice Field	19,750.3	20.47
Mixed garden	35,879.0	37.18
Coconut plantation	15,533.0	16.10
Rubber plantation	2,276.0	2.36
Secondary forest	811.3	0.84
Protection forest	93.4	0.09
Teak forest	5,397.0	5.59
Swamp	82.1	0.08
Settlement	5,030.3	5.21
Mangrove forest	98.2	0.10
Dry land agriculture	9,736.7	10.09
Bush	1,810.9	1.88
Bare land	11.2	0.01
TOTAL	96,500.0	100.00

Approximately 37 % of the area consists of mixed gardens with coconut (*Cocos mucifera*) being the most dominant tree crop. Other crops include mango (*Mangifera indica*), cloves (*Eugenia aromatica*), coffee (*Coffea robusta*), durian (*Durio zibethinus*), and rambutan (*Nephelium lapacum*). This mapping unit also includes a significant proportion of dryland upland agriculture, agro-forestry, and general mixed cropping.

The largest and most productive proportion of dryland agriculture is located in *Kabupaten* Ciamis. There, in its southern part, are well maintained, diversified tree crop plantations which provide good examples of appropriate soil conservation practices.

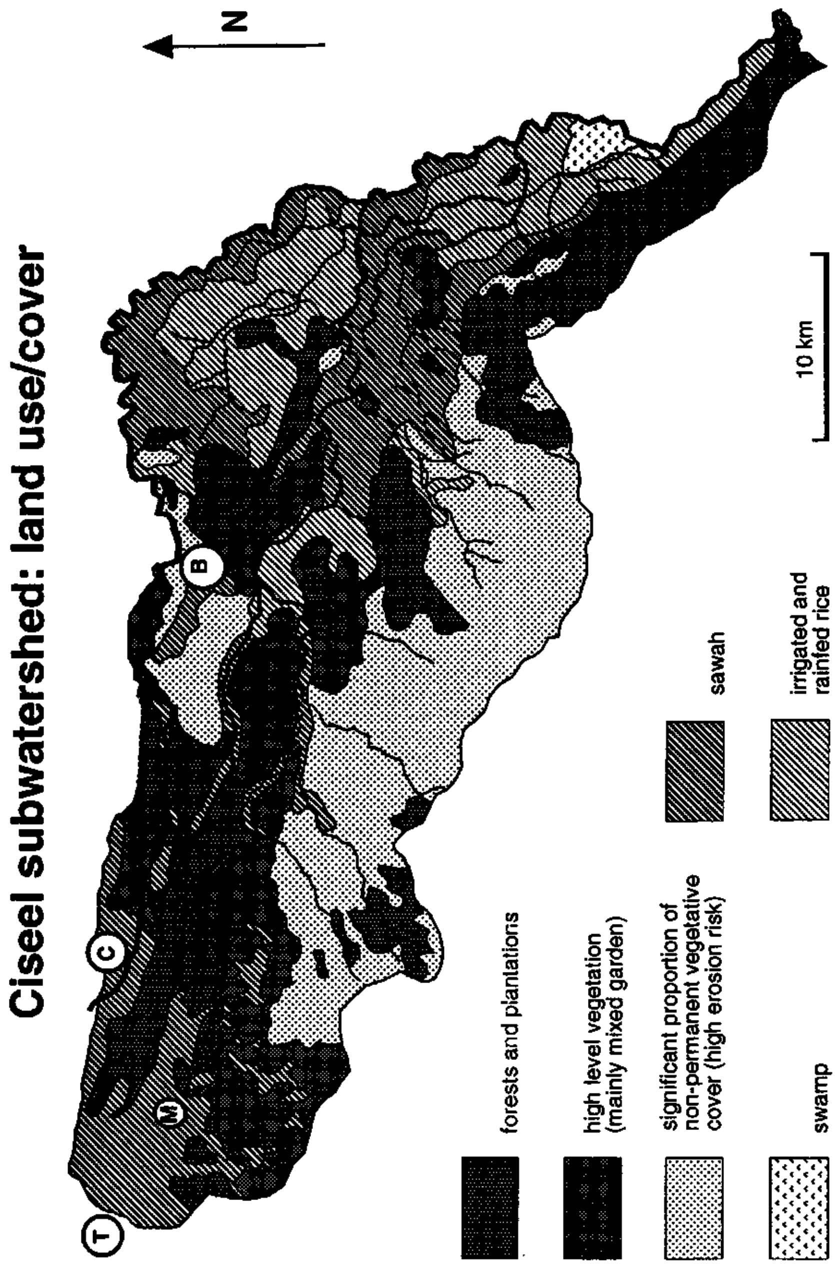


Figure 14: General landuse and ground cover in the Ciseel Subwatershed

Forests were mapped on 6.302 hectares or 6.5% of the basin's total land area, including about 99 ha (0.1%) of coastal mangrove swamp forest. A summary of the established land use mapping units is presented in Table 5 which contains a breakdown of present land use by land use classes for the entire study area.

5.2 Data sets used

Geographical data

During the implementation of the Citanduy II Project (1982-1988), an interdepartmental cooperation project (cofinanced by USAID) conducted various studies including a natural resources inventory in the entire Citanduy Watershed.

In the year 1986/87, rapid field checking and the external testing method developed by Chrisman (1984) was carried out by the FEDP Team. This testing included various data on soil type, land use, administrative boundary, geology, slope steepness, geomorphology, and terrain classification used in the establishment of the FEDP. For the purpose of the study these datasets were collected, digitised and stored in a GIS.

The main sources of those maps are (i) topographical maps at scale 1 : 50.000 produced by US-ARMY in 1955, (ii) aerial photo interpretation, at scales 1: 15 000 and 1: 50 000 surveyed by Suravia Jaya Ltd. made in 1982.

Remote sensing data

The high resolution RS data are Landsat Thematic Mapper (TM) images that were acquired in the years 1986, 1987 and 1991. They are all very good (relatively cloud free) images.
A combination of several RS i.e. Landsat TM 121/65 images is considered very convenient for updating existing landuse data. The application of limited bands was considered most efficient. From experience it is apparent that an optimal band combination to observe natural vegetation consists of two visible and a near infrared band (Epema, 1986). In this study the combination of band 2,3, and 4 is used.

The application of RS data either from TM or SPOT images in identifying terrain objects e.g. terrain system is discussed in Chapters 6, 7 and 8.

Software used

The ERDAS 7.5 Software was used in processing Landsat Thematic Mapper data (ERDAS, 1991). In addition to this IDRISI, ILWIS and Arc/Info were used. To formulate special commands the ERDAS tool kit was used. This combination of software allows the facilitating of data integration, aggregation and conversion from raster to vector and vice versa.

Chapter 6

The establishment of functional observation units: TMU classes for describing terrain object characteristics at the sub-regional level

6.1 The establishment of TMUs: the application of aerial photo and SPOT image interpretation

This section describes a method of the observation of TMUs with their classes and class composition. This method was based on the concept of spatial object models as discussed in Chapter 3.

The interpretation processes were carried out on panchromatic black and white aerial photographs at scales 1: 15 000 and 1: 50 000 (1982) and SPOT standard film (1987) for both monoscopic and stereo analysis interpretation. The ITC system of geomorphological mapping methodology and techniques as developed by Van Zuidam et al. (1985) was applied to assist the aerial photo interpretation procedure in differentiation and delineation of units according to the following aspects of the terrain:(i) main origin of the landforms and relevant specific origin e.g. mass wasting within the denudational origin; (ii) geology, emphasizing lithology; (iii) morphometry; internal relief and degree of dissection; (iv) internal relief-forms (most units can be differentiated at 1:100.000 scale which are composed of smaller, more homogeneous subunits or catena elements); (v) actual processes and hazards. This study emphasizes erosional processes, but for fluvial units flooding has been included.

Adopting the general methodology as described above, and using photomorphic characteristics of the SPOT and other satellite images (see Appendix 2 for legend for Landsat-based terrain classification, as used in Indonesia), the establishment and the extraction of terrain object characteristics associated with TMUs was obtained through processes as depicted in Figure 15.

Figure 15 shows that the entire process consists of: (i) definition and the identification of TMUs as terrain objects representing the basic spatial object for terrain characteristics inventory at sub-regional level; (ii) using the interpretation of satellite imagery and aerial photo keys for the determination of the terrain division with homogeneous relief and size of the units followed by fieldwork; (iii) data extraction and definition of TMU characteristics, i.e., observing spatial pattern; (iv) the differentiation and classification of TMU classes and composite classes; (v) TMU mapping in terms of homogeneous relief and terrain characteristics. A detailed discussion on these processes is given in Sections 6.2, 6.3 and 6.4.

Related to studies as discussed in Chapters 2 and 4, Figure 15 explains that the aerial photo interpretation at difffrent scales may serve the following: (i) the identification of specific information on active processes, e.g., soil erosion that may exist in the photo and which can be used for further processing and classification purposes; (ii) the evaluation of the satellite image interpretation results and other relevant existing thematic maps; (iv) sub-regional scale representation.This shows that the interpretation of aerial photos plays a very important role in delineating boundaries and the aspects of the terrain.

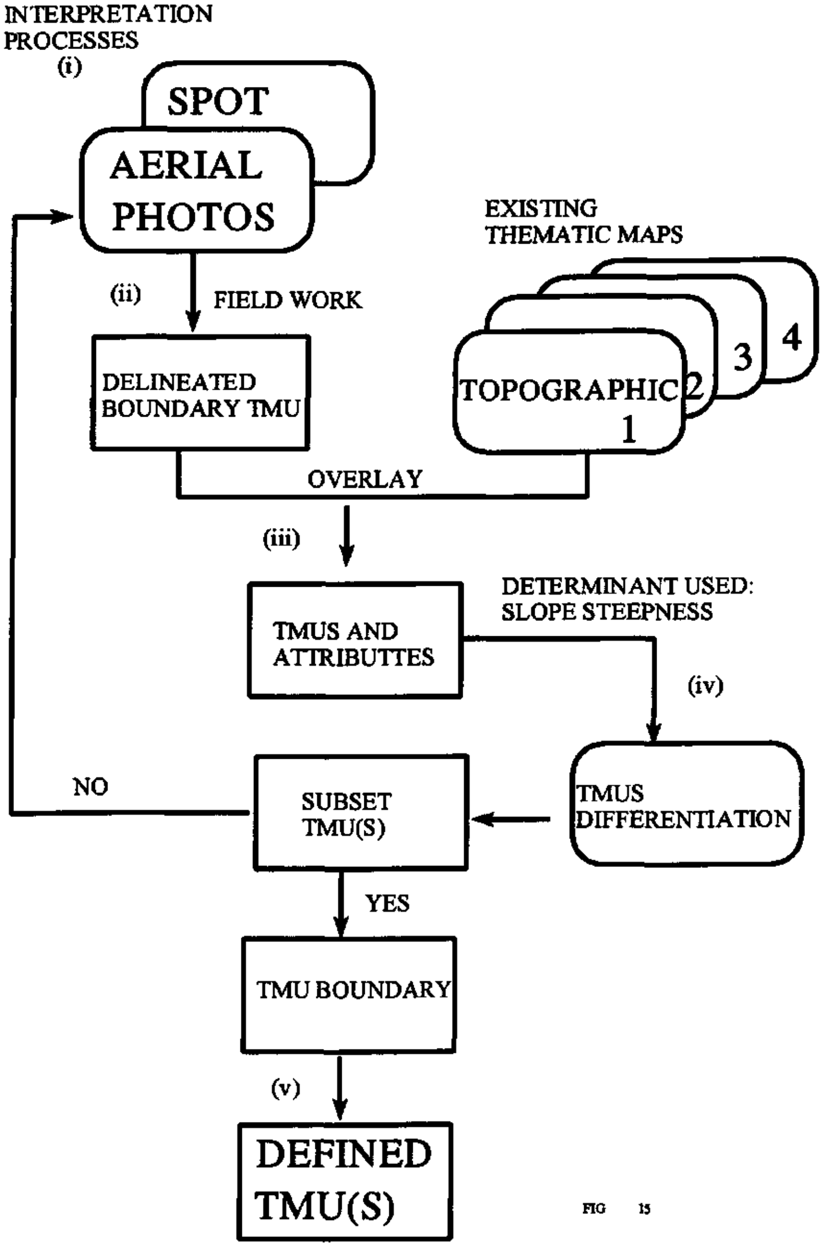


Figure 15: TMU identification procedure

It should be remarked that the definition and object classification adopted by the traditional aerial photo interpretation in general and the traditional establishment of TMUs in particular, are not recognised as independent processes: they are inseparable and interrelated as depicted in Figure 16 (b) below. A risky consequence of this procedure, is that confusion between object classification, e.g., soil class and object definition of soil units inevitably occurs.

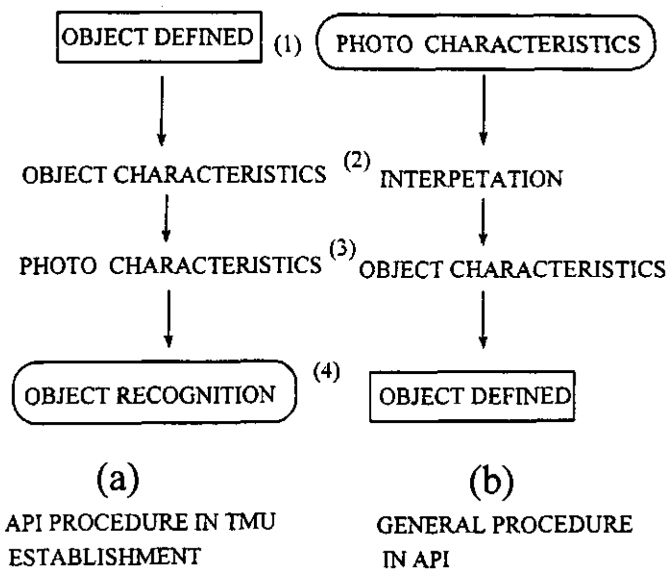


Figure 16: (a) Aerial photo interpretation (API) procedure adopted in the establishment of TMUs
(b) Traditional aerial photo interpretation (After Huising, 1993)

As shown in Figure 16 (b) in the traditional procedure knowledge of the image representation of the possible object TMUs is necessary, as well as knowledge of the region, e.g., through field checking and incorporating ancillary data. Aerial photo interpretation includes the subjectivity of the interpreter. Therefore, the knowledge incorporated in the interpretation process contains some limitations particularly in terms of object definition. There may be problems of confusion, misclassification, misaggregation and scale dependence as explained further below.

In the traditional procedure it is very likely that object class definition (*step 4*) may be confused with the definition of terrain objects (*step 1*). Take for example the application of aerial photo interpretation to soil investigation used for agricultural development. In the traditional procedure soil units or soil classes are classified on the basis of attribute values obtained during image interpretation and combined with several field observations (*step 1*). In this regard, the soil scientist starts to observe and interpret photomorphic characteristics which can be recognised on the image and which will be associated with the most probable soil classes (*step 2 and 3*). By analysing and grouping or classifying these observations a particular soil unit or soil class is established (*step 4*).

The decision rules used in grouping and classifying observations adopted by the traditional approach are very seldomly formulated explicitly. Often object classes are not defined *a priori* and the process of class definition is not properly explained either. Thus, the interpretion units produced by the traditional procedure do not represent individual and independent spatial objects (entities). As a consequence, the traditional procedure has a very poor repeatability or reproducibility. Considering this, the aerial photo interpretation which was completed in this study particularly in the identification of terrain object TMUs is a step beyond the geomorphology mapping method developed by Verstappen and Van Zuidam (1975).

The object based TMU is addressed to obtaining geometrical aspects, i.e., boundary of the terrain object. As a natural division of the terrain, the interpreted physical boundaries of TMUs are considered to be observable, fixed or repeatable boundaries and attribute values within this boundary are uniform. The interpretation of aerial photos and SPOT images for both monoscopic and stereo analysis was aimed mainly at the delineation of homogeneous relief units TMUs and to identify the spatial distribution of soil erosion in the study area. Accepting this concept, the terrain object can also be combined to capture other terrain object characteristics (Meijerink, 1988). Moreover, the interpretation of aerial photos for the identification of terrain mapping units is independent of the incorporated knowledge of the interpreter during the interpretation process. In other words, the final result of the interpretation or the object recognition (*step 4*) is dependent only on the defined object characteristics visible in the photos (*steps 2 and 3* in Figure 13 (a)), there is no subjective inference on the part of the interpreter.

6.2 TMU defined for data acquisition, data capture and observation procedure

6.2.1 Introduction

This section is a descriptive analysis of the object orientation concept applied to managing the complexity associated with the state and behaviour of TMUs. It includes a method by which TMUs can be defined and placed within class and aggregation hierarchies. It is intended to provide a solid foundation in establishing a hierarchical method for data acquisition and for producing the integrated information categories at targetted scales.

Three fundamental object-oriented concepts apply to the TMU. These concepts include object aggregation, class generalisation or specialisation and association (Molenaar, 1991; Webster, 1991; Smith and Smith, 1977; Braspenning, Uiterwijk, Bakker and Van Leeuwen, 1991). A detailed discussion on this issue is given in Chapter 7.

Object definition at the targetted scale is the most important aspect to be able to describe properly terrain objects. A clear object definition leads to a clear classification of terrain objects under discussion. Adopting a geo-information theoretical approach as discussed in Chapter 3, this section is intended to underline the process in defining TMU for data acquisition, data capture and producing soil erosion information categories at regional scales.

6.2.2 TMU defined

Van Zuidam (1985) has mentioned spatial entities or units of analysis at different hierarchical levels of a terrain system as shown in Table 6. See also Dayat (1993). This hierarchy was constructed on the basis of different aspects of geomorphology and the potential use of terrain classification and it represents different aggregation levels. The relations between these units can be modelled through the aggregation and classification structure of the terrain objects as discussed in Section 7.3.

A TMU can be described by static (attribute values) and dynamic properties (object-specific operations). The state of the TMU is given by its attribute values and its behaviour (specific operation) which are encapsulated within the description of the TMU. See Figure 17. Thus, TMU spatial objects are defined at different levels of complexity and have unique indentifiers.

Table 6: Units of analysis within a hierarchy of terrain systems (modified from Van Zuidam, 1985)

Unit of Observation	Systems	Unit of analysis / utilisation
Terrain provinces	Regional or watershed	Terrain systems / Master Plan
Terrain systems/TMU	Sub-regional or sub watershed	Homogeneous land form / feasibility of general land development, Long Term Plan
Homogeneous land form/slope level	Terrain unit or sub division or small catchment of the terrain	Farmer's field level / semi detail, Medium Term Plan
Farmer's field level	Terrain component or sub-sub division of the terrain	Detailed terrain characteristics / engineeering, Short Term Plan

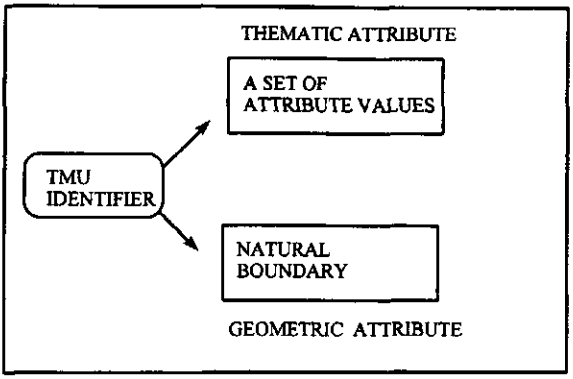
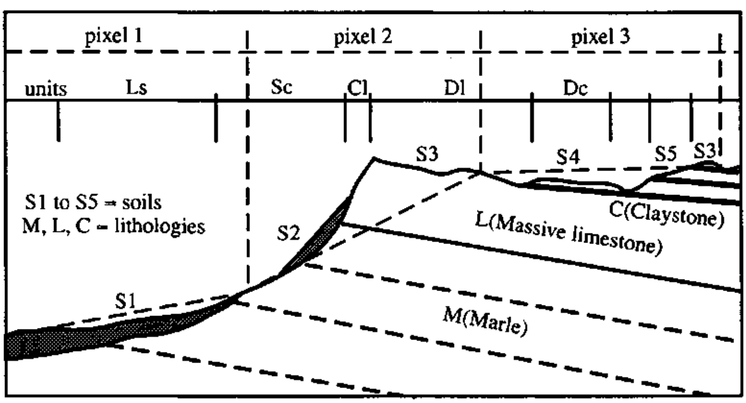


Figure 17: TMUs as terrain objects described by thematic and geometric attributes

The TMUs in this thesis are considered physiographically as homogeneous units, in which *soil erosion* may occur. Based on this understanding Meijerink (1988) suggests defining the TMU as a natural division of the terrain as illustrated in Figure 18. Each division of the terrain represents observed patterns of a characteristic distribution of relief elements with a certain size, inclination, orientation, roughness, granularity of upper layer or surface deposit and with a certain ground cover.

The natural division of the terrain depends therefore on the combination of terrain characteristics, say TMU influencing factors including *morphometry (slope steepness, degree of dissection), lithology and geology*. This reveals the complexity of thematic and geometric attributes associated with TMUs.

On the basis of attribute values such as size of the unit, the TMUs can also be considered as *nested terrain objects* which can be hierarchically ordered. An example expressing the hierarchical structure of TMUs is the situation where a TMU consists of a complex set of aspects which exhibit a characteristic spatial pattern e.g. the strongly dissected middle slope of a volcano. With increased scale of observation or mapping, this type of TMU can be specified in a hierarchical ordering system into independent sub units and sub-sub TMUs. Thus, the hierarchical ordering entails that each sub-entity, i.e., sub or sub-sub TMU is *a_part* or member of a particular main unit (TMU). This implies that TMUs allow the identification of terrain characteristics at different aggregation levels. A detailed discussion on this subject is given in Section 7.2.



Legend:
Dc = Dipslope on shale ; Dl = Dipslope on limestone; C = Cliff (outcrop); Sc = Scree; Ls = Lower front-slope; Pixel 1...3 = Pixel of raster network; S1...S5 = Soils; M = Marle; L= Massive limestone; C = Claystone.

Figure 18: Escarpment slope facets of a particular terrain system (adapted from Meijerink, 1988)

Figure 18 shows the cross section of the escarpment of a particular terrain system consisting of several units i.e. lower front slope (Ls), scree (Sc), cliff/out crop (Cl), dipslope on limestone (Dl), and dipslope on shale (Dc). This figure also shows the complexity of the terrain. For example the unit Ls is formed as an association of soil type S_1, geology M and slope (R_1). In other words, each of these units is composed of a particular soil, slope steepness and geology. This concept implies that each delineated unit has its own terrain characteristics and differs from other surrounding delineated units. This is the subject of Section 6.5. Above all, it can be used as a basic mapping unit for data acquisition, data capture and for producing composite information as illustrated in Figure 19 (a). For comparison data acquisition by digitising an existing map is also presented. See Figure 19 (b).

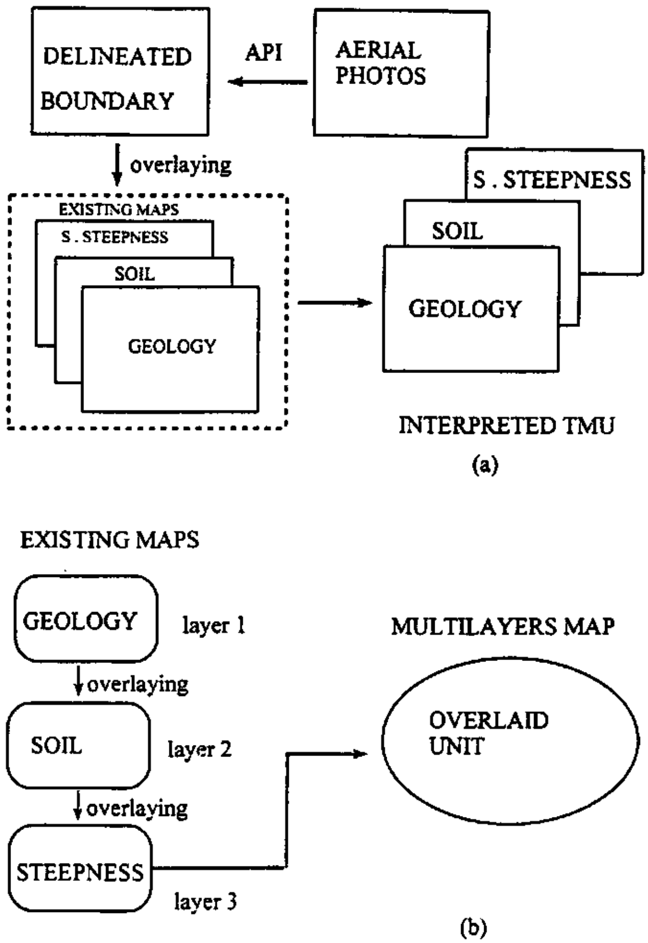


Figure 19: Data acquisition techniques:
(a) using the concept of interpreted TMU;
(b) digitising/overlaying of existing thematic maps

Though the inputs and outputs are in fact the reverse, both techniques in establishing the TMU are closely related to each other. Both may contain similar attributes. Figure 19 (a) shows that each interpreted unit of the TMU carries data on geology, soil and slope steepness that can be derived in the raster network of a particular GIS environment.

The procedure associated with the TMU produced by map overlay is slightly different to that of the interpreted TMU. See Figure 19 (b). In this case, single layer independent thematic maps are overlaid to create a particular mapping unit. As stated in Subsection 1.1.3, the integration of the independent thematic maps remains the problem because of different scales and different types of data.

As shown in Figure 19 (a), by applying simple data base operations the interpreted units can then be linked and used for data acquisition, data capture and in producing information categories. In addition to this, analytical aerial photo interpretation leads to the final differentiation as well as ensuring a proper structure and to have the benefit of information derived by a geomorphological analysis.

6.3 The identification of homogeneous relief TMUs and spatial distribution of accelerated erosion at sub-regional scale

6.3.1 Introduction

This part of the thesis discusses the application of object orientation in identifying a pattern associated with a TMU. This is a method prerequisite to the classification of terrain objects and is a step further than object definition as discussed in Section 6.2. and referred to as *steps 2 and 3* in Figure 16 (a). The identification of TMUs refers to a process or method by which TMUs can be recognised and detected on the images.

Under a stereo viewer the use of aerial photos or SPOT images facilitates identifying TMUs, e.g., by the presence of uniform relief unit. It is also used in the recognition and data extraction related to terrain object characteristics finally leading to the definition of the terrain object TMU.

The delineated homogeneous photomorphic units of TMUs in this study were interpreted using observable characteristics such as degree of dissection and slope steepness. In this regard, each delineated unit represents the combination of slope steepness, degree of dissection, lithology, geology and may include dominant vegetation cover. The interpretation of SPOT images and aerial photos at scale 1 : 50 000 and 1 : 15 000 allows the identification of the combination of these parameters forming a spatial pattern associated with each TMU. The combination of these parameters makes this unit identifiable and distinct in the image including the differences in shape, pattern, slope gradient and soil.

6.3.2 Spatial and specific patterns associated with TMUs

Photo interpretation key and pattern associated with TMU identification

With regard to the TMU a photo pattern element or photo interpretation keys for images are recognised as distinctive patterns representing the appearance of TMUs and their associated characteristics in the images. Considering this, criteria used in the identification of TMUs are *the homogeneous relief units of the terrain.* This is considered as identical to the term *face* as described by Huising (1993); Molenaar (1995 b).

For example a complex alluvial fan covered dominantly with rice fields is very distinctive compared to a hilly area covered with upland agricultural activity. Each pattern exhibits a particular spatial arrangement of the various features in a repeated sequence (order) within the interpreted area. Considering the photomorphic properties related to the obtaining of TMUs, panchromatic SPOT images (1987), aerial photos at scales 1 : 50 000 and 1:15 000 produced in the year 1982 were interpreted and the TMU map of the study area was drawn.

HR terrain mapping unit and erosion process

As stated previously the particular pattern associated with each TMU is identifiable on aerial photos at scale 1 : 50 000 and 1: 15 000. Through the stereo viewer we can evidently obtain knowledge on the spatial distribution and areal percentage of each TMU affected by accelerated erosion, which is functionally related to the arrangement of other terrain objects, e.g., ground cover and active processes.

The accelerated erosion processes which occur in a particular TMU are active processes and are the result of peculiar genesis and development compared to other geological environments. This allows such active processes to be identifiable in the Earth observation remote sensing images (Bergsma, 1982; Pickup and Chewing, 1986; Hill, 1993). The characteristics of radiation from a material are a function of material properties, observation of soil reflectance can provide data on the properties and the state of the topsoil (Epema, 1986). Both progressive and regressive pedogenesis cause alterations of the soil surface which, to a certain extent, are spectrally detectable (Hill, Smith and Alther 1993).

Erosion processes wash and transport the fertile layer including organic matter, clay minerals and soil particles. As an effect the infertile layer is brought nearer to the surface and the reflectance becomes higher (Bergsma, 1982). According to a study done by Seubert, Baumgardner and Weismiller (1979) in Northern Indiana, the class with the highest reflectance is correlated with severely eroded upland soil. On this basis and under a certain percentage of ground cover density, the soil erosion process is also recognisable on the aerial photos and SPOT images.

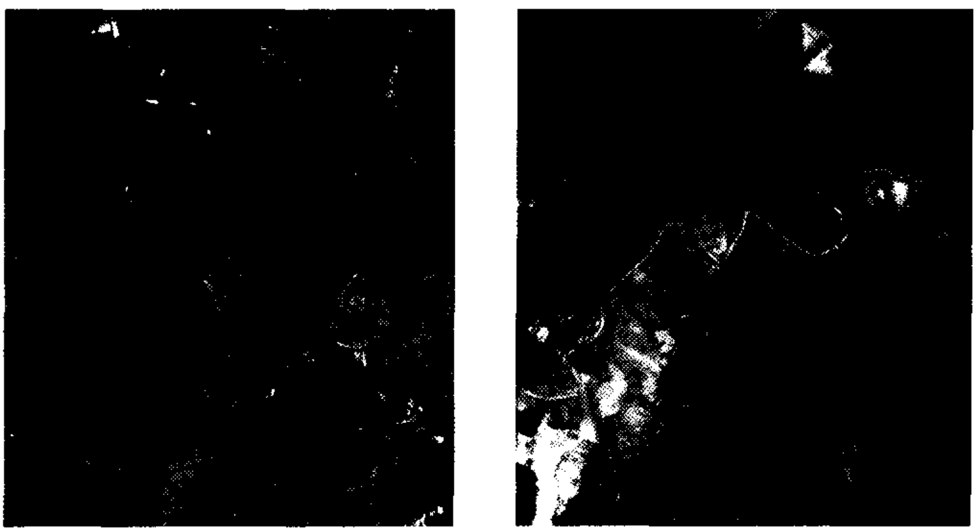
In this study, the interpretation of aerial photos included the identification and delineation of TMUs that are affected by sheet and rill erosion as well as the length and depth of gullies. See Appendix 7. The result of aerial photo interpretation shows that the most affected areas occur in the sloping or undulating volcanic areas associated with upland colluvium and in faulted and folded areas in volcanic mountainous terrain units. The erosion severity classes of the entire area range between very slighty (lightly affected) and very severely (heavily affected) eroded as described further below. The information on the spatial distribution obtained from this stage of the study was used as basis inputs to conduct erosion survey at FFL as discussed in Chapter 8.

Spatial patterns in the study area

A set of characteristic patterns for a specific area usually can be defined and used in the identification of Terrain Mapping Units (TMUs). As an illustration some spatial patterns in the study area are described below.

Homogeneous relief unit of area along rivers in alluvium systems

A unit of physiography typical of this area is found along both sides of the Ciseel Subwatershed. There are no extreme differences in terms of internal relief. It is a generally flat area. Dominant land use/cover is irrigated rice field, indicated on the aerial photos by a smooth light grey tone. See Figure 20 (b).This pattern also shows that in flat valley paddy lands and adjacent low hills that are protected with a good cover of permanent type vegetation little or no erosion is evident in the stereo photo coverage. Hillside areas of any slope that are protected by a dense native forest cover show no evident erosion on the stereo photo coverage.



(a) **(b)**

Figure 20: Examples of specific patterns:
(a) interhill valley pattern
(b) area along river

Homogeneous relief unit of interhill valley in upland hilly terrain systems

As we can see in Figure 20 (a) the physiographical unit of interhill valley is a narrow flat area between two hill complexes. Dominant land use/cover is irrigated rice field. This is indicated on the photos as light grey with slightly rough texture. Little erosion is evident in the stereo photo coverage.

Homogeneous relief unit of volcanic area associated with upland colluvium

Physiographically this area is found on the upper parts of fault (broken) areas with steep and very steep slopes. Because of high difference of internal relief this type of spatial pattern has an irregular grey tone and rough texture. Up to 25% of the area may be in poor permanent cover or lacking an adequate conservation system to control erosion. Most hillsides are under 40% slope. Moderate erosion is found in stereo photo coverages.

Homogeneous relief unit of fault area in volcanic mountainous terrain
The physiographical unit of this spatial pattern is found at the back of river terraces along the sides of the Ciseel Subwatershed. It is identified as having a very extreme internal relief and very steep slopes. The land cover consists of mixed garden, upland rice, and cassava cultivation. Up to 50% of the area may be in poor permanent cover or lacking an adequate conservation to control erosion. Many hill sides are over 40% slope. Severe erosion is evident in the stereo photo coverages. On aerial photos these areas show irregular tone and texture.

Homogeneous relief unit of folded area in volcanic mountainous terrain
This physiographical unit varies from steep to very steep or extremely steep. This type of spatial unit is found in almost the entire study area. Because of variation in slope steepness (internal relief) this spatial unit can be identified on the photo by extreme differences in tone (shadows) and rough texture. Over 50% of the area may be in poor permanent cover or lacking an adequate conservation system to control erosion. Many hillsides are over 40% slope. There is very severe erosion evident in the stereo photo coverage.

6.3.3 Conclusion
The (*complex*) variation of the units in morphometry, lithology, geology and dominant vegetation cover forms distinctive patterns representing the photomorphic units of TMUs and their associated characteristics in the images. In turn these are associated with particular erosion classes.

The most affected areas occur in the sloping/undulating to rolling units of volcanic areas associated with upland colluvium and in faulted and foulded areas associated with volcanic mountainous terrain units.

Based on image interpretation and field checking it was identified that slope hydrology has a very strong influence on the occurrence of a particular type of erosion, e.g., sheet and gully erosion. In volcanic units sub surficial flows are dominant as the erosion producing agent. Abandoned fields have a strong association with the initiation of gully erosion.

6.4 The differentiation of classes of TMUs at sub-regional level

6.4.1 Method used
As explained in Chapter 3, Subsection 3.3 classes of terrain objects should be mutually exclusive (distinct). The analysis of variance (ANOVA) (Kendall and Stuart, 1968; Snedecor and Cohran, 1989) which follows is intended to provide a better insight into a quantitative description of a terrain object in general and is used in the differentiation or classification of an independent object TMU resulting from aerial photo interpretation in particular.

As stated previously visual aerial photo interpretation allows one to identify and differentiate TMUs. From visual observation and stereo photo interpretation each type of TMU can be differentiated by relief attributes, as shown in Table 7.

Table 7: Relief attributes associated with TMUs

Grouped TMUs	Number of observations	Average slope length (m)	Average slope steepness (%)
I	13	694.45	2.15
II	8	491.18	5.25
III	5	294.45	11.40
IV	12	367.65	20.50
V	14	425.00	32.70
VI	5	676.45	59.40

Using a 1:50 000 topographical map with 25 m contour interval combined with aerial photos at scale 1:15 000, relief attributes including slope steepness as shown in Table 7 have been systematically measured for each representative TMU. Each figure shows the steepest part of a profile, i.e., a true down slope line to the base of slope. See further in Subsection 9.5.1.

In order to be able to perform the differentiation analysis, the individual measurement of the slope steepness associated with each TMU has to be scaled according to the slope steepness classes as shown in Table 8.

The differences between defined TMUs in the study area were statistically differentiated using one-way analysis of variance (ANOVA), Bartlett's (B) test and Scheffe's (S) test as discussed further below.

Table 8: Slope steepness classes on the basis of individual measurements

Slope steepness (%)	Description
0 - 3	Nearly level
3.1 - 8	Gently undulating
8.1 - 15	Undulating
15.1 - 25	Steeply rolling
25.1 - 45	Hilly
> 45	Steep

6.4.2 Results and discussion

Regardless of the genesis and lithology of a TMU the relief attributes, e.g. *the slope steepness of each identified TMU, was taken into account in the classification of each group of TMUs.* Slope steepness is considered as one of the most important factors affecting the utilisation of a piece of land for a particular use (FAO,1976 and see also Section 5.1). In a mature topographic development, slope steepness is considered as an independent morphological manifestation of a steady state of the terrain (Schumm and Mosley, 1973). By this is meant that it is not influenced by any other relief attributes. For example it is sometimes not related to the size of the area concerned. However, it can be an indicator of causative (active) forces such as soil erosion. The effect of slope steepness in the classification of TMUs is discussed further below. The result of the ANOVA apply to the classification of TMU is given in Table 9.

Table 9: One-way analysis of variance of slope steepness attributes

Source of variance	Degrees of freedom (df)	Sum of squares of slope steepness	Means of squares of slope steepness	F calculated	F table
Between groups	6-1= 5	16272.22	3254.44	107.58*)	3.17
Within the groups	57-6=51	1542.55	30.25	-	-
TOTAL	56	-	-	-	-

Note:

*) = yields significant different at $F_{(0.01, Df)}$

F calculated = $\dfrac{\text{mean squares among the groups}}{\text{mean squares within the group}}$

Means of squares = (sum of squares/Df)

From Table 9 it is clear that the calculated value of F is greater than the critical F value of 3.17 with *confidence interval (CI)* 95%. A slope steepness with calculated F of 107.58 yields a very significant difference. It means that classified terrain object TMUs are statistically distinct if they belong to different groups. In this regard, slope steepness associated with each identified TMU is considered as the best significant difference attribute in the differentiation of TMUs in the study area. Thus, on the basis of slope steepness the classification of TMUs can be performed.

Furthermore, Bartlett's (B) test was used in order to evaluate the homogeneity of the variance. The values associated with all attributes (B calculated = 69.30) are less than the critical value (B table = 73.29) for 56 degrees of freedom (df) of *Chi-square* (χ^2), 95% of CI. This shows that one condition of analysis of variance was completely fulfilled. See also Snedecor and Cohran (1989).

The ANOVA as discussed above leads us to the question which group of TMUs yields this significant difference. It means that further statistical testing is required. For this purpose, the Scheffe's (F) test or the Least Significant Difference (LSD) as described also by Kendall and Stuart (1968) was applied. The result of this test was aimed at facilitating the establishment of subsets (groups) of homogeneous TMUs.

The Scheffe's F test is expressed by an F value and it was used as *a measure of differentiation, range, dispersion or overlapping of two or more observation means.* This can be done by ranking the group of means in declining order. The test then can be executed and the difference (F value) is observed. In this case the higher the F-value the larger the overlap and *vice versa.* Figure 21 shows the situation graphically. The bars indicated by letters a, b, c and d show the significant mean of slope steepness differences between TMU groups at the 0.05 level. Only the mean of slope difference between group I and group II is not significant at this level. In other words, the Scheffe's F test shows the overlapping classes within the proposed TMU classes, i.e., between group I and II. Considering this, they were separated and rescaled into five different slope steepness classes as presented in Table 10.

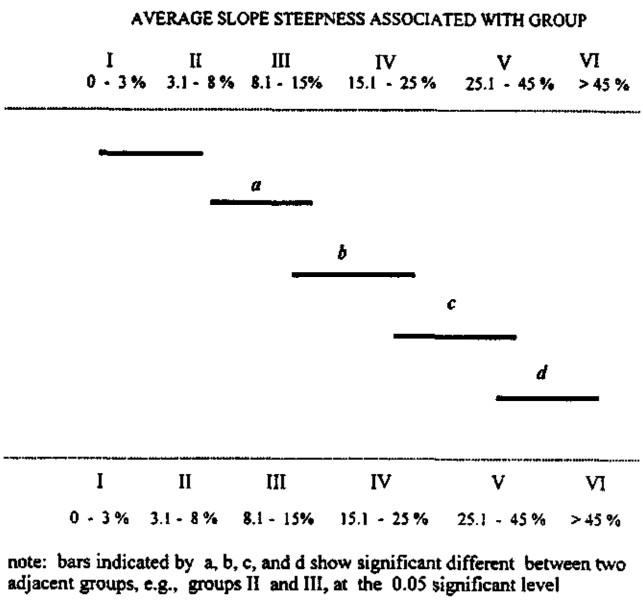


Figure 21: Diagram of separability of TMUs according to slope steepness (%)

Table 10: Regrouped and rescaled slope steepness classes

Regrouped TMU	Slope steepness (%)	Description
I	0 - 8	Flat to undulating
II	8 - 15	Undulating to gently rolling
III	15 - 25	Gently rolling to steeply rolling
IV	25 - 45	Hilly
V	> 45	Very steep

In this exercise this test was also applied to differentiate the mean values of slope steepness compared to the slope length. The result of this test is presented in a graph as shown in Figure 22.

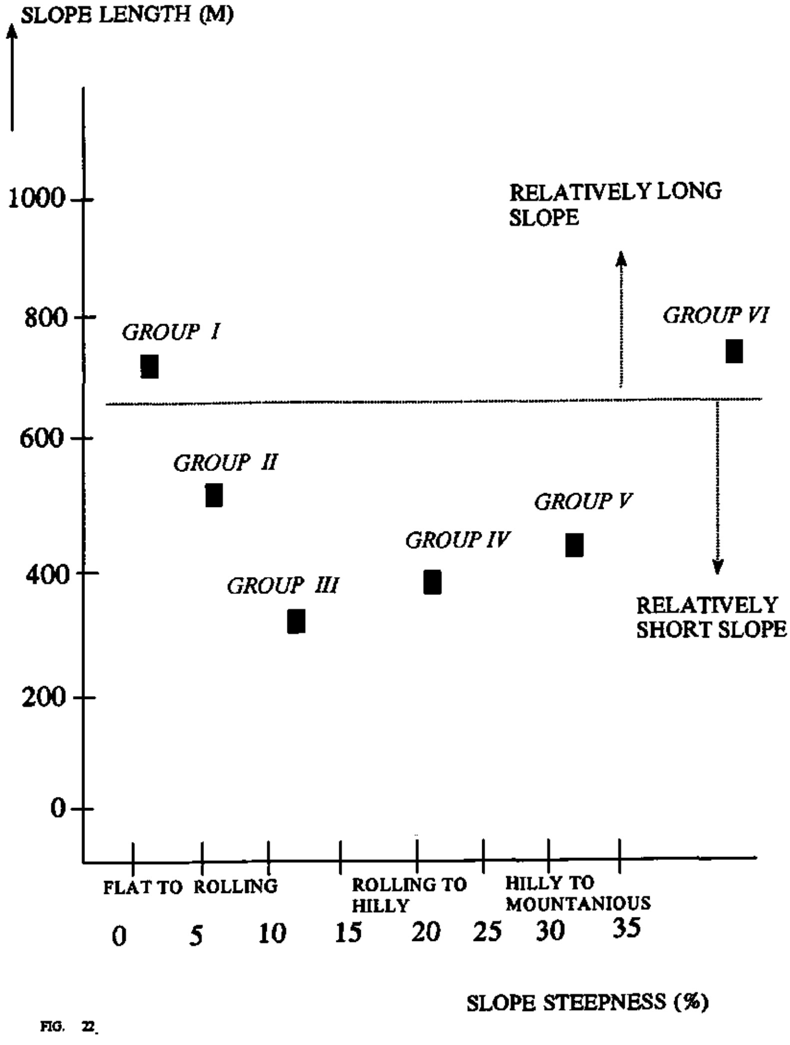


FIG. 22

Figure 22: Means of slope steepness (%) mapped into slope length (m) using Scheffes' F test

Figure 22 below shows that there is no correlation between slope steepness and slope length. This fact is basically explained by the long and steep slopes of volcanic origin and the relatively gentle and short denudational slopes on sedimentary rocks. The rest of the steep and gentle slopes consist of summit surfaces.

Figure 22 also shows that a differentiation between flat, undulating and undulating to rolling groups is no longer possible based on relief amplitude. The flat to rolling group plot in the area of low steepness of the graph and only can be differentiated by slope length, the boundary lying between 600 and 700 meters. In general volcanic flows plot in the portion of longer slope. Plains and most of the footslopes also plot together in the area of longer slopes. The rolling to hilly and hilly group plots in the steeper but shorter part of the graph.

The hilly group is formed by volcanic scarps, the denudational slopes on volcanoes and the relatively younger lava flows. It is clear that the younger the flow, the flatter its top and steeper the edges. The sedimentary group and the plain plot completely apart. The distinctive morphologic characteristics of both units are clearly shown. The structural sedimentary slopes and high relief amplitude for the former, flatness or nearly flatness and low relief amplitude for the latter.

The flat and undulating units have very gentle slopes and very low relief amplitude. The undulating to rolling group has slightly steeper slopes and larger relief amplitude. Both groups are well differentiated from the steeper rolling to hilly group, though the relief amplitude remains the same. The hilly group shows an increase in steepness and partly in amplitude except extreme cases (flat to undulating and hilly to mountainous groups). Therefore, there is a large group of TMUs that can be differentiated by slope steepness.

6.5 The repeatability of the boundary of TMUs

The importance of having a stable basic mapping unit used in a GIS environment has been discussed. In addition, both the advantages and disadvantages of techniques to obtain this unit were discussed (see again Subsection 1.1.3).

As often occurs during aerial photo or other image interpretation, the interpreted TMU is affected by the subjectivity of the interpreter in defining terrain objects particularly in delineating the natural boundary of the TMU. Without overemphasising the drawbacks this and this part of this thesis was written to provide better insight into the objectivity aspect of the interpreted TMU.

The objectivity of the interpreted objects is closely related to the stability of their boundaries and attribute values as mentioned by Middelkoop (1990) and Edwards (1994). In other words, the stability of terrain objects is related to the sensitivity of boundaries to changes of thematic and geometrical aspects of terrain objects. This part of the thesis discusses the stability of the boundary in terms of geometrical statement (Molenaar, 1995 a), i.e., boundary aspect including size, the actual shape of the object drawn by different interpreters and at different times.

The stability of the basic mapping unit in this study refers to "*the repeatability of the boundary of the units.*" In this regard, whenever interpreters are assigned to interpret and delineate the terrain, the predefined object, e.g., TMUs will be yielded firmly with almost similar ("*coincident*") boundaries. In other words, the difference in geometrical aspect particularly in the metric sense is negligible. The repeatability is required for evaluating and monitoring of changes in the static and dynamic aspects of terrain objects.

The repeatability of terrain object boundaries varies and depends on the type of terrain object. For example using satellite imagery Janssen (1993) researched terrain objects which correspond to the ownership boundary of agriculture fields (parcels). This type of boundary is a typical man made boundary and is not a natural boundary. Therefore it was repeatable due to easy changes in thematic and geometrical aspects of the field (parcel). It is consistent with the result of active process mapping done by Carara (Carara, 1992). He has studied units representing the repeatability of the interpreted landslide hazard assessment in La Honda basin, California, found by several surveyors.

On the basis of the dynamics of land use, Huising (1993) studied the repeatability of boundaries of land use zones in Costa Rica. Using maximum and minimum percentage of changes (changes in structure and thematic categories), he comes to the conclusion that the boundaries of agricultural land use zones (areas) are relatively unrepeatable compared to natural vegetation and forested areas.

Taking the above experiences into account, in this research the repeatability of the boundary of TMUs was evaluated. These typical boundaries usually are identifiable, repeatable and stationary. Stable TMU boundaries in this study was obtained using either geological, lithological or morphometric boundaries.

As stated in Chapter 3, Subsection 3.3 of this thesis the *attributes of terrain objects include the description of the size or area* of corresponding terrain objects. The bigger the differences of the area the less repeatable is the boundary of the object. Considering this, a convenient method of assessing the repeatability of TMU boundaries is to compare the results of two or more different, independent interpreters produced at different dates. A simple way to do this is, for a specific region, to measure the area of each particular category of TMU on each interpretation, e.g., TMU 23, 44 and 51, then to find what area differences are present. If these differences are small, then we can assume that *the method produces repeatable boundaries*, and that subjective influences are negligible. As an example two terrain classification maps produced by the University of Padjadjaran (UNPAD,1983) and RMI (1986) are superimposed and analysed. The superimposed polygon map is presented in Figure 23 and the area difference is presented in Table 11.

Table 11: An example of area difference between two corresponding TMU maps

TMU ref.nr.	Area by RMI (Ha)	Area by UNPAD (Ha)	Mean (Ha)	Difference (%)
51	9.90	9.81	9.86	0.47
23	26.80	27.51	27.16	1.30
44	4.19	4.35	4.27	3.70

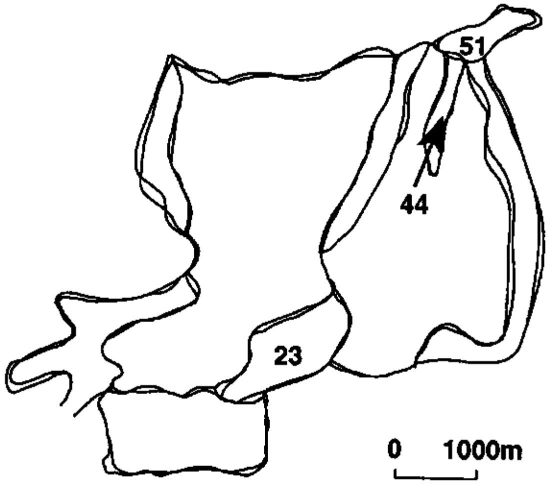


Figure 23: TMU boundaries produced by UNPAD and RMI.

Figure 23 shows that the interpreters identified and delineated the same targetted objects. The small area differences (see Table 11 as examples) were considered mainly due to errors in interpretation and delineating boundaries and were not caused by actual changes of thematic or geometrical structure of the TMUs. They can be considered as geometrical error associated with TMUs which is not treated in this thesis. See Offermans (1986).

Chapter 7

The hierarchical method for data acquisition and for representing erosion information at local and sub-regional level

7.1 General view

The selection of a proper level of detail for any problem to be studied is important to optimise functioning of the disciplinary expertise (O'Neil, 1988; Bergkamp, 1995; Bouma and Hoosbeek, 1996). This chapter should in this respect provide a better insight into especially the hierarchical aspects associated with inductive erosion modelling based on TMU.

An IEM in a GIS environment needs to be supported by the availability of a stable spatial observational unit and unit of analysis which can be hierarchically ordered and can also be used in the identification of the relationship between soil erosion influencing factors and a set of erosion severity classes (in Section 2.3). The nature of TMUs does provide this.

The hierarchical level of observation units adopted by an IEM represents the interlinkages between point observations at farmer's field level, homogeneous land forms and larger parts of the terrain (Suryana and Molenaar, 1995). In the light of a geo-information theoretical approach (Molenaar, 1995b), these units are related to the hierarchical level of terrain objects in which erosion severity classes are also considered as active attributes. As an implication, the abstraction of soil erosion severity class information follows the abstraction of terrain mapping unit information.

The above statement implies that the implementation of an IEM based on TMU for modelling erosion severity class at sub-regional level depends on the following knowledge.

(i) Spatial distribution and possible degree of existing soil erosion which may occur on a particular TMU. This was obtained using the interpretation of remote sensing images including aerial photos at scales 1: 50 000 and 1: 15 000 (in Chapter 6).

(ii) Spatial relationship between soil erosion influencing factors and particular erosion severity class. Knowledge of this particular aspect of an IEM can be used as a basis in constructing the inference rules. This can be obtained only after conducting more detailed study, i.e., erosion survey at FFL. This is the subject of Chapter 8.

(iii) Linking between (i) and (ii). This is the subject of Sections 7.2, 7.3 and 7.4. Subsection 7.2 discusses the observational role of the TMU in the identification of soil erosion influencing factors and soil erosion severity class. Subsection 7.3 discusses the aggregation and class hierarchy of TMUs. Subsection 7.4 provides a better insight into the relationship between TMU and an IEM.

7.2 The observational role of the TMU in identification of soil erosion influencing factors

7.2.1 Introduction

This section is aimed at providing better insight into the dynamic and static properties of TMUs, leading to better understanding of the classification and aggregation hierarchy of TMUs as discussed conceptually in Subsection 3.5. The class hierarchy of TMUs is used further as a basis for constructing the relationship between TMU and the proposed IEM in predicting soil erosion at different hierarchical levels.

TMUs carry implicitly information on the static and dynamic soil erosion influencing factors. The relatively static erosion influencing factors include, e.g., soil type and slope steepness. The ground cover, soil structure and soil erodibility are the dynamic erosion influencing factors. This information can be structured at several hierarchical levels. Observational processes can be structured according to these levels.

The dynamic aspects of TMUs can also be expressed through the thematic attributes pertaining to the active processes. The information about these processes completely depends on the mapping and observation scale. Increasing or decreasing the scale of mapping and observation affects the information content attached to a particular class. The most dynamic aspects may be found at the lowest level of the hierarchy, i.e., subsub TMUs, which carry detailed attributes of the terrain. This is the subject of Section 7.3.

In addition to the observational procedures adopted in this study, the visual and digital interpretation of remote sensing images including SPOT and aerial photos at scales 1: 50 000 and 1 : 15 000 which was cross checked by observation conducted at farmers' field level also provide information on active processes, e.g., erosion processes, landslide and erosion features that may occur within the units. Adopting TMUs as observation units allows the establishment of the relationship between soil erosion influencing factors and erosion severity class at local and sub-regional level. This is the subject of Section 7.4.

7.2.2 The observational characteristics as the dynamic and static properties of TMU

An example of a thematic description of a TMU is given in Table 12. This table shows that TMUs carry data about terrain characteristics which are relevant inputs to the proposed IEM.

Each TMU has its own attribute values which belong only to a particular TMU. As discussed in Section 6.4 each TMU can be differentiated from surrounding TMUs on the basis of these thematic properties. The measurement of these properties can be obtained using the method developed by Meijerink (1988).

Table 12: Observational characteristics of terrain mapping units

Name of TMU	
Code of TMU	
Nr. of TMU	
Nr. of Sub-sub TMU	
I. ORIGIN	
II. MORPHOMETRY	
1. Maximum altitude (m)	
2. Minimum altitude (m)	
3. Internal relief (m)	
4. Slope steepness (%)	
5. Slope length (m)	
6. Steepness/relief class	
III. MORPHOGRAPHY	
1. Slope position	
2. Slope form	
IV. MORPHODHYDROLOGY	
1. Drainage density (m/km^2)	
2. Drainage pattern	
V. ROCK TYPE	
1. Substratum/age	
2. Surficial deposit	
3. Duricrust	
4. Weathering depth	
5. Fracturing	
VI. MORPHODYNAMIC	
1. Erosion type/degree	
2. Mass wasting (type)	
VII. SOIL CHARACTERISTICS	
1. Dominant USDA/FAO class	
2. Depth	
3. Texture	
VIII. DOMINANT VEGETATION COVER	

Section 6.4 presents the determination of TMUs in the study area. Six main TMU classes were interpreted from aerial photos and their differentiation was tested. This analysis demonstrates that the thematic properties, e.g., slope steepness attached to TMU can be used as TMU class determinants which can begrouped into the primary, secondary and tertiary determinants (see Table 13). Each group of determinants consists of a set of interrelated terrain characteristics.

Table 13: TMU determinant order

Order or hierarchy of determinant	Description
A. Superclass TMU or Level +1 or primary determinants*)	Represents the origin of landforms or complexes. It implies the endogenous and exogenous processes responsible for their establishment including lithology, morphometry, valley density. It may also include vegetation cover.
B. Class TMUs or Level 0 or secondary determinants*)	Represents the specific origin, lithology, soil type, slope class, slope range, type of active processes. It may include vegetation compositions.
C. Elementary object TMUs or Level -1 or tertiary determinants*)	Represents the slope steepness, slope length,soil pedon, active proceses and vegetation types.

Remark:
*) the relationship between class and aggregation hierarchies is implicit

Each group of these determinants from Table 13 represents clearly the degree (level) of complexity and the stability of their attribute values. The attribute values associated with the primary determinants are considered more stable compared to the secondary and the tertiary determinants. The tertiary determinants contain more transient (unstable) attribute values. This is related to the different degree or scale of observation.

From Table 13 and the result of TMU differentiation analysis described in Section 6.4, thematic attributes associated with the primary determinants, i.e., morphometry, origin of landforms, lithology, and vegetation covers, are considered as the most appropriate determinants in the disaggregation of main unit TMU into sub units. However, the sub unit and catena elements which are determined by the secondary and tertiary determinants, e.g., slope steepness, vegetation types may not be representable. The implication of this statement is that the information on these sub units and catena could not be visualised on the map but will be maintained in the data base.

7.3 Hierarchical structure for TMUs

7.3.1 Class and aggregation hierarchies of TMUs

The structuring and representing of TMUs for large areas with the objectives of this study in mind, is an interactive endeavour. In a new area with different physical characteristics, the pre-conceived classification scheme may have to be adjusted because one encounters new combinations of landforms, lithologies and erosional processes.

Consideration of time and scale, or the spatial resolution of the geometrical part of the GIS imposes restrictions. For example, actively eroding, deep valleys are important terrain features; however, they may be too small to be mapped individually, and so they have to be aggregated as *part_of* larger units in an aggregation hierarchy.

This section is intended to discuss an approach to TMU modelling which can be implemented into a computerized information system. It sets up an efficient TMU classification hierarchy without reducing the completeness of the information. For this purpose, an object oriented approach including generalization, specialization, aggregation and association is adopted. Thus, reducing the complexity associated with TMUs as mentioned above does not always result in losing too much information. There are techniques for managing the complexity and multiscaled dependency of hierarchically structured attributes (Ward, 1963; Coad and Yourdon, 1991; Bergkamp, 1995; Bouma and Hoosbeek, 1996). These methods involve the establishment of a data base management structure including how TMUs can be stored as well as how to provide correct and timely relevant information. This can be augmented through a proper data modelling including selection of the most appropriate criteria (attributes) for classification, scale, hierarchical structure and functional units in describing a particular system of interest. The first two have been discussed in Chapter 6 and the rest are given in this chapter.

TMUs with similar characteristics can be grouped into a class with a specific attribute structure. That means that TMUs within such a class have the same description structure, their attribute values may very though (see Figure 24 (a)). To be able to assign TMUs to a class, we need determinants, these should be based on the value of the same attributes. This implies a class hierarchy where at the highest level all TMUs have the same attributes. The determinants at this level assign TMUs to subclasses. Each sub class will have its own attributes and per subclass determinants can be defined to assign TMUs to sub-subclasses etc. (see Figure 24 (b)). The assignment of TMUs to sub classes implies here a disaggregation of TMUs into sub TMUs and into sub-sub TMUs (see again Subsection 3.5).

The above statement is consistent with the hierarchy theory related to global change developed by O'Neill (1988). He has mentioned that the system of interest such as terrain mapping unit can be divided into three different hierarchy levels namely level +1, level 0 and level -1. Here the dynamic of the upper level usually appears as constant in the lower and the lowest level. In this way, TMUs (at level +1 or superclass level) can be classified into sub units (at level 0 or class level) and sub-sub units (at level -1) as elementary object level or instances. See again Figure 24 (b) below.

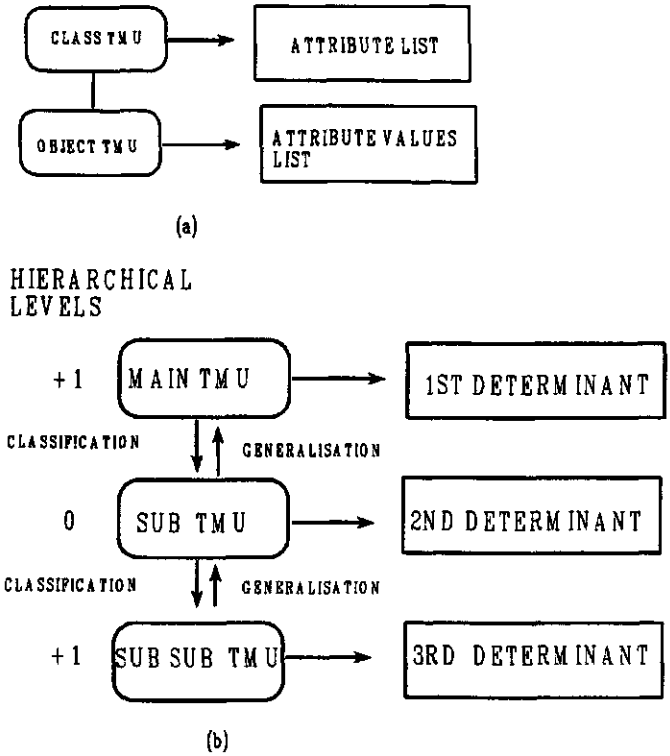


Figure 24: (a) Attributes defined at class level and evaluated at object level;
(b) Determinant order related to different aggregation levels

As discussed in Section 6.5, the repeatability of the boundary TMUs and Figure 24 above show that the terrain object TMU at the highest level (+1) is considered as the best functional unit in describing and predicting soil erosion at sub-regional level. In this regard, elementary objects at levels (-1) and (0) were used as basic units in the implementation of the process of aggregation and class generalisation at superclass level (see again Section 3.5). Figure 24 also shows that at the top of the hierarchy, general erosion influencing factors associated with prime determinants are used in conjunction with the highest level TMUs, referred to as level +1. Lower down the hierarchy, using sub (level 0) and sub-sub TMUs (level -1), the erosion influencing factors must be specified in more detail.

7.3.2 Structuring of the units

The classification hierarchy of TMUs as stated earlier was done on the basis of three different determinants i.e. primary, secondary and tertiary determinants (in Table 13). These determinants as stated in Section 7.2.2 are composed of general attribute structures, e.g., morphometry, and are subsequently related to each level of classification hierarchy of the terrain.

TMU superclass and main unit TMU
Superclass TMUs and main unit TMUs are located at the highest or superclass (+1) level in classification and aggregation hierarchy. They are established according to their partly common attributes (prime determinant). From Table 12 in Section 7.2 the main geomorphological units are differentiated according to their origin and used further as main functional units. Thus, the main units are mappable units and composed by objects at a lower level of the aggregation hierarchy. Further differentiation into geomorphologcal units was done on the basis of specific landforms, which are indicated by means of a symbol, e.g., W for mass wasting, J for steep, erosive valley, or by means of a description and coding of the subunits and/or catena. Each geomorphological or main unit is accompanied by a physiographic description to satisfy general needs and has a specified morphometric characteristic.

TMU subclass and sub TMU
Many main units will consist of an aggregation of different subunits. These subunits, with complete attributes structure (secondary determinants), are located at class (zero) level. Because of their small size these units are not mappable units on small scale maps but using a specific TMU identifier they are inventorised, included and linked into main units in the data base. The soil and land use information will be tied up with description of the small units.

The forms of the catena elements are coded, to save lengthy descriptions. Crests can be sharp, narrow, broad convex, flat etc. Valleys may be incised, terraced flat, irregular, etc. Slope types can be straight, convex, concave, irregular stepped.

Subsub TMU or catena elements
These subsub units are located at the lowest level at the classification and aggregation hierarchy. These units are not mappable units at small scale. Through their identifier they can be linked to larger units through *part-of relation.*

Table 14: The interpreted TMUs in the study area

Terrain systems	Main units	Symbols	Subunits	Dominant slope (%)	Dominant cover
I. Alluvium	1. Recent	**A1**	a. River alluvium	0 - 2	Rice field
			b. Flood plain	0 - 2	Rice field
	2. Subrecent	**A2**	River alluvium	0 - 2	Rice field
II. Alluvio-colluvium	Recent/sub-recent alluvio-colluvium deposit	**A4**	Slope wash and alluvium	0 - 8	Mixed garden/dryland/ ricefield
III. Upland colluvium sedimentary	1. Upland hilly terrain system	**SC1**	a. Plateaus and gently sloping hill	0 - 4	Mixed garden/ bush
			b. Sloping and dissected hill	4 - 25	Mixed garden/ bush
	2. Mountainous terrain system	**SC2**	a. Mountainous strongly sloping and dissected	15 - 45	Dryland farming with terrace
			b. Mountainous terrain very steep	45 - 85	Mixed garden
IV. Upland colluvium volcanic	1. Volcanic plain planeze or lahar	**UCV1**	a. Flat residual deposit	0 - 2	Mixed garden
			b. Uneven colluvial plain	2 - 8	Mixed garden
	2. Volcanic hilly terrain lava flow	**UCV2**	Flat to undulating volcanic flow	4 - 35	Mixed garden
V. Upland colluvium	Volcanic mountainous terrain	**VC1**	a. Strongly sloping or/and dissected slope	25 - 85	Mixed garden
			b. Dissected slope	> 85	Mixed garden

7.3.3 The interpreted TMUs in the study area

On the basis of interpretation parameter keys associated with major divisions and origin of the terrain as presented in Appendix 2, the TMUs in the study area were identified, delineated and verified in the field. The structure approach of Subsection 7.3.2 has been applied. The determinants to classify the main units and the determinants for disaggregation these into sub units are given in Table 14. These subunits have been assigned to land use classes, which have been given in Appendix 2. Table 15 gives the total areas per main TMU class.

Considering the hierarchical ordering of terrain determinants associated with a particular TMU (in Subsection 7.3.1) the main unit of the terrain still can be desaggregated into two or more subunits. In this regard, the thematic and geometric description associated with a subunit is necessarily a refinement of the description of the major division. It can be used also to describe simultaneously other terrain characteristics, e.g., eroded severity class at different scales.

Table 15: Cover percentages of the main unit types in the study area

Main terrain division	Area (Ha)	Area (%)
Alluvium	30,531	31.63
Recent and subrecent alluvio-colluvium deposit	9,768	10.12
Upland colluvium sedimentary	19,690	20.40
Upland colluvium volcanic	34,254	35.49
Upland colluvium	2,257	2.36
Total	96,500	100.00

Table 15 also shows that 58% of the study area is composed of areas very prone to erosion. This is consistent with geological material composition which are suceptible to erosion as presented in Tables 2 and 4 (in Chapter 5).

7.4 The relationship between TMU and an IEM in erosion modelling at local and sub-regional levels

The entire IEM modelling procedure depends on a close relationship between TMUs and soil erosion influencing factors at all levels of the hierarchy. The aerial photo interpretation processes and field checking for delineation of TMUs allow the identification of soil erosion that may take place within the TMUs (see Section 6.4). This was used further as a basis for constructing relationships between soil erosion severity classes and their associated erosion influencing factors and in inferring soil erosion at different aggregation levels. Based on this, a strong relationship is expected to exist between soil erosion severity class on the one hand and a set of erosion influencing factors associated with each TMU on the other. In this regard, each TMU carries information on influencing factors including geology, lithology, morphometry and soil which are used partly as inputs to the proposed IEM. Overlaid with existing rainfall and vegetation maps, TMUs were used as the unit of analysis in inferring soil erosion severity class (see Figure 25).

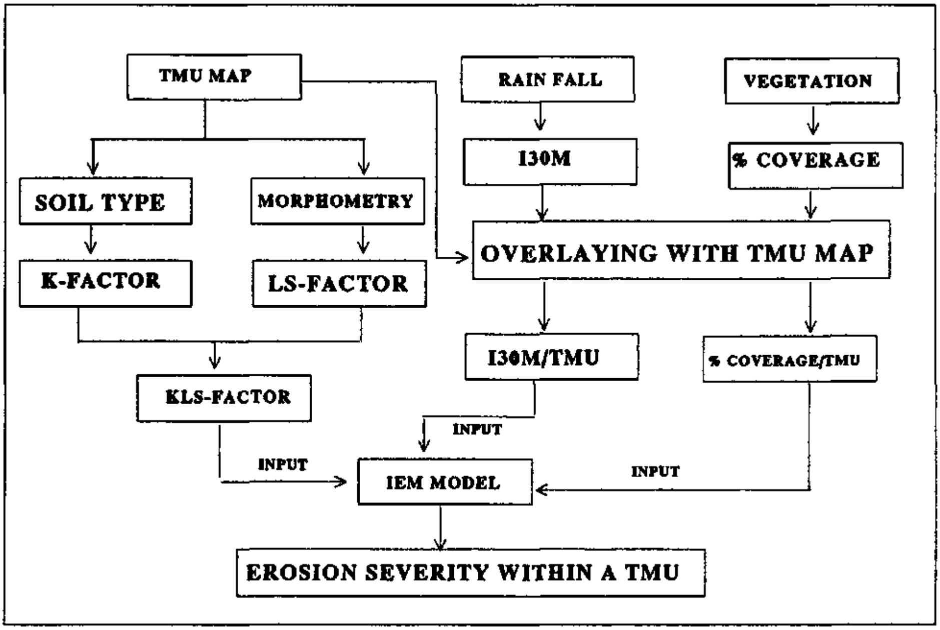


Figure 25: The general relationship between the TMU and the proposed IEM

The proposed IEM will be formulated as rules to infer erosion class from TMU properties. See also Suryana and Molenaar, (1995). This will allow GIS users to infer particular erosion classes on the basis of given influencing factors at different aggregation levels of TMUs. In this regard, either presence or absence of soil erosion depends on the composition of these environmental characteristics which are associated with each TMU. The more erodible environmental characteristics which are observed the more severe the expected soil erosion class in a particular TMU, and vice versa. Figure 26 shows how the three aggregation levels are related and the content of statements about erosion for these levels is different. They follow accordingly the generalisastion and classification hierarchy line of TMU. As consequence, the higher the hierarchy or scale of the prediction, the more generalised the content of soil erosion information and vice versa.

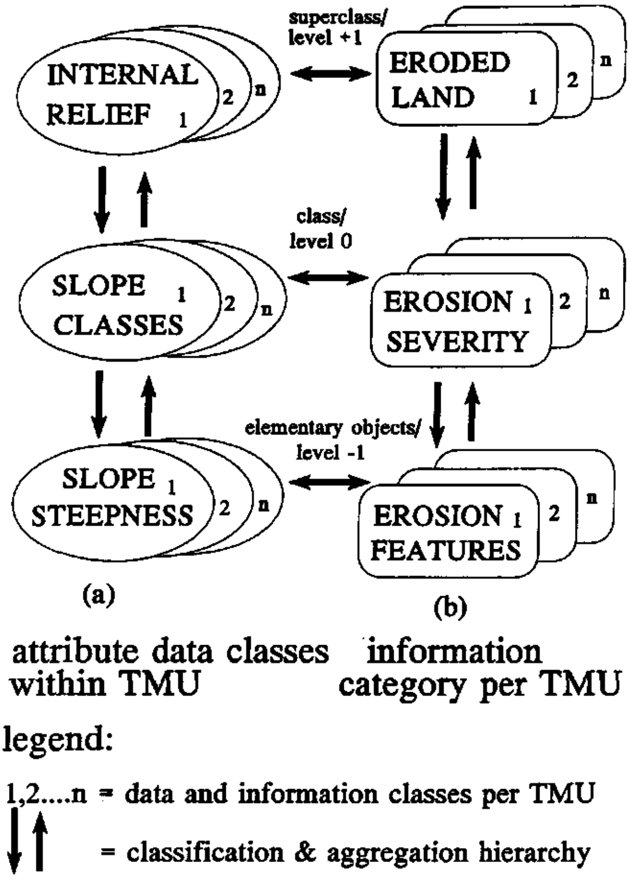


Figure 26: Stepwise induction steps at different aggregation levels

Although the thematic content of the data at the three levels is different according to Figure 26, we see in Figure 27 that the syntaxtic at the three descriptive levels is similar. This observation is important for the definition of data base operations transfering data between these levels.

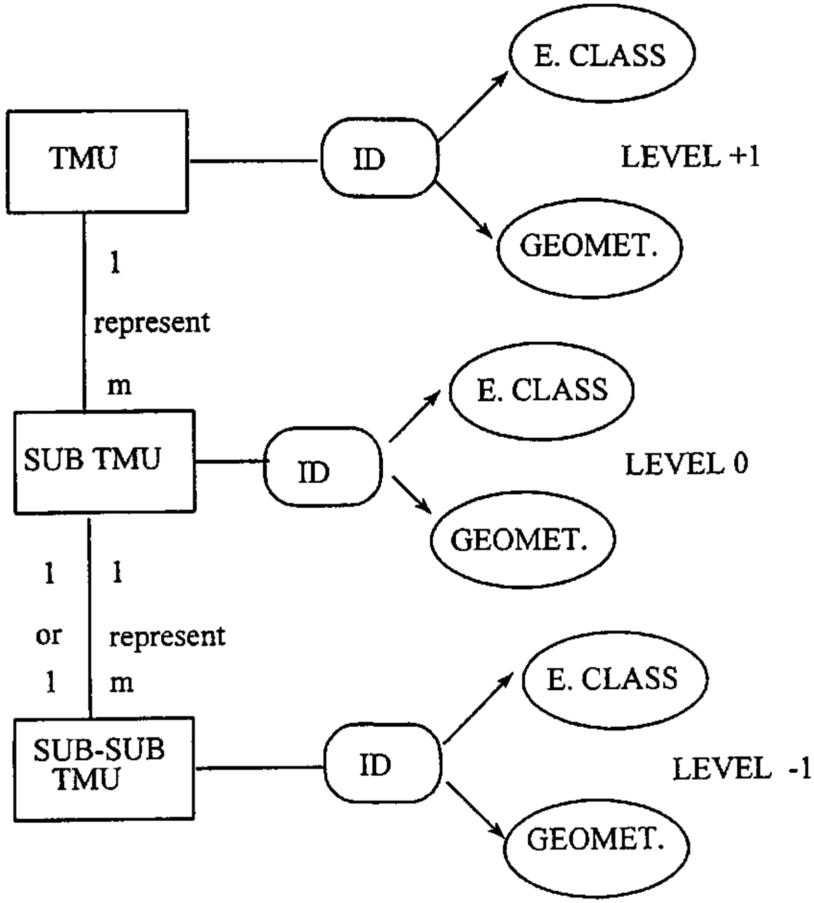


Figure 27: **The syntax of data models for aggregation TMUs at different aggregation levels**

Chapter 8

Data requirements, methods and techniques used in data acquisition and in building an Inductive Erosion Model

8.1 Introduction

The strongest characteristic of an Inductive Erosion Model (IEM) arises from the way it is built on the heuristic dialogue between expert, model and the underlying process obtained from a set of field observations at FFL (as training areas) which at the end yields a set of decision rules for inferring a specific erosion severity class on a specific terrain mapping unit (Suryana, 1992; Suryana and Molenaar, 1995).

With regard to building the proposed IEM, this part of the research explains the role of observational procedure and data acquisition at FFL referred to as the lowest level of the aggregation and classification hierarchy of TMUs. The main input to this part of the study is the spatial distribution of soil erosion as discussed and obtained from aerial photos and SPOT image interpretation as discussed in Chapters 6 and 7.

The discussion of this part includes:
(i) the identification of an IEM; (ii) the definition and synthesing of an IEM; (iii) the identification of an IEM requirements; (iv) sample area selection; (v) erosion surveys at FFL; (vi) developing an IEM as discussed further in Sections 8.2 to 8.5. The general procedure of the inductive modelling approach is presented in Figure 28.

8.2 The identification and the definition of an Inductive Erosion Model

8.2.1 Assumption and approach

Considering the complexity of the terrain and to be able to build the proposed IEM, this part of the study adopts the following assumption and approach.

(1) The occurrence of a known soil erosion severity class on a particular TMU is always closely related to the associated attributes, e.g., attributes of each soil erosion influencing factor. See again Chapters 6 and 7.

(2) The specific relationship between erosion severity classes and their influencing factors at specific location can be considered as an observed pattern. This can be obtained from observations made in particular sample areas. This subject is treated in Sections 8.3 and 8.4.

(3) Based on observed patterns a set of decision rules representing the relationship as mentioned in (2) can be derived, induced and tested. This is the subject of Sections 8.5 and 8.6.

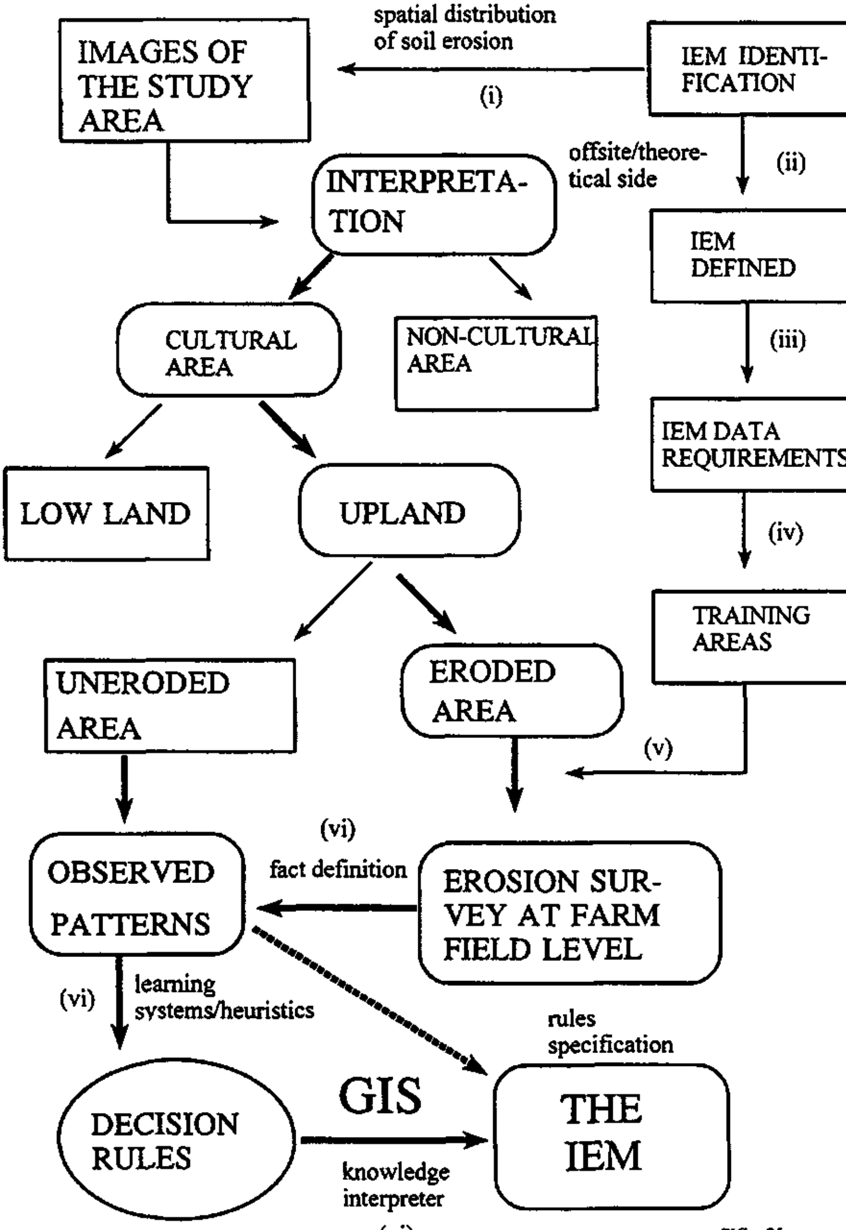


FIG. 25

Figure 28: **An inductive erosion modelling procedure**

8.2.2 The identification of an IEM

The concept of an IEM, the bottom-up approach erosion model, arose out of dissatisfaction with the USLE top-down model that cannot handle local variability of erosion influencing factors (in Subsection 1.1.1).

On this basis an inductive approach which involves an Inductive Erosion Model (IEM) is proposed as an alternative strategy for solving this particular problem (Subsection 1.1.2 and Section 2.3 discusses the definition of the concept associated with an IEM). It may not be exact but it should be usable, in an acceptable period and operational conditions, by those who are concerned with erosion problems. The intention of an IEM *is not to describe the processes or effects of erosion, but to predict where erosion is taking place and its degree, together with an estimate of the uncertainty of the prediction.*

With regard to the problems of the variability of typical upper watersheds, as well as problems associated with obtaining good quality sample datasets *"before and after"* the modelling process, the implementation of the proposed IEM is facilitated using terrain object TMUs as discussed in Chapters 6 and 7. The differentiation of TMUs at sub regional scale was tested (see Section 6.4). The result of the analysis shows that the classified TMUs at sub-regional scale are distinct if they belong to different groups. In this study therefore TMU are used to describe terrain object characteristics at this scale.

The delimitation of the upper watershed into TMUs allows the establishment of different classification and aggregation hierarchies. It implies the ability to attach to different aggregation levels a particular label or attribute value. Once the aggregation hierarchy level of a TMU is established, it facilitates the application of hydrological knowledge from sites at which records have been collected to other areas where data are required but are unavailable or incomplete (Hendriks, 1990; Mosley, 1981). Considering the hierarchical approach as discussed by (O'Neil, 1988; Bergkamp, 1995; Bouma and Hoosbeek, 1996), an IEM involves interlinkages among three different aggregation levels, namely point observations at FFL, slope levels and larger parts of the terrain (Suryana and Molenaar, 1995). These are referred to as sub-sub division, sub division and main division of the terrain systems (see Table 6 in Chapter 6).

Experts' knowledge of spatial erosion severity classes and erosion influencing factors and their interrelationships, obtained through *"onsite (practical side)* and *offsite (theoretical side)"* studies of erosion processes is very important for building an IEM model. It helps the heuristic dialogue between the expert, the model and the underlying soil erosion processes (Bonnet, 1985; Mortimer, 1985; Keller, 1987). This allows an IEM to be easily modified and contributes greatly to the flexibility of the approach.

8.2.3 Defining and synthesising an IEM

As discussed in Section 2.1, soil erosion is a set of processes involving detachment, transportation and deposition of soil particles which are determined by locational erosion influencing factors (Morgan, 1974). Erosion severity classes are affected generally by type and percentage of vegetation cover and slope steepness. As might be expected, the higher the percentage of vegetation cover and the gentler the slope the smaller the risk of erosion and vice versa. The relative importance of these two factors may be deduced from the fact that many units with a high percentage of vegetation cover and steep slopes have a low risk of erosion, showing the great importance of the vegetation (Wiersum and Ambar, 1981; Wiersum, 1990; Bergsma and Kwaad, 1992; Stroosnijder and Eppink, 1993).

The building of the proposed IEM in this study was formulated from the perspective of erosion influencing factors. By this is meant that the process of soil erosion on a particular terrain mapping unit is determined by various influencing factors. The relationship between the influencing factors and hypothetical erosion severity class as established by experts is a nondeterministic and often uncertain causality. Often there is insufficient testing or validation by experts. As suggested by Stocking, Chakela and Elwel (1988), therefore soil erosion classes predicted by the proposed IEM are measured and expressed as abstract indication of erosion severity, e.g., severely eroded, rather than as quantified estimate of soil loss in ton/ha/year. Considering this and using notation as discussed in Subsection 1.1.2 and Section 2.3, the proposed IEM infers the effect of each erosion influencing factor and the causal relationship of these factors on particular soil erosion severity classes as formulated as follows:

$$E_j = f(D_i) \qquad \text{....eq.(18)}$$

where:

E_j	=	erosion severity class
D_i	=	erosion influencing factors
D_i	=	$d_1, d_2, d_3, d_4, ..d_n$

The erosion severity class (E_j) is an information category which is determined by d_1 d_2....d_n, which are attribute values associated with erosion influencing factors or data classes (D_i). See Section 2.3. In other words, each combination of a set of attribute values relates to one value of E_j. Established erosion influencing factors in this study include rainfall erosivity (R-factor = d_1), soil types (ST-factor = d_2), slope steepness (SS-Factor = d_3), slope length (SL-factor = d_4), percentage ground cover (V-factor = d_5).

From the above equation and erosion processes, obviously an IEM adopts *commutativity*. This implies that (i) an IEM model prediction should not depend on the order in which soil erosion influencing factors are inserted in the model; (ii) the more influencing factors inserted into an IEM model, the better the confidence in the model prediction. With regard to (i), different orders of insertion of the above established erosion influencing factors do not affect the final result of an IEM's prediction. In this case, an IEM always infers the same erosion severity class. With regard to (ii), the occurrence of a particular erosion severity class (E_j) is determined by various factors. This impplies that there are other factors than the above mentioned erosion influencing factors. Therefore, the more influencing factors are inserted, the better the confidence in an IEM's prediction. In addition, the proposed IEM may also include surface roughness (SR-factor = d_6).

The incorporation of an IEM into a GIS environment involved two main procedures. Firstly, it involved observing, selecting and taking a *set of erosion influencing factors for a set of known erosion severity classes on selected observation points made at FFL*. As supported also by Evans (1990); Evans and Boardman (1994); Bergsma and Kwaad (1992), at this level of observation the effect of each locationally variable erosion influencing factor was observed, analysed, classified and controlled. This becomes the strength of an IEM modelling procedure (Suryana, 1992). Once this has been done, the builder of an inductive strategy has to establish *a set of decision or inference rules* explaining the observed pattern, i.e., the relationship between the erosion severity class and erosion influencing factors. The established rules were contructed constantly in the form of *"IF conditions THEN conclusion "* structure (see also FAO, 1976; Negoita, 1985; Drummond, 1991; Rodrigue, 1995). The application of these rules can be regarded as the process of inferring general relationships after a set of decision rules derived from a set of observation on observed patterns is established (see again Section 2.3). It becomes apparent that the relationships between the premises and the conclusion which allow the criteria of correctness of inference rules to be established, depend on the number of the observed cases.

Secondly, the process involved looking for a suitable, systematic method of predicting to which erosion class any particular object belongs. This procedure can be regarded as a learning or inducing process in which decision rules are developed according to the analysis of the relationship between a set of objects and their attributes as explained below.

Referring to the second procedure, equation (18) and Subsection 1.1.2 and Section 2.3, the definition of an IEM in GIS can be restated mathematically as function of *inference model* R that relates a set of erosion severity classes E_j with the degree of uncertainty of U_{ij} to a set of terrain object descriptions D_i and with the degree of uncertainty S_i.

The mathematical formulation above conludes that *the known* erosion influencing factors of an *unknown* erosion severity class associated with a terrain object TMU can be found to fit a *known observed pattern*, belonging to a set of erosion severity classes. By inference, the terrain object itself is then classified as belonging to the erosion severity class with that known observed pattern of erosion influencing factors. A particular strength of this approach is that classifications of objects can often be made even when attribute data are not complete. Of course, as more data are collected, the classification becomes more reliable. The degree of reliability can be quantified.

8.2.4 Dataset requirements for an IEM

The datasets consist of the erosion influencing factors and erosion indicators which are obtained from each sample area. The required datasets to build and operate an IEM have been grouped into (i) percentage of ground cover; (ii) erosion class at FFL; (iii) land degradation; (iv) terrain and climate data. The most recent (1982) aerial photographs at scales 1: 50 000 and at 1: 15 000 were used for the pre-selection of relevant sites for collecting data on landcover and landuse, slope length, slope steepness, slope form, and slope exposure.

The preliminary landcover and landuse types classification, i.e., dryland agriculture, bare land, plantations, bush and mixed garden were based on remote sensing image characteristics including stereoscopically derived height measurement, pattern, tone, texture, shape and planting direction. The datasets on slope length and slope steepness, surface roughness and ground cover types were obtained directly from each observation point.

Landcover and landuse data

To get better insight into the effect of vegetation (V-factor) including the aspect of cover types on the observed erosion, representative landcover type data collected by the FEDP team (1986) were studied. For comparison, the estimation of ground cover aspects, at sample points, includes the following: (i) basal cover is the percentage cover of live vegetation at the soil surface, usually consisting of stems, leaves, twigs of creeping herbs; (ii) litter cover which is the percentage of dead plants or organic material on the soil surface; (iii) stoniness which is the percentage cover of rocks and surface gravel. The estimating of canopy cover within each stratum includes (i) the height of each stratum of vegetation within each site; (ii) the percentage canopy cover within each stratum. From the aerial photos the landcover aspects were observed according to the ground cover and aerial cover (canopy). Ground cover percentage was estimated using the FAO guidelines for estimating proportion of mottles and fragments.In this study landuse data represents the most dominant crop, and dominant cropping systems practised for a sufficiently long time (>10 years).

Erosion data

Erosion identification at FFL was focussed on observing *field erosion indicators* which include: (i) topsoil condition; (ii) splash erosion features; (iii) sheet erosion; (iv) overland flow or runoff indicators; (v) rill erosion features (Clark, 1980). For this purpose, the method and guidelines from Ditkontan (1985), Morgan (1988), Ambar and Syarifudin (1977) were adopted.

(i) Topsoil condition

Topsoil condition was used as an indicator of the presence of erosion processes taking place in a particular farmer's field within a particular terrain mapping unit (see Figure 29). Topsoil condition was observed through: (a) presence of horizon B, (b) topsoil structure, (c) biological activities.

The presence of a thick layer of bare soil aggregates, the ease of breakage of a soil clod by shaking usually leads to the ranking of the topsoil as crumb, especially when low biological activities may not reflect the original top soil conditions. Topsoil categorized as hard and difficult to break is ranked as compact, which is an indication of an advanced stage of erosion. Topsoil in between crumb and compact is classified as intermediate.

The presence of subsoil material (horizon B) is an indicator of disappearance of topsoil by erosion. It was assessed by comparing depth of soil horizon with similar profile characteristics of known undisturbed profiles, particularly variation in surface soil color relative to color differences known to occur through the profile. Biological activities were assessed by the presence of wormcasts, ants, and other microfauna.

Figure 29: Eroded and compacted soil surface

Figure 30: Pedestal

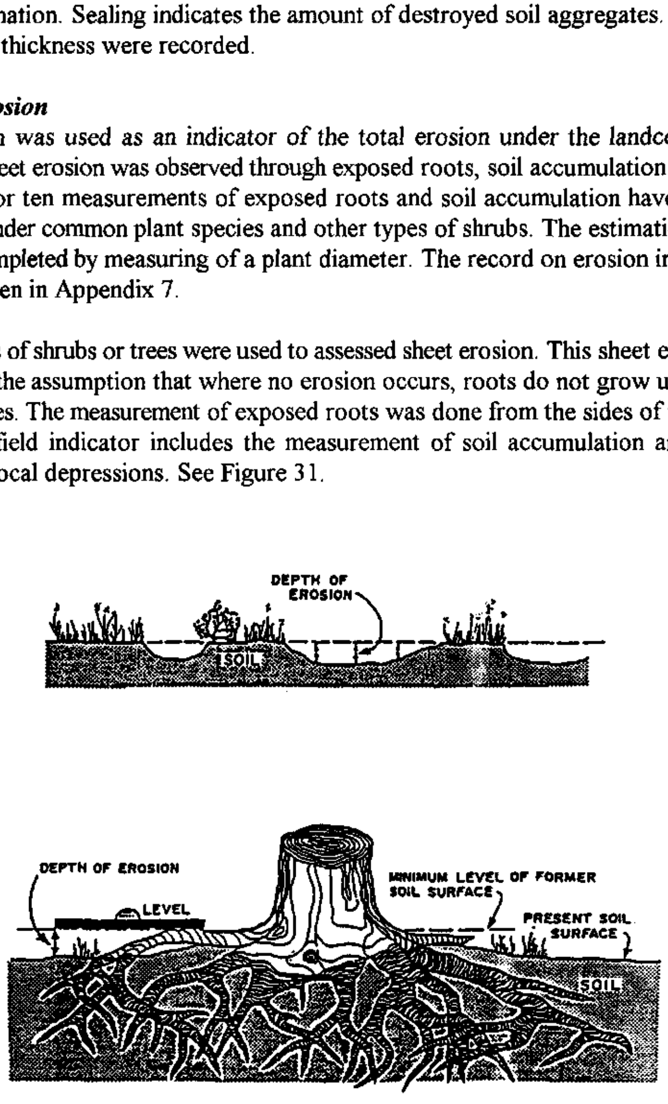
(ii) Splash erosion features

Splash erosion features were assessed and estimated through: pedestal, aggregate and sealing. These three indicators reflect the effect of rainsplash on topsoil. Pedestals are earth columns protected by stones or gravel from the effect of rainsplash (Figure 30).Their height and percentage area were recorded. The size and percentage of cover of aggregates on bare soil were assessed by visual estimation. Sealing indicates the amount of destroyed soil aggregates. The percentage coverage and thickness were recorded.

(iii) Sheet erosion

Sheet erosion was used as an indicator of the total erosion under the landcover types. The presence of sheet erosion was observed through exposed roots, soil accumulation and deposition. About eight or ten measurements of exposed roots and soil accumulation have been done per sample area under common plant species and other types of shrubs. The estimation of the age of plants was completed by measuring of a plant diameter. The record on erosion indicators (i), (ii) and (iii) is given in Appendix 7.

Exposed roots of shrubs or trees were used to assessed sheet erosion. This sheet erosion indicator was based on the assumption that where no erosion occurs, roots do not grow usually on top of the soil surfaces. The measurement of exposed roots was done from the sides of the stems of the plants. This field indicator includes the measurement of soil accumulation and soil material deposited in local depressions. See Figure 31.



(a)

(b)

Figure 31: Field (sheet) erosion indicators
(a) Sheet erosion between vegetative remnants
(b) Exposed roots
(adapted from Dunne, 1977)

(iv) Overlandflow indicator

It was observed by the distribution of litter on the soil surface in depressions, behind stems, and those litter that are directed in the downslope direction.

Rill erosion data

Rill erosion was observed under particular cover types such as fallow. The degree of rill erosion varies depending on the depth of incision and the distribution or density of rills per unit area under investigation. For the purpose of conversion to soil loss per hectare, the observed parameters included in this study are depth and width of rills and the number occurring in a 10 m square. They were measured in the middle section of the slope.

Soil data

The result of a previous study on soil classification in the study area conducted by RMI (1986) was used to get a better impression of the soil. The imformation on soil erodibility was obtained by applying the quick test of of Bergsma (Bergsma, 1990). This observation includes rating structural stability, infiltration and sealing, sensitivity of soil surface and simulation of a small test plot.

Terrain data

Terrain in the study areas was identified. This observation includes lithology, slope percentage, slope length, slope form, and slope exposure to assess their effect on erosion that may occur under different landcover/landuse types.This information was gathered using existing terrain data.

8.2.5 Selection of sample areas

Sample areas are required for building, validating and verifying the proposed IEM. Sample areas used for model building are not used further in validating and verifying the model. The sample areas were selected according to the main geological formations including their variation in soil type and lithology (some samples areas are presented in Appendix 4).

As stated earlier priority of selection was given to areas with a long agricultural landuse history and with variation in cover. Multitemporal remote sensing images, i.e., TM 1987 and 1991 and black and white aerial photographs at scales 1: 50 000 and 1: 15 000 were used, with special reference to change of land use.

Using inputs from Chapters 6 and 7, the selected training areas for building and testing of an IEM model were sloping agriculture lands under long agricultural use with special reference to the most erodible slopes, i.e. slope steepness greater than 10%. Considering an IEM datasets requirements as mentioned in Subsection 8.2.4 and accessibility led to preliminary selection of about 210 training areas, 90 for model building and 120 for model testing.

8.3 Erosion survey at the FFL and observed patterns

8.3.1 Introduction

The previous erosion studies show that farmers and their farming practices are the most active agents in the erosion process. The erosion survey was therefore devoted to agricultural land, especially upland agriculture areas, where erosion is likely to be more severe than in lowland areas.

To study the relationship of locational attributes to soil erosion, the erosion survey was conducted at FFL. The FFL was considered to be the most suitable basic functional unit to describe erosion severity class at local and regional level (in Chapter 7). The study was an observational approach aimed mainly at identifying and quantifying the observed pattern per observation point within a particular terrain system. At FFL, a set of *"learning samples or observed patterns"* was determined. That is the set of objects of *"known erosion severity class"* and their *"locational erosion influencing factors"* as described further in Section 8.4.

In order to obtain *observed patterns* as mentioned above, field work was carried out in two stages. The first stage was intended to select sample areas and to build an IEM. To obtain learning sample data sets or the observed patterns at FFL, the observation of locational erosion severity classes and locational key erosion influencing factors including crop type, percentage of ground cover, slope length, slope steepness, soil texture and soil erodibility was undertaken. The observation of soil erosion severity class was completed using erosion feature identification, e.g., surface erosion, sheet erosion, soil surface condition, rill erosion per 10 metre square, and by measuring the effect of each significant formation factor. The presence of erosion indicators, e.g., pedestal, sealing, root exposure, soil accumulation were collected. The classification of soil erosion was completed using combination of qualitative and quantitative measurements. These were followed by the preliminary selection of the most significant or influental combination of specific site erosion influencing factors.

During the field work period, the above bio-physical data were collected, processed and analysed using methods and techniques as developed by Ambar and Syafrudin (1979); Dissmeyer and Foster (1981); Stocking (1981); Ditkontan (1985); Meijerink (1988); Evan (1990) and Bergsma (1990). The second second stage was concentrated on developing and verifying the proposed IEM. The verification of an IEM as well as data quality assessment were carried out using external testing as discussed by Greendland, Socher and Thomson (1985); Chrisman (1984); Hoord and Brooner (1976) and Drummond (1977).

8.3.2 Method and techniques used

The role of vegetation in surface soil erosion is determined by the effect of canopy and ground cover. According to studies conducted by Thornes (1995), canopy cover reduces erosion through the effect on the amount and kinetic energy of rainfall that reaches the soil surface. Ground cover is more effective in reducing erosion than canopy cover, because (i) no remaining fall height so the direct impact of raindrops on the soil surface is completely eliminated; (ii) it slows runoff, increases infiltration, and reduces the transporting capacity of runoff; (iii) aspects of groundcover (litter, tree trunks, residues etc.) can create small reservoirs of ponded runoff in which detached particles are trapped; (iv) by protection against rainsplash, it prevents soil surface pores from being clogged , and thus sealing is prevented.

Different vegetation types have different effects on erosion class. Forests, well-established plantations and mixed gardens have been identified as most protective against erosion. Burning increases erosion, the amount of which depends on the intensity of burning. Intense burning removes most of the litter and soil erodibility characteristics are influenced. Pore volume, infiltration and the percentage of water stable aggregates will strongly decrease after intense burning (Imeson, 1995). When canopy and ground cover are removed the exposed soil surfaces are subject to the effect of rainsplash and erosion. As found in most sample areas, open areas generate the highest erosion.

The effects of agricultural landuse on erosion depend on the type of crop grown and land management. The protective value of the different annual and perennial cropping systems depends on the length of the period needed to provide good cover (Suryana, 1980). It also depends on the total biomass production and height of the crop, which determine rainfall interception rates, dripfall height, and protection against rainsplash (de Jong and Riezebos, 1994). Mixed cropping or mixed garden reduces erosion more than single cropping (Sub Balai RLKT, 1986 and Kwaad, 1994). This is because of the increased rate of canopy development and the availability of residue as ground cover produced by the harvested crops. The effect of perennial crops on erosion depends on the type of crop and management. Frequent clean weeding, e.g., in rubber and teak plantations, generally increases erosion.

Considering this, the identification (observational) procedure to obtain a learning sample set (observed patterns) was completed using the following steps (see again Figure 28):

(1) Interpretation was carried out of panchromatic black and white aerial photographs at scales 1 : 50 000 and 1: 15 000 (1982) and of satellite images including Landsat TM of June (1988) and SPOT standard film 1987 for both monoscopic and stereo analysis interpretation. Other relevant information used included topographic maps and geology and landuse maps at scale 1:50 000.

(2) Six sampling areas of 100 x 100 pixels were made with consideration to the coverage of all reflectance variations of the image. An unsupervised (clustering) algorithm was applied to all bands in each window and each class was defined per sampling area. The result was visually inspected and the best classes (best coverage, good differentiation of cover) were selected.

(3) The result of this interpretation allows the study area to be classified into less variable units with two major land use or land cover types, i.e., agricultural land and non-agricultural land.

(4) The area classified as agricultural land was grouped according to the type of agricultural practices, either as upland or lowland agricultural practices. Related to accelerated erosion process much attention was devoted to upland agriculture.

(5) The identification, discrimination and classification of upland agriculture either into eroded or uneroded upland agriculture were done by using SPOT image classification registered to the image of a Landsat TM.

(6) The classification of eroded and non-eroded upland agriculture for the entire area was accomplished using a maximum likelihood algorithm with all selected spectral classes. Through this entire process the study area was stereoscopically interpreted and the least variable unit map was drawn.

(7) To obtain learning sample data sets at farmers' field level the determination was undertaken of learning samples and observation parameters of locational key attributes including crop type, percentage of ground cover and other relevant soil erosion influencing factors. This was followed by the preliminary selection of the most significant or influental combination of specific site erosion influencing factors. It was undertaken according to erosion feature identification, e.g., surface erosion, sheet erosion, soil surface condition, rill erosion per 10 metre square, and by measuring the effect of each significant influencing

factor. The classification of soil erosion was measured as a combination of qualitative and quantitative measurements.

(8) At FFL level, the particular influencing factors considered were: slope steepness; slope length; percentage ground cover; top soil condition; soil erodibility (called soil reconsolidation effect); residual binding effects of fine roots in the top 30 - 50mm of soil; organic matter content. Information on how these data are processed and analysed are given in Appendices 6 and 7.

8.3.3 Data analysis

Field data collected at the FFL have been analysed to obtain information on erosion classification, landcover classification, the estimation of V-factor, soil erodibility, slope length and slope steepness at each sample area. Values for these data (factors) are supported by the existing values which are derived from tables and published information. A detailed description of the analysis of these data analysed is given in Appendix 6.

8.4 Results and observed patterns in the study area

8.4.1 Erosion classification under different cover types

As was mentioned in Section 8.2.2, the erosion classification was observed only within the sites under (semi-) natural and nonnatural vegetation cover types. The classification of erosion from field observation was performed and the occurrence of erosion classes (very slightly eroded, slightly eroded, moderately eroded and severely eroded) were identified within each of the cover types in the training areas as presented in Table 16.

Table 16: The occurrence of erosion classes under different cover types

Erosion severity class*)	Number of cases of erosion classes/cover type					
	Dryland agric.	Mixed garden	Rubber plant.	Bush	Sec. forest	Bare land
Very slightly eroded	0	5	2	1	3	0
Slightly eroded	5	3	1	2	3	5
Moderately eroded	9	0	1	0	0	13
Severely eroded	14	0	0	0	0	23
Total observations	28	8	4	3	6	41

Remark: *) As a result of erosion assessment and erosion classification at FFL as explained in Section 8.3. Each erosion severity class as mentioned above was defined as a function of the class of splash, sheet and rill erosion per each observation point. See Annex 7.

From Table 16 it can be observed that higher erosion classes, e.g., severely and moderately eroded, occur within the fallow vegetation associated with dryland agriculture and bare land cover types. In the case of dryland agriculture accelerated erosion may be caused by the reduction of ground cover by tillage, weeding and burning of the fallow vegetation (mainly Imperata Cylindrica and Pennisetum), by farmers in order to (i) control the density of weed regrowth and for easy ploughing, e.g., by tractor when the land is being prepared for cultivation; (ii) provide young palatable grass for rainy season grazing when the valleys are being use for rice cultivation.

The burning of fallow vegetation associated with dryland agriculture is seasonal, coinciding with the normal burning operations, which take place at the beginning of the rainy season. The burning is usually very intense leaving the fallow fields completely bare except for an insignificant percentage of basal cover, which provides no protection against erosion. Between February and April, the cover development is very low due to the low rainfall and high temperatures. In most of the sample sites in April, the aerial cover was about 10% and the groundcover was about 2-5%. Because of this low cover the fallow fields are exposed to the eroding impact of the usually heavy sporadic rains in April-May and erosion can be severe.

Of the other vegetation types, the secondary forests and bush cover types fall within the very slightly to slightly eroded class. This could be due to the generally dense aerial cover and the high groundcover percentages within the bush and secondary forest cover types. The mixed gardens and rubber plantations fall within the slightly to moderately eroded classes. This could occur because of the low shrub cover, ground cover and canopy cover in both cover types. From field observations it was found that the erosion classes under different plantations vary. The different methods of management in terms of the type of weeding. The rubber plantations measured are generally old, mostly esablished during the Dutch colonial period. Those that fall within the very slightly to slightly eroded classes are those in which slash weeding is conducted without any soil disturbance, and which have a high percentage of ground cover after weeding. The plantations within slightly to moderately eroded classes are those in which hoe weeding is done between the rows of trees in April-May.

8.4.2 Erosion features under the cover types

The measured indicators for sheet and rill erosion were discussed in Subsection 8.2.4. The measurements of exposed roots, soil accumulation and rills are very useful indicators of the occurrence of soil erosion. Table 17 shows that the total soil losses from field observation are higher than the soil losses predicted the USLE. It is caused by the rain erosivity values used (see further Subsection 8.4.4).

When the total root exposure of and soil accumulation behind plants of different ages are compared, it is observed that the values are higher for older plants. This can be explained by the fact that sheet erosion is higher on young fallow land than older fallow on which a complete ground cover has developed. Therefore, we can conclude that the older a succession of ground cover types the lower the annual soil loss by sheet erosion. Within the other vegetation types the highest sheet erosion indicator by young plants occurs in the grassy shrublands. This could be due to the effect of generally low cover.

Rill erosion occurs mostly in the cultivated farmers' field (not further quantified) associated with dryland agriculture and badland materials or bare land cover types (see also Imeson and Verstraten, 1988). This could also be related to the carryover effect of tillage from cropping. Rills are generally obliterated during weeding. However there might remain some areas with relatively loose soil in the plough path which can act as flow channels for water during the next rainy season. Rill erosion, like soil accumulation and root exposure, is more active during the first year than other years of different cultvations. This is reflected in the annual soil loss by rill erosion, being more in the younger dryland agricuture than in the older ones. This could be due to the increase in the ground cover on older sites. Rill erosion in the other vegetation types is insignificant. It could probably occur because of no tillage, or the increase in cover especially litter.

Table 17: Erosion features under different cover types

Erosion features/total observed erosion/predicted and erosion class	Observed and predicted erosion under different cover types (ton/ha/year)*)					
	Dryland agric.	**Mixed garden**	**Rubber plant.**	**Bush**	**Sec. forest**	**Bare land**
Observed erosion: A. Root exposure	57	34	35	28	34	76
B. Rill erosion	22	2	4.7	2.5	0	18
Total observed erosion (A+B)	79	36	39.7	30.5	34	94
predicted erosion**)	13	5	6	3	4	17
Erosion class*) observed / predicted by the USLE**	IV/III	II/I	II/I	II/I	II/I	IV/III

Remarks

*) converted from erosion class at each point observation

**) predicted by the USLE at the same location, estimated by FEDP Team;

***) I = very slightly eroded; II = slightly eroded; III = moderately eroded; IV = severely eroded.

8.4.3 The use of indicators for erosion assessment

Indicators (Subsection 8.2.4) are very useful for assessing the presence and absence of sheet and rill erosion, which take place on training areas. These features were observed in most of the sites in the training areas, but in some instances exposed roots and rills were significantly absent even though soil accumulation occured. Soil accumulation per se can only provide a general impression of the amount of the sheet erosion process. Rills and exposed roots may be used to derive average quantitative soil loss values (see again Table 17). This however has to be confirmed by other studies. The absence of rills and exposed roots simplifies the accomplishment of a rapid erosion assessment. This implies that the absence of these indicators shows no erosion. Where soil accumulation occurs, but no exposed roots or rills are observed, it is difficult to make quick assessments of erosion rates.

From sample areas, we found that root exposure is closely related to certain plant species measured. In some sample areas, no exposed roots were observed on some sites where *Imperata* *cylindrica* is growing, even though soil accumulation occurred. This was caused by: (i) the possible coverage of originally exposed roots by accumulating soil; (ii) the growth characteristic of plants like *Imperata*, which grow from root stalks or old plant bases, well below the soil surface. This therefore could lead to an underestimation of the amount of sheet erosion, based on the method used, on sites where imperata was measured.

The other observed plant species measured (*Eupatorium, Pennisetum, and Bamboo*), mostly showed exposed roots. This can lead to the conclusion that some plants are more suitable than others for the measurement of root exposure and this should be considered when applying this methodology for similiar studies elsewhere.

8.4.4 Soil loss estimations

The estimated values of the average soil losses (from field observation), and erosion classes under the different landcover and landuse types were mentioned earlier in Subsections 8.4.1 and 8.4.2.

The estimated soil loss from field observations was based on the assumption that the height of root exposure may correspond with soil loss. The values for rills could be assumed to be more-representative.

Moreover, it was found that the highest observed soil erosion occurs within the fallow fields associated with dryland agriculture, although the values vary according to the age of vegetation (see Table 17). From field observation it is clear that due to the past effect of tillage, erosion diminishes as the vegetation gets older. Here the older the vegetation types the greater the percentage of ground covers. In comparison soil losses under the other cover types are lower. This could be related to the generally higher ground cover and absence of tillage.

Generally there is a strong correlation between the predicted and the observed erosion classes (Figure 32). The highest classes fall within the dryland agriculture category, and lower classes in the other vegetation types. Under other cover types, the predicted soil loss by the USLE made by the FEDP Team is correlated with the highest predicted soil loss under dryland agriculture.

However, the observed soil loss is generally higher than that predicted. This is caused by (i) the rainfall erosivity value (2 225 mm) used in the predicted soil loss may be too low which could have accounted for the lower predicted soil loss; (ii) the height of root exposure associated with plant indicators may correspond with soil loss. One of the plant indicators, e.g., *Imperata cylindrica*, used for assessing erosion by field observations could have overestimated the total classification. As stated earlier, the growth characteristics of *Imperata sp.*, some measurement of root exposure could have been over estimation, because young plants grow from old root stalks.

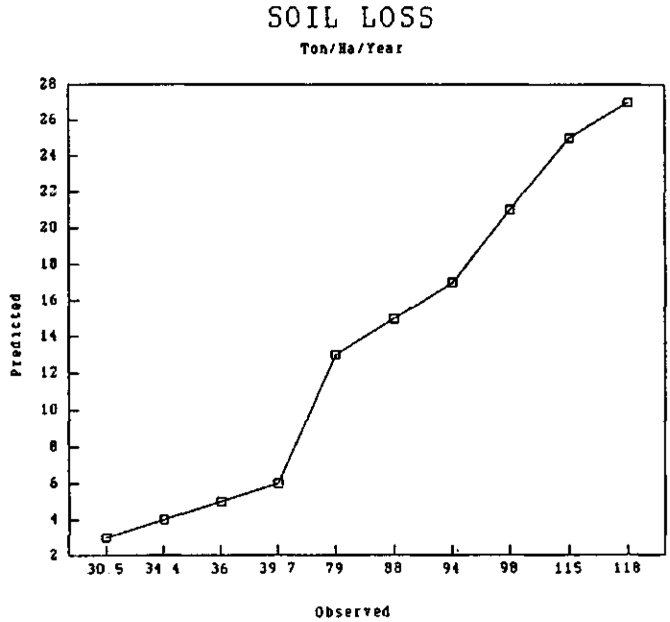


Figure 32: The relationship between the estimated and observed erosion class

As has been stated *Imperata sp.* is a poor indicator, and this could have caused the stronger variation. In other training areas, Pennisetum was the most common plant used for the measurement of root exposure in the fallow. This implies that Pennisetum is a better indicator for erosion assessment.

Converting the average total annual observed soil loss per land cover land use types to millimeter lowering of the land surface per year, using Table 18 and the calculated value of 1mm=13 tons/ha, gives the following approximate values per sample area, which also reflects the differences between the observed and the USLE soil loss.

Table 18: Observed and predicted soil losses stated in mm/year

Type of estimation of soil erosion	Soil loss cover type (mm/year)					
	Dryland agric.	Mixed garden	Rubber plant.	Bush	Sec. forest	Bare land
Observed erosion	6	2.8	3	2.3	2.6	7.2
Predicted erosion by the USLE	1	0.4	0.5	0.2	0.3	1.3

8.4.5 The relationship between erosion and locational attributes

This section is a descriptive analysis of the relationship between erosion severity classes and locational erosion influencing factors. It is intended to provide better insight into the establishment of the observed patterns. The analysis was based on a graphical analysis as described below.

Relationship between erosion and terrain

Both the observed and predicted erosion were used in this correlation. From field observations there is considered to be no significant linear relationship between terrain characteristic, i.e., soil erodibility (K) factor, the observed erosion and the erosion predicted by the USLE (see Figures 33a and 33b). The relationship between SL factor, the observed and USLE predicted erosion, was, however, slightly stronger (in Figure 34a and 34b). This demonstrates that within the training areas soil erodibility and slope length are of limited importance for predicting erosion values. In other words, this shows the importance of the role of vegetation cover in reducing erosion.This also implies that the causative effect of erosion influencing factors is different. In many cases one factor may accelerate the processes while another reduces them. In addition, there are other factors that have not been considered but which affect the total result.

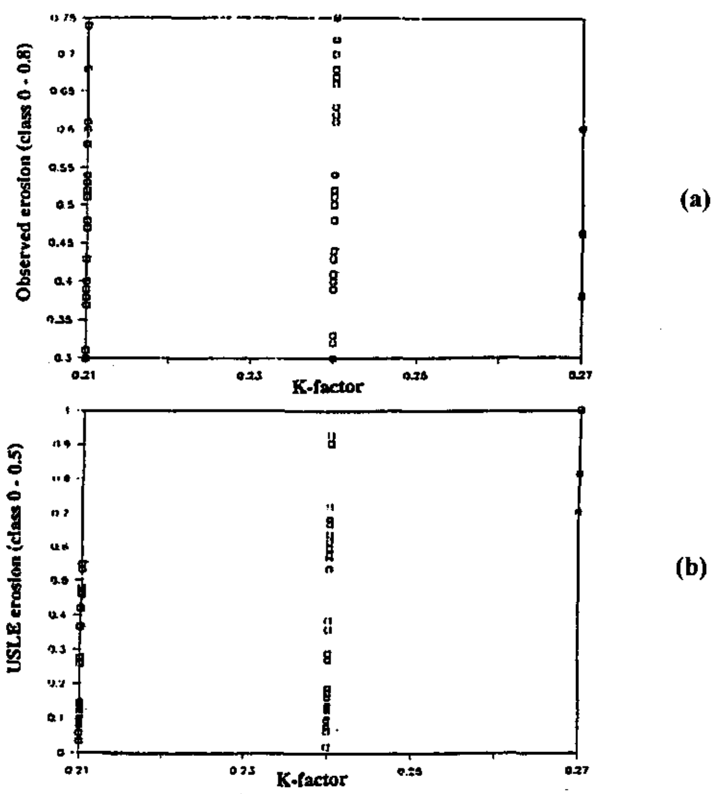


Figure 33: **(a) The relationship between K-factor and observed erosion;**
(b) The relationship between K-factor and predicted erosion by the USLE

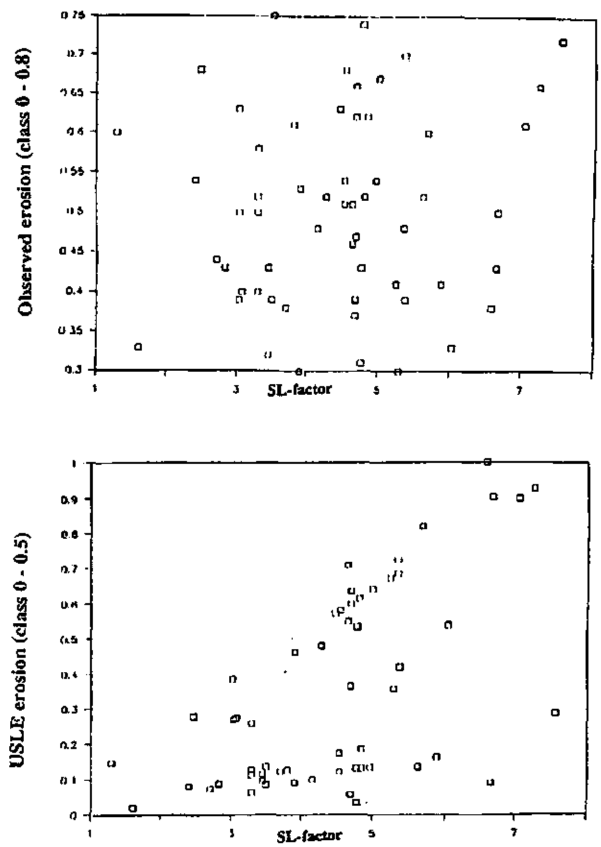


Figure 34: **(a) The relationship between SL-factor and observed erosion; (b) The relationship between SL-factor and predicted erosion by the USLE**

Relationship between ground cover, root exposure and bare soil

The percentage of bare soil associated with a particular ground cover was used in assessing this relationship and was recorded during fieldwork. See Appendix 7. High bare soil percentages occurred in the dryland agriculture compared to the other cover types, because of burning of fallows. The annual range of bare soil for fallows is also higher than for the other cover types. Therefore it can be assumed here that the relationship between the ground cover and bare soil is mainly attributable to the inclusion or exclusion of the stone cover in which affect ground cover computations (see also Thornes, 1995). It also does not reflect the actual percentage bare soil recorded at the time of fieldwork.

The effect of ground cover in reducing erosion has been under investigation by researchers. The general finding of this research is that the lower the ground covers percentages the higher the soil loss (in Figure 35).

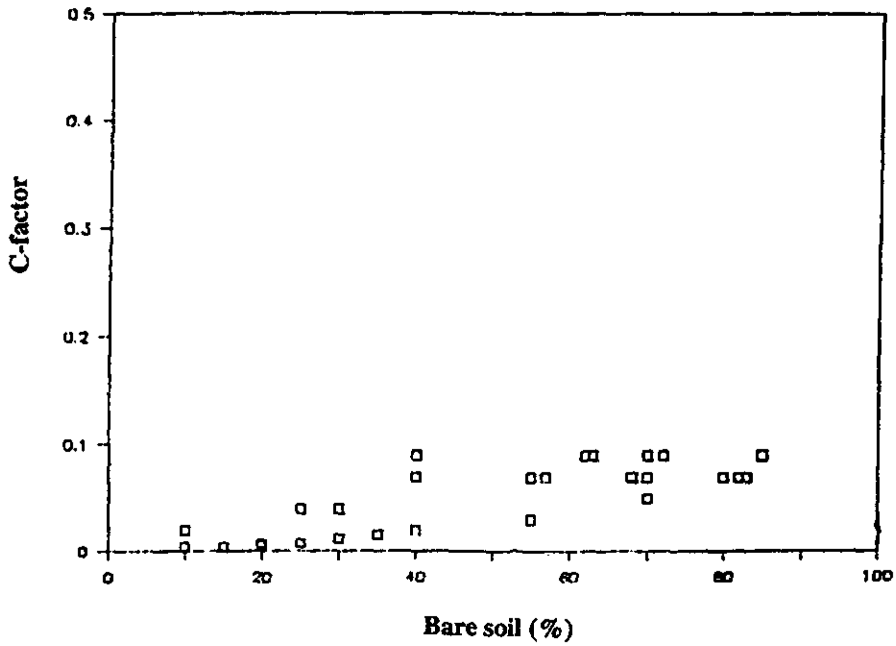


Figure 35: The relationship between bare soil and C-factor

In the sample areas, sites with low percentage of bare soil were associated with low root exposures, especially in the secondary forests and bush. There are however some discrepancies where root exposures occur on sites with a very low percentage (5%) of bare soil. A possible explanation of this follows.

(i) Erosion is not just a function of bare soil or total groundcover, but also a function of distribution. According to Wischmeier and Smith (1978), Dissmeyer and Foster, (1980) " a random distribution is assumed in most surveys, concentrations and interconnections of bare patches may cause larger runoff and soil losses than originally was expected." Where there are no interconnecting bare surfaces and no exposed surfaces around plant bases erosion would not be expected. (ii) Erosion could have occurred because of water running under freshly fallen leaves on the surfaces (see also Poesen, 1995). Conversely no root exposure occurs on some sites with a high percentage of bare soil. This could be due to washed soil covering exposed roots, or the growth characteristics of plants like *Imperata Cylindrica.*

Relationship between ground cover and observed erosion

There is a strong correlation between the ground cover and the highest observed erosion recorded in each sample area (in Figure 36). The low coverage for dryland agriculture was derived as a result of the surface stone cover incorporated in the calculation.

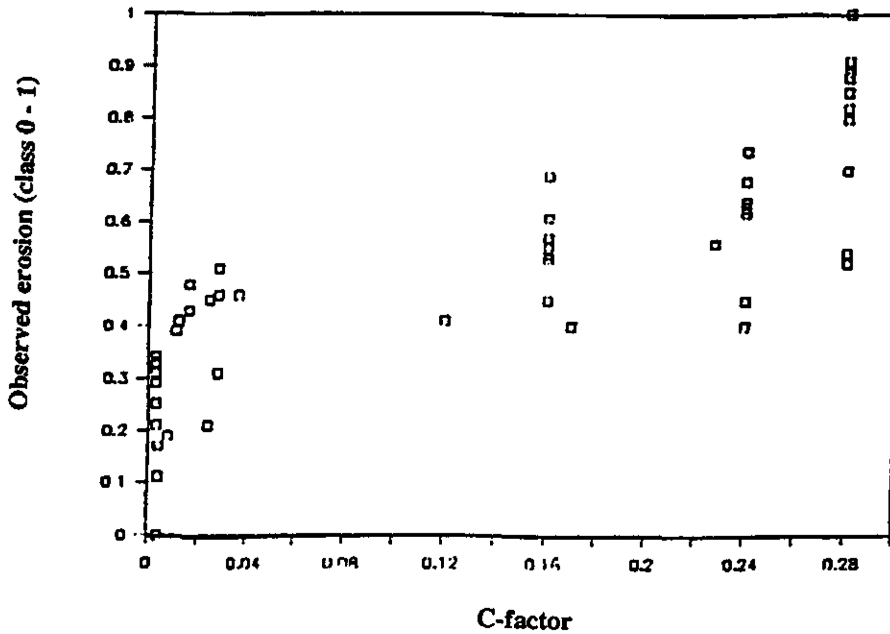


Figure 36: The relationship between C-factor and observed soil erosion

The observed erosion however is much higher than what one can expect from the effect of ground cover. This is probably due to the fact that surface stones are not as effective in reducing erosion as previously thought and this could be due to the following reasons: (i) the soil underneath the surface stones is usually protected from sealing. In some cases the effect of surface stones can be compared to the effect of stem flow under vegetation, where relatively large amounts of runoff from the stones can flow to small interspaces and cause erosion (from Kooiman 1987 after De Ploey 1985); (ii) depending on the rock fragment position, size and on fine earth porocity (Poesen and Ingelmo-Sanchez, 1992; Poesen 1995).

The above conclusions however do not imply that the soil loss estimated from field observations is completely correct. Overestimations could have occurred, especially with regard to the measurement of root exposure as previously mentioned.

In some sample areas the relationship between the ground cover and observed erosion is very strong. This can be seen by the big differences of soil loss between predicted and observed under mixed garden or dryland agriculture cover types (see Table 17). It was found also by Sub Balai RLKT(1985) that the highest effect of ground cover of 0.5 also falls within the range of the highest observed erosion.

The observed patterns in the study area

The observed pattern (the relationship between identified erosion class and its locational attributes) was established at each observation point. It represents the occurrence of a particular erosion process at the FFL. An example of the typical observed pattern is given in Table 19, showing knowledge expressing that at observation point 13 the process of soil erosion is present. By using erosion indicators (a), (b) and (c), observation 13 is classified as severely eroded with erosion classification degree (ECD) of 0.80 and associated with erosion influencing factors or locational attributes (1), (2), (3), (4), and (5).

From observed patterns and their causal locational relationship (e.g., from Table 19), a set of decision rules for every single observation both for categorical and non-categorical attributes at farmers' field level was established to predict the erosion class of any uniform area (i.e., a TMU) within the same geographical region. See Subsection 8.6.2. Inducing these decision rules into other uniform areas, the soil erosion process was predicted or classified and finally regionalised on the map. Figure 37 depicts the induced erosion classes namely: (i) soil erosion is absent (class 0); (ii) very slightly eroded (class 1); (iii) slightly eroded (class 2); (iv) moderately eroded (class 3); (v) severely eroded (class 4).

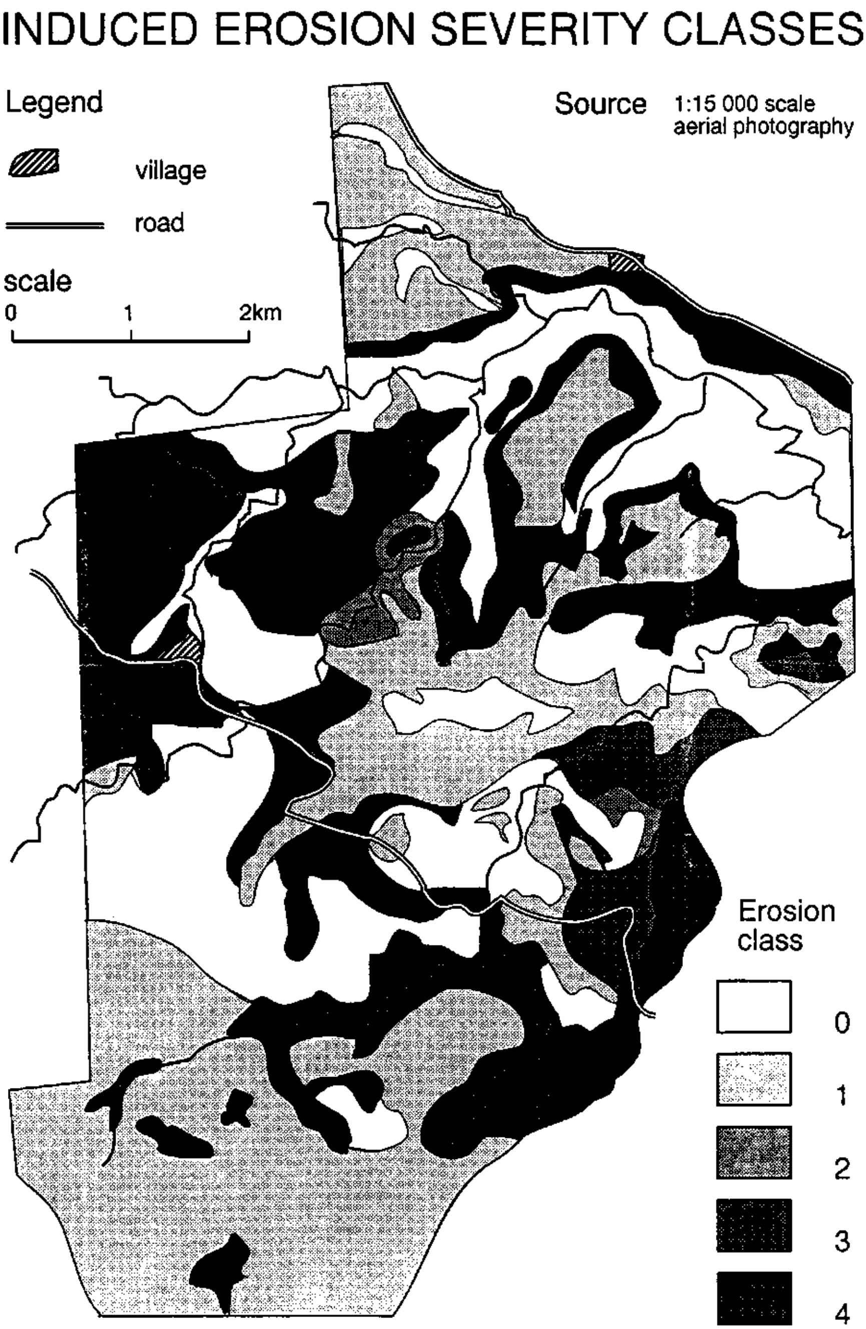


Figure 37: Induced erosion classes in the study area

Table 19: An observed pattern and locational attributes at observation point 13

Observation point nr.	13
Terrain mapping unit nr.	115
Erosion identification	***PRESENT***
Erosion classification degree (ECD)	0.8
Erosion class	***SEVERELY ERODED****
EROSION INDICATORS	
(a) Splash erosion degree#)	0.78
* sealing degree	0.78
* pedestal degree	0.78
(b) sheet erosion degree	0.81
* soil accumulation degree	0.81
* root exposure	0.81
(c) rill erosion degree	0.95
(d) soil surface condition	0.66
LOCATIONAL ATTRIBUTES@	
(a) slope steepness (S)	26%
(b) slope length (L)	20m
*** SL-factor#)**	1.55
(c) soil erodibility (K)	
* soil residual degree	1.01
* soil reconsolidation	1.14
* organic matter	1.00
*** K-factor**	1.15
(d) ground cover (C)	
* mixed garden with young cassava	
*** C-factor**	0.28
(e) EI_{30}	9.20%

note:#) = weighting value derived from figure, published information and field observation;it is treated as dimensionless;

(*) = observed erosion severity class obatined using erosion indicators;

@) = used for comparison between observed (*)and estimated soil erosion;

8.5 Testing and developing an Inductive Erosion Model

8.5.1 Introduction

Sections 8.2 to 8.4 discuss the entire procedure of building the proposed IEM and the results of erosion survey at FFL. This part of the thesis discusses the performance of the proposed IEM in predicting the occurrence of locational erosion severity classes at different aggregation levels of TMUs.

In order to know the robustness, i.e., how the model performs, and the sensitivity of the model to the several datasets, the proposed IEM model was tested in the entire study area. The evaluation includes the stages of obtaining datasets and certainty factors (CF) values. A more elaborate discussion on CF and the quality of information produced by an IEM in a GIS is given in Chapter 10.

The evaluation has been completed by comparing predicted erosion severity classes yielded by an IEM and USLE to the visually observed classes which are referred to as the ground truth in each observation point.

Observation points were randomly selected and not used in an IEM model building (see Section 6.5). Using similar erosion indicators per observation point and procedures as applied earlier the USLE model was compared to an IEM.

8.5.2 Methods and techniques used

Methods and techniques used involve a check procedure that includes the comparison between observed (true) and predicted classes. The implementation of this evaluation involved also field checking at selected observation points (check locations).

Soil erosion maps predicted by the USLE and an IEM were prepared. The results of the erosion study as described in Section 8.3 were particularly obtained from an IEM model building which identifies four erosion severity classes. The study involved a minimum of 30 samples per class as recommended by Hay (1979). On a topographical map at scale 1: 50 000, 120 point observation units were randomly selected and spread throughout the entire study area.

Maintaining the objectivity of evaluation is very important. To achieve this, four groups of three persons consisting of two soil conservationists and one assistant surveyor were assigned independently to observe, identify, quantify and classify erosion severity class and erosion influencing factors (ground truth) at each selected observation point.

On the same day each observation point was observed alternately by four groups of observers. The data including the most influential factors of ground cover, soil erodibility, intensity of rain within 30 minutes (EI_{30}), slope length and slope steepness were collected and analysed. The EI_{30} measure was assumed to be uniform. The results of field observation and erosion classes predicted by an IEM and the USLE are compared as presented in Section 8.5.3.

8.5.3 Results

The result of the check procedure is presented in the form of an error matrix which purports to show the probability (reliability) that a particular check location is severely eroded, moderately eroded, slightly eroded, or very slightly eroded. The result appears very promising as shown in the following error matrix.

Table 20: Error matrix of an IEM compared to visually observed (true) class

Visually observed class	Erosion severity classes predicted by an IEM				Total
	I	II	III	IV	
I	_107_	13	0	0	120
II	7	_101_	13	0	121
III	0	8	_112_	0	120
IV	0	0	0	_119_	119
Total	114	122	125	119	480

Note:

(i). underlined figures are correctly classified

(ii). I = very slightly eroded; II = slightly eroded; III = moderately eroded; IV = severely eroded

(iii). the visually observed class was obtained using methods and techniques used as explained in Section 8.3

From Table 20 we obtain the overall accuracy (reliability) of an IEM model prediction to be 91% (439/480 * 100%). This 91% figure shows that from 480 randomly selected check locations 439 are correctly classified. Confusion classes (only about 10%) are found between slightly eroded (II) and moderately eroded (III) classes as well as between very slightly eroded (I) and slightly eroded (II). The main cause of this is explained in Section 8.4.3.

From Table 21 we obtain the overall accuracy (reliability) of the USLE model prediction, to be 54% (261/480 * 100%).

From the results of this comparison we conclude that the proposed IEM model gives a much better prediction than the USLE.

From Table 20 below we also obtain the following:

% IEM identified class I and USLE identified class I	=	44%
% IEM identified class II and USLE identified class II	=	48%
% IEM identified class III and USLE identified class III	=	71%
% IEM identified class IV and USLE identified class IV	=	53%

Table 21: Error matrix of the USLE compared to an IEM class

Class predicted by an IEM	Erosion severity classes predicted by the USLE				Total
	I	II	III	IV	
I	<u>53</u>	33	14	20	120
II	4	<u>58</u>	38	20	120
III	0	8	<u>86</u>	26	120
IV	13	13	30	<u>64</u>	120
Total	70	112	168	130	480

Note:

(i). underlined figures are correctly classified

(ii). I = very slightly eroded; II = slightly eroded; III = moderately eroded; IV = severely eroded

8.6 The incorporation of an IEM into a Geographical Information Systems environment

8.6.1 Introduction

The procedure of induction is similar to the procedure of regionalisation in hydrology. It is considered as a part of a learning system, in which decision rules are developed from an examination of a set of known objects and their associated attributes, called the observed pattern. Using these rules, unknown objects, i.e. areas of unknown erosion severity class, but having known erosion foramtion factors, can be classified (induced) into one of the predetermined erosion classes. In this regard, the robustness of the rules associated with the proposed IEM was tested in Section 8.5.

This part of the thesis, however, is intended not only to demonstrate how to classify terrain objects but is also aimed at providing better insight into how induction operates. Using inputs from Subsection 8.2.3. and with regard to Subsection 1.1.2, the major component of this part describes how the induction of decision rules or the incorporation of expert's inference has been added to a GIS environment and how the process of induction is carried out.

The second component of this part describes the incorporation of decision rules into the prediction of the occurrence of erosion severity classes at different hierarchical levels. With regard to the proposed IEM, these hierarchies are closely related to the linking between observation points at FFL (level -1), slope level (Level 0) and larger parts of the terrain (level +1). See Sections 7.3 and 7.4.

8.6.2 The incorporation of induction task (rules) into a GIS

From erosion survey at FFL, a soil conservation expert (see again Section 8.4.) gains knowledge on locational soil erosion classes and their spatial influencing factors. This knowledge can be formulated in the form of a proper structure and is expressed in the form of decision rules (FAO, 1976; Negoita, 1985; Mortimer, 1985; Bonnet, 1985; Rodrigue, 1995). These rules are usually constructed constantly in the form of *"If.....Then"* structures representing a part of the problem solving knowledge (chain of reasoning) of the expert. Thus, the *If...* part can be regarded as a list of conditions e.g. locational erosion influencing factors, and the *Then* ......part is a list of conclusions concerning underlying processes e.g. erosion severity classes (see again an example of observed pattern in Table 19).

Looking further at the *"If.....Then"* structure, one has to understand that each rule associated with this production system is an independent piece of knowledge, containing all conditions required for its application. Therefore, it can be created according to a certain condition activated at any moment and represented as a domain specific set of conditional rules. Decision rules adopted by an IEM may also have an overall goal of confirming or discarding hypotheses on the basis of the facts introduced by users. As stated ealier, this mechanism has to have the capability to cope with competing hypotheses and may lead to conflict resolution (Suryana, Molenaar and Imeson, 1996).

As mentioned in Sections 2.3 and 4.5 each erosion influencing factor included in the proposed IEM stands for an independent hypothesis related to a particular erosion class. From erosion survey at FFL it was found that different combinations of influencing factors (conditions) may yield the same erosion severity class (a conclusion). Therefore, an important step in the incorporation (computerisation) of the proposed IEM into a GIS is to select the set of observed patterns, i.e. erosion severity classes and their associated locational attributes obtained from Subsection 8.4.6. This was obtained using regression or correlation analysis. The bigger the correlation coefficient (r) the stronger the relationship between erosion influencing factors and a particular erosion class.

A method called the Classification Regression Tree (CART) from Breiman (Ward, 1963; Stanat and Allisteter, 1977; Breiman, 1984) was adopted in this study. This method facilitates the establishment of a set of decision rules attached to each node of a decision tree. Each node of the tree represents one state (hypothesis) on each possible causal relationship between erosion influencing factor and erosion class, and each edge represents a possible transition from one state to another. See Figure 38. Imposing a set of decision rules on one or more data attributes one can detect a particular erosion class at the end of a particular node. The entire procedure is outlined below.

The program starts with one of the erosion influencing factors (partition), e.g., rainfall intensity within 30 minutes (R-factor) and goes down to the second partition, e.g., slope steepness (SS-factor) and the third slope length (SL-factor). It then examines the uniformity of the descendant of these erosion influencing factors particularly if these partitions consist of a range of more than one value, e.g., slope steepness consists of several erodible slope steepness classes. The inference process continues until the end of the erosion influencing factors involved. Thus, each observation down the decision tree represents a set of decision rules under which the most probable model prediction is made. In other words, each connected node within a particular decision tree represents the significant relationship between erosion class and associated locational attributes.

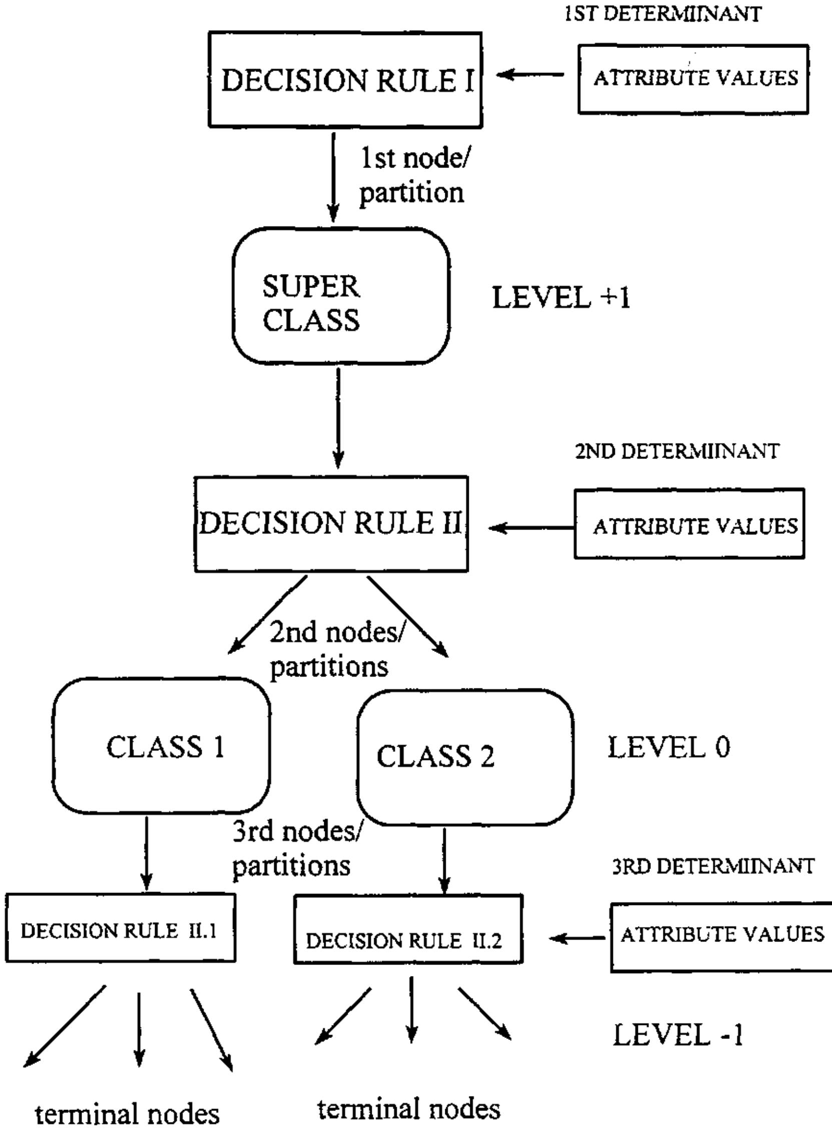


Figure 38: Decision tree adopted by an IEM (modified from Huising, 1993)

The procedure then repeats itself with each descendant partition. Partioning continues until all objects have been correctly classified: a terminal node has been reached. Finally, one of the best combinations (the best tree) is selected. A simple Pascal program was written to perform this operation (Appendix 8). After syntax checking, translation, and creating an executable program, the rule is stored in a GIS environment, e.g., GIS-ILWIS system under the option of inductive spatial modelling.

Further the established rule is used in the prediction of the occurrence of erosion severity classes in other similar areas. An unknown erosion severity class at Sub TMU level but with known locational attributes associated with particular erosion class, is allocated by an IEM to this erosion class.

From the decision tree (Figure 38) it becomes apparent that the accuracy of an IEM model prediction increases with the number of terminal nodes, which have to be labelled. Finally, many terminal nodes may have the same label. In other words, a combination of more erosion influencing factors (data) reduces the danger of making a misclassification. But it should be remarked that an additional partition step (additional data) may not result in a very significant improvement in an IEM model prediction, because a very big decision tree may possibly only produce one classified object. Too short a decision tree may result in an unacceptability high level of misclassification. This statement implies that finding the best combination of erosion influencing factors is very crucial in incorporating an IEM into GIS. For this Walker and More (1988) suggest to use the rule of tree selection.

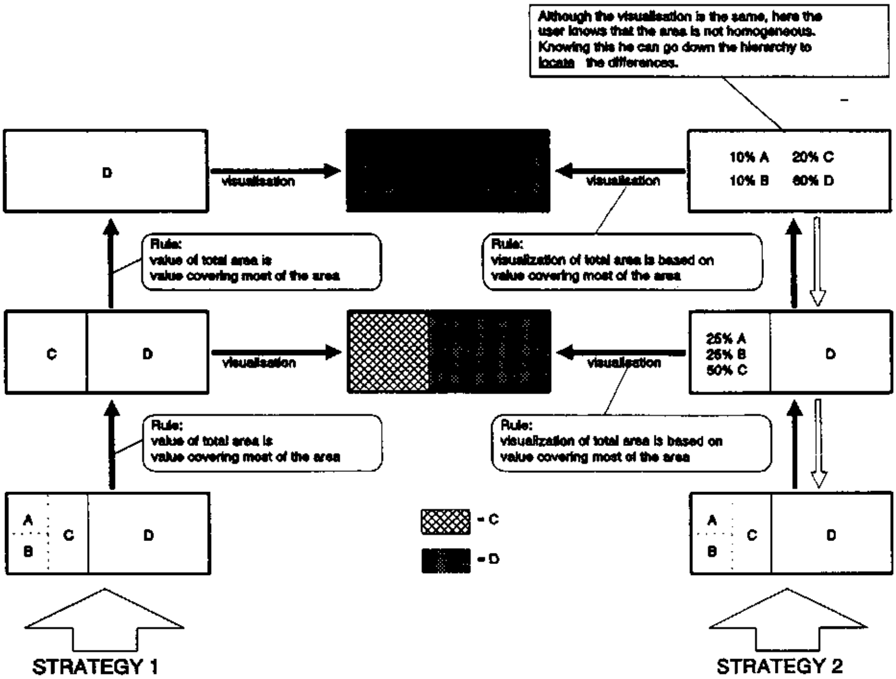
8.6.3 Class generalisation and object aggregation adopted by an IEM

This part of this thesis discusses the application of decision rules adopted by the proposed IEM in transforming data classes into erosion information categories at different aggregation levels. As discussed previously in this thesis, the terrain object description, e.g. severely eroded TMU, is considered as a function of users' context (requirements) and terrain object definition at different aggregation hierarchy levels. The hierarchical order of TMU has been established (see Chapter 7).

In this regard, the hierarchical order of soil erosion information categories follows the hierarchical order of TMUs. This will imply that each aggregation level of TMUs holds its own data classes (formation factors) and erosion information categories (erosion severity class). The necessary generalisation of the TMUs should also generalise correctly the soil erosion information. In actual fact each level of the TMU hierarchy includes information on the soil erosion classes. Therefore, attribute e.g. soil erosion class generalisation is implicit.

As stated earlier the conceptual approach for erosion modelling at regional level was based on the establishment and mapping of eroded mappable homogeneous terrain mapping units. In addition to this, the spatial characteristics of the TMUs must also be generalised as scale decreases. This is not a problem for adjacent areas belonging to the same higher level TMU: they are simply merged. However, it may occur that there is a patchwork of low level TMU's belonging to different classes at higher level. See Subsections 3.5 and 3.6.

A cartographer would begin by omitting all very small areas. Information is therefore lost on the map, but of course it is maintained in the data base (see Figure 39). However, regional planners want general information only.



(a) **(b)**

Figure 39: Erosion generalisation strategies adopted by an IEM

In the context of the erosion model there is no problem if the omitted areas belong to the same erosion class as the areas with which they are merged. For example the TMUs *"A"* and *"B"* in Figure 39 are adjacent and are classified as severely eroded with area cover of 4.19 Ha and 9.90 Ha respectively. Using dissolution of these adjacent boundaries a homogeneous severely eroded TMU *"C"* at higher level can be established.

If, however, adjacent merged TMUs belong to a different class (see Figure 39(b)), the erosion class of the merged area may require to be adjusted. To overcome this problem, two strategies which can be done manually or by computer algorithm, depending for example on the relative area (percentage) of the omitted sub(sub) TMUs have been proposed (Suryana and Molenaar, 1995).

Strategy 1 omits small, unmappable classified TMUs, and merges them with the adjacent or surrounding larger TMUs (see again Figure 39a). This strategy is similar to the traditional analogue generalisation techniques i.e. creation of new spatially mixed (sub) TMU classes. The advantage of this procedure is simplicity. The disadvantage is that the information on the fixed relationship, at higher levels, between (sub) TMUs and erosion class is lost. This will have an adverse affect on the top-down modelling process.

Adopting *strategy 2* it may be possible in a computer GIS to specify a range or percentage of each erosion class invoved for each classified (sub) TMU at the higher level in the hierarchy to accommodate these spatial considerations, and still maintain consistency in an IEM modelling procedure. Thus, the advantage of *strategy 2* is maintenance of original information i.e. the percentage of each classified TMU in the data base. Its disadvantage is increased complexity. *Strategy 2* should be adopted for a top-down modelling procedure.

Figure 39 shows that although *strategy 2* retains more information in the GIS, its visualisation may appear the same as the visualisation of the result of *strategy 1*.

8.7 Conclusions

The information on the spatial distribution of soil erosion obtained from Chapters 6 and 7 of this thesis provides very useful inputs for building the proposed IEM. The results of erosion survey at farmers' field level provide information on erosion influencing factors, erosion severity class, erosion features and their relationship. Considering the hierarchical level associated with TMU as well the relationship between TMU and an IEM this information can be used further as a basis for the establishment of decision rules for the prediction of soil erosion at different aggregation levels.

The process of soil erosion is a phenomenological process. It is determined by interrelated influencing factors. By this is meant that the degree of erosion severity is not determined by only a single dominant influencing factor but it is caused by joint effects of different independent factors. The interrelationship between factors involved is different. One factor may accelerate the process of soil erosion leading to more severe erosion but other factors may decelerate it.

The process of soil erosion is region specific and follows a specific pattern. Thus, the variation of soil erosion on a specific site is not caused by the effect of random variables but it is determined by locational factors.

Considering the specific character of soil erosion process, the deductive approach model, i.e., USLE is not fully applicable. Thus, it requires a particular inductive approach in handling local variability of soil erosion.

The concept of induction set up in a GIS environment can be established in the form of GIS inference capability. This concept is required in order to increase the added value of a GIS in handling a particular task.

Chapter 9

The attribute accuracy aspects of the interpreted aerial photo TMUs used in the FEDP for the land rehabilitation and soil conservation programme in Indonesia

9.1 Introduction

This part of this thesis was written in connection with uncertainty assessment associated with data inputs into a GIS environment in general and into the computerised FEDP in particular. The discussion includes the preprocessing stage, data acquisition and data type (continuous or noncontinuous data). It proposes approaches that may be useful for handling data quality aspects associated with data gathered from different sources, used in the GIS environment.

The uncertainty associated with data inputs was analysed firstly by looking at how these inputs are collected either using field measurement (survey) or using remote sensing image interpretation (Section 9.2). Secondly the analysis was done based on uncertainty aspects of GIS, particularly the attribute aspect (Section 9.3).

Using standard deviation of the attribute data derived from continuous classified polygon maps and topographic maps and check procedures used to evaluate the attribute information in categorical coverage maps are given special emphasis and practical examples are given.

The first part of the attribute accuracy assessment is concentrated on quality assessment associated with the situation when all possible techniques in deriving probability values can be performed. In other words, the data quality is obtained and represented using formal statistical parameters (Sections 9.4 and 9.5). The alternative method and technique, i.e., using certainty factor (CF) is discussed more elaborately in Chapter 10.

9.2 Data sources for data input, data capture and error in GIS

As stated earlier the FEDP for land rehabilitation and soil conservation involves various sources of data from different areas of expertise. The quality of the data sources used in the FEDP effects the quality of resulting information handled by the GIS. This is because no single data input (map) is completely error free.

The main sources of data inputs into the FEDP and into GIS are: (i) existing topographical maps; (ii) existing categorical-coverage maps; (iii) visually interpreted imagery; (iv) digital classification of remote sensing imagery; (v) direct measurements. These main sources of data inputs are gathered in the field or in the office, interpreted processed and assessed intuitively. Thus, these data inputs contain error.

Moreover, when these data are entered into a GIS uncertainty is also produced during the transformation of these sources into digital form. For example a soil depth map as input into the FEDP is mainly established through interpolation from point observations. The categorical data inputs, e.g., land use map do not often represent the reality due to fuzziness. In this regard, the boundary of land use categories is usually presented using a crisp boundary but the boundary of these types of data is often gradual.

Between the data gatherers and GIS users of geographic data as stated above, the topographic scientists have established standards for describing data quality. For example this group of scientists has considerable experience in the use of models for processing their data, making good use of variance propagation to estimate error associated with these models (Topping, 1979).

Moreover, the links between data gathering, data capture, processing and representations of resulting information from various areas of expertise are different. This implies that knowledge of data and the processing models used on those data are strongly linked in some disciplines but less linked in others. For example in the topographical area of expertise they may cover all aspects of defining the shape or form of entities, labelling surface features and presenting the resulting information to the public. Thus, topographic science includes both knowledge of the data capture methods and the models used to provide information for the public.

However, this is different from an Earth scientist, e.g. a soil scientist, who is assigned to evaluate potential areas for a specific use. For this purpose this scientist operates the models used using data that may be gathered partly by different groups of specialists. Thus, soil scientists may have less ability in deducing error in the original data than the topographic experts. In Chapter 10 a simple example shows the difficulties faced by soil conservation specialist in deducing error originating from a simple model, namely the inductive erosion model.

9.3 Uncertainty aspect of GIS data inputs

The attribute aspect of the FEDP may take different forms, either categorical (discontinuous) or non categorical (continuous) data types. From this apparently since the deductive method is used to determine errors associated with data inputs and resulting information, a good knowledge of the data inputs and processing model used to produce the resulting information is required. Resulting information is processed and stored in information system using dedicated software. The software may affect indirectly the quality of the resulting information. However, this quality aspect is not considered in this study.

It has been suggested by several researchers (e.g., Moellering, 1986; Chrisman, 1984) to look at five aspects of geographic data quality as follows:

(1) attribute quality of the data;
(2) positional quality of the data;
(3) lineage of the data;
(4) completeness of the data;
(5) logical consistency of the data;

The attribute quality of the data in this study is only related to values, substance and characteristics and is associated with geographic objects including man-made objects or coverages. These attributes in a GIS must be labelled. However, such attributes may be completely right, partially right or wrong.

Considering the second aspect, i.e., the positional quality of the data, it has been found that there is great variety between GIS developers and users. For users operating at large scales like in public works, civil engineering and cadastre, a GIS may store information about well defined man-made objects, for example the location of demonstration plots or small check dams. These objects are usually defined in two or three-dimensional coordinates. These measurements are associated always with *standard deviation* (*sd*) derived from repeating measurements (Wheeler and Lyday, 1990). Therefore, *sd* can be used to assess the positional quality of data (Drummond, 1987; Drummond, 1991; Chrisman, 1984). Users of small scale GIS like in Earth sciences are more concerned with natural objects, especially coverages that are usually represented using polygons. In this case the coverage may not be located easily and the *sd* of each boundary may not be relevant.

The third aspect of quality, i.e., the lineage is the description of the date, processing methods used, and sources to generate particular data sets. Thus, this aspect of quality describes the historical background of data input to a GIS. Therefore, these aspects cannot be used to test data for their lineage but can be used to provide input for deducing aspects of their quality only.

The remaining two aspects of data quality are outside the scope of this thesis. In Appendix 9. generalised flow charts for handling data quality are given.

9.4 Quality assessment of existing categorical maps: non-continuous data type

9.4.1 The application of external testing

As stated in Section 4.3, the application of external testing requires adequate sampling, for example randomised sampling, with a sufficient sample size. A minimum of 30 samples per class is required in determining % correctly classified per class.

In this exercise 530 sampling locations were randomly selected. The existing present land use map made by the FEDP Team and carried out in the year 1986 was updated using TM 121/65 (1991) and used to examine the classification accuracy of 11 classes out of 13 classes of land use/ land cover (see Table 5 in Chapter 5 of this thesis). The methods and techniques as discussed in Subsection 4.3.2 were use. The result is presented in the following error matrix (Table 22).

Table 22: The error matrix of the present landuse

Visual/ observed classes	**Interpreted classes**											
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**Total**
A	*60*	-	-	-	-	-	-	3	-	-	-	63
B	-	*8*	-	-	-	-	-	-	-	-	2	10
C	-	-	*4*	-	-	1	-	-	-	-	1	6
D	-	-	1	*81*	11	5	9	3	-	-	-	110
E	-	-	1	-	*69*	-	-	-	-	-	6	76
F	-	-	1	-	3	*28*	-	-	-	6	-	38
G	-	-	2	3	-	-	*3*	-	-	-	-	8
H	15	-	-	9	-	-	-	*51*	-	6	-	81
I	-	-	-	-	-	-	-	-	*6*	-	-	6
J	3	-	-	6	-	-	3	-	-	*106*	-	118
K	-	4	-	-	-	-	-	2	2	-	*6*	14
Total	78	12	9	99	83	34	15	59	8	118	15	530

Note:

(i) underlined italic figures are correctly classified

(ii) A=rice field;B=protection forest;C=rubber plantation;D=coconut plantation; E=teak forest; F=mixed garden; G= settlement;H=dryland agriculture;I=bush;J=bareland;K= secondary forest

The matrix clearly shows the following:

(1) From the 530 sample check points on the updated landuse map, 422 were correct. In other words, by applying equation (3) in Subsubsection 4.3.2.2 a classification accuracy of 79 - 80 % for the whole survey can be determined. This represents the overall accuracy estimate (Pov) at coverage level.

(2) Using equation (2) as described in Subsubsection 4.3.2.2 the class probability estimate for entity levels can be determined as follows (A actually classified as A and so on):
A 95 %; B 80 %; C 67 %; D 74 %; E 95 %; F 74 %; G 38 %; H 63 %; I 100%; J 89 %; K 43 %

9.4.2 The application of Kappa statistic

As stated in Subsubsection 4.3.2.2 we need to remove the effect of chance from the overall attribute accuracy estimate (Pov). The Kappa statistic is addressed to handle the problem associated with the percentage correctly classified per class or for the whole set due to chance.

From the 530 observation points as presented in the error matrix above, we obtain the following specification:

From the matrix it can be seen also that the observed accuracy estimate (Pov) is 79 - 80 % which is greater than the expected probability estimate of 15 % (Pe). The Kappa statistic (K) is a function of both observed accuracy estimate (Pov) and the expected probability estimate (Pe). Based on equation (5) in Subsubsection 4.3.2.2 the expected probability due to chance (Pe), particularly after random classification is 0.15. Subsituting Pe of 0.15 in the equation (4) we obtained the Kappa statistic (K) of 76%. From this can be concluded that after removing chance the overall accuracy estimate has declined from 79 - 80 % to 76 %.

9.4.3 The application of internal testing and the percentage correctly classified at stated confidence level

This method involves the determination of the consistency of attributes assigned to polygons during survey for a categorical coverage map.

The determination of consistency requires repeated classification of polygons as found in the resurveyed area by several independent workers using the sampling procedure usually used in soil and vegetation survey.

To explore an example of the application of internal testing procedure the landuse map (see landuse/landcover map in Chapter 5) was resurveyed by superimposing 10 (ten) aerial photo interpretation results of the same part of the study area. About 150 randomly selected check points on each map were selected. In this exercise the classification accuracy of nine landcover classes was estimated.

In this exercise, the true class is defined as the majority class found by different interpreters at the same observation point in a particular landuse class done by 10 professional aerial photo interpreters. Therefore, the matrix used in internal testing is not the same as a misclassification matrix (error matrix) used in external testing as discussed previously. It can be used to provide a classification accuracy statement, or probability that a polygon is assigned its correct attribute on one of the maps in that class, for example the true map accuracy range at the 99.70 % confidence level.

Table 23: The result of internal testing of 150 randomly selected observation points on 10 landuse/landcover maps

True or majority classes	Photo interpreted classes									
	A	B	C	D	E	F	G	H	I	Classific-ation accuracy*)
A	*62*	0	0	3	0	8	0	0	6	78%
B	5	*609*	0	0	0	0	0	0	13	97%
C	0	0	*75*	0	1	4	4	0	11	79%
D	0	3	1	*255*	35	14	0	0	0	83%
E	1	0	2	3	*66*	4	0	0	0	97%
F	12	0	0	4	9	*69*	0	0	0	73%
G	0	0	0	0	0	0	*100*	0	0	100%
H	1	0	2	0	0	3	0	*4*	2	33%
I	4	4	5	0	1	3	0	9	*83*	76%

Note:

(i) underlined italic figures are correctly classified

(ii) *)=calculated as the percentage of true class correctly classified by photo interpretation

(iii) A=rice field; B=protection forest; C=rubber plantation; D=coconut plantation; E=teak forest; F=mixed garden; G= settlement; H=dryland agriculture; I=bush.

Considering the class of landcover protection forest (B), if confidence level is 95% or 95 % probability, *Za* (probability value derived from the normal distribution curve) equals 1.96 and the total sample size (N) equals 627 as stated in the Table 23, with the percentage correctly classified class (x) of 97 %, then by applying the equation (6) in Subsubsection 4.3.2.2 the lower limit (y_l) of 95.33 % and the upper limit of (y_2) is 98.09 %. Based on this calculation, we can inform GIS users the true map accuracy, e.g., landcover protection forest (B), is between the range of 95.33% and 98.09% correctly classified, with a 95% confidence level for sample accuracy of 97% from 627 checkpoints.

9.5 Attribute accuracy assessment associated with classified continuous data

9.5.1 Introduction

This section is intended to assess the accuracy of classified continuous attribute data. An example as found in this study concerns the allocation of TMUs to slope steepness classes.

At the class level of the TMU hierarchy, each TMU has a single slope steepness class attribute (see again Sections 6.4 and 7.2). The first step in classification of the slope data is to decide slope steepness classes. Having done so, the next step is to estimate the accuracy of the slope steepness data, e.g., correct ranking, and therefore derive a general measure of uncertainty in the allocation of each TMU to a particular slope steepness class. Slope steepness for each TMU was calculated from the contour spacing on the 1: 50 000 scale topographic maps. Suitable slope class boundaries and their description (see Tables 7 and 8 in Section 6.4) were derived and the class polygon boundaries were drawn, with reference to the contour spacing.

Several sources of error can arise associated with the establishment of a slope steepness map. These includes (i) probability of misclassification particularly for observations falling near slope steepness class boundaries; (ii) the description of slope steepness classes, e.g., "nearly level" and "gently undulating" implies the fuzziness or gradual transition of their boundaries which cannot be sharply defined; (iii) the number of classes may be incorrectly chosen.

In the case of (i), the probability of misclassification near class boundaries can be found using standard deviation of slope steepness, interpreted from maps, assuming a normal distribution of errors. In other words, this is the precision measure.

In the case of (ii), the problem is how to relate fuzzy verbal descrption, e.g., gently undulating, to crisp class boundaries, e.g., 15 - 25%. This is solved by means of membership functions.

In the case of (iii), if there are too many narrow classes errors will increase, if there are too few broad classes then the usefulness (for erosion modelling) decreasees. Also, the classes themselves must coincide with typical slopes found in the terrain, with class boundaries coinciding ideally with breaks of slope.

In this research, the measure of dispersion *"the standard deviation (sd)"* in equation (7) was obtained and the slope steepness map was tested by comparing the map class values with slope measurements made in the field at nearly 150 locations (often more than one per TMU).

In the following sections, the derivations of suitable class boundaries and classification precision are discussed in detail. Both of these aspects of classification should ideally be included in a GIS package.

9.5.2. Practical examples

In order to provide a better insight into the application of the above methods and techniques used in handling uncertainty associated with continuous data inputs contributing to the generation of the FEDP, a typical classified polygon map, i.e., *"field checked"* slope steepness map at scale 1: 25 000 produced by UNPAD (1983), was used as a higher accuracy map, this is *a control map*.

A similar map at scale 1: 50 000 produced by the FEDP Team (1986) has been used as *a test map*. These maps are stored in GIS-ILWIS environment depicting the slope steepness class as also presented in Tables 10 (in Chapter 5) and 24.

Table 24: Slope steepness and interval class

Slope steepness	Interval class
1	0 - 8%
2	8 - 15%
3	15 - 25%
4	25 - 45%
5	> 45%

In this practical example the interval class of 15 - 25% slope steepness from similar points (geographical) locations has been selected. Point observations from both maps are given below.

The higher accuracy source the UNPAD map: *control* map.

Table 25: Observation points of the higher accuracy (*control*) map

X	Y	Point nr.	Value
483.61	213.19	9	16.75
1009.25	318.29	28	17.00
1498.07	319.49	21	14.75
601.28	877.60	13	20.25
1029.16	796.65	19	18.00
1503.28	842.48	20	17.50
532.19	1390.00	24	21.50
1001.25	1425.00	16	21.50
1600.49	1398.69	18	22.50

where X, Y = point coordinates; point nr. = observation point; Value = slope steepness value

Table 26: Observation points of the FEDP (*test*) map

X	Y	Point nr.	Value
485.58	212.00	9	16.25
1001.00	316.00	28	21.00
1498.07	317.89	21	15.25
615.08	873.60	13	21.50
1026.16	794.25	19	17.75
1501.18	840.08	20	16.50
534.19	1379.00	24	24.50
1011.25	1424.00	16	24.75
1598.49	1399.69	18	24.00

where

X, Y = point coordinates; Point nr. = observation point; Value = slope steepness value

Using these data, the computation of measure of dispersion (sdv), the overlap probability (Pov) and the probability of correct ranking (Pvi) as discussed in Subsection 4.4.2 has been performed as described below.

(i) Measure of dispersion (sd)

The standard deviation (sd) of slope steepness maps have been used as a measure of the fuzziness (inexactness) of ranking between 15 - 25%.

* Standard deviation of all observations (sd_{all}) = 1.39

On the basis of overall sd, observation points were grouped into three groups: (i) observations falling outside of the leftside transition zone, which are possibly misclassified with the lower adjacent class; (ii) observations falling outside of the rightside transition zone which are possibly misclassified with the upper adjacent class; (iii) correctly classified observations falling within (i) and (ii). See again Figure 10 in Chapter 4.

Applying equation (7) the standard deviation of slope steepness of these groups of observations has been calculated as follows:

* Standard deviation (sd_1) of observations (i) = 0.39
* Standard deviation (sd_2) of observations (ii) = 0.31

(ii) The probability of correct ranking of sample value (Pov_1) and membership degree of sample value (MFv_1) falling within (around) the dispersion of lower bound:

Observation 21 is around the lower bound. By applying equations (10) and (11) the probability (Pov_1) and the membership degrees (MFv_1) of correct ranking have been calculated as presented in Table 27.

Table 27: The probability and the membership degree of observations falling around the lower boundary of the class

Lower bound of ranked interval	Point nr.	Values	Pov_1 (probability of misclassified into adjacent class)	Mfv_1 (membership degree of belonging to the class)
15 %	21	15.25	0.24	0.88

(iii) The probability of correct ranking of sample value (Pov_2) and membership degree of sample value (MFv_2) falling within (around) the dispersion of higher bound:
Observations 16 and 24 are around the upper bound. By applying equations (13) and (14) the probability (Pov_2) and the membership degrees (MFv_2) of correct ranking have been calculated as presented in Table 28.

Table 28: The probability and the membership degree of observations falling around the upper boundary of the class

Upper bound of ranked interval	Point nr.	Values	Pov_2 (probability of misclassified into adjacent class)	Mfv_2 (membership degree of belonging to the class)
25%	16	24.75	0.73	0.65
	24	24.50	0.20	0.82

(iv) The probability of correct ranking of sample value (Pov_{1_2}) and membership degree of sample value (MF_{1_2}) falling in between the dispersion of the kernel of the class:
Observations 9, 13, 18, 19, 20 and 28 are between the upper and the lower bounds. As stated in equation (12) the probability (Pov_{1_2}) and the membership degree (MFv_{1_2}) are equal to 1.

Table 29: The probability and the membership degree of observations falling within the kernel of the class

Ranked interval	Point nr.	Values	Pov_{1-2}(probability of correctly classified)	Mfv_{1-2}(membership degree of belonging to the class)
15 - 25 %	9	16.50	1	1
	13	21.50	1	1
	18	24.00	1	1
	19	17.75	1	1
	20	16.50	1	1
	28	21.00	1	1

(v) Overlap probability (Pov)

Using the map precision value as obtained from (i) above, the interval perfomance test has been applied to the range 15 - 25%, defined by the FEDP Team. From observations obviously the dispersion limit to the right of the lower bound is less than the dispersion limit to the left of the upper bound. This implies that the range between 15 - 25% is a good ranking with the overlap probability (from eq.(8)) of 0 (zero) within 99.7% significant level. Considering this and (i), we can inform users that the 15 - 25% ranking is a correct ranking (see also Table 29). In other words, the range beween 15 - 25% is neither a too narrow nor a too broad class interval. This implies that the classes coincide with typical slopes found in the terrain, with class boundaries coinciding ideally with breaks of slope.

From Table 29, observation values near the centre of each slope steepness class have a very high probability of being correctly classified. The probability and the membership degree become 1(certainty).

From Tables 27 and 28, for observations falling toward the class boundaries, there is an increasing probability of the observation values being incorrectly classified. The smaller the probability of misclassification the greater is the membership degrees of observations belonging to the class and vice versa. It can be deduced that the greatest probability of incorrect classification occurs for observations falling exactly on class boundaries.

(vi) The overall probability of correct ranking ($P_{\_ALL}$) and membership degree of observation ($MF_{\_ALL}$) associated with TEST map

After obtaining the probability ranking as discussed in (iii), (iv) and (v), a GIS user may want to know the average of the probability of correct ranking. The result is presented below.

* P_{-ALL} of correct ranking = 0.66

* $MFv_{\_ALL}$ of all samples = 0.78

9.6 Conclusions

There is no single data input into GIS that is completely error free. The knowledge on source of data inputs into GIS is very important for estimating the associated errors. Uncertainty aspect of data inputs includes (i) attribute and positional quality of data; (ii) lineage, completeness and logical consistency of data.

The attribute quality of data is related to value, substance and characteristics. Generally, data inputs into GIS can be classified into continuous and non-continuous data types. Attribute accuracy assessment associated with continuous, e.g., slope steepness data type can be represented using the standard deviation from which the probability value and membership degree, used in the representation of data quality, are derived. Attribute accuracy assessment associated with non-continuous data, e.g., landuse type can be handled using external testing and internal testing.

In addition to the results of the attribute accuracy assessment, it becomes apparent that the observational procedure based on TMU does provide both the stability of boundary (Section 6.5) and thematic attributes (Chapter 9). In other words, the TMU is considered as a stable or repeatable observational unit for data acquisition, data capture, producing and presenting information outputs.

Chapter 10

Handling uncertainty-ambiguity associated with the Inductive Erosion Model in Geographic Information Systems

10.1 Introduction

This chapter discusses the application of evidence theory (plausibility reasoning) in a GIS for handling uncertainty-ambiguity associated with experts' inference in predicting the occurrence of soil erosion severity classes in particular areas (Subsection 8.6.2).

Based on how a model used in a GIS is constructed, the proposed IEM as discussed in Chapter 8 can be classified as a logical or inference processing model (Drummond, 1990; Suryana, 1993). By this is meant that this logical model is not based on a mathematical formula and does not admit to the application of error propagation theory. In this regard, the proposed IEM more closely represents the working of a soil conservation expert than does the USLE mathematical model. However, it is different to a logical crop suitability model as explained in Section 4.5 and Subsection 10.2.2.

Datasets on locational erosion class and locational attributes obtained from the study area (Sections 8.3 and 8.4), were used to provide an example of how the uncertainty-ambiguity associated with an IEM can be handled in a GIS environment.

10.2 An approach to the model quality assessment associated with an IEM

10.2.1 General view

A probability values have been used for a long time to represent uncertainty associated with field observations, and interpretation of remotely sensed data. Several programs have successfully modelled the diagnostic process, relying on the use of statistical decisions as found in Bayes' Theorem in manipulating conditional probabilities. It has however been proved by researchers that this approach requires a very large amount of good quality data, numerous approximations and assumptions (Heckerman, 1988). Considering these heavy requirements and to overcome the dificulties associated with deriving probability value an alternative solution to the problem is required.

Inexact reasoning is common in science (Negoita, 1985). Generally it is associated with the art of good guessing, hunching, feeling or good scientific judgement without losing too much of the accuracy (Drummond, 1990; Shortliffe et al., 1984; Bonnet, 1985).

Within the framework of this research an attempt was made to investigate the application of CF in geographical information manipulation particularly in handling uncertainty associated with a typical logical or inference processing model, e.g., an inductive erosion model.

The general understanding of the concept of CF has to be viewed mainly as experts' expression particularly in inferring conclusions from some evidences. These ideas come naturally to experts and can be expressed verbally and handled numerically within a GIS as discussed below.

10.2.2 Plausibility reasoning and an IEM

The process of erosion is a natural phenomenon of which the degree of occurrence is determined by its various causal factors and their relationships. Compared to an original hypothesis concerning the erosion class based on a single erosion influencing factor, each subsequently observed factor can either increase (positive) or reduce (negative) the expected degree of erosion. The likelihood that an area belongs to a particular erosion class therefore changes as the evidence (observations of erosion factors) accumulates.

In plausibility reasoning, the Certainty Factor attached to the expert's expression therefore can be regarded as *a measure of change in belief* after each piece of evidence is collected. Thus, in plausibility reasoning as applied to an IEM, the weights or CFs can be regarded as subjective *"changes"* in degree of belief as new evidence (influencing factor) is gathered. See also Suryana (1993). The change in belief is related to the situation when the experts infer the effect of combined independent evidences including causal relationships. In the latter case confidence in this inference is either increased progressively or shown to be untenable as more evidence is collected.

On the other hand a plausibility measure (degree of belief) is related to the situation when the expert infers the effect of a single independent measurement (evidence), e.g., slope steepness on soil erosion.

It is to be expected, then, that the certainty factor approach is a very useful method for reasoning or managing uncertainty in specific questions which are determined by chances of different possible answers. This is especially true of situations in which probabilistic reasoning, would require large amounts of data and high computional demands. It is also worth noting that because the processes involved may not be physically independent, an IEM model is of a type which does not suit the application of overlay techniques such as Fuzzy Conjunction or Fuzzy Disjunction, which depend on independent measurements not related to phenomenonological processes.

10.2.3 Synthesizing an IEM in the light of plausibility reasoning perspective

Erosion types (sheet or rill erosion) to which an IEM is addressed were affected generally by type and percentage of vegetation cover and slope steepness. As might be expected, the higher the percentage of vegetation cover and the gentler the slope the smaller the risk of erosion and vice versa. As stated earlier in Chapter 8, the relative importance of these two factors in the study area may be deduced from the fact that many units with a relatively high percentage of vegetation cover and steep slopes have a low risk of erosion, showing the great importance of the factor vegetation in the study area.

Established resilience and inertia specific site erosion influencing factors include rainfall erosivity (R-factor = E_1), soil erodibility (ST-factor = E_2), slope steepness (SS-Factor = E_3), slope length (SL-factor = E_4), percentage ground cover and soil conservation practices (VP-factor = E_5). For each of these, a CF is derived for each hypothesis, i.e. for each erosion class. These classes have to be decided in advance, and the number of classes will most likely not exceed five or six.

In the case of soil erosion, the combination of factors is a case of parallel combination. For example, of the factors mentioned in the previous paragraph, some are closely linked, e.g., soil erodibility with percentage ground cover and with soil conservation practices. According to Shortliffe and Buchanan (1984), the Certainty Factor model creators, the relevant formulae for parallel combination are based on the axiom that there is some function g such that:

$$CF(H,E_1E_2) = g(CF(H,E_1),CF(H,E_2)) \qquad \text{......eq.(19)}$$

where: H = original hypothesis
E_1 = first evidence (i.e. influencing factor 1)
E_2 = second evidence (i.e. influencing factor 2)

In other words, two updates for the same hypothesis can be combined to yield a single update using the formulae for parallel combination as developed by Shortliffe et al. (1984).

From equation (19) it becomes apparent that parallel combination as stated above should not depend on the order in which evidence is considered. In other words, the order of combination should have no effect on the final result, and the calculations can be continued as new evidence is collected.

With regard to equation (19), Grosof (1986), Horvitz and Heckerman (1986) have found a gross inconsistency between the definition that CF is belief update (eq.17 in Section 4.5) and the necessity of parallel combination of functions. These studies show that parallel combination is neither commutative nor associative. This inconsistency, for example in erosion study using the proposed IEM, at first appears harmless. But the proposed IEM involves many erosion influencing factors, and *the end result should not depend on the order* in which they are considered or updated.

Horvitz and Heckerman (1986) suggest that it is desirable that certainty factors include fundamental properties of belief update, and properties of parallel and sequential combinations. Considering this and probability interpretation of certainty factors, definitions of CF (H,E_1E_2) were *reformulated* as follows:

$$CF\,(H,E_1E_2) = \frac{CF\,(H,E_1)\ +\ CF\,(H,E_2)}{1 + \{CF\,(H,E_1) * CF\,(H,E_2)\}} \qquad \text{....eq.(20)}$$

remark: $E_1 ....E_n$; *H,* E_1 *and* E_2 *have the same meaning as in equation (19).*

Formula (20) is considered relevant to modelling uncertainty associated with resulting information yielded by the proposed IEM. For example the proposed IEM considers that soil erosion severity class is a linear and associative function of the combination of soil erosion influencing factors. In this regard, the more factors which are involved (unlimited and not dependent on order), the more reliable the IEM model prediction. In addition to this, an IEM takes also the causal relationship (causative affect i.e. confirming and disconfirming factor) of each erosion influencing factor involved into account.

10.2.4 Certainty Factor

The inference rule associated with an IEM can be expressed verbally using *a fuzzy identifier* such as *likely, extremely likely, extremely unlikely* representing the degrees of likelihood and can be represented by Certainty Factors or Certainty Statistics (Bonnet 1985; Negoita, 1985; Drummond, 1989; Klir and Folger, 1988). The term Certainty Factor (CF) was adopted in this study and a method for obtaining CF in GIS environment is outlined as follows.

Erosion severity classes information input to a GIS may originate from a terrain mapping unit map providing terrain characteristics as described in Chapters 6 and 7. However, a soil conservationist specialist may be accessible to the manager of the GIS, who could then ask questions concerning the certainty of the classification . An example is given below.

TMU 115 carries the following environmental properties: slope steepness : 26%; slope length: 20m and ground cover: mixed garden with young cassava. The specialist could be asked to consider each environmental property separately, and for each to state the likelihood of an original allocation to a soil erosion class, assuming the given environmental property was presented as new evidence. In other words, he has to state his change in belief in the given class. Take for example the case of slope steepness 26%, in TMU 115 and the likelihood of the original class being correct after being presented with slope data.

Original class	*Belief in original class after presentation of new evidence*
severely eroded	very likely
moderately eroded	likely
slightly eroded	unlikely
very slightly eroded	very unlikely

A similar list is produced for the other environmental properties.

The soil conservationist has used d*iffererent terms referring to degrees of likelihood*. With the assistance of the soil conservationist specialist these terms should be ranked: any missing terminology which an expert might wish to insert should also be included in the ranking, and synonyms identified. A final ranking of this particular likelihood terminology could then be presented in the left column of Table 30.

The CF used for change in belief may take any value between -1 and +1 (Heckerman et al., 1988; Buchanan, 1988; Bonnet, 1988). In the particular case of an IEM, the expert can then be asked to grade these likelihoods in terms of certainty factor, which can range anywhere between +1 and -1. This may result in the right column of Table 30.

Table 30: Typical occurrence likelihood of erosion and corresponding CFs

Occurrence likelihood class of erosion*)	**Certainty factors (CFs)**)**
Absolutely likely	1
Extremely likely	0.9
Very likely	0.6
Likely	0.3
Neither likely nor unlikely	0
Unlikely	-0.3
Very unlikely	-0.6
Extremely unlikely	-0.9
Absolutely unlikely	-1

note *) Experts' expression of occurrence likelihood of class of erosion

**) Established as translation of experts' expression in allocation of effect of each additional erosion influencing factor on belief in original hypothesis

Table 30 shows that -1 represents definite exclusion, +1 represents definite allocation (complete confidence), while value 0 represents no change in opinion. See Subsection 4.4.2. Therefore the allocation of objects to a certain class is determined finally by the overall CF closest to +1, starting with all possible classes, then applying the CF for each observed or measured erosion influencing factor operating independently.

In this case, the effect of each factor on the erosion class is different. Depending on the erosion class first assumed, the same factor can cause either an increase (CFs range between 0 and +1) or a decrease (CF range between 0 and -1) in belief. Take for example the extreme situation where a terrain mapping unit (TMU) is covered by forest and has a slope steepness between 15% and 35%. In this example forest has a reducing effect (CF is negative) on the process of soil erosion while the effect of slope steepness is to accelerate the process (CF is positive). If for example the original inference was that this TMU has high erosion, then a CF value of -1 can be assigned to forest and a CF value of about +0.6 can be given to slope steepness. High erosion is now excluded as a possibility, despite the steep slope (CF value +0.6), because field observation shows that there is never high erosion in densely forested areas (CF value -1). A more elaborate discussion is given below.

10.2.5 Practical example

The first step in the procedure to assess the ambiguity of an IEM is to analyse observed patterns collected during erosion survey at the FFL. The second step is to tabulate the effects of the erosion influencing factors found by previous erosion studies and published. Some of these are given here in Tables 31 to 34. The knowledge obtained at this stage of the study is used further as basic knowledge for the assignment of CFs to each of the erosion influencing factors as summarised below.

The effects of vegetation cover on erosional processes especially on surface erosion are numerous and varied depending on type of vegetation cover, density, undergrowth cover and litter. These determine the interception loss, absorption of kinetic energy and water infiltration. Land with good cover allows soil retardance to overland flow. The effect of vegetation cover as presented in Table 31; C*P is the combination of crop factor and land management factor.

Slope steepness and slope length have a strong relationship to erosional processes. Therefore, both of them can be used for quantitative evaluation and input in inductive erosion modelling. The values of slope steepness and slope length as present in the study area are presented in Tables 32 and 33.

The soil type (ST) in a specific site is closely related to landform, soil characteristics and their susceptibility to surface erosional processes. For example on volcanic rocks the soil is more resistant to soil erosion than on sedimentary rocks. In other words, the soil erodibility value (K-factor in Table 34) of each soil type indicates the susceptibility of the soil to erosion: it was obtained using the reconsolidation factor of Dissmeyer and Foster (1981). It is calculated according to the soil reconsolidation and residual building value related to time elapsed after a particular piece of land is tilled.

Table 31: Vegetation cover, soil conservation measure (in Java)

Ground cover/soil conservation measure	C * P value#
Undisturbed forest	0.01
with undergrowth	0.03
without undergrowth & mulch	0.50
Undisturbed bush	0.01
mixed with grass	0.10
mixed garden woodlot	0.02
Garden	0.07
Yard	0.20
Estate, fully covered	0.20
partly covered	0.01
Completely grass covered	0.01
mixed with Imperata sp.	0.02
Imperata burned once per year	0.06
mixed with Citronella grass	0.65
Agricultural crops: tubers	0.51
cereals	0.36
beans	0.43
mixed	0.43
irrigated rice	0.02
Upland crop rotated 1 year fallow	0.28
rotated 2 year fallow	0.19
Agricultural soil conservation techniques	
mulching	0.14
bench terrace	0.04
counterridges	0.14

Note that in Table 31 # crop factor (C-factor) and soil conservation measure (P-factor) are treated as dimensionless. The bigger the C*P values the bigger the effect.

Table 32: Slope steepness value

Slope steepness (%)	SS value#
0 - 2	0.1
2 - 8	0.5
8 - 15	1.4
15 - 25	3.1
25 - 40	6.1
40 - 65	11.9

Note:# SS value is treated as dimensionless.
The steeper the slope, the more prone to erosion.

Table 33: Average slope length overland flow

Average slope length overlandflow (m)	L value#
<50	1.5
<75	1.8
<150	2.7
<300	3.7

Note:# Slope length (L-factor) is treated as dimensionless.
The longer the slope, the more prone to erosion

Table 34: Soil type, soil erodibility

Soil type	K value*
Oxic distropept	0.12
Typic haplorthox	0.26
Typic tropodult	0.23
Lytic troporthent/dystropet	0.27
Oxic dystropept	0.16
Tropoquept	0.13
Typic dystropept	0.15
Typic tropoquept	0.13

Note:# Soil erodibility (K-factor) is treated as dimensionless and obtained using nomograph of Weischmeier.

The next stage in the procedure is to assign CFs to each of the erosion influencing factors. Several soil conservation specialists were interviewed, relating to these factors at each observation point (Appendix 5). The average values of the estimated CF they gave, based on their experience, were then calculated. Table 35 gives an example for CFs related to observation point 13, Table 19 Chapter 8, using figures relating to four classes of erosion. This particular area has the vegetation cover mixed garden with young cassava and slope length in the class of 20 m. Both of these factors would lead us to expect moderately eroded. But the soil type is typic dystropept, which has high erodibility, and the slope steepness is in the class greater than 25%, that is steep. Taking first the hypothesis that this area has very slightly eroded, and combining the first two certainty factors according to the equation (20) as given above, we get:

$$\{-0.6 + (0.3)\}/\{1+[(-0.6)*(0.3)]\} = -0.3/0.82 = -0.37$$

Now combining this result with the third certainty factor we get:

$$\{-0.37 - 0.9\}/\{1 + [(-0.37)*(-0.9)]\} = -1.27/1.33 = -0.95$$

Then, combining this result with the fourth certainty factor we get:

$$\{-0.95 + (-0.8)\}/\{1+[(-0.95)*(-0.8)]\} = -1.75/1.76 = -0.99$$

Finally, combining this result with the fifth certainty factor we get:

$$\{(-0.99) + (-0.9)\}/\{1 + [(-0.99)*(-0.9)]\} = -1.89/1.89 = -1$$

This leads us to reject absolutely the hypothesis that the area is very slightly eroded.

Table 35: Example of CFs at observation point 13 (see also Sections 8.3 and 8.4)

Erosion influencing factors	CFs per hypothetical erosion class*)			
	I	II	III	IV
Slope steepness: 26% (very erodible slope)	-0.6 *(very unlikely)*	+0.1 *(neither likely nor unlikely)*	+0.7 *(very likely)*	+0.8 *(extr. likely)*
Slope length: 20m (low erodibility)	+0.3 *(likely)*	-0.3 *(unlikely)*	-0.7 *(very unlikely)*	-0.9 *(extr. unlikely)*
Soil type: Typic distropept, soil erodibility: 1.15 (very erodible)	-0.9 *(extr. unlikely)*	-0.7 *(very unlikely)*	+0.6 *(very likely)*	+0.7 *(very likely)*
Ground cover: Mixed garden with young cassava, C-factor: 0.28 (very erodible)	-0.8 *(extr. unlikely)*	-0.5 *(very unlikely)*	+0.5 *(very likely)*	+0.8 *(extr. likely)*
Rainfall intensity within 30 minutes (EI-30): 9.20 (high erosivity index)	-0.9 *(extr. unlikely)*	-0.7 *(very unlikely)*	+0.6 *(very likely)*	+0.8 *(extr. likely)*

Note:
I = very slightly eroded; II = slightly eroded; III = moderately eroded;IV = severely eroded
*) = Obtained using the procedure described in Subsection 10.2.4.

Repeating the calculations for slightly, moderately and severely eroded, we find overall CFs of -0.98, +0.96 and +0.99 respectively. The last value is very high (close to +1) and leads us to accept the hypothesis of severely eroded. In other words, there is absolute likelihood that at observation point 13 the class is severely eroded. The hypothesis of moderately eroded also has a high overall CF, however. If we find by field observation that the area actually falls into this class, this leads us to question our original estimates of the CFs. By applying an iterative procedure, we adjust the CFs accordingly. Finally, we allocate the sample areas to their correct erosion class on the basis of the model, by choosing the class with the highest CF. It is also possible to display the CFs for each erosion class separately, as an indication of the certainty of the allocated class.

Figure 40 shows erosion classes predicted by the USLE, Figures 41a, 41b,41c, 41d show the CFs related to the severity classes as obtained by an IEM.

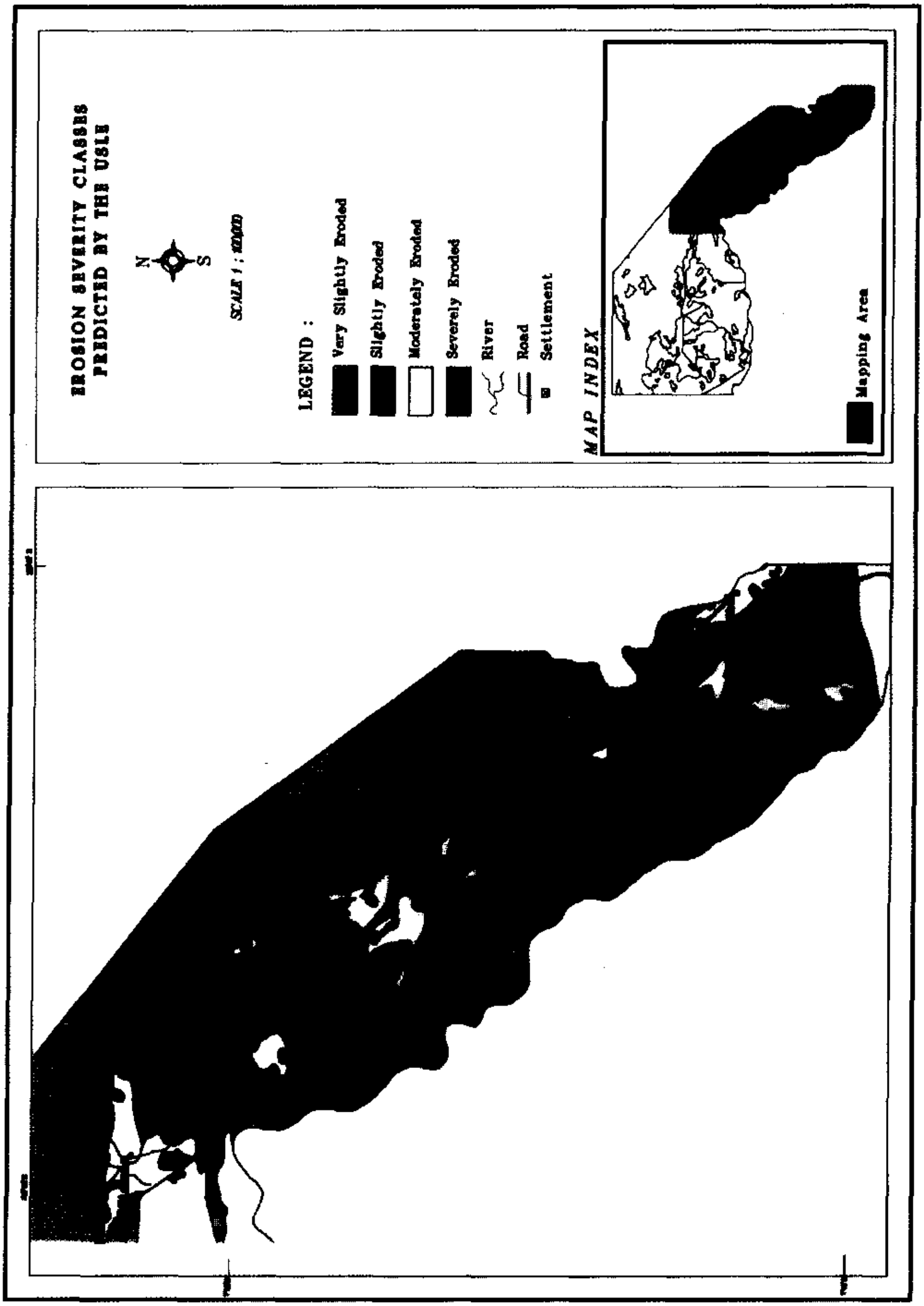


Figure 40: Erosion severity classes predicted by the USLE

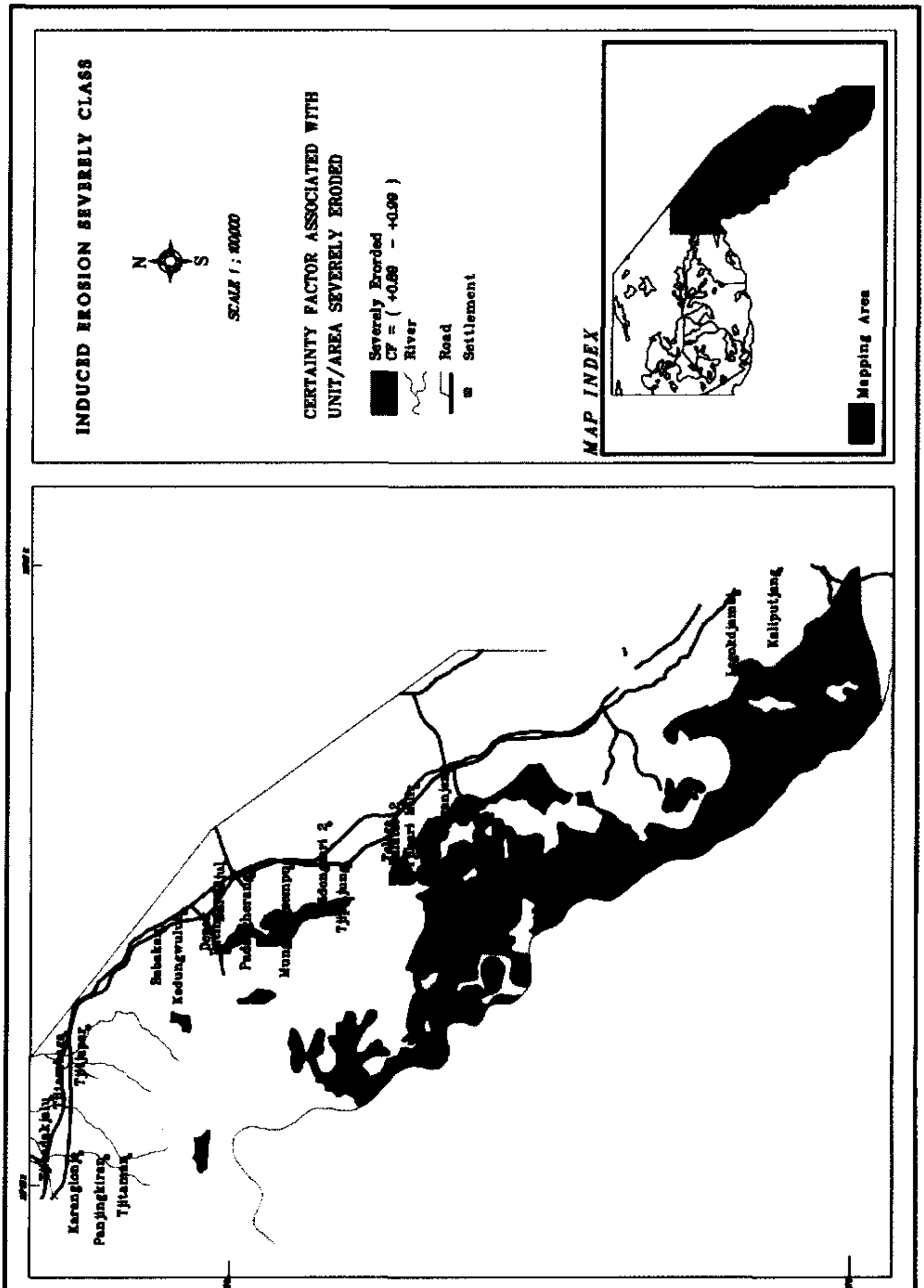


Figure 41 (a): Induced erosion severity class: Severely eroded and CFs

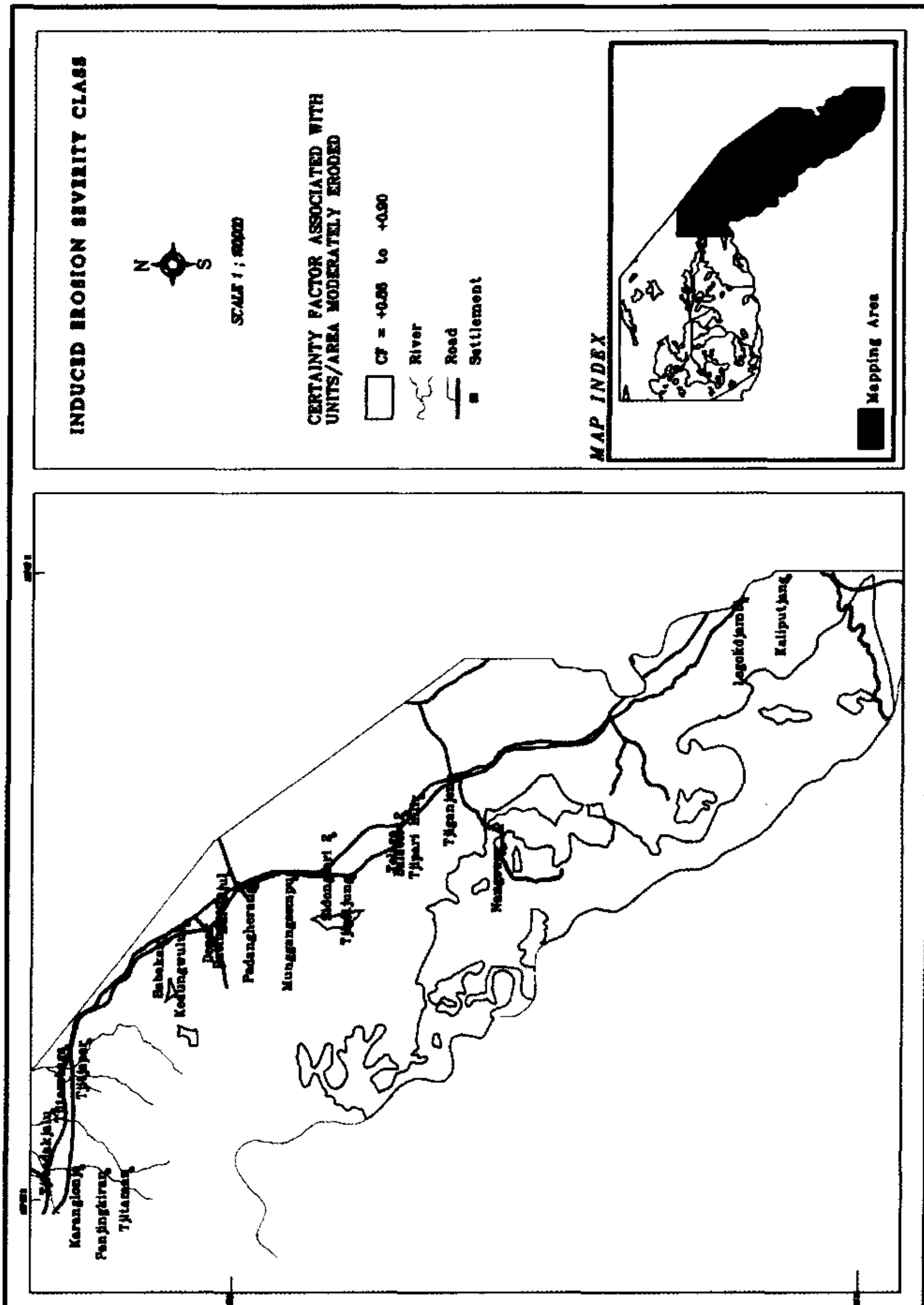


Figure 41 (b): Induced erosion severity class: Moderately eroded and CFs

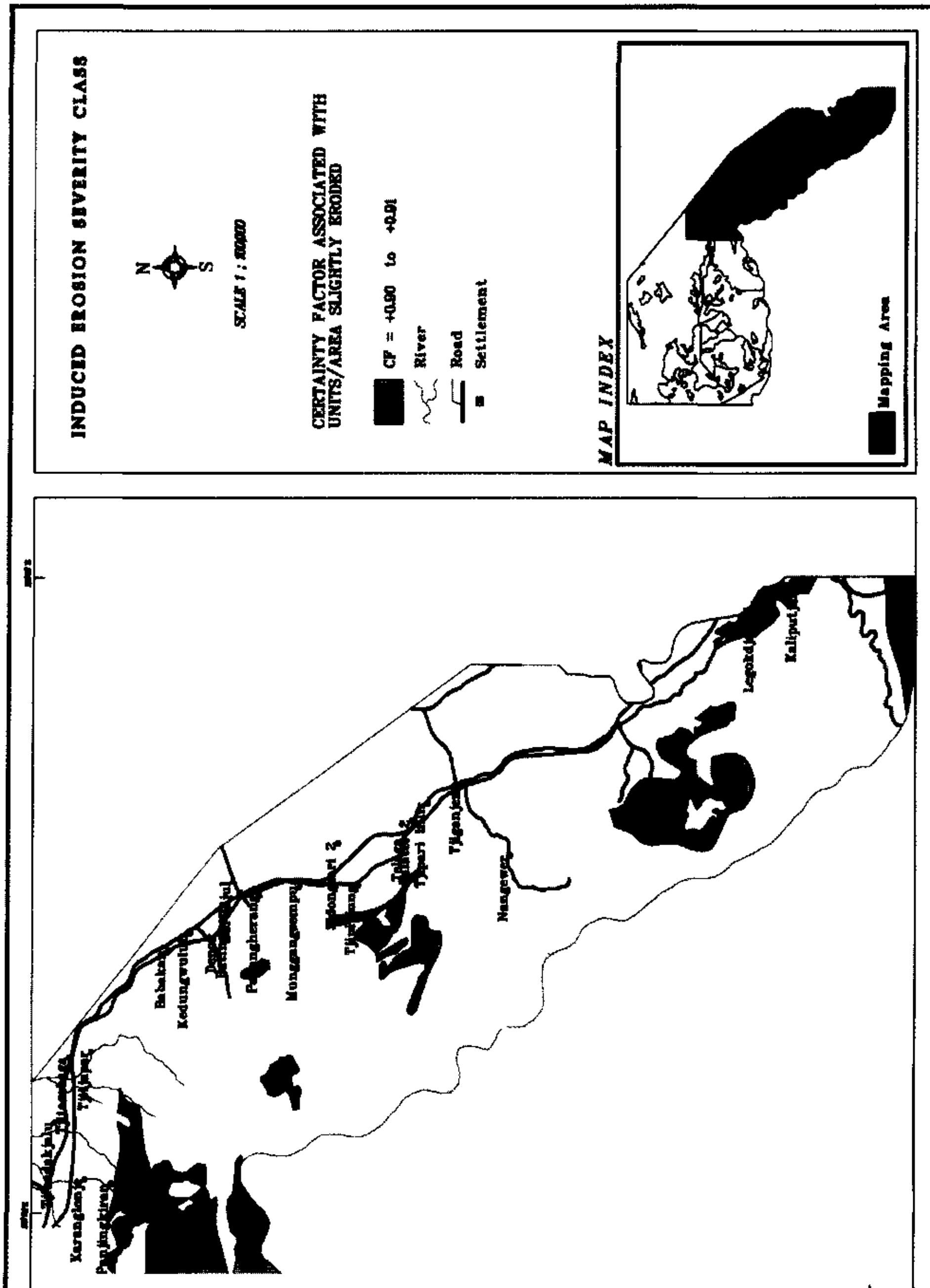


Figure 41 (c): Induced erosion severity class: Slightly eroded and CFs

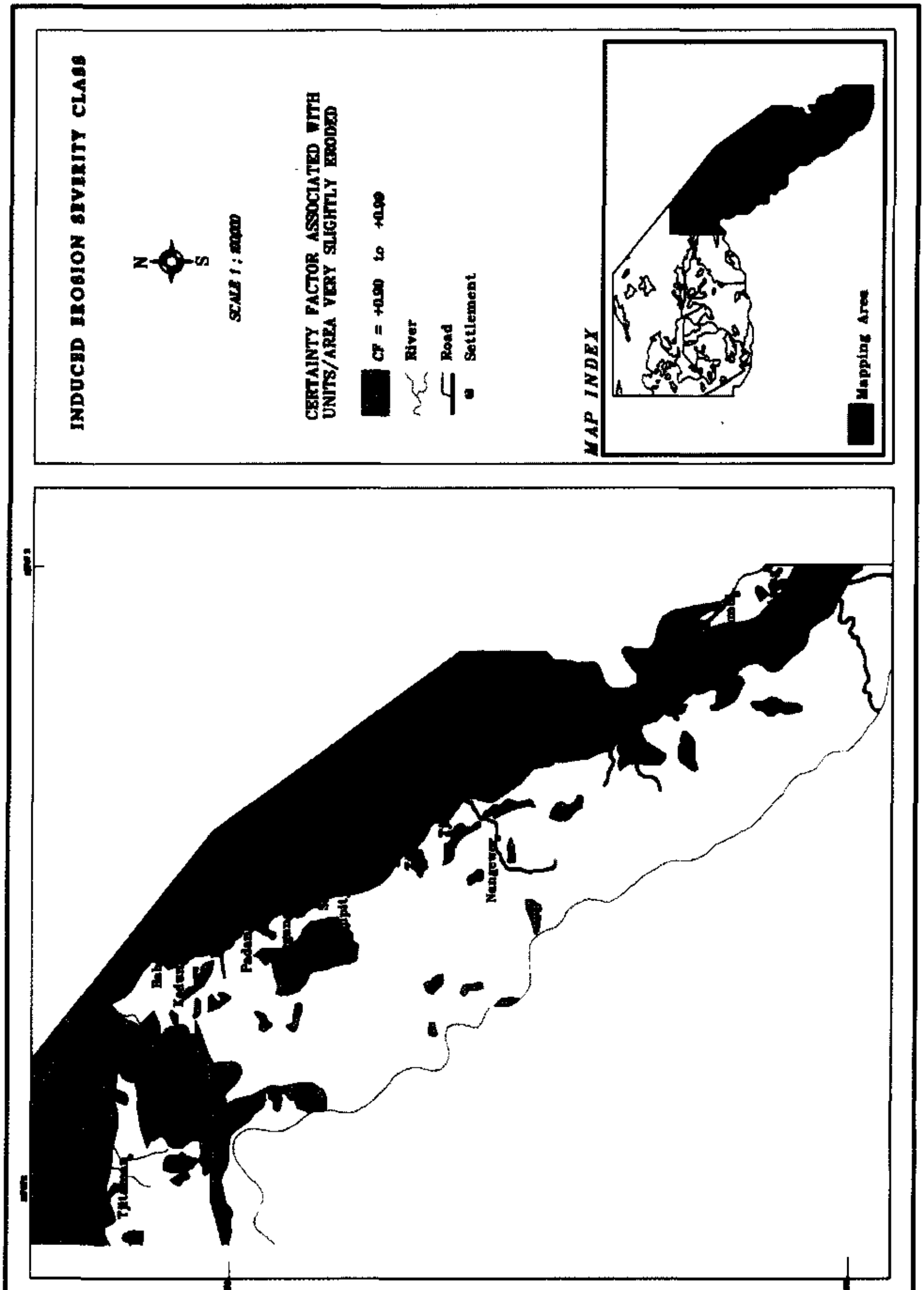


Figure 41 (d): Induced erosion severity class: Very slightly eroded and CFs

10.3 Conclusions

In Chapter 4 it was pictured that a wide range of approaches for handling uncertainty already exists in geographic science. Some have been tested in handling data inputs into a GIS environment in general and into the establishment of the FEDP in particular (as discussed in Chapter 9).

In Chapter 8 it was clearly shown that an IEM is a logical processing (expert's inference) model and not a mathematical processing model. This model was constructed in the form of *If...Then* structure like many other logical models such as the crop suitability model.

Considering the process of soil erosion, which takes the causal relationship between soil erosion and erosion influencing factors into acount, the assessment of an IEM model quality can be handled neither by the error propagation techniques nor by the fuzzy subset theory. This is caused by the characteristic of an IEM which is more concerned with the ambiguity and less concerned with the vagueness (fuzziness) of the underlying concept. In this regard, the fuzzy measures adopted by an IEM are associated with the situation when one has to search for perfect evidence to decide an underlying process, e.g., erosion class, in which full membership in one and only one is allowed. Additionally this chapter also explores the application of plausibility reasoning (evidence theory) in handling uncertainty associated with the typical IEM inference model and represented using certainty factor.

CF in an IEM model does not represent membership degree but represents change in belief in a particular hypothesis on the basis of given evidence. CF adopted in an IEM is established based on subjective judgement by experts, and its value depends on the original hypothesis.

Chapter 11

Conclusions and recommendations

In the Introduction part, three major research objectives were defined. These are restated briefly as follows:

(i) the establishment of functional observation units for observing terrain characteristics from which erosion classes can be inferred at different aggregation levels;

(ii). the building and developing of a rule-based inductive erosion model (IEM) in a GIS environment as an alternative to the USLE;

(iii) the establishment of a framework for handling uncertainty associated with a GIS environment.

The above research objectives have led to the following major conclusions and recommendations:

11.1 Conclusions related to the establishment of the hierarchical and observational procedure using TMU

Chapters 6, 7 and 9 demonstrated the suitability of the terrain mapping unit (TMU) as the basic mapping unit. The contribution of this particular research was the application of the object orientation concept to the traditional TMU mapping system, which produces stable (repeatable) observational units that can be used in data acquisition, data capture and data processing. Using the object orientation concept, TMUs can be used to produce predefined information at different aggregation levels. In this study, the occurrence of soil erosion, which is an active process, was considered to be an attribute of a TMU.

The TMUs were established by means of interpretation of aerial photos at scales 1: 50 000 and 1 : 15 000 and of SPOT images, supported by topographic maps.The esssential concept associated with the establishment of TMUs was the integration of relief, constituent materials and forming process or genesis of landforms. The basic step in the identification of TMUs was the delineation of areas of homogeneous relief, using image interpretation (mono and stereo) and topographic maps to establish these boundaries. The (complex) variation in morphometry, lithology, geology and dominant vegetation cover and land use in the area, with their distinctive patterns in the images, was found to correspond to the TMUs. In other words, each TMU could be considered to be uniform in all these aspects. The same was true of the active process of soil erosion. The soil erosion classes were also identified from the images. Finally, all this work was checked and finalised in the field. It is worth noting here that the interpretation and classification procedures were considered to be independent and separable processes.

A development from the basic concept of the TMU was to consider the TMUs as nested objects which are organised in a hierarchy, in three levels: farmer's field level, slope level and terrain systems. Following the hierarchical line, these levels are related to the subsub division of TMU, subdivision of TMUand subregional TMU. The aggregation and classification hierarchy was based on three diffrerent determinants (see Chapter 7). This implies that a particular spatial unit TMU can be defined on the basis of a homogeneous description at class level. Other unit TMUs have a homogeneous description at superclass level (see Figure 24). Because the boundaries are stable at the class level of the hierarchy (different interpreters found coincident boundaries), they are also stable when TMUs are aggregated at higher levels. Aggregation of TMUs implies that their attributes are also aggregated. In the context of this research, the soil erosion attribute was examined in detail, and the aggregation principle applied successfully to it.

Three determinants have been introduced for the identification of TMUs at three different aggregation levels. The first determinant segemented the study area into TMUs at subregional level. The second and third determinants were used to disaggregated these TMUs into units at the next two lower levels.

The TMU was found to be a very useful carrier of information on terrain characteristics in a GIS environment. These can be used for example as inputs in modelling soil erosion at different class and aggregation levels. The repeatability or stability of the boundaries as well as the thematic attributes is required for evaluating and monitoring changes in the static and dynamic aspects of terrain objects. The first objective of the research, the establishment of functional observation units for describing soil erosion at different levels of the classification and aggregation hierarchy, was therefore achieved.

11.2 Recommendations relating to the first research objective

The approach developed in this thesis, where TMUs are defined at three aggregated levels allows the user to realise terrain description at different levels of generalisation. This allows users to select at each level those details of the terrain system which are needed for a particular application, e.g., erosion study and to ignore other details which are still kept in the data base. This is a means for managing complexity of the terrain system without losing too much information. From the result it was clear that the establishment of a hierarchical structure sustains all relevant details in a controlled way. Two kinds of abstraction were adopted that are fundamentally important in establishing the hierarchical structure of functional TMUs as well as in the mapping strategy adopted by an IEM. These are the object aggregation and class generalisation concepts (Chapters 7 and 8). It was found that object aggregation and class generalisation work semi independently. In other words, class generalisation and object aggregation are two different steps but not necessarily independent when used in a generalisation processes. By adopting this concept mixture a rich variety of terrain systems can be clearly defined.

The integration of terrain description at small scale and large scale was established by compiling and managing data from the lowest to the highest aggregation level. In other words, objects at one level are the details for those at related (next) higher levels. Therefore, TMU at class level was considered as an obvious level of integration at sub-regional scale. However, an attempt sould be made to develop or improve methods for describing the hierarchical procedures in which levels of abstraction are related in a functional and quantitative manner. This aspect of the study should continue and be a high priority for future research. Additionally, the establishment of database abstraction which can support step wise aggregation and generalisation rules in wider applications has still to be investigated.

11.3 Conclusions related to building and developing the proposed IEM in a GIS environment

The research was intended to demonstrate the possibility of incorporating the proposed IEM into a GIS as an acceptable means of estimating soil erosion in a large catchment in a fast, easy and reliable way.

The process of soil erosion can be considered as a phenomenological process. From erosion study at farmer's field level (FFL), we concluded that the degree of soil erosion on a specific site was the result of locational factors. This led to the development of an inductive approach in handling local variability of soil erosion. It was found that erosion from field observations and predicted by the USLE gave comparable results in indicating relative soil loss value. Thus, the results of *onsite* erosion estimation can be used as a substitute of the USLE. The differences in the estimation are caused by the rainfall erosivity values used and the growth characteristics of certain plants measured.

Using correlation analysis as mentioned in Chapter 8, it was found that the effect of the influencing factors involved in soil erosion were different. One factor may accelerate the process of soil erosion leading to more severe erosion but other factors may decelerate it. Soil erosion predicted by the proposed IEM is measured and expressed in terms of classes. From the result of erosion studies as discussed in Chapter 8, an IEM was proposed as a bottom-up erosion model adopting a reasoning method based on forward chaining. As concluded earlier, each point observation made at farmer's field level and located within a TMU provided information on observed patterns, i.e., the relationship between erosion influencing factors and possible erosion classes.

The concept of induction was incorporated in the form of GIS inference capability which increases the added value of a GIS in handling a particular task. Decision rules of an IEM represent a *condition - conclusion couples*, meaning that whenever a certain condition is encountered the associated conclusion is drawn.

An IEM has therefore three components:

(i) the rule base consisting of the set of production rules associated with a particular erosion class;
(ii) data on the known facts as observed and classified at FFL; and
(iii) the interpreter of these facts which decides which rule to apply and therefore initiates the corresponding conclusion.

The advantages of the decision rules method adopted by the proposed IEM can be explained as follows. Because of its considerable degree of modularity an IEM can easily be modified and updated and as it evolves, the established rules are not affected by the addition, deletion or modification of other rules.

The achievement of the hierarchical observational procedure (first objective) supported the implementation of the inductive concept in modelling soil erosion at different aggregation levels. In other words, the information on the spatial distribution of soil erosion obtained from aerial photo and SPOT image interpretation provided very useful inputs for building the proposed IEM.

Aggregation levels are related to observations made at different scales. The proposed IEM took three different, but interlinked levels of functional units of observation, namely point observations at farmer's field level (FFL), homogeneous landforms and larger parts of the terrain system.

Point observations for the proposed IEM are conducted at FFL, which was found to be the most appropriate basic or elementary functional unit to describe erosion over a larger area. When tested, an IEM was found to predict erosion in the study area much better (91% reliability) than the USLE (54% reliability), which is the method used in the Field Engineering Design Plan (FEDP) in Indonesia. We can therefore conclude that the second objective of the research was achieved.

11.4 Recommendations relating to the second research objective

Erosion survey at FFL provides better insight into the relationship between erosion influencing factors and erosion class. To avoid overestimation of soil erosion, for further study particularly in performing erosion classification care must be taken in using the growth characteristics of plants, finding a proper root exposure measurement and selecting appropriate soil erosion indicators. In addition to soil erosion measurement using root exposure it is recommended that this be supported by plot studies at the same location.

Erosion severity class is one of the active attributes associated with a particular TMU. In this research, the occurrence of soil erosion at different aggregation levels was properly modelled and predicted by adopting the concept of induction. This was established using and following the inheritance line of TMU information.

From the results of the research, it becomes clear that the utility of a simplified description of the terrain system TMU and erosional process depends on how well the linkages between terrain components are described. Thus, the research established the contiguity relationship which was found to be highly dependent on the spatial distribution of topographic parameters at different scales which influence the occurrence of soil erosion. In this regard, the integration of small scale observation with large scale units was based on soil erosion and terrain component relationships. Furthermore, vegetation cover, soil texture and soil structure were identified as dynamic aspects of terrain characteristics.The soil conservation measures provided to farmers were found as a function of degraded soil physics characteristics, e.g, changes in soil structure and reducing water holding capacity, which in turn were identified as the major effect of soil erosion on soil and plant relationship. From this analysis arose a new research question: how can we approach these changes of terrain characteristics at different scale of observations, from which appropriate soil conservation measures can be formulated.

The concept of contiguity and the bottom-up approach adopted by the proposed IEM and implemented in a GIS environment is considered as a workable means for predicting or monitoring these active processes and their effect at different level of the aggregation hierarchy. However, this remains to be investigated in more detail.

It was explained in Chapter 8 that the proposed IEM used continuous and non-continuous data inputs. An IEM provides better prediction compared to the model curently used in the FEDP, i.e., the USLE. It was found that vegetation cover is the most influental factor for the process of erosion. It is clearly explained that the incorporation of a wider range of terrain characteristics data into the data base and modelling will improve the decision rules. However, the current research did not identify those attribute errors to which the process is most sensitive. Based on these considerations, the establishment of a sensitivity analysis which can fit to an IEM characteristics is required.

11.5 Conclusions related to handling uncertainty in a GIS environment

With regard to the erosion study, the introduction of GIS into the FEDP offered some advantages. A computerised GIS system allows one to conduct erosion assessment in order to produce timely and newly predefined erosion severity information used at watershed and subwatershed levels of the planning. In the context of the FEDP we provide soil conservation options to farmers. In this regard, uncertainty matters in relation to the consequences of making any wrong decision on the basis of wrong information.

Chapter 9 demonstrated that there was no single data input into GIS in general and into the FEDP in particular that was completely error free. From this exercise, knowledge on the source of data inputs into a GIS is very important for estimating the associated errors. The uncertainty aspects of data inputs including attribute and positional quality, lineage, completeness and logical consistency were identified. All these aspects need to be incorporated into a prototype uncertainty subsystem, as illustrated in Appendix 9. In this research it was identified that uncertainy associated with information produced by an IEM depends on the rules rather than the data.

The attribute quality of data is related to value, substance and characteristics. In this research data inputs into a GIS can be classified into continuous and non-continuous data types. It was found that attribute accuracy assessment associated with continuous, e.g., slope steepness data type, can be sufficiently represented using the standard deviation (*sd*) from which the probability value and membership degree, used in the representation of data quality, were derived. This research differentiated membership degree from probability as explained further below.

The procedure associated with the establishment of a slope steepness map was found to include the following:(i) probability of misclassification particularly for observations falling near slope steepness class boundaries; (ii) membership function to represent the membership degree of observations belonging to a particular description of slope steepness classes, e.g., nearly level and gently undulating includes the fuzziness or gradual transition of their boundaries which cannot be sharply defined; (iii) probability of correct ranking to represent the fact that the number of classes and slope steepness class boundaries may be incorrectly chosen.

As a rule it can be stated that control point values near the centre of each slope steepness class have a very high probability of being correctly classified. If the classes are sufficiently broad, this probability becomes 1(*one*), i.e., certainty. Toward the class boundaries, there was found to be an increasing probability and decreasing membership degree of the control point values being incorrectly classified and correctly classified respectively. The greatest probability of incorrect (confusion) classification occurred for values falling exactly on class boundaries. Chapter 9 discussed attribute accuracy assessment associated with non-continuous data, e.g., landuse type. This was handled sufficiently using external testing and internal testing.

Some decision rules of an IEM allowed weight values to be introduced. Such weights are often mentioned as plausibility descriptions. In this connection, this research adopted the terminology of Certainty Factor (CF). This proved to have no specific statistical meaning. Also, the plausibilities that are used are often not all objective, but represent, e.g., the experience of an expert in the field who was preprared to assign weights to rare events although he had no statistical basis for doing so. However, it should be remarked that the plausibility reasoning adopted by an IEM differentiates facts from data. Facts are a set of permanent knowledge or information which are incorporated into the program, while the data are related to a particular problem.

Therefore, the CF model adopted by an IEM is not concerned with the propagation of error associated with data inputs (see eqs. 19 and 20, in Subsection 10.2.3). This was caused by the characteristic of the plausibility reasoning which is more concerned with the uncertain information (e.g. inference rules established by experts) and is not concerned with uncertainty associated with the data. In this regard, the plausibility reasoning adopted by an IEM refers to the situation when one has to select perfect evidence to support a hypothesis , e.g., erosion class, in which full certainty in one and only one hypothesis is allowed. Thus, the overall CF in an IEM model does not represent membership degree but represents change in belief in a particular hypothesis on the basis of given evidence.

Chapter 10 explored the application of plausibility reasoning (evidence theory) as an acceptable means for handling uncertainty associated with the typical IEM inference model and represented using CF.

11.6 Recommendations relating to the third research objective

As a *"general conclusion"*, the application of a GIS, e.g., in erosion study, is determined or generated by the models used and data inputs gathered from observation units at different levels. Thus, the quality of resulting information is determined by the availability of good quality data, the reliability of the model used and a method to represent the information categories. However, the case of an inductive erosion modelling based on TMUs is different. Chapter 6, Subsection 6.5 and Chapter 9 discussed the repeatability or stability of TMUs boundaries and the thematic attributes. This shows the stability of both geometrical and thematical aspects of TMUs. Considering this, data quality associated with data inputs (see Chapter 9) does not effect the final result of an IEM model prediction (see again Chapter 8, Subsection 8.5). In this research, it would appear that the uncertainty may be caused mainly by inference rules associated with an IEM. In other words, an IEM does not propagate the individual uncertainty of data inputs. To ascertain this, a method for evaluating the effect of data quality into an IEM model prediction is required. Additionally, consideration should be given to improving the display of information quality.

The standard deviation associated with the contributing continuous variables was treated as a characteristic of normally distributed error. It was used as a basis for deriving probability and membership degree of observations. An attempt should also be made to develop methods for parameterisation of different fuzzy models, as well as exploring a method to combine multiple probability and fuzzy models.

Related to the selection of a suitable recommended soil conservation option in the FEDP, for example TMU 115 in Figure 41 was identified as being severely eroded with the CF value of 0.99. This CF value can be interpreted as a risk in making a wrong decision on the basis of incomplete information. In conjunction with the computerised FEDP, the establishment of a soil conservation decision support system with special emphasis on risk analysis is highly recommended as a topic for future research.

11.7 Final conclusions

(1). Terrain Mapping Units (TMUs) as stable observational units were identified, delineated and differentiated on the basis of homogeneous morphometry or relief of the terrain. The nature of TMUs allows the establishment of hierarchical order. In this regard, TMUs were identified as good carriers of observational terrain characteristics at different hierarchical levels and used as inputs to an IEM.

(2). Compared to the top down, traditional erosion model, the Universal Soil Loss Equation (USLE), the proposed IEM is a "*region specific model*" which provides a better prediction of the spatial distribution of soil erosion. It was built on the basis of a set of decision rules established by soil conservation specialists. Each decision rule represents the relationship between an erosion severity class and its influencing factors. This construction allows an IEM to be easily modified and updated.

(3). No single data inputs either of the continuous or the discontinuous data types are error free. In this research, six uncertainty aspects associated with data inputs were identified. Special emphasis was laid on handling uncertainty associated with the quality of attribute of continuous and discontinuous data.

(4). With special reference to the application of data poor environment, the plausibility reasoning was found to be an acceptable means for handling uncertainty associated with information yielded by an IEM.

11.8 Final recommendations:

(1). For wider application in GIS environment, there is an urgent need to investigate and explore the establishment of data bases which can support stepwise generalisation and aggregation of terrain objects. In addition to the establishment of TMUs, there is also a need to find an improved method for describing the hierarchical procedure.

(2). In addition to IEM model building and different data inputs into IEM, there is a requirement to investigate a suitable means for conducting sensitivity analysis of IEM.

(3). There is also need to explore an inductive approach in modelling dynamic aspects of the terrain characteristics.

(4). Considering the probability and membership degree of observations, there is a need to develop methods for parameterisation and there is a need to explore methods to combine multiple probability and fuzzy models.

(5). Regarding the fifth conclusion, there is a need to explore a method to compare the overall quality information yielded by plausibility reasoning with other types of quality information.

Appendices

Appendix 1: Mathematical equations associated with empirical and process-based erosion model

The equation of the rainstorm parameter

$$R = EI_{30} \qquad eq....(21)$$

where E(total storm energy) is calculated by:

$$E = a + b\log_{10}I \qquad eq....(22)$$

where:

E	: total storm energy	[MJ ha^{-1}]
I	: rainfall intensity	[mm hr^{-1}]
I_{30}	: the maximum 30-min intensity	[mm hr^{-1}]
a	: model parameter	[-]
b	: model parameter	[-]

GAMES is a modification to estimate erosion over an average season. The model provides estimates of soil loss by water erosion and subsequent delivery of sediment yield from agricultural catchments. The USLE equation is converted into the following:

$$A_s = R_s * K_s * L_s * S * C_s * P_s \qquad eq....(23)$$

where:

A_s	: computed soil loss per unit area for the selected season	[tons ha^{-1} yr^{-1}]
R_s	: seasonal erosivity factor	[MJ mm ha^{-1} h^{-1}]
K_s	: seasonal soil erodibility factor	[tons MJ^{-1} mm^{-1}]
L_s	: slope length factor	[-]
S	: slope gradient factor	[-]
C_s	: seasonal land use or management factor	[-]
P_s	: seasonal supporting practice factor	[-]

The percentage of the potential field soil loss that is delivered to the main stream channel and out of the watershed is determined from the following delivery ratio expression:

$$DR_s = \alpha(\frac{1}{n_s} * S^{1/2} * Hc_s * \frac{1}{L_s})^{\beta} \qquad eq...(24)$$

where:

DR_s	: seasonal delivery ratio between two selected points, and $0 \leq DR \leq 1$	[-]
n_s	: seasonal surface roughness	[-]
S	: slope steepness factor	[-]
Hc_s	: seasonal hydrologic coefficient, an index of the amount of overland flow	[-]
L_s	: seasonal length of the overland flow path factor	[-]
α, ß	: calibrated parameters (in the original experiments ranging from 10.1 to 13.0 (α) and from 0.55 to 0.98 (ß))	[-]

The spatial component of GAMES is added for the event that runoff travels across downslope fields prior to entering the stream. The characteristics of different land cells are incorporated in the equation:

$$DR_s = \alpha[1/\Sigma^{m}_{j=1}(n_{sj}\frac{1}{S_j^{1/2}} * \frac{1}{Hc_{sj}} * L_{sj})^{\beta} \qquad eq..(25)$$

where:

m	: number of downslope fields	[-]
n_{sj}	: seasonal roughness of the j*th* field	[-]
S_j	: slope of the j*th* field	[-]
Hc_{sj}	: seasonal hydrological condition of the j*th* field	[-]
L_{sj}	: seasonal length of the j*th* field cell	[-]

The main equation governing both the overland flow and channel elements is the steady-state continuity equation for sediment transport:

$$dG/dx = D_f + D_s \qquad eq...(26)$$

where:

G	: sediment load	[$kgm^{-1}s^{-1}$]
x	: distance	[m]
D_f	: detachment or deposition rate	[$kg\ m^{-2}\ s^{-1}$]
D_s	: delivery rate of sediment to flow from lateral areas	[$kg\ m^{-2}\ s^{-1}$]

The WEPP erosion model computes estimates of net detachment and deposition using a steady state sediment continuity equation:

$$dG/dx = D_f + D_i \qquad eq...(27)$$

where:

x	: distance downslope	[m]
G	: sediment load	[kg $s^{-1}m^{-1}$]
D_i	: interrill erosion rate	[kg $s^{-1}m^{-2}$]
D_f	: rill erosion rate	[kg $s^{-1}m^{-2}$]

The net soil detachment in rills, i.e. rill erosion rate, is calculated for the case when hydraulic shear stress exceeds the critical shear stress of the soil and when sediment load is less than sediment transport capacity:

$$D_f = D_c[1 - G/T_c] \qquad eq...(28)$$

where:

D_f	: rill erosion rate	[kg $s^{-1}m^{-2}$]
D_c	: detachment capacity by flow	[kg $s^{-1}m^{-2}$]
T_c	: sediment transport capacity in the rill	[kg $s^{-1}m^{-1}$]

The detachment capacity D_c is expressed as follows, but only when the hydraulic shear stress exceeds critical shear stress for the soil:

$$D_c = K_r(\tau_f - \tau_c) \qquad \tau_f > \tau_c \qquad eq...(29)$$

where:

D_c	: detachment capacity by flow	[kg $s^{-1}m^{-2}$]
K_r	: rill soil erodibility parameter	[s m^{-1}]
τ_f	: flow shear stress acting on the soil	[kg $s^{-2}m^{-1}$]
τ_c	: rill detachment threshold parameter, or critical shear stress, of the soil	[kg $s^{-2}m^{-1}$]

When sediment load G is greater than sediment transport capacity T_c net deposition is calculated:

$$D_f = [V_f/q][T_c - G] \qquad eq...(30)$$

where:

D_f	: rill erosion rate	[kg $s^{-1}m^{-2}$]
V_f	: effective fall velocity for the sediment	[m s^{-1}]
q	: flow discharge per unit width	[m^2s^{-1}]
T_c	: sediment transport capacity	[kg $s^{-1}m^{-1}$]
G	: sediment load	[kg $s^{-1}m^{-1}$]

Three hydrological variables are needed in the WEPP model, peak runoff P_r, effective runoff duration t_r and effective rainfall intensity I_e. P_r was assigned the value equal to that of the peak runoff. t_r is calculated as:

$$t_r = V_t/P_r \qquad eq...(31)$$

where:

V_t	: total runoff volume	[m s^{-1}]
P_r	: peak runoff rate	[m s^{-1}]

Effective rainfall intensity I_e which is used to estimate interrill soil loss, is calculated with the following equation:

$$I_e = [(\int I^2 dt)/t_e]^{1/2} \qquad eq...(32)$$

where:

I	: rainfall intensity	[mm h^{-1}]
t	: time	[s]
t_e	: total time during which the rainfall rate exceeds infiltration rate	[s]

The processes involved in rainfall detachment rate represented here are defined by the following equations:

$$e_i = aC_e P^p/I \qquad eq...(33)$$

where:

e	: rainfall detachment rate	[kg m^{-2} h^{-1}]
a	: measure of detachability of soil by rainfall	[kg m^{-2} mm^{-1}]
C_e	: fraction of soil not protected by direct cover	[-]
P	: rainfall rate	[mm h^{-1}]
p	: model parameter (in the range from 1.3 to 2)	[-]

$$d_i = v_i c_i \qquad eq...(34)$$

where:

d_i	: deposition rate of size range class i	[kg m^{-2} s^{-1}]
v_i	: settling velocity of sedimentary units of size range class i	[m s^{-1}]
c_i	: sediment concentration in size range class i	[kg m^{-3}]

and

$$r_i = (\rho g S K R_l / I)(\gamma_i - \frac{v_i}{R_l}\frac{x_*}{x}) + \frac{\delta}{\delta t}(Dc_i) \qquad eq....(35)$$

where:

r_i	: rate of sediment entrainment of size range class i	
ρ	: density of water	[kg m^{-3}]
g	: acceleration due to gravity	[m s^{-2}]
S	: slope of the plane (sine of the slope angle)	[-]
K	: efficiency of bed load transport (η) corrected for density of sediment in relation to water (0.276η)	[-]
R_l	: runoff rate per unit plane area	[m s^{-1}]
I	: number of sediment size ranges	[-]
γ_i	: $(1 + v_i/R_l)$	
v_i	: settling velocity of sedimentary units of size range class i	[m s^{-1}]
x	: distance downslope from the top of the plane	[m]
x_*	: value of x beyond which $r>0$	[m]
D	: approximation to depth of overland flow	[m]
c_i	: sediment concentration in size range class i	[kg m^{-3}]

The governing equation for the erosion component of the model is developed by Foster and Meyer (1972, in Beasley et al., 1980):

$$\frac{\delta G_F}{\delta x} = R_{DT} + D_F \qquad eq....(36)$$

where:

G_F	: rate of sediment movement in flow	[kg m^{-1} s^{-1}]
x	: distance along flow surface	[m]
R_{DT}	: rainfall detachment rate	[kg m^{-2} s^{-1}]
D_F	: flow detachment rate	[kg m^{-2} s^{-1}]

The detachment of soil particles by raindrop impact is calculated using the relationship described by Meyer and Wischmeier (1969):

$$D_R = 0.027CKA_I I^2 \qquad eq....(37)$$

where:

D_R	: rainfall impact detachment rate	[kg min^{-1}]
C	: cropping and management factor (from USLE)	[-]
K	: soil erodibility factor (from USLE)	[tons MJ^{-1} mm^{-1}]
A_I	: area increment	[m^2]
I	: rainfall intensity	[mm min^{-1}]

The detachment of soil particles by overland flow is calculated by the following equation (Foster, 1976):

$$D_F = 0.018CKA_I SQ \qquad eq....(38)$$

where:

D_F	: overland flow detachment rate	[kg min^{-1}]
C	: cropping and management factor (from USLE)	[-]
K	: soil erodibility factor (from USLE)	[tons MJ^{-1}]
A_I	: area increment	[m^2]
I	: rainfall intensity	[mm min^{-1}]
S	: slope steepness factor	[-]
Q	: flow rate per unit width	[m^2 min^{-1}]

The transport capacity is calculated by either two of the equations, depending on the flow rate:

$$T = 146(SQ^{1/2}) \quad (for\ Q \leq 0.046\ m^2/\text{min}) \qquad eq....(39)$$

$$T = 14600(SQ^2) \quad (for\ Q > 0.046\ m^2/\text{min}) \qquad eq....(40)$$

where:

T	: transport capacity of flow	[kg $min^{-1}m^{-1}$]
Q	: flow rate per unit width	[m^2 min^{-1}]
S	: slope steepness factor	[-]

Appendix 2: Land use classification as used by RMI

LEGEND INTEGRATED LAND USE CLASSIFICATION
LANDSAT SATELLITE THEMATICAL MAPPING

SYSTEM	SUB-SYSTEM	MAPPING SYMBOL	PRESENT LAND USE	DESCRIPTION OF CROPPING PATTERN, VEGETATION TYPE OR LAND UTILIZATION
CULTIVATED FIELD CROPS	SAWAH PADI	S2 Pl	Sawah rice 2 x Palawija	Two crops of rice followed by one crop of palawija a year
		S3	Sawah 3 x	Three crops of rice a year
		S2	Sawah rice 2 x	Two crops of rice
		S1 Pl	Sawah rice x Palawija	One crop of rice followed by one crop of palawija a year
	SAWAH, TAMBAK OR KOLAM IKAN	S2I1	Sawah rice 2x Ikan(fish)	Two crops of rice followed by one crop of fish (ikan tambak) a year
		S2I2	Sawah rice 2x Ikan(fish)	One crop of rice followed by two crops of fish (ikan tambak) a year
		I2/I3	Ikan (fish) x 2/3	Ikan tambak (fish), two or three harvests a year, or kolam ikan.
		T	Tegal dryland crops terraced	Cassava, paddy gogo, corn,soybean, cowpeas or other root or fiber crops planted in open fields. May include some Alang-alang.
CULTIVATED TREE CROPS	ESTATE, SMALL-HOLDER UP-LAND AND ORCHARD CROPS	Kc	Kebun campuran Mixed orchard usually terraced-simple terraces	Predominantly tree crops and agro-forestry including coconut, banana, cloves, mango, coffee or other fruit trees. Often some perennial crops included,some bamboo near streams or some house lots.
		Kk	Kebun kelapa Coconut orchard	Predominantly coconut trees. Usually orchard terraces constructed.
		Kn	Kebun cengkeh Clove orchard	Predominantly cloves. Usually orchard terraces constructed
		Kr	Kebun karet Rubber orchard	Predominantly rubber. Usually orchard terraces constructed.
		Ko	Kebun cokelat Cocoa orchard	Predominantly cocoa. Usually orchard terraces constructed.
CULTIVATED FORESTRY	PRODUCTION FOREST	Hj	Hutan jati Teak forest	Managed production teak forest
		Hp	Hutan pinus Pine forest	Managed production pine forest
NON-CULTIVATED LAND	PROTECTION FOREST	B	Belukar - Bush	Bush, shrub or thicket formation
		Fm	Mangrove forest	Coastal mangrove forest, some shrimp,fish or charcoal production
		Hn	Hutan lindung Protection forest	Natural primary forest
		Hs	Hutan skunder Secondary forest	Natural secondary forest. Partially logged
	GRASS-LAND AND WATER	G	Grass land	Grass land usually Alang-alang
		R	Rawa - Swamp	Lowlying areas inundated or with water table at or near the surface for most of the year. Grassland or aquantic vegetation is common. May be bunded with some local variety rice grown in the dry season in places.
		A	Salt evaporation ponds	Salt production

NON-CULTIVATE	GRASS-LAND AND WATER	G	Grass land	Grass land usually Alang-álang
		R	Rawa - Swamp	Lowlying areas inundated or with water table at or near the surface for most of the year. Grassland or aquantic vegetation is common. May be bunded with some local variety rice grown in the dry season in places.
		A	Salt evaporation ponds	Salt production
QUALIFYING SYMBOLS		(w)	Water accumulation or flooding	Areas subjected to regular annual flooding. Usually with local rice variety or flood damage to conventional varieties.
		x	Significant proportion of land may suffer from erosion hazard.	Significant areas shown on LANDSAT imagery as without vegetative cover or with signs of possible erosion problem.

Example of Mapping Unit:

8 2
S2P1 / S3

Sawah field producing 2 crops of rice and 1 crop of palawija annually on 80% of the area, remaining 20% produces three crops of rice.

Appendix 3: Landsat image-based terrain classification as used in Indonesia

LEGEND FOR LANDSAT SATELLITE BASED TERRAIN CLASSIFICATION

SYSTEM AND MAPPING SYMBOL			SUBSYSTEM	MAPPING SYMBOLS	MAJOR LAND FACETS (LANDFORMS) INCLUDED	DOMINANT SLOPE %
ALLUVIUM	RECENT ALLUVIUM A1	(A1)	River Alluvium	A1p	Low Alluvial Cover Plain	0-1
				A1m	Meander Complex	0-2
				A1mt	Meander Terrace	0-2
				A1mtt	Low Meander Terrace	0-2
			Flood Plain	A1l	Levee	0-2
				A1e	Recent River Bank Deposits	0-2
				A1d	Local Minor Depression	0-2
				A1b	Backswamp Shallow	0-2
				A1b1	Backswamp Moderately Deep	0-2
				A1b2	Backswamp Deep	0-2
				A1b3	Backswamp Very Deep	0-2
				A1c	Former River Channel	0-2
			Valleys		Alluvial Flat	0-2
					River Valley-Simple	1-4
					River Valley-Complex	0-8
					Flat Valley Bottom	0-2
	SUB-RECENT ALLUVIUM A2	(A2)	River Alluvium	A2p	Alluvial Cover Plain	0-2
				A2f	Alluvial Fan	0-2
	OLD ALLUVIUM A3	(A3)	River Alluvium	A3t	Old Terrace	0-2
				A3t	Terrace Dissected	0-4
ALLUVIO COLLUVIUM	RECENT & SUBRECENT ALLUVIO-COLLUVIAM DEPOSITS A4	(A4)	Slopewash and Alluvium	A4b	Flat Valley Bottom	0-4
				A4v	Interhill Valley-Complex	4-8
				A4f	Colluvial Fan	4-6
				A4s	Interhill-Simple	2-4
ALLUVIO MARINE	RECENT ALLUVIO MARINE DEPOSITS A5	(A5.1)	River Delta Deposits	A51s	Recent Delta Plain Deposits (swampy)	0-2
			Castal Plain Deposits	A51m	Recent Marine Beach Deposits	0-2
	SUB-RECENT ALLUVIO MARINE DEPOSITS A5	(A5.2)	Delta and Castal Plain Deposits	A5.2p	Castal Plain	0-2
				A5.2d	Deltaic Deposits	0-2
UPLAND COLLUVIUM SEDIMENTARY	UPLAND PLAIN DEPOSITS SC	(SC1)	Flat Residual Deposit	U1p	Flat Plain	0-2
			Uneven and Gently Sloping Plains	U1u	Undulating Plain	2-4
				U1r	Rolling Plain	4-8
				U1s	Sloping Plain	6-8
	UPLAND HILLY TERRAIN SYSTEM SC	(SC2)	Plateaus and Gently Sloping Hills	U2f Slopes	Interhill Plateaus or Gently Sloping Side of Interhill Valleys	0-4
			Sloping and Dissected Hills	U2u	Hilly Undulating Terrain	4-8
				U2r	Hilly Rolling Terrain	8-15
				U2s	Hilly Simple Sloping	15-25
				U2d	Hilly Dissected Terrain	15-25
	MOUNTAINOUS TERRAIN SYSTEM SC	(SC3)	Mountainous Strongly Sloping and Dissected	U31	Mountainous Simple	15-35
				U3d	Mountainous Dissected	35-45
		(SC4)	Mountainous Terrain Very Steep	U3s	Mountainous Strongly Sloping	45-65
				U3e	Mountainous Extremely Sloping	65-85
				U3x	Mountainous Very Extremely Sloping	>85

UPLAND COLL[UVIUM]	MOUNTAINOUS TERRAIN SYSTEM SC	(SC3)	Mountainous Strongly Sloping and Dissected	U3l U3d	Mountainous Simple Mountainous Dissected	15-35 35-45
		(SC4)	Mountainous Terrain Very Steep	U3s U3e U3x	Mountainous Strongly Sloping Mountainous Extremely Sloping Mountainous Very Extremely Sloping	45-65 65-85 >85
UPLAND COLLUVIUM VOLCANIC	VOLCANIC PLAIN PLANEZE OR LAHAR* VC	(VR1)	Flat Residual Deposits	V1p	Flat Plain	0-2
		(VC1)	Uneven Colluvial Plains	V1u V1r V1s	Undulating Plain Rolling Plain Sloping Plain, Dissected	2-4 4-8 6-8
	VOLCANIC HILLY TERRAIN AND LAVA FLOWS VC	(VC2)	Flat to Undulating Volcanic Flows	V2s V2u V2r V2d V2t V2p	Volcanic Sloping Hill Side Even Volcanic Hilly Undulating Volcanic Hilly Rolling Volcanic Dissected Toe of a lava flow Interhill Plateau	4-15 4-8 8-15 15-35 15 2-4
UPLAND COLLUVIUM	VOLCANIC MOUNTAINOUS TERRAIN VC	(VC3)	Strongly Sloping or/and Dissected Slopes	V3l V3g V3d	Volcanic Mountainous Simple Volcanic Mountainous Sloping Volcanic Mountainous Dissected; Steep	25-35 25-35 35-45
		(VC4)		V3s V3e V3x	Volcanic Mountainous Dissected or Strongly Sloping Volcanic Extremely Sloping Mountainous Very Extremely Sloping	45-65 65-85 >85
	VOL-CANIC CRA-TERS VC	(VC)	Hillocky to Mountainous Terrain	V4t V4v V4c	Crater Vent Old Crater Dissected	0-6 * 4-8

Note: * For the appropriate symbol and slope see the hilly or mountainous classes as applicable. Lahar will have a special subscript <u>l</u>, recent lava flow <u>r</u> and ancient lava flow <u>a</u>

Where dissected, includes subscript <u>d</u> within the appropriate mapping symbol, usually associated erosion and slope instability hazard.

SCM......Predominantly miocene origin with significant proportion of marine clastic rocks;
SCV,.....Predominantly pleistocene volcanic rocks.
SC.......predominantly pliocene sedimentary rocks.

Qualifying symbol for sedimentary rocks.

<u>Example:</u>

Upland Colluvium

SC_M3d

Sedimentary

Dissected terrain usually indicates high erosion and/or slope instability hazard.

Mountainous terrain with general slope range 25-45 percent.

Miocene sedimentary rocks, significant proportion of marine sediments.

 ..... Most recent sediments of Segara Anakan area.

Appendix 4: Examples of selected sample areas for building the Induction Erosion Model

CONTOH SEBAGIAN DAFTAR TITIK-TITIK PENGAMATAN

Kelas bahaya erosi	Nomor unit lahan	Lokasi
1	2	3
SUB DAS CISEEL		
I	1068	Banjarsari
	514	Pamarica
II	548	Banjaranjar
	744	Bangunsari
III	572	Kalijaga
	1047	Pasirlawang
IV	1268	Sindangwangi
	587	Pasawahan
V	1271	Sindangwangi
	1215	Padaherang
SUB DAS CITANDUY HULU		
I	351	Sukaresik
	352	Panaragan
II	347	Damar Caang
	348	Damar Caang
III	464	Padamulya
	220	Bungursari
IV	914	Citamba
	940	Dirgahayu
V	937	Dirgahayu

Kelas bahaya erosi	Nomor unit lahan	Lokasi
1	2	3
SUB DAS CIMUNTUR		
I	664	Sukahurip/Bangunharja
	462	Cipaku
II	700	Bojonggedang
	437	Selamanik/Selagi
III	640	Sidamulya
	53	Lumbungsari
IV	608	Tanjung Sukar
	475	Gunungsari
SUB DAS CIJOLANG		
I	434	Panulisan
	570	Legok Herang
II	189	Jatisari
	490	Dayeuh Luhur
III	560	Cilumping
	356	Bangunharja
IV	561	Sindangharja/Cijeruk
	216	Jalatrang/Cilebak
V	162	Gunung Aci
	151	Begawat

Kelas bahaya erosi	Nomor unit lahan	Lokasi
1	2	3
SUB DAS SAGARA ANAKAN		
I	992	Kali Jeruk
	481	Prapagan
II	524	Jeruk Legi Kulon
	443	Sarwodadi
III	454	Citepus
	948	Tritiswetan
IV	380	Cisumena
	151	Dermaji
V	456	Citepus
	210	Ringkang

Remarks:

Kelas	I	=	< 15 ton/ha/th
	II	=	15 - <60 ton/ha/th
	III	=	60 - <180 ton/ha/th
	IV	=	180 - <480 ton/ha/th
	V	=	>480 ton/ha/th

Appendix 5: Erosion survey form

FORMULIR ISIAN UNTUK SURVEY EROSI
PERSATUAN UNIT LAHAN
(erosion survey form per unit lahan)

A. IDENTIFIKASI WILAYAH
(area identification)

Lokasi petunjuk *(ref. point)*	....	Tanggal *(date)*	....
Nr. Unit Lahan *(nr. unit lahan)*	....	Pengamat *(surveyor)*	....

Atribut lokasi	**Dari RTL**	**Certainty factor (CF)**	**Pengamatan visual**
(locational attributes)	*(from FEDP)*	*(CF)*	*(visual observation)*
Kenampakan kelas erosi *(apparent erosion class)*	....	....	....
Kemiringan lereng (%) *(slope steepness)*	....	....	....
Penutup lahan *(ground cover)*	....	....	....
Panjang lereng (m) *(slope length)*	....	....	....
Curah hujan *(rainfall)*	....	....	....
Erodibilitas tanah *(soil erodibility)*	....	....	....

B. **PENGAMATAN EROSI**
(erosion observation)

(1) **Keadaan permukaan tanah**
(soil surface condition)

Keadaan gembur *(crumb)*	....
Keadaan padat *(compacted)*	....
Transisi gembur-padat *(transition)*	....

(2) **Erosi percikan**
(splash erosion)

Sealing	ada/tdk ada *(present/absent)*	bila ada *(if present)*
	Ketebalan *(thickness)*	.... cm
	Sealing yang paling tebal *(thickest sealing)*	.... cm
Pedestal	ada/tdk ada *(present/absent)*	bila ada *(if present)*
	Ketinggian *(height)*	.... cm
	Pedestal tertinggi *(highest pedestal)*	.... cm
	Umur tanaman *(age)*	.... tahun *(yrs)*

Skets
(sketch)

(3) Erosi permukaan
(surface erosion)

Penimbuan tanah tererosi *(soil accumulation)*	ada/tdk ada *(present/absent)*	bila ada *(if present)*
	Terjadi pada kemiringan *(at slope steepness)*	.... %
	Umur tanaman *(age of crop)*	.... tahun *(yrs)*
	Penimbunan tanah *(soil accumulation)*	.... cm

Skets
(sketch)

Kemunculan akar tanaman *(root exposure)*	ada/tdk ada *(present/absent)*	bila ada *(if present)*
	Tinggi kemunculan akar *(height of root exposure)*	.... cm
	Kemunculan akar tertinggi *(highest root exposure)*	.... cm
	Umur tanaman/pohon *(age of crop/trees)*	.... tahun *(yrs)*

(4) **Erosi alur**
(rill erosion)

Kenampakan *(presence)*	ada/tdk ada *(present/absent)*	bila ada *(if present)*
	Kedalaman alur *(depth of rill)*	.... cm
	Lebar alur *(width of rill)*	.... cm
	Jumlah alur/10m *(nr. of rills/10m)*	.... buah *(nr.)*

BD tanah 1.3
(bulk density)

C. PENGAMATAN PENUTUPAN TANAH
(ground cover observation)

(1) Jenis penutupan tanah *(type of ground cover)*	Perkiraan (%) *(estimation)*	Pengukuran (%) *(measurement)*	Penghitungan dari photo (%) *(from air photo)*
Tanah terbuka *(bare land)*	....	....	....
Penyebaran *(distribution)*			
- acak *(random)*			
- terkonsentrasi *(concentrated)*			
Penutup *(cover)*			
Serasah *(mulch)*	....	....	....
- ketebalan *(thickness)* .... cm			

Penutupan basal *(basal cover)*	....	....	....
Permukaan batuan *(surface stones)*	....	....	....

Tingkatan	**Penutup tanah**	**Ketinggian**	**Daya tutup**
(type)	*(ground cover)*	*(height)*	*(coverage)*
rumput/herb *(grass/herb)*	....	....	....
Shrub	....	....	....
Pohon *(trees)*	....	....	....

Skets
(sketch)

Appendix 6: Data analysis for building the Inductive Erosion Model

1. IEM development

This part of the study was aimed at identifying the robustness of the proposed IEM model and data quality in a GIS data base. This was achieved using the following steps:

1 establishing the relationship between crop yield and soil erosion in the study area;
2 establishing the relationship between soil erosion, rainfall, soil texture, ground cover, surface roughness, slope length, and gradient;
3 establishing the decision rules to predict erosion class, which will form a main component of an IEM soil erosion model;
4 establishing a means to represent data quality in the GIS database;
5 using the established data quality values and performing a sensitivity analysis on both an IEM and USLE models to determine each model's sensitivity to error in the contributing variables, thereby identifying the most robust model (note that this step was done for only one data type);
6 establishing the quality of the models, using sampling locations not used in model building;
7 evaluating the usefulness of remote sensing in estimating erosion hazard;
8 evaluating the usefulness of data quality information in correctly predicting erosion class.

2 Land cover (vegetation) classification

Landcover classification was based on height and cover percentages of the various stratum data collected from sample areas. For this classification purposes, the vegetation structure cube devloped by ITC (1987) was used.

The most probable landcover in the study area was classified according to the following scheme:

Landcover class	Areal cover (%)		
	Tree	Shrub	Grass
Open forest	60 - 80	20 - 30	-
Woodlands	30 - 60	0 - 30	-
Bushlands	<10	50 - 100	-
Grassy shrublands	<10	30 - 50	<10
Grass (fallow)	either <30	or <30	>40

3 Erosion classification at sample areas

The erosion classification was performed using several indicators as specified below. The final erosion classification was done after each indicator used was separately classified.

(1) Top soil classification (TSC)
The following scheme was used:

Topsoil condition	Score
Compact	3
Intermediate	2
Crumb	1

(2) Splash erosion classification (SPC)
Classified using indicators i.e. sealing and pedestals.

(2a) Sealing classification
Sealing is a dynamic process, changing with wetting and drying. Sealing was classified according to the thickness (ST) and the cover (SC) of sealing. These indicators were obtained using the following formula:

$$ST = \frac{\text{Sealing thickness/site}}{\text{Thickest seal}}$$

$$SC = \frac{\text{Sealing cover/site(\%)}}{\text{Highest cover (\%)}}$$

$$\text{Sealing classification } SC = \frac{STC + SC/site}{2}$$

(2b) Pedestal
Pedestals are features, which if not obliterated by tillage, can increase in height with successive rainfall events. To get an impression of meaningful classification the pedestal heights were divided by the age of vegetation.

Final pedestal classification was estimated using the following formula:

$$PC = \frac{\text{Pedestal height/yr/site}}{\text{Highest pedestal/year}}$$

The splash erosion (SE) was classified using the formula:

$$SE = \frac{SC + PC}{2}$$

(3) Sheet erosion classification (SHC)

Sheet erosion classification was based on soil accumulation and root exposure classification as described below.

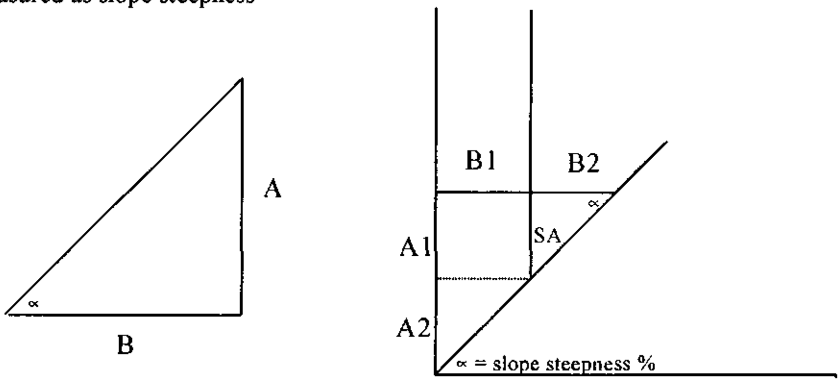
(3a) Soil accumulation classification

Soil accumulation classification was estimated using the formula (see also Figure 42):

$$B(B1+B2) = \frac{A(A1+A2)}{tg\,\alpha} \rightarrow A1 = tg\,\alpha * (A/tg\,\alpha - B1) \text{ or } A1 = A - tg\,\alpha * B1$$

A = measured soil accumulation (cm)

$tg\,\alpha$ = measured as slope steepness



note: SA = soil accumulation

Figure 42 : Soil accumulation estimation

A1 is the corrected soil accumulation estimation and with this formula the correct soil accumulation behind all the plants measured on each site can be found. The estimated soil accumulation per year can be obtained using the following formula:

$$\text{Soil accumulation/year} = \frac{A1}{\text{estimated ages of plants}}$$

In order to obtain a reliable soil accumulation per year (i) the ratio between soil accumulation behind one year old plants and older plants on the same site as well as the ratio for two older plants occurring on the same site was used as correction factor, (ii) in the situation where two older plants exist on the same site the ratio from (i) for younger plants is multiplied by the present ratio, to derive the age correction factor for the oldest plant.

A reliable soil accumulation per year then can be obtained by multiplying the age correction factor and the corrected soil accumulation using the following formula:

$$\text{CSA/year} = \frac{\text{SAC correct/plant}}{\text{age of plant}} * \text{age correction factor}$$

note: CSA = Corrected soil accumulation

The average soil accumulation per year on sites where plants of different ages were measured, was calculated by adding all corrected soil accumulation and dividing by the number of different plant ages measured.

Considering the whole process as described above then the soil accumulation classification (SAC) can be performed using the following formula:

$$SAC = \frac{\text{correct soil accumulation/year/site}}{\text{highest correct soil accumulation/year}}$$

(3b) Root exposure classification

The root exposure classification (REC) was done using the following formula:

$$REC = \frac{\text{root exposure cm/year/site}}{\text{highest root exposure cm/year}}$$

Considering (a) and (b) then the sheet erosion classification (SHC) was done using the formula:

$$SHC = \frac{SAC + REC}{2}$$

(4) Rill erosion classification (RLC)

Rill erosion was measured under various landcovers/landuses. The parameters of rill erosion including depth, width, number per 10 metres were taken into account. Using these parameters allowed the estimation of volume soil loss in tons/ha using soil bulk density equal 1.3.

$$\text{Soil loss (ton/ha)} = \text{depth*width* Nr.per 10 metres*1.3}$$

To estimate soil loss/ha/year the ages of vegetation as described in the previous part were used using the following formula:

The rill soil loss tons per hectare was calculated by dividing the soil loss/ha by the age of the fallow, and younger plants within the natural vegetation.

$$\text{Rill soil loss} = \frac{\text{Rill soil loss tons/ha/yr./site}}{\text{Highest rill soil loss tons/ha/yr.}}$$

Considering (1), (2), (3) and (4) the final soil erosion classification on a particular sample area i.e. observed pattern was determined based on:

$$\text{Erosion} = \frac{TSC + SPC + SHC + RLC}{4}$$

To allow a classification from 0 - 1, soil erosion on a particular sample area was divided by the highest classification. Soil erosion on a particular site was then classified following the scheme below:

Classification	Ranking	Description
0 - 0.25	1	Very slightly eroded
0.26 - 0.50	2	Slightly eroded
0.51 - 0.75	3	Moderately eroded
0.76 - 1.0	4	Severely eroded

4 Estimation of the C-factor

C-factor estimation was applied to the (semi-) natural vegetion (cover) type and to the annual cropping sytems.

The method to calculate C-factor from Kooiman (1987) was used. For different cover types and cropping system the C-factor worked out as follows:

(i) through the cover characteristics during certain periods of the year and corresponding rainfall intensity during thirty minutes (EI_{30}) for that period

(ii) the C-factor for the other types will be worked out directly from the cover data to the sub factor.

The sub factors were estimated as follows:

(4a) Canopy subfactor:
This was worked out according to the cover stratum within each vegetation type. It was divided according to the height of 0.1 - 1.5 m; 1.5 - 5 m; and > 5 m. The effective canopy cover percentage of each height class was first calculated and the canopy sub-factor then obtained based on the Figure 43. Then the sub-factor of different strata was obtained by multiplying this canopy subfactor to obtain the final canopy subfactor.

Based on a random distribution of litter and canopy that can be observed in the sample areas, the effective canopy percentage was calculated using the following formula from Kooiman (1987):

(1) Effective lowest canopy % = (% canopy cover - basal cover)*% bare soil

(2) Effective canopy of the 2nd layer = %2nd canopy*(100% - % 1st canopy).

(3) Effective canopy of 3rd layer = %3rd canopy * (100% - % 2nd canopy * (100% - % 1st canopy).

Equation (1) was used as correction factor in the estimation of (2) and (3).

(4b) Groundcover subfactor:
The ground cover subfactor was calculated using the estimation from USLE

$$Fg = e^{-b * Gc}$$

The (-b) is 0.025 for moldboard.

(4c) Erodibility subfactor

This was calculated by multiplying the reconsilidation subfactor, residual landuse subfactor, and organic matter subfactor.

The reconsolidation and landuse subfactor were obtained using Figures 43, 44 and 45.

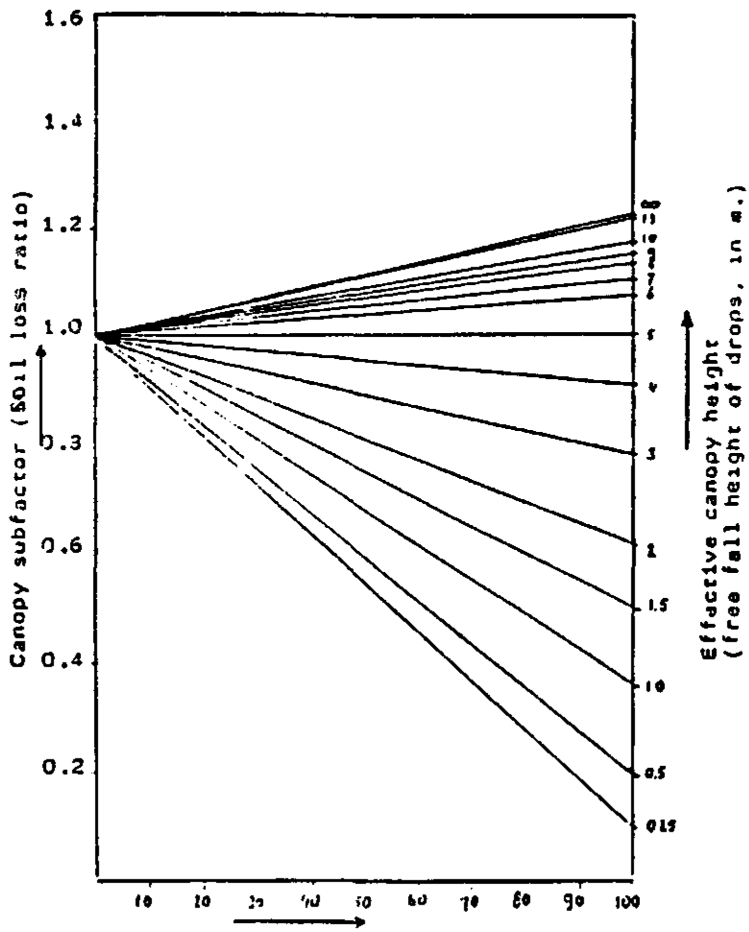


Figure 43: Influence of vegetal canopy height and cover percentage on kinetic energy of throughfall per unit rainfall, assuming bare soil underneath the canopy (Dissmeyer and Foster, 1981)

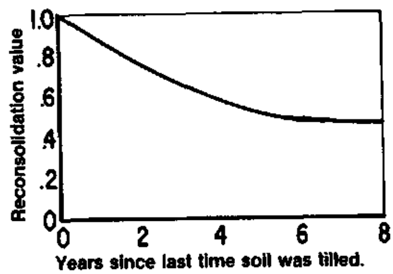


Figure 44: **Soil reconsolidation effects on erodibility in the form of a reconsolidation factor (Dissmeyer and Foster, 1981)**

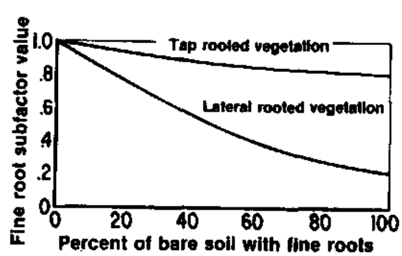


Figure 45: **Effect of fine roots in top 30-50mm of soil erodibility, in the form of a fine root subfactor value (Dissmeyer and Foster, 1981)**

The organic matter was calculated using the method of Dissmeyer and Foster (1981), i.e, a value 0.7 for forest. Higher values were given to vegetation types assumed to have lower organic matter content.

5 The C-factor for (semi-) natural cover type and perennial cropping

1 The canopy subfactor was calculated following the method that was outlined in 4a. The effective canopy was assumed to be 0.5 times the maximum height.

2 The -b for this type of cover is 0.05 (undisturbed soil).

3 The reconsolidation was estimated as follows:
- Plantation with an average of <2 times hoe weeding in Dec - Jan = 0.
- Plantation with slash weeding and no soil disturbance = 0.45
- Other natural cover type = 0.45

4 The organic matter subfactor of 0.7 was used for open forest, 1.0 for fallows and 0.9 for other cover types and plantation.

6 The C-Factor for fallows

1 The average canopy and groundcover percentage of sample areas was calculated by adding all recorded cover percentages and dividing by the number of sample areas.
2 The canopy subfactor was calculated using the method outlined in 4.
3 The b value was used (assumed no cultivation)
4 The reconsolidation factor was worked out according to age of fallow. It was estimated as follows:

1 year	= 0.84
1.1 - 2 year	= 0.74
2.1 - 3 year	= 0.62
3.1 - 4 year	= 0.58
4.1 - 5 year	= 0.47

5 The residual effect was estimated to be 1 (assuming the area to have been cleared for quite some time).
6 The organic matter subfactor of 1 was used.

7 The C-Factor for annual cropping systems

This was calculated using the method developed by Harper 1987 and was adjusted by C-factor calculation suitable to the management system in Indonesia.

The calculation was done according to the development of crop calender for crops, farm management (including weeding, residue treatment, groundcover) and the cover development in time by which the effective ground cover (mulch,litter), and canopy cover (crops and weeds) were used.

The crop stages developed by Wischmeier and Smith 1978 were used as follows:
1 Period F (rough fallow)- from inverson ploughing to secondary tillage.
2 Period SB (seedbed)- From secondary tillage for seedbed preparation until crop has developed 10 % canopy.
3 Period 1 from end of seedbed until the crop has developed 50 % cover.
4 Period 2 from end of 1 until crop has developed 75 % canopy.
5 Period 3 maturing crop from end of 2 until crop harvest.
6 Period 4 from harvest to newsoil preparation.

Information on crop calenders was obtained through interviews. The cover development was obtained using estimation and observations.

The observation of cropping system was concentrated on the most longterm cropping.

The C-factor was calculated per crop stage i.e. based on groundcover, canopy percentage and multiplying by %EI30 for the normal rainfall.

The final C-factor for the cropping systems was calculated by adding all the C-factors per crop stage throughout the year. The following assumptions were used:

1. The selected b-value was 0.025 for maize and the first crop of vegetable because of mechanical tillage, and 0.04 for mungbean and the second crop of vegetables because of hoe tillage, and 0.05 for short fallows because of no tillage.
2. The litter distribution is random so the effective canopy cover was calculated as for other vegetation types.
3. The effective height (average free fall height of throughfall drops) was 0.5 for all crops.
4. The reconsolidation, residual and organic matter sub factors were ignored.

8 The soil erodibility factor (K)

The erodibility of soil was determined according to the following soil properties: organic matter percentage; % silt and very fine sand; % sand ; soil structure; permeability.

The quick test method from Bergsma (1987) was adopted.

9 The slope length and slope steepness factor

The effect of slope steepness was estimated using the following formula adopted from Gandasasmita, 1987.

$$LS = (L/22.1)^{0.79} * (6.432*\sin S\propto^{0.79} * \cos S)$$

L = slope length in metres

S = slope steepness in %

10 Land degradation assessment

(1) Analysis of soil sample

It is aimed at obtaining % organic carbon content. The organic carbon enrichment ratio was worked out by dividing the percentage organic carbon from the valleys and foot slopes, by the percentage organic carbon from the sideslopes.

(2) Calculation of declining yield

The crop yield was obtained by observing crop yield in the study area. This was also done by interviewing farmers who cultivate land in the sample areas. Considering the fluctution of crop yields in the sample areas, a steadily declining crop yield assessement was found from the first year of cultivation.

Appendix 7: Erosion classification at FFL

Appendix 7.1

Splash erosion classification

Sample - ID.	Splash erosion classification								
	Sealing					Pedestal			
	TSC class	Cover (%)	Cover class	Thick (cm)	Thick class	Hght. (mm)	Hght (mm /year)	Class PC	Splashe ros. class
NW 2	0.23	40	0.5	0.1	0.25	10	2.25	0.20	0.32
NW 4	0.67	40	0.5	0.3	0.75	10	10.0	0.91	0.72
NW 7	0.33	80	1.0	0.1	0.25	10	5.0	0.45	0.57
NW 8	0.33	70	0.88	0.2	0.5	9	9.0	9.00	0.82
NW 10	0.33	60	0.75	0.1	0.25	9	6.0	0.55	0.52
NW 11	0.33	40	0.50	0.1	0.25	10	4.76	0.43	0.39
NW 13	0.33	55	0.69	0.3	0.75	9	4.5	0.41	0.62
NW 19	0.33	50	0.63	0.3	0.75	11	5.5	0.50	0.63
NW20	0.33	10	0.13	0.2	0.5	11	2.2	0.2	0.28
NW 21	0.33	55	0.69	0.3	0.75	0	0.0	0.0	0.72
NW 26	0.33	40	0.50	0.1	0.25	9	9.0	0.82	0.52
NW 28	0.33	50	0.63	0.3	0.75	8	5.0	0.45	0.61
NW 29	0.33	40	0.50	0.1	0.25	0	0.0	0.0	0.38
NW 33	0.69	25	0.31	0.2	0.5	11	11.0	1.0	0.60
NW 34	0.33	40	0.50	0.1	0.25	8	8.0	0.73	0.49
NW 35	0.67	65	0.81	0.2	0.5	9	9.0	0.82	0.71
NW 36	0.67	65	0.81	0.3	0.75	10	10.0	0.91	0.82
NW 38	0.67	40	0.50	0.1	0.25	0	0.0	0.00	0.38
NW 40	0.33	20	0.25	0.2	0.50	8	2.67	0.24	0.33
NW 42	0.33	30	0.38	0.1	0.25	9	3.00	0.27	0.30
NW 43	0.33	45	0.56	0.1	0.25	9	4.50	0.41	0.41
NW 45	0.33	70	0.88	0.1	0.25	9	3.60	0.33	0.48
NW 46	0.33	53	0.66	0.3	0.75	10	10.0	0.91	0.77

NW 52	0.67	30	0.38	0.20	0.50	10	5.56	0.51	0.46
NW 54	0.67	30	0.38	0.20	0.50	0	0.00	0.00	0.44
NW 57	0.33	45	0.56	0.10	0.25	0	0.00	0.00	0.41
NW 58	0.33	46	0.58	0.30	0.75	8	8.00	0.73	0.68
NW 9	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
NW 23	0.33	10	0.13	0.20	0.50	0	0.00	0.00	0.31
NW 27	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
NW 32	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 39	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 41	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.22
NW 44	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 48	0.33	5	0.06	0.20	0.00	6	0.00	0.00	0.25
NW 60	0.33	0	0.00	0.00	0.00	0	0.00	0.00	0.00
NW 64	0.33	10	0.13	0.10	0.25	6	1.50	0.14	0.17
NW 67	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
NW 17	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 47	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
NW 49	0.33	0	0.00	0.20	0.50	0	0.00	0.00	0.25
NW 50	0.33	20	0.25	0.20	0.50	0	0.00	0.00	0.38
NW 59	0.33	10	0.13	0.10	0.25	8	8.00	0.73	0.37
NW 61	0.33	0	0.00	0.00	0.00	0	0.00	0.00	0.00
NW 62	0.33	0	0.00	0.00	0.00	0	0.00	0.00	0.00
NW 65	0.33	25	0.31	0.20	0.50	6	1.43	0.13	0.31
NW 66	0.33	10	0.13	0.10	0.25	7	3.18	0.29	0.22
NW 68	0.33	15	0.19	0.20	0.50	0	0.00	0.00	0.34
NW 71	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 1	0.33	25	0.31	0.20	0.50	0	0.00	0.00	0.41
NW 14	0.33	25	0.31	0.10	0.25	10	3.57	0.32	0.30
NW 51	0.33	35	0.44	0.20	0.50	0	0.00	0.00	0.47
NW 53	0.33	25	0.31	0.10	0.25	0	0.00	0.00	0.28
NW 55	0.33	15	0.19	0.10	0.25	8	8.00	0.73	0.39
NW 56	0.33	30	0.38	0.20	0.50	0	0.00	0.00	0.44

NW 63	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
NW 69	0.33	40	0.50	0.10	0.25	0	0.00	0.00	0.38
SE 3	0.67	60	0.75	0.30	0.75	0	0.00	0.00	0.75
SE 4	0.33	35	0.44	0.20	0.50	0	0.00	0.00	0.47
SE 8	0.33	65	0.81	0.30	0.75	0	0.00	0.00	0.78
SE 12	0.67	45	0.56	0.30	0.75	9	4.29	0.39	0.57
SE 20	0.67	60	0.75	0.30	0.75	0	0.00	0.00	0.75
SE 22	0.33	00	1.00	0.40	1.00	0	0.00	0.00	1.00
SE 23	0.33	50	0.63	0.10	0.25	0	0.00	0.00	0.44
SE 27	0.33	80	1.00	0.30	0.75	6	6.00	0.55	0.77
SE 28	0.67	65	0.81	0.30	0.75	0	0.00	0.00	0.78
SE 29	0.33	70	0.88	0.30	0.75	0	0.00	0.00	0.81
SE 32	0.33	30	0.38	0.40	1.00	0	0.00	0.00	0.69
SE 33	0.33	65	0.81	0.30	0.75	0	0.00	0.00	0.78
SE 35	0.33	60	0.75	0.30	0.75	0	0.00	0.00	0.75
SE 37	0.33	20	0.38	0.20	0.50	10	8.30	0.75	0.54
SE 38	0.33	55	0.69	0.20	0.50	0	0.00	0.00	0.59
SE 39	0.33	60	0.75	0.20	0.50	0	0.00	0.00	0.63
SE 40	0.33	40	0.50	0.10	0.25	0	0.00	0.00	0.38
SE41	0.33	35	0.44	0.20	0.50	0	0.00	0.00	0.47
SE 42	0.33	60	0.75	0.10	0.25	0	0.00	0.00	0.50
SE 43	0.33	75	0.94	0.20	0.50	0	0.00	0.00	0.72
SE 44	0.33	70	0.88	0.20	0.50	0	0.00	0.00	0.69
SE 46	0.33	60	0.75	0.10	0.25	0	0.00	0.00	0.50
SE 51	0.33	15	0.19	0.20	0.50	0	0.00	0.00	0.34
SE 52	0.33	35	0.44	0.30	0.75	0	0.00	0.00	0.59
SE 53	0.33	80	1.00	0.30	0.75	0	0.00	0.00	0.88
SE 65	0.33	70	0.83	0.40	1.00	0	0.00	0.00	0.94
SE 66	0.33	55	0.81	0.40	1.00	0	0.00	0.00	0.91
SE 67	0.67	40	0.50	0.30	3.75	0	0.00	0.00	0.63
SE 68	0.33	60	0.75	0.20	0.50	0	0.00	0.00	0.63
SE 69	0.33	40	0.50	0.20	0.50	0	0.00	0.00	0.50

SE 74	0.67	20	0.25	0.10	0.25	0	0.00	0.00	0.25
SE 1	0.33	0	0.00	0.00	0.00	0	0.00	0.00	0.00
SE 6	0.33	45	0.56	0.40	1.00	0	0.00	0.00	0.78
SE 30	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
SE 31	0.33	20	0.25	0.10	0.25	0	0.00	0.00	0.25
SE 70	0.33	40	0.50	0.30	0.75	0	0.00	0.00	0.63
SE 5	0.33	30	0.38	0.20	0.50	0	0.00	0.00	0.44
SE 48	0.33	40	0.50	0.10	0.25	10	5.00	0.45	0.40
SE 49	0.33	40	0.50	0.10	0.75	0	0.00	0.00	0.63
SE 50	0.33	25	0.31	0.10	0.25	0	0.00	0.00	0.28
SE 56	0.33	30	0.38	0.10	0.25	7	3.50	0.32	0.31
SE 57	0.33	30	0.38	0.10	0.25	0	0.00	0.00	0.31
SE 71	0.33	25	0.31	0.20	0.50	10	5.00	0.45	0.42
SE 72	0.33	35	0.44	0.10	0.25	0	0.00	0.00	0.34
SE 75	0.33	28	0.35	0.10	0.25	0	0.00	0.00	0.30
SE 47	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
SE 54	0.33	0	0.00	0.00	0.00	0	0.00	0.00	0.00
SE 55	0.33	15	0.19	0.20	0.50	0	0.00	0.00	0.34
SE 60	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
SE 61	0.33	5	0.06	0.10	0.25	0	0.00	0.00	0.16
SE 62	0.33	8	0.10	0.10	0.25	0	0.00	0.00	0.18
SE 63	0.33	10	0.13	0.10	0.25	0	0.00	0.00	0.19
SE 64	0.33	15	0.19	0.10	0.25	0	0.00	0.00	0.22
SE 58	0.33	10	0.13	0.20	0.50	0	0.00	0.00	0.00
SE 59	0.33	30	0.38	0.20	0.50	0	0.00	0.00	0.44

Appendix 7.2

Sheet erosion classification (exemplar 1)

Samp-ID.	Sheet erosion classification (1)									
	Soil accu. (cm)	Plant diam. (cm)	Slope steep. (%)	Soil acc. corr. (cm)	plant spe.	Esti. age (yrs)	Soil acc. age cm/yr	Soil acc. corrf.	Soil acc. corrf. cm/yr	Soil acc. class
NW 2	3.1	4.1	16	2.44	impe.	3.5	0.7	2.00	1.40	0.42
NW 4	2.1	1.0	22	1.88	eupa.	1.0	1.88	1.00	1.88	0.00
NW 7	2.8	2.5	24	2.20	penn.	2.0	1.14	1.40	1.60	0.47
NW 8	0.0	1.60	16	0.00	penn.	1.0	0.00	1.00	0.00	0.00
NW 10	3.7	2.10	40	2.86	penn.	2.0	1.43	1.40	2.00	0.00
NW 11	2.2	2.00	16	1.84	penn.	2.1	0.88	1.40	1.23	0.37
NW 13	4.3	2.10	23	3.82	bamb	2.1	1.82	1.40	2.54	0.00
NW 19	4.1	3.10	38	2.92	penn.	2.0	1.46	1.40	2.05	0.62
NW20	5.6	4.80	28	4.26	penn.	5.0	0.85	3.00	2.55	0.77
NW 21	2.5	1.80	23	2.09	eupa.	1.30	1.38	1.10	1.10	0.00
NW 26	2.0	1.60	22	1.65	penn.	1.00	1.65	1.00	1.65	0.50
NW 28	0.0	2.50	10	0.00	penn.	1.60	0.00	1.10	0.00	0.00
NW 29	2.8	2.20	13	2.38	penn.	1.00	2.38	1.00	2.38	0.72
NW 33	2.6	1.60	22	2.25	eupa.	1.00	2.25	1.00	2.25	0.00
NW 34	3.3	1.80	26	2.83	penn.	1.00	2.83	1.00	2.83	0.86
NW 35	3.8	1.80	36	3.15	penn.	1.00	3.15	3.15	1.00	0.95
NW 36	3.6	1.60	24	3.22	eupa.	1.00	3.22	1.00	3.22	0.79
NW 38	3.10	2.60	13	2.76	penn.	1.70	1.62	1.40	2.27	0.69
NW 40	3.00	3.80	12	2.54	penn.	3.00	0.85	2.00	1.70	0.52
NW 42	5.20	4.00	36	3.76	penn.	3.00	1.25	2.00	2.51	0.76
NW 43	4.50	2.80	38	3.44	penn.	2.00	1.72	1.40	2.41	0.00
NW 45	5.10	3.50	34	3.91	penn.	2.50	1.56	1.90	2.97	0.00
NW 46	3.00	1.80	23	2.59	eupa.	1.30	1.99	1.00	1.99	0.00
NW 52	4.20	2.80	28	3.42	penn.	1.80	1.90	1.40	2.66	0.66
NW 54	3.60	2.60	26	2.92	penn.	1.50	1.95	1.10	2.14	0.65

NW 57	3.00	1.30	28	2.64	bamb	1.30	2.03	1.10	1.10	0.00
NW 58	2.60	1.60	27	2.17	eupa.	1.00	1.00	1.00	2.17	0.00
NW 9	3.80	3.10	22	3.12	bamb	3.00	1.04	2.00	2.08	0.63
NW 23	3.20	3.70	22	2.39	bamb	3.70	2.00	2.00	1.29	0.00
NW 27	3.10	2.20	40	2.27	grass	2.20	1.01	1.40	1.41	0.43
NW 32	2.80	3.20	22	2.10	bamb	3.20	0.65	2.00	1.31	0.40
NW 39	3.70	2.00	37	2.96	eupa.	1.00	1.48	1.40	2.07	0.63
NW 41	4.20	2.60	26	3.52	shrub	2.60	1.36	1.90	2.58	0.78
NW 44	3.10	2.50	16	2.70	bamb	2.50	1.08	1.90	2.05	0.62
NW 48	3.20	3.00	34	2.18	bamb	3.00	0.73	2.00	1.45	0.00
NW 60	3.20	2.80	24	2.53	bamb	2.80	0.90	2.00	1.81	0.00
NW 64	4.60	4.00	36	3.16	bamb	4.00	0.79	2.20	1.74	0.53
NW 67	3.00	3.00	28	2.16	bamb	3.00	0.72	2.00	1.44	0.44
NW 17	4.80	3.00	40	3.60	bamb	3.00	1.20	2.00	2.40	0.73
NW 47	3.10	1.60	33	2.57	eupa.	1.00	2.57	1.00	2.57	0.78
NW 49	2.80	1.40	22	2.49	eupa.	1.40	1.78	1.10	1.96	0.59
NW 50	3.10	3.20	28	2.20	bamb	3.20	0.69	2.00	1.38	0.42
NW 59	2.70	1.30	23	2.40	eupa.	1.00	2.40	1.00	2.40	0.73
NW 61	2.30	4.10	26	1.23	bamb	4.10	0.30	2.20	0.66	0.20
NW 62	0.00	1.40	22	0.00	eupa.	1.00	0.00	0.00	0.00	0.00
NW 65	4.20	4.20	34	2.77	bamb	4.20	0.66	2.20	1.45	0.44
NW 66	10.00	19.30	38	2.67	tree	19	0.14	3.00	0.42	0.00
NW 68	2.20	2.00	36	1.48	eupa.	2.00	0.74	1.40	1.04	0.32
NW 71	2.70	1.60	32	2.19	shore	1.60	1.37	1.10	1.50	0.45
NW 1	3.70	4.60	25	2.55	bamb	4.60	0.55	3.00	1.66	0.50
NW 14	4.00	2.80	40	2.88	bamb	2.80	1.03	2.00	2.06	0.62
NW 51	11.20	19.80	26	6.05	tree	19	0.32	3.00	0.96	0.00
NW 53	11.00	20.80	22	6.42	tree	20.0	0.32	3.00	0.96	0.00
NW 55	2.00	1.40	36	1.50	eupa.	1.00	1.50	1.00	1.50	0.00
NW 56	3.10	2.60	32	2.27	crot.	2.60	0.87	1.90	1.66	0.00
NW 63	3.60	3.00	32	2.64	bamb	3.00	0.88	2.00	1.71	0.52
NW 69	4.20	2.60	32	3.37	bamb	2.60	1.30	1.90	2.46	0.75

SE 70	2.10	3.70	16	1.51	bamb	3.70	0.41	2.20	0.90	0.27
SE 3	2.40	1.30	15	2.21	impe.	1.00	2.27	1.00	2.27	0.69
SE 4	5.00	2.50	40	4.00	impe.	2.50	1.60	1.90	3.04	0.82
SE 8	4.00	3.40	16	3.46	impe.	2.50	1.38	1.90	2.63	0.80
SE 12	3.60	3.10	22	2.92	bamb	2.10	1.39	1.40	1.95	0.00
SE 20	2.80	1.60	24	2.42	impe.	1.00	2.42	1.00	2.42	0.73
SE 22	3.20	2.60	24	2.58	impe.	1.50	1.72	1.10	1.89	0.57
SE 23	5.20	3.40	33	4.08	penn.	2.50	1.63	1.90	3.10	0.94
SE 27	3.60	1.60	33	3.07	impe.	1.00	3.27	1.00	3.27	0.99
SE 28	2.20	0.80	11	2.11	mimo	1.00	2.11	1.00	2.11	0.00
SE 29	3.40	1.00	33	3.07	mimo	1.00	3.70	1.00	3.07	0.00
SE 32	2.80	3.60	16	2.22	impe.	2.70	0.82	1.90	1.57	0.48
SE 33	3.00	1.00	26	2.58	impe.	1.00	2.58	1.00	2.58	0.81
SE 35	2.60	1.40	22	2.29	impe.	1.00	2.29	1.00	2.29	0.69
SE 37	0.00	0.00	18	0.00	impe.	1.00	0.00	1.00	0.00	0.00
SE 38	3.10	1.80	27	2.61	impe.	1.10	2.38	1.00	2.38	0.72
SE 39	3.60	1.60	34	3.06	impe.	1.00	3.06	1.00	3.06	0.93
SE 40	1.60	1.20	22	1.34	impe.	1.00	1.34	1.00	1.34	0.41
SE41	3.20	3.00	21	2.57	impe.	2.00	1.29	1.40	1.80	0.55
SE 42	4.60	2.20	30	3.94	impe.	1.40	2.81	1.10	3.10	0.94
SE 43	3.80	1.40	36	3.30	impe.	1.00	3.30	1.00	3.30	1.00
SE 44	5.60	3.20	36	4.45	impe.	2.20	2.02	1.40	2.83	0.86
SE 46	3.80	1.80	38	3.12	impe.	1.00	3.12	1.00	3.12	0.95
SE 51	3.80	4.20	26	2.71	impe.	3.50	0.77	2.20	1.70	0.52
SE 52	2.80	2.00	28	2.24	impe.	1.20	1.00	1.00	1.87	0.57
SE 53	2.60	2.00	15	2.30	penn.	1.10	2.04	1.00	2.09	0.63
SE 65	2.80	1.60	12	2.61	impe.	1.00	2.61	1.00	2.61	0.79
SE 66	3.10	1.00	28	2.82	impe.	1.00	2.82	1.00	2.82	0.85
SE 67	2.00	1.80	24	1.57	eupa.	1.00	1.97	1.00	1.97	0.59
SE 68	2.90	1.40	32	2.45	impe.	1.00	1.00	1.00	2.45	0.74
SE 69	5.00	3.00	36	3.92	impe.	2.00	1.96	1.40	2.74	0.83
SE 74	4.40	3.50	18	3.77	impe.	2.50	1.51	1.90	2.87	0.81

SE 1	0.00	1.60	16	0.00	penn.	1.00	0.00	1.00	0.00	0.00
SE 6	2.20	1.30	18	1.97	impe.	1.00	2.06	1.00	2.06	0.62
SE 30	0.00	3.60	26	0.00	impe.	2.50	0.00	1.90	0.00	0.00
SE 31	0.00	3.80	28	0.00	impe.	2.50	0.00	1.90	0.00	0.00
SE 70	0.00	2.30	22	0.00	impe.	1.50	0.00	1.10	0.00	0.00
SE 5	2.80	1.60	18	2.52	bamb	2.00	1.26	1.40	1.76	0.69
SE 48	3.00	2.00	18	2.64	grass	2.00	1.32	1.40	1.85	0.54
SE 49	2.80	2.00	22	2.36	eupa.	1.40	1.69	1.10	1.85	0.00
SE 50	3.80	3.00	22	3.14	impe.	2.00	1.57	1.40	2.20	0.61
SE 56	3.20	1.40	20	2.92	impe.	3.00	1.00	1.00	3.00	0.91
SE 57	3.00	2.00	15	2.70	eupa.	2.00	1.35	1.40	1.92	0.58
SE 71	4.30	2.00	20	3.90	bamb	2.00	1.95	1.40	2.73	0.00
SE 72	2.00	1.40	14	1.80	impe.	1.00	1.85	1.00	1.85	0.56
SE 75	2.60	1.30	16	2.39	impe.	1.00	2.39	1.00	2.39	0.72
SE 47	4.60	3.20	33	3.54	bamb	3.00	1.11	2.00	2.21	0.67
SE 54	3.00	4.00	38	1.48	bamb	4.00	0.37	2.20	0.81	0.00
SE 55	4.80	3.00	40.00	3.60	bamb	3.00	1.20	2.00	2.40	0.00
SE 60	5.30	5.30	40.00	3.46	crato.	1.0	2.20	1.00	2.20	0.68
SE 61	10.20	22.50	38.00	1.65	tree	22.0	0.08	3.00	0.24	0.00
SE 62	5.20	4.00	36.00	3.76	bamb	4.00	0.94	2.20	2.07	0.00
SE 63	3.12	2.00	32.00	2.48	crato.	2.00	1.24	1.40	1.74	0.53
SE 64	5.20	3.00	34.00	4.18	bamb	3.00	1.39	2.00	2.79	0.00
SE 58	6.30	6.30	14.00	2.74	pinus	25.0	0.11	3.00	0.33	0.10
SE 59	4.30	3.00	38.00	3.16	bamb	3.00	1.05	2.00	2.11	0.00

Appendix 7.2

Sheet erosion classification (continued)

Sample -ID.	Sheet erosion classification (2)							
	Root height (cm)	Plant diam. (cm)	Esti. age (yr)	Plant spe.	Esti. age (yrs)	Root expo. age cm/yr	Root expo. class.	Sheet erosi. class
NW 2	1.60	4.60	-	impe.	4.0	0.4	0.27	0.35
NW 4	1.10	1.20	-	eupa.	1.20	0.32	0.00	0.00
NW 7	1.40	3.00	2.0	penn.	1.60	0.75	0.00	0.00
NW 8	1.00	1.00	-	penn.	1.0	0.00	0.00	0.00
NW 10	1.80	3.00	-	penn.	2.0	0.90	0.00	0.00
NW 11	1.40	2.00	1.0	penn.	1.00	1.40	0.93	0.65
NW 13	1.60	2.10	-	bamb	2.10	0.76	0.00	0.00
NW 19	1.80	3.00	2.00	bamb	2.00	0.90	0.60	0.61
NW20	2.00	4.80	5.00	penn	5.00	0.40	0.27	0.52
NW 21	1.00	1.20	-	eupa.	1.20	0.83	0.00	0.00
NW 26	1.00	1.20	-	penn.	0.70	1.00	0.67	0.58
NW 28	0.00	2.50	1.00	penn.	1.50	0.00	0.00	0.00
NW 29	0.90	2.20	1.00	penn.	1.00	0.30	0.60	0.65
NW 33	0.00	1.60	1.00	eupa.	1.00	0.00	0.00	0.00
NW 34	1.10	1.80	1.00	penn.	1.00	1.10	0.73	0.70
NW 35	0.80	1.60	0.60	penn.	1.00	0.80	0.53	0.70
NW 36	1.30	1.80	1.00	penn.	1.00	1.30	0.87	0.91
NW 38	0.80	2.00	-	penn.	1.10	0.80	-	0.00
NW 40	1.20	2.60	2.00	penn.	1.60	0.75	0.50	0.59
NW 42	1.60	3.60	3.00	penn.	2.80	0.57	0.38	0.45
NW 43	2.00	4.00	3.00	penn.	3.00	0.67	0.45	0.60
NW 45	1.80	2.80	2.00	penn.	2.00	0.80	0.00	0.00
NW 46	1.60	3.50	-	eupa.	2.50	0.64	0.48	0.57
NW 52	1.70	2.20	-	bamb	2.29	9.77	0.62	0.73

NW 54	0.80	2.50	-	bamb	1.30	0.00	0.00	0.00
NW 57	0.70	1.80	-	penn.	1.00	1.40	0.93	0.80
NW 58	1.30	1.80	-	penn.	1.00	1.30	0.80	0.75
NW 9	1.20	4.00	-	bamb	4.00	0.30	0.33	0.00
NW 23	0.97	1.80	-	crato.	1.80	0.54	0.20	0.33
NW 27	0.80	2.00	-	grass	2.00	3.40	0.27	0.35
NW 32	0.00	3.20	-	bamb	3.20	0.00	0.00	0.20
NW 39	0.50	1.00	-	eupa.	1.00	0.50	0.33	0.48
NW 41	0.00	2.60	-	shrub	0.00	0.00	0.00	0.39
NW 44	1.50	3.00	-	bamb	3.00	0.50	0.33	0.48
NW 48	0.90	0.30	-	bamb	3.00	0.30	-	-
NW 60	0.00	2.80	-	bamb	2.80	0.00	0.00	0.00
NW 64	0.80	1.80	-	bamb	2.00	0.40	0.27	0.40
NW 67	0.60	3.30	-	bamb	3.30	0.20	0.13	0.28
NW 17	1.60	3.00	-	bamb	3.00	0.37	0.25	0.49
NW 47	0.50	1.60	-	eupa.	1.00	0.50	0.33	0.50
NW 49	0.00	1.40	-	eupa.	1.40	0.00	0.00	0.30
NW 50	1.10	3.20	-	bamb	3.20	0.34	0.23	0.32
NW 59	0.40	1.80	-	eupa.	1.00	0.40	0.22	0.47
NW 61	0.00	4.10	-	bamb	4.00	0.00	0.00	0.10
NW 62	0.00	1.40	-	eupa.	1.20	0.00	0.00	0.00
NW 65	1.90	4.10	-	bamb	4.00	0.46	0.31	0.37
NW 66	6.00	19.30	-	tree	19.0	0.32	0.00	0.00
NW 68	0.00	2.00	-	eupa.	2.00	0.00	0.00	0.16
NW 71	0.80	1.60	-	shore	1.60	0.50	0.33	0.39
NW 1	0.00	4.60	-	bamb	4.60	0.00	0.00	0.25
NW 14	1.00	2.80	-	bamb	2.80	0.35	0.23	0.43
NW 51	7.20	19.80	-	tree	19.0	0.36	0.00	0.00
NW 53	7.00	20.80	-	tree	20.0	0.35	0.00	0.00
NW 55	0.60	1.40	-	eupa.	1.00	0.60	0.00	0.00
NW 56	1.60	2.60	-	crot.	2.50	0.62	0.00	0.00
NW 63	1.80	2.80	-	bamb	2.80	0.64	0.00	0.00

NW 69	1.30	2.80	-	bamb	2.80	0.46	0.31	0.53
NW 70	1.75	3.00	-	bamb	3.00	0.58	0.39	0.33
SE 3	1.00	1.30	1.00	impe.	1.00	1.00	0.67	0.68
SE 4	1.60	1.60	1.00	impe.	1.00	1.60	0.00	0.00
SE 8	1.50	2.60	2.00	impe.	1.50	1.00	0.67	0.73
SE 12	0.00	3.10	2.00	impe.	1.50	0.00	0.00	0.00
SE 20	0.90	1.60	-	impe.	4.60	0.00	0.60	0.67
SE 22	1.30	3.40	-	penn.	2.50	0.53	0.00	0.29
SE 23	1.20	1.30	-	impe.	1.00	1.50	0.35	0.65
SE 27	0.80	1.20	1	mimo	1.00	1.33	1.00	1.00
SE 28	1.30	1.00	1	mimo	1.00	1.30	0.00	0.00
SE 29	0.00	3.60	-	impe.	2.70	0.00	0.00	0.79
SE 32	0.00	1.60	-	impe.	1.00	0.00	0.00	0.24
SE 33	1.30	1.10	-	impe.	1.00	1.30	1.30	0.87
SE 35	0.00	1.60	-	impe.	1.20	0.00	0.00	0.00
SE 37	1.10	1.60	-	impe.	1.00	1.10	0.73	0.73
SE38	1.30	1.80	-	impe.	1.00	1.30	0.87	0.90
SE 39	0.00	1.20	-	impe.	1.00	0.00	0.00	0.20
SE 40	1.40	3.60	-	impe.	2.60	0.81	0.54	0.54
SE41	2.10	2.20	-	impe.	1.50	0.93	0.62	0.78
SE 42	0.00	1.60	1.00	impe.	1.00	0.60	0.10	0.50
SE 43	1.50	3.20	2.00	impe.	2.20	0.68	0.45	0.66
SE 44	1.40	1.00	-	impe.	1.00	1.40	0.93	0.94
SE 46	1.60	4.00	4.00	impe.	3.50	0.46	0.31	0.41
SE 51	0.00	2.00	-	impe.	1.20	0.00	0.00	0.28
SE 52	1.00	2.00	1.00	penn.	1.10	1.00	0.67	0.65
SE 53	0.80	1.60	1.00	eupa.	1.00	0.80	0.53	0.66
SE 65	1.30	1.00	-	impe.	1.00	1.30	0.87	0.86
SE 66	1.20	1.80	1.00	impe.	1.00	1.20	-	0.00
SE 67	1.00	1.00	-	eupa.	1.00	1.00	0.73	0.66
SE 68	1.40	1.20	-	impe.	1.00	1.40	0.00	0.37
SE 69	1.80	3.00	2.00	impe.	2.00	0.69	0.46	0.65

SE 74	0.00	3.50	-	impe.	2.50	0.00	0.00	0.43
SE 1	0.00	1.60	-	impe.	1.00	0.00	0.00	0.00
SE 6	0.80	0.80	0.60	impe.	0.60	0.80	0.53	0.58
SE 30	0.00	3.60	-	impe.	2.50	0.00	0.00	0.00
SE 31	0.00	3.80	-	impe.	2.50	0.00	0.00	0.00
SE 70	0.70	1.30	1.00	impe.	1.00	0.70	0.47	0.23
SE 5	0.80	1.60	-	bamb	1.00	1.14	0.00	0.00
SE 48	1.00	1.60	-	impe.	1.10	1.25	-	-
SE 49	1.30	2.00	-	eupa.	1.40	0.92	-	-
SE 50	0.90	1.40	-	impe.	1.00	1.13	0.75	0.83
SE 56	1.10	2.00	-	impe.	1.20	0.79	0.53	0.00
SE 57	0.00	1.60	-	imper	1.00	0.00	0.00	0.00
SE 71	1.10	2.00	-	bamb	2.00	0.55	-	-
SE 72	0.70	1.10	-	impe.	0.70	1.00	0.67	0.61
SE 75	1.00	2.30	-	impe.	1.60	0.63	0.46	0.59
SE 47	0.96	3.20	-	bamb	3.20	0.30	0.20	0.43
SE 54	0.00	2.00	-	eupa.	2.00	0.00	0.37	0.31
SE 55	0.00	1.80	-	polyg	1.20	0.00	0.00	0.36
SE 60	0.00	1.00	-	crato.	1.00	0.00	0.24	0.46
SE 61	0.00	22.50	-	tree	22.0	0.32	0.00	0.00
SE 62	4.00	15.60	-	tree		0.32	0.00	0.00
SE 63	1.20	3.00	-	crato.	2.00	0.45	0.30	0.41
SE 64	0.00	25.40	-	pinus	25.0	0.00	0.00	0.00
SE 58	2.30	3.00	-	bamb	3.00	0.77	0.00	0.00
SE 59	0.90	1.20	-	shore	1.20	0.75	0.00	0.00

Appendix 7.3

Rill erosion classification

Sample -ID.	Rill erosion classification					
	Rill depth (cm)	Rill width. (cm)	Rill inten/1 0msq	Soil loss ton/ ha	Soil loss ton/ ha/yr	Rill erosi. class
NW 2	9.30	15.80	6.00	114	29	0.33
NW 4	0.00	0.01	9.00	0	-	-
NW 7	0.00	0.90	0.00	0	-	-
NW 8	6.80	9.00	4.00	31	31	0.36
NW 10	0.00	0.00	0.00	0	-	-
NW 11	5.30	12.00	6.00	49	49	0.56
NW 13	0.00	0.00	0.00	0	-	-
NW 19	9.50	12.30	6.00	91	46	0.52
NW20	7.30	17.00	4.00	64	13	0.15
NW 21	0.00	0.00	0.00	0	-	-
NW 26	7.50	9.60	7.00	65	65	0.75
NW 28	6.90	8.30	4.00	29	18	0.21
NW 29	6.30	10.30	4.00	33	33	0.38
NW 33	7.80	9.30	5.00	47	47	0.54
NW 34	8.30	10.00	5.00	54	54	0.62
NW 35	5.60	16.50	6.00	72	72	0.83
NW 36	6.70	13.10	7.00	79	46	0.53
NW 38	7.40	13.60	5.00	65	21	0.24
NW 40	9.20	12.10	6.00	86	28	0.32
NW 42	0.00	0.00	0.00	0	-	-
NW 43	0.00	0.00	0.00	0	-	-
NW 45	6.30	9.70	6.00	47	26	0.30
NW 46	8.00	11.00	7.00	80	53	0.61
NW 52	0.00	0.00	0.00	0	-	-
NW 54	0.00	0.00	0.00	0	-	-

NW 57	0.00	0.00	0.00	0	-	-
NW 58	0.00	0.00	0.00	0	-	-
NW 9	0.00	0.00	0.00	0	-	-
NW 23	0.00	0.00	0.00	0	-	-
NW 27	0.00	0.00	0.00	0	-	-
NW 32	0.00	0.00	0.00	0	0	0.00
NW 39	0.00	0.00	0.00	0	0	0.00
NW 41	0.00	0.00	0.00	0	0	0.00
NW 44	0.00	0.00	0.00	0	0	0.00
NW 48	0.00	0.00	0.00	0	0	0.00
NW 60	0.00	0.00	0.00	0	0	0.00
NW 64	0.00	0.00	0.00	0	0	0.00
NW 67	0.00	0.00	0.00	0	0	0.00
NW 17	0.00	0.00	0.00	0	0	0.00
NW 47	0.00	0.00	0.00	0	0	0.00
NW 49	0.00	0.00	0.00	0	0	0.00
NW 50	0.00	0.00	0.00	0	0	0.00
NW 59	0.00	0.00	0.00	0	0	0.00
NW 61	0.00	0.00	0.00	0	0	0.00
NW 62	0.00	0.00	0.00	0	0	0.00
NW 65	0.00	0.00	0.00	0	0	0.00
NW 66	0.00	0.00	0.00	0	0	0.00
NW 68	0.00	0.00	0.00	0	0	0.00
NW 71	0.00	0.00	0.00	0	0	0.00
NW 1	0.00	0.00	0.00	0	0	0.00
NW 14	0.00	0.00	0.00	0	0	0.00
NW 51	0.00	0.00	0.00	0	0	0.00
NW 53	0.00	0.00	0.00	0	0	0.00
NW 55	0.00	0.00	0.00	0	0	0.00
NW 56	0.00	0.00	0.00	0	0	0.00
NW 63	0.00	0.00	0.00	0	0	0.00
NW 69	0.00	0.00	0.00	0	0	0.00

NW 70	0.00	0.00	0.00	0	0	0.00
SE 3	6.00	10.00	5.00	39	39	0.45
SE 4	0.00	0.00	0.00	0	-	-
SE 8	5.70	9.60	7.00	49	19	0.22
SE 12	0.00	0.00	0.00	0	-	-
SE 20	9.80	10.10	6.00	77	51	0.59
SE 22	8.30	11.30	5.00	60	24	0.28
SE 23	9.80	10.10	6.00	77	77	0.89
SE 27	0.00	0.00	0.00	0	-	-
SE 28	0.30	13.10	5.00	79	-	-
SE 29	8.20	12.80	5.00	81	30	0.34
SE 32	0.00	0.00	0.00	0	-	-
SE 33	8.80	12.20	6.00	83	83	0.95
SE 35	7.50	12.20	6.00	71	71	0.82
SE 37	9.80	10.20	4.00	51	43	0.49
SE 38	9.30	12.10	6.00	87	87	1.00
SE 39	7.80	10.00	6.00	60	60	0.69
SE 40	8.60	10.10	5.00	56	56	0.64
SE41	0.00	0.00	0.00	0	0	0.00
SE 42	6.00	15.00	3.00	35	23	0.26
SE 43	6.70	12.00	7.00	73	73	0.84
SE 44	9.40	10.80	5.00	65	32	0.37
SE 46	0.00	0.00	0.00	0	0	0.00
SE 51	7.30	12.00	4.00	45	12	0.14
SE 52	0.00	0.00	0.00	0	0	0.00
SE 53	4.50	13.40	6.00	47	47	0.54
SE 65	5.60	11.30	5.00	41	41	0.47
SE 66	6.80	12.60	7.00	77	77	0.89
SE 67	0.00	0.00	0.00	0	-	-
SE 68	9.60	11.00	5.00	68	68	0.78
SE 69	10.00	11.00	5.00	71	35	0.40
SE 74	8.00	12.00	5.00	62	24	0.28

SE 1	0.00	0.00	0.00	0	0	0.00
SE 6	0.00	0.00	0.00	0	0	0.00
SE 30	0.00	0.00	0.00	0	0	0.00
SE 31	0.00	0.00	0.00	0	0	0.00
SE 70	0.00	0.00	0.00	0	0	0.00
SE 5	0.00	0.00	0.00	0	0	0.00
SE 48	0.00	0.00	0.00	0	0	0.00
SE 49	0.00	0.00	0.00	0	0	0.00
SE 50	0.00	0.00	0.00	0	0	0.00
SE 56	0.00	0.00	0.00	0	0	0.00
SE 57	0.00	0.00	0.00	0	0	0.00
SE 71	0.00	0.00	0.00	0	0	0.00
SE 72	0.00	0.00	0.00	0	0	0.00
SE 75	0.00	0.00	0.00	0	0	0.00
SE 47	0.00	0.00	0.00	0	0	0.00
SE 54	0.00	0.00	0.00	0	0	0.00
SE 55	0.00	0.00	0.00	0	0	0.00
SE 60	0.00	0.00	0.00	0	0	0.00
SE 61	0.00	0.00	0.00	0	0	0.00
SE 62	0.00	0.00	0.00	0	0	0.00
SE 63	0.00	0.00	0.00	0	0	0.00
SE 64	0.00	0.00	0.00	0	0	0.00
SE 58	0.00	0.00	0.00	0	0	0.00
SE 59	0.00	0.00	0.00	0	0	0.00

Appendix 7.4

Total erosion class on each observation point

Spl-ID	Total Erosion Class							
	Cover type	Estim. age (year)	Top soil class	Splash erosion class	Sheet erosi. class	Rill erosi. class	Final class	Rank 1 - 8
NW 2	F	4.00	0.23	0.32	0.35	0.33	0.33	4
NW 4	F	1.00	0.67	0.72	0.57	0.39	0.59	6
NW 7	F	2.00	0.33	0.57	0.48	0.17	0.39	4
NW 8	F	1.00	0.33	0.82	0.00	0.36	0.38	4
NW 10	F	1.50	0.33	0.52	0.59	0.74	0.55	6
NW 11	F	1.00	0.33	0.39	0.65	0.56	0.48	5
NW 13	F	2.00	0.33	0.62	0.60	0.38	0.48	5
NW 19	F	2.00	0.33	0.63	0.61	0.52	0.52	6
NW20	F	5.00	0.33	0.28	0.52	0.15	0.32	4
NW 21	F	1.50	0.33	0.72	0.54	0.57	0.54	6
NW 26	F	1.00	0.33	0.52	0.58	0.75	0.54	6
NW 28	F	1.50	0.33	0.61	0.00	0.21	0.29	3
NW 29	F	1.00	0.33	0.38	0.64	0.38	0.43	5
NW 33	F	1.00	0.33	0.67	0.70	0.38	0.50	5
NW 34	F	1.00	0.33	0.49	0.70	0.54	0.52	6
NW 35	F	1.00	0.67	0.71	0.91	0.62	0.73	8
NW 36	F	1.00	0.67	0.82	0.70	0.83	0.76	8
NW 38	F	1.70	0.33	0.38	0.59	0.53	0.46	5
NW 40	F	3.00	0.33	0.33	0.45	0.24	0.34	4
NW 42	F	3.00	0.33	0.30	0.60	0.32	0.39	4
NW 43	F	2.00	0.33	0.41	0.57	0.34	0.41	5
NW 45	F	2.50	0.33	0.48	0.73	0.55	0.52	6
NW 46	F	1.00	0.33	0.77	0.64	0.76	0.62	7
NW 52	F	1.80	0.33	0.46	0.66	0.30	0.44	5
NW 54	F	1.80	0.33	0.44	0.50	0.61	0.47	5
NW 57	F	1.30	0.33	0.41	0.80	0.45	0.50	5

NW 58	F	1.00	0.33	0.68	0.75	0.69	0.61	7
NW 9	OF	n	0.33	0.19	0.42	0.00	0.24	3
NW 23	OF	n.a	0.33	0.31	0.33	0.00	0.24	3
NW 27	OF	n	0.33	0.19	0.35	0.00	0.22	3
NW 32	OF	n	0.33	0.22	0.20	0.00	0.19	2
NW 39	OF	n	0.33	0.22	0.48	0.00	0.26	3
NW 41	OF	n	0.33	0.19	0.39	0.00	0.19	2
NW 44	OF	n	0.33	0.22	0.48	0.00	0.26	3
NW 48	OF	n	0.33	0.25	0.32	0.00	0.23	2
NW 60	OF	n	0.33	0.00	0.27	0.00	0.26	3
NW 64	OF	n	0.33	0.17	0.40	0.00	0.23	2
NW 67	OF	n	0.33	0.19	0.28	0.00	0.15	2
NW 17	B	n	0.33	0.22	0.49	0.00	0.26	3
NW 47	B	n	0.33	0.19	0.50	0.00	0.26	3
NW 49	B	n	0.33	0.25	0.30	0.00	0.22	2
NW 50	B	n.a.	0.33	0.38	0.32	0.00	0.26	3
NW 59	B	n	0.33	0.31	0.47	0.00	0.28	3
NW 61	B	n	0.33	0.00	0.10	0.00	0.11	1
NW 62	B	n.a	-	0.00	-	0.00	-	-
NW 65	B	n	0.33	0.31	0.37	0.00	0.25	3
NW 66	B	n.a	0.33	0.22	0.37	0.03	0.24	2
NW 68	B	n	0.33	0.34	0.16	0.00	0.21	2
NW 71	B	n	0.33	0.22	0.39	0.00	0.24	2
NW 1	W	n	0.33	0.41	0.25	0.00	0.25	3
NW 14	W	n	0.33	0.30	0.43	0.00	0.27	3
NW 51	W	n.a	0.33	0.47	0.37	0.09	0.31	3
NW 53	W	n	0.33	0.28	0.51	0.00	0.28	3
NW 55	W	n	0.33	0.39	0.47	0.00	0.30	3
NW 56	W	n	0.33	0.44	0.51	0.03	0.33	3
NW 63	W	n.a	0.33	0.22	0.57	0.00	0.28	3
NW 69	W	n	0.33	0.19	0.53	0.00	0.26	3
SE 70	W	n	0.33	0.38	0.33	0.00	0.26	3

SE 3	F	1.00	0.67	0.75	0.68	0.45	0.64	6
SE 4	F	2.50	0.33	0.47	0.84	0.00	0.41	4
SE 8	F	2.50	0.33	0.78	0.73	0.22	0.51	5
SE 12	F	2.10	0.67	0.57	0.23	0.37	0.46	5
SE 20	F	1.00	0.67	0.75	0.67	0.00	0.52	5
SE 22	F	1.50	0.33	1.00	0.29	0.59	0.55	6
SE 23	F	2.50	0.33	0.44	0.65	0.28	0.42	4
SE 27	F	1.00	0.33	0.77	1.00	0.89	0.75	8
SE 28	F	1.00	0.67	0.78	0.74	0.54	0.68	7
SE 29	F	1.20	0.33	0.81	0.79	0.00	0.48	5
SE 32	F	2.70	0.33	0.69	0.24	0.34	0.40	4
SE 33	F	1.00	0.33	0.78	0.40	0.95	0.62	6
SE 35	F	1.00	0.33	0.75	0.78	0.82	0.67	7
SE 37	F	1.20	0.33	0.54	0.00	0.49	0.34	3
SE 38	F	1.00	0.33	0.59	0.73	1.00	0.66	7
SE 39	F	1.00	0.33	0.63	0.90	0.69	0.64	6
SE 40	F	1.00	0.33	0.38	0.20	0.64	0.39	4
SE41	F	2.60	0.33	0.47	0.54	0.00	0.34	3
SE 42	F	1.50	0.33	0.50	0.78	0.26	0.47	5
SE 43	F	1.00	0.33	0.72	0.50	0.84	0.60	6
SE 44	F	2.00	0.33	0.69	0.66	0.37	0.51	5
SE 46	F	1.00	0.67	0.50	0.94	0.00	0.53	5
SE 51	F	0.33	0.34	0.41	0.14	0.30	0.41	4
SE 52	F	1.20	0.33	0.59	0.28	0.00	0.30	3
SE 53	F	1.00	0.67	0.88	0.65	0.54	0.69	7
SE 65	F	1.00	0.33	0.94	0.66	0.47	0.60	6
SE 66	F	1.00	0.33	0.91	0.86	0.89	0.75	8
SE 67	F	1.00	0.67	0.63	0.66	0.78	0.69	7
SE 68	F	1.00	0.33	0.63	0.67	0.00	0.41	4
SE 69	F	2.00	0.33	0.50	0.65	0.40	0.47	5
SE 74	F	2.50	0.67	0.25	0.43	0.28	0.41	5
SE 1	PL	18.00	0.33	0.00	0.00	0.00	0.00	1

SE 6	PL	7.00	0.33	0.78	0.58	0.00	0.42	4
SE 30	PL	7.00	0.33	0.19	0.00	0.00	0.13	1
SE 31	PL	7.00	0.33	0.25	0.00	0.00	0.15	2
SE 70	PL	7.00	0.33	0.63	0.23	0.00	0.30	3
SE 5	GS	n	0.33	0.44	0.58	0.00	0.34	3
SE 48	GS	n	0.33	0.40	0.59	0.07	0.35	4
SE 49	GS	n	0.33	0.63	0.58	0.00	0.39	4
SE 50	GS	n	0.33	0.83	0.00	0.36	0.48	5
SE 56	GS	n	0.33	0.31	0.29	0.00	0.23	2
SE 57	GS	n	0.33	0.00	0.00	0.16	0.21	2
SE 71	GS	n	0.33	0.42	0.58	0.06	0.35	4
SE 72	GS	n	0.33	0.34	0.61	0.00	0.32	4
SE 75	GS	n	0.33	0.30	0.59	0.00	0.31	3
SE 47	OF	n	0.33	0.22	0.43	0.00	0.25	3
SE 54	OF	n	0.33	0.00	0.31	0.00	0.16	2
SE 55	OF	n	0.33	0.34	0.36	0.00	0.26	3
SE 60	OF	n	0.33	0.19	0.46	0.00	0.25	3
SE 61	OF	n	0.33	0.16	0.25	0.00	0.19	2
SE 62	OF	n	0.33	0.18	0.33	0.03	0.22	2
SE 63	OF	n	0.33	0.19	0.41	0.00	0.23	2
SE 64	OF	n	0.33	0.22	0.43	0.00	0.25	3
SE 58	W	n	0.33	0.00	0.00	0.00	0.08	1
SE 59	W	n	0.33	0.44	0.40	0.00	0.29	3

Appendix 8: Pascal program for erosion hazard assessment using an IEM

An IEM was built in the computer in pascal using the following procedures:

```
Procedure erosion;
Var
  E    : Real;
  EF   : Real;
  TMUNo : integer;
  i    : array [1...7] of real;
  j    : integer;
  val  : integer;
Begin
  E    := 1;
  dst  := askreg('output to',2);
  clear(dst);
  TMUNo := askI('Terrain Mapping UnitNo:',4);
  For j:= 1 to 7 do
    Begin
      Writeln('Map-Factor:');
      Readln(i[j]);
    End;

  For j:= 1 to 7 do
    E:= E * i[j];
  With vec do
   If TMU^[TMUNo].upad then
    Begin
     Writeln('Palected Terrain Mapping Unit');
     TMUras(TMUNo,255,dst);
    End;
  Writeln('Eroded terrain Mapping Unit:',E:10);
  Paupa;
End;
```

After syntax checking, translation, and creating an executable program, an IEM procedure was stored in the memory i.e. IGP program under Vector main menu and an IEM submenu as described schematically in the following figure.

```
  IGP
.... ..... Vector
.... ..... ..... .... ...

      IEM ..... ..
```

Erosion severity classification scheme:

E Value	Erosion Severity Class
0.76 - 1.00	Severely eroded
0.51 - 0.75	Moderately eroded
0.26 - 0.50	Slightly eroded
0.00 - 0.25	Very slightly eroded

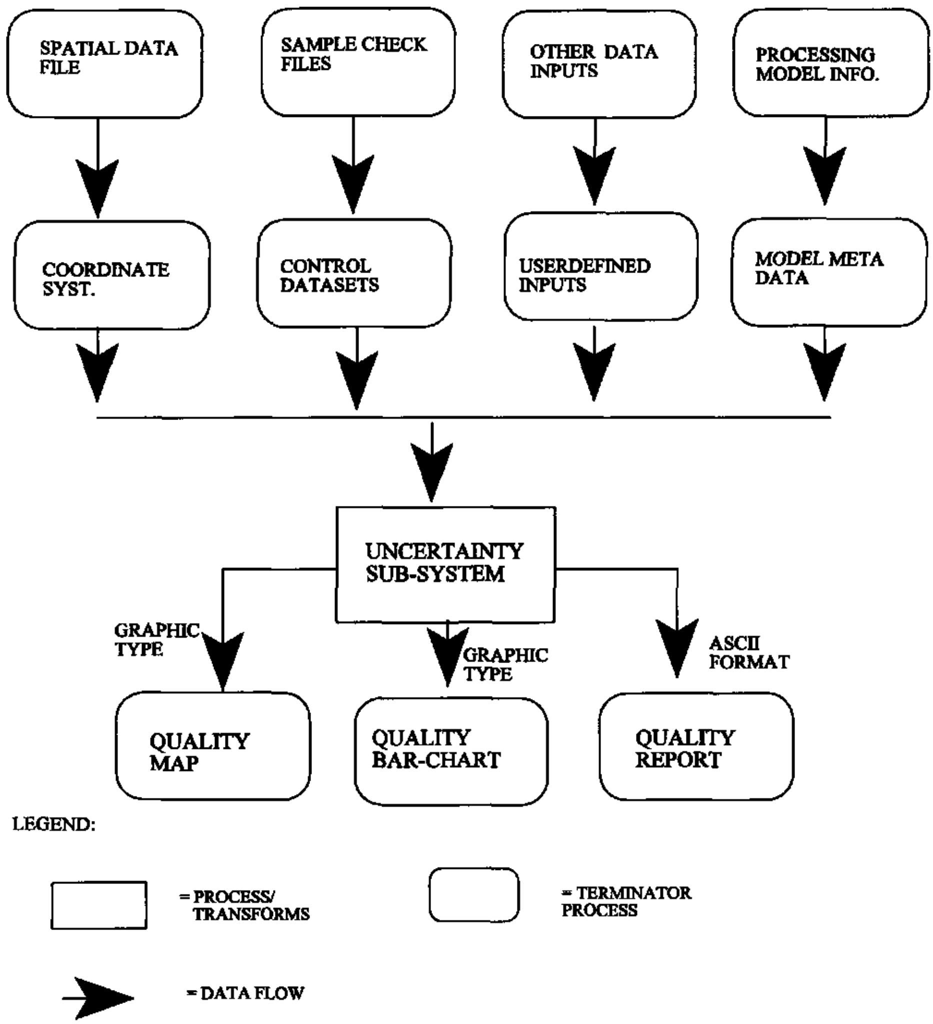
The steps in the erosion paverity aspassment using an IEM model are as follows:

1 Run IGP programme;
2 Select Vector Menu;
3 Select submenu : Load polygon: inpart TMU.pol
4 Select PoltoRas menu;
5 Select LUT: Choose standard;
6 Select Vector Menu;
7 Select submenu IEM;
   a. Output (register 1,...4):[2]
   b. Clearing.....
   c. PolNo:[4] ....
   d. Map-Factor1:.....
   e. Map-Factor2:.....
   ....
   ....
   i. Map-Factor7:.....
   j. Selected polygon....
   k. Erosion polygon:.........E.....

Appendix 9

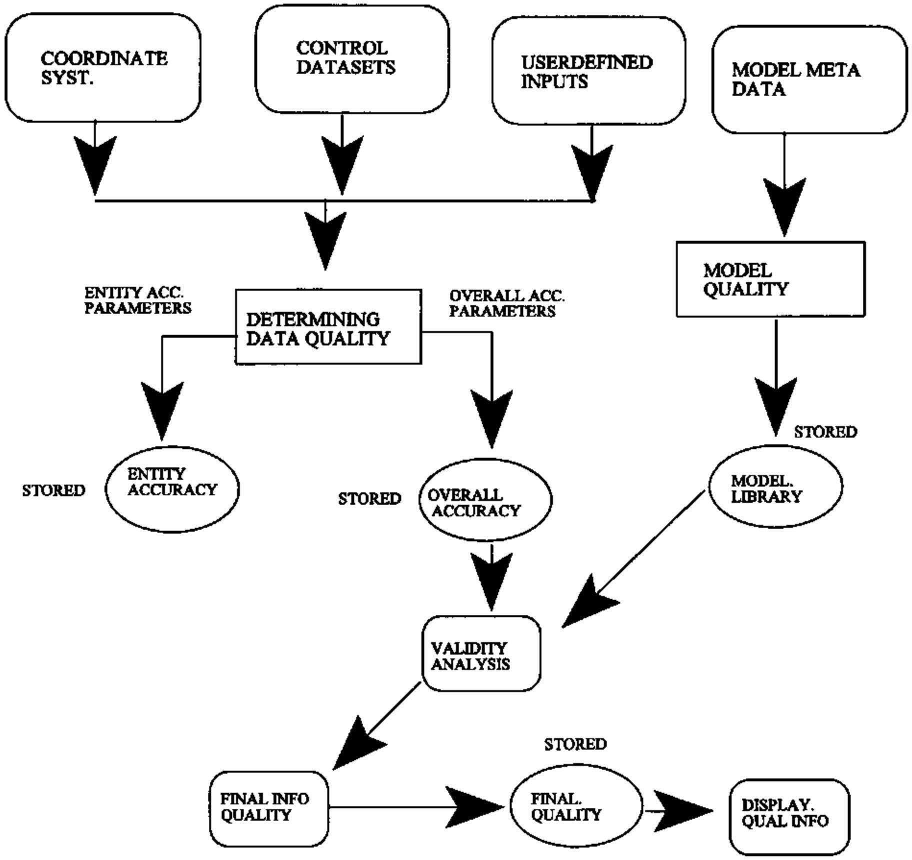
Appendix 9.1

The prototype uncertainty subsystem (1)



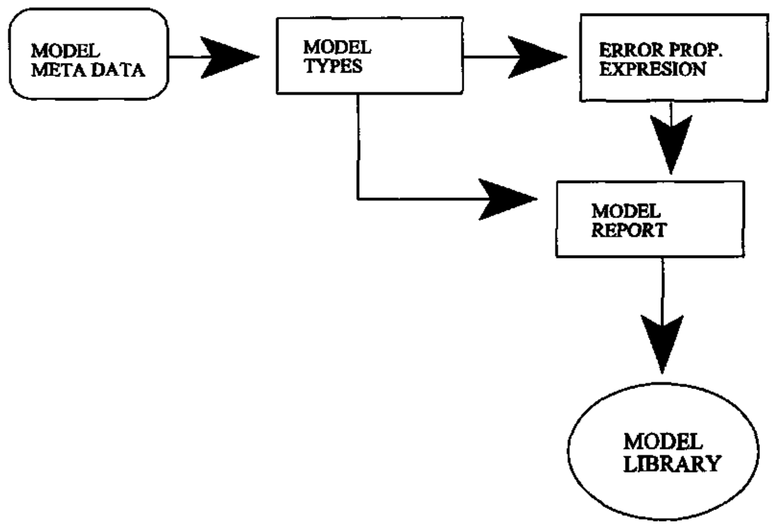
Appendix 9.2

The prototype uncertainty subsystem (2)



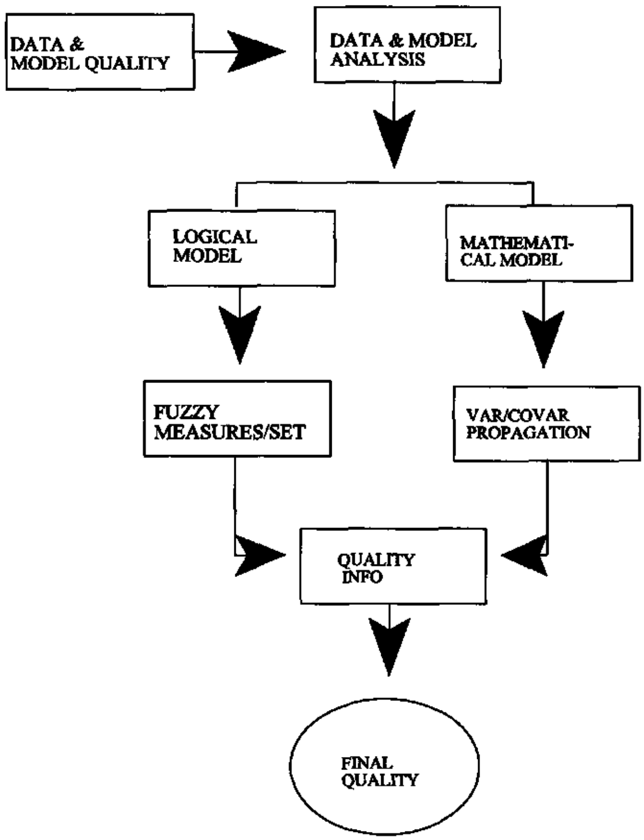
Appendix 9.3

Model information quality file



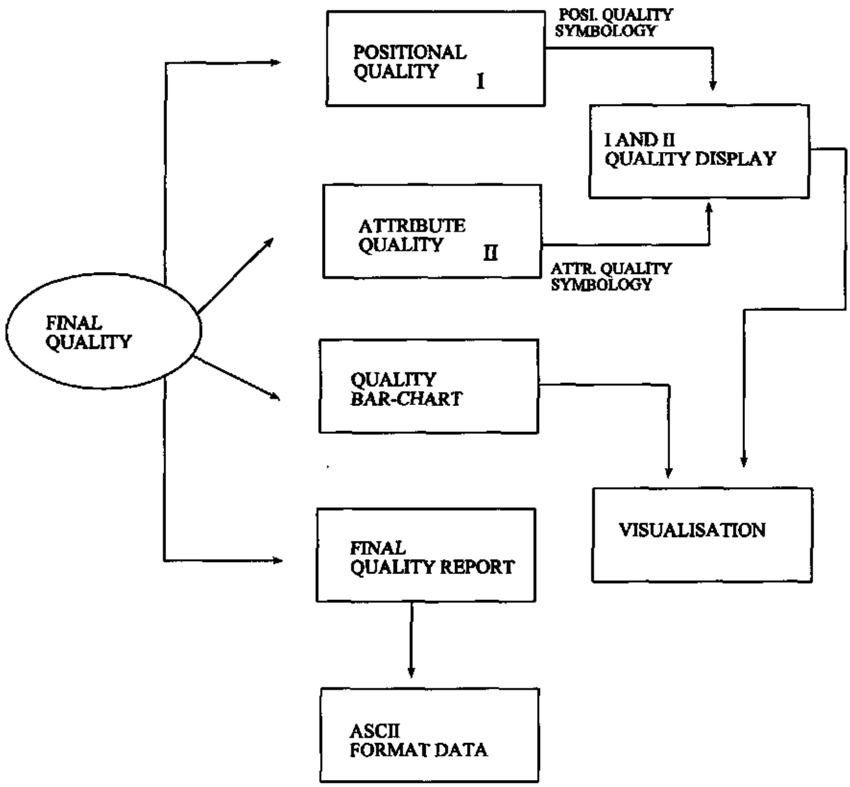
Appendix 9.4

Final information quality file



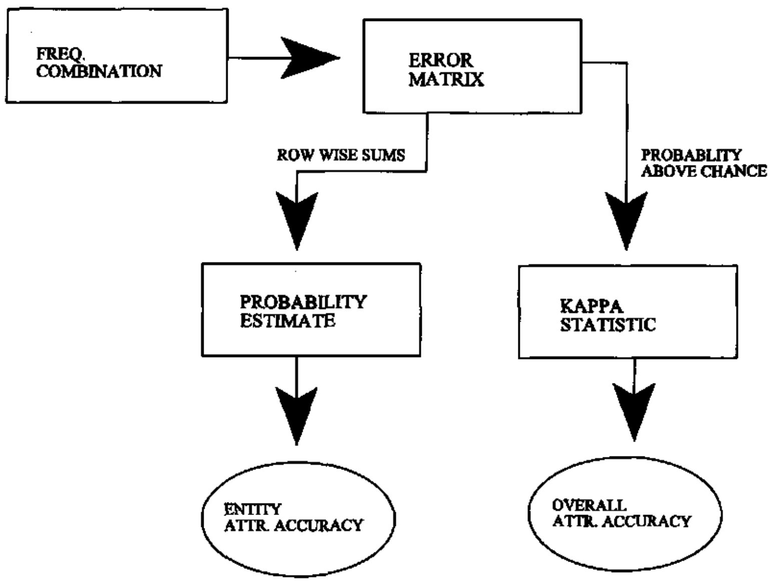
Appendix 9.5

Display quality information



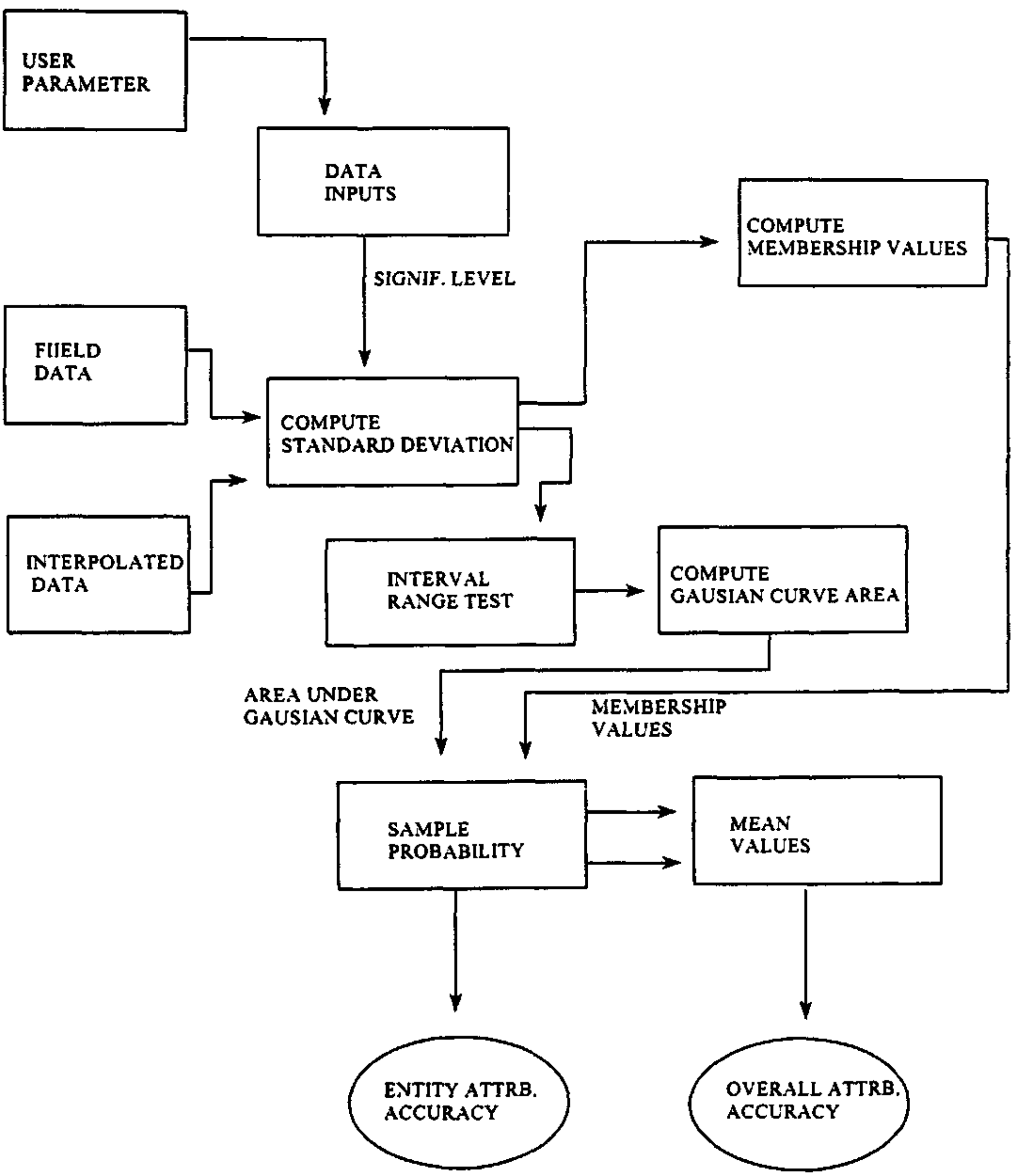
Appendix 9.6

Discontinuous attribute accuracy



Appendix 9.7

Continuous attribute accuracy



ABBREVIATIONS AND GLOSSARY

ANSWERS	=	the Areal Nonpoint Source Watershed Environment Response Simulation.
CREAMS	=	the Chemical Runoff and Erosion from Agricultural Management Systems.
EPIC	=	the Erosion Productivity Index Calculator.
GAMES	=	the Guelph model for evaluating the effects of Agricultural Management Systems.
SLEMSA	=	the Soil Loss Estimation Model for Southern Africa.
SOTER	=	SOil and TERrain.
SWEAP	=	SOTER Water Erosion Assessment Programme.
SBRLKT	=	Sub Balai Rehabilitasi Lahan and Konservasi Tanah is the Indonesian land rehabilitation and soil conservation subcenter.
WEPP	=	the Water Erosion Prediction Project.
RMI	=	The Resources Management International is an international consulting company.
UNPAD	=	Universitas Padjadjaran is a state university in Bandung.

Aggregation groups multiple individuals to a new (complex) object (adopts part_of_relation).

Algorithm is a step-by-step procedure for solving a mathematical problem. For example, the conversion of data in one map projection to another map projection requires that the data be processed through an algorithm of precisely defined rules or mathematical rules.

Accuracy
(1) If applied to paper maps or map data bases, represents degree of conformity with a standard or accepted value. It relates to the quality of a result.
(2) If applied to data collection device such as digitisers, it represents the degree of obtaining the correct value.

Ambiguity describes the difference between maximum and all other possibilities to indicate specifity and diversity of possible allocation.

Chaining is a reasoning technique using inference rule in which the truth of the condition of one rule leads to another rule for which that conclusion is a premise (*forward chaining*), or in which the need to prove the truth of a premise of one rule leads us to another rule for which that premise is a conclusion (*backward chaining*)

Certainty Factor (CF) is an informal measure of the likelihood that the premise or conclusion of an inference rule is true. It can be a degree of certainty or an expert expression in inferring a particular conclusion on an underlying process based on given particular conditions. CF is expressed numerically in two forms: absolute belief, CF values between 0 and 1; change in belief value between -1 and +1.

Conflict resolution is the process of determining which of two or more production rules are to be used when each may be applicable to a set of circumstances.

Confusion describes the difference between maximum and second possibility value.

Crisp set is a set whose boundaries are crisp or abrupt, e.g., degree of certainty 1.0 and 0.0. For the definition of a crisp set only a threshold value is required.

Classification is the abstraction from individuals with common properties to a class (instance_of_a relation).

Connectivity is a topological construct.

Coordinate system is a system to measure horizontal and vertical distance on a planimetric map. In a GIS, it is the system whose units and characteristics are defined by a map projection. A common coordinate system is used to spatially register geographic data for the same area.

Contiguity is the topological identification of adjacent polygons rcording the left and right polygons of each arc.

Continuous data in a GIS usually refers to numerical grid or raster data representing surfaces such as elevation. In this instance, the data can be any value, positive, or negative. Sometimes referred to as real data.

Contour is a line connecting points of equal value. Often in reference to a horizontal datum such as mean sea level.

Digital is referred to as data that are in computer-readable format.

Domain is an area of study or activity.

Data is a general term to denote any or all facts, numbers, letters and symbols that refer to or describe an object, idea, condition, situation or other factors. May be line graphic, imagery and or alphanumerics. They can also be basic elements of information which can be processed, stored and produced by a computer.

Expert system is a computer system whose goal is to make decisions or plan as well as or better than an expert in a particular domain.

Farmer's Field Level (FFL) is a level where farmers conduct their agricultural activities.

Field Engineering Design Plan (FEDP) is a medium term plan (five years plan) for land rehabilitation and soil conservation programme in Indonesia.

Fuzziness or vagueness describes the overall deviation of all available possible values related to absolute truth and falsehood.

Fuzzy set is a set whose boundaries are characterised by a gradual (transition) zone. For the definition of a fuzzy set values for threshold and dispersion and a proper membership function are required.

Fuzzy boundary is associated with the gradual transition from membership to non membership of an object in a fuzzy set.

Fuzzy subset is a finite supporting subset of a crisp set with fuzzy members.

Feature is a representation of a geographic entity, such as a point, line or polygon.

Generalisation is the combination of several classes to a more general superclass (is_a-relation).

GIS is an organised collection of computer hardware, software, geographic data and personnel designed to efficiently capture, store, update, manipulate, analyse, and display all forms of geographically referenced information.

Heuristics is a rule of thumb, or guide line, that can be applied to making a decision when we are not sure which path to take; after applying the guideline, we may still not know if the correct path was taken.

Homogeneous relief Terrain Mapping Unit (TMU) is the natural division of a terrain and is considered to be homogeneous in relief surface.

Inductive Erosion Model (IEM) is a rule-based inference model, built after observing observed patterns at FFL.

Information is data related to answering a specific question.

Is_ a_ relationship is the upward relationship of a classification hierarchy. This relation links each particular object to a class and to a super class. This relation represents that an object is an instance of a class and that a class is a special case of more general super class.

Inference is the act of reaching a conclusion based on a set of logical rules.

Inference rule is a primary way of representing knowledge for use in an expert system. The inference rule has one or more conditions (*the IF part*) followed by a conclusion (*the THEN part*) that is true if the premises are true.

Possibility is membership degree of the truth between 0 and 1 indicating the membership degree of a fuzzy set for a particular data value. Memberships of a set of fuzzy sets are mutually dependent.

Polygon is a vector representation of an enclosed region, described by a sequential list of vertices or mathematical functions.

Part_of relationship is the upward relationship of an aggregation hierarchy. This relation links a particular set of objects to a specific composite object and on to a specific more complex object.

Rule-based is a collection of inference rules that forms the knowledge base for an expert system.

Scheffe's test is a conservative method of testing the significance of one or more comparisons of mean values arising in analysis of variance where the comparison are selected by inspection as being of interest.

Sliver polygon is a relatively narrow feature commonly occurring along the borders of polygons following the overlay of two or more geographic data sets. Also occurs along map borders when two maps are joined as a result of inaccuracy of the coordinates in either or both maps.

Transition zone is a range of data values which are assigned possibility values between 0 and 1 by a specific membership function.

Topology is a spatial relationship between connecting or adjacent coverage features, e.g., arcs, nodes, polygons and points.

Terrain object are a user-defined phenomena that can be modelled or represented using geographic data sets. Examples of terrain objects include parcels.

References

Ambar, S. and Syafrudin. 1979. Pemetaan erosi DAS Jatiluhur. Seminar masalah erosi DAS Jatiluhur. Lembaga Ecology, Universitas, Padjadjaran.Bandung. September 1979, 34 pp.

Aspinall, R.1992. An inductive modelling procedure based on Bayes's Theorem for analysis of pattern in spatial data. International Journal of Geographical Information Systems. Vol 6, No.2. Taylor and Francis. London, New York and Philadelphia, pp.105-121.

Beasley, D. B., L. F. Huggins, E.J. Monke. 1980. ANSWERS: A model for watershed planning. American Society of Agricultural Engineers. Transaction of the ASAE, Vol. 23(4), pp. 938-944.

Bergkamp, G. 1995. A hierarchical approach for desertification assessment. In Environmental Monitoring and Assessment. Special issue: Desertification in developed countries. Vol 37 Nos 1-3. Eds. G.B. Wiersma, D.A. Mouat and C.F. Hutchinson. U.S. Environmental Protection Agency, National Helath and Environmental Effects Laboratory, Corvalis, OR, USA, pp. 59-78.

Bergsma, E. 1982. Aerial photo-interpretation for soil erosion and conservation surveys. Part III: Erosion mapping. ITC-lecture note SOL15, Enschede, 58 pp.

Bergsma, E. 1990. Approximation of soil erodibility using simple test. Lecture note. ITC, Enschede. 22 pp.

Bergsma, E. and F. J. P. M. Kwaad. 1992. Rain erosion hazard evaluated from soil as well as from features of soil surface micro topography. Paper presented at the VII th Conference of the International Soil Conservation Organisation, ISCO, Sydney, Australia. 17pp.

Bocco, G. and C. R. Valenzuela. 1988. Integration of GIS and image processing in soil erosion studies using ILWIS. ITC Journal. 1988-4. ITC, Enschede, The Netherlands, pp. 309-319.

Bocco, G.1990. Gully erosion analysis using remote sensing and geographic information systems. A case study in Central Mexico. Doctorate Thesis. University of Amsterdam, Amsterdam, pp.129.

Bonnet, A.1985. Artificial intelligence. Promise and performance. Prentice/Hall International, 221 pp.

Bouma, J. and M. R. Hoosbeek. 1996. The contribution and importance of soil scientists in interdisciplinary studies dealing with land. In The Role of Soil Science in Interdisciplinary Research. Special publication 45. American Society of Agronomy and Soil Science of America, Madison, USA, pp. 1- 15.

Braspenning, P. J., J. W. H. M. Uiterwijk and H. Bakker. 1989. Object-georienteerd modelleringsgereedschap voor complexe objecten. In Proceeding of The Nederlandstalige Artifical Intelligence Conferentie (NAIC), Enschede, 1989. Eds. L. Siklossy and R.R. Bakker. Academic service, pp. 121-128.

Brimicombe, A. J. 1993. Responsibility, uncertainty, and fitness-for use implication for geo-information theory. Paper presented at ISPRS, Commission III, Washington, 9 pp.

Burgess,T. M. and R. Webster. 1983. A probabilistic approach to mapping soil boundaries. Soil information technology. Proceeding of the 6th meeting of the ISSS working group on soil information systems. Bolkesjo, Norway. February 1983, pp 121-126.

Burrough, P. A. 1989. Matching spatial databases and quantitative models in land resource assessment. In Soil Use and Management. ILRI, Vol. 5, Nr. 1, March 1989, pp. 3 - 8.

Burrough, P. A., R. A. MacMillan and W. van Deursen. 1992. Fuzzy classification methods for determining land suitability from soil profile observations and topography. Journal of Soil Science, 1992,43, pp. 193-210.

Burrough, P. A., G. B. M. Heuvelink. 1992. The sensitivity of boolean and continuous (fuzzy) logical modelling to uncertain data. In Proceeding EGIS 92, Vol.2 The Third European Conference on Geographic Information Systems, Munich, Germany. Eds. J. Harts, Ottens, H.P.L., Scholten, H.J. EGIS Foundation, Utrecht, The Netherlands, pp. 1032-1041.

Carrara, A. 1992. Landslide hazard assessment. In 1er Simposio Internacional Sobre Sensores Remotos y Sistemas de Information Geographica (SIG) Para el Estudio de Riesgos Naturales. Memorias.Ed. J. B. Alzate B.Instituto Geografico Agustin Codazzi, Bogota, pp. 329 - 355.

Chrisman, N. 1984. Method for spatial analysis based on error in categorical maps. Unpublished doctorate thesis. University of Bristol, UK, 260 pp.

Coad, P. and E. Yourdon. 1991. Object-oriented analysis.Yourdon Press. Second edition.Prentice Hall. Englewood Cliffs, New Jersey, 233 pp.

Dayat, D. N. S. 1993. Organisational efficiency in managing the regreening project as an example of natural resource management in Indonesia. A case study at Sukabumi district, with special reference to the Cibodas Biosphere Reserve, West Java, Indonesia. Unpublished Msc. thesis, ITC, Enschede, 73 pp.

De Hoop, S. 1993. Feature-oriented data models for GIS: concepts and potentials. Proc. 3rd Int. User forum, geographical information systems new perspective, Siemens Nixdorf Information Systems, Munich, 10 pp.

De Jong, S.M. 1994. Application of Reflective Remote Sensing for Land Degradation Studies in a Mediterranean Environment. Doctorate thesis. The Netherlands Geographical Studies, The State University of Utrecht, Utrecht, pp. 237.

De Jong, S. M. and H. Th. Riezebos. 1994. Imaging spectroscopy, geostatistics and soil erosion models. In conserving soil resources. European perspective. Selected papers from the First International Congres of the European Society for Soil Conservation. Ed. R.J. Rickson. CAB International, Willingford, Oxon, UK, pp. 232 - 233.

De Man, E. 1988. Establishing a geographical information system in relation to its use. A process of strategy choices. International Journal of Geographical Information Systems. Vol 2, No.3. Taylor and Francis. London, New York and Philadelphia, pp. 245-261.

De Roo, A. P. J., L. Hazelhoff and P.A. Burrough. 1989. Soil erosion modelling using ANSWERS and geographical information systems. Earth Surface Processes and Landform, Vol. 14, pp. 517-532.

De Roo, A. P.J. 1993. Modelling surface runoff and soil erosion in catchment using geographical information systems. Doctorate thesis. The Netherlands Geographical Studies, The State University of Utrecht, Utrecht, pp. 295.

Dickinson, W. T., G. J. Wall and R. P. Rudra. 1990. Model building for predicting and managing soil erosion transport. In Soil Erosion on Agricultural Land. Eds. J. Boardman, I. D. L. Foster and J. A. Dearing. John Wiley and Sons, Ltd., London, UK, pp. 415-445.

Dissmeyer, G. E. and E. Foster. 1981. A guide for predicting sheet and rill erosion on forest lands. USDA, Forest Service, Georgia, USA, pp. 40.

Dissmeyer, G. E. and E. Foster. 1985. Modifying the universal soil loss equation for forest land. In Soil Erosion and Conservation. Eds. S.A. El-Swaify, C. Moldenhauer and A. Lo, Soil Conservation Soc. of America, pp. 480-495.

Ditkontan. 1985. Petunjuk Pelaksanaan Penyusunan RTL-RLKT. Direktorat Konservasi Tanah, Ditjen RRL, Department Kehutanan, pp. 47.

Doyle, J. 1983. Methodological simplicity in expert system construction: The case of judgement and reasoned assumption. AI magazine. 4 (2). Summer, 1983, pp. 33-43.

Drummond, J. 1987. A framework for handling error in geographic data manipulation. ITC Journal 1987-1. ITC, Enschede, The Netherlands, pp. 73-82.

Drummond, J. 1989. Model and data quality handled in a Dbase/AutoCad GIS. ITC Journal 1989-3/4. ITC, Enschede, The Netherlands, pp. 205-211.

Drummond, J. 1991. Determining and processing quality parameters in geographic information systems. Unpublished doctorate thesis. Department of Surveying, University of Newcastle Upon Tyne. 247 pp.

Dunne, T. 1977. Evaluation of erosion conditions and trends. In Guidlines for Watershed Management 1. FAO Conservation Guide, Rome. FAO.

Edward, G. 1994. Modelling fuzzy data. Aggregation and disaggregation of fuzzy polygons for spatial-temporal modelling. In advanced geographic data modelling. Spatial data modelling and query languanges for 2D and 3D. Eds. M. Molenaar and S. de Hoop. Netherlands Geodetic Commission. Publication on Geodesy, new series, nr. 40, pp. 141 - 154.

Ehlers, J., G. Edwards and Y. Bedard, 1989. Integration of remote sensing with geographic information systems: a necessity evolution. Photogrametric Engineering and Remote Sensing, Vol.55, No.11, pp 1619-1627.

Elsenbeer, H. and D. K. Cassel and W. Tinner. 1993. A daily rainfall erosivity model for Western Amazonia. Journal Soil and Water Conservations, Vol. 48(5), pp. 439-444.

Epema, G. 1986. Processing Thematic Mapper data for mapping in Tunisia. ITC Journal, 1986-1, Enschede, pp. 30-42.

Eppink, L. A. A. J. 1985. Soil conservation and erosion control. Lecture notes. Department of land and water use. Wageningen Agricultural University, pp. 73.

Erdas, 1991. Erdas field guide. Version 7.5, July 1991, Atlanta GA, USA, 394 pp.

Evans, R. 1990. Water Erosion in British farmer's fields - some causes, impacts, prediction. Progress in physical geography. Edward Arnold, Mill Road, Sevenoaks, Kent, pp. 199-219.

Evans, R. and J. Boardman. 1994. Assessment of water erosion in farmers' field in the UK. In conserving soil resources. European perspective. Selected papers from the First International Congres of the European Society for Soil Conservation. Ed. R.J. Rickson. CAB International, Willingford, Oxon, UK, pp. 232 - 233.

Eweg, Rik. 1994. Computer supported reconnaisance planning. Implementing a planning methodology with a geographical information system in Noord-Brabant, The Netherlands. Doctorate thesis. Wageningen Agricultural University, Wageningen, p.227.

FAO. 1976. A framework for land evaluation, International Institute for land reclamation and improvement (ILRI). Publication 22, 87pp.

Freebairn, D. M., D. M. Silburn and R.J. Loch. 1989. Evaluation of three soil erosion models for clay soils. Australian Journal Soil Res. Vol. 27, pp. 199-211.

Field Engineering Design Plan (FEDP) Team. 1986. Rencana teknik lapangan rehabilitasi lahan dan konservasi tanah (RTL-RLKT). Sub daerah aliran sungai Ciseel. Buku II dan III: Data dan Peta. Sub BRLKT Citanduy-Cisanggarung, Pamalayan, Ciamis. Departemen Kehutanan, Ditjen RRL, BRLKT wilayah IV. 192 pp.

Fisher, P. F. 1992. Probable and fuzzy models of the viewshed operation. In Innovation in GIS Eds. M.F. Worboy. Taylor and Francis, London. pp. 161-174.

Forstner, W. 1993. A future for photogrammetric research. NGT geodesia, No 8, pp 372- 383.

Foster, G. R. 1991. Advances in wind and water erosion prediction. Journal of Soil and Water Conservation, Vol. 46(1), pp.27-29.

Fox, J. and J. Chow. 1988. Geographic information systems for rural development: appropriate technology or white elephant? Working paper. The East West Environment and Policy Insitute. East-West Center, Honolulu, Hawaii. 34pp.

Goodchild, M. and S. Goval. 1989. Accuracy of spatial databases. Taylor and Francis, London, UK. 289 pp.

Greenland, A., R. M. Socher and M.R. Thomson. 1985. Statistical evaluation of accuracy for digital cartographic data bases. Proceeding Aerocarto 7, ASP-ACSM, Washington DC.

Grumazescu, H., G. Stancalie and C. Ungureanu.1990. Remote sensing techniques for determining the regionally variable characteristics of drainage basins. In regionalization in hydrology. Eds. M.A. Beran, M. Brilly, A. Becker, O. Bonacci. IAHS Press, Inst. of Hydrology, Oxfordshire. UK, pp. 3-10.

Hay, A. M. 1979. Sampling designs to test land use map accuracy. Photogrammetric Engineering and Remote Sensing, Vol. 45, no.4, pp. 529-533.

Heckerman, D. 1988. Probabilistic interpretations for Mycin's Certainty Factor. In uncertainty in artifical intelligence. Eds. L.N. Kanal and J.F. Lemmer. Machine intelligence and pattern recognition 4. North-Holland, Amsterdam, New York, Oxford, Tokyo, pp.167-195.

Hendriks, M. R. 1990. Regionalisation of hydrological data. Effects of lithology and landuse on storm runoff in East Luxembourg. Doctorate Thesis. Free University. Amsterdam, pp.168.

Heuvelink, G. B. M., P. A. Burrough and A. Stein. 1989. Propagation of errors in spatial modelling with GIS. International Journal of Geographical Information Systems, Vol. 3, no.4. Taylor and Francis. London, New York and Philadelphia, pp. 303-322.

Heuvelink, G. B. M. 1993. Error propagation in quantitative spatial modelling. Application in geographical information systems. PhD Thesis. The Netherlands Geographical Studies, The State University of Utrecht, Utrecht, p.151.

Hill, J., M. O. Smith, W. Mehl and M. Altherr. 1993. The use of aviris data for land degradation and soil erosion hazard mapping in a mediterrean Ecosystem. In the Proceeding the 25th International Symposium, Remote Sensing and Global Environment Change. Graz, Austria, pp. I(747-758).

Hill, J. 1993. Land degradation and soil erosion hazard mapping in Mediterranean environments with operational Earth observation satellites. In Proceedings of the International Symposium "Operationalisation of Remote Sensing", ITC, Enschede, pp. 41 -51.

Hill, J. 1995. Land degradation mapping and monitoring. The role of remote sensing with Earth observation satellites. Final report. Eds. R. Fantechi, P.Balabanis and J.L. Rubio. Directorate General Science, Research and Development, pp.247-261.

Hirschi, M. C. and B. J. Barfield. 1988. KYERMO-A physical based research erosion model part I. Model development. American Society of Agricultural Engineers. Transactions of the ASAE, pp. 804-813.

Hord, R. M. and W. Brooner. 1976. Land use map accuracy criteria. Photogrammetric Engineering and Remote Sensing, Vol. 42 no 5.

Hudson, N. 1971. Soil conservation. B.T. Batsford Limited, London, pp. 324.

Huising, J. E. 1993. Land use zone and land use patterns in the atlantic zone of Costa Rica. Doctorate Thesis. Wageningen Agricultural University, the Netherlands, 222 pp.

Imeson, A. C. 1984. Desertification in Europe. In Proceedings of the Information Symposium in the EEC Programme on Climatology. Eds. R. R. Fanthechi and N.S. Margaris. D. Reidel Publishing Company. Dordrecht, The Netherlands, pp. 110 - 125.

Imeson, A. C. and J. M. Verstraten, 1988. Rills on badlands slopes: a physico-chemically controlled phenomenon. Catena Suplement 12. Catena verlag. Brunschweig, pp. 139-150.

Imeson, A. C., L. H. Cammeraat and M. Boer. 1992. Mediterranean desertification and land use. Final report Amsterdam group. Landscape and Environmental Research Group. University of Amsterdam, Amsterdam.

Imeson, A. C. 1995. Soil erosion in Mediterranean environment. In desertification in a European context:physical and socio-economic aspects. Final report. Eds. R. Fantechi, P.Balabanis and J.L. Rubio. Directorate-General Science, Research and Development, pp.153-168.

International Cartographic Association (ICA).1973. Multilingual dictionary of technical terms in Cartography. Franz Steiner Verlag GMBH. Wiesbaden, Germany, pp. 108-109.

Janssen, L. L. F. 1993. Methodology for updating terrain object data from remote sensing data. The application of landsat TM data with respect to agricultural fields. Doctorate thesis. Wageningen Agricultural University. Wageningen, 171 pp.

Kandel, A. 1986. Fuzzy mathematical techniques with application. Massachusetts: Addison-Wesley.

Keller, R. 1987. Expert system technology. Development and application. Yourdon Press. A Prentice Hall Company, Engelwood, Cliff, New Jersey, 246pp.

Kendall, M. G. and A. Stuart. 1968. The advanced theory of statistics. Vol. 3, Design and analysis and time-series. Charless Griffin and Co, Ltd., London, UK, 690 pp.

Kendall, M. G. and W. R. Buckland. 1988. A dictionary of statistical terms. Oliver, Boyd, Edinburgh, Scotland, UK, 166pp.

Kwaad, F. J. P. M. 1994. A splash delivery ratio to characterize soil erosion events. In conserving soil resources. European perspective. Selected papers from the First International Concres of the European Society for Soil Conservation. Ed. R.J. Rickson. CAB International, Willingford, Oxon, UK, pp. 264 - 272.

Laflen, J.M., W. J. Elliot, J. R. Simanton, C. S. Holzhey and K. D. Kohl. 1991. WEPP soil erodibility experiments for rangeland and cropland soils. Journal of Soil and Water Conservation, pp. 39-44.

Lane, L. J., E.D. Shirley and V. P. Singh. 1988. Modelling erosion on hillslopes. In Modelling Geomorphological Systems. Ed. M.G. Anderson. John Wiley and Sons Ltd., pp. 287-308.

Lemmens, M. J. P. M.1993. Kwaliteitsbeheer van ruimtelijke informatie. NGT Geodesia 93-9, pp. 406-412.

Loch, R. J., D. M. Silburn and D. M. Freebairn. 1989. Evaluation of the CREAMS model. II. Use of rainulator data to derive soil erodibility parameters and prediction of field soil losses using derived parameters. Australian Journal Soil Resources, Vol. 27, pp. 563-576.

Meijerink, A. M. J.1988. Data acquistion and data capture through terrain mapping units. ITC Journal 1988-1, Enschede, The Netherlands, pp. 23-44.

Middelkoop, H. 1990. Uncertainty in a GIS: a test for quantifying interpretation output. ITC Journal, 1990-3. Enschede, pp 225-235.

Millington, A. C. 1986. Reconaissance scale soil erosion mapping using a simple GIS in humid tropics. In Siderius, W. (Ed.), 1986: Land Evaluation for Landuse Planning and Conservation in Sloping Areas. ILRI Publication Nr. 40. Wageningen, pp. 64-81.

Molenaar, M. 1989(a). Toward a geographic information theory. ITC Journal, 1989-1, Enschede, pp. 5 - 11.

Molenaar, M. 1989 (b). Single valued vector maps. A concept in GIS. Geo-Information Systems, 2(1), pp. 18 - 26.

Molenaar, M. 1991. Status and problems of geographical information systems. The necessity of geoinformation theory. ISPRS Journal of Photogrammetry and Remote Sensing, 46 (1991), pp.85 - 103.

Molenaar, M. 1993. Object hierarchies and uncertainty in GIS or why is standardisation so difficult. Geo-Information Systemen, Vol. 6, No. 4. pp. 22-28.

Molenaar, M. 1993. Terrain objects, data structures and query space. In Geoinformatik, Schilder, M. (ed.), Munich, Germany, pp 53 - 70.

Molenaar, M. and L. L. F. Janssen, 1994. Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing. In Remote Sensing and Geographical Information Processing. Concepts and application for land use monitoring and agriculture. A selection of papers of the research programme VF 89.24. Ed. M. Molenaar, Wageningen Agricultural University, Wageningen, pp 1-6.

Molenaar, M. 1995 (a). The relationship between the extensional and geometric uncertainty of spatial data. Paper presented at GIS data modelling. European Science Foundation (ESF), Lisbon, Portugal.

Molenaar, M. 1995 (b). An introduction into the theory of topologic and hierarchical object modelling in geo information systems. Departmental Lecture Note. In preparation for publication.

Morgan, R. P. C. 1974. Estimating regional variation in soil erosion hazard in Peninsular Malaysia. Malay National Journal 28 (2).

Morgan, R. P. C. 1986. Soil erosion and conservation. Longman, England, pp. 298.

Mortimer, H., 1985. The logic of induction. Ellis Horwood Limited. Chichester. pp.182.

Mosley, M. P. 1981. Delimitation of New Zealand hydrological regions. Journal of Hydrology 49, pp.173 - 192.

Mulders, M. A. and G. F. Epema, 1986. The Thematic Mapper: a new tool for soil mapping in arid areas. ITC Journal, 1986-1. Enschede, pp. 24-29.

Nearing, M. A., G.R. Foster, L.J. Lane and S. C. Finkner. A process-based soil erosion model for USDA-water erosion prediction project technology. American Society of Agricultural Engineers. Transaction of the ASAE. Vo. 32(5), pp. 1587-1593.

Negoita, C. V. 1985. Expert systems and fuzzy systems. The Benjamin/Cumming Publishing Company, Inc. Menlopark, California. pp.189.

Nortcliff, S. 1986. Soil loss estimation. Progress in Physical Geography, Vol.10, pp. 1-12.

Offermas, D. 1986. Kroonschuwheid in grootschalige luchtopnamen van piptadeniastrum Africanum (Mimosaceae). Scriptie 1986 - 2. Vakgroep Landmeetkunde en Teledetectie, LUW. Wageningen. pp.51.

O'Neill, R. 1988. Hierarchy theory and global change. In scale and global change. Eds. T. Rosswall, R. G. Risser. Scientific committee on problems of the environment (SCOPE). John Wiley and Sons. ltd. pp. 29-45.

Onstad, C.A. and G.R. Foster. Erosion modelling on a watershed. 1975. Transactions of the ASAE, Vol. 18(2), pp. 288-292.

Openshaw, S. 1987. An automated geographical analysis system. Environment and Planning, A, 18, pp. 143.

Peuquet, D. J. 1984. A conceptual framework and comparison of spatial data models. Cartographica, Vol. 21, No.4, pp. 66-113.

Pickup, G. and V. H. Chewings. 1986. Mapping and forecasting soil erosion patterns from Landsat on a microcomputer-based image processing facility. In Aust.Rangel. Journal. 8(1). 1986. pp. 57-62.

Poesen, J. and F. Ingelmo-Sanchez. 1992. Runoff and sediment yield from topsoils with different porosity as affected by rock fragment cover and position. Catena 19, pp. 451-474.

Poesen, J. 1995. Soil erosion in a Mediterranean environment. In Desertification in a European context:physical and socio-economic aspects. Final report. Eds. R. Fantechi, P. Balabanis and J. L. Rubio. Directorate-General Science, Research and Development, pp.123-151.

Resources Management International (RMI).1986. Resources inventory of the Citanduy river basin. Draft final report submitted to the Ministry of Public Work, Ministry of Forestry, Ministry of Home Affair, Ministry of Agriculture Republic Indonesia. Vol. II, maps of the basin at 1: 250 000 scale.

Rodrigue, J. P. 1995. The heuristic classification of functional land use: A knowledge-based approach. In Geographical Systems. Vol. 2. Amsterdam, pp. 103-120.

Rose, C.W., J. R. Williams, G. C. Sander and D. A. Barry. 1983. A mathematical model of soil erosion and deposition processes: I. Theory for a plane land element. Soil Science Society of America. Vol. 47, pp. 991-995.

Rose, C. W. and D. M. Freebairn. 1985. A new mathematical model of soil erosion and deposition processes with applications to field data. In Soil Erosion and Conservation. Eds. S.A. El-Swaify, C. Moldenhauer and A. Lo, Soil Conservation Soc. of America, pp. 549-557.

Schum, S. A. and M. P. Mosley. 1973. The role of creep and rainwash on the retreat badlands slope. In slope morphology. Eds. S. A. Schum and M. P. Mosley. Benchmark Books Publishing Programme. Dowden, Hutchinson and Ross Inc. Stadsburg, Pennsylvania, pp. 287-300.

Seubert, C. E., M. F. Baumgardner and R. A. Weismiller. 1979. Mapping and estimating areal extent of severely eroded soil of selected sites in northen Indiana. In Proceeding Symposium Machine Processing of Remotely Sensed Data, pp. 234-238.

Shortliffe, E. H. and B. G. Buchanan. 1984. A model of inexact reasoning in medicine. In Rule-Based Expert Systems. The Mycin experiments of the standford heuristic programming project. Eds. B. G. Buchanan and E. H. Shortliffe. Addison-Wesley Publishing Company. Reading, Massachusetts, Sydney, pp. 233 - 262.

Silburn, D. M. and R. J. Loch. 1989. Evaluation of CREAMS Model. Sensitivity analysis of the soil erosion/sedomentation component for aggregated clays sois. Australian Journal Soil Res., Vol. 27, pp. 545-561.

Smith, J. M. and D. C. P. Smith. 1977. Database abstractions: Aggregation and generalization. Association for Computing Machinery (ACM) transaction on Database Systems, Vol. 2, no. 2, pp. 105-133.

Smith, J. W. F. and I. A. Campbell. 1989. Error in polygon overlay processing of geomorphologic data. In Earth Surface and Landforms, Vol. 14. John Wiley and Sons, Ltd., pp. 703 - 717.

Soemarwoto, O. 1985. Ekologi, Lingkungan hidup dan pembangunan. Penerbit Djambatan. Anggota IKAPI. Jakarta, Indonesia, pp.357.

Stanat, D. F. and D. F. McAllister. 1977. Discrete mathematics in computer science. Prentice Hall, Inc., Englewood Cliffs, N. J., pp. 401.

Stocking, M. A. 1981. A working model for the estimation of soil loss suitable for underdeveloped areas. Occasional paper No. 15, School of Development Studies, UEA, Norwich, UK, pp. 56.

Stocking, M. A. 1984. Erosion and soil productivity: A review. Soil conservation programme land and water development division. Consultants' working paper No. 1 AGLS, FAO in cooperation with UEA Norwich, UK, pp. 102 .

Stocking, M. A., Q. Chakela and H. A. Elwell. 1988. An improved method for erosion hazard mapping. Part I: The technique. Geografiska Annaler Vol. 70 (A.3), pp. 169-180.

Sub Balai RLKT Citanduy-Cisanggarung. 1985. Laporan penelitian Erosi Pada Plot Standard (Termasuk Penetapan Nilai C and P) di wilayah DAS Citanduy Tahun 1985/1986. Petikan Hasil Penelitian Komponen Paket Teknologi Pertanian di Bagian Hulu DAS Citanduy. Ed.Sekretariat Team Koordinasi Pusat Proyek Pengembangan Wilayah DAS Citanduy, 1988.Departement Kehutanan, pp. 93 - 104.

Sukresno. 1991. Evaluation of USLE, RUSLE and MUSLE on steep watershed in Java Indonesia. MSc Thesis. The Graduate College at The University of Nebraska. Lincoln, Nebraska, pp.122.

Suryana, N. 1980. The effect of soil tillage and different vegetable crops on soil erosion and yield. Unpublished thesis. Faculty of Agriculture. Department of Soil Sciences. Padjadjaran University. Bandung, Indonesia, 123 pp.

Suryana, N. 1991. The use of multiplying effect of ancillary data in erosion severity classification; a method for improving data quality of multispectral classification. In Proceedings of EGIS 91, Vol.2, the decond European Conference on geographic information system, Brussels, Belgium. Eds. J. Harts, Ottens, H.P.L., Scholten, H.J.. EGIS Foundation, Utrecht, The Netherlands, pp.1046-1057.

Suryana, N. 1992. The use of a basic mapping unit and an inductive erosion model (IEM) in geographical information systems. In proceedings of EGIS 92, Vol.2, the third European Conference on geographic information system, Munich, Germany. Eds. J. Harts, H. P.L. Ottens, H.J. Scholten. EGIS foundation, Utrecht, The Netherlands, pp. 1012-1020.

Suryana, N. 1993. The application of evidence theory and fuzzy subset in handling uncertain information and imprecision data handled by geographical information systems. In Proceedings of the Seventh Annual Symposium GIS 93. British Columbia, Vancouver, Canada, Feb 15 -18, 1993, pp. 872-882.

Suryana, N. and S. de Hoop. 1994. The hierarchical structure of terrain mapping unit. In Proceeding of EGIS 94, Vol.1, the Fifth European Conference on geographic information systems, Paris, France. Eds. J. Harts, H.P.L. Ottens, H. J. Scholten. EGIS Foundation, Utrecht, The Netherlands, pp. 869-877.

Suryana, N. and M. Molenaar. 1995. An observational approach to erosion modelling at local and sub-regional levels. In the proceeding of GIS AM/FM Asia '95. 21 -24 August 1995. Bangkok, Thailand, pp.d1(1-10).

Suryana, N., M. Molenaar and A. C. Imeson. 1996. An approach to building and handling uncertainty associated with the inductive erosion model in geographical information systems (GIS). The International Journal of Geographical Information Analysis, Theory, and Decision. Gordon Breach Science Publisher, UK. (Accepted for publication).

Suwardjo and S. Abunyamin. 1983. Crop residue mulch for conserving soil in upland agriculture. Paper presented in Malama Aina, pp. 13.

Snedecor, G. M. and W. G. Cochran. 1989. Statistical methods. 8th edition: AMES-IOWA State University Press, USA, 503 pp.

Stroosnijder, L. and L. A. A. J. Eppink (ed.), 1993. Principles of Soil and Water Conservation. Lecture Notes of Course K200-500/510. Wageningen Agr. Univ. Wageningen, The Netherlands.

Thornes, J. B.1995.Mediterranean desertification and the vegetation cover. In Desertification in a European context:physical and socio-economic aspects. Final report. Eds. R. Fantechi, P.Balabanis and J.L. Rubio.Directorate-General Science, Research and Development, pp.169-194.

Topping, J. 1979. Error of observation and their treatment. Fourth Edition. Chapman nad Hall. London, pp.119.

Universitas Padjadjaran (UNPAD). 1983. Survai dan pemetaan tanah semi detail (kesesuaian lahan) daerah Sub DAS Cimuntur, Ciseel, Cikawung dan Citanduy Hulu wilayah DAS Citanduy. Final report. Vol. II lampiran peta-peta.

Utomo, W. H. and N. Mahmud. 1984. The possibility of using the universal soil loss equation in mountanious areas of East Java with humus rich Andosol. In Proceedings of Soil Conservation 5th Conference, Bangkok, Thailand, pp. E5(1-13).

Van der Laan, F. B. 1988. Improvement of classification results from a satellite image using context information from a topographic map. Proc. 16th ISPRS concress, Vol.27, B8, Cie7, pp 98-108.

Van Zuidam, R. A. 1985. Aerial photo interpretation in terrain analysis and geomorphologic mapping. Smith Publ., The Hague, The Netherlands.pp.442.

Wadge, G. 1988. The potential of GIS modelling of gravity flows and slope instabilities. International journal of. geographical information systems. Vol 2, No.2. Taylor and Francis. London, New York and Philadelphia.pp. 143-152.

Walker, P. A. and D. M. Moore. 1988. SIMPLE. An inductive modelling and mapping tool for spatially-oriented data. International journal of geographical information systems. Vol.2, No. 4. Taylor and Francis. London, New York and Philadelphia, pp. 347 - 363.

Ward, J. H. 1963. Hierarchical grouping to optimise an objective function, Journal of the American Statistical Association, pp. 236-244.

Webster, C. 1991. The object oriented paradigm in GIS. Working paper. Wales and South West England Regional Research Laboratory, Department of City and Regional Planning, Univ. of Wales College of Cardiff, UK, 23 pp.

Wesseling, C. G. and G. B. M. Heuvelink. 1991. Semi-automatic evaluation of error propagation in GIS operations. In Proceeding EGIS 91, Vol.2, the decond European Conference on geographic information system, Brussels, Belgium. Eds. J. Harts, H. P. L. Ottens, H. J. Scholten. EGIS Foundation, Utrecht, The Netherlands, pp.1228-1237.

Wheeler, D. J. and R. W. Lyday.1990. Evaluating the measurement process. 2nd edition. Addison-Wesley publishing company, Wokingham, UK, pp.106.

Wiersum, K. F. and S. Ambar. 1981. Tropical ecological forest research in Indonesia. Vegetation and erosion in Jatiluhur area. Final Report of PUO/NUFFIC Project LHW/8. The Institute of Ecology, Bandung, Indonesia. Department of Sylviculture. Wageningen Agricultural University, pp. 43.

Wiersum, K. F. 1990. Soil erosion and conservation in agroforestry systems. In Biophysical Research for Asian Agroforestry. Eds. M.E. Avery, M.G.R. Cannell and C.K. Ong. Winrock International, USA and South Asia Books, USA, pp. 209 - 230.

Wischmeier, W. H. 1976. Use and misuse of the universal soil loss equation. Journal of soil and water conservation. Edision January - February, 1976. pp. 5-9.

Wischmeier, W. H. 1978. Estimating the soil loss equation, cover and management factors for undisturbed areas. Agric. Res. Service USDA.

Wischmeier, W. H. and D. D. Smith. 1978. Predicting rainfall erosion losses. Agricutural handbook no.537, USDA-science and education administration, pp. 58.

Samenvatting

Suryana, N., 1996. *Een geo-informatieve theoretische benadering van inductieve erosiemodellering gebaseerd op terreinkarteringseenheden.* Proefschrift ter verkrijging van de graad van Doctor, Landbouw Universiteit Wageningen, Wageningen, Nederland, (xxvi) + 235 pp.

Drie belangrijke aspecten waarop het onderzoek zich gericht heeft, namelijk het begrip object-oriëntatie, de ontwikkeling van het *Inductive Erosion Model* (IEM) en het ontwikkelen van een stramien voor het omgaan met onzekerheid in de gegevens of de informatie van een GIS, zijn met elkaar verweven in dit proefschrift. Met het eerste en het tweede aspect van het proefschrift zijn tegelijkertijd de toepassing van een basiskarteringseenheid en een IEM in een GIS-omgeving besproken. Het doel van het bestuderen van deze aspecten was om een flexibele en gemakkelijk bij te houden, alternatieve oplossing te bieden voor problemen die samenhangen met gegevensinwinning, gegevensinvoer en het produceren van geaggregeerde informatie voor een GIS.

Met het derde aspect is de toepassing van formele behandeling van onzekerheid besproken, zoals het gebruik van de standaarddeviatie, de kans op verkeerde classificatie, mate van lidmaatschap en evidentietheorie voor het omgaan met fouten in gegevens of onzekerheid die samenhangt met de invoer van gegevens in een GIS in het algemeen, en in het *Indonesian Field Engineering Design Plan* (FEDP) in bijzonder. Het doel was met name om een raamwerk te creëren voor het weergeven van onzekerheid bij de manipulatie van geografische gegevens. GIS classificatiemodellen, de kenmerken van GIS-modellen, toepasbaarheid, soorten onzekerheid inclusief onnauwkeurigheid en onzekerheid als gevolg van variabiliteit (fouten) of vaagheid en een voorstel voor een conceptueel raamwerk gebaseerd op het concept van zekerheidsfactoren zijn besproken.

Bij het onderzoek is uit gegaan van de vorming van stabiele basiseenheden voor de kartering, die het mogelijk maken om herhaalbare observatieprocedures te definiëren. Deze oplossing werd vooral van toepassing geacht in situaties waar hoogwaardige programmatuur en gegevens van goede kwaliteit niet voorhanden zijn. In het onderzoek zijn *terrain mapping units* (TMU's) gedefinieerd als een combinatie van geologische, geomorfologische, morfometrische en bodemkundige kenmerken, meestal verkregen uit interpretatie van luchtfoto's of SPOT-beelden. Terrein beelden met homogen reliëf wordt geïdentificeerd, afgebakend en in het veld geverifieerd. De afgebakende TMU's vertegenwoordigen natuurlijke terreineenheden met vaak duidelijke grenzen.

De attributen die aan de gevormde TMU's waren verbonden zijn geselecteerd en gebruikt om de TMU's te classificeren. Er is een classificatiehiërarchie gevormd met het oog op object-georiënteerde modellering, inclusief abstractie, overerving, aggregatie en associatie van terreinobjecten. De hiërarchie heeft drie niveaus, namelijk niveau +1 (superklasse niveau of TMU), 0 (sub-TMU-niveau) en -1 (elementair object ofwel subsub-TMU). Een lager niveau in de classificatiehiërarchie vertegenwoordigt meer gedetailleerde of specialistische informatie.

Het bekende deductieve erosiemodel de *Universal Soil Loss Equation* (USLE) is incompleet in het voorspellen van ruimtelijke erosieprocessen. Verfijndere modellen (b.v. *CREAMS, ANSWERS, EPIC, WEPP, GAMES*) zijn er niet in geslaagd om de complexiteit van erosieprocessen mee te nemen en er zijn geen mogelijkheden om voorspellingen volgens deze modellen te bevestigen. Als alternatief voor het probleem werd een inductieve benadering (van onderop) voorgesteld. Deze heeft tot een *Inductive Erosion Model* geleid, dat is opgebouwd uit waarnemingen met dynamische (herstellingsvermogen) en statische (inertie), locatiespecifieke erosievormende factoren in een of meer trainingsgebieden, die ter plekke zijn gemaakt op het niveau van de akker van de boer: de beste functionele eenheid om een erosieklasse op lokaal niveau te beschrijven. Het IEM is daarom regio-specifiek. Als het IEM eenmaal is geformuleerd en getest voor elk type TMU, dan kan het binnen een GIS-omgeving worden gebruikt als een aanvaardbaar middel om veilig de hevigheid van de bodemerosie voor het gehele studiegebied te voorspellen. De door het IEM voorspelde erosie-hevigheidsklassen zijn beschouwd als actieve of dynamische attributen van de gevormde TMU's. Per definitie voorziet de TMU inherent in de erosievormende factoren, de zogenaamde terreinkarakteristieken inclusief de morfometrie, geologie, bodem en grondbedekking. Het voorgestelde IEM is bedoeld om homogene erosie-hevigheidsklassen te kunnen voorspellen, die zijn gerelateerd aan TMU's op verschillende aggregatie- of hiërarchische niveaus. De aggregatieniveaus zijn gerelateerd aan puntwaarnemingen, *farmer's field level* (FFL) en grotere delen van het terrein. De discussie over dit aspect spitste zich toe op de rol van de geïnterpreteerde thematische karteringseenheid (TMU), gedefinieerd als een natuurlijk terreineenheid die in de waarnemingsprocedure invoer levert aan het *Inductive Erosion Model.*

De vastgestelde hiërarchische karteringseenheden vormen de basis voor de inductieve modellering van erosie, waarbij tevens gebruik gemaakt wordt van op specialistische kennis gebaseerde, logisch afgeleide regels. Er is een multi-schaal benadering gevolgd en de inductieve erosiemodellering is geïmplementeerd in een GIS-omgeving. Toepassing van de begrippen regionalisatie, waargenomen patroon, en beslissingsregels om te voorspellen en te modelleren zijn besproken. Op regionaal niveau zijn patronen die geassocieerd worden met de belangrijkste eroderende processen, zoals laag-, ril-, geul- en ravijnerosievormen, meestal nog herkenbaar op luchtfoto's op schaal 1:50 000. Meer gedetailleerde informatie over dit soort actieve processen op lokaal niveau kan echter alleen worden verkregen uit nadere studie, d.w.z. bestudering van de erosie op het FFL. In dit opzicht is het FFL beschouwd als een geschikte functionele basiseenheid om erosie op lokaal niveau te beschrijven.

In plaats van kansberekening toe te passen - waarbij aan statische eisen moet worden voldoen - staan de productieregels het invoeren van een gewicht, de *Certainty Factor* (CF) toe om met onzekerheid om te gaan in zowel de gegevens als in GIS-modellen. De CF kan worden verkregen uit een subjectieve beoordeling door specialisten; hij komt vanzelfsprekend voort uit óf wel het afleiden van onderliggende processen, óf wel het schatten van de kwaliteit van de gebruikte gegevens en modellen. De toepassing van het begrip CF is met name gericht op de situatie waarbij het niet waarschijnlijk is dat procedures en technieken voor kansberekening en het verkrijgen van kwantitatieve informatie kunnen worden uitgevoerd.

In het licht van de evidentietheorie is het IEM voor het voorspellen van de hevigheid van de erosie in een specifiek TMU geformuleerd als een functie van verschillende zekerheidsfactoren van ruimtelijke erosievormende factoren. De zekerheidsfactor heeft een waarde tussen -1 en +1, en geeft de geschatte verandering aan in het geloof dat een TMU in een bepaalde erosieklasse valt als bewijs is verzameld (van kaarten, luchtfoto's, veldwaarnemingen etc.) voor elk van de bijdragende factoren. De erosie-hevigheidsklasse waaraan een TMU uiteindelijk wordt toegekend is degene met een totale zekerheidsfactor die het dichtst bij +1 ligt. Dit is voorgesteld als een methode om om te gaan met onzekere informatie veroorzaakt door incompleetheid, en omvat gevolgtrekkingen die zijn vastgesteld door specialisten en zijn verkregen uit een serie waarnemingen waarbij ook het effect van causale relaties tussen verschillende bewijzen van onzekerheid wordt betrokken.

Trefwoorden: functionele eenheden op basis van waarneming, inductieve modellering, benadering van onderaf, geografische informatiesystemen, waargenomen patronen, fenomenologische processsen, fouten en onzekerheid, vaagheid, maten voor vaagheid.

About the author

Nanna Suryana Warsitakusumah (Nanna) was born in Tasikmalaya regency, West Java, Indonesia, on June 13th, 1956. The study area includes the greater part of this regency. He graduated at the Department of Soil Sciences, Faculty of Agriculture, Padjadjaran University (UNPAD), Bandung in the year 1980, with the degree of Agricultural Engineer (Ir), with distinction. He took soil and water engineering as a major subject. Since 1979-1980, he worked as a laboratory technician, an assistant lecturer at the same faculty and as a freelance soil surveyor for a private consulting company. Thereafter, he worked as Soil Scientist and Landuse Specialist for an American company, the Black and Veatch International Consulting Company. For this job he was stationed in South Kalimantan Province. This position involved him in soil investigation, soil mapping for agricultural development and transmigration settlements, forest inventory, swampy area development and feasibility studies in most provinces of Kalimantan.

From the end of 1982 until 1984 he worked for the Directorate of Soil Conservation, Directorate General Reforestation and Land Rehabilitation, Ministry of Forestry. This job involved him in watershed management development planning, located in most provinces of Indonesia.

In the year 1985 he went to The Netherlands to study at the International Institute for Aerospace Survey and Earth Sciences (ITC), Enschede. In the year 1987, he obtained there a MSc degree in Survey Integration for Resources Development Planning. The research for this degree involved considerable work with the Oracle and ITC GIS-System, ILWIS. A couple of months after his return to Indonesia, he was promoted as Chairman of civil engineering soil conservation techniques in the Directorate. In the year 1988 and in cooperation with the East West Centre, Hawaii, he was involved in testing the effectiveness of the IDRISI GIS software in establishing the Field Engineering Design Plan (FEDP) for the Land Rehabilitation and Soil Conservation Programme in Indonesia.

In 1989, with his research interest in his mind, he took the status of unpaid leave of absence from his organisation and went back to The Netherlands to continue his studies. In the period until 1991 he followed selected courses in the Management of Agricultural Knowledge Systems (MAKS) offered by the Department of Communication and Innovation Study, Wageningen Agricultural University (WAU). At the end of 1991 he was admitted officially to conduct doctorate research in the Department of Geographic Information Processing and Remote Sensing on a self sufficiency basis. His general research interest is the uncertainty and efficiency of GIS in natural resources and erosion hazard assessment. The particular research of which this thesis is a product, was done under the supervision of Prof. M. Molenaar (WAU), Prof. R. Atmawidjaja (BAKOSURTANAL, Indonesia) and Prof. A. Imeson (University of Amsterdam), with much assistance from Dr. J.E. Drummond (now of Glasgow University, UK, but stationed in Indonesia, for ITC, during the author's field research period). The author has published findings related to his ongoing work in international journals, and has presented papers at several international conferences. In the same period he found it of great interest to be involved in lecturing in GIS Theory for MSc students at the WAU.